

基于深度学习的生物显微图像超分辨率算法综述思考

自 1590 年世界上最早显微镜诞生以来，显微镜分辨率随技术 进 步 不 断 提 升。17 世纪科学家在光学显微镜下能够观察到红血细胞、细菌、酵母菌以及游动的精子。1873 年 德国物理学家恩斯特·阿贝第一次发现光学成像具有衍射限制现象，即显微技术受制于光线波长限制，0.2um 成为显微镜难以突破的瓶颈。自 20 世纪八九十年开始，埃里克·白兹格教授、威廉姆·艾斯科·莫尔纳尔教授和斯特凡·w·赫尔教授经过不懈努力突破了衍射极限，建成了可以观察纳米级的超分辨率荧光显微镜^[1]。

目前采用光学装置的方法超高分辨率荧光显微技术大体可分为两类，一类基于图像照明，即利用光学非线性效应修饰照明光路，使荧光发光点突破衍射极限，如受激发射损耗（STED）技术^[2]、可逆饱和荧光跃迁（RESOLFT）技术^[3]。另一类基于光控荧光探针和单分子定位，即通过光控荧光分子获得超分辨成像，如光敏定位显微（PALM）技术^[4]、随机光学重构显微（STROM）技术^[5]。

不同于显微技术的突破衍射极限的超分辨方法，在通用图像中 Tsai 和 Huang 于 1984 年提出了超分辨率重建^[6]的新概念，即将同一场景下获得的低分辨率图像序列作为样本，重建一幅空间分辨率较高的图像。随后，学者们从不同的研究角度提出了许多算法。根据不同的重建方法，超分辨率重建算法可以分为三类：基于插值、基于重建和基于学习^[7]。

而在基于学习的方法中的深度学习算法自被首次应用于图像超分辨率以来，诸多网络模型被提出，这些模型根据构建网络模型的方法可以分为：基于线性的卷积神经网络模型^[8]、基于残差网络的模型^[9]、基于密集网络的模型^[10]、基于生成对抗网络的模型^[11]。

基于线性的卷积神经网络模型

基于线性的卷积神经网络的超分辨率模型不同于一般的卷积神经网络模型，其中一般不包含池化层和全连接层。SRCNN^[12] 是首次被用于解决图像超分辨率重建问题的深度学习模型。ESPCN^[13] 和 FSRCNN^[14] 在重建方法和计算效率对 SRCNN 进行改进。ESPCN 通过引入了亚像素卷积层高效的完成重建时的上采样。FSRCNN 是针对 SRCNN 的卷积层进行改进设计，并且引入了反卷积层^[15] 用来实现上采样。基于卷积神经网络的模型一般网络结构较浅，简单的加深卷积神经网络的深度，在训练时往往会变得难以收敛。其中 SRCNN 只包含了三层卷积层，网络感受野大小为 13。较小的感受野会使得在重建时只有小部分区域的低分辨率图像信息可以作为参考信息，这在很大程度上制约重建效果。

表 1.SRCNN、EPSCN、FSRCNN 模型损失函数

模型	损失函数
SRCNN	$L(\theta) = \frac{1}{n} \sum_{i=1}^n \ F(Y_i; \theta) - X_i\ ^2$
EPSCN	$\ell(\omega_{1:l}, b_{1:l}) = \frac{1}{r^2 HW} \sum_{x=1}^{rH} \sum_{y=1}^{rW} (I_{x,y}^{HR} - f_{x,y}^l(I^{LR}))^2$
FSRCNN	$L(\theta) = \min_{\theta} \sum_{i=1}^n K \ F(Y_s^i; \theta) - X^i\ _2^2$

在 SRCNN, ESPCN 和 FSRNN 中都采用均方误差形式的函数作为网络训练的损失函数, 如表 1 所示, 其中 L 和 ℓ 表示损失函数, Θ , θ 表示网络参数, $F()$, $f()$ 表示网络的输出函数, Y_i , X_i 和 ILR , IHR 都分别表示训练图像对中的低分辨率图像和相应真实高分辨率图像, r , W , H 分别表示尺度变换因子, 特征图谱的宽和高, $\omega_{1:l}$, $b_{1:l}$ 分别表示网络中每层的权值和偏置。均方误差是反映估计量与被估计量之间差异程度的一种度量, 可以在一定程度上衡量重建图像与真实高分辨率图像之间的差异, 并且可以使网络取得较好的 PSNR 评价指标。

基于残差网络的模型

浅层的网络模型一般感受野都很小, 这会制约重建效果, 而深层的卷积神经网络又难以训练。由于在图像超分辨率时, 低分辨率图像和高分辨率图像之间存在大量相同信息, 同时残差网络具有一定的抗退化性能, 所以基于残差学习的模型充分利用这些特点, 使得网络只需要学习低分辨率图像和高分辨率图像之间的残差映射关系, 这相比于直接学习低分辨率图像到高分辨率图像之间的映射关系, 降低了网络参数的复杂性, 在一定意义上降低了学习难度, 所以基于残差学习的模型^[16-17]一般具有较深的网络结构。VDSR^[18]提出的网络结构中引入了全局残差学习, 并通过结构可调梯度裁剪技术成功将网络层数加深至 20 层, 其感受野增加到 41×41 (相比 SRCNN 的 13×13)。EDSR^[19]通过堆叠多个残差单元加深网络, 并在网络中引入了全局残差学习, 最终搭建了对某一特殊重建放大尺度的网络模型。同时, 作者还在 EDSR 的基础上, 通过在网络的前后两端加入不同重建放大尺度的处理模块构建了实现多尺度重建的 MDSR 网络。除此之外, DRCN^[20]和 DRRN^[21]通过不断递归残差单元加深网络, 通过递归的方式可以使得在加深网络的同时不增加网络参数, 其中 DRCN 包含了 16 个递归层, 整个网络的感受野达到 41×41 。DRRN 通过不断递归残差网络块将网络结构加深至 52 层。

基于密集网络的模型

低分辨率图像相对于高分辨率图像丢失了很多图像高频信息, 并且每一个像素点的修复都与其周边的图像像素信息密切相关, 所以, 在重建高分辨率图像时我们期望能尽可能多的提供低分辨率图像信息, 这不仅需要网络具有更大的感受野, 还需要充分利用网络中提取的分层特征信息。基于密集网络的模型通过在网络中引入密集跳步连接, 充分融合利用网络中的分层特征信息, 为重建高分辨率图像提供了更多更丰富的特征信息, 这有助于网络进行更加准确的图像重建。

MenNet^[22]以 Memory block 为网络单元, 对网络中 Memory block 进行密集跳步连接, 目的就是为了解决网络自适应学习的连续记忆功能, 同样在 Memory block 中把每一个残差单元都连接到最后的 Gate Unit, 与前面的 Memory block 输出进行特征融合。SRDenseNet^[23]以密集网络块作为网络的基本单元, 在通过跳步连接将每一个密集单元的输出连接到后面的卷积层进行分层特征融合。RDN^[24]在密集单元的基础引入残差学习, 构建了残差密集网络作为网络的基本单元, 最后同样地将每一个残差密集网络单元的跳步连接至一个 1×1 的卷积层进行全局特征融合。基于分层特征融合的模型针对特征信息在网络局部和全局进行融合, 使得深层网络中各层的特征信息能更加充分的融合利用, 为重建提供了更多的特征信息, 这有助于网络实现更好的重建效果。

基于生成式对抗网络的模型

基于对一般的深度学习网络模型进行训练学习, 建立的低分辨率图像到高分辨率图像之间的非线性映射关系, 由于人为设置的基于均方误差的损失函数常常使得网络最后学习到的重建结果往往过于平滑, 会丢失一些真实高分辨率图像中的细节和高频信息, 虽然在 PSNR 指标评价取得

了好的效果，但是却不一定能达到良好的人类视觉感受。生成式对抗网络通过生成器和鉴别器的相互博弈，可以使得生成器生成的图像更加接近真实的高分辨率图像，能更加准确的重建出图像的高频细节部分。SRGAN^[25] 中通过堆叠多个残差模块构建了深层的生成网络，同时构建了包含 8 个卷积层的鉴别网络。整个网络通过生成器和鉴别器的博弈交替优化，最终实现两个网络的纳什平衡。其最终目的就是实现 $V(G,D)$ 函数的优化，其中 G 、 D 分别表示生成函数和鉴别函数， z 、 x 分别表示训练图像对中的低分辨率图像和相应真实高分辨率图像， p_{data} 和 p_z 分别表示真实高分辨率的概率分布和低分辨率的概率分布。

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))]$$

训练生成模型使其生成的图像能够混淆鉴别模型，使鉴别模型难以区分图像是真实图像还是生成图像。而训练鉴别模型的目的就是使其尽可能区分图像是否为真实图像。此外，作者针对生成器的损失函数进行了优化，提出了感知损失函数 (perceptual loss)，由三部分组成：内容损失 (content loss)、对抗损失 (adversarial loss) 以及正则化损失 (regularization loss)，这使得整个生成式对抗模型产生的图像更加接近真实图像。

AttnGAN^[26] 提出了细粒度图像生成，借助文本描述生成包含充分细节重建图像。利用多模态相似性为目标，协同优化特征提取，使得 GAN 性能得到提升。Adrian Bulat^[27] 等人针对训练时常常人为对高分辨率图像进行双线性下采样以获得用于训练的低分辨率图像-高分辨率图像对。这使得网络在处理真实低分辨率图像时难以取得良好的重建效果，提出了先利用未配对的高分辨率图像-低分辨率图像训练生成式对抗网络，经过训练学习到图像的退化和下采样方法，再把这个网络的输出用于训练进行重建的生成式对抗网络。这使得生成对抗式网络能更好的学习到真实图像的退化过程，能使重建结果更加真实。

基于深度学习的显微超分辨应用

Elisa Nehme 等^[28]通过结合全卷积神经网络和 STORM 的技术相结合提出了 Deep-STORM 模型，将量子点集通过 ImageJ 的插件 ThunderSTORM 处理得到高分辨率的图像作为金标准，并采用 Deep-STORM 得到衍射受限的量子点到处理后高分辨率的映射模型，该模型可以提高 STORM 技术提高显微图像的效率并具有较高的准确性。Hao zhang 等^[29]采用 GAN 网络模型构建了一个 RFGANM 超分辨率工作流，该研究采用退化模型得到高低分辨率图像对，并在宽场显微、光片荧光等显微图像中进行了测试都得到了可观的效果。Louis-Emile Robitaille 等^[30]将深度学习的技术应用用于对 STED 技术得到的高分辨率图像质量的评估，该研究将网络模型与随机基准的评估模型进行比较，结果表明，深度学习有更好的结果，但作者也表示，由于现有数据的有限，得到的结果无法说明有显著性差异。Hongda Wang 等^[31]采用 GAN 网络实现了跨模态荧光显微成像超分辨率技术。该研究主要的思路是将分辨率较差的技术产生的图片映射到分辨率较好的技术生成的图片，其主要应用深度学习对三对硬件实现的超分辨率技术的映射关系分别建模，包括不同倍率的宽场荧光成像之间的映射、共聚焦到 STED 的映射、TIRF 到 TIRF-SIM 的映射。分别构建的模型用于相对应的映射关系的图像的转换，从而达到实现无参数预设、精准的超分辨率技术。Ruud. J.G. van Sloun^[32]等将深度学习的技术应用用于超声显微超分辨率的研究，提出了 Deep-ULM 模型，该模型采用基于 u-net 的网络机构将超声造影数据减少衍射限制的影响，得到接近采用 PFS 技术进行超分辨率得到的高分辨率的结果，该模型可实时的得到准确的超分辨的图像。

思考

经过对文献的阅读，基于深度学习的技术层出不穷，在生物显微的超分辨率领域主要从两个方面切入，第一，应用层面，大多数现有的将深度学习应用在显微超分辨率的方法都是从各类硬件技术的超分辨水平参差不齐入手，通过将一种低水平的超分辨率技术训练模型得到与高水平超分辨技术相当的结果。第二，算法层面，将各种经典的网络模型思想进行混合，并根据使用的场景引入适当的数学模型进行优化改进。目前需要尝试的是结合这两个方面进行探索，并重点倾向于算法层面的改进，发挥工科的优势。

参考文献

- [1] 超分辨率荧光显微技术---2014 年诺贝尔化学奖
- [2] Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy
- [3] Fluorescence nanoscopy: Breaking the diffraction barrier by the RESOLFT concept
- [4] Imaging intracellular fluorescent proteins at nanometer resolution
- [5] Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)
- [6] Multiple frame image restoration and registration
- [7] 基于医学图像的超分辨率重建算法综述
- [8] Backpropagation applied to handwritten zip code recognition
- [9] Deep Residual Learning for Image Recognition
- [10] Densely Connected Convolutional Networks
- [11] generative adversarial network
- [12] Image super-resolution using deep convolutional networks
- [13] Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network
- [14] Accelerating the super-resolution convolutional neural network
- [15] Adaptive deconvolutional networks for mid and high level feature learning
- [16] Wavelet residual network for low-dose CT via deep convolutional framelets
- [17] Handwritten Chinese character recognition with spatial transformer and deep residual networks
- [18] Accurate image super-resolution using very deep convolutional networks
- [19] End-to-end image super-resolution via deep and shallow convolutional networks
- [20] Deeply-recursive convolutional network for image super-resolution
- [21] Image super-resolution via deep recursive residual network

- [22] MemNet: a persistent memory network for image restoration
- [23] Image super-resolution using dense skip connections
- [24] Residual dense network for image restoration
- [25] Photo-realistic single image super-resolution using a generative adversarial network
- [26] AttnGAN: fine-grained text to image generation with attentional generative adversarial networks
- [27] To learn image super-resolution, use a GAN to learn how to do image degradation first
- [28] Deep-STORM: super-resolution single-molecule microscopy by deep learning
- [29] High-throughput, high-resolution registration-free generative adversarial network microscopy
- [30] Learning to Become an Expert: Deep Networks Applied To Super-Resolution Microscopy
- [31] Deep learning enables cross-modality super-resolution in fluorescence microscopy
- [32] Super-resolution Ultrasound Localization Microscopy through Deep Learning