



Views & Comments

利用证候要素架构中医人工智能诊疗模型

王阶, 段练, 李洪峥, 刘金垒, 陈恒文

Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing 100053, China

1 引言

中西医结合在慢性病和疑难病的治疗方面具有优势。在此次新冠疫情中,湖北省确诊病例的中医药使用率和治疗总有效率均超过了90%,中西医结合的疗效得到了肯定。但中医诊疗思维的传承与中西医结合的未来发展仍面临着严峻挑战。从20世纪50年代起,人们试图运用信息化技术进行中医四诊的客观化探索;70年代起,国内外尝试研发具备专家临床诊疗思维的中医专家系统,但始终无法较好地模拟中医师辨证论治过程;90年代以来,随着人工智能(artificial intelligence, AI)技术的不断突破,中医专家系统开始运用神经网络模糊逻辑、关系数据库等技术方法,挖掘中医大数据,并进行客观化、规范化、量化研究。但是完全基于数据驱动的方式学习模型参数,难以将中医师长期临证总结形成的经验和知识融入模型,故而难以对模型学习过程施加有效约束。

近年来,善于处理非线性问题的人工神经网络及善于处理高维数据的决策树和随机森林算法都被应用于中医模型构建[1-2]。但不同于现代医学具有近乎统一的明确的诊疗规则,多元化的传统中医诊疗理论之间存在交叉,不同中医学派及不同中医师的诊疗规则也存在交叉,导致难以储备相同规则体系的高质量中医大数据。面对小样本训练甚至零训练样本等真实现状,可以考虑利用规则化可解释模型完成病证结合下的知识+AI融合。

2. 利用证候要素可统一不同诊疗规则

证候是中医学辨证论治理论体系构建与临床应用的核心,是通过病因、病位、病性对人体疾病病理、生理变化的整体概括[3],是“辨证”(总结归纳症状群特点)的结果和“论治”(决定治疗原则与具体治疗方法)的依据[4],具有多层次性、复杂性、抽象性、高维度性的特点[5]。但是每种辨证纲领的使用范围存在区分界限模糊,应用场景重叠度高,辨证内容交叉等问题[6]。这导致基于中医专家系统开具处方的准确性不足,给诊断规范化及中医智能化研究带来很大困难。

有鉴于此,在病证结合的人工智能研究中,证候要素可作为基石联系诊疗过程中症状、治法、中药、方剂等内在要素。任何复杂的证候,都是由病位、病性、邪正关系等具体维度交叉确定,每一维度的具体表现即可认为是该辨证维度的要素[7]。研究发现,通过降维升阶的方式,可分离出60项左右的证候基本辨证要素,对基本辨证要素进行排列组合,可覆盖中医证候类型[8]。就像字母是构成单词的最小单元一样,中医领域中的证候要素,是组成证候的最小单元。不同于证候的复杂、高维,每一证候要素均具备具体的症状群,也具备不同于其他要素的特异性,通过组合、叠加,可形成临床所见的证候,即证候要素具有低维度、可叠加、可组合等特点。而每一证候要素均有其对应的治法、中药或药物组合,明确了证候要素及证候要素组合形态,即可相应确定治法、方药[9]。

以冠心病为例，对证候要素为核心的中医诊疗流程进行分析。冠心病患者的所有临床表现均可归于血瘀、气虚、痰浊、阴虚、气滞、阳虚、寒凝、热蕴八大证候要素[10]，通过叠加、组合可成为气虚血瘀、气阴两虚、痰瘀互结、气滞血瘀、痰阻热蕴和阳虚寒凝等证候类型[11]。根据证候要素与治法的对应关系，可知冠心病的中医治疗方法可包含活血、益气、化痰、滋阴、理气、温阳、散寒、清热等；根据治法可获得对应的方剂、药对或单味中药，如冠心2号方等方剂，党参-黄芪、葶苈-高良姜等药对，以及丹参、三七等单味中药；再根据证候要素组合形态进行药物组合，可对应血府逐瘀汤、瓜蒌薤白半夏汤等冠心病治疗的代表方药[12]。由此可知，证候要素的叠加组合是药物方剂变化加减的依据和关键，可整合统一不同来源的中医诊疗规则。

3. 利用算法集成适合小样本中医数据模型

算法是人工智能领域的重要技术，是研究中医人工智能系统的重要路径。在中医人工智能算法的选择上，我们应该谨慎考虑其适用性。对于四诊证候信息的识别[13-14]，考虑病例数量，可以选择贝叶斯网络和支持向量机，因为其在小规模数据上表现良好，对数据噪声较为鲁棒。但该算法应用于真实世界的大数据量样本中则计算难度增大故而影响实用性。

中医证候体系的复杂性决定其研究应从非线性设计着手，人工神经网络以其优秀的非线性问题解析能力，可以模拟人脑神经网络的结构和功能，对数据进行有效处理，识别出数据中隐藏的复杂规律，对确定中医证候问题具有优势[15]。但神经网络算法对数据的标准化水平有较高的要求，并且训练需要数据规模较大，而目前中医数据的标准化水平不足，导致缺乏大量的训练数据，这是制约此算法发挥作用的重要因素。

中医诊疗过程的完全模拟和多维、复杂数据的处理一直是中医人工智能研究的难题。决策树和随机森林算法都善于处理高维数据，探究数据的特征之间的相互作用关系，符合数据的要求。但决策树容易过拟合，可导致准确率下降。随机森林算法是用随机的方式建立多个决策树，擅长处理不平衡数据。运用随机森林算法对慢性疲劳中医证候要素建立预测模型的研究[16]，已取得较高的预测准确率。但其可控性不足，模型对于小规模数据不能较好的分类，应用于中医领域同样存在数据来源受限的问题。

尽管人工智能已经在中医学领域取得了一些有价值的成果，但主要存在两个方面的问题。一是标准化、客观化

数据缺乏的问题。人工智能计算需要大量的标准化、客观化数据，但受限于中医行业的特殊性，收集的数据流派化、主观化特征显著，使得中医客观化进程缓慢，导致上述人工神经网络等人工智能新算法缺乏大量数据支撑，难以取得突破。二是已有研究的应用场景单一，仅尝试应用一个算法虽可实现基于某种特征的数据的预测模型构建，但通用性与可迁移性普遍不足，无法形成多病种通用的、实现症状-证候-治法-方药完整诊疗流程的系统化研究成果。因此，依据中医人工智能研究的现状，在解决数据标准化问题的基础上，将多种算法结合，实现症状到证候的准确计算，从而实现症状-证候-治法-方药这一全诊疗流程的中医人工智能系统，是当前迫切需要解决的问题。

4. 中西医结合人工智能的研究策略和方向

在新一代人工智能算法的基础上，我们提出基于证候要素探索规则化中医人工智能研究模型的思路：以中医人工智能为研究对象，以证候要素为切入点，以规则集成为路径，以行业诊疗指南赋予的症状积分为权重，通过症状组合、疾病识别、证候推导、治法治则、方剂生成、加减药味、药量赋值等过程，形成推荐处方（图1）。此外，利用小样本、高质量中医诊疗数据训练模型以修正证候计算的准确性。还可利用反馈机制，允许中医师在输出结果的基础上进行调整，并将这种调整作为新规则加入规则数据集。

其次借助中医人工智能模型可学习性与拓展性，考虑加入证候、证候要素、中药、方剂、临床病例、古典医籍与医案等内容，使不同来源的中医诊疗知识体系完成知识体系融合、建立症状-证候要素-证候-治法-方剂-中药对应关系，提高辨证指向的准确性。在具体实现方面，可以考虑以症状和证候要素作为节点，以症状-证候要素关联性作为边界，构建诊疗规则的知识图谱构建。使用单一疾病行业诊疗指南赋予的症状积分作为权重，通过构建邻接矩阵、度矩阵计算出拉普拉斯矩阵（Laplacian matrix），来表示不同症状在证候要素下的权重，进而使用卷积操作加权求和输出结果，从而建立可视化模型，在此基础上实现规则的可视化显示。

在上述设计思路基础上，致力于形成符合行业标准、经典古籍、专家意见的中医人工智能系统（图1）。为标准化、客观化的中医人工智能诊疗系统研究提供新方向。

5. 展望

我们以冠心病为单病种中西医结合+AI的研究突破

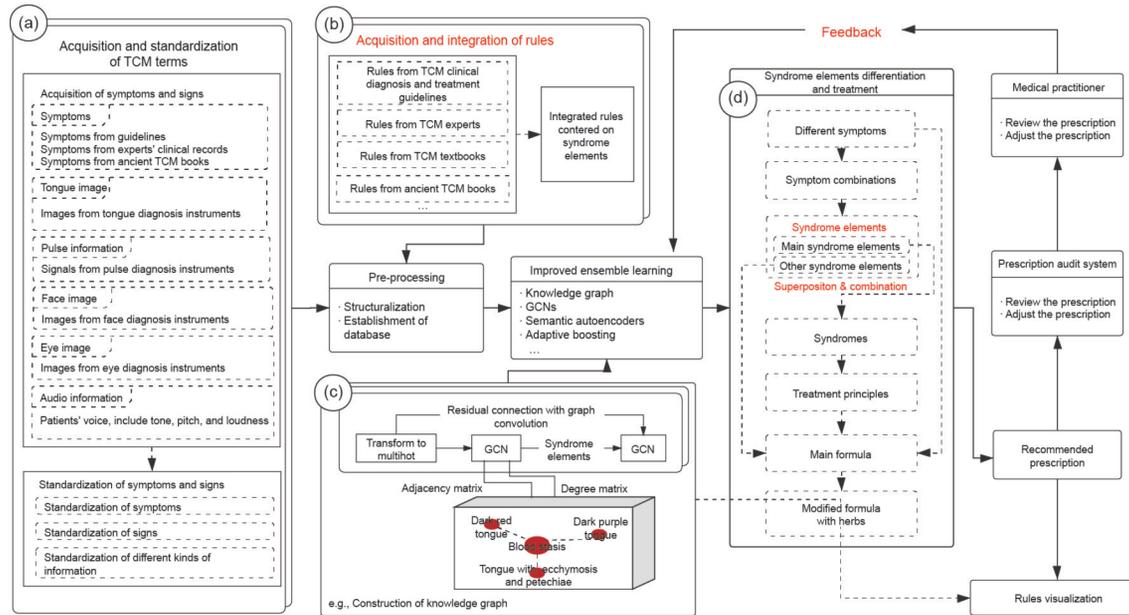


图1 单病种中西医结合人工智能模型设计路线。(a) 中医术语的获取和标准化。通过获取不同来源的症状、体征术语信息，采集患者的舌象、脉象、面象、目诊图像、声音等信息，并将以上信息内容进行数据标注，进而完成数据标准化工作。(b) 中医诊疗规则获取。中医诊疗规则可来自于中医临床诊疗指南、中医专家经验、中医学教材、古代中医典籍等，通过证候要素关联并集成以上不同来源的诊疗规则，从而形成以证候要素为核心的集成诊疗规则。(c) 以图卷积神经网络和知识图谱的构建为例，说明集成算法规则，形成可视化模型方案。构建中医人工智能知识图谱，以证候要素和症状作为节点，以症状-证候要素关联性作为边界，以疾病行业诊疗指南赋予的症状积分为权重基础，表示不同症状在证候要素下的权重，进而使用卷积操作加权求和输出结果，并支持可视化展示。(d) 集成各来源的诊疗规则形成辨别证候要素输出处方的过程。以标准化症状为输入层，首先完成标准化症状对应证候要素的判断，再经由证候要素的叠加组合形成证候，通过证候-治法的对应关系，输出主要证候要素组合对应的方剂、中药；除却主要证候要素外，再由其他证候要素及症状，输出对应的中药，形成针对主要方剂的药物加减处方作为输出层。(e) 基于知识图谱和处方推荐结果实现规则可视化展示。

口，引入人工先验知识，通过临床症状及舌脉等信息分析其可能的证候要素，再以证候要素为核心获得可靠方药，由此形成了基于证候要素的中医人工智能诊疗模型设计，利用小样本诊疗数据对此模型进行训练，可实现较“中医大数据”形成的诊疗模型更高的AI处方精度。

我们预计中医诊疗系统会实现由单病种诊疗模型到多病种诊疗模型的发展，由纯中医诊疗系统到真正的中西医结合诊疗系统的跨越。我们希望不久的将来它可以简化就诊流程，提高诊疗效率，让患者便捷地接受基于稀缺的高水平中医专家诊疗经验的人工智能诊疗服务。即使距离医院较远，也可以在家就收到最优质的道地药材，获得最优质的医疗服务。

致谢

本研究由岐黄工程首席科学家项目和中医药管理局循证中医药项目支持（60103）。

References

[1] Zhang F, Wu C, Jia C, Gao K, Wang J, Zhao H, et al. Artificial intelligence

- based discovery of the association between depression and chronic fatigue syndrome. *J Affect Disorders* 2019;250:380–90.
- [2] Wang Z, Li L, Song M, Yan J, Shi J, Yao Y. Evaluating the traditional Chinese medicine (TCM) officially recommended in China for COVID-19 using ontology-based side-effect prediction framework (OSPF) and deep learning. *J Ethnopharmacol* 2021;272:113957.
- [3] Sun XL, Jiang WW, Yu DL, Lin X, Ding BG, Liu MA. Research on basic path and inherent law in the process of Chinese medicine treatment based on syndrome differentiation. *J Tradit Chin Med* 2016;57(4):289–94. Chinese.
- [4] Wang YY. Suggestions on improving the system of traditional Chinese medicine syndrome differentiation. *J Tradit Chin Med* 2004;45(10):729–31. Chinese.
- [5] Li S. Discussion on the characteristics of TCM syndromes from the concepts of dimension and ranks: approaches to the standardization of TCM syndromes. *J Beijing University Tradit Chin Med* 2003;26(3):1–4. Chinese.
- [6] Zhang ZB, Wang YY. Establishment of a new TCM syndrome differentiation system. *J Beijing University Tradit Chin Med* 2005;28(1):1–3. Chinese.
- [7] Zhang ZB, Wang YY. Research on TCM syndrome nomenclature and classification: review and hypothesis. *J Beijing University Tradit Chin Med* 2003;26(2):1–5. Chinese.
- [8] Zhu WF. Standardization research of differentiation system of symptoms and signs and ‘‘syndrome’’ in TCM. *Tianjin J Tradit Chin Med* 2002;19(5):1–4. Chinese.
- [9] Wang J, Tang YL, He QY, Xiong XJ. Thinking of prescriptions corresponding to syndromes in construction of syndrome differentiation system. *J Tradit Chin Med* 2009;24(7):837–9. Chinese.
- [10] Wang J, Li J, Yao KW, Zhong JB. Study on syndrome elements and their combination laws in patients with angina pectoris. *J Tradit Chin Med* 2007;48(10):920–2. Chinese.
- [11] Li J, Wang J. Literature case analysis of 5099 cases of syndrome elements and syndrome combinations of angina pectoris of coronary heart disease. *Chin J Basic Med Tradit Chin Med* 2007;13(12):926–7. Chinese.
- [12] China Association of Chinese Medicine. Guidelines for diagnosis and

- treatment of stable angina pectoris of coronary heart disease (T/CACM1325–2019). *J Tradit Chin Med* 2019;60(21):1880–90. Chinese.
- [13] Zhang H, Ni W, Li J, Zhang J. Artificial intelligence-based traditional Chinese medicine assistive diagnostic system: validation study. *JMIR Med Inform* 2020;8(6):e17608.
- [14] Xu F, Xu ZX, Xu WJ, Wang YW, Liu T, Xia CM, et al. Classification of TCM syndromes in 835 cases of coronary heart disease: on the basis of Bayesian networks principle. *Shanghai J Tradit Chin Med* 2014;48(1):10–3. Chinese. [15] ZhangNL, YuanS, ChenT, WangY. Statistical validation of traditional Chinese medicine theories. *J Altern Complem Med* 2008;14(5):583–7.
- [16] Hong YZ, Zhou CL, Zhang ZF, Xu JT. Selection of characteristic symptoms of chronic fatigue syndrome elements based on random forest method. *J Tradit Chin Med* 2010;7:634–8. Chinese.