# 基于深度学习建立颈椎病 MR 诊断模型研究

朱逸峰,李雨师,孙兆男,张耀峰,李家轮,张晓东,王可欣,王霄英

【摘要】目的:探索利用深度学习方法在建立颈椎病 MR 诊断模型的可行性。方法:回顾性搜集 本院 2020年10月-2023年3月诊断为颈椎病患者的 MR 图像 514例,使用已有颈椎分割模型在轴面 T<sub>2</sub>WI上分别预测硬膜囊、脊髓、椎间盘、后纵韧带和黄韧带,在矢状面 T<sub>1</sub>WI 和 T<sub>2</sub>WI 上预测颈椎椎体 和椎间盘。由一位低年资放射科医生(阅片经验2年)修改标注,另一位高年资放射科医生(阅片经验≥ 15年)对低年资医师的标注进行复核。按照颈椎病的不同诊断要点分别进行 3D 或 2D 深度学习分类 模型训练,包括①颈椎椎体增生模型;②颈椎椎体滑脱模型;③颈椎间盘突出分类模型;④后纵韧带增厚 模型;⑤黄韧带增厚模型。将模型输出结果导入 R 软件进行混淆矩阵分析及 ROC 曲线绘制,采用正确 率、灵敏度、特异度、阳性预测值、阴性预测值以及 ROC 曲线下面积等评价 5 种模型的分类效能。 结果:5 种分类模型中诊断效能最好的是颈椎间盘突出分类模型,正确率 0.90,灵敏度 0.95,特异度 0.85,ROC 曲线下面积 0.982。颈椎椎体增生和滑脱的正确率分别为 0.81 和 0.80,灵敏度为 0.74 和 0.76,特异度为 0.84 和 1.00,ROC 曲线下面积分别为 0.855 和 0.905。后纵韧带和黄韧带增厚的模型正 确率分别为 0.82 和 0.77,灵敏度为 0.78 和 0.84,特异度为 0.86 和 0.70,ROC 曲线下面积分别为 0.902 和 0.929。结论:本部分研究采用深度学习方法建立了颈椎病 MR 的自动分类诊断模型,对颈椎椎体增 生、滑脱、椎间盘突出、后纵韧带及黄韧带增厚进行了分类模型训练,证明深度学习方法可以用于颈椎病 MR 的辅助诊断,为未来进一步探索建立颈椎病 MR 自动诊断模型及结构化报告的植入奠定了基础。

【关键词】 磁共振成像; 颈椎病; 深度学习

【中图分类号】R445.2;R274.9;R-056 【文献标识码】A 【文章编号】1000-0313(2024)06-0779-09 DOI:10.13609/j.cnki.1000-0313.2024.06.011 开放科学(资源服务)标识码(OSID): 驟



A preliminary study on establishing MRI diagnosis model of cervical spondylosis based on deep learning ZHU Yi-feng,LI Yu-shi,SUN Zhao-nan, et al.Department of Radiology,Peking University First Hospital,Beijing 100034,China

**[Abstract]** Objective: To explore the feasibility of using deep learning methods in establishing MRI diagnostic models for cervical spondylosis. Methods: A retrospective collection of 514 MR images of patients were diagnosed with cervical spondylosis in our hospital from October 2020 to March 2023. Use existing cervical spine segmentation models to predict the subarachnoid space, spinal cord, intervertebral disc, posterior longitudinal ligament, and ligamentum flavum on axial  $T_2$  WI, and predict the cervical vertebral bodies and intervertebral disc on sagittal  $T_1$  WI and  $T_2$  WI.A junior radiologist (with 2 years of experience) revised the labeling, and another senior radiologist (with  $\geq$ 15 years of experience) reviewed the labels. According to the different radiologic sign of cervical spondylosis, 3D or 2D U-net deep learning classification model training is carried out, including cervical vertebral hyperplasia model, cervical spondylolisthesis model, cervical disc herniation classification model. The output results of the models were imported into R programming software for confusion matrix analysis and ROC curve drawing. And the classification performance of the five models (accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and area under the ROC curve) was evaluated. **Results**:

作者单位:100034 北京,北京大学第一医院医学影像科[朱逸峰(博士研究生在读期间所在医院),孙兆男,张晓东,王霄英]; 116027 辽宁,大连医科大学附属第二医院放射科[朱逸峰(现工作单位)、李雨师];100011 北京,北京赛迈特锐医学科技有限公 司(张耀峰、李家轮);100069 北京,首都医科大学(王可欣)

作者简介:朱逸峰(1993-),男,辽宁大连人,博士研究生,住院医师,主要从事人工智能新技术在骨肌影像学中应用的研究。 通迅作者:王霄英,E-mail:cjr.wangxiaoying@vip.163.com

Among the five classification models, the cervical disc herniation classification model had the best diagnostic performance, with a correct rate of 0.90, sensitivity of 0.95, specificity of 0.85, and area under the ROC curve of 0.982. The correct rate of cervical vertebral hyperplasia and spondylolisthesis also reached 0.81 and 0.80, the sensitivity was 0.74 and 0.76, the specificity was 0.84 and 1.00, and the area under the ROC curve was 0.855 and 0.905 respectively. The correct rates of the models for thickening of the posterior longitudinal ligament and ligamentum flavum were 0.82 and 0.77, respectively, with sensitivities of 0.78 and 0.84, specificities of 0.86 and 0.70, and areas under the ROC curve of 0.902 and 0.929, respectively. **Conclusion**: In the study, deep learning method based on MRI was used to establish the automatic classification diagnosis model of cervical spondylosis, and the classification model was trained for cervical vertebral body hyperplasia, slippage, intervertebral disc herniation, thickness of posterior longitudinal ligament and ligamentum flavum. The study proved that deep learning method based on MRI can be used in the computer aided diagnosis of cervical spondylosis, which laid the groundwork and foundation of the development of structured reports of cervical spine.

[Key words] Magnetic resonance imaging; Cervical spondylosis; Deep learning

MRI 因其出色的软组织分辨力而普遍地应用于 颈椎和脊髓疾病的诊断<sup>[1,2]</sup>。临床工作中颈椎病的影 像学诊断多是基于 MRI 图像进行定性的主观判读和 定性诊断,分类诊断应用较少。由于颈椎区域的骨骼、 韧带、血管和神经等解剖结构较多,简便、客观地量化 分析颈椎和脊髓的结构有一定困难<sup>[3]</sup>。影像科日常工 作量繁重,对每一例患者进行颈椎和脊髓 MRI 图像的 手工测量并不现实,如果能够从 MRI 图像中自动获取 颈椎和脊髓的形态学指标并进行自动分类诊断将会为 临床工作提供强而有效的帮助。

随着 AI 技术的发展,应用 DL 等方法对图像进行 自动分割和测量在肺<sup>[4]</sup>、肝脏<sup>[5]</sup>、乳腺<sup>[6]</sup>、前列腺<sup>[7]</sup>、肾 脏<sup>[8]</sup>等方面都已经取得进展。本课题组既往对成人颈 椎 MR 各结构进行了自动分割的研究,建立了成人颈 椎 MR 分割模型,具备良好的准确性与稳定性<sup>[9]</sup>。本 部分研究将在前期研究的基础上开发基于深度学习的 颈椎病 MR 诊断模型,为进一步构建颈椎病 MR 自动 诊断模型和结构化报告的植入奠定基础。

#### 材料与方法

#### 1. 研究队列

本部分研究已获得本单位伦理审查委员会批准 (2019-170),对入组患者豁免知情同意,按照颈椎病的 不同诊断要点分别进行回顾性病例搜集并进行模型训 练,包括①颈椎椎体增生模型;②颈椎椎体滑脱模型; ③颈椎间盘突出分类模型;④后纵韧带增厚模型;⑤黄 韧带增厚模型。具体纳入标准:①颈椎椎体骨质增生、 伴或不伴滑脱;②颈椎间盘膨出(<5 mm)或突出(中 央型、旁中央型和椎间孔型);③后纵韧带和/或黄韧带 增厚。纳入的图像序列为颈椎病患者常规 MR 平扫 序列,不包含增强序列及其他功能成像序列。排除标 准:①行 MR 检查前有颈椎手术史、外科治疗史;②颈 椎骨折、畸形、感染等疾患;③不能配合、图像质量不佳 者。

### 2. 扫描设备参数

MRI 扫描设备包括 GE Signa Excite 3.0T、GE Discovery HD 750 3.0T、Siemens Aera 1.5T、Philips Ingenia 3.0T、Philips Achieva 3.0T、Philips Multiva 1. 5T 和联影医疗 uMR 790 3.0T。扫描序列包括常规矢状面 T<sub>1</sub>WI、矢状面 T<sub>2</sub>WI、轴面 T<sub>2</sub>WI 以及脂肪抑制 T<sub>2</sub>WI(FS-T<sub>2</sub>WI),详细扫描序列及参数如表 1。

3. 数据标注(图 1)

将入组患者的矢状面 T<sub>1</sub>WI、T<sub>2</sub>WI 及轴面 T<sub>2</sub>WI 的 DICOM 格式图像转换为 NiFTI 格式。使用本课 题组既有颈椎 MR 分割模型对其进行预测<sup>[9]</sup>,在轴面 T<sub>2</sub>WI 上分别预测硬膜囊、脊髓、椎间盘、后纵韧带和 黄韧带,在矢状面 T<sub>1</sub>WI 和 T<sub>2</sub>WI 上预测颈椎椎体和 椎间盘。由一位低年资放射科医生(阅片经验 2 年)修

表1 颈椎 MRI 常规检查序列及参数

序列名称	TR (ms)	TE (ms)	层厚 (mm)	层间距 (mm)	FOV (cm)	矩阵
$T_1WI$ , Sag	$400 \sim 850$	$6 \sim \! 16$	$3\!\sim\!4$	3.3~4.4	$8 \sim \! 15$	≥250×190
$T_2$ WI, Sag	$2000 \sim 3500$	$60\!\sim\!140$	$3 \sim 4$	3.3~4.4	$10\!\sim\!16$	$\geqslant 250 \times 180$
$T_2$ WI, Axi	$2000 \sim 5000$	$80 \sim \! 140$	$3 \sim 4$	3.3~5.5	$7\!\sim\!10$	$\geqslant$ 212 $\times$ 168
FS-T <sub>2</sub> WI, Axi	220	6.5	3	3.3	$7 \sim 8$	$\geq 252 \times 249$

注:TR, Repetition Time(重复时间); TE, Echo Time(回波时间); FOV, Field of View(视野); FS, Fat Suppression(脂肪抑制)



图 1 颈椎椎体和椎间盘的定位过程。a)在矢状面 T<sub>2</sub>WI 图像对椎体进 行分割;b)将椎体标签按不同连通域进行分割,从 C2 椎体向下依次生成 各椎体定位;c)确定椎间盘的测量层面,图中蓝线和绿线代表轴面 T<sub>2</sub>WI 扫描平面,其中绿线为通过椎间盘水平的测量平面;d)将矢状面 T<sub>2</sub>WI 的定位信息与轴面 T<sub>2</sub>WI 进行匹配,从而得到每个椎间盘的位置,图示 为 C5/6 椎间盘。

改标注,另一位高年资放射科专家(阅片经验≥15年) 对低年资医师的标注进行复核,如有不一致的情况以 高年资放射科专家的标注为金标准。修改标注后,在 矢状面椎体分割结果上自动按照 C2~C7 的顺序进行 椎体定位,将矢状面定位信息匹配至轴面 T₂WI,得到 各个椎间盘的定位(C2/3~C6/7)。

4. 椎体增生与滑脱模型训练

在定位完成后,由低年资放射科医生在颈椎正中 矢状面 T1WI 图像中对认为存在增生的椎体进行标 注,标注标准为在矢状面 T<sub>1</sub>WI 中 判断椎体边缘是否存在骨质增生 变尖等异常信号,标注点位于椎体 的中心位置,直径10像素,没有增 生的椎体不予标注(即归为正常 组)。由另一位高年资放射科专家 进行复核,标注过程如图2所示。 通过连通域分割的方法将已经定 位好的椎体取出,最终将矢状面 T<sub>1</sub>WI 图像、椎体的图像标签以及 椎体增生的标签共同输入 3D-ResNet 模型,以每一个椎体为一 个数据进行模型训练,模型训练使 用的硬件为 GPU NVIDIA Tesla P100 16G,软件为 Python3.6、Pytorch 0.4.1、Opency、Numpy 和 SimpleITK 等, Adam 为训练优化 器。模型的训练参数为图像预处 理参数.size =  $32 \times 32 \times 64$ (z.v. x),自动窗宽、窗位。共计 874 个 数据按照 8:1:1 的比例随机分为 训练集(train dataset,n=698)、验 证集(validation dataset,n=88)和 测试集(test dataset,n=88),每次 读取的图像数量(batch size)为 20, 学习率(learning rate)为 0.0001,训练次数(epoch)为500。

对于颈椎椎体滑脱的标注,同 样由低年资放射科医生在颈椎正 中矢状面 T<sub>1</sub>WI 图像中对滑脱的 椎体进行标注,由另一位高年资放 射科专家进行复核。滑脱标准以 下位椎体为基准进行4等分,上位 椎体的下缘前滑小于1/4下位椎 体的上缘即为 I°前滑脱,后滑小于 1/4下位椎体的上缘即为 I°后滑 脱,依次类推,滑脱标签标注于椎 体的中心位置,直径10 像素,对没

有滑脱的椎体不予标注(即归为正常组),标注过程如 图 3 所示。将矢状面 T1WI 图像、椎体的图像标签以 及椎体滑脱的标签共同输入 3D-ResNet 模型,以每一 个椎体为一个数据进行模型训练,模型的训练参数为 图像预处理参数:size= $32 \times 96 \times 96(z,y,x)$ ,自动窗 宽、窗位。共计 239 个数据按照 8:1:1 的比例随机分 为训练集(train dataset,n=189)、验证集(validation dataset,n=25)和测试集(test dataset,n=25),每次读



图 2 颈椎椎体增生标注。a)颈椎正中矢状面 T<sub>1</sub>WI 图像;b)分割模型自动预测的椎体标签;c)增生椎体的标注,在有骨质增生的 C3~C7 椎体中心位置进行标注,直径 10 像素,颜色为红色。 图 3 颈椎椎体滑脱标注。a)颈椎正中矢状面 T<sub>1</sub>WI 图像;b)分割模型自动预测的椎体标签;c)滑脱椎体的标注,分别在后滑脱的 C3 和 C5 椎体中心位置进行标注,直径 10 像素,颜色为红色。

取的图像数量(batch size)为4,学习率(learning rate) 为 0.0001,训练次数(epoch)为 500。

5. 椎间盘突出分类模型训练

由低年资放射科医生在轴面  $T_2$ WI上对显示椎间 盘的层面进行突出标签的标注,在突出的椎间盘图像 正中位置标注直径 10 像素的圆点,正常或膨出 < 5 mm的椎间盘不予标注(即归为正常组),后由高年资 放射科专家进行复核,分类标准①椎间盘正常或膨出: 椎间盘后缘不超过或均匀超过椎体后缘;②椎间盘中 央型突出:指髓核突出位于椎间盘后方正中央,较大时 可以压迫两侧的神经根;③椎间盘旁中央型突出:指突 出髓核位于椎间盘后方中央偏于一侧,即位于椎管中 央和椎间孔之间,一般压迫一侧的神经根;④椎间盘椎 间孔型突出:指突出的髓核位于椎间孔或椎间孔外,主 要引起神经根损害,标注过程如图 4 所示。标注完成 后将轴面 T<sub>2</sub>WI 图像、椎间盘的图像标签及正常/膨出 或突出的分类标签共同输入 2D-Unet 模型,以每一个 椎间盘层面为一个数据进行模型训练,模型的训练参 数为图像预处理参数:size= $224 \times 224(y,x)$ ,自动窗 宽、窗位。共计 379 个数据按照 8:1:1 的比例随机分 为训练集(train dataset, n=301)、验证集(validation dataset, n=39)和测试集(test dataset, n=39),每次读 取的图像数量(batch size)为 10,学习率(learning rate)为 0.0001,训练次数(epoch)为 500。

6. 后纵韧带、黄韧带增厚模型训练

由低年资放射科医生在轴面 T<sub>2</sub>WI 上对显示后纵 韧带的层面进行增厚标签的标注,在增厚的后纵韧带 层面标注直径3 像素的圆点,后由高年资放射科专家





黄韧带增厚标注。a)颈椎轴面  $T_2$  WI 图像;b)分割模型自动预测的后纵韧带标签(红色为增厚的后纵韧带); c)分割模型自动预测的黄韧带标签(绿色为增厚的黄韧带);d)在增厚的后纵韧带及黄韧带上进行标注,其中 后纵韧带增厚的标签为直径 3 像素的红色圆点,黄韧带增厚的标签为直径 3 像素的绿色圆点。

进行复核,标注标准为后纵韧带增厚超过 2 mm,对未 增厚的后纵韧带不予进行标注(即归为正常组),过程 如图 5 所示。标注完成后将轴面 T<sub>2</sub>WI 图像、后纵韧 带的图像标签及增厚分类的标签共同输入 2D-Unet 模型,以每个后纵韧带层面为一个数据进行模型训练, 模型的训练参数为图像预处理参数:size= $224 \times 224$ (y,x),自动窗宽、窗位。共计 1892 个数据按照 8:1:1 的比例随机分为训练集(train dataset,n=1510)、验证 集(validation dataset,n=191)和测试集(test dataset, n=191),每次读取的图像数量(batch size)为 10,学习 率(learning rate)为 0.0001,训练次数(epoch)为 500。

黄韧带增厚的标注,由低年资放射科医生在轴面 T<sub>2</sub>WI上对显示黄韧带增厚超过2mm的层面进行标 注,在增厚的黄韧带层面标注直径 3 像素的圆点,后由 高年资放射科专家进行复核,对未增厚的黄韧带不予 进行标住(即归为正常组),过程如图 5 所示。标注完 成后将轴面 T<sub>2</sub>WI 图像、黄韧带的图像标签及增厚分 类的标签共同输入 2D-Unet 模型,以每个黄韧带层面 为一个数据进行模型训练,模型的训练参数为图像预 处理参数:size= $224 \times 224(y,x)$ ,自动窗宽、窗位。共 计 343 个数据按照 8:1:1 的比例随机分为训练集 (train dataset, n=271)、验证集(validation dataset, n=36)和测试集(test dataset, n=36),每次读取的图 像数量(batch size)为 10,学习率(learning rate)为 0.0001,训练次数(epoch)为 500。

6. 模型评价

将5个模型输出的原始数据输入R4.2.0软件,绘制混淆矩阵图及受试者操作特性曲线(receiver operating characteristic curve,ROC)。根据混淆矩阵的结 果评价5个分类模型的正确率(accuracy,ACC)、灵敏 度(真阳性率)(true positive rate,TPR)、特异度(真阴 性率)(true negative rate,TNR)、阳性预测值(positive predictive value,PPV)、阴性预测值(negative predictive value,NPV)、阳性似然比(positive likelihood ratio,PLR)、阴性似然比(negative likelihood ratio, NLR)、ROC 曲线下面积(area under curve,AUC)以 及F1-Score。

#### 结 果

#### 1. 一般资料

5 种分类模型纳入的患者一般资料详见表 2,共搜 集 514 例患者的颈椎 MR 图像,年龄范围 28~96 岁, 平均年龄 58.6±12.9 岁,其中男 229 例,女 285 例,男、 女比例为 1:1.2。

表 2 五种分类模型纳入的病例数据汇总

模型名称	MR 检查时间	男	女	合计
颈椎椎体增生	2020/10 - 2023/2	50	60	110
颈椎椎体滑脱	2020/10 - 2022/1	24	29	53
颈椎间盘突出	2020/10 - 2021/9	22	38	60
后纵韧带增厚	2020/10 - 2023/3	89	103	192
黄韧带增厚	2020/10 - 2023/3	44	55	99
合计	—	229	285	514

2. 五种分类模型的评价

5种分类模型的训练集、验证集及测试集数据分 组情况如表 3 所示。5种分类模型的测试集混淆矩阵 图及 ROC 曲线如图 6~10 所示,各模型的 ACC、 TPR、TNR、PPV、NPV、PLR、NLR 以及 F1-Score 如 表 4 所示。5种分类模型测试集中效能最好的是颈椎 间盘突出分类模型,正确率 0.90,灵敏度 0.95,特异度 0.85,ROC 曲线下面积 0.982(95% CI:0.958~ 1.000)。颈椎椎体增生和滑脱的正确率分别达到 0.81 和 0.80,灵敏度为 0.74 和 0.76,特异度为 0.84 和 1.00, 测试集 ROC 曲线下面积分别为 0.855(95% CI: 0.772~0.939)和 0.905(95% CI:0.772~1.000)。后 纵韧带和黄韧带增厚的模型正确率分别为 0.82 和 0.77,灵敏度为 0.78 和 0.84,特异度为 0.86 和 0.70,测 试集 ROC 曲线下面积分别为 0.902(95% CI:0.871~ 0.932)和 0.929 (95% CI:0.831~1.000)。

表 3 五种分类模型的训练集、验证集及测试集数据分组

模型名称/分类	训练集	验证集	测试集	合计
颈椎椎体增生				
正常	244	31	31	306
增生	454	57	57	568
颈椎椎体滑脱				
正常	156	20	21	197
滑脱	33	5	4	42
颈椎间盘突出				
正常/膨出	145	19	19	183
突出	156	20	20	196
后纵韧带增厚				
正常	1378	172	172	1722
增厚	1435	224	196	1855
黄韧带增厚				
正常	964	109	117	1190
增厚	641	112	112	865
合计	5606	769	749	7124

讨 论

颈椎病的发病率近年来逐渐升高,且有明显低龄 化的趋势。随着 AI 与 DL 的发展与进步,AI 与 DL 技术在颈椎 MR 诊断的应用逐渐增多,除在颈椎的分 割和测量方面以外,亦有部分研究关注了 AI 与 DL 在 颈椎疾病诊断中的应用如椎管受压的分类<sup>[10]</sup>、脊髓损 伤和间盘退变的检测<sup>[11]</sup>等。本研究在既有颈椎 MR 分割模型的基础上<sup>[9]</sup>尝试对颈椎病的几种主要影像诊 断要点(椎体增生、滑脱、椎间盘突出、韧带增厚)进行 了全新分类模型的训练,5 种模型的诊断效能目前均 能够满足临床需求,证明 DL 方法可以用于颈椎病 MR 的辅助诊断,可以减轻放射科医生的工作负担,提 高日常工作的效率。

既往也有研究对 AI 在脊柱退行性疾病诊断中的 应用进行研究, Michopoulou 等<sup>[12]</sup>使用贝叶斯二元分 类器算法开发了基于 T<sub>2</sub>WI 图像上椎间盘退变的自动 分类系统,该系统根据椎间盘的信号将椎间盘按照正 常或退变两种分类进行输出。Raja'S 等<sup>[13]</sup>也使用 30 个 MR 的数据集进行了类似的研究。Ghosh 等<sup>[14]</sup>使 用 SVM 进行了类似的研究,他们纳入了 35 个数据, 该模型的最终准确度在 80%~94%之间,优于贝叶斯

表 4 五种分类模型的诊断效能

模型名称	ACC	TPR	TNR	PPV	NPV	PLR	NLR	F1-Score
颈椎椎体增生	0.81	0.74	0.84	0.72	0.86	4.70	0.31	0.73
颈椎椎体滑脱	0.80	0.76	1.00	1.00	0.44	4.20	0.24	0.86
颈椎间盘突出	0.90	0.95	0.85	0.86	0.94	6.32	0.06	0.90
后纵韧带增厚	0.82	0.78	0.86	0.83	0.82	5.91	0.24	0.81
黄韧带增厚	0.77	0.84	0.70	0.75	0.81	2.87	0,21	0.79

注:ACC,正确率;TPR,灵敏度;TNR,特异度;PPV,阳性预测值;NPV,阴性预测值;PLR,阳性似然比;NLR,阴性似然比



图 6 颈椎椎体增生分类模型混淆矩阵和 ROC 曲线图。a)混淆矩阵,Nor-中对颈椎病 MR 的影像诊断流 mal 为正常椎体,Intersection 为增生椎体;b)ROC 曲线。 图 7 颈椎椎体 程,具备真正应用于放射科临床工 滑脱模型混淆矩阵和 ROC 曲线图。a)混淆矩阵,Normal 为无滑脱椎体,作的潜力。这种整合了多模型 DL Slippage 为滑脱椎体;b)ROC 曲线。 图 8 颈椎间盘突出分类模型混淆矩 的方法目前在国内外研究中未见 阵和 ROC 曲线图。a)混淆矩阵,Normal 为正常或膨出的颈椎间盘,Abnor-明确报道,具备一定的临床应用前 mal 包括中央型突出、旁中央型突出及椎间孔型突出 3 种类型的颈椎间盘;景。 b)ROC 曲线。

二元分类器算法。Hao 等<sup>[15]</sup>提出了一种基于 SVM 的方法,除了信号强度和纹理信息外,还纳入了椎间盘 的形状信息,以便将其分类为退变或未退变,该方法准 确率高达 91.6%。Oktay 等<sup>[16]</sup>通过纳入椎间盘 T<sub>1</sub>WI 图像的信息进一步改进了这种方法。Ruiz-Espana 等<sup>[17]</sup>和 Castro-Mateos 等<sup>[18]</sup>的研究则进一步取得了

突破,他们将分类标准由二分类调 整为临床实践中普遍使用的 Pfirrmann 5 分类标准<sup>[19]</sup>,这两项研 究都包括自动分割椎间盘并提取 椎间盘信号强度和形状的特征,将 其输入分类器后交给 ANN 进行 分类输出。Jamaludin 等<sup>[20]</sup>的研 究介绍了一些改进和创新,例如搜 集大量椎间盘图像用于训练和测 试,该研究纳入了 2009 名患者的 12018个椎间盘进行模型训练,该 方法与医生的一致性达到 70.1%, 这个结果与不同放射科专家之间 的准确度相当(70.4%)。此外这 种方法在椎体终板病变和骨髓病 变中也得到了应用。Niemever 等<sup>[21]</sup>使用了深度 CNN 网络并进 一步增加了训练集的大小,最终使 自动 Pfirrmann 分类系统的准确 率达到 97%。上述的既往研究均 为单一的分类模型,尽管部分研究 有较大规模的数据量做支撑,但应 用大多局限在单一的椎间盘退变 分类方面。从广义上来看,既往的 相关研究结果临床意义较局限,很 难实际应用到临床工作中。

本部分研究的创新之处在于 除了采用深度学习方法建立颈椎 间盘突出的分类模型之外,还尝试 引入了颈椎椎体增生、椎体滑脱、 后纵韧带及黄韧带增厚的分类模 型,更模拟和接近放射科实际工作 距降,Nor-中对颈椎病 MR 的影像诊断流 颈椎椎体 程,具备真正应用于放射科临床工 骨脱椎体,作的潜力。这种整合了多模型 DL 建型混淆矩 的方法目前在国内外研究中未见 急,Abnor-明确报道,具备一定的临床应用前 频椎间盘;景。

> 本研究也存在明显的局限性。 首先,本研究目前仍为初步研究, 仅初步探索了使用 DL 方法建立

颈椎病 MR 诊断模型的可行性,由于入组数据较少, 尚不能使模型达到极高标准的准确率和稳定性。在未 来的研究中应进一步增加数据量来迭代模型,同时还 需要进行大规模、前瞻性的外部验证以更进一步提高 模型的诊断准确性、稳定性和泛化能力。此外,还需要



图 9 后纵韧带增厚分类模型混淆矩阵和 ROC 曲线图。a)混淆矩阵, Normal 为正常后纵韧带, Thickness 为增厚的后纵韧带; b) ROC 曲线。 图 10 黄韧带增厚分类模型混淆矩阵和 ROC 曲线图。a) 混淆矩阵, Nor-

mal 为正常黄韧带, Thickness 为增厚的黄韧带; b) ROC 曲线。

进一步进行放射科医师的读片实验来对比 DL 模型的 效能与放射科医师的诊断效能。第二,目前本研究训 练的 DL 模型尚属于定性诊断模型,在颈椎病乃至其 他脊柱疾病中定量诊断亦非常重要,当脊髓出现髓内 肿瘤、血管畸形、炎症等疾病或椎管和脊髓神经系统发 育异常时,其形态学上通常会出现相应的膨大或萎缩 改变,此时脊髓的形态和其在硬膜囊中的相对位置均 会出现变化,此时定量测量及诊断能够更加客观、准确 地描述疾病的发生发展规律,也有助于疾病的鉴别诊 断,既能够对脊髓中枢神经系统的发育进行定量评价, 亦可对颈椎和脊髓疾病及其内科保守治疗的疗效进行 定量诊断和动态评估。目前本部分研究纳入的数据量 相对较少,在颈椎病异常情况中模型自动分割生成标 签的准确率尚不能完全满足自动或半自动测量的需 求,因此将来应进一步增加数据量进行模型的迭代,在 提高模型分割能力的基础上继续探索研究基于深度学 习方法的颈椎病 MR 影像自动测量模型。第三,颈椎 病及相关疾病的诊断不仅包括形态学评估,还应评估

信号变化、演进过程,将来应在自 动分割和定量测量的基础上进一 步自动获得解剖结构的信号强度, 并通过信号变化来诊断和鉴别诊 断颈椎及脊髓疾病,这亦是 MR 高软组织分辨力成像的关键优势 所在。

综上所述,本部分研究采用深 度学习方法建立了颈椎病 MR 的 自动分类诊断模型,对颈椎椎体增 生、滑脱、椎间盘突出、后纵韧带及 黄韧带增厚进行了分类模型的训 练,种模型的诊断效能目前均能够 满足临床需求,证明深度学习方法 可以用于颈椎病 MR 的辅助诊 断,为未来进一步探索建立颈椎病 MR 自动诊断模型及结构化报告 的植入奠定了基础。

#### 参考文献:

- Boden SD, McCowin PR, Davis DO, et al. Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subjects. A prospective investigation
  [J]. J Bone Joint Surg Am, 1990, 72 (8):1178-1184.
- [2] 葛雅丽,郑敏文,张劲松,等.MRI 全脊 柱移床扫描技术在脊髓及椎体疾病诊 断中的应用[J].实用放射学杂志, 2002,18(11):971-973.
- [3] Kopjar B, Bohm PE, Arnold JH, et al. Outcomes of surgical decompression in patients with very severe degenerative cervical myelopathy[J].Spine (Phila Pa 1976).2018.43(16):1102-1109.
- [4] 沈慧聪,王淼淼,李智勇,等.人工神经网络(ANN)在孤立性肺结 节鉴别诊断中的应用初探[J].实用放射学杂志,2004,20(10):3.
- [5] 张波,姜建辉,贾娟,等.人工神经网络肝癌 CT 影像辅助诊断模型的建立[J].实用放射学杂志,2006,22(9):4.
- [6] Ma M,Gan L,Liu Y, et al.Radiomics features based on automatic segmented MRI images:prognostic biomarkers for triple-negative breast cancer treated with neoadjuvant chemotherapy[J].Eur J Radiol,2022,146:110095.
- [7] Ma S, Xie H, Wang H, et al. MRI-based radiomics signature for the preoperative prediction of extracapsular extension of prostate cancer[J].J Magn Reson Imaging, 2019, 50(6): 1914-1925.
- [8] 崔应谱,孙兆男,刘想,等.用深度学习和阈值算法自动检出 CT 平 扫图像中肾结石的可行性研究[J].中华放射学杂志,2020,54(9): 869-873.
- [9] 朱逸峰,赵凯,郭丽,等.基于深度学习模型实现颈椎 MR 图像上 各结构的自动分割[J].放射学实践,2021,36(12):1558-1562.
- [10] Merali Z, Wang JZ, Badhiwala JH, et al. A deep learning model

for detection of cervical spinal cord compression in MRI scans [J].Sci Rep,2021,11(1):10473.

- [11] Ma S, Huang Y, Che X, et al. Faster RCNN-based detection of cervical spinal cord injury and disc degeneration[J].J Appl Clin Med Phys, 2020, 21(9):235-243.
- [12] Michopoulou SK, Costaridou L, Panagiotopoulos E, et al. Atlasbased segmentation of degenerated lumbar intervertebral discs from MR images of the spine[J].IEEE Trans Biomed Eng, 2009, 56(9):2225-2231.
- Raja'S A, Corso JJ, Chaudhary V, et al. Desiccation diagnosis in lumbar discs from clinical MRI with a probabilistic model[C]// 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro.IEEE, 2009:546-549.
- [14] Ghosh S.Raja'S A.Chaudhary V.et al.Computer-aided diagnosis for lumbar MRI using heterogeneous classifiers[C]//2011 IEEE international symposium on biomedical imaging; from nano to macro.IEEE,2011:1179-1182.
- [15] Hao S, Jiang J, Guo Y, et al. Active learning based intervertebral disk classification combining shape and texture similarities[J]. Neurocomputing, 2013, 101:252-257.
- [16] Oktay AB, Albayrak NB, Akgul YS. Computer aided diagnosis of

degenerative intervertebral disc diseases from lumbar MR images [J].Comput Med Imaging Graph,2014,38(7),613-619.

- [17] Ruiz-Espana S, Arana E, Moratal D. Semiautomatic computer-aided classification of degenerative lumbar spine disease in magnetic resonance imaging[J].Comput Biol Med, 2015, 62:196-205.
- [18] Castro-Mateos I.Pozo JM, Lazary A, et al.2D segmentation of intervertebral discs and its degree of degeneration from T<sub>2</sub>-weighted magnetic resonance images [C]//Medical Imaging 2014: Computer-Aided Diagnosis.SPIE, 2014, 9035:310-320.
- [19] Pfirrmann CW, Metzdorf A, Zanetti M, et al. Magnetic resonance classification of lumbar intervertebral disc degeneration[J]. Spine (Phila Pa 1976), 2001, 26(17):1873-1878.
- [20] Jamaludin A, Lootus M, Kadir T, et al. ISSLS Prize in Bioengineering Science 2017: Automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist[J].Eur Spine J,2017,26:1374-1383.
- [21] Niemeyer F, Galbusera F, Tao Y, et al. A deep learning model for the accurate and reliable classification of disc degeneration based on MRI data[J].Invest Radiol, 2021, 56(2):78-85.

(收稿日期:2023-07-26 修回日期:2023-11-24)

## 《请您诊断》栏目征文启事

《请您诊断》是本刊 2007 年新开辟的栏目,该栏目以临床上少见或容易误诊的病例为素材,杂志在 刊载答案的同时配发专家点评,以帮助影像医生更好地理解相关影像知识,提高诊断水平。栏目开办 13 年来受到广大读者欢迎。《请您诊断》栏目荣获第八届湖北精品医学期刊"特色栏目奖"。

本栏目欢迎广大读者踊跃投稿,并积极参与《请您诊断》有奖活动,稿件一经采用稿酬从优。

《请您诊断》来稿格式要求:①来稿分两部分刊出,第一部分为病例资料和图片;第二部分为全文,即 病例完整资料(包括病例资料、影像学表现、图片及详细图片说明、讨论等);②来稿应提供详细的病例资 料,包括病史、体检资料、影像学检查及实验室检查资料;③来稿应提供具有典型性、代表性的图片,包括 横向图片(X线、CT或MRI等不同检查方法得到的影像资料,或某一检查方法的详细图片,如CT平扫 和增强扫描图片)和纵向图片(同一患者在治疗前后的动态影像资料,最好附上病理图片),每帧图片均 需详细的图片说明,包括扫描参数、序列、征象等,病变部位请用箭头标明。

具体格式要求请参见本刊(一个完整病例的第一部分请参见本刊正文首页,第二部分请参见2个月 后的杂志最后一页,如第一部分问题在1期杂志正文首页,第二部分答案则在3期杂志正文末页)

栏目主持:石鹤 联系电话:027-69378385 15926283035