

放疗影像自动分割效果评估中几何参数与剂量学参数之间的关联性

余行, 刘欢, 傅玉川
四川大学华西医院放疗科, 四川 成都 610041

【摘要】目的:通过分析感兴趣区域(ROI)的几何参数与剂量学参数之间的关联性,探讨放疗影像自动分割效果评估时联合使用几何参数与剂量学参数的必要性。**方法:**利用卷积神经网络构建的分割模型对18例宫颈癌术后患者的危及器官与靶区进行自动分割,把自动分割结果与医生手动勾画结果进行比较,用于评估的几何参数包括基于体积/面积的Dice相似性系数、相对体积差与基于距离的几何参数:最大Hausdorff距离、95% Hausdorff距离、质心差,剂量学参数包括针对危及器官的平均剂量差、针对靶区的 ΔD_{95} 和 ΔD_{98} 。采用线性回归方法研究不同分割方式下ROI几何学参数与剂量学参数间的关系,并使用Spearman相关性分析获得几何参数间的相关性及医生勾画与自动分割间剂量学的相关性。**结果:**所有ROI的几何参数与剂量学参数间的关系均较弱(63.3%的 $R^2 < 0.4$)且不稳定;同时几何参数间的相关系数 $|r|$ 不超过0.625,互为弱相关或不相关。**结论:**在对放疗领域的图像分割结果进行评估时,应该同时考虑到几何参数与剂量学参数。选择几何参数时,应联合基于面积/体积的评估方式与基于距离的评估方式。

【关键词】放射治疗;自动分割;几何参数;剂量学参数

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2021)05-0540-05

Correlation between geometric parameters and dosimetric parameters in the evaluation of image auto-segmentation for radiotherapy

YU Hang, LIU Huan, FU Yuchuan

Department of Radiotherapy, West China Hospital of Sichuan University, Chengdu 610041, China

Abstract: Objective To investigate and discuss the necessity of combining geometric parameters and dosimetric parameters in the evaluation of image auto-segmentation for radiotherapy by analyzing the relationship between geometric parameters and dosimetric parameters of regions of interest (ROI). **Methods** An auto-segmentation model established by convolutional neural network was used for the auto-segmentation of organs-at-risk and target areas for 18 patients who received postoperative radiotherapy for cervical cancer, and then the auto-segmentation results were compared with manual segmentation results. The geometric parameters used for evaluation included volume/area-based parameters (Dice similarity coefficient, relative volume difference) and distance-based parameters (maximum Hausdorff distance, 95% Hausdorff distance, centroid difference), while the dosimetric parameters used for evaluation included the average dose difference for organs-at-risk, as well as the ΔD_{95} and ΔD_{98} for target area. The relationships between geometric parameters and dosimetric parameters of ROI under different segmentation methods were analyzed by linear regression method; and the correlation between geometric parameters and the dosimetric correlation between manual segmentation and automatic segmentation were obtained by Spearman correlation analysis. **Results** The relationship between geometric parameters and dosimetric parameters of all ROI was weak (63.3% of $R^2 < 0.4$) and unstable; and meanwhile, the correlation coefficient $|r|$ between geometric parameters did not exceed 0.625, indicating weakly correlated or not correlated with each other. **Conclusion** Both geometric parameters and dosimetric parameters should be concerned when evaluating the results of image segmentation for radiotherapy, and the combination of area/volume-based evaluation method and distance-based evaluation method should be used for the selection of geometric parameters.

Keywords: radiotherapy; auto-segmentation; geometric parameter; dosimetric parameter

【收稿日期】2021-01-08

【基金项目】四川省科技计划重点研发项目(2020YFS0274)

【作者简介】余行, 硕士, 技师, 研究方向: 医学物理, E-mail: 506640120@qq.com

【通信作者】傅玉川, 博士, 主任技师, 研究方向: 医学物理, E-mail: ychfu@hotmail.com

前言

靶区及危及器官的勾画是放射治疗流程中重要的一步,相较于手动勾画,自动分割具有提高勾画效率、降低勾画差异等优势^[1]。目前,有基于图谱的自动分割商用软件可应用于临床,在对感兴趣区域(Region of Interest, ROI)(主要是危及器官)进行自动分割后,由医生对自动分割结果进行修改并确认^[2]。近年来,基于深度学习的医学影像自动分割方法迅速发展并开始运用于放射治疗靶区及危及器官的自动分割^[3-4],并以其鲁棒性和灵活性显示出巨大的潜力^[5]。

无论是基于图谱的自动分割,还是基于深度学习的自动分割,对分割结果的评估和验证都是将其应用于临床必不可少的环节之一。在评估过程中使用最广泛的轮廓指标是几何指数,包括体积差异、质心距离、一致性指数和边缘分析等^[6]。尽管几何指数计算方便,并且可以选择多种几何参量来进行较为全面的评估,但它们缺乏临床相关意义,不能直接反映勾画差异对靶区和危及器官剂量分布的影响^[7]。

本文利用深度学习模型自动分割生成的靶区及危及器官轮廓与医生手动勾画轮廓进行多参量的对比评估,探讨对自动分割结果进行评估时将几何参数与剂量学参数联合使用的必要性与可行性,同时还分析了几种常用几何评估参数之间的相关性。

1 资料与方法

1.1 影像资料
图像数据来自于2015年2月至2019年7月在四

川大学华西医院放疗科接受治疗的180例宫颈癌患者,分期为IB~IIA期,均于术后接受放射治疗,其中162套CT数据用于训练,18套CT数据用于测试。所有患者均在CT扫描前憋尿以保证膀胱充盈,扫描时采用仰卧位,双手举过头顶,使用头颈肩热塑性面罩固定,并注射对比剂。图像使用西门子模拟定位CT Siemens Somato (Siemens Healthcare, Forchheim, Germany)进行采集,图像重建分辨率为512×512,层厚为3 mm。危及器官包括膀胱、左右股骨头、小肠以及直肠。临床靶区(Clinical Target Volume, CTV)分为原发灶的CTV和盆腔淋巴结的CTV,原发灶CTV包括宫颈肿块、整个宫颈及宫体、宫旁组织及阴道,盆腔淋巴结CTV包括髂总、髂内、髂外、骶前及闭孔淋巴引流区^[8]。

1.2 自动分割模型及结果

使用全卷积神经网络FCN-8s^[9]构建我们的自动分割模型,采用有公开代码和参考模型的开源深度学习仿真工具Caffe^[10],基于ImageNet数据库进行迁移学习完成预训练。使用Python作为编程语言对医学影像DICOM RT文件进行相关的转换、处理及计算。将162例患者的CT图像用于模型的训练,采用预训练与数据扩增避免过拟合,在NVIDIA GTX 1080TI GPU上进行模型训练。利用训练好的模型对18例测试患者的CT图像数据进行靶区和危及器官的自动分割,再将自动分割结果与医生手动勾画轮廓进行比较,表1列出了以几何指数表示的分割结果,每一个几何参数的定义将在第1.4节中详细说明。

表1 以几何参数表示的分割结果
Tab.1 Segmentation results evaluated by geometric parameters

参数	膀胱	左股骨头	右股骨头	直肠	小肠	CTV
DSC	0.92±0.04	0.87±0.04	0.90±0.03	0.76±0.08	0.67±0.09	0.73±0.05
CD	3.63±3.23	3.01±2.61	2.44±1.17	5.38±3.00	27.03±13.93	6.85±4.39
Max HD/cm	4.02±2.52	1.97±2.67	1.26±1.00	3.73±5.78	5.77±2.95	5.00±0.96
95% HD/cm	0.84±0.88	0.65±0.68	0.48±2.38	1.00±4.09	2.86±1.24	1.52±0.42
RVD	-0.05±0.20	-0.08±0.13	0.05±0.08	0.12±0.20	0.36±0.20	-0.09±0.23

DSC: Dice相似性系数; CD: 质心差; Max HD: 最大 Hausdorff 距离; 95% HD: 95% Hausdorff 距离; RVD: 相对体积差

1.3 计划设计
所有计划均在 Raystation 4.7 计划系统上进行设计,计划靶区体积(Planning Target Volume, PTV)处方剂量为 50.4 Gy/28 次,危及器官的剂量限值分别为膀胱、直肠: $V_{40}<60\%$ 、 $V_{50}<50\%$; 左右股骨头: $V_{50}<5\%$; 小肠: $V_{30}<40\%$ 。计划均采用容积旋转调强

(Volumetric Modulated Arc Therapy, VMAT) 计划,采用双全弧照射($180^{\circ}\sim181^{\circ}$, $181^{\circ}\sim180^{\circ}$),对于靶区而言,至少 95% 的靶区体积接受 100% 的处方剂量,靶区热点剂量不超过处方剂量的 107%。

1.4 几何评估参数

参考各类评估方法,我们选取了 5 种广泛应用的

几何评估参数指标^[6,11-12],基于面积/体积的评估方法包括 Dice 相似性系数 (Dice Similarity Coefficient, DSC) 和相对体积差 (Relative Volume Difference, RVD);基于距离的评估方法包括最大 Hausdorff 距离 (Max Hausdorff Distance, Max HD)、95% Hausdorff 距离 (95% Hausdorff Distance, 95% HD) 和质心差 (Centroid Difference, CD)。以医生手动勾画为标准来评价自动分割效果,本文中几何参数的计算都基于开源软件 3D Slicer^[13]。

1.4.1 DSC DSC 是医学图像分割中常用的指标,其计算公式如下所示:

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \quad (1)$$

其中, $|A|$ 代表医生手动勾画的分割区域, $|B|$ 代表自动分割区域, $|A \cap B|$ 代表 A 和 B 的交集。DSC 的数值范围介于 0 到 1 之间,其中 1 表示图像之间完全重叠,0 表示没有重叠,数值越高代表分割效果越好。

1.4.2 RVD RVD 表示分割图像二者的体积差异,其计算公式如下所示:

$$RVD = \frac{|A| - |B|}{|B|} \times 100\% \quad (2)$$

其中, $|A|$ 表示自动分割结果, $|B|$ 表示医生手动勾画结果。

1.4.3 Max HD 对于两个给定的集合 A 和 B , Max HD 定义如下所示:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (3)$$

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (4)$$

其中, $\|a - b\|$ 是欧几里得距离, a 和 b 分别是 A 和 B 边缘上的点,同时 $h(A, B)$ 通常被称为直接 HD。

1.4.4 95% HD 95% HD 与 Max HD 相似,但 95% HD 基于 A 和 B 中边界点之间距离的第 95 个百分位的计算,其目的在于消除极端点对于分割结果的影响。

1.4.5 CD CD 为两组勾画轮廓质心之间的距离,用于评价两组轮廓之间的总体位置偏差,其定义如式(5)所示:

$$CD = \sqrt{(X_{\text{auto}} - X_{\text{manual}})^2 + (Y_{\text{auto}} - Y_{\text{manual}})^2 + (Z_{\text{auto}} - Z_{\text{manual}})^2} \quad (5)$$

其中, $(X_{\text{auto}}, Y_{\text{auto}}, Z_{\text{auto}})$ 为自动分割结构的质心坐标, $(X_{\text{manual}}, Y_{\text{manual}}, Z_{\text{manual}})$ 为医生手动勾画结构的质心坐标。

1.5 剂量评估参数

本研究为回顾性研究,利用已实施的放疗计划,分别获得医生手动勾画轮廓和自动分割轮廓的剂量-体积直方图,如图 1 所示,并提取了多个剂量学参数。对于 CTV,剂量学参数为 ΔD_{95} (95% 靶区体积接受的最小照射剂量差) 与 ΔD_{98} ,对于危及器官,剂量学参

数为 ΔD_{mean} ,其公式如下所示:

$$\Delta D_n = \left| \frac{(D_{n, \text{manual}} - D_{n, \text{auto}})}{D_{\text{prescription}}} \right| \quad (n=95, 98) \quad (6)$$

$$\Delta D_{\text{mean}} = \left| \frac{(D_{\text{mean, manual}} - D_{\text{mean, auto}})}{D_{\text{mean, manual}}} \right| \quad (7)$$

其中, $D_{n, \text{manual}}$ 为医生手动勾画 CTV 剂量, $D_{n, \text{auto}}$ 为自动分割 CTV 剂量, $D_{\text{mean, manual}}$ 为医生手动勾画危及器官剂量, $D_{\text{mean, manual}}$ 为自动分割危及器官剂量。

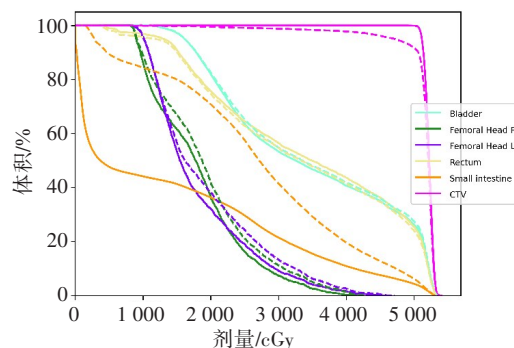


图 1 针对同一位患者医生手动勾画轮廓(实线)和自动分割轮廓(虚线)的剂量-体积直方图

Fig.1 Dose-volume histogram of manual segmentation contours (solid lines) and automatic segmentation contours (dashed lines) for a patient

1.6 统计学方法

所有数据均使用 SPSS 20.0 版 (SPSS, Chicago, IL, USA) 进行分析,计量资料采用均数±标准差表示,线性回归分析探索靶区与危及器官剂量学参数与几何参数的关系,采用 Spearman 相关性分析探究几何参数间的相关性及医生勾画与自动分割剂量学间的相关性。 $P < 0.05$ 为差异有统计学意义。

2 结果

图 2 显示了对于不同 ROI,医生手动勾画轮廓和自动分割轮廓之间的平均剂量对比。几何参数与剂量学参数间关系如图 3 所示。首先,几何参数与剂量学参数之间的关系较弱,如图 3a 所示几何参数与剂量学参数的关系研究中,66.7% (20/30) 的关联系数 $R^2 < 0.6$,63.3% (19/30) 的关联系数 $R^2 < 0.4$ 。靶区及危及器官的轮廓一旦确定,其几何参数就不会改变,而对于剂量学参数而言,它的参数选择较广,我们以靶区为例,分别选择了 ΔD_{95} 、 ΔD_{98} 两个参数作为剂量学参数 (图 3a、图 3b),结果表明剂量学参数的选择影响了关联系数的数值大小,但其关联关系仍然较弱。

其次,几何参数与剂量学参数之间的关系不稳

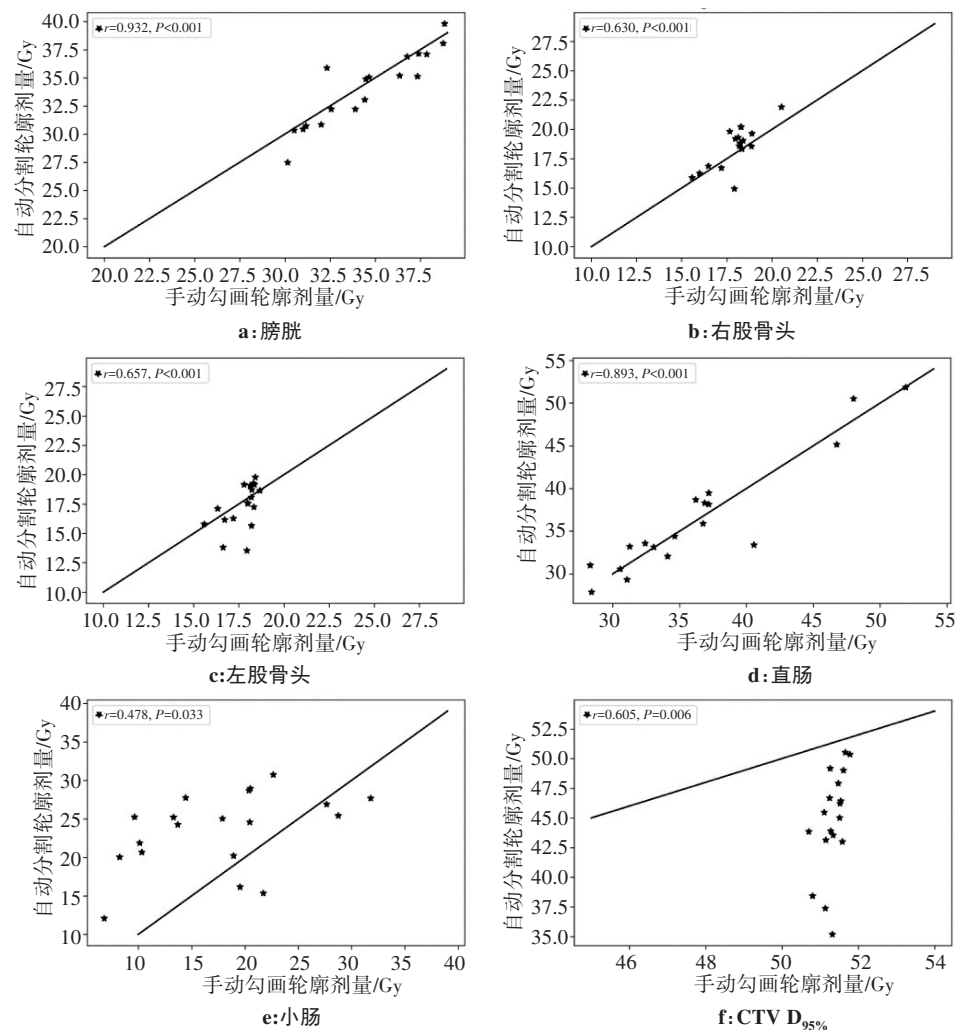


图2 医生手动勾画与自动分割的剂量对比

Fig.2 Comparison of dose parameters between manual segmentation and automatic segmentation

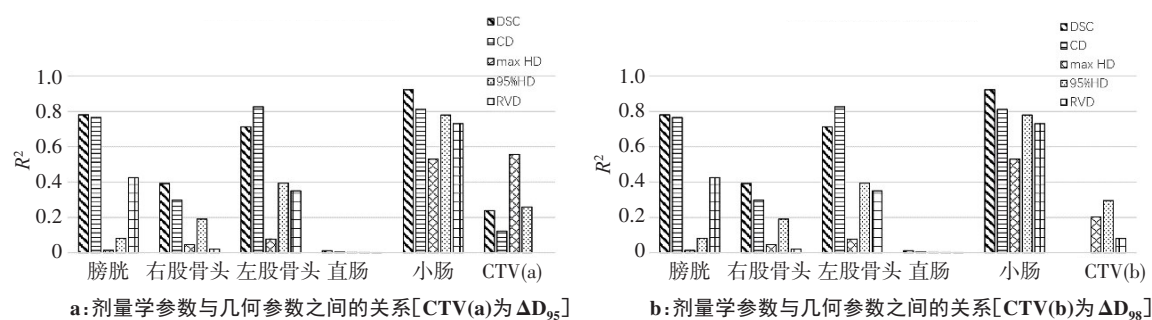


图3 剂量学参数与几何参数之间的关系

Fig.3 Correlation between geometric parameters and dosimetric parameters

定,没有单一的几何参数与剂量学参数始终显示出较高或者较低的相关性(图3a、图3b),小肠的平均剂量 ΔD_{mean} 与DSC的关联系数 R^2 为最高值0.922,而同样对于DSC而言,CTV的 ΔD_{95} 与其关联系数 R^2 仅为0.239,低于CTV与Max HD的关联系数0.556。

表2所示为几何参数间的关系,相关系数 $|r|$ 的数值不超过0.625,CD与95% HD的关系系数 $|r|$ 最大,其余几何参数间的 $|r|$ 均小于0.5。

表2 几何参数间的相关系数r

Tab.2 Correlation coefficient r between geometric parameters

参数	DSC	CD	Max HD	95% HD	RVD
DSC	-	-0.353	-0.098	-0.414	0.386
CD	-0.353	-	0.168	0.625	0.126
Max HD	-0.098	0.168	-	0.133	-0.158
95% HD	-0.414	0.625	0.133	-	0.004
RVD	0.386	0.126	-0.158	0.004	-

3 讨论

本研究利用基于caffe搭建的全卷积神经网络FCN-8s对宫颈癌外照射放疗中的危及器官和靶区进行自动分割,就常用的几何评估常数而言,其分割结果优于临床上常用的ABAS分割结果^[14],因此用该模型的分割结果进行评估参数的关联性研究是可行的。

在这项研究中,我们发现大多数几何指数与关键剂量学终点之间关联较弱且不一致,这表明使用几何参数估算剂量学参数是不可接受的。剂量与几何评估方法呈弱相关;几何评估方法容易出现假阳性结果,这一般发生在危及器官上,即对于几何学参数而言,分割结果不可以接受,然而就剂量学而言,其满足临床要求。一般来说,当DSC大于0.7时,可认为分割结果是可以接受的^[15]。Kaderka等^[16]研究表明,有时候即使DSC值很低,其剂量学参数也符合临床要求,特别是对于小体积的危及器官。而对于靶区,评估容易出现假阴性结果,即靶区具有可以接受的几何结果但其剂量结果不符合临床要求,这是由于在靶区周围具有较高的剂量梯度,相同程度的轮廓改变,其剂量影响大于对危及器官的影响。

通常在对图像分割进行评估时,会选用不同的几何参数对图像分割进行评估。在我们的研究中,选取了评价标准中常见的5个参数^[17-20],它们可以分为两大类,一种是基于体积/面积的评估方式例如DSC、RVD;另一种是基于距离的评估方式如Max HD、95% HD及CD。DSC容易受结构体积/面积的影响,而对边缘信息不敏感。例如两组结构有相同的重叠体积/面积,而总体积/面积的不同,将会导致其DSC值不同。基于距离的评估方式如Max HD对错误的点非常敏感^[17]。我们的研究分析了几种不同几何评估参数间的关系,结果表明几种参数间为弱相关(95% HD与CD)或者不相关。因此在对分割结果进行评估时,可以分别使用不同的参数。

本研究的不足之处在于,对于危及器官,我们仅提取了平均剂量作为剂量学参数,尽管对于靶区而言,我们选择了 ΔD_{95} 、 ΔD_{98} 作为剂量学参数,但是没有临床上常见的CI指数。我们的结果表明,不同参数的选择会对几何参数与剂量学参数的相关系数造成影响,但是其趋势及相关性强弱大致不变。因此,在后续研究中,将考虑纳入更多的剂量学参数。

4 结论

在对放疗领域的图像分割结果进行评估时,应该同时考虑到几何参数与剂量学参数。选择几何参数时,应联合使用基于面积/体积的评估方式与基于距离的评估方式。

【参考文献】

- [1] WU D, MAYR N, KARATAŞ Y, et al. Interobserver variation in cervical cancer tumor delineation for image-based radiotherapy planning among and within different specialties[J]. J Appl Clin Med Phys, 2005, 6(4): 106-110.
- [2] 阴晓娟, 胡彩容, 张秀春, 等. 基于图谱库的ABAS自动勾画软件在头颈部肿瘤中的可行性研究[J]. 中华放射肿瘤学杂志, 2016, 11(25): 1233-1237.
- [3] YIN X J, HU C R, ZHANG X C, et al. The feasibility study of atlas-based autosegmentation (ABAS) software in head-and-neck cancer[J]. Chinese Journal of Radiation Oncology, 2016, 11(25): 1233-1237.
- [4] MEN K, CHEN X, ZHANG Y, et al. Deep deconvolutional neural network for target segmentation of nasopharyngeal cancer in planning computed tomography images[J]. Front Oncol, 2017, 7: 315.
- [5] KAZEMIFAR S, BALAGOPAL A, NGUYEN D, et al. Segmentation of the prostate and organs at risk in male pelvic CT images using deep learning[J]. Biomed Phys Eng Express, 2018, 4(5): 055003.
- [6] 傅玉川, 余行. 医学影像自动分割技术在放射治疗中的应用及其发展趋势[J]. 中国医疗器械杂志, 2020, 44(5): 420-424.
- [7] FU Y C, YU H. Application and development trend of medical image automatic segmentation technology in radiation therapy[J]. Chinese Journal of Medical Instrumentation, 2020, 44(5): 420-424.
- [8] HANNA G G, HOUNSELL A R, O'SULLIVAN J M. Geometrical analysis of radiotherapy target volume delineation: a systematic review of reported comparison methods[J]. Clin Oncol, 2010, 22(7): 515-525.
- [9] LIM T Y, GILLESPIE E, MURPHY J, et al. Clinically oriented contour evaluation using dosimetric indices generated from automated knowledge-based planning[J]. Int J Radiat Oncol Biol Phys, 2019, 103(5): 1251-1260.
- [10] SMALL J W, MELL L K, ANDERSON P, et al. Consensus guidelines for delineation of clinical target volume for intensity-modulated pelvic radiotherapy in postoperative treatment of endometrial and cervical cancer[J]. Int J Radiat Oncol, 2008, 71(2): 428-434.
- [11] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans Pattern Anal, 2014, 39(4): 640-651.
- [12] JIA Y, SHELHAMER E, DONAHUE J, et al. Caffe: convolutional architecture for fast feature embedding[C]//Proc ACM Int Conf Multimedia. 2014: 675-678.
- [13] ROACH D, JAMESON M G, DOWLING J A, et al. Correlations between contouring similarity metrics and simulated treatment outcome for prostate radiotherapy[J]. Phys Med Biol, 2018, 63(3): 035001.
- [14] FECHTER T, ADEBAHR S, BALTAS D, et al. Esophagus segmentation in CT via 3D fully convolutional neural network and random walk[J]. Med Phys, 2017, 44(12): 6341-6352.
- [15] KIKINIS R, PIEPER S D, VOSBURGH K G. 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support[M]. New York: Springer, 2014.
- [16] 王金媛, 郑庆增, 张慧娟, 等. 宫颈癌危及器官MIM-Maestro和RayStation软件自动勾画优势对比研究[J]. 中华肿瘤防治杂志, 2019, 26(24): 1889-1894.
- [17] WANG J Y, ZHENG Q Z, ZHANG H J, et al. Comparative study of two software tools on the atlas-based auto-segmentation of organs-at-risk in cervical cancer[J]. Chinese Journal of Cancer Prevention and Treatment, 2019, 26(24): 1889-1894.
- [18] ZIJDENBOS A, DAWANT B, MARGOLIN R, et al. Morphometric analysis of white matter lesions in MR images[J]. IEEE Trans Med Imaging, 1994, 13(4): 716-724.
- [19] KADERKA R, GILLESPIE E, MUNDT R, et al. Geometric and dosimetric evaluation of atlas based auto-segmentation of cardiac structures in breast cancer patients[J]. Radiother Oncol, 2018, 131: 215-220.
- [20] 何奕松, 蒋家良, 余行, 等. 影像分割中Dice系数和Hausdorff距离的比较[J]. 中国医学物理学杂志, 2019, 36(11): 1307-1311.
- [21] HE Y S, JIANG J L, YU H, et al. Comparison of Dice similarity coefficient and Hausdorff distance in image segmentation[J]. Chinese Journal of Medical Physics, 2019, 36(11): 1307-1311.
- [22] TAHAA A A, HANBURY A. Metrics for evaluating 3D medical image segmentation: analysis, selection and tool[J]. BMC Med Imaging, 2015, 15: 29.
- [23] YEGHIAZARYAN V, VOICULESCU L. Family of boundary overlap metrics for the evaluation of medical image segmentation[J]. J Med Imaging, 2018, 5(1): 015006.
- [24] WANG Z, WANG E, ZHU Y. Image segmentation evaluation: a survey of methods[J]. Artif Intell Rev, 2020, 53(1): 5637-5674.

(编辑:陈丽霞)