深度学习在生物显微图像中的应用综述

叶学华 黄钢

**摘要** 近年来深度学习技术的发展已经慢慢的渗入了各行各业之中，尤其是计算机视觉领域，深度学习凭借其强大的图像高维信息的处理能力，将人类所不能理解的信息进行提取及自动化分析。随着生物医学技术的不断发展，研究者们不断的深入了解生物的本质，生物显微技术在其中就起到了十分重要的作用，而随之而产生的大量的生物显微图像，如数字病理图、荧光细胞图像等，正好与数据驱动的深度学习技术相契合，使生物显微图像的分析得到快速的发展。该文章先简要对深度学习技术进行了介绍，然后根据分类、目标检测、分割、超分辨率等子任务对深度学习的应用及创新进行详细的论述，最后展望了深度学习在生物显微图像上的发展趋势及所面临的挑战。

**深度学习在显微图像分类上的应用**

图像分类，即将给定的图像定义一个标签，来决定该图像的类别。如根据荧光显微图像来分类蛋白质的表达是在细胞质内还是细胞核内，来深度学习作为强监督的学习算法，对数据的提取与分析，与传统的学习算法相比有显著的提升，在生物显微图像领域，深度学习也是生物显微图像的分类中十分强大的工具。通过深度学习的方法进行分类可以直接使用神经网络作为分类器，得出预测的结果。也可以将神经网络作为特征提取器，该提取器一般为在大量数据下训练后的预训练模型，采用预训练模型从数据中提取有效的特征，再结合其他的分类器来得到一个更好的分类模型，或者对于小数据集的分类问题，采用迁移学习的方法对数据进行微调。

Gao[]等采用CNN作为分类器将HEp-2细胞分为六类，该研究通过实验比较了CNN超参数、数据增广、图像前景模板对图像分类表现的影响，另外该研究表明通过使用大量数据训练过的模型作为预训练模型，并微调模型来训练较小的数据集，可以得到比使用小数据集直接训练更好的精准度。实际上，使用预训练模型更符合真实场景中主要为小数据量的情况。Carneiro[]等采用了深度学习的模型推测出人体鳞状细胞癌中的微循环供应单元（MCSU）中类别的数量及占比。对于一对输入的图像，该研究首次四种不同的分类器来生成多层的输出映射，然后再利用CNN来进行预测MCSU的种类。预训练的CNN模型也可以作为特征提取器，将高维的信息提取出来，再采用其他工具进行分类。在[94]中，作者采用ImageNet数据集训练后的预训练模型，将其网络模型的倒数第二层的结果作为高维的特征图，并进行特征池化并提取出特征，最后采用线性支持向量分类器来分类多形性成胶质细胞瘤和低级别胶质瘤。Nils Gessert等[]对VGG-16、InceptionV3、Densenet21、SE-Resnext这四个经典的架构进行了迁移学习的比较，研究中采用这四个预训练的模型分别都对共聚焦激光显微图像下的直肠、腹膜进行了三个二分类的任务，实验表明这些在ImageNet中预训练的模型，迁移到共聚焦激光显微图像上是可行的，但事实上所得到的结果还有有限制的，文章表明深度学习在一定程度是有效的，但不可能在所有的共聚焦显微图像任务中有效，对特定的任务还是得寻找特定得方法。

**深度学习在显微图像目标检测上的应用**

目标检测，即对目标物体进行分类及定位，该任务关注的是特定物体在图像中的位置信息，便于对特定目标物体的研究，因此对感兴趣目标的检测在生物显微图像中也有十分重要的意义，如对细胞及个细胞器的检测，为目标物体的计数、分割和跟踪提供研究基础。目标物体的位置一般用检测到的质心附近的点来标记，即目标检测可以被模拟成像素级的分类问题，对于输入的图像，经过神经网络输出概率图，进而在概率图中寻找局部的最大值，最终得到检测的目标位置。通常为了优化最终的结果，采用非极大值抑制来将干扰的结果去除。

不同于将深度学习视为像素级的分类任务，Ciresan等[]在斑纹鱼宽场显微细胞图像的检测中，采用了SVM分类器检测细胞的候选区域，然后再通过CNN网络来区分候选区域为背景还是细胞。实验表明该方法相对与采用CNN对每个像素进行分类，准确率有了显著的提升。对于有丝分裂的检测中，A.Shkolyar[]等也采用了类似的方法，采用CNN来对有丝分裂的候选块进行分类，但将SVM替换成了简单的预处理操作，如滤波、阈值处理等。Albarqouni[]等提出了一个多尺度的AggNet框架，用于从大量的病例中检测H&E染色乳腺癌细胞中的有丝分裂。该研究中引入了一个附加的众包层，AggNet可以直接处理学习过程中的数据聚合，从而使非专家标注的图像可以应用于生物医学当中。具体来说，在AggNet中首先采用金标准的标注图像训练多个CNN网络，然后利用多个CNN对新的数据进行检测，通过Web平台将检测出来的候选区域给网络中参与众包的人标注，最后将收集回的标注用来优化模型并生成金标准，该框架同时也避免噪声标注的干扰。相对于先将检测出来的候选区域，再将其进行分类的二段式算法，YOLO等一段式的算法可以实现端到端的检测模型，并且由于去除了生成候选区域的步骤使计算的速度大大加快，并且能够充分地使用到图像的全局信息，因此在检测的准确率上有较大的提升。Thorsten Wagner等[]提出了CrYOLO模型，将YOLO框架应用于低温电子显微单粒子的检测，在40个数据集中进行了训练，并且在未知的大分子中也做了测试，实现了专家级的全自动化的单粒子检测模型，并且为了弥补YOLO在小物体上检测不精确的缺点，将原图进行了下采样，提高了一定的精度。

**深度学习在显微图像分割上的应用**

图像分割，即从原图中提取分割对象所占的区域。生物显微图像中，特别是对细胞及细胞核的分割，是许多图像分析任务的基础，如细胞形态学计算，特征量化，细胞识别等。因此对生物显微图像的精准分割也是一个研究的重点，传统的分割算法对显微图像也有一定的准确度，如水平集、阈值分割、分水岭等算法，但深度学习的CNN模型的出现使分割的结果得到了更大的提升。CNN模型解决了一个像素级的分类问题，即将输出的每个像素分为背景或分割对象，最后输出分割后的结果。而由于CNN像素级的分类使计算量大大提升，为了解决这个问题，J.Long等[]提出FCN将CNN中的全连接层使用卷积层代替，从而使输出的结果为一个概率图，最终实现一个端到端的神经网络。

基于FCN的思想，Ronneberger等[]提出了U-Net网络模型，该模型在2015年电子显微图像分割竞赛中获得了第一名。U-Net网络考虑到越深的网络层次所提取的高维信息会失去纹理信息，因此在网络设计时构建了收缩路径和拓展路径的对称结构，收缩路径提取高维的特征，而在拓展路径中，将卷积的结果和与之相对称的收缩路径中的信息复制级联，最终输出一个分割结果的概率图。该神经网络的设计将在不断的利用更深的网络提取高维的信息时，有效的结合最开始几层中的较低维的特征图，使信息得到更充分的利用。另外该研究还提出使用弹性变换来对数据进行增强，有利于小数据集分割的研究。Chen等[]在分割神经元结构的显微图像时，改善并通过引入多层次的纹理信息和辅助的分类器对FCN进行了拓展，也将模型分为收缩路径和拓展路径，收缩路径用来将语义信息进行分层，而拓展路径包括卷积和反卷积层，将收缩路径中每层的高维特征图拓展为原始图像大小，最后将各层的拓展路径相加求和并输入到softmax层中。为了降低梯度消失发生的概率，并提高中间层对特征区分的能力，引入了辅助的分类器来进行端到端的训练。Assfaf Arbelle等[]受GAN网络的启发提出了对抗式的神经网络Rib Cage，该网络由预估器与判别器组成，不同于GAN网络是该模型不是从随机噪声产生图像，而是预估图像中的底层的变量，通过预估器生成某些分割图像，判别器则区分专家标注图像与生成的分割图像。判别器就是最小化输入两类图像的分类损失，如最小化标注图像与生成图像的相似度，那么预估器就可以是最大化生成图像与标注图像的相似度。rib cage网络将GAN网络拓展到了细胞分割的任务上，结合了分割图像和原图像之间的交叉联系，并且该研究还证明在低数据量的情况下也有不错的效果，减少了人工标注的负担。assfaf等[]结合了LSTM和U-Net结构各自的优势，提出了新型的网络用于时序的细胞序列图像分割，LSTM多用于自然语言处理的任务，擅长捕捉动态变化的关联信息，而通过该两种模型的结合将动态的细胞序列图像的时序变化和空间信息充分利用，从而达到提高分割准确率的作用。

**深度学习在显微图像超分辨率上的应用**

超分辨率技术，即提高图片的辨识度。随着医学技术的发展，组织病理学、细胞学等学科研究对高分辨率图像的需求逐渐提高，但由于硬件设备的限制，得到更高分辨率的图像就必须得选择更小的视场，并且在特定的情况下，还会产生衍射限制的现象，使更微观层次的探索难以进行。虽然可以通过升级硬件，但所需成本却十分高昂，而深度学习算法凭借其强大的学习能力，通过学习高分辨率图像与低分辨率图像之间高频细节的分布差异，可直接推理低输入的低分辨率的图像得到高分辨率的图像，从而在相对较差的硬件条件下得到精确而又清晰的图像。

目前采用深度学习算法研究超分辨率技术主要的思路就是：找到两种或多种合适的不同分辨率的显微成像技术，通过FCN、GAN等主流的网络来构建模型。Elisa Nehme等[28]通过结合全卷积神经网络和受激发损耗（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。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。Hao zhang等[29]也采用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。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技术进行超分解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。Louis-Emile Robitaille等[30]将深度学习的技术应用于对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。