

doi: 10.12452/j.fxcxb.26011901

# 基于机器学习模式和气相色谱-质谱联用技术的特征刻画方法分析卡西酮类新精神活性物质

王洋<sup>1</sup>, 陈学国<sup>1,2\*</sup>, 王琳娜<sup>1,2</sup>, 刘春庚<sup>1</sup>, 刘诗怡<sup>1</sup>

(1. 中国刑事警察学院 刑事科学技术学院, 辽宁 沈阳 110854; 2. 辽宁省禁毒技术重点实验室, 辽宁 沈阳 110854)

**摘要:** 为解决新精神活性物质结构多样化、传统检测技术谱库依赖性强、未知物识别弱的问题, 该文基于机器学习模式和气相色谱-质谱联用技术, 以卡西酮类新精神活性物质为例, 选择气相色谱-质谱联用法分析得到的谱图特征为基础, 结合分子结构与质谱分析裂解规律, 通过总结、提炼碎片离子规律, 构建二级分类特征, 借助Python与PySide6, 开发建立了新精神活性物质精准鉴识智能平台, 实现了卡西酮类新精神活性物质的特征刻画。结果显示, 建立的用于刻画卡西酮类新精神活性物质特征的智能分析系统, 模型准确率为98.41%, 精确率为94.03%, 召回率为100%, 特异度为97.88%, F1分数为96.92%, 平衡准确率为98.94%。该方法兼具高识别精度与强可解释性等特点, 不仅弥补了传统检测技术与模型的短板, 而且为新精神活性物质鉴定提供了新的策略。

**关键词:** 卡西酮类; 气相色谱-质谱; 机器学习; 特征刻画; Python

**中图分类号:** O657.7; D919.2 **文献标识码:** A **文章编号:** 1004-4957(XXXX)XX-0001-13

## A Characterization Method for Synthetic Cathinone New Psychoactive Substances Based on Machine Learning and Gas Chromatography-Mass Spectrometry Techniques

WANG Yang<sup>1</sup>, CHEN Xue-guo<sup>1,2\*</sup>, WANG Lin-na<sup>1,2</sup>, LIU Chun-geng<sup>1</sup>, LIU Shi-yi<sup>1</sup>

(1. College of Forensic Science and Technology, Criminal Investigation Police University of China, Shenyang 110854, China; 2. Key Laboratory of Narcotics Control Technology of Liaoning Province, Shenyang 110854, China)

**Abstract:** Aiming to address the challenges posed by the structural diversity of synthetic cathinones, the strong library dependence of traditional detection techniques, and their weak capability in identifying unknown compounds, this study proposed an intelligent analysis system based on machine learning integrated with gas chromatography-mass spectrometry (GC-MS) technology. Leveraging spectral features derived from GC-MS analysis of synthetic cathinones, combined with insights from molecular structures and mass spectrometric fragmentation rules, secondary classification features were constructed by systematically summarizing and refining fragment ion patterns. An intelligent analysis and precise identification platform for the characterization of synthetic cathinones was developed to precisely characterize their features utilizing Python and PySide6. Experimental results demonstrated that the developed intelligent analysis system for characterizing synthetic cathinones achieved a model accuracy of 98.41%, precision of 94.03%, recall of 100%, specificity of 97.88%, F1 score of 96.92%, and balanced accuracy of 98.94%. The approach not only boasted high identification accuracy and strong interpretability, but also effectively overcame the limitations of traditional detection techniques and models, thereby offering a novel strategy for the identification of new psychoactive substances.

**Key words:** cathinones; GC-MS; machine learning; characterization; Python

收稿日期: 2026-01-19; 修回日期: 2026-04-15

基金项目: 辽宁省重点研发项目(2024JH2, 102500079); 中国刑事警察学院科研重点攻关项目(D2025011); 2025年度研究生创新能力提升项目(2025YCYB70); 辽宁省科学技术计划项目(2025080101-JH2/1013); 公安部科技计划项目(2025JSM11)

\* 通讯作者: 陈学国, 博士, 教授, 研究方向: 法医毒物分析, E-mail: chenxueguo@cipuc.edu.cn

作为一种毒品, 卡西酮源自卫矛科巧茶属植物恰特草<sup>[1]</sup>, 是 $\beta$ -酮苯丙胺类似物<sup>[2]</sup>。作为一类重要的新精神活性物质, 合成卡西酮类物质是不法分子以卡西酮为母核合成的多种结构类似的卡西酮类物质, 亦称作“卡西酮类策划药”<sup>[3]</sup>。该类物质通过口服、吸入或注射摄入后, 具有类似苯丙胺、可卡因、摇头丸的毒性作用, 可直接作用于中枢神经系统<sup>[4]</sup>, 诱发滥用者依赖及多种病症<sup>[5]</sup>, 严重威胁社会安全与公众健康<sup>[6]</sup>。

质谱法是检测卡西酮类新精神活性物质的有力工具, 其中电子轰击离子源质谱法(EI-MS)应用广泛<sup>[7]</sup>, 其可与气相色谱法等技术联用进行定性<sup>[8]</sup>、定量分析<sup>[9]</sup>, 具有高灵敏度、高特异性及分析用量少等优势<sup>[10]</sup>。图1为卡西酮类物质的来源、结构与EI-MS分析原理示意图。

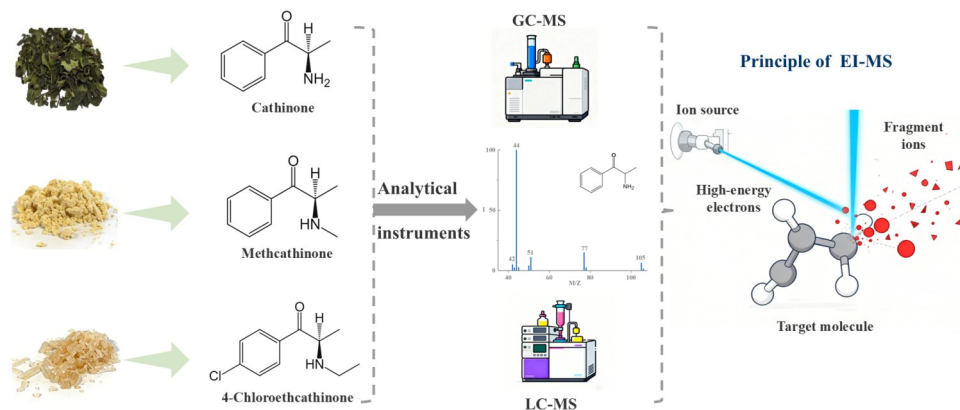


图1 卡西酮类物质: 来源、结构与EI-MS分析原理示意图  
Fig. 1 Cathinones: origin, structure, and EI-MS analytical principle

随着科学技术的快速发展, 质谱技术在卡西酮类新精神活性物质研究中的应用日益广泛<sup>[11]</sup>。Davidson等<sup>[12]</sup>借助多种质谱技术解析了卡西酮类新精神活性物质的裂解机理, 系统研究了 $\alpha$ -吡咯烷苯丙酮类合成卡西酮的裂解行为, 发现不同电离激活方式会诱导该类物质产生特定裂解途径<sup>[13]</sup>, 但规律总结单一。李雅文等<sup>[14]</sup>用便携式质谱仪建立了卡西酮类新精神活性物质的筛查库, 发现该类毒品均有 $\alpha$ 位C-N键裂解, 但过度依赖谱库匹配。Lee等<sup>[15]</sup>基于机器学习模型和相似性搜索算法开发了新精神活性物质分类软件, 对未知管制物质与新精神活性物质识别开展研究, 准确率最高可达94.3%, 但可解释性不足。

本文以包括211种卡西酮类新精神活性物质在内的842种新精神活性物质的EI-MS数据为基础, 结合化合物分子结构与质谱分析裂解规律之间的关系, 通过统计提炼化合物质谱分析碎片特征离子, 构建了二级分类体系, 采用机器学习模式, 结合Python与PySide6, 开发了基于机器学习模式和气相色谱-质谱联用技术的新精神活性物质特征刻画智能系统。与传统方法相比, 本文开发的系统不仅能够不依赖现有谱库的前提下, 对已知卡西酮类物质进行快速确证, 实现对谱库未收录的卡西酮类似物的有效识别, 而且突破了传统机器学习方法仅能分类、无法解析结构的局限, 可进一步对母核与支链结构进行精准定位和预测。另外, 依托PySide6的模块化、脚本化设计, 可实现系统快速更新、迭代与维护, 有效解决了传统谱库更新滞后、维护成本高、扩展能力弱等难题, 为包括卡西酮类在内的新精神活性物质鉴识提供了一种新策略。

## 1 实验部分

### 1.1 数据收集

基于美国国家标准与技术研究院数据库(NIST)、缉获药物分析科学工作组数据库(SWGDRUG), 结合相关文献报道及本实验室测定的卡西酮及非卡西酮类新精神活性物质的气相色谱-质谱联用分析数据, 共收集了211个卡西酮类新精神活性物质、631个其他类新精神活性物质的EI-MS数据。

### 1.2 数据预处理

按照收集的卡西酮类新精神活性物质的化学结构特征, 赋予相应的标签, 0代表非卡西酮类新精神活性物质, 包含631个化合物; 1代表卡西酮类新精神活性物质, 包含211个化合物, 其中10代表结

构与卡西酮类似的化合物，包含 5 个样本；11 代表结构与甲卡西酮类似的化合物，包含 26 个样本；12 代表结构与乙卡西酮类似的化合物，包含 20 个样本；13 代表含亚甲二氧基结构的卡西酮类新精神活性物质，包含 34 个样本。将所有数据集按照分层抽样法，以 7: 3 的比例划分为训练集与验证集。

### 1.3 特征提取

EI-MS 数据与物质的相对分子量、结构有关，碎片离子质荷比和丰度值可作为结构特异性标识。卡西酮类新精神活性物质样本特征提取时，选取丰度值较大的 10 组碎片离子作为分析对象，按照离子丰度值从高到低的顺序，将碎片离子依次标记为 MS1~MS10，并统一数据格式，供后续构建特征矩阵使用。

### 1.4 一级分类标准

首先需要确定是否为卡西酮类物质。本文从多维度进行判定，其中“氮规则<sup>[16]</sup>”用于确定化合物分子量为奇数时其分子中氮原子数必为奇数；如图 2 所示，卡西酮类物质中结构最简单的卡西酮可用于确定分子量下限；分子稳定性与空间位阻限定上限；EI-MS 分析常用 70 eV 电离能远高于多数共价键键能，可促使分子碎裂充分，且卡西酮中的特征活性位点  $\alpha$  位 C-N 键断裂后失去的含氮片段也多为低质荷比离子<sup>[17]</sup>；丰度最强的碎片离子（MS1）则可最大限度保留物质关键结构信息。因此，结合卡西酮类新精神活性物质特征，确立样本的一级分类标准如下：

(1) 分子量特征：相对分子质量 ( $M$ ) 为奇数，且  $163 \leq M \leq 317$ ，其中 163 为天然卡西酮的分子量，317 为已知最大卡西酮类物质的分子量上限；

(2) 碎片离子下限：离子丰度值前 10 的碎片离子中，必存在至少 1 个质荷比  $\leq 77$  的碎片；

(3) 碎片离子上限：离子丰度值前 10 强的碎片离子中，最大质荷比  $\leq 206$ ；

(4) 离子丰度值最高的碎片（MS1）特征：符合“1.5 二级分类特征”中的任意一类支链规律；

(5) 辅助验证条件：离子丰度值前 10 的碎片离子中，不存在质荷比 188、189、175、177、136、153 的碎片离子。

### 1.5 二级分类特征

基于训练集中的 EI-MS 数据，在一级分类判定基础上，结合卡西酮类物质的  $\beta$ -酮安非他明分子结构特性，按照“理论推导-样本验证-规律提炼”的逻辑，构建了卡西酮类物质结构解析图，见图 2。

卡西酮类物质的二级分类中的母核取代基解析以化学键特性与样本验证为核心逻辑进行。图 3 为含亚甲二氧基取代的卡西酮类物质 EI-MS 分析可能的裂解途径。从图 3 中可以看出，亚甲二氧基 ( $-O-CH_2-O-$ ) 与苯环的共轭体系在 EI 电离下易发生 C-O 键断裂，生成含苯环与亚甲二氧基片段的质荷比为 149 的特征离子，再对训练集中所有含该取代基的样本进行碎片离子逐一核对、统计，发现该类母核均存在质荷比为 149 的碎片离子，据此完成“理论裂解-样本验证-统计分析”，确立专属对应关系，其余母核按此范式验证并得到分类结果。

卡西酮类物质的支链结构解析基于离子丰度值前 2 的碎片离子（MS1、MS2）进行。因在 EI 电离下，卡西酮类物质优先经  $\alpha$  位 C-N 键或羰基  $\alpha$  位 C-C 键断裂生成含氮片段或烷基链的碎片离子，离子丰度值较高，居前 2 位，兼具强结构代表性与低干扰优势。本文通过对训练集样本 MS1、MS2 的逐一归属，结合支链碳链长度、环系类型及取代位置，将卡西酮类物质支链结构划分为 5 大类。表 1 为卡西酮类物质的支链类型与特征离子对照表。

基于部分卡西酮子类物质的 EI-MS 数据，选取离子丰度值前 10 的碎片离子，结合对应子类标签，构建的三维特征分布可视化图见图 4。由图 4 可以看出，不同卡西酮子类的特征离子呈现明显的聚类趋势，子类间的特征离子差异与表 1 中的支链结构规律一致：如卡西酮的 MS1 集中在 44，甲卡西酮对应 MS1 为 58，乙卡西酮对应 MS1 为 72、MS2 为 44。

### 1.6 卡西酮类新精神活性物质精准鉴识平台开发

结合卡西酮类新精神活性物质的 EI-MS 数据特征与规律，基于 Python（版本 3.13，2024）与 PySide6 开发库，构建了卡西酮类新精神活性物质精准鉴识平台。该平台的工作流程图如图 5 所示。

在卡西酮类新精神活性物质精准鉴识平台的模型模块中，首先使用 pandas 库完成 Excel 格式数据的

读取,同时自动校验数据是否包含“编号”“分子量”“MS”“丰度值”等必要信息,若存在信息缺失,系统会即时弹出报错窗口提示补充;随后,依据“1.4”确立的一级分类标准,通过if not条件判断语句对输入数据进行前置校验,确保后续分析的基础数据符合要求。完成数据校验后,谱库匹配环节将通过for循环语句遍历库中所有已收录数据,直至分子量及前10强碎片离子的质荷比数值比对成功,若是浮点数则采用近似匹配规则,若是整数则执行完全匹配。对于经一级分类判定为卡西酮类、但未通过谱库匹配的未知物质,则参照“1.5”的结构规律,将预设的母核与支链特征离子标记为特征点,构建特征分析数据集,例如:质荷比149对应母核的亚甲二氧基结构,MS1=44,MS2=100对应支链R<sub>3</sub>的叔丁基取代;随后,逐点比对待分析物质的碎片离子,统计并记录所有匹配的特征点,最终输出匹配特征点列表与匹配比例,实现对未知卡西酮类物质结构的推断。

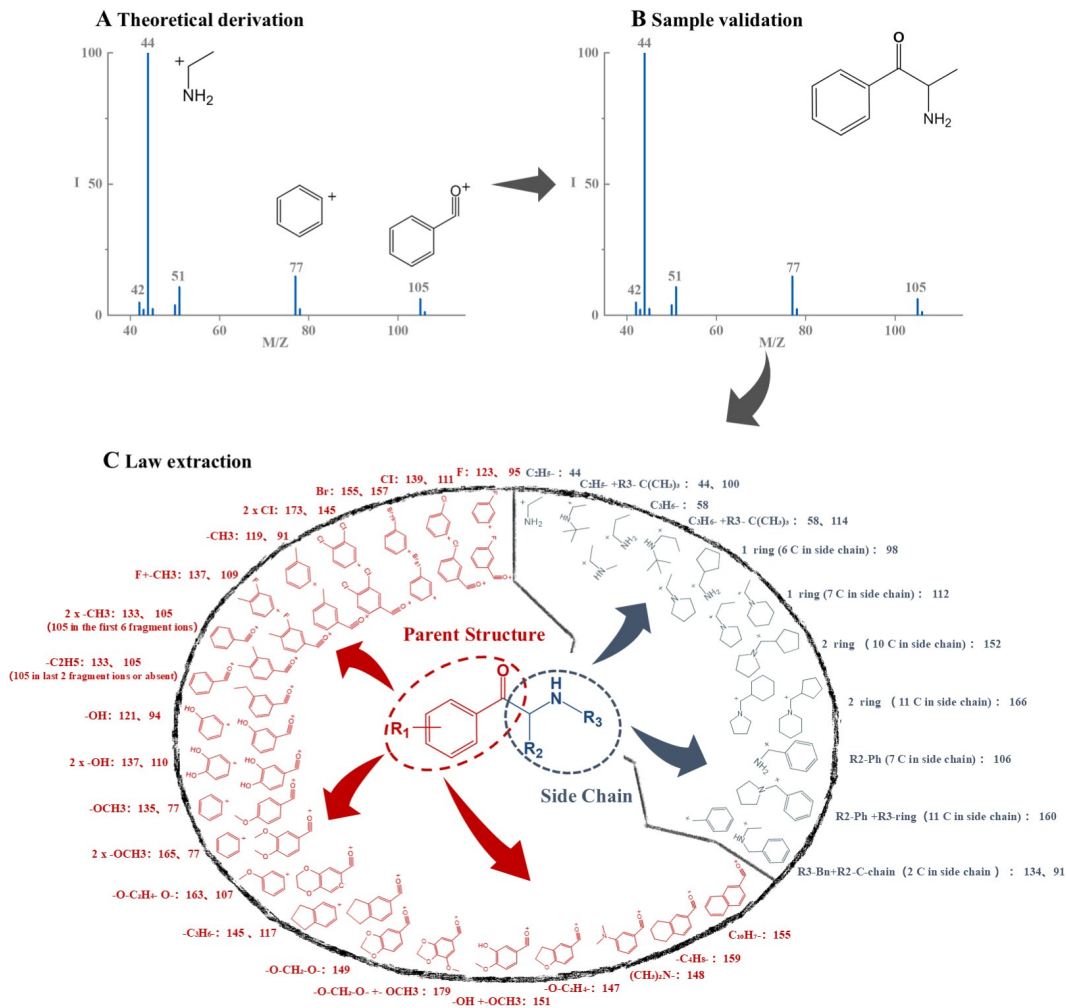


图2 卡西酮类物质结构解析图

Fig. 2 Structural analysis diagram of cathinones

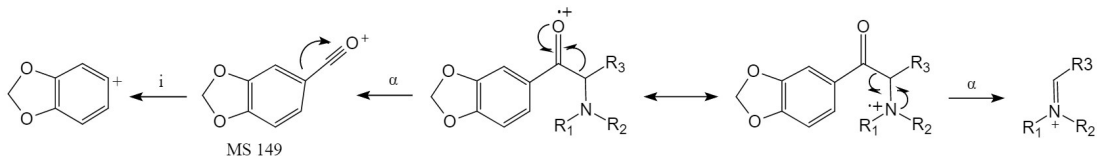


图3 含亚甲二氧基卡西酮可能的裂解途径

Fig. 3 Possible fragmentation pathways of cathinones containing methylenedioxy group

在对比模型部分,先使用train\_test\_split函数按照7:3的比例将预处理后的数据集划分为训练集和测试集,随后基于scikit-learn库(版本1.3.0)分别调用支持向量机(SVM)、K近邻算法(KNN)与随机森林(RF)3种传统模型:其中SVM模型选择poly作为核函数,设置惩罚参数C=100.0、gamma

参数为“scale”；KNN模型设定近邻数 $n\_neighbors=7$ 、距离度量参数 $p=1$ ，并采用“distance”加权方式；Random Forest模型则保持最大深度 $max\_depth$ 为None，最小分裂样本数 $min\_samples\_split=2$ ，同时将决策树数量 $n\_estimators$ 设为300。还通过xgboost库调用XGBoost模型，设置迭代次数 $n\_estimators=200$ 、最大深度 $max\_depth=3$ 、学习率 $learning\_rate=0.05$ ，以及通过lightgbm库调用LightGBM模型，设置叶节点数 $num\_leaves=63$ 、学习率 $learning\_rate=0.05$ ；为优化模型性能，对传统模型采用网格搜索法，对集成模型采用随机搜索法，结合5折交叉验证完成参数调优，代码参见Github链接<https://github.com/Dylan111233/A-Classification-Model-for-Cathinone-Substances>。

表1 卡西酮类物质的支链类型与特征离子对照表

Table 1 Comparison table of side chain types and characteristic ions for cathinones

Side chain	Characteristic ion (MS) pattern	Corresponding structure	Example
普通碳链型	①MS1=44, 58, 72... (+14×n, n为非负整数, 下同)	基础碳链取代	甲卡西酮, MS1=58
	②MS1=44, MS2=100; MS1=58, MS2=114... (+14×n)	R <sub>3</sub> 为叔丁基取代	3-氯叔丁基卡西酮, MS1=44, MS2=100
单饱和环型	MS1=98, 112, 126... (+14×n)	支链含1个饱和环	吡咯丙酮, MS1=98
双饱和环型	MS1=152, 166, 180... (+14×n)	支链含2个饱和环	2-(吡咯烷基)-1-苯基-2-(环己基)乙酮, MS1=166
单苯环型	① MS1=106, 120, 134... (+14×n)	支链R <sub>2</sub> 为苯环, R <sub>3</sub> 为碳链	1, 2-二苯基-2-氨基-1-乙酮, MS1=120
	② MS1=91	支链R <sub>2</sub> 为碳链, R <sub>3</sub> 为苄基	苄基卡西酮, MS1=91
苯环-饱和环混合型	MS1=160, 174, 188... (+14×n)	支链含1个苯环、1个饱和环	1, 2-二苯基-2-(1-吡咯烷基)乙酮, MS1=160

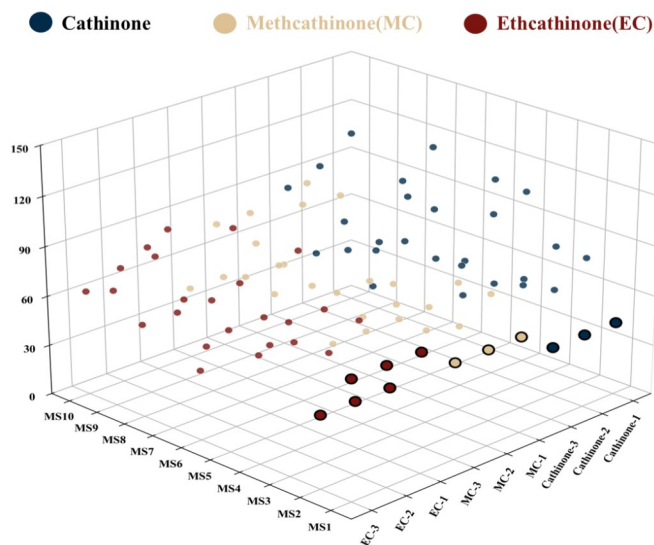


图4 常见卡西酮子类特征碎片离子散点图

Fig. 4 scatter plot of characteristic fragment ions for common cathinone subclasses

平台的可视化交互界面则依托PySide6库（版本6.6.0）开发，界面中通过QTableWidget控件清晰展示待分析物质的离子丰度值前10的碎片离子信息，同时借助QPushButton控件设置“数据导入”“结构预测”等功能按钮，用户点击对应按钮即可触发相应操作；该界面支持Windows10/11、macOS13及以上版本、Linux Ubuntu 22.04等主流操作系统，保障跨平台稳定运行，便于不同环境下的用户快速使用。

## 2 结果和讨论

### 2.1 模型结果与评估

为了从多个维度综合衡量模型的各项性能指标，本文采用混淆矩阵、准确率（Accuracy）、精确率（Precision）、召回率（Recall）、F1分数（F1-score）、特异度（Specificity）及平衡准确率（Balanced accuracy）等参数评估模型，计算方法见式（1）~（6），其中TP代表真阳性，FP代表假阳性，TN代表真阴性，FN代表假阴性。

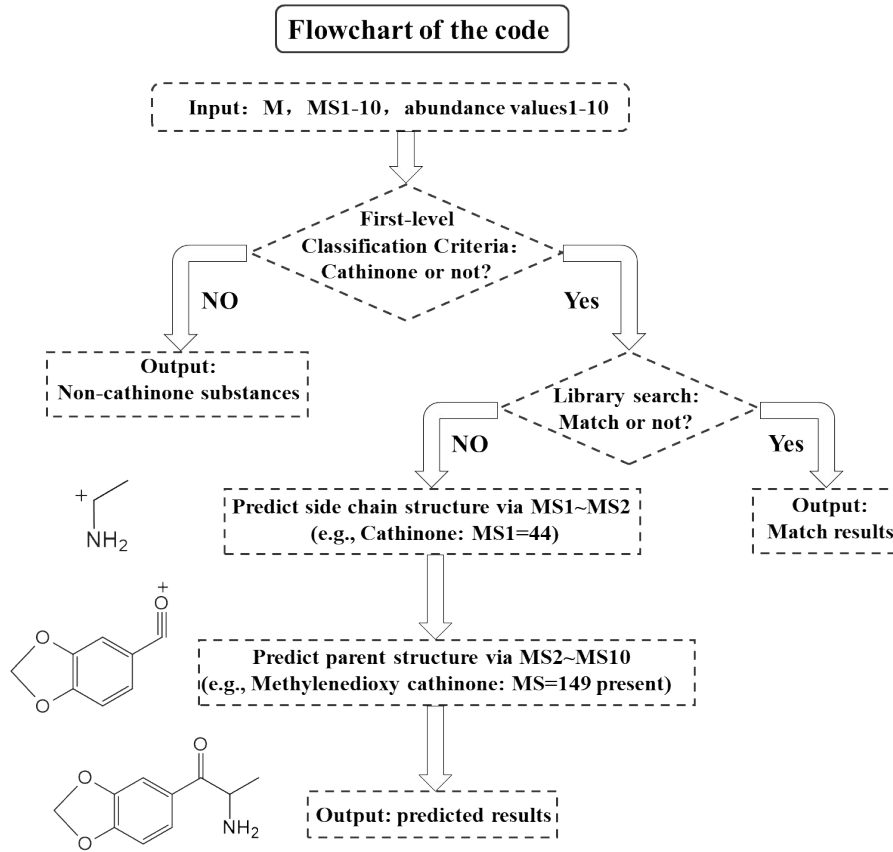


图5 卡西酮类物质结构判定程序框图

Fig. 5 Flowchart for structure determination of cathinone substances

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 \text{ score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

$$Specificity = \frac{TN}{TN + FP} \quad (5)$$

$$Balanced \text{ Accuracy} = \frac{1}{2} \times (Recall + Specificity) \quad (6)$$

图6为不同模型在卡西酮类物质鉴识任务中的混淆矩阵可视化结果,是基于63个卡西酮类样本与189个非卡西酮类样本的统一验证集预测结果统计绘制,每个子图对应不同模型的分类表现,矩阵内的数值直观反映了各模型真阳性、假阳性、真阴性、假阴性的数量分布。从图6可以看出,KNN、SVM等传统模型均存在不同程度的漏判或误判,而卡西酮精准鉴识模型(PIM-C)的表现尤为突出,对验证集中的63个卡西酮类样本实现了100%的判定准确率,达成零漏判,针对189个非卡西酮类样本也仅出现4例误判;结合量化性能指标来看,该模型的整体分类准确率为98.41%、精确率为94.03%、召回率高达100%、特异度为97.88%、F1分数为96.92%、平衡准确率为98.94%,各项指标均处于较高

水平，充分验证了该模型的科学性、应用价值与鲁棒性。

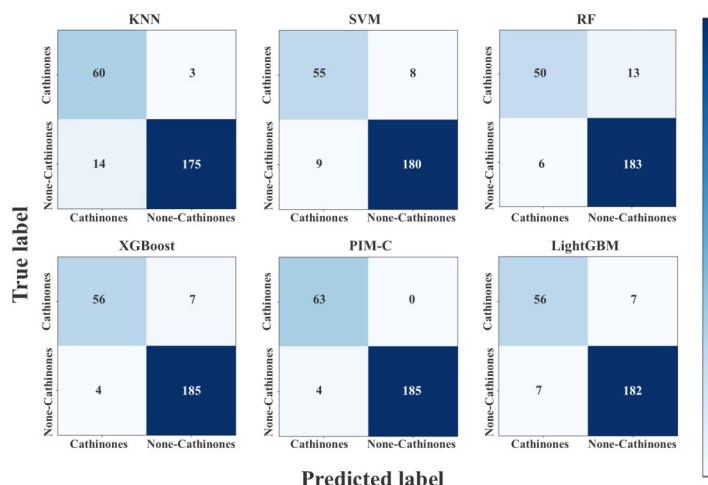


图 6 各模型的卡西酮类物质判定混淆矩阵对比图

Fig. 6 Confusion matrix comparison for cathinones with various models

基于统一验证集，涵盖准确率、精确率、召回率、F1 分数、特异度、平衡准确率 6 项核心指标绘制的不同分类模型在卡西酮类物质鉴识任务中的性能指标条形图对比结果见图 7，图中不同颜色的条形分别对应 KNN、SVM、RF、XGBoost、PIM-C、LightGBM 模型。由图 7 可知，PIM-C 在各项指标中均呈现明显优势，是唯一在召回率指标上达到 100% 且其余指标均处于较高水平的模型。相较之下，KNN 的精确率不足 85%，对目标类（卡西酮类）的识别准确性欠佳；SVM 的召回率、F1 分数、精确率均低于 90%，整体表现偏弱；Random Forest 的召回率不足 80%，对真实卡西酮类样本的检出能力较弱；XGBoost 与 LightGBM 的表现虽优于传统模型，但平衡准确率等指标未达最优。进一步分析：KNN 的短板或是因为卡西酮类与部分非卡西酮类样本特征空间重叠，干扰了近邻选择；SVM 的不足推测是 poly 核函数对特征尺度敏感导致的拟合偏差；RF 的召回率偏低，是因集成过程对少数类样本关注度不足、决策阈值偏高，造成漏判较多；XGBoost 与 LightGBM 虽性能更优，但平衡准确率不足，反映出其在类别不平衡场景下的均衡分类能力有限，且“黑箱”特性使其可解释性不及基于明确结构特征规则构建的 PIM-C 模型。

## 2.2 卡西酮类 NPS 自动识别平台建设

平台以“数据支撑-谱库匹配-智能预测”为核心逻辑，采用模块化设计，由数据查看模块、谱库匹配模块和结构预测模块等三大核心功能单元组成，集成了数据处理、已知物质快速识别、未知物质结构预测等全流程功能，实现了从数据导入到结果输出的自动化、智能化鉴识。

**2.2.1 数据查看模块** 本模块为软件数据输入与基础查看入口，旨在接收、核对、解析并可视化待分析数据：用户点击左侧顶端按钮，导入 Excel 格式的待分析物质数据，数据将显示于左侧表格；點選表格对应行时，右侧同步展示该物质前十强质谱数据的质谱图及表格形式的数据信息；当界面呈现需查询、预测结构的目标物质时，点击右下角按钮可跳转至相应界面；若导入数据存在缺失或错误，系统将弹窗提醒以避免误差（见图 8）。

**2.2.2 谱库匹配模块** 本模块为谱库匹配与一级分类标准执行部分，旨在实现已知卡西酮类物质的快速识别与未知物质的初步判定：待分析数据将与谱库内已知物质逐一精准校验，通过相对分子量与前 10 强碎片离子质荷比的双重完全对应，确保已收录物质 100% 识别。成功匹配后，展示该物质名称、结构、标准谱图等信息；反之，将弹窗提示“谱库中未查找到该物质，是否进行结构预测？”。點選确定后，执行一级分类判定：通过即跳转至后续分析流程，未通过则弹窗提示“该物质为非卡西酮类物质”（见图 9）。

**2.2.3 结构预测模块** 本模块针对谱库未匹配但经一级判定为卡西酮类的未知化合物，按“先母核、后支链”逻辑分步预测：母核预测基于特征碎片离子，例如亚甲二氧基母核 MS=149，比对筛选出匹配度前三名的候选母核；支链预测依据丰度最强的 MS1、MS2 碎片离子，筛选匹配度前三名的候选支链。

结果展示于界面右侧表格,表格中标注匹配度,且点击结果时数据会以变色标注,辅助用户验证预测合理性(见图10)。

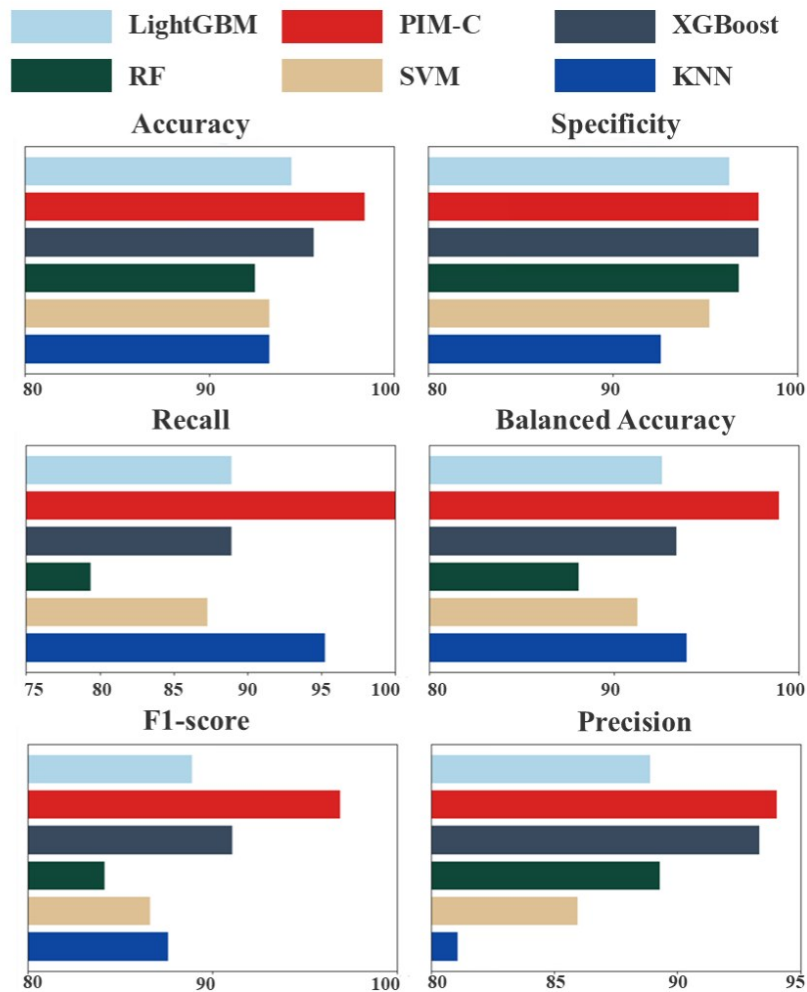


图7 各模型卡西酮类物质的判定多参数性能对比图

Fig. 7 Multiple parameters comparison of various models for cathinones determination

### 2.3 实例验证

为验证所构建智能系统的实战适用性与可靠性,采用QP 2020 NX型气相色谱-质谱联用仪,参照GA/T 1920-2021《疑似毒品中211种麻醉药品和精神药品检验 气相色谱-质谱法》方法<sup>[18]</sup>,对XX市公安局查获的疑似毒品中2种卡西酮类物质(4-氯乙卡西酮、2-(乙基氨基)-1-苯基-1-戊酮)进行验证测试。

**2.3.1 已知毒品检测** 将4-氯乙卡西酮样品按上述方法进行进样分析,获得其EI-MS质谱信息。提取该物质质谱图中离子相对丰度排名前10的特征碎片离子质荷比及对应丰度数据,导入本研究开发的智能识别系统,系统输出识别结果界面如图11所示。经系统自动匹配与算法判别,可快速、精准识别目标物质为4-氯乙卡西酮,检测结果与样品实际结构完全吻合。该结果表明,本系统对已知卡西酮类物质具有良好的快速检测与精准确证能力。

**2.3.2 未知毒品检测** 以实际缉获的未知疑似毒品样品(主要成分为2-(乙基氨基)-1-苯基-1-戊酮)为研究对象,按上述相同仪器条件与前处理方法进样检测,获得其EI-MS质谱信息。首先,采用NIST 2.0现有谱库对其进行检索识别,未获得有效匹配结果,无法实现结构判定,检索结果见图12。然后,将该品质谱图中相对数据导入本文开发的智能系统,进行结构推测分析,系统输出的识别与结构预测结果如图13所示。结果表明,本文开发的系统可突破传统谱库仅能识别已知物质的局限,实现对新型、未收录卡西酮类物质的有效判别与结构预测。

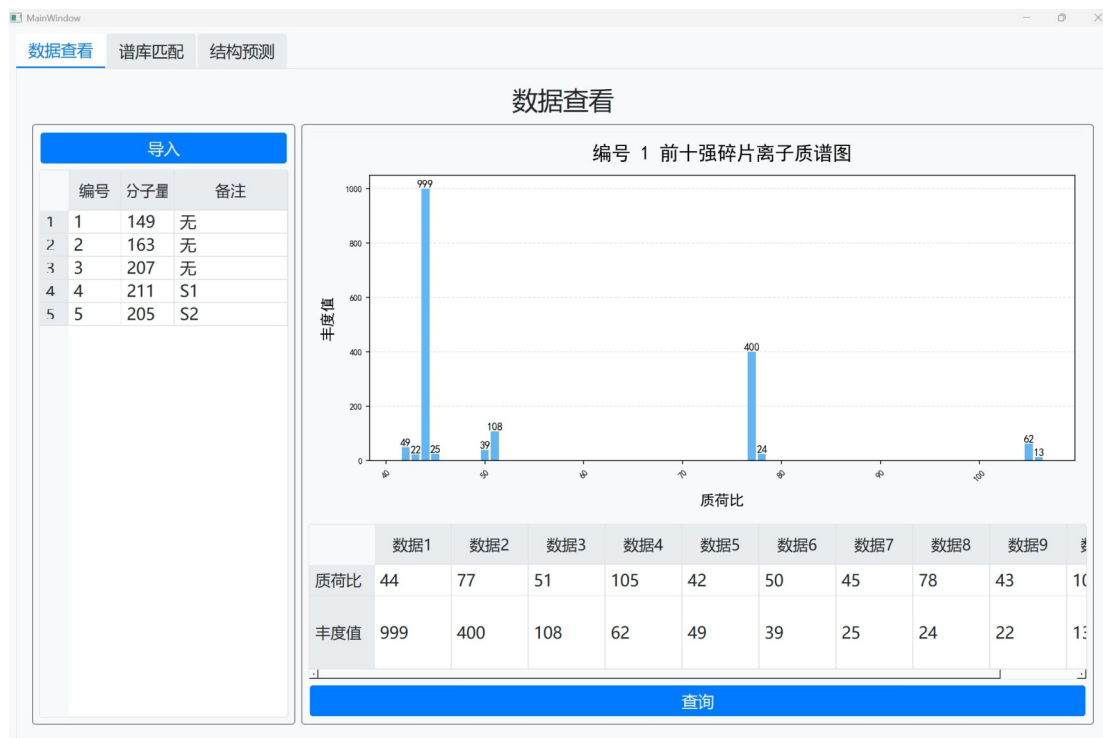


图8 卡西酮类物质自动识别平台的数据查看模块

Fig. 8 Data viewer of the automatic identification platform for cathinones

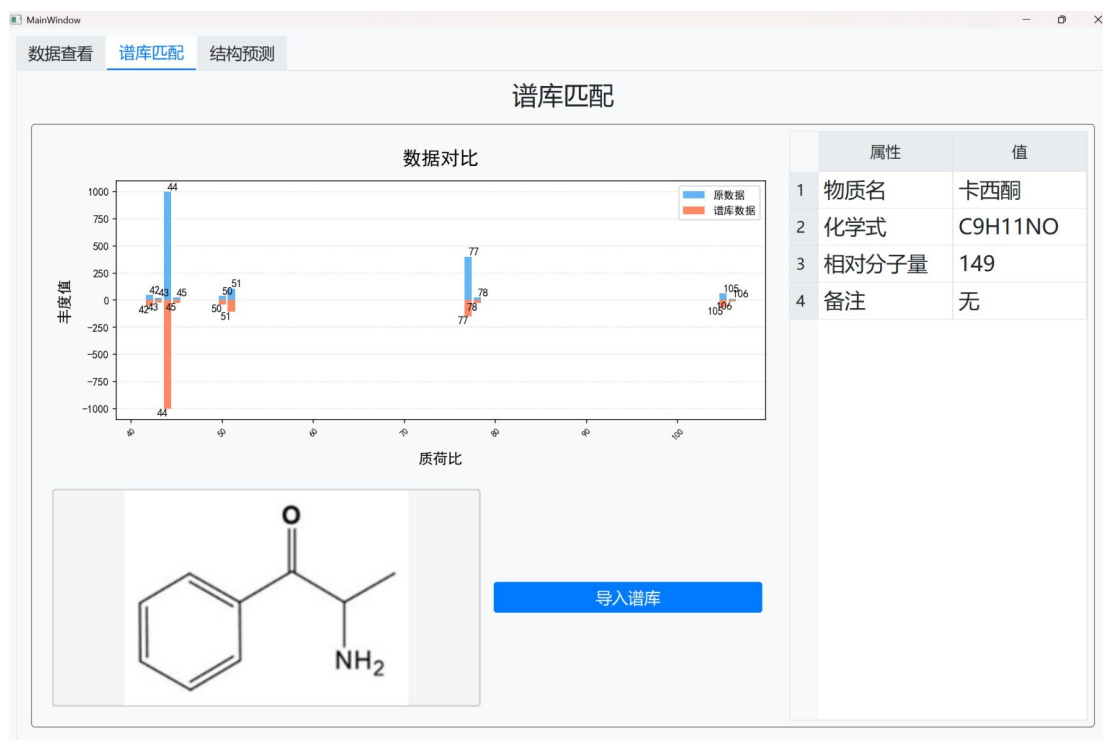


图9 卡西酮类物质自动识别平台的谱库匹配模块

Fig. 9 Library matcher of the automatic identification platform for cathinones

### 3 结 论

本文以 842 个化合物的 EI-MS 数据为基础，按“裂解推导-样本验证-统计学提炼”逻辑提炼碎片离子规律构建二级分类体系，开发 Python-PySide6 智能预测系统，导入待检物质相对分子量、前十强碎片离子质荷比及丰度值即可快速分析预测结构。本文建立的用于刻画卡西酮类新精神活性物质特征的智能分析系统在验证集与分类任务中均表现优异，较 SVM、KNN 等机器学习模型优势显著：不仅能

够摆脱对现有谱库的依赖,同时实现已知卡西酮类物质检测和未知的新型卡西酮类似物识别,而且突破了传统机器学习仅能分类、无法解析结构的局限,可对未知卡西酮类物质的母核与支链结构进行定位推测。此外,基于 Python-PySide6 的设计使系统的更新、迭代与维护更加便捷,避免了传统谱库更新滞后、维护成本高、扩展能力弱的问题,且支持跨平台使用,具有更强的适用性与拓展性,为包括卡西酮类在内的新精神活性物质的快速检测提供了新策略。未来,相关研究将持续扩充化合物数据库,纳入更多新型类似物数据,并依托系统模块化架构,将开发的方法推广至更多类别的新精神活性物质鉴别,不断提升模型的泛化能力与检测覆盖范围,以更好满足一线实战中日益复杂的检测需求。

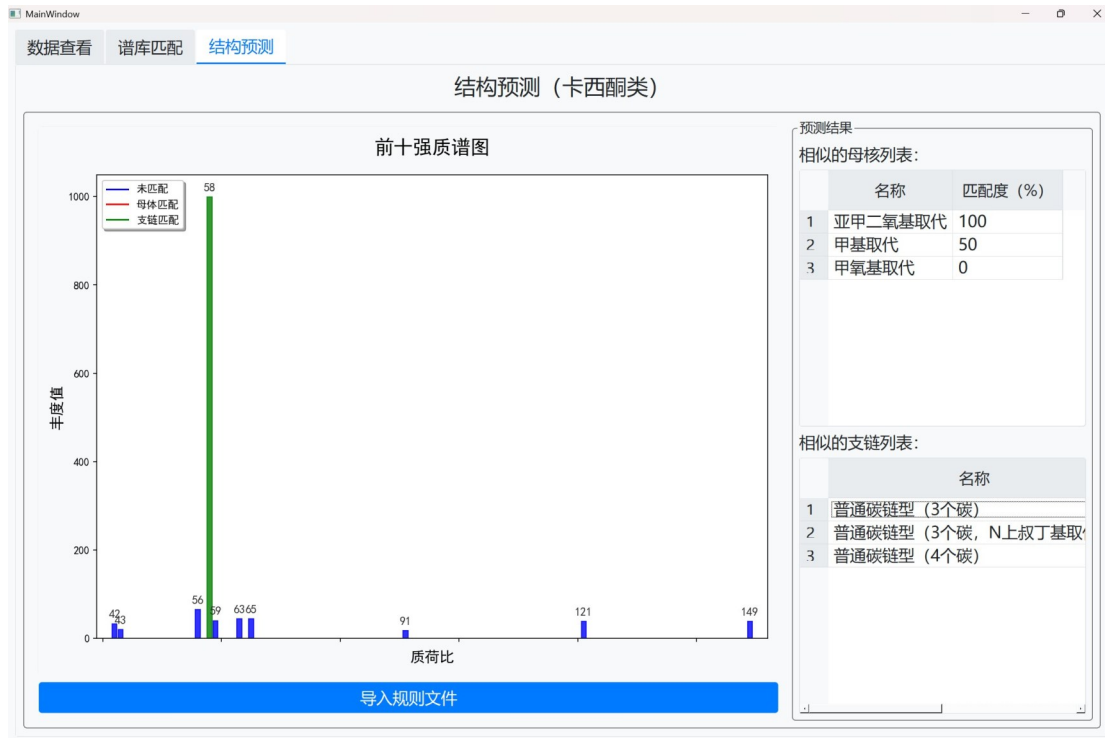


图 10 卡西酮类物质自动识别平台的结构预测模块

Fig. 10 Structure predictor of the automatic identification platform for cathinones

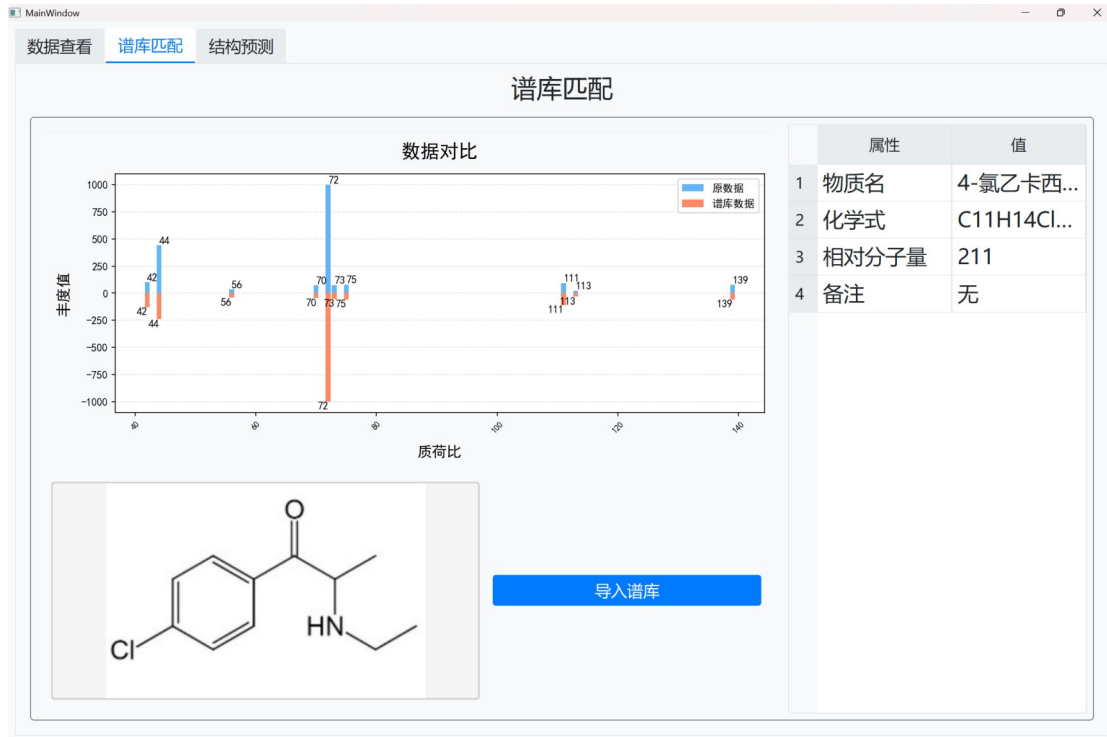


图 11 4-氯乙卡西酮的识别结果图  
Fig. 11 Identification result of 4-chloroethcathinone

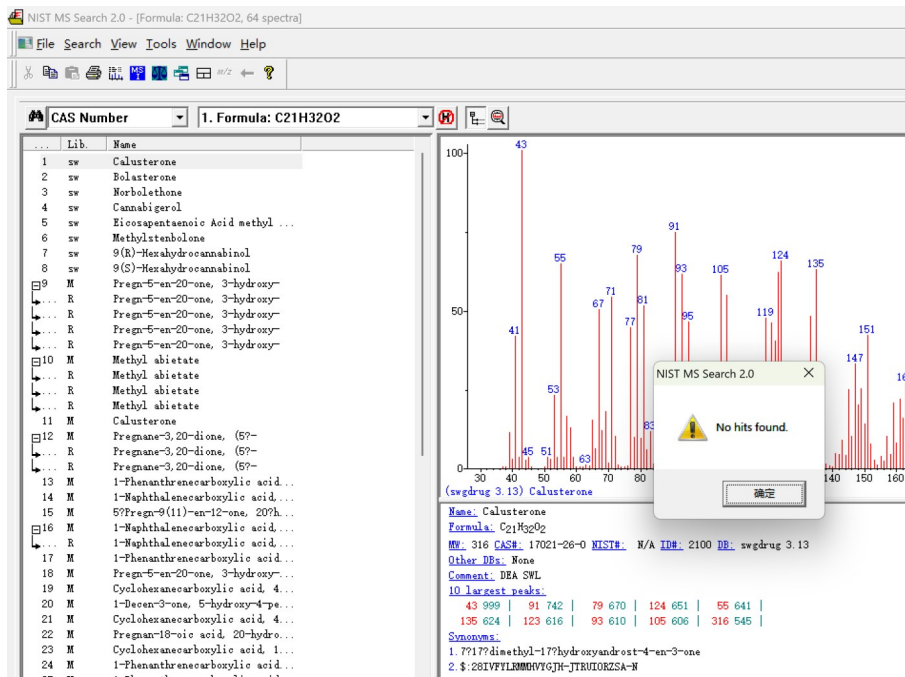


图 12 NIST 2.0 谱库识别结果  
Fig. 12 Identification result of NIST 2.0 library

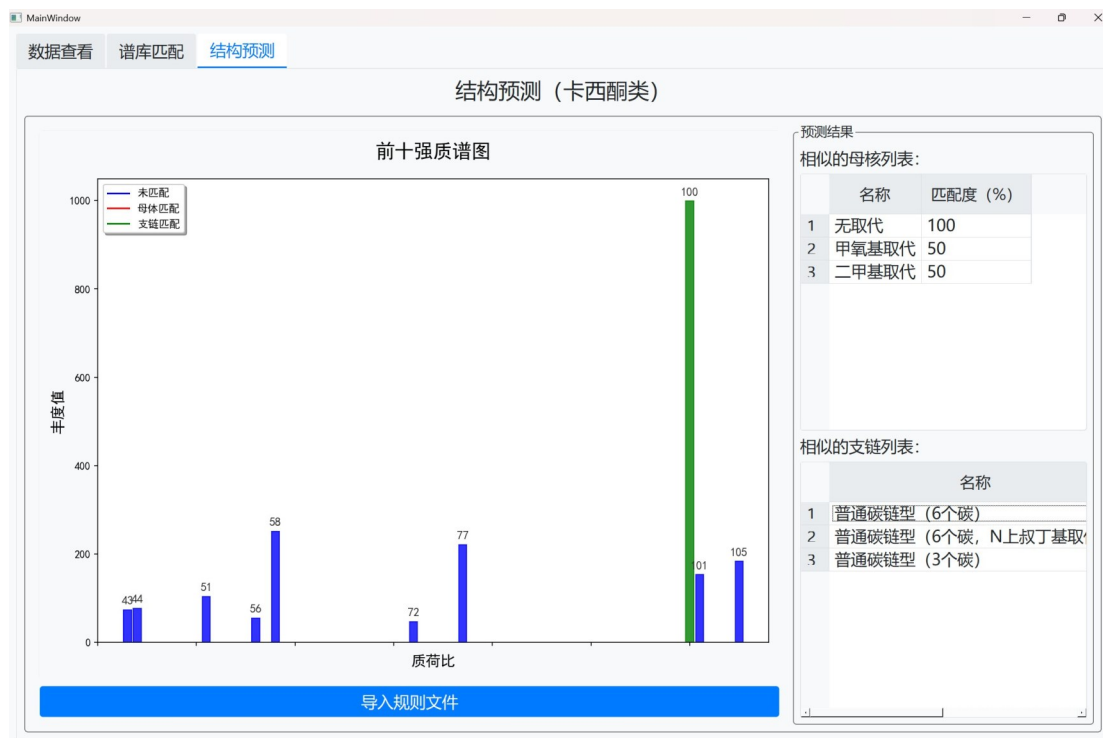


图 13 卡西酮类物质自动识别平台识别结果

Fig. 13 Identification result of the automatic identification platform for cathinones

## 参考文献:

- [1] Shi H F, Xu B P, Xu C X, Zhou X Q, Zhou Y F, Liu K Y, Guo M X. *J. Instrum. Anal.* (史洪飞, 徐伯, 徐成鑫, 周修齐, 周亦飞, 刘柯宇, 郭明鑫. 分析测试学报), **2025**, 44 (7): 1406-1412.
- [2] Xu B P, Shi H F, Xu C X. *Forensic Sci. Technol.* (徐伯, 史洪飞, 徐成鑫. 刑事技术), **2024**, 49 (1): 72-79.
- [3] Yin Q H, Yang L H, Chen J, Yang C, Zhu Y Q. *Phys. Test. Chem. Anal.: Chem. Anal.* (殷勤红, 杨黎华, 陈瑾, 杨成, 朱艳琴. 理化检验-化学分册), **2023**, 59 (2): 241-248.
- [4] Adamowicz P. *Clin. Toxicol.*, **2021**, 59 (7): 648-654.
- [5] Nadal-Gratacós N, Pazos M D, Pubill D, Camarasa J, Escubedo E, Berzosa X, López-Arnau R. *ACS Pharmacol. Transl. Sci.*, **2024**, 7: 2588-2603.
- [6] Fan Y N, Li L. *Chin. J. Drug Abuse Prev. Treat.* (范语诺, 李莉. 中国药物滥用防治杂志), **2024**, 30 (10): 1779-1784.
- [7] Jin B S, Wang P, Liu X Y, Guan H Q, Luo D H, Zhou H, Liang G Q. *Forensic Sci Technol* (金彬书, 王平, 刘晓云, 关浩泉, 罗德杭, 周航, 梁桂巧. 刑事技术), **2025**, 50 (2): 175-181.
- [8] Deng Q Y, Sun W J, He S Y, Wang K D, Chen Y S, Wang W, Liang C. *J. Forensic Med.* (邓乾亚, 孙文娟, 何思阳, 王跨陡, 陈永生, 王威, 梁晨. 法医学杂志), **2022**, 38 (4): 473-477.
- [9] Zhang Y C, Chen X G, Zhao D, Yao W X, Song H. *J. Instrum. Anal.* (张一辰, 陈学国, 赵丹, 姚伟宣, 宋辉. 分析测试学报), **2021**, 40 (4): 560-564.
- [10] Wang D, Chen X G, Ming Z, Jiang L M, Zhou Y. *Separations*, **2022**, 9 (1): 3.
- [11] Xing J, Guo Z J, Wei C S, Zheng J L. *J. Anal. Sci.* (邢警, 郭子健, 魏春生, 郑吉龙. 分析科学学报), **2025**, 41 (6): 715-721.
- [12] Davidson J T, Piacentino E L, Sasiene Z J, Abiedalla Y, DeRuiter J, Clark C R, Berden G, Oomens J, Ryzhov V, Jackson G P. *Forensic Chem.*, **2020**, 19: 100245.
- [13] Davidson T, Sasiene Z J, Abiedalla Y, DeRuiter J, Clark C R, Jackson G P. *Int. J. Mass Spectrom.*, **2020**, 453: 116343.
- [14] Li Y W, Lian R, Wang K D, Xia P, Bu J X, Qiu B L, Zhang Y R. *J. Chin. Mass Spectrom. Soc.* (李雅文, 连茹, 王跨陡, 夏攀, 卜杰洵, 邱秉林, 张玉荣. 质谱学报), **2021**, 42 (4): 427-438.
- [15] Lee S Y, Lee S T, Suh S, Ko B J, Oh H B. *J. Anal. Toxicol.*, **2022**, 46 (7): 732-742.
- [16] Silverstein R M, Bassler G C. *J. Chem. Educ.*, **1962**, 39 (11): 546.
- [17] Li G H, Feng Y, Cheng Y J, Sun W J, Chen L, Feng T Y, Wang X H, Yang Z. *Vacuum Cryogenics* (李冠华, 冯焱, 成永军, 孙雯君, 陈联, 冯天佑, 王星辉, 杨. 真空与低温), **2025**, 31 (2): 148-155.
- [18] GA/T 1920-2021. Forensic Science—Analysis of 211 Narcotic Drugs and Psychotropic Substances in Suspected Drugs—

Gas Chromatography-Mass Spectrometry Method. Public Security Industry Standard of the People's Republic of China (法庭科学 疑似毒品中 211 种麻醉药品和精神药品检验 气相色谱-质谱法. 中华人民共和国公共安全行业标准).

(责任编辑:)