

## 基于神经网络对CT上腹部脏器边界识别及其 临床应用中的检测效能\*

夏春潮, 张 凯, 刘秀民, 蒲 进, 彭婉琳, 徐 旭, 曾令明, 曾 文, 李真林<sup>△</sup>

四川大学华西医院 放射科(成都 610041)

**【摘要】目的** 探讨基于神经网络模型和不同层面组合对CT上腹部影像主要脏器边界进行识别,并检测其临床实际应用效能。**方法** 纳入2018年3月-2019年3月在我院行或包括上腹部增强CT检查且图像质量满足临床诊断要求的病例2 000例,对上腹部8个主要脏器边界层面(肝上下缘、脾上下缘、左肾下缘、右肾下缘、胃下缘和胆囊下缘)进行标注,利用不同神经网络方法和不同层面组合进行模型训练(训练集、验证集和测试集),获得边界识别的准确率,并通过识别50例临床数据检测其准确率和临床应用效能。**结果** 两种模型按照不同权重比例整合的融合模型准确率最高,EfficientNet-b3模型次之,Xception模型最低。各模型中5层图像识别边界的准确率要高于3层图像的准确率,1层图像的准确率最低。融合模型连续5层图像获得肝上缘、肝下缘、脾上缘、脾下缘、左肾下缘、右肾下缘、胃下缘和胆囊下缘的识别准确率分别是91%、87%、92%、85%、92%、95%、76%、74%。融合模型对50例效能检测数据进行识别,获得8个层面准确率为88%、86%、88%、80%、82%、80%、69%、65%;满足临床应用各边界的准确率为98%、98%、95%、98%、99%、98%、80%、77%。**结论** 通过增加边界在连续层面的变化逻辑,按照不同权重比例整合的融合模型识别上腹部CT各边界的准确率最高,并在临床实践中获得较高的临床检查效能。

**【关键词】** 神经网络 深度学习 CT 上腹部 边界识别

### The Clinical Effectiveness of Neural Network-based Boundary Recognition of Upper Abdominal Organs on CT

**Images** XIA Chun-chao, ZHANG Kai, LIU Xiu-min, PU Jin, PENG Wan-lin, XU Xu, ZENG Ling-ming, ZENG Wen, LI Zhen-lin<sup>△</sup>. Department of Radiology, West China Hospital, Sichuan University, Chengdu 610041, China

<sup>△</sup> Corresponding author, E-mail: HX\_lizhenlin@126.com

**【Abstract】 Objective** To assess the clinical effectiveness of boundary recognition of upper abdomen organs on CT images based on neural network model and the combination of different slices. **Methods** A total of 2 000 patients who underwent upper abdomen enhanced CT scans from March 2018 to March 2019 were included in the study. The quality of the CT images met the requirements for clinical diagnosis. Eight boundary layers (the upper and lower edge of liver, the upper and lower edge of spleen, the lower edge of left kidney, the lower edge of right kidney, the lower edge of the stomach and the lower edge of the gallbladder) of the main organs in the upper abdomen were labeled. The model training (training set, verification set and test set) based on different neural network methods and combinations of different slices were then performed to assess the accuracy of boundary recognition. Furthermore, clinical data from 50 cases were used as test group for assessing the accuracy and clinical effectiveness of this model. **Results** The fusion model created by integrating the two models according to different weight ratios yielded the highest accuracy, and then followed the EfficientNet-b3 model, with the Xception model showing the lowest accuracy. In each model, the boundary recognition accuracy of 5-slice image is higher than that of 3-slice image, and that of 1-slice image is the lowest. The recognition accuracy of fusion model of the 5-continuous-slice image for upper edge of liver, lower edge of liver, upper edge of spleen, lower edge of spleen, lower edge of left kidney, lower edge of right kidney, lower edge of stomach and lower edge of gallbladder was 91%, 87%, 92%, 85%, 92%, 95%, 76% and 74%, respectively. The fusion model was checked with the effectiveness data of 50 cases, yielding 88%, 86%, 88%, 80%, 82%, 80%, 69%, and 65% accuracy for 8-slice image, respectively, and the accuracy of meeting clinical application requirement was as high as 98%, 98%, 95%, 98%, 99%, 98%, 80% and 77%, respectively. **Conclusion** By increasing boundary change logics in the continuous slices, the fusion model integrating different weight proportions demonstrates the highest accuracy for identifying the boundary of upper abdominal organs on CT images, achieving high examination effectiveness in clinical practice.

**【Key words】** Neural networks Deep learning CT Upper abdomen Boundary recognition

目前,上腹部CT检查已经广泛应用于临床,成为上

腹部疾病检查的重要手段之一。2006年,原卫生部办公厅要求有条件的医院间逐步开展医学影像检查结果互认工作<sup>[1]</sup>。紧接着,国务院和各省份又推出多项通知和指导意见促进检查互认的实施<sup>[2]</sup>。但影像学检查重复率仍较

\* 四川省科技计划项目(No. 2019YFS0522)和四川大学华西医院学科卓越发展1.3-5工程项目(No. ZYGD18019)资助

<sup>△</sup> 通信作者, E-mail: HX\_lizhenlin@126.com

高,不同等级医院间医疗检查结果互认程度较低<sup>[3]</sup>。目前,人工智能(artificial intelligence, AI)在上腹部CT影像学中的研究主要集中在器官分割、病变识别和鉴别诊断。肝脏分割是肝病医学图像诊断中的关键基础技术,是计算机辅助诊断领域中肝脏病灶目标检测的基础<sup>[4-5]</sup>,也是三维重建的关键步骤<sup>[6-7]</sup>,分割结果的准确度直接影响及决定后续更高阶的图像处理操作质量<sup>[8]</sup>。图像分割的准确基于完整的CT图像数据,脏器边界的准确识别是确保上腹部CT影像数据完整的第一步。到目前为止,国内外未见通过AI在上腹部CT影像边界识别来保证扫描时影像数据完整性的相关研究。本研究基于不同神经网络模型对CT上腹部各主要脏器的边界进行识别和检测其效能,并探讨其在临床应用中的价值。

## 1 资料与方法

### 1.1 研究对象

模型建立: 纳入2018年3月-2019年3月于我院放射科因病情需要行或包括上腹部CT检查的患者2000例,其中男性1198例,女性802例,年龄18~97岁,平均(56±14)岁。模型效能测试: 纳入2019年4月-2019年5月在我院放射科因病情需要行或包括上腹部CT检查的患者50例,其中男性29例,女性21例,年龄19~80岁,平均年龄(57±15)岁。排除标准: ①年龄<18周岁; ②无法配合屏气; ③腹腔巨大肿物; ④有胸腹腔积液; ⑤未包全8个主要脏器层面。本研究经四川大学华西医院生物医学伦理委员会批准,批准号为2019年审(206)号。

### 1.2 扫描方法

采用西门子SOMATOM Definition Flash CT和SOMATOM Definition AS+ CT成像仪,患者仰卧位,双手上举,头先进或脚先进。扫描范围包括膈顶到肝下缘,管电压为120 kV,管电流使用CARE Dose4D技术,视野(FOV)为(320~400) mm×(320~400) mm,螺距为0.9,转速0.5 s/圈,卷积核采用软组织卷积核(B40r medium),层厚和层间距均为7 mm。

### 1.3 数据标记

**1.3.1 图像获取** 标准: ①三维容积DICOM数据,对患者的姓名、年龄、住院号等涉及隐私的信息进行脱敏; ②图像清晰,满足临床诊断要求,无明显的运动伪影; ③上腹部主要脏器的边界清晰可见; ④包括8个主要脏器边界及每个边界层的上两层和下两层图像。

**1.3.2 标记方法** 由经验丰富的技师对上腹部CT数据进行肝上下缘、脾上下缘、左肾下缘、右肾下缘、胃下缘和胆囊下缘8个主要脏器的边界层面进行标注,包全上腹部

各个主要器官,对有异议的图像进行协商确定。利用DICOM信息中的层数信息,自动确定各层面的上下层。

### 1.4 模型训练流程

**1.4.1 图像类型转换** 将DICOM图像转换成JPG图像: 同一幅图的CT值按不同窗宽、窗位生成三个灰度图转换成对应的三幅JPG图,图像矩阵为350×350,保留图像的绝大部分信息。基于一般彩色图像由RGB三个通道组成的原理,使用OpenCV视觉库中的图像处理的相关算法,将单层CT按不同窗宽窗位生成的A(300, 40)、B(210, 210)、C(70, 70)三张灰度图像,分别对应RGB三个通道,合成彩色图D(图1)。

在显存为16G的GPU的电脑上,直接使用分辨率350×350作为网络输入,单张图像预测的时间为0.03 s左右,达到实时运算的速度。

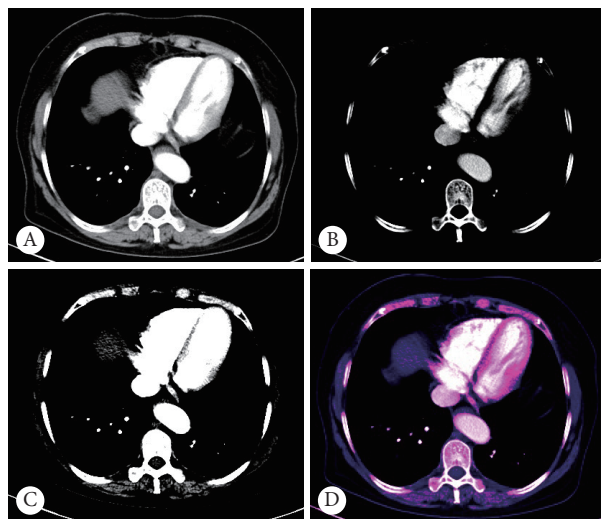


图1 A、B、C为同一幅CT图像不同窗宽窗位下转换成的灰度图, D为利用A、B、C三个通道(红、黄、蓝)合成的彩图

Fig 1 A, B and C are the grayscale CT images in the same location with different window width and level. D is the generated color image of red, yellow and blue derived from A, B and C

**1.4.2 层面逻辑组合** 将JPG图根据图像的连续性及组织器官边界的渐变规律转换为A、B、C三组。其中A组: 标记层(0); B组: 将标记层的上一层(-1)、标记层(0)及标记层的下一层(1)组合; C组: 将标记层的上上层(-2)、标记层的上一层(-1)、标记层(0)、标记层的下一层(1)和标记层的下下层(2)组合。窗宽窗位为350/40。B、C组的组合方式也是基于一般彩色图像由RGB三个通道组成的原理,使用OpenCV视觉库中的图像处理的相关算法,将多层CT按照不同层面分别对应RGB三个通道,合成彩图D(图2)。3组图像分成训练集、验证集、测试集,样本例数分配比例为7:2:1。

**1.4.3 模型建立和测试** 采用基于深度学习的Xception

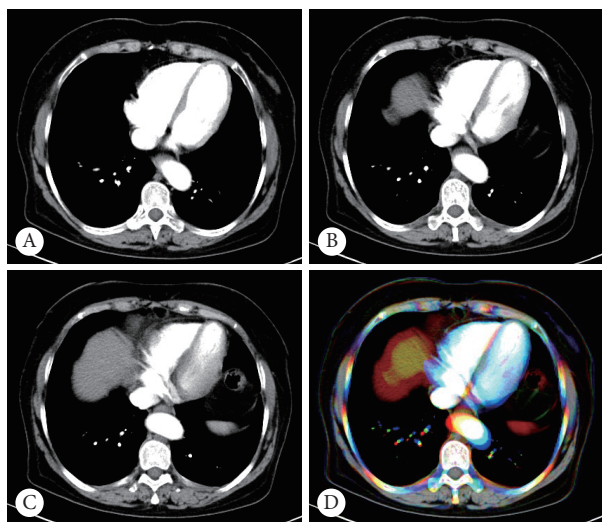


图 2 A 为边界层的上一层, B 为边界层, C 为边界层的下一层, D 为 A、B、C 利用 JPG 图像分三个通道(红、黄、蓝)合成的彩图

Fig 2 B is the boundary slice, while A and C were the slice above and the slice under, respectively. D is the generated color image of red, yellow and blue derived from A, B and C

模型<sup>[9]</sup>和 EfficientNet-b3 模型<sup>[10]</sup>分别对 8 个主要层面进行建模,两种模型按照不同权重比例整合成获得准确率最高的模型。以标注图像为金标准,计算模型对各层面识别的准确率。准确率=正确识别某一边界数量÷标记某一边界的总数量×100%

**1.4.4 模型效能验证** 利用准确率最高的模型对 50 例上腹部增强 CT 影像数据进行边界识别,同时上缘边界识别为边界层及其上一层,下缘边界识别为边界层及其下一

层认定为满足临床应用,计算其准确率和满足临床应用的准确率。准确率=正确识别某一边界数量÷标记某一边界的总数量×100%。

**1.4.5 统计学方法** 数据采用  $\bar{x} \pm s$  表示。计量资料的比较采用独立样本  $t$  检验,计数资料的比较采用卡方检验,  $P < 0.05$  为差异有统计学意义。

## 2 结果

### 2.1 患者特征

本研究共纳入 2 000 例上腹部 CT 检查增强图像,其中用于训练集 1 400 例,验证集 400 例,测试集 200 例。用于模型效能测试患者 50 例。性别、年龄方面差异均无统计学意义 ( $P > 0.05$ )。

### 2.2 深度学习模型测试结果

卷积神经网络的 Xception 算法、EfficientNet-b3 算法和两种算法按照不同权重比例整合而成的合成算法,结合不同层面逻辑关系,分别完成上腹部各主要脏器边界识别模型的建立。3 种模型在单层、3 层和 5 层的图像建立的模型的准确度呈上升趋势(图 3)。Xception 模型识别的准确率相对较低, EfficientNet-b3 模型有明显的提升,两种模型按照不同权重比例整合的合成的模型准确率最高(表 1)。利用连续 5 层图像识别边界的准确率要高于连续 3 层图像的准确率,单层图像的准确率最低。合成模型通过连续 5 层图像的识别方法获得肝上缘、肝下缘、脾上缘、脾下缘、左肾下缘、右肾下缘、胃下缘和胆囊下缘的

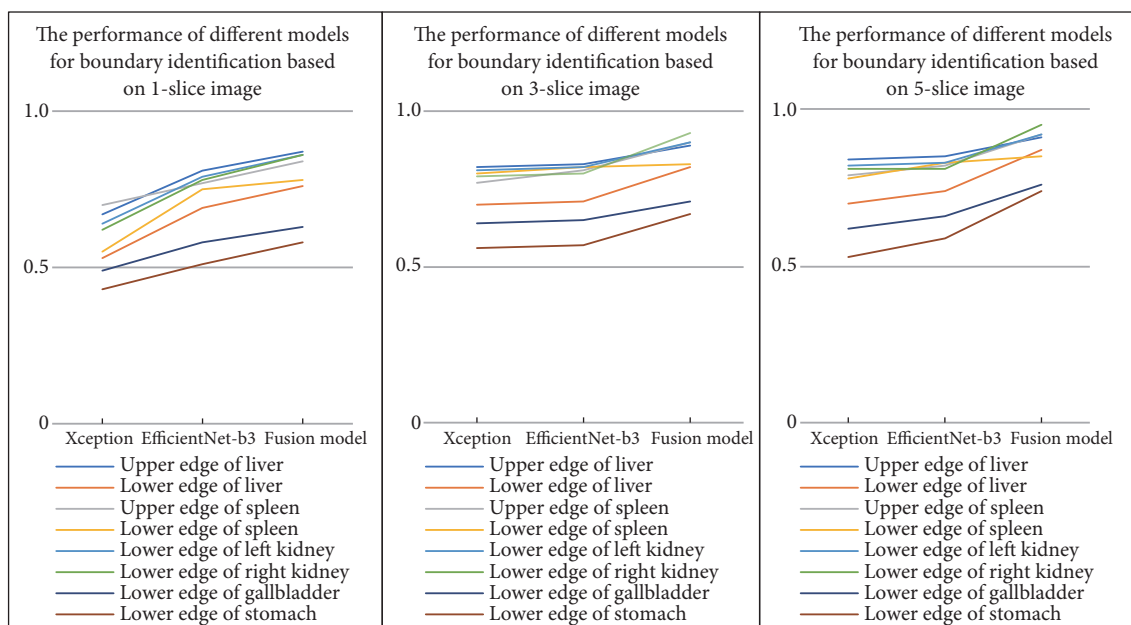


图 3 基于单层、3 层和 5 层图像建立的不同模型对各边界识别的准确率

Fig 3 The performance of different models for boundary identification based on 1-, 3- and 5-slice images

Accuracy ranged 0-1.0, on the left side of each figure.

表 1 不同模型对2000例上腹部各主要脏器边界识别的准确度

Table 1 The diagnostic performance of different models for identifying the boundary of organs in upper abdomen of 2 000 patients

| Edge                       | Case/layer  | Accuracy |         |         |                 |         |         |              |         |         |
|----------------------------|-------------|----------|---------|---------|-----------------|---------|---------|--------------|---------|---------|
|                            |             | Xception |         |         | EfficientNet-b3 |         |         | Fusion model |         |         |
|                            |             | 1-slice  | 3-slice | 5-slice | 1-slice         | 3-slice | 5-slice | 1-slice      | 3-slice | 5-slice |
| Upper edge of liver        | 2 000/6 000 | 0.67     | 0.82    | 0.84    | 0.81            | 0.83    | 0.85    | 0.87         | 0.89    | 0.91    |
| Lower edge of liver        | 2 000/6 000 | 0.53     | 0.70    | 0.70    | 0.69            | 0.71    | 0.74    | 0.76         | 0.82    | 0.87    |
| Upper edge of spleen       | 2 000/6 000 | 0.70     | 0.77    | 0.79    | 0.77            | 0.81    | 0.82    | 0.84         | 0.90    | 0.92    |
| Loweredge of spleen        | 2 000/6 000 | 0.55     | 0.80    | 0.78    | 0.75            | 0.82    | 0.83    | 0.78         | 0.83    | 0.85    |
| Lower edge of left kidney  | 2 000/6 000 | 0.64     | 0.81    | 0.82    | 0.79            | 0.82    | 0.83    | 0.86         | 0.90    | 0.92    |
| Lower edge of right kidney | 2 000/6 000 | 0.62     | 0.79    | 0.81    | 0.78            | 0.80    | 0.81    | 0.86         | 0.93    | 0.95    |
| Lower edge of gallbladder  | 2 000/6 000 | 0.49     | 0.64    | 0.62    | 0.58            | 0.65    | 0.66    | 0.63         | 0.71    | 0.76    |
| Lower edge of stomach      | 2 000/6 000 | 0.43     | 0.56    | 0.53    | 0.51            | 0.57    | 0.59    | 0.58         | 0.67    | 0.74    |

识别准确率分别高达是91%、87%、92%、85%、92%、95%、76%、74%。

2.3 融合模型的检测效能

融合模型对50例临床上腹部增强数据,各主要边界分别为150个层面进行效能检测,识别的具体层面数见表2。肝上缘、肝下缘、脾上缘、脾下缘、左肾下缘、右肾下缘、胃下缘和胆囊下缘8个层面的进行效能检测后的准确率为88%、86%、88%、80%、82%、80%、69%、65%。满足临床实际应用的各边界准确率为98%、98%、95%、98%、99%、98%、80%、77%。

表 2 融合模型对50例(150层)三维数据识别的具体层面

Table 2 The specific slice of 3-D data recognition of the fusion model in 50 patients (150 layers)

| Edge                       | Total/layer | Upper slice/layer | Boundary slice/layer | Lower slice/layer |
|----------------------------|-------------|-------------------|----------------------|-------------------|
| Upper edge of liver        | 150         | 15                | 131                  | 4                 |
| Lower edge of liver        | 150         | 17                | 130                  | 3                 |
| Upper edge of spleen       | 150         | 7                 | 132                  | 11                |
| Loweredge of spleen        | 150         | 4                 | 120                  | 26                |
| Lower edge of left kidney  | 150         | 1                 | 123                  | 26                |
| Lower edge of right kidney | 150         | 3                 | 120                  | 27                |
| Lower edge of gallbladder  | 150         | 30                | 103                  | 17                |
| Lower edge of stomach      | 150         | 34                | 98                   | 18                |

3 讨论

我国医疗现况存在医疗资源、医疗质量的异质化的问题<sup>[1]</sup>。随着医学影像云平台和5G网络的推广和应用,医学影像报告和数据随处可查,给检查互认、远程诊断和技术指导带来新的挑战 and 机会。影像资料互认是基础,

规范影像检查技术,确保影像资料质量。目前,质量控制管理模式中图像质量评价主要是影像医师和技师的主观评价,通过plan-do-check-action(PDCA)循环实现<sup>[12]</sup>。由于医师或技师间存在质控标准不一致和主观偏见,导致图像质量可能不满足临床及远程诊断。AI在医学影像领域展现美好前景的同时,也面临许多挑战和瓶颈,如数据质量问题,国内尚未就影像图像质量和规范实现完全的同质化,不同种类和质量的数据势必会影响AI模型的准确性和普适性<sup>[11]</sup>。如何获取高质量的影像数据是当前行业发展的关键。上腹部CT扫描包涵多个器官,标准化的上腹部DICOM数据是影像诊断的关键和前提。但由于扫描技师的技术水平参差不齐,临床工作量大,导致部分图像不能达到影像诊断要求。脏器边界识别属于图像分类,卷积神经网络作为AI最主要的计算模式。本研究将基于卷积神经网络的Xception(Extreme Inception)和EfficientNet-b3算法按照不同权重比例整合成的合成模型对各脏器主要器官边界进行识别。

本研究各模型分别对边界单层、上下三层和上下五层的不同层面组合进行建模,发现在同一种层面组合中,EfficientNet-b3算法获得模型的准确率较Xception算法获得模型的准确率高,按照不同权重比例整合的融合模型的准确率最高。其原因是EfficientNet模型是Google公司通过机器学习搜索算法得来,通过仔细地平衡网络深度(depth)、宽度(width)和分辨率(resolution),以便更有效地扩展,获得更好的性能。鉴于16G GPU的电脑的运算速度,将分辨率为512×512的DICOM图像转换为分辨率为350×350的JPG图像,既保存大部分有效影像信息,又能实现快速运算。在EfficientNet的多种算法中,预实验测试发现EfficientNet-b3算法获得的准确率最高。而

Xception是Google对Inception v3的另一种改进,主要是采用深度可分离卷积(depthwise separable convolution)来替换原来Inception v3中的卷积操作。其特征提取部分由36层卷积层组成,包含14个模块,线性Residual connection也在这14个模块中。合成模型采用了集成学习中对多个模型的预测结果做加权平均的方法。集成学习方法是指组合多个模型,以获得更好的效果,使集成的模型具有更强的泛化能力。而实际应用中,模型的精度往往有较大差异,通常情况下会更重视准确率好的模型的结果,并赋予其更大的权重值,将多个弱学习模型提升为强学习模型。本研究发现在同一种模型中,通过增加连续层面数,利用OpenCV的通道切分合并获得的图像含有各层的信息,即增加了景深。只对边界层一层进行建模发现准确率一直不高,原因可能是由于人体各上腹部脏器结构存在个体差异、临界层和上下层的图像比较相近、平扫和增强的特征变化等。本研究利用边界层及其上下相关层边界连续变化的逻辑关系,发现脏器临界层识别精度有所提高。

本研究利用融合模型对50例上腹部增强CT的患者数据图像进行模型效能检测。发现8个层面的准确率稍低于模型的准确率,边界识别都在边界层、边界层上一层和边界层下一层。在临床实际应用中,为确保包全整个上腹部,上腹部CT扫描的上下范围会多扫几层。脏器上缘的边界层边界及上一层和下缘边界及下一层均可认为是满足临床边界识别要求,各边界识别的准确率为77%~99%。单张图片识别耗时0.03 s,能及时反馈上下界的识别情况。在上腹部CT检查过程中,技师可根据边界识别情况及时调整扫描范围,获得更加准确的定位,减少辐射剂量,降低重复扫描;同时可以确保远程传输图像的完整性,提高医疗效率,有利于促进影像检查互认。

本研究也存在一定的局限性。第一,本研究只对涉及上腹部上下范围的主要脏器的8个边界进行识别,并未对各脏器的所有上下边界进行识别。第二,本研究纳入的都是满足临床诊断要求的数据,未对病变、强化效果和伪影等进行识别,在未来的研究中需进一步实现。第三,效能测试的数据量相对较少,未来会继续增加病例数,以

获得更加客观的模型效能。

## 参 考 文 献

- [1] 卫生部办公厅关于医疗机构间医学检验、医学影像检查互认有关问题的通知. 中华人民共和国卫生部公报, 2006(3): 68[2019-07-19]. <https://kns.cnki.net/kcms/detail/detail.aspx?DBCode=CJFD&DBName=CJFDN0508&fileName=WSGB200603033>.
- [2] 肖晓华, 梁恒斌, 谢欣睿, 等. 我国医疗检查结果互认制度实施现状. 中国医院管理, 2019, 39(5): 12-13.
- [3] 肖晓华, 苏茹茹, 潘振威, 等. 二三级医院影像学检查结果互认现状及影响因素研究. 中国医院管理, 2019, 39(5): 20-22.
- [4] CHLEBUS G, SCHENK A, MOLTZ J H, *et al.* Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing. *Sci Rep*, 2018, 8(1): 15497[2019-07-13]. <https://www.nature.com/articles/s41598-018-33860-7>. doi: 10.1038/s41598-018-33860-7.
- [5] LI X, CHEN H, QI X, *et al.* H-DenseUNet: hybrid densely Connected UNet for liver and tumor segmentation from CT volumes. *IEEE Trans Med Imaging*, 2018, 37(12): 2663-2674.
- [6] WANG Z, ZHANG C, JIAO T, *et al.* Fully automatic segmentation and three-dimensional reconstruction of the liver in CT images. *J Healthc Eng*, 2018, 2018: 6797102[2019-08-08]. <https://doi.org/10.1155/2018/6797102>.
- [7] ZOU H, XUE H, TAO Y. Liver Three-dimensional reconstruction accurately predicts remnant liver volume for HBV-related hepatocellular carcinoma prior to hepatectomy. *Indian J Surg*, 2017, 5: 1-6[2019-08-08]. <https://link.springer.com/article/10.1007/s12262-017-1645-8>. doi: 10.1007/s12262-017-1645-8.
- [8] 林天武, 邹春莉, 吴佩琪. CT图像处理中肝脏分割技术研究进展. 中国继续医学教育, 2019, 11(21): 71-73.
- [9] CHOLLET F. Xception: deep learning with depthwise separable convolutions. *arXiv E-prints*, 2016(2016-10-07)[2019-08-08]. <https://arxiv.org/abs/1610.02357>. doi: 10.1109/CVPR.2017.195.
- [10] TAN M, LE Q V. EfficientNet: rethinking model scaling for convolutional neural networks. 2019(2019-05-28)[2019-08-08]. <https://arxiv.org/abs/1905.11946>.
- [11] 刘士远, 萧毅. 基于深度学习的人工智能对医学影像学的挑战和机遇. 中华放射学杂志, 2017, 51(12): 899-901.
- [12] 关策, 杜伟, 庄惠丹. 影像科管理中的质量控制与质量保证. 中国医院管理, 2007, 27(12): 49.

(2020-10-13收稿, 2021-01-31修回)

编辑 吕 熙