

基于深度学习的数字病理图像分割综述与展望^{*}

宋杰¹, 肖亮², 练智超², 蔡子贇¹, 蒋国平¹

¹(南京邮电大学 自动化学院、人工智能学院, 江苏 南京 210023)

²(南京理工大学 计算机科学与工程学院, 江苏 南京 210094)

通讯作者: 肖亮, E-mail: xiaoliang@mail.njust.edu.cn



摘要: 数字病理图像分析对于乳腺癌、前列腺癌等良恶性分级诊断具有重要意义,其中,组织基元的形态和目标测量是量化分析的重要依据.然而,由于病理数据多样性和复杂性等新特点,其分割任务面临着特征提取困难、实例分割困难等挑战.人工智能辅助病理量化分析将复杂病理数据转化为可挖掘的图像特征,使得自动提取组织基元的定量化信息成为可能.特别是随着计算机计算能力的快速发展,深度学习技术凭借其强大的特征学习、设计灵活等特性在数字病理量化分析领域取得了突破性成果.系统概述目前代表性深度学习方法,包括卷积神经网络、全卷积神经网络、编码器-解码器模型、循环神经网络、生成对抗网络等方法体系,总结深度学习在病理图像分割等任务中的建模机理和应用,并梳理了现有方法的方法理论、关键技术、优缺点和性能分析.最后讨论了未来数字病理图像分割深度学习建模的开放性挑战和新趋势.

关键词: 数字病理;组织基元;实例分割;特征表示学习;深度模型

中图分类号: TP391

中文引用格式: 宋杰,肖亮,练智超,蔡子贇,蒋国平.基于深度学习的数字病理图像分割综述与展望.软件学报,2021,32(5): 1427–1460. <http://www.jos.org.cn/1000-9825/6205.htm>

英文引用格式: Song J, Xiao L, Lian ZC, Cai ZY, Jiang GP. Overview and prospect of deep learning for image segmentation in digital pathology. Ruan Jian Xue Bao/Journal of Software, 2021, 32(5): 1427–1460 (in Chinese). <http://www.jos.org.cn/1000-9825/6205.htm>

Overview and Prospect of Deep Learning for Image Segmentation in Digital Pathology

SONG Jie¹, XIAO Liang², LIAN Zhi-Chao², CAI Zi-Yun¹, JIANG Guo-Ping¹

¹(College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

²(School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: The quantitative analysis of digital pathology images plays a significant role in the diagnosis of benign and malignant diseases such as breast cancer and prostate cancer, in which the morphology measurements of histologic primitives serve as a basis of quantitative analyses. However, the complex nature of digital pathology data, such as diversity, present significant challenges for such segmentation task, which might lead to difficulties in feature extraction and instance segmentation. By converting complex pathology data into minable image features using artificial intelligence assisted pathologist's analysis, it becomes possible to automatically extract

* 基金项目: 国家自然科学基金(62001247, 61871226, 61571230, 62006127, 61873326, 61672298); 江苏省社会发展重点研发计划(BE2018727); 江苏省自然科学基金(BK20190728); 江苏省高等学校自然科学研究面上项目(20KJB520005); 南京邮电大学引进人才科研启动基金(NY219152, NY218120)

Foundation item: National Natural Science Foundation of China (62001247, 61871226, 61571230, 62006127, 61873326, 61672298); Jiangsu Provincial Social Developing Project (BE2018727); Natural Science Foundation of Jiangsu Province (BK20190728); Natural Science Foundation for Colleges and Universities in Jiangsu Province (20KJB520005); Introduction of Talent Research Start-up Fund of Nanjing University of Posts and Telecommunications (NY219152, NY218120)

收稿时间: 2020-08-15; 修改时间: 2020-09-27; 采用时间: 2020-11-18; jos 在线出版时间: 2020-12-02

quantitative information of individual primitives. Machine learning algorithms, in particular deep models, are emerging as leading tools in quantitative analyses of digital pathology. It has exhibited great power in feature learning with producing improved accuracy of various tasks. This survey provides a comprehensive review of this fast-growing field. Popular deep models are briefly introduced, including convolutional neural networks, fully convolutional networks, encoder-decoder architectures, recurrent neural networks, and generative adversarial networks, and current deep learning achievements in various tasks are summarized, such as detection and segmentation. This study also presents the mathematical theory, key steps, main advantages and disadvantages, and performance analysis of deep learning algorithms, and interprets their formulations or modelings for specific tasks. In addition, the open challenges and potential trends of future research are discussed in pathology image segmentation using deep learning.

Key words: digital pathology; histologic primitive; instance segmentation; feature representation learning; deep model

全视野数字切片是一种高分辨率的数字病理图像,其所包含的巨大信息为数字病理的定量分析提供了可靠的基础^[1].病理量化分析对于乳腺癌、前列腺癌、肺癌等良恶性分级诊断具有重要意义,其中,组织基元形态与目标测量是病理量化分析的重要依据.国际上,以 Lin Yang^[2,3]、Erik Meijering^[4]、Badrinath Roysam^[5]、Nasir M. Rajpoot^[6]、Bahram Parvin^[7]为代表的众多学者对数字病理数据进行了大量的建模统计和量化分析,为当代数字病理的发展奠定了重要的理论和实践基础.在国内,包括清华大学^[8]、浙江大学^[9]、复旦大学^[10]、武汉大学^[11]、湖南大学^[12]、四川大学华西医院^[13]、南京信息工程大学^[14]等在内的多家科研机构均在病理图像处理领域进行了较深入的研究,极大地推动了数字病理在国内的发展和应用.

在病理诊断和分析过程中,显微解剖结构的分割起着关键作用.分割结果可被用于可视化组织基元形态、计算与病变相关的量化指标、定位感兴趣区域(region of interest,简称 ROI)以及辅助制定外科手术方案等.随着数字病理的发展,临床和医学研究迫切需要能够进行多个器官和疾病状态的病理图像高精度分割方法.然而,由于病理图像背景结构复杂,组织基元高密度分布、粘连重叠等,精准实例分割是一个挑战性问题(如图 1 所示).

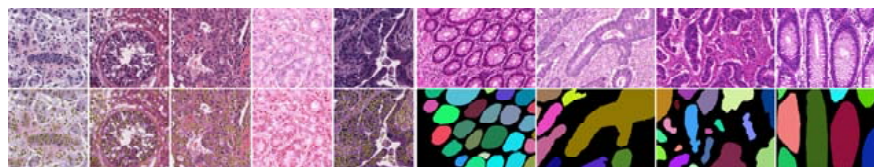


Fig.1 Digital pathology images from different tissues and disease states

图 1 不同组织器官和疾病状态下的数字病理图像

目前,基于数据与模型驱动的深度学习的分割方法是研究热点.传统人工特征选择依赖于大量专业知识,有限的表达能力难以涵盖图像的全面信息^[2].深度学习可分层抽取原始图像的抽象特征,为解决表现多样性的图像分割问题提供了有效途径.代表性网络包括卷积神经网络(convolutional neural network,简称 CNN)^[15]、全卷积网络(fully convolutional networks,简称 FCN)^[16]、编码器-解码器^[17]、循环神经网络(recurrent neural networks,简称 RNN)^[18]以及生成对抗网络(generative adversarial network,简称 GAN)^[19].深度学习模型采用逐层训练方式,完成从底层到高层的特征学习过程,利用深层特征表示学习来获取数据的高层语义信息,强调了模型对特征的学习能力,甚至可通过微调对特征进行调整,以便得到更有效的特征表示.近年来,基于深度学习技术,大量方法被提出来,广泛应用于目标检测^[20]、图像分割^[21]、自然语言处理(natural language processing,简称 NLP)^[22]、医学成像^[23,24]、计算生物学^[25]等领域.通过自动发现潜在数据结构,深度学习在处理不同领域的不同数据时,也能达到极佳的性能^[26-28].

数字病理图像量化分析领域积累了数以千计的原始文献资料,以若干综述加以贯穿.例如:Gurcan 等人^[29]概述了计算机辅助诊断的病理图像分析方法,包括预处理、细胞核与腺体分割、特征提取和分类;Veta 等人^[30]专题论述了乳腺癌组织病理图像分析方法,并对有丝分裂检测和增殖评估进行了额外讨论;Irshad 等人^[31]针对特定染色的组织病理图像,总结了细胞核检测、分割、特征提取以及分类方法.最近,来茂德等人^[10]结合数字病理图像,对细胞、腺体等组织基元的检测与分割算法进行了补充.

本文正是在此背景下,围绕多个器官和疾病状态的组织基元深度学习图像分割问题,通过对相关领域有代表性或前沿性文献的归纳总结,系统概述目前代表性深度学习方法,包括卷积神经网络、全卷积网络、编码器-解码器模型、循环神经网络、生成对抗网络等方法体系,总结深度学习在组织基元检测和实例分割等任务中的建模机理和应用,并梳理了现有方法的方法理论、关键技术、优缺点和性能分析.最后,本文讨论了未来数字病理图像分割深度学习建模的开放性挑战和新趋势.

1 数字病理相关学科背景

病理学是研究疾病病因、发病机制、形态结构改变以及由此而引起的功能变化的一门基础医学与临床医学之间的桥梁学科.近年来,由于生物学、微生物学、生物化学、免疫学和分子生物学等的迅猛发展,以及透射电镜、扫描电镜、图像分析及流式细胞仪等的研制成功,病理诊断已经成为当今疾病诊断、治疗和预防的重要途径和金标准,而病理切片的数字化被认为是病理学和病理诊断发展过程中的重要转折点^[1].

1.1 病理切片染色及成像模式

数字病理切片的制作首先需要组织染色.为了突出切片中特定的细胞核和腺体特征,限定并检查组织,通常使用染色剂来增强组织成分间的对比度,主要包括苏木精-伊红(hematoxylin-eosin,简称 H&E)和免疫组织化学(immuno histo chemical,简称 IHC),如图 2 所示.H&E 是最常用的染色方法,其中,苏木精为碱性染料,可使细胞核内的染色质与胞质内的核酸着紫蓝色;而伊红为酸性染料,使细胞质和细胞外基质中的成分着粉红色.组织结构对不同染料的结合程度不同,由此显示组织切片和细胞标本中的成分以及病变的一般形态结构特征^[31].与 H&E 常规染色相比,IHC 染色利用抗原抗体的特异性结合反应来检测和定位组织和细胞中的某些化学物质,具有较高的敏感性,可将形态学改变与功能代谢变化相结合,从而能够鉴别、诊断和治疗恶性肿瘤.

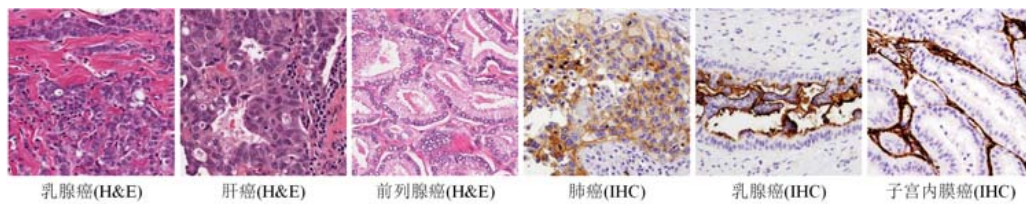


Fig.2 Examples of H&E and IHC staining for different types of digital pathology images

图 2 不同类型数字病理切片的 H&E 和 IHC 染色示例

一般而言,染色之后,可利用数字显微镜或放大系统在低倍物镜下对玻璃切片进行逐幅扫描采集成像.图像压缩和存储软件将高分辨数字图像自动进行无缝拼接处理,制作生成整张全视野数字切片(多层图像,如图 3 所示).病理学家和研究者通过对可视化数据进行任意比例放大或缩小以及任意方向移动和浏览,可实现高效病理分析和评估,包括腺体形态、有丝分裂数目、细胞核外观、周围细胞质颜色以及实质与间质的比例等^[5].需要注意的是:由于受显微镜设置、样品厚度、染色深浅和渗漏的影响,成像数据往往呈现不均匀照明.为了减少照明差异,大多数的扫描装置利用标准软件包来归一化和校正光谱及空间照明差异.

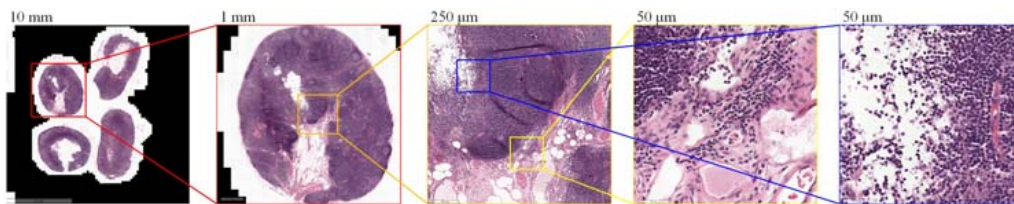


Fig.3 Example of whole slide images from CAMELYON17^[32]

图 3 全视野数字切片示例,来源于 CAMELYON17 数据集^[32]

1.2 肿瘤的组织学分级

肿瘤分级^[33]反映了在组织结构和细胞形态上肿瘤组织与正常组织细胞不同程度的形态差异,可用来判断组织是否癌变.目前,许多类型肿瘤的组织学分级与患者预后的关系已经引起病理学家的重视,包括:

(1) 乳腺癌组织学分级:乳腺癌病理组织形态较为复杂,类型众多,且在同一块癌组织甚至同一张切片中,往往存在两种以上类型.因此,临床实践需要结合病理类型进行组织学分级.一般而言,乳腺癌组织学分级主要从 3 个方面进行评估:腺管形成的程度、细胞核的多形性以及细胞核分裂计数.

(2) 前列腺癌组织学分级:在前列腺癌的组织学分级中,Gleason 评分系统根据前列腺组织的细胞、腺体等结构的形态特征,将前列腺癌分为 5 个等级,等级越高,恶性程度越高,预后也就越不理想.

(3) 结肠直肠癌组织学分级:大多数的结肠直肠癌是腺癌,其组织学分级建议采用 4 级分类法,与组织学分型高、中、低和未分化相对应.

(4) 胃腺癌组织学分级:消化道是腺癌主要的发生部位,根据形态特征及参照 WHO 标准,通常把胃腺癌分为高、中、低分化这 3 个等级.

(5) 肝细胞癌组织学分级:肿瘤细胞具有肝细胞的基本形态特征,呈索状或团块状排列.按肿瘤细胞的异型性及其组织结构,肝细胞癌分为 I 级(高分化)、II 级(中分化)、III 级(低分化)和 IV 级(未分化)这 4 个级别.

(6) 肾细胞癌组织学分级:肾细胞癌的组织学分级采用适合临床实际应用的 Fuhrman 细胞核分级标准.核分级对于透明细胞性肾细胞癌而言,是除临床分期外最重要的预后指标,该标准根据细胞核的大小、轮廓和核仁的特点,将肾细胞癌分为 4 级.

(7) 膀胱癌组织学分级:膀胱癌的复发、浸润与组织学分级密切相关.尿路上皮癌的形态学分级方案有多种,基本原则建立在细胞形态上,并兼顾结构和浸润状态,但在浸润性癌中较为强调分期和组织学变型,缺乏组织学分级标准;而在非浸润性癌组织学分级标准,则采用 WHO 分类的分级标准.

2 病理图像特点与分割挑战

迄今为止,研究者主要针对计算机断层扫描(computer tomography,简称 CT)、核磁共振(magnetic resonance image,简称 MRI)、正电子发射断层扫描(positron emission tomography,简称 PET)、X 射线、超声、各种显微镜等不同成像原理的医学图像分析任务开展了一系列的深度学习研究工作.CT 图像不能提供清晰的软组织和病灶影像,精确分割软组织器官和肿瘤 CT 图像十分困难.MRI 能够更好地可视化肿瘤和正常软组织器官,但不能提供骨性组织解剖结构影像,组织器官之间的空隙容易导致伪影,且难以校正.PET 图像通常缺乏组织器官的解剖结构影像,且其空间分辨率远低于 CT 和 MRI.X 射线图像中各组织器官和病灶信息重叠,自动识别比较困难.而超声图像常用来鉴别胎儿发育是否正常,判断内脏器官的形态是否有异常,确定病灶范围和物理性质.

相比之下,数字病理图像具有独特的微观结构特征,包括组织基元部分成像、高密度分布、重叠粘连和噪声等.其中,组织基元部分成像受着色、切割均匀度、切片厚度和重叠程度的影响,而缺失部分不仅出现在边缘附近,也可能发生在目标中心附近;组织基元高密度分布和噪声也会导致分割性能受到限制,从而出现过检测、漏检测和误检测等突出问题;另外,由于染剂的不均匀吸收,组织基元内外会有灰度(或颜色)和纹理差异,从而降低了目标前景和图像背景间的对比度,使其边缘模糊,导致过分割;与体外培养的正常基元不同,恶性肿瘤基元的接触抑制特性降低或丧失,导致其重叠生长和异质性.基元间的重叠度越高,对比度就越低,遮挡边缘就越难察觉,从而导致欠分割.病理图像特征提取通常需要显微学专家认真的分析和严格的理论推导,面对如今表现多样性、染质不均匀、背景结构复杂的数字病理图像,相应的图像分析工作难度陡增,有限的专家资源远不能满足对其进行有效特征提取的需求.

不同类型的病理图像呈现出在特定环境下不同形态的颗粒状、管状结构.正常情况下,腺通常由形成内部管状结构的管腔区域和细胞质周围的上皮细胞核构成,如图 1 所示.相比之下,由腺上皮引起的恶性肿瘤,即腺癌,会使腺的形态呈现出不规则的退化表现,例如乳腺癌、前列腺癌、结肠癌等.个体细胞分割具有如下挑战:一方面,对于病理状态(例如增生或某种癌症亚型)图像而言,由于细胞核增大,并呈现染质浓集贴边,即核内染质较

浅,而边缘附近染质较深,而着色较深的核仁也大量出现在核内;另一方面,由于病理图像中往往细胞密度高、细胞间出现重叠和成团等突出问题,加剧了个体细胞分割的难度.与颗粒状目标分割不同,神经纤维等管状结构的完整形态追踪和重构是描述神经元类型的关键参数.然而,一方面,由于受成像和着色局限性的影响,使得管状和树状纤维的灰度及厚度发生急剧变化,导致其形态不规则;另一方面,纤维扭曲、纤维间出现缠绕等问题,加剧了个体纤维追踪的难度.高效的分割算法需要在提取图像语义特征的同时,建模实例的底层图像信息.一方面,现有方法是问题导向的,使其可迁移性较差,导致分割性能受到限制;另一方面,由于神经纤维扭曲、缠绕,细胞高密度分布、粘连重叠等,建模过程容易丢失底层图像信息,加剧了病理图像实例分割的难度.

此外,小样本问题是指数据规模大但其中标注样本很少.数字病理图像的人工标注往往需要专业人士的指导,不仅耗时耗力,而且存在主观差异性.因此在实际操作中,由于受少量可训练样本的影响,使得具有强表达能力的分割模型易产生过拟合结果.

表 1 列出了现有公开病理数据集.

Table 1 Overview of histopathological image sets publicly available

表 1 现有公开的组织病理图像数据集概述

数据及相关文献	图像尺寸(像素)	图像数	染色方法	疾病类型	额外数据	用途
TCGA ^[34]	400×400	18 462	H&E	肾细胞癌	基因组、转录组和表观基因组	细胞核检测与分割
Warwick-QU ^[6]	574~775×430~522	165	H&E	结肠直肠癌	腺区域掩膜	腺分割
CAMELYON17 ^[32]	131072~195072×80384~90112	1 000	H&E	乳腺癌(淋巴结转移)	癌症区域掩膜	乳腺癌转移检测与分类
KIMIA Path960 ^[35]	308×168	960	H&E, IHC	各类型组织	—	上皮细胞分割
Bio-Segmentation ^[36]	896×768, 768×512	58	H&E	乳腺癌	—	细胞分割、微管追踪
DIADEM ^[37]	3752~8040×2280~4343, 1024×1024, 512×512	302(图像切片层数)	—	各类型神经元	中轴线和体素坐标	神经元、投射纤维、爬行纤维、轴突追踪与重构
Kather ^[38]	1000×1000	100	IHC	结肠直肠癌	血管计数	血管检测
MITOS-ATYPIA-14 ^[39]	1539×1376, 1663×1485	4 240	H&E	乳腺癌	有丝分裂坐标、核异型评估准则	有丝分裂检测、细胞核异型分类
Kumar ^[40,41]	1000×1000	44	H&E	各类型癌症	细胞核边缘坐标	细胞核分割
DEEP LEARNING ^[42]	2000×2000, 100×100, 1000×1000, 2000×2000, 775×522	681	H&E	乳腺癌、结肠直肠癌	有丝分裂坐标、淋巴细胞坐标、上皮细胞掩膜、细胞核掩膜、腺区域掩膜	有丝分裂检测、淋巴细胞检测、上皮细胞分割、细胞核分割、腺分割
Gertych ^[43]	1200×1200	210	H&E	前列腺癌	腺区域掩膜	腺分割
MEGA ^[44]	1280×800	300	H&E	结肠癌	结肠癌区域掩膜	病理图像分割
AMIDA2016 ^[45]	2000×2000	621	H&E	乳腺癌	ROI 区域掩膜	病理 ROI 区域检测
DRIVE, STARE, HRF ^[46]	565×584, 700×605, 3504×2336	40,20,45	—	糖尿病视网膜病变、脉络膜病变等	血管区域掩膜、视场角掩膜	视网膜血管分割

因此,病理图像分割具有如下新挑战.

(1)

由于组织病理图像中组织基元高密度分布、基元间出现重叠、缠绕,使得传统基于图像分析方法的检测与分割很难获得正确的结果.

(2)

对于病理状态图像而言,由于染质稀疏、背景杂斑干扰、表观多样等,基于有效特征提取的个体基元精准实例分割是一个挑战性问题.

(3)

对于不同类型的病理图像而言,由于图像特性存在显著差异,如何设计一个普适的分割方法是一个重要且具有挑战性的研究问题.

- (4) 大多数已有的分割方法只局限于重叠或接触基元的简单分割,精确的完整轮廓推断能够显著改进个体基元的形态计算精度.
- (5) 在已有方法中,检测与分割是两个相互独立的过程,如何设计一个统一的框架,联立执行检测与分割是亟待解决的问题.
- (6) 已有深度学习方法需要较大规模的训练数据集以经验拟合深层网络参数,如何设计更加高效模型提升表征学习能力是当前的研究热点之一.

3 深度学习方法及软硬件概述

3.1 卷积神经网络

CNN 是近年来深度学习能够在各个领域取得突破性成果的基石^[47],它主要包括 4 类网络层,如图 4 所示.

- (1) 卷积层,旨在学习一组卷积核,并通过在输入图像上滑动不同的卷积核计算多张特征图.其中,每张特征图对应一个卷积核,这意味着同一张特征图共享权重和偏置项.
- (2) 非线性层,通过在特征图上逐元素地应用激活函数,增强判定函数和整个网络的非线性,使得网络能够对非线性函数进行建模.
- (3) 池化层,利用池化核对特征图中不同位置的特征进行聚合统计^[48],大大降低了特征维数并增加了对局部变换的鲁棒性.
- (4) 全连接层,在网络中起到分类器的作用,将学习到的特征映射到样本标记空间.其中,每个单元都与前层中的所有单元相连接,并且最后一层通过使用损失函数惩罚预测值与真实值间的差异以实现精确分类的目的.

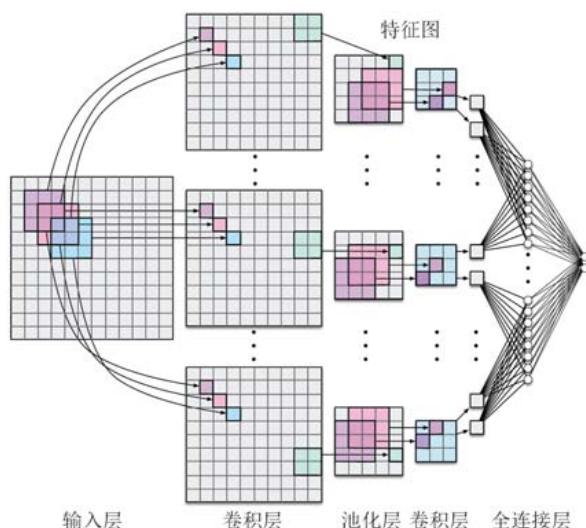


Fig.4 Architecture of CNN

图 4 CNN 体系结构

在 CNN 中,影响元素前向计算的所有可能的输入区域称为该元素的感受野(receptive field).可通过更深的网络使特征图中单个元素的感受野变得更加广阔,从而捕捉输入上更大尺寸的特征.与传统的全连接神经网络相比,CNN 的主要优势是通过局部感受野、权值共享以及时间或空间子采样减少网络中自由参数的个数,获得某种程度上的位移、尺度和形变不变性.比较流行的 CNN 网络结构包括:AlexNet^[49]、VGGNet^[50]、ResNet^[51]、GoogLeNet^[52]、MobileNet^[53]和 DenseNet^[54].

CNN 网络采用随机梯度下降和误差反向传播方法对模型参数进行更新,以实现模型训练的目的.CNN 网络

为了避免过拟合,通常也将随机丢弃(dropout)技术^[55]应用到网络模型学习.此外,CNN 通过使用数据扩充机制增加训练样本数量及多样性,进而既避免了过拟合又提升了模型性能.

在分割应用中,CNN 在卷积层之后会接上若干个全连接层,将卷积层产生的特征图映射成一个固定长度的特征向量,以实现图像级的分割任务^[49].基于 CNN 的分割方法为了对一个像素分类,通常使用该像素周围的一个图像块作为 CNN 的输入用于训练和预测.由于只能提取一些局部的特征,从而导致分割的性能受到限制.近年来,随着网络变得越来越深,CNN 的特征表达能力不断提升,例如 ResNet 和 DenseNet,但其仍然无法避免从固定输入到固定输出的学习本质.因此,传统基于 CNN 的图像分割方法采取的策略是使用预训练好的 CNN 网络模型作为编码器用于生成系列特征图,并结合额外的预处理或后处理步骤来完成特定的分割任务,具体细节请读者参考 CNN 源码链接(<https://github.com/weiaicunzai/pytorch-cifar100>).

3.2 全卷积网络

FCN^[16]是 CNN 的变种网络模型,其主要思想是采用卷积神经网络实现从图像像素到像素类别的变换,例如 VGG16、GoogLeNet.与 CNN 不同,FCN 网络使用卷积层代替全连接层,并通过转置卷积(transposed convolution)层将中间层特征图的高和宽变换回输入图像的尺寸,从而令预测结果与输入图像在空间维上一一对应.因此,在网络实现中,FCN 可以接受任意尺寸的输入图像,并生成相同尺寸的分割图.图 5 显示了 FCN 网络示例.

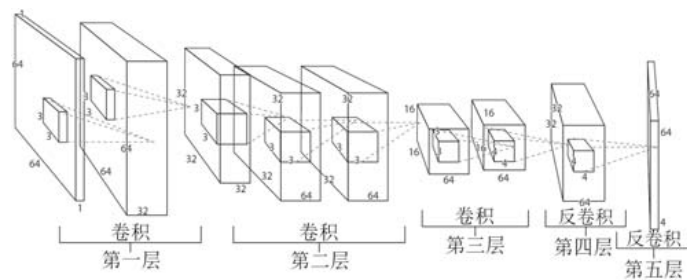


Fig.5 Architecture of FCN

图 5 FCN 体系结构

FCN 通过使用跳跃连接结构对网络模型最后层的特征图进行上采样,并与较浅层的特征图进行融合,将高层语义信息和底层图像信息组合起来,以生成精细的分割结果.目前,FCN 在 PASCAL VOC、NYUDv2 等多个测试数据集上均取得了较传统方法更优的分割性能.

FCN 为图像分割带来了里程碑式的影响,能够从抽象的特征中恢复出每个像素所属的类别,即从图像级别的分类进一步延伸到像素级别的分类,从而可实现端到端的模型训练和语义分割结果.然而,FCN 仍存在一些不足,包括:(1) 实时推理能力有待加强;(2) 没有充分考虑全局上下文信息;(3) 难以推广到三维(3D)及更高维图像.近年来,基于 FCN 网络模型,很多改进的方法被提出来广泛应用于语义分割中.例如,Liu 等人提出的 ParseNet 模型通过使用一层的平均特征增广每个位置上的特征,以实现补充全局语义信息的目的^[56].至于其他模型,请读者参考 FCN 源码链接(https://github.com/tangzhenyu/semanticsegmentation_dl).

3.3 编码器-解码器与自编码器模型

另一类用于图像语义分割的代表性网络是编码器-解码器模型^[57].编码器-解码器模型利用对称网络结构将数据点从输入域映射到输出域,其主要思想是:利用编码函数 $z=g(x)$ 所定义的编码器(encoder)将输入压缩成特征空间表示,再利用解码器(decoder) $y=f(z)$ 从特征空间表示中预测输出结果.其中,特征表示能够捕获输入潜在的语义信息,有助于预测输出.编码器-解码器模型是目前比较流行的图像到图像转换以及自然语言处理方法,如图 6 所示.这类模型通过最小化重构误差损失 $L(y, \hat{y})$ 进行逐层训练,可以实现端到端的学习.其中, y 为人工分割样本, \hat{y} 为预测的分割结果.

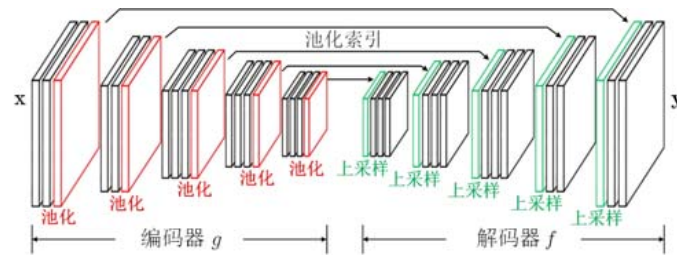


Fig.6 Architecture of a classical encoder-decoder model

图6 经典编码器-解码器模型体系结构

基于编码器-解码器的分割方法在编码器之后接上解码器,利用对称网络结构进行图像语义解析.例如,DeconvNet^[58]由VGG16网络和反卷积网络构成:前者起到编码器作用,用于提取图像特征;而后者使用特征向量作为反卷积层和上池化层的输入用于逐像素分类和预测分割掩膜.类似地,SegNet^[59]利用最大池化索引执行非线性上采样.与DeconvNet相比,该学习模型能够避免上采样的学习过程,因而参数规模较小.至于更多基于编码器-解码器的分割方法,请读者参考源码链接(<https://github.com/mrgloom/awesome-semantic-segmentation>).

自编码器属于编码器-解码器模型的一个特例,这是一种无监督的特征学习网络,其主要思想是:通过将输入作为学习目标,对其进行特征表示学习.近年来,基于自编码器,很多改进的模型被提出来广泛应用于数据分析、模式识别等领域.代表性模型包括栈式去噪自编码器(stacked denoising auto-encoder,简称SDAE)^[60]和变分自编码器(variational auto-encoder,简称VAE)^[61]:前者将自编码器进行堆叠用于图像去噪;后者将一个先验分布引入特征表示,由此从给定数据分布中生成真实样本.

3.4 循环神经网络

如图7所示,RNN及其变种网络^[62,63]可描述动态时间行为,显式地建模序列数据和多维数据中的复杂依赖关系,其主要思想是利用独特设计结构:“门”,将输入序列编码为隐藏状态,并通过更新隐藏状态记忆数据中的重要信息,实现对长期依赖关系的预测.

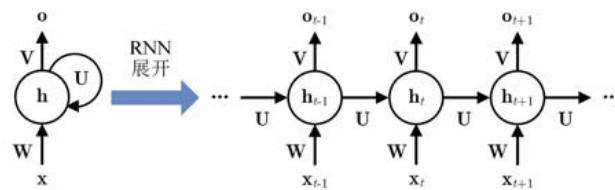
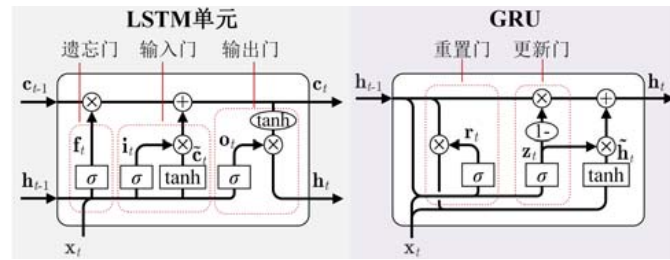


Fig.7 Schematic diagram of RNN

图7 RNN 示意图

而隐藏状态的更新策略包括3类:双曲正切(tanh)^[18]、长短期记忆(long short-term memory,简称LSTM)^[62]和门控循环单元(gated recurrent unit,简称GRU)^[63].其中,LSTM单元和GRU通过设计门内特殊交互机制调节信息流,有选择地丢弃和添加信息,从而保留重要特征,并保证其在长期传播过程中不会丢失,见图8.在训练过程中,门控机制也在一定程度上缓和了传统RNN的退化问题,有助于理解模型的决策依据.

基于RNN的分割方法通常采用RNN作为CNN模型的独立层用于建模图像局部和全局上下文信息以提取像素序列特征,从而可实现精准语义分割.门控RNN使用LSTM单元和GRU代替传统tanh模块,虽然提升了上下文特征学习的能力,但由于单元中序列输入的拓扑结构大多是预定义的^[64],只能学习固定的上下文特征,缺乏对结构化信息的考虑,因而当目标结构变化时,分割结果会有较大差异.为此,人们提出了若干改进方法,例如梁小丹等人结合图论改进输入拓扑以引导门控RNN自适应地学习节点间的语义相关性^[65,66].至于其他方法,可参考RNN源码链接(<https://github.com/element-research/rnn>).

Fig.8 Structures of LSTM^[62] and GRU^[63] modules图 8 LSTM 单元^[62]和 GRU^[63]模块结构

3.5 生成对抗网络

在各种不同的深度网络模型中,GAN 是一种简单有效的数据建模工具,在其整个训练过程中,都无需提前对数据做任何假设,因而受到了众多学者的青睐^[67].此种网络结构通常包含两个部分,即生成器(generator)和判别器(discriminator).如图 9 所示,判别器试图区分由生成器产生的伪数据和真实数据,而生成器则想要产生与真实数据更加相似的伪数据去扰乱判别器,此过程类似于博弈论中的二人零和博弈(zero-sum game).GAN 的训练实质是寻找零和博弈的一个纳什均衡解.最终,生成器能够产生与真实数据拥有相同分布的伪数据,也即学习到了真实数据的潜在分布.

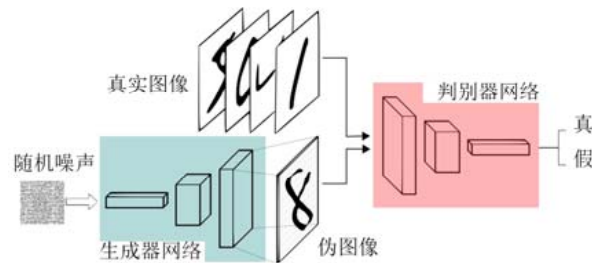


Fig.9 Network architecture of GAN

图 9 GAN 基本网络架构

近年来,基于 GAN 网络模型,很多改进的方法被提出来,广泛应用于图像到图像的转换(image-to-image translation)中.例如:Radford 等人采用反卷积层代替生成器网络中的全连接层对图像进行生成^[68];Arjovsky 等人将 Wasserstein 距离引入损失函数,由此估计真实样本分布和生成样本分布不重叠情况下的距离^[69].此外,研究人员发现:GAN 通过半监督方式添加图像类别标签进行训练,可极大提高生成样本的质量.基于此,Mirza 等人提出了一个带条件约束的 GAN 模型,称为条件 GAN(conditional GAN,简称 CGAN)^[70].至于其他相关工作,请读者参考 GAN 源码链接(<https://github.com/hindupuravinash/the-gan-zoo>).

3.6 硬件和软件

深度学习算法相关文献急剧增加的主要原因之一,来自于图形处理器(graphics processing unit,简称 GPU)和 GPU 计算库的可用性.GPU 是高度并行的计算引擎.与传统中央处理器(central processing unit,简称 CPU)相比,GPU 的运算速度要高出一个数量级甚至以上,使其对运行分析和深度学习算法尤其有用.

除了硬件,随着开源软件包在深度学习领域的普及应用,很多深度学习库被提出来,广泛应用于诸如卷积操作等高效 GPU 实现中.深度学习库提供了一系列深度学习组件.用户通过直接调用函数接口定义新的算法,无需从复杂的神经网络开始编写代码.代表性的深度学习库包括:

- (1) Tensorflow^[71]:该框架最初由谷歌大脑团队开发,用于 Google 的研究和生产,提供 C++接口和 Python 接口.

- (2) PyTorch^[72]:由 Facebook 人工智能研究院基于 Torch 研发,提供 Python 接口.
- (3) Caffe^[73]:由加州大学伯克利分校的贾扬清博士研发,提供 C++接口和 python 接口.
- (4) Theano^[74]:由蒙特利尔学习算法研究所开发,提供 Python 接口.

在硬件和开发环境方面,Keras 和 Lasagne(<https://keras.io/>,<https://github.com/Lasagne/Lasagne>)支持多操作系统下的多 GPU 并行计算,可以作为 Tensorflow,Theano 等框架的高阶应用程序接口,以进行深度学习模型的设计、调试、评估、应用和可视化.

4 基于深度学习的病理图像分割

病理图像分割是当前数字病理分析深度学习算法研究的主要应用领域.有限的特征表达能力,使得传统方法的分析性能遭受限制.相比之下,以人工神经网络为架构的深度学习算法对海量病理数据的特征提取能力更强.这得益于算法在自学习、非线性建模、优化解寻找等方面的提升^[47].因此,基于深度学习的病理图像量化分析研究既是数字病理的大势所趋,也是学术界和医学界共同努力的方向.

然而,一方面,不同成像原理的病理图像分割与计算机视觉领域中的自然图像分割存在较大差别;另一方面,由于肿瘤异质性、病理图像可视化特征难寻、数据集有限等限制,深度学习模型辅助病理诊断在临床上的应用能力还十分有限.至今为止,国内外学者主要针对明视野(bright field)、荧光(fluorescence)、相位差(phase contrast)、微分干涉差(differential interference contrast,简称 DIC)、共聚焦(confocal)以及电子(electron)等不同显微成像原理的病理分析任务开展了系列深度学习研究工作^[3].另外,借助苏木精-伊红和免疫组织化学等染色处理,研究者已在 CAMELYON17、Warwick-QU 和 TCGA 等多种数字病理图像集(camelyon17.grand-challenge.org/, warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/download/, nucleisegmentationbenchmark.weebly.com/, becklab.hms.harvard.edu/software/diademchallenge.org/data_sets.html, kimialab.uwaterloo.ca/kimia/index.php/data-and-code/, bioimage.ucsb.edu/research/bio-segmentation/mitos-atypia-14.grand-challenge.org/, andrewjanowczyk.com/deep-learning/, mega.nz/#F!BiZAELJa,tupac.tue-image.nl/node/3,biomisa.org/index.php/downloads/, brainiac2.mit.edu/SNEMI3D/)上进行了深度模型性能实验,并取得了较好的结果.表 2 列出了现有适用于检测、分割及分类任务的深度网络模型.

Table 2 Overview of existing models using representative deep networks for detection, segmentation, and classification in pathology images

表 2 现有适用于病理图像检测、分割及分类任务等代表性深度网络模型概述

相关文献	任务	染色方法或成像模式	方法
Ronneberger 等人 ^[57]	细胞分割	电子显微镜	结合 U-Net 及弹性变形的数据扩充方法
Song 等人 ^[75]	细胞核和细胞质分割	H&E	基于多尺度 CNN 和图划分算法
Xie 等人 ^[76]	细胞核检测	H&E	基于 CNN 的结构回归模型
Albarqouni 等人 ^[77]	有丝分裂检测	H&E	将众包(crowdsourcing)层引入 CNN 框架
Chen 等人 ^[78]	有丝分裂检测	H&E	深度回归网络模型
Sirinukunwattana 等人 ^[79]	细胞核检测及分类	H&E	基于空间约束回归的 CNN 框架
Xing 等人 ^[80]	细胞核分割	H&E、IHC	CNN 及稀疏形状建模方法
Xu 等人 ^[81]	细胞核检测	H&E	栈式稀疏自编码器
Yang 等人 ^[82]	神经胶质细胞分割	双光子显微镜	结合 FCN 及迭代 k-终端割算法
BenTaieb 等人 ^[83]	腺分割	H&E	基于多损失函数的 FCN 模型
Chen 等人 ^[84]	神经元膜分割	电子显微镜	结合 U-Net 及双向 LSTM 网络的分割模型
Xie 等人 ^[85]	肌束膜分割	H&E	空间时钟频率驱动的 RNN 网络模型
Xie 等人 ^[86]	细胞核检测	H&E	深度投票机制
Fakhry 等人 ^[87]	神经元分割	电子显微镜	结合 CNN 及分水岭算法模型
Stollenga 等人 ^[88]	神经元分割	电子显微镜	多维 LSTM 网络
Xu 等人 ^[89]	腺分割	H&E	结合深度监督网及整体嵌套边缘检测算法模型
Chen 等人 ^[90]	神经元分割	电子显微镜	深度上下文网络模型
Chen 等人 ^[91]	腺实例分割	H&E	深度轮廓感知网络模型
Song 等人 ^[92]	细胞分割	H&E	多尺度 CNN
Fakhry 等人 ^[93]	神经元分割	电子显微镜	残差反卷积网络模型

Table 2 Overview of existing models using representative deep networks for detection, segmentation, and classification in pathology images (Continued)**表 2** 现有适用于病理图像检测、分割及分类任务等代表性深度网络模型概述(续)

相关文献	任务	染色方法或成像模式	方法
Li 等人 ^[94]	HEp-2 样本图像分类	免疫荧光	全卷积残差网络
Li 等人 ^[95]	三维神经元重构	电子显微镜	三维全卷积网络
Zeng 等人 ^[96]	三维神经突分割	电子显微镜	将 Inception-残差模块引入深度 CNN 框架
Khoshdeli 等人 ^[97]	三维细胞核分割	共聚焦显微镜	多层编码器-解码器网络模型
Saha 等人 ^[98]	细胞膜和细胞核分割及分类	IHC、HER2 单抗	将梯形 LSTM 引入编码器-解码器模型
Senaras 等人 ^[99]	病理图像生成	IHC,Ki-67	条件 GAN 网络模型
Januszewski 等人 ^[100]	神经突追踪与重构	电子显微镜	泛洪网络
Graham 等人 ^[101]	细胞核实例分割	H&E	染质感知的 FCN 网络模型
Yan 等人 ^[46]	视网膜血管分割	检眼镜	联合逐像素和逐段损失优化的编码器-解码器
Mahmood 等人 ^[102]	细胞核分割	H&E	对抗回归网络框架
Yan 等人 ^[103]	视网膜血管分割	检眼镜	三阶段 FCN 网络模型
Xue 等人 ^[104]	细胞检测	H&E、IHC	结合 CNN 及压缩感知的端到端模型
Jin 等人 ^[105]	视网膜血管分割	检眼镜	将可变形卷积引入 U-Net 框架
Chakravarty 等人 ^[106]	细胞核分割、视盘、杯分割	H&E、检眼镜	水平集驱动的 RNN 网络模型
Wollmann 等人 ^[107]	细胞分割	荧光显微镜	结合 GRU 及稠密模块的分割模型
Yi 等人 ^[108]	神经细胞实例分割	视频显微镜	将注意力机制引入 FCN 框架
Naylor 等人 ^[109]	细胞核分割	H&E	深度距离图回归网络模型
Payer 等人 ^[110]	细胞实例分割及跟踪	H&E、视频显微镜	循环栈式沙漏网络模型
Shin 等人 ^[111]	微血管分割	血管造影、检眼镜	结合 CNN 及图神经网络模型
Qu 等人 ^[112]	细胞核和腺实例分割	H&E	将方差约束的交叉熵损失引入全分辨率卷积网
Xie 等人 ^[113]	腺实例分割	H&E	结合 SegNet 及校正网络模型
Luna 等人 ^[114]	粘连重叠细胞核分离	H&E	孪生神经网络
Koohbanani 等人 ^[115]	细胞核实例分割	H&E	空间感知深度网络模型
Mou 等人 ^[116]	管状结构分割	光学相干层析成像	通道与空间注意力网络模型
Fang 等人 ^[117]	血管中轴线提取	血管造影	结合 CNN 及拓扑优化算法模型
Kumar 等人 ^[40,41]	细胞核分割	H&E	基于 CNN 的三分类模型
Hashimoto 等人 ^[118]	癌症亚型分类	H&E	多尺度域对抗实例深度学习框架
Hayashida 等人 ^[119]	细胞跟踪	相位差显微镜	基于 CNN 的运动和位置图联合表示学习
Liu 等人 ^[120]	细胞核实例分割	H&E、荧光显微镜	循环一致性全景域自适应 Mask R-CNN 模型
Pandey ^[121]	细胞核分割	荧光显微镜	两阶段 GAN 网络模型
Yan 等人 ^[122]	腺实例分割	H&E	形状感知的对抗学习框架
Zhao 等人 ^[123]	淋巴结转移检测	H&E	结合变分自编码器、GAN 及图卷积网络模型

实例分割是一个挑战性问题,需要考虑在实现语义分割的同时保留实例空间几何信息,以对相同类别不同实例进行区分.大多数已有的实例分割学习算法局限于候选区域(region-proposal)网络的训练与测试,例如 Mask R-CNN^[124],或特定编码器—解码器结构,例如 U-Net^[57].虽然后者通过融入上下文信息能较好地避免病理图像复杂背景杂斑的干扰,但由于网络架构的影响,使其仍然对染质稀疏的目标内细节不够敏感.

目前,基于深度学习的病理图像实例分割研究主要集中在以下几种途径:(1) 改进网络结构以获取更多有意义的上下文信息,例如信息聚合网络模型^[125];(2) 将辅助输出层引入网络,由此建模实例轮廓及轮廓间交互关系,例如深度轮廓感知网络模型^[91];(3) 将经典建模理论和深层特征表示学习机制进行有机结合,使得分割性能得到更有效的提升^[126-129].因此,本节将从深度学习的视角,总结和分析目前代表性病理图像分割方法体系.

4.1 基于CNN的病理图像检测与分割方法

针对不同类型的数字病理图像,研究者们通过构建 CNN 网络模型进行逐层训练和预测,可输出一个概率图,其中,每个像素值表示该像素是种子的概率;然后,根据生成的概率图寻找局部最大值,从而实现组织基元的准确定位.目前,基于 CNN 分类,研究者提出了许多病理图像目标检测方法,其中,应用领域包括肾细胞癌^[34]、肺癌^[130]、乳腺癌^[77,80]、结肠直肠癌^[131]等.然而,一方面,基于 CNN 分类的方法为了对一个像素分类,通常使用该像素周围的一个图像块作为 CNN 的输入用于训练和预测,由于存储开销大且计算效率低下,从而导致检测的性能受到限制;另一方面,为解决错误分类问题,CNN 也尝试采用回归策略将空间拓扑引入网络模型^[79,86,104]以赋

值目标中心邻近像素更高的概率值,进而实现精确检测目的,例如深度回归网^[76,78,79]、深度投票网^[86](如图 10 所示)、局部敏感深度网^[79]、CNN 和压缩感知相结合的网络模型^[104](如图 11 所示)。

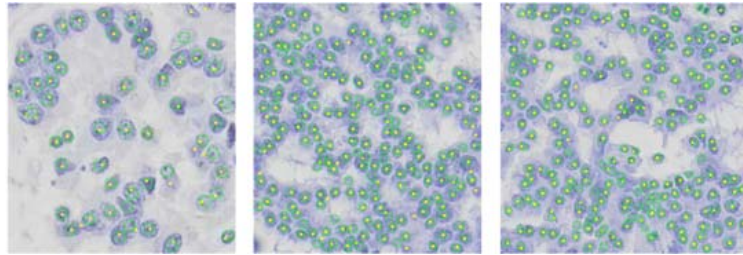


Fig.10 Nucleus detection using deep voting network

图 10 基于深度投票网络的细胞核检测结果

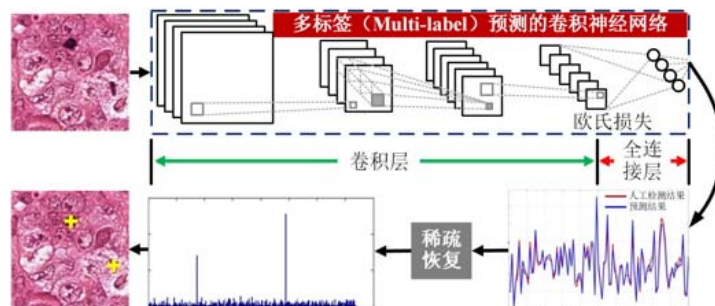


Fig.11 Cell detection model by combining CNN and compressed sensing

图 11 CNN 和压缩感知相结合的细胞检测模型

而在病理图像分割方面,基于 CNN 预测的分割模型也是应用较为成功的技术之一.CNN 深层特征表示学习有效提高了模型对图像噪声和数据非均匀的鲁棒性,为解决目标表现多样性的病理图像分割问题提供了有效途径.这类模型通常使用自动学习得到的表现信息来估计目标区域和非目标区域的概率图,并依靠特定的后处理来获得 ROI 中实例分割结果.后处理一般采取基于图像分析的方法,完成从底层视觉处理到中层视觉特征提取的过程.复杂背景病理图像实例分割是在成千上万个个体目标汇集的图像中逐个分割出具有相对完整边界的目标实例,而底层视觉处理到中层视觉特征提取可充分利用图像底层信息,较好地弥补了传统 CNN 在图像细节捕获能力上的缺陷.例如,杨林等人^[80]针对组织病理细胞核图像,构建了两个卷积-池化层对和两个全连接层构成的 CNN2 模型,实现了端到端的模型训练,然后根据输出结果应用水平集模型,通过最小化能量函数完成对测试图像重叠细胞核的分割.在此基础上,Kumar 等人^[40,41]提出了更深的 CNN3 模型(如图 12 所示),实现了更广泛病理图像的分割.此外,彭汉川等人^[87]进一步改进基于 CNN 网络的分割框架,并在改进的网络框架上应用分水岭合并树,进而提升神经元分割精度.

以上研究主要采用 CNN 逐像素预测作为初始结果,结合基于图像分析的后处理方法实现分割过程个体目标分离.由于这两个过程相互独立,从而导致这类方法可迁移性较差.当前研究进展表明,基于 CNN 的分割模型最关键的问题在于尝试通过端到端的模式将经典方法理论引入网络训练过程.其中,代表性的端到端模型包括泛洪网络^[100]、图网络^[111]以及其他新型网络模型.例如,Albarqouni 等人^[77]在 CNN 中引入了一种新型众包 (crowdsourcing) 层,通过反向传播方式,直接利用噪声标签实现了端到端的训练.该方法可获取不同标注者的可信度和偏差,并在不同组织病理数据集上取得最优结果.另外,针对三维神经突分割问题,Zeng 等人^[96]提出了一种新的基于 Inception-残差学习的 CNN 网络模型.研究表明:Inception 模块施加残差学习策略可以更加高效地学习多尺度上下文信息,从而可实现不同厚度目标边缘的准确预测.

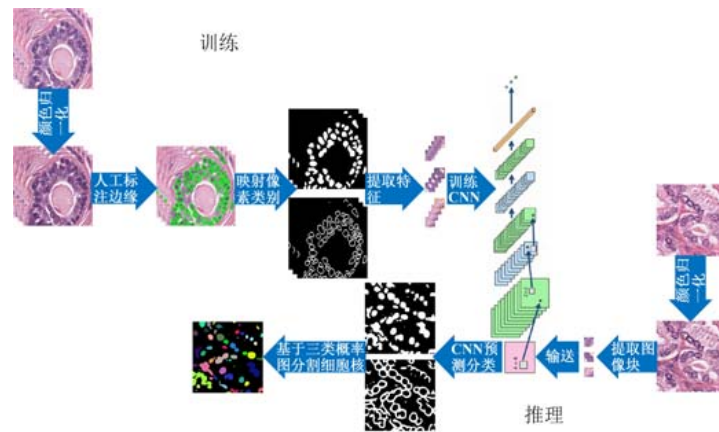


Fig.12 CNN3 model for nuclear segmentation

图 12 用于细胞核分割的 CNN3 模型

4.2 基于FCN的病理图像分割方法

基于 FCN 的病理图像分割方法通常使用像素级别的人工分割样本作为训练数据集,进而通过逐像素计算损失进行学习,例如交叉熵损失,虽然提升了传统 CNN 网络的分割性能,但由于受下采样的影响,使得上采样的特征图很难保留有意义的空间信息,例如用于分离重叠细胞的边缘信息^[3].另外,由于传统像素级损失函数同等对待图像中每个像素点,因而当大小、厚度不一致时,分割和追踪结果会有较大差异^[82,95],如图 13 所示.

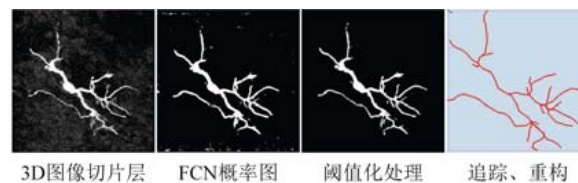


Fig.13 FCN-based 3D neuron segmentation and reconstruction

图 13 基于 FCN 的三维神经元分割与重构

已经有学者证明:FCN 除了改进网络结构和训练学习的方法,还可以通过定义不同的损失函数解决病理实例的分割问题^[91,101,103,108,109].损失函数通常作为学习准则与优化问题相联系,即通过优化算法有限次迭代网络参数,来尽可能降低损失函数的值.而基于反向传播算法的自动求梯度,极大地简化了优化算法的实现.考虑到边缘信息的空间特性,Chen 等人^[91]将一个轮廓数据拟合项引入交叉熵损失函数,由此,通过随机梯度下降建模腺体间交互关系,如图 14 所示;而 Naylor 等人^[109]则将分割任务建模为监督回归问题,由此通过自适应矩估计 (adaptive moment estimation,简称 Adam)解决重叠细胞核分割问题.考虑到病理细胞核内外染质的差异性,Graham 等人^[101]提出加权交叉熵损失函数,通过对核内染质较浅的像素及核外染质较深的像素均赋值更高的权值,实现病理细胞核的正确分割.考虑到血管厚度的不一致性,Yan 等人^[103]将血管骨架划分为若干片段,然后根据片段设计损失函数,通过最小化新的损失函数完成对粗、细血管的正确分割;最近,Yi 等人^[108]采用统一 FCN 框架(如图 15 所示)对检测和分割任务进行建模,通过定义并最小化检测损失函数和分割损失函数,实现对不同类型细胞的准确识别.类似于文献[108]的思想,何凯明等人^[124]在区域卷积神经网络(region-based CNN,简称 R-CNN)的基础上提出 Mask R-CNN.Mask R-CNN 通过使用 ROI 对齐(RoIAlign)层代替先前模型中的 ROI 池化层,以实现精确定位和分割目标区域的目的.这些基于损失函数变种的分割方法为提升 FCN 的分割精度提供了另一条路径,但仍然依赖于逐像素比较差异的机制,使其对空间几何信息约束的能力十分有限.此外,它们需要较大规模的训练数据集以经验拟合深层网络参数,且计算复杂度高.

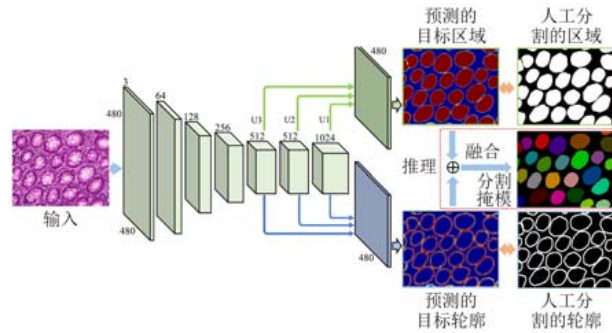


Fig.14 DCAN model for gland and nucleus segmentation

图 14 用于腺和细胞核分割的 DCAN 模型

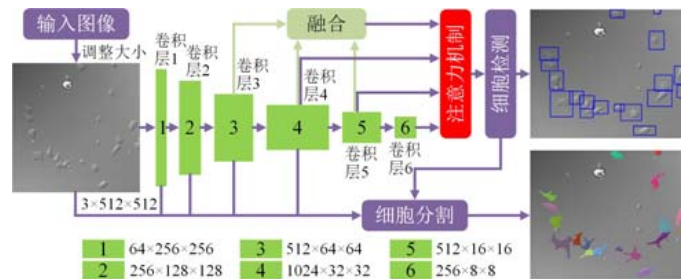


Fig.15 FCN model for neural cell detection and segmentation

图 15 用于神经细胞检测与分割的 FCN 模型

随着深度图像先验建模研究的发展,研究者充分意识到:需利用经典建模理论来学习病理图像实例的有效表示,并发展出一类损失泛函驱动的深度分割方法.其关键是建立经典建模理论和学习机制间的相关性,代表性方法为深度主动轮廓模型^[127-129].主动轮廓能够将更多的图像信息引入能量泛函以提高分割的精度,且优化理论基础坚实,为 FCN 网络的空间几何信息约束提供了新的思路.例如:文献[127]提出的端到端学习网络将 Chan-Vese 模型引入损失函数,由此建模图像中目标的空间细节;而在文献[128]中,作者使用总变差(total variation,简称 TV)代替偏微分方程(partial differential equation,简称 PDE),通过获取 Chan-Vese 损失泛函的全局最小值,实现了心脏磁共振图像的精准分割.相比之下,病理图像背景结构复杂,实例高密度分布、实例间出现重叠、缠绕,因此仅仅使用简单二相或多相模型来约束损失函数学习的能力十分有限.虽然 Snake 模型通过融入几何先验能较好地解决多目标实例分割问题^[129],但由于初始轮廓位置的影响,使其自动分割性能遭受大幅度的下降.此外,水平集函数在演化过程中需保持为符号距离函数,因此,周期性地重新初始化水平集函数不仅耗时,而且会增大数值误差.虽然李纯明等人^[132]通过在能量泛函中引入一个惩罚项避免了重新初始化的过程,但由于外部能量项的权重系数不能根据图像信息自适应地调整大小,因而限制了主动轮廓捕获多目标边界的能力,尤其是待分割目标存在重叠的情况.随着病理数据中更多挑战的引入,如何在深度网络框架下,设计并引入先验约束项以建模挑战,同时实现联合优化,仍需进一步深入研究.

综上,基于深度先验驱动的轮廓演化及推理在病理图像实例分割应用,尚有大量工作可做.

4.3 基于编码器-解码器模型的病理图像分割方法

目前,编码器-解码器模型在计算机视觉和人工智能领域得到了广泛应用.与传统 FCN 相比,编码器-解码器模型在病理图像分析的应用逐步引起广泛关注^[57,133].在现有的编码器-解码器模型中,U-Net^[57](见图 16)是病理图像分割中应用最为广泛的技术之一,其主要思想是:在收缩路径上捕获全局特征和在扩展路径上实现精确定位,从而较好解决了复杂神经元结构的分割问题.然而,U-Net 没有充分考虑像素间的局部依赖关系,使其易受目

标外部特征的影响,因此对图像细节特征捕获和紧致表达能力有待加强,尤其是待分割目标存在弱边缘和稀疏染质的情况.为此,沈琳琳等人^[94]利用全卷积网络架构对 ResNet 进行拓展,通过在更深的网络中解决梯度消失问题,以实现精确分割癌细胞的目的;进一步,Jégou 等人^[134]将 DenseNet 应用到语义分割,通过增强特征传播,以实现降低外部特征影响的目的.V-Net^[133]是另一种版本的 3D U-Net,与 U-Net 相比,该模型可以实现端到端的 3D 深度监督学习,并通过引入残差学习模块和 Dice 损失函数,避免了前背景像素类别数量不平衡的情况.

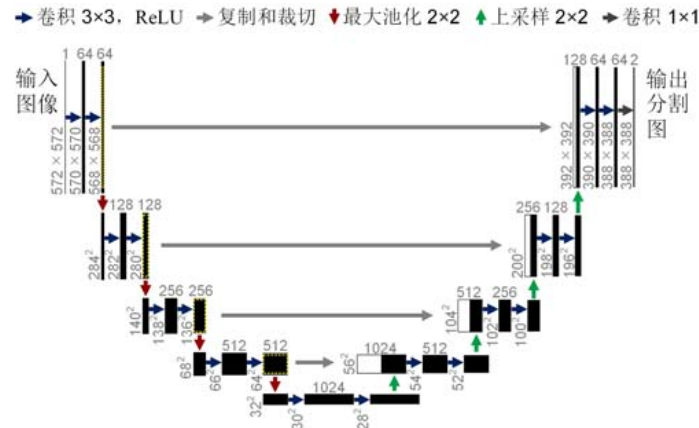


Fig.16 U-Net architecture for pathology image segmentation

图 16 用于病理图像分割的 U-Net 体系结构

除了主流的 U-Net, V-Net 及其变种网络,其他类型的编码器-解码器模型^[90,93,98,105]也已经应用到病理图像的语义分割.例如:Chen 等人^[90]和 Fakhry 等人^[93]针对大脑神经元图像,分别提出了深度上下文分割网络和残差反卷积分割网络(如图 17 所示);Saha 等人^[98]针对人表皮生长因子受体-2(HER2)乳腺癌细胞图像,提出了 Her2Net 分割网络;Jin 等人^[105]针对眼底视网膜血管图像,提出了可变形 U-Net 分割网络;Xie 等人^[113]针对结直肠癌图像,提出了深度分割校正网络.这些编码器-解码器模型通过解析训练图像的内容,在分割图像的同时获取图像所有分割区域甚至每个像素的语义类别,从而得到图像基于内容的标注,但仍缺少对病理图像底层信息的充分利用,导致最终的分割结果比较粗糙.编码器-解码器模型与注意力机制的复合是另外一条技术路线.与经典池化方法相比,注意力机制通过赋值权重,能够对不同位置特征的重要性进行评估,进而建模局部特征的上下文依赖关系.方法主要包括逆向注意力网络^[135]、金字塔注意力网络^[136]、对偶注意力网络^[137]等.受上述方法启发,研究者将复合模型应用于病理图像语义分割.代表性方法包括注意力神经网络分割模型^[108]、通道与空间注意力分割模型^[116]等.

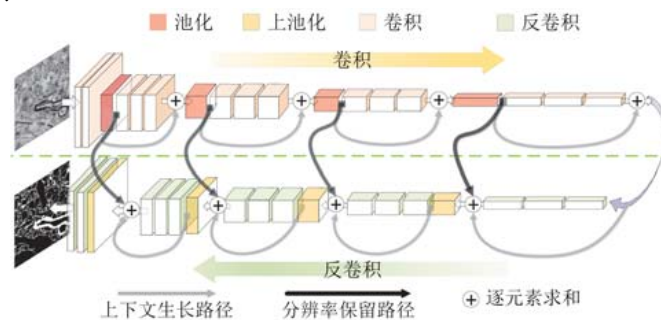


Fig.17 Residual decovolutional network model for neuron segmentation

图 17 用于神经元分割的残差反卷积网络模型

而对于大量无标签病理图像块的处理问题,许多研究者更倾向于使用自编码器模型.研究发现,可利用自编

码器模型学习性能增强的结构化语义特征提取器^[81,138,139].其中,文献[81]和文献[139]均提出了一种基于栈式稀疏自编码器(stacked sparse autoencoder,简称 SSAE)的框架,分别用于检测细胞核(如图 18 所示)和分类病理图像.后者利用自编码器去重建信号,然后将一个学习得到的线性分类映射添加到自编码器之后,用于分类.实验表明:在同等条件下,该模型在若干重建指标上优于压缩感知技术,且计算复杂度低,具有较好的实时性.

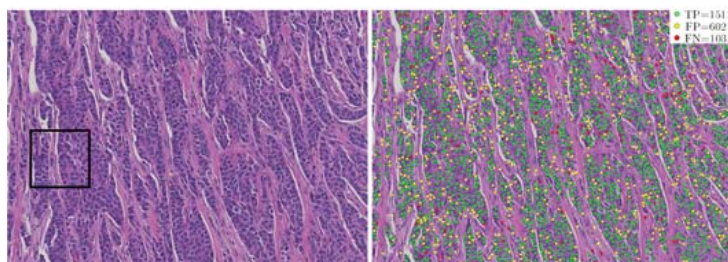


Fig.18 SSAE-based nuclear detection

图 18 基于栈式稀疏自编码器的细胞核检测

4.4 基于RNN的病理图像分割方法

CNN 和 FCN 没有充分考虑相似像素间长期的、全局的依赖关系,因此对分割标记的空间和表现一致性的捕获能力有待加强^[140].而在病理图像分割应用中,这种缺陷会导致分割结果目标失真、边界模糊、区域杂小等突出问题.RNN 网络结构和训练方式为解决上述分割问题提供了有效途径^[84,85,88,98,107,110].例如,Xie 等人^[85]首次将多维循环神经网络(multi-dimensional RNN,简称 MDRNN)应用于骨骼肌显微图像的肌束膜分割中.其主要思想是:在时空维度上,使用 RNN 的反馈连接学习显微数据内的关系.其中,每个循环单元当前时间步的隐状态由该时间步的像素输入和上一个时间步各方向上的像素隐状态决定,因而可获取图像中其他所有像素特征信息.为了在不影响性能的情况下大大降低计算复杂度,Stollenga 等人^[88]针对经典 RNN 网络的有向性,进一步改进 MDRNN 中反馈连接的拓扑结构,采用逐像素金字塔结构对信息进行传递,从而实现复杂神经元图像和三维脑图像的分割.

与 CNN 和 FCN 模型相比,RNN 在病理图像分割领域,研究者关注较少,但对于后续改进建模等依然有许多借鉴之处.一般而言,RNN 利用自身循环结构渐进地训练多个弱预测器,可迭代改进分割结果,增加对局部错误的鲁棒性,且模型参数规模较小.然而,在基于 RNN 的分割方法中,尤其是待分割目标存在重叠粘连的情况,如何同时捕捉空间区域内各元素强度的相关性以及它们各自随时间的演变规律,一直是亟待解决的核心难题之一.针对这一问题,研究者提出了系列 CNN 和 RNN 级联的病理图像分割模型^[84,106,107].其中,一类模型将 CNN 或 FCN 与 RNN 进行堆叠,这样在每个时间步中,上层卷积网络特征编码的输出为下层 RNN 序列建模的输入,例如 Chakravarty 等人提出的 RACE-Net 模型^[106](分割结果如图 19 所示)和 Wollmann 等人提出的 GRUU-Net 模型^[107](分割框架如图 20 所示).而在文献[84]中,作者采用深度双向 RNN(bi-directional RNN),通过捕获前后时间步的上下文信息,实现对三维神经元结构的准确分割.除了上述级联方法,其他类型的组合模型也已经应用到病理图像的分割^[140].这类模型通过在 RNN 体系结构中使用 CNN 代替循环单元,可将经典平均场迭代推理任务嵌入 RNN,并建模为 CNN 卷积计算问题,进而利用反馈连接实现精准语义分割.

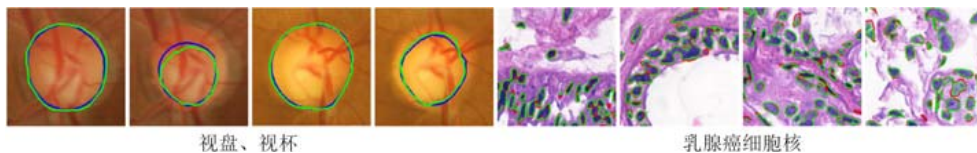


Fig.19 RACE-Net-based object contour inferences in pathology images

图 19 基于 RACE-Net 的病理图像目标轮廓推理

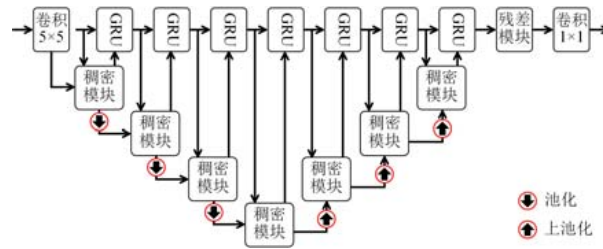


Fig.20 GRUU-Net architecture for cell segmentation

图 20 用于细胞分割的 GRUU-Net 模型结构

4.5 基于GAN的病理图像分割方法

自 2014 年 GAN 网络出现以来,关于其在病理学领域的应用研究文献不断涌现,其中包括图像生成^[99,141]、病灶检测^[123,142,143]、癌症亚型分类^[118]、腺和细胞分割^[102,121,122,144,145](如图 21 所示)等.在实际操作中,由于不需要预先建模,生成器和判别器可以选择为任何结构的神经网络,也未考虑建模数据的专有属性.此时,GAN 模型通常会过于自由、训练过程可控性太差、模型稳定性不足.因此,在利用 GAN 模型学习大规模病理数据的分布时,需要考虑病理数据自身特性,以增强模型及其训练过程的稳定性.一般而言,病理图像更倾向于使用 CGAN 作为基础模型用于图像生成^[99].至于“将 GAN 应用于病灶检测”,研究者则通过学习正常病理图像的概率分布检测异常,即具有不同分布的病理图像将被判别器认为是病灶.例如,Alex 等人^[142]训练了一种 GAN 模型,通过在生成器网络中建模正常病理图像块的分布和在判别器网络中计算测试图像块的后验分布,较好地解决了大脑损伤的检测问题.在此基础上,Chen 等人^[143]提出了对抗自编码器模型,实现了更广泛病理图像的异常检测.

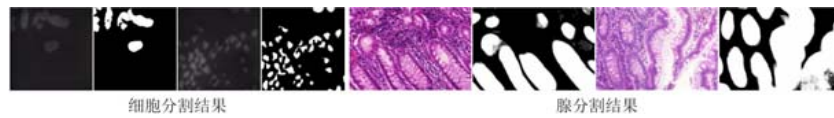


Fig.21 GAN model-based cell and gland segmentation

图 21 基于 GAN 模型的细胞和腺分割

正如第 3.1 节和第 3.2 节所述,CNN 和 FCN 通常使用像素级别的人工分割样本作为训练数据集,进而通过逐像素计算损失进行学习,例如交叉熵损失.为了得到正确的细胞实例分割结果,网络需要进一步结合后处理步骤来提升其对空间邻近性的度量能力,如 CNN-CRF 模型.但这类模型往往受限于成对势能的使用,由于只能利用数据的二阶统计特性,导致其分割的性能受到限制.相比之下,GAN 通过对抗训练将更高阶势能引入网络模型,从而更好地刻画细胞的复杂本质结构^[102,121,122].其中,判别器起到形状正则化作用,用于加强目标形状的紧致表达能力.另外,类似于感知损失^[146],通过使用对抗训练方式代替传统人工设计的损失函数,GAN 能够检测到预测分割和人工分割结果间高阶统计量的细微差异^[144].尤其随着 CGAN 的引入,对抗分割方法可通过增加感受野的大小学习基于上下文感知的结构化网络损失^[102],从而解决复杂背景结构下的细胞分割问题,如图 22 所示.

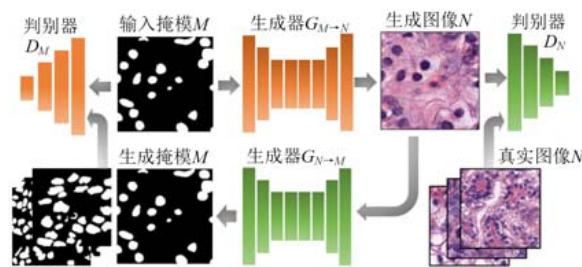


Fig.22 CGAN model for nuclear segmentation

图 22 用于细胞核分割的 CGAN 模型

5 评价指标

数字病理图像检测与分割算法性能的客观评价是算法应用于实际病理诊断的关键,检测与分割结果的好坏需要定性和定量的评价.有效的定量评价应该包含像素级和目标级两个方面,且评价标准需要惩罚以下 4 种错误.

- (1) Ground Truth(人工标注)目标的漏检测;
- (2) 伪目标的误检测;
- (3) 正确检测目标的欠分割;
- (4) 正确检测目标的过分割.

常用的目标检测度量指标包括敏感度(sensitivity,简称 Se)、特异性(specificity,简称 Sp)、准确率(accuracy,简称 Acc)、曲线下的面积(area under curve,简称 AUC)和 $F1$ 分数($F1$ score,简称 $F1$).不妨将 i 索引的 ground truth 目标记为 G_i , j 索引的分割(或检测)目标记为 S_j .通过计算真正(true positive,简称 TP)、真负(true negative,简称 TN)、假正(false positive,简称 FP)和假负(false negative,简称 FN),可定义 Se, Sp, Acc 以及 $F1$ 如下:

$$Se = \frac{TP}{TP + FN} \quad (1)$$

$$Sp = \frac{TN}{TN + FP} \quad (2)$$

$$Acc = \frac{TP + TN}{N} \quad (3)$$

$$F1 = \frac{2 \cdot PR \cdot RC}{PR + RC} = \frac{2TP}{2TP + FP + FN} \quad (4)$$

其中, TP 表示有对应 S_j 的所有 G_i 的个数; FP 表示无对应 G_i 的所有 S_j 的个数; FN 表示无对应 S_j 的所有 G_i 的个数, $N = TN + TP + FN + FP$; Se 表示正确分类的血管像素占真实像素的百分率; Sp 表示正确分类的非血管像素占真实非血管像素的百分率; Acc 表示正确分类血管和非血管像素占整个图像总像素的百分率.

虽然 $F1$ 能够有效评估目标间的关联程度,但它往往没有考虑像素级的错误.因此,为了计算 G_i 与 S_j 间的形状和尺寸一致性,研究者提出了系列像素级评价指标,包括 Jaccard 相似性系数(Jaccard index,简称 JI)^[147]、Dice 相似性系数(Dice similarity coefficient,简称 DSC)^[148]以及平均边缘距离(average boundary distance,简称 ABD)^[149].其中, ABD 是基于像素级边缘的评价方法,它表示 S_j 的边缘 S_j^b 与 G_i 边缘 G_i^b 间的平均距离.

$$ABD = \frac{1}{2C} \left(\frac{\sum_{u \in S_j^b} d(u, G_i^b)}{|S_j^b|} + \frac{\sum_{v \in G_i^b} d(v, S_j^b)}{|G_i^b|} \right) \quad (5)$$

其中, $d(u, G_i^b)$ 表示 S_j^b 上的像素 u 到 G_i^b 的最小距离, $d(v, S_j^b)$ 表示 G_i^b 上的像素 v 到 S_j^b 的最小距离, C 表示细胞(或细胞核)的个数. JI 和 DSC 是基于像素级面积的评价指标,通常被用于测量两个像素集合之间的重叠面积.其中,前者通常也称为交并比(intersection over union,简称 IoU).假设 G_i 的像素集合为 \mathcal{G}_i , S_j 的像素集合为 \mathcal{S}_j , 则 JI (或 IoU)和 DSC 可定义如下:

$$IoU = JI = \frac{|\mathcal{G}_i \cap \mathcal{S}_j|}{|\mathcal{G}_i \cup \mathcal{S}_j|}, DSC = \frac{2|\mathcal{G}_i \cap \mathcal{S}_j|}{|\mathcal{G}_i| + |\mathcal{S}_j|} \quad (6)$$

其中,两者的取值范围均为 0~1. $mIoU$ 是 IoU 的变种指标,表示所有类别的 IoU 均值,被广泛应用于现在的深度学习分割算法评估中. AJI (aggregated Jaccard index)是推广的 JI 指标,其通过最大化 JI ,将 ground truth 与检测结果进行匹配.这样, AJI 就对应于这些匹配成分的交并比的基数和的比率,并将所有不匹配的检测到的成分添加到分母中.与 IoU 相比, AJI 更加适合于评估细胞核分割结果,其定义如下:

$$AJI = \frac{\sum_{i=1}^L |\mathcal{G}_i \cap \mathcal{S}_k^*(i)|}{\sum_{i=1}^L |\mathcal{G}_i \cup \mathcal{S}_k^*(i)| + \sum_{l \in U} |\mathcal{S}_l|} \quad (7)$$

其中,分割得到的连通成分 $S_k^*(i)$ 利用 G_i 得到 \mathbf{JI} 最大值.

6 深度学习分割模型性能分析

本节首先在目前比较流行的分割基准公共数据集上对最新的深度学习算法的性能进行统计分析,然后,在此基础上对不同类型的数字病理图像的分割性能进行评估和总结.需要注意的是:虽然大多数已有模型基于标准分割指标均得到了综合性能评测,但是依然有一些模型未能做到,因此难以对它们全面比较.

表 3 总结了在 PASCAL VOC、Cityscapes 和 MS COCO 这 3 个自然图像数据集上代表性深度学习算法实现的定量分割结果.可以看出:FCN 的网络概念引入,大大提升了模型的分割精度,同时也衍生出系列语义分割和实例分割的编码器-解码器代表性体系.在 PASCAL VOC 数据集和 Cityscapes 数据集上,最新发展的深度分割模型的 mIoU 值分别超出原始 FCN 模型大约 27 个百分点和 20 个百分点,而在更具挑战性的 MS COCO 数据集上,虽然不同模型的分割性能遭受不同程度的下降,但与早期方法相比,最新的深度学习技术已经逐渐实现令人满意的分割性能.总之,网络结构的优化、网络层操作的改进以及经典理论的引入,显著提升了深度学习的表达和建模能力,降低了参数规模,减少了计算与内存资源,扩大了跨域迁移学习范围.更多的后续方法期待进一步算法开源和可复原研究,并在统一的数据集上进行严格的综合性能测试.

Table 3 Performance statistics of various deep learning algorithms for semantic segmentation and instance segmentation on several natural image datasets

表 3 不同深度学习算法在多个自然图像数据集上的语义分割和实例分割性能统计

PASCAL VOC 数据集			Cityscapes 数据集			MS COCO 数据集		
方法	基础网络	mIoU	方法	基础网络	mIoU	方法	基础网络	mIoU
FCN ^[16]	VGG-16	62.2	SegNet basic ^[59]	—	57.0	RefineNet ^[155]	ResNet-101	33.6
CRF-RNN ^[140]	—	72.0	FCN-8s ^[16]	—	65.3	CCN ^[181]	DenseNet-101	35.7
CRF-RNN* ^[140]	—	74.7	DPN ^[152]	—	66.8	DANet ^[177]	ResNet-50	37.9
BoxSup* ^[150]	—	75.1	Dilation10 ^[171]	—	67.1	DSSPN ^[182]	ResNet-101	37.3
Piecewise ^[151]	—	75.3	DeeplabV2 ^[153]	ResNet-101	70.4	EMANet ^[168]	ResNet-50	37.5
Piecewise* ^[151]	—	78.0	RefineNet ^[155]	ResNet-101	73.6	SGR ^[183]	ResNet-101	39.1
DPN ^[152]	—	74.1	FoveaNet ^[172]	ResNet-101	74.1	OCR ^[157]	ResNet-101	39.5
DPN* ^[152]	—	77.5	Ladder DenseNet ^[173]	DenseNet-169	73.7	DANet ^[177]	ResNet-101	39.7
DeepLab-CRF ^[153]	ResNet-101	79.7	GCN ^[154]	ResNet-101	76.9	EMANet ^[168]	ResNet-50	39.9
GCN* ^[154]	ResNet-152	82.2	Wide ResNet ^[156]	WideResNet-38	78.4	AC-Net ^[179]	ResNet-101	40.1
RefineNet ^[155]	ResNet-152	84.2	PSPNet ^[158]	ResNet-101	85.4	OCR ^[157]	HRNetV2-W48	40.5
Wide ResNet ^[156]	WideResNet-38	84.9	BiSeNet ^[174]	ResNet-101	78.9			
OCR ^[157]	ResNet-101	84.3	DFN ^[162]	ResNet-101	79.3			
PSPNet ^[158]	ResNet-101	85.4	PSANet ^[160]	ResNet-101	80.1			
DeeplabV3 ^[159]	ResNet-101	85.7	DenseASPP ^[175]	DenseNet-161	80.6			
PSANet ^[160]	ResNet-101	85.7	SPGNet ^[176]	2xResNet-50	81.1			
EncNet ^[161]	ResNet-101	85.9	DANet ^[177]	ResNet-101	81.5			
DFN ^[162]	ResNet-101	82.7	CCNet ^[178]	ResNet-101	81.4			
DFN* ^[162]	ResNet-101	86.2	DeeplabV3 ^[159]	ResNet-101	81.3			
Exfuse ^[163]	ResNet-101	86.2	DeeplabV3 ^[169]	Xception-71	82.1			
SDN ^[164]	DenseNet-161	83.5	AC-Net ^[179]	ResNet-101	82.3			
SDN* ^[164]	DenseNet-161	86.6	OCR ^[157]	ResNet-101	82.4			
DIS ^[165]	ResNet-101	86.8	GS-CNN ^[180]	WideResNet	82.8			
DM-Net ^[166]	ResNet-101	84.4	HRNetV2+OCR ^[157]	HRNetV2-W48	83.7			
DM-Net* ^[166]	ResNet-101	87.1						
APC-Net ^[167]	ResNet-101	84.2						
APC-Net* ^[167]	ResNet-101	87.1						
EMANet ^[168]	ResNet-101	87.7						
DeeplabV3+ ^[169]	Xception-71	87.8						
Exfuse ^[163]	ResNeXt-131	87.9						
MSCI ^[170]	ResNet-152	88.0						
EMANet ^[168]	ResNet-152	88.2						
DeeplabV3+* ^[169]	Xception-71	89.0						

表 4 列出了在 Kumar 细胞核数据集上,不同团队实现的深度学习分割性能,并总结了各团队在实现过程中所使用的建模和模型优化方法.在模型框架选择中,22 支团队采用 U-Net 及其变种网络,其中,9 支团队模型基于

原始 U-Net 框架,4 支团队模型基于 U-Net+VGGNet 框架,两支团队模型基于 U-Net+FCN 框架,两支团队模型基于 U-Net+ResNet 框架,一支团队模型基于 U-Net+DenseNet 框架,一支团队模型基于 U-Net+VGGNet+FCN 框架,3 支团队模型基于 U-Net+ResNet+Mask R-CNN 框架;而剩余团队则分别采用 Mask R-CNN、CNN 以及定制网络作为基础网络模型,其中,定制网络包括深层聚合、TernausNet 和 PANet.另外,模型优化根据预测的期望输出确定损失函数.其中,28 支团队利用逐像素交叉熵损失预测二元或三元概率图,12 支团队分别使用 Dice 及其变种损失(包括 JI 损失和 IoU 损失).为了提升网络逐像素回归能力,12 支团队又分别引入 L_1 损失和 L_2 损失.通过端到端的模型训练,大多数团队在不同器官病理图像和整个数据集上均取得了较好的分割结果.与传统网络相比,这些网络或损失的集成有利于互补信息,进而提升细胞核图像分割算法性能.因此,即使当图像出现核内染质稀疏、背景杂斑干扰和细胞核重叠等情形时,它们也仍然有效.

Table 4 Segmentation performance comparison of deep learning techniques
on Kumar nucleus dataset

表 4 深度学习技术在 Kumar 细胞核数据集上的分割性能比较

相关团队	模型及基础网络								
	U-Net	Mask R-CNN	CNN	FCN	PANet	ResNet	VGGNet	DenseNet	距离图
CUHK&IMSIGHT	✓	—	—	✓	—	—	—	—	—
BUPT.J.LI	—	—	—	—	—	—	—	—	—
pku.hzq	✓	✓	—	—	—	✓	—	—	—
Yunzhi	✓	—	—	—	—	✓	—	—	—
Navid Alemi	—	—	—	—	—	—	—	—	—
xuhuaren	—	✓	—	—	—	✓	—	—	—
aetherAI	—	✓	—	—	✓	✓	—	✓	—
Shuang Yang	—	✓	—	—	✓	✓	—	—	—
Bio-totem&SYSUCC	—	✓	—	—	—	✓	—	—	—
Amirreza Mahbod	✓	—	—	—	—	—	—	—	✓
CMU-UIUC	✓	—	—	—	—	✓	—	—	—
Graham&Vu	—	✓	—	—	—	✓	—	✓	✓
Unblockabulls	✓	—	—	—	—	—	—	—	—
Tencent AI Lab	✓	✓	—	—	—	✓	—	—	—
DeepMD	✓	—	—	—	—	—	✓	—	—
Antanas Kascenas	✓	—	—	—	—	—	✓	—	✓
Johannes Stegmaier	✓	—	—	—	—	—	—	—	✓
Yanping	✓	✓	—	—	—	✓	—	—	—
Philipp Gruening	✓	—	—	—	—	—	✓	—	—
Agilent Labs	✓	—	—	—	—	—	—	—	✓
Konica Minolta Lab EU	✓	—	—	—	—	—	✓	—	—
OnePiece	✓	—	—	—	—	—	—	—	—
Junma	—	—	—	—	—	—	—	—	—
Biosciences R&D, TCS	✓	—	—	—	—	—	—	—	—
Azam Khan	—	✓	—	—	—	✓	—	—	—
CVBLab	✓	—	—	—	—	—	—	—	—
Linmin Pei	✓	—	—	—	—	—	—	—	—
Peter Naylor	✓	—	—	✓	—	—	—	—	—
Kumar	—	—	✓	—	—	—	—	—	—
DB-KR-JU	—	—	—	—	—	—	—	—	—
VISILAB	—	—	—	—	—	—	✓	—	—
Sabarinathan	✓	—	—	—	—	—	—	—	—
Fuyong Xing	—	—	✓	—	—	—	—	—	—
Silvers	✓	—	—	—	—	—	—	✓	—
TJ	✓	—	—	✓	—	—	✓	—	—

Table 4 Segmentation performance comparison of deep learning techniques on Kumar nucleus dataset (Continued)

表 4 深度学习技术在 Kumar 细胞核数据集上的分割性能比较(续)

相关团队	损失函数				乳腺	肾	前列腺	膀胱	结肠	肺	整体
	交叉熵	Dice 损失	L ₁ 损失	L ₂ 损失	AJI	AJI	AJI	AJI	AJI	AJI	AJI
CUHK&IMSIGHT	✓	✓	—	—	0.646	0.707	0.706	0.731	0.641	0.696	0.691
BUPT.J.LI	✓	—	—	✓	0.644	0.701	0.705	0.723	0.615	0.696	0.687
pku.hzq	✓	—	✓	—	0.646	0.698	0.702	0.708	0.611	0.696	0.685
Yunzhi	✓	—	—	—	0.632	0.694	0.704	0.695	0.619	0.695	0.679
Navid Alemi	✓	✓	—	—	0.621	0.692	0.703	0.712	0.612	0.684	0.678
xuhuaren	✓	—	✓	—	0.608	0.680	0.696	0.677	0.606	0.688	0.664
aetherAI	✓	—	✓	—	0.628	0.677	0.660	0.707	0.602	0.672	0.663
Shuang Yang	✓	✓	—	—	0.624	0.680	0.672	0.688	0.583	0.672	0.662
Bio-totem&SYSUCC	✓	—	✓	—	0.617	0.677	0.684	0.686	0.592	0.679	0.662
Amirreza Mahbod	✓	✓	—	—	0.628	0.673	0.680	0.687	0.601	0.666	0.657
CMU-UIUC	✓	✓	—	—	0.612	0.669	0.666	0.694	0.580	0.685	0.656
Graham&Vu	✓	—	✓	✓	0.624	0.674	0.652	0.639	0.580	0.682	0.653
Unblockabulls	✓	—	—	—	0.594	0.663	0.665	0.714	0.557	0.648	0.651
Tencent AI Lab	✓	—	✓	—	0.596	0.668	0.659	0.684	0.574	0.651	0.646
DeepMD	✓	✓	—	—	0.576	0.643	0.679	0.682	0.602	0.651	0.633
Antanas Kascenas	✓	—	—	—	0.627	0.653	0.664	0.704	0.577	0.674	0.633
Johannes Stegmaier	✓	—	—	—	0.547	0.646	0.665	0.664	0.610	0.659	0.623
Yanping	—	—	—	—	0.586	0.625	0.644	0.680	0.577	0.633	0.623
Philipp Gruening	✓	—	—	—	0.570	0.644	0.612	0.683	0.540	0.614	0.621
Agilent Labs	—	—	—	✓	0.570	0.650	0.655	0.667	0.560	0.647	0.618
Konica Minolta Lab EU	—	✓	—	—	0.569	0.623	0.615	0.583	0.552	0.657	0.611
OnePiece	✓	—	—	—	0.538	0.617	0.635	0.673	0.551	0.604	0.606
Junma	—	✓	—	—	0.541	0.602	0.601	0.629	0.516	0.595	0.593
Biosciences R&D, TCS	—	—	—	—	0.565	0.602	0.626	0.685	0.541	0.629	0.578
Azam Khan	✓	—	✓	—	0.514	0.610	0.614	0.606	0.546	0.515	0.575
CVBLab	—	✓	—	—	0.526	0.562	0.612	0.640	0.523	0.570	0.574
Linmin Pei	—	—	—	✓	0.551	0.517	0.592	0.608	0.565	0.582	0.562
Peter Naylor	✓	—	—	✓	0.561	0.553	0.628	0.597	0.436	—	0.559 8
Kumar	✓	—	—	—	0.539	0.573	0.434	0.522	0.529	—	0.508
DB-KR-JU	✓	—	—	—	0.486	0.475	0.302	0.507	0.443	0.479	0.455
VISILAB	✓	—	—	—	0.422	0.406	0.419	0.378	0.468	0.553	0.444
Sabarinathan	✓	✓	—	—	0.331	0.462	0.467	0.397	0.424	0.553	0.444
Fuyong Xing	✓	—	—	—	0.426	0.407	0.456	0.319	0.308	—	0.348
Silvers	✓	—	—	—	0.445	0.136	0.171	0.394	0.073	0.368	0.278
TJ	✓	—	—	—	0.111	0.096	0.089	0.122	0.077	0.284	0.130

为了更好地启发读者了解和研究病理图像深度分割方法,下面将详细介绍排名前五的方法.至于更多算法的参数设置细节,请读者参考 MoNuSeg 稿件:<https://monuseg.grand-challenge.org/Manuscripts/>.

- (1) CUHK&IMSIGHT:将细胞核分割任务划分为细胞核分割和边缘分割.受 FCN 和 U-Net 的启发,提出了轮廓信息聚合网络(CIA-Net)模型,并基于 ResNet-50 基础网络,实现了细胞核和细胞核边缘的联合分割.其中,模型训练过程通过结合细胞核及其边缘的标注误差引导交叉熵损失函数进行逐层训练,实现了端到端的学习.该算法对于一部分细胞核目标依然存在误检测和过分割问题.
- (2) BUPT.J.LI:将深层聚合网络应用于细胞核分割任务,包括检测核内像素、估计核内像素的几何中心以及为每个核内像素估计一个指向几何中心的中心向量.该模型能够处理细胞核重叠或接触等突出问题,但由于没有显式地建模细胞核边缘信息,因而易受高曲率边缘的影响.
- (3) pku.hzq:通过使用 U-Net 预测三元概率图,并结合 Mask R-CNN 实现自顶向下的细胞核实例分割.该模型依靠额外的形态学腐蚀操作和随机游走算法,完成从预测结果到实例分割结果的转化,因而对重叠细胞核间的边缘细节不够敏感.
- (4) Yunzhi:通过级联两个 U-Net 实现细胞核实例分割结果.前一个 U-Net 预测核内像素及到核中心的单位向量,后一个 U-Net 在此基础上预测细胞核边缘.分割结果依赖于指向核中心单位向量的计算,因此

容易产生过分割和欠分割现象.

- (5) Navid Alemi:将多尺度特征共享网络与跳跃连接模式相结合,提出基于 SpaghettiNet 的细胞核分割模型.训练过程使用 Dice 和交叉熵损失加权组合,以实现精确检测和分割的目的.由于该模型在预测结果上应用分水岭变换,得到最终的实例分割结果,因而易产生过分割现象.

针对 Warwick-QU 腺数据集,代表性深度学习分割模型主要包括 CUMedVision1、CUMedVision2、CVML、ExB、Freiburg 和 vision4GlaS,分别采用 FCN 与腺区域掩膜多级特征表示相结合的方法、深度轮廓感知网络模型、两个卷积-池化层对和两个全连接层构成的 CNN 模型、多路卷积神经网络模型、U-Net 网络模型以及 CNN 与主动轮廓相结合的方法.而 ExB 和 Freiburg 根据后处理的使用进一步分类为 ExB1、ExB2、ExB3 和 Freiburg1、Freiburg2.

在实现过程中,以上方法主要采用两种分割策略:(1) 识别和分组腺目标像素;(2) 对腺候选目标做有无腺体的分类.针对前者:CVML、ExB 和 vision4GlaS 分别构建了一个 CNN 分类器,并使用像素周围的一个图像块作为 CNN 分类器的输入用于训练和预测;而 CUMedVision 和 Freiburg 则均基于 FCN 网络框架从抽象的特征中恢复出每个像素所属的类别,完成像素级的预测.为了解决腺实例分割问题,CVML 和 vision4GlaS 将轮廓信息推理引入端到端的网络学习过程,由此建模腺实例轮廓间的交互关系;ExB 通过额外的网络训练获取实例分割结果;而 CUMedVision 和 Freiburg 则显式地定义基于边缘的损失函数项.另外,针对后者,LIB 首先识别出部分腺目标(例如管腔或上皮细胞核),然后对部分目标进行分类和分割.需要注意的是:ExB、LIB 和 vision4GlaS 分割方法均引入腺的组织学分级,例如 ExB 所使用的边缘检测参数值取决于图像的组织学分级预测值、vision4GlaS 的像素分类取决于组织学分级信息等.

模型分割性能的排名综合考虑了两组评估标准:(1) glas 挑战赛评估标准,即 $F1$ 分数+Dice 相似性系数+Hausdorff 距离;(2) 分割精度评估标准,即 $F1$ 分数+调整兰德系数+Hausdorff 距离.通过 Warwick-QU 多组腺病理图像的分割实验结果表明:CUMedVision2、ExB1 和 ExB3 在多个指标上均优于其他分割方法,可以更好地避免过分割和欠分割,对于腺图像分割具有较好应用前景.此外,通过结合深度轮廓感知网络模型的使用,CUMedVision2 在良恶性腺分割上均超出 CUMedVision1.而 ExB2 和 ExB3 相比,ExB2 通过分离边界,也较好地解决了恶性腺的分割问题.

深度学习模型在视网膜血管分割方面也得到了广泛应用.与传统方法相比,深度学习在 DRIVE、STARE 和 CHASE_DB 这 3 个数据集上对敏感度、特异性、准确率和 AUC 这 4 个指标值的改进明显,见表 5.表中列举了近 5 年最具代表性深度学习方法的分割性能.在视网膜血管分割算法中,有监督算法的性能要普遍优于无监督算法.另外,对比得出 Soomro 等人、Rodrigues 等人 and Wu 等人的算法在保证高准确率的同时,具有较好的敏感度.这意味着分割结果能够尽可能地保留血管信息,分割得到的血管连续完整.其中,Soomro 等人将步长卷积引入编码器-解码器网络结构中,由此提升分割效率;Rodrigues 等人提出了多模态网络框架;Wu 等人通过设计网络中网络,实现对血管的精确分割.

Table 5 Segmentation performance comparison of deep learning algorithms on DRIVE, STARE, and CHASE_DB retinal vessel datasets

表 5 深度学习算法在 DRIVE、STARE 和 CHASE_DB 视网膜血管数据集上的分割性能比较

方法	年份	DRIVE 数据集				STARE 数据集				CHASE_DB 数据集			
		敏感度	特异性	准确率	AUC 值	敏感度	特异性	准确率	AUC 值	敏感度	特异性	准确率	AUC 值
Li 等人 ^[184]	2016	0.756 9	0.981 6	0.952 7	0.973 8	0.772 6	0.984 4	0.962 8	0.987 9	0.750 7	0.979 3	0.958 1	0.971 6
Fu 等人 ^[185]	2016	0.760 3	—	0.952 3	—	0.741 2	—	0.958 5	—	0.713 0	—	0.948 9	—
Liskowski 等人 ^[186]	2016	0.781 1	0.980 7	0.953 5	0.979 0	—	—	0.974 0	0.994 3	—	—	0.962 8	0.982 3
Orlando 等人 ^[187]	2017	0.789 7	0.968 4	—	—	0.768 0	0.973 8	—	—	0.727 7	0.971 5	—	—
Dasgupta 等人 ^[188]	2017	0.769 1	0.980 1	0.953 3	0.974 4	—	—	—	—	—	—	—	—
Yan 等人 ^[146]	2018	0.765 3	0.981 8	0.954 2	0.975 2	0.758 1	0.984 6	0.961 2	0.980 1	0.763 3	0.980 9	0.961 0	0.978 1
Srinidhi 等人 ^[189]	2018	0.864 4	0.966 7	0.958 9	0.970 1	0.832 5	0.974 6	0.950 2	0.967 1	0.829 7	0.966 3	0.947 4	0.959 1
Alom 等人 ^[190]	2018	0.779 2	0.981 3	0.955 6	0.978 4	—	—	—	—	0.775 6	0.982 0	0.963 4	0.981 5

Table 5 Segmentation performance comparison of deep learning algorithms on DRIVE, STARE, and CHASE_DB retinal vessel datasets (Continued)

表 5 深度学习算法在 DRIVE、STARE 和 CHASE_DB 视网膜血管数据集上的分割性能比较(续)

方法	年份	DRIVE 数据集				STARE 数据集				CHASE_DB 数据集			
		敏感度	特异性	准确率	AUC 值	敏感度	特异性	准确率	AUC 值	敏感度	特异性	准确率	AUC 值
Wu 等人 ^[191]	2018	0.784 4	0.980 7	0.956 7	0.981 9	—	—	—	—	0.753 8	0.984 7	0.963 7	0.982 5
Yan 等人 ^[103]	2018	0.763 1	0.982 0	0.953 8	0.975 0	0.773 5	0.985 7	0.963 8	0.983 3	0.764 1	0.980 6	0.960 7	0.977 6
Jin 等人 ^[105]	2019	0.796 3	0.980 0	0.956 6	0.980 2	0.759 5	0.987 8	0.964 1	0.983 2	0.815 5	0.975 2	0.961 0	0.980 4
Soomro 等人 ^[192]	2019	0.870 0	0.985 1	0.956 1	—	0.848 0	0.986 0	0.968 0	—	0.886 1	0.982 2	0.976 1	—
Adeyinka 等人 ^[193]	2019	0.760 3	—	0.952 3	—	0.741 2	—	0.958 5	—	0.713 0	—	0.948 9	—
Shin 等人 ^[111]	2019	0.925 5	0.938 2	0.927 1	0.980 2	0.935 2	0.959 8	0.937 8	0.987 7	0.936 4	0.946 3	0.937 3	0.983 0
Wu 等人 ^[194]	2019	0.803 8	0.980 2	0.957 8	0.982 1	—	—	—	—	0.813 2	0.981 4	0.966 1	0.986 0
Guo 等人 ^[195]	2019	0.789 1	0.980 4	0.956 1	0.980 6	—	—	—	—	0.788 8	0.980 1	0.962 7	0.984 0
Ma 等人 ^[196]	2019	0.791 6	0.981 1	0.957 0	0.981 0	—	—	—	—	—	—	—	—
Wang 等人 ^[197]	2019	0.794 0	0.981 6	0.956 7	0.977 2	—	—	—	—	0.807 4	0.982 1	0.966 1	0.981 2
Wang 等人 ^[198]	2020	0.799 1	0.981 3	0.958 1	0.982 3	0.818 6	0.984 4	0.967 3	0.988 1	0.823 9	0.981 3	0.967 0	0.987 1
Rodrigues 等人 ^[199]	2020	0.898 3	0.981 6	0.974 0	0.993 6	0.942 6	0.986 2	0.982 7	0.994 6	0.878 2	0.985 2	0.977 8	0.992 3
Wu 等人 ^[200]	2020	0.799 6	0.981 3	0.958 2	0.983 0	0.796 3	0.986 3	0.967 2	0.987 5	0.800 3	0.988 0	0.968 8	0.989 4

7 研究方向与开放问题

(1) 建立更具挑战性的病理图像数据集

病理图像语义分割及实例分割领域已经建立了带有人工标注的训练样本集,如 Kumar^[40,41]、Warwick-QU^[6],但仍然需要更具挑战性的数据集.一般而言,不同显微成像和染色处理的病理数据仅能反映组织或器官特定的解剖和功能信息,且各有优缺点,而当前,临床研究和应用通常需要多种成像原理和模态的病理图像的综合分析.具体地,病理诊断报告是病理学中的重要临床数据.除了肉眼观察,病理专家还将活检组织的观察结果以自然语言形式记录在病理报告中,例如细胞的位置、数量等.因此,大规模地获取多样化和多模态的可训练病理数据集将非常有意.另外,随着 3D 图像分割的普及应用,生物医学研究也迫切需要更具挑战性的 3D 病理图像数据集.

(2) 模型的可解释性

深度学习技术虽然凭借其强大的特征学习、设计灵活等特性在数字病理量化分析领域取得了突破性成果,但是其主要问题是模型的决策过程难以理解,例如深度模型学习什么?如何解释模型学习得到的特征?在给定数据集上实现特定分割精度所需最优网络架构等.相对于大规模病理数据集的建立,深度模型决策过程的理解对于数字病理的发展更为重要,为病理学领域的新发现提供契机.至今为止,一些研究已经尝试提供解决方案,包括隐层分析法^[201,202]、模拟模型方法^[203,204]、注意力机制^[205]等,但这些方法本质上仍是通过隐层特征的可视化与分析获取某种程度的局部可解释性.因此,建立融合可解释性技术与模型的病理图像处理与分析新方法,是值得进一步研究的方向.

(3) 多模信息融合的语义增强

深度模型的发展促进了多模信息融合与语义增强研究的进展,建立了图像视觉特征提取和文字处理间的连接,使得深度模型对图像的理解从简单的独立标签拓展至连贯的文字描述,进而通过多模态特征交互赋予计算机对复杂病理图像细节更强的感知能力.多模态数据融合应用时,深度网络对不同模态数据共享特征和互补性特征的表达能力需要提高.而在数字病理领域,如何将这解决方案嵌入到新型深度网络的设计中,挖掘多模态病理数据中的关键特征,自适应编码特征间的多模态交互,提升深度学习技术对病理学上下文特征的提取能力,是亟待解决的一个关键科学问题.

(4) 小样本学习问题

应用弱监督(weakly-supervised)甚至无监督学习(unsupervised learning)解决图像分割任务,是计算机视觉领

域的发展趋势.迁移学习和微调^[206]是其中一个代表性技术,其主要思想是:使用预训练模型作为监督学习网络的初始值,然后利用任务相关的有限标注数据集,以监督方式对网络参数进行微调,调整网络中的若干或全部层.例如,Bayramoglu 等人^[207]通过在不同深度的 CNN 模型上使用迁移学习,探究预训练网络参数对细胞核分类性能的影响.自监督学习(self-supervised learning)^[208]是另外一条技术路线,这类模型通常使用更小规模的分割样本作为网络的输入用于训练和预测,可以从无标注数据中学习得到有意义的图像信息.例如,Zhang 等人^[209]将自监督学习机制引入 CNN,由此精细识别人体部位.综上,深度学习对较大规模训练数据集的依赖性,与数字病理图像分割任务中的小样本问题相矛盾.如何在特征挖掘过程中完成对有意义图像信息的自行建模和推理,降低小样本问题对分割的影响,是数字病理图像分割的重要内容.

(5) 提升高通量检测与鲁棒性分割能力

对于高通量显微镜产生的海量病理数据,传统肉眼检查和人工方法不仅耗时耗力,而且存在主观差异性,例如从成千上万个体细胞汇集的图像中逐个分割出具有相对完整边界的细胞,由此派生出许多自动化检测与分割工具,例如 CellProfiler^[210]、Neuromantic^[211]等.一方面,这些工具通过组合预定义模块实现不同任务,例如细胞计数,但其整体性能易受参数规模的影响,尤其是当多个模块均执行复杂操作的情况.因此,如何实现分割性能与计算代价的最优折中是一个挑战性问题.另一方面,分割结果为后续形态计算和病理特征提供定量分析,而分割过程引入的错误会严重影响计数、识别等更高级分析任务的精度.因此,提升高通量检测与鲁棒性分割能力是当前数字病理应用中提高精准预测和应用能力的迫切需求.

8 结 论

本文总结了 100 多种基于深度学习算法的最新病理图像检测与分割模型,这些模型在不同类型病理图像分割任务和基准测试中均取得了突破性成果.目前的方法体系主要分 5 类:卷积神经网络、全卷积网络、编码器-解码器模型、循环神经网络以及生成对抗网络.在此基础上,梳理了现有适用于检测、分割及分类任务的公开病理数据集,以及不同方法的方法理论、关键技术、优缺点和性能分析.最后讨论了未来几年病理图像分割领域所面临的一些开放挑战和潜在研究方向.

References:

- [1] Weinstein RS, Graham AR, Richter LC, Barker GP, Krupinski EA, Lopez AM, Erps KA, Bhattacharyya AK, Yagi Y, Gilbertson JR. Overview of telepathology, virtual microscopy, and whole slide imaging: Prospects for the future. *Human Pathology*, 2009, 40(8):1057–1069.
- [2] Xing FY, Yang L. Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: A comprehensive review. *IEEE Reviews in Biomedical Engineering*, 2016, 9:234–263.
- [3] Xing FY, Xie YP, Su H, Liu FJ, Yang L. Deep learning in microscopy image analysis: A survey. *IEEE Trans. on Neural Networks and Learning Systems*, 2018, 29(10):4550–4568.
- [4] Ulman V, Maška M, Magnusson KEG, Ronneberger O, Haubold C, Harder N, Matula P, Matula P, Svoboda D, Radojevic M, Smal I, Rohr K, Jaldén J, Blau HM, Dzyubachyk O, Lelieveldt B, Xiao PD, Li YX, Cho SY, Dufour AC, Olivo-Marin JC, Reyes-Aldasoro CC, Solis-Lemus JA, Bensch R, Brox T, Stegmaier J, Mikut R, Wolf S, Hamprecht FA, Esteves T, Quelhas P, Demirel Ö, Malmström L, Jug F, Tomancak P, Meijering E, Muñoz-Barrutia A, Kozubek M, Ortiz-de-Solorzano C. An objective comparison of cell-tracking algorithms. *Nature Methods*, 2017, 14(12):1141–1152.
- [5] Isse K, Lesniak A, Grama K, Roysam B, Minervini MI, Demetris AJ. Digital transplantation pathology: Combining whole slide imaging, multiplex staining and automated image analysis. *American Journal of Transplantation*, 2012, 12(1):27–37.
- [6] Sirinukunwattana K, Pluim JPW, Chen H, Qi XJ, Heng PA, Guo YB, Wang LY, Matuszewski BJ, Bruni E, Sanchez U, Böhm A, Ronneberger O, Cheikh BB, Racoceanu D, Kainz P, Pfeiffer M, Urschler M, Snead DRJ, Rajpoot NM. Gland segmentation in colon histology images: The GlaS challenge contest. *Medical Image Analysis*, 2017, 35:489–502.
- [7] McCann MT, Ozolek JA, Castro CA, Parvin B, Kovacevic J. Automated histology analysis: Opportunities for signal processing. *IEEE Signal Processing Magazine*, 2014, 32(1):78–87.

- [8] Xu G, Song ZG, Sun Z, Ku C, Yang Z, Liu CC, Wang SH, Ma JP, Xu W. CAMEL: A weakly supervised learning framework for histopathology image segmentation. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2019. 10682–10691.
- [9] Xu Y, Tang Y, Yan W, Zhang YZ, Lai MD. Current status and prospect of artificial intelligence in pathology. Chinese Journal of Pathology, 2017,46(9):1–4 (in Chinese with English abstract).
- [10] Wang SY, Xiang JB, Li ZY, Lu SH, Hu J, Gao X, Yu L, Wang L, Wang JP, Wu Y, Chen ZY, Zhu HG. A plasma microRNA panel for early detection of colorectal cancer. Int'l Journal of Cancer, 2015,136(1):152–161.
- [11] Xiao SY. Pathology Atlas of Inflammatory Bowel Disease. Beijing: Peking Union Medical College Press, 2016 (in Chinese).
- [12] Tian JX, Liu GC, Gu SS, Ju ZJ, Liu JG, Gu DD. Deep learning in medical image analysis and its challenges. Acta Automatica Sinica, 2018,44(3):401–424 (in Chinese with English abstract).
- [13] Hu ZH, Zhao C, Bao J, Bu H. Application of whole slide imaging in diagnostic cytology. Chinese Journal of Pathology, 2017, 46(8):581–585 (in Chinese with English abstract).
- [14] Luo XF, Xu J, Chen JM. A deep convolutional network for pixel-wise segmentation on epithelial and stromal tissues in histologic images. Acta Automatica Sinica, 2017,43(11):2003–2013 (in Chinese with English abstract).
- [15] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc. of the IEEE, 1998, 86(11):2278–2324.
- [16] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2015. 3431–3440.
- [17] Sutskever I, Vinyals O, Le QV. Sequence to sequence learning with neural networks. In: Proc. of the Conf. on Neural Information Processing Systems. 2014. 3104–3112.
- [18] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Cognitive Modeling, 1986,323: 533–536.
- [19] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Proc. of the Int'l Conf. on Neural Information Processing Systems. 2014. 2672–2680.
- [20] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2014. 580–587.
- [21] Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D. Image segmentation using deep learning: A survey. arXiv preprint arXiv:2001.05566, 2020.
- [22] Deng L, Yu D. Deep learning: Methods and applications. Foundations and Trends® in Signal Processing, 2014,7(3-4):197–387.
- [23] Shen DG, Wu GR, Suk HI. Deep learning in medical image analysis. Annual Review of Biomedical Engineering, 2017,19: 221–248.
- [24] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. Medical Image Analysis, 2017,42:60–88.
- [25] Min S, Lee B, Yoon S. Deep learning in bioinformatics. Briefings in Bioinformatics, 2017,18(5):851–869.
- [26] Pan YS, Liu MX, Xia Y, Shen DG. Neighborhood-correction algorithm for classification of normal and malignant cells. In: Proc. of the IEEE Int'l Symp. on Biomedical Imaging. 2019. 73–82.
- [27] Payer C, Štern D, Bischof H, Urschler M. Integrating spatial configuration into heatmap regression based CNNs for landmark localization. Medical Image Analysis, 2019,54:207–219.
- [28] Chen H, Wang XY, Huang YJ, Wu XY, Yu YZ, Wang LS. Harnessing 2D networks and 3D features for automated pancreas segmentation from volumetric CT image. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2019. 339–347.
- [29] Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: A review. IEEE Reviews in Biomedical Engineering, 2009,2:147–171.
- [30] Veta M, Pluim JPW, van Diest P, Viergever MA. Breast cancer histopathology image analysis: A review. IEEE Trans. on Biomedical Engineering, 2014,61(5):1400–1411.
- [31] Irshad H, Veillard A, Roux L, Racocanu D. Methods for nuclei detection, segmentation, and classification in digital histopathology: A review-current status and future potential. IEEE Reviews in Biomedical Engineering, 2014,7:97–114.

- [32] Bándi P, Geessink O, Manson Q, van Dijk M, Balkenhol M, Hermesen M, Bejnordi BE, Lee B, Paeng K, Zhong AX, Li QZ, Zanjani FG, Zinger S, Fukuta K, Komura D, Ovtcharov V, Cheng SH, Zeng SQ, Thagaard J, Dahl AB, Lin HJ, Chen H, Jacobsson L, Hedlund M, Çetin M, Halıcı E, Jackson H, Chen R, Both F, Franke J, Küsters-Vandeveld H, Vreuls W, Bult P, van Ginneken B, van der Laak J, Litjens G. From detection of individual metastases to classification of lymph node status at the patient level: The CAMELYON17 challenge. *IEEE Trans. on Medical Imaging*, 2019,38(2):550–560.
- [33] Lai MD. *Histologic Atlas Grading of Common Tumor*. Beijing: People's Medical Publishing House, 2009 (in Chinese).
- [34] Irshad H, Montaser-Kouhsari L, Waltz G, Bucur O, Nowak JA, Dong F, Knoblauch NW, Beck AH. Crowdsourcing image annotation for nucleus detection and segmentation in computational pathology: Evaluating experts, automated methods, and the crowd. In: *Proc. of the Pacific Symp. on Biocomputing*. 2015. 294–305.
- [35] Kumar MD, Babaie M, Zhu SJ, Kalra S, Tizhoosh HR. A comparative study of CNN, BoVW and LBP for classification of histopathological images. In: *Proc. of the IEEE Symp. Series on Computational Intelligence*. 2017. 1–7.
- [36] Gelasca ED, Byun JY, Obara B, Manjunath BS. Evaluation and benchmark for biological image segmentation. In: *Proc. of the IEEE Int'l Conf. on Image Processing*. 2008. 1816–1819.
- [37] Brown KM, Barrionuevo G, Canty AJ, De Paola V, Hirsch JA, Jefferis GSXE, Lu J, Snippe M, Sugihara I, Ascoli GA. The DIADEM data sets: Representative light microscopy images of neuronal morphology to advance automation of digital reconstructions. *Neuroinformatics*, 2011,9:143–157.
- [38] Kather JN, Marx A, Reyes-Aldasoro CC, Schad LR, Zöllner FG, Weis CA. Continuous representation of tumor microvessel density and detection of angiogenic hotspots in histological whole-slide images. *Oncotarget*, 2015,6(22):19163–19176.
- [39] Li C, Wang XG, Liu WY, Latecki LJ, Wang B, Huang JZ. Weakly supervised mitosis detection in breast histopathology images using concentric loss. *Medical Image Analysis*, 2019,53:165–178.
- [40] Kumar N, Verma R, Sharma S, Bhargava S, Vahadane A, Sethi A. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Trans. on Medical Imaging*, 2017,36(7):1550–1560.
- [41] Kumar N, Verma R, Anand D, Zhou YN, Onder OF, Tsougenis E, Chen H, Heng PA, Li JH, Hu ZQ, Wang YZ, Koohbanani NA, Jahanifar M, Tajeddin NZ, Gooya A, Rajpoot N, Ren XH, Zhou SH, Wang Q, Shen DG, Yang CK, Weng CH, Yu WH, Yeh CY, Yang S, Xu SY, Yeung PH, Sun P, Mahbod A, Schaefer G, Ellinger R, Ecker O, Smedby CL, Wang B, Chidester TV, Ton MT, Tran J, Ma MN, Do S, Graham QD, Vu JT, Kwak A, Gunda R, Chunduri I, Hu C, Zhou XY, Lotfi D, Safdari R, Kascenas A, O'Neil A, Eschweiler D, Stegmaier J, Cui YP, Yin BC, Chen KL, Tian XM, Gruening P, Barth E, Arbel E, Remer I, Ben-Dor A, Sirazitdinova E, Kohl M, Braunewell S, Li YX, Xie XP, Shen LL, Ma J, Baksi KD, Khan MA, Choo J, Colomer A, Naranjo V, Pei LM, Iftekharruddin KM, Roy K, Bhattacharjee D, Pedraza A, Bueno MG, Devanathan S, Radhakrishnan S, Koduganty P, Wu ZH, Cai GY, Liu XJ, Wang YQ, Sethi A. A multi-organ nucleus segmentation challenge. *IEEE Trans. on Medical Imaging*, 2020, 39(5):1380–1391.
- [42] Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, 2016,7:Article No.29.
- [43] Gertych A, Ing N, Ma ZX, Fuchs TJ, Salman S, Mohanty S, Bhele S, Velásquez-Vacca A, Amin MB, Knudsen BS. Machine learning approaches to analyze histological images of tissues from radical prostatectomies. *Computerized Medical Imaging and Graphics*, 2015,46:197–208.
- [44] Xu Y, Jia ZP, Wang LB, Ai YQ, Zhang F, Lai MD, Chang EIC. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics*, 2017,18:Article No.281.
- [45] Veta M, Heng YJJ, Stathonikos N, Bejnordi BE, Beca F, Wollmann T, Rohr K, Shah MA, Wang DY, Rousson M, Hedlund M, Tellez D, Ciompi F, Zerhouni E, Lanyi D, Viana M, Kovalev V, Liauchuk V, Phoulady HA, Qaiser T, Graham S, Rajpoot N, Sjöblom E, Molin J, Paeng K, Hwang S, Park S, Jia ZP, Chang EIC, Xu Y, Beck AH, van Diest PJ, Pluim JPW. Predicting breast tumor proliferation from whole-slide images: The TUPAC16 challenge. *Medical Image Analysis*, 2019,54:111–121.
- [46] Yan ZQ, Yang X, Cheng KT. Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation. *IEEE Trans. on Biomedical Engineering*, 2018,65(9):1912–1923.
- [47] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015,521:436–444.
- [48] LeCun Y, Kavukcuoglu K, Farabet C. Convolutional networks and applications in vision. In: *Proc. of the IEEE Int'l Symp. on Circuits and Systems*. 2010. 253–256.

- [49] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Proc. of the Int'l Conf. on Neural Information Processing Systems. 2012. 1097–1105.
- [50] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2015.
- [51] He KM, Zhang XY, Ren SP, Sun J. Deep residual learning for image recognition. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2016. 770–778.
- [52] Szegedy C, Liu W, Jia YQ, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2015. 1–9.
- [53] Howard AG, Zhu ML, Chen B, Kalenichenko D, Wang WJ, Weyand T, Andreetto M, Adam H. MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861, 2017.
- [54] Huang G, Liu Z, van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2017. 4700–4708.
- [55] Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR. Improving neural networks by preventing coadaptation of feature detectors. arXiv preprint arXiv:1207.0580, 2012.
- [56] Liu W, Rabinovich A, Berg AC. ParseNet: Looking wider to see better. arXiv preprint arXiv:1506.04579, 2015.
- [57] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2015. 234–241.
- [58] Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2015. 1520–1528.
- [59] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2017,39(12):2481–2495.
- [60] Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol PA. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. Journal of Machine Learning Research, 2010,11(12):3371–3408.
- [61] Kingma DP, Welling M. Auto-encoding variational Bayes. In: Proc. of the Int'l Conf. on Learning Representations. 2014. 1–14.
- [62] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation, 1997,9(8):1735–1780.
- [63] Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. 2014. 1724–1734.
- [64] van den Oord A, Kalchbrenner N, Kavukcuoglu K. Pixel recurrent neural networks. In: Proc. of the Int'l Conf. on Int'l Conf. on Machine Learning. 2016. 1747–1756.
- [65] Liang XD, Shen XH, Feng JS, Lin L, Yan SC. Semantic object parsing with graph LSTM. In: Proc. of the European Conf. on Computer Vision. 2016. 125–143.
- [66] Liang XD, Lin L, Shen XH, Feng JS, Yan SC, Xing EP. Interpretable structure-evolving LSTM. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2017. 2175–2184.
- [67] Ho J, Ermon S. Generative adversarial imitation learning. In: Proc. of the Conf. on Neural Information Processing Systems. 2016. 4565–4573.
- [68] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2016.
- [69] Arjovsky M, Chintala S, Bottou L. Wasserstein GAN. arXiv preprint arXiv:1701.07875, 2017.
- [70] Mirza M, Osindero S. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, 2014.
- [71] Abadi M, Agarwal A, Barham P, Brevdo E, Chen ZF, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia YQ, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mane D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viegas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng XQ. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467, 2016.
- [72] Liao XY. Introduction of Deep Learning with PyTorch. Beijing: Publishing House of Electronics Industry, 2017 (in Chinese).

- [73] Jia YQ, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T. Caffe: Convolutional architecture for fast feature embedding. In: Proc. of the ACM Int'l Conf. on Multimedia. 2014. 675–678.
- [74] Bastien F, Lamblin P, Pascanu R, Bergstra J, Goodfellow I, Bergeron A, Bouchard N, Bengio Y. Theano: New features and speed improvements. arXiv preprint arXiv:1211.5590, 2012.
- [75] Song YY, Zhang L, Chen SP, Ni D, Lei BY, Wang TF. Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. IEEE Trans. on Biomedical Engineering, 2015,62(10):2421–2433.
- [76] Xie YP, Xing FY, Kong XF, Su H, Yang L. Beyond classification: Structured regression for robust cell detection using convolutional neural network. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2015. 358–365.
- [77] Albarqouni S, Baur C, Achilles F, Belagiannis V, Demirci S, Navab N. AggNet: Deep learning from crowds for mitosis detection in breast cancer histology images. IEEE Trans. on Medical Imaging, 2016,35(5):1313–1321.
- [78] Chen H, Wang X, Heng PA. Automated mitosis detection with deep regression networks. In: Proc. of the IEEE Int'l Symp. on Biomedical Imaging. 2016. 1204–1207.
- [79] Sirinukunwattana K, Raza SEA, Tsang YW, Snead DRJ, Cree IA, Rajpoot NM. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans. on Medical Imaging, 2016,35(5):1196–1206.
- [80] Xing FY, Xie YP, Yang L. An automatic learning-based framework for robust nucleus segmentation. IEEE Trans. on Medical Imaging, 2016,35(2):550–566.
- [81] Xu J, Xiang L, Liu QS, Gilmore H, Wu JZ, Tang JH, Madabhushi A. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. IEEE Trans. on Medical Imaging, 2016,35(1):119–130.
- [82] Yang L, Zhang YZ, Guldner IH, Zhang SY, Chen DZ. 3D segmentation of glial cells using fully convolutional networks and k -terminal cut. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2016. 658–666.
- [83] BenTaieb A, Kawahara J, Hamarneh G. Multi-loss convolutional networks for gland analysis in microscopy. In: Proc. of the IEEE Int'l Symp. on Biomedical Imaging. 2016. 642–645.
- [84] Chen JX, Yang L, Zhang YZ, Alber M, Chen DZ. Combining fully convolutional and recurrent neural networks for 3D biomedical image segmentation. In: Proc. of the Conf. on Neural Information Processing Systems. 2016. 3036–3044.
- [85] Xie YP, Zhang ZZ, Sapkota M, Yang L. Spatial clockwork recurrent neural network for muscle perimysium segmentation. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2016. 185–193.
- [86] Xie YP, Kong XF, Xing FY, Liu FJ, Su H, Yang L. Deep voting: A robust approach toward nucleus localization in microscopy images. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2015. 374–382.
- [87] Fakhry A, Peng HC, Ji S. Deep models for brain EM image segmentation: Novel insights and improved performance. Bioinformatics, 2016,32(15):2352–2358.
- [88] Stollenga MF, Byeon W, Liwicki M, Schmidhuber J. Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation. In: Proc. of the Conf. on Neural Information Processing Systems. 2015. 2998–3006.
- [89] Xu Y, Li Y, Liu MY, Wang YP, Lai MD, Chang EIC. Gland instance segmentation by deep multichannel side supervision. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2016. 496–504.
- [90] Chen H, Qi XJ, Cheng JZ, Heng PA. Deep contextual networks for neuronal structure segmentation. In: Proc. of the AAAI Conf. on Artificial Intelligence. 2016. 1167–1173.
- [91] Chen H, Qi XJ, Yu LQ, Dou Q, Qin J, Heng PA. DCAN: Deep contour-aware networks for object instance segmentation from histology images. Medical Image Analysis, 2017,36:135–146.
- [92] Song YY, Tan EL, Jiang XD, Cheng JZ, Ni D, Chen SP, Lei BY, Wang TF. Accurate cervical cell segmentation from overlapping clumps in PAP smear images. IEEE Trans. on Medical Imaging, 2017,36(1):288–300.
- [93] Fakhry A, Zeng T, Ji SW. Residual deconvolutional networks for brain electron microscopy image segmentation. IEEE Trans. on Medical Imaging, 2017,36(2):447–456.
- [94] Li YX, Shen LL, Yu SQ. HEp-2 specimen image segmentation and classification using very deep fully convolutional network. IEEE Trans. on Medical Imaging, 2017,36(7):1561–1572.
- [95] Li RJ, Zeng T, Peng HC, Ji SW. Deep learning segmentation of optical microscopy images improves 3-D neuron reconstruction. IEEE Trans. on Medical Imaging, 2017,36(7):1533–1541.

- [96] Zeng T, Wu B, Ji SW. DeepEM3D: Approaching human-level performance on 3D anisotropic EM image segmentation. *Bioinformatics*, 2017,33(16):2555–2562.
- [97] Khoshdeli M, Winkelmaier G, Parvin B. Multilayer encoder-decoder network for 3D nuclear segmentation in spheroid models of human mammary epithelial cell lines. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops*. 2018. 2320–2326.
- [98] Saha M, Chakraborty C. Her2Net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation. *IEEE Trans. on Image Processing*, 2018,27(5):2189–2200.
- [99] Senaras C, Sahiner B, Tozbikian G, Lozanski G, Gurcan MN. Creating synthetic digital slides using conditional generative adversarial networks: application to Ki67 staining. In: *Proc. of the Medical Imaging 2018: Digital Pathology*. 2018. [doi: 10.1117/12.2294999]
- [100] Januszewski M, Kornfeld J, Li PH, Pope A, Blakely T, Lindsey L, Maitin-Shepard J, Tyka M, Denk W, Jain V. High-precision automated reconstruction of neurons with flood-filling networks. *Nature Methods*, 2018,15:605–610.
- [101] Graham S, Rajpoot NM. Sams-Net: Stain-aware multi-scale network for instance-based nuclei segmentation in histology images. In: *Proc. of the IEEE Int'l Symp. on Biomedical Imaging*. 2018. 590–594.
- [102] Mahmood F, Borders D, Chen RJ, McKay GN, Salimian KJ, Baras A, Durr NJ. Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *IEEE Trans. on Medical Imaging*, 2020,39(11):3257–3267.
- [103] Yan ZQ, Yang X, Cheng KT. A three-stage deep learning model for accurate retinal vessel segmentation. *IEEE Journal of Biomedical and Health Informatics*, 2019,23(4):1427–1436.
- [104] Xue Y, Bigras G, Hugh J, Ray N. Training convolutional neural networks and compressed sensing end-to-end for microscopy cell detection. *IEEE Trans. on Medical Imaging*, 2019,38(11):2632–2641.
- [105] Jin QG, Meng ZP, Pham TD, Chen Q, Wei LY, Su R. DUNet: A deformable network for retinal vessel segmentation. *Knowledge-based Systems*, 2019,178:149–162.
- [106] Chakravarty A, Sivaswamy J. RACE-Net: A recurrent neural network for biomedical image segmentation. *IEEE Journal of Biomedical and Health Informatics*, 2019,23(3):1151–1162.
- [107] Wollmann T, Gunkel M, Chung I, Erfle H, Rippe K, Rohr K. GRUU-Net: Integrated convolutional and gated recurrent neural network for cell segmentation. *Medical Image Analysis*, 2019,56:68–79.
- [108] Yi JR, Wu PX, Jiang ML, Huang QY, Hoepfner DJ, Metaxas DN. Attentive neural cell instance segmentation. *Medical Image Analysis*, 2019,55:228–240.
- [109] Naylor P, Laé M, Reyat F, Walter T. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Trans. on Medical Imaging*, 2019,38(2):448–459.
- [110] Payer C, Štern D, Feiner M, Bischof H, Urschler M. Segmenting and tracking cell instances with cosine embeddings and recurrent hourglass networks. *Medical Image Analysis*, 2019,57:106–119.
- [111] Shin SY, Lee S, Yun ID, Lee KM. Deep vessel segmentation by learning graphical connectivity. *Medical Image Analysis*, 2019,58: 1–14.
- [112] Qu H, Yan ZN, Riedlinger GM, De S, Metaxas DN. Improving nuclei/gland instance segmentation in histopathology images by full resolution neural network and spatial constrained loss. In: *Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention*. 2019. 378–386.
- [113] Xie YT, Lu H, Zhang JP, Shen CH, Xia Y. Deep segmentation-emendation model for gland instance segmentation. In: *Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention*. 2019. 469–477.
- [114] Luna M, Kwon M, Park SH. Precise separation of adjacent nuclei using a siamese neural network. In: *Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention*. 2019. 577–585.
- [115] Koohbanani NA, Jahanifar M, Gooya A, Rajpoot N. Nuclear instance segmentation using a proposal-free spatially aware deep learning framework. In: *Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention*. 2019. 622–630.
- [116] Mou L, Zhao YT, Chen L, Cheng J, Gu ZW, Hao HY, Qi H, Zheng YL, Frangi A, Liu J. CS-Net: Channel and spatial attention network for curvilinear structure segmentation. In: *Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention*. 2019. 721–730.

- [117] Fang HH, Ai DN, Cong WJ, Yang SY, Zhu JJ, Huang Y, Song H, Wang YT, Yang J. Topology optimization using multiple-possibility fusion for vasculature extraction. *IEEE Trans. on Circuits and Systems for Video Technology*, 2020,30(2):442–456.
- [118] Hashimoto N, Fukushima D, Koga R, Takagi Y, Ko K, Kohno K, Nakaguro M, Nakamura S, Hontani H, Takeuchi I. Multi-scale domain-adversarial multiple-instance CNN for cancer subtype classification with unannotated histopathological images. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2020. 3852–3861.
- [119] Hayashida J, Nishimura K, Bise R. MPM: Joint representation of motion and position map for cell tracking. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2020. 3823–3832.
- [120] Liu DN, Zhang DH, Song Y, Zhang F, O'Donnell L, Huang H, Chen M, Cai WD. Unsupervised instance segmentation in microscopy images via panoptic domain adaptation and task re-weighting. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2020. 4243–4252.
- [121] Pandey S, Singh PR, Tian J. An image augmentation approach using two-stage generative adversarial network for nuclei image segmentation. *Biomedical Signal Processing and Control*, 2020,57:1–10.
- [122] Yan ZQ, Yang X, Cheng KT. Enabling a single deep learning model for accurate gland instance segmentation: A shape-aware adversarial learning framework. *IEEE Trans. on Medical Imaging*, 2020,39(6):2176–2189.
- [123] Zhao Y, Yang F, Fang YQ, Liu HL, Zhou NY, Zhang J, Sun JR, Yang S, Menze B, Fan XJ, Yao JH. Predicting lymph node metastasis using histopathological images based on multiple instance learning with deep graph convolution. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2020. 4837–4846.
- [124] He KM, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. In: *Proc. of the IEEE Int'l Conf. on Computer Vision*. 2017. 2961–2969.
- [125] Zhou YN, Onder OF, Dou Q, Tsougenis E, Chen H, Heng PA. CIA-Net: Robust nuclei instance segmentation with contour-aware information aggregation. In: *Proc. of the Int'l Conf. on Information Processing in Medical Imaging*. 2019. 682–693.
- [126] Tajbakhsh N, Jeyaseelan L, Li Q, Chiang JN, Wu ZH, Ding XW. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. *Medical Image Analysis*, 2020,63:Article No.101693.
- [127] Kim Y, Kim S, Kim T, Kim C. CNN-based semantic segmentation using level set loss. In: *Proc. of the IEEE Winter Conf. on Applications of Computer Vision*. 2019. 1752–1760.
- [128] Chen X, Williams BM, Vallabhaneni SR, Czanner G, Williams R, Zheng YL. Learning active contour models for medical image segmentation. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2019. 11632–11640.
- [129] Zhang LS, Bai M, Liao RJ, Urtasun R, Marcos D, Tuia D, Kellenberger B. Learning deep structured active contours end-to-end. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2018. 8877–8885.
- [130] Xu Z, Huang JZ. Detecting 10 000 cells in one second. In: *Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention*. 2016. 676–684.
- [131] Kashif MN, Raza SEA, Sirinukunwattana K, Arif M, Rajpoot N. Handcrafted features with convolutional neural networks for detection of tumor cells in histology images. In: *Proc. of the IEEE Int'l Symp. on Biomedical Imaging*. 2016. 1029–1032.
- [132] Li CM, Xu CY, Gui CF, Fox MD. Distance regularized level set evolution and its application to image segmentation. *IEEE Trans. on Image Processing*, 2010,19(12):3243–3254.
- [133] Milletari F, Navab N, Ahmadi SA. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In: *Proc. of the IEEE Int'l Conf. on 3D Vision*. 2016. 565–571.
- [134] Jégou S, Drozdal M, Vazquez D, Romero A, Bengio Y. The one hundred layers tiramisu: Fully convolutional DenseNets for semantic segmentation. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops*. 2017. 11–19.
- [135] Huang Q, Xia CY, Wu CH, Li SY, Wang Y, Kuo CCJ. Semantic segmentation with reverse attention. In: *Proc. of the British Machine Vision Conf.* 2017. 1–13.
- [136] Li HC, Xiong PF, An J, Wang LX. Pyramid attention network for semantic segmentation. In: *Proc. of the British Machine Vision Conf.* 2018. 1–13.
- [137] Fu J, Liu J, Tian HJ, Li Y, Bao YJ, Fang ZW, Lu HQ. Dual attention network for scene segmentation. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2019. 3146–3154.

- [138] Zhang XF, Dou H, Ju T, Xu J, Zhang ST. Fusing heterogeneous features from stacked sparse autoencoder for histopathological image analysis. *IEEE Journal of Biomedical and Health Informatics*, 2016,20(5):1377–1383.
- [139] Gogna A, Majumdar A, Ward R. Semi-supervised stacked label consistent autoencoder for reconstruction and analysis of biomedical signals. *IEEE Trans. on Biomedical Engineering*, 2017,64(9):2196–2205.
- [140] Zheng S, Jayasumana S, Romera-Paredes B, Vineet V, Su ZZ, Du DL, Huang C, Torr PHS. Conditional random fields as recurrent neural networks. In: *Proc. of the IEEE Int'l Conf. on Computer Vision*. 2015. 1529–1537.
- [141] Yang QS, Yan PK, Zhang YB, Yu, HY, Shi YY, Mou XQ, Kalra MK, Zhang Y, Sun L, Wang G. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans. on Medical Imaging*, 2018, 37(6):1348–1357.
- [142] Varghese A, Mohammed SKP, Sai Saketh C, Ganapathy K. Generative adversarial networks for brain lesion detection. In: *Proc. of the Medical Imaging 2017: Image Processing*. 2017. [doi: 10.1117/12.2254487]
- [143] Chen XR, Konukoglu E. Unsupervised detection of lesions in brain MRI using constrained adversarial auto-encoders. *arXiv preprint arXiv:1806.04972*, 2018.
- [144] Dou Q, Ouyang C, Chen C, Chen H, Heng PA. Unsupervised cross-modality domain adaptation of convnets for biomedical image segmentations with adversarial loss. In: *Proc. of the Int'l Joint Conf. on Artificial Intelligence*. 2018. 691–697.
- [145] Bokhorst JM, Pinckaers H, van Zwam P, Nagtegaal I, van der Laak J, Ciompi F. Learning from sparsely annotated data for semantic segmentation in histopathology images. In: *Proc. of the Int'l Conf. on Medical Imaging with Deep Learning*. 2019. 84–91.
- [146] Shan HM, Zhang Y, Yang QS, Kruger U, Kalra MK, Sun L, Cong WX, Wang G. 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network. *IEEE Trans. on Medical Imaging*, 2018,37(6):1522–1534.
- [147] Jaccard P. The distribution of the flora in the alpine zone. *New Phytologist*, 1912,11(2):37–50.
- [148] Dice LR. Measures of the amount of ecologic association between species. *Ecology*, 1945,26(3):297–302.
- [149] Song J, Xiao L, Lian ZC. Boundary-to-marker evidence-controlled segmentation and MDL-based contour inference for overlapping nuclei. *IEEE Journal of Biomedical and Health Informatics*, 2017,21(2):451–464.
- [150] Dai JF, He KM, Sun J. BoxSup: Exploiting bounding boxes to supervise convolutional networks for semantic segmentation. In: *Proc. of the IEEE Int'l Conf. on Computer Vision*. 2015. 1635–1643.
- [151] Lin GS, Shen CH, van den Hengel A, Reid I. Efficient piecewise training of deep structured models for semantic segmentation. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2016. 3194–3203.
- [152] Liu ZW, Li XX, Luo P, Loy CC, Tang XO. Semantic image segmentation via deep parsing network. In: *Proc. of the IEEE Int'l Conf. on Computer Vision*. 2015. 1377–1385.
- [153] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2017,40(4): 834–848.
- [154] Peng CX, Zhang XY, Yu G, Luo GM, Sun J. Large kernel matters—Improve semantic segmentation by global convolutional network. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2017. 4353–4361.
- [155] Lin GS, Milan A, Shen CH, Reid I. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2017. 1925–1934.
- [156] Wu ZF, Shen CH, van den Hengel A. Wider or deeper: Revisiting the ResNet model for visual recognition. *Pattern Recognition*, 2019,90:119–133.
- [157] Yuan YH, Chen XL, Wang JD. Object-contextual representations for semantic segmentation. *arXiv preprint arXiv:1909.11065*, 2019.
- [158] Zhao HS, Shi JP, Qi XJ, Wang XG, Jia JY. Pyramid scene parsing network. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. 2017. 2881–2890.
- [159] Chen LC, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.

- [160] Zhao HS, Zhang Y, Liu S, Shi JP, Change Loy C, Lin DH, Jia JY. PSANet: Point-wise spatial attention network for scene parsing. In: Proc. of the European Conf. on Computer Vision. 2018. 267–283.
- [161] Zhang H, Dana K, Shi JP, Zhang ZY, Wang XG, Tyagi A, Agrawal A. Context encoding for semantic segmentation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2018. 7151–7160.
- [162] Yu CQ, Wang JB, Peng C, Gao CX, Yu G, Sang N. Learning a discriminative feature network for semantic segmentation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2018. 1857–1866.
- [163] Zhang ZL, Zhang XY, Peng C, Xue XY, Sun J. Exfuse: Enhancing feature fusion for semantic segmentation. In: Proc. of the European Conf. on Computer Vision. 2018. 269–284.
- [164] Fu J, Liu J, Wang YH, Zhou J, Wang CY, Lu HQ. Stacked deconvolutional network for semantic segmentation. arXiv preprint arXiv:1708.04943, 2019. [doi: 10.1109/TIP.2019.2895460]
- [165] Luo P, Wang GR, Lin L, Wang XG. Deep dual learning for semantic image segmentation. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2017. 2718–2726.
- [166] He JJ, Deng ZY, Qiao Y. Dynamic multi-scale filters for semantic segmentation. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2019. 3562–3572.
- [167] He JJ, Deng ZY, Zhou L, Wang YL, Qiao Y. Adaptive pyramid context network for semantic segmentation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2019. 7519–7528.
- [168] Li X, Zhong ZS, Wu JL, Yang YB, Lin ZC, Liu H. Expectation-maximization attention networks for semantic segmentation. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2019. 9167–9176.
- [169] Chen LC, Zhu YK, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proc. of the European Conf. on Computer Vision. 2018. 801–818.
- [170] Lin D, Ji YF, Lischinski D, Cohen-Or D, Huang H. Multiscale context intertwining for semantic segmentation. In: Proc. of the European Conf. on Computer Vision. 2018. 603–619.
- [171] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. arXiv preprint arXiv:1511.07122, 2015.
- [172] Li X, Jie ZQ, Wang W, Liu CS, Yang JM, Shen XH, Lin Z, Chen Q, Yan SC, Feng JS. Foveanet: Perspective-aware urban scene parsing. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2017. 784–792.
- [173] Krapac J, Šegvic IKS. Ladder-style densenets for semantic segmentation of large natural images. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2017. 238–245.
- [174] Yu CQ, Wang JB, Peng C, Gao CX, Yu G, Sang N. BiSeNet: Bilateral segmentation network for real-time semantic segmentation. In: Proc. of the European Conf. on Computer Vision. 2018. 325–341.
- [175] Yang MK, Yu K, Zhang C, Li ZW, Yang KY. DenseASPP for semantic segmentation in street scenes. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2018. 3684–3692.
- [176] Cheng BW, Chen LC, Wei YC, Zhu YK, Huang ZL, Xiong JJ, Huang TS, Hwu WM, Shi HH. SPGNet: Semantic prediction guidance for scene parsing. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2019. 5218–5228.
- [177] Fu J, Liu J, Tian HJ, Li Y, Bao YJ, Fang ZW, Lu HQ. Dual attention network for scene segmentation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2019. 3146–3154.
- [178] Huang ZL, Wang XG, Huang LC, Huang TS, Wei YC, Liu WY. CCNet: Criss-cross attention for semantic segmentation. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2019. 603–612.
- [179] Fu J, Liu J, Wang YH, Li Y, Bao YJ, Tang JH, Lu HQ. Adaptive context network for scene parsing. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2019. 6748–6757.
- [180] Takikawa V, Acuna D, Jampani V, Fidler S. Gated-SCNN: Gated shape CNNs for semantic segmentation. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2019. 5229–5238.
- [181] Ding HH, Jiang XD, Shuai B, Qun Liu A, Wang G. Context contrasted feature and gated multi-scale aggregation for scene segmentation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2018. 2393–2402.
- [182] Liang XD, Zhou HF, Xing EP. Dynamic-structured semantic propagation network. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2018. 752–761.

- [183] Liang XD, Hu ZT, Zhang H, Lin L, Xing EP. Symbolic graph reasoning meets convolutions. In: Proc. of the Conf. on Neural Information Processing Systems. 2018. 1853–1863.
- [184] Li QL, Feng BW, Xie LP, Liang P, Zhang HS, Wang TF. A cross-modality learning approach for vessel segmentation in retinal images. *IEEE Trans. on Medical Imaging*. 2016,35(1):109–118.
- [185] Fu HZ, Xu YW, Wong DWK, Liu J. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In: Proc. of the IEEE Int'l Symp. on Biomedical Imaging. 2016. 698–701.
- [186] Liskowski P, Krawiec K. Segmenting retinal blood vessels with deep neural networks. *IEEE Trans. on Medical Imaging*. 2016,35: 2369–2380.
- [187] Orlando JI, Prokofyeva E, Blaschko MB. A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images. *IEEE Trans. on Biomedical Engineering*, 2016,64:16–27.
- [188] Dasgupta A, Singh S. A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation. In: Proc. of the IEEE Int'l Symp. on Biomedical Imaging. 2017. 248–251.
- [189] Srinidhi CL, Aparna P, Rajan J. A visual attention guided unsupervised feature learning for robust vessel delineation in retinal images. *Biomedical Signal Processing and Control*, 2018,44:110–126.
- [190] Alom MZ, Hasan M, Yakopcic C, Taha TM, Asari VK. Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation. *arXiv preprint arXiv:1802.06955*, 2018.
- [191] Wu YC, Xia Y, Song Y, Zhang YN, Cai WD. Multiscale network followed network model for retinal vessel segmentation. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2018. 119–126.
- [192] Soomro TA, Afifi AJ, Gao JB, Hellwich O, Zheng LH, Paul M. Strided fully convolutional neural network for boosting the sensitivity of retinal blood vessels segmentation. *Expert Systems with Applications*, 2019,134:36–52.
- [193] Adeyinka AA, Adebisi MO, Akande NO, Ogundokun RO, Kayode AA, Oladele TO. A deep convolutional encoder-decoder architecture for retinal blood vessels segmentation. In: Proc. of the Int'l Conf. on Computational Science and Its Applications. 2019. 180–189.
- [194] Wu YC, Xia Y, Song Y, Zhang DH, Liu DN, Zhang CY, Cai WD. Vessel-Net: Retinal vessel segmentation under multi-path supervision. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2019. 264–272.
- [195] Guo S, Wang K, Kang H, Zhang YJ, Gao YQ, Li T. BTS-DSN: Deeply supervised neural network with short connections for retinal vessel segmentation. *Int'l Journal of Medical Informatics*, 2019,126:105–113.
- [196] Ma WN, Yu S, Ma K, Wang JX, Ding XH, Zheng YF. Multi-task neural networks with spatial activation for retinal vessel segmentation and artery/vein classification. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2019. 769–778.
- [197] Wang B, Qiu S, He HG. Dual encoding U-Net for retinal vessel segmentation. In: Proc. of the Int'l Conf. on Medical Image Computing and Computer Assisted Intervention. 2019. 84–92.
- [198] Wang DY, Haytham A, Pottenburgh J, Saeedi O, Tao Y. Hard attention net for automatic retinal vessel segmentation. *IEEE Journal of Biomedical and Health Informatics*, 2020,24(12):3384–3396. [doi: 10.1109/JBHI.2020.3002985]
- [199] Rodrigues EO, Conci A, Liatsis P. ELEMENT: Multi-modal retinal vessel segmentation based on a coupled region growing and machine learning approach. *IEEE Journal of Biomedical and Health Informatics*, 2020,24(12):3507–3519. [doi: 10.1109/JBHI.2020.2999257]
- [200] Wu YC, Xia Y, Song Y, Zhang YN, Cai WD. NFN+: A novel network followed network for retinal vessel segmentation. *Neural Networks*, 2020,126:153–162.
- [201] Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: Proc. of the European Conf. on Computer Vision. 2014. 818–833.
- [202] Samek W, Binder A, Montavon G, Lapuschkin S, Müller KR. Evaluating the visualization of what a deep neural network has learned. *IEEE Trans. on Neural Networks and Learning Systems*, 2017,28(11):2660–2673.
- [203] Hou BJ, Zhou ZH. Learning with interpretable structure from gated RNN. *IEEE Trans. on Neural Networks and Learning Systems*, 2020,31(7):2267–2279.

- [204] Song J, Xiao L, Lian ZC. Cascade sparse convolution and decision tree ensemble model for nuclear segmentation in pathology images. *Acta Automatica Sinica*, 2020,47(2):378–390 (in Chinese with English abstract). [doi: 10.16383/j.aas.c190672]
- [205] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. In: *Proc. of the Int'l Conf. on Learning Representations*. 2015. 1–15.
- [206] Baxter J. *Theoretical Models of Learning to Learn*. Boston: Kluwer Academic Publishers, 1998.
- [207] Bayramoglu N, Heikkilä J. Transfer learning for cell nuclei classification in histopathology images. In: *Proc. of the European Conf. on Computer Vision*. 2016. 532–539.
- [208] Doersch C, Gupta A, Efros AA. Unsupervised visual representation learning by context prediction. In: *Proc. of the IEEE Int'l Conf. on Computer Vision*. 2015. 1422–1430.
- [209] Zhang PY, Wang FS, Zheng YF. Self supervised deep representation learning for fine-grained body part recognition. In: *Proc. of the IEEE Int'l Symp. on Biomedical Imaging*. 2017. 578–582.
- [210] Kametsky L, Jones TR, Fraser A, Bray MA, Logan DJ, Madden KL, Ljosa V, Rueden C, Eliceiri KW, Carpenter AE. Improved structure, function and compatibility for CellProfiler: Modular high-throughput image analysis software. *Bioinformatics*, 2011, 27(8):1179–1180.
- [211] Myatt DR, Hadlington T, Ascoli GA, Nasuto SJ. Neuromantic—From semi-manual to semi-automatic reconstruction of neuron morphology. *Frontiers in Neuroinformatics*, 2012,6(4):1–14.

附中文参考文献:

- [9] 许燕,汤焯,闫雯,张益肇,来茂德.病理人工智能的现状和展望.中华病理学杂志,2017,46(9):1–4.
- [11] 肖书渊.炎症性肠病病理图谱.北京:中国协和医科大学出版社,2016.
- [12] 田娟秀,刘国才,谷珊珊,鞠忠建,刘劲光,顾冬冬.医学图像分析深度学习方法研究与挑战.自动化学报,2018,44(3):401–424.
- [13] 胡佐鸿,赵春,包骥,步宏.全片数字化图像在诊断细胞病理学中的应用.中华病理学杂志,2017,46(8):581–585.
- [14] 骆小飞,徐军,陈佳梅.基于逐像素点深度卷积网络分割模型的上皮和间质组织分割.自动化学报,2017,43(11):2003–2013.
- [33] 来茂德.常见肿瘤组织学分级图谱.北京:人民卫生出版社,2009.
- [72] 廖星宇.深度学习入门之 PyTorch.北京:电子工业出版社,2017.
- [204] 宋杰,肖亮,练智超.级联稀疏卷积与决策树集成的病理图像细胞核分割方法.自动化学报,2020,47(2):378–390 [doi: 10.16383/j.aas.c190672]



宋杰(1986—),男,博士,讲师,CCF 专业会员,主要研究领域为生物医学图像处理,深度学习,机器学习与模式识别.



蔡子贇(1987—),男,博士,讲师,主要研究领域为深度学习,计算机视觉,模式识别,域的自适应.



肖亮(1976—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为信号处理,生物医学图像处理,机器学习与模式识别,计算机视觉.



蒋国平(1966—),男,博士,教授,博士生导师,主要研究领域为智能系统与复杂网络.



练智超(1983—),男,博士,副教授,主要研究领域为计算机视觉,人工智能系统安全,工业互联网.