

Artificial intelligence knowledge graph and image classification for quality control of chest posterior-anterior position X-ray radiograph

WANG Qian¹, SONG Liangliang¹, HAN Xiao², LIU Ming^{3,4}, ZHANG Biao^{1,4}, ZHAO Shibo²,
GU Zongyun^{2,4}, HUANG Lili^{3,4}, LI Chuanfu^{2,4,5}, LI Xiaohu^{1,5}, YU Yongqiang^{1,5*}

(1. Department of Radiology, The First Affiliated Hospital of Anhui Medical University, Hefei 230022, China;
2. College of Medical Information Engineering, Anhui University of Chinese Medicine, Hefei 230012, China;
3. College of Computer Science and Technology, Anhui University, Hefei 230093, China; 4. Artificial
Intelligence Research Institute of Hefei Comprehensive National Science Center, Hefei 230088,
China; 5. Anhui Provincial Imaging Diagnosis Quality Control Center, Hefei 234099, China)

[Abstract] **Objective** To observe the value of artificial intelligence (AI) knowledge graph and image classification for quality control (QC) of chest posterior-anterior position X-ray radiograph (abbreviated as chest film). **Methods** Totally 9 236 chest films from 595 medical institutions in Anhui province imaging cloud platform were retrospectively enrolled. QC knowledge graph containing 21 classification labels were constructed. Firstly, QC of chest films based on the above knowledge graph were performed by 10 technicians for 2 rounds of single person and 1 round of multi person, and the results were recorded as A, B and C, respectively. Then AI algorithms were used to classify and evaluate based on knowledge graph, and the result was recorded as D. Finally, a QC expert reviewed results C and D to determine the final QC results and taken those as references to analyze the efficiency of the above 4 QC. **Results** The area under the curve (AUC) of AI algorithm for QC of chest films were all ≥ 0.780 , with an average value of 0.939. The average precision of QC for chest films of A, B, C and D was 81.15%, 85.47%, 91.65% and 92.21%, respectively. **Conclusion** AI knowledge graph and image classification technology could be effectively used for QC of chest posterior-anterior position X-ray radiograph.

[Keywords] thorax; artificial intelligence; quality control; X-rays

DOI:10.13929/j.issn.1003-3289.2024.06.026

人工智能知识图谱和图像分类用于胸部 后前位 X 线片质量控制

王倩¹, 宋亮亮¹, 韩啸², 刘明^{3,4}, 张标^{1,4}, 赵士博², 谷宗运^{2,4},
黄莉莉^{3,4}, 李传富^{2,4,5}, 李小虎^{1,5}, 余永强^{1,5*}

(1. 安徽医科大学第一附属医院放射科, 安徽合肥 230022; 2. 安徽中医药大学医药信息工程学院, 安徽合肥 230012;
3. 安徽大学计算机科学与技术学院, 安徽合肥 230093; 4. 合肥综合性国家科学中心人工智能实验室,
安徽合肥 230088; 5. 安徽省影像诊断质控中心, 安徽合肥 234099)

[摘要] **目的** 观察人工智能(AI)知识图谱和图像分类对胸部后前位 X 线片(简称胸片)质量控制(QC)的价值。 **方法**

[基金项目] 安徽高校协同创新项目(GXXT-2022-031)。

[第一作者] 王倩(1998—), 女, 安徽安庆人, 在读硕士。研究方向: 医学影像学人工智能。E-mail: 2867471752@qq.com

[通信作者] 余永强, 安徽医科大学第一附属医院放射科, 230022; 安徽省影像诊断质控中心, 234099。E-mail: cjr.yuyongqiang@vip.163.com

[收稿日期] 2023-12-22 [修回日期] 2024-01-26

回顾性分析安徽省影像云平台中 595 家医疗机构共 9 236 幅胸片,构建包含 21 个分类标签的 QC 知识图谱。先由 10 名技师据此对胸片进行 2 轮单人 QC 和 1 轮多人 QC,分别将结果记为 A、B、C;再以 AI 算法进行分类评估,将结果记为 D。最后由 1 名 QC 专家对 C、D 进行审核并确定最终 QC 结果,以之为参考评估上述 4 种 QC 效果。结果 AI 算法用于胸片 QC 的曲线下面积(AUC)均 ≥ 0.780 ,平均 AUC 为 0.939。A、B、C、D 胸片 QC 的平均精确率分别为 81.15%、85.47%、91.65%、92.21%。结论 AI 知识图谱和图像分类技术可有效用于胸部后前位 X 线片 QC。

[关键词] 胸部;人工智能;质量控制;X 线

[中图分类号] R56;R814.3 [文献标识码] A [文章编号] 1003-3289(2024)06-0922-06

胸部后前位 X 线片(以下简称胸片)是常用影像学检查之一,其图像质量直接影响诊断准确性。对胸片图像质量进行评估和控制极为重要,为此需建立完整且有效的图像质量控制(quality control, QC)体系。目前我国已建立相对公认的医学图像质量评价体系,但其实际执行过程仅限于人工层面,具有一定主观性。本研究观察人工智能(artificial intelligence, AI)知识图谱和图像分类技术用于胸片 QC 的价值。

1 资料与方法

1.1 研究对象 回顾性收集 2018 年 1 月—2022 年 12 月安徽省影像云平台(<http://www.imagingunion.com>)中来自 595 家医疗机构 9 236 例患者的 9 236 幅胸片,其中男 5 359 例、女 3 877 例,年龄 4~99 岁、平均(61.8 \pm 19.6)岁。纳入标准:①常规胸片;②来自 >3 岁的门诊和住院患者。按照《信息安全技术—个人信息安全规范》对数据进行脱敏处理。本研究经医院伦理委员会免除伦理审查。

1.2 QC 标准 根据《放射科管理规范与质控标准(2017 版)》^[1],经 10 名放射科医师共同讨论并确定的 8 项胸片 QC 标准见表 1。其中,对序号 1 项目通过研究人员前期研发的 X 线检查项目分类算法进行判断^[2],对序号 5 通过常规编程读取图像 DICOM 属性

值进行判断,对序号 8 通过深度学习将图像清晰度质量分级评估转换为回归问题进行判断;故本研究仅就其余 5 项(序号 2,3,4,6,7)进行分析。

1.3 构建知识图谱 基于 QC 标准构建包括“类别(Category)-特征(Feature)-细节(Detail)”3 层结构的胸片 QC 知识图谱(表 2),并通过纵向连接 3 层结构的不同节点构成包含 31 个节点和 30 条关系的可视化 QC 知识图谱(图 1)。图谱中所有节点均包含名称、id 标识号及层级共 3 个基本属性;其中,叶子节点(即末级节点)构成本研究关注的 21 个 QC 问题标签。采用 Neo4j 的 Cypher 语句将知识图谱存储于图数据库,可随时扩展或更新^[3],并可对 QC 节点增加 QC 评分等属性,以调用知识图谱接口程序而生成 QC 评分报告。

1.4 QC 分析

1.4.1 人工标注 将胸片以 DICOM 格式上传至本团队开发的数据标注平台。由 10 名具有 10 年以上工作经验并曾受培训的主管技师根据对应节点标注判断标准(表 2)及 QC 问题示例(图 2)勾选知识图谱叶子节点,并进行 3 轮标注;前 2 轮标注均由系统将图像随机分配,分别将标注结果记为 A、B;之后在标注界面上将 A 与 B 不一致的节点标为红色,并将这些图像再次随机分配进行标注,将结果记为 C。

1.4.2 AI 算法 AI 算法由 TResNet^[4] 主干和 ML-Decoder^[5] 分类头构成;其中的主干可提取图像分类细节特征,分类头可对提取特征进行分类解码。采用多标签分类,即 1 幅图像可与多种 QC 问题分类标签相关联。按 7:1:2 将 9 236 幅胸片分为训练集($n=6 462$)、验证集($n=927$)及测试集($n=1 847$),见表 3。首先在 ImageNet 数据集上进行预训练,通过迁移学习

表 1 胸部后前位 X 线片 QC 标准

序号	项目	内容	判断方法
1	拍摄目的	判断是否为胸部后前位片	AI
2	拍摄范围	判断拍摄范围是否满足 QC 要求,即上界为肩部软组织上方 3~5 cm、下界为肋膈角,左右肋骨及软组织完整	AI
3	投照中心线	判断投照中心线是否居中,即经 T6 垂直入射,可见 T6 位于图像正中	AI
4	体位	判断拍摄体位是否标准,即双侧肩胛骨内缘投影于肺野外,两侧锁骨平坦、锁骨外缘水平等高	AI
5	DICOM 参数	判断图像信息是否完整,即是否完整包含患者姓名、性别、年龄、影像号、检查部位、检查日期、检查时间、管电压及管电流等	常规编程
6	左右标识	判断图像是否有左(L)或右(R)标识	AI
7	图像伪影	判断图像是否存在体外异物伪影或设备伪影	AI
8	图像清晰度	以 5 个级别判断图像清晰度,最高级别为隐约可见左心影内肺纹理、膈肌以上肋骨整体显示清晰、正常结构与异常病变对比良好	AI

表 2 胸部后前位 X 线片 QC 知识图谱实体设计

类别(第 1 层)	特征(第 2 层)	细节(第 3 层)	叶子节点判断内容
左右标识	缺少左右标识	—	无左(L)或右(R)标识
范围	范围过大	—	显示下颌之上或髂嵴之下
	范围过小	肺尖 肋膈角 肩部上方	双侧肺尖区域不完整 双侧肋膈角区域不完整 至少一侧肩部软组织上方<3 cm
投照中心线	图像偏左	—	脊柱与图像左缘距离小于与右缘距离
	图像偏右	—	脊柱与图像右缘距离小于与左缘距离
	图像偏上	—	中心线靠下, 整个肺野偏上
	图像偏下	—	中心线靠上, 整个肺野偏下
体位	肩胛骨在肺野之内	左侧 右侧	左侧肩胛骨内缘在左肺野内 右侧肩胛骨内缘在右肺野内
	双侧肩部高低不齐	—	两侧锁骨外缘与图像上缘距离不相等
	图像伪影	体外异物伪影	内衣扣、内衣饰品等异物影 纽扣、外衣饰品、衣服褶皱等异物影 密度高的金属项链等异物影 拉链和拉链头等异物影 皮带等异物影 其他 耳环、回形针或螺丝帽等异物影
	设备伪影	裂纹 射线 其他	裂纹样设备伪影 长条散射线样设备伪影 图像四周白线条样设备伪影等

微调网络参数, 设置算法训练初始学习率 0.000 01、Batch Size 16、Adam 优化器, 直至训练 50 个 epoch 损失值不再减小后自动停止训练。之后, 输入测试集胸片, 由 AI 算法输出 0 到 1 之间的数字, 以代表各个 QC 分类问题的概率; 将 AI 算法 QC 分析结果记为 D。

1.5 参考标准 系统用红色标注 C 与 D 不同的节点, 由 1 名 QC 专家采用双盲法对此进行审核并确定最终 QC 结果, 并以之为参考标准。

1.6 评价指标 基于训练集及验证集绘制受试者工作特征(receiver operating characteristic, ROC)曲线, 计算曲线下面积(area under the curve, AUC), 评估 AI 算法的分类性能。基于测试集计

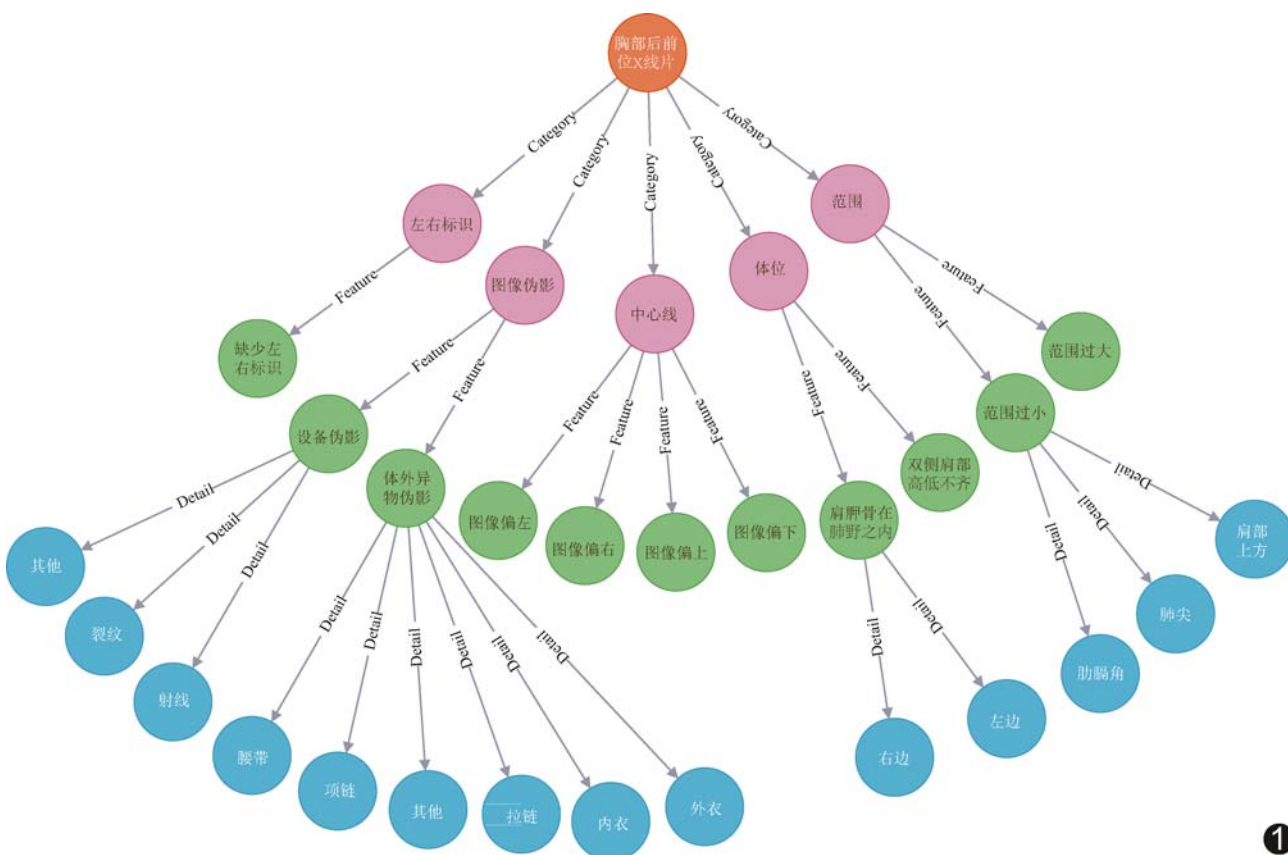


图 1 胸片可视化 QC 知识图谱

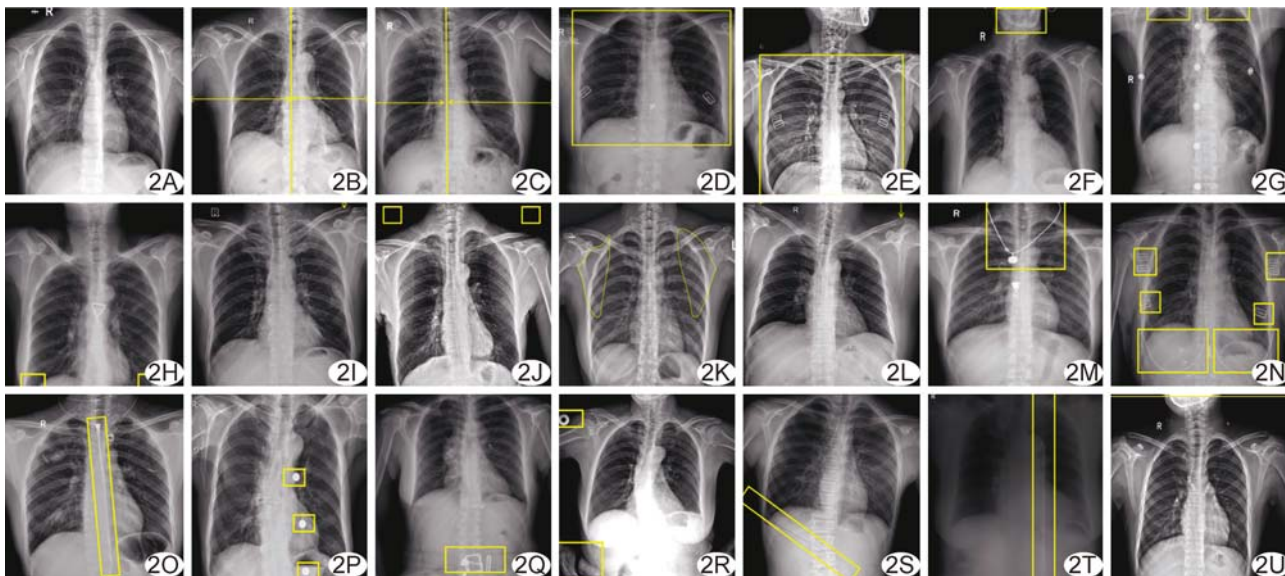


图 2 胸部后前位 X 线片标准参考图及 QC 问题示例图 A. 参考图; B. 图像偏左; C. 图像偏右; D. 图像偏上; E. 图像偏下; F. 范围过大; G. 范围过小-肺尖; H. 范围过小-肋膈角; I. 范围过小-肩部上方; J. 缺少左右标识; K. 肩胛骨在肺野之内; L. 双侧肩部高低不齐; M. 体外异物伪影-项链; N. 体外异物伪影-内衣; O. 体外异物伪影-拉链; P. 体外异物伪影-外衣; Q. 体外异物伪影-腰带; R. 体外异物伪影-其他; S. 设备伪影-裂纹; T. 设备伪影-射线; U. 设备伪影-其他

表 3 胸部后前位 X 线片数据集划分详细信息

特征(-细节)	训练集 (n=6 462)	测试集 (n=1847)	验证集 (n=927)
缺少左右标识	1 418	354	204
范围过大	1 742	512	235
范围过小-肺尖	171	52	25
范围过小-肋膈角	117	34	17
范围过小-肩部上方	2 673	794	366
图像偏左	1 884	509	263
图像偏右	2 051	585	289
图像偏上	297	86	36
图像偏下	2 336	668	316
肩胛骨在肺野之内-左边	5 653	1 603	822
肩胛骨在肺野之内-右边	5 435	1 545	783
双侧肩部高低不齐	2 745	798	391
体外异物伪影-内衣	571	155	83
体外异物伪影-外衣	1 062	292	145
体外异物伪影-项链	410	123	61
体外异物伪影-拉链	330	84	52
体外异物伪影-腰带	142	35	22
体外异物伪影-其他	1 207	355	164
设备伪影-裂纹	227	64	33
设备伪影-射线	88	25	13
设备伪影-其他	939	259	123
阴性	75	22	11

算精确率(Precision), 评估 A、B、C、D 的 QC 效果, 公式如下:

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \times 100\%$$

其中, TP 为真阳性(true positive), FP 为假阳性(false positive)。

2 结果

2.1 人工标注 前 2 轮人工 QC 标注结果中, 仅 635 幅(635/9 236, 6.88%)胸片 QC 分类节点完全一致, 8 601 幅(8 601/9 236, 93.12%)至少 1 个分类节点标注不一致。

2.2 AI 算法 AI 算法用于胸片 QC 的 AUC 均 ≥ 0.780 , 平均 AUC 为 0.939。见图 3。

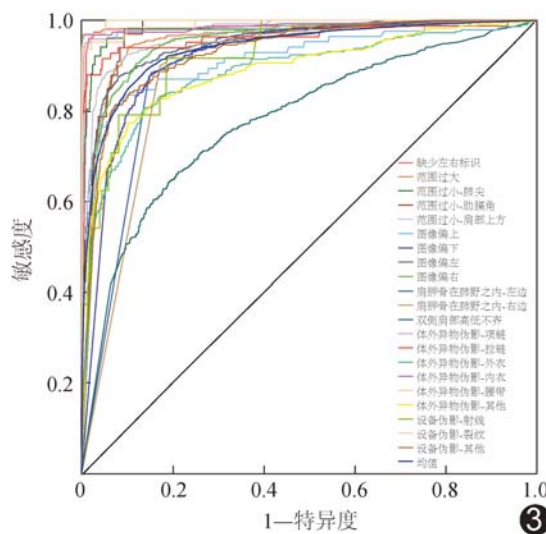


图 3 AI 算法用于胸片 QC 的 ROC 曲线

2.3 结果比较 胸片 A、B、C、D QC 的平均精确率分别为 81.15%、85.47%、91.65%、92.21%，见表 4。

表 4 AI 算法与人工标注用于胸片 QC 的精确率

QC 问题分类标签(-细节)	A	B	C	D
缺少左右标识	92.00%	91.04%	97.18%	95.99%
范围过大	73.70%	80.17%	88.48%	90.59%
范围过小-肺尖	67.69%	84.31%	90.38%	84.62%
范围过小-肋膈角	40.91%	56.25%	58.82%	60.00%
范围过小-肩部上方	92.16%	91.95%	93.45%	98.61%
图像偏左	92.28%	85.89%	92.73%	97.27%
图像偏右	96.23%	91.07%	93.85%	96.25%
图像偏上	40.23%	65.06%	88.37%	89.29%
图像偏下	95.05%	90.98%	94.46%	96.68%
肩胛骨在肺野之内-左边	98.51%	97.91%	99.44%	99.54%
肩胛骨在肺野之内-右边	98.88%	98.04%	99.74%	99.33%
双侧肩部高低不齐	89.60%	87.09%	96.37%	92.38%
体外异物伪影-内衣	98.55%	95.36%	100%	100%
体外异物伪影-外衣	84.39%	87.03%	95.55%	90.29%
体外异物伪影-项链	96.43%	97.56%	99.19%	98.39%
体外异物伪影-拉链	87.80%	94.74%	96.43%	90.41%
体外异物伪影-腰带	78.05%	94.12%	100%	97.14%
体外异物伪影-其他	80.90%	81.41%	98.46%	97.44%
设备伪影-裂纹	93.02%	92.59%	95.31%	100%
设备伪影-射线	14.49%	39.13%	48.00%	63.64%
设备伪影-其他	93.21%	93.21%	98.46%	98.62%
均值	81.15%	85.47%	91.65%	92.21%

3 讨论

近年来，卷积神经网络(convolutional neural network, CNN)模型在医学图像处理领域的应用发展迅猛^[6]，主要用于实现目标检测、分割和分类等任务，如 AlexNet^[7]、VGGNet^[8]、GoogLeNet^[9] 及 ResNet^[10]。TRResNet 是在 ResNet50 基础上加以改进和优化的高效图形处理单元架构，通过改进 stem 设计和下采样方式而显著提高了分类准确率，已成为图像分类领域的经典算法^[4]。知识图谱是以符号形式描述不同概念及其相互关系的结构化语义知识库，能将异构数据转换为计算机可识别的语言，以更好地指导影像学 AI 研究^[11]。执行医学影像 QC 有助于准确诊断疾病，但目前尚处于人工阶段，耗时、费力且存在观察者间差异。

数据标注是 AI 算法训练的基础，高质量标注可确保算法能更好地捕捉细节特征而提高其准确率。本研究采用 3 轮标注方式，其中第 3 轮用于解决前 2 轮标注存在的分歧，技师在该轮次中仔细查看前 2 轮标注的不同点并补充未发现细节，可有效减少因视觉疲劳、认知水平不足等所致误标或漏标。本研究前 2 轮标注中仅 6.88% 结果完全一致，足以表明进行第 3 轮标注的必要性。

贾晓茜等^[12]以 CNN 将胸片质量分为优、良、中、差 4 个等级，其分类总准确率为 75.93%，但为单中心研究，且未分析导致图像质量较差的原因。孟宇等^[13]开发的基于 CNN 的胸片异物检测算法的分类性能亦较佳，但仅限于异物伪影，QC 内容相对局限。本研究收集安徽省 595 家医疗机构数据，并将图像 QC 问题分为 21 个问题标签，可更明确地追踪投照技术质量问题所在；所用 AI 算法用于胸片 QC 的 AUC 均 ≥ 0.780，其中 20 个问题标签的 AUC > 0.8，表明利用 TRResNet 算法可精准评估多数 QC 问题；且 AI 算法 QC 的精确率几乎全部高于单人工 QC（即结果 A、B）、部分高于多人 QC（即结果 C），整体而言其 QC 效果可媲美多人 QC。

综上所述，AI 知识图谱和图像分类技术可有效用于胸片 QC。本研究 AI 算法对部分节点的 QC 精确率较低，可能与该类别样本量、标注质量及算法敏感度有关，有待后续加以完善。

利益冲突：全体作者声明无利益冲突。

作者贡献：王倩研究实施、统计分析、撰写和修改文章；宋亮亮数据采集、研究实施；韩啸、刘明研究设计；张标、赵士博、谷宗运和黄莉莉研究实施、修改文章；李传富研究设计、经费支持；李小虎研究实施、修改文章；余永强研究设计。

[参考文献]

- [1] 中华医学会放射学分会. 放射科管理规范与质控标准(2017 版) [M]. 北京: 人民卫生出版社, 2017: 57-210.
- [2] 韩啸, 谷宗运, 赵士博, 等. 基于图像分类技术在医学影像数据治理过程中的研究与应用[J]. 中国医疗设备, 2023, 38(4): 78-83.
- [3] GUIA J, SOARES V G, BERNARDINO J. Graph databases: Neo4j analysis//19th international conference on enterprise information systems[C]. Philadelphia: ICEIS, 2017: 351-356.
- [4] RIDNIK T, LAWEN H, NOY A, et al. TRResNet: High performance GPU-dedicated architecture//IEEE/CVF winter conference on applications of computer vision [C]. Waikoloa: IEEE Computer Society, 2021: 1400-1409.
- [5] RIDNIK T, SHARIR G, BEN-COHEN A, et al. ML-decoder: Scalable and versatile classification head//IEEE/CVF winter conference on applications of computer vision [C]. Waikoloa: IEEE Computer Society, 2023: 32-41.
- [6] 宋若晨, 褚相乐, 黄勇华, 等. 基于颅脑 T1WI 对比增强图像构建卷积神经网络模型鉴别肺癌与乳腺癌脑转移[J]. 中国医学影像技术, 2023, 39(7): 982-986.
- [7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks//Advances

- in neural information processing systems[J]. Cambridge: MIT, 2012:25.
- [8] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition//ImageNet large scale visual recognition challenge[J]. Los Alamos: arXiv, 2014: 1409, 1556.
- [9] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions//IEEE conference on computer vision and pattern recognition[C]. Boston: IEEE Computer Society, 2015:1-9.
- [10] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition//IEEE conference on computer vision and pattern recognition[C]. Las Vegas: IEEE Computer Society, 2016:770-778.
- [11] ZHANG Y, WANG X, XU Z, et al. When radiology report generation meets knowledge graph//AAAI conference on artificial intelligence[C]. Los Alamos: arXiv, 2020:12910-12917.
- [12] 贾晓茜, 张向利, 刘哲, 等. 卷积神经网络在辅助胸部数字 X 线图像质量控制工作的应用价值[J]. 西安交通大学学报(医学版), 2019, 40(5):784-788.
- [13] 孟宇, 马之骋, 阮敬儒, 等. 基于 Faster R 卷积神经网络构建胸部 X 线片异物智能检测模型的可行性研究[J]. 中华放射学杂志, 2022, 56(12):1359-1364.

《中国介入影像与治疗学》杂志 2024 年征订启事

《中国介入影像与治疗学》杂志创刊于 2004 年,是由中国科学院主管,中国科学院声学研究所主办的国家级学术期刊,主编为肖越勇教授、李肖教授、金龙教授。刊号:ISSN 1672-8475, CN 11-5213/R。为中国精品科技期刊(第 2 届)、中国科技论文统计源期刊(中国科技核心期刊)、《中文核心期刊要目总览》收录期刊、荷兰《医学文摘》收录源期刊、WHO《西太平洋地区医学索引》(WPRIM)来源期刊、《日本科学技术振兴机构中国文献数据库》(JSTChina)收录期刊。

《中国介入影像与治疗学》杂志以报道介入影像与治疗学、介入超声学、介入材料学、药物学与护理学等方面的临床研究、基础研究及医、理、工结合的成果与新进展为主,是介入影像、介入治疗专业人员学习、交流的平台。

《中国介入影像与治疗学》为月刊,64 页,大 16 开,彩色印刷。单价:20 元,全年定价 240 元。订户可随时向当地邮局订阅,邮发代号:80-220;亦可向编辑部直接订阅,免邮寄费(银行、支付宝转账均可,附言栏请注明订阅杂志名称)。

网 址 www.cjiit.com

编辑部地址 北京市海淀区北四环西路 21 号,中科院声学所语声楼 201 室 邮编 100190

联系人 杜老师 联系电话 010-82547903 传真 010-82547903

银行账户名 《中国医学影像技术》期刊社有限公司

开户行 招商银行股份有限公司北京清华园科技金融支行

账 号 110907929010201

支付宝账号 cjmit@mail.ioa.ac.cn(账户名同银行账户名)

