

doi: 10.12452/j.fxcxb.26020901

基于高光谱成像与深度学习的滩羊PGK1和PKM2酶水平预测研究

徐海龙¹, 于建国¹, 王彦云¹, 崔佳锐², 李海峰¹, 王松磊^{1*}

(1. 宁夏大学 食品科学与工程学院, 宁夏 银川 750021; 2. 宁夏大学 动物科技学院, 宁夏 银川 750021)

摘要: 宰后肌肉糖酵解速率异常是异质肉(PSE/DFD)形成的重要驱动因素之一。磷酸甘油酸激酶1(PGK1)与丙酮酸激酶M2(PKM2)分别催化糖酵解过程中两步关键底物水平磷酸化反应,可在一定程度上表征宰后能量代谢状态。该研究以滩羊肌肉为对象,测定PGK1与PKM2的免疫反应性水平,构建可见-近红外高光谱成像(Vis-NIR HSI)结合二维相关光谱(2D-COS)的无损评估框架以增强关键谱段解析与模型可解释性。2D-COS识别出5个敏感特征波段(476、562、605、715和800 nm),其光谱响应可能与肌红蛋白状态变化以及水分结合状态/组织结构差异相关。进一步系统比较了多种预处理与特征提取方法在偏最小二乘回归(PLSR)、随机森林(RF)以及卷积神经网络(CNN)中的预测效果。结果表明,结合变量组合群体分析-迭代保留信息变量(VCPA-IRIV)的CNN模型表现最佳:PGK1预测集 $R_p^2=0.8920$, RMSEP=38.3653, RPD=3.0935; PKM2预测集 $R_p^2=0.9030$, RMSEP=18.1778, RPD=3.2659。基于最优模型实现像素级伪彩色可视化,直观呈现了不同贮藏阶段两种酶水平的空间异质性与动态变化趋势。本研究表明,Vis-NIR HSI结合2D-COS与深度学习方法可实现羊肉PGK1和PKM2水平的无损预测,为代谢相关指标的快速分级与过程监测提供技术支撑。

关键词: 可见-近红外高光谱; 滩羊; 磷酸甘油酸激酶1; 丙酮酸激酶M2; 二维相关光谱; 深度学习; 无损检测

中图分类号: O657.3; TS251.53 文献标识码: A 文章编号: 1004-4957(2026)06-0001-11

Nondestructive Prediction of PGK1 and PKM2 Enzyme Levels in Tan Sheep Using Hyperspectral Imaging and Deep Learning

XU Hai-long¹, YU Jian-guo¹, WANG Yan-yun¹, CUI Jia-rui², LI Hai-feng¹, WANG Song-lei^{1*}

(1. School of Food Science and Engineering, Ningxia University, Yinchuan 750021, China; 2. School of Animal Science and Technology, Ningxia University, Yinchuan 750021, China)

Abstract: Abnormal postmortem glycolytic rates are among the primary drivers of meat quality defects, such as pale, soft, exudative(PSE) and dark, firm, dry(DFD) conditions. Phosphoglycerate kinase 1(PGK1) and pyruvate kinase M2(PKM2) catalyze two key substrate-level phosphorylation steps in glycolysis and can therefore partially reflect postmortem energy metabolism status. In this study, Tan sheep muscle was used as the research object. The immunoreactivity levels of PGK1 and PKM2 were determined, and a nondestructive evaluation framework integrating visible-near-infrared hyperspectral imaging (Vis-NIR HSI) with two-dimensional correlation spectroscopy (2D-COS) was developed to enhance sensitive band interpretation and model explainability. Five sensitive wavelengths (476, 562, 605, 715, and 800 nm) were identified by 2D-COS, and their spectral responses were plausibly associated with changes in myoglobin state and differences in water-binding status and/or tissue-structure-related scattering. Furthermore, multiple preprocessing and feature extraction strategies were systematically compared across partial least squares regression (PLSR), random forest (RF), and convolutional neural network (CNN) models. The best performance was achieved by the CNN model coupled with variable combination population analysis-iteratively retaining informative variables (VCPA-IRIV), yielding for PGK1: $R_p^2=0.8920$, RMSEP=38.3653, RPD=3.0935; and for PKM2: $R_p^2=0.9030$, RMSEP=18.1778, RPD=3.2659 in the prediction set. Pixel-wise pseudo color visualization based on the optimal models intuitively revealed the spatial

收稿日期: 2026-02-09; 修回日期: 2026-03-23

基金项目: 国家自然科学基金资助项目(32260625)

* 通讯作者: 王松磊, 博士, 教授, 研究方向: 农产品加工与无损检测, E-mail: wangsonglei163@126.com

heterogeneity and dynamic evolution of the two enzyme levels across different storage stages. These results demonstrate that Vis-NIR HSI combined with 2D-COS and deep learning enables nondestructive prediction of PGK1 and PKM2 enzyme levels in mutton, providing technical support for rapid grading and process monitoring of metabolism-related indicators.

Key words: visible-near infrared hyperspectral imaging; Tan sheep; phosphoglycerate kinase 1; pyruvate kinase M2; two-dimensional correlation spectroscopy; deep learning; nondestructive detection

宁夏盐池滩羊因独特风味与营养特征在国内羊肉市场中具有较高认可度^[1],但在生产与流通环节仍存在明显的个体差异与批次波动,给品质分级与在线监控带来困难。羊肉品质异质性与宰后肌肉代谢过程及其调控机制紧密相关,常表现为可识别的品质缺陷类型,如苍白、软、渗出(pale, soft, exudative, PSE)和暗、紧、干(dark, firm, dry, DFD)。屠宰后肌肉供血终止,代谢模式转向无氧糖酵解,且受到几种关键酶的调控,Huang等^[2]提出并验证了糖酵解通路中的关键蛋白可作为羊肉品质的稳健生物标志物,磷酸甘油酸激酶1(PGK1)与丙酮酸激酶M2(PKM2)属于底物水平磷酸化的关键物质,与宰后能量耗竭速率密切相关^[3]。基于此,以PGK1和PKM2等关键酶开展快速检测有望为羊肉品质异质性的早期识别与监控提供关键信息。

现有酶蛋白检测手段,如比色法、蛋白免疫印记和酶联免疫吸附试验(ELISA)等普遍存在费力、样品具有破坏性和耗时等限制,难以满足肉类加工产业对在线、快速与高通量检测的需求。因此,无损检测技术受到广泛关注。近年来,可见-近红外高光谱成像(Vis-NIR HSI)因可同步获取光谱与空间分布信息,在肉类无损检测方面显示出应用潜力^[4],已被证明可用于嘌呤^[5]、挥发性盐基氮(TVB-N)^[6]和脂肪与蛋白质氧化^[7]检测。Vis-NIR HSI对PGK1和PKM2水平的检测并非直接识别,而是通过糖酵解过程所引起的肉色、水分和组织结构变化之间的耦合关系进行间接表征。两种酶的水平变化通常伴随乳酸积累、pH值下降、肌红蛋白氧化还原平衡状态及肌肉保水性的改变,这些因素共同影响样品的光谱响应,特别是与O—H、N—H、C—H等含氢基团吸收和组织散射特性相关的信号,从而为利用HSI间接预测酶水平提供了理论依据。此外,PGK1位于糖酵解中段,主要参与早期能量转换;而PKM2位于糖酵解末端,与丙酮酸生成和乳酸代谢过程密切相关。因此,尽管二者均与宰后代谢强度相关,但变化幅度并不完全一致,在HSI响应中表现出可区分的特征,因而可分别进行建模。

然而,酶蛋白水平变化往往伴随肌肉微环境、蛋白溶解性与结构状态改变,并以综合效应体现在光谱吸收中。另一方面,光谱数据由于样品表面状态和采集条件的差异存在强共线性与冗余信息。因此,在使用合理的预处理和特征波长提取策略来提升模型稳健性的同时,还需要提高模型的可解释性。二维相关光谱(2D-COS)为此提供了重要参考模式:它能在外部扰动条件下解析光谱信号的同步/异步相关结构,识别对抗扰动最敏感的波段^[8],并借助Noda规则揭示不同波段响应的协同关系与先后顺序,为光谱与指标之间的机理解释提供参考^[9]。2D-COS已被广泛应用于包括高分子、食品、纳米材料和微塑料等多个领域^[9]。Liu等^[10]使用2D-COS识别与抗氧化酶相关的光谱特征,揭示了612-799-473-708-559 nm特征峰下化学键的变化顺序。Hao等^[11]使用2D-COS识别对多环芳烃含量变化敏感的特征峰,获得了4个有较强相关性的荧光发射峰。Cui等^[12]使用2D-COS提取出对番茄有机酸变化最敏感且与多酚等关键组分存在显著相关结构的波段。

因此,本研究以滩羊为研究对象,测定PGK1和PKM2免疫反应性水平,获取其HSI数据,并通过比较不同预处理、特征波长提取对模型性能的提升程度,探究两种酶水平最优的无损预测模型。此外,引入2D-COS从“扰动响应”的角度识别敏感谱段并解析波段间协同/先后变化关系,增强模型与谱段的可解释性。本研究可为羊肉品质异质性的早期识别与过程监控提供新的技术方案,也为HSI在肉类代谢相关指标无损检测中的应用拓展提供参考。

1 实验部分

1.1 样本制备与贮藏

羊肉样本购自宁夏盐池滩羊产业发展集团有限公司。选取年龄和体重相近的滩羊60只,宰后取前

腿、后腿和背最长肌作为研究样本。其中, 0 h样品为尽可能保持初始代谢状态, 采集后经液氮速冻带回实验室, 其余样品置于冷藏保温箱中, 并辅以冰袋维持低温, 运输至宁夏大学肉品加工与质量安全控制实验室。样品前处理及制备时间控制在1 h左右, 以尽量减少运输和制样过程对宰后代谢进程的额外干扰。到达实验室后, 使用不锈钢手术刀将肌肉切割为平整的肉块(30 mm×25 mm×10 mm), 同时剔除筋膜和可见脂肪组织, 共制备180份样品。为模拟商业冷鲜肉零售条件, 将每份样品单独装入聚乙烯袋, 在4 °C条件下贮藏。为表征酶水平与光谱响应的变化, 设置0、12、24、48、72和120 h共6个贮藏时间点, 每个时间点30个样本。本研究样品未进行排酸处理, 而是在宰后取样并制备完成后直接进入4 °C贮藏, 因此0~120 h的样品变化过程本身涵盖了宰后排酸及成熟早期阶段的动态变化。不同部位样品在各贮藏时间点中尽可能均衡分配。本研究未对前腿、后腿和背最长肌进行单独比较, 而是将所有样品作为整体样本集, 重点分析PGK1和PKM2水平及其高光谱响应随贮藏时间的变化特征。到达各时间点后取出样品, HSI扫描后立即取样并液氮速冻, 随后于-80 °C保存。

1.2 高光谱扫描与酶测定

1.2.1 高光谱图像采集 采用中国双利合谱公司的GaiaField pro-V10版本Vis-NIR HSI仪器进行光谱采集。该光谱仪器分辨率为3.5 nm, 波长范围为400~1 000 nm, 可获得176个波长的光谱数据。数据采集前, 打开光谱仪预热30 min, 以消除基线漂移等造成的误差。经预实验, 扫描参数设置如下: 起始位置1.0 cm, 扫描距离20.0 cm, 前向扫描速度0.9 cm/s, 后向扫描速度0.20 cm/s, 曝光时间10 ms, 增益为1。为减弱光照分布不均和暗电流噪声的影响, 进行黑白校正, 公式如下:

$$R = \frac{R_r - R_b}{R_w - R_b} \quad (1)$$

式中: R 为校正后的图像; R_r 为原始图像; R_b 为黑校正图像; R_w 为白校正图像。

1.2.2 酶免疫反应性水平测定 采用ELISA试剂盒(ml952410、ml962502; 上海酶联生物科技有限公司, 中国)测定PGK1和PKM2水平, 结果以试剂盒标定单位(U/mL)表示。按照说明书要求对肌肉组织提取液进行处理, 并使用酶标仪(Infinite F50, Tecan, 瑞士)在450 nm读取吸光度(OD)值。以标准品浓度为横坐标、OD值为纵坐标绘制标准品线性回归曲线, 并根据曲线方程计算样品浓度。

1.3 光谱数据处理

使用ENVI5.6软件提取光谱图像的感兴趣区域并获取平均光谱。采用基于联合X-Y距离的样本划分算法(SPXY)按3:1:1划分为校正集(108个)、验证集(36个)和预测集(36个)。为减少光散射与有效光程差异引起的基线漂移与散射变异, 对光谱数据进行预处理: 分别评估了Savitzky-Golay平滑(S-G)、基线校正(Baseline)以及标准正态变量变换(SNV), 以减弱样品摆放差异及采集条件波动对光谱信号的影响^[13]。在此基础上, 为实现降维去冗余并提升模型稳健性, 进一步采用竞争性自适应重加权采样(CARS)、区间变量迭代空间收缩分析(iVISSA)以及变量组合群体分析—迭代保留信息变量(VCPA-IRIV)筛选信息波长。

1.4 2D-COS分析

2D-COS可用于分析样品在外部扰动作用下的光谱响应规律。该方法通过将一维光谱随扰动变量变化所包含的信息分解为同步相关谱与异步相关谱, 从而揭示不同谱变量之间的协同变化关系以及潜在的先后响应顺序, 为特征变化的解释提供补充依据。设外部扰动变量为 t , 在 $T_{min} \leq t \leq T_{max}$ 范围内采集得到光谱强度 $y(v, t)$ 。以样品体系的参考光谱 $\bar{y}(v)$ 为基准, 可构建动态光谱 $\tilde{y}(v, t)$, 公式如下:

$$\tilde{y}(v, t) = \begin{cases} y(v, t) - \bar{y}(v) & \text{for } T_{min} \leq t \leq T_{max} \\ 0 & \text{other} \end{cases} \quad (2)$$

其中, 若共有 m 条扰动序列光谱, 则可将变量 v 在各扰动点的动态强度组成向量, 公式如下:

$$\tilde{y}(v) = [\tilde{y}(v, t_1), \tilde{y}(v, t_2), \dots, \tilde{y}(v, t_m)]^T \quad (3)$$

根据Noda二维相关理论, 同步谱强度 $\Phi(v_1, v_2)$ 可表示为不同谱变量 (v_1, v_2) 的动态光谱向量的内积, 公式如下:

$$\Phi(v_1, v_2) = \frac{1}{m-1} \tilde{y}(v_1)^T \cdot \tilde{y}(v_2) \quad (4)$$

异步谱强度 $\Psi(v_1, v_2)$ 则由 Hilbert-Noda 变换引入的矩阵 N 计算, 公式如下:

$$\Psi(v_1, v_2) = \frac{1}{m-1} \tilde{y}(v_1)^T \cdot N \cdot \tilde{y}(v_2) \quad (5)$$

其中 N 为 Hilbert-Noda 矩阵, 公式如下:

$$N_{jk} = \begin{cases} 0, & j = k \\ \frac{1}{\pi(k-j)}, & j \neq k \end{cases} \quad (6)$$

在同步相关谱中, 对角线上的峰为自峰, 通常为正值, 用于表征该谱变量在扰动作用下的总体变化强度。非对角线处出现的峰为交叉峰, 其符号反映两个谱变量变化方向的一致性: 同步谱交叉峰为正表示两变量随扰动同向变化, 为负则表示反向变化。异步谱关于主对角线反对称分布, 因此主对角线上信号为零, 不出现自峰。一般而言, 只有当两个谱变量对扰动的响应不同步时, 异步谱中才会出现交叉峰^[14]。结合同步谱与异步谱交叉峰的符号, 可进一步利用 Noda 规则推断变化的先后顺序:

- (1) 若 $\Phi(v_1, v_2)$ 和 $\Psi(v_1, v_2)$ 符号相同, 则 v_1 变化发生在 v_2 之前。
- (2) 若 $\Phi(v_1, v_2)$ 和 $\Psi(v_1, v_2)$ 符号相反, 则 v_1 变化发生在 v_2 之后。

1.5 模型建立

本研究分别构建了偏最小二乘回归(PLSR)、随机森林回归(RF)以及卷积神经网络回归(CNN)3种模型。PLSR 作为一种经典的线性回归方法, 核心思想是将原始光谱变量压缩为少数潜变量, 最大化保留光谱信息的同时增强与变量的相关性, 在样本量相对有限的同时仍能保持稳定的拟合能力^[15]。RF 通过多棵回归树的集成平均实现预测, 可以有效刻画光谱与变量间的非线性关系, 对异常值与噪声具有较强鲁棒性。CNN 属于深度学习方法, 能够通过卷积核在光谱序列上提取局部谱段特征, 学习复杂的非线性映射关系。本研究构建的 CNN 模型架构如图 1 所示。简单来说, 采用两层卷积提取光谱的局部特征, 并结合批量归一化(Batch normalization)与 ReLU 激活函数提升训练稳定性与非线性表达能力; 同时引入最大池化(Max pooling)以降低维度、抑制噪声并增强特征。为进一步缓解过拟合, 网络在第二个卷积模块后加入 Dropout 层(Dropout=0.1)。随后, 将卷积提取到的高维特征通过全连接层映射到单一输出节点, 最终由回归层(Regression layer)输出目标变量的预测值。详细模型参数见表 1。

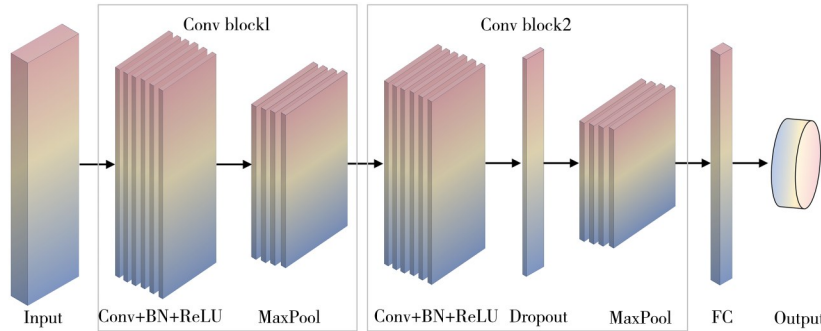


图 1 CNN 模型的结构流程图

Fig. 1 Structural diagram of the CNN model

Input: input layer; Conv: convolution layer with 2×1 kernel and 32 filters; BN: batch normalization layer; ReLU: activation layer; Pool: max pooling layer with 2×1 window; Dropout: dropout rate of 0.1; FC: fully connected layer with one output neuron; Output: output layer (regression)

1.6 模型预测结果可视化

基于最优预测模型对高光谱图像进行像素级预测, 获得 PGK1 与 PKM2 的伪彩色图像, 用于直观展示不同贮藏时间点样品的空间分布及其变化趋势。

1.7 数据处理与统计分析

模型性能采用决定系数(R^2)、校正集、验证集和预测集的均方根误差(分别为 RMSEC、RMSEV 和 RMSEP)以及相对分析误差(RPD)进行评价。一般而言, R^2 越高, RMSE 越低且 RPD 越大, 表明模型拟

合效果与预测性能越好。

表1 CNN模型结构与训练参数
Table 1 Structural and training parameters of the CNN model

Parameter	Value	Parameter	Value
Input data format	(176, 1, 1)	Initial learning rate	0.001
Kernel size	(2, 1)	Max epochs	600
Number of filters	32	Mini-batch size	32
Pooling type	max	Shuffle strategy	every-epoch
Pooling size	(2, 1)	Dropout rate	0.1
Pooling stride	2	L2 regularization	0.1
Activation function	ReLU	Learn rate schedule	piecewise
Batch normalization	True	Learn rate drop period	100
Optimizer	Adam	Learn rate drop factor	0.5
Validation frequency	20	Early stopping patience	30

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

$$SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2} \quad (9)$$

$$RPD = \frac{SD}{RMSEP} \quad (10)$$

其中, \bar{y} 为参考值的平均值, n 为样本数量, y_i 与 \hat{y}_i 分别表示第 i 个样本的实测值与预测值, SD 为参考值的标准差。

进行数据处理及建模的计算机配置如下: Intel Core i5-13500HX, 16 GB RAM, NVIDIA GeForce RTX 4060, Windows 11 64-bit。数据分析与模型构建在 MATLAB R2025a 与 The Unscrambler X 10.4 中完成, 图形绘制使用 Origin 2022a。深度学习模型基于 MATLAB Deep Learning Toolbox 实现, 并通过 Parallel Computing Toolbox 启用 GPU 加速。所有数据均以均值±标准差表示。采用单因素方差分析对不同贮藏时间下 PGK1 和 PKM2 水平进行显著性分析; 当总体差异显著时, 进一步采用 Tukey 多重比较方法对组间差异进行检验。显著性水平设定为 $p < 0.05$ 。

2 结果与讨论

2.1 PGK1 和 PKM2 免疫反应性水平分析

表2显示, PGK1 与 PKM2 在宰后贮藏过程中总体呈现“先降后升”的变化趋势, 但两者的变化幅度并不一致。屠宰后血供中断、氧供应迅速终止, 肌肉能量代谢由有氧氧化转向以无氧糖酵解为主, 伴随 ATP 周转、乳酸累积与 pH 值下降, 这是驱动宰后品质形成的关键生化基础之一^[16]。PGK1 与 PKM2 均为糖酵解中与 ATP 生成直接相关的核心酶, 其水平变化可能影响糖原消耗、ATP 生成及 pH 值变化进程, 并进一步关联肌原纤维蛋白降解等肉质变化^[17]。宰后早期 PGK1 与 PKM2 水平同步降低, 可能与 pH 值下降及环境变化引起部分肌浆蛋白构象改变, 从而降低提取率与抗体识别效率有关^[18]。随着贮藏推进, 蛋白氧化/变性与蛋白降解过程持续发生, 肌肉结构逐步松解, 蛋白分布与可溶性发生再调整^[19], 可能使部分代谢酶在可溶组分中的“可检测性”提高, 因而在 72~120 h 出现回升。此外, PGK1 下降幅度显著大于 PKM2, 提示两者在宰后过程中对环境变化的敏感性、结构稳定性, 或与肌原纤维的相互作用方式可能存在差异。因此, PGK1 与 PKM2 水平变化所伴随的色素状态、水分分布和组织散射响应并不完全一致, 这使其在 HSI 中表现出不同的光谱映射关系, 为分别建立两种酶的预测模型提供了理论依据。

2.2 原始光谱数据分析

图2展示了原始光谱曲线及6个贮藏时间点的平均光谱。总体而言, 样品在 Vis-NIR 波段的光谱轮廓基本一致, 但不同样本间反射强度存在一定波动。该现象在肉品光谱研究中较为常见, 可能与个体

差异有关,也可能反映宰后贮藏过程中肌肉成分与微观结构持续变化^[13];同时,样品表面状态与采集条件差异亦可能引入测量波动^[10]。随着贮藏时间延长,平均光谱在可见光区和近红外区均呈现出可区分的时序差异:0 h的整体反射率相对偏低,随后随时间推移逐渐变化,说明宰后生化过程与组织结构演变共同影响光谱响应^[20]。

表2 PGK1和PKM2的酶免疫反应性水平(U/mL,试剂盒定义)
Table 2 Enzymatic immunoreactive level(U/mL, kit-defined) of PGK1 and PKM2

Enzyme	Enzyme protein level(U/mL, kit-defined)					
	0 h	12 h	24 h	48 h	72 h	120 h
PGK1	464.01 ± 60.44 ^a	401.40 ± 17.02 ^b	334.60 ± 16.85 ^c	118.68 ± 16.77 ^e	267.51 ± 19.45 ^d	315.07 ± 51.06 ^c
PKM2	473.47 ± 12.86 ^a	414.13 ± 29.10 ^b	360.69 ± 30.41 ^c	301.08 ± 15.06 ^d	360.11 ± 30.34 ^c	399.04 ± 61.35 ^b

Different superscript letters within the same row indicate significant differences among storage times ($p < 0.05$) (同一行中不同上标字母表示不同贮藏时间之间差异显著 ($p < 0.05$))

从吸收归属来看, Vis-NIR 区间的光谱特征主要由水分、蛋白质与脂质相关的 C-H、N-H、O-H 基团吸收以及组织散射效应共同决定^[21]。在可见光区, 约 415 nm 及 540~580 nm 附近的吸收通常与血红素色素相关(如不同形态肌红蛋白/血红蛋白的吸收差异)^[22]; 此外, 430~560 nm 区间的谱形还可能受到组织结构差异所引起的散射与基线变化叠加影响。在近红外区, 约 720~760 nm 附近的响应往往同时受水的弱吸收、色素吸收与散射变化的影响, 而 980 nm 附近的特征通常归因于水的 O-H 倍频或合频吸收^[23]。值得注意的是, 48 h 时 730~980 nm 区间反射率相对升高, 提示组织水分结合状态及空间分布可能发生调整, 该现象与宰后酸化及系水力变化趋势相吻合^[24]。

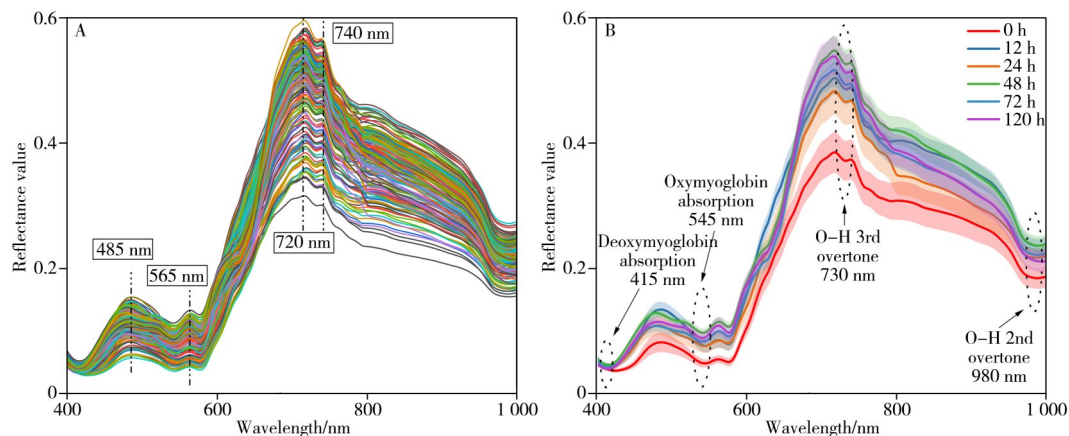


图2 原始光谱(A)及6个贮藏时间点的平均光谱(B)

Fig. 2 Raw spectra(A) and mean spectra at six storage-time points(B)

2.3 2D-COS解析

本研究分别以PGK1与PKM2水平对样品光谱进行排序,构建扰动序列 t ,并计算同步与异步二维相关谱。结果显示,两种排序方式得到的二维谱特征高度一致,主要表现为自相关峰位置及主要交叉峰的分布基本相同。这可能是因为二者都是糖酵解关键酶,均与宰后肌肉的能量代谢、酸化程度及其引发的肉色变化和水分迁移等密切相关,本质上都反映同一宰后代代谢过程驱动的综合光谱响应,所以具有较高一致性。因此正文以PGK1驱动二维谱作为代表进行解析(图3A~D)。根据Noda二维相关理论,二维相关谱的横纵坐标均为波长变量,图中等高线/颜色表示相关强度。同步谱对角线上的自相关峰用于表征相应波段对扰动的敏感程度,而非对角线处的交叉峰反映两波段在扰动作用下的协同变化关系及其变化方向。异步谱交叉峰用于表征两波段变化的先后顺序。

在同步2D-COS中(图3A),可观察到5个显著自相关峰,分别位于476、562、605、715和800 nm,表明这些波段对酶水平变化最为敏感。PGK1作为糖酵解过程中与ATP生成密切相关的关键酶,其水平变化通常伴随肌肉pH值、氧化还原状态及蛋白质构象的调整,进而影响肌红蛋白不同形态之间的转化、水分结合状态以及组织结构的散射特性。具体而言,476 nm位于蓝光区,主要反映肌红蛋白体系氧化还原程度变化下引起的光谱响应,对色素状态变化较敏感;562 nm通常与氧合肌红蛋白、脱氧肌红蛋白与高铁肌红蛋白之间的相对比例相关,可综合表征肌红蛋白形态转化对可见光反射的影响;

605 nm 位于红光敏感区, 更强调与肌红蛋白氧化导致的褐变趋势; 715 nm 常用于表征色素吸收减弱后由光散射主导的谱形变化, 同时也可能受到样本含水状态与组织结构散射差异的影响^[25]; 800 nm 通常与水分含量及结合状态有关, 其响应主要来源于O—H、N—H和C—N等含氢基团的倍频与合频吸收。由此可见, PGK1水平变化与宰后代谢强度变化密切相关, 并进一步影响肉色、肌红蛋白状态及水分结构特征, 最终在可见-近红外区域形成可检测的综合光谱响应。

异步2D-COS(图3B)中, 在(476~562)、(476~715)、(605~715)nm处出现明显的负交叉峰, 而(562~605)、(562~800)、(715~800)nm的交叉峰为正。异步交叉峰的出现表明相应波段对扰动的响应并非完全同步。结合同步/异步交叉峰符号并依据Noda规则, 可推断各敏感波段随酶水平变化的响应顺序为: 562 nm~605 nm~715 nm~800 nm~476 nm(表3)。该结果表明, 宰后代谢变化可能首先引起与肉色相关的主导吸收特征变化, 随后伴随肌肉微结构松解和水分重新分布, 最终表现为色素氧化还原平衡的进一步调整。这表明, PGK1表征的代谢变化与色素、水分及结构特征的响应具有一定的阶段性特征。

图3C与图3D分别为同步与异步相关谱的对应等高线图的立体化表达。三维图在保留峰位与符号信息一致性的基础上, 更直观展示了相关强度的空间起伏特征。总体而言, 2D-COS不仅能够识别对酶水平变化最敏感的特征波段, 还可揭示不同波段之间的协同变化及先后响应关系, 从而更清晰地刻画光谱在扰动条件下的动态演变模式。

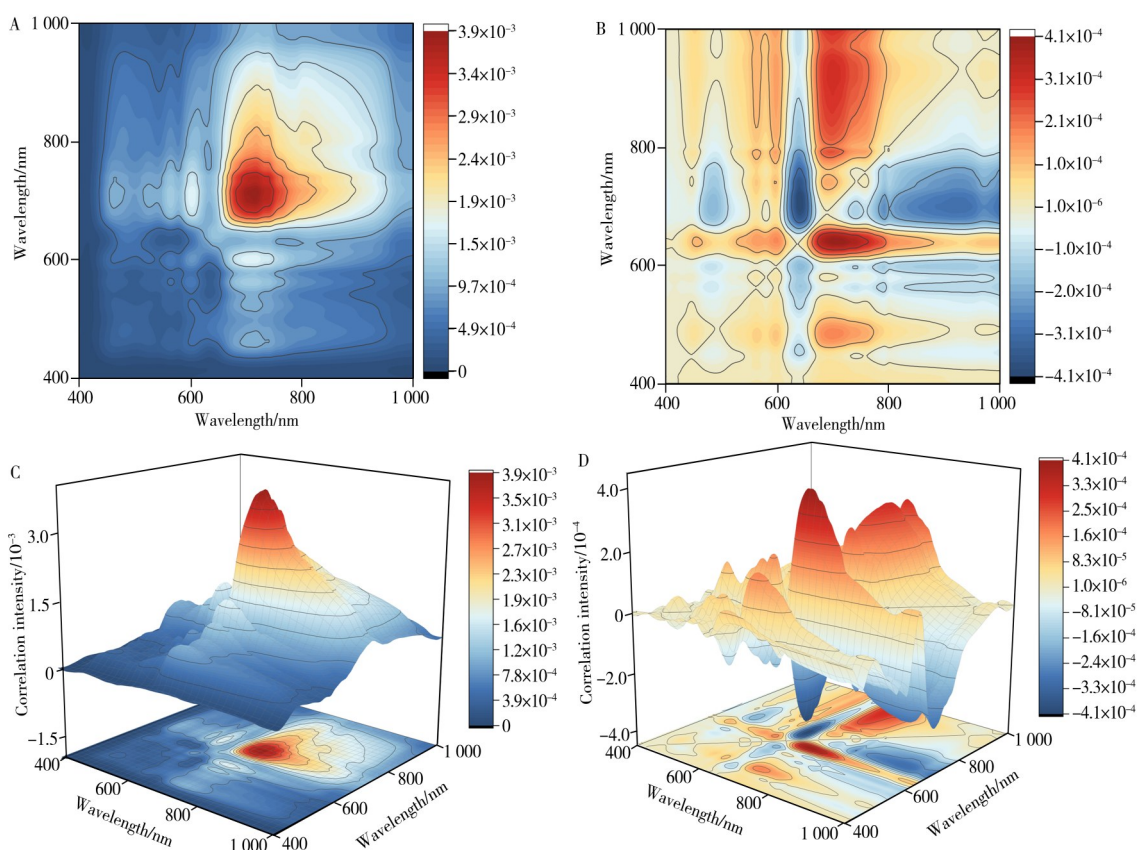


图3 2D-COS结果

Fig. 3 Results of 2D-COS

A. 2D contour map of the synchronous correlation spectrum; B. 2D contour map of the asynchronous correlation spectrum; C. 3D surface plot of the synchronous correlation spectrum; D. 3D surface plot of the asynchronous correlation spectrum(A. 同步相关谱二维等高线图; B. 异步相关谱二维等高线图; C. 同步相关谱三维表面图; D. 异步相关谱三维表面图)

2.4 光谱预处理

为降低仪器噪声、基线漂移及散射效应对建模的影响, 本研究比较了经3种常用预处理方法SG、Baseline与SNV预处理后的光谱图像, 如图4所示。在不同预处理与建模方法组合下, PGK1与PKM2的预测性能存在差异(表4)。总体而言, 3类模型在两种酶指标上均获得了较好的预测效果, 除个别组

合外, 预测集 R_p^2 主要分布在 0.82~0.88, RPD 多数高于 2.4, 表明模型具备一定的定量预测能力。由于不同模型对光谱干扰的敏感性不同, 为各模型匹配更合适的预处理策略是必要的^[26]。

表 3 同步二维相关谱中交叉峰的符号(Φ , 括号外)及异步二维相关谱中交叉峰的符号(Ψ , 括号内)
Table 3 Signs of cross peaks in synchronous(Φ , outside brackets) and asynchronous spectra(Ψ , inside brackets)

Synchronous		Asynchronous			
476	476	562	605	715	800
476	+	+(-)	+(-)	+(-)	+(-)
562		+	+(+)	+(+)	+(+)
605			+	+(+)	+(+)
715				+	+(+)
800					+

对于 PGK1, 3 种模型对应的最优预处理方法并不一致, 说明其光谱信号易受多种因素干扰。具体而言, SNV 更偏向削弱散射与光程差异带来的尺度变化, 有助于增强 PLSR 的线性稳定性^[27]; RF 对整体基线漂移非常敏感, 而 Baseline 能显著降低模型风险, 改善模型泛化^[6]; 噪声会干扰 CNN 的卷积特征提取, 而 SG 有利于提高数据稳定性并提升预测表现^[28]。相比之下, PKM2 在 3 类模型中均表现出 Baseline 优于其他预处理方法的趋势, 表明其预测性能主要受基线漂移影响。尤其是 RF 模型在预处理后模型 R_p^2 从 0.676 0 提高到 0.831 0。总的来说, 由于 PGK1 与 PKM2 的水平值与光谱信号之间的关联模式不同, 预处理对“有效信息保留”与“无关干扰抑制”的作用在不同目标变量上表现出差异。

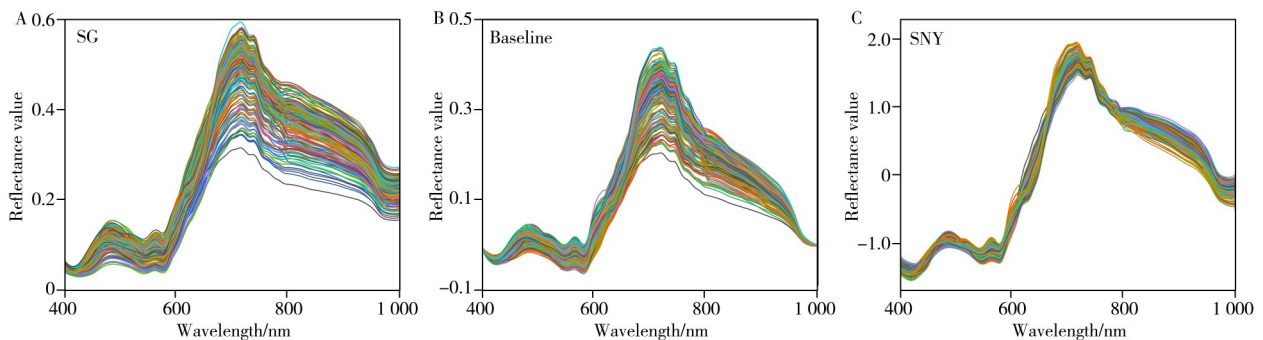


图 4 不同预处理方法下的光谱对比

Fig. 4 Spectral comparison under different preprocessing methods

表 4 原始光谱与最佳预处理方法下 3 种模型对 PGK1 和 PKM2 的预测性能比较

Table 4 Comparison of prediction performance for PGK1 and PKM2 using three models with raw spectra and the optimal preprocessing method

Enzyme	Model	Method	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
PGK1	PLSR	Raw	0.856 9	41.679 8	0.851 2	47.406 2	0.857 5	44.100 6	2.686 9
		SNV	0.864 5	40.565 7	0.846 6	48.127 1	0.864 9	42.945 2	2.759 2
	RF	Raw	0.852 5	42.314 1	0.802 9	54.546 8	0.823 0	49.150 1	2.410 9
		Baseline	0.891 3	36.331 9	0.849 5	47.664 2	0.869 3	42.244 1	2.805 0
	CNN	Raw	0.902 1	34.265 1	0.841 6	48.848 0	0.864 6	42.946 2	2.765 1
		SG	0.924 8	30.153 8	0.848 4	47.798 9	0.870 7	41.994 6	2.825 0
PKM2	PLSR	Raw	0.904 6	17.674 3	0.849 1	19.688 9	0.853 9	22.330 4	2.654 0
		Baseline	0.888 4	19.116 0	0.846 4	19.868 0	0.881 8	20.081 5	2.951 2
	RF	Raw	0.888 6	19.103 0	0.742 2	25.738 5	0.676 0	33.247 0	1.782 6
		Baseline	0.925 1	15.660 0	0.805 1	22.380 1	0.831 0	24.014 8	2.467 8
	CNN	Raw	0.900 9	17.970 0	0.854 5	19.326 6	0.868 2	21.139 2	2.821 9
		Baseline	0.924 7	15.647 8	0.875 6	17.862 8	0.882 0	20.048 2	2.960 9

2.5 特征波长提取

由于光谱具有高维、强共线性和信息冗余等特点, 且肉样易受表面散射、基线波动和噪声的影响从而引入与目标无关的信息, 直接使用全波段建模易导致模型复杂度升高, 泛化能力下降。因此, 本研究采用 CARS、iVISSA 和 VCPA-IRIV 三种特征提取方法筛选关键信息波长, 实现降维去冗余、提升模型稳健性和可解释性。

对 PGK1 来说(图 5A), 3 种特征提取方法在变量压缩强度上差异明显, 分别筛选了 67、29 和 31 个

波长,说明不同算法对如何保留有用信息和去除数据冗余的权衡不同。更关键的是,3种方法的共同重叠波长共有11个(505.71、525.42、528.72、541.95、555.24、698.25、701.75、740.31、772.20、855.01、898.99 nm),说明PGK1的有效光谱信息在不同筛选条件下仍能被反复识别,为后续提高模型性能提供了理论支撑。PKM2的筛选结构呈现不同的特点(图5B):CARS仅筛选了15个波长,iVISSA和VCPA-IRIV均筛选了29个波长,但共同重叠波长仅有3个(528.72、715.71、851.38 nm),说明PKM2的特征提取对算法策略更敏感,不同方法能从不同波段组合中获取有用信息。总体来说,PGK1具有更强的关键波段一致性,而PKM2的筛选结果差异更大,表明两类酶与光谱信号的关联结构存在差别。此外,3种特征提取方法所关注的波段区间与上述2D-COS所识别的特征峰大体一致,进一步说明特征选择的合理性。

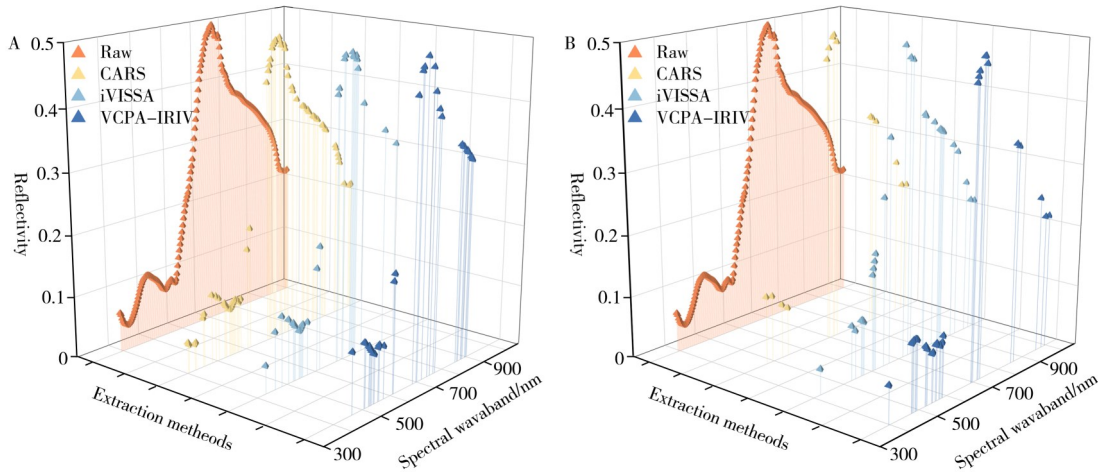


图5 不同特征提取方法筛选波长的可视化对比

Fig. 5 Visualization of selected wavelengths obtained by different feature selection methods
A. PGK1; B. PKM2

2.6 特征波长建模结果

基于特征提取后的光谱变量建模结果显示,PGK1与PKM2均能实现较好的定量预测,但不同模型对特征提取策略的响应存在显著差异。总体上,PLSR在降维后性能波动较小,体现出较稳定的线性拟合能力;CNN在筛选关键波长后往往能够进一步提升预测精度与泛化表现;而RF更依赖全光谱信息,变量过度压缩常导致预测能力下降,提示对集成型非线性模型而言,保留更完整的谱段信息可能更有利于捕获复杂的光谱与指标间的映射关系。

对于PGK1(表5),CNN与PLSR的整体预测性能高于RF,且CNN对特征提取方法更为敏感。与全谱建模相比,3种特征提取方法均在不同程度上提高了CNN的预测性能,其中VCPA-IRIV-CNN表现最优,获得最高 R_p^2 (0.8920)和最低RMSEP(38.3653),RPD提升至3.0935;CARS-CNN也取得较优结果($R_p^2=0.8858$,RMSEP=39.4380,RPD=3.0105)。这表明筛选后的特征波长能够帮助CNN更聚焦于与PGK1相关的关键谱段,从而提升预测稳定性。PLSR在特征提取后性能得到提升,且整体结果较为稳健,其中CARS-PLSR达到 $R_p^2=0.8904$,RMSEP=38.6884,RPD=3.0628。相比之下,RF在全光谱数据上表现最好,特征筛选后反而下降,说明过度压缩变量可能削弱其对非线性信息的表征能力。

对于PKM2(表6),不同模型对特征提取的响应趋势与PGK1基本一致,CNN仍是综合表现较优的算法,且最优组合同样出现在CNN与VCPA-IRIV的结合上($R_p^2=0.9030$,RMSEP=18.1778,RPD=3.2659)。PLSR中CARS相比全谱仅呈小幅改善($R_p^2=0.8851$,RMSEP=19.8040,RPD=2.9926),而iVISSA的性能提升相对较弱。与PGK1一致,RF在PKM2上仍以全光谱组合最优,经过特征提取后性能明显下降,再次说明RF对全谱变量的保留更为依赖。

综合来看,特征提取对不同模型性能的影响不同:在两种酶的预测中,CNN与筛选后的特征波长(尤其是VCPA-IRIV)适配性最高,可稳定提升预测精度,体现出较强的预测能力;PLSR在特征维度降低后通常仍能保持较稳健的线性拟合;而RF更适合利用全光谱数据,过强的变量筛选可能造成有效

信息丢失并削弱模型泛化能力。

表5 基于特征提取后光谱变量的 PGK1 预测模型性能比较

Table 5 Comparison of prediction performance for PGK1 models using feature-extracted spectral variables

Enzyme	Model	Method	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
PGK1	PLSR	Full spectrum	0.864 5	40.565 7	0.846 6	48.127 1	0.864 9	42.945 2	2.759 2
		CARS	0.858 7	41.424 7	0.853 3	47.069 7	0.890 4	38.688 4	3.062 8
		iVISSA	0.867 6	40.094 9	0.869 3	44.428 8	0.885 6	39.520 9	2.998 3
		VCPA-IRIV	0.837 7	44.394 2	0.878 9	42.754 1	0.886 5	39.354 4	3.010 9
	RF	Full spectrum	0.891 3	36.331 9	0.849 5	47.664 2	0.869 3	42.244 1	2.805 0
		CARS	0.864 1	40.629 6	0.857 3	46.416 3	0.847 4	45.642 3	2.596 1
		iVISSA	0.862 0	40.931 8	0.819 6	52.196 2	0.809 0	51.062 1	2.320 6
		VCPA-IRIV	0.824 4	46.178 6	0.771 3	58.765 1	0.779 2	54.896 9	2.158 5
	CNN	Full spectrum	0.924 8	30.153 8	0.848 4	47.798 9	0.870 7	41.994 6	2.825 0
		CARS	0.941 1	26.674 1	0.889 1	40.820 6	0.885 8	39.438 0	3.010 5
		iVISSA	0.905 2	33.851 0	0.869 5	44.335 0	0.874 9	41.315 5	2.869 4
		VCPA-IRIV	0.904 8	33.525 2	0.855 6	46.404 4	0.892 0	38.365 3	3.093 5

表6 基于特征提取后光谱变量的 PKM2 预测模型性能比较

Table 6 Comparison of prediction performance for PKM2 models using feature-extracted spectral variables

Enzyme	Model	Method	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
PKM2	PLSR	Full spectrum	0.888 4	19.116 0	0.846 4	19.868 0	0.881 8	20.081 5	2.951 2
		CARS	0.879 8	19.841 4	0.879 4	17.601 0	0.885 1	19.804 0	2.992 6
		iVISSA	0.837 3	23.085 6	0.779 0	23.829 8	0.859 6	21.889 3	2.707 5
		VCPA-IRIV	0.871 6	20.506 4	0.867 3	18.467 4	0.882 5	20.025 3	2.959 5
	RF	Full spectrum	0.925 1	15.660 0	0.805 1	22.380 1	0.831 0	24.014 8	2.467 8
		CARS	0.866 7	20.892 9	0.644 9	30.205 9	0.732 8	30.196 1	1.962 7
		iVISSA	0.893 2	18.703 2	0.656 6	29.703 5	0.782 7	27.228 0	2.176 6
		VCPA-IRIV	0.903 2	17.803 7	0.696 7	27.917 9	0.807 4	25.634 8	2.311 9
	CNN	Full spectrum	0.924 7	15.647 8	0.875 6	17.862 8	0.882 0	20.048 2	2.960 9
		CARS	0.886 9	19.220 5	0.808 6	22.168 0	0.881 8	20.032 8	2.973 6
		iVISSA	0.914 0	16.762 8	0.816 5	21.704 2	0.890 4	19.330 0	3.068 5
		VCPA-IRIV	0.904 9	17.634 1	0.805 0	22.371 9	0.903 0	18.177 8	3.265 9

2.7 PGK1 和 PKM2 水平可视化

图6展示了不同贮藏时间羊肉样品中PGK1与PKM2水平的可视化预测分布。根据色标,红色表示预测水平较高,蓝色表示较低。两种酶的空间分布均呈现明显的非均匀性,样品内部高、低值区域交错分布,表明相关光谱响应在肌肉组织中具有显著空间异质性。这一现象与肉类基质的复杂性密切相关:肌肉内部的水分、蛋白质、脂质和结缔组织等成分分布不均,导致生化信号在空间上呈现不均一特征。同时提示单点取样难以充分反映整体水平,而成像结果能够更直观地揭示其空间变化规律。随贮藏时间由0 h延长至120 h,两种酶的预测值整体呈现“先降低后回升”的变化趋势。该趋势可能与宰后贮藏期间一系列生化与理化过程的共同作用有关,例如糖酵解进程推进、蛋白质结构变化以及水分迁移等,这些因素均可能影响酶相关信号强度及其空间分布格局。总体而言,像素级可视化方法能够提供更丰富的空间信息,有助于揭示样品中位置特异性的变化,为理解贮藏过程中指标的区域差异与演变提供了直观依据。

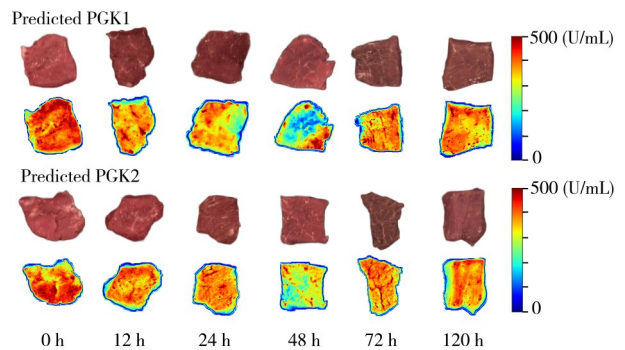


图6 不同贮藏时间羊肉中PGK1和PKM2水平的可视化分布图

Fig. 6 Visualization maps of PGK1 and PKM2 levels in lamb at different storage times

3 结论

本研究验证了利用HSI结合深度学习无损预测滩羊肉中PGK1和PKM2免疫反应性水平的可行性。通过2D-COS识别出5个敏感特征波段(476、562、605、715和800 nm),解析出光谱响应与酶水平变化之间的关键关联:一方面表现为肉色相关的肌红蛋白状态变化,另一方面体现为水分结合状态与组

织结构散射差异的综合影响,为后续特征提取与建模提供了可解释的波段依据。经过合适的预处理方法与特征提取后,VCPA-IRIV-CNN在PGK1($R_p^2=0.8920$, RMSEP=38.3653, RPD=3.0935)和PKM2($R_p^2=0.9030$, RMSEP=18.1778, RPD=3.2659)预测上均取得了最佳结果。进一步进行伪彩色可视化,直观展示了酶水平的空间异质性与变化趋势。值得注意的是,2D-COS解析得到的关键敏感波段与特征提取结果大体一致,说明基于扰动响应识别的谱段与数据驱动筛选的有效变量具有一致性,从而增强了模型的可解释性。总体而言,本研究证明了无损预测糖酵解酶水平的可行性,为滩羊肉代谢关键酶的快速无损评估与品质过程监测提供了技术路径。后续可在更复杂的实际生产场景中进一步扩展样本与验证模型稳健性,以推动方法的工程化应用。

参考文献:

- [1] Zhang L L, Gao S, Ma M Y, Xiao Y Y, Yu T, Jiang Q J, Li Y L, Luo R M. *Food Sci.* (张玲玲, 高爽, 马梦瑶, 肖玉莹, 于婷, 江启健, 李亚蕾, 罗瑞明. 食品科学), **2026**, 47(5): 50-61.
- [2] Huang C Y, Zhang D Q, Wang Z Y, Zhao Y X, Blecker C, Li S B, Zheng X C, Chen L. *Food Chem.*, **2023**, 427: 136739.
- [3] Huang C Y, Xiang C, Wang F Z, Blecker C, Wang Z Y, Chen L. *Food Front.*, **2024**, 5(4): 1629-1641.
- [4] Ma X, Li N, Wang R F, Ma J X, Zhu N H, Li T T, Zhang Z X, Li H F, Wang S L, Zhang H H. *Chemom. Intell. Lab. Syst.*, **2026**, 268: 105575.
- [5] Liu S, Cui J R, Lv Y, Wang S L. *Food Res. Int.*, **2025**, 205: 116000.
- [6] Xi Q B, Chen Q M, Ahmad W, Pan J, Zhao S G, Xia Y, Ouyang Q, Chen Q S. *Food Chem.*, **2025**, 481: 143997.
- [7] Yi W G, Zhao X Y, Yun X Y, Wang S L, Dong T. *Food Res. Int.*, **2025**, 203: 115905.
- [8] Noda I. *Appl. Spectrosc.*, **1993**, 47(9): 1329-1336.
- [9] Penf S, Wang F P, Wei D B, Wang C P, Ma H J, Du Y G. *J. Environ. Sci.*, **2025**, 147: 200-216.
- [10] Liu S J, Dong F J, Hao J, Guo J H, Wang S L, Luo R M, Lv Y, Cui J R. *Spectrochim. Acta A*, **2023**, 291: 122342.
- [11] Hao J, Cui J R, Liu S J, Lv Y, Zhang Z X, Wang S L. *Food Chem.*, **2025**, 493: 145727.
- [12] Cui J R, Cao S B, Hao J, Zhang Y, Gao S, Li J S, Fatimah, Wu L G. *Ind. Crops Prod.*, **2025**, 236: 121909.
- [13] León-Ecay S, Insausti K, López-Maestresalas A, Prieto N. *Meat Sci.*, **2025**, 228: 109893.
- [14] Dong F J, Bi Y Z, Hao J, Liu S J, Yi W G, Yu W J, Lv Y, Cui J R, Li H, Xian J H, Chen S C, Wang S L. *Food Chem.*, **2024**, 440: 138040.
- [15] Xiao D, Yu P, Guo X, Luan M Q, Li F X, Yang J, Zhou J H. *J. Instrum. Anal.* (肖丹, 于苹, 郭雪, 栾美琪, 李飞羽, 杨健, 周骏辉. 分析测试学报), **2026**, 45(3): 591-599.
- [16] Ma C, Zhang W G, Zhang J, Zhou L, Xing L J, Liu R. *J. Adv. Res.*, **2025**.
- [17] Huang C Y, Zhang D Q, Blecker C, Zhao Y X, Xiang C, Wang Z Y, Li S B, Chen L. *Food Chem. X*, **2024**, 21: 101125.
- [18] Yang N, Liang X L, Cao J X, Zhang Q, Tan Y Z, Xu B T, Yang Y L, Wang Y, Yang Q J, Liu H, Liu J. *Meat Sci.*, **2022**, 185: 108723.
- [19] Guo Z B, Chen C, Ma G Y, Yu Q, Zhang L. *Food Chem. X*, **2023**, 20: 100891.
- [20] Ismail A, Park S, Kim H J, Hong G, Kim G, Jo C. *LWT-Food Sci. Technol.*, **2025**, 216: 117302.
- [21] Ren W J, Jiang Q Q, Qi W L. *Chem. Biol. Technol. Agric.*, **2025**, 12(1): 32.
- [22] Zheng X C, Chen L, Li X, Zhang D Q. *Foods*, **2023**, 12(2): 300.
- [23] Feng Y J, Dong F J, Wu C, Cui J R, Wang G X, Zhang Z X, Cao S M, Liu S J, Li H, Shi Y K, Yu J G, Wang S L. *Food Res. Int.*, **2025**, 204: 115877.
- [24] Ma C, Zhang J, Zhang R, Zhou L, Ni L X, Zhang W G. *Food Chem. X*, **2024**, 24: 101913.
- [25] Waga M, Nakade K. *Meat Sci.*, **2025**, 219: 109664.
- [26] Liu Z Y, Li L N, Xu W M, Xu X S, Cui Z C, Jia L C, Lv W H, Shen Z H, Shu R. *Remote Sens.*, **2023**, 15(13): 3311.
- [27] Zhang W, Kasun L C, Wang Q J, Zheng Y J, Lin Z P. *Sensors*, **2022**, 22(24): 9764.
- [28] Wang Z T, Zhang T S, Ding H Y, Zhao S Y, Fan Y, Qi D, Dai Y J, Liu Z Y. *Microchem. J.*, **2025**, 218: 115161.

(责任编辑: 盛文彦)