

基于人工智能的数字 X 线机自助检查系统的设计

杨欣, 陈碧, 王辉, 罗江, 曾轩, 朱昱昊, 陈丽惠, 池显辰
(徐州医科大学医学影像学院, 江苏徐州 221000)

摘要: 为了实现数字 X 线机检查的体位识别智能化与摆位自助化, 采用自主研发的基于堆叠沙漏网络的体位识别装置、应用动作相关关系模型的体位判断装置和加入语音提示功能的数字 X 线机“自助检查”装置, 解决了如今影像摄片体位精准度不高、影像科技人员数量不足、传统摄片耗时较长等问题; 经多次实验结果表明, 系统将原本人均将近 6 分钟的摄影时间缩短到大约三分半, 将甲级片出产率的不足 40% 提升到将近 80%, 进而将误诊率减少 30%, 同时能够有效地减少医患接触, 减低院内感染的风险, 系统实现了缩短摄片时间、提高摄影质量、降低院内感染的目标。

关键词: 人工智能; 体位识别; 摆位不当; 堆叠沙漏网络; 计算机视觉技术

Design of Digital X-ray Machine Self Examination System Based on Artificial Intelligence

YANG Xin, CHEN Bi, WANG Hui, LUO Jiang, ZENG Xuan, ZHU Yuhao,
CHEN Lihui, CHI Xianchen

(College of Medical Imaging, Xuzhou Medical University, Xuzhou 221000, China)

Abstract: In order to realize the intelligent position recognition and self-help positioning of digital X-ray machine examination, the self-developed position recognition device based on stacked hourglass network, the position judgment device applying action correlation model and the “self-service examination” device of digital X-ray machine with voice prompt function are adopted to solve the problems, the position accuracy of image photography is low, the technicians in the imaging department is insufficient, and the traditional photography takes a long time. The experimental results show that the original photography time is reduced from nearly 6 minutes to about 3.5 minutes, the first-grade film productive rate increases from less than 40% to nearly 80%, and then the misdiagnosis rate reduces by 30%. At the same time, it can effectively reduce the contact between doctors and patients and the nosocomial infective risk. The system is realized by shortening the photography time, improving the photography quality and achieving the goal of reducing nosocomial infection.

Keywords: artificial intelligence; posture recognition; improper positioning; stacked hourglass network; computer vision technology

0 引言

摆位作为影像摄片中的核心环节, 一直以来都是医学生学习的重难点, 而摆位不当也一直是影响医生诊断、导致误诊漏诊的罪魁祸首^[1]。近年来, 影像科医学生成长速度慢、临床影像科医生水平良莠不齐、临床诊断对影像摄片的依赖性增高等问题逐渐凸显, 摆位不当、医疗资源紧张和影像科技人员压力大成为传统摆位的三大弊端。因此, 研究出一款能够实现辅助受检者摆位、提高摄影准确性、

缩短摄影时间这三大功能的产品具有十分重要的现实意义。

近年来, 新冠肺炎疫情在世界范围内蔓延, 发现并控制疫情传染源至关重要, 尤其是出现无症状感染者时, 利用数字 X 线机精准筛查可疑人员就成为医院排查的一项重要措施, 但由于部分地区防疫物资的不足, 医护人员的保护不尽人意。2019 年底全球大规模爆发的新型冠状病毒肺炎 (NCP) 和 2003 年的非典型肺炎 (SARS), 对人类生命和健康产生了重大的威胁, 医护人员也付出了沉重的代

收稿日期: 2021-08-17; 修回日期: 2021-09-09。

基金项目: 2021 年江苏省大学生创新训练计划省级一般项目 (202110313090Y)。

作者简介: 杨欣 (2000-), 女, 湖北孝感人, 主要从事医学影像学、医学影像技术学方向的研究。

通讯作者: 陈碧 (1981-), 男, 江苏徐州人, 医学博士, 副主任医师, 副教授, 主要从事呼吸系统疾病方向的研究。

王辉 (1981-), 男, 江苏徐州人, 高级实验师, 主要从事医学虚拟仿真教学、生物医学工程学等方向的研究。

引用格式: 杨欣, 陈碧, 王辉, 等. 基于人工智能的数字 X 线机自助检查系统的设计[J]. 计算机测量与控制, 2022, 30(3): 151-155.

价^[2]。此外,临床上摆位不当还是影响 X 线摄片质量、造成非甲级片出现的主要原因^[3],非甲级片往往会干扰影像医师做出准确判断,有时甚至会误导医师,造成耽误患者病情延误治疗时间的严重后果。另一方面,据不完全统计,影像技师每天的工作量平均 60~80 份磁共振或者 120~150 个超声检查^[4]。由于数字 X 线机在医院临床任务重,同时又具有射线辐射,不适合用于医学生实践教学。由此造成相关专业实践教学效果并不理想,医学生动手机会少,动手能力差,特别是在乡镇医院里低年资医师实际操作的机会较少,专科医师培养周期较长^[5]。据悉,每年国内放射科医疗工作量增速达 18.8%,而放射科医生增速为 4.1%,两者增长速率相差甚远,医疗资源相对不足间接激化医患矛盾^[6]。

基于人工智能的医用数字 X 线机自助检查系统就是针对上述需求,把人工智能应用于在数字 X 线机检查中,开发了体位识别装置、体位判断装置和语音提示纠正装置,对受检者实现体位自动识别,自主纠错,自我调整。而影像科工作人员只需在监视系统下按下曝光按钮便可完成检查,真正达到了方便、安全、高效的目的。

随着我国积极推进医疗现代化的发展,目前医用数字 X 线机的发展日趋完善,而人工智能与影像结合正在成为医疗检测和诊断领域的新趋势。AI 技术在医疗领域的布局与应用一直受到国内外科技巨头的高度重视,其中“医学影像+AI”是未来人工智能医学发展中最被看好部分,也是最有可能率先实现商业化的领域^[7]。人工智能诊断在临床上已有经开始了大面积应用,人工智能辅助摆位将成为下一个热点。美国的五大顶尖医院如梅奥、克里夫兰等都开始与人工智能公司合作,希望人工智能为医学影像技术助力,成为人工智能医疗应用领域的领跑者,实现疾病检查、诊断、治疗和管理的智能化。相比于美国等发达国家,国内企业在医学 AI 发展上紧跟时代步伐。北京航空航天大学生物医学工程的高精尖创新中心在荧光分子断层成像方面应用了 AI 技术,研究了一种基于深度学习技术的可以高精度、超快速成像的方法^[8]。深度学习技术在辅助诊断方面同样可以发挥重要作用。利用大数据与人工智能的相关技术来辅助放射科医生进行医学图像识别,不仅能提高医学影像数据识别的准确率,还能减轻放射科医生的工作压力,以达到疾病早发现、早治疗的目的^[9]。随着信息技术和知识经济的快速发展,人工智能的成果已经得到越来越多的应用。可以预见,在不久的将来人工智能的发展和普及必将势不可挡,未来人工智能将会成为医务工作者的绝佳助手。这些智能化发展也为人类的就医方式带来了巨大的转变,同样,病情诊断和疾病治疗的高效与准确也给人们的生活健康带来前所未有安全感^[10]。

1 “自助检查”系统的结构及原理

1.1 “自助检查”系统的结构

“自助检查”系统由图像采集装置、体位识别装置、体位判断装置和语音提示纠正装置 4 个部分组成。整套系统

可以快速准确辅助受检者“自助检查”,减轻影像科工作压力。“自助检查”系统结构如图 1 所示。

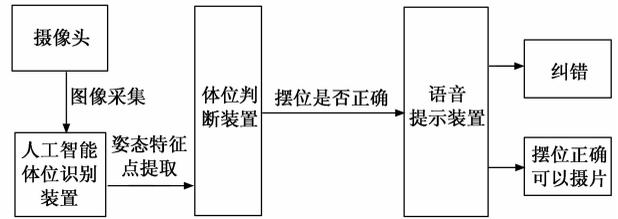


图 1 “自助检查”系统结构图

1.2 “自助检查”系统的原理

该系统采用软硬件结合技术,首先利用摄像头对受检者进行图像采集,接着利用基于堆叠沙漏模型的体位识别系统检测受检者姿态关键点,然后通过体位判断装置,将大数据中的正确体位与受检者体位对比,通过语音提示装置提示影像技术人员可以摄片或者提示受检者纠正体位。

2 基于人工智能的数字 X 线机“自助检查”系统的设计与实现

2.1 “自助检查”系统的设计

“自助检查”系统设计主要有一下几个步骤:1) 图像采集:首先令受检者在诊断床上或荧光板前就位,用 PC 端摄像头对受检者体位进行拍摄采集,最后将得到的图像命名并存储到特定文件夹。2) 体位识别:提取采集到的图像,利用添加了可变形卷积和注意力机制的堆叠沙漏模型检测受检者头、颈、胸、腹及四肢等大约 20 个姿态特征点并反馈坐标信息。3) 体位判断:提取受检者姿态关键点及坐标信息后,对比标准体位数据库里的姿态关键点信息,在 95% 的容错率下,判断受检者体位是否标准。4) 语音提示纠正装置将体位判断的结果通过播放器播放,从而对患者的不当体位进行实时纠错或者提示影像技师可以摄片。

2.1.1 图像采集

采用一个正对诊断床或荧光板的 PC 端摄像头对受检者体位进行拍摄采集,利用 OpenCV 得到照片的灰度图,以及再对灰度图进行二值化处理和腐蚀处理,最后将得到的图像命名并存储到特定文件夹。

2.1.2 体位识别

人工智能体位识别系统采用添加了可变形卷积和自注意力机制的堆叠沙漏网络(Stacked Hourglass Networks),通过提取采集到的图像,精确检测受检者头、颈、胸、腹及四肢等大约 20 个 X 线摄片姿态特征点并反馈坐标信息。

1) 堆叠沙漏网络:

堆叠沙漏模型的目标是在推理阶段就尽量在所有尺寸的特征图上捕获信息。首先通过若干个池化层和卷积层执行下采样程序,并通过反卷积进行上采样处理。然后将这两个过程中尺寸相同的特征图相加,正如 ResNet 中所做的相加一样,重复这种“先下后上”的处理来建立“堆叠沙漏”网络,并且在每个“沙漏”后进行中间监督训练。

传统的堆叠沙漏网络结构如图 2 所示,其中的每个沙

漏形状的结构即代表一个沙漏模块 (Hourglass Network), 多个这种结构堆叠起来, 称作堆叠沙漏 (Stacked Hourglass)^[11]。单个沙漏模块的网络结构如图 3 所示, 其中每一个方块即代表了一个残差模块, 方块的大小代表着输入特征图的尺寸, 残差模块网络结构如图 4 所示。沙漏模块的特点是具有对称的拓扑结构, 其体现在下采样使用卷积处理, 而上采样使用反卷积处理^[12-13]。首先是下采样过程中进行若干次卷积和最大值池化操作得到最小尺寸的特征图。在每次最大值池化操作后网络会产生一个包含着若干卷积的分支, 用于下一阶段的特征图融合。此分支的输入是最大值池化操作之前的特征图^[14]。得到最小尺寸图之后, 网络开始进行上采样和特征融合过程。上采样使用反卷积操作, 将上采样后的特征图与网络分支上同尺寸的特征图进行元素级的相加, 即可得到融合的特征图。残差模块使得 1×1 和 3×3 步长为 1 的卷积层不改变输入特征图的尺寸^[14]。图 4 表示输出卷积核大小和特征纬度, 其中跳层连接用虚线表示。中间监督 (intermediate supervision) 过程如图 5 所示, 虚线代表跳层连接, 沙漏图形代表了一个沙漏模块^[14], 右侧深色的方块代表着生成的 HeatMap, 训练阶段, 将这些 HeatMap 和真实 HeatMap 计算损失 MSE, 并求和, 得到损失; 推断阶段, 使用的是最后一个 Hourglass 的 HeatMap^[11]。

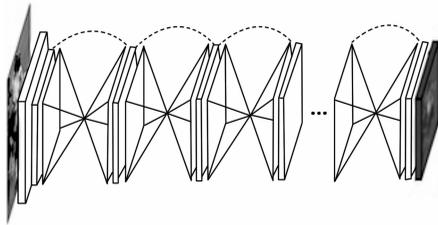


图 2 堆叠沙漏网络模型

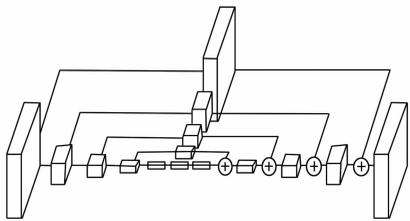


图 3 单个沙漏模块

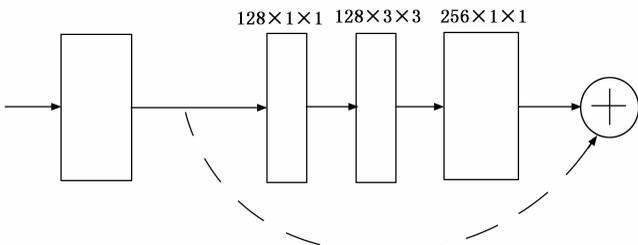


图 4 残差结构

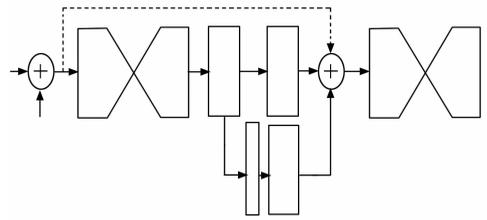


图 5 中间监督模块

2) 可变形卷积:

可变形卷积 (deformable convolution) 是指在常规的网格状空间采样中加入 2D 偏移 (offset), 即在标准卷积的基础上给每个卷积核的参数添加一个方向向量, 使我们的卷积核可以变为任意形状, 从而使采样点分布到我们感兴趣的区域, 更好地提取输入的特征, 达到增强 CNNs 对于变形目标的建模能力^[15]。

2D 卷积包含两个步骤, 首先用常规的 grid R 在输入特征图上 x 进行采样, 接着将这些采样结果以权重 w 进行归并。grid R 定义了空洞率和感受野的大小, 如 $R = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}$ 就定义了一个空洞率为 1、感受野尺寸为 3×3 的核^[16]。

在标准卷积中 p_n 是 R 中有限个位置的枚举, 对于输出特征图 y 中的每一个位置的 p_0 都有:

$$y(p_0) = \sum_{p_n \in R} w(p_n) x(p_0 + p_n)$$

在可变形卷积中, 常规网格 R 经过了偏移 $\{\delta p_n = n = 1, \dots, N\}$, 其中 $N = |R|$ 后, 对于输出特征图 y 中的每一个位置的 p_0 都有:

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \delta p_n) \quad (2)$$

因此采样就在不规则且有偏移的位置 $p_n + \delta p_n$ 上进行, 由于存在偏差 δp_n , 故可以通过双线性差值实现, 其中 p 表示任意的位置 (公式 (2) 中 $p = p_0 + p_n + \delta p_n$), q 是从特征图 x 中选择出的:

$$x(p) = \sum_q G(q, p) \cdot x(q) \quad (3)$$

因为 $G(\cdot, \cdot)$ 是二维的双线性插值的核函数, 故可以分成两个核, 其中 $g(qx, px) = \max(0, 1 - |qx - px|)$, $g(qy, py) = \max(0, 1 - |qy - py|)$:

$$G(q, p) = g(qx, px) \cdot g(qy, py) \quad (4)$$

如图 6 所示, 首先输入图像, 利用标准卷积核提取特征图, 接着把得到的特征图作为输入, 对特征图再施加一个卷积层, 从而得到可变形卷积的偏差, 输出的偏差域和输入特征图的尺寸相同, 通道维度的 $2N$ 对应 N 的二维偏移, 同时对生成偏移和输出特征的卷积核进行训练, 从而实现更准确的姿态估计^[16]。

3) 自注意力机制:

自注意力机制 (self-attention mechanism) 与人类观察外界事物机制很类似, 当人类观察外界事物的时候, 往往倾向于根据需要进行选择性地捕捉被观察事物的某些重要部分, 例如当看到图 7 的大量水果时, 人类往往能够迅速找

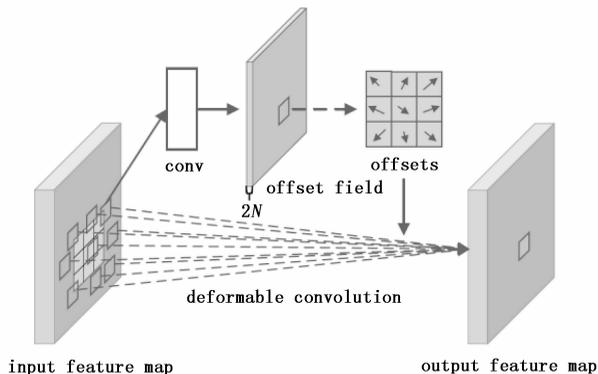


图 6 可变形卷积

出某种特定的水果，这正是因为人类可以选择性地捕捉信息，并对信息进行筛选^[17]。同理，自注意力机制可以对输入特征的不同部分赋予不同的权重，从关键区域中捕捉更重要的信息，从而减少了模型计算与储存的开销，同时使识别结果更为准确^[18]。



图 7 水果堆

自注意力机制可以视为一个特征提取层，输入特征 I ，提取部分特征 a^1, a^2, \dots, a^n ，经过特征提取层 (self-attention layer)，对不同的特征进行加权提取，得到新的特征 b^1, b^2, \dots, b^n ，通过融合新特征，输出特征 O ^[19]。输入特征 I 与输出特征 O 的区别在于，输出图像 O 放大了关键信息而忽略了次要信息，从而提高神经网络处理信息的能力。具体原理如图 8 所示，首先设输入特征为 I ，然后分别将其乘以 W_q 、 W_k 和 W_v 得到 Q (query) 矩阵、 K (key) 矩阵和 V (value) 矩阵，接着用 Q 矩阵与 K 矩阵相乘得到注意力 A 矩阵，又通过归一化处理得到 \hat{A} 矩阵，最后将归一化后的注意力矩阵 \hat{A} 矩阵乘上 V 矩阵，即可得输出特征 O ^[19]。

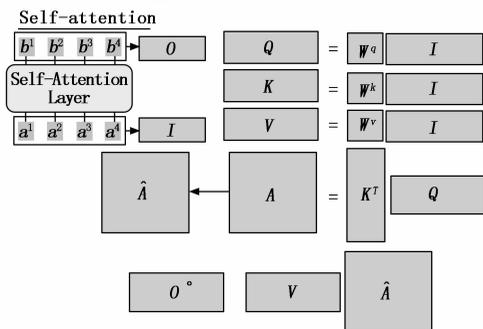


图 8 自注意力机制原理

2.1.3 体位判断

体位判断装置在 PC 端提取受检者姿态关键点及坐标信息后，调用标准体位数据库里的姿态关键点信息，并将二者进行对比，在 95% 的容错率下，判断受检者体位是否标准以及误差纠正。其中标准体位数据库是以 MPII 和 LSP 数据集为基础，与多家医院影像科达成合作、在专家的指导下自主研发。针对关键点信息的对比，本装置利用动作相关关系模型、关键点相对位置模型^[20]以及余弦角度模型进行错误姿势判断，整体算法基于改进后的堆叠沙漏模型，可以对照片中受检者姿态进行正误判断以及误差纠正。若姿态正确，则输出“CORRECT”，若不正确，则输出“ERROR”，并根据上述 3 个模型的判断结果，对错误姿势的纠正进行简单说明。

2.1.4 语音提示纠正装置

语音提示纠正装置将体位判断的结果通过播放器播放，从而在正式摄片前对患者的不当体位进行纠错或者提示影像技术人员可以摄片。

具体过程如下：首先把语音文案转成 MP3 格式语音文件，接着把语音文件提供给芯片方，定制语音；将焊接到单片机上，然后根据芯片规格书写控制时序程序，每条语音都对应了一个地址；最后将地址数据写入芯片，即可自动播放指定语音。

3 实验结果与分析

3.1 实验步骤

首先征集 20 名志愿者作为受检者。令受检者在诊断床上或荧光板前就位，摆出摄片体位，接着由图像采集系统采集并储存图像，用堆叠沙漏模型采集受检者姿态关键点，同时有可变形卷积优化关键点识别效果、注意力机制较少模型计算与储存的开销，然后用体位判断装置将受检者姿态关键点与标准体位数据库中姿态关键点进行对比，判断受检者体位是否标准以及误差纠正，若受检者体位不当，则通过语音提示装置对受检者进行纠错指导；若受检者体位正确，则提示影像技术人员可以摄片。图 9 是自助检查系统流程图。

系统测试实验结果主要包括两部分，一部分是对甲级片出片率、摄片时间和院内感染 3 个指标定量的测试，另一部分是对诊断错误率、受检者和医生满意度 3 个指标近似定性的测试。前三者的结果通过比较使用“自助检查”系统前后的医疗数据得出，后三者的结果通过对医生及受检者进行问卷的形式得出。

3.2 实验结果

根据实际情况，该系统只对部分实验数据结果进行了展示。图 10 展示了在应用数字 X 线机自助检查系统后，某三甲医院影像科在诊断错误率、甲级片出片和摄片时间等 6 项指标上的变化以及变化的幅度。

3.3 结果分析

实验结果表明，该系统在诊断错误率、摄片时间、院内感染的发生方面较未使用“自助检查”系统时均有减少；

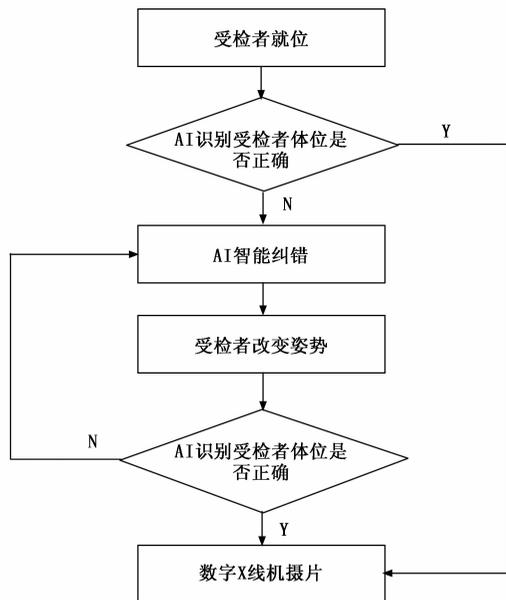


图 9 “自助检查”系统流程图

指标	某三甲医院影像科
诊断错误率	↓ 26% - 33%
甲级片出片率	↑ 30% - 36%
摄片时间	↓ 44% - 60%
院内感染	↓ 5% - 12%
受检者满意度	↑ 27% - 38%
医生满意度	↑ 40% - 66%

图 10 实验结果

甲级片出片率、受检者及医生满意度均有提升。摄片时间缩短了 50% 左右, 将甲级片出片率从不足 40% 提升到将近 80%, 进而将误诊率减少了至少 30%, 同时能够有效地减少医患接触, 减低了院内感染的风险, 系统实现了缩短摄片时间、提高摄影质量、降低院内感染以及提高医生和受检者满意度的目标。

4 结束语

传统摄片带来的一系列问题随着人们需求的不断增长而日益显现, 摆位不当、医疗资源紧张和影像科技技术人员压力大是目前亟待解决的问题。基于人工智能的数字 X 线机体位识别系统, 利用自主研发的人工智能体位识别装置、体位轮廓投射控件和贴心的语音提示功能, 帮助受检者自助检查, 缩短摄片时间, 同时提高摄片精确度, 减少误诊漏诊, 还能避免院内感染, 保护一线医务人员。医疗设备的更新换代, 可以挽救人们的生命; 而医疗模式的智慧化、升级化, 则可以改变人们的生活。

参考文献:

[1] PROMMA S, SRITARA C, WIPUCHWONGSAKORN S, et al. Errors in patient positioning for bone mineral density assessment by dual X-ray absorptiometry: effect of technologist re-

training [J]. Journal of Clinical Densitometry, 2018, 21 (2): 252 - 259.

[2] 张 鸽. 新型冠状病毒肺炎疫情下对医务人员院内感染管理的反思及其对策 [J]. 中华结核和呼吸杂志, 2021, 44 (1): 74 - 76.

[3] 冯希超. 浅谈质量控制管理对基层医院数字化 X 线摄影效果的影响 [J]. 影像研究与医学应用, 2019, 3 (8): 237 - 238.

[4] 刘志鹏, 侯瑞刚, 李 雯, 等. AI 医学影像学诊断新模式前景的调查研究 [J]. 山东医学高等专科学校学报, 2020 (3): 224 - 226.

[5] 周瑞泉, 纪洪辰, 刘 荣. 智能医学影像识别研究现状与展望 [J]. 第二军医大学学报, 2018, 39 (8): 917 - 922.

[6] 于杰夫. 人工智能识别 CT 颅内血肿影像及脑出血混合征相关算法的研究 [D]. 沈阳: 中国医科大学, 2018.

[7] 中华医学会核医学分会分子影像人工智能工作委员会. 分子影像人工智能专家共识 (2019 版) [J]. 中华核医学与分子影像杂志, 2019, 39 (12): 748 - 751.

[8] GUO L, LIU F, CAI C, et al. 3D deep encoder - decoder network for fluorescence molecular tomography [J]. Optics letters, 2019, 44 (8): 1892 - 1895.

[9] 周进凡. 基于大数据的肺部 X 光图像的分析与研究方法 [D]. 贵州: 贵州大学, 2019.

[10] 邵语千. 人工智能的现状与发展趋势 [J]. 科技传播, 2018 (19): 106 - 107.

[11] NEWELL A, YANG K, DENG J. Stacked hourglass networks for human pose estimation [C] //European conference on computer vision. Springer, Cham, 2016: 483 - 499.

[12] TOMPSON J J, JAIN A, LECUN Y, et al. Joint training of a convolutional network and a graphical model for human pose estimation [J]. Advances in neural information processing systems, 2014, 27: 1799 - 1807.

[13] LONG J, SHELHAMER E, Darrell T. Fully convolutional networks for semantic segmentation [C] //Proceedings of the IEEE conference on computer vision and pattern recognition, 2015: 3431 - 3440.

[14] 刘昊宇, 韩梓轩. 人体关键点检测模型研究 [D]. 哈尔滨: 哈尔滨工业大学, 2019.

[15] MCNEELY-WHITE, DAVID G. Same data, same features: modern ImageNet-trained convolutional neural networks learn the same thing [Z]. 2020.

[16] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks [C] //Proceedings of the IEEE international conference on computer vision, 2017: 764 - 773.

[17] 黄海新, 王瑞鹏, 刘孝阳. 基于 3D 卷积的人体行为识别技术综述 [J]. 计算机科学, 47 (11): 139 - 144.

[18] 邱锡鹏. 神经网络与深度学习 [M]. 北京: 机械工业出版社, 2020.

[19] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C] //Advances in neural information processing systems, 2017: 5998 - 6008.

[20] 解 辉. 基于视觉的人体动作矫正系统 [D]. 济南: 山东师范大学, 2020.