

DOI:10.16781/j.0258-879x.2018.08.0923

• 综述 •

## 人工智能技术在组织和细胞形态学评估中的应用

王 斐<sup>1</sup>, 魏培莲<sup>2</sup>, 潘 军<sup>3</sup>, 武 清<sup>2</sup>, 于观贞<sup>2\*</sup>

1. 海军军医大学(第二军医大学)长海医院普外三科, 上海 200433
2. 上海中医药大学附属龙华医院肿瘤七科, 上海 200032
3. 解放军 81 医院全军肿瘤中心肿瘤内科, 南京 210002

**[摘要]** 形态学评估和特殊染色评分系统是目前基础和临床科研的重要组成部分, 对于判断药物疗效和基因干预效果非常重要。然而, 目前的视觉评分系统主观性强、重复性差、准确率低, 极易出现漏诊和误诊。基于深度学习的人工智能技术有望克服这一问题, 利用卷积神经网络能准确提取与肿瘤患者疗效和预后相关的内部特征, 如肿瘤-间质比、神经侵犯和淋巴细胞空间分布; 图像化和数字化显示药物干预疾病进展的疗效, 同时能对与临床治疗、分型和预后相关的分子标志物进行定量化和自动化评分。人工智能技术应用于组织和细胞形态学评估后, 将会推动临床药物评价和基础科研评价的一致性、重复性和准确性, 有望进一步推动医学科研的发展。

**[关键词]** 人工智能; 肿瘤-间质比; 神经侵犯; 免疫组织化学

**[中图分类号]** R 329-33 **[文献标志码]** A **[文章编号]** 0258-879X(2018)08-0923-05

### Application of artificial intelligence technology in tissue and cell morphology assessment

WANG Fei<sup>1</sup>, WEI Pei-lian<sup>2</sup>, PAN Jun<sup>3</sup>, WU Qing<sup>2</sup>, YU Guan-zhen<sup>2\*</sup>

1. Department of Surgery (III), Changhai Hospital, Navy Medical University (Second Military Medical University), Shanghai 200433, China
2. Department of Oncology (VII), Longhua Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai 200032, China
3. Department of Medical Oncology, Cancer Center of PLA, No. 81 Hospital of PLA, Nanjing 210002, Jiangsu, China

**[Abstract]** Presently morphological evaluation and special staining scoring system are important components of basic and clinical research, and are very important for judging the efficacy of drugs and gene intervention. However, the current visual scoring system has some disadvantages such as strong subjectivity, poor repeatability and low accuracy, and is prone to missed diagnosis and misdiagnosis. Artificial intelligence technology based on deep learning is expected to overcome these problems. In our study, we found that the convolutional neural network can be used to accurately extract internal features related to the treatment and prognosis of tumors, such as tumor-stroma ratio, nerve invasion and spatial distribution of lymphatic cells in tumor specimens, visualizing and digitalizing the curative effect of drug intervention on disease progression, and can quantify and automatic evaluate the expression of molecular biomarkers related to clinical treatment, classification and prognosis. The application of artificial intelligence technology in tissue and cell morphology assessment will promote the consistency, repeatability and accuracy of clinical drug evaluation and basic scientific research evaluation, and is expected to further promote the development of medical research.

**[Key words]** artificial intelligence; tumor-stroma ratio; neural invasion; immunohistochemistry

[Acad J Sec Mil Med Univ, 2018, 39(8): 923-927]

医学发展源远流长, 经历了几千年的发展历史, 其由传统的“救死扶伤、防病治病”转变为

满足人类多层次和多元化的健康需求, 重在延长人类寿命并提高生活质量。在这一历程中, 科学技术

**[收稿日期]** 2018-06-21 **[接受日期]** 2018-07-12

**[基金项目]** 国家自然科学基金(81572856), 上海中医药大学附属龙华医院高层次人才引进项目(LH02.51.002)。Supported by National Natural Science Foundation of China (81572856) and High-level Talent Introduction Program of Longhua Hospital Affiliated to Shanghai University of Traditional Chinese Medicine (LH02.51.002).

**[作者简介]** 王 斐, 博士, 主治医师。E-mail: drwangfei@126.com

\*通信作者(Corresponding author)。Tel: 021-64385700, E-mail: qiaoshanqian@aliyun.com

的进步是推动医学发展的原动力,科学理论或技术的每一次重大突破都会促进医学救治能力和医疗服务水平的发展。医学发展与科技创新和前沿学科交叉融合并密切相关。医学的科技创新体现在两个方面,一是临床问题引导的基础医学创新,二是前沿技术的交叉融合创新。

基于临床问题的基础医学研究主要集中于细胞学、分子生物学、动物实验和人体的组织样本,上述技术的应用加深了对医学的了解,也推动了基础-临床转化。在这些科研项目中,经过干预的肿瘤细胞、动物样本和人体样本的形态学会发生相应改变,这种改变需通过特殊的方式予以显示和统计,但既往研究对于形态学的观察主要集中于肉眼和显微镜,必要时行免疫组织化学或免疫荧光检测协助判断,而后拍照进行人工计数或借助软件统计。上述方法极具主观性,易产生假阳性,重复性差,亟需一种新的手段评价形态学变化。

深度学习是人工智能技术的重要领域,通过构建多隐层模型和学习海量训练数据,获取数据有用的特征。深度学习采用层次网络结构进行逐层特征变换,将样本特征表示变换到一个新的特征空间,使分类或预测更容易。深度学习驱动图像识别已取得成功,包括对病理<sup>[1]</sup>和医学影像<sup>[2]</sup>的识别。斯坦福大学 Esteva 团队采用包含 2 032 种不同皮肤病的 129 450 张临床图像数据集,利用谷歌 TensorFlow 和 Inception\_V3 模型进行迁移学习,其对皮肤恶性黑素瘤的诊断准确率均达到 90% 以上<sup>[3]</sup>。ISBI (International Symposium on Biomedical Imaging) 举办的研究者挑战赛 (CAMELYON16) 评估了深度学习算法检测乳腺癌患者淋巴结转移病理切片中转移灶的潜力,结果显示深度学习算法诊断的曲线下面积 (area under curve, AUC) 为 0.556~0.994,病理医师诊断的 AUC 为 0.724 [95% 置信区间 (95% CI) 0.643~0.804],其中深度学习最佳算法在诊断模拟中的表现优于病理医师 [AUC: 0.994 (95% CI 0.983~0.999) vs 0.810 (95% CI 0.738~0.884)],  $P < 0.001$ <sup>[4]</sup>。

基于基础研究中形态学评估的局限性和深度学习的先进性,能否利用人工智能技术评估组织和细胞的形态学特征,特别是药物干预后的形态学特征变化呢?本课题组进行了大量尝试,取得了阳性结果,现总结如下。

## 1 机体组织样本中内部特征的量化分析与临床评价

人工智能技术为数字病理诊断带来了技术革新,有助于提高病理医师的诊断效率。目前普遍认为病理医师培养周期长、病理医师短缺、病理工作环境恶劣,期待通过人工智能技术协助诊断解决上述难题。由于病理诊断的金标准属性,期望人工智能技术代替病理医师是不可能的,但人工智能技术在挖掘疾病样本中的内部特征方面仍极具优势。

1.1 人工智能技术识别肿瘤-间质比 (tumor-stroma ratio, TSR) TSR 是指肿瘤组织内肿瘤细胞与间质成分的比值,主要通过术后病理切片评估获得。在结肠癌、非小细胞肺癌、乳腺癌、食管鳞癌、鼻咽癌、宫颈癌、肝细胞癌等实体瘤中,TSR 是影响肿瘤患者预后的独立危险因素<sup>[5-6]</sup>。既往主要由医师通过显微镜下观察肉眼判断 TSR,大多以 50% 作为间质丰富或缺乏的界定值。超过 50% 定义为间质丰富型,肿瘤侵袭能力更强;而低于 50% 定义为间质缺乏型,肿瘤侵袭能力稍差。然而这种评判标准存在许多问题,一是医师经验决定了 TSR 的精准性,二是 50% 的临界值是否准确。应用人工智能技术可准确量化 TSR,如果肿瘤细胞判断准确,TSR 可以精确到个位数。例如,本研究小组利用人工智能技术判读某张肿瘤组织病理切片的 TSR,可见肉眼判读的 TSR 为 30%~50% (图 1A),而通过人工智能技术判读的 TSR 为 27.3% (图 1B),表明人工智能技术在识别肿瘤样本内部特征方面具有明显优势。

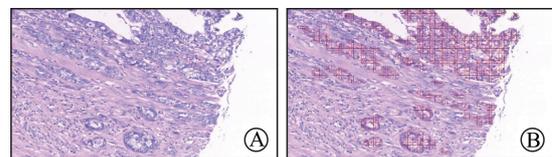


图 1 人工智能技术判读肿瘤-间质比

Fig 1 Tumor-stroma ratio identified by artificial intelligence technology

A: Tumor tissues; B: Tumor cells identified by artificial intelligence. H-E staining. Original magnification:  $\times 100$

1.2 人工智能技术识别肿瘤内部淋巴细胞的数量和空间分布 肿瘤浸润淋巴细胞 (tumor infiltrating lymphocyte, TIL) 是指从肿瘤组织中分离出来的浸润淋巴细胞,富含肿瘤特异性细胞毒性 T 淋巴

细胞和自然杀伤 (natural kill, NK) 细胞。TIL 在调节对化学治疗的反应和改善恶性肿瘤的临床结局中起着重要作用。如三阴性乳腺癌最易发生 >50% TIL 的肿瘤, 每增加 10% 的 TIL 都能增加生存获益; 所有乳腺癌亚型都有低、中等或高 TIL 肿瘤, 高 TIL 肿瘤中程序性死亡蛋白配体 1 (programmed cell death 1 ligand 1, PD-L1) 的表达上调, 导致此类三阴性乳腺癌对免疫检查点抑制剂治疗表现出强烈的临床反应<sup>[7]</sup>。因此, 鉴定和评价肿瘤内部的 TIL 对于判断预后和指导治疗具有重要价值。然而, 对于如何定量 TIL 以及分析它们的空间分布, 传统的基于 H-E 染色或免疫组织化学染色的分析极具主观性, 且耗时费力、准确性差, 而基于卷积神经网络的深度学习技术能解决这一问题。Saltz 等<sup>[8]</sup>利用肿瘤基因组图谱 (The Cancer Genome Atlas, TCGA) 数据库, 提出了基于 13 种 TCGA 肿瘤类型的 H-E 图像的 TIL 映射。这些 TIL 映射通过计算染色得到, 使用训练好的卷积神经网络对图像进行分类, 揭示了 TIL 模式的局部空间结构, 并与总体生存时间进行关联。TIL 图谱结构模式<sup>[8]</sup>采用标准的组织病理学参数进行分组, 丰富了从分子测量中衍生而来的 T 细胞亚群。TIL 的密度和空间结构在肿瘤类型、免疫亚型和肿瘤分子亚型间差异明显, 说明空间浸润状态可以反映特定的肿瘤细胞畸变状态。

**1.3 人工智能技术识别神经侵犯** 淋巴结转移是判断肿瘤患者预后和指导治疗的决定性因素, 人工智能技术用于识别淋巴结内的肿瘤细胞已有报道<sup>[4]</sup>。神经侵犯是肿瘤细胞转移的一种特殊方式, 如肝门部胆管癌和胰腺癌<sup>[9]</sup>。然而目前评价神经侵犯的方式仍是显微镜下肉眼观察, 易漏诊, 且不能反映整张切片的神经侵犯状态。本研究小组采用深度学习技术对肝门部胆管癌肿瘤细胞和神经组织分别进行学习和识别, 显示了肿瘤细胞侵犯神经组织的全过程, 包括肿瘤细胞首先向神经组织聚集, 而后侵犯神经鞘膜, 然后侵蚀神经纤维, 最终沿着神经转移 (图 2)。

## 2 细胞和动物组织样本的量化分析与药效关系

基础研究和临床药效评价会使用细胞和动物模型, 药物或基因干预手段对机体和肿瘤的治疗效果和不良反应需通过形态学方法予以展示和评价。传统的基于 H-E 染色或特殊染色的显微镜下肉眼

观察和判读具有局限性, 可采用深度学习技术学习细胞和动物病变样本的形态学表现, 如坏死、出血、淋巴细胞反应、纤维增生、肿瘤形成和数目、血管形成等。这些表现均极具特征性和规律性, 因此利用人工智能技术进行药效评价的可操作性强。本研究小组前期构建了胆管癌动物模型<sup>[10]</sup>, 采用不同药物进行干预, 然后利用人工智能技术学习该疾病特征, 结果表明人工智能技术可清晰显示疾病的发生过程和临床疗效, 见图 3。

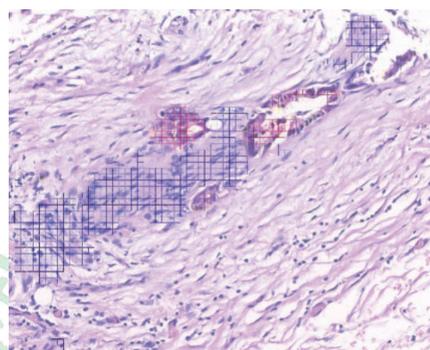


图 2 人工智能技术识别神经侵犯

Fig 2 Neural invasion identified by artificial intelligence technology

Blue: Nerve tissue; Red: Tumor cells. Original magnification:  $\times 200$

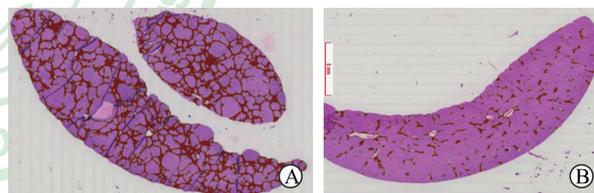


图 3 人工智能技术评估药物疗效

Fig 3 Drug efficacy assessed by artificial intelligence technology

A: Liver without intervention; B: Liver lesion after intervention. Red represents liver lesion recognized by artificial intelligence. Original magnification:  $\times 4$

## 3 细胞识别与分选

细胞学实验是基础和临床转化研究的基石, 但少有研究集中于细胞的形态学变化。伦敦癌症研究所的 Chris Bakal 教授和 Julia Sero 博士使用珀金埃尔默公司的 Opera<sup>®</sup> 高内涵成像分析系统获取图像, 采用类似于神经网络研究的方法分析了不同治疗条件下数以千计的个体乳腺癌细胞的形态和纹理特征, 并对线粒体群体的变化和趋势进行了检测<sup>[11]</sup>。该模型中对称性、阈值紧性、轴向、径向和

轮廓 (symmetry, threshold compactness, axial, radial and profile, STAR) 算法和鞍-边缘-脊 (saddle-edge-ridge, SER) 纹理起着重要作用。SER 纹理特征反映了像素强度的局部模式, 揭示了线粒体的结构, 且可以区分点状、管状线粒体结构。STAR 算法通过综合描述细胞内信号的对称性、信号密度、轴向特性、放射性及轮廓, 为统计表型提供了一组稳定可靠的参数, 例如: 线粒体是平均分布于细胞内还是按一定极性分布? 是否按细胞的质量中心分布? 是致密还是松散? 是靠近核膜、胞质还是外膜? 该研究将会在表型筛选和未知药物作用机制的研究中发挥作用。

Ota等<sup>[12]</sup>研发了一种新的细胞识别和分选系统, 称之为鬼影细胞测定仪 (ghost cytometry)。该系统将一种新的成像技术与人工智能技术结合, 用于识别和分选细胞, 不需使用空间分辨探测器即能产生物体图像, 基于单像素探测器的无图像超快荧光“成像”细胞术从细胞运动中获取空间信息, 而后被压缩转换成在单像素探测器上连续到达的信号。鬼影细胞测定仪以每秒 1 万多个细胞的速度识别细胞, 以每秒数千个细胞的速度对细胞进行分类。此外, 时间波形与随机模式强度分布的组合使之能在计算机上重建细胞形态, 可以直接在压缩波形上应用机器学习而不用进行图像重构, 实现高效的基于图像的无形态学细胞检测。这种方法将用于识别和分选患者血液中的循环肿瘤细胞, 加速药物发现和改进基于细胞疗法的疗效。

#### 4 特殊染色结果的量化分析和临床治疗与预后

评价药物或基因干预效果时, 除可使用 H-E 染色外, 还可借助特殊染色辅助判别, 包括免疫荧光技术和免疫组织化学技术。其中免疫组织化学技术因具有经济、方便、快速和高通量等特点而应用广泛, 但由于技术水平差异和评价体系的局限性, 其结果标准性较差。由于免疫组织化学染色图像也属于二维图像, 特征性明显, 人工智能技术非常适合对其结果进行判读和一致性评价。组织细胞的染色部位分为细胞核、细胞质和细胞膜, 鉴于细胞质染色评估较为容易, 本文重点介绍细胞核和细胞膜的人工智能自动评分系统。

4.1 人工智能自动评估人表皮生长因子受体 2 (human epidermal growth factor receptor 2, HER2) 染色 HER2 是诊断和评估胃癌和乳腺癌

的重要指标, 在临床实践中被公认为肿瘤预测和预后的标志。目前通过免疫组织化学对浸润性乳腺癌中 HER2 染色的评估也是通过显微镜下肉眼观察, 视觉评分较为主观, 容易出现观察者间的差异。鉴于 HER2 染色评分的预后和治疗意义, 需要一种更客观的评价方法。2016 年 6 月在诺丁汉举行的自动 HER2 染色评分比赛旨在推进基于人工智能技术的自动 HER2 染色评分方法。86 例数字化浸润性乳腺癌全切片同时行 H-E 染色和 HER2 免疫组织化学染色, 人工智能算法预测 HER2 免疫组织化学染色的评分, 并与两位专家进行比较, 结果显示人工智能算法的评估效果优于病理学专家, 证明了自动算法在帮助病理学家进行目标免疫组织化学评分方面具有巨大潜力<sup>[13]</sup>。另外一项基于深度学习技术的 HER2 染色自动评分算法对 71 例乳腺肿瘤切除样本进行了评分, 结果显示自动评分与病理学家的一致性为 83%; 对 12 例不一致的病例进行了独立检查, 发现 8 例患者的初步病理诊断存在问题, 分析诊断不一致的原因可能是 HER2 染色异质性高, 在评估 HER2 的表达时存在视觉差异<sup>[14]</sup>。

4.2 人工智能自动评估 Ki-67 染色 Ki-67 是一种常用的肿瘤细胞增殖标志物, 在判断预后和指导分型方面具有重要价值, 然而其评估较为困难, 目前的人工评分方法具有很高的误差, 因为对于同一肿瘤不同观察者可能会选择不同区域进行评分。基于深度学习技术的 Ki-67 染色自动评分方法可通过使用全玻片分析来识别和评估增殖率最高的肿瘤区域, 消除选择偏差。一项研究对 328 例雌激素受体/孕激素受体阳性、HER2 阴性乳腺癌患者早期切除的全切片行 Ki-67 染色自动评分, 发现其区分 Oncotype Dx 高风险和低风险患者的准确度为 97%、灵敏度为 98%、特异度为 80%。Ki-67 染色自动评分成功解决了既往评分一致性、重现性和准确性差的问题, 可以集成到病理学实验室的工作流程中<sup>[15]</sup>。基于深度学习技术的免疫组织化学自动算法同样也能应用于类似的基础和临床研究中, 其评估结果具有高度一致性、重复性和准确性。

4.3 人工智能自动评估多重免疫组织化学染色 基于深度学习技术的算法能根据组织图像中的颜色和空间信息对多重染色的生物标志物进行量化研究, 自动定位和量化整张切片任意选定区域中阳性表达的生物标志物。此类算法为卷积神经网络, 可根据所表达的生物标志物分别对每个细胞核

进行分类。Sheikhzadeh 等<sup>[16]</sup>收集并使用了免疫组织化学染色的组织切片, 根据每个细胞核中表达的生物标志物 (p16 阳性、Ki-67 阳性、p16 阳性/Ki-67 阳性, p16 阳性/Ki-67 阴性), 人工标记细胞核图像用于训练, 将训练好的算法扩展到全玻片, 结果显示该算法与人工标记结果高度一致。

## 5 总结和展望

基于深度学习的人工智能技术有望克服组织学和细胞学视觉评分系统主观性强、重复性差、准确率低等问题, 避免或减少漏诊率和误诊率。利用人工智能技术能准确提取与肿瘤预后和疗效相关的内部特征, 如 TSR、神经侵犯和淋巴细胞空间分布; 图像化显示药物干预疾病进展的疗效, 定量化和自动化评分与临床治疗、分型和预后相关的分子标志物。人工智能技术将会极大推动临床药物评价和基础科研评价的一致性、重复性和准确性, 有望进一步促进医学科研的发展。

## 【参考文献】

- [1] YE J J. Artificial intelligence for pathologists is not near—it is here: description of a prototype that can transform how we practice pathology tomorrow[J]. *Arch Pathol Lab Med*, 2015, 139: 929-935.
- [2] SETIO A A A, TRAVERSO A, DE BEL T, BERENS M S N, BOGAARD C V D, CERELLO P, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge[J]. *Med Image Anal*, 2017, 42: 1-13.
- [3] ESTEVA A, KUPREL B, NOVOA R A, KO J, SWETTER S M, BLAU H M, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. *Nature*, 2017, 542: 115-118.
- [4] EHTESHAMI BEJNORDI B, VETA M, JOHANNES VAN DIEST P, VAN GINNEKEN B, KARSSMEIJER N, LITJENS G, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer[J]. *JAMA*, 2017, 318: 2199-2210.
- [5] ERIKSEN A C, SORENSEN F B, LINDEBJERG J, HAGER H, DEPONT CHRISTENSEN R, KJÆR-FRIFELDT S, et al. The prognostic value of tumour stroma ratio and tumour budding in stage II colon cancer. A nationwide population-based study[J/OL]. *Int J Colorectal Dis*, 2018. doi: 10.1007/s00384-018-3076-9. [Epub ahead of print].
- [6] VAN PELT G W, SANDBERG T P, MORREAU H, GELDERBLUM H, VAN KRIEKEN J H J M, TOLLENAAR R A E M, et al. The tumour-stroma ratio in colon cancer: the biological role and its prognostic impact[J]. *Histopathology*, 2018, 73: 197-206.
- [7] STANTON S E, DISIS M L. Clinical significance of tumor-infiltrating lymphocytes in breast cancer[J/OL]. *J Immunother Cancer*, 2016, 4: 59. doi: 10.1186/s40425-016-0165-6.
- [8] SALTZ J, GUPTA R, HOU L, KURC T, SINGH P, NGUYEN V, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images[J]. *Cell Rep*, 2018, 23: 181-193.
- [9] YU G, YU W, JIN G, XU D, CHEN Y, XIA T, et al. PKM2 regulates neural invasion of and predicts poor prognosis for human hilar cholangiocarcinoma[J/OL]. *Mol Cancer*, 2015, 14: 193. doi: 10.1186/s12943-015-0462-6.
- [10] 王芳芳, 陈颖, 魏培莲, 赵行, 李超富, 高云姝, 等. 阿司匹林中断硫代乙酰胺诱发大鼠胆管上皮癌变[J]. *第二军医大学学报*, 2018, 39: 285-290.  
WANG G F, CHEN Y, WEI P L, ZHAO X, LI C F, GAO Y S, et al. Aspirin interrupts bile duct carcinoma in rats induced by thioacetamide[J]. *Acad J Sec Mil Med Univ*, 2018, 39: 285-290.
- [11] PerkinElmer, Inc. Phenotypic characterization of mitochondria in breast cancer cells using morphology and texture properties [R/OL]. [2018-06-18]. [http://www.perkinelmer.com.cn/CMSResources/Images/CST\\_Phenotypic\\_Characterization\\_Mitochondria\\_BreastCancer.pdf](http://www.perkinelmer.com.cn/CMSResources/Images/CST_Phenotypic_Characterization_Mitochondria_BreastCancer.pdf).
- [12] OTA S, HORISAKI R, KAWAMURA Y, UGAWA M, SATO I, HASHIMOTO K, et al. Ghost cytometry[J]. *Science*, 2018, 360: 1246-1251.
- [13] QAISER T, MUKHERJEE A, REDDY P B C, MUNUGOTI S D, TALLAM V, PITKÄÄHO T, et al. HER2 challenge contest: a detailed assessment of automated HER2 scoring algorithms in whole slide images of breast cancer tissues[J]. *Histopathology*, 2018, 72: 227-238.
- [14] VANDENBERGHE M E, SCOTT M L, SCORER P W, SÖDERBERG M, BALCERZAK D, BARKER C. Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer[J/OL]. *Sci Rep*, 2017, 7: 45938. doi: 10.1038/srep45938.
- [15] THAKUR S S, LI H, CHAN A M Y, TUDOR R, BIGRAS G, MORRIS D, et al. The use of automated Ki67 analysis to predict Oncotype DX risk-of-recurrence categories in early-stage breast cancer[J/OL]. *PLoS One*, 2018; 13: e0188983. doi: 10.1371/journal.pone.0188983.
- [16] SHEIKHZADEH F, WARD R K, VAN NIEKERK D, GUILLAUD M. Automatic labeling of molecular biomarkers of immunohistochemistry images using fully convolutional networks[J/OL]. *PLoS One*, 2018, 13: e0190783. doi: 10.1371/journal.pone.0190783.