

# 基于语音识别的中医临床诊疗智能语音助理设计与实现\*

郭哲韬 石文丽 杨 涛

(南京中医药大学人工智能与信息技术学院 南京 210023)

[摘要] 介绍语音识别在国内外医疗领域应用情况, 阐述中医临床诊疗智能语音助理设计和实现方法, 包括技术架构、功能架构、系统架构, 分析相关问题, 指出该系统能够实现对中医临床诊疗语音降噪录入和准确识别, 对提高中医诊疗效率、推动中医临床智能化发展具有一定应用价值。

[关键词] 中医临床; 语音识别; 语音助理

[中图分类号] R-058 [文献标识码] A [DOI] 10.3969/j.issn.1673-6036.2022.09.012

**Design and Implementation of the Intelligent Voice Assistant for Clinical Diagnosis and Treatment of Traditional Chinese Medicine Based on Speech Recognition** GUO Zhetao, SHI Wenli, YANG Tao, School of Artificial Intelligence and Information Technology, Nanjing University of Chinese Medicine, Nanjing 210023, China

[Abstract] The paper introduces the applications of speech recognition in the medical field at home and abroad, expounds the design and implementation methods of the intelligent voice assistant for clinical diagnosis and treatment of Traditional Chinese Medicine (TCM), including technical architecture, functional architecture and system architecture, analyzes relevant issues, and points out that the system can realize the voice noise reduction input and accurate recognition of TCM clinical diagnosis and treatment, which has certain application value to improve the efficiency of TCM diagnosis and treatment and promote the development of TCM clinical intelligence.

[Keywords] Traditional Chinese Medicine (TCM) clinical practice; speech recognition; voice assistant

[修回日期] 2021-11-01

[作者简介] 郭哲韬, 本科生; 通信作者: 杨涛, 博士, 副教授。

[基金项目] 江苏省高等学校大学生实践创新训练计划“中医临床智能语音助理的设计及实现”(项目编号: 202010315039Z); 江苏高校哲学社会科学一般项目“基于医案数据挖掘的名老中医临床经验数字化传承模式构建及方法研究”(项目编号: 2021SJA0333)。

## 1 引言

在信息技术飞速发展背景下国家提出中医药信息化发展战略。中医药传承与发展离不开新技术支撑。其中中医智能辅助诊疗系统是中医药信息化建设的重要引擎和技术支撑<sup>[1]</sup>。语音识别技术是将语音作为一种输入内容让计算机自动进行识别和分析, 涉及模式识别、信号处理和计算机接口等<sup>[2]</sup>, 是一门包括声学、概率论、信息论和计算机科学的交叉学科<sup>[3]</sup>。随着医院信息化发展, 中医电子病历应用越来越广泛。然而传统采用计算机录入病历的

方法效率不高,特别是名老中医由于不熟悉电脑操作,在患者较多的情况下为提高诊疗效率其医案记录往往较为简单,患者临床信息记录不全,为后续中医经验总结和传承带来不便。提高中医医案数据采集与整理效率是推动中医传承和研究尤其是名老中医临床经验传承的关键问题之一。利用智能语音识别技术可以提供有效的解决方案,将诊疗原始语音进行准确录制存档,同时利用训练好的中医诊疗语音识别模型进行准确识别并显示和存档文本,将传统诊疗中冗余繁杂的诊疗过程进行优化,记录形式上变“写病历”为“说病历”,可以大幅提高诊疗服务效率。

## 2 语音识别在国内外医疗领域的应用探索

### 2.1 国外

国外智能语音识别技术在医疗领域的应用已较为成熟。Nuance 公司的临床语音识别终端为全球超过 50 万名医生及 1 万多个医疗机构等提供技术支持。数据显示,在美国使用 Nuance 产品 Ambient Clinical Intelligence 的医生平均效率可提高 30%<sup>[4]</sup>。Amazon 公司研发的语音助手能够高效采集电子病历信息,为医生的快速准确诊疗提供方便<sup>[5]</sup>。谷歌公司的相关模型借助深度学习技术可对电子病历进行语音转写,使患者及家属对于自身病情有更加精确的了解<sup>[6]</sup>。MediaInterface 公司为欧洲多国医疗机构提供语音技术解决方案,其核心产品 SpeaKing 作为一种利用人工智能(Artificial Intelligence, AI)技术支持医疗文档语音识别的高效工具被 600 多家医院、700 多家医疗机构、75 000 多名用户使用<sup>[7]</sup>。

### 2.2 国内

国内医疗智能语音识别技术虽然起步较晚,但发展迅速。中科汇能科技有限公司将大量医疗文本用于模型训练,致力于语言电子病历的研究开发,其开发的产品“医语通”识别准确率可达 97%。以语音转录及人工智能技术为核心的科大讯飞公司采用软硬件结合的方法实时记录诊疗信息并即时生成结构化电子病历<sup>[8]</sup>。从 2017 年开始科大讯飞口腔科语音电子病

历系统在北京协和医院投入使用。云知声智慧医疗语音录入系统可以为临床诊疗语音录入提供便捷的软硬件一体化解决方案,在医疗人工智能技术和大数据分析技术的支持下,智能语音交互和病历查阅、健康风险预测和对患者分群分析得以实现。百度、东软和阿里健康等公司也在智慧医疗语音识别领域进行研究和实践。目前还没有专门针对中医药领域的语音助理或语音电子病历。中医和西医的诊疗是两个不同的理论系统,中医讲究“望闻问切”四诊合参,具有独特的术语体系,因而中医诊疗语音模型无法采用传统诊疗语料进行训练,需要采集大量专门的中医临床语料进行训练,同时结合中医知识库构建适用于中医领域的专用语音助理。

## 3 中医临床智能语音助理设计与实现

### 3.1 技术架构

3.1.1 概述 中医临床智能语音助理核心是语音识别模型。通过采集中医临床语音数据训练适用于中医临床的语音识别模型,从中医领域特定的句子分割、词性标注、实体识别等方面解决中医学术语识别问题。通过设计软硬件集成平台实现语音采集、分析和识别,最终构建中医智能语音助理系统(以下简称系统)。

3.1.2 中医语言识别过程 包括语音信息预处理、特征提取、模式匹配等核心步骤。首先对获取的中医诊疗语音进行诊疗音频信号处理、语言识别与智能降噪,再利用梅尔倒谱系数(Mel Frequency Cepstrum Coefficient, MFCC)特征提取算法得到关键信息并将其与语料库中中医诊疗语音识别模型进行匹配,通过修正语言信息获取不全、多音字等问题,最后得到语言识别结果,见图 1。

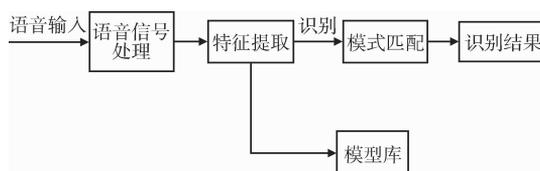


图 1 语言识别模型构建过程

3.1.3 系统构成 包括前端和后端两个基本模

块。前端模块主要功能是端点检测、降噪、特征提取等；后端模块主要功能是利用训练好的声学模型或者语言模型对诊疗语音特征向量进行统计模式识别（又称解码），得到其包含的文字信息。此外后端模块中还存在一个自适应反馈模块，可以对用户临床诊疗语音进行自学习，从而对声学模型和语音模型进行矫正完善，进一步提升语音识别准确率。

### 3.2 功能架构

3.2.1 系统主要功能 包括中医诊疗的语音读取、语音处理、语音识别、语音分析、病历生成、病历存档功能。语音读取功能可以录制中医诊疗对话音频；语音处理功能可以对诊疗语音进行初步降噪处理；语音识别功能可以实现中医临床诊疗语音文件转写以及中医诊疗实时语音转写；语音分析功能可以对语音识别结果进行基于中医临床诊疗模型的评估；病历生成功能可以提取出原始诊疗音频中的关键信息；病历存档功能可以将提取出的关键信息保存至数据库。此外还有用户注册、使用说明、客服反馈、记录信息等基本功能，见图 2。

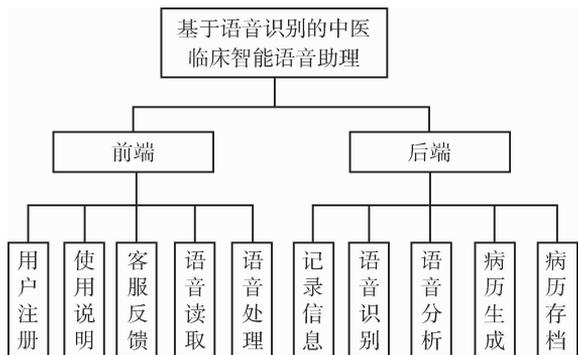


图 2 软件功能架构

3.2.2 结构化医案及其他功能 通过语音识别和自然语言处理技术将非结构化医疗文本转化为结构化数据，有利于诊疗信息存储、学习与分享。结构化医案供医师及科研人员追踪改善诊疗方案和学习借鉴，从而提高中医传承效率。此外平台支持人工在线客服线上咨询功能以及常见使用问题指导方案，患者可以进行初步问诊和问题反映。对中医临床诊治信息进行智能语音识别可以将病历信息进行完整存储与场景再现。

### 3.3 模型构建

采集大量原始诊疗语音数据并进行预处理，提取长度适当的诊疗语音制作诊疗模型训练集。通过集成标准的讯飞语音识别控件，结合 Sphinx-4 语音识别引擎部署中医语音识别平台。在经过初始化配置信息、特殊发音识别、语音模型配置等前期准备后进行声学模型训练，使用充足、同规格中医诊疗语音数据中的中医术语语音与语音词典进行映射，形成术语语音到文字的对应表，完善 Sphinx-4 模型对应表，最终得到中医诊疗语音识别模型，见图 3。

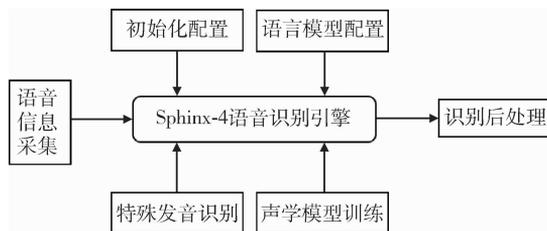


图 3 模型构建流程

### 3.4 系统架构

3.4.1 总体架构 采用客户机/服务器（Client/Server, C/S）架构设计中医临床诊疗智能语言助理，利用 Python 作为前后端处理语言，借助丰富的语音处理框架和较为成熟的中医语音识别模型进行音频数据处理。采用可视化网页编程技术，设计相应前端使用界面和框架，用于音频数据采集、归档、识别等。采集的音频进行压缩编码后传递到服务端进行语音识别解析；识别完成得到较为准确的结果后通过网络传递到后端，后端再将识别文本传递到前端显示。系统在使用时部分功能需协同作用，因此在界面设计过程中将语音读取功能与语音处理功能、语音识别功能与语音分析功能、病历生成功能与病历存档功能进行合并，形成 3 个核心模块。

3.4.2 中医诊疗语音读取与处理模块 该模块将诊疗语音进行录制与存档，录制过程会进行语音预处理以提升信噪比，减少噪声对实验数据的影响，录制完成后自动本地存档。

3.4.3 中医诊疗语音识别与分析模块 该模块将中医临床诊疗语音进行文件转写或实时转写，经过中医诊疗语音模型转写后中医术语识别准确率可达

90% 以上, 准确率将随模型完善进一步提升, 转写完成后可供医生和患者进行后续分析。

3.4.4 中医诊疗病历生成与存档模块 该模块可以在语音识别完成后提取识别结果中的关键词并进行存档, 便于医生与患者快速获取诊疗关键信息, 并可优化二次诊疗的数据。

## 4 分析及讨论

### 4.1 语音识别技术实际应用中的噪声处理相关问题

本研究将语音识别技术应用到中医临床诊疗领域, 实现高效便捷的临床诊疗信息录入, 为医生带来便利。由于在实际诊疗过程中外部环境较为嘈杂, 语音识别干扰项较多, 包括方言和口音问题等, 可能导致语音识别速度较慢和语音识别无效等, 对中医临床诊疗带来不便, 利用特征提取的方法可以有效去除声音样本中无效的冗余信息。目前在语音识别领域梅尔倒谱系数是较为广泛应用的特征参数之一<sup>[9]</sup>, 该参数基于人耳感知的特性, 在抗噪声能力和鲁棒性方面表现优异<sup>[10]</sup>。此外研究适用于平稳和非平稳两类噪声环境的语音端点检测技术, 通过排除噪声干扰可以有效提升语音识别速度。同时在语音识别的预处理阶段采用语音增强算法, 进一步增强语音可理解程度和信噪比。对于实际诊疗中出现的名老中医方言识别问题等, 将利用深度学习算法进行优化, 对不同方言进行语言识别训练和设置容错机制。本研究还将对算法抗干扰能力进行优化以应对诊疗环境嘈杂和中医医学专业术语语音输入难等核心问题, 高效辅助医师诊疗。

### 4.2 中医语音识别平台与其他系统互联问题

目前中医语音识别平台可以实现语音的有效识别, 灵活设置停顿符号和换行。但未实现与中医院内医生开方管理系统、舌苔检查图片管理系统、药房药库管理系统等信息系统的对接融合, 导致信息分散。为了实现真正智能化, 将中医语音识别平台与信息化系统进行对接融合, 实现信息共享, 建立模板<sup>[11]</sup>, 通过语音识别与分析进行智能填写和模板导入导出, 为医生带来便利。中医临床各学科差异较大, 门类多且专业性强, 中药、方剂等术语较

多, 中医药特殊剂量、特殊诊法、中药材炮制方法等不尽相同, 还需持续深入研究探索, 对中医临床诊疗语料库不断进行更新、扩充<sup>[12]</sup>, 并与医学知识库进行深度融合。

## 5 结语

通过自然语言处理技术对中医诊疗过程中的相关信息进行详细记录, 实现高效、便捷的诊疗, 提高中医临床诊疗效率, 为患者提供详细的临床信息记录, 不仅具备即时性、准确性和可存储性, 而且可实现对医患双方的有效保护。未来中医语音识别技术和平台将越来越成熟, 更好地为中医临床提供支撑。

## 参考文献

- 1 珠海横琴博远中医药科技发展有限公司. 一种中医诊疗辅助信息管理系统: CN202010342551.1 [P]. 2020-08-25.
- 2 詹新明, 黄南山, 杨灿. 语音识别技术研究进展 [J]. 现代计算机 (专业版), 2008 (9): 43-45.
- 3 黄子君, 张亮. 语音识别技术及应用综述 [J]. 江西教育学院学报, 2010, 31 (3): 44-46.
- 4 李雨晨. 语音识别巨头 Nuance 推出 ACI 技术“无感体验”让医生摆脱临床文档负担 [EB/OL]. [2021-07-20]. <https://www.leiphone.com/category/aihealth.html>.
- 5 徐青. 语音助手兴起 [EB/OL]. [2021-07-22]. <https://www.cn-healthcare.com/article/20170927/content-495992.html>.
- 6 川陀学者. 语音识别中的 End2End 模型: CTC, RNN-T 与 LAS [EB/OL]. [2021-07-22]. <https://zhuanlan.zhihu.com/p/62836549>.
- 7 Appen. 医疗语音识别技术的应用案例 [EB/OL]. [2021-07-24]. <https://www.appen.com.cn/blog/what-you-need-to-know-speech-recognition-in-healthcare/>.
- 8 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.
- 9 毕灶荣, 董东兵, 陈巧玉. 基于快速 MFCC 计算的说话人识别系统的设计 [J]. 电子科技, 2018, 31 (4): 25-28.
- 10 周春晖, 卢荣, 潘姿蓉. 说话人识别特征参数 MFCC 的提取与分析 [J]. 电子技术与软件工程, 2016 (22): 90.
- 11 张海波, 周民伟, 刘晓辉, 等. 智能语音识别技术在医院临床的探索与应用 [J]. 中国卫生信息管理杂志, 2017, 14 (5): 660-663.
- 12 赵媛媛. 基于智能语音识别技术在医疗病历录入领域的应用探索 [J]. 科技经济导刊, 2021, 29 (14): 48-49.