

数据融合技术在中药分析领域中的应用进展

赵倩^{1,2}, 缪培琪^{3,4}, 李小莉^{1,2}, 于洋^{1,2}, 刘长青^{3,4*}, 李正^{1,2,5*}

1. 天津中医药大学中药制药工程学院, 天津 301617
2. 组分中药国家重点实验室, 天津 301617
3. 天津现代创新中药科技有限公司, 天津 300380
4. 国家地方共建现代中药创新中心, 天津 300392
5. 天津现代中医药海河实验室, 天津 301617

摘要: 由于中药成分复杂, 依据单一的检测分析技术往往无法获得充分的样本化学信息, 信息表征不全面导致中药质量检测精度较低。数据融合技术通过对多种检测信息结果进行关联综合分析, 可从信息互补的角度增强检测精度、降低检测误差。目前, 已经有相关学者将数据融合技术应用到中药检测分析中并取得了较好的结果。通过从中药原料、生产加工过程及终产品质量控制 3 个方面出发, 论述了目前数据融合技术在中药领域全过程生产质量检测和控制中的应用进展, 展现了数据融合技术与中药检验及质量控制结合的应用潜力, 为数据融合技术在中药分析领域中的进一步深入应用提供参考。

关键词: 数据融合; 中药; 分析领域; 生产过程; 质量检测; 质量控制

中图分类号: R282 **文献标志码:** A **文章编号:** 0253-2670(2023)11-3706-09

DOI: 10.7501/j.issn.0253-2670.2023.11.032

Application progress on data fusion technology in analysis field of traditional Chinese medicine

ZHAO Qian^{1,2}, MIAO Pei-qi^{3,4}, LI Xiao-li^{1,2}, YU Yang^{1,2}, LIU Chang-qing^{3,4}, LI Zheng^{1,2,5}

1. College of Pharmaceutical Engineering of Traditional Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China
2. State Key Laboratory of Component Traditional Chinese Medicine, Tianjin 301617, China
3. Tianjin Modern Innovative TCM Technology Co., Ltd., Tianjin 300380, China
4. National and Local Joint Innovation Center for Modern Chinese Medicine, Tianjin 300392, China
5. Haihe Laboratory of Modern Chinese Medicine, Tianjin 301617, China

Abstract: Due to the complex composition of traditional Chinese medicine (TCM), it is often impossible to obtain sufficient chemical information from samples based on a single detection and analysis technique. The situation of incomplete detection of information leads to low accuracy of TCM quality detection. Data fusion technology can enhance detection accuracy and reduce detection errors from the perspective of complementary information by correlating and synthesizing the results of multiple detection information. At present, some scholars have applied data fusion technology to the analysis of TCM detection and achieved good results. By summarizing the research on three aspects of data fusion technology in the quality control of raw materials, production and processing processes, and final products of TCM, this article discusses the application of data fusion technology in the detection and control of product quality during the whole production process of TCM, showing the potential of data fusion technology combined with TCM inspection and quality control, and providing a reference for the further in-depth application of data fusion technology in the field of TCM analysis.

Key words: data fusion; traditional Chinese medicine; analysis field; production processes; quality testing; quality control

收稿日期: 2023-01-07

基金项目: 国家自然科学基金面上项目 (82074276); 国家中医药管理局创新团队与人才培养计划资助 (ZYXCXTD-D-202002)

作者简介: 赵倩, 硕士研究生, 研究方向为中药智能制造。E-mail: zhao03129983@163.com

*通信作者: 刘长青, 高级工程师, 从事中药智能装备研发和车间智能化等应用研究。E-mail: cq_liu@cntmic.com

李正, 研究员, 从事中药智能制造关键技术及装备的研发与产业化应用基础研究。E-mail: lizheng@tjutcm.edu.cn

中药成分复杂、基体元素多样的特点导致了中药产品质量可控性较差的问题^[1]。依照《药品生产质量管理规范》中对药品质量提出的要求，需要对中药生产过程中的每个环节进行严格的质量检测与控制以保证最终中药产品的质量均一性。因此，加强从原料采集、中药生产加工过程及终产品生成这3个大环节的质量检测和控制是保障中药产品质量的关键手段。

目前大多数的中药质量检测方法多针对分析需求，选择单一数据源的检测结果进行分析。但由于检测过程中存在其他因素的干扰，导致依据片面的中药信息进行分析检测的精度较低^[2-3]。因此需要从宏观整体的检测角度上进行中药质量的检测与评价以实现中药产品质量的精准控制^[4-6]。为此，研究学者们提出一种可以结合多种信息用作整体分析的技术——数据融合。数据融合是近些年分析化学领域内逐渐发展起来的一种分析技术，通过将多种传感器检测信息进行互补来满足精准全面检测的要求^[7-8]。无论是通过多数据源检测弥补单一数据源的检测精度，还是将不同原理分析仪器的检测信息进行融合达到整体质量评价的目的，数据融合技术相比传统的单一检测分析技术都具有应用优势。目前，数据融合已经广泛应用于矿业^[9-11]、航空^[12-13]、食品^[14-15]、农业^[16]等领域的检测分析领域，面对高精度的目标检测需求，数据融合技术的应用取得了良好的效果。在现代中药各个生产环节中运用数据融合，同样对于实现中药全面精确的检测效果具有重要的意义。

随着数据融合技术的不断发展，越来越多的研究者开始尝试将不同检测仪器获取的数据信息进行融合，其中以指纹图谱信息融合在中药分析领域中的应用最为广泛^[17-20]。通过融合不同检测原理的光谱仪器获取相应的指纹图谱数据信息，结合化学计量学进行分析，实现多源信息的融合，达到中药质量整体检测与评价的目的。目前，研究者主要就中药原料检测、中药生产过程的实时检测分析及中药产品的质量评价，对数据融合在中药检测分析领域中的应用进行研究。相关研究成果表明数据融合在产地鉴别、伪劣品鉴定、药材品质评价、中药生产过程质量评价、中药制药过程在线质量控制、中药制剂和产品包装质量检测等方面具有应用优势和潜力。因此，本文拟对目前数据融合技术在中药检测领域中的研究现状进行综述，并就如何在中药生

产加工全过程中充分应用展开讨论，为数据融合技术在中药分析领域中的深入应用提供指导和参考。

1 数据融合技术

数据融合技术主要是通过将由多个数据源获取的数据与来自相关数据库的信息进行结合，以获得对于检测目标而言质量更高或更相关的数据信息。它的独特之处在于其数据可从不同分析技术中独立生成，然后组合用于模型的建立及结果的计算和预测。与单个数据源相比，它可以实现更高的精度和更具体的分析判断^[21]。因此，数据融合往往被用于建立高精度和高稳定性的预测模型以满足研究人员的检测需求，完成更准确的数据信息分析^[22]。如图1所示，数据融合可根据不同信息源的融合级别分为低级、中级和高级数据融合3类^[23]。

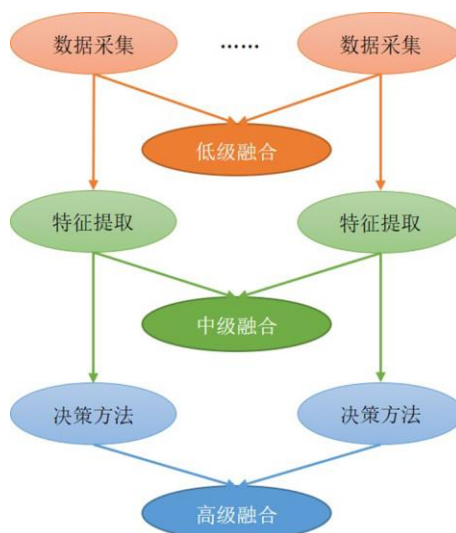


图1 不同层次数据融合的实现结构

Fig. 1 Implementation structure for different levels of data fusion

其中，低级数据融合是简单的数据层融合，通过将来自不同数据源的全部信息通过首尾相连的方式直接串联。然而研究学者发现这种融合方式常常会把一些冗余的数据也应用到后续建模过程中，使数据处理的工作量变大^[24]，故而常补充采用中级融合以达到精简模型的目的。

中级数据融合是分别从每个数据源中通过特定的分析方法提取相关特征，然后将它们组合成新的特征数据矩阵，因此又称为特征层融合。这一融合方式成功实现了数据的降维，突出了有效信息，有利于快速建模，在多数情况下相比低级数据融合效果较好。但是由于数据量的减少，它所构建的模型

在精度上往往会受到一定的影响,因此高级数据融合往往被应用于建立更高精度的模型^[25]。

高级数据融合又称决策层融合,基于不同分析仪器数据所构建的分析模型,此融合以一定的方法对其分别赋予不同的权重,从而构建出一种新的决策方式来实现建模分析。因此高级数据融合往往具有高灵活性和强包容性,在简化模型的同时能够取得高精度的检测结果。然而合适的决策方法探索相对比较困难,所以高级数据融合的应用相比于低级和中级数据融合较为复杂^[26]。

近年来,数据融合技术凭借其可以组合多个传感器信息进行更加准确全面的检测优势得到了不同领域研究人员的青睐。多个检测领域的研究结果表明相比较单个数据源建立的检测模型而言,经过数据融合技术建立的模型体现出更强的检测精度和实用性^[27-32]。

中药的成分复杂,其化合物种类和含量变化都会影响产品质量。但是中药质量检测中多以单一仪器的检测结果进行质量评价,导致存在分析精度较低的检测缺陷,需要应用数据融合技术对中药样本的信息进行深度挖掘。通过将生产过程中工艺参数、物料属性、多传感器中检测出的中药化学信息等进行数据融合,实现生产过程重要环节的准确检测。

2 数据融合用于精准管控中药原料质量

中药原料的优劣直接影响中成药产品质量。中国食品药品监督管理局 2018 年发布的最新传统药物开发指南鼓励药材同质化,以确保最终制剂质量的稳定性。原料药材的品质通常受原药材品种、采收、产地、部位等多种因素的影响^[33]。其中,药材品种和产地通常是影响中药原料质量的 2 个重要因素。近年来,已有研究人员就中药原料的产地鉴别、伪劣品鉴定、品质评价 3 方面进行了数据融合检测的相关研究,为数据融合与中药后续生产结合奠定了基础。

2.1 产地鉴别

产地鉴别有助于筛选出符合国家相关标准要求的道地药材。这类药材由于特定舒适的生长环境通常具有比其他产地同种药材更好的品质和药效,为中药原料的质量控制提供了更便捷的选择方案。研究者可以通过数据融合技术弥补单一检测技术信息不足的缺陷实现精准产地判别。

吴雪梅^[34]利用紫外可见光谱、衰减全反射傅里叶变换红外光谱和超高效液相色谱,采用低级、中

级和高级数据融合结合化学计量学建模实现了 177 份不同产地的野生滇重楼的来源鉴别,主要涉及的数据处理分析软件为 OMNIC 8.2、UVProbe 2.34 及 SIMAC-P⁺ 13.0。其中低级数据融合的数据矩阵为紫外可见光谱、衰减全反射傅里叶变换红外光谱数据预处理后直接串联;中级数据融合的数据矩阵为采用主成分分析分别提取两光谱特征变量后组合的新矩阵;高级数据融合先分别建立紫外可见光谱和衰减全反射傅里叶变换红外光谱的随机森林分类模型,再根据模糊集合理论的决策方法对分类结果进行融合得到决策层融合结果。通过比较单光谱、数据融合建立的产地识别模型结果可知高级数据融合模型的效果最佳,预测正确率高达 98%。同样,Pei 等^[35]采用傅里叶变换中红外光谱和近红外光谱结合低级、中级、高级数据融合,基于 OMNIC 9.7.7、SIMAC-P⁺ 13.0、MATLAB R2017a 和 RStudio 3.5.2 软件进行数据处理和建模分析,通过建立偏最小二乘判别分析和随机森林分类模型成功实现了 196 种野生重楼的地理溯源,其中由主成分特征变量提取的高级数据融合建立的模型产地识别准确率为 100%,证明这 2 种光谱融合用于野生云南重楼精准产地溯源的可行性。

此外,还有一些其他学者在茯苓^[36]、三七^[37]、杜仲叶^[38]、牛肝菌^[39]等中药材上成功应用数据融合实现了产地鉴别。除了应用各种光谱、色谱信息,图像特征信息亦被用于产地鉴别的信息补充。殷文俊等^[40]基于高光谱成像技术在可见-近红外波段和短波红外波段中提取光谱和图像特征,从可见近红外-短波红外、光谱-图像和全数据 3 个维度上进行数据融合,通过采用 HIS Analyzer、ENVI 5.3、MATLAB 2017a 和 Spyder (Python 3.7) 4 个软件进行数据校正与分析处理,建立了不同产地甘草药材的分类预测模型,结果表明不同维度的数据融合模型增强了单数据源模型的分类性能。以上研究证明了数据融合相对单个检测源数据分析可以取得更好的中药产地鉴别结果,但是大量的分析软件的应用往往会造成数据融合分析流程繁琐,在某一数据处理软件中实现完整数据融合分析流程,通过精简分析过程将会进一步促进数据融合技术在中药分析领域中的实际应用。

2.2 伪劣品鉴定

伪劣品鉴定也是数据融合技术应用在中药原料检测上的重要组成部分,常见为重量掺伪、形似掺

伪、染色掺伪、纯加工品掺伪、非药用部位掺伪等。通过数据融合进行中药材的伪劣品鉴定,可为原料药材提供质量保障。Yang等^[41]利用不同等级三七掺杂样本的近红外和中红外光谱进行低级、中级、高级数据融合构建了不同混合比例的三七分类模型。主要分析软件 Unscrambler 10.4、OriginPro 9.0 和 MATLAB 2018a 的数据分析结果显示,与单光谱分析结果相比,数据融合技术的应用成功提高了三七粉中掺假物的识别能力。其中低级数据融合效果不理想,考虑到可能是原始信息中包含大量的干扰和无用信息,利用高级数据融合建立的随机森林特征变量选择模型在精简模型的同时取得了最好的分类结果。此外,在加工品掺伪中最常见的为直接伪造真品。这种伪品与真品相似,通常需要专业的技术人员才能够进行辨别。如用新疆虫草代替冬虫夏草、用红花代替西红花、用桔梗根代替人参及用羊角藤代替巴戟天等。孙飞等^[9]通过近红外和中红外光谱技术,基于 SIMCA-P11.5 和 Unscrambler 7.0 软件进行数据处理,采用单一光谱数据和中级数据融合分别建立起姜半夏和伪品姜虎掌南星的偏最小二乘判别分析模型,结果显示采用单一近红外或中红外建立的检测模型分类准确率分别为 84.62%、92.41%,而采用数据融合进行建模的准确率为 100%,且样品在模型的空间分布呈现明显的分类聚集状态。这一结果表明数据融合后模型性能均优于单一光谱,证明近红外和中红外的信息互补可以提高姜半夏的真伪鉴定结果。综上,数据融合可以显著提高中药材伪劣品的鉴别准确率,但目前采用近红外仪器进行数据融合技术鉴定中药真伪的应用较多,缺乏其他不同原理检测信息源的数据融合分析,从扩充数据源的角度入手可以进一步实现数据融合对中药真伪全面精准的鉴别分析。

2.3 药材品质评价

药材品质评价也可以通过数据融合技术得以实现。当选用原料药材进行中药产品的制备时,通常需要根据《中国药典》2020年版^[42]进行相关有效成分的含量检测以控制原料药材的质量。这一环节中通常会涉及多种仪器进行检测,依据数据融合技术实现全面检测通常可以取得良好的品质判断效果。杨海玲^[43]以不同等级的川芎作为研究对象,基于感官评价、顶空气相色谱-质谱联用与近红外的检测数据,从气味、颜色、化学成分上进行综合分析,结合化学计量学建立不同质量等级的川芎识别模型。

经过 SPSS 21.0、OMNIC 9.7.7、SIMCA-P 14.1.0、MATLAB 2016a 和 Python 3.7.3 软件处理的数据分析结果表明基于极限学习机算法建立的分类模型可以很好地区分不同等级的川芎。毛胜楠^[44]通过将高效液相色谱法测定的余甘子中没食子酸等多种有效成分的信息与傅里叶变换近红外光谱采集的余甘子整体的化学指纹图谱进行融合,基于 OMNIC 和 SIMCA-P 13.0 数据分析软件,通过结合多部分变量特征建立偏最小二乘判别分析算法模型和 *t*-分布邻域嵌入算法模型,成功鉴别了不同储藏年份不同品质的余甘子,相比单一的傅里叶变换近红外鉴别模型,融合模型分类准确率由 89.8%提升至 100%。证明数据融合用于增强不同储藏年限余甘子分类的有效性。Qi等^[45]基于高效液相色谱获得的定量和定性代谢谱结合傅里叶变换近红外和傅里叶变换中红外评估不同产地黄连的质量。研究中,基于高效液相色谱测定的 8 种生物碱,筛选出可作为黄连品质鉴定的指标化合物,参考指标化合物含量数据建立单一技术模型和光谱融合模型。结果表明将光谱技术数据融合后结合偏最小二乘回归算法可以有效确定指标化合物含量,取得更好的预测结果,融合技术可以为黄连提供更全面有效的质量评估。由此可见,数据融合技术可以通过融合不同检测仪器的检测信息,多角度全方面了解中药样本信息,精准评价中药品质。此外,以上研究中多采用传统的偏最小二乘算法建立融合分析模型,多种检测分析模型的应用可从整体上进行比较分析,选出最优的质量检测模型,将数据融合更好地应用于中药实际质量评估。

3 数据融合用于实时控制中药生产过程

如何有效控制生产过程中中药质量保证产品的质量均一性一直是中药质量检测和领域中的难点。目前中药生产过程往往通过提前设定的工艺参数进行控制,实时的过程信息却没能被很好地加以应用,使得建立的分析模型不太理想。充分利用生产过程信息,弥补工艺检测模型的信息缺口,采用数据融合从实时过程分析控制的角度充分实现应用是未来中药生产过程质量控制的方向之一。数据融合技术在中药生产加工中主要被用于中药生产过程质量评价及中药制药过程在线质量控制 2 方面。

3.1 中药生产过程质量评价

中药生产过程涉及多方面的影响因素,为保证中药生产产品的批间一致性,提取、浓缩、色谱等

重要工艺环节在中药生产过程质量控制中具有重要作用。通过数据融合技术对中药生产重要过程的质量标志物 (quality marker, Q-Marker) 进行实时分析, 可以有效实现中药生产过程质量评价。张娜等^[46]以中药丹参饮片乙醇提取过程为研究对象, 采用近红外和高效液相色谱实时分析丹参提取液中有效成分的含量变化, 建立基于提取过程中提取液的偏最小二乘质量预测模型 (包括以丹参的指标性成分建立工艺参数过程模型、工艺参数-物料属性融合模型以及工艺参数-物料属性-过程质量参数融合模型)。其中 PLS Toolbox 2.1 结合 MATLAB 7.0 软件的数据分析结果表明多源信息融合模型比常规模型具有更好的校正和预测能力, 数据融合技术可以有效提高提取液质量的可控性。可见近红外光谱可以有效反映提取过程中的工艺状态变化信息, 填补丹参醇提过程中的工艺质量控制的空白。Zhang 等^[47]以小儿消积止咳口服液的提取过程为研究对象, 通过测定提取过程中 7 种关键质量属性浓度, 基于近红外和中红外这 2 种光谱的低级和中级数据融合建立偏最小二乘模型, 其中低级数据融合数据块为简单的原始光谱串联, 中级数据融合为基于波段选择的特征变量相连。MATLAB 2019a 软件的数据分析结果表明数据融合模型对 7 种关键质量属性的预测优于单光谱模型, 可以克服提取中信息检测不全面的检测缺陷, 快速准确的监测小儿消积止咳口服液的提取过程。同时, 面对中药柱层析吸附过程中存在着吸附时间过长导致柱层析产品中存在有效成分比例异常的问题, 蒋程^[48]通过近红外和紫外的光谱低级和中级数据融合分别建立对三七柱层析吸附流出液 5 种皂苷质量浓度的实时定量分析模型, MATLAB 7.5.0 软件的数据分析结果表明, 相比单一光谱建立的模型, 数据融合模型预测误差大幅降低, 可显著增强信噪比, 提升模型准确度和稳健性, 这一方法实现了三七吸附过程的质量实时监控。李建宇^[49]采集纯化过程不同阶段洗脱液的近红外数据, 将光谱与过程工艺参数融合, 采用 SIMCA P⁺ 11.5 和 MATLAB 7.0 软件进行光谱预处理和建模分析, 基于融合数据建立起大孔吸附树脂纯化栀子提取物的过程监测模型。结果显示将经过化学计量学处理后的光谱数据与工艺参数数据融合建模后发现与单纯的工艺参数模型对比, 采用数据融合建模可以在一定程度上提高过程监测模型的性能。综上, 将关键质量属性或工艺参数和仪器检测信息通过数据融合的方式进行

融合建模, 可以充分填补目前中药生产重要过程产品质量无法实时监控分析的检测缺口。因此, 数据融合技术在中药重点工艺环节过程质量评价分析上具有明显的应用优势。

3.2 中药制药过程在线质量控制

在线质量控制是中药生产加工中的重要环节, 而提高数据采集处理的实时性则是满足在线质量检测的要求的一个重要方法。相关研究者采用数据融合技术在这方面进行了研究。以干燥过程为例: 干燥是中药生产中常见的重点操作环节之一, 目前普遍采用取样监测物料含水率的方法进行中药生产干燥过程终点的判断, 这样无法保证对水分的实时控制。现在已经发展的水分在线检测技术可以最大限度实现中药的增效保质。然而依靠单一信息建立的在线分析模型往往不能够全面反映物料含水率的真实状态。数据融合技术可以用来对影响检测目标的因素进行整体分析, 最大限度消除检测过程的不确定因素, 提供准确的数据处理分析结果。彭彦昆等^[50]基于多点同时采集的近红外光谱融合数据, 采用 MATLAB 7.0 软件进行数据分析, 在优化在线检测装备设计和参数的基础上构建了水分在线分级模型, 相比静态离线检测模型, 多点数据融合模型提升了含水率的预测精度。考虑到水分在线检测在实施过程中亦可受干燥环境的温度、湿度等信息影响, 王学成等^[51]指出可以在多点同步检测的同时, 采集温度、湿度、密度等易获取环境参数, 得到综合多维数据融合分析和环境因素的物料含水率预测模型进一步提高检测精度。由此可见, 数据融合技术在中药生产中的在线质量控制上应用前景广阔, 有待进一步深入探索。

4 数据融合用于有效评价中药产品

中药产品的质量检测与评价是对中成药整个生产过程最后把关。目前, 数据融合技术质量检测和包装质量检测中同样具有一定的应用研究。

4.1 中药制剂质量检测

中药制剂质量检测是保障其安全可控的基础, 检测方法目前逐渐由单组分检测转变成多组分和整体特征指纹图谱检测并取得了良好的结果^[52]。Yan 等^[53]采用紫外光谱指纹、多波长融合高效液相指纹图谱及抗氧化活性测定结合, 基于内部开发的指纹超信息特征数字化评估系统 4.0 和 SIMCA 13.0 分析软件的数据分析结果, 从定性和定量的角度成功对 26 批次的六味地黄丸的质量一致性进行评价。2 种

指纹图谱从不同角度反映该药的真实质量,为中药制剂的质量评估提供了一种快速有效的方法。Wang等^[54]首次采用物理化学中常用的燃烧热代表三黄片整体化学信息结合紫外和傅里叶变换红外的光谱定量指纹图谱对30批次不同来源的三黄片的质量一致性进行了评价,数字化评估系统软件的分析结果显示基于单种分析技术评价的黄连质量等级具有较大的差异,从侧面反映采用多信息融合来进行中药质量信息评价的必要性。综上,数据融合通过对多源检测信息的优化组合,成功实现了更准确的中药制剂质量检测结果,为中药制剂质量提供了新的高效检测方案。

为了更加有效的评价质量一致性,需要采用合适方法进行生物学评估。目前,体外溶出度实验已经广泛应用于固体制剂的质量评价中,但在中药制剂中的应用较少。张婧^[55]通过近红外光谱和拉曼光谱对不同来源的盐酸青藤碱缓释片进行溶出行为的模型化研究,研究中采用MATLAB R2018a和SIMCA 14.1数据分析软件,首先通过2种光谱仪器采集盐酸青藤碱缓释片压片前粉末数据,再基于预处理和变量选择确定光谱最优主成分数,最后结合制片处方配比和工艺参数进行多源数据融合建立偏最小二乘和反向传播神经网络模型溶出行为模型。结果显示基于单光谱建立的溶出模型预测能力一般,融合模型可以较好的预测溶出曲线。此研究提出了一种快速有效的中药固体制剂生物学一致的质量一致性评价方法,为中药固体制剂的质量评价提供了新思路。

4.2 中药产品包装质量检测

中药产品包装质量检测是中药产品不可忽略的重要环节。目前,往往采用人工或某些底层的检测装置对产品包装的完成度实现检测,往往缺乏统一固定标准并且存在很强的主观能动性。因此,通过采用一些特定传感器进行包装数据信息获取实现对中药产品的包装检测具有应用前景,不仅可以严格把控流入市场中药产品的合格率,还可以最大限度地提升工厂的生产效率和智造水平。沈正福^[56]基于多传感器的信息融合实现了对产品包装的缺陷检测,降低了基于单一数据源决策的导致的错误率,并基于Python语言采用决策层的融合方法,使用登姆普斯特-谢弗证据理论搭建工位系统数据融合模型,为缺陷检测判别行为的优化提供了方案。

5 数据融合技术的综合评价

综上所述,在面临高精度检测需求时,虽然数据融合技术具有明显的应用优势,但其本身也存在一定的局限性。目前,采用数据融合进行信息数据的处理上缺乏一定的标准化规范,往往依据分析人员的思路或软件设定程序进行操作分析。现在数据融合普遍采用低级、中级、高级数据融合策略进行分析,然而融合层级高的数据融合策略在某些特定的数据和应用场景上可能不能取得比低层级的数据融合更好的结果^[57]。因此在研究中需要依据具体的分析数据和场景选择适宜的融合策略进行分析,避免盲目的“择高效应”。

当采集信息数据量庞大时,应用低级数据融合技术进行检测分析会耗费大量的时间,并且难以满足实时性分析的检测需求。然而,当信息数据量较小、信息维度较低及实时性检测需求不高时,由于低级数据融合处理框架简单、检测精度高,它的应用往往比中级和高级数据融合技术更加快速便捷。处理计算强度大或复杂程度高的数据信息,应首先考虑采用中级数据融合寻找数据中的特征信息,依据部分特征变量分析检测对象可以极大的缩减变量数,减少计算时间,同时满足实时性分析的检测需求。由于决策方法的探索通常会耗费大量的时间,高级数据融合的构建较为复杂,但它的应用可以解决中级数据融合由于缩减变量导致的精度下降问题,适合具有高精度实时分析要求的应用场景。

当然,对于低级数据融合本身存在的检测缺陷,可尝试采用先进的统计学理论知识优化建模,解决由于高数据量导致的计算复杂耗时问题。面对中级数据融合精度下降问题,可以用神经网络及深度学习算法替代传统的机器学习实现建模,通过智能学习实现模型精度优化。高级数据融合由于其本身决策方法探索的困难程度较高、针对性较强,通过迁移学习的方式扩大决策方法的实用性可能是增加其应用场景和检测对象的一个突破点。

6 结语与展望

数据融合技术凭借对传感器信息互补,扩大检测样本信息透明度以实现更充分的目标样本检测的技术特点取得了广泛的应用。在中药原料药材方面,已经成功用于产地鉴别、真伪品鉴定、品质评价3方面的检测研究;在中药生产过程方面,成功应用于生产中重要工艺环节的质量检测评价和制药过程在线控制;在中药终产品质量检测方面,可用于制

剂评价和包装检测等方面。未来,在改进数据融合技术的同时需要加强数据融合技术在中药制药工业全流程多环节的应用研究。在中药材种植方面,可利用多传感器结合,在实现温度、湿度控制的同时,利用数据融合分析技术模拟控制该药材道地产区的生长环境,在农残、重金属限量等方面采用数据融合技术进行全面检测,加强中药材品质控制;在中药生产过程中,继续对各个重点生产环节(包括精制、浓缩、成型等)采用多传感器进行多位点数据采集,通过融合数据分析找到最佳工艺参数,实现过程精准控制;在终产品检测环节,除了利用现代化装备结合数据融合对产品进行化学指标定性定量检测和生物指标评价(如体外溶出度检测),还可以加强对后续包装中异物检测等方面的应用。综上所述,需要进一步针对中药领域特殊需求展开研究,重点使用数据融合技术突破中药生产过程中检测难点,同时与实际应用相结合,完成实验室研究向工厂实际应用研究的转化,提升中药产品质量。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 李天娇,包永睿,王帅,等. 中药质量控制与评价创新方法研究进展及应用 [J]. 中草药, 2022, 53(20): 6319-6327.
- [2] 林兆洲. 中药高维数据融合分析与评价方法研究 [D]. 北京: 北京中医药大学, 2015.
- [3] 王琴琴,沈涛,左智天,等. 基于数据融合和多指标定量对滇龙胆产地鉴别和质量评价 [J]. 中国中药杂志, 2018, 43(6): 1162-1168.
- [4] 孙昱,徐敢,马双成. 中药质量整体评价研究思路探讨 [J]. 药学学报, 2021, 56(7): 1749-1750.
- [5] 王露露,李冰,王圳伊,等. 基于“整体观”系统生物学技术在中药研究中的应用进展 [J]. 中草药, 2020, 51(19): 5053-5064.
- [6] 杨立伟,王海南,耿莲,等. 基于标准汤剂的中药整体质量控制模式探讨 [J]. 中国实验方剂学杂志, 2018, 24(8): 1-6.
- [7] Doeswijk T G, Smilde A K, Hageman J A, et al. On the increase of predictive performance with high-level data fusion [J]. *Anal Chim Acta*, 2011, 705(1/2): 41-47.
- [8] Whittle M, Gillet V J, Willett P, et al. Analysis of data fusion methods in virtual screening: Similarity and group fusion [J]. *J Chem Inf Model*, 2006, 46(6): 2206-2219.
- [9] 王瑞. 开采沉陷监测多源数据融合技术及应用研究 [D]. 徐州: 中国矿业大学, 2022.
- [10] Wang Q Y, Li F S, Jiang X Y, et al. On-stream mineral identification of tailing slurries of tungsten via NIR and XRF data fusion measurement techniques [J]. *Anal Methods*, 2020, 12(25): 3296-3307.
- [11] Lorenz S, Seidel P, Ghamisi P, et al. Multi-sensor spectral imaging of geological samples: A data fusion approach using spatio-spectral feature extraction [J]. *Sensors*, 2019, 19(12): 2787.
- [12] 刘君强,谢吉伟. 基于性能退化数据融合的航空发动机剩余寿命预测方法 [J]. 武汉理工大学学报: 交通科学与工程版, 2020, 44(6): 994-998.
- [13] Yan H S, Zuo H F, Sun J Z, et al. Two-stage degradation assessment and prediction method for aircraft engine based on data fusion [J]. *Int J Aerosp Eng*, 2021, 2021: 1-16.
- [14] Márquez C, López M I, Ruisánchez I, et al. FT-Raman and NIR spectroscopy data fusion strategy for multivariate qualitative analysis of food fraud [J]. *Talanta*, 2016, 161: 80-86.
- [15] Yu H D, Qing L W, Yan D T, et al. Hyperspectral imaging in combination with data fusion for rapid evaluation of tilapia fillet freshness [J]. *Food Chem*, 2021, 348: 129129.
- [16] Barrile V, Simonetti S, Citroni R, et al. Experimenting agriculture 4.0 with sensors: A data fusion approach between remote sensing, UAVs and self-driving tractors [J]. *Sensors*, 2022, 22(20): 7910.
- [17] Fu H Y, Shi Q, Wei L N, et al. Rapid recognition of geoherbals and authenticity of a Chinese herb by data fusion of near-infrared spectroscopy (NIR) and mid-infrared (MIR) spectroscopy combined with chemometrics [J]. *J Spectrosc*, 2019, 2019: 1-9.
- [18] Jiang C, Liu Y, Qu H B. Data fusion strategy based on near infrared spectra and ultraviolet spectra for simultaneous determination of ginsenosides and saccharides in Chinese herbal injection [J]. *Anal Methods*, 2013, 5(17): 4467-4475.
- [19] 孙飞,陈雨,王凯洋,等. 基于红外光谱数据融合的姜半夏鉴别方法研究 [J]. 北京中医药大学学报, 2019, 42(10): 862-868.
- [20] Zhou Y H, Zuo Z T, Xu F R, et al. Origin identification of *Panax notoginseng* by multi-sensor information fusion strategy of infrared spectra combined with random forest [J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2020, 226: 117619.
- [21] Zhou L, Zhang C, Qiu Z J, et al. Information fusion of emerging non-destructive analytical techniques for food quality authentication: A survey [J]. *Trac Trends Anal Chem*, 2020, 127: 115901.

- [22] Willett P. Combination of similarity rankings using data fusion [J]. *J Chem Inf Model*, 2013, 53(1): 1-10.
- [23] Castanedo F. A review of data fusion techniques [J]. *Sci World J*, 2013, 2013: 704504.
- [24] Zhao M, Markiewicz-Keszycska M, Beattie R J, et al. Quantification of calcium in infant formula using laser-induced breakdown spectroscopy (LIBS), Fourier transform mid-infrared (FT-IR) and Raman spectroscopy combined with chemometrics including data fusion [J]. *Food Chem*, 2020, 320: 126639.
- [25] Obisesan K A, Jiménez-Carvelo A M, Cuadros-Rodríguez L, et al. HPLC-UV and HPLC-CAD chromatographic data fusion for the authentication of the geographical origin of palm oil [J]. *Talanta*, 2017, 170: 413-418.
- [26] Borràs E, Ferré J, Boqué R, et al. Data fusion methodologies for food and beverage authentication and quality assessment—A review [J]. *Anal Chim Acta*, 2015, 891: 1-14.
- [27] Xian G M. Multiple microphone speaker recognition system for second language based on biomimetic pattern recognition with big data fusion [J]. *IOP Conf Ser: Mater Sci Eng*, 2020, 790(1): 012144.
- [28] Ihnaini B, Khan M A, Khan T A, et al. A smart healthcare recommendation system for multidisciplinary diabetes patients with data fusion based on deep ensemble learning [J]. *Comput Intell Neurosci*, 2021, 2021: 4243700.
- [29] Fang J Q, Hou H P, Bi Z M, et al. Data fusion in forecasting medical demands based on spectrum of post-earthquake diseases [J]. *J Ind Inf Integr*, 2021, 24: 100235.
- [30] Nachev V G, Titova T P, Kosturkov R D. Instrumental data fusion for food analysis application [J]. *IFAC-PapersOnLine*, 2019, 52(25): 58-63.
- [31] 冯伟, 谭美亭. 基于多源异构数据融合的化工安全风险动态量化评估方法 [J]. *化工管理*, 2022(27): 62-65.
- [32] Zhao L Q, Wang R T, Wang J L, et al. Nonlinear state estimation with delayed measurements using data fusion technique and cubature Kalman filter for chemical processes [J]. *Chem Eng Res Des*, 2019, 141: 502-515.
- [33] 程海军. 保障中药原料质量稳定的环节与因素分析 [J]. *深圳中西医结合杂志*, 2014, 24(3): 180-181.
- [34] 吴雪梅. 数据融合策略结合化学计量学对滇重楼产地溯源研究 [D]. 昆明: 云南中医药大学, 2019.
- [35] Pei Y F, Zuo Z T, Zhang Q Z, et al. Data fusion of Fourier transform mid-infrared (MIR) and near-infrared (NIR) spectroscopies to identify geographical origin of wild *Paris polyphylla* var. *yunnanensis* [J]. *Molecules*, 2019, 24(14): 2559.
- [36] Wang L, Wang Q Q, Wang Y Z, et al. Comparison of geographical traceability of wild and cultivated *Macrophysoria cocos* with different data fusion approaches [J]. *J Anal Methods Chem*, 2021, 2021: 5818999.
- [37] Li Y, Zhang J Y, Wang Y Z. FT-MIR and NIR spectral data fusion: A synergetic strategy for the geographical traceability of *Panax notoginseng* [J]. *Anal Bioanal Chem*, 2018, 410(1): 91-103.
- [38] 王朝勇. 基于多源信息融合的杜仲叶产地溯源及品质评价 [D]. 吉首: 吉首大学, 2020.
- [39] Yao S, Li T, Liu H G, et al. Geographic characterization of *Leccinum rugosiceps* by ultraviolet and infrared spectral fusion [J]. *Anal Lett*, 2017, 50(14): 2257-2269.
- [40] 殷文俊, 茹晨雷, 郑洁, 等. 基于高光谱成像技术融合光谱和图像特征鉴别不同产地的甘草 [J]. *中国中药杂志*, 2021, 46(4): 923-930.
- [41] Yang X D, Song J, Peng L, et al. Improving identification ability of adulterants in powdered *Panax notoginseng* using particle swarm optimization and data fusion [J]. *Infrared Phys Technol*, 2019, 103: 103101.
- [42] 中国药典 [S]. 一部. 2020: 4-402.
- [43] 杨海玲. 基于感官识别-特征成分-机器学习的川芎品质评价研究 [D]. 成都: 成都中医药大学, 2020.
- [44] 毛胜楠. 余甘子立地条件及品质特征的研究 [D]. 成都: 成都中医药大学, 2019.
- [45] Qi L M, Ma Y T, Zhong F R, et al. Comprehensive quality assessment for *Rhizoma Coptidis* based on quantitative and qualitative metabolic profiles using high performance liquid chromatography, Fourier transform near-infrared and Fourier transform mid-infrared combined with multivariate statistical analysis [J]. *J Pharm Biomed Anal*, 2018, 161: 436-443.
- [46] 张娜, 徐冰, 贾帅芸, 等. 丹参提取过程多源信息融合建模方法研究 [J]. *中草药*, 2018, 49(6): 1304-1310.
- [47] Zhang J, Xu X H, Li L, et al. Multi critical quality attributes monitoring of Chinese oral liquid extraction process with a spectral sensor fusion strategy [J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2022, 278: 121317.
- [48] 蒋程. 三七柱层析过程质量控制方法研究 [D]. 杭州: 浙江大学, 2013.
- [49] 李建宇. 梔子提取物大孔树脂纯化过程监测及建模研究 [D]. 北京: 北京中医药大学, 2016.
- [50] 彭彦昆, 杨清华, 王文秀. 基于近红外光谱的猪肉水分在线检测与分级 [J]. *农业机械学报*, 2018, 49(3): 347-353.
- [51] 王学成, 王雅琪, 李远辉, 等. 水分在线检测技术及其在中药干燥领域应用展望 [J]. *中国中药杂志*, 2021,

- 46(1): 41-45.
- [52] Wei X C, Cao B, Luo C H, *et al.* Recent advances of novel technologies for quality consistency assessment of natural herbal medicines and preparations [J]. *Chin Med*, 2020, 15: 56.
- [53] Yan B, Sun G X. Monitoring quality consistency of Liuwei Dihuang Pill by integrating the ultraviolet spectroscopic fingerprint, a multi-wavelength fusion fingerprint method, and antioxidant activities [J]. *J Sep Sci*, 2018, 41(6): 1182-1191.
- [54] Wang Y, Sun G X, Jin Y, *et al.* A comprehensive strategy to quantify the complex system by ultraviolet and infrared spectra analyses coupled with combustion heat for recognizing the quality consistency of San-Huang Tablets [J]. *Am J Anal Chem*, 2017, 8(10): 668-680.
- [55] 张婧. 中药固体制剂溶出行为模型化方法研究 [D]. 天津: 天津中医药大学, 2021.
- [56] 沈正福. 基于机器视觉的尺寸测量与包装缺陷检测工位的设计 [D]. 南京: 南京理工大学, 2020.
- [57] Wang Q Q, Huang H Y, Wang Y Z. Geographical authentication of *Macrohyporia cocos* by a data fusion method combining ultra-fast liquid chromatography and Fourier transform infrared spectroscopy [J]. *Molecules*, 2019, 24(7): 1320.

[责任编辑 赵慧亮]