

智能云平台在芬太尼类似物质谱鉴定中的应用与性能研究

姜延祚, 王晟升, 钱志恒, 段炼, 马乔

(杭州谱育科技发展有限公司, 浙江 杭州 311305)

摘要: 芬太尼类似物因分子结构的高度多样性与快速迭代的特性, 对传统分析检测技术体系构成严峻挑战。现有分析方案存在数据处理流程依赖人工判读、本地数据库覆盖范围有限、跨区域数据共享机制缺失三方面技术瓶颈。本研究提出了一种基于云平台的芬太尼类似物智能检测系统, 搭建了大模型图像与语义理解、特征匹配检索与联邦搜索算法的协同分析框架。该系统通过大模型实现原始质谱数据、检测报告文本及质谱图像等多源信息的结构化特征提取, 并集成一键式智能分析 workflow, 可自动完成从数据预处理到相似化合物匹配的全流程操作。在不同算力配置的云主机环境中, 通过对系统的检索响应速度、资源占用率及并发吞吐量等关键性能指标进行系统性测试与对比, 各项参数均能满足实际分析场景的应用需求。将联邦搜索算法引入鉴定系统架构, 构建云端-客户端协同计算模式, 通过分布式节点的协同检索机制, 为毒品鉴定数据孤岛问题提供了技术方案。本研究实现了检测模式从本地离线分析向现场实时智能识别的范式转变, 为跨区域协同检测网络构建提供了完整的系统级解决方案。

关键词: 芬太尼类似物; 质谱图鉴定; 云平台; 大模型; 联邦搜索

中图分类号: O657.63 文献标志码: A 文章编号: 1004-2997(2026)02-0197-07

DOI: 10.7538/zpxb.2025.0086 CSTR: 32365.14.zpxb.2025.0086

Implementation and Performance Assessment of an Intelligent Cloud-Enabled Mass Spectrometry Recognition System for Fentanyl Analogs

JIANG Yan-zuo, WANG Sheng-sheng, QIAN Zhi-heng, DUAN Lian, MA Qiao

(Hangzhou EXPEC Technology Co., Ltd., Hangzhou 311305, China)

Abstract: The proliferation of fentanyl analogs, marked by their structurally diverse and rapidly evolving molecular frameworks, presents formidable and urgent challenges to traditional detection methods. The existing techniques based on gas chromatography-mass spectrometry (GC-MS) are evidently limited by the subjective judgment of operators. The existing solutions all have obvious shortcomings. First, the manual judgment leads to the low efficiency of the entire data processing workflow. Second, the limited coverage and frequent delay of database make it unable to promptly incorporate newly emerging molecules. Third, the entirely single-point detection fails to form effective linkage, resulting in a severe issue of data silos. To overcome these deficiencies, this study specifically introduced an intelligent detection system based on a scalable cloud platform. The system combines three advanced techniques: large-scale artificial intelligence (AI) models, proprietary high-speed spectral matching algorithms, and an innovative federated retrieval framework. This integration

enables a fully automated, one-click analytical workflow that fundamentally redefines the detection process. It supports multimodal inputs including raw analytical data, analytical reports, and spectral images captured by smartphones. Once data have been input, the system will automatically perform high-precision information extraction, feature analysis, and unambiguous substance identification without manual intervention. This study also conducted comprehensive and systematic testing in order to verify the overall performance of the developed system. Core parameters, including analytical accuracy, processing speed, sample throughput, and computational resource utilization, were quantitatively evaluated in the testing process. The results confirm that the system is fully capable of meeting the diverse requirements of rapid on-site sample testing. Additionally, a series of performance comparison tests were examined on the cloud platform hosts with varied performance levels. These hosts allow users to select servers tailored to specific needs, thus balancing performance and cost-effectiveness. This newly-designed system also incorporates a federated search algorithm, thereby establishing a cloud-client collaborative computing mode for detection tasks. When the algorithm is activated, the central processor searches the main path library, while distributing retrieval and matching tasks to associated clients, then the clients join in the matching and searching process. These processes achieve rapid coverage of newly emerged types of unknown controlled samples—especially under the circumstance that the main library lacks real-time updated sample data. This node-distributed network retrieval model can avoid sharing sensitive proprietary raw data, while breaking down the data silos under secure conditions. However, the operation of the algorithm will increase the overall system retrieval time, so it is necessary to carefully consider the actual needs and use the algorithm selectively. This intelligent system provides an effective solution for transforming the traditional mode of drug analysis into an intelligent multi-point collaborative detection approach, while presenting immediate and powerful tools for curbing the spread of fentanyl analogs. It can also offer some similar solutions for other scenarios related to drug control. This study effectively contributes to the integration of artificial intelligence with analytical chemistry, demonstrating the advent of an upcoming intelligent and data-driven era of analysis.

Key words: fentanyl analogs; mass spectrometry identification; cloud platform; large model; federated search

1960年合成的芬太尼作为强效镇痛剂(镇痛效果约是吗啡的100倍)在临床领域应用广泛^[1]。但近年来,芬太尼类物质作为合成阿片类毒品,因其结构多样性和快速迭代的特性,已成为全球公共安全与药物管制领域的重大威胁^[2-3]。因此,构建高效精准的检测体系是禁毒工作的核心难题。

目前,芬太尼及其类似物的鉴定通常采用色谱、质谱、拉曼光谱等技术。其中,色谱-质谱联用技术因分离和定性定量能力突出,已成为主流检测方法,其标准流程包括样品前处理、仪器分离检测、谱图与标准谱库比对3个关键环节^[4-5]。然而,传统检测方案过度依赖人工判读经验,容易造成漏检与误检,因此有必要借助其他技术手段辅助人工完成检测判定。

近年来,随着计算机技术的不断发展,计算机辅助鉴别技术为弥补上述不足提供了解决方案。Koshute等^[6]提出了一种基于手动筛选特征与机器学习相结合的方法,用于辅助鉴定芬太尼类似物质谱图,达到了较高的准确性。许情等^[7]将随机森林等机器学习模型应用于6类新精神活性药物分类预测场景,证明了人工智能技术在毒品预测问题中具有较大潜力。徐平等^[8]应用连体神经网络与谱库搜索结合的方法,在芬太尼类似物归类上得到了较高的分类精度。

然而,上述工作仍存在明显不足:其一,相关模型多为离线模型,当出现新型样本或需要补充已有样本的谱图信息时,需更新静态模型,无法实现动态优化,智能化程度不足,难以适应毒品

种类的演变;其二,模型训练完成后,对输入形式与数据结构有严格限制,目前研究多聚焦于原始数据层面,对常见的图像、文本报告等形式支持不足,导致缺乏数据处理经验的现场基层人员使用意愿较低;其三,数据孤岛问题突出,不同端侧的客户端难以共享自身数据集,无法借助多源数据的关联信息提升检测鲁棒性,协同效率欠佳,仪器间的数据互通性较差,难以有效联动进行检测。

基于上述不足,本研究提出一种基于多模态大模型的芬太尼类似物智能识别云平台鉴定方案。利用大模型提取并处理原始数据、分析测试报告及相机拍摄质谱图像等多源信息,以提升综合分析能力。引入联邦搜索算法以实现跨机数据检索,破解数据孤岛与共享难题,旨在为芬太尼类似物等毒品鉴定场景提供一套快捷方便智能化解决方案。

1 实验方法

1.1 实验条件

1.1.1 仪器条件 3500系列色谱-质谱联用仪;杭州谱育科技有限公司产品。电子电离(EI)源,电离能70 eV,解吸温度300℃,气质接口温度250℃,进样口温度280℃,传输线温度250℃,离子阱温度150℃,质量扫描范围 m/z 40~400。

1.1.2 云平台条件 云主机A:腾讯云标准型服务器,CPU8核16线程,32GBRAM,50GSSD;云主机B:腾讯云轻量应用服务器,CPU4核4线程,8GBRAM,120GSSD。

1.1.3 软件环境 系统基于Python语言开发测试,涉及到的软件或工具包对应的版本信息列于表1。

1.2 数据处理

本研究数据库共包含274种芬太尼类似物对应的共计2794张质谱图,谱库中一部分来源于公开数据集——缉获药物分析科学工作组数据库(The Scientific Working Group for the Analysis of Seized Drugs, SWGDRUG)和开曼(Cayman)数据库,另一部分为3500系列色谱-质谱联用仪采集得到。数据集中的每条数据保留质量数范围 m/z 41~352,对齐后强度统一归一化至1。对于待测物谱图数据,在送入算法端时也会执行与数据预处理相同的步骤,以确保数据对齐并实现正

表1 开发使用的软件或工具包及其对应的版本信息

Table 1 Software or toolkits used for development and their corresponding version information

软件/工具包 Software/toolkits	版本 Version
Python	3.9.6
Flask	2.2.5
Gunicorn	23.0.0
Pandas	1.3.5
Numpy	1.21.4
Scikit-learn	0.24.1
Matplotlib	3.4.6
Scipy	1.6.0
Openai	1.39.0

确匹配。随机抽取数据集中的50个样本,分析并观测其特征分布,示于图1。可见,每个芬太尼类似物样本均可通过其特有的几个特征峰进行判别。

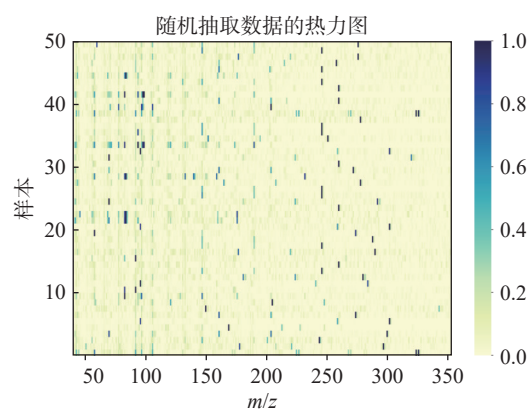


图1 随机抽取数据集中50个样本的数据特征分布

Fig. 1 Data feature distribution of 50 randomly selected samples

1.3 匹配算法建模

本研究在对比测试过程中使用3种谱图鉴定匹配算法,其简要思路介绍和性能优劣列于表2。

1.4 大模型的选取与调用

本研究选取多款典型的国产大模型进行测试,包括通义千问系列(QVQ-72B-Preview、QVQ-Max)、豆包(1.5-vision-pro)、DeepSeek(v12)、智谱清言(glm-4v-Plus)等多家主流模型。这些大模型在处理某些专门领域时具有特定优势,针对本研究涉及的谱图信息提取处理问题,根据测试结果,最终选用通义千问-QVQ-72B-Preview作为调用的接口模型。其他模型,通义-QVQ-max、豆

表2 本研究使用的3种谱图鉴定算法的对比分析

算法名称 Algorithm name	基本介绍 Basic introduction	优劣势分析 Advantages and disadvantages analysis
余弦距离度量 (Cosine)	将质谱图转化为向量,在向量空间内,通过计算向量间余弦夹角衡量谱图相似性。	优势:度量准确性高,容易理解,对谱图整体特征把握良好; 劣势:计算复杂度高,在数据量大的前提下效率低。
局部敏感哈希 (LSH)	利用哈希函数将相似谱图映射到同一桶中,通过桶内检索实现快速匹配。	优势:适合大量数据的快速定性检索,时间复杂度低; 劣势:存在哈希碰撞误差,可能遗漏部分正确谱图。
基于规则的匹配 (Rule)	初筛与细查结合,提取待测谱图与谱库中前5个最强峰进行匹配,若交集特征峰 ≥ 3 ,才会进入后续匹配候选物中进行二次匹配计算。	优势:简单易实现,初筛匹配可过滤大量不相关谱图,匹配速度快; 劣势:仅依赖强峰进行匹配,忽略了谱图总体特征,有可能产生误判。

包模型的精读尚可,但引入了思维链机制,导致生成时间过长,调用成本较高;DeepSeek模型输出信息的稳定性不足,给算法的后端信息提取匹配带来了难题,增加了数据处理复杂程度;而智谱模型整体识别效果欠佳,频繁出现错误,较难满足实际应用场景的需求。综合考虑响应时间、提取信息质量、调用成本等多方面要素,本研究认为QVQ-72B-Preview是当前在谱图信息提取与处理领域性价比方面最优的多模态大模型。

2 结果与讨论

2.1 大模型识别效果

本研究选取一个典型测试图像及其识别结果(质量数大于550的谱峰识别结果受限于图像大小未列出),示于图2。可以看到,该质谱图仅展示了 m/z 信息,未标注具体的对应强度值,采用传统算法(如光学符号识别OCR等)很难估算强度信息,而大模型为此类难以识别的图像提供了一种新的识别范式。

需要注意的是,现有模型对未同时包含 m/z 值与强度信息的谱峰识别能力有限,尤其在处理同位素峰的识别问题时仍存在不足。因此,本研究将识别结果设计为可增删改的编辑模式,用户可根据实际需求对识别结果进行完善;同时,后续工作中,本团队会尝试将此类难以判别的图像作为本地知识库送入模型中进行迭代训练。

除支持上述图像信息外,模型还支持原始数据、实验测试报告等形式,无需用户对输入信息进行预处理和提取。根据多轮测试结果,模型识别单张谱图耗时在30~80s之间(取决于网络状况和输入文件的复杂程度)。对于结构化数据和文本形式这类大模型擅长的输入类型,识别正确率极高。后期随着谱图向量化的完善以及大模

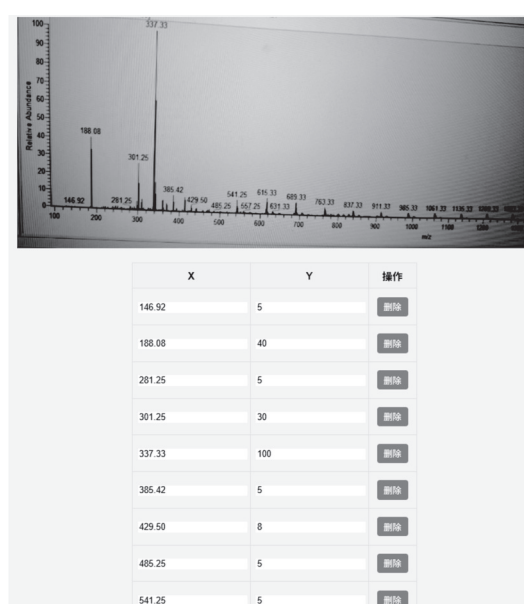


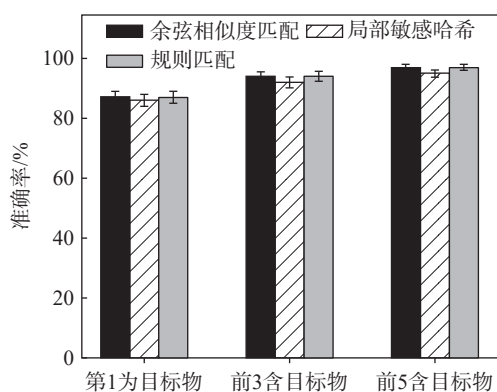
图2 利用大模型识别由手机拍摄的质谱图结果
Fig. 2 Results of recognizing mass spectrum taken by mobile phones using large models

型自身性能的提升,整体提取识别精度将得到进一步提升。

2.2 算法准确性对比

本研究在服务器上对比了所开发的几种匹配检索算法的运行准确性,结果示于图3。

可知,基于余弦匹配相似度的算法准确率略高于基于规则的匹配算法和局部敏感哈希算法,表明余弦相似度匹配在单最佳匹配时更精准,能够通过全局相似性计算定位最优解;受局部敏感哈希算法固有近似性的影响,排在首位(rank1)的目标化合物匹配误差呈小幅增大趋势。随着候选集规模的扩大,各算法的识别正确率均有明显提升,排名前5(rank5)的物质中含有目标物分子的准确性均在94%以上。余弦匹配



注: 基于五折交叉验证

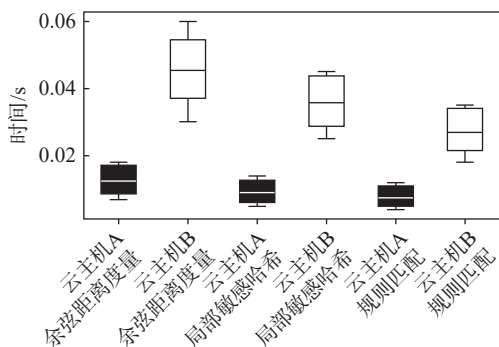
图3 3种匹配算法的准确率对比

Fig. 3 Accuracy comparison of three matching algorithms

算法在匹配中的良好表现得益于其对谱图整体特征的精确建模,能够在多候选中持续输出高匹配度结果。相比之下,基于规则匹配的算法会漏掉部分正确物质,需要使用者在时间成本和匹配精度之间做出权衡考量(目前数据规模不大,影响较小,后期扩充时则需要权衡);局部敏感哈希算法整体表现略差,但在rank5的场景下与另外2个算法基本持平。在检测芬太尼类似物时,鉴于数据集规模整体不会过大,可结合耗时等要素做出取舍,而当数据规模大幅拓展时(如需解决整个新精神活性毒品的鉴定问题),采用遍历匹配的方案显然不是明智的选择。

2.3 检索速度对比

为评估检索算法在云端运行后是否存在较明显的延时现象,本研究对比了几种典型的匹配算法在云端运行的检索速度,结果示于图4。可



注: 5次交叉验证

图4 从云端请求访问算法接口时的检索速度对比

Fig. 4 Comparison of retrieval speeds for requesting access to algorithm interface from cloud platform

知,通过云端请求访问算法时,无论使用何种云主机,响应时间均为毫秒级,但云主机A(黑色箱线)的性能整体显著优于云主机B(白色箱线),且在云端处理和本地运行中均表现更优(如云主机A的余弦距离度量云端耗时中位数0.012s,云主机B为0.047s,比后者快3倍多)。相较于本地匹配检索模式,耗时主要集中于网络传输、资源调度及部署开销,因此在部署过程中建议服务器所处位置与使用地点不宜相距过远,以免通信延迟过高影响使用体验。

2.4 资源占用

为评估云主机在算法访问过程中的CPU利用率、内存占用情况,本研究进行了性能测试。通过监测算法调用时的资源峰值波动与持续占用情况,可有效评估云主机配置是否存在计算资源瓶颈等问题,测试结果示于图5。

在质谱图匹配算法的资源占用测试中,云主机A与B的CPU占用差异显著:A的平均占用率为17.5%,远低于B的51.1%。A的CPU占用集中在低区间且波动幅度小,表明其处理算法时资源消耗低且运行稳定,能高效支持计算密集型任务;B的CPU平均占用率和整体水平较高,表明轻量云因硬件配置较低,处理同类任务时对CPU资源需求更为迫切,更适用于轻量级场景,且不适合多线程运行大型任务。内存占用方面,A和B的平均占用值分别为178.2、166.4MB,差异约为7%。A的内存分布较宽,可容纳更大波动;B的内存分布更紧凑,使用更稳定,但受CPU高消耗限制,整体资源调度偏向“轻量”。两者内存占用值均不大,主要开销在数据集加载、索引构建和匹配结果缓存等方面。当预算充足时,选择高性能云主机可获得更好体验,支持更优调度匹配;而预算有限时,使用轻量云主机也能完成任务,但需注意任务调度,避免CPU占用过高导致死机。

2.5 吞吐量测试

吞吐量是衡量系统数据处理能力的核心指标,本研究针对云端本地运行与客户端远程访问场景进行测试,旨在明确部署模式(本地 vs 远程)与硬件配置差异(中等性能 vs 轻量级)对计算效能的影响。通过量化不同并发线程下的性能表现,可基本了解算法能够处理的并发量上限水平,避免短时间内大量谱图执行查询操作,

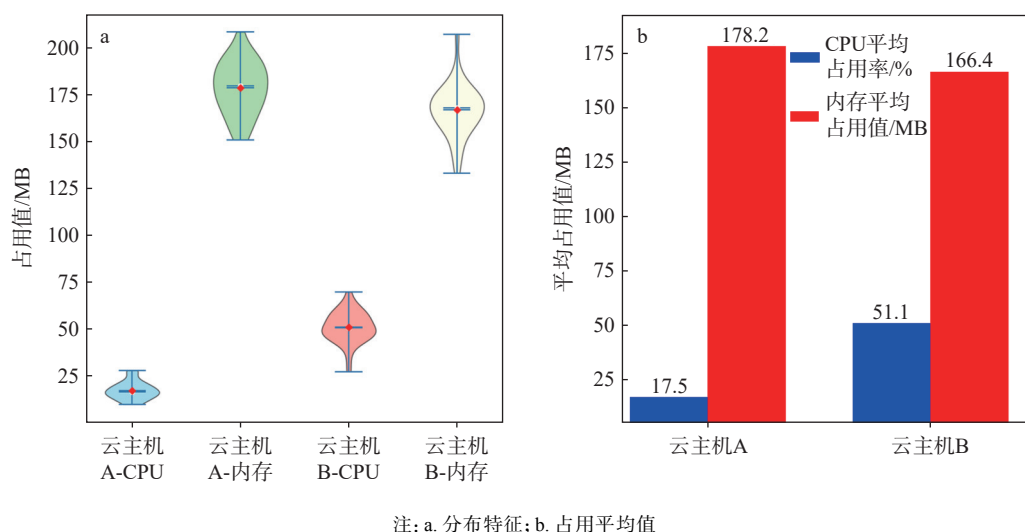


图5 算法运行中云主机的资源占用对比

Fig. 5 Comparison of cloud resource occupancy during algorithm running

导致平台宕机而影响使用,具体测试结果示于图6。

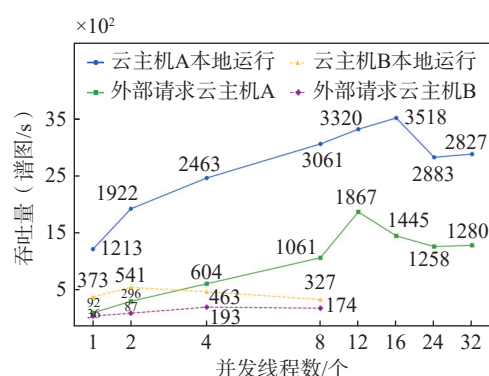


图6 吞吐量对比测试

Fig. 6 Throughput comparison test

在硬件配置维度,云主机A(8核16线程)吞吐量显著优于B(4核4线程):本地运行时,A的吞吐量随并发线程增长持续攀升,并发线程数为16时达到3518谱图/s峰值;B吞吐量始终较低,最高仅为541谱图/s,在高并发场景下增长乏力。该差异源于A的多核多线程架构对并行计算的支撑,可通过线程级并行降低资源竞争;而B受限于CPU核心数量与线程数,并发提升时易陷入资源瓶颈。对比本地与外部请求场景,云端外部访问吞吐量均低于本地(如并发线程数为12时,A本地吞吐量为3320谱图/s,外部请求仅为1867谱图/s),主因是网络开销与跨节点通信成本,包括网络传输、协议解析、中间件转发等环节带来的I/O延迟与服务调度损耗。

2.6 联邦搜索模块

联邦搜索算法采用“云端-客户端”协同检索模式,系统默认检测模式沿用传统服务器端单源检索逻辑,当用户启用联邦搜索功能后,则会同步触发关联客户端的本地数据查询功能。服务器在查询中央数据库的同时,会将查询数据与命令同时下发至与之相连的其他关联客户端进行查询,待完成结果整合汇总后再输出。通过这种分布式检索、计算汇总的模式,能够解决单一服务器的数据覆盖局限,尤其适用于解决中心节点数据缺失场景下的检索完整性问题,例如,部分客户端可能未将采集的谱图数据及时更新至总库。相较传统方案,该方案在异构数据环境中具有更强的适应性。

使用示例示于图7。中央服务器放置于上海市,在南京市和绍兴市分别设有2台与主机关联的客户端。用户在杭州市通过云平台对丁酰芬太尼物质进行谱库检索时,若采用传统的单源检测方案,则获取的排名第一物质为:N-(3,4-亚甲二氧基苯丙胺基)芬太尼,匹配度仅为73.2%,并非正确目标物;但如果系统关联的客户端(如南京市客户端)此前已检测过该物质并更新至本地数据库,启用联邦检索时则能以92.74%的高匹配度,准确识别出分析物为丁酰芬太尼。

但联邦搜索算法存在较为明显的“木桶效应”,即整体响应时间除服务器自身响应外,还依赖于与服务器端相连的其他客户端的响应水平。在其他子客户端完成响应后,云端主机对结

芬太尼质谱图分析系统



The screenshot shows a web interface for the Fentanyl Mass Spectrum Analysis System. At the top, there is a search bar with a file selection button labeled '选择文件 Image.png', a dropdown menu for '服务器计算模型', a checked '联邦检索' button, and an '开始分析' button. Below the search bar, the title '候选物质排名' is displayed above a table with the following data:

排名	中文名称	英文名称	CAS	质量数	匹配度	数据来源
1	丁基芬太尼	Butyryl fentanyl	1443-52-3	350	0.9274	江苏省南京市
2	异丁基芬太尼	Isobutyryl fentanyl	117332-90-8	350	0.8996	江苏省南京市、浙江省绍兴市
3	N-(MDA) 芬太尼	N-(MDA) Fentanyl	-	394	0.7320	上海市(中央服务器)、江苏省南京市、浙江省绍兴市
4	N-(6-APDB) 芬太尼	N-(6-APDB) Fentanyl	-	392	0.7273	上海市(中央服务器)、江苏省南京市、浙江省绍兴市
5	N-(DOC) 芬太尼	N-(DOC) Fentanyl	-	444	0.7260	上海市(中央服务器)、江苏省南京市、浙江省绍兴市

图7 采用联邦搜索可在相连客户端辅助进行数据查找

Fig. 7 Federated search for assisted data retrieval among connected clients

果的加工汇总也需要耗费一定时间。如在本实验中,开启联邦搜索功能后,单谱图平均检索时间由12 ms增至574 ms,系统匹配耗时的增加,主要是因为通讯开销及边缘侧性能有限。因此,将联邦搜索方案设置为可勾选启用的模式更为合理,用户可根据自身需求和关联端的实际状况综合评估后,再决定是否启用该算法。

3 结论

本研究提出了一种基于多模态大模型的芬太尼类似物智能识别云平台鉴定方案:前端输入应用大模型完成信息提取,支持多模态输入,有效降低了分析人员的使用门槛。同时,对开发的云端鉴定系统的准确性、检索时间、吞吐量等性能指标进行表征分析,结果表明,该系统具备物质鉴定准确率高(匹配候选排名前5的物质中含目标物的准确性>94%)、检索时间短(高性能主机单张谱图匹配检索可在12 ms左右完成)、并发处理能力强(高性能主机每秒最多可处理3518张谱图)的优点。所引入的联邦搜索算法为打破数据孤岛和共享难题提供了解决方案。本研究通过融合大模型与云平台,推动检测分析模式从传统实验室离线分析向现场实时智能识别的变革,不仅为禁毒提供了兼具先进性与操作性的技术方案,同时对生命科学、环境监测等领域具有借鉴价值。

参考文献:

[1] ALLAN L, HAYS H, JENSEN N H, de WAROUX B L, BOLT M, DONALD R, KALSO E. Randomised crossover trial of transdermal fentanyl and sustained release oral morphine for treating chronic non-cancer pain[J]. *BMJ*, 2001, 322(7 295): 1 154-1 158.

[2] MOUNTENEY J, GIRAUDON I, DENISSOV G, GRIF-FITHS P. Fentanyls: are we missing the signs? Highly potent and on the rise in Europe[J]. *The International Journal on Drug Policy*, 2015, 26(7): 626-631.

[3] United States Drug Enforcement Administration (DEA). Fentanyl[EB/OL]. (2020-04-01)[2021-03-07]. <https://www.dea.gov/factsheets/fentanyl>.

[4] KANAMORI T, IWATA Y T, TSUJIKAWA K, KUWAYAMA K, YAMAMURO T, SEGAWA H, INOUE H. Simultaneous analysis of 18 compounds of fentanyl and its analogues by TLC, GC/MS and LC/MS[J]. *Japanese Journal of Forensic Science and Technology*, 2016, 21(2): 139-147.

[5] 燕瑾, 花镇东, 贾薇, 刘翠梅. 芬太尼类化合物的质谱特征[J]. *法医学杂志*, 2019, 35(2): 216-223.
YAN Jin, HUA Zhendong, JIA Wei, LIU Cuimei. Mass spectral character of fentanyl analogues[J]. *Journal of Forensic Medicine*, 2019, 35(2): 216-223(in Chinese).

[6] KOSHUTE P, HAGAN N, JAMESON N J. Machine learning model for detecting fentanyl analogs from mass spectra[J]. *Forensic Chemistry*, 2022, 27: 100 379.

[7] 许情, 吕敏, 邓虹霄, 胡驰, 向平, 陈航. 基于电子电离质谱数据和机器学习的新精神活性物质分类预测模型构建[J]. *质谱学报*, 2024, 45(5): 640-646.
XU Qing, LYU Min, DENG Hongxiao, HU Chi, XIANG Ping, CHEN Hang. Construction of prediction models for classification of new psychoactive substances based on EI-MS data and machine learning[J]. *Journal of Chinese Mass Spectrometry Society*, 2024, 45(5): 640-646(in Chinese).

[8] XU P, ZHAO Y, XUE L, LIU Y, YAN M, ZHU L, WENG L, HU S, WEN L. Fentanyl analogs classification via Siamese network and mass spectral library searching[J]. *Expert Systems with Applications*, 2023, 217: 119 534.

(收稿日期: 2025-06-03; 修回日期: 2025-08-15)