

结合高光谱图像的光谱和纹理信息预测羊肉 可溶性蛋白和 GSH 含量

乔 芦, 王松磊*, 郭建宏, 贺晓光

宁夏大学食品与葡萄酒学院, 宁夏 银川 750021

摘 要 可溶性蛋白和谷胱甘肽(GSH)是羊肉重要的生理生化指标,是衡量机体抗氧化能力大小的重要因素,传统检测方法程序复杂,检测费时。为此应用可见-近红外(400~1 000 nm)高光谱成像技术实现可羊肉可溶性蛋白和还原性谷胱甘肽(GSH)含量无损、快速检测。首先,对采集的180个羊肉样本的原始光谱信息采用4种方法进行预处理,再运用竞争自适应加权算法(CARS)、区间变量迭代空间收缩算法-迭代和保留信息变量法(iVISSA-IRIV)进行特征波段的提取。同时使用灰度共生矩阵法(GLCM)提取贡献率最高的主成分图像的纹理信息。最后将优选出的预处理方法和特征波长信息作为光谱信息和光谱-纹理融合信息分别结合多元线性回归(MLR)、最小二乘支持向量机(LS-SVM)模型建立羊肉可溶性蛋白和谷胱甘肽含量的预测模型。结果显示未经预处理的原始光谱建立的羊肉可溶性蛋白含量 PLSR 模型效果最佳,其 R_c 和 R_p 分别为 0.875 7和0.854 7;采用 SNV 法预处理后光谱建立的羊肉 GSH 含量 PLSR 模型效果最佳,其 R_c 和 R_p 分别为 0.804 8和0.826 5。利用 iVISSA-IRIV 共筛选出 31 个特征波长,建立的羊肉可溶性蛋白 LS-SVM 模型的 R_c 和 R_p 最优,分别为 0.914 6和0.881 8;同时利用 iVISSA-IRIV 筛选出 29 个特征波长,建立的羊肉 GSH-MLR 模型的 R_c 和 R_p 最优,分别为 0.844 6和0.870 5。最终经光谱特征信息和图谱信息融合模型对比发现,建立 iVISSA-IRIV-LS-SVM 模型对羊肉可溶性蛋白预测效果最佳,其 R_c 和 R_p 分别为 0.914 6和0.881 8;利用 SNV-iVISSA-IRIV 法提取的光谱特征信息与纹理信息融合建立的 MLR 模型为预测羊肉 GSH 含量的最优模型,其 R_c 和 R_p 分别为 0.849 5和0.890 4。利用最优 iVISSA-IRIV-LS-SVM 和 iVISSA-IRIV-MLR 模型和成像处理方法,结合伪彩色图像直观地表示羊肉样本的可溶性蛋白和 GSH 含量的空间分布情况。研究结果表明利用高光谱图像的光谱和纹理信息能够用来预测羊肉可溶性蛋白和 GSH 含量。

关键词 高光谱成像技术;特征波长筛选;可溶性蛋白和 GSH 含量;纹理特征;可视化

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)01-0176-08

引 言

谷胱甘肽(GSH)^[1-2]是羊肉体内重要的水溶性抗氧化剂,是一种普遍存在的内源性生物硫醇,在动物体内提供解毒代谢的作用,因而 GSH 是维持生命系统抗氧化能力的重要指标,大多数酶类的代谢都要可溶性蛋白质参与,同时也是测定 GSH 含量的一个重要指标,因此可以将可溶性蛋白和 GSH 含量当做肉类新鲜度检测指标。当前常规检测方法为分光光度计法,此方法虽然可靠,但耗时、费力。故选用一种快速准确的方法来定量检测羊肉中可溶性蛋白和 GSH

含量非常必要。

高光谱成像技术可以提供样本的光谱和空间信息,它作为一种无损、实时的食品质量分析和检测工具越来越受到人们的重视^[3-4]。其纹理信息在肉品领域应用逐渐变广。Xiong^[5]等应用光谱和纹理数据相结合区分散养鸡和肉鸡,分类正确率达 93.33%。孙宗保等^[6]采用光谱和纹理信息鉴别冰鲜和冻融三文鱼及其冻融次数,取得了较好的效果。Lin 等^[7]利用高光谱图谱融合技术对猪肉肌内脂肪含量进行预测。Yang^[8]等采用高光谱图谱合一的特点,提取熟牛肉样本 30 个纹理信息结合反向人工神经网络(back propagation-artificial neural networks, BP-ANN)模型对熟牛肉的存储时间

收稿日期: 2020-12-23, 修订日期: 2021-03-09

基金项目: 国家自然科学基金项目(31660484)资助

作者简介: 乔 芦, 1995 年生, 宁夏大学食品与葡萄酒学院硕士研究生 e-mail: qiaol22@163.com

* 通讯作者 e-mail: wangsonglei163@126.com

和含水率进行预测。Zhang 等^[9]利用信息融合的方法预测鲑鱼的脂肪和水分, 校正预测相关系数均高达 0.95 以上。这些研究显示融合信息均优于仅光谱或图像纹理信息的结果, 这表明光谱和图像纹理数据结合不仅可以对肉品质进行定性判别, 也能够用来对理化指标定量预测分析。但据了解, 近几年国内外学者利用光谱技术对植物酶类有研究, 但尚未见光谱数据和图像纹理信息融合来确定羊肉可溶性蛋白和 GSH 含量的相关研究报道。

本研究的主要目的是利用光谱和图像纹理信息相结合来预测羊肉可溶性蛋白含量和 GSH 含量。通过采集羊肉样本的可见-近红外(400~1 000 nm)高光谱数据, 利用 4 种预处理方法 2 种变量选择工具筛选出最优预处理方法和特征波长方法; 并使用灰度共生矩阵法提取羊肉样本纹理信息, 通过基于光谱数据及其光谱和纹理融合信息建立线性和非线性预测模型预测羊肉可溶性蛋白含量和 GSH 含量, 并利用伪彩色图绘制其空间分布图。

1 实验部分

1.1 羊肉样本采集与处理

羊肉样本采自宁夏盐池县鑫海食品有限公司, 共计 60 只羊。经屠宰后, 将其胴体冷藏于 4 °C 条件下进行 48 h 排酸处理, 排酸完成后, 取 60 只羊胴体前腿肉、后腿肉、背最长肌各一块肉, 最终得到羊肉样本 180 个。为保持样本的一致性, 将羊肉制备成体积大小为(35 mm×25 mm×10 mm)的肉块, 并于 4 °C 冰箱储藏备用为后续光谱采集和化学值测定。

1.2 高光谱数据采集

高光谱图像由可见-近红外(波长范围 400~1 000 nm)高光谱成像系统获取。为消除图像模糊以及传感器暗电流的影响, 需在图像采集前进行对焦及黑白校正, 经多次实验确定的采集参数为: 扫描线长度 60 mm, 物距 380 mm, 电控位移平台初始速度 1.5 mm·s⁻¹。

1.3 羊肉可溶性蛋白和 GSH 含量测定

在获取高光谱图像后, 测定羊肉可溶性蛋白和 GSH 含量。称取 1.000 0 g 羊肉样本, 制备组织上清液, 利用紫外分光光度计测定吸光度在 595 和 420 nm 处可溶性蛋白和 GSH 含量。

1.4 光谱数据分析

1.4.1 光谱预处理

在光谱采集过程中, 由于外部环境和仪器的影响导致光谱信号容易受到大量随机噪声的干扰, 不利于提高预测模型的准确性, 故需要对光谱进行预处理^[10]。为筛选出较好的预处理方法, 采用卷积平滑(savitzky-golay, SG)、标准正态变量变换(standard normal variate, SNV)、正交信号校正(orthogonal signal correction, OSC)、去趋势法(detrend)共 4 种预处理方法。

1.4.2 特征波长提取

有效的特征波长提取能够删除与样本成分无关的信息, 提取最能代表肉样的有效信息, 缩短计算时间^[11]。采用

CARS 和 iVISSA-IRIV 提取特征波长方法来简化模型, 提升模型性能。CARS 是一种新的变量选择方法, 它可以在消除新的变量的同时有效减少共线变量对模型的影响^[12]。区间变量迭代空间收缩法(interval variable iterative space shrinkage approach, iVISSA)是一种基于交叉验证均方根误差(RMSECV)的波长间隔选择方法。它经过多次迭代, 交替使用全局和局部过程来优化间隔位置, 直至优选出最佳特征变量^[13]。迭代和保留信息变量法(iteratively retaining informative variables, IRIV)是基于二进制矩阵变换滤波器的思想提出的一种新的变量选择方法, 该方法采用 BMS 作为变量空间采样方法, 迭代保留信息变量, 直到不存在干扰变量。利用 IRIV 法对 iVISSA 法筛选出的波长进行二次提取, 形成了一种变量选择的混合方法, 解决了变量数量多的问题。

1.4.3 图像纹理信息提取

纹理是图像应用中最重要特征之一, 表示的是相邻像素之间的强度关系, 表示灰度像素之间关系的常用方法叫做灰度共生矩阵法(GLCM), 其描述了不同灰度组合在一幅图像中同时出现的频率, 常被用来提取图像纹理信息。由于不同方向和距离的图像可以构成不同的灰度共生矩阵, 故本研究中距离设置为 1, 依次取不同方向(0°, 45°, 90°和 135°)的能量(energy)、熵(entropy)、同质性(homogeneity)和相关性(correlation)四个特殊相互独立的特征来描述共生矩阵数据^[9]。从每幅图片中共选取 16 个纹理参数用于后续建模。

2 结果与讨论

2.1 可溶性蛋白和 GSH 含量描述统计

如图 1 所示, 可溶性蛋白和 GSH 含量在不同部位之间存在显著差异。羊肉背最长肌的可溶性蛋白含量最高, 而 GSH 含量最低, 这种差异可能是由不同部位之间的理化特征引起的。也可能与羊不同部位活动量不同有关, 由于前腿和后腿相对于背最长肌运动较多, 运动会增加有氧代谢, 故 GSH 含量相对较高。

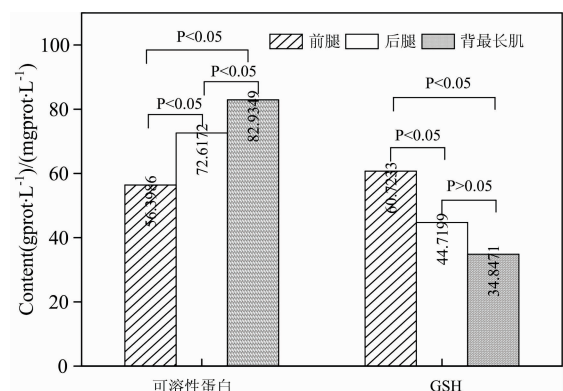


图 1 不同部位可溶性蛋白和 GSH 含量统计分析图

Fig. 1 Statistical analysis of soluble protein and GSH contents in different parts

对提取到的光谱数据进行建模, 需要仔细选择校正集, 使其能够代表用于校准的样本。因此, 180 个羊肉样本经别

除 9 个异常样本后,按照 3 : 1 比例分为校正集($n=128$)和预测集($n=43$)两个子集共计 171 个羊肉样本,利用校正集构建模型,预测集检验模型性能。由表 1 可知,校正集样品

的范围涵盖了预测集范围,故可使用预测集中的样本来检验最终模型性能。此外,可溶性蛋白和 GSH 分布值范围较宽,表明化学分析结果具有代表性,可用于后续模型。

表 1 羊肉可溶性蛋白和 GSH 含量样本集划分统计结果
Table 1 Statistical of soluble protein and GSH contents of mutton of sample sets

样本集	校正集				预测集			
	样本量	测量范围	平均值	标准差	样本量	测量范围	平均值	标准差
可溶性蛋白	128	19.96~110.87	70.55	18.29	43	38.81~93.55	71.77	15.83
GSH	128	17.34~172.75	46.43	24.51	43	17.81~169.95	44.08	26.38

2.2 羊肉样本光谱特征分析

图 2(a)为 171 个羊肉样本的原始光谱图像,429 和 540 nm 附近的光谱反射率可能与氧合血红蛋白有关^[14];479,723 和 799 nm 附近的高反射率可能与蛋白质组中—NH₃—基团吸收以及 C—H 倍频和脂肪中 C—O 倍频的综合作用有关;在 900~1 000 nm 波长范围内的光谱峰值与羊肉样本水分子吸收带有关^[6]。这些光谱差异可用于后续预测羊肉样本可溶性蛋白和 GSH 含量。图 2(b)为羊肉样本前腿、

后腿、背最长肌的光谱值求平均值并作出平均光谱图。图 2 (b)中羊肉样本不同部位光谱曲线呈现相似趋势,其中在 450~525 和 600~1 000 nm 波长范围之间,前腿的光谱反射率均为最高,后腿最低;525~600 nm 波长范围之间,背最长肌光谱反射率最高,后腿反射率最低。可以发现,后腿的在全波段光谱反射率均为最低。这些差异可能与羊肉样本的主要化学成分有关。

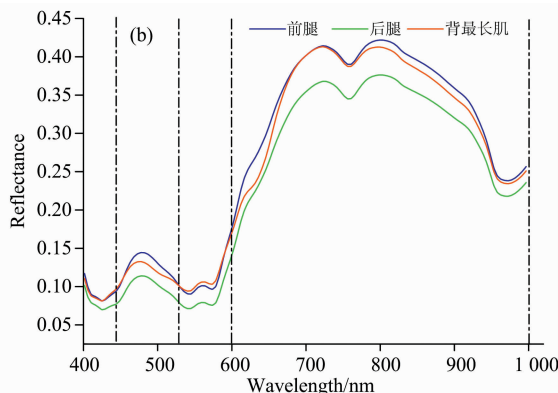
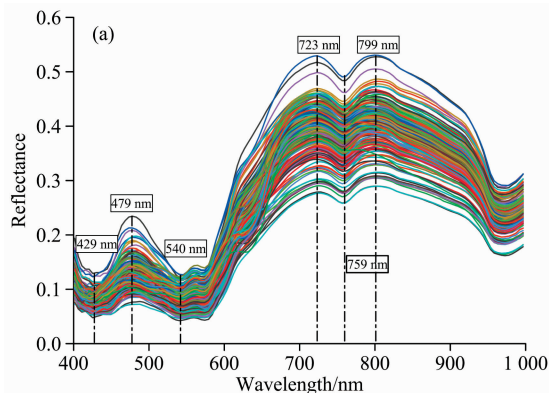


图 2 羊肉样本的原始光谱图(a)与不同部位平均光谱图(b)

Fig. 2 Raw (a) and average spectra (b) of mutton samples

2.3 不同预处理方法的高光谱建模分析

如表 2 所示,未经预处理的原始光谱建立的可溶性蛋白含量模型建模效果最好, R_c 和 R_p 分别为 0.875 7 和 0.854 7,并且 RMSEC 和 RMSEP 均为最低。SNV 方法预处理后建立的羊肉样本 GSH 含量模型的 R_c 和 R_p 分别为 0.804 8 和 0.826 5,说明 SNV 方法能够去除与羊肉 GSH 含量无关的冗余信息,提高了模型的预测性能。因此在后续的变量筛选中,采用原始光谱建立可溶性蛋白含量模型,SNV 作为建立 GSH 含量模型所用预处理方法。

2.4 特征波长提取

2.4.1 CRAS 法特征波长提取

采用 CARS 法、iVISSA-IRIV 2 种方法提取特征波长。经过多次试验,设置 CARS 算法参数为:主成分数为 10,数据选用“center”处理方式,十折交叉验证,其中,可溶性蛋白蒙特卡洛采样次数设置为 350,GSH 的采样次数为 500。如图 3 所示为可溶性蛋白和 GSH 的 CARS 算法挑选特征波长过程。图 3A(a)表示随着抽样运行次数的增加,抽样波长数

表 2 不同预处理方法的羊肉指标 PLSR 模型

Table 2 PLSR models for mutton indicators by different pretreatment methods

建模对象	预处理方法	LVs	校正集		预测集	
			R_c	RMSEC	R_p	RMSEP
可溶性蛋白	Raw	17	0.875 7	8.830 0	0.854 7	9.827 3
	SG	16	0.857 8	9.398 0	0.814 9	10.584 5
	SNV	11	0.828 2	10.248 1	0.854 2	9.726 8
	OSC	13	0.848 6	9.674 2	0.846 8	9.919 2
	Detrend	10	0.830 1	10.195 9	0.881 4	9.074 5
GSH	Raw	17	0.810 0	14.375 0	0.780 5	13.564 0
	SG	16	0.751 7	16.168 4	0.732 2	18.155 8
	SNV	17	0.804 8	14.049 6	0.826 5	12.355 8
	OSC	11	0.862 1	12.423 1	0.702 0	20.074 0
	Detrend	14	0.816 0	14.170 7	0.745 8	16.678 1

的变化。图 3A(b)表示为抽样运行期间的十折交叉验证值 RMSECV 变化,在运行至第 83 次时该值最小,此时变量个

数逐步减少导致部分有效信息被消除，模型精度下降；图 3A (c)为每次采样过程中回归系数的变化路径，最左边的粗竖线代表当采样运行次数为 21 次时 RMSECV 值最小，此时与

羊肉可溶性蛋白无用的信息被剔除，最终共选取有效信息最多的 48 个特征波长变量。CARS 方法提取羊肉 GSH 含量样本的特征波长情况如图 3(b)所示，共提取出 19 个特征波长。

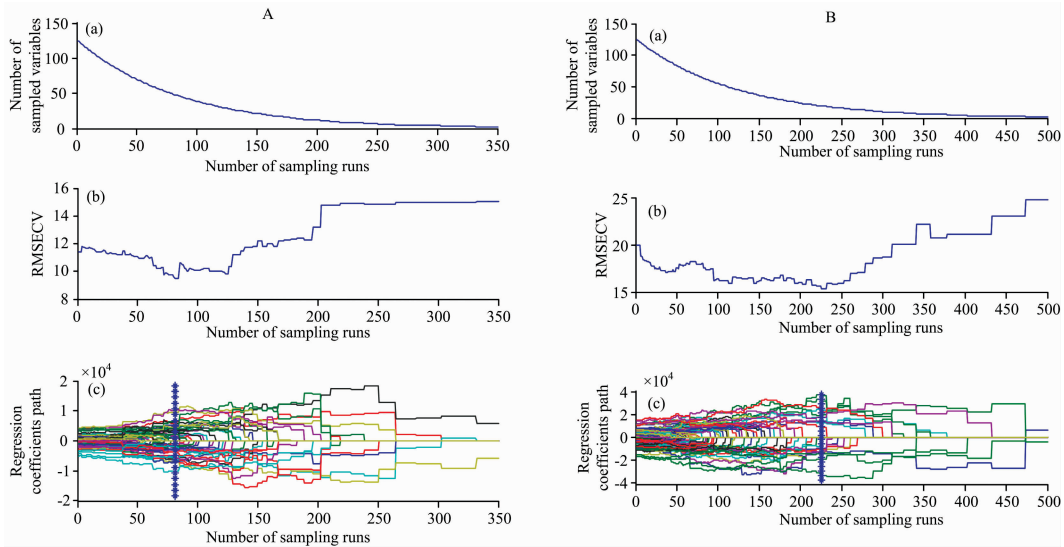


图 3 CARS 法特征波长的筛选

A: 可溶性蛋白含量; B: GSH 含量

Fig. 3 Characteristic wavelengths selected by CARS algorithm

A: soluble protein content; B: GSH content

2.4.2 iVISSA-IRIV 特征波长选取

利用 iVISSA-IRIV 融合方法对光谱特征波长进行筛选，采样次数设置为 50。图 4(a)为 iVISSA 算法在迭代过程中各个波长的采样权重值随迭代次数的变化情况。此三维立体图 X, Y, Z 坐标轴分别代表波段数、迭代次数、权重，结合色度带上方绿色部分采样权重接近 1，此时有效信息被保留，下端紫色部分则采样权重值接近 0，这部分波段为无效信息被剔除，而采样权重处于 0~1 之间时，相应的那些波长会重新在迭代循环中继续被选择，直到所有特征波长选择完毕，

经过 25 次迭代，采样权重基本不变，共筛选出 52 个特征波长，特征波长数仍然占总波长数的 41.6%，会造成信息冗余，增加计算难度和时间，故需再通过 IRIV 算法二次提取特征波长。如图 4(b)所示羊肉可溶性蛋白含量 iVISSA-IRIV 法特征波长筛选图，柱条表示提取的 31 个特征波长。而同理可得，经 iVISSA-IRIV 算法提取的 GSH 含量样本的特征波长结果如图 5 所示，图 5(b)为羊肉 GSH 含量 iVISSA-IRIV 法特征波长筛选图，共提取 29 个特征波长。

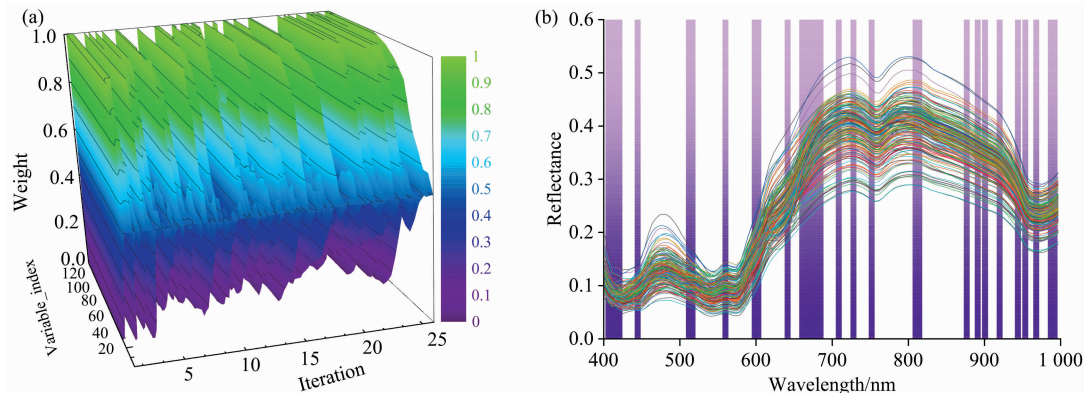


图 4 可溶性蛋白含量 iVISSA-IRIV 法筛选特征波长

Fig. 4 Selection of characteristic wavelengths using iVISSA-IRIV algorithm for soluble protein content

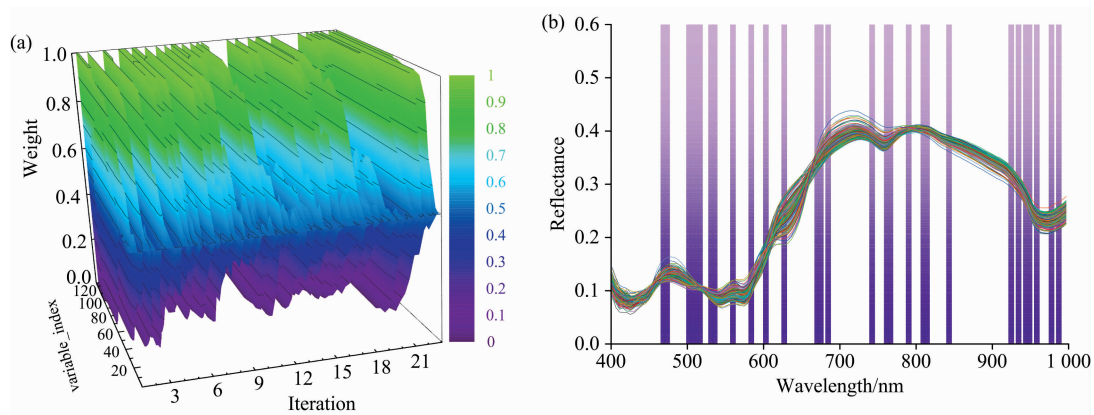


图 5 GSH 含量 iVISSA-IRIV 法筛选特征波长

Fig. 5 Selection of characteristic wavelengths using iVISSA-IRIV algorithm for GSH content

2.5 建模结果分析

采用 MLR 和 LS-SVM 两种模型对提取出的特征波长进行有效性评价。如表 3 所示, 与全波段相比, 提取特征波长建立的 MLR 和 LS-SVM 模型预测性能良好。两种特征波长提取方法都能够达到降维目的, 提取有用信息, 但 iVISSA-IRIV 法经过两次提取特征波长, 解决变量数多的同时更多的保留有效信息, 故经 iVISSA-IRIV 法提取特征波长建模效果较优。羊肉可溶性蛋白 iVISSA-IRIV-LS-SVM 为最优预测模型, 其 R_c 和 R_p 分别为 0.914 6 和 0.881 8, 且均方根误差均最低。表明基于非线性的 LS-SVM 模型的性能优于线性

MLR 模型, 说明羊肉可溶性蛋白含量与光谱吸收之间具有较强的非线性关系。羊肉 GSH 含量 iVISSA-IRIV-MLR 为最优预测模型, 其 R_c 和 R_p 分别为 0.844 6 和 0.870 5, 比原始光谱建模 R_c 和 R_p 高了 0.046 和 0.097 1, 说明在建模过程中, 羊肉 GSH 含量不利于非线性预测模型的建立。综上所述, 羊肉光谱信息经 iVISSA-IRIV 法提取特征波长后, 建立的羊肉可溶性蛋白和 GSH 含量预测模型取得了较好的建模效果, 表明 iVISSA-IRIV 法对光谱数据进行降维简化了定量分析模型的复杂性, 提高了模型的预测能力。

表 3 不同特征波长和模型的羊肉可溶性蛋白和 GSH 含量预测结果

Table 3 Prediction results for soluble protein and GSH content in mutton using different characteristic wavelengths and models

建模对象	模型	特征波长数	特征波长	校正集		预测集	
				R_c	RMSEC	R_p	RMSEP
可溶性蛋白	MLR	125	Raw	0.872 1	10.214 2	0.832 1	10.987 3
		48	CARS	0.918 4	9.210 9	0.860 4	10.822 0
		31	iVISSA-IRIV	0.904 2	9.020 5	0.881 6	9.454 6
	LS-SVM	125	Raw	0.891 8	8.971 3	0.867 0	8.423 8
		48	CARS	0.929 6	6.806 1	0.885 2	8.640 1
		31	iVISSA-IRIV	0.914 6	7.422 5	0.881 8	8.527 2
GSH	MLR	125	Raw	0.798 6	15.864 5	0.773 4	20.062 3
		19	CARS	0.841 0	14.440 1	0.795 5	24.017 4
		29	iVISSA-IRIV	0.844 6	14.998 5	0.870 5	15.517 7
	LS-SVM	125	Raw	0.777 2	15.430 6	0.724 2	46.348 1
		19	CARS	0.855 8	12.689 3	0.729 7	27.272 9
		29	iVISSA-IRIV	0.936 4	9.102 8	0.708 2	30.307 2

2.6 光谱特征和纹理信息融合建模分析

为了进一步分析羊肉内部属性和外部属性对可溶性蛋白和 GSH 含量预测模型的影响, 采用提取的最优特征波长和纹理信息进行融合, 建立基于光谱信息和纹理特征融合信息的 MLR 模型以及 LS-SVM 模型。因为高光谱成像在某个特定波长下图像也反映样本化学成分基团对光子吸收后的反射光谱图像, 其对某个特征也会有较显著的反映, 因此图像纹

理特征在一定程度上能够反映化学成分以及结构方面的差异。如图 6 所示为对高光谱图像进行主成分分析后提取的主成分图像贡献率 99.50% 的 3 个主成分图像, 主成分贡献率依次为 98.59%, 0.79% 和 0.12%, 故采用第一张最优主成分图像进行后续纹理提取, 将提取的图像纹理信息和提取的光谱特征信息进行融合建模分析。

如图 7 所示为经全波段、iVISSA-IRIV 特征波段和光谱

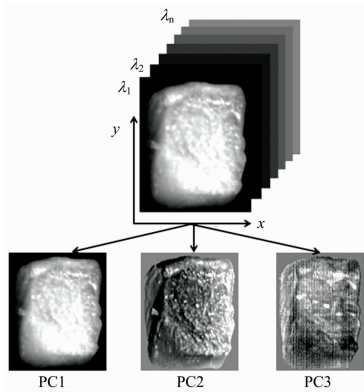


图 6 羊肉样本的前三个主成分图像

Fig. 6 The first three principal component images of mutton samples

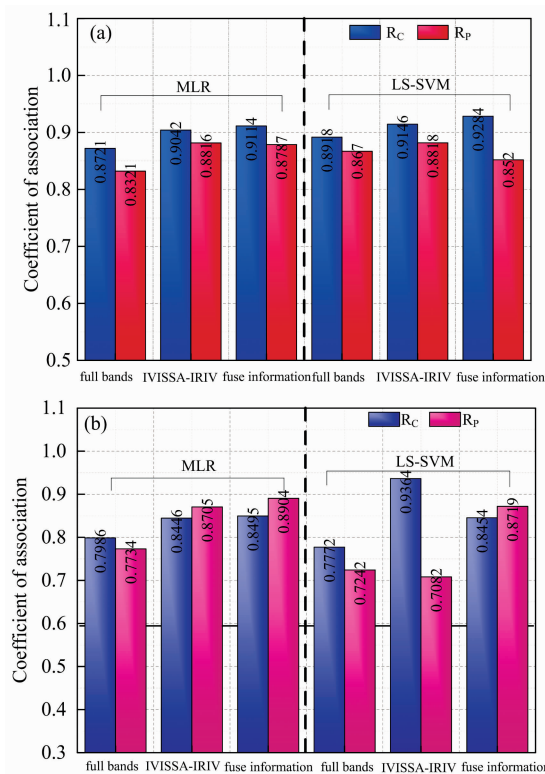


图 7 模型结果对比图

(a): 羊肉可溶性蛋白含量模型对比图;

(b): 羊肉 GSH 含量模型对比图

Fig. 7 Comparison of model results

(a): Comparison chart for soluble protein content in mutton;

(b): Comparison chart for GSH content in mutton

纹理融合信息建立的 MLR 和 LS-SVM 模型。可以看出羊肉可溶性蛋白和 GSH 含量光谱纹理融合信息建立的 MLR 和 LS-SVM 模型取得了较好的预测效果。如图 7(a)所示,利用光谱和纹理信息融合建立的羊肉可溶性蛋白含量 MLR 模型预测性能较好,其 R_c 和 R_p 达到 0.911 4 和 0.878 7;但比 iVISSA-IRIV 光谱信息的 LS-SVM 模型略低 0.003 2 和

0.003 1, 故利用光谱特征信息建立的 LS-SVM 为羊肉可溶性蛋白最佳模型,这可能是利用光谱特征信息建立 LS-SVM 模型比纹理信息携带的信息更多,羊肉样本的内部成分比其外部属性的联系更加紧密,也可能是现在所选用的纹理特征提取方法不适用与当前的羊肉样本的图像,但融合信息建立的模型精度仅仅比光谱信息建模精度略低,说明利用图像纹理信息还是可行的。

如图 7(b)所示为羊肉 GSH 含量模型对比图,光谱特征波长与纹理信息融合建立的 MLR 模型和 LS-SVM 模型均有较好的预测效果,均比全波段和 iVISSA-IRIV 法提取的特征波段建模结果好,光谱和纹理信息融合建立的 MLR 模型效果更好,比 iVISSA-IRIV-MLR 模型的 R_c 和 R_p 值高 0.004 9 和 0.019 9。证明结合羊肉样本的内部成分和外部属性可以很好的预测羊肉 GSH 含量,因此光谱与纹理特征融合信息建立的 MLR 模型为预测羊肉 GSH 含量的最优模型。

2.7 羊肉可溶性蛋白和 GSH 含量空间分布可视化

利用最佳 iVISSA-IRIV-MLR 和 iVISSA-IRIV-LS-SVM 模型生成羊肉可溶性蛋白和 GSH 含量分布图。通过提取羊肉高光谱图像中每个像素的光谱值,然后结合最佳预测模型获取与化学值相关的权重系数,以此来通过伪色彩图像的不同区域颜色差异及深浅来表示羊肉可溶性蛋白和 GSH 含量的分布情况。从图 8 分布图可以看出,随着羊肉可溶性蛋白和 GSH 含量逐渐减少,红色逐渐减弱,证明通过 iVISSA-IRIV 法选出的特征波段有很好的显示可溶性蛋白和 GSH 含量的能力,并且通过可视化可以直观的看出其含量分布,这是高光谱成像技术相对于常规光谱和成像技术的最大优势。

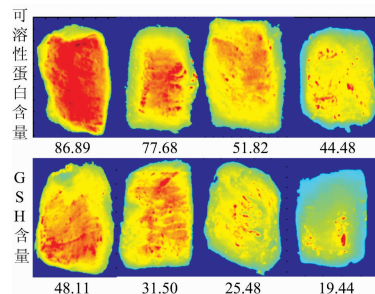


图 8 羊肉可溶性蛋白和 GSH 含量空间分布可视化图

Fig. 8 Visualizations of spatial distributions of soluble protein and GSH contents in mutton

3 结 论

探讨了利用高光谱图像的光谱和纹理信息快速预测羊肉可溶性蛋白和 GSH 含量的可行性。主要结论如下:(1)通过对羊肉样本原始光谱进行 4 种预处理方法比较,发现未经预处理的原始光谱建立的可溶性蛋白 PLSR 模型效果最好,采用 SNV 法预处理后建立的羊肉 GSH 含量 PLSR 模型为最佳预测模型。(2)利用 CARS 法和 iVISSA-IRIV 法对经预处理后的光谱提取特征波长,其中 iVISSA-IRIV-LS-SVM 模型对羊肉可溶性蛋白含量预测有最佳效果,iVISSA-IRIV-MLR 模型为羊肉 GSH 含量光谱数据最佳预测模型。(3)进一步利

用光谱特征和图像纹理融合信息建立 MLR 和 LS-SVM 模型对比发现,通过光谱信息建立的 iVISSA-IRIV-LS-SVM 模型对羊肉可溶性蛋白含量的预测效果最好,其 R_c 和 R_p 分别为 0.914 6 和 0.881 8;通过光谱和图像纹理融合信息建立的 MLR 模型对羊肉 GSH 含量的预测效果最佳, R_c 和 R_p 分别为 0.849 5 和 0.890 4。综上所述,羊肉可溶性蛋白含量与光谱之间具有较强的非线性关系,建立的 LS-SVM 模型具有较

好的预测性能,基于光谱和图像纹理融合信息比单独光谱预测羊肉 GSH 含量更有效,通过选取最佳光谱预测模型生成可视化分布图可直观看见羊肉可溶性蛋白和 GSH 含量分布,结果表明利用高光谱图像的光谱和纹理信息对羊肉可溶性蛋白和 GSH 含量的预测是可行的,可进一步改进该方法去预测肉品新鲜度。

References

- [1] Zhu Y, Wu J, Wang K, et al. *Talanta*, 2020, 224: 121852.
- [2] Chen J, Li Y, Feng X, et al. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2021, 246: 119041.
- [3] Ke J, Rao L, Zhou L, et al. *Infrared Physics & Technology*, 2020, 105: 103185.
- [4] Lu Y, Wang W, Ni X, et al. *Infrared Physics & Technology*, 2020, 111: 103509.
- [5] Xiong Z, Sun D, Pu H, et al. *LWT - Food Science and Technology*, 2015, 60(2, Part 1): 649.
- [6] SUN Zong-bao, LIANG Li-ming, LI Jun-kui, et al(孙宗保, 梁黎明, 李君奎, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2020, 40(11): 3530.
- [7] Lin H, Zhao J, Chen Q, et al. *Food Research International*, 2013, 54(1): 821.
- [8] Yang D, He D, Lu A, et al. *Infrared Physics & Technology*, 2017, 83: 206.
- [9] Zhang H, Zhang S, Chen Y, et al. *Journal of Food Composition and Analysis*, 2020, 92: 103567.
- [10] DUAN Yu-fei, WANG Qiao-hua(段宇飞, 王巧华). *Food Science(食品科学)*, 2020, 41(12): 273.
- [11] Zhang C, Wu W, Zhou L, et al. *Food Chemistry*, 2020, 319: 126536.
- [12] Li H, Zhu J, Jiao T, et al. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2020, 243: 118765.
- [13] Zhang R, Zhang F, Chen W, et al. *Chemometrics and Intelligent Laboratory Systems*, 2019, 184: 132.
- [14] Cozzolino D, Murray I. *LWT—Food Science and Technology*, 2004, 37(4): 447.

Combination of Spectral and Textural Informations of Hyperspectral Imaging for Predictions of Soluble Protein and GSH Contents in Mutton

QIAO Lu, WANG Song-lei*, GUO Jian-hong, HE Xiao-guang

School of Food and Wine, Ningxia University, Yinchuan 750021, China

Abstract Soluble protein and glutathione (GSH) are important physiological and biochemical indicators of mutton, which are also significant in measuring the body's antioxidant capacity. However, the traditional detection methods are complicated and time-consuming. This report applied visible-near-infrared (400 ~ 1 000 nm) hyperspectral imaging technology to achieve nondestructive and rapid detection of soluble protein and glutathione (GSH) content in mutton. Four methods are used to preprocess the original spectral information of the collected 180 mutton samples, and then use the competitive adaptive weighting algorithm (CARS), the wavelength space iterative shrinkage algorithm-iteration and retained information variable method (iVISSA-IRIV) method for characteristics band extraction. At the same time, the gray level co-occurrence matrix method (GLCM) is used to extract the texture information of the principal component image with the highest contribution rate. Finally, the optimized preprocessing method and the characteristic wavelength information are combined with multiple linear regression (MLR), and least squares support vector machine (LS-SVM) prediction models respectively, as spectral information and spectral-texture fusion information, to establish the prediction models of soluble protein and glutathione content of mutton. The results illustrate that the PLSR model of mutton soluble protein content established by the original spectrum without pretreatment has the best effect, and its R_c and R_p are 0.875 7 and 0.854 7, respectively; the PLSR model of mutton GSH content established by the spectra after pretreatment with SNV method work best, with R_c and R_p of 0.804 8 and 0.826 5, respectively. A total of 31 characteristic wavelengths were screened using iVISSA-IRIV, and the R_c and R_p of the established mutton soluble protein LS-SVM model were 0.914 6 and 0.881 8 respectively, which are the best. The meanwhile, 29

characteristic wavelengths were screened using iVISSA-IRIV, and the R_c and R_p of the established mutton GSH-MLR model were optimal, 0.844 6 and 0.870 5, respectively. The comparison of the spectral feature information and the fusion model of the map information revealed that the establishment of the iVISSA-IRIV-LS-SVM model was the best for the prediction of soluble protein in mutton, with R_c and R_p of 0.914 6 and 0.881 8, respectively. The MLR model established by fusion of the spectral feature information extracted by SNV-iVISSA-IRIV method with the texture information is the optimal model for predicting the GSH content of mutton, and its R_c and R_p are 0.849 5 and 0.890 4, respectively. The optimal iVISSA-IRIV-LS-SVM and iVISSA-IRIV-MLR models and imaging processing methods visually represented the spatial distribution of soluble protein and GSH contents of mutton samples in combination with pseudo-color images. The current study demonstrated that the spectral and textural information from hyperspectral images could predict soluble protein and GSH content of mutton.

Keywords Hyperspectral imaging technology; Characteristic wavelength selection; Soluble protein and GSH contents; Textural features; Distribution visualization

(Received Dec. 23, 2020; accepted Mar. 9, 2021)

* Corresponding author

本 刊 声 明

近期以来,一些不法分子假冒《光谱学与光谱分析》期刊社名义,以虚假网站等形式欺骗广大作者、读者。这些虚假网站公然假冒《光谱学与光谱分析》期刊名义进行大肆的征稿并骗取作者的审稿费和版面费。经部分作者及读者举报,现有关部门已就此介入调查。本刊将通过法律途径向假冒者追究相应的责任,维护本刊权利。

本刊官方网站已正式开通,网址为

<http://www.gpxygpx.com/>

在此郑重声明,本网站为《光谱学与光谱分析》期刊唯一开通运行的官方网站。本刊从未授权任何单位或个人以任何形式(包括网上网下)代理本刊征稿、审稿等项业务。

希望广大读者和作者切实维护好自身的合法权益,防止受骗上当。

《光谱学与光谱分析》期刊社

2019年3月15日