DOI：10.12287/j.issn.2096-8965.20220310

人工智能在中医药的研究现状及展望

李心怡[#1]，罗思言[#1]，徐常胜[2]，丁贵广[3]，张勇东[4]，田贵华[1]*

（1. 北京中医药大学东直门医院，北京 100700；
2. 中国科学院自动化研究所，北京 100864；
3. 清华大学软件学院，北京 100084；
4. 中国科学技术大学信息学院，安徽 合肥 230026）

【摘要】中医药在世界范围内难以被广泛接受的主要问题是缺少足够客观定量的数据支撑和完备而自洽的理论体系。此外，传统中医术语的模糊性、理论知识的难以理解性、治疗思维的抽象性、中医医案的繁杂性也对中医现代化发展带来了极大的挑战。随着信息技术的日新月异，大数据和人工智能技术为规范中医诊疗数据、构建智能中医诊疗体系以及突破传统中医诊疗模式提供了新方法，进一步推动了中医药智能化的发展。经过半个多世纪的发展，中医与人工智能技术的融合逐步深入，取得了一定的应用成果，如中医药中医信息化数据库、中医四诊采集设备、中医辅助诊疗系统以及智慧中医健康管理等。但就目前而言，中医药智能化发展中仍存在数据标准欠缺、相关制度不够完善、交叉人才匮乏等问题，未来还需进一步建立规范的数据标准，完善数据共享、知识产权、伦理规范等法律法规，加速培养学科交叉复合型人才，创新思维革新医疗模式等，促进人工智能背景下的中医药创新发展。本文从知识发现与机器学习的角度，对人工智能在中医药的研究进展概括总结，以期为中医药智能化提供助力。

【关键词】人工智能；中医；机器学习；知识发现

中图分类号：TP18，R2　　**文献标识码**：A　　**文章编号**：2096-8965（2022）03-0069-05

Current status and future prospects of artificial intelligence in traditional Chinese medicine

Li Xinyi[#1], Luo Siyan[#1], Xu Changsheng[2], Ding Guiguang[3], Zhang Yongdong[4], Tian Guihua[1]*

(1. Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing 100700, China;
2. Institute of Automation of the Chinese Academy of Sciences, Beijing 100864, China;
3. School of Software, Tsinghua University, Beijing 100084, China;
4. School of Information, University of Science and Technology of China, Hefei 230026, Anhui, China)

【**Abstract**】Traditional Chinese Medicine (TCM) has not been widely accepted in the world for the lack of objective data and complete theoretical system. In addition, the ambiguous of language, the difficulty in understanding theoretical knowledge, the abstractness of therapeutic mind, and the complexity of medical records have also brought great challenges to the development of TCM modernization. With the rapid development of information technology, big data and Artificial Intelligence (AI) technology have provided new methods for

收稿日期：2022-09-01；**修回日期**：2022-09-06

基金项目：国家中医药管理局青年岐黄学者支持项目（国中医药人教发〔2020〕7号）

通讯作者：田贵华（1983- ），山东人，研究员，主要从事中医药防治慢性疾病的智能诊疗、疗效评价和效应机制研究。

E-mail：rosetgh@163.com

#为共同第一作者

standardizing TCM diagnosis and treatment data, building an intelligent TCM diagnosis and treatment system and breaking through TCM diagnosis and treatment model, further promoting the development of TCM intellectualization. After more than half a century of development, the integration of TCM and AI technology has gradually deepened and certain application results have been achieved, such as the information database of TCM, four diagnosis collection equipment of TCM, the clinical auxiliary diagnosis and treatment system of TCM and intelligent health management of TCM, etc. However, for now, there are still some problems in the intellectualized development of TCM, such as the lack of data standards, the insufficient improvement of related systems and the lack of talents with interdisciplinary backgrounds, etc. The focus of future work is to establish normative data standards, improve relevant laws and regulations on data sharing, intellectual property rights, and cord of ethics, so as to promote the innovative development of TCM in the AI background. The study summarizes the progress of AI in TCM from the perspective of knowledge discovery and machine learning, in order to provide assistance for the intellectualization of TCM.

【**Keywords**】Artificial Intelligence; Traditional Chinese medicine; Machine learning; Knowledge discovery

前言

人工智能（Artificial Intelligence，AI）特指一类通过仿生学、神经学、统计学等原理模仿人类感知及行动的行为体[1]，人工智能从20世纪70年代初开始应用于医学领域[2]，用于辅助临床诊断及规划治疗方案。近年来，由于并行计算软硬件的革命性进步及大数据时代的到来，医学人工智能迎来了飞速发展[3]，中医与人工智能的有机结合也迎来了新的契机[4]。中医坚持“以人为本，重视预防”，在传承发展中医特色的同时，充分结合现代科技技术，发挥中医自身优势，从而在做到“整体观念”“辨证论治”的基础上同时推进中医客观化、标准化、准确化，具体应用包括：自然语言处理技术被运用于古籍及医案的数据挖掘当中，计算机视觉技术被运用于舌诊等领域，小样本机器学习被运用在辨证论治等领域，无监督聚类算法被运用在药理分析等领域。

知识发现是一种广义的“数据挖掘”说法，指从大量结构化或非结构化数据中提取出有实用价值模式的非平凡过程。相较于一般的统计学分析，知识发现具备一定智能性和自主行为能力。机器学习是一类算法的总称，目的是通过先验知识形成训练集进行学习，具有预测相同类型数据的能力。并且，当训练集数据量增加，机器学习可自适性提高预测准确性，而不仅仅是依赖于固定的方程模型。方药作为中医疗法的重要组成部分，在中医现代化领域占据重要部分。通过对海量中医数据库和相关文献的学习，知识发现可以挖掘中医数据库、书籍、文献之中的潜在模式，对药与药、药与方和方与证之间进行多维度的关联分析[5]。机器学习则可形成具有中医思维的人工智能模型，既能探索方药组成规律，又能在不知道组成规律的情况下对方药进行预测分类[6]。知识发现和机器学习共同应用有利于结构化中医药数据库的建立，实现中医药规范化。本文立足于方药，从方药药理分析和方-证关联维度具体分析人工智能应用于中医方药的研究进展，以在大数据时代的背景下，为中医智能化与现代化提供参考依据。

1 人工智能在中医学的应用现状

人工智能技术为中医药现代化发展提供了有力的技术支持，通过对CNKI数据库发表的中医人工智能相关文献进行检索，利用vosviewer软件（https://www.vosviewer.com/）构建了智能中医关键词聚类图和关键词时间叠加网络图，可视化重现中医药与人工智能结合的应用现状（见图1）。研究发现，智能中医关键词聚类主要围绕“人工神经网络”“人工智能”“数据挖掘”“深度学习”和“机器人”等五大聚类。在智能中医应用领域，“人工神经网络”与“模式识别”等为初期应用的常用方法，“深度学习”“机器学习”“卷积神经网络”“网络药理学”“自然语言处理”和“大数据”等为目前热点方法。此外，从智能中医药研究内容来看，早期多涉及对“中药的质量”“中药成分”及“专家系统”等研究，而“名老中医经验”“中医诊断”“用药规律”“健康管理”与“中医诊疗设备”等中医临床智能诊疗相关研究逐渐成为目前的研究热点。

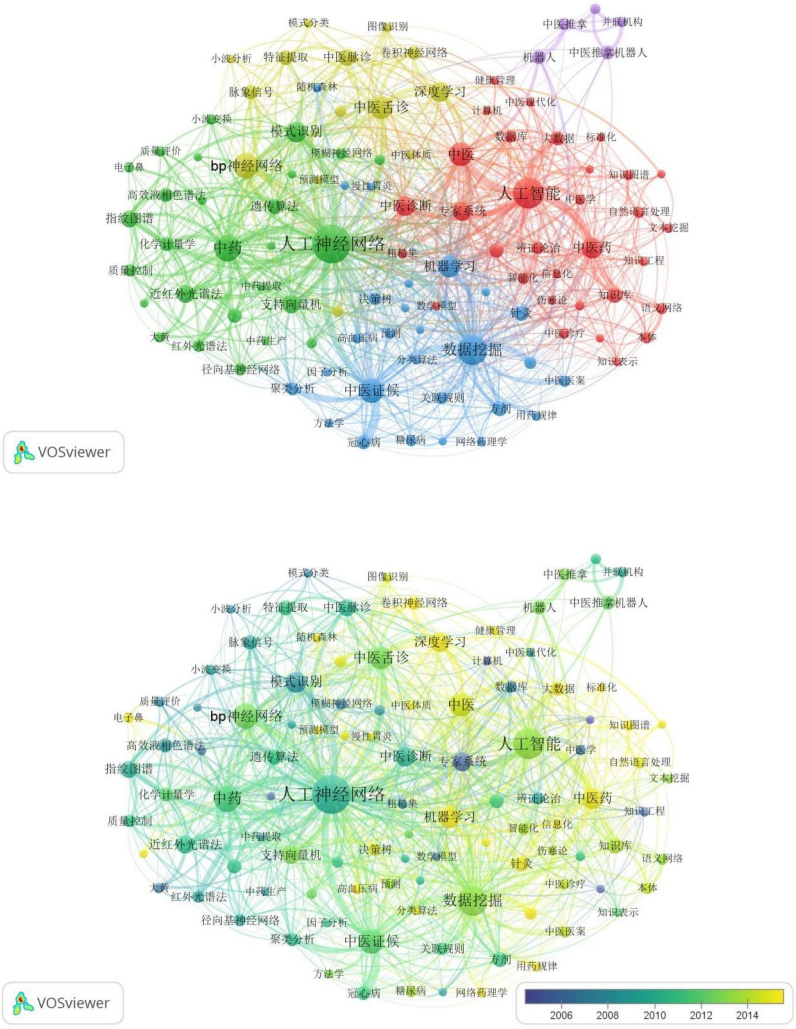

图1　人工智能在中医领域的关键词聚类分析图和时间网络叠加图

2　人工智能在中医药中的运用

2.1　AI与中医方药药理分析

中药分为四气五味、升降浮沉和归经。四气即寒、热、温、凉四种不同药性，又称为四性，是由于药物作用于人体产生的不同反应和疗效归纳总结而来[7]。五味即辛、甘、酸、苦、咸五种药味，既源于中药本身的味道，也与中药本身作用有关[8]。升降浮沉是与疾病的相对趋向性而言，与中医基础理论有关[9]。归经则是指中药对于作用部位有选择作用，以脏腑经络为基础，以药物疗效为依据[10]。根据中国古代哲学和长期经验，各类的中药进行配伍联合使用，符合“君臣佐使”规律，增强疗效、减少毒副作用、兼顾次症，被称为方剂。近年来开始进行对照组验证及有效成分探究[11-13]，这从一定程度上解释了中药和方剂的有效性。但是方剂组成的基本规律，即“君臣佐使”，具有盲目性、模糊性、主观性、片面性[14]的特点。从而导致了方剂的随意性和缺乏普适性，进而造成了科学研究方剂的困扰[15]。

而AI的介入为解决中医药现代化和科学化问题提供了新的途径。目前急待处理的数据大致可分为两类，一类为中医几千年传承所得的文献与临床经验[16]，其中蕴藏着大量基于长期临床实践得来的经验性知识。该类数据庞大的规模及缺乏统一标准的用词很大程度上制约了它的充分利用。另一类为近年来为促使中医科学化与中西医结合所做的临床研究和数据整理，此类数据本身经过处理，但缺乏统一标准并极少公开所有数据[17, 18]。相较于人工对上述数据进行学习与整理，知识发现与机器学习技术能更好地汇总两类数据、加速中医数据库融合并减少人力资源的投入。其优越的运算性能够在少量

时间内推算出有参考性的结论，包括中药组方的核心探究、药对使用规律、方剂使用原则、方剂实施的安全性和规范化、中药方剂使用经验和传承等。

基于人工智能的知识发现可应用于探索中医药中，方剂错综复杂的组合规律，包括安全性评价、中药对机制探索、中草药的核心药材辨识、方剂中的草药配比和处方验证等。关联规则挖掘、支持向量机（Support Vector Machine, SVM）、贝叶斯网络（Bayesian Network, BN）、组合规则、马尔可夫决策过程（Markov Decision Process, MDP）、文本聚类分析和文献计量学在该领域应用广泛。Yao等[5]以3 090例中医临床病例数据及中医先验知识为基础，建立了可以有效提出相应治疗方案的知识发现系统，并进一步对中药的机制进行了探索。Yang等[19]则将偏头痛处方作为模型训练集，以无监督分层模式发现算法学习规则关联，提出了新的偏头痛处方，并利用川芎为监督学习信号进一步对处方进行优化。Sheng等[20]则更进一步地建立了证、病、药联合分析的条件概率模型，通过挖掘三者之间的潜在模式为临床应用提供了依据。

机器学习中的分类及聚类算法也是探索方剂组合规律方面的关键，包括概率神经网络（Probabilistic Neural Network, PNN）、K-近邻算法（K-Nearest Neighbor, KNN）等。Ung等[6]在不知道配对规律的前提下，根据已知中药对所属药性，利用机器学习算法对随机中药组合和中药对进行预测和区分，其准确率在90%以上。姜皓等[21]则将神经网络、SVM、朴素贝叶斯和随机森林四种机器学习算法应用在了妊娠期中药“禁忌慎”判别上，并取得了最高0.885的F1分数。Li等[22]则将代谢组学与SVM算法相结合，建立了一种快捷、灵敏、特异性强的中药配伍心脏毒性预测方法，并将此方法用于验证附子配伍的心脏毒性。此外，深度学习与文本挖掘技术也被用于进行中药方剂分类与表征学习当中[23, 24]。

2.2 AI与中医方-证关联

正如前言所说，中医历来有“同病异治，异病同治”的观念。具体指病相同而证不同时需用不同方剂进行治疗，而病不同证却相同时可用相同方剂进行治疗。这不仅仅说明了辨证论治的重要性，也说明了方和证是存在关联的，如果希望达到良好的临床疗效，必须在辨证的基础上，给予对应的方剂处方。方与证的内在联系立足于方剂临床疗效总结，具体受到中药本身药性和中药间配伍影响。目前来说，方-证关联的主要知识来源为古文献的传承学习[25]和近代医家在临床使用中通过加减化裁，形成自己的经验方并加以验证后确有疗效[26, 27]。

由于方-证之间的关系纷繁复杂，相较于传统基于经验的方法，现代数据挖掘、机器学习技术将海量古籍以及临床数据库统一为结构化数据库，并从统计学角度对元知识之间的潜在模式进行挖掘，从中获取到的客观化、定量化的方-证关联知识可为其他中医师临床治疗同一证时提供参考依据。自动编码器、MDP、*K*-均值聚类算法、SVM、BN、异构因子图模型（Heterogeneous Factor Graph Model, HFGM）等数据挖掘及机器学习算法被广泛用于该领域。如Zhang等[28]利用基于证型的自动编码器模型将临床病例自动分类，并分析了典型方剂对不同类别中高血压各症状的疗效；Meng等[29]基于粒级属性偏序结构图（Attribute Partial Ordered Structure Diagram, APOSD），通过对中医临床数据库中脾虚证治疗方剂的分析，提取出对脾虚证治疗具有特定效果的药材，并可在APOSD中进行数据关联性可视化。知识发现理论也可根据患者特定证型对方剂进行个性化的调整，以获得更好的临床疗效，如Zhao等[30]根据115位中风患者的生理指标与中医师对其开具的处方为训练样本，基于流形学习理论建立了中成药治疗脑卒中的辅助决策模型。

3 结语

中医药在世界范围内难以被广泛接受，主要问题是缺少足够客观定量的数据支撑和完备而自洽的理论体系。与此同时，中医语言的模糊性、理论知识的难以理解性、治疗思维的抽象性、中医医案的繁杂性阻碍着中医药的客观化与规范化。然而，AI的出现弥补了中医药与现代科技的裂痕，基于统计学原理的数据挖掘与知识发现方法将中医药内部纷繁复杂的联系转化为不同变量在隐层空间的数学关系；基于自然语言处理与文本挖掘的AI技术将海量中医文献和繁杂的科研成果，规范了中医用词模糊性和主观性，推进了中医药理论体系标准化、理解化；聚类算法与无监督学习理论则将方剂、证型分为多个具有隐式内部关联的簇。上述方法所建立的具备中医思维与大数据知识的模型将成为连接AI和中医的重要桥梁，并为世界接受中医药体系的传承与创新、中医循证医学的发展做出巨大贡

献，而AI的加入对于医疗数据库建立、中医理论体系规范化、临床诊断精确性、教学传承标准性、科研的科学性都大有裨益。

随着人工智能技术的成熟，人工智能技术与中医的交叉融合为突破中医药现代化的发展瓶颈、探索中医药科技创新发展提供了可能。2021年，国内首部智能中医学专著《智能中医学概论》[31]的发布，在行业内反响巨大，书中首次明确了“智能中医学”的概念，并对其科学内涵进行了全面的阐释，进一步促进了智能中医学学科的形成和发展。目前，中医数字化的统一标准尚无定论，仍需各个研究之间达成共识。总之想要达到畅想的未来，还需不同国家AI领域和中医领域研究人员进一步加强合作，共同建立起全球接受的AI辅助中医理论体系。

参 考 文 献

[1] FRANK P. Introduction to knowledge systems[J]. Artificial Intelligence in Medicine, 1997, 9(2): 201-203.

[2] MARIO S. The socio-organizational age of artificial intelligence in medicine[J]. Artificial Intelligence in Medicine, 2001, 23(1): 25-47.

[3] 李明珠, 李静然, 储成顶, 等. 人工智能在我国医学科技领域的研究现状及趋势分析[J]. 中国医学装备, 2022, 19(2): 176-180.

[4] 田赛男, 刘琦, 夏帅帅, 等. 人工智能技术在中医药领域中的应用与思考[J]. 时珍国医国药, 2021, 32(11): 2740-2742.

[5] YAO L, ZHANG Y, WEI B, et al. Discovering treatment pattern in traditional Chinese medicine clinical cases by exploiting supervised topic model and domain knowledge [J]. Journal of Biomedical Informatics, 2015, 58: 260-267.

[6] UNG C Y, LI H, CAO Z W, et al. Are herb-pairs of traditional Chinese medicine distinguishable from others? Pattern analysis and artificial intelligence classification study of traditionally defined herbal properties[J]. Journal of Ethnopharmacology, 2007, 111(2): 371-377.

[7] 刘晓燕, 崔亚东, 田合禄. 中医四气五味理论与脏腑补泻关系的探讨[J]. 世界中医药, 2021, 16(1): 121-124, 129.

[8] 陈铜锁, 何永生. 四气五味制方原理初探[J]. 湖南中医杂志, 2019, 35(5): 112-114.

[9] 陈龙. 浅析中药升降浮沉药性理论[J]. 广西中医药, 2020, 43(2): 44-47.

[10] 郜贺, 庞敏, 于睿, 等. 白长川教授谈归经药与引经方[J]. 中华中医药学刊, 2022, 40(1): 66-69.

[11] 王菁斐. 白虎加桂枝汤治疗类风湿关节炎湿热痹阻证的临床疗效及可行性分析[J]. 中国现代药物应用, 2022, 16(2): 204-206.

[12] 侯明. 慢性荨麻疹患者联合应用桂枝汤与针刺放血治疗的临床疗效及对免疫指标的影响[J]. 黑龙江医药, 2021, 34(6): 1388-1389.

[13] 王春茜, 高旅, 吴茵, 等. 基于内源性代谢产物解析桂枝汤解热作用的分子机制[J]. 中华中医药杂志, 2021, 36(9): 5169-5175.

[14] 王付. 重新厘定从君臣佐使角度研究方剂的理论意义及临床价值[J]. 中医药通报, 2020, 19(3): 25-27.

[15] 龚轩.《方剂学》本科教学中“君臣佐使”的处理[J]. 教育现代化, 2018, 5(35): 205-207.

[16] 荣士琪, 杨朝晖. 中医古籍数据库数字化建设发展综述[J]. 黑龙江科技信息, 2015, (11): 174-175.

[17] 翁晓兰, 申力, 贾佳, 等. 中医医案数据库的构建及其应用研究[J]. 浙江中医杂志, 2016, 51(8): 613-614.

[18] 于琦, 李敬华, 高宏杰, 等. 中医养生数据库建设[J]. 中国数字医学, 2015, 10(8): 73-75.

[19] YANG H, CHEN J, TANG S, et al. New drug r&d of traditional Chinese medicine: role of data mining approaches[J]. Journal of Biological Systems, 2009, 17(3): 329-347.

[20] SHENG W, HUANG E W, ZHANG R, et al. A conditional probabilistic model for joint analysis of symptoms, diseases, and herbs in traditional Chinese medicine patient records[C]: 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2016.

[21] 姜皓, 张冰, 张晓朦, 等. 基于4种机器学习算法的妊娠期中药“禁忌慎”判别[J]. 中草药, 2021, 52(24): 7596-7605.

[22] LI Y, ZHOU H, XIE J, et al. A novel method for evaluating the cardiotoxicity of traditional Chinese medicine compatibility by using support vector machine model combined with metabonomics[J]. Evidence-Based Complementary and Alternative Medicine, 2016(2016): 6012761.

[23] CHENG N, CHEN Y, GAO W, et al. An improved deep learning model: s-textBL CNN for traditional Chinese medicine formula classification[J]. Front Genet, 2021, 12: 807825.

[24] WEI L, ZHENG Y. Distributed representation for traditional Chinese medicine herb via deep learning models[J]. 2017.

（下转第92页）