

DOI: 10.13288/j.11-2166/r.2021.12.006

中医学学术客观化、智能化传承的思考及初步实现

王丽婷^{1,2}, 刘长松^{1,2}, 魏玮³, 赵强^{1,2}, 陈彦^{1,2}, 申文涛^{1,2}, 独思静³

1. 清华大学电子工程系, 北京市海淀区清华园 1 号, 100084; 2. 北京信息科学与技术国家研究中心; 3. 中国中医科学院望京医院

摘要 名老中医的学术思想和临床经验具有非常重要的学术价值, 传承名老中医经验工作势在必行。针对中医经验传承非确定因素和主观成分多、客观化和标准化程度低等问题, 从解决名老中医传承客观化、智能化、可扩展普及等实际需求出发, 以人工智能基础理论和基本方法研究为核心, 将人工智能的前沿技术融合至中医传承典型应用过程中。提出客观化人工智能传承系统的初步实现过程: 1) 在医生诊疗时, 同步采集患者脉搏波信息和舌象等原始信息以及医生给出的处方; 2) 人工智能学习中医处方, 通过端到端的深度学习方法, 得到该医生对应的人工智能模型; 3) 作为医生的人工智能助手, 该人工智能模型可提供辅助诊断治疗方案; 4) 整合信息, 建设名老中医传承病案数据库。

关键词 中医传承; 人工智能; 辅助诊疗

中医学是中华民族在长期的生产与生活实践中认识生命、维护健康、战胜疾病的宝贵经验的总结, 是中华文明的瑰宝。中医学作为中华民族原创的医学科学, 与现代医学形成互补, 能够对人类健康做出持续性的贡献。中医药的振兴发展已经上升到国家战略层面^[1], 同时也承载着中国传统文化复兴的历史使命。名老中医学术思想和临床经验的传承是中医药发展创新工作的重要环节, 中医数千年“师带徒”的传承方式^[2]在当代社会被之以院校教学方式, 在扩大教学规模的同时, 也存在一定弊端。中医客观化研究领域的蓬勃兴起, 促进了中医学与现代科学的接轨。本文探讨利用人工智能工具及技术手段, 通过对舌象、脉象这类含有丰富信息的客观原始图像和波形数据的分析, 使用机器学习算法辅助中医诊断和经验传承, 实现不依赖于人工建立的规则, 直接通过对案例样本的学习获得客观数据与诊断信息的推演关系, 以期为客观化、人工智能模式的中医传承提供思路。

1 中医学学术客观化、智能化传承的必要性

1.1 传统中医传承专家系统的不足

传统中医传承专家系统往往以问诊为主, 智能

化程度不高, 经常需要人工辨证, 使用的是一些简单特征并套用一些现有规则^[3]。存在的问题主要有中医经验传承的非确定因素和主观成分多、客观化和标准化程度低。不同患者的具体情况不会完全相同, 即使是同一个患者, 不同医生诊疗判断的结果也不同^[4-5]。中医四诊信息具有复杂、多面、量大的特点, 各信息之间的关系藏而不显, 传统的人脑处理模式和原始的数据统计方法无法对客观化中医数据进行实时的识别、提取、分析与处理; 同时中医学缺乏完整、准确、真实的信息图谱, 无法统一诊断标准、疗效标准并进行临床推广应用。此外, 专家系统往往无法体现感觉, 模仿医生效果差, 临床实用价值受限。

1.2 人工智能应用于中医学传承的必要性

当代中医药的发展, 需要尊重中医学的核心本质和历史文化, 同时应积极引入现代科学技术, 尤其是大数据、人工智能、信息技术、电磁技术等, 让中医学呈现出可解释的综合优势, 体现中医生命科学的完整体系。人工智能技术在图像识别、语音识别、自然语言处理等方面取得了极大成功, 在解决许多特定问题方面甚至超过了人^[6-9]。人工智能是帮助人类解决难题的有力工具, 如何用它来辅助中医学传承, 是目前中医药界必须要解决的问题。具体需要解决中医经验传承非确定因素和主观成分

基金项目: 国家自然科学基金(12050001); 佛山—清华产学研合作协同创新专项(佛山创新专项) 资金(2020THFS0110)

✉ 通讯作者: lcs@tsinghua.edu.cn

多、客观化和标准化程度低等困难,从满足名老中医传承客观化、智能化、可扩展普及等实际需求出发,以人工智能基础理论和基本方法研究为核心,将人工智能的前沿技术融入至中医传承典型应用过程中。

2 客观化人工智能中医传承方案的提出

我们提出一种客观化人工智能中医传承方案,其基本思想是:通过客观化健康数据获取,在医生诊疗疾病时,用仪器配合软件记录患者完整信息以及医生在诊断过程中所需要收集的信息,同步采集患者的脉搏波信息和舌象等原始信息,以及记录医生解决方案(处方)最终的药方;采用人工智能学习中医药方,通过深度学习的方法,得到该医生对应的人工智能模型,该人工智能模型作为医生的人工智能助手(全客观数据运算、无需主观判断和标注)提供辅助诊断治疗方案。

客观化人工智能中医传承方案对比传统中医传承方案有以下四点不同:1)智能化程度高,不需要人工辨证,客观地对中医数据进行实时的识别、提取、分析与智能化处理,替代传统的人脑处理模式和原始的数据统计方法;2)客观化特征提取方法,以名老中医处方中的药物为特征的中医健康指标体系和建模、标注方法;3)利用最新的人工智能和深度学习技术,学习产生概率模型,对传统的规则推理模型进行补充;4)客观记录诊断的原始数据(包括患者完整信息以及医生在诊断过程中所需要收集的信息),避免人工整理规则造成信息缺失,且随着数据积累的增多,可以反复迭代提高。

3 客观化人工智能中医传承的初步实现

客观化中医传承一直缺乏有效的技术手段。过去的传承方法往往以整理医案为主,信息记录不全面,主观描述存在歧义问题,已有的一些中医四诊仪因其功能不足、使用不便、价格过高等因素,并未被大量应用于中医传承研究中。随着信息技术、大数据、精准医学、生物信息学、人工智能等技术的蓬勃兴起,从信息与系统的角度解决上述问题,可为中医客观化传承带来新机遇。

我们在客观化人工智能中医传承方案的基础上,对中医人工智能传承完成了初步实现。中医人工智能传承除了全面、精准地记录诊断信息以外,涉及的关键技术是如何用人工智能方法训练得到医生对应的人工神经网络模型。人工智能学习需要患者端的客观化数据、医生端的诊断结果,以及一定

的样本数量。经过五年临床跟诊,我们积累了数万例病例,在中医人工智能传承领域积累了一定经验。其中,卷积神经网络(CNN)在图像识别方面取得了极大的成功,是目前最主要的深度学习模型^[1]。但我们探索使用CNN建立舌象与药方间的关系,发现简单的直接使用CNN无法得到有意义的结果,存在的问题包括数据归一化难、样本量不足、处方标签学习方法单一等。我们做出若干改进以解决上述问题,包括调整网络结构、迁移学习、数据增强、标签均衡、多标签学习等。

中医通过四诊合参来掌握患者病情资料,其中舌诊和脉诊具有独特的中医药特色,其可反映脏腑虚实、气血盛衰、病位深浅、预后好坏等,以思外揣内、见微知著、以常衡变。然而,传统的中医传承系统常缺乏舌脉信息的客观化录入,即使辅助中医四诊仪,也面临舌脉信息处理机械化和平面化的瓶颈,尤其是脉象方面。因此,针对脉象,我们提出将每个脉搏波切割下来作为图像中的一行,从上到下顺序排列每一行图像,从而把一维的脉搏波信号转化为二维图像,称为二维脉图。一般的脉诊研究只关注一个脉搏波周期^[2],实际上人体每个脉搏波周期的时间长度和波形都是变化的,这种变化与自主神经活动以及呼吸有关,二维脉图中可见明显的周期性条纹,就是呼吸频率的表现。相比于二维脉图,一维脉搏波能够表达更丰富的身体信息,与常用的CNN模型匹配度也更好。

建立中医人工智能神经网络识别系统,输入的是经过自动定位的舌象^[3]和二维脉图,输出的是根据医生的用药习惯选定的每组药物出现的概率。药物组的选定通过采集大量中医临床医师实际诊断处方进行聚类分析得到。通过采集特定中医临床医师治病时的舌象/脉图和对应的治疗处方大数据,直接学习舌象/脉图与药物组合之间的映射关系。这里隐含的假设有两点:一是处方要全面正确反映患者的状况,二是处方与舌象或脉象的某种特点要有一致和稳定的对应关系。处方具体与什么特点有关不需要人工描述,CNN通过学习算法可以自动找到该特点。也就是说,该系统要求参与学习的处方中,对于某种舌象的用药习惯始终一致。

该系统为机器学习中典型的端到端的过程,即只要给定输入与输出,中间判断完全由算法自动学习得到。端到端的学习方法能够全面利用图像反映出来的特征,能够有效避免人工选择特征不全面、不易定量、不专业、不易与治疗方案结合的弊

端。端到端的学习采用一般的深度学习基本流程,能够达到一定的准确率,但是达到实用还远远不够;同时深度学习需要海量的样本,这些样本必须覆盖每种可能的情况,且每种情况都保证有相当多的数量,才能取得比较理想的效果。

神经网络输出的每个概率值分别可以用来表示舌象加脉象的某种特点,把它们组合成特征向量,作为舌脉信息的总体描述。用特征向量表达的中医特征可以进行数学计算,例如,通过计算两两病例之间的特征距离并排序可以搜索最接近的病例,也可以计算特征向量与每个处方的距离,从而找到最接近的处方。观察每个特征值对应的舌象特点,可以了解神经网络自动学到的特征,实现可视化和一定的可解释性。当前已经有一些工具可以帮助观察神经网络内部的细节^[13],随着人工智能传承平台上数据的增加以及技术的不断进步,可解释性也将不断提高。在舌诊/脉诊联合分析问题方面,因很难得到足够多的样本,故还需要通过技术手段结合舌诊专业知识来改善系统准确度。通过与专家的沟通交流,学习专家根据舌图上的哪些特征决定了某种用药,再对识别不准确的舌象进行分析统计,找到进一步提高的方法。

4 人工智能中医传承平台的作用

人工智能中医传承平台在采集到大量的客观化数据并建立起人工智能模型的基础上,通过建设人工智能传承系统,可发挥中医服务与客观化传承两方面的作用。具体体现在:1) 案例的客观记录。包括数据可视化、提供数据指标分析结果。2) 案例辅助传承教学,对中医教育模式进行补充。平台拥有大量案例,年轻医生可以通过查阅平台上的案例自主学习,快速成长。3) 学习多位名医案例。平台是开放性的,不是一个流派的专家系统,而是集成了多位中医流派和名医的案例,博采众长,可以在平台上同时学习多位名医的案例。4) 传承与服务并举。人工智能模型可复制、服务大量人群,在传承的基础上可以进行诊疗服务;同时,每个流派每位名医一套人工智能神经网络,会形成多人工智能会诊的局面,而多人工智能会诊可以提供更丰富的治疗思路。

5 小结

本文探讨了对于舌象、脉象这类含有丰富信息的客观原始图像和波形数据,使用机器学习算法辅助中医诊断和传承的方法,这种方法的思想同样也

可以扩展到面诊、闻诊等其他测量信息的处理。目前人工智能中医传承虽然获得了一些经验,未来发展却依然难度很大。未来方案需要克服的问题有:样本少、缺乏客观测量指标;流派众多,治疗方法多样,缺乏标准;同一个病症可以用多种药物,同一个药物可治疗多种疾病,样本标注困难;用药习惯随时间演变等。

本文所涉及的方法只是对于传统望、闻、问、切中部分信息的处理方法的探索,该方法的局限性包括:对案例质量要求高、工程实现难度大、对信号采集质量要求严格等。今后需要结合望、闻、问、切中未能包含的其他信息,甚至时间、空间、环境等全面信息,才能实现最佳的中医诊断和传承。

通过机器学习和数学计算得到的结果是客观化的,但这是一个“黑箱系统”,有可能会给出无法预知的结果,因此,我们将其定位为医生辅助系统,给出可能性较大的判断供医生参考。对于学习得到的人工智能模型,其能力和适用范围需要更多的研究来确定;同时也需要用可视化的方法研究其做判决的机理,这可以帮助我们发现现有理论中忽略的信息,也可以发现其中存在的错误。

参考文献

- [1] 中共中央国务院关于促进中医药传承创新发展的意见 [N]. 人民日报, 2019-10-27(001).
- [2] 徐里. 中医传承教育中“师带徒”模式的体会 [J]. 中医药管理杂志, 2015, 23(4): 26-27.
- [3] 王瑞祥, 刘晓玉. 一种基于粗集简约的中医专家系统设计和实现 [J]. 时珍国医国药, 2013, 24(10): 2476-2477.
- [4] 李世辉. 综合集成中医专家系统研制的可能途径: 典型信息法简介 [J]. 首都师范大学学报(社会科学版), 2008, 36(S1): 97-112.
- [5] 梁建庆, 何建成. 基于数字化系统的中医问诊诊断客观化研究 [J]. 中华中医药杂志, 2014, 29(5): 1534-1538.
- [6] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.
- [7] WANG M, DENG WH. Deep face recognition: A survey [J]. Neurocomputing, 2021, 429: 215-244. doi: 10.1016/j.neucom. 2020. 10. 081.
- [8] ATMA PS, RAVINDRA N, SANTOSH K. A survey: speech recognition approaches and techniques [C] // Department of CSE Maharishi University of IT Lucknow India, Department of CSE UIET CSJM University Kanpur India. 5th IEEE Uttar Pradesh Section International Conference on

Electrical, Electronics and Computer Engineering. Gorakhpur (IN): Department of CSE Maharishi University of IT Lucknow India, Department of CSE UIET CSJM University Kanpur India, 2018; 1–4.

[9] YOUNG T, HAZARIKA D, PORIA S, et al. Recent trends in deep learning based natural language processing [J]. IEEE Comput Intell Mag, 2018, 13(3): 55–75.

[10] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [J]. NIPS, 2012, 25(2): 1097–1105.

[11] 费兆馥. 现代中医脉诊学 [M]. 北京: 人民卫生出版社, 2003: 55–76.

社, 2003: 55–76.

[12] YUAN W, LIU CS. Cascaded CNN for real-time tongue segmentation based on key points localization [C] // Institute for Artificial Intelligence Tsinghua University. IEEE International Conference on Big Data Analytics. Suzhou (CN): Institute for Artificial Intelligence Tsinghua University, 2019: 303–307.

[13] WANG ZJ, ROBERT T, OMAR S, et al. CNN explainer: learning convolutional neural networks with interactive visualization [J/OI]. (2020-08-28) [2021-01-01]. <https://arxiv.org/pdf/2004.15004v3.pdf>.

Artificial Intelligence-based Objectification of Inheritance of Traditional Chinese Medicine

WANG Liting^{1,2}, LIU Changsong^{1,2}, WEI Wei³, ZHAO Qiang^{1,2}, CHEN Yan^{1,2}, SHEN Wentao^{1,2}, DU Sijing³

1. Department of Electronic Engineering, Tsinghua University, Beijing, 100084; 2. Beijing National Research Center for Information Science and Technology; 3. Wangjing Hospital, Chinese Academy of Chinese Medical Sciences

ABSTRACT Considering the important value of the academic thoughts and clinical experience of the famous traditional Chinese medicine (TCM) doctors, it is imperative to inherit and pass on this treasure in good ways. The problems in the inheritance of TCM experience are found to be having many uncertain factors and subjective components, as well as low degree of objectification and standardization. To meet the actual needs of objectification, intelligentization and generalization of inheriting TCM experience, the cutting-edge technology of artificial intelligence (AI) is integrated into the typical application of TCM inheritance by taking the basic theories and methods of AI as the core. The AI-based objectification process of TCM experience inheritance is preliminarily designed as follows: First, through the objectification of the health data, the original data of the patients' pulse and tongue manifestations as well as the final prescription can be synchronously collected during the medical visits; Second, AI learning of TCM prescriptions adopts the end-to-end deep learning method to obtain the AI model of the doctors; Third, as an facilitator, this AI model can provide auxiliary diagnosis and treatment schemes; Fourth, through information integration, the medical records database of the famous TCM doctors is constructed.

Keywords inheritance of traditional Chinese medicine; artificial intelligence; auxiliary diagnosis and treatment

(收稿日期: 2020-03-04; 修回日期: 2021-03-19)

编辑: 柴倩云