

· 专题:双清论坛“虚拟生理人体与医学应用” ·

病理图像分析的深度学习研究方法综述*

宋国利¹ 陈 杰^{1,2**}

1. 鹏城实验室,深圳 518066
2. 北京大学 信息工程学院,深圳 518055

[摘 要] 数字病理在临床诊断中的普及为人工智能在病理学中的应用创造了条件。由于强大的建模表征能力,深度学习技术在计算病理学中得到广泛应用,在改善疾病诊断方面展现了巨大潜力。本文回顾了深度学习和病理图像分析相结合的应用,并概述了三个关键任务的领域进展,包括分类、分割和检测。针对每一个任务,介绍了相关的临床价值、技术难点、以及主流的算法设计等。目前病理人工智能算法已经取得了一些令人瞩目的成果,但仍未达到临床应用的标准。本文分析了将人工智能技术从研究转化为临床应用时面临的挑战以及未来研究方向。

[关键词] 病理图像分析;人工智能;深度学习;分类;分割;检测

1 从病理学到人工智能

病理图像包含丰富的表型信息,可用于诊断和监测癌症患者的病情,预测存活率,为癌症个性化治疗提供参考依据。传统病理学诊断是通过显微镜评估玻片上的活体组织切片来完成的。然而,这一过程往往繁琐且耗时,世界上大多数地区都面临着病理学家严重短缺问题。同时,受到病理医师经验水平和主观因素的影响,分析结果易产生较大差异性,甚至造成假阴性诊断结果,从而影响患者的治疗方案。此外,人工视觉评估也阻碍了分析结果的重现性,尤其不利于严重依赖定量组织病理学的指标判断,如基于 Ki67 免疫组织化学的增殖评分、肿瘤浸润淋巴细胞检测等。

数字显微镜的快速发展为数字病理的图像分析带来了机遇,实现了患者病理切片的数字化管理,能够减少切片识别错误,降低损坏和丢失的风险。使用人工智能(Artificial Intelligence, AI)技术分析组织切片通常被称为计算病理学^[1],允许基于机器的定量图像分析,辅助病理医师实现精准、客观、可重复的诊断。计算病理的早期研究可追溯到在细胞图



陈杰 北京大学深圳研究生院副教授,博士生导师。主要研究方向为计算机视觉和医学图像分析,在 *IEEE Transactions on Pattern Analysis and Machine Intelligence*、*International Journal of Computer Vision* 等杂志及 *IEEE Conference on Computer Vision and Pattern Recognition* 等会议发表高水平学术论文 100 余篇。近五年承担国家自然科学基金、广东重点和应急以及科技部重大科技基础设施建设等项目,获国家科技进步奖二等奖两次。担任国际期刊 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 和 *International Journal of Computer Vision* 客座编辑等。



宋国利 鹏城实验室助理研究员,主要研究方向包括医学图像分析、计算机视觉、多模态学习等。主持国家自然科学基金青年科学基金项目,参与多项国家自然科学基金项目。在相关国际期刊和国际会议上已发表学术论文 10 余篇。

像的初步应用^[2],对血涂片中的细胞个体根据定量的细胞特征,例如大小、形状和染色质分布,进行亚型分类。

收稿日期:2021-12-30;修回日期:2022-03-04

* 本文根据第 296 期“双清论坛”讨论的内容整理。

** 通信作者,Email:chenj@pcl.ac.cn

本文受到国家自然科学基金项目(61972217,32071459)、广东省基础与应用基础研究基金项目(2019B1515120049)的资助。

近年来,全切片数字扫描技术将病理组织切片转换为高分辨率的数字图像,即全视野切片图像(Whole-Slide Image, WSI)^[1],其所包含的巨大信息为机器学习(Machine Learning, ML)技术在病理学中的应用创造了条件。早期病理图像分析模型依赖于人类对特征的工程化,需精心设计和生物模式或结构匹配的特征,进而设计算法学习特征和特定任务之间的复杂关系。而深度学习(Deep Learning, DL)强大的特征表达能力促进了端到端学习,成为病理图像分析的主流方法。深度学习模型已被广泛用于肿瘤的检测、分级和分型。相关工作^[3-5]表明深度学习可以显著提高淋巴结转移检测、乳腺癌 Ki67 评分、前列腺癌 Gleason 分级,有效预测临床特征和某些分子标记物的状态。基于深度学习的预后模型也在肺癌、黑色素瘤和胶质瘤等多种疾病中得到验证,甚至可以取得优于病理学家的结果^[2, 5]。大量研究成果已表明深度学习技术可以为医生提供诊断参考,促进临床诊断效率和准确率的提升。

2 深度学习简介

深度学习是机器学习领域的一个分支,是一种以神经网络(Neural Network)为架构,对数据进行表征学习的算法^[6]。受生物神经网络的信息处理方式的启发,神经网络采用了分层结构,对观测数据进行多层次抽象^[7]。层次越深,特征的抽象程度越高。神经网络的结构和特性有助于学习数据中高度复杂和非线性的模式,在计算机视觉、自然语言处理、生物信息等许多领域都取得了优异的成果。

深度学习模型的基础是深度神经网络(Deep Neural Network, DNN),包含输入层、隐藏层和输出层,其中每层由并行的单元组成。DNN 可以包含多个隐藏层,增加隐藏层可以提升模型的表达能力,但过多的隐藏层会增加训练时间甚至产生过拟合问题。神经网络模型训练的目标是最小化适应任务的损失函数,可以通过随机梯度下降和误差反向传播方法学习模型参数。针对不同的任务和应用,可以

构建不同的网络结构,如针对图像、语音数据处理的卷积神经网络,针对自然语言处理、机器翻译和时序数据的循环神经网络。前馈神经网络是最常见的一种神经网络类型,它采用单向多层结构,参数从输入层向输出层单向传播。一些常用的神经网络层的基本构建块有卷积、池化、激活函数、全连接、残差、空洞卷积、注意力机制等^[8, 9]。卷积神经网络(Convolutional Neural Network, CNN)是一种具有局部连接、权重共享等特性的深层前馈神经网络(图 1),主要用来处理图像信息,在病理图像的方法研究中被广泛采用。

3 深度学习在病理图像分析中的应用

基于深度学习的病理切片图像的定量分析研究逐年增多。从病理图像一些基本元素(如细胞、腺体等)的分类、分割、检测等任务,到对整张病理图像不同区域的分割、检测,对肿瘤恶化程度的评级,以及提供生存分析和个性化治疗参考等,基于深度学习的方法都得到了广泛应用并取得了令人瞩目的成果。本文将侧重于总结深度学习技术应用于病理图像分析时面临的挑战和解决方法。更多关于深度学习模型的总结和这些模型在相关任务上的性能评价,可参考相关综述文献^[2, 10-12]。

在病理图像分析过程中应用深度学习方法的典型流程如图 2 所示,主要包括数据处理、模型训练和模型测试^[8, 13]。目前,医学领域的 AI 应用仍以监督学习为主^[2],数据需要先由专家手动标注,提供标签(如临床分类或图像标注等)用于模型训练。数字病理图像一般尺寸较大,难以直接加载到内存中,并且原始图像存在大量对于建模无意义的区域(如背景、污染区域等)。因此,需要将病理图像降采样或拆解为小的图像块(Patch)。进一步地,可以对图像进行归一化降低不同染色条件的影响,利用数据增强(如旋转,翻转,图像亮度、饱和度、对比度变化)等预处理操作增加训练数据的多样性,从而提高模型的泛化性。处理后的数据可以分为训练集和验证集,用

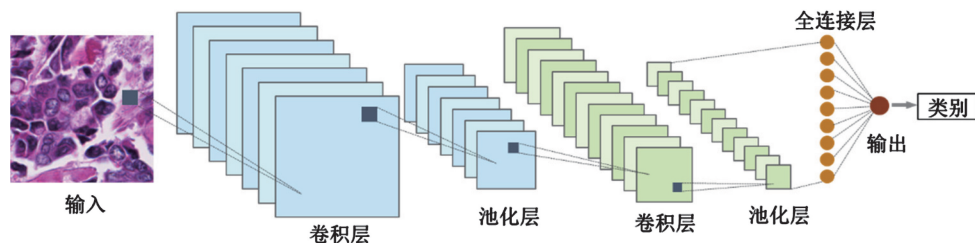


图 1 基于卷积神经网络的分类模型

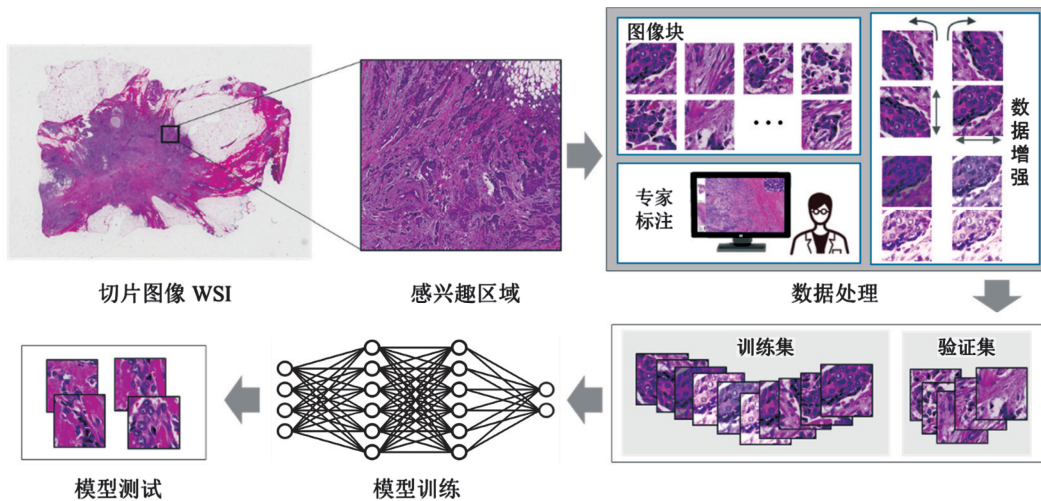


图 2 基于深度学习的病理图像分析方法的典型流程

于深度学习模型的训练优化。训好的模型可以部署到新的数据上进行测试。

3.1 分类模型

作为癌症诊断的金标准,病理诊断是病理图像的主要临床应用之一。而图像分类是深度学习技术应用最成功的领域之一。特别地,在数字组织病理图像分类方面,深度学习的发展使肺癌、乳腺癌、脑肿瘤、食道癌、胃癌等疾病的精确分类成为可能,分类模型还被用于癌症的亚分型、评级和预后等^[5]。国内外学者利用卷积神经网络(图 1)已经构建了很多成功的分类网络架构(如 VGGNet、InceptionNet、ResNet、EfficientNet 等),通过逐层提取图像特征,可以快速得到准确的诊断结果^[2, 4]。最近,基于 Transformer 的模型(如 ViT、Swin Transformer 等)通过对图像块序列的预测实现了图像分类,正被越来越多的研究者应用到医学图像领域^[11, 14]。研究表明目前的深度学习模型已经可以实现专家水平的疾病诊断,甚至在某些方面的表现还优于经验丰富的临床医生^[15]。但不能忽视的是基于深度学习的病理分类算法的研究仍然存在很多问题亟待解决。

(1) 数字组织病理图像分辨率过大,直接用于机器识别难以实现。

一张典型的 WSI 的分辨率约为 $100\ 000 \times 100\ 000$,在目前的计算机容纳能力下,CNN 难以对其直接计算。如果对图像进行降采样,则可能丢失一些有判别力的细节信息。为此,目前主流的方法采用两阶段的方式:先从 WSI 选取多个小区域块(如 224×224),使用深度学习网络对这些图像块进行训练;然后利用图像块的特征或预测,通过机器学

习等方法得到最终的全切片的诊断结果。针对第一步图像块的选择,研究者们提出了多种方法^[4, 13]。比如最简单的网格采样,采用滑动窗口法提取所有图像块或者采样一部分图像块进行 CNN 训练。随机采样的方式可能产生一些对疾病诊断无效或无信息的图像块。为了提高模型的计算效率和准确率,可以先由专家标注或者算法识别感兴趣区域,再针对这些区域进行采样。比如只在细胞核密度高的地方提取图像块而绕过脂肪和基质区域。

将原始的高分辨率图像分割成小块并分别处理的方式,需要对每个图像块进行标注,并且模型可能会在无信息的图像块上浪费计算资源。最近的工作表明,神经网络与其他方法(如注意力机制^[14]、自监督学习^[16]等)的结合可以增加深度学习解决方案在数字病理学的可行性。例如,Xu 等人^[17]提出一种基于注意力机制的决策网络可以自动筛选图像中有用的区域,并用于乳腺癌病理图像的分类。Katharopoulos 等人^[18]提出注意力采样方法,利用注意力分布找到图像中信息丰富的区域,实现了用一个 CPU 或 GPU 处理几百万像素的图像。DiPalma 等^[19]将知识蒸馏和 CNN 结合,将在高分辨率 WSI 上训好的模型中的知识迁移到在低分辨率 WSI 上训练的模型中,从而使模型在低分辨率的病理组织图像上也能取得较好的分类结果,同时提升了模型的计算效率。

(2) 病理图像的专家标注成本较高,模型训练欠缺有标注的数据。

基于深度网络的监督学习需要大量的训练数据和相应的标签,但对于十亿级的病理全切片图像,要获取精细的标签是十分困难的。一种可以降低标注

成本的策略是直接对 WSI 数据进行端到端的模型训练。Chen 等人^[20]提出一种统一的内存机制,可以直接对 $21\,500 \times 21\,500$ 分辨率的 WSI 进行端到端训练,实现了肺癌病理切片的 CNN 分类模型。但这种方法仍面临高分辨率数据处理和模型计算资源的挑战。针对人工标注中存在噪声的问题,研究者们也开展了噪声标签学习 (Noisy Label Learning)^[21] 方面的相关研究。比如 Yu 等人^[22]设计了一种噪声纠正的损失函数,用于降低噪声标签对肝癌细胞识别的影响。

同时,现有的病理图像数据集样本数较小,难以捕捉临床病例的广泛差异。相比于像素级或者图像块级的标签,图像级的标签(如是否存在肿瘤、是否恶性等)是比较容易获取的,故弱监督学习方法得到了越来越多研究者的关注。Campanella 等人^[23]基于多示例学习 (Multi-instance Learning) 构建了一种能够达到临床水平的深度学习系统。该系统在涵盖了前列腺癌、基底细胞癌、淋巴结转移乳腺癌的 44 732 张 WSI 切片数据集上进行分类实验。结果表明相比小数据集上的全监督训练,使用大规模数据的弱监督训练可以获得更好的泛化性能。为了充分利用有限的像素级标签和大量的图像级标签, Li 等人^[24]提出一种混合监督学习方法用于 WSI 分类,模型在训练过程中可以利用图像标签来改进生成的像素级伪标签。随着预训练模型在图像分类任务取得的成功,只需无标签数据的自监督预训练模型开始被应用到医学图像领域。Azizi 等人^[25]提出一种多示例的对比学习方法,利用每个病例的多幅图像进行数据增强以实现自我监督学习。

(3) 目前基于深度学习的分类模型主要针对单一尺度的病理图像。

不同倍率下的图像特征对于实际诊断都是至关重要的,在低倍率下主要看组织形态的特征,在高倍率下主要看细胞级别的特征。因此,临床实际诊断中,病理医生是在显微镜或者终端屏幕上不断移动评阅,并在不同倍率下分析病理图像,而不是针对局部或固定倍率。但目前的分类算法主要是针对单一尺度下的病理图像的切片进行分类分析。因此,为了提高分类网络的准确性和可解释性,一些算法提出结合不同尺度的病理图像特征进行分析,以期与临床实际诊断的方法相匹配。Li 等人^[26]利用 CNN 提出一个两阶段模型,包括低倍镜下的癌症检测阶段和高倍镜下可疑区域的癌症分类阶段。Wang 等人^[27]结合细胞核层面和图像块层面的多尺

度的结构信息,提出一种层次化的图卷积神经网络用于癌症生存预测。为了模拟病理学家的实际诊断过程,Hashimoto 等人^[28]将多示例学习、领域对抗 (Domain Adversarial) 和多尺度学习 (Multi-scale Learning) 技术与 CNN 有效地结合起来,并在恶性淋巴瘤分型任务中取得了优异性能。

针对单一倍率下的病理图像,也存在不同尺度的信息,比如有些图像异常区域较大,而有些异常区域却很小。为此,在网络设计方面,需要考虑如何处理这些多尺度信息。Song 等人^[29]基于经典的分类网络 ResNet 和多尺度特征学习网络 DeepLab v3,实现了临床适用的胃癌组织病理学诊断系统。Chen 等人^[30]提出一种结合 Transformer 和 CNN 的多尺度模型,用于提取图像中的全局信息和局部信息,实现了胃组织病理图像的分类。

3.2 分割模型

在病理分析和诊断过程中,从组织病理学图像中准确分割一些基本元素和组织区域具有重要作用。分割结果可用于可视化组织基元形态、计算形态学指标,是评估肿瘤恶性程度的前序步骤。例如,细胞核的形态和外观特征,包括大小和形状特征、密度、细胞核与细胞质之比、多态性等,可以有助评估癌症级别和预测治疗效果。腺体的形态学特征,包括其外部结构和腺体的形成过程,是病理学家诊断肠癌的重要标准。

随着数字病理研究的发展,需要能够面向多个器官和不同疾病的病理图像高精度分割方法。例如,为了保证细胞核形态的多样性,分割数据集 MoNuSeg^[31] 的训练集涵盖了来自乳房、肝脏、肾脏、前列腺、膀胱、结肠和胃七个不同器官的良性和病变组织样本。然而,不同组织基元的形态和大小差异很大,形状结构不规则,呈现高密度分布和粘连重叠的现象。此外,病理切片图像特征易受到切片的厚度、切割均匀度、杂质、染色深浅等因素的影响。为了克服这些挑战,研究者提出了不同类型的深度学习方法^[10],致力于对感兴趣区域的轮廓或整个内部进行像素级描绘以实现准确的病理图像分割。

基于 CNN 的分割模型被广泛应用于病理图像分割领域,其中最具影响力的是基于全卷积网络 (Fully Convolutional Network, FCN) 的模型^[32]。不同于传统的 CNN 分类模型,FCN 在网络架构上不再使用全连接层,而是用卷积层和池化层进行替代,并采用上采样得到原图像大小的稠密像素预测,实现了端到端的像素级的分割网络训练。受 FCN

的启发,Ronneberger 等人^[33]为生物学图像专门开发了 U-Net 分割网络,通过采用一个编码器捕获上下文和一个解码器来恢复空间分辨率(图 3),并使用一系列跨层连接来融合低层级的细节信息和高层级的语义特征,大幅提升了分割的精细度。

进一步地,通过考虑上下文信息或专家知识,研究者们提出了一系列的基于深度学习网络的病理图像分割模型,这些模型主要从以下几个方面进行设计:

(1) 改进网络结构,设计新的网络层或模块(如注意力机制、空间金字塔池化、空洞卷积等)。例如 FullNet^[34]提出全分辨率的网络结构,用 dense block 和空洞卷积代替池化层和上采样,以提升实例的定位准确率。CE-Net^[35]提出一种密集的空洞卷积模块和残差多核池化模块,更好地提取高层特征和多尺度上下文信息。TransUNet^[36]将 Transformer 用作医学图像分割任务的编码器,通过结合 U-Net 与 Transformer 的优点,利用 CNN 的高分辨率空间特征和注意力机制来提取的全局上下文信息以实现精确的定位。

(2) 辅助任务或辅助输出层引入分割网络,或者设计新的损失函数(如 Weighted Cross Entropy Loss、Focal Loss、Dice Loss、Tversky Loss、Combo Loss 等)。例如 DCAN^[37]利用腺体轮廓信息构建多任务学习框架,以改善稠密实例的分割效果。边界损失(Boundary Loss)^[38]是一种在边界空间构造的距离度量函数,可以作为区域损失函数(如 Dice Loss)的补充。也可以通过挖掘不同损失函数间的关系构造新的损失目标,比如 Yeung 等人^[39]提出一种 Unified Focal Loss,同时扩展了 Dice loss 和 Cross Entropy Loss,可以更好地处理分割任务中类别不平衡的数据。

(3) 将深度学习网络和经典建模理论进行有机结合。例如 CDNet^[40]提出将向心方向特征与 U-Net 分割网络结合,可以更好地建模相邻像素间的

空间关系。Li 等人^[41]利用 CNN 中的等变性(Equivariance)对无标签数据提出一种变换一致性策略,并应用于半监督分割任务。Karimi 等人^[42]提出直接降低或最小化 Hausdorff 距离来训练 CNN 分割模型。

(4) 针对病理图像的分割任务设计新的建模方式。例如 Naylor 等人^[43]针对相邻或重叠的细胞核的边界模糊问题,将像素层面的实例分割任务形式化为距离图的回归问题。为了有效利用全局语义特征和局部纹理特征处理细胞边界模糊的问题,Zhang 等人^[44]引入 Body-edge 分支提供边缘位置信息,并提出一种多分枝的 Transformer 结构。针对训练数据的多样性问题,Mahmood 等人^[45]利用对抗网络生成病理图像,并将细胞核的分割任务转换为图像翻译(Image-to-image Translation)问题。

3.3 检测模型

检测任务的目的是发现和定位病理切片中的感兴趣区域,对癌症筛查具有重要意义^[2, 13]。比如淋巴细胞的准确检测可以帮助设计新的生物标志物,用于监测癌症进展和治疗效果。有丝分裂检测用于量化肿瘤细胞的分化程度,是癌症分级中最重要的预后因素之一。

基于深度学习的目标检测技术根据有无候选框生成阶段可分为:两阶段目标检测和单阶段目标检测^[46],如图 4 所示。两阶段算法在准确度上有优势,而单阶段算法在速度上有优势。两阶段算法将检测问题划分为两个阶段,首先使用主干网络产生候选区域(Region Proposal),然后对候选区域分类。此类算法的典型代表是 R-CNN 系列算法。单阶段算法是基于回归的算法,只需一次提取特征,直接预测图像中感兴趣目标的类别和位置信息。代表性算法包括 YOLO 系列算法、SSD、Retina-Net 等。这两类目标检测算法都已经被应用于病理图像领域^[13],比如 Faster R-CNN 框架可以成功检测相差显微镜图像中的粘附细胞,SSD 用于解决精确的神经细胞

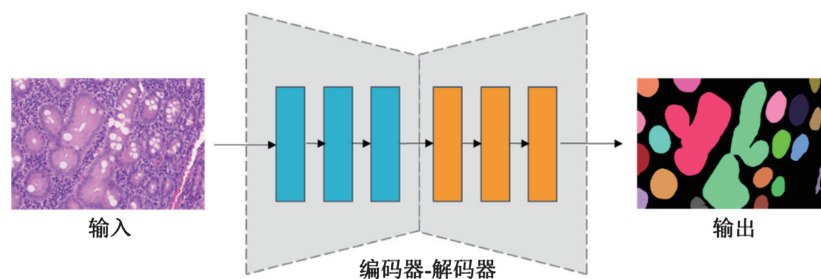


图 3 基于编码器—解码器网络结构的分割模型示意图

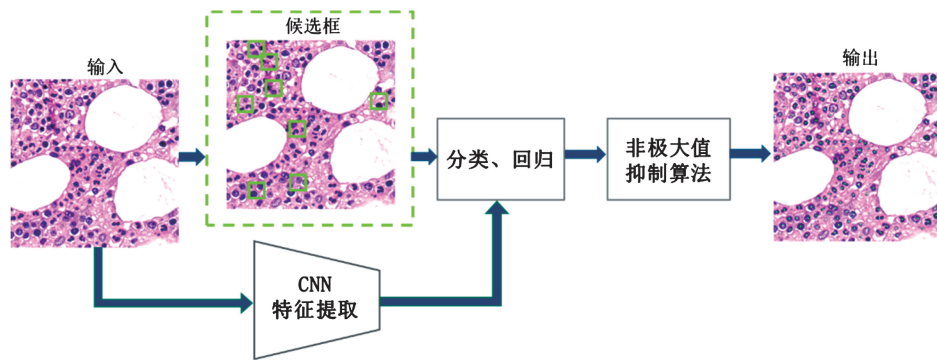


图4 检测模型典型示意图

检测任务。

最近一些目标检测工作从常用的基于候选框（Anchor-based）的方式转换为基于关键点（Keypoint-based）的方式^[46]。例如，CornerNet使用角点对表示每个目标，但该方法主要关注目标的边界信息，缺乏对目标内部信息的获取，容易造成误检。为此，CenterNet在角点对的基础上加入了中心关键点，同时兼顾了目标的边界信息和内部信息。CircleNet^[47]则是针对医学图像提出了一种圆形框的表示方式，其具有更好的旋转不变性，因而更适用于肾小球和细胞核等目标的检测。

随着DETR在目标检测领域的成功，一些结合CNN和Transformer的方法也被逐渐应用到医学图像的检测任务中^[14]。例如，COTR利用卷积层和Transformer层相结合组成编码器用于特征提取，实现了端到端的息肉检测。TR-Net则将Transformer用于多平面重建图像中冠状动脉狭窄的检测。但目前针对病理图像的基于Transformer的目标检测工作仍然较少，需要研究者们进一步探索。

然而，上述算法的研究主要聚焦于自然图像中通用目标的定位识别，在广泛的病理图像领域仍无法获得令人满意的性能。针对病理组织图像中存在的视觉挑战，仍需考虑检测目标的特异性，设计相关的深度学习解决方案。本文主要介绍两种常见的病理图像中的检测任务：细胞检测和有丝分裂检测。

(1) 细胞检测

在组织病理学图像中，细胞或细胞核检测有助于分析细胞空间分布，也是细胞或细胞核的实例分割、跟踪和形态学测量等一系列后续任务的必要前提。细胞或者细胞核一般存在大量重叠现象，是细胞检测任务面临的主要挑战。

由于目标检测任务要解决分类和定位的问题，

所以很多细胞检测的工作利用CNN分类算法和分割算法解决检测问题^[48, 49]。但这些工作没有考虑细胞层面的拓扑约束，忽略了细胞密度和细胞核一般呈圆形的假设。为了充分利用目标中心的像素和邻近目标中心像素之间的拓扑关系，许多学者将细胞或细胞核检测任务建模为回归问题^[2]。这类方法不是预测标签类别，而是直接预测细胞或细胞核的位置。比如定位细胞中心，利用回归模型预测每个像素值到最近细胞中心的距离得分。这类方法一般假设热图上概率最高的点对应于细胞的中心，代表性工作是非极大值抑制（Non-Maximum Suppression, NMS）算法，在细胞检测等任务上已取得了很好的效果^[13, 48]。中心点检测算法DCNet^[50]通过预测每种免疫细胞的中心点热图，很好地处理了细胞壁重叠或边界模糊的问题。RetinaNet可以利用细胞核内的点标注预测细胞的边界框，已经被应用于乳腺癌病理图像中上皮细胞的检测^[51]。

(2) 有丝分裂检测

细胞分裂对于维持生命至关重要，而快速生长的肿瘤往往具有更多的有丝分裂象，对细胞有丝分裂活动的准确监测对癌症的分级与诊断意义重大。不同于稠密细胞的检测，有丝分裂细胞在空间上分布比较分散，但有丝分裂的检测和计数既费时又困难，极富挑战性。一方面，有丝分裂是一个复杂的过程，具有多个分裂期，在时间和空间上都具有随机性；另一方面，有丝分裂细胞形态变化多端，胞核大小、形状、染色均不一致，并且大多数有丝分裂的细胞核与一般的细胞核难以区分。

随着深度学习技术的发展，研究人员开始探索使用神经网络实现有丝分裂细胞的检测和计数。Ciresan等人^[52]最早将CNN作为像素级分类器用于检测组织学图像中的有丝分裂。该工作完成了2012年ICPR和AMIDA 2013有丝分裂检测中的

挑战,并改变了整个数字组织病理学领域。这种方法通常计算昂贵且耗时,在临床环境中并不实用。此后,一些工作提出基于图像块的方法,使用 CNN 输出热图,计算每个像素属于有丝分裂细胞的概率。比如 SegMitos^[53] 利用语义分割网络预测每个像素的类别,输出分割结果概率图,然后直接从分割图定位有丝分裂细胞。还有一些研究者提出一种包括检测和判别的两步法:第一步先检测出所有细胞核,第二步判别每个细胞核是否为有丝分裂。比如,DeepMitosis^[54] 将目标检测框架 Faster R-CNN 应用于有丝分裂检测任务,并提出一种基于深度验证网络的分类模型,用于进一步验证检测结果。

4 病理人工智能的挑战与方向

深度学习等人工智能算法大大推动了病理图像自动诊断技术的发展,许多工作已经达到甚至超过临床医师的水平^[2, 5]。然而,目前大多计算机辅助诊断模型仍未能真正进入临床应用。由于医疗数据的限制以及临床准确率和可靠性的要求,病理人工智能的研究仍面临许多挑战,需要人工智能专家与病理学家的通力合作。

(1) 标注数据的稀缺性

费时费力的人工标注极大阻碍了深度学习算法在病理图像分析领域的广泛应用,该领域仍欠缺类似于 ImageNet 的数据集。一些工作通过众包平台来获取大量的图像标注,但最终仍需病理医师审核和校准。利用弱监督学习技术可以缓解对标注数据的依赖问题^[23],除了图像层面的标注,一些工作开始研究如何提取并利用病理报告和诊断记录的临床标注。近年来,半监督学习和自监督学习成为深度学习领域的热门方向。半监督学习是监督学习与无监督学习相结合的一种学习方法,该方法只需要少量带标签的样本和大量无标签的样本。自监督学习则完全不需要数据集提供标签。在病理图像分析中,基于半监督学习和自监督学习的 CNN 模型已经得到了研究者的关注^[55, 56]。

此外,医疗数据往往分散在不同机构,因其敏感性而很难被开放共享。患者隐私的保护和获得使用数据的授权是医学图像研究需要遵守的必要前提。伦理法规的要求可能会限制人工智能算法的数据选择,使得研究者需要在数据保护和模型创新之间取得平衡。目前实现隐私计算的一种方式联邦学习^[57],通过中央服务器使用分散在各地的数据,确保在模型训练过程中不会泄漏,从而使患者数据能

保留在当地机构或个人手中,避免将数据集中在一起的问题。

(2) 模型的泛化性

模型较差的泛化性是导致深度学习技术难以在临床中大规模应用的重要瓶颈。目前用于研究的病理数据集并不能代表临床实践中遇到的数据类型。不同医疗中心或地区的患者群体不同,不同的扫描仪或染色技术会造成病理数据的差异。因此,训练数据集不能涵盖足够丰富的数据,会导致深度学习病理分析模型的泛化性较差,即模型即便在训练集上表现很好,在另一个数据源上的测试结果仍可能很差。

迁移学习可以用来提高模型的稳定性和可泛化性,通过在大型和多样化的数据集上训练或微调模型减少泛化误差。针对泛化性问题,在病理图像处理过程中也常常应用数据增强和数据归一化等技术。数据增强是对数据引入额外的多样性,比如利用颜色增强模拟不同实验室的染色差异。数据归一化则是通过标准化和严格控制流程,将图像变换到统一的标准,最小化数据差异对模型的影响。最近,领域泛化(Domain Generalization, DG)是非常热门的一个方向,它研究的问题是从若干个具有不同分布的数据集中学习一个泛化能力强的模型,使其在未知的测试集上取得较好的效果^[58]。

(3) 模型的可解释性

基于深度学习的人工智能模型被认为是“黑箱”模型,其决策过程难以解释和理解。由此产生的高风险是阻碍深度学习产品获得临床实践和监管许可的主要问题之一。因此,在 AI 辅助诊断方面,深度学习产品更多被应用于检测或分割这些不需要解释决策的过程。例如,第一个获得美国食品和药品监督管理局批准的基于人工智能的病理学辅助产品 Paige Prostate,主要用于辅助检测前列腺活检中潜在的癌症细胞。

目前,可解释的 AI 是一个非常活跃的研究领域,研究者们也提出了一些解决方案。例如,一种基于算法融合的策略先使用深度学习算法进行初始任务,随后利用领域知识手工设计机器学习方法并用于预测,从而确保该方法的生物学可解释性^[59]。另一种提供解释性的典型方式是利用注意力机制或显著图,通过可视化图像中的判别性区域,为病理学家提供临床可解释的依据^[60]。这些方法部分缓解了可解释性问题,但仍然缺乏病理学或临床价值方面的知识。

5 总 结

人工智能在病理图像的分析应用中不断发展,对于辅助临床诊断意义重大,可以有效提升诊断效率和准确率。目前的研究表明,深度学习技术在某些基础任务上可以与人类专家的表现相媲美,但病理人工智能的研究仍然存在局限和挑战。首先,病理图像的分析计算对硬件和模型都有着较高的要求,现有设备对大尺寸病理切片的处理仍有困难,模型算法应具备较好的重复性、泛化性、可解释性,能提供真实可信的治疗建议;其次,在将人工智能技术从研究转化为临床应用时,须克服监管和隐私保护方面的挑战,模型和设备应证明是安全有效的。最后,为了在计算病理学领域取得实质性进展,将深度学习范式成功应用于病理分析任务,需要人工智能专家和病理学家之间的通力合作。

参 考 文 献

- [1] van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. *Nature Medicine*, 2021, 27(5): 775—784.
- [2] Srinidhi CL, Ciga O, Martel AL. Deep neural network models for computational histopathology: a survey. *Medical Image Analysis*, 2021, 67: 101813.
- [3] Acs B, Rantalainen M, Hartman J. Artificial intelligence as the next step towards precision pathology. *Journal of Internal Medicine*, 2020, 288(1): 62—81.
- [4] Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 2017, 42: 60—88.
- [5] Echle A, Rindtorff NT, Brinker TJ, et al. Deep learning in cancer pathology: a new generation of clinical biomarkers. *British Journal of Cancer*, 2021, 124(4): 686—696.
- [6] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015, 521(7553): 436—444.
- [7] Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(8): 1798—1828.
- [8] Moen E, Bannon D, Kudo T, et al. Deep learning for cellular image analysis. *Nature Methods*, 2019, 16(12): 1233—1246.
- [9] Alom MZ, Taha TM, Yakopcic C, et al. A state-of-the-art survey on deep learning theory and architectures. *Electronics*, 2019, 8(3): 292.
- [10] Asgari Taghanaki S, Abhishek K, Cohen JP, et al. Deep semantic segmentation of natural and medical images: a review. *Artificial Intelligence Review*, 2021, 54(1): 137—178.
- [11] Shamshad F, Khan S, Zamir SW, et al. Transformers in medical imaging: a survey. *arXiv preprint*, 2022, arXiv: 2201.09873.
- [12] Fernando T, Gammulle H, Denman S, et al. Deep learning for medical anomaly detection—A survey. *ACM Computing Surveys*, 2022, 54(7): 1—37.
- [13] Salvi M, Acharya UR, Molinari F, et al. The impact of pre- and post-image processing techniques on deep learning frameworks: a comprehensive review for digital pathology image analysis. *Computers in Biology and Medicine*, 2021, 128: 104129.
- [14] He K, Gan C, Li Z, et al. Transformers in Medical Image Analysis: A Review. *arXiv preprint*, 2022, arXiv: 2202.12165.
- [15] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 2017, 542(7639): 115—118.
- [16] Graham S, Epstein D, Rajpoot N. Dense steerable filter CNNs for exploiting rotational symmetry in histology images. *IEEE Transactions on Medical Imaging*, 2020, 39(12): 4124—4136.
- [17] Xu BL, Liu JX, Hou XX, et al. Attention by selection: a deep selective attention approach to breast cancer classification. *IEEE Transactions on Medical Imaging*, 2020, 39(6): 1930—1941.
- [18] Katharopoulos A, Fleuret F. Processing megapixel images with deep attention-sampling models// *International Conference on Machine Learning*. PMLR, 2019: 3282—3291.
- [19] DiPalma J, Suriawinata AA, Tafe LJ, et al. Resolution-based distillation for efficient histology image classification. *Artificial Intelligence in Medicine*, 2021, 119: 102136.
- [20] Chen CL, Chen CC, Yu WH, et al. An annotation-free whole-slide training approach to pathological classification of lung cancer types using deep learning. *Nature Communications*, 2021, 12: 1193.
- [21] Algan G, Ulusoy I. Image classification with deep learning in the presence of noisy labels: a survey. *Knowledge-Based Systems*, 2021, 215: 106771.
- [22] Yu X, Feng Z, Song M, et al. Tendentious Noise-rectifying Framework for Pathological HCC Grading// *British Machine Vision Conference*. 2021.
- [23] Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*, 2019, 25(8): 1301—1309.
- [24] Li JH, Chen W, Huang XD, et al. Hybrid supervision learning for pathology whole slide image classification. *Medical Image Computing and Computer Assisted Intervention-MICCAI 2021*, 2021: 309—318.
- [25] Azizi S, Mustafa B, Ryan F, et al. Big self-supervised models advance medical image classification// *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal, QC, Canada: IEEE, 2021: 3458—3468.

- [26] Li JY, Li WY, Sisk A, et al. A multi-resolution model for histopathology image classification and localization with multiple instance learning. *Computers in Biology and Medicine*, 2021, 131: 104253.
- [27] Wang ZC, Li JY, Pan ZF, et al. Hierarchical graph pathomic network for progression free survival prediction. *Medical Image Computing and Computer Assisted Intervention-MICCAI 2021*. Cham: Springer International Publishing, 2021; 227—237.
- [28] Hashimoto N, Fukushima D, Koga R, et al. Multi-scale domain-adversarial multiple-instance CNN for cancer subtype classification with unannotated histopathological images. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA. IEEE, 2020; 3851—3860.
- [29] Song Z, Zou S, Zhou W, et al. Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nature Communications*, 2020, 11: 4294.
- [30] Chen H, Li C, Li X, et al. GasHis-Transformer: A multi-scale Visual Transformer approach for gastric histopathology image classification. *arXiv preprint*, 2021, arXiv: 2104.14528.
- [31] Kumar N, Verma R, Anand D, et al. A multi-organ nucleus segmentation challenge. *IEEE Transactions on Medical Imaging*, 2020, 39(5): 1380—1391.
- [32] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015; 3431—3440.
- [33] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015*, 2015; 234—241.
- [34] Qu H, Yan ZN, Riedlinger GM, et al. Improving nuclei/gland instance segmentation in histopathology images by full resolution neural network and spatial constrained loss. *Medical Image Computing and Computer Assisted Intervention-MICCAI 2019*, 2019; 378—386.
- [35] Gu ZW, Cheng J, Fu HZ, et al. CE-net: context encoder network for 2D medical image segmentation. *IEEE Transactions on Medical Imaging*, 2019, 38(10): 2281—2292.
- [36] Chen J, Lu Y, Yu Q, et al. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint*, 2021, arXiv: 2102.04306.
- [37] Chen H, Qi XJ, Yu LQ, et al. DCAN: deep contour-aware networks for accurate gland segmentation. *2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA. IEEE, 2016; 2487—2496.
- [38] Kervadec H, Bouchtiba J, Desrosiers C, et al. Boundary loss for highly unbalanced segmentation. *Medical Image Analysis*, 2021, 67: 101851.
- [39] Yeung M, Sala E, Schönlieb CB, et al. Unified focal loss: generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized Medical Imaging and Graphics*, 2022, 95: 102026.
- [40] He HL, Huang ZY, Ding Y, et al. CDNet: centripetal direction network for nuclear instance segmentation// *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal, QC, Canada. IEEE, 2021; 4006—4015.
- [41] Li XM, Yu LQ, Chen H, et al. Transformation-consistent self-ensembling model for semisupervised medical image segmentation. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(2): 523—534.
- [42] Karimi D, Salcudean SE. Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks. *IEEE Transactions on Medical Imaging*, 2020, 39(2): 499—513.
- [43] Naylor P, Lae M, Reyal F, et al. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Transactions on Medical Imaging*, 2019, 38(2): 448—459.
- [44] Zhang YL, Higashita R, Fu HZ, et al. A multi-branch hybrid transformer network for corneal endothelial cell segmentation. *Medical Image Computing and Computer Assisted Intervention-MICCAI 2021*. Cham: Springer International Publishing, 2021; 99—108.
- [45] Mahmood F, Borders D, Chen RJ, et al. Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *IEEE Transactions on Medical Imaging*, 2020, 39(11): 3257—3267.
- [46] Liu L, Ouyang WL, Wang XG, et al. Deep learning for generic object detection: a survey. *International Journal of Computer Vision*, 2020, 128(2): 261—318.
- [47] Nguyen EH, Yang HC, Deng RN, et al. Circle representation for medical object detection. *IEEE Transactions on Medical Imaging*, 2022, 41(3): 746—754.
- [48] Swiderska-Chadaj Z, Pinckaers H, van Rijthoven M, et al. Learning to detect lymphocytes in immunohistochemistry with deep learning. *Medical Image Analysis*, 2019, 58: 101547.
- [49] Huang ZY, Ding Y, Song GL, et al. BCData: a large-scale dataset and benchmark for cell detection and counting. *Medical Image Computing and Computer Assisted Intervention-MICCAI 2020*, 2020; 289—298.
- [50] Lee SMW, Shaw A, Simpson JL, et al. Differential cell counts using center-point networks achieves human-level accuracy and efficiency over segmentation. *Scientific Reports*, 2021, 11: 16917.
- [51] Mercan C, Balkenhol M, van der Laak J, et al. From point annotations to epithelial cell detection in breast cancer histopathology using RetinaNet// *International Conference on Medical Imaging with Deep Learning-Extended Abstract Track*. 2019.

- [52] Cireşan DC, Giusti A, Gambardella LM, et al. Mitosis detection in breast cancer histology images with deep neural networks. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2013*, 2013; 411–418.
- [53] Li C, Wang XG, Liu WY, et al. Weakly supervised mitosis detection in breast histopathology images using concentric loss. *Medical Image Analysis*, 2019, 53: 165–178.
- [54] Li C, Wang XG, Liu WY, et al. DeepMitosis: Mitosis detection via deep detection, verification and segmentation networks. *Medical Image Analysis*, 2018, 45: 121–133.
- [55] Marini N, Otálora S, Müller H, et al. Semi-supervised training of deep convolutional neural networks with heterogeneous data and few local annotations: an experiment on prostate histopathology image classification. *Medical Image Analysis*, 2021, 73: 102165.
- [56] Koohbanani NA, Unnikrishnan B, Khurram SA, et al. Self-path: self-supervision for classification of pathology images with limited annotations. *IEEE Transactions on Medical Imaging*, 2021, 40(10): 2845–2856.
- [57] Rieke N, Hancox J, Li W, et al. The future of digital health with federated learning. *Npj Digital Medicine*, 2020, 3: 119.
- [58] Zhang HR, Dullerud N, Seyyed-Kalantari L, et al. An empirical framework for domain generalization in clinical settings//*Proceedings of the Conference on Health, Inference, and Learning*. 2021: 279–290.
- [59] Wang X, Janowczyk A, Zhou Y, et al. Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from digital H&E images. *Scientific Reports*, 2017, 7: 13543.
- [60] Zhang Z, Chen P, McGough M, et al. Pathologist-level interpretable whole-slide cancer diagnosis with deep learning. *Nature Machine Intelligence*, 2019, 1(5): 236–245.

A Review of Deep Learning Methods for Pathology Image Analysis

Song Guoli¹ Chen Jie^{1,2*}

1. *Peng Cheng Laboratory, Shenzhen 518066*

2. *School of Electronic and Computer Engineering, Peking University, Shenzhen 518055*

Abstract Recent advances of digital pathology in clinical diagnosis make it possible for the application of artificial intelligence to pathology. Due to the powerful learning ability in dealing with complex patterns, deep learning algorithms are widely being applied to computational pathology and have great potential to improve diagnosis. Here we review the intersection between deep learning and pathology image analysis, and survey the progress of three key tasks: classification, segmentation, and detection. For each task, we cover its clinical value, technical challenges, and state-of-the-art algorithms. Despite the promising results, very few artificial intelligence algorithms have reached clinical deployment. We describe the challenges that still need to be addressed before transforming artificial intelligence technologies from research to clinical practice and suggest directions for future work.

Keywords pathology image analysis; artificial intelligence; deep learning; classification; segmentation; detection

(责任编辑 吴征天)

* Corresponding Author, Email: chenjj@pcl.ac.cn