

利用 U-net 算法在头 CT 平扫图像上分割脑梗死的初步探究

李晓庆, 王可欣, 额·图娅, 李昌欣, 王祥鹏, 张晓东, 王霄英

【摘要】 目的:基于深度学习方法训练辅助诊断模型,探究对头 CT 平扫(NCHCT)图像中脑梗死病灶自动分割的可行性。**方法:**搜集自 2018 年 5 月 1 日—2020 年 8 月 31 日影像诊断报告中诊断印象包含“梗死”关键词的 NCHCT 连续病例 1185 例,筛选出最终证实为脑梗死的数据 362 例。由两位医师结合最终结果手工标注脑梗死区域。将数据按 8:1:1 的比例随机分为训练集(training set, n=288)、调优集(validate set, n=37)和测试集(test set, n=37 例)。训练 2D U-net 模型分割脑梗死病灶,分割结果自动输入到结构化报告中。对测试集的预测结果和人工标注的结果进行比较,统计梗死病灶体积、径线的差异,使用 Dice 相似系数(DSC)、体积相似度(VS)和 Hausdorff 距离(HD)评价模型的预测效能。使用 Bland-Altman 评价模型预测的病灶体积、径线和 CT 值与手工标注的一致性。**结果:**测试集中平均 DSC 为 0.66 (95% CI:0.60~0.72),平均 VS 为 0.75 (95% CI:0.69~0.82),平均 HD 为 39.69 mm (95% CI:32.38~47.01)。Bland-Altman 图显示模型预测与手工标注对病灶大小和 CT 值测量的一致性较高,体积、径线和 CT 值数据点位于 95% 一致性界限(95% limits of agreement, 95% LoA)外的数据为 2.8%~11.1%。**结论:**基于深度学习的辅助诊断模型可用于分割 NCHCT 中的脑梗死病灶,并自动生成报告,对患者分诊有一定作用。

【关键词】 深度学习; 体层摄影术, X 线计算机; 脑梗死; 研究报告

【中图分类号】 R-056; R814.4; R743.33; R-05 **【文献标识码】** A

【文章编号】 1000-0313(2022)06-0669-07

DOI:10.13609/j.cnki.1000-0313.2022.06.001

开放科学(资源服务)标识码(OSID):



A preliminary study of cerebral infarction segmentation on CT images based on U-Net algorithm

LI Xiao-qing, WANG Ke-xin, E Tu-ya, et al. Department of Radiology, Peking University First Hospital, Beijing 100034, China

【Abstract】 Objective: To explore the feasibility of cerebral infarction segmentation on non-contrast head CT (NCHCT) images by using deep learning algorithms. **Methods:** Totally 1185 NCHCT cases, diagnosed as "infarction" in radiologic reports, were retrospectively collected from May 1, 2018, to August 31, 2020. In these cases, 362 cases were finally included in our study, which clinically confirmed as cerebral infarction. First, the infarction ranges of these cases were manually labeled by two experienced radiologists. Then, these cases were randomly allocated to the training set (n=288), validation set (n=37), and test set (n=37), and were trained by 2D U-net model. Meanwhile, the prediction results were automatically output to the structured report of the patient. Finally, the infarction volume, diameter and CT value, which predicted by 2D U-net model, were compared with the manual labeled results, by using Bland-Altman plot. Besides, the dice similarity coefficient (DSC), volume similarity (VS), and Hausdorff distance (HD) were used to evaluate the efficacy of the model. **Results:** The Bland-Altman plot showed the results predicted by 2D U-net model were mostly consistent with manual labeled results, and 2.8% to 11.1% data within the 95% limits of agreement (95% LoA). Besides, the average DSC, VS, and HD in the test set were 0.66 (95% CI:0.60~0.72), 0.75 (95% CI:0.69~0.82), and 39.69mm (95% CI:32.38~47.01), respectively. **Conclusion:** It is feasible to segment cerebral infarction on NCHCT by using deep learning algorithms. And the automatic structured reporting could be clinically used for the triage of patients with stroke.

【Key words】 Deep learning; Tomography, X-ray computed; Brain infarction; Research report

我国现患脑卒中人数世界第一,脑卒中已成为导致成年居民发死亡和残疾的首位病因^[1],其中又以脑梗死(缺血性脑卒中, ischemic stroke)占比为多。颅脑 CT 平扫(non-contrast head CT, NCHCT)是目前疑似脑卒中患者的首选影像学检查方法^[2],也是医学影像科最常见的检查项目之一。在急诊或常规非急诊场景下,对于临床疑似存在脑血管事件风险的患者,在首先除外出血性脑卒中之后,使用辅助诊断软件在大量 NCHCT 受检者中快速筛选出脑梗死阳性发现患者并增加其诊疗优先级,提高影像科医生诊断效率,为临床医师尽快选择治疗干预决策提供帮助,有利于优化实际医疗工作中的诊疗流程。近年来,人工智能(artificial intelligence, AI)技术在医学影像中的应用探索越来越多,对图像分类^[3,4]、脏器分割^[5,6]、病灶检出^[7]等均有较多应用,在提升诊断准确性的同时提高工作效率。

本研究拟使用深度学习算法研发影像智能辅助诊断工具,对疑诊脑卒中患者的 NCHCT 中存在的脑梗死病灶进行分割,自动生成结构化报告,有助于患者快速分诊,进而优化临床卒中管理模式。

材料与方 法

本研究为回顾性研究,获得了医院伦理委员会的批准,按照本单位人工智能 AI 模型训练规范执行研究方案。

1. 用例定义

根据本院人工智能(artificial intelligence, AI)项目管理方法,首先定义研发 CT 颅脑平扫分割 AI 模型

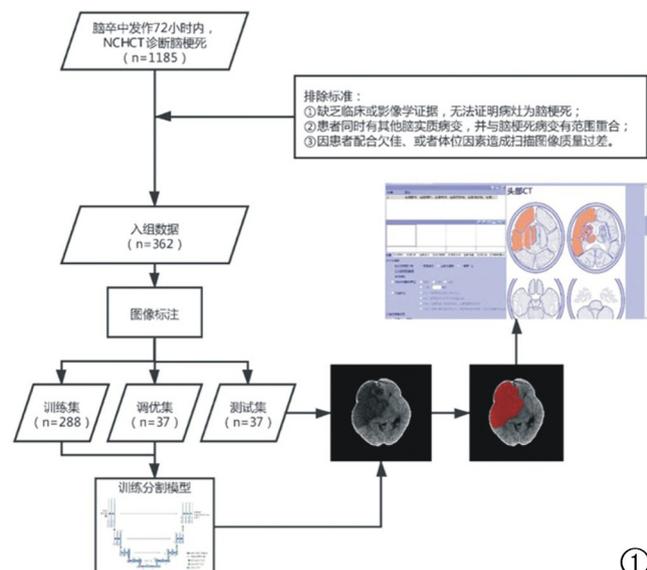


图 1 研究数据入组和训练流程。

的用例,包括: AI 模型的名称(identity document, ID)、临床问题、场景描述、模型在实际工作中的调用流程、模型输入输出数据结构等。AI 生成的结果自动返回到结构化报告的“病灶列表”(图 1)。

2. 回顾队列建立

从本院影像归档和通讯系统(picture archive and communication system, PACS)搜集检查时间 2018 年 5 月 1 日—2020 年 8 月 31 日共 1185 例影像学诊断中包含“梗死”的颅脑 CT 平扫图像,由两名影像医师再次读片,结合近期 MR 或 CT 随访、临床病程证实为脑梗死。入组标准:脑卒中发作时间 <72 h。排除标准:①缺乏临床或影像学证据,无法证明病灶为脑梗死;②患者同时有其他脑实质病变,并与脑梗死病变有范围重合;③因患者配合欠佳、或者体位因素造成扫描图像质量过差。经上述标准筛选病例,最终共有 362 例 NCHCT 检查纳入本研究。头颅 CT 平扫图像采集于不同的 CT 设备(表 1)。

表 1 CT 图像采集参数

设备制造商	Siemens	GE Medical Systems	GE Medical Systems
设备型号	Light Speed VCT	SOMATOM Definition Flash	Discovery CT750 HD
图像性质	CT	CT	CT
管电压(kV)	120.0	120.0	120.0
层厚(mm)	5.0	5.0	5.0
层间距(mm)	5.0	5.0	5.0

3. 图像标注

在进行图像标注前,所有医学数字成像和通信(digital imaging and communications in medicine, DICOM)格式的 CT 图像都被转换为神经信息技术协议(neuroimaging informatics technology initiative, NIFTI)格式。由一位具有 2 年工作经验的影像科医生使用 ITK-SNAP(version3, Philadelphia, PA)软件对脑梗死病灶进行标注,并由一位有 15 年以上工作经验的影像医师修改和确认。标注时结合 CT 图像特征并参考患者临床综合信息以确定脑梗死病灶的位置。CT 图像特征包括可识别的早期脑梗死征象,如脑实质低密度、灰白质界限消失、脑回肿胀、脑沟变浅、岛带征,豆状核模糊及大脑中动脉高密度征^[8]等。临床综合信息包括 NCHCT 检查时的临床表现、CT 灌注成像(computed tomography perfusion imaging, CTP)、CT 血管造影(computed tomography angiography, CTA),或 NCHCT 检查后 1 个月内的 MR 或 CT 复查影像,以及患者的临床诊断、治疗记录等。对于发病时影像中见的陈旧脑梗死病灶或软化灶也给予标注。

4. 模型训练

将 362 例脑梗死患者的 NCHCT 检查随机按 8:1:1 的比例随机分为训练集 (288 例)、调优集 (37 例) 和测试集 (37 例)。深度学习的模型为 2D U-Net^[9]。将 NIFTI 格式的 NCHCT 图像分辨率设置为 256×256 。采用随机噪声、平移、左右翻转、仿射变换等图像扩增方法。使用 ADAM 梯度下降优化算法, 初始学习率 (learning rate) 为 1×10^{-4} , 每次读取的图像数量 (batch size) 为 24。训练次数 (Epoch) 为 200, 研究所使用的程序语言为 Python。模型训练的硬件为 GPU NVIDIA Tesla P100 16G, 软件包括 Python3.6、Pytorch 0.4.1、Opencv、Numpy、SimpleITK 等。

5. 模型评价

模型输出结果包括脑梗死病灶区域, 使用最小体积包围盒算法^[10]计算的病灶的三维径线, 基于预测标签计算的病灶体积和 CT 值, 生成的结果自动输入到结构化报告中。

以 Dice 相似系数 (dice similarity coefficient, DSC) 评估模型预测结果与手工标注的重合度, 计算公

式: $DSC(A, B) = 2 \cdot |A \cap B| / |A| + |B|$, 其中 A 是手工标注的体素数, B 是模型分割的体素数。以体积相似度 (volume similarity, VS) 评估模型预测的病灶体积与手工标注病灶体积的一致性, 以 Hausdorff 距离 (Hausdorff distance, HD) 评价模型预测的病灶与手工标注病灶形态相似度。

6. 统计方法

使用 R4.1.0 软件进行统计分析。符合正态分布的计量资料以均值 ± 标准差表示, 不符合正态分布的连续变量表示为中位数 (四分位间距) 及中位数的 95% 置信区间。使用 Bland-Altman 评价模型预测的病灶体积、径线和 CT 值与手工标注的一致性。

结果

1. 研究数据特征

所有 362 例患者平均年龄 (68.9 ± 13.7) 岁, 均为后续能证实存在脑梗死病变的患者。患者的年龄、性别等临床特征如表 2。

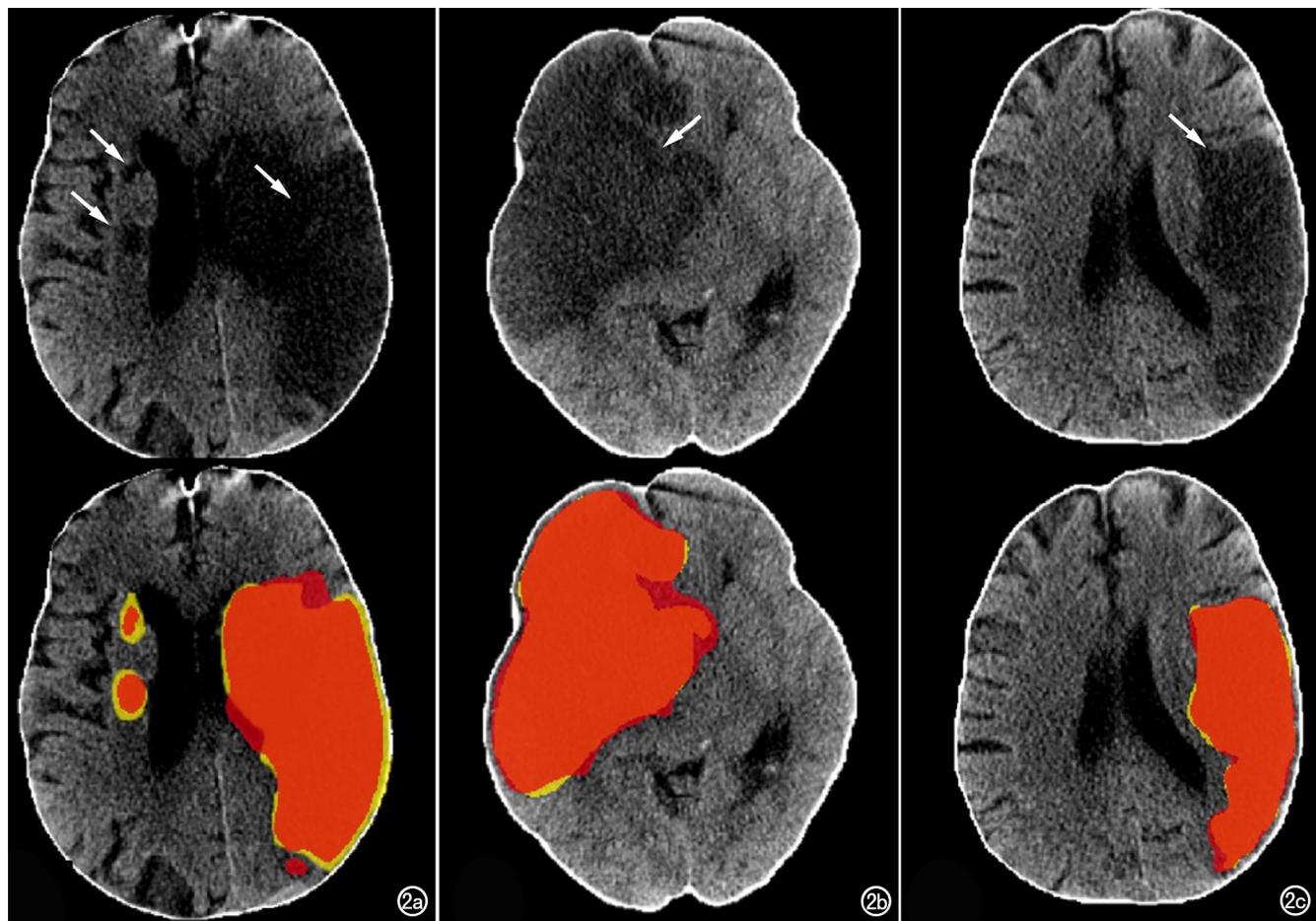


图 2 测试集中的脑梗死分割结果。白箭指示脑梗死病灶, 黄色区域为手工标注范围, 红色区域为模型预测范围。a) 左侧额顶叶、放射冠大面积梗死及右侧放射冠小面积梗死, 主要病灶均被识别, 预测范围与手工标注范围相差较小; b) 右侧额顶颞岛叶、基底节区、胼胝体大面积脑梗死, 主要病灶被识别, 预测范围与手工标注范围相差较小; c) 左侧额顶叶、放射冠大面积梗死, 主要病灶被识别, 预测范围与手工标注范围相差较小。

表2 入组病例临床信息

参数	训练集	调优集	测试集
检查数	288	37	37
平均年龄(岁)	68.7±13.8	70.5±12.9	68.6±13.8
按P去重性别比(男:女)	156:95	25:12	23:14

2. 模型预测结果

在测试集共 37 例检查图像中, 31 例 DSC 值 ≥ 0.5 , 主要脑梗死病灶均被识别, 且预测范围与手工标注范围相差相对较小(图 2)。另外的 6 例 DSC 值 < 0.5 , 其中(图 3): ①4 例预测病变范围相对手工标注范围小, 或者存在多个病灶仅能预测其中部分(图 3a); ②3 例将邻近的脑白质脱髓鞘病变包含入脑梗死病灶

预测范围(图 3b); ③2 例将病变临近脑沟低密度包含入脑梗死病灶预测范围(图 3c); ④1 例将贴近颅底的低密度错误预测为脑梗死病灶(图 3d); ⑤1 例将不对称的单侧基底节区结构错误预测为脑梗死病变(图 3e)。此外, 尚有 1 例数据未能检出梗死灶(图 3f)。

在测试集所包含的 37 例病例中, 依照连通域计算共 97 个病灶, 其中有 82 个病灶位于前循环供血区, 占测试集数据的 84.5%; 15 个病灶位于后循环供血区, 占测试集数据的 15.5%。上述病灶具体累及各脑区的分布如表 3。

3. 模型定量评估

测试集中平均DSC为0.66(95%CI:0.60~0.72),

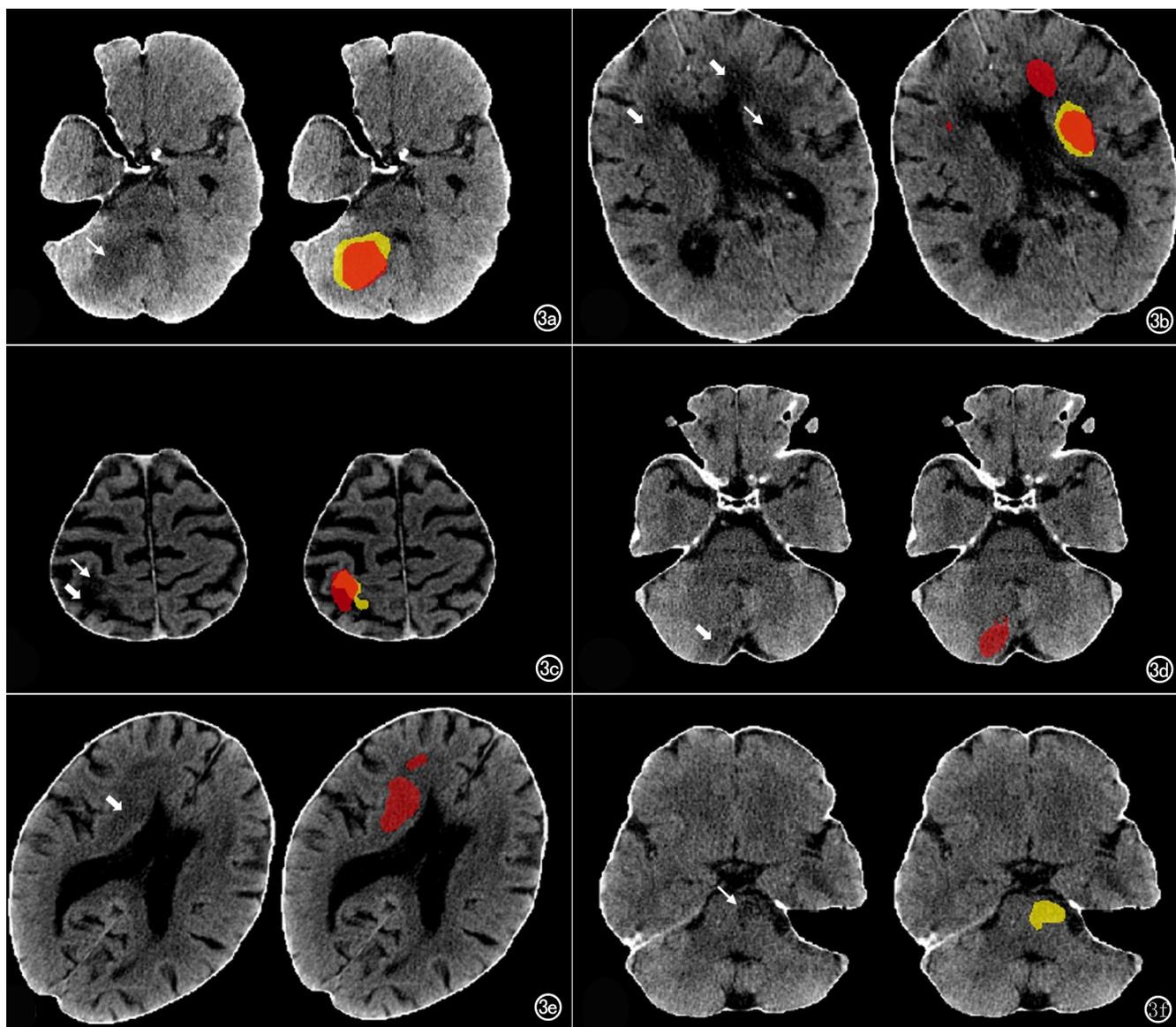


图3 测试集中DSC值 < 0.5 的6例脑梗死分割结果。细白箭指示脑梗死病灶,粗白箭指示非脑梗死病灶的低密度区域,黄色区域为手工标注范围,红色区域为模型预测范围。a)模型预测范围相对手工标注范围小;b)模型将邻近的脑白质脱髓鞘病变包含入脑梗死病灶预测范围;c)模型将病变临近脑沟低密度包含入脑梗死病灶预测范围;d)模型将贴近颅底的低密度误预测为脑梗死病灶;e)模型将不对称的单侧基底节区结构错误预测为脑梗死病变;f)模型未能检出梗死灶。

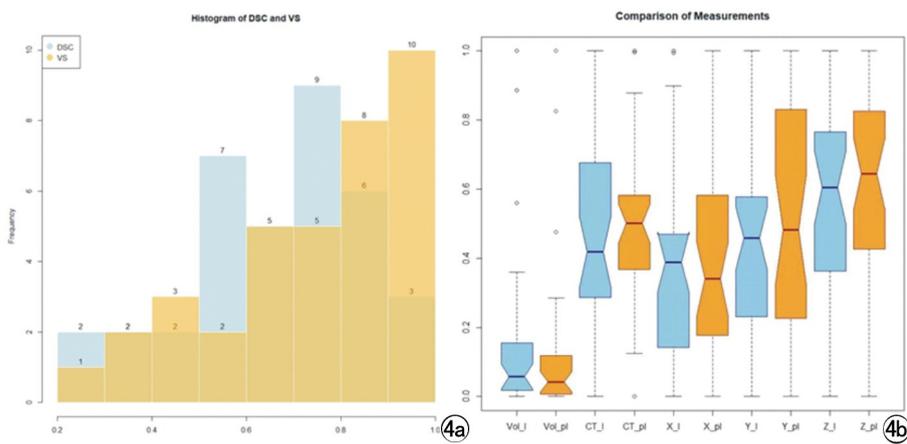


图4 测试集定量评估。a)DSC和VS的分布。b)测量值的缺口箱线图。从左到右的5组数据分别代表病灶体积、CT值、最小径、中径、最长径的参考值(蓝色)和预测值(黄色)的统计量;这5组数据的预测值和参考值中位数的95%置信区间有重叠。

表3 测试集中病灶的分布

病灶累及的脑区	病灶数量(个)
额叶	29
顶叶	21
颞叶	14
岛叶及基底节区	47
丘脑	3
脑干	3
小脑	6
枕叶	3

平均 VS 为 0.76 (95% CI: 0.69 ~ 0.82), 平均 HD 为 39.69mm (95% CI: 32.38 ~ 47.01)。计算模型输出的病灶体积、径线和 CT 值与相对应的手工分割测量值的中位数及其 95% 置信区间, 缺口箱线图 (图 4) 显示两组间各个观察值中位数的 95% 置信区间有重叠^[11]。Bland-Altman 分析显示模型预测与手工标注对病灶大小和 CT 值测量的一致性较高, 2.8% ~ 11.1% 的测量值位于 95% 一致性界限 (95% limits of agreement, 95% LoA) 外 (图 5)。

讨论

脑卒中中具有发病时间早、发病人数多和患病人群庞大等特点^[12], 随着老龄化加剧, 脑卒中成为我国医疗系统乃至社会需要面对的重要问题之一。NCHCT 是脑卒中的首选检查方法^[13], 影像医生应从 NCHCT 图像中尽早识出脑梗死征象, 从而辅助临床医师及时制定治疗干预决策, 这对改善病人预后具有积极影响^[14]。

本研究在 NCHCT 上分割脑梗死病灶, 是在前期



图5 测试集中手工标注和模型预测的 Bland-Altman 分析。a) 病灶体积 B-A 分析结果; b) 病灶 CT 值 B-A 分析结果; c) 病灶最小径 B-A 分析结果; d) 病灶中径 B-A 分析结果; e) 病灶最长径 B-A 分析结果。

NCHCT 正常、异常二分类的基础上^[15],进一步用 AI 方法检出脑梗死病变,为使用 AI 技术全面评估 NCHCT 图像进行阶段性探索。本研究结果显示使用深度学习模型对于脑梗死的早期识别诊断具有可行性,值得进一步深入研究。研究者观察到模型在部分病例中预测的范围与手工标注的范围有出入,除外手工标注的问题后,我们认为这可能与模型自身的局限性相关;此外,部分病例模型预测存在误诊,模型将脑室周围脱髓鞘病变、脑沟或者不对称的基底节区结果误认为脑梗死,研究者分析认为可能与这些病变与梗死在 NCHCT 中同样显示为相对低密度,且与主要病灶位置相近有关。对于不能成功预测的 1 例病灶,研究者推测与病灶位于脑干,邻近硬化伪影重相关;此外,对于位于颅底、头顶的病灶模型的预测也存在偏差,研究者考虑除伪影等因素外,与病变自身像素占图层比例小可能有关。

本研究目前的定位是作为辅助分诊工具,在扫描完成后即刻筛查出有明确梗死灶的患者,发出提示信息,并将结果生成到结构化报告中。在结构化报告脑梗死重点病灶评估的“CT 平扫表现”模块中,左侧分类列项提示梗死病灶在 NCHCT 中可能存在的影像学表现,包括脑实质密度减低、灰白质界限模糊、脑沟消失、大脑中动脉致密征及出血转化等,而位于右侧的图示将同时提供在听眦线层面、鞍上池层面、室间孔层面、半卵圆中心等主要层面的病变位置信息(图 1)。这一流程有利于缩短报告时间,提高救治效率。

近年来,关于脑梗死的影像辅助诊断研究是本领域的热点问题之一。Aneta Lisowski 等^[16,17]使用卷积神经网络(convolution neural network, CNN)方法检测 NCHCT 中的缺血性病变,认为使用该方法检测 NCHCT 中的缺血性病变是可行的。Wu 等^[18]使用了一种机器学习(machine learning, ML)方法对 NCHCT 中的急性缺血性脑卒中病变进行分割,证实了该方法在 NCHCT 图像中分割结果与 DWI 序列显示的急性缺血性卒中病灶具有较好的一致性。Paul Bentley 等^[19]的 ML 研究基于影像预测中风治疗结果,提示 ML 预测卒中预后较传统方法有更高的预测性能。Manuel 等^[20]对于 NCHCT 中脑梗死进行分类,研究了数据集偏差可能存在的影响。与既往类似研究比较,本研究同样验证了用 AI 检出 NCHCT 中缺血性脑卒中病变的可行性,提示了 AI 在医学影像识别方面可能的广袤前景。本研究关注重点是病灶的形态学分割,并得到基于体素的定量测量值;本研究与既往类似研究相比数据量较大,分割结果较好;并且本研究中模型生成的结果未来可自动输入到结构化报告中,对促进临床流程、优化方法有一定的作用。

本研究存在许多不足之处。第一,脑梗死病灶的选取和标注有一定局限性。尽管本研究在可行范围内严格规范了入组病例的入选及排除标准,结合影像与临床信息对病灶进行标注,但本研究的参考标准并非真正的病理诊断,可能对研究结果造成一定影响,这也是多数同类研究的共同局限^[14,21,22]。同时,由于我们回顾性搜索并选择影像学诊断包含“梗死”的病例,因而最终纳入研究的脑梗死病灶不包含人眼难以识别的超急性期脑梗死,对于这部分病变的识别我们的模型仍存在欠缺。第二,部分患者出现神经系统症状时的 CT 图像中不仅有急性脑梗死病灶,且同时存在陈旧梗死病灶。后者可能与当前症状无关,但考虑到这些病灶也属于图像识别的任务,并可以反映患者脑血管的背景状态,所以最终研究将其一并纳入标注。因此,模型输出的结果既有新鲜梗死,也有陈旧梗死,未来研究者将对此开展工作,进一步将二者区分开来,更有利于临床治疗决策的制定。第三,对急性脑梗死的评估,临床需要基于定位信息完成半定量评估指标(如 Alberta Stroke Program Early CT score, ASPECTS 评分)^[23],本研究未给出 ASPECTS 评分,这是下一步研究的内容。最后,本研究作为初步研究,仅使用了回顾性数据验证模型的效能,未来还应进行外部验证,以进一步评估模型的泛化性能。

总之,基于深度学习算法训练的辅助诊断模型可用于分割 NCHCT 中的脑梗死病灶并自动生成报告,对脑卒中患者影像检查后分诊有一定作用。

致谢:感谢北京赛迈特锐医学科技有限公司张虽虽在结构化报告工作中的贡献。

参考文献:

- [1] 王亚楠,吴思缈,刘鸣.中国脑卒中 15 年变化趋势和特点[J].华西医学,2021,36(6):803-807.
- [2] 中华医学会神经病学分会,中华医学会神经病学分会脑血管病学组.中国急性缺血性脑卒中诊治指南 2018[J].中华神经杂志,2018,51(9):666-682.
- [3] 张晓东,邢倩,韩超,等.基于机器学习的低剂量胸部 CT 肺结节分类和预后随访的研究[J].临床放射学杂志,2020,39(10):1962-1966.
- [4] 额·图娅,郭小超,王可,等.基于深度学习训练上腹部 DCE-MRI 扫描期相分类模型的可行性[J].放射学实践,2021,36(8):1052-1058.
- [5] 孙兆男,崔应谱,林志勇,等.U-Net 模型在 CT 图像实现肾实质和肾窦分割及体积和径线测量[J].放射学实践,2020,35(10):1303-1309.
- [6] 王可,姜原,黄嘉豪,等.基于深度学习 MR 体部脂肪组织的自动分割和定量测量研究[J].放射学实践,2021,36(6):792-798.
- [7] 刘想,谢辉辉,许玉峰,等.AI 软件自动检出胸部 CT 图像上肋骨骨折的诊断效能研究[J].临床放射学杂志,2021,40(7):1369-1374.
- [8] 国家卫生健康委员会脑卒中防治工程委员会神经影像专业委员

- 会,中华医学会放射学分会神经学组.脑血管病影像范化应用中国指南[J].中华放射学杂志,2019,53(11):916-940.
- [9] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [EB/OL]. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234-241, DOI: 10.1007/978-3-319-24574-4_28, 2015.
- [10] Barequet G, Har-Peled S. Efficiently approximating the minimum-volume bounding box of a point set in three dimensions [J]. J Algorithms, 2001, 38(1):91-109.
- [11] Nuzzo RL. The box plots alternative for visualizing quantitative data [J]. PM&R, 2016, 8(3):268-272.
- [12] Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the global burden of disease study 2019 [J]. The Lancet, 2020, 396(10258):1204-1222.
- [13] Mccarty JL, Leung LY, Peterson RB, et al. Ischemic infarction in young adults: a review for radiologists [J]. Radiographics, 2019, 39(6):1629-1648.
- [14] Potter CA, Vagal AS, Goyal M, et al. CT for treatment selection in acute ischemic stroke: A code stroke primer [J]. Radiographics, 2019, 39(6):1717-1738.
- [15] 张晓东, 刘想, 谢辉辉, 等. 使用深度学习模型对头颅 CT 平扫图像进行二分类的初步探讨 [J]. 实用放射学杂志, 2020, 36(10):1670-1675.
- [16] Lisowska A, O Neil A, Dilys V, et al. Context-aware convolutional neural networks for stroke sign detection in non-contrast CT scans [C]. Cham: Springer International Publishing, 2017: 723, 494-505.
- [17] 娄昕, 何建风, 马林. 脑血管病医学影像人工智能研究现状与展望 [J]. 中华放射学杂志, 2020, 54(5):389-390.
- [18] Qiu W, Kuang H, Teleg E, et al. Machine learning for detecting early infarction in acute stroke with non-contrast-enhanced CT [J]. Radiology, 2020, 294(2):638-644.
- [19] Bentley P, Ganesalingam J, Carlton Jones AL, et al. Prediction of stroke thrombolysis outcome using CT brain machine learning [J]. NeuroImage: Clinical, 2014, 4:635-640.
- [20] Schultheiss M, Noël P B, Riederer I, et al. Towards subject-level cerebral infarction classification of CT scans using convolutional networks [J]. PLOS ONE, 2020, 15(7):e0235765.
- [21] Wardlaw JM, Mielke O. Early signs of brain infarction at CT: observer reliability and outcome after thrombolytic treatment-systematic review [J]. Radiology, 2005, 235(2):444-453.
- [22] Kanekar SG, Zacharia T, Roller R. Imaging of stroke: Part 2, pathophysiology at the molecular and cellular levels and corresponding imaging changes [J]. Am J Roentgenol, 2012, 198(1):63-74.
- [23] Soun JE, Chow DS, Nagamine M, et al. Artificial intelligence and acute stroke imaging [J]. Am J Neuroradiol, 2021, 42(1):2-11.

(收稿日期:2021-08-30 修回日期:2022-01-30)