

智慧口腔正畸助手设计与原型实现

李娜¹, 郭磊², 王杰¹, 王玉玲¹, 刘海英¹

(1. 山东第一医科大学第一附属医院/山东省千佛山医院口腔医学, 山东 济南 250014;

2. 山东第一医科大学/山东省医学科学院现代教育技术中心, 山东 济南 250014)

摘要:智能技术在正畸领域应用前景广阔,本研究通过全面分析正畸诊疗全流程业务场景,综合采用人工智能技术,设计了智慧助手的分层技术架构,给出了各层应提供的服务集合,实现了多终端信息系统、语音识别引擎、人脸识别引擎等核心系统原型。实验结果表明,智慧助手设计科学合理,系统语音识别和人脸识别成功率较高,能满足正畸业务需要,可提高正畸医生诊疗效率,提升患者就医体验。

关键词:口腔正畸;智慧助手;语音交互;图像识别

中图分类号:R783.5

文献标识码:B

DOI:10.3969/j.issn.1006-1959.2022.13.006

文章编号:1006-1959(2022)13-0034-05

Design and Prototype Implementation of Intelligent Orthodontic Assistant

LI Na¹, GUO Lei², WANG Jie¹, WANG Yu-ling¹, LIU Hai-ying¹

(1. Department of Stomatology, the First Affiliated Hospital of Shandong First Medical University/

Shandong Provincial Qianfoshan Hospital, Jinan 250014, Shandong, China;

2. Modern Educational Technology Center, Shandong First Medical University/Shandong Academy of Medical Sciences,

Jinan 250117, Shandong, China)

Abstract: Intelligent technology has broad application prospects in the field of orthodontics. In this study, through a comprehensive analysis of the whole process business scene of orthodontic diagnosis and treatment, the hierarchical technical architecture of intelligent assistants is designed by using artificial intelligence technology, and the service set of each layer is given. The core system prototypes such as multi-terminal information system, speech recognition engine and face recognition engine are realized. The experimental results show that the intelligent assistant design is scientific and reasonable, and the success rate of speech recognition and face recognition is high. It can meet the needs of orthodontic business, improve the efficiency of orthodontic diagnosis and treatment, and improve the patient's medical experience.

Key words: Orthodontic; Intelligent assistant; Voice interact; Image recognition

语音交互、图像识别、机器人、专家系统是人工智能(artificial intelligence)领域的研究热点。矩阵计算性能的大幅提升及深度学习技术的快速发展,为人工智能的应用奠定了坚实的技术基础,并在医学领域的研究与应用进展迅速^[1,2]。口腔正畸诊疗周期长、复诊次数多,病历记录复杂,牙齿模型测量分析、诊疗方案设计、效果模拟与动态调整专业性强,过度依赖医生直接参与,智慧应用水平不高。人工智能在口腔正畸领域广泛应用,可提高正畸诊疗的效率,提升医患交互体验,应用前景广阔。本文主要对智慧口腔正畸助手设计与原型实现进行概述,以期临床应用提供参考。

1 需求分析

1.1 应用现状 目前,口腔正畸诊疗智慧应用严重不足,主要表现为:①缺乏正畸专用信息系统,导致正畸病历仍需手工记录,患者影像、图片、数字建模资料及病历分散存储,查询分析效率低下;②缺乏智慧

正畸方案设计与模型可视化应用,导致正畸模型测量、方案设计、弓丝弯制等仍需要全手工操作,测制精度不高,正畸效果可预见性差;③缺乏智慧医患交互体验应用,导致患者不能充分参与其诊疗过程,不能享有人工智能技术带来的便捷智慧的就诊体验。

已有应用研究如语音电子病历在口腔医院的应用^[3-5]、机器人辅助弯制正畸弓丝^[6-8]、基于规则推理的专家系统用于诊疗方案设计^[9]、口腔正畸模型的病历管理、测量管理和专家系统^[10-12]、具备智能提醒功能的口腔信息管理系统^[13,14]、口腔修复全程信息化管理系统^[15]等,为提升口腔正畸诊疗的自动化、智慧化水平进行了有益探索,但大都聚焦在人工智能技术应用的某一方面,缺乏对人工智能应用的系统性、整体性研究。

1.2 建设目标 近年来,随着以深度学习为代表的人工智能技术飞速发展,传统领域智慧赋能产生了很多创新性应用。本文设计了一款智慧口腔正畸助手,给出了分层可扩展的系统架构,列出了全应用场景,实现了核心系统原型,对人工智能技术在口腔正畸中的应用进行了全面的探索,以期工业级的系统建设和应用推广提供借鉴。

基金项目:山东省中医药科技发展计划(编号:2019-0357)

作者简介:李娜(1984.8-),女,山东聊城人,硕士,主治医师,主要从事口腔正畸学及人工智能研究

2 体系架构设计

此正畸智慧助手具有五层技术架构,分别是交互设备层、智慧应用服务层、语音交互与图像处理引擎、人工智能与大数据处理平台、信息技术基础设施。

擎层、人工智能与大数据平台层和信息技术基础设施层,各层之间协同工作、轻耦合连接。智慧助手体系架构见图 1。

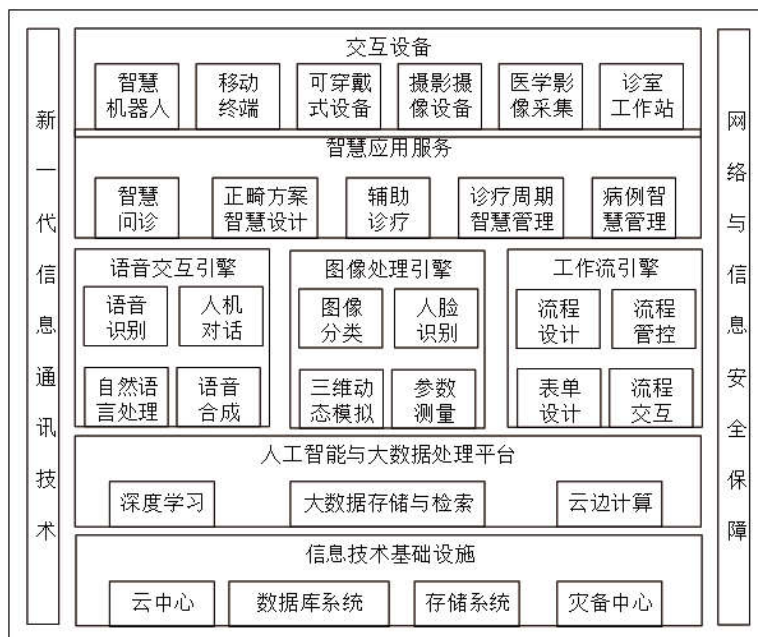


图 1 智慧助手体系架构

2.1 交互设备层 交互设备层是智慧助手与正畸医生和患者沟通的媒介,负责文字、语音、图片与影像信息的交互。智慧机器人是一个综合交互式终端,除了支持语音交互与图片影像显示外,还支持隔空手势操作,自带的智能耗材柜可以根据指令自动打开,最大程度解放正畸医生双手。摄像扫描设备提供各类正畸图片如正面像、正面微笑像、侧面像等,也可扫描正畸模型。医学影像采集设备自动采集患者 X 光片及 CT 片。移动终端(手机、PAD)、智能穿戴设备(手表、眼镜等)也可用来完成语音交互与图像采集。诊室工作站提供诊室相关数据快速存取、边缘计算、网络中继等功能。

2.2 智慧应用服务层 智慧应用服务层提供智慧助手的核心服务功能,包括智慧问诊、正畸方案智慧设计、辅助诊疗、诊疗周期智慧管理、病历智慧管理等。

2.2.1 智慧问诊 采用语音合成技术可以使智慧助手“克隆”医生的音色,代替正畸医生进行问诊,在就诊前即可完成患者个人情况、主诉及病史、正畸期望、复诊原因等信息的采集;现场拍摄的 X 光片、CT 片、面像照片经图像自动识别出图像类型后存入对应患者的工作站数据库中,不需手动命名分类识别。

2.2.2 正畸方案智慧设计 正畸模型经扫描仪、相机拍摄、三维 CT 扫描等手段数字化,在进行图像矫正

与定点后,自动进行模型测量,得出牙列拥挤度、Bolton 指数、牙弓长度宽度等数据。自动分析数字化的 X 线头颅侧位片得出牙颌面软硬组织结构的角度、线距等测量值,并对牙颌面软硬组织自动综合测量分析。根据测量参数自动调用 Bonwill-Hawley 方法绘制理想牙弓形态图并打印,另根据弓形参数自动弯制特定型号的弓丝,以及不同正畸方案下牙齿三维动态演变与表征。

2.2.3 辅助诊疗 诊疗前同步对比患者面像和口内像,根据语音指令自动检索患者不同时期的影像图片进行对比分析,优化复诊方案。此外,诊疗中智慧助手机器人可根据指令,将所需弓丝、钳子等器械耗材送到医生手中。

2.2.4 诊疗周期智慧管理 根据正畸方案智慧设计结果,调用 workflow 引擎,自动生成个性化正畸诊疗工作流程图,每 1 个流程节点代表一次诊疗活动,活动列表给出每次诊疗的日期、预计诊疗的内容、所需的正畸工具与耗材、预计诊疗所需时间与费用。预约管理可以根据流程执行状态自动生成预约信息推送给医生和患者。

2.2.5 病历智慧管理 病历智慧管理负责病历智慧录入、病历检索、病历分析、病历归档等功能。智慧录入功能可通过智慧助手机器人、医生智能手表、移动麦

克风等设备采集医生口述、医患对话、患者自述等语音信息自动填充并生成病历,诊疗与病历记录同步进行。病历检索、分析可根据医生指令或手势操作进行病历检索和对比分析。此外,病历可直接归档到工作站与云中心,可自动打印生成纸质病历和语音电子病历。

2.3 语音交互与图像处理引擎层

2.3.1 语音交互引擎 语音交互引擎可实现语音识别、自然语言处理、人机对话、语音合成等功能,且可通过口述指令如“小美、小美”唤醒智慧助手,医生患者的语音数据首先通过语音识别模块处理成文字;文字再经自然语言处理模块识别成真实语义被智慧助手所理解,可以记录成病历,也可以作为医生的某项指令执行。人机对话可以支持医生与智慧助手之间的指令交互与病历记录,支持患者与智慧助手之间的自动问诊、病历查询、诊疗效果互动展示等。同时,语音合成可以对特定医生音色进行学习“克隆”,增强患者与智慧助手交互的真实情感。

2.3.2 图像处理引擎 图像处理引擎可实现人脸识别、图像分类、参数测量、三维模拟等。人脸识别可根据初诊采集的患者面部图片信息实现患者自动识别,患者可“人脸”挂号复诊、支付及与智慧助手交互。图像分类可自动实现采集图像的类型,如头颅侧位片、曲面断层片、正面像、下颌像等的自动标签分类。参数测量可根据三维模型、影像资料为正畸方案设计提供相关参数,也可根据侧位片的脊柱关节形态确定骨龄。三维模拟可实现全诊疗周期牙齿三维模型动态可视化,提供给医生、患者虚拟现实交互体验,也可根据正畸方案参数调整实时模拟诊疗效果变化。

2.3.3 工作流引擎 流程设计可以自动生成诊疗流程图、诊疗活动清单、定义活动输入与约束;流程管控可实现对整个正畸过程的自动管理,正畸医生可根据诊疗需要进行流程启动、挂起、变更等操作;表单设计可根据每次诊疗活动的特点设计通用或个性化的病历内容。流程交互支持主流程与子流程的协同,以及不同流程间的数据与任务共享。

2.4 人工智能与大数据处理平台

2.4.1 深度学习 利用 PaddlePaddle、TensorFlow、Pytorch 等开源平台,为引擎层提供深度学习算法支持、模型训练以及推理框架和基础模型库。在此基础上,针对正畸专业和图像影像特点,优化生成正畸专业模型库,提高语音与图像识别准确度。

2.4.2 大数据存储与检索 支持患者信息、诊疗信息、

设备与耗材信息等结构化数据,电子病历、面部照片、口腔影像等非结构化数据的分布式存储,同时也支持诊室工作站快速缓存检索与云中心检索,以及支持面向语音与图片的智慧检索。

2.4.3 云边计算 支持边缘计算和云计算双模计算方式,部分轻量化图像识别与语音计算模型部署在智慧手机机器人等终端设备上,提高系统响应速度,发挥终端设备计算能力,降低云中心负载。

2.5 信息技术基础设施 信息技术基础设施为系统提供计算与数据存储功能,包括高性能 GPU 集群、分布式数据库与数据仓库系统,高速存储与容灾备份系统。

3 系统原型实现

本研究实现了智慧助手的系统原型,包括多终端信息系统、语音识别引擎、人脸识别引擎。

3.1 多终端信息系统 实现多终端信息系统是智慧助手的基础与核心。客户端采用 Python3.9 开发、PyQt5 渲染 GUI 界面,移动端采用 Android Studio 开发,后台数据库采用开源分布式 MySQL 实现,见图 2。多终端信息系统主要功能包括患者管理、诊断与治疗、病历记录与查询、材料使用与库存、系统管理等模块,实现患者信息与口腔检查结果采集与录入、模型分析与影像测量结果采集与录入、诊断与矫治计划生成、病历一键生成与打印、自动绘制 Bonwill-Hawley 弓形图等功能,见图 3。

3.2 语音识别引擎 实现语音识别采用基于深度学习平台 PaddlePaddle 的 Deep Speech 2^[16],实现端到端自动语音识别^[17]。计算采用 NVIDIA CUDA^[18] 10.2 平台,训练数据集分两种,一种为 openSLR 共享的 Free ST American English Corpus、Free STChinese-MandarinCorpus、THCHS-30^[19]、Aishell^[20]等 4 个通用数据集;一种为正畸专业数据集,选取 10 份完整正畸病历,由 3 位正畸医生在安静房间内,分别阅读病历,同时使用 Android 手机,以频率 16 kHz 录制,每个短句生成一个音频文件和一个脚本,每个病历生成约 150 个音频文件和脚本文件。采用数据增强方法进行模型训练,首先得到一个初步通用模型,然后用正畸专业数据集对初步模型进行迁移学习,最终生成正畸专业预测模型。最后进行模型验证,由第 4 位医生,在同样条件下阅读第 11 份病历,生成 152 个音频文件和脚本文件,将音频文件逐一导入预测模型输出文本结果,对比对应的脚本文件,用字符错误率来评价模型的性能。

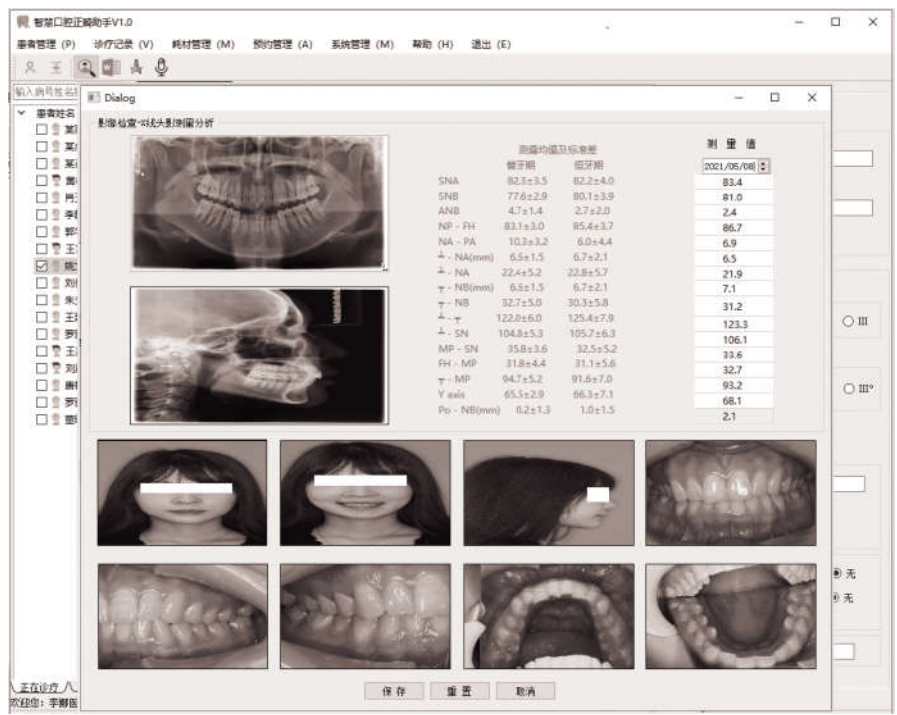


图 2 智慧助手多终端信息系统

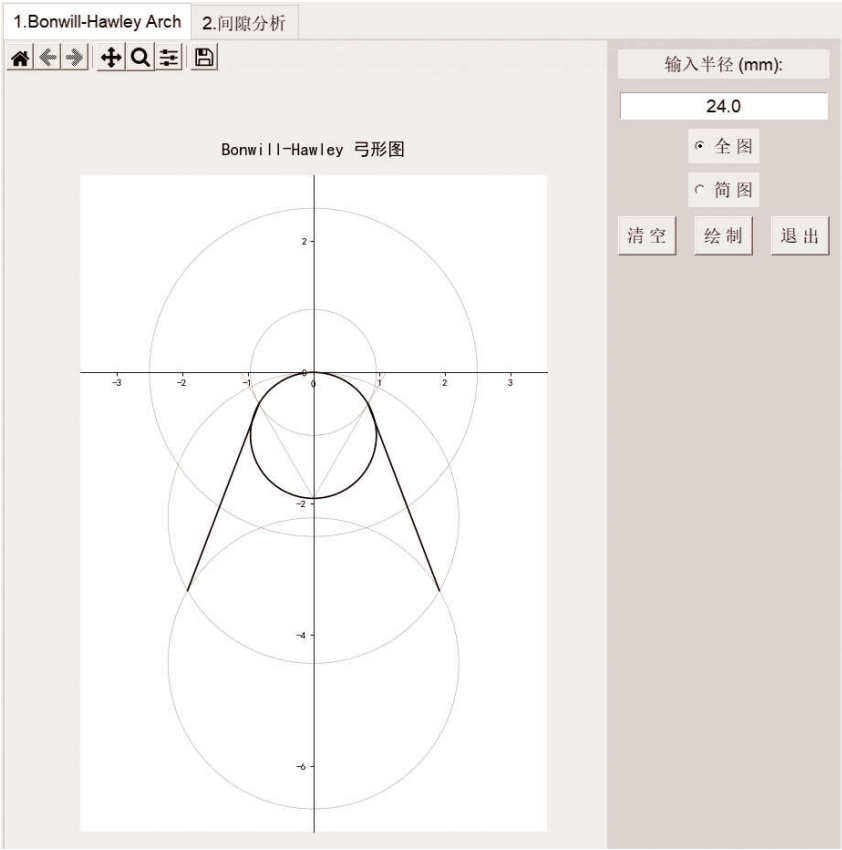


图 3 智慧助手自动绘制 Bonwill-Hawley 弓形图

3.3 人脸识别引擎 实现人脸识别基于深度学习平台 PaddlePaddle 实现，计算采用 NVIDIA CUDA^[18] 10.2 平台。采用数据集包括 CASIA-WebFace 基础数据集和 200 名患者的正面像和正面微笑像作为补充数据集，同时所有患者的正面像作为患者身份照

存入注册库。首先对所有图像进行裁剪、对齐、数据增强处理，然后采用残差神经网络^[21]Resnet-101 进行模型训练，模型输出为人脸特征向量，同时将该特征向量存入对应的患者图像注册库。当患者来就诊时，通过智慧助手自动采集患者人脸图像，调用预训

练模型得出人脸特征向量,逐一与注册库中患者特征向量进行比对,计算余弦相似度,余弦相似度大于0.9视为同一患者。

4 应用效果

系统实现的语音识别与人脸识别两大人工智能基础引擎,识别成功率分别为91.5%、94.5%,可为人工智能技术全面应用提供基础支撑。应用智慧病历管理服务,医生可快速查找患者诊疗记录、综合分析各类影像资料,动态调整诊疗方案,提高诊疗效果。应用辅助弓丝弯制服务可提高弓丝弯制精度,减少医生测量耗时。采用多终端信息系统,就诊患者进入诊室就可在移动终端上交互动完成身份识别、基本信息与面部照片等信息录入,大幅降低医生问诊时间,提升医患交互体验。

5 总结

本研究设计了一款正畸智慧助手,给出了其分层技术架构,详细阐述了各层架构应提供的服务,同层之间内聚性强、相邻层间松散耦合,设计合理便于扩展,同时实现了多终端信息系统、语音识别引擎、人脸识别引擎等技术架构部分核心系统原型。因实验计算平台性能限制,系统响应时间略长,但系统语音识别和人脸识别成功率高,能满足实际业务需要。为进一步提升自动识别成功率和响应速度,设计并实现产品级的实体智慧助手机器人和可穿戴设备是下一步工作方向。

参考文献:

- [1]刘蓬然,霍彤彤,陆林,等.人工智能在医学中的应用现状与展望[J].中华医学杂志,2021,101(44):3677-3683.
- [2]毕小琴,赵佛容.人工智能技术在口腔专科治疗及护理中的应用[J].华西口腔医学杂志,2018,36(4):452-456.
- [3]王鹏程,杨辉,张颖钊.语音电子病历在口腔医院的应用[J].医学信息,2020,33(19):19-20.
- [4]蒋盼,王捷,徐冬,张志宏.基于语音入口的电子病历在口腔科中的应用与改进[J].中国数字医学,2018,13(10):75-77.
- [5]于茵茵,王旭东.语音识别技术在口腔门诊病历系统中的应用[J].河南科技,2019(23):36-38.
- [6]Rigelsford J. Robotic bending of orthodontic archwires[J]. Industrial Robot: An International Journal, 2004, 65(4): 321-335.
- [7]张永德,蒋济雄.正畸弓丝弯制特性分析及实验研究[J].中国

机械工程,2011,22(15):1827-1831.

- [8]张永德,左思浩,姜金刚,等.个性化正畸弓丝曲线交互调整方法研究[J].仪器仪表学报,2017,38(7):1616-1624.
- [9]谢晓秋,王林,王阿明.专家系统在口腔正畸等口腔医学领域中的应用[J].现代口腔医学杂志,2009,23(3):315-319.
- [10]周传涛.口腔正畸模型的计算机辅助测量分析系统研究[D].重庆:重庆大学,2007.
- [11]姜逸菲.基于三维电子散斑干涉技术口腔正畸力学特性实验研究[D].南京:东南大学,2020.
- [12]宋红芳,吴宝平,赵丽颖,等.三维重建在口腔生物力学中的应用[J].医学信息,2018,31(7):24-25.
- [13]黄文静,赵浩宇,熊宇,等.智能化的口腔管理信息系统设计与应用[J].中国卫生信息管理杂志,2021,18(6):814-818.
- [14]蒋冰.口腔医院信息系统的设计与实现[D].济南:山东大学,2017.
- [15]徐晓明,魏晓岚,印黎涛.口腔修复全程信息化管理的探索与实践[J].中国卫生信息管理杂志,2018,15(5):579-583.
- [16]Deep speech 2: end-to-end speech recognition in English and mandarin[C]//International Conference on Machine Learning. hgpu.org, 2016.
- [17]He Y, Sainath TN, Prabhavalkar R, et al. Streaming End-to-end Speech Recognition for Mobile Devices [C]//ICASSP 2019 -2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019.
- [18]NVIDIA. NVIDIA CUDA Compute Unified Device Architecture: Programming Guide [EB/OL]. https://developer.download.nvidia.cn/compute/cuda/1.0/NVIDIA_CUDA_Programming_Guide_1.0.pdf, 2007-06-23/2021-04-25.
- [19]Wang D, Zhang XY, Zhang ZY. THCHS-30: A Free Chinese Speech Corpus [EB/OL]. <https://arxiv.org/pdf/1512.01882v1.pdf>, 2015-12-07/2021-04-25.
- [20]Bu H, Du JY, NA XY, et al. AISHELL-1: An open-source Mandarin speech corpus and a speech recognition baseline [C]//2017 IEEE Conference of the Oriental Chapter of the International Coordinating Committee on Speech Databases & Speech I/o Systems & Assessment, 2017: 1-5.
- [21]He KM, Zhang XY, Ren SQ, et al. Deep Residual Learning for Image Recognition [EB/OL]. <https://arxiv.org/pdf/1512.03385.pdf>, 2015-12-10/2021-04-25.

收稿日期:2022-05-05;修回日期:2022-05-23

编辑/杜帆