

## 基于深度学习的荧光显微性能提升 (特邀)

熊子涵<sup>1,2</sup>, 宋良峰<sup>1,2</sup>, 刘欣<sup>1</sup>, 左超<sup>1</sup>, 郜鹏<sup>1\*</sup>

1. 西安电子科技大学物理学院, 陕西 西安 710071;
2. 西安电子科技大学杭州研究院, 浙江 杭州 311200)

**摘要:** 荧光显微镜具有对样品损伤小、可特异性成像等优点, 是生物医学研究的主流成像手段。随着人工智能技术的快速发展, 深度学习在逆问题求解中取得了巨大成功, 被广泛应用于诸多领域。近年来, 深度学习在荧光显微成像中的应用掀起了一个研究热潮, 为荧光显微技术发展提供了性能上的突破与新思路。基于此, 首先介绍了深度学习的基本网络模型, 然后对基于深度学习的荧光显微成像技术在荧光显微的空间分辨率、图像采集及重建速度、成像通量和成像质量提升方面的应用进行阐述。最后, 对目前深度学习在荧光显微成像中的研究进行总结与展望。

**关键词:** 荧光显微成像; 深度学习; 超分辨; 超分辨显微成像; 图像重建  
**中图分类号:** O439 **文献标志码:** A **DOI:** 10.3788/IRLA20220536

## Performance enhancement of fluorescence microscopy by using deep learning (*invited*)

Xiong Zihan<sup>1,2</sup>, Song Liangfeng<sup>1,2</sup>, Liu Xin<sup>1</sup>, Zuo Chao<sup>1</sup>, Gao Peng<sup>1\*</sup>

1. School of Physics, Xidian University, Xi'an 710071, China;
2. Hangzhou Institute of Technology, Xidian University, Hangzhou 311200, China)

**Abstract:** Fluorescence microscopy has the advantage of minimal invasion to bio-samples and visualization of specific structures, and therefore, it has been acting as one of mainstream imaging tools in biomedical research. With the rapid development of artificial intelligence technology, deep learning that has outstanding performance in solving sorts of inverse problems has been widely used in many fields. In recent years, the applications of deep learning in fluorescence microscopy have sprung up, bringing breakthroughs and new insights in the development of fluorescence microscopy. Based on the above, this paper first introduces the basic networks of deep learning, and reviews the applications of deep learning in fluorescence microscopy for improvement of spatial resolution, image acquisition and reconstruction speed, imaging throughput, and imaging quality. Finally, we summarize the research on deep learning in fluorescence microscopy, discuss the remaining challenges, and prospect the future work.

**Key words:** fluorescence microscopy imaging; deep learning; super-resolution; super-resolution microscopy imaging; image reconstruction

收稿日期: 2022-08-01; 修订日期: 2022-08-20

基金项目: 国家自然科学基金 (62075177); 国家重点研发计划 (2022YFE0100700, 2021YFF0700300); 中国轻工业五粮液浓香型白酒固态发酵重点实验室开放基金 (2019 JJ012); 中央高校基本科研业务费专项资金 (QTZX22039); 中波科技人员交流项目 (2021-2022); 瞬态光学与光子技术国家重点实验室开放基金 (SKLTOP202001)

作者简介: 熊子涵, 女, 硕士生, 主要从事深度学习方面的研究。

导师(通讯作者)简介: 郜鹏, 男, 教授, 博士生导师, 博士, 主要从事超分辨光学显微与定量相位显微方面的研究。

## 0 引言

在人类探索微观世界的进程中,显微镜起了决定性作用,它的发明与发展推动了生物学、医学、化学、物理学的进步。最初,人们用肉眼仅能分辨 0.1 mm 的物体。16 世纪末,显微镜的问世使得生命的基本组成部分——细胞被发现,从此打开了人们认识复杂微观生物世界的大门。精准快速观测细胞结构和其动态过程是探索生物过程的重要一步,但由于细胞结构大部分呈透明态,传统的显微技术无法观察<sup>[1-2]</sup>。相衬显微技术于 1934 年由荷兰物理学家 Zernike 提出,将光波的相位分布转化为强度分布,可为透明样品提供高衬度图像。随着数字感光器件 (CCD/CMOS) 和相移技术的问世,传统相衬显微技术逐渐被定量相位显微技术所取代。后者通过对光波相位的精确测量,不仅可以提高透明样品的成像对比度,而且可以获得样品的三维形貌或折射率分布<sup>[3-6]</sup>。然而,相衬显微和定量相位显微技术仍然缺失对细胞中特殊功能/化学结构的选择性成像能力。1911 年,Heimstadt 发明了荧光显微镜<sup>[7]</sup>。荧光显微镜借助荧光标记技术(如免疫细胞化学、原位杂交或荧光蛋白标签),可以在化学固定或活体样品中有选择地“点亮”人们感兴趣的亚细胞结构、蛋白质或基因组序列<sup>[8]</sup>。该技术不仅克服了透明样品成像衬度低的问题,还实现了对特定生物分子的选择性成像<sup>[7,9-13]</sup>,长期以来是生物医学研究的重要工具。然而,由于光学衍射极限的存在,传统荧光显微的空间分辨率只有波长的 1/2 (200 nm)。20 世纪 90 年代,人们提出了超分辨光学成像 (Super-resolution Optical Microscopy) 技术<sup>[14-17]</sup>,是光学显微成像领域最重大的突破之一,打破了光学显微镜的分辨率极限,为生命科学研究提供了先进的研究工具。然而,每种技术都会存在一些自身局限性,传统的荧光显微镜在空间分辨率、成像速度、光毒性和样品荧光漂白等方面受到限制<sup>[18-22]</sup>,有待于结合新的技术获得进一步的提高。

深度学习是一种利用深度神经网络<sup>[23]</sup>来进行自动分类、预测和学习的技术。它的概念源于人工神经网络的研究,其中人工神经网络是受生物大脑启发,基于模拟生物大脑的结构与功能,采用数学和物

理方法进行研究而构成的一种信息处理系统。与传统“基于物理模型”的算法不同,深度学习是一种“数据驱动”方法,具有处理各种非结构化数据的特点,在图像处理、语音识别等领域有着出色的表现。近年来,一些专家学者也将深度学习引入荧光超分辨领域,并取得了重要的创新成果<sup>[24]</sup>。例如实现空间分辨率的增强<sup>[25-30]</sup>、加快图像采集及重建速度<sup>[31-39]</sup>、提高成像通量<sup>[40-44]</sup>、改善成像质量<sup>[45-46]</sup>等。

综上,文中将首先简要介绍深度学习的基础理论与发展,然后从基于深度学习实现荧光显微成像空间分辨率增强、图像采集与重建速度提升、成像通量提高和质量改善等方向论述研究进展,最后对结合深度学习的荧光显微成像技术的局限性和发展方向进行总结与展望。

## 1 深度学习基本介绍与发展

深度学习,是机器学习方法中最受欢迎的算法<sup>[47]</sup>。1943 年,McCulloch 等人提出了第一个脑神经元抽象模型——麦克洛克—皮茨 (M-P) 模型,由此打开人工神经网络的大门<sup>[48]</sup>。随后,为解决非线性分类与学习的问题,Rumelhart 和 McClelland 于 1986 年提出了多层网络学习的误差反向传播学习算法 (BP 算法),引起神经网络的第二次热潮<sup>[49]</sup>。但在 1991 年 BP 算法被指出存在梯度消失问题,直至 2006 年,Hinton 提出深度神经网络 (Deep Neural Networks, DNN) 模型,使深度学习再次崛起<sup>[50]</sup>。2012 年,AlexNet 在 ImageNet 竞赛中的优异表现,引起了公众的全面关注<sup>[51]</sup>。

近年来,深度学习已被证明是一种通过训练大量成对图像来解决各种逆问题的有力工具<sup>[52]</sup>。深度学习通过利用大量成对的“输入-输出”数据集对神经网络进行训练,使得参数优化后的神经网络能够显性表达“输入”和“输出”之间的非线性映射关系。在网络训练过程中,输入图像与输出图像之间的映射函数称为目标函数,表示如下:

$$R_{\text{learn}} = \arg \min_{R_{\theta}, \theta \in \Theta} \sum_{n=1}^N F[x_n, R_{\theta}(y_n)] + g(\theta) \quad (1)$$

式中:  $R_{\text{learn}}$  是权重参数的最优解;  $x$  和  $y$  是训练集;  $F$ <sup>[53]</sup> 是具有适当误差度量的损失函数;  $g(\theta)$  是参数  $\theta$  的

正则化项,以避免过拟合<sup>[54]</sup>。利用神经网络的输出  $R_\theta(y_n)$  与目标图像  $x_n$  之间的误差,通过反向传播获取梯度,再根据梯度下降法来实现网络参数更新。因此,只要通过数据统计完成模型的建立,就可应用于成像中逆问题求解。

### 1.1 卷积神经网络

卷积神经网络 (Convolutional Neural Network, CNN) 是一类包含卷积计算且具有深度结构的前馈神经网络<sup>[55]</sup>。它利用局部连接与权重共享的网络结构,降低了网络模型的复杂度,减少了权值的数量。CNN 主要由卷积层 (Convolution Layer)、池化层 (Pooling Layer) 以及全连接层 (Fully-Connected Layer) 组成 (如

图 1(a) 所示),通常用于图像处理。其中,卷积层是构建 CNN 的核心,通过卷积运算来提取输入目标的不同特征;池化层主要作用是下采样 (Down Sampling),把数据降维,可以有效地避免过拟合;最后的全连接层将提取的特征图连接起来,根据不同任务输出我们想要的结果。CNN 中的卷积与池化操作对图像的平移、变形具有不变性,使其在生物图像分割、计算机视觉等领域得到了快速发展<sup>[56-57]</sup>。

U-Net 作为一种基于全卷积网络 (Fully Convolutional Networks, FCN) 的经典图像分割网络于 2015 年被提出<sup>[56]</sup>,在生物医学图像处理领域表现优异<sup>[59]</sup>。如图 1(b) 所示,U-Net 由编码器 (Encoder)、解码器 (Decoder)

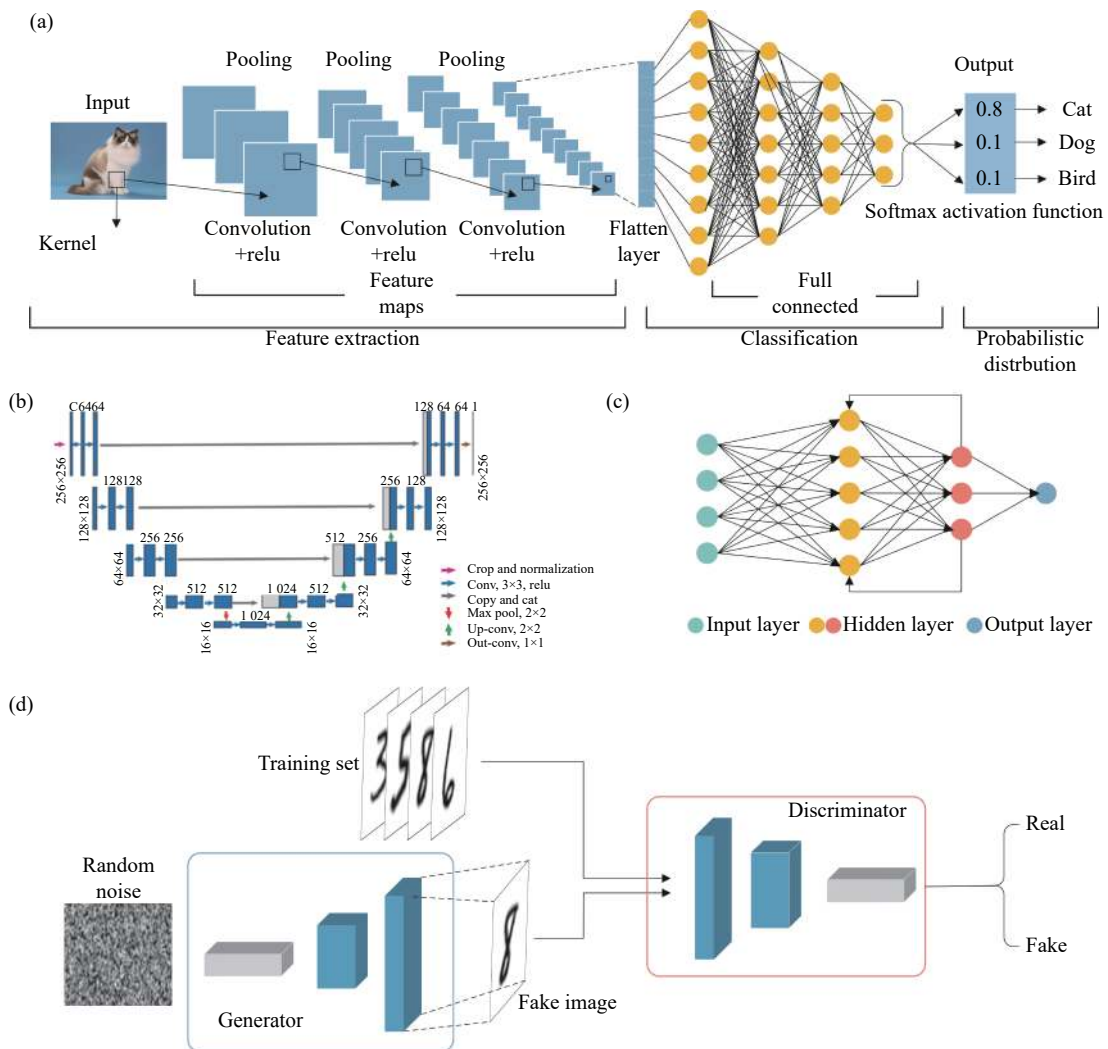


图 1 常用的神经网络结构图。(a) 卷积神经网络; (b) U-Net 结构<sup>[31]</sup>; (c) 循环神经网络; (d) 生成对抗网络<sup>[58]</sup>

Fig.1 Diagram of the structure of a commonly used neural network. (a) Convolutional neural network; (b) U-Net architecture<sup>[31]</sup>; (c) Recurrent neural network; (d) Generative adversarial networks<sup>[58]</sup>

和跳跃连接 (Skip Connection) 组成。原始图像在 U-Net 网络左侧完成下采样过程, 图像在网络右侧完成上采样过程。另外, U-Net 左右两侧中的层由 Skip Connection 连接, 通过浅层信息与深层信息融合来有效防止信息丢失, 解决了模型性能随构架深度增加而下降的问题。

## 1.2 循环神经网络

循环神经网络 (Recurrent Neural Network, RNN) 是一类具有短期记忆能力的神经网络, 它通过自反馈的神经元来处理任意长度的时序数据<sup>[60]</sup>。如图 1(c) 所示, RNN 由输入层 (Input Layer)、隐含层 (Hidden Layer) 以及输出层 (Output Layer) 组成, 用于处理和预测序列数据。例如, 在语音识别中, 将声音转换为字符序列。在 RNN 中, 隐含层之间的节点是有连接的, 这使输出值不仅取决于当前的输入值, 还与上一次隐含层的取值有关。在每一时刻, RNN 在处理输入序列的同时对之前的历史信息进行保存并应用于当前输出计算中。由于学习长期依赖, 为防止梯度消失/爆炸 (Gradient Vanishing/Exploding) 现象产生, 网络需要在更长时间内保留状态信息。为解决长期记忆问题, 可通过增加一个长期保持其状态的记忆单元实现。采用特殊隐式单元以长期保存输入的长短期记忆网络 (Long Short-Term Memory Networks, LSTM)<sup>[61]</sup> 被首先提出。与传统 RNN 相比, LSTM 能有效解决短期依赖瓶颈、缓解梯度消失/爆炸问题, 在语音识别、机器翻译等任务中具有优势<sup>[62-65]</sup>。

## 1.3 生成对抗网络

生成对抗网络 (Generative Adversarial Networks, GAN) 是一种无监督深度学习模型, 由 Goodfellow 等人于 2014 年提出<sup>[66]</sup>。受博弈论启发, GAN 由生成器 (Generator) 与判别器 (Discriminator) 两个模块所组成 (如图 1(d) 所示), 生成器生成接近真实的数据, 鉴别器判断生成数据的真实性, 两部分网络相互博弈以提高网络性能。与传统方法相比, GAN 的创新是将生成模型与判别模型相结合, 基于一种博弈与对抗关系从训练分布中生成数据, 在样本数据生成、图像生成以及提高图像分辨率等有着广泛的应用前景<sup>[67-70]</sup>。同时, GAN 还可应用于迁移学习和强化学习等领域,

不同学科领域之间的交叉融合极大程度上促进了人工智能的发展。

## 2 深度学习在荧光显微中的应用

荧光显微镜具有对样品损伤小、可对特定功能结构进行特异性成像等优点, 在生物医学等领域发挥着决定性作用。随着人工智能和深度学习技术的快速发展, 基于训练数据和物理模型的神经网络为全面提升荧光显微技术的性能提供了全新的途径。

以下将对深度学习在荧光显微成像中增强空间分辨率、提升成像和重建速度、提高荧光显微成像通量与成像质量及应用进行介绍。

### 2.1 增强空间分辨率

光学衍射极限于 1873 年由 Abbe 首次提出<sup>[71]</sup>, 也被称为阿贝极限, 其数值约为可见波波长的 1/2。阿贝衍射公式如下:

$$\sigma = \frac{\lambda}{2n \sin \alpha} = \frac{\lambda}{2NA} \quad (2)$$

式中:  $\lambda$  为波长;  $NA$  为数值孔径,  $NA = n \sin \alpha$ ,  $n$  为成像介质的折射率,  $\alpha$  为物镜孔径角的一半。当显微镜物方介质为空气时,  $NA_{\max} = 1$ ; 可见光范围内, 蓝紫光波长最短在 400 nm 左右, 据公式 (2) 可知, 传统光学显微镜分辨率约为 200 nm<sup>[72]</sup>, 不满足观测样品细节的需求。近年来, 一些超分辨率 (Super Resolution, SR) 显微技术<sup>[73]</sup> 应运而生, 如受激发射损耗荧光显微 (STED)<sup>[15]</sup>、结构光照明显微 (SIM)<sup>[16]</sup>、单分子定位显微 (STORM/PALM)<sup>[74]</sup> 等。但目前的这些技术还存在光路复杂、设备昂贵、需要记录大量原始数据等不易实现的问题, 深度学习技术的加入使这些问题迎刃而解<sup>[19, 74-76]</sup>。以下列举并简要介绍了目前几种用于增强荧光显微成像分辨率的网络模型。

加州大学 Wang 等<sup>[25]</sup> 于 2019 年提出了基于 GAN 的深度网络框架, 在无需预先了解样品的成像过程 (数值建模与点扩散函数 (Point Spread Function, PSF) 估计) 的情况下, 将在低倍物镜、共聚焦显微镜以及全内反射荧光显微镜下所获取的低分辨率 (Low Resolution, LR) 图像转换为高分辨 (High Resolution, HR) 图像, 并具有大视场 (Field of View, FOV)、高信

噪比 (Signal-to-Noise Ratio, SNR) 与高自由度等特点。作者采用了一个 U-Net 神经网络 (图 2(a)), 完成从 LR 图像到 HR 图像的转换。该网络选用 2625 对 LR-HR 数据集 (分别采用  $10\times/0.4$  NA 物镜与  $20\times/0.75$  NA 物镜拍摄获得) 作为训练集, 使用未在训练集中出现过

的 LR 图像作为测试集。为验证该网络性能, 首先对牛肺动脉内皮细胞 (BPAEC) 结构进行成像, 将使用低数值孔径 ( $10\times/0.4$  NA) 物镜所获取的宽场图像转换为分辨率增强图像, 如图 2(b1)~ (b3) 所示。然后, 使用基于 F-actin 数据集训练好的模型对荧光小球进行成

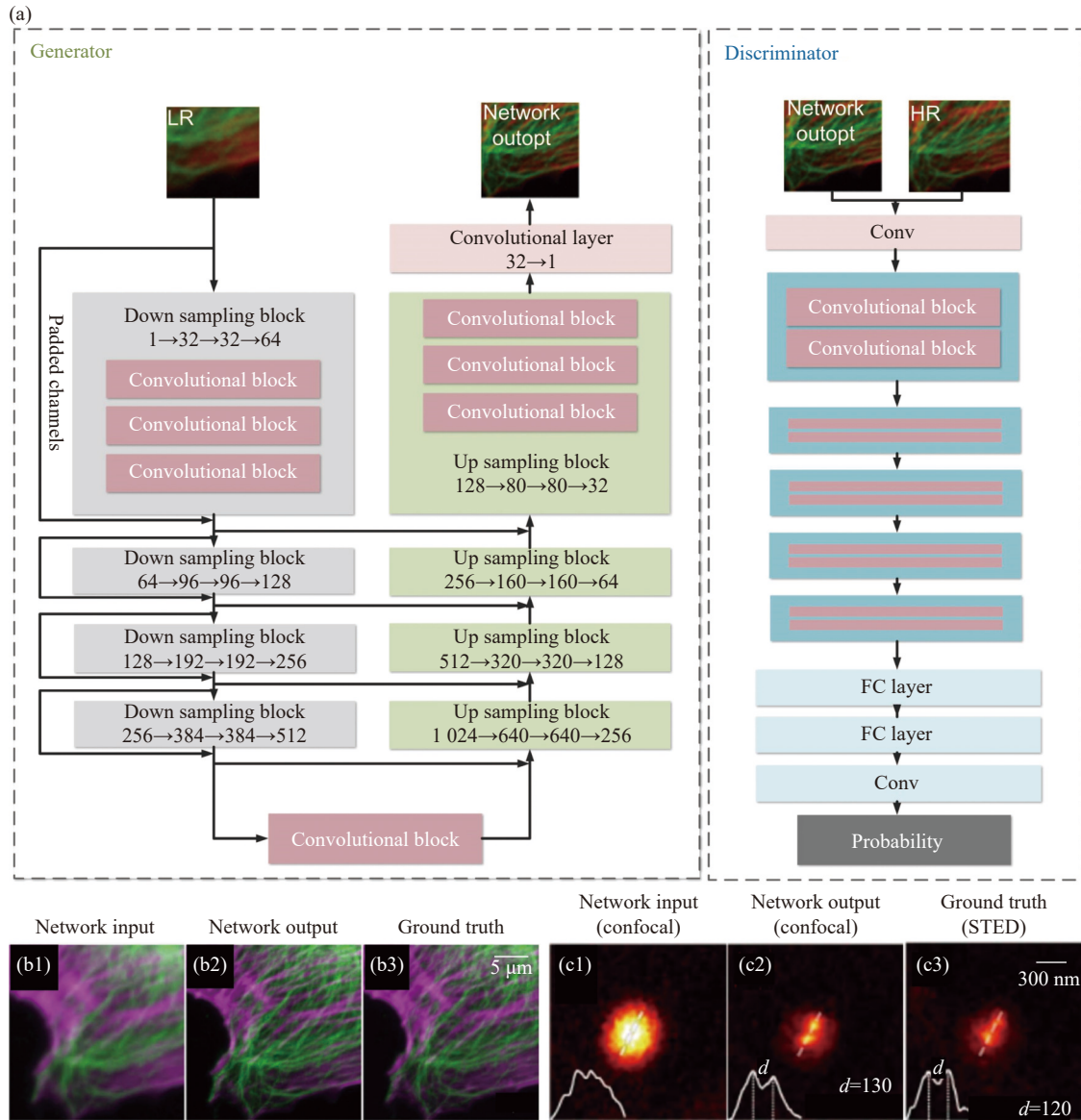


图 2 深度学习网络提高宽场和共聚焦显微镜的成像空间分辨率<sup>[25]</sup>。(a) 用于图像超分辨率的生成对抗网络 (GAN) 构架; (b1)~(b3) 网络对牛肺动脉内皮细胞 (BPAECs) 的成像结果, (b1) 使用  $10\times/0.4$  NA 物镜采集的网络输入图像, (b2) 网络输出图像, (b3) 使用  $20\times/0.75$  NA 物镜采集的图像; (c1) 荧光小球的共聚焦图像; (c2) 利用网络获得的跨模态图像; (c3) STED 所获得的超分辨图像

Fig.2 Deep learning network for spatial resolution enhancement of wide-field imaging and confocal microscopy<sup>[25]</sup>. (a) Architecture of generative adversarial network (GAN) used for image super-resolution; (b1)-(b3) Imaging results of bovine pulmonary artery endothelial cells (BPAECs), (b1) Network input image acquired with a  $10\times/0.4$  NA objective lens, (b2) Network output image, (b3) Image acquired with a  $20\times/0.75$  NA objective lens; (c1) Confocal images of fluorescent spheres; (c2) Cross-modality images obtained using the network; (c3) Super-resolution images obtained by STED

像,通过计算其半高全宽测定了该方法的空间分辨率的增强倍率,并验证了模型的通用性。另外,文中还实现了跨模态成像,将共聚焦显微镜图像转换为 STED 显微图像,使得空间分辨率得到显著提升。作者利用荧光小球作为样品对该方法进行了测试,在作为网络输入的共聚焦图像中(图 2(c1)),间距为 130 nm 的两个小球无法分辨。经过网络跨模态转换的共聚焦图像与 STED 超分辨图像(图 2(c2)和(c3))可以分辨这两个荧光小球。该深度网络框架的提出是超分辨率成像领域向前迈出的重要一步。

同年 Zhang 等<sup>[26]</sup>也采用 GAN 实现从 LR 图像到 HR 图像的转换。该网络将利用宽场显微镜和光片显微镜单次采集的 LR 图像重建为大视场 HR 图像,实现了对人类病理切片、荧光标记的成纤维细胞以及转基因小鼠大脑深层组织的高分辨成像。2020 年,华中科技大学的 Zhou 等<sup>[27]</sup>提出基于 3D 深度学习的双生成对抗网络(Dual GAN),无需精准图像配准,就能实现 LR 图像转换为 HR 图像。该网络对光片显微镜采集的神经元图像进行验证,研究结果证明从采集的粗配准  $5\times/0.16$  NA 体积图像中能够恢复  $20\times/1.0$  NA 体积图像,即将 LR 体积图像中的一个体素所包含的信息成功分配到预测的 HR 体积图像的  $3\times 3\times 3$  体素中。另外, Dual GAN 预测的 HR 图像的 PSNR 和 SSIM 值分别为 27.77 和 0.83,其定量衡量指标数值均优于上采样方法。2020 年, Wang 课题组<sup>[28]</sup>提出了一种深度对抗网络(Deep Adversarial Network, DAN),以一个 STED 图像作为输入,输出一个对应更高分辨率的 STED 图像。通过模拟数据测试表明,该方法可将 STED 的空间分辨率由 60 nm 提高到 30 nm,标签和输出结果之间的结构相似性指数(Structural Similarity, SSIM)达到 0.98 左右,成功实现超分辨率与抑制噪声的优异性能。2021 年, Christensen 等<sup>[29]</sup>提出 ML-SIM 重建方法,其使用端到端深度残差神经网络。该模型通过从 DIV2K(1000 张高分辨图像)图像数据集获取高质量的图像,然后计算得到模拟 SIM 原始图像作为训练数据。ML-SIM 网络将一组原始 SIM 图像(9 张)重建成一个单一的超分辨率图像,达到了与传统重建算法(OpenSIM、FairSIM、CC-SIM)相当的分辨率。同

年, Shah 等<sup>[30]</sup>提出基于深度学习的 SR-SIM 图像重建方法 RED-fairSIM (Residual Encoder-Decoder fairSIM),将 15 张  $512\times 512$  的原始 SIM 图像作为输入,重建后获得  $1024\times 1024$  的 SR-SIM 图像,在低 SNR 条件下获得超分辨率图像,并有效地抑制了噪声。

## 2.2 提升图像采集与重建速度

成像速度是衡量光学显微镜性能指标的另一个重要指标。SR 荧光显微技术通常需要大量、高质量的原始数据以及复杂的重建算法来获得高分辨的图像<sup>[77-78]</sup>,这导致了成像速度减慢,不利于活体动态样品的观测。以下对目前具有代表性的用于提升荧光显微成像时间分辨率的网络模型进行列举。

结构光照明显微(SIM)技术,通过记录样品在不同方向和相移量条纹结构光照明下的荧光图像,可以重构出空间分辨率超越光学衍射极限的荧光图像。SIM 的超分辨图像重建需要采用方向互成  $120^\circ$  三组条纹结构光照明样品,并且每组条纹结构光通过平移,分别进行三次(五次)相移操作,故限制了其成像速度。为解决这个问题, Jin 等<sup>[31]</sup>于 2020 年提出基于 U-Net 网络的 DL-SIM 方法,将超分辨率 SIM 所需的原始图像数量减少 5 倍,以提高 SIM 的时间分辨率。Jin 等依次使用 15 张/3 张 SIM 原始图像与相应的 15 张常规 SIM 重建结果分别作为输入和真值(Ground Truth, GT),训练得到 U-Net-SIM15/U-Net-SIM3,如图 3(a)所示。经过训练的 U-Net 对 4 种不同的亚细胞结构(微管、粘附、线粒体和 F-肌动蛋白)图像进行恢复重建,如图 3(b)所示。其中,微管、粘附、线粒体和 F-肌动蛋白图像中沿虚线的强度分布如图 3(c)所示。从图 3(c)中可见 U-Net-SIM3 所恢复的图像与常规 SIM 重建和 U-Net-SIM15 所恢复出的图像分辨率相当,证明 DL-SIM 方法可以以较少的原始图像数量来实现快速超分辨 SIM 成像。

随后,研究人员使用不同的网络架构通过减少原始图像数量加快 SIM 重建速度。深圳大学袁小聪教授课题组<sup>[32]</sup>于 2020 年提出 CycleGAN 深度神经网络。该网络无需任何样本结构的先验知识,将 SIM 超分辨图像重建所需的原始图像帧数由 9 帧减少至 3 帧,实现快速 SIM 超分辨成像。

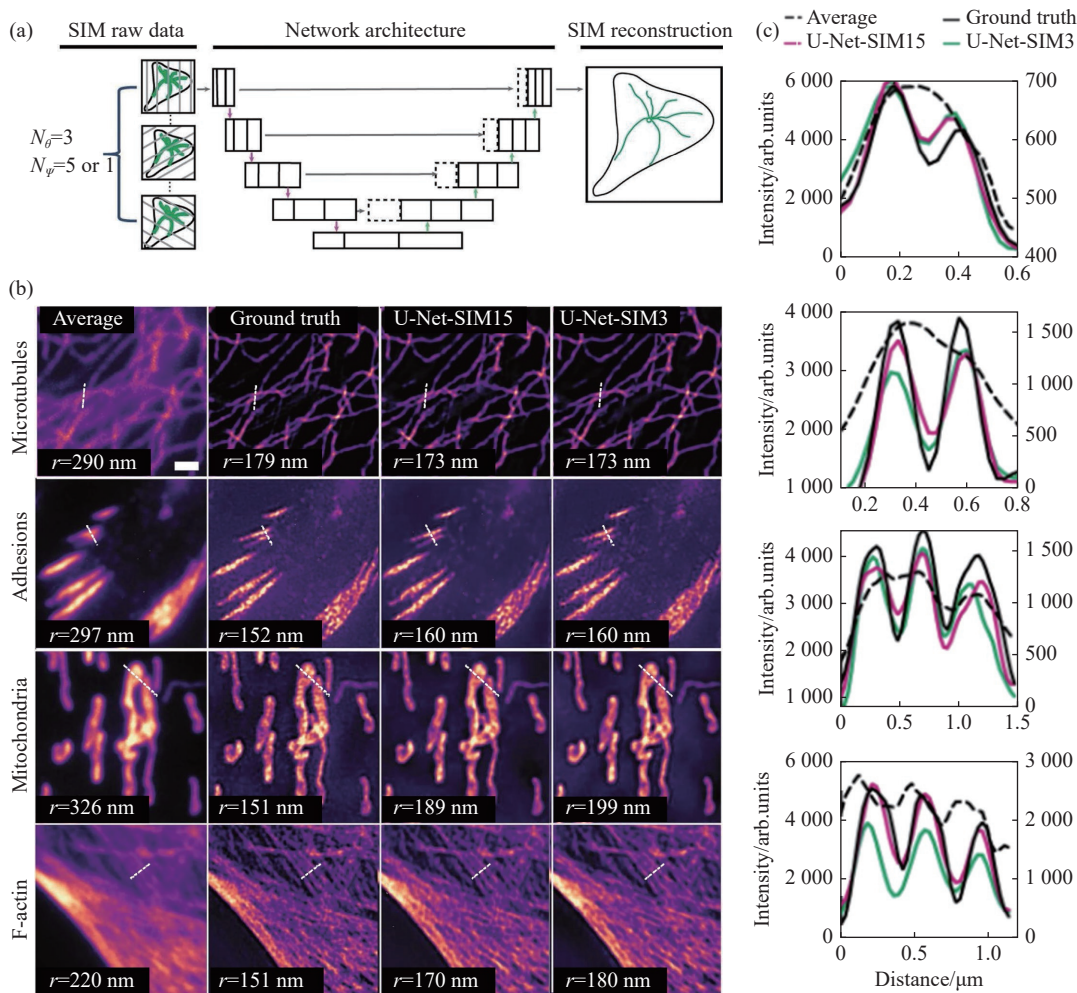


图 3 使用 U-Net 的超分辨率 SIM 成像<sup>[31]</sup>。(a) 使用 15 或 3 张 SIM 原始数据图像作为输入, 并将 15 张图像中相应的 SIM 重建作为训练 U-Net 的标签图像; (b) 不同亚细胞结构的重建结果; (c) 图 (b) 每张图像中沿虚线的强度分布。图中, Average 显示在右侧 y 轴上, 其他共享在左侧 y 轴

Fig.3 Super-resolution SIM imaging with U-Net<sup>[31]</sup>. (a) Fifteen or three SIM raw data images were used as input and the corresponding SIM reconstructions from 15 images were used as the ground truth to train the U-Net; (b) Reconstruction results for different subcellular structures; (c) Intensity distribution of Figure (b) along the dashed line in each image. In the plot, the average is shown on the right y-axis and all others share is the left y-axis

单分子定位显微技术 (SMLM) 也是目前另一种常用的超分辨荧光显微技术。该技术在每一帧图像中仅激活稀疏分布的几个荧光分子, 然后通过高斯拟合来精确定位这些荧光分子的位置。经过持续多次循环后, 细胞内大多数荧光分子被精确定位。将这些分子的图像合成到一张图上, 最后可以得到比传统光学显微镜分辨率高 10 倍的显微图像。传统 SMLM 需要记录上万张原始图像才能重构出单张超分辨显微图像, 成像和重构过程均耗时较长。2018 年, Boyd 等<sup>[33]</sup> 使用 DeepLoco 方法将 SMLM 中活跃的荧光团

位置生成简短列表, 代替了原本需要对每一帧分析的复杂步骤, 然后使用 CNN 改善传统算法复杂、耗时的问题, 完成对帧中每个活跃荧光分子的快速 3D 定位。2018 年, Shechtman 课题组<sup>[34]</sup> 将 CNN 与 SMLM 相结合, 提出了一种快速、精确和无参数的算法, 避免 SMLM 中发光分子稀疏性约束导致定位时间较长的问题, 实现 SMLM 的快速超分辨率图像重建。由于 SMLM 的标准分析算法需要发光的荧光分子稀疏分布 (在空间分布上不能重叠), 限制了荧光标记密度和成像速度。为解决此问题, Speiser 等<sup>[35]</sup> 于 2021

年开发了 DECODE(Deep Context Dependent) 模型,可在发光样品分子局部重叠的情况下实现单分子定位,将成像时间减少了一个数量级,实现快速活细胞 SMLM。

此外,2018年,Weigert等<sup>[36]</sup>提出了基于深度学习的内容感知图像恢复(CARE)网络。CARE模型使用U-Net构架,在轴向切片数量减少为原来数量10%的情况下获得各向同性的空间分辨率。该方法有效解决了光片荧光显微镜(LSFM)轴向采样率低的问题,并使其能在更短的曝光时间和更低的光强度情况下工作。2021年,华中科技大学费鹏教授课题组<sup>[37]</sup>提出视道深度神经网络(View-Channel-Depth Neural Network, VCD-Net),以解决光场显微镜(LFM)在捕获毫秒量级生命过程中空间分辨率低且不均匀的限制。该团队将VCD网络和LFM相结合,实现了对斑马鱼幼虫心脏跳动的实时成像(图4(a),(b)),展示了视道深度神经网络光场荧光显微镜(VCD-LFM)在瞬态生物动态过程成像中的巨大潜力。

力。荧光显微中成像通量往往受到物镜(分辨率与成像视场)、相机的像素个数等的限制,不能满足人们的需求。尤其在三维(3D)荧光显微成像<sup>[79-80]</sup>中,高的成像通量变得十分迫切。共聚焦显微镜通过3D扫描聚焦光斑,并记录焦点所激发的荧光,可以获得样品的3D图像。共聚焦显微镜中的针孔可以滤除样品上的离焦信息,从而获得高的层析能力。然而传统扫描方法耗时过长,限制了成像速度与通量。常用的改善方法有多焦点成像<sup>[81-82]</sup>、光场显微镜<sup>[83]</sup>和压缩感知<sup>[84]</sup>等。此外,结构光照明和光片照明也可以用来对样品进行3D层析成像。然而,以上这些方法在重建图像上相对耗时,同时需要用户定义准确模型,普适性差。近年来,研究者们提出了各种深度学习模型用于提高荧光显微成像通量这一性能指标。

2021年,美国加州大学Huang等<sup>[40]</sup>设计了一个基于深度学习的体积图像推理框架——Recurrent-MZ(图5),改善了3D成像中对样本体积连续扫描限制成像通量的问题。该方法采用基于编码器-解码器卷积循环网络(CRNN),通过一系列级联的编码器-解码器对,更好地利用且合并来自不同轴向位置上的不同尺度的特征图。作者使用63×/1.4 NA物镜采集3张轴向间隔为6 μm的线虫样品图像(1024 pixel×1024 pixel)作为网络输入,同时采集轴向间隔为0.2 μm的91张图像作为GT,结果表明网络预测结果与GT图像质量相当,标准均方根误差(Normalized Root Mean Square Error, NRMSE)均值为4、峰值信噪比(Peak Signal to Noise Ratio, PSNR)均值为34。另外,由Recurrent-MZ推断的单个纳米小球的半峰全宽(FWHM)分布(平均FWHM=0.4401 μm)与GT结果(平均FWHM=0.4428 μm)非常匹配,证明了重建的3D图像具有高保真度。相较于Deep-Z框架<sup>[41]</sup>将单图像作为输入,Recurrent-MZ还突破了对于有效体积空间带宽的限制。该方法将传统3D成像所需要的轴向切片数量减少了30倍,成像速度和吞吐量得到显著提升,并且拓展了成像系统景深。

2019年,Ozcan教授课题组<sup>[41]</sup>提出基于GAN的荧光显微数字图像重聚焦框架——Deep-Z。该方法从单个2D宽场图像对荧光样品进行3D重建,最终

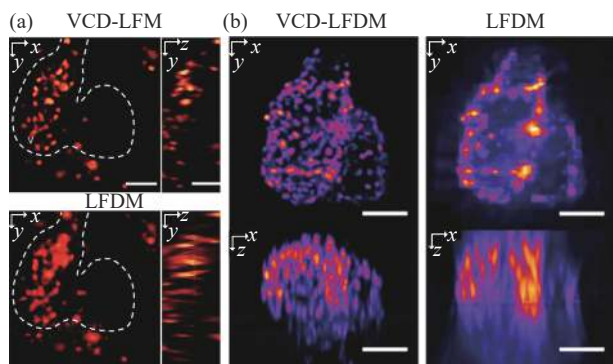


图4 深度学习提高光场显微镜的成像分辨率<sup>[37]</sup>。(a)利用VCD-LFM和光场反卷积显微镜(LFDm)对斑马鱼心脏跳动中的红细胞成像;(b)利用VCD-LFM和LFDm对斑马鱼心脏跳动中的心肌细胞核成像

Fig.4 Deep learning for resolution enhancement of light-field microscopy<sup>[37]</sup>. (a) Imaging of RBCs in the beating zebrafish heart using VCD-LFM and LFDm; (b) Imaging of cardiomyocyte nuclei in the beating zebrafish heart using VCD-LFM and LFDm

### 2.3 提高荧光显微成像通量

成像通量是光学显微另外一个重要特性,它描述了成像系统在单位时间内能获得光学信息总量的能



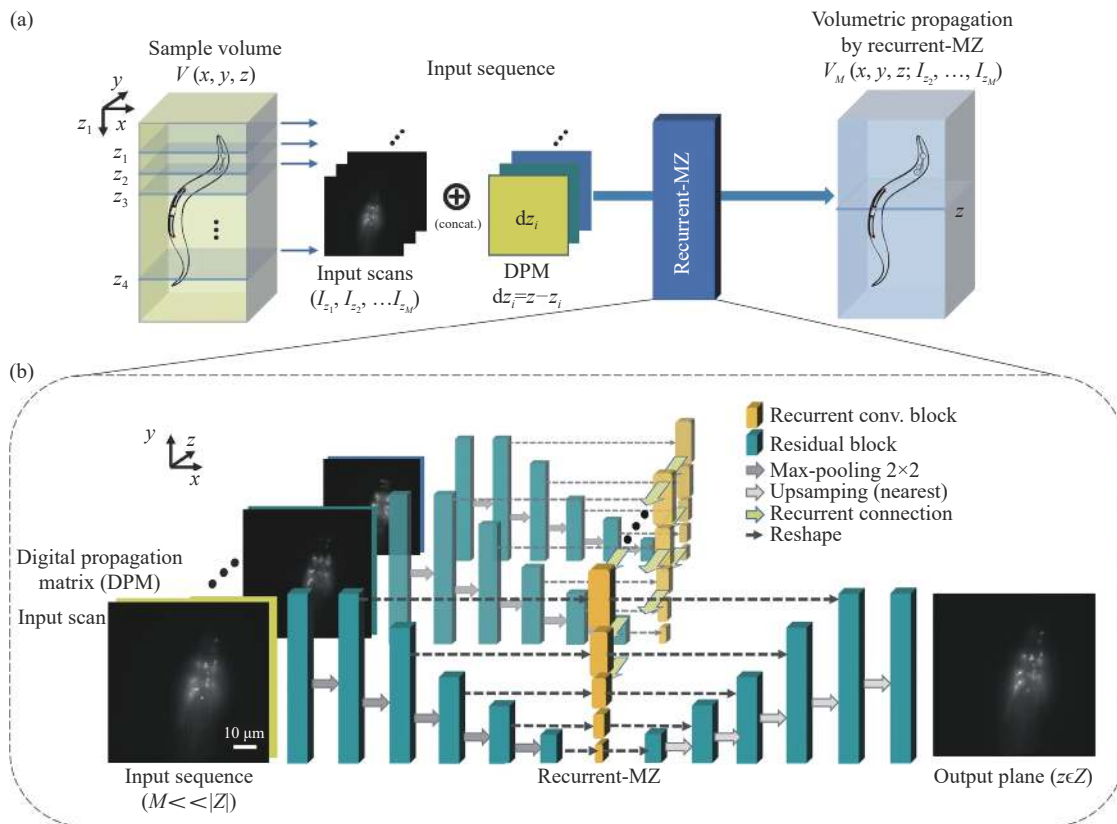


图 5 通过 Recurrent-MZ 的体积成像<sup>[40]</sup>。(a) Recurrent-MZ 体积成像框架。 $M$  是输入扫描 (2D 图像) 的数量, 每个输入扫描与其对应的 DPM(数字传播矩阵) 配对; (b) Recurrent-MZ 网络结构

Fig.5 Volumetric imaging through Recurrent-MZ<sup>[40]</sup>. (a) Recurrent-MZ volumetric imaging framework.  $M$  is the number of input scans (2D images), and each input scan is paired with its corresponding DPM (Digital Propagation Matrix); (b) Recurrent-MZ network structure

实现在  $\pm 10 \mu\text{m}$  的轴向范围内重聚焦, 将自然景深 (Depth of Field, DOF) 扩展约 20 倍, 克服传统技术成像速度和吞吐量的限制, 并减少了对样品的曝光量与光损伤。2020 年, 华中科技大学的 Ning 等<sup>[42]</sup> 提出了一种基于深度学习的荧光显微光学切片断层扫描 (DL-fMOST) 的成像方法, 首次使用深度学习进行全脑成像。其中, 使用 CNN 实现对宽场图像的实时光学切片, 提高了全脑成像的空间分辨率与通量。作者以  $0.32 \mu\text{m} \times 0.32 \mu\text{m} \times 2 \mu\text{m}$  分辨率 (体像素) 完成了对单个小鼠大脑的数据采集, 采集时间由 3 天缩短为 1.5 天, 该数据集大约包含 28 万个图像光学切片 (z-stack)。与广域大体积断层扫描 (WVT) 系统相比, DL-fMOST 将每个 z-stack 的平均时间由 780 ms 缩短为 262 ms, 总成像时间由 60 h 缩短为 20 h<sup>[85]</sup>, 大大提高了成像吞吐量。2020 年, 中国科学院西安光学精密机械研究所姚保利课题组<sup>[43]</sup> 提出了一个用于转盘式

共聚焦显微镜 (SDCM) 三维成像的深度学习方法——SRRF-Deep, 通过减少扫描次数与重建时间, 显著提高了利用超分辨率径向波动 (Super-Resolution Radial Fluctuations, SRRF) 方法成像的吞吐量, 同时减轻了光漂白与光毒性的影响。SRRF-Deep 将所需的扫描次数减少至 SRRF 的 1/100, 同时将  $1024 \text{ pixel} \times 1024 \text{ pixel}$  大小的切片的重建时间由 3 s 减少至约 0.1 s, 将重建速度提升了约 30 倍。2020 年, 华中科技大学费鹏教授课题组<sup>[44]</sup> 提出了一种双级互反馈 (DSP-Net) 深度学习 (图 6), 通过网络端到端非迭代推理突破 3D 荧光显微镜通量的极限。该方法可在 6 min 内完成小鼠大脑的 3D 成像, 成像速度比传统成像方法提升了两个数量级; 在  $3300 \mu\text{m} \times 830 \mu\text{m} \times 55 \mu\text{m}$  大视场下, 以 30 Hz 的体积速率对秀丽隐杆线虫活动进行追踪, 提高了捕获瞬时生物过程的时间分辨率, 有效改善了 LSFM 光学吞吐量限制。

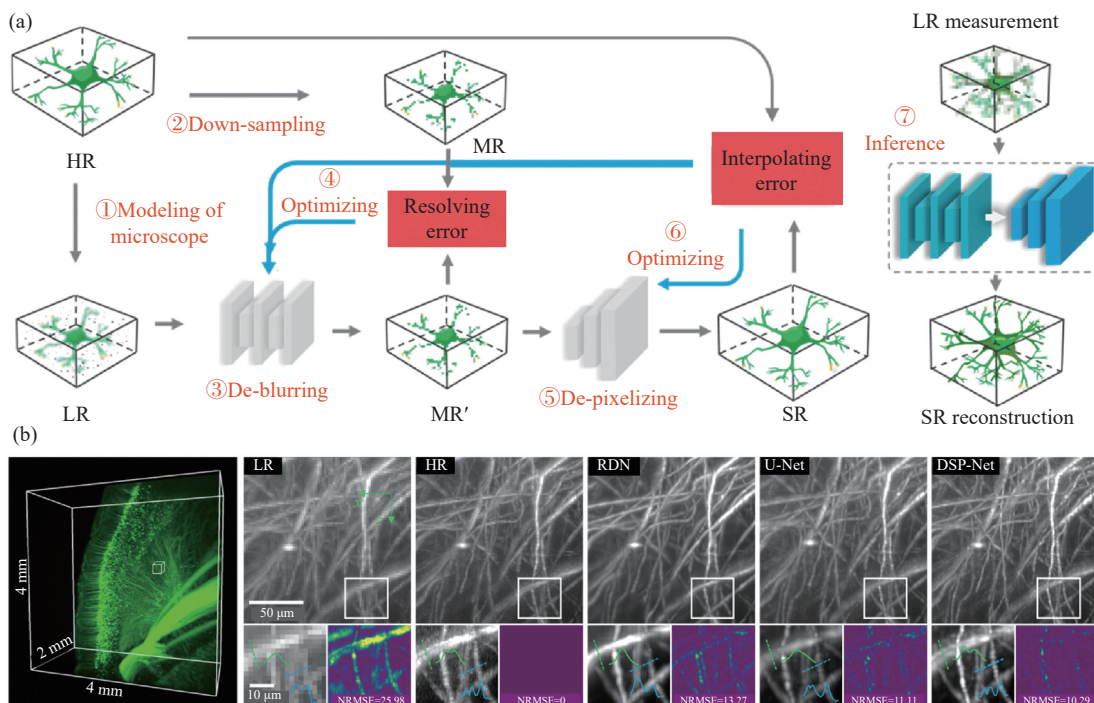


图 6 基于双级互反馈神经网络 (DSP-Net) 的小鼠大脑 3D 成像<sup>[44]</sup>。(a) DSP-Net 的工作流程图; (b) DSP-Net 结合 LSFM 对小鼠大脑中荧光标记神经元图像的处理结果

Fig.6 DSP-Net for 3D imaging of mouse brain<sup>[44]</sup>. (a) Work flow of DSP-Net; (b) DSP-Net-enabled LSFM results of the fluorescence-labelled neurons in the mouse brain

### 2.4 改善荧光显微成像质量

相比于透射式或反射式明场显微镜, 荧光显微通过收集样品上的荧光信号来获得样品的图像, 所采集图像的 SNR 较低。为此, 早期研究者们采用压缩感知<sup>[86]</sup>、图像自相似性<sup>[13]</sup>和小波变换后的阈值处理<sup>[87]</sup>等方法来改善荧光图像的质量。由于所建立模型精准性的限制, 这些方法往往达不到理想的效果, 并且重建效率较低, 因此阻碍其进一步发展。随后, 学者们利用各种深度学习模型来降低噪声干扰、减少伪影等。以下列举出了几种性能优良的算法以求解像质下降这一逆问题<sup>[88-90]</sup>。

首先, 噪声是影响荧光图像质量下降的主要因素, 抑制噪声是改善荧光显微成像质量的首要手段。2019 年, Zhang 等<sup>[91]</sup>构建一种用于泊松-高斯去噪的数据集——荧光显微镜去噪 (Fluorescence Microscopy Denoising, FMD) 数据集, 如图 7 所示。该数据集由 12000 张真实荧光显微图像组成, 主要包含了共聚焦、双光子、宽场显微镜和具有代表性生物样本如细胞、斑马鱼和小鼠脑组织。该方法使用图像平均来获取 60000 张不同噪声水平的图像和对应的标签。该

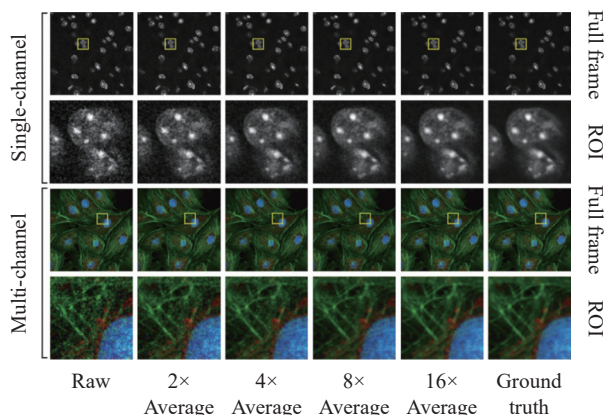


图 7 具有不同噪声水平的荧光显微图像数据集。用双光子显微镜在固定的小鼠脑组织上获取单通道 (灰色) 图像。用双光子显微镜在固定的 BPAE 细胞上获得多通道 (彩色) 荧光图像。真实图像 (GT) 是通过对 50 张有噪声的原始图像进行平均估计得到的<sup>[91]</sup>

Fig.7 Fluorescence microscopy image datasets with different noise levels. The single-channel (gray) images are acquired with two-photon microscopy on fixed mouse brain tissues. The multi-channel (color) images are obtained with two-photon microscopy on fixed BPAE cells. The ground truth images are estimated by averaging 50 noisy raw images<sup>[91]</sup>

数据集的建立极大地填补了该领域在原始数据上的空白,一定程度上推进了深度学习与荧光显微镜的结合。此外,上述数据集中的图像为 8 位,存在图像质量有限的问题。为了进一步推进,2021 年,科罗拉多大学的 Hagen 等<sup>[92]</sup>开发了针对同一图像同时提供高 SNR 和低 SNR 的数据集,该数据集涵盖了广泛的样本类型和成像模式,并且用基于 U-Net 架构的 Noise2Void(N2V) 的监督学习方法进行了验证。

其次,荧光显微通过收集样品上的荧光信号来获得样品的图像,每个像素上有限的光子数降低了图像的亮度和图像质量。CNN 作为一种重要的深度学习方法,为增强荧光显微图像提供了有效的支持<sup>[93]</sup>。2021 年,浙江大学的 Hu 等<sup>[45]</sup>提出了一种尺度循环环差矫正网络 (SRACNet) 来增强荧光显微图像,并将像差先验知识引入到深度学习中,无需附加额外的波前校正。SRACNet 网络 (如图 8(a) 所示) 包含参数共享的 3 个尺度循环训练: 64 pixel×64 pixel 的图像作为第 1 个尺度输入,经过网络训练后得到 128 pixel×128 pixel 的

图像,再将输出图像与预先缩放后的 128 pixel×128 pixel 图像进行拼接作为第 2 个尺度的输入图像,最后在第 3 个尺度调整为 256 pixel×256 pixel 的图像输出,完成网络的整个训练。另外,3 个尺度网络均基于 U-Net 结构,由卷积层、激活层以及修改后的残差块组成。其中,修改后的残差块 (图 8(b)) 在残差连接基础上加入 2 个具有不同卷积核大小的分支,为每个水平提供不同的感受野,同时并行化处理,提高网络训练速度。网络采用 RMSE 与 SSIM 相结合的损失函数来执行训练,在提高网络输出图像质量的同时,增强网络性能。通过共聚焦显微镜与光子显微镜采集的图像对 SRACNet 性能进行验证,结果表明:该网络可以有效增强图像细节,降低图像噪声,抑制伪影以及像差修正,且处理速度达到 40.70 ms/幅 (512 pixel×512 pixel),适用于实时生物应用。此外,该方法的 PSNR 和 SSIM 值分别为 29.29 和 0.83,而 U-Net 对应的值分别为 28.74 和 0.82,前者成像质量明显优于后者。

此外,厚样品中的光散射在很大程度上限制了荧

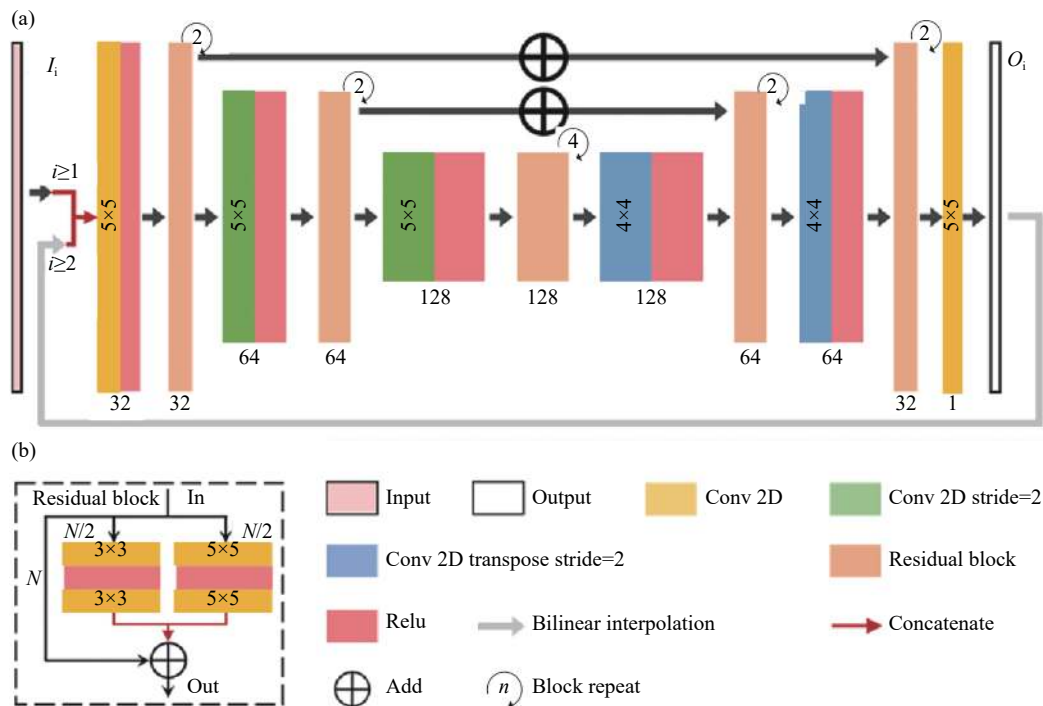


图 8 用于荧光显微图像增强的尺度循环环差矫正网络 (SRACNet)<sup>[45]</sup>。(a) SRACNet 网络的训练过程。圆圈箭头内的数字代表残差块的重复次数,每层下的数字代表通道数,下标  $i$  代表不同的尺度;(b) 修改后的残差块结构示意图。“ $N$ ”和“ $N/2$ ”代表通道数

Fig.8 Architecture of the SRACNet for fluorescence micrograph enhancement<sup>[45]</sup>. (a) Training processes of SRACNet architecture. The numbers inside the circle arrows represent the repeat times of the residual blocks. The numbers under each level represent the number of channels. The subscript  $i$  represents different scales; (b) Schematic diagram of the modified residual block architecture. “ $N$ ” and “ $N/2$ ” represent the number of channels

光显微镜 (如光片荧光显微镜) 对生物样本的 3D 成像。基于此, 2020 年, 华中科技大学的 Xiao 等<sup>[46]</sup> 提出了一种与光片荧光显微镜 (LSFM) 相结合的深度学习方法——ScatNet, 能够直接从散射图像中预测高

质量的图像, 扩展了 LSFM 的成像深度。通过对小鼠脑组织进行实验 (图 9), 其结果显示, 该方法提高了小鼠大脑中超出光子弹道区域的神元元的 SNR 和分辨率, 还将 LSFM 200  $\mu\text{m}$  传统成像深度提高至 300  $\mu\text{m}$ 。

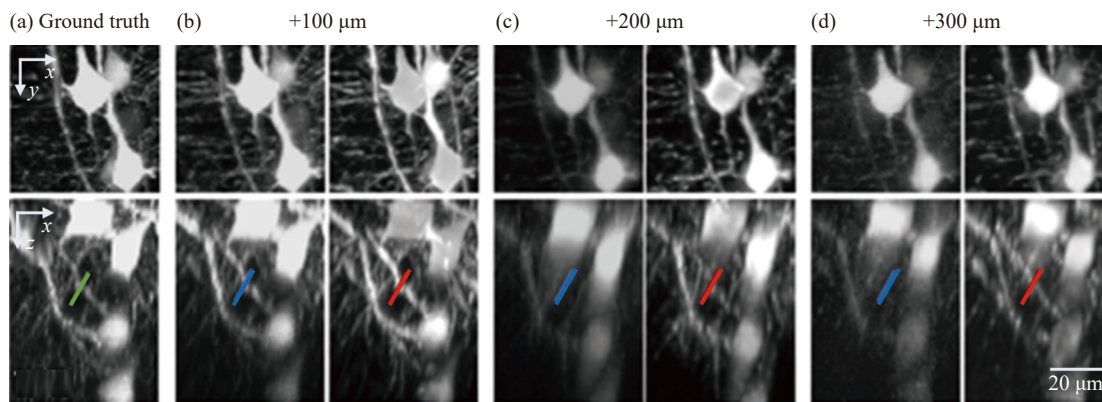


图 9 ScatNet 恢复的由双光子激发荧光显微镜成像的小鼠脑切片 (Thy1-YFP-M) 中神经元的信号<sup>[46]</sup>。(a) 深度为 0  $\mu\text{m}$  (ground-truth) 平面上脑切片的成像结果; (b) 左: 深度为 100  $\mu\text{m}$  平面上脑切片的成像结果, 右: 对应左图中的散射信号恢复结果; (c) 左: 深度为 200  $\mu\text{m}$  平面上脑切片的成像结果, 右: 对应左图中的散射信号恢复的结果; (d) 左: 深度为 300  $\mu\text{m}$  平面上脑切片的成像结果, 右: 对应左图中的散射信号恢复的结果

Fig.9 ScatNet restoration of neuronal signals in mouse brain slices (Thy1-YFP-M) imaged by TPfEM<sup>[46]</sup>. (a) The raw imaging results through 0  $\mu\text{m}$  (ground-truth) brain slices; (b) Left: the raw imaging results through 100  $\mu\text{m}$  brain slices, Right: the results restored from the scattered signals in Left; (c) Left: the raw imaging results through 200  $\mu\text{m}$  brain slices, Right: the results restored from the scattered signals in Left; (d) Left: the raw imaging results through 300  $\mu\text{m}$  brain slices, Right: the results restored from the scattered signals in Left

## 2.5 其他应用

深度学习除了提升荧光显微空间/时间分辨率、成像通量和质量外, 还在目标检测、模式识别及成像质量定量评价等领域拥有巨大潜力。

在利用荧光显微镜观察生物样本前, 首先需要对本样本进行荧光标记, 通常, 一个样本需要多个荧光标记, 这不但增加了样品制备及荧光标记工作的复杂性, 甚至将对细胞造成致死影响, 同时由于光谱重叠限制会对观察细胞带来困难。2018 年, 谷歌公司<sup>[94]</sup> 提出深度学习方法——In Silico Labeling (ISL), 样品无需进行荧光染色, 运用 ISL 模型就可从未标记的固定或活体样本的透射光的 z-stack 图像中预测多种荧光标记, 如图 10(a) 所示。其中, 训练集中包含了 3 种细胞 (多功能干细胞诱导的人类运动神经元、小鼠大脑皮层组织培养细胞和人类乳腺癌细胞) 样品图像。3 种细胞样品的不同结构的大量透射光 z-stack 图像以及样品进行不同的荧光标记之后, 在不同的显微镜

(明场、相差、微分干涉相差) 下获得像素匹配的荧光图像 (作为标签)。利用以上透射图像和荧光图像训练网络, 测试经过训练的网络, 可得到 z-stack 图像中每个像素的荧光标签, 准确预测细胞核的位置、细胞的活性状态、细胞类型及亚细胞结构类型。图 10(b)~(d) 展示了包含多种细胞的样本的透射光图像、使用 ISL 模型预测荧光标记及真实荧光标记图像。并且此网络还可进行迁移学习适应新的数据集, 加快网络学习收敛速度, 使用最少的训练集对新型荧光标记进行预测, 以实现细胞结构的可视化。

Ounkomol 等<sup>[95]</sup> 于 2018 年采用 U-Net 架构成功从透射光图像中预测三维 (3D) 荧光活细胞图像, 对  $64 \times 64 \times 32$  的 3D 图像预测仅需 1 s, 实现快速、无损伤的 3D 荧光图像的无标签预测。随后, Khater 等<sup>[96]</sup> 在 2019 年使用 CNN 可以从 SLM 数据中自动识别生物结构。实验利用多视图 CNN (MVCNN) 网络模型对前列腺癌细胞 (PC-3) 细胞膜上的小窝结构进行识

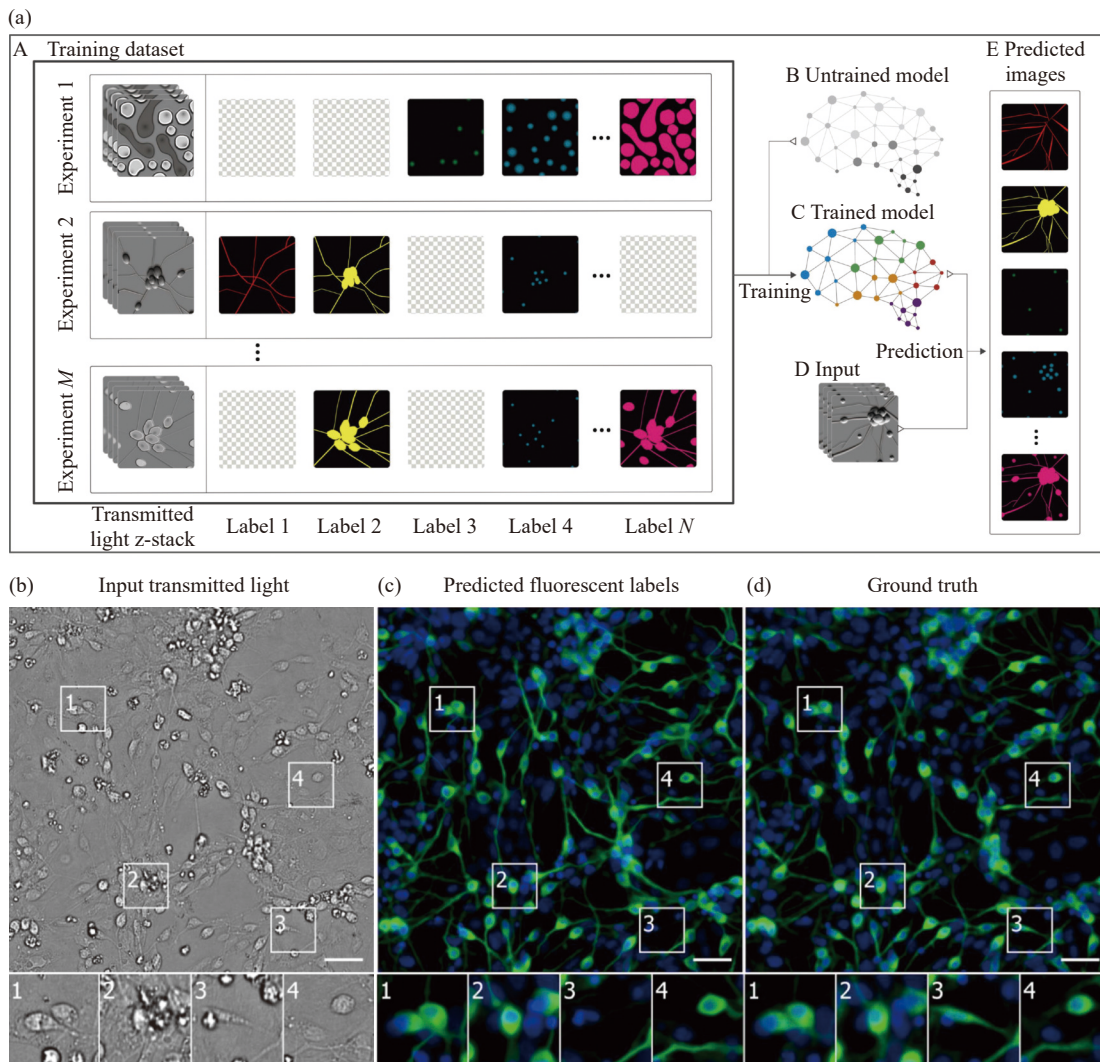


图 10 基于深度学习实现虚拟荧光标记<sup>[94]</sup>。(a) 从未标记图像中预测荧光标记的深度学习训练过程；(b)~(d) 相同细胞的透射光图像、荧光图像，以及用 ISL 模型预测的荧光标记

Fig.10 Deep learning-based implementation of virtual fluorescent labeling<sup>[94]</sup>. (a) Deep neural network train process for predicting fluorescent markers from unlabeled images; (b)~(d) Transmitted light images of the same cells, fluorescent images, and fluorescent markers predicted with the ISL model

别分类并在测试集中评估分类器性能,分类准确率达到到了 94%。2021 年,清华大学戴琼海课题组<sup>[97]</sup>提出无监督图像变换网络框架 (UTOM),不需要通过配对训练数据对网络进行训练,就能找到训练对之间的映射关系。如图 11 所示,提取人类大肠组织作为样本,完成从无标记的组织微阵列 (TMA) 自体荧光成像图像到苏木精和伊红 (H&E)染色图像的转换,实现与真实 H&E 染色图像相同的效果。作者为验证 UTOM 虚拟荧光标记性能,将神经元的相差图像转换为荧光图像,实现不同成像模式之间的高保真转换。这一过程

为医生在手术中快速诊断癌症、肿瘤等提供便捷。同年,Popescu 教授课题组<sup>[98]</sup>将深度学习与激光扫描定量相位成像 (QPI) 系统相结合,实现了从相位图像到荧光共聚焦图像的转换,采用无标记成像复制荧光共聚焦显微的优势:实现共聚焦水平的层析切片,并解决了光漂白和光毒性问题。针对非专业人士无法区分 STED 显微镜下活细胞中的微小结构,对图像质量评估困难的问题。2018 年,Robitaille 等<sup>[99]</sup>提出了一个基于神经网络对 STED 图像质量的定量评价系统,网络输入需要评价的图像,输出对应图像的质量

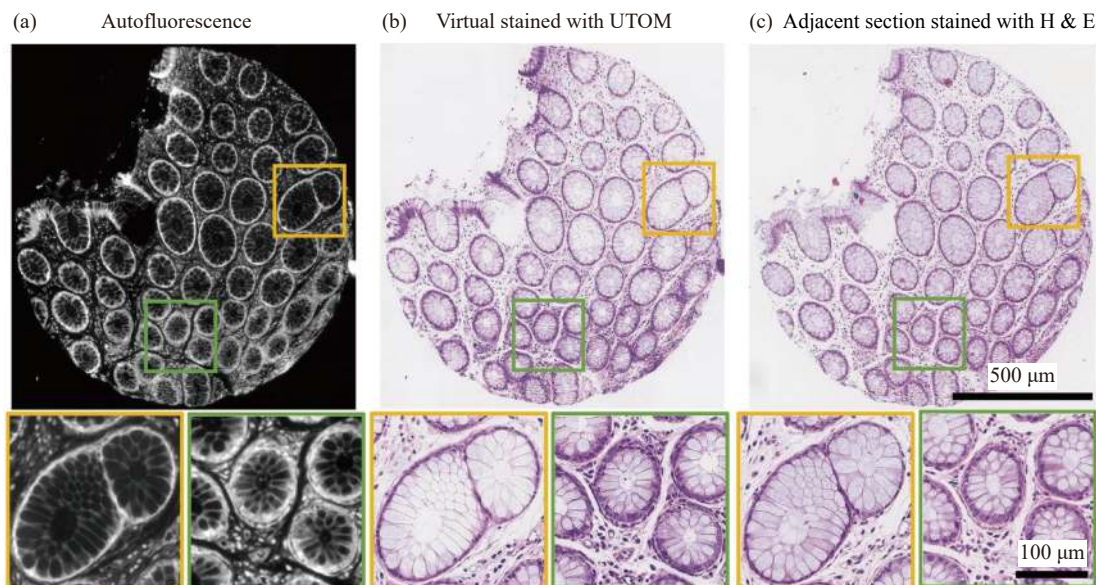


图 11 无监督图像变换网络框架的虚拟荧光成像<sup>[97]</sup>。(a) 无标记组织核心的自体荧光图像; (b) UTOM 预测的 H&E 染色; (c) 用于比较的相应相邻 H&E 染色部分的明场图像

Fig.11 Unsupervised image transformation network framework for virtual fluorescence imaging<sup>[97]</sup>. (a) Autofluorescence image of unlabeled tissue cores; (b) H&E staining predicted by UTOM; (c) Bright-field image of the corresponding adjacent H&E-stained section for comparison

量分数, 其结果是输入样本结构分辨率、SNR、光漂白及光氧化等的综合评价。可应用于预测生物中微小结构及自动化分析评估 STED 图像质量中。

### 3 总结与展望

荧光显微因其操作简便、成本低、对样品损伤小、可特异性成像等优点, 一直是生物医学研究的主流工具。借助荧光显微技术, 人类清晰地观察活体细胞的内部结构及其动态变化。近年来, 随着深度学习技术的不断完善, 并被逐渐应用于荧光显微, 为荧光显微技术发展提供了新思路。首先, 深度学习解决了传统荧光显微存在的光毒性、光漂白等方面的问题。其次, 深度学习还提高了传统荧光显微的空间分辨率、成像速度、吞吐量、信噪比等, 在保留荧光显微成像灵敏度高、特异性强的基础上实现更高质量的成像。

深度神经网络利用其强大的特征学习能力解决了荧光显微成像的不足, 但其也存在一定的自身局限性。例如, 在深度学习中, 往往需要先采集大量且高质量的数据集并对其进行标注, 然后用海量数据集对搭建的深度神经网络模型进行训练, 最后使用训练好的模型测试数据。其次, 深度学习的泛化能力有待于进一步增强。目前, 基于数据驱动的深度神经网络模型对标注数据有较大的依赖性。这就要求采集

的数据集不仅需要具备清晰、特征明显、符合逻辑的特点, 还需具有广泛性特征, 使其种类和特征能够覆盖将来潜在的测试样品。此外, 对于深度学习在计算成像领域中的应用来说, 其缺乏可靠合理的理论解释。神经网络就好比是一个“黑盒子”, 原本可以明确进行过程解析的成像过程, 目前在不同层级网络结构上又失去了可解释性。其理论支撑的薄弱, 使生物医学领域研究者无法完全信服它能更好地代替传统荧光显微技术。但近几年, 深度学习在荧光显微领域中的成果令人欣喜, 尤其是通过将数字全息显微真实物理模型融入到神经网络中, 可以有效减少所需的训练集数量并提供神经网络的通用性和可靠性。在之后的科学研究中, 人们能够将深度学习作为传统方法的分析辅助工具, 揭示生命科学的奥秘。

在大数据时代, 深度学习在荧光显微中尽管取得了一定的进展, 但仍然有很长的一段路要走。未来可通过 FPGA 编程将深度学习算法集成到显微镜硬件中去, 开发出基于深度学习的荧光显微镜, 成为更多领域成像和分析的核心技术。

#### 参考文献:

[1] Hamilton N. Quantification and its applications in fluorescent microscopy imaging [J]. *Traffic*, 2009, 10(8): 951-961.

- [2] Shi R, Jin C, Xie H, et al. Multi-plane, wide-field fluorescent microscopy for biodynamic imaging in vivo [J]. *Biomed Opt Express*, 2019, 10(12): 6625-6635.
- [3] Goodman J W, Lawrence R. Digital image formation from electronically detected holograms [J]. *Applied Physics Letters*, 1967, 11(3): 77-79.
- [4] Fan Y, Li J, Lu L, et al. Smart computational light microscopes (SCLMs) of smart computational imaging laboratory (SCILab) [J]. *PhotonIX*, 2021, 2(1): 1-64.
- [5] Gao P, Yuan C. Resolution enhancement of digital holographic microscopy via synthetic aperture: a review [J]. *Light: Advanced Manufacturing*, 2022, 3(1): 105-120.
- [6] Gao Peng, Wen Kai, Sun Xueying, et al. Review of resolution enhancement technologies in quantitative phase microscopy [J]. *Infrared and Laser Engineering*, 2019, 48(6): 0603007. (in Chinese)
- [7] Lichtman J W, Conchello J A. Fluorescence microscopy [J]. *Nature Methods*, 2005, 2(12): 910-919.
- [8] Conchello J A, Lichtman J W. Optical sectioning microscopy [J]. *Nature Methods*, 2005, 2(12): 920-931.
- [9] Murfin K E, Chaston J, Goodrich-Blair H. Visualizing bacteria in nematodes using fluorescent microscopy [J]. *Journal of Visualized Experiments*, 2012, 68: e4298.
- [10] Mickoleit M, Schmid B, Weber M, et al. High-resolution reconstruction of the beating zebrafish heart [J]. *Nature Methods*, 2014, 11(9): 919-922.
- [11] Giepmans B N, Adams S R, Ellisman M H, et al. The fluorescent toolbox for assessing protein location and function [J]. *Science*, 2006, 312(5771): 217-224.
- [12] Palmer A E, Tsien R Y. Measuring calcium signaling using genetically targetable fluorescent indicators [J]. *Nature Protocols*, 2006, 1(3): 1057-1065.
- [13] Boulanger J, Kervrann C, Bouthemy P, et al. Patch-based nonlocal functional for denoising fluorescence microscopy image sequences [J]. *IEEE Transactions on Medical Imaging*, 2009, 29(2): 442-454.
- [14] Betzig E, Patterson G H, Sougrat R, et al. Imaging intracellular fluorescent proteins at nanometer resolution [J]. *Science*, 2006, 313(5793): 1642-1645.
- [15] Hell S W, Wichmann J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy [J]. *Optics Letters*, 1994, 19(11): 780-782.
- [16] Gustafsson M G. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy [J]. *Journal of Microscopy*, 2000, 198(2): 82-87.
- [17] Gao P, Prunsche B, Zhou L, et al. Background suppression in fluorescence nanoscopy with stimulated emission double depletion [J]. *Nature Photonics*, 2017, 11(3): 163-169.
- [18] Klar T A, Jakobs S, Dyba M, et al. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission [J]. *Proc Natl Acad Sci USA*, 2000, 97(15): 8206-8210.
- [19] Shroff H, Galbraith C G, Galbraith J A, et al. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics [J]. *Nat Methods*, 2008, 5(5): 417-423.
- [20] Mertz J. Optical sectioning microscopy with planar or structured illumination [J]. *Nature Methods*, 2011, 8(10): 811-819.
- [21] Icha J, Weber M, Waters J C, et al. Phototoxicity in live fluorescence microscopy, and how to avoid it [J]. *Bioessays*, 2017, 39(8): 1700003.
- [22] Helmerich D A, Beliu G, Matikonda S S, et al. Photobleuing of organic dyes can cause artifacts in super-resolution microscopy [J]. *Nature Methods*, 2021, 18(3): 253-257.
- [23] Wang S-C. Artificial Neural Network [M]/Interdisciplinary Computing in Java Programming. Boston, MA: Springer, 2003: 81-100.
- [24] Shin H C, Roth H R, Gao M, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning [J]. *IEEE Trans Med Imaging*, 2016, 35(5): 1285-1298.
- [25] Wang H, Rivenson Y, Jin Y, et al. Deep learning enables cross-modality super-resolution in fluorescence microscopy [J]. *Nature Methods*, 2019, 16(1): 103-110.
- [26] Zhang H, Fang C, Xie X, et al. High-throughput, high-resolution deep learning microscopy based on registration-free generative adversarial network [J]. *Biomed Opt Express*, 2019, 10(3): 1044-1063.
- [27] Zhou H, Cai R, Quan T, et al. 3D high resolution generative deep-learning network for fluorescence microscopy imaging [J]. *Optics Letters*, 2020, 45(7): 1695-1698.
- [28] Li M Z, Shan H M, Pryshchep S, et al. Deep adversarial network for super stimulated emission depletion imaging [J]. *Journal of Nanophotonics*, 2020, 14(1): 016009.
- [29] Christensen C N, Ward E N, Lio P, et al. ML-SIM: Universal reconstruction of structured illumination microscopy images using transfer learning [J]. *Biomedical Optics Express*, 2021, 12(5): 2720-2733.
- [30] Shah Z H, Müller M, Wang T C, et al. Deep-learning based denoising and reconstruction of super-resolution structured illumination microscopy images [J]. *Photonics Research*, 2021, 9(5): B168-B181.

- [31] Jin L, Liu B, Zhao F, et al. Deep learning enables structured illumination microscopy with low light levels and enhanced speed [J]. *Nature Communications*, 2020, 11(1): 1934.
- [32] Ling C, Zhang C L, Wang M Q, et al. Fast structured illumination microscopy via deep learning [J]. *Photonics Research*, 2020, 8(8): 1350-1359.
- [33] Boyd N, Jonas E, Babcock H, et al. DeepLoco: Fast 3D localization microscopy using neural networks [Z/OL]. bioRxiv, (2018-02-26)[2022-08-01]. <https://doi.org/10.1101/267096>.
- [34] Nehme E, Weiss L E, Michaeli T, et al. Deep-STORM: super-resolution single-molecule microscopy by deep learning [J]. *Optica*, 2018, 5(4): 458-464.
- [35] 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.
- [36] Weigert M, Schmidt U, Boothe T, et al. Content-aware image restoration: pushing the limits of fluorescence microscopy [J]. *Nature Methods*, 2018, 15(12): 1090-1097.
- [37] Wang Z, Zhu L, Zhang H, et al. Real-time volumetric reconstruction of biological dynamics with light-field microscopy and deep learning [J]. *Nature Methods*, 2021, 18(5): 551-556.
- [38] Zhang X, Chen Y, Ning K, et al. Deep learning optical-sectioning method [J]. *Optics Express*, 2018, 26(23): 30762-30772.
- [39] Bai C, Liu C, Yu X H, et al. Imaging enhancement of light-sheet fluorescence microscopy via deep learning [J]. *IEEE Photonics Technology Letters*, 2019, 31(22): 1803-1806.
- [40] Huang L, Chen H, Luo Y, et al. Recurrent neural network-based volumetric fluorescence microscopy [J]. *Light Sci Appl*, 2021, 10(1): 62.
- [41] Wu Y C, Rivenson Y, Wang H D, et al. Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning [J]. *Nature Methods*, 2019, 16(12): 1323-1331.
- [42] Ning K, Zhang X, Gao X, et al. Deep-learning-based whole-brain imaging at single-neuron resolution [J]. *Biomedical Optics Express*, 2020, 11(7): 3567-3584.
- [43] Bai C, Yu X, Peng T, et al. 3D imaging restoration of spinning-disk confocal microscopy via deep learning [J]. *IEEE Photonics Technology Letters*, 2020, 32(18): 1131-1134.
- [44] Zhang H, Zhao Y, Fang C, et al. Exceeding the limits of 3D fluorescence microscopy using a dual-stage-processing network [J]. *Optica*, 2020, 7(11): 1627-1640.
- [45] Hu L, Hu S, Gong W, et al. Image enhancement for fluorescence microscopy based on deep learning with prior knowledge of aberration [J]. *Optics Letters*, 2021, 46(9): 2055-2058.
- [46] Xiao L, Fang C, Zhu L, et al. Deep learning-enabled efficient image restoration for 3D microscopy of turbid biological specimens [J]. *Optics Express*, 2020, 28(20): 30234-30247.
- [47] Lecun Y, Bengio Y, Hinton G. Deep learning [J]. *Nature*, 2015, 521(7553): 436-444.
- [48] McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity [J]. *The Bulletin of Mathematical Biophysics*, 1943, 5(4): 115-133.
- [49] Lecun Y. A theoretical framework for back-propagation[C]// Proceedings of the 1988 Connectionist Models Summer School, 1988: 21-28.
- [50] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks [J]. *Science*, 2006, 313(5786): 504-507.
- [51] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[C]//NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems, 2012, 1: 1097-1105.
- [52] Ongie G, Jalal A, Metzler C A, et al. Deep learning techniques for inverse problems in imaging [J]. *IEEE Journal on Selected Areas in Information Theory*, 2020, 1(1): 39-56.
- [53] Ghosh N, Bhattacharya K. Cube beam-splitter interferometer for phase shifting interferometry [J]. *Journal of Optics*, 2009, 38(4): 191-198.
- [54] Mccann M T, Jin K H, Unser M. Convolutional neural networks for inverse problems in imaging: A review [J]. *IEEE Signal Processing Magazine*, 2017, 34(6): 85-95.
- [55] O'shea K, Nash R. An introduction to convolutional neural networks [EB/OL]. (2015-11-26)[2022-08-01]. <https://arxiv.org/abs/1511.08458>.
- [56] Pang S, Du A, Orgun M A, et al. Beyond CNNs: exploiting further inherent symmetries in medical images for segmentation [EB/OL]. (2020-05-08)[2022-08-01]. <https://arxiv.org/abs/2005.03924>.
- [57] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing And Computer-assisted Intervention, 2015: 234-241.
- [58] Feizabadi M M, Shujjat A M, Shahid S, et al. Interactive latent interpolation on MNIST dataset [EB/OL]. (2020-10-15)[2022-08-01]. <https://arxiv.org/abs/2010.07581>.
- [59] Zhu Linlin, Han Lu, Du Hong, et al. Multi-active contour cell segmentation method based on U-Net network [J]. *Infrared and Laser Engineering*, 2020, 49(S1): 20200121. (in Chinese)



- [60] Medsker L R, Jain L. Recurrent neural networks [J]. *Design and Applications*, 2001, 5: 64-67.
- [61] Hochreiter S, Schmidhuber J. Long short-term memory [J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [62] Vinyals O, Toshev A, Bengio S, et al. Show and tell: A neural image caption generator[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, 2015: 3156-3164.
- [63] Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks [EB/OL]. (2014-09-10)[2022-08-01]. <https://arxiv.org/abs/1409.3215>.
- [64] Graves A. Generating sequences with recurrent neural networks [EB/OL]. (2013-08-04)[2022-08-01]. <https://arxiv.org/abs/1308.0850v5>.
- [65] Sajjad M, Kwon S. Clustering-based speech emotion recognition by incorporating learned features and deep BiLSTM [J]. *IEEE Access*, 2020, 8: 79861-79875.
- [66] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [EB/OL]. (2014-06-10)[2022-08-01]. <https://arxiv.org/abs/1406.2661>.
- [67] Isola P, Zhu J-Y, Zhou T, et al. Image-to-image translation with conditional adversarial networks[C]//Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition, 2017: 1125-1134.
- [68] Brock A, Donahue J, Simonyan K. Large scale GAN training for high fidelity natural image synthesis [EB/OL]. (2018-09-28)[2022-08-01]. <https://arxiv.org/abs/1809.11096v2>.
- [69] Cao J, Hou L, Yang M-H, et al. Remix: Towards image-to-image translation with limited data[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 15018-15027.
- [70] Wang X, Yu K, Wu S, et al. Esgan: Enhanced super-resolution generative adversarial networks[C]//Proceedings of the European Conference On Computer Vision (ECCV) Workshops, 2018.
- [71] Abbe E. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung [J]. *Archiv für Mikroskopische Anatomie*, 1873, 9(1): 413-418.
- [72] Pawley J. Handbook of Biological Confocal Microscopy [M]. New York: Springer Science & Business Media, 2006.
- [73] Ji Wei, Xu Tao, Liu Bei. Super-resolution fluorescent microscopy: A brief introduction to the Nobel Prize in Chemistry 2014 [J]. *Chinese Journal of Nature*, 2014, 36(6): 404-408. (in Chinese)
- [74] Rust M J, Bates M, Zhuang X J N M. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM) [J]. *Nature Methods*, 2006, 3(10): 793-796.
- [75] Heintzmann R, Huser T. Super-resolution structured illumination microscopy [J]. *Chemical Reviews*, 2017, 117(23): 13890-13908.
- [76] Tam J, Merino D. Stochastic optical reconstruction microscopy (STORM) in comparison with stimulated emission depletion (STED) and other imaging methods [J]. *J Neurochem*, 2015, 135(4): 643-658.
- [77] Huang B, Bates M, Zhuang X. Super-resolution fluorescence microscopy [J]. *Annu Rev Biochem*, 2009, 78: 993-1016.
- [78] Schermelleh L, Heintzmann R, Leonhardt H. A guide to super-resolution fluorescence microscopy [J]. *J Cell Biol*, 2010, 190(2): 165-175.
- [79] Nguyen J P, Shipley F B, Linder A N, et al. Whole-brain calcium imaging with cellular resolution in freely behaving *Caenorhabditis elegans* [J]. *Proc Natl Acad Sci USA*, 2016, 113(8): E1074-1081.
- [80] Juette M F, Gould T J, Lessard M D, et al. Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples [J]. *Nature Methods*, 2008, 5(6): 527-529.
- [81] Prabhat P, Ram S, Ward E S, et al. Simultaneous imaging of different focal planes in fluorescence microscopy for the study of cellular dynamics in three dimensions [J]. *IEEE Transactions on NanoBioscience*, 2004, 3(4): 237-242.
- [82] Johnson C, Exell J, Kuo J, et al. Continuous focal translation enhances rate of point-scan volumetric microscopy [J]. *Optics Express*, 2019, 27(25): 36241-36258.
- [83] Li H, Guo C, Kim-Holzappel D, et al. Fast, volumetric live-cell imaging using high-resolution light-field microscopy [J]. *Biomedical Optics Express*, 2019, 10(1): 29-49.
- [84] Pascucci M, Ganesan S, Tripathi A, et al. Compressive three-dimensional super-resolution microscopy with speckle-saturated fluorescence excitation [J]. *Nature Communications*, 2019, 10(1): 1327.
- [85] Gong H, Xu D, Yuan J, et al. High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level [J]. *Nature Communications*, 2016, 7(1): 1-12.
- [86] Carlton P M, Boulanger J, Kervran C, et al. Fast live simultaneous multiwavelength four-dimensional optical microscopy [J]. *Proceedings of the National Academy of Sciences*, 2010, 107(37): 16016-16022.
- [87] Luisier F, Blu T, Unser M. Image denoising in mixed Poisson-Gaussian noise [J]. *IEEE Trans Image Process*, 2011, 20(3): 696-708.
- [88] Soubies E, Soulez F, Mccann M T, et al. Pocket guide to solve

- inverse problems with GlobalBioIm [J]. *Inverse Problems*, 2019, 35(10): 104006.
- [89] Arigovindan M, Fung J C, Elnatan D, et al. High-resolution restoration of 3D structures from widefield images with extreme low signal-to-noise-ratio [J]. *Proceedings of the National Academy of Sciences*, 2013, 110(43): 17344-17349.
- [90] Setzer S, Steidl G, Teuber T. Deblurring Poissonian images by split Bregman techniques [J]. *Journal of Visual Communication and Image Representation*, 2010, 21(3): 193-199.
- [91] Zhang Y, Zhu Y, Nichols E, et al. A poisson-gaussian denoising dataset with real fluorescence microscopy images[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 11710-11718.
- [92] Hagen G M, Bendesky J, Machado R, et al. Fluorescence microscopy datasets for training deep neural networks [J]. *GigaScience*, 2021, 10(5): giab032.
- [93] Belthangady C, Royer L A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction [J]. *Nat Methods*, 2019, 16(12): 1215-1225.
- [94] 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): 792-803e719.
- [95] 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.
- [96] Khater I M, Aroca-Ouellette S T, Meng F, et al. Caveolae and scaffold detection from single molecule localization microscopy data using deep learning [J]. *PLoS One*, 2019, 14(8): e0211659.
- [97] Li X, Zhang G, Qiao H, et al. Unsupervised content-preserving transformation for optical microscopy [J]. *Light Sci Appl*, 2021, 10(1): 44.
- [98] Chen X, Kandel M E, He S, et al. Artificial confocal microscopy for deep label-free imaging [EB/OL]. (2021-10-28)[2022-08-01]. <https://arxiv.org/abs/2110.14823>.
- [99] Robitaille L É, Durand A, Gardner M-A, et al. Learning to become an expert: Deep networks applied to super-resolution microscopy[C]//Thirty-Second AAAI Conference on Artificial Intelligence, 2018.