

DOI:10.16781/j.0258-879x.2018.08.0917

• 综述 •

智能医学影像识别研究现状与展望

周瑞泉^{1,2}, 纪洪辰¹, 刘 荣^{1,2*}

1. 解放军总医院肝胆外科, 北京 100853

2. 南开大学医学院, 天津 300071

[摘要] 智能医学影像识别是基于人工智能技术,对 X 线片、计算机断层扫描、磁共振成像等常用医学影像学技术扫描图像和手术视频进行分析处理的过程,其发展方向主要包括智能影像诊断、影像三维重建与配准、智能手术视频解析等。智能影像诊断和影像三维重建与配准可提高影像识别的效率和质量,为疾病诊断和治疗提供帮助;智能手术视频解析可帮助外科医师学习、理解外科手术,并进一步指导手术过程。目前,对该领域的研究已取得一定的进展,正在逐步走向临床应用。本文就智能医学影像识别取得的进展进行总结,并对该领域的发展前景进行展望。

[关键词] 人工智能;智能医学;影像识别;机器学习;卷积神经网络

[中图分类号] R 445 **[文献标志码]** A **[文章编号]** 0258-879X(2018)08-0917-06

Intelligent medical image recognition: progress and prospect

ZHOU Rui-quan^{1,2}, JI Hong-chen¹, LIU Rong^{1,2*}

1. Department of Hepatobiliary and Pancreatic Surgical Oncology (II), General Hospital of PLA, Beijing 100853, China

2. School of Medicine, Nankai University, Tianjin 300071, China

[Abstract] Based on artificial intelligence technology, the intelligent medical image recognition refers to the analysis and process of medical images scanned by medical imaging technologies such as X-ray films, computed tomography and magnetic resonance imaging, and surgical video. Major trends in intelligent medical image recognition include intelligent image diagnosis, three-dimensional reconstruction and registration, intelligent surgery video parsing and so on. Intelligent image diagnosis and three-dimensional reconstruction and registration can improve the efficiency and quality of image recognition, and provide a helpful method for clinical diagnosis and treatment; intelligent surgery video parsing can help surgeons learn and understand surgical procedures, and further guide the operation process. Now the research of intelligent medical image recognition has gained some theoretical and technological achievement and gradually been applied in clinic. In this paper, we summarized the progress of intelligent medical image recognition and put forward the development prospect in this field.

[Key words] artificial intelligence; intelligent medicine; image recognition; machine learning; convolutional neural network

[Acad J Sec Mil Med Univ, 2018, 39(8): 917-922]

近年来,人工智能的浪潮席卷了诸多领域,在智能医学影像识别方面亦取得了长足进步。智能医学影像识别是基于人工智能技术,对 X 线片、计算机断层扫描(computed tomography, CT)、磁共振成像(magnetic resonance imaging, MRI)等常用的医学影像学技术扫描图像和手术视频进行分析处理的过程,其发展方向主要包括智

能影像诊断、影像三维重建与配准、智能手术视频解析等。智能影像诊断可利用机器学习的方法对疾病进行病理分类,提高诊断准确率;智能器官识别、血管分割等方法可为影像三维重建提供构架,为疾病的诊断和治疗提供帮助;智能手术视频解析能够帮助外科医师学习、理解外科手术,进一步指导手术过程。本文就目前在智能医学影像识别研究

[收稿日期] 2018-06-21 **[接受日期]** 2018-07-16

[作者简介] 周瑞泉, 博士, 住院医师. E-mail: ruiquanzhou@yeah.net

*通信作者(Corresponding author). Tel: 010-66937591, E-mail: liurong301@126.com

方面取得的进展作一综述,分析目前存在的问题,并对该领域的发展前景进行展望。

1 智能医学影像识别发展现状

1.1 器官识别 智能图像处理的基本思路是目标识别—目标分割—后续分析。无论是影像诊断还是三维重建配准(图1),其第1个步骤都是从图像中识别目标区域。目前,完成目标器官识别的主要方法是采用经可靠人工标注的CT或MRI图像作为训练集训练机器学习模型,获得器官的形状及位置,然后对目标器官进行分割,这在很大程度上简化了算法的流程,也提高了分割的准确性。该方法对大器官的识别非常有效,能获得很高的准确率;但对小器官的识别效果欠佳,例如胰腺体积约占腹腔的0.5%,上述方法的识别效果就不甚理想^[1]。



图1 基于冠状动脉计算机断层扫描的三维重建
Fig 1 Three-dimensional reconstruction based on multi-slices spiral computed tomography of coronary artery

A: A cross profile of multi-slices spiral computed tomography of coronary artery; B: Three-dimensional reconstruction model of coronary artery

目前对于小器官的识别效果欠佳问题有2种解决方案。一种方案是首先对小器官(目标器官)邻近的大器官(定位器官)进行定位,然后根据两者的相对位置关系确定包含目标器官的感兴趣容积(volume of interest, VOI)区域,进而在更小、更准确的VOI区域内进行目标器官的分割操作。为了使获得的VOI区域尽量准确,选择的定位器官应该与目标器官有尽可能稳定的位置关系。例如在胰腺的识别中,肝脏是一个很好的定位器官^[2],可以通过获得肝脏的位置找到胰腺的VOI区域(图2)。该方法的不足之处在于由于器官形态、位置的差异性,框定的VOI区域通常比实际位置大,这在一定程度上降低了图像识别的效率。

另一种方案是使用迭代的方法^[1]。训练2个深度神经网络,一个用于对原图像进行粗略分割,另一个用于对分割结果周围的局部图像进行分割细化,通过不断迭代后一个深度网络来优化分割结果。该方法的缺点非常明显,即运算量过大,对硬件的要求非常高。

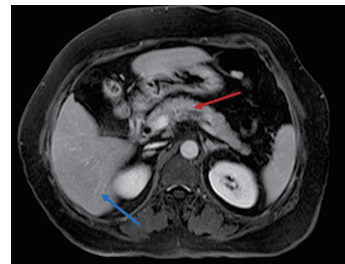


图2 腹部磁共振成像上肝脏(蓝色箭头)与胰腺(红色箭头)的位置关系

Fig 2 Relative position between liver (blue arrow) and pancreas (red arrow) in abdominal magnetic resonance imaging

随着深度学习技术的不断发展,智能医学影像识别在器官识别分割中的应用越来越广泛,提高了模型的识别准确性^[3-4]。但是,器官识别仍然存在结果易受背景像素影响、未充分利用先验知识等问题,有待进一步研究解决。

1.2 血管分割 提取并重建器官周围的血管能使医师对器官周围的血管有清晰、直接的认知,并辅助医师在实施侵入性治疗前规划更佳的路径,避免血管损伤。

血管分割在医学图像处理领域中非常重要,已经成为一个独立的发展方向,发展出了许多面向特定任务和特定模型的分割算法。血管是一种管腔结构,传统的管腔分割方法大致可以分为自上向下(近心端—远心端)和自下向上(远心端—近心端)两类^[5]。自上向下的方法分为2步,第1步是通过检测得到种子点,第2步是从种子点开始迭代地对图像进行分割,直至算法运行至血管末梢或满足预设的终止条件。该方法对图像中小血管的分割难度较大、效果较差。自下向上的方法则主要是利用图像全局信息,并考虑了形状、位置以及尺度等约束因素的影响。因此该类算法可以很好地解决存在血管分叉和低对比度情况的分割问题^[6],实用性较强,并且可以兼容许多解决图像优化的算法。但该算法消耗的计算资源多,对硬件配

置要求较高。

当前, 深度学习算法也被应用到医学图像的血管分割中, 如使用人工提取特征进行学习的冠状动脉分割方法^[7], 三维卷积神经网络 (convolutional neural network, CNN; 图 3) 可以实现对胸腔 MRI 图像中胸部肌肉和计算机断层扫描血管造影 (computed tomography angiography,

CTA) 图像中冠状动脉的分割^[8], 如果将机器学习方法与形状模型的方法结合起来, 可以获得更高的分割精度^[9]。但是受限于当前计算水平, 三维 CNN 的计算复杂度仍然较高, 而血管在图像中所占体积过小, 其提取到的特征易受背景元素的影响, 从而降低分割效果。

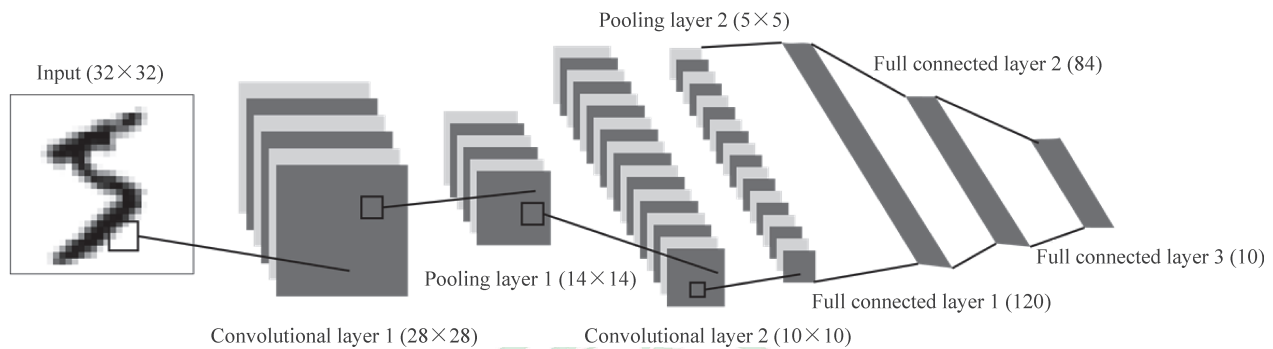


图 3 卷积神经网络模式图

Fig 3 Schematic diagram of convolutional neural network

1.3 基于医学影像的疾病诊断 疾病的病理分类是当前智能医学影像研究的重要问题。早期的病理分类通常分为 3 步, 第 1 步是在影像中人工标注目标区域, 第 2 步是对分割出来的区域进行识别分类, 第 3 步是对整个诊断结果进行宏观的判断。Goldbaum 等^[10]最早阐述了这一基本思路, 随后很多学者在这一思路的基础上进行基于医学影像的病理分类, 构建了基于浅层结构的分类器^[11-13]。随着 CNN 的不断发展, 分类器越来越强大^[14], 新的算法可以直接端到端 (图像端到结果端) 地对图像进行分类并检测物体。但目前可利用的医学影像图像数据量通常较少, 为深度神经网络的训练带来了难题。

迁移学习的理论是解决该方法之一。迁移学习是指在一种数据库中训练得到的较为普适、低层次的特征在不同模型之间进行迁移, 使之可以在一定的基础上完成不同的任务, 从而大大降低需要的样本量。在迁移过程中, 锁定低层次的网络参数, 而仅对高层次的参数进行反向传播训练。Keremany 等^[15]采用这一手段对视网膜的光学相干断层扫描 (optical coherence tomography, OCT) 图像进行了分析, 并在一系列儿科胸片中对算法进行了测试, 结果显示该人工智能系统能够在 30 s 内判断患者是否需要接受治疗, 且准

确率达到了 95% 以上。这一成果被认为是计算机自动病理分类应用的重要突破, 并被 *Cell* 杂志以封面文章的形式进行报道。

1.4 手术视频解析 手术视频解析是智能外科的重要组成部分, 是智能手术的基础。对手术视频的内容进行解析, 让机器理解当前手术视频中的操作, 可以使计算机帮助医师在手术中做出合理的选择, 协助医师规划下一步的手术操作, 并通过比对数据库中的内容揭示医师手术中各个操作的细节。尽管手术视频解析起步较晚, 目前只能应用于一些简单的手术 (如胆囊切除术) 中, 但其已经具备成熟的技术思路和方法。

手术视频解析的研究内容主要包括: (1) 将一段手术的视频根据手术步骤进行分段, 也称为流程分析; (2) 对视频中的特定动作或任务进行检测和识别; (3) 对视频中的器具进行识别、分割和跟踪等。在手术流程分析方面, 可以采集额外的传感器信息, 例如采用额外的相机并利用人工标注提供器具信息、使用无线射频识别设备 (radio frequency identification device, RFID) 标签和电磁学传感器提供器具使用信号^[16]; 或者采用器具的移动轨迹作为特征^[17]。这类算法依赖额外的传感器或特殊的设备, 尽管在技术原理上并不复杂, 但实际投入应用仍有一定困难。Klank 等^[18]利用

强化学习的方法,应用基于支持向量机(support vector machine, SVM)的算法直接从图像中提取特征,在胆囊切除术流程分析实验中达到了50%的准确率,尽管准确率并不理想,但这是最早的一项直接采用图像特征进行流程分析的研究。Dergachyova等^[19]提出了基于人工设计的252维描述子(descriptor)联合复合隐马尔科夫模型(hidden Markov model, HMM)的方法取得了接近90%的准确率,但该准确率依赖于已经得到的器具信息。尽管人工描述子效果尚可,但它们仅能提取专家已经定义过的特征,而无法提取隐藏在图像和时间序列中的深层特征。Twinanda等^[20]提出了一种名为EndoNet的CNN结构,将深度学习提取特征的模式引入到该领域,并使用迁移学习的方法,将预训练的Alexnet模型迁移到医疗领域使用,并配合复合HMM最终取得了92.2%的离线准确率。

在特定动作或任务检测和识别方面,目前应用前景最好的是机器人手术,其主要原因为机器人手术具有详细的器具运动轨迹记录,方便进行特征提取。Lea等^[21]采用了CNN代替人工提取特征,在融合了中、低尺度的视觉信息的同时,HMM融合了大尺度的时间信息,从而形成了多尺度特征融合,其在对特定操作的检测中取得了74%左右的准确率,是目前该领域的基准。特定操作识别中的一个重要问题是关键动作的定义在不同手术中有所不同,因此研究者们开始尝试使用无监督学习的方法,来提取更加普适、可迁移的动作特征。Murali等^[22]采用里程碑学习的方法训练能够提取动作特征的网络。该算法不需要制定具体的动作,也不需要人工标签,将特征向不同的具体任务迁移后分别获得了74%和66%的准确率。Rupprecht等^[23]提出了一种不同的思路,其训练了预测器具三维姿态的模型,再通过预测结果进行动作和任务识别,获得了60%~77%的准确率,几乎与直接利用已知器具运动轨迹的预测结果相当,同时其准确率超过了仅仅使用视觉特征识别的最优算法。

在器具跟踪、检测方面,颜色和色彩梯度是重要的器具识别特征。早在1994年就有研究者报道了直接使用RGB颜色特征对器具进行跟踪和检测的方法^[24],2015年Alshekhali等^[25]使用L*a*b*色彩空间对颜色进行编码并改进了算法。手术器具

的质地纹理和特定的形状是重要的检测依据和特征。人工设计的特征在2016年前后逐渐被机器学习获得的特征取代。

2 智能医学影像识别的展望

智能医学影像识别无论是在疾病诊断还是治疗方面,都有着广泛的应用前景。在疾病的诊断方面,影像诊断占据着重要位置。目前,智能影像诊断在很多疾病中的诊断准确率上已超过临床医师^[15,26-27]。在未来完全可以依靠独立智能影像诊断系统对医学影像及病理切片进行可靠诊断,这将有利于节省医院影像科及病理科的人力成本,并可提高影像诊断和病理诊断的质量和效率。智能影像诊断水平提高及实用化有望实现不同条件下诊断的标准化和优质化。与医师不同,数学模型具备在不同终端无差别传输“诊断经验”的能力。也就是说,无论是在三甲医院还是在基层卫生机构,都可以通过机器学习模型获得同样的诊断。这对于解决我国目前医疗水平差异巨大的现状、促进医疗资源的合理分配以及落实分级诊疗制度等方面都具有重要的意义。

在手术视频解析领域,尽管利用高效、准确的视频解析指导手术尚未实现,但手术视频解析仍有很大的应用空间。例如在外科医师的教育中,低年资医师实际操作的机会较少,外科医师培养周期较长。在手术视频解析的基础上,一方面可配合三维打印等技术,利用模拟手术对外科医学生进行现场教学、评估;另一方面结合现已高度完善的互联网平台,有望发展出集合直播、点播、自主学习等多项功能的在线手术教学多功能平台,从而缩短外科医师培养周期。从远期来看,智能外科的目标是实现智能手术,即完全使用机器人自主实施手术,人类只起到监督决策的作用。当前,在机器视觉的基础上,已有很多相关的研究,如智能手术导航、三维模型与实体器官配准等^[28]。而让计算机“理解”手术是智能手术的第1步,手术视频分析技术的进步将为智能手术的实施打下坚实的基础。

尽管智能医学影像识别有着广泛的应用前景,但其存在的问题依然无法被忽视,如疾病谱窄、诊断准确率仍需提高等。就目前来说,智能医学影像识别存在的问题很大程度上归因于训练数据缺乏且获取难。医学数据具有获取途径、获取条件

单一的特点, 因此具有很高的知识产权价值和经济价值。目前医疗数据的传播大多局限于一所医院、甚至是一个科室。如何整合资源, 如何解决知识产权中存在的问题, 如何满足人工智能技术所需要的大数据量, 是智能医学影像识别研究面临的紧要问题。随着这一问题的解决, 智能影像识别技术必将快速从实验室走向临床, 惠及越来越多的患者。

[参 考 文 献]

- [1] ZHOU Y, XIE L, SHEN W, WANG Y, FISHMAN E K, YUILLE A L. A fixed-point model for pancreas segmentation in abdominal CT scans[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2017: 693-701.
- [2] KARASAWA K, ODA M, KITASAKA T, MISAWA K, FUJIWARA M, CHU C, et al. Multi-atlas pancreas segmentation: atlas selection based on vessel structure[J]. *Med Image Anal*, 2017, 39: 18-28.
- [3] ROTH H R, FARAG A, LU L, TURKB EY E B, SUMMERS R M. Deep convolutional networks for pancreas segmentation in CT imaging[C/OL]//Proc SPIE 9413, Medical Imaging 2015: Image Processing (2015-03-20) [2018-06-20]. doi: 10.1117/12.2081420.
- [4] ROTH H R, LU L, FARAG A, SHIN H C, LIU J, TURKB EY E B, et al. DeepOrgan: multi-level deep convolutional networks for automated pancreas segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015: 556-564.
- [5] SCHNEIDER M, HIRSCH S, WEBER B, SZÉKELY G, MENZE B H. Joint 3-D vessel segmentation and centerline extraction using oblique Hough forests with steerable filters[J]. *Med Image Anal*, 2015, 19: 220-249.
- [6] BREITENREICHER D, SOFKA M, BRITZEN S, ZHOU S K. Hierarchical discriminative framework for detecting tubular structures in 3D images[C]//International Conference on Information Processing in Medical Imaging. IPMI 2013: Information Processing in Medical Imaging. Springer, 2013: 328-339.
- [7] ZHENG Y, LOZICZONEK M, GEORGESCU B, ZHOU S K, VEGA-HIGUERA F, COMANICIU D. Machine learning based vesselness measurement for coronary artery segmentation in cardiac CT volumes[C/OL]. Proc SPIE 7962, Medical Imaging 2011: Image Processing (2011-03-11) [2018-06-20]. doi: 10.1117/12.877233.
- [8] MOESKOPS P, WOLTERINK J M, VAN DER VELDEN B H M, GILHUIJS K G A, LEINER T, VIERGEVER M A, et al. Deep learning for multi-task medical image segmentation in multiple modalities[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2016: 478-486.
- [9] LUGAUER F, ZHENG Y, HORNEGGER J, KELM B M. Precise lumen segmentation in coronary computed tomography angiography[C]//Medical computer vision: algorithms for big data: International Workshop, MCV 2014, held in conjunction with MICCAI 2014, Cambridge, MA, USA, September 18, 2014, revised selected papers. Springer, 2014: 137-147.
- [10] GOLDBAUM M, MOEZZI S, TAYLOR A, CHATTERJEE S. Automated diagnosis and image understanding with object extraction, object classification, and inferencing in retinal images[C]//The 3rd IEEE International Conference on Image Processing. IEEE, 1996: 695-698.
- [11] CHAUDHURI S, CHATTERJEE S, KATZ N, NELSON M, GOLDBAUM M. Detection of blood vessels in retinal images using two-dimensional matched filters[J]. *IEEE Trans Med Imaging*, 1989, 8: 263-269.
- [12] HOOVER A, GOLDBAUM M. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels[J]. *IEEE Trans Med Imaging*, 2003, 22: 951-958.
- [13] HOOVER A, KOUZNETSOVA V, GOLDBAUM M. Locating blood vessels in retinal images by piece-wise threshold probing of a matched filter response[J]. *Proc AMIA Symp*, 1998: 931-935.
- [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems. Association for Computing Machinery, 2012: 1097-1105.
- [15] KERMANY D S, GOLDBAUM M, CAI W, VALENTIM C C S, LIANG H, BAXTER S L, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning[J/OL]. *Cell*, 2018, 172: 1122-1131.e9. doi: 10.1016/j.cell.2018.02.010.
- [16] BARDRAM J E, DORYAB A, JENSEN R M, LANGE P M, NIELSEN K L G, PETERSEN S T. Phase recognition during surgical procedures using embedded and body-worn sensors[C]//2011 IEEE International Conference on Pervasive Computing and Communications (PerCom). IEEE, 2011: 45-53.
- [17] WEEDE O, DITTRICH F, WÖRN H, JENSEN B, KNOLL A, WILHELM D, et al. Workflow analysis and surgical phase recognition in minimally invasive surgery[C]//2012 IEEE International Conference on Robotics and Biomimetics (ROBIO). IEEE, 2013: 1080-1074.
- [18] KLANK U, PADOY N, FEUSSNER H, NAVAB N.

- Automatic feature generation in endoscopic images[J]. *Int J Comput Assist Radiol Surg*, 2008, 3(3/4): 331-339.
- [19] DERGACHYOVA O, BOUGET D, HUAULMÉ A, MORANDI X, JANNIN P. Automatic data-driven real-time segmentation and recognition of surgical workflow[J]. *Int J Comput Assist Radiol Surg*, 2016, 11: 1081-1089.
- [20] TWINANDA A P, SHEHATA S, MUTTER D, MARESCAUX J, DE MATHELIN M, PADOY N. EndoNet: a deep architecture for recognition tasks on laparoscopic videos[J]. *IEEE Trans Med Imaging*, 2017, 36: 86-97.
- [21] LEA C, REITER A, VIDAL R, HAGER G D. Segmental spatiotemporal CNNs for fine-grained action segmentation[C]//European Conference on Computer Vision. ECCV 2016: Computer Vision. Springer, 2016: 36-52.
- [22] MURALI A, GARG A, KRISHNAN S, POKORNY F T, ABBEEL P, DARRELL T, et al. TSC-DL: unsupervised trajectory segmentation of multi-modal surgical demonstrations with deep learning[C]//2016 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2016: 4150-4157.
- [23] RUPPRECHT C, LEA C, TOMBARI F, NAVAB N, HAGER G D. Sensor substitution for video-based action recognition[C]//2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2016: 5230-5237.
- [24] LEE C, WANG Y F, UECKER D R, WANG Y. Image analysis for automated tracking in robot-assisted endoscopic surgery[C]//The 12th International Conference on Pattern Recognition. IEEE, 2002: 88-92.
- [25] ALSHEAKHALI M, YIGITSOY M, ESLAMI A, NAVAB N. Surgical tool detection and tracking in retinal microsurgery[C/OL]//Proc. SPIE 9415, Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling. (2015-03-18) [2018-06-20]. doi: 10.1117/12.2082335.
- [26] YU K H, ZHANG C, BERRY G J, ALTMAN R B, RÉ C, RUBIN D L, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features[J/OL]. *Nat Commun*, 2016, 7: 12474. doi: 10.1038/ncomms12474.
- [27] ESTEVA A, KUPREL B, NOVOA R A, KO J, SWETTER S M, BLAU H M, et al. Corrigendum: Dermatologist-level classification of skin cancer with deep neural networks[J]. *Nature*, 2017, 542: 115-118.
- [28] HASHIMOTO D A, ROSMAN G, RUS D, MEIRELES O R. Artificial intelligence in surgery: promises and perils[J]. *Ann Surg*, 2018, 268: 70-76.

[本文编辑] 孙岩