

深度学习技术在中药质量控制中的应用与探讨

汪舒琦^{1,2}, 申蕾^{1,2}, 宋立亚^{1,2}, 耿朴玉^{1,2}, 邱宁宁^{1,2}, 刘文龙^{1,2*}, 张喜利^{1,2*}

1. 湖南中医药大学药学院, 湖南长沙 410208

2. 中药成药性与制剂制备湖南省重点实验室, 湖南长沙 410208

摘要: 中药质量控制方法的科学性是制约中药全球高质量发展发展的关键因素, 同时在现有中药质量标准控制下的中药质量优劣也是中医药治疗疾病效果的关键因素。在人工智能和大数据快速发展的背景下, 深度学习在中药质量控制研究中获得了广泛关注, 相较于传统方法其展现出明显优势, 并在中药外观识别、成分分析、安全性评估及工艺质量控制等方面取得了显著成果。从深度学习技术的基本框架和常见模型出发, 系统梳理了深度学习在中药质量控制中的应用模式, 深入分析了其在中药质量检验中面临的挑战, 并基于已有研究成果提出了针对性的解决措施。总结了深度学习在中药质量控制领域的应用差距及未来发展趋势。以期为中药质量控制体系的升级和现代化发展提供新的思路。

关键词: 中药质量控制; 人工智能; 深度学习; 中药识别; 大数据; 算法模型

中图分类号: TP18; R282.6 文献标志码: A 文章编号: 0253-2670(2025)16-6048-10

DOI: 10.7501/j.issn.0253-2670.2025.16.032

Application and discussion of deep learning techniques in quality control of traditional Chinese medicines

WANG Shuqi^{1,2}, SHEN Lei^{1,2}, SONG Liya^{1,2}, GENG Puyu^{1,2}, QIU Ningning^{1,2}, LIU Wenlong^{1,2}, ZHANG Xili^{1,2}

1. School of Pharmacy, Hunan University of Chinese Medicine, Changsha 410208, China

2. Hunan Key Laboratory of Druggability and Preparation Modification of Traditional Chinese Medicine, Changsha 410208, China

Abstract: The scientific nature of traditional Chinese medicine (TCM) quality control methods is a key factor that restricts the development of global high-quality exhibition of TCM, and at the same time, the quality of TCM under the existing quality control is also a key factor in the effectiveness of TCM in treating diseases. Under the background of rapid development of artificial intelligence and big data, deep learning has gained wide attention in the research of quality control of TCM, which shows obvious advantages compared with traditional methods and has achieved remarkable results in the recognition of the appearance of TCM, composition analysis, safety assessment and process quality control. Therefore, starting from the basic framework and common models of deep learning technology, this paper systematically combs through the application modes of deep learning in the quality control of TCM, deeply analyzes the challenges it faces in the quality inspection of TCM, and puts forward targeted measures to solve the problems based on the existing research results. Finally, the application gaps and future development trends of deep learning in the field of quality control of TCM are summarized. In order to provide new ideas for the upgrading and modernization of the quality control system of TCM.

Key words: quality control of traditional Chinese medicine; artificial intelligence; deep learning; traditional Chinese medicine identification; big data; algorithmic model

收稿日期: 2025-05-11

基金项目: 国家自然科学基金项目(81874344); 湖南省重点领域研发计划(2023SK2046); 湖南省卫生健康高层次人才重大科研专项(R2023139); 湖南省自然科学基金项目(2023JJ60474); 长沙市自然科学基金项目(kq2208148, kq2208191); 湖南创新型省份建设专项(2024RC8110)

作者简介: 汪舒琦, 硕士研究生, 研究方向为中药质量。E-mail: wangsq0615@163.com

*通信作者: 张喜利, 副教授, 硕士研究生, 研究方向为中药质量。E-mail: xiaoli610@126.com

刘文龙, 教授, 博士, 研究方向为中药质量。E-mail: dragon5240@126.com

作为中国传统医学理论体系的物质载体，中药凭借其多成分、多靶点的治疗特性，在慢性病防治及全球补充替代医学领域占据战略地位。然而，中药因其化学成分的复杂性（单味药材就含有超过 200 种化合物），以及质量属性的非线性关联，使得其质量控制面临诸多挑战，如过度依赖经验、量化标准不足等科学难题。如今广泛应用的中药质量控制方法虽然在一定程度上保障了中药的安全性和有效性，但在面对大规模、多样化的中药产品时，显示出明显的局限性。根据国家中医药管理局发布的《中医药信息化发展“十三五”规划》，明确指出信息化方法是推

动中医药现代化的关键驱动力量。近年来，人工智能作为一种新兴技术，开始渗透社会各个领域，人工智能子领域深度学习技术的兴起为中药质量控制带来了新的机遇。深度学习技术，特别是结合了传统中医学和现代分析技术的可解释机器学习方法，通过处理和分析从各种分析仪器获得的大量数据，能够有效识别中药组分之间的复杂关系和模式。这为中药品质的实时、快速和准确评估提供了新的可能性，从而提高检测效率和准确性。本文旨在综述深度学习在中药质量控制中的应用，探讨其优势与挑战，并展望未来的发展方向（图 1）。

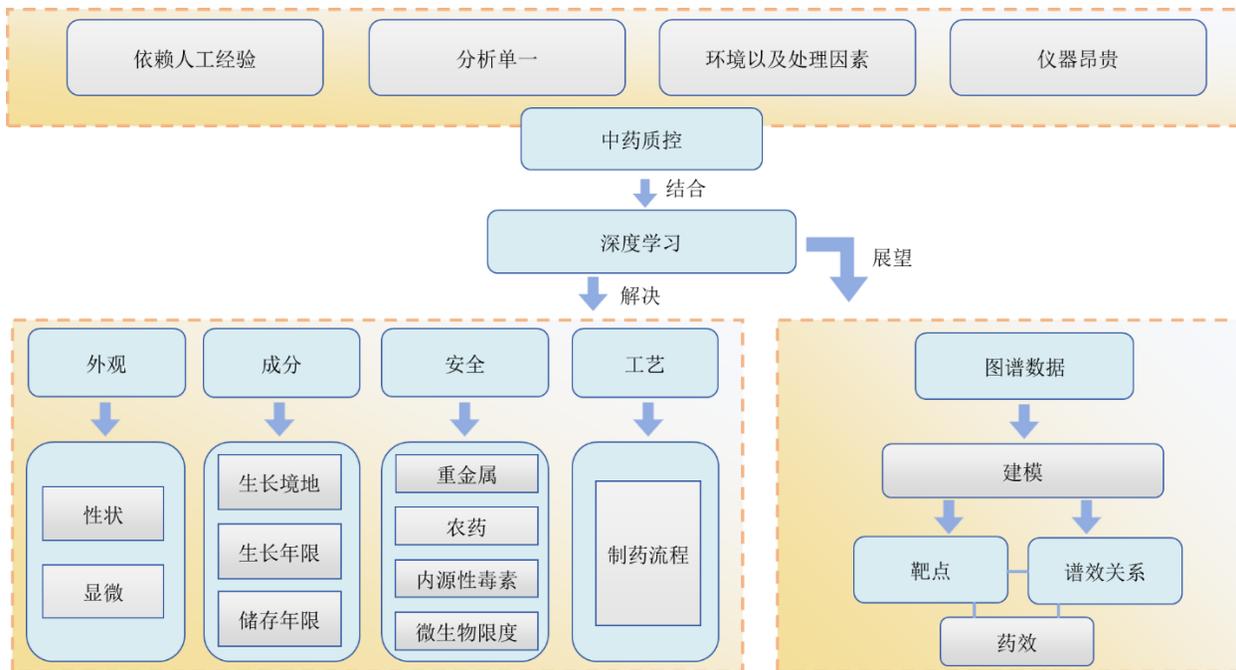


图 1 基于深度学习技术的中药质量控制

Fig. 1 Quality control of traditional Chinese medicine based on deep learning technology

1 中药质量控制的技术需求与深度学习适配性

1.1 深度学习的概述

近年来，深度学习作为机器学习领域的重要分支及人工智能的关键组成部分，取得了显著的研究成果（图 2）。深度学习的基础是神经网络，神经网络旨在通过其基本计算单元（也称为“神经元”）学习输入和输出之间的非线性映射。通常情况下，神经网络必须有 3 个以上的层才能被认为是深度神经网络（deep neural network, DNN），即深度学习，深层结构允许模型逐层提取更高层次的特征抽象^[1]。机器学习的发展过程主要分为浅层学习和深度学习 2 类。2006 年，其研究方向主要集中在用于数据处理的浅层学习结构上。相较于深度



图 2 人工智能与深度学习的关系

Fig. 2 Relationship between artificial intelligence and deep learning

学习, 浅层学习对非线性特征转换层的数量有所限制, 通常不超过2层。至今, 浅层学习在处理具有多重约束的问题时表现出高效性, 但在处理自然图片、复杂视觉场景等方面则显得力不从心。深度学习技术因其强大的数据处理和特征提取能力, 已在图像识别、语音识别和自然语言处理等领域得到了广泛应用, 并在解决传统算法难以处理的复杂问题方面展现出显著优势。深度学习算法中, 算法通过从自身错误中学习来自动提取特征^[2], 无需人工设计特征, 直接从数据中学习(如图像的边缘→纹理→物体部件)。与传统的机器学习模型相比, 层的嵌入带来了更高效的学习体验。如今经典深度学习模型有卷积神经网络(convolutional neural network, CNN)、递归神经网络(recurrent neural network, RNN)、变换器(Transformer)等。

1.2 传统方法的局限性与智能化转型契机

1.2.1 性状鉴别 传统的性状鉴别是指利用人的感官去判断, 这种方法优势在于快速、简便, 能够在短时间内对大量药材进行初步筛选。但中药的外观性状常因种属、光照、加工方式和采收时间的不同而存在较大的差异^[3], 受描述主观性影响大^[4]。尽管人工智能电子感官技术(如电子眼)通过图像处理提升了识别效率, 仍面临环境敏感(如光线、背景干扰)、特征提取单一及设备通用性不足等局限^[5]。然而深度学习基于神经网络的深度结构能够提取图像的高级特征表达, 实现从原始像素到高层语义的渐进式特征解耦, 对模糊等扰动具有相对较好的容忍度, 其本质在于构建干扰不变性表示, 从而在复杂环境中保持对视觉本质结构的鲁棒感知^[6]。

1.2.2 显微鉴别 显微鉴定是指用显微镜对药材(饮片)切片、粉末、解离组织或表面制片及含饮片粉末的制剂中饮片的组织、细胞或内含物等特征进行鉴别的一种方法^[7]。然而该技术高度依赖操作者专业经验, 无法进行化学成分分析, 且对复杂掺伪手段鉴别能力有限。而通过深度学习可构建的多层特征金字塔, 从微观纹理(如细胞核染色质分布)到宏观结构(如组织排列)的多层次特征^[8]。可分层解析显微结构中的多尺度特征, 相较于人工分析, 其效率显著提升, 从而弥补了高度依赖专业人员的不足。

1.2.3 成分检测 如今用来检测中药成分的方法主要有色谱法、光谱法、分子检测法等, 这些方法具有良好的客观性和量化性, 并且衍生出了一系列

创新性成果。例如, 用于表征中药化学特征的中药指纹图谱, 以及能够体现遗传差异的DNA条形码等。这些技术利用中药分子内部基团振动的倍频和合频吸收来实现对中药的快速鉴别, 同时在中药质量安全检测中亦发挥核心作用, 如重金属、农药残留、微生物限度等检测^[9-11]。然而, 实际操作过程中需要大量具有代表性且化学值已知的样品来建立模型, 因此, 对于小批量样品的分析, 采用上述技术显得不太切实际。此外, 中药成分复杂, 不同成分间存在复杂的相互作用; 许多现代分析技术需要使用高精度、高成本的仪器设备。现传统机器学习方法对以上问题有所改进, 常用的传统机器学习算法如支持向量机等, 它主要依赖于计算机和软件工具, 不需要昂贵的实验设备和复杂的仪器。传统机器学习方法能够处理复杂体系, 通过特征工程和模型训练做到精确预测^[12]。然而, 该方法通常依赖于手工特征提取, 既耗时又耗力, 因此在处理大规模数据时受到限制。深度学习因自动提取特征, 在处理大规模数据时具有显著优势, 因为具有更多的网络层数表示更复杂的函数, 能更好拟合数据的复杂非线性关系^[13]。与传统的机器学习方法相比, 深度学习不仅简化了数据处理流程与模型构建步骤, 还大幅度提升了模型的预测准确性。

1.2.4 制药工艺质量控制 中药生产工艺直接关系到中药制剂的质量与疗效。中药制剂的生产工艺决定了其安全、有效的物质基础^[14], 生产过程的控制水平决定了药物制剂的质量和功效的稳定^[15], 中药制剂的生产过程涉及多个工艺过程参数, 且不同单元操作、不同参数间还可能相互影响, 从而导致产品质量的差异^[16]。然而, 目前人工操作难以精准地控制生产参数, 这将使中药制剂的质量与疗效得不到保证。深度学习技术通过对大量的历史数据和实时监测数据可深入分析中药提取、浓缩、干燥等工艺参数与产品质量的关联性, 能够准确预测生物制品的产量、纯度和质量, 从而优化生产过程中的参数, 提高整体生产效率和产品质量^[17]。

1.3 深度学习技术模型与中药数据特性匹配性

中药质量控制是一个复杂的过程, 其涉及多维度、多模态的数据。这些数据涵盖图像数据, 诸如显微结构和外观形态, 其高分辨率图像能达 $1\ 000\times 1\ 000$ 像素, 因此数据量庞大, 细节丰富, 对数据存储、处理及分析的能力提出了高要求。光谱数据如近红外光谱和拉曼光谱, 由数千个波长点组

成，构建了一个高维特征空间。此外，时序数据同样重要，如炮制过程中温度、湿度的动态变化，这些因素直接影响着中药成分的含量。中药含有多种化学成分，这些成分间及其与药效之间存在着复杂的非

线性关系，因此，中药的质量控制难以通过简单的线性模型或单一指标进行准确评估。深度学习技术模型与中药数据特性具有良好的匹配性，3 种深度学习模型在中药质量控制中的应用对比见表 1。

表 1 3 种深度学习模型的对比

Table 1 Comparison of three deep learning models

模型	核心能力	擅长数据	中药质控主要应用点	优势	劣势
CNN	空间局部特征提取、平移不变性	高维图像/谱图数据	图像鉴别(性状/显微)、光谱/色谱图像分析	图像特征提取高效、参数共享、平移不变性好	处理长序列依赖差、难以建模全局上下文
RNN	时序依赖性建模、动态预测	时序数据	生产工艺时序监控/预测	建模序列时序依赖	并行性差、训练慢、长程依赖仍可能有问题
Transformer	全局依赖捕捉、多模态融合	多模态数据	谱效关系建模、图像鉴别	全局依赖建模最强、并行计算快	计算资源消耗大、数据需求量大、可解释性相对复杂

1.3.1 CNN CNN 是深度学习中最常见的一类模型，特别适用于处理图像数据。CNN 通过卷积层提取图像的局部特征，并通过池化层进行特征降维，最终通过全连接层进行分类或回归卷积操作，能够有效捕捉图像中的空间层次结构，且无论物体在图像中的位置如何都能对其进行识别，具有平移不变性^[18]。其擅长处理图像数据，所以其在计算机视觉领域广泛应用，如图像分类、目标检测和图像生成。CNN 的衍生模型，如残差神经网络 (residual network, ResNet)、You only look once (YOLO) 和视觉几何组 (visual geometry group, VGG)，在多个领域中得到了广泛应用，例如，在卫星影像处理中，VGG 模型被用于微调以检测影像镶嵌过程中的错位点；在人脸识别技术中，VGG 和 ResNet 模型被用来提高识别准确性。

1.3.2 RNN RNN 在处理序列数据 (如时间序列、文本) 方面表现出色。RNN 具有记忆功能，能够对序列中的历史信息进行建模，其通过循环连接，使得网络能够记忆和利用序列中的前后文信息，适用于动态变化的数据^[19]。经典的 RNN 存在长时依赖问题，而长短期记忆网络 (long short-term memory, LSTM) 通过特殊的门控机制解决了这一问题，设计用于处理序列数据，生产工艺参数是典型的时间序列数据，具有强时序依赖。特别是 LSTM 能够高效地建模长时间依赖关系，从而准确预测未来状态或识别出异常模式，广泛用于自然语言处理、时间序列预测等领域。

1.3.3 Transformer 2017 年，Vaswani 等提出了一种基于自注意力机制的深度学习架构

Transformer。Transformer 模型摒弃了传统的 RNN 和 CNN，通过自注意力机制动态地关注输入序列中的关键信息，有效捕捉长距离依赖关系，以及并行处理序列数据^[20]。在融合不同类型数据 (图像、文本、序列) 方面展现出强大能力，能学习跨模态的联合表示，适合处理多模态数据。

2 深度学习在中药质量控制中的应用

2.1 中药外观质量控制

2.1.1 中药原药材与中药饮片的性状鉴别 中药原药材与饮片通过性状鉴别成为保障质量的首要手段，凭借形态、色泽、纹理等直观特征能够迅速辨别真伪、验证炮制水平并评判品质优劣，这些特征与药效物质密切相关，结合深度学习技术，鉴别效率与准确性得以显著提升。在中药质量控制领域，图像识别作为深度学习的重要应用领域之一，特别是 CNN，近年来日益受到重视，并展现出巨大的应用潜力。孙鑫等^[21]使用 VGG16 的模型测试了 50 种中药饮片的图像，在图像材料多数为多个饮片相互遮蔽并带有复杂背景的条件下，依旧能成功识别。吴冲等^[22]通过获取中药外在性状特征，建立 ResNet50 模型来对 11 种饮片 (生山楂、炒山楂、焦山楂、山楂炭、松贝、青贝、炉贝、生半夏、姜半夏、法半夏、清半夏) 进行图像识别，炮制品与中药种类均能准确区分。CNN 等深度学习技术已成功应用于人参、黄芪、当归等百余种中药材的自动分类，其分类准确率均达到 90% 以上^[23]。中国科学院在 2021 年提出设计了一个双流 (CNN+Transformer) 协同架构，使其在视觉识别任务中同时获得了强大的局部细节

建模能力和全局上下文理解能力^[24]。张静^[25]提出了一种基于改进 Swin Transformer 模型的中草药图像识别方法。结合了 Transformer 和 CNN 的优点，既能够捕捉全局的依赖关系，又能够关注局部的细节特征，从而更加准确地捕捉到中草药的特征。总之，深度学习技术的应用极大地提升了中药材及其饮片的识别效率，显著提高了中药饮片“辨状论质”的精度与效率。

2.1.2 显微鉴别 中药显微鉴别通常应用于性状特征不显著或难以辨识的药材鉴别、多基原药材的种属区分、破碎中药材及粉末中药材、用饮片粉末制成的制剂的鉴别以及纯度检查等，但其对检验人员的经验和知识水平要求较高，如今引入深度学习技术在细胞图像的分割中，取得了较好的

分割效果^[26]。陈一峰等^[27]基于显微图像处理技术结合 YOLO 模型，设计并开发了一种中药自动鉴别系统。该系统通过采集中药材的显微图像数据，构建特征分析模型，实现了对药材显微结构的自动化识别与分类。Ma 等^[28]发现 CNN 能通过厚朴、黄柏、肉桂、黄连 4 种易混淆中草药的石细胞图像从而准确区分中草药。孙霞^[29]提出的 CNN 模型能够有效地识别出不同中草药粉末的显微特征，为中草药的鉴定提供了一种新的技术手段。将深度学习技术与显微技术相结合已成为中药鉴别领域的一项重要创新，降低了检测人员观察难度，显著提升了对中药显微图像的识别效率与准确性。

对于深度技术如何解决传统方法在外观质量控制上的局限性见表 2。

表 2 对于外观质量控制局限深度学习的解决方式

Table 2 Solution to limitations of appearance quality control through deep learning

应用场景	传统方法痛点	深度学习解决方案	文献
品种鉴别	外观相似	ResNet 精准聚焦麦冬的细微形态差异	30
		GoogLeNet 融合人参与西洋参饮片的图像特征	31
		CNN 的局部特征提取和 Transformer 的全局关系建模	32
炮制程度判定	主观性强	1D-CNN 直接从红外光谱中学习炮制品的光谱指纹	33
		CNN 有效识别红外光谱中不同炭化成分的蒲黄的特征向量	34
等级分级	肉眼分拣一致性差	YOLOv8 模型通过识别中药饮片面积	35
显微鉴别	依赖专家经验	CNN 识别中药粉末显微图像的特征，确定成分种类	36
		CNN 对 15 种中药材粉末显微导管特征图像进行识别	37
		添加双重注意力机制模块的 ResNet-50 自动学习和关注图像中最重要的特征区域，从而提高对细微特征的捕捉能力	38

2.2 中药成分质量控制

国际医药市场对药品的质量要求非常严格，只有加强中药成分质量控制，才能使中药产品符合国际标准。中药多为多成分体系，传统分析方法难以全面解析其化学指纹图谱及其功效。深度学习（如 CNN）可处理高维光谱、色谱（HPLC、GC-MS）等图谱数据，自动提取关键特征，识别微量成分或异常信号，从而分析中药生物功效的变化。例如，人工智能技术在金陵药业的中药配方优化中展示了其潜力，通过深度挖掘中药成分之间的联系，优化配方并识别有效成分的组合。通过深度学习与现有分析方法的协同应用，能够对中药进行精确分析鉴别，确保所选中药品质上乘，从而构建出更为高效、精确的中药质量保障体系。

不同的生长境地对中药的生长和成分积累有

显著影响。中药的生物活性成分组成及含量会随着环境条件的改变而发生变化。因此，生长境地的快速鉴定可为药材的质量控制提供有价值的参考^[39]。Zhang 等^[40]采用 HPLC-MS 对不同产地甘草的活性成分（如甘草酸、甘草苷等）进行综合评价，发现不同来源的甘草在活性成分含量和生物活性上存在显著差异。Ma 等^[41]以安徽、广东、河北、河南和江苏 5 省桑叶为材料，利用拉曼光谱结合 CNN 建模，证明不同产地桑叶有效成分有差异，可实现产地预测，筛选优质桑叶。厦门大学发明了一种结合高光谱成像和核磁共振的深度学习算法的枸杞产地溯源方法^[42]。在中药资源中，多年生中药占据了相当大的比例，同一品种在不同生长年限时，外观、成分和药效差异大，对其进行准确鉴别与质量控制是中药质量评价体系的关键内容。人参中的人参皂甙

含量随生长年限不同而有所变化^[43], 连翘若提前采收, 其有效成分含量不足, 药效会大打折扣^[44]。粉葛的生长年限影响其质量, 不同生育期的葛根素含量不同, 含量越高则质量越好。Zhang 等^[45]利用 CNN 结合高光谱成像技术, 开发了一种快速识别粉葛生长年份的分类框架。其次储存年限也是一个影响中

药质量的重要因素, 柑橘的气味和药性经过多年的老化才能形成具有老化价值的品种——陈皮。所以有研究建立了 CNN-LSTM 分类模型, 能快速、无损地区分不同年龄组柑橘产物型并进行了定量预测^[46]。深度学习技术用于中药成分质量控制的其他典型研究见表 3。

表 3 深度学习技术在中药成分质量控制的应用

Table 3 Application of deep learning technology in quality control of traditional Chinese medicine ingredients

类型	中药	方法技术	研究结果	文献
中药生长境地鉴别	三七	近红外同步二维相关光谱与 ResNet	区分准确度达到 100%	47
	茯苓	电子鼻与 DCNN-LSTM	分类准确率达到 88.9%	48
中药生长年限鉴别	重楼	傅里叶变换红外光谱与 CNN 优化模型残差神经网络	实现不同生长年限的识别, 准确率达到 100%	49
	大黄	高光谱成像与 CNN	鉴别模型准确率达到 98.10%	50
中药储存年限鉴别	枸杞	FF-EEM 荧光光谱与 CNN	正确分类率在 98% 以上	51
	普洱茶	电子舌与一维深度 CNN 结合 ELM 模型	准确率和模型泛化性能好	52

2.3 中药安全性质量控制

中药材及饮片的原材料主要来自自然环境和人工培育条件下的植物、动物或矿物等。由于土壤和水源污染、过量喷洒或误用不当农药、自身蓄积伴生等原因, 可能导致部分药材中的重金属、农药残留、微生物限度及内源性毒素偏高, 增加了用药风险。化学法、光谱法、色谱法、质谱法等技术手段可用于检测中药成分和含量以保障安全。随着技术的进步和检测要求的提高, 多技术融合检测方法越来越多地用于中药材及饮片的检测中, 如液相色谱-质谱联用, 并成为未来的主要发展方向。然而, 鉴于传统方法存在的局限性, 多技术融合检测方法亟需进一步的优化与升级。深度学习技术作为一种强大的人工智能手段, 凭借其对海量标注及未标注数据的深度学习能力, 能够构建出精确识别中药材中各类有害成分的模型。目前, 深度学习技术在中药质量安全检测上已有一定的应用。Lu 等^[53]提出了一种深度学习模型预测东南景天重金属。Zhou 等^[54]将高光谱成像的三维数据优势与深度 3D-CNN 强大的空间-光谱联合特征提取能力相结合, 实现了对生菜中多种重金属复合污染的高精度定量预测。有研究提出将手持光谱学与深度学习相结合, 将光谱数据输入到 1D-ResNet 模型中, 对金橘上农药残留进行了检测^[55]。Yang 等^[56]结合了 2 种深度学习网络, 即 Wasserstein 生成对抗网络 (wasserstein generative

adversarial network, WGAN) 和残差神经网络 (ResNet), 基于太赫兹光谱检测多菌灵, 分类准确率达到 91.4%。此外, 在微生物限度检查上, 赵培文等^[57]采集 30 批中药制剂中间体样本的表面增强拉曼散射光谱, 将拉曼光谱与 CNN 结合, 构建快速预判微生物限度模型, 该模型能够有效判断样本是否符合微生物限度标准。对于内源性毒素如马兜铃酸、吡咯里西啶生物碱等, 深度学习可辅助建立风险评估模型, 评估其毒性^[58]。迄今为止, 深度学习模型在中药质量安全检测领域的应用仍处于初步探索阶段, 相关的研究报道相对较少, 然而, 深度学习模型在处理和复杂数据以及复杂数据方面具有显著优势。鉴于中药的多组分、多靶标特性, 深度学习模型能够深入挖掘中药成分的结构和生物学特征, 在预测中药有害成分方面具有巨大的潜力, 有望为中药的安全性评价和临床合理应用提供有力的技术支持。

2.4 中药制药工艺质量控制

在现代中药制药工业领域, 常围绕中药材提取进行生产过程工艺参数优化, 可以有效保障中药材的提取效率、批次一致性、药效及安全性等。均匀设计、正交实验设计等传统的实验设计方法常被用于探索工艺条件, 但它们存在局限性, 如可探索的因素水平有限, 往往只能得出局部最优解, 难以满足更广泛的需求。借助数字建模分析, 深度学习能

够在中药生产过程中,利用丰富的历史数据和实时监测信息,对提取制备工艺中的多参数实施随机非线性优化,进而保障生产过程的高效性、稳定性,并提升产品质量。谢勇等^[59]利用 LSTM 模型对监测的温湿度、风速数据进行多步预测,从而动态调整干燥设备参数,最终实现当归出料含水率的标准偏差精确控制在 0.5%以内。苏州茂可科技有限公司发明了一种基于大数据的生物制药智能纯化工艺控制系统,该系统通过构建 CNN-LSTM 算法提高系统纯化效果^[60]。CNN 可以从关键参数(如温度、压力等)的变化数据中提取出有用的特征,为后续的预测提供高质量的输入,结合 LSTM 预测时序变化趋势,提高对生产过程的智能化控制和优化能力。He 等^[61]使用近红外光谱并建立 CNN-LSTM 算法精准地对竹茹煎剂浓缩过程中的多指标成分进行快速检测。Tian 等^[62]提出了一种结合注意力机制的新型 CNN-LSTM 模型,用于基于近红外光谱实现流化床制粒过程中水分含量的在线监测。若将投料物料质量参数和工艺过程参数综合考虑,则它们可构成的矩阵图便可使用 CNN 模型分析数据,仲恠等^[63]以去壁灵芝孢子粉制药过程为研究对象,对引发质量波动的关键参数与关键工艺过程参数进行分析,然后进一步引入 CNN 模型,构建了关键参数与关键质量属性之间的定量模型。阎续^[64]将近红外光谱与 CNN 结合,构建定量校正模型来监测洗脱过程及终点判断。

3 结语与展望

3.1 面临的挑战及建议

深度学习技术在中药质量控制中的应用虽然取得了一定的进展,但也存在一些问题和挑战,需要进一步研究和解决。

3.1.1 数据收集与检测标准化体系建设 深度学习模型高度依赖于大规模且高质量的数据集,然而在中药领域,数据稀缺与质量参差不齐成为一大挑战。中药材的来源复杂多样,产地、批次及炮制方法均对中药的各项质量数据产生显著影响。加之数据获取需耗费大量人力、物力,使得中药数据的收集尤为困难。此外,中药生产过程中数据采集的不完整性进一步加剧了数据质量与数量难以满足模型训练需求的困境。常见药材数据多,珍稀药材数据少,导致模型泛化能力不足。我国中医药标准化工作存在起步较晚、基础相对薄弱以及任务艰巨繁重的现状,而数据标准的缺失问题尤为突出,已成

为制约中医药与人工智能融合发展的关键瓶颈^[65]。需加强数据基础设施建设,统一中医药数据标准,构建高质量中药数据集。该数据集应涵盖中药材及饮片的分类、产地溯源、质量评价及药效评价等,确保各环节均有明确且可操作的标准,制定统一的数据标准和格式,如建立中药材质量检测数据的国家标准、规范检测方法和数据记录格式。加强中医药数据的互联互通,利用大数据等技术,搭建中药材及饮片检测数据共享平台,实现数据的存储、处理、分析和共享等功能。同时,探索多模态数据融合分析,将药材种植环境数据、生长状态图像、质量检测报告等多源信息进行整合分析,为检测标准的制定和完善提供更全面的依据,满足中药材及饮片行业的数据需求。构建“来源可溯、质量可控、数据可信、标准可通”的现代中药数字生态系统,推动中医药从经验科学向数据驱动的精准医学转型。

3.1.2 多模态数据融合困境 中药质量控制技术从主观到客观,从成分到疗效,每个方向都有自己模态的数据体系,有数值性、图片性、行为记录、图像性、结构化表格等各种模态数据类型。由于不同模态的数据在尺度上存在较大差异,且它们的采样时间与空间难以统一,因此同一样品的多种模态数据往往容易缺失,这给多模态数据的融合带来了不小的挑战。深度学习可开发跨模态对齐技术,如针对不同模态数据的时空对齐、将不同特征的模型数据嵌入数据共享空间的特征级对齐。还可采用广义相加模型或者鲁棒表征对缺失模态数据进行预测与补齐。

3.1.3 模型的可解释性 神经网络的黑箱化特性使其与传统中医药理论之间存在显著的解释鸿沟。即便模型能够准确预测,也难以明确其判断依据,如模型究竟是基于哪些特征来判定某个饮片不合格的。这在高度强调安全性和需要明确质量依据的药品监管中是一个重大障碍,难以满足质量控制中透明化、可追溯的监管要求,如 FDA 对药品生产数据的审查标准。为此,可以优化模型架构,嵌入注意力机制[如梯度加权类激活映射(gradient-weighted class activation mapping, Grad-CAM)]实时可视化决策依据(如定位药材霉变区域或关键色谱峰),并融合中药知识图谱约束特征学习方向,在神经网络中融入基于规则的模块,使其决策过程与人类可理解的规则相联系,从而便于解读模型是如何基于输入信息做出判断的。

3.2 展望

随着信息化时代的到来,深度学习技术在中药鉴定领域的应用日益广泛。通过融合性状鉴别、光谱色谱等多源数据构建指纹图谱和预测模型,实现了对中药内外质量特征的全面精准分析,显著提高鉴别准确性和药理机制研究层次,有力推动中药的国际化进程。同时,深度学习赋能中药质量检查和生产流程优化(结合工业4.0),高效处理海量数据,提升分析效率与精度,推动中药产业向自动化、智能化、高质量和可持续发展迈进。

随着人工智能技术的迭代升级,深度学习与中药质量控制的深度融合已从单一鉴别向多维度智能决策系统演进。传统的中药质量评价多侧重于对化学成分的定性和定量分析,而谱效关系研究将药效与化学成分谱图相结合,使质量评价不再局限于成分的有无或含量的高低,而是更加关注成分与药效之间的内在联系。中药谱效关系的难点在于3点:数据的高维性(色谱峰太多)、成分间相互作用复杂、药效评价主观。已有研究结合图谱数据利用主成分分析法等方法探究了中药的药效物质基础,为中药的质量控制提供了新的方法^[66-67]。但对于中药中各种化学成分的相互作用对中药药效的影响研究较少。深度学习技术擅长处理高维、非线性、存在交互作用的复杂数据,能自动提取特征,避免人工选择特征的偏差。因此,深度学习技术能全面分析中药材中的多种成分及其相互作用,避免仅关注单一成分的限制性。其次,根据数据特点和研究目标,选择合适的深度学习模型,如CNN用于处理谱图数据,RNN及其变体LSTM用于处理序列数据,如LSTM建模药物代谢时序过程,通过血药浓度的变化去预测药效的变化;Transformer适合处理多模态数据,即多源数据并列处理。将化学指纹图谱和药效数据输入模型,然后通过神经元的非线性变换提取数据的特征和模式从而预测药效。进一步构建中药成分-靶点数据库[如炎症成分与环氧合酶-2(cyclooxygenase-2, COX-2)炎症靶点],要整合已有的药物-靶点数据库,结合网络药理学,用大量数据去训练深度学习模型。将相应活性成分作靶点与靶效关系联合研究,将能更加全面、系统地对中药药效机制进行剖析,助力中药高质量发展。

深度学习模型融合了传统药性理论与现代分析技术,为药材鉴定、新药开发提供数据驱动的新范式。深度学习将加速中药质量控制从经验驱动向

数据驱动的范式转型。未来需加强复方配伍影响研究,建立标准化多模态数据库以提升预测准确性,强化计算机科学、合成生物学与中医药理论的交叉创新,最终构建“智能鉴别-精准制药-个体化用药”的全链条质控生态,为中药现代化与国际化提供核心驱动力。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Zhu Y, Wang M, Yin X, *et al.* Deep learning in diverse intelligent sensor based systems [J]. *Sensors: Basel*, 2022, 23(1): 62.
- [2] Ahmed S F, Alam M S B, Hassan M, *et al.* Deep learning modelling techniques: Current progress, applications, advantages, and challenges [J]. *Artif Intell Rev*, 2023, 56(11): 13521-13617.
- [3] 黄建萍, 夏泽冰, 井亚江, 等. 基于“辨状论质”理论分析不同采收期黄芩外观性状与内在质量的相关性 [J]. *中草药*, 2024, 55(10): 3487-3496.
- [4] 王皓南, 田滢琦, 刘大会, 等. 中药“辨状论质”的历史、发展与应用 [J]. *中药材*, 2021, 44(3): 513-519.
- [5] 江如蓝, 雷结语, 陈文礼, 等. 基于人工智能感官技术的中药质量控制方法研究进展 [J]. *药学前沿*, 2024, 28(11): 550-556.
- [6] Uličný M, Lundström J, Byttner S. Robustness of deep convolutional neural networks for image recognition [A] // *Intelligent Computing Systems* [M]. Cham: Springer International Publishing, 2016: 16-30.
- [7] 中国药典 [S]. 四部通则. 2020: 265.
- [8] Campanella G, Hanna M G, Geneslaw L, *et al.* Clinical-grade computational pathology using weakly supervised deep learning on whole slide images [J]. *Nat Med*, 2019, 25(8): 1301-1309.
- [9] 张怡婧, 张书乐, 陈海洋, 等. 中药材中重金属的监测与控制研究进展 [J]. *中南药学*, 2025, 23(3): 717-724
- [10] 周婷婷, 刘芫汐, 金红宇, 等. 农药残留快速检测技术研究进展及其在中药中的应用 [J/OL]. *中国现代中药*, [2025-05-15]. <https://doi.org/10.13313/j.issn.1673-4890.20241008003>.
- [11] 吴伟. 中药饮片微生物限度检查方法探究及质量控制 [J]. *实验室检测*, 2025, 3(1): 149-151.
- [12] 夏铭泽. 基于改进支持向量机的产品质量预测系统的研究 [D]. 北京: 机械科学研究总院, 2020.
- [13] Le Cun Y, Bengio Y, Hinton G. Deep learning [J]. *Nature*, 2015, 521(7553): 436-444.
- [14] 陈霞, 阳长明, 陈浩, 等. 基于中药复方制剂特点的中药复方制剂生产工艺研究 [J]. *中草药*, 2021, 52(19): 5807-5813.

- [15] 申向荣, 阳长明, 王若瑾, 等. 中药在加拿大注册情况及中药国际监管协调对策建议 [J]. 中药药理与临床, 2024, 40(8): 2-9.
- [16] 熊诗慧, 邱婷, 王学成, 等. 基于质量标志物 (Q-Marker) 的中药制剂质量过程控制方法与策略 [J]. 中草药, 2021, 52(9): 2534-2540.
- [17] 薛启隆, 苗坤宏, 于洋, 等. 基于深度强化学习的中药制药过程自主优化决策方法研究 [J]. 中国中药杂志, 2023, 48(2): 562-568.
- [18] Le Cun Y, Bottou L, Bengio Y, *et al.* Gradient-based learning applied to document recognition [J]. *Proc IEEE*, 1998, 86(11): 2278-2324.
- [19] 袁冰清, 陆悦斌, 张杰. 神经网络与深度学习基础 [J]. 数字通信世界, 2018(5): 32-33.
- [20] Zhang Y, Liu C, Liu M, *et al.* Attention is all you need: utilizing attention in AI-enabled drug discovery [J]. *Briefings Bioinform*, 2023, 25(1): 1-22.
- [21] 孙鑫, 钱会南. 基于深度卷积网络的中药饮片图像识别 [J]. 世界科学技术—中医药现代化, 2017, 19(2): 218-222.
- [22] 吴冲, 谭超群, 黄永亮, 等. 基于深度学习算法的川贝母、山楂及半夏饮片的智能鉴别 [J]. 中国实验方剂学杂志, 2020, 26(21): 195-201.
- [23] 胡晓东. 基于改进深度学习算法的中药饮片图像识别研究 [D]. 长春: 吉林农业大学, 2023.
- [24] Peng Z L, Huang W, Gu S Z, *et al.* Conformer: Local features coupling global representations for visual recognition [A] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV) [C]. Montreal: IEEE, 2021: 357-366.
- [25] 张静. 基于改进 Swin Transformer 模型的中草药图像识别研究 [J]. 云南师范大学学报: 自然科学版, 2024, 44(6): 44-47.
- [26] 夏平, 王爽, 胡蓉, 等. 全卷积深度学习网络的细胞显微图像分割 [J]. 计算机仿真, 2022, 39(4): 133-141.
- [27] 陈一锋, 庞光垚, 黄宏本. 基于显微图像处理技术的中药自动鉴别系统设计 with 实现 [J]. 智能城市, 2024, 10(8): 1-3.
- [28] Ma Y Y, Zhong Y M, Su Q, *et al.* Study on identification algorithm of traditional Chinese medicinals microscopic image based on convolutional neural network [J]. *Medicine*, 2023, 102(25): e34085.
- [29] 孙霞. 基于深度学习的中草药粉末显微特征研究 [D]. 北京: 北方工业大学, 2020.
- [30] 秦竹媛, 吴浩忠, 谭代庆, 等. 基于多尺度 ResNet 融合注意力机制的麦冬细粒度识别 [J]. 计算机与现代化, 2023(7): 105-111.
- [31] 徐飞, 孟沙, 吴启南, 等. 基于卷积神经网络的人参西洋参饮片鉴别方法研究 [J]. 南京中医药大学学报, 2018, 34(6): 621-624.
- [32] 谭代庆, 肖志鸿, 吴浩忠, 等. 融合改进残差网络和 Transformer 的易混淆中药饮片识别研究 [J]. 中国数字医学, 2023, 18(6): 42-50.
- [33] Shi Y B, He T Y, Zhong J J, *et al.* Classification and rapid non-destructive quality evaluation of different processed products of *Cyperus rotundus* based on near-infrared spectroscopy combined with deep learning [J]. *Talanta*, 2024, 268(Pt 1): 125266.
- [34] 陈承武, 王天舒, 胡孔法, 等. 基于卷积神经网络与投票机制的蒲黄炮制品近红外判别方法 [J]. 光谱学与光谱分析, 2022, 42(11): 3361-3367.
- [35] 彭彪. AI 技术在中药饮片 (以羌活、当归为例) 识别与分级中的应用探究 [D]. 济南: 齐鲁工业大学, 2024.
- [36] 山东省食品药品检验研究院. 基于显微图像的中药粉末智能鉴别方法和系统: 中国, CN202010585564.1 [P]. 2020-09-25.
- [37] 王一丁, 石铎, 李耀利, 等. 基于 SqueezeNet 深度网络的中药材粉末显微特征图像识别研究 [J]. 电子显微学报, 2019, 38(2): 130-138.
- [38] Zhu X Y, Pang G Y, He X, *et al.* A segmentation-combination data augmentation strategy and dual attention mechanism for accurate Chinese herbal medicine microscopic identification [J]. *Front Plant Sci*, 2024, 15: 1442968.
- [39] Yang Y, Wu Y J, Li W L, *et al.* Determination of geographical origin and icariin content of *Herba Epimedii* using near infrared spectroscopy and chemometrics [J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2018, 191: 233-240.
- [40] Zhang X M, Guo X H, Zhao P, *et al.* Chemometric analysis of active compounds and antioxidant and α -glucosidase inhibitory activities for the quality evaluation of licorice from different origins [J]. *Biomed Chromatogr*, 2021, 35(12): e5215.
- [41] Ma Z W, Tang J W, Liu Q H, *et al.* Identification of geographic origins of *Morus alba* Linn. through surfaced enhanced Raman spectrometry and machine learning algorithms [J]. *J Biomol Struct Dyn*, 2023, 41(23): 14285-14298.
- [42] 厦门大学. 一种结合 HSI 与 NMR 技术的深度学习算法的枸杞产地溯源方法: 中国, CN202411211170.4 [P]. 2024-12-06.
- [43] 余江锋, 李育平, 何伟, 等. 吉产不同生长年限人参中 8 种主要人参皂苷与人参皂苷 Rg_1 比值的规律研究 [J]. 中国药房, 2019, 30(1): 31-35.
- [44] 王明喜. 浅谈中药的采收与质量的关系 [J]. 中医临床

- 研究, 2016, 8(13): 116-117.
- [45] Zhang L, Guan Y, Wang N, *et al.* Identification of growth years for *Puerariae Thomsonii Radix* based on hyperspectral imaging technology and deep learning algorithm [J]. *Sci Rep*, 2023, 13: 14286.
- [46] Qin Y W, Zhao Q, Zhou D, *et al.* Application of flash GC e-nose and FT-NIR combined with deep learning algorithm in preventing age fraud and quality evaluation of pericarpium citri reticulatae [J]. *Food Chem X*, 2024, 21: 101220.
- [47] Liu C L, Zuo Z T, Xu F R, *et al.* Study of the suitable climate factors and geographical origins traceability of *Panax notoginseng* based on correlation analysis and spectral images combined with machine learning [J]. *Front Plant Sci*, 2023, 13: 1009727.
- [48] 徐昊, 章检明, 王中鹏, 等. 基于电子鼻的深度卷积神经网络茯苓产地分类方法 [J]. 传感器与微系统, 2023, 42(12): 138-141.
- [49] Yue J Q, Li Z M, Zuo Z T, *et al.* Study on the identification and evaluation of growth years for *Paris polyphylla* var. *yunnanensis* using deep learning combined with 2DCOS [J]. *Spectrochim Acta Part A Mol Biomol Spectrosc*, 2021, 261: 120033.
- [50] 郑州大学. 基于多模态图谱融合的大黄年限鉴别方法: 中国, CN202310295861.6 [P]. 2023-06-27.
- [51] Yan X Q, Wu H L, Wang B, *et al.* Front-face excitation-emission matrix fluorescence spectroscopy combined with interpretable deep learning for the rapid identification of the storage year of Ningxia wolfberry [J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2023, 295: 122617.
- [52] 杨正伟, 张鑫, 李庆盛, 等. 基于电子舌及一维深度 CNN-ELM 模型的普洱茶贮藏年限快速检测 [J]. 食品与机械, 2020, 36(8): 45-52.
- [53] Lu Y, Nie L J, Guo X Y, *et al.* Rapid assessment of heavy metal accumulation capability of *Sedum alfredii* using hyperspectral imaging and deep learning [J]. *Ecotoxicol Environ Saf*, 2024, 282: 116704.
- [54] Zhou X, Sun J, Tian Y, *et al.* Hyperspectral technique combined with deep learning algorithm for detection of compound heavy metals in lettuce [J]. *Food Chem*, 2020, 321: 126503.
- [55] Dai Q F, Luo Z, Li Z, *et al.* Field-based, non-destructive, and rapid detection of pesticide residues on kumquat (*Citrus japonica*) surfaces using handheld spectrometer and 1D-ResNet [J]. *Agronomy*, 2025, 15(3): 625.
- [56] Yang R Z, Li Y, Qin B Y, *et al.* Pesticide detection combining the Wasserstein generative adversarial network and the residual neural network based on terahertz spectroscopy [J]. *RSC Adv*, 2022, 12(3): 1769-1776.
- [57] 赵培文, 李芷瑶, 刘艺丹, 等. 基于机器学习的中药制剂中间体微生物限度快速预判研究 [J]. 分析测试学报, 2024, 43(11): 1725-1734.
- [58] 颜彩琴, 范睿琦, 宁雨坪, 等. 深度学习模型在中药毒性预警中的应用和前景 [J]. 中国药理学与毒理学杂志, 2022, 36(3): 231-240.
- [59] 谢勇, 李良庭. 基于长短期记忆人工神经网络的中草药干燥含水率感知预测模型 [J]. 智能物联技术, 2025, 57(2): 136-140.
- [60] 苏州茂可科技有限公司. 基于大数据的生物制药智能纯化工艺控制系统: 中国, CN202411194465.5 [P]. 2025-02-11.
- [61] He T Y, Shi Y B, Cui E Z, *et al.* Rapid detection of multi-indicator components of classical famous formula Zhuru Decoction concentration process based on fusion CNN-LSTM hybrid model with the near-infrared spectrum [J]. *Microchem J*, 2023, 195: 109438.
- [62] Tian G, Zhao J, Qu H B. A novel CNN-LSTM model with attention mechanism for online monitoring of moisture content in fluidized bed granulation process based on near-infrared spectroscopy [J]. *Spectrochim Acta Part A Mol Biomol Spectrosc*, 2025, 340: 126361.
- [63] 仲怿, 范晓辉, 李振皓. 基于精益理念的中药制药工业数据挖掘及生产工艺持续改进策略研究: 以去壁灵芝孢子粉制药过程为例 [J]. 中国中药杂志, 2023, 48(3): 829-834.
- [64] 阎续. 机器学习在制药过程分析中的应用基础研究 [D]. 杭州: 浙江大学, 2020.
- [65] 林静怡, 李诗翩, 郭义, 等. 人工智能助力中医药发展现状、问题及建议 [J]. 世界中医药, 2022, 17(6): 864-867.
- [66] 谭小青, 唐红珍, 高红伟, 等. 黄精 GC-MS 指纹图谱及体外抗病毒活性的谱效关系研究 [J]. 中药药理与临床, 2021, 37(1): 116-120.
- [67] 黄李璐, 郑雨, 王希, 等. 木芙蓉叶抗氧化活性成分谱效关系研究 [J]. 中国现代应用药学, 2022, 39(19): 2489-2497.

[责任编辑 潘明佳]