周梦蕾, 赵新城, 李佳燕,等. 基于 GAM 和 Transformer 的近红外-拉曼光谱融合的药物 API 定量分析模型[J]. 智能计算机与应用,2026,16(8):121-128. DOI:10.20169/j.issn.2095-2163.26060804

基于 GAM 和 Transformer 的近红外-拉曼光谱融合的药物 API 定量分析模型

周梦蕾[1], 赵新城[1], 李佳燕[1], 李启恩[2], 邹北骥[3], 刘 芳[2], 刘青萍[1]
(1 湖南中医药大学 信息科学与工程学院, 长沙 410208; 2 湖南中医药大学 药学院, 长沙 410208;
3 中南大学 计算机学院, 长沙 410083)

摘 要: 药品质量控制中活性成分(API)的快速无损检测对制药工业具有重要意义。针对近红外光谱(NIR)与拉曼光谱(Raman)融合定量分析中存在的量纲差异大、特征冗余度高以及小样本建模易过拟合等问题,本文提出了一种基于全局注意力机制(GAM)与 Transformer 结合的定量分析模型(TFGA)。该模型采用数据级融合策略将 NIR 与 Raman 原始光谱直接拼接,利用 GAM 强化关键光谱特征并抑制背景噪声,通过 Transformer 建模跨模态光谱序列的长程依赖关系。同时,引入基于高斯噪声、基线漂移和强度缩放的物理驱动在线数据增强策略,以提高模型鲁棒性。片剂 API 含量预测实验表明,TFGA 在预测集上的决定系数(R_p^2)为 0.966 6,预测均方根误差(RMSEP)为 0.205 3,相对分析误差(RPD)为 5.585 6,整体性能显著优于 PLSR、SVM、RF 和 1D-CNN 模型。输入梯度显著性分析证实,该模型能够准确识别与 API 浓度及 C≡N 等特征化学键相关的敏感波段。结果表明,TFGA 模型可为多源异构光谱的高精度、可解释定量分析提供有效方法。

关键词: 近红外光谱; 拉曼光谱; 数据级融合; 定量分析; 全局注意力机制; Transformer; 数据增强
中图分类号: TP391 **文献标志码**: A **文章编号**: 2095-2163(2026)08-0121-08

Quantitative analysis of pharmaceutical APIs using a GAM and Transformer-based NIR-Raman spectral fusion model

ZHOU Menglei[1], ZHAO Xincheng[1], LI Jiayan[1], LI Qien[2], ZOU Beiji[3], LIU Fang[2], LIU Qingping[1]
(1 School of Informatics, Hunan University of Chinese Medicine, Changsha 410208, China;
2 School of Pharmacy, Hunan University of Chinese Medicine, Changsha 410208, China;
3 School of Computer Science and Engineering, Central South University, Changsha 410083, China)

Abstract: Rapid, non-destructive determination of active pharmaceutical ingredients (API) is crucial for the pharmaceutical industry. To address scale discrepancies, feature redundancy, and small-sample overfitting in near-infrared (NIR) and Raman spectral fusion, we propose TFGA, a quantitative model combining a Global Attention Mechanism (GAM) and a Transformer. TFGA directly concatenates raw spectra, utilizing GAM to enhance key features and suppress background noise, while the Transformer captures cross-modal long-range dependencies of spectral sequences. A physics-driven online data augmentation strategy is introduced to boost robustness. For tablet API prediction, TFGA achieved an R_p^2 of 0.966 6, RMSEP of 0.205 3, and RPD of 5.585 6, significantly outperforming PLSR, SVM, RF, and 1D-CNN models. Saliency analysis confirms it accurately identifies sensitive bands linked to API concentration and characteristic C≡N bonds. Overall, TFGA provides an effective approach for high-precision, interpretable quantitative analysis of heterogeneous multi-source spectra.

Key words: near-infrared spectroscopy; Raman spectroscopy; data-level fusion; quantitative analysis; Global Attention Mechanism; Transformer; data augmentation

0 引 言

药品质量控制中活性成分(API)含量的测定,是固体制剂生产、过程监控及成品放行的核心环节,对保障药品的有效性和安全性至关重要[1-2]。传统理化分析方法虽具有较高准确度,但其繁琐的样品

基金项目: 国家重大科技专项(2018AAA0102100);2025 年湖南中医药大学研究生科研创新项目(2025CX123)。
作者简介: 周梦蕾(2001—),女,硕士研究生。主要研究方向:医学人工智能,光谱分析。
通信作者: 刘 芳(1980—),女,博士,教授,硕士生导师。主要研究方向:中药药理,中药分析。Email:fliu0825@126.com; 刘青萍(1982—),女,博士,教授。主要研究方向:医学人工智能,智能信息处理。Email:liuliu@hnucm.edu.cn。
收稿日期: 2026-06-08

前处理步骤和冗长的分析周期,难以完全满足现代制药工业对快速、无损及在线检测的迫切需求[3]。近红外光谱(NIR)与拉曼光谱(Raman spectroscopy)作为两种重要的分子振动光谱技术,具有制样简单、检测速度快、非侵入性等优势,已广泛应用于药物分析、过程分析等领域[4-5]。其中,NIR 主要反映含氢基团,如 C-H、O-H、N-H 的倍频与合频吸收信息,适合表征样品的整体组成与物理性质[6];Raman 则基于分子振动引起的极化率变化,其特征峰形尖锐,对分子骨架结构高度敏感,能够解析关键的化学结构信息[7],适用于药物固体的定性与定量分析,具有快速、无损、结构信息丰富等特征优势[8]。在药片 API 的定量分析中,NIR 与 Raman 联用能够提供更全面的物质信息,如 Dyrby[9] 等研究发现,即便目标组分含氰基(C≡N)等 Raman 活性的官能团,其定量结果仍可能受到 Raman 局部采样特性的限制,而 NIR 透射光谱则更能反映样品的整体平均信息。研究表明,在多源光谱融合建模时,可以通过合理的融合策略和建模方法促进两种光谱信息的互补优势,从而提升 API 定量分析的准确性与可靠性。

目前,光谱定量分析常用的方法包括偏最小二乘判别分析(Partial Least Squares Discriminant Analysis, PLS-DA)[10]、偏最小二乘回归(Partial Least Squares Regression, PLSR)[11]、支持向量机(Support Vector Machine, SVM)[12]、随机森林(Random Forest, RF)[13]等传统化学计量学与机器学习方法。面对复杂基体以及多源光谱融合带来的非线性响应,传统线性模型在深层特征提取和复杂模式表达方面存在一定局限。Zhang[14] 等指出,机器学习方法已广泛应用于 NIR 光谱的预处理、变量选择和定量分析任务,但模型性能仍受到样本规模、变量维度及特征表达方式等因素影响。随着深度学习的快速发展,卷积神经网络(Convolutional Neural Network,CNN)逐渐被用于光谱定量分析[15]。于水[16] 等通过并联卷积模块提取多尺度光谱特征,提升了模型预测性能与可解释性。不过,CNN 主要依赖局部卷积核提取特征,在建模远距离波段之间的关联关系时仍存在不足。相比之下,Transformer 架构中的自注意力机制能够建模序列中任意位置之间的全局依赖关系,不受固定感受野限制[17]。因此,将 Transformer 引入 NIR-Raman 融合光谱建模,有助于捕捉跨波段、跨模态的全局关联信息。在多源光谱融合策略方面,数据级融合通过直接拼接原始光谱变量,能够最大程度保留初始信息。Szabó-Szücs[18] 等指出融合策略的选择会显著影响模型的预测性能和稳健性。对于小样本、高维度的 NIR-Raman 数据集而言,如何在保留原始光谱信息的同时,从大量冗余变量中提取关键特征,并抑制噪声对模型训练的影响,仍是当前研究中的关键问题。

针对上述挑战,本文提出一种基于全局注意力机制(Global Attention Mechanism,GAM)与 Transformer 结合的融合定量分析模型,简称 TFGA 模型。基于数据级融合框架,本模型以级联后的 NIR-Raman 原始光谱为输入端。为避免光谱直接级联带来的特征冗余问题,模型在自注意力机制前置入 GAM 模块,通过通道与位置双维度的特征解耦与动态交互,使模型在 Transformer 编码前能够优先关注与 API 含量变化相关的敏感区域。随后,借助 Transformer 对长距离依赖关系的建模能力,进一步捕捉融合光谱中跨波段、跨模态信号之间的全局关联,从而实现对 NIR-Raman 融合光谱深层特征的有效表征。与传统机器学习模型相比,TFGA 模型能够更充分地提取融合光谱中的非线性特征;与 CNN 类模型相比,其具备更强的全局依赖建模能力;与单纯 Transformer 相比,引入 GAM 有助于提高模型对关键光谱区域的敏感性。因此,TFGA 模型在小样本、高维 NIR-Raman 药片 API 定量分析任务中,能够兼顾局部特征增强、全局信息建模和双光谱融合表达,为多源光谱定量分析提供了一种可行的建模思路。

1 实验部分

1.1 数据来源

本研究使用的数据集源自 H. Lundbeck A/S 公司的艾司西酞普兰(Escitalopram™)片剂质量控制相关研究[19]。数据获取地址:https://eigenvector.com/resources/data-sets/。样品来自实验室、试点及全规模生产批次,总计 31 批次。片剂主要由微晶纤维素和常规辅料组成,部分样品具有涂层。NIR 光谱在 4 000-14 000 cm^{-1} 范围内采用透射模式测量,同时采集范围为 3 600-200 cm^{-1} 的 FT-Raman 光谱,以充分捕获含 C≡N 键的 API 特征信号。分别采集 310 个样本的 NIR 光谱、120 个样本的 FT-Raman 光谱数据,记录样本编号、活性成分含量、样本类型等。NIR 与 Raman 光谱数据的样本匹配情况见表 1。采用 NIR-Raman 双光谱融合建模,仅保留具有 NIR 和 Raman 光谱的共同样本,最终获得 120 个融合样本。数据集的预测目标为药片中活性成分含量,单位为%(w/w)。融合后样本同时包含

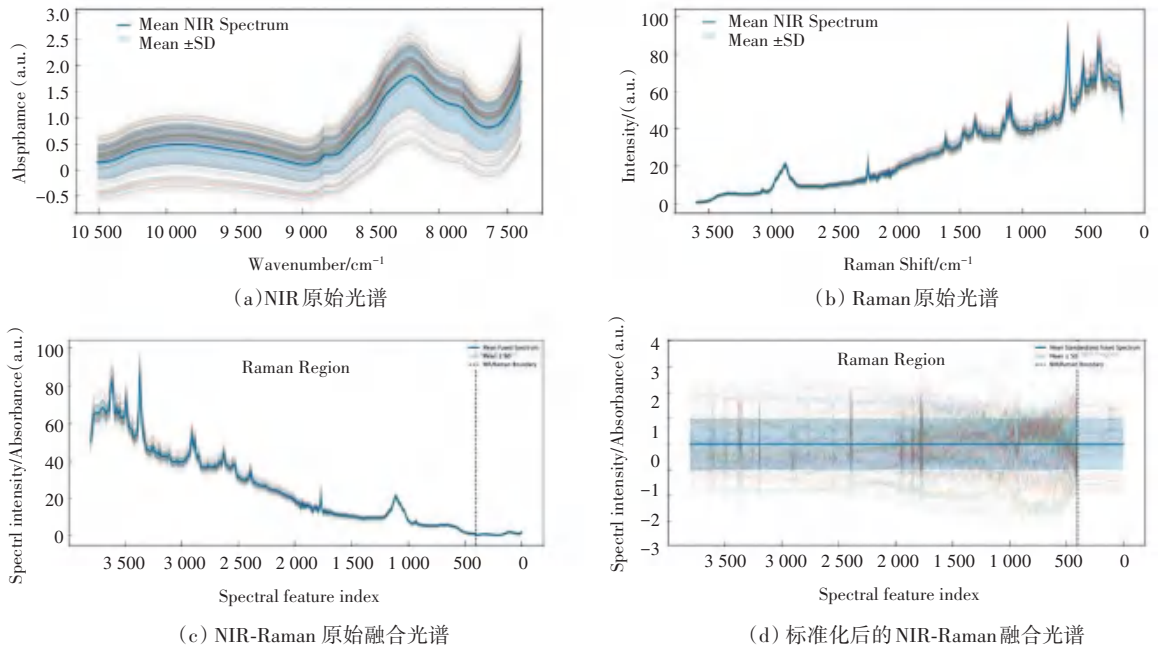
NIR 与 Raman 两类光谱变量，构成小样本、高维度的多源光谱数据。NIR 光谱主要表征样品整体组成和物理性质[20]，Raman 光谱则基于分子振动引起的极化率变化，对分子骨架结构高度敏感，能够解析关键的局部化学结构信息[21-22]，两者在定量分析中具有潜在的互补优势。

表 1　原始光谱数据与融合后样本数量统计

Table 1　Statistics of sample quantities for original and fused spectral data

数据类型	原始样本数	匹配后样本数
NIR	310	120
Raman	120	120
NIR+Raman 融合	120	120

图 1 展示了片剂样本的 NIR、Raman 及融合光谱预处理结果。其中，图 1(a)为 NIR 原始光谱，整体表现为较宽的吸收带，峰形重叠较明显；图 1(b)为 Raman 原始光谱，具有较尖锐的特征峰，能够提供局部分子振动信息；图 1(c)所示原始融合光谱，采用数据级融合方式将 NIR 与 Raman 光谱在特征维度上直接拼接形成，能充分利用两种光谱的互补信息；采用 Standard Scaler 对融合光谱进行标准化处理，得到图 1(d)所示的标准化融合光谱，可以解决 NIR 与 Raman 光谱在信号强度和量纲上存在的差异化问题，降低不同光谱变量尺度差异对模型训练的影响。



(a)NIR 原始光谱　(b) Raman 原始光谱

(c) NIR-Raman 原始融合光谱　(d) 标准化后的 NIR-Raman 融合光谱

图 1　片剂样本的原始光谱、融合光谱及标准化处理结果

Fig. 1　Raw spectra, fused spectra and standardized spectra of tablet samples

1.2　数据预处理与在线数据增强

本文将融合后的 NIR-Raman 光谱样本按照 8∶2 的比例随机划分为训练集和预测集。标准化处理采用 Standard Scaler。考虑到融合样本数量有限，为提高模型训练稳定性和泛化能力，本文在训练阶段引入基于光谱物理扰动机制的动态数据增强策略，通过模拟基线漂移、强度扰动、波长偏移及噪声变化等方式扩充训练样本分布[23-24]。对于标准化后的融合光谱向量 $\boldsymbol{x}$，增强后的光谱 $\tilde{x}$ 可表示为：

$$\tilde{x} = (x + \epsilon + \delta) \cdot s \qquad (1)$$

其中，ϵ 表示服从零均值高斯分布的随机噪声；δ 表示小幅度基线偏移；s 表示接近 1 的随机强度缩放因子。本文设置噪声标准差为 0.005，基线漂移最大幅度为 0.01，强度缩放最大幅度为 0.01。

1.3　模型结构

TFGA 模型以数据级融合后的 NIR-Raman 光谱作为输入，通过局部光谱片段嵌入、全局注意力增强和 Transformer 全局依赖建模，实现对药片 API 含量的回归预测。模型整体结构包括 5 个部分：光谱数据级融合模块、Patch Embedding 模块、全局注意力增强模块、Transformer 编码模块和回归预测模块，如图 2 所示。

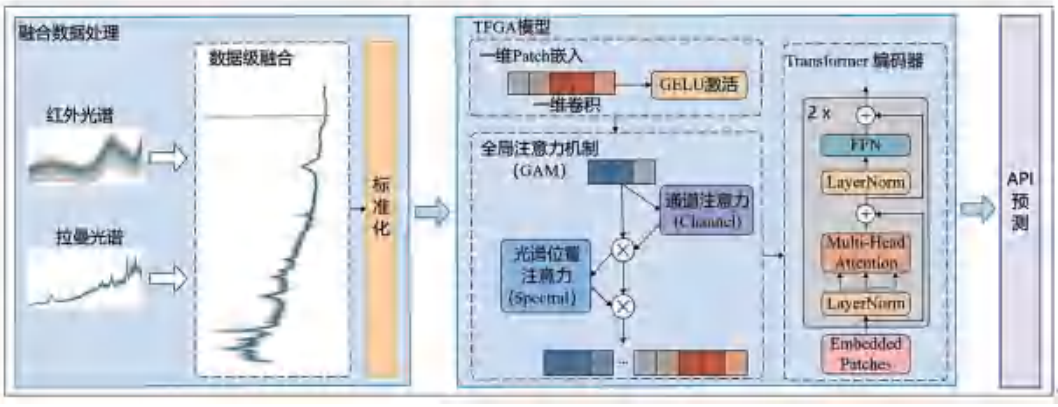


图 2 TFGA 模型工作流程

Fig. 2 Workflow of the TFGA model for fused NIR-Raman spectra

首先,对匹配后的 NIR 光谱和 Raman 光谱进行数据级融合。对于第 i 个样本,其近红外光谱向量记为 $\boldsymbol{X}_{\mathrm{NIR}}^{(i)}$,拉曼光谱向量记为 $\boldsymbol{X}_{\mathrm{Raman}}^{(i)}$,则融合光谱可表示为:

$$\boldsymbol{X}_{\mathrm{fusion}}^{(i)} = [X_{\mathrm{NIR}}^{(i)}, X_{\mathrm{Raman}}^{(i)}] \tag{2}$$

其中,$[\boldsymbol{X}_{\mathrm{NIR}}^{(i)}, \boldsymbol{X}_{\mathrm{Raman}}^{(i)}]$ 表示在特征维度上的拼接。由于 NIR 与 Raman 光谱在信号强度、变量尺度和峰形结构上存在差异,本文在模型训练前对融合光谱进行标准化处理,使不同光谱变量处于相对一致的数值范围,以降低量纲差异对模型训练的影响。

在 Patch Embedding 模块中,标准化后的融合光谱首先被输入一维卷积层。该模块的作用是将连续的高维一维光谱划分为若干局部光谱片段,并映射到统一的特征维度。卷积嵌入后进一步加入 GELU 激活函数,以稳定特征分布并增强非线性表达能力[25-26]。该过程可表示为:

$$Z = \mathrm{GELU}(\mathrm{Conv1D}(X_{\mathrm{fusion}})) \tag{3}$$

其中,X_{fusion} 表示标准化后的融合光谱输入;Conv1D 表示一维卷积式 Patch Embedding 操作;Z 表示经过局部片段嵌入后的光谱 token 序列。

随后,本文在 Transformer 编码前引入 GAM 对光谱 token 进行特征增强。通过 GAM 模块,模型能够在进入自注意力建模前优先突出对定量预测更有贡献的光谱区域。经过 GAM 增强后的 token 序列被输入 Transformer 编码模块。本文在 token 序列前加入一个可学习的 CLS token,用于聚合整个融合光谱的全局表示;同时引入位置编码,以保留光谱片段之间的顺序信息。Transformer Encoder 由多头自注意力层、前馈神经网络、残差连接和 Layer Normalization 组成[27-28]。其核心作用是建模不同光谱片段之间的长距离依赖关系,尤其适用于 NIR-Raman 融合后存在跨波段、跨模态关联的信息表征。相比仅依赖卷积核局部感受野的 CNN 结构,Transformer 可以直接建立远距离光谱片段之间的关联,从而增强模型对融合光谱全局结构的学习能力。

最后,模型将 Transformer 输出的 CLS token 表征输入回归预测模块。回归头由全连接层、GELU 激活函数和 Dropout 组成,最终输出药片 API 含量的预测值:

$$\hat{y}^{(i)} = f_{\mathrm{reg}}(h_{\mathrm{CLS}}^{(i)}) \tag{4}$$

其中,$h_{\mathrm{CLS}}^{(i)}$ 表示第 i 个样本经过 Transformer 编码后的全局光谱表征,$\hat{y}^{(i)}$ 为模型预测的 API 含量。模型训练过程中采用均方误差(Mean Squared Error,MSE)作为损失函数:

$$L = \frac{1}{N}\sum_{i=1}^{N}(y^{(i)} - \hat{y}^{(i)})^2 \tag{5}$$

其中,$y^{(i)}$ 为真实 API 含量;$\hat{y}^{(i)}$ 为预测值;N 为训练样本数。

综上,TFGA 模型的核心思想是在数据级融合阶段保留 NIR 与 Raman 的原始信息,首先通过 Patch Embedding 提取局部光谱片段特征,再利用 GAM 对关键特征进行增强,最后通过 Transformer 建模融合光谱中的全局依赖关系。该结构兼顾了局部特征提取、关键波段增强和跨波段全局建模能力,适用于小样本、高维度 NIR-Raman 融合光谱的药片活性成分定量预测。

1.4 模型评价

采用预测集决定系数 R_p^2、预测均方根误差 RMSEP、相对分析误差 RPD 和平均绝对百分比误差 MAPE 作为主要评价指标[29-30]。

各评价指标计算公式如下:

$$R_p^2 = 1 - \frac{\sum_{i=1}^{N_p}(y_{p,i} - \hat{y}_{p,i})^2}{\sum_{i=1}^{N_p}(y_{p,i} - \bar{y}_p)^2} \quad (6)$$

$$\mathrm{RMSEP} = \sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(\hat{y}_{p,i} - y_{p,i})^2} \quad (7)$$

$$\mathrm{RPD} = \frac{SD_p}{\mathrm{RMSEP}} \quad (8)$$

$$\mathrm{MAPE} = \frac{100\%}{N_p}\sum_{i=1}^{N_p}\left|\frac{y_{p,i} - \hat{y}_{p,i}}{y_{p,i}}\right| \quad (9)$$

其中，N_p 表示预测集样本数；$y_{p,i}$、$\hat{y}_{p,i}$ 分别表示预测集中第 i 个样本的真实 API 含量和预测 API 含量；$\bar{y}_p$ 表示预测集真实值的平均值；SD_p 表示预测集真实值的标准差。通常情况下，较高的 R_p^2 和 RPD，以及较低的 RMSEP 和 MAPE，表明模型具有更好的预测精度和泛化能力。

2　结果与讨论

2.1　基于 TFGA 模型的贡献度波长分析

采用输入梯度显著性分析方法，计算模型输出对输入光谱变量的局部敏感程度，并将归一化显著性权重用于表征不同光谱特征点对预测结果的相对贡献，以分析 TFGA 模型在 NIR-Raman 融合光谱定量预测过程中的特征敏感区域。该方法反映模型在当前训练条件和测试样本下对输入变量的梯度敏感性，因此其结果主要用于解释模型关注的重要区域。由图 3 可以看出，TFGA 模型在 NIR-Raman 融合光谱上并不是均匀关注所有输入变量，而是表现出明显的局部敏感特征。上半部分展示了 NIR 与 Raman 的原始平均融合光谱，下半部分展示了对应的归一化输入梯度显著性权重。整体来看，NIR 区域前段出现较高显著性响应，说明该区域的光谱变化对 API 含量预测较为敏感。同时，Raman 区域也存在多个局部显著性高峰，尤其在若干尖锐峰附近显著性权重明显升高，表明 Raman 光谱中的局部分子振动信息也参与了模型预测过程。结合图 4 进一步分析可以发现，TFGA 模型的主要贡献特征集中在 NIR 区域约 7 460-7 637 cm^{-1} 附近，其中 7 544-7 568 cm^{-1} 附近的显著性权重最高，提示 NIR 光谱在本研究的 API 含量预测中提供了重要的定量信息。由于 NIR 光谱主要来源于 C-H、O-H、N-H 等含氢基团的倍频与合频吸收，且峰形通常较宽、重叠较强，因此该区域反映了药片中 API 与辅料整体组成、物理状态或基体效应的综合变化，不能直接归属于某一个单一官能团峰。从 Raman 区域来看，图 4 表明高贡献特征主要分布在 613-636 cm^{-1} 和 2 181-2 193 cm^{-1} 附近，其中 613-636 cm^{-1} 区域出现多个连续高显著性特征，提示该区域峰的信息对模型预测具有一定贡献，该区域与药物分子或辅料中的骨架振动、环变形振动等低波数结构信息有关；2 181-2 193 cm^{-1} 区域也表现出较高的显著性，结合该药片数据集的背景可知，目标活性成分含有 Raman 活性的 C≡N 官能团，该高显著性区域可能与 C≡N 相关伸缩振动信息有关。

结合图 3 和图 4 可知，NIR 区域的高贡献特征分布更集中，说明整体吸收变化是 API 含量预测的重要基础；Raman 区域的贡献表现为多个局部峰，说明其提供了补充性的分子结构信息。该结果与 TFGA 的模型设计思路一致，即通过数据级融合保留 NIR 与 Raman 的原始信息，再利用 GAM 和 Transformer 对关键光谱区域及跨波段关系进行建模。因此，贡献度波长分析表明，TFGA 模型能够从高维融合光谱中筛选出与 API 含量变化相关的敏感区域，为模型预测结果提供可解释性。

2.2　不同建模方法分析

选取传统机器学习模型、深度学习基线模型与本文所提出的 TFGA 模型进行对比分析。不同模型的实验结果见表 2。由表 2 可知，TFGA 模型在数据集上取得 R_p^2 = 0.966 6、RMSEP = 0.205 3、RPD = 5.585 6 和 MAPE = 2.40%，在所有模型中表现最优。

表 2　不同模型在预测集上的定量分析性能对比

Table 2　Comparison of quantitative prediction performance of different models on the prediction set

模型	R_p^2	RMSEP	RPD	MAPE/%
PLSR	0.534 1	0.766 1	1.496 6	6.59
SVM	0.742 4	0.569 6	2.012 8	5.58
RF	0.799 7	0.502 2	2.282 7	4.30
1D-CNN	0.760 1	0.549 8	2.085 5	5.99
1D-CNN+GAM	0.811 0	0.487 9	2.349 9	5.34
Transformer	0.931 7	0.293 2	3.909 8	3.26
TFGA	0.966 6	0.205 3	5.585 6	2.40

结果表明，传统机器学习模型虽然具有建模流程简单、可解释性较强等优点，但在面对 NIR-Raman 高维融合光谱时，对复杂非线性关系和跨模态互补信息的提取能力不足。1D-CNN 和 1D-CNN+GAM 能够增强局部光谱特征表达，但对长距离波段依赖

关系的建模能力仍有限。相比之下,TFGA 通过数据级融合保留 NIR 与 Raman 的原始光谱信息,利用 GAM 增强关键光谱区域,并通过 Transformer 捕捉长距离依赖关系,因此在预测精度、误差控制和模型表征能力方面均表现更优。结果说明,基于全局注意力机制和 Transformer 的融合建模方法更适合本研究中的小样本、高维度 NIR-Raman 光谱 API 含量定量分析任务。

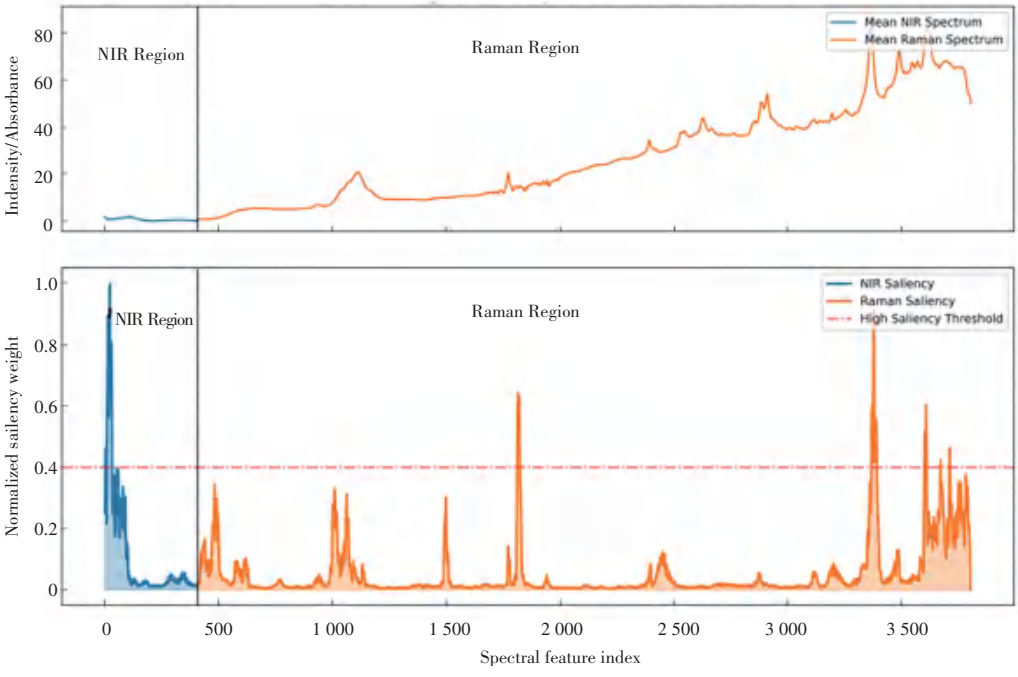


图 3 TFGA 模型的融合光谱显著性分析图

Fig. 3 Saliency analysis of fused spectra using the TFGA model

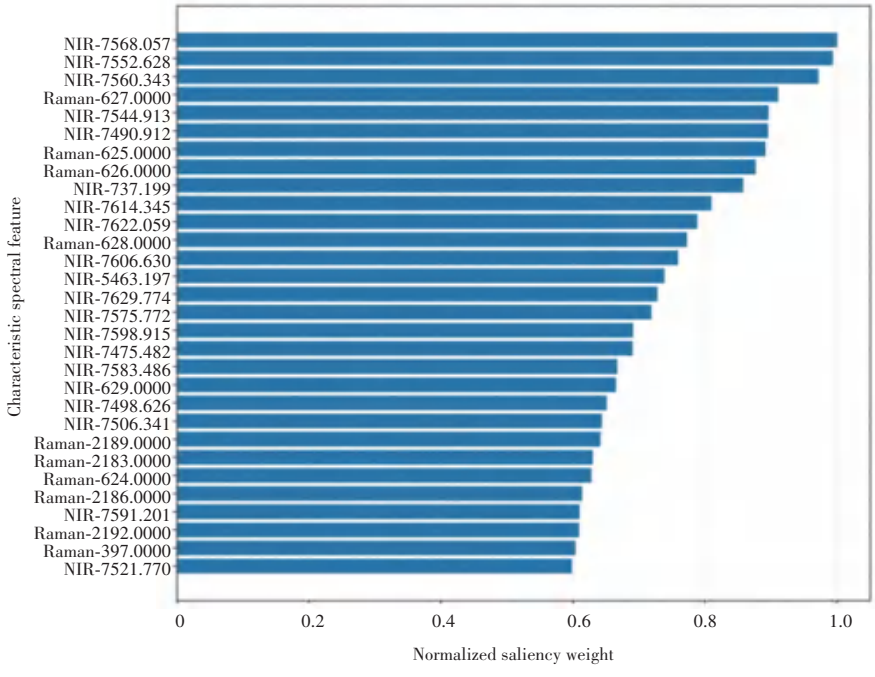


图 4 TFGA 模型 Top30 高贡献光谱特征

Fig. 4 Top 30 salient spectral features identified by the TFGA model

2.3 消融实验

为进一步验证 TFGA 模型中关键模块对预测性能的贡献,本文以 Transformer 作为基础模型,考察 GAM 对模型预测性能的影响。由表 3 可见,与

Transformer 相比，引入 GAM 后，TFGA 的预测性能进一步提升，R_p^2 提高至 0.966 6，RMSEP 降低至 0.205 3，MAPE 降低至 2.40%。该结果表明，结合 GAM 与 Transformer，能够同时发挥关键光谱特征增强和全局依赖建模的优势，从而提高 NIR-Raman 融合光谱对片剂 API 含量的定量预测能力。

表 3　TFGA 模型关键模块消融实验结果

Table 3　Ablation study results of key modules in the TFGA model

模型	R_p^2	RMSEP	RPD	MAPE/%
Transformer	0.931 7	0.293 2	3.909 8	3.26
TFGA	0.966 6	0.205 3	5.585 6	2.40

3　结束语

本文提出了一种基于全局注意力机制和 Transformer 的近红外光谱与拉曼光谱融合定量分析模型 TFGA，用于片剂活性成分 API 含量的快速预测。该模型首先采用数据级融合方式，将 NIR 与 Raman 光谱在特征维度上直接拼接，以保留两种光谱的互补信息；随后利用 GAM 对关键光谱区域进行增强，并通过 Transformer 捕捉融合光谱中的长距离依赖关系，建立融合光谱与 API 含量之间的非线性映射关系。实验结果表明，TFGA 模型在预测集上取得了较好的定量分析性能，TFGA 能够更充分地挖掘 NIR 与 Raman 光谱之间的互补信息。消融实验进一步证明，GAM 是提升 TFGA 模型性能的关键模块，该模块能够有效增强与 API 含量预测相关的关键光谱特征。此外，基于输入梯度的显著性分析结果表明，TFGA 模型对 NIR 区域前段的部分特征关注更多，同时 Raman 区域的若干局部特征也对预测结果产生贡献，提示模型并非均匀地依赖全部融合光谱变量，而是从高维光谱数据中主动筛选出对 API 含量变化更敏感的关键区域，从而在一定程度上提升了模型的可解释性。

综上所述，TFGA 模型能够有效融合 NIR 与 Raman 光谱信息，并在小样本、高维度的片剂 API 含量预测任务中表现出较好的预测能力和稳定性。该方法为药品质量控制中的快速、无损定量检测提供了可行的思路。但本文仍存在样本量有限、外部验证不足等问题，后续研究可进一步扩大样本规模，结合独立验证集和更深入的化学峰归属分析，以提升模型的泛化能力和实际应用价值。

参考文献

[1] Nikam N A, Shinde S R, Redasani V K. In process and quality control tests for tablets: A review[J]. International Journal of Pharmaceutical Research and Applications, 2025, 10(1): 1113-1120.

[2] Goodwin D J, Van den Ban S, Denham M, et al. Real time release testing of tablet content and content uniformity[J]. International Journal of Pharmaceutics, 2018, 537(1/2): 183-192.

[3] 李伟俊，褚小立，陈瀑，等. 现代智能光谱分析技术的演进、进展及展望[J]. 化学试剂，2026，48(1)：1-15. DOI：10.13822/j.cnki.hxsj.2026.0043.

[4] Roggo Y, Chalus P, Maurer L, et al. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies[J]. Journal of Pharmaceutical and Biomedical Analysis, 2007, 44(3): 683-700.

[5] 秦晨，朱东亮，李海龙，等. 光谱技术在药物分析中的研究进展[J]. 广州化工，2026，54(2)：12-16. DOI：10.20220/j.cnki.1001-9677.2026.02.004.

[6] Ozaki Y. Near-infrared spectroscopy—its versatility in analytical chemistry[J]. Analytical Sciences, 2012, 28(6): 545-563.

[7] Ren J, Mao S, Lin J, et al. Research progress of Raman spectroscopy and Raman imaging in pharmaceutical analysis[J]. Current Pharmaceutical Design, 2022, 28(18): 1445-1456.

[8] Strachan C J, Rades T, Gordon K C, et al. Raman spectroscopy for quantitative analysis of pharmaceutical solids[J]. Journal of Pharmacy and Pharmacology, 2007, 59(2): 179-192.

[9] De Man A, De Souter L, Shi Z, et al. Evaluating the improvement of blend potency measurements in the feed frame of a rotary tablet press using combined NIR and Raman spectroscopy[J]. Analytical Chemistry, 2024, 96(26): 10586-10593.

[10] Yin Y, Wei X, Song Y, et al. Quality evaluation of Astragalus membranaceus based on polysaccharide fingerprinting and chemometric analysis[J]. International Journal of Biological Macromolecules, 2025, 339: 149726.

[11] Brereton R G. Partial least squares estimates and regression coefficients[J]. Journal of Chemometrics, 2026, 40(5): e70071.

[12] Igne B, Drennen J K, Anderson C A. Improving near-infrared prediction model robustness with support vector machine regression: A pharmaceutical tablet assay example[J]. Applied Spectroscopy, 2014, 68(12): 1348-1356.

[13] Sow A, Traore I, Diallo T, et al. Comparison of Gaussian process regression, partial least squares, random forest and support vector machines for a near infrared calibration of paracetamol samples[J]. Results in Chemistry, 2022, 4: 100508.

[14] Zhang W, Kasun L L C, Wang Q J, et al. A review of machine learning for near-infrared spectroscopy[J]. Sensors, 2022, 22(24): 9764.

[15] Wang N, Li Y, Hou K, et al. Quantitative analysis of near-infrared spectroscopic data based on adaptive feature optimized CNN[J]. Measurement, 2025, 247: 116817.

[16] 于水，宦克为，刘小溪，等. 并联卷积神经网络的近红外光谱定量分析模型[J]. 光谱学与光谱分析，2024，44(6)：1627-1635.

[17] Koyun O C, Keser R K, Şahin S O, et al. RamanFormer: A transformer-based quantification approach for Raman mixture components[J]. ACS Omega, 2024, 9(22): 23241-23251.

[18] Szabó-Szücs B, Ficzere M, Péterfi O, et al. Simultaneous prediction of the API concentration and mass gain of film coated

tablets using near-infrared and Raman spectroscopy and data fusion [J]. International Journal of Pharmaceutics, 2025, 668: 124957.
[19] Dyrby M, Engelsen S B, Nørgaard L, et al. Chemometric quantitation of the active substance in a pharmaceutical tablet using near- infrared transmittance and NIR FT-Raman spectra [J]. Applied Spectroscopy, 2002, 56(5): 579-585. .
[20]王家琪, 唐千慧, 李荣强, 等. 基于双向门控单元与门控注意力卷积的近红外光谱定量分析模型[J]. 智能计算机与应用, 2025, 15(12): 74-82. DOI: 10. 20169/j. issn. 2095-2163. 25092201.
[21]Kalatzis D, Nega A, Kiouvrekis Y. Raman spectra classification of pharmaceutical compounds: A benchmark of machine learning models with SHAP-based explainability[J]. Engineering, 2025, 6(7): 145.
[22]Liu Y, Chen S, Xiong X, et al. Artificial intelligence guided Raman spectroscopy in biomedicine: Applications and prospects [J]. Computers in Biology and Medicine, 2025, 170: 107982.
[23]Zheng B, Yang J, Li J, et al. Advances in deep learning for intelligent non-destructive spectral analysis: from preprocessing and data augmentation to interpretability, fusion, and transfer learning[J]. TrAC Trends in Analytical Chemistry, 2026, 197: 118741.
[24] Blazhko U, Shapaval V, Kovalev V, et al. Comparison of augmentation and pre-processing for deep learning and chemometric classification of infrared spectra[J]. Chemometrics and Intelligent Laboratory Systems, 2021, 215: 104367.
[25]Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[C]// International Conference on Learning Representations, 2021.
[26]Liu Z, Mao H, Wu C Y, et al. A ConvNet for the 2020s[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11976-11986.
[27]Han K, Wang Y, Chen H, et al. A survey on vision transformer [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(1): 87-110.
[28] Xu P, Zhu X, Clifton D A, et al. Multimodal learning with transformers: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(10): 12113-12132.
[29]李志豪, 肖金凤, 张洪明, 等. 融合特征空间注意力的卷积神经网络用于酒石酸近红外在线检测[J]. 光谱学与光谱分析, 2025, 45(12): 3358-3365. .
[30]Liu H, Chen Y, Shi C, et al. FT-IR and Raman spectroscopy data fusion with chemometrics for simultaneous determination of chemical quality indices of edible oils during thermal oxidation [J]. LWT-Food Science and Technology, 2020, 119: 108906.