

# 基于 LIBS 技术结合 PCA-SVM 机器学习对猪肉部位的识别研究

许毓婷<sup>1</sup>, 孙浩然<sup>2</sup>, 高 勋<sup>1\*</sup>, 郭凯敏<sup>3\*</sup>, 林景全<sup>1</sup>

1. 长春理工大学理学院, 吉林 长春 130022
2. 长春工业大学机电工程学院, 吉林 长春 130012
3. 包头师范学院物理科学与技术学院, 内蒙古 包头 014030

**摘 要** 近年来, 激光诱导击穿光谱(LIBS)与算法相结合分类、识别生物组织逐渐兴起。由于猪肉各部位组织光谱特性相似, 仅通过分析光谱信息很难达到准确识别的效果, 采用来自于同一个体、四个不同部位的猪肉进行研究, 并将其进行切片、压平, 应用 LIBS 技术对 4 种部位的组织(里脊, 梅花, 前腿, 五花)进行了样品光谱的采集, 每种样品采集 100 幅光谱进行分析, 选取 Ca, Na, K 等 6 条谱线进行了初步光谱分析, 观测谱线发现除脂肪含量较多的五花组织 C—N 以及 C 含量较其他组织高以外, 其他组织很难区分, 进一步对这 6 个成分进行主成分分析(PCA), 得到 PC1, PC2, PC3 累计贡献率达到 95%, 通过特征分数作为支持向量机(SVM)模型输入源, 建立 SVM 分类模型, 得到几种部位样品的混淆矩阵图, 通过观察混淆矩阵可以清楚分辨出每个种类样品的分类准确率, 发现四种样品准确率分别为 96%, 98%, 97%, 100%, 平均准确率达到了 97%以上。研究证明 LIBS 结合 PCA-SVM 可作为一种快速鉴别猪肉不同组织部位的检测方法。

**关键词** 激光诱导击穿光谱; 主成分分析; 支持向量机; 猪肉组织

**中图分类号:** O433.4 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)11-3572-05

## 引 言

猪肉制品营养成分丰富, 在人类饮食中占有重要地位, 由于近年来猪肉制品价格水涨船高。不同部位的猪肉组织较为类似, 在某些猪肉产品销售市场上, 一些不良商家将廉价部位的猪肉以次充好高价售卖, 让消费者无法分辨。猪肉部位的传统检测方法主要有物理学方法(显微检测法)<sup>[1]</sup>、免疫学方法(蛋白质检测)<sup>[2]</sup>和分子生物学方法(检测 DNA 分子)<sup>[3]</sup>, 但这些方法操作繁琐, 检测周期长, 需专业技术人员操作。基于分子光谱学的猪肉部位检测方法如近红外光谱技术和拉曼光谱技术<sup>[4]</sup>, 方法虽便捷, 由于各部分猪肉成分组成接近且猪肉组织接近, 造成猪肉部位分类识别的识别率有待提高。

激光诱导击穿光谱技术(laser-induced breakdown spectroscopy, LIBS)<sup>[5]</sup>是一种利用高能激光脉冲激发待测样品并产生等离子体, 通过分析样品的等离子体发射光谱, 最终识别待测物质成分的分析方法。近年来, LIBS 作为一种新兴的

光化学元素分析技术, 具有快速、绿色、无损检测等优点<sup>[6]</sup>, 因此 LIBS 技术在冶金、航天、土壤污染、水体污染、食品安全等多个领域进行推广和应用<sup>[7-8]</sup>。在食品安全领域, 王彩虹<sup>[9]</sup>等基于 LIBS 技术和线性回归方法对猪肉中的 Cr 元素进行定量分析, 分析了单变量和多变量特征谱线对猪肉中 Cr 元素的定量分析影响。Gonca Bilge<sup>[10]</sup>等将猪肉、牛肉、鸡肉碾碎混合, 基于 LIBS 技术并结合主成分分析(principal component analysis, PCA)机器学习, 对肉类的分类和识别准确率达到 83.37%。汪海燕<sup>[11]</sup>等提出了 LIBS 技术结合 KNN 和支持向量机(support vector machines, SVM)机器学习对猪油、猪皮、里脊等软组织进行分类识别, 识别准确率和灵敏度分别达到了 99.83%和 99.5%。

研究上述表明 LIBS 技术结合机器学习对肉类组织以及肌肉组织的分类和识别是可行的, 但在猪肉部位分类和识别方面的研究还未见报道。本工作通过 LIBS 技术与 PCA-SVM 机器学习相结合, 对同一头猪的不同猪肉部位开展研究, 实现猪肉部位的分类和识别。

收稿日期: 2020-10-23, 修订日期: 2021-02-04

基金项目: 国家自然科学基金项目(61575030), 吉林省科技厅项目(20180201046YY, 20180101283JC, 20200301042RQ), 吉林省教育厅项目(JJKH20190539KJ)资助

作者简介: 许毓婷, 女, 1996年生, 长春理工大学理学院硕士研究生 e-mail: 2742592493@qq.com

\* 通讯作者 e-mail: lasercust@163.com

## 1 实验部分

### 1.1 装置

猪肉部位分类和识别 LIBS 实验装置示意图如图 1 所示, 实验均在一个大气压、环境温度 23 °C、相对湿度 40% 的洁净空间下完成。激光光源为输出波长 1 064 nm 的 Nd: YAG 激光器(SLII-10, Continuum), 光束模式 TEM<sub>00</sub>, 脉宽为 10 ns, 重复频率为 10 Hz, 光束直径约为 6 mm, 最大输出脉冲能量为 250 mJ。激光光束经二分之一玻片和格兰棱镜组成的能量衰减系统, 经 1 064 nm HR 镜和焦距为 120 mm 的石英透镜聚焦猪肉组织表面, 焦点位于猪肉样品表面内 0.1

mm, 诱导击穿猪肉的激光脉冲能量为 60 mJ。诱导击穿产生的等离子体由 45° 放置、焦距为 75 mm 的透镜收集经光纤耦合传输到配有 ICCD 探测器的中阶梯光谱仪(Mechelle 5000, Andor, 波长范围 200~850 nm, 光谱分辨率为 0.05 nm) 进行探测。ICCD 探测器和激光输出同步触发由数字脉冲延时发生器 DG645 控制, ICCD 探测器相对激光脉冲延时为 1 μs, 积分时间 1 μs。由于猪肉组织样品表面不均匀, 为使猪肉样品不被击穿打在载玻片上产生误差, 在实验中同时保证 PI 平移台处于以 0.1 mm·s<sup>-1</sup> 的速度移动, 采集 100 个脉冲的 LIBS 光谱并进行平均, 降低脉冲能量抖动对 LIBS 光谱的稳定性影响, 且在同一实验条件下每个猪肉部位样品采集 100 幅光谱。

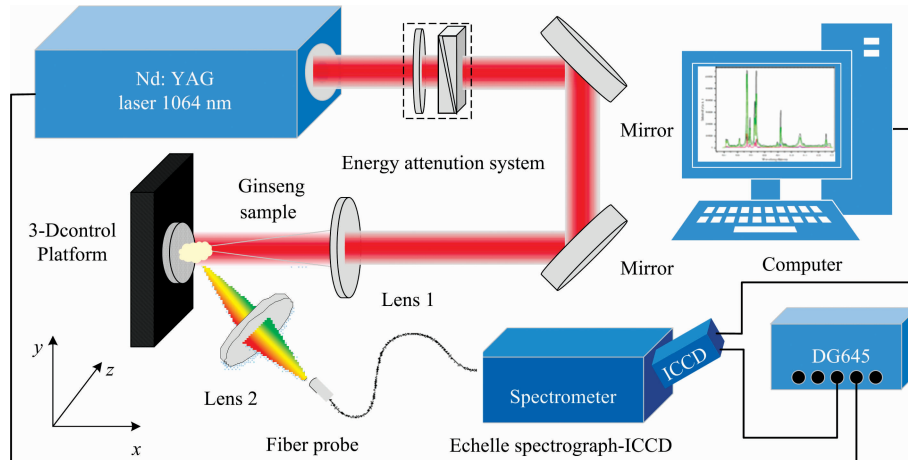


图 1 LIBS 实验装置图

Fig. 1 LIBS experimental setup

### 1.2 样品制备

实验样品购自长春市莲花山猪肉市场的同一头肉猪的前腿肉、梅花肉、里脊肉、五花肉。由于猪肉组织成分的表面分布不同于固体金属表面均匀稳定。本实验为了减小样品的随机性, 对同一种类的猪肉采集多个切片; 样品形态、表面温度、干燥程度均会对实验结果产生影响, 为减小外界因素对实验结果产生影响, 实验开始前通过人工切片的方式将猪肉切成 7 mm×2 mm×1 mm 片状样品, 并放置于零下 20 °C 的保鲜柜中, 为反映实验结果真实, 不做干燥处理。实验开始前将样品取出, 将样品放置在三维平移台上, 确保每次激光作用于新的激光烧蚀点。

## 2 结果与讨论

四种猪肉部位降噪处理后的 LIBS 光谱如图 2 所示。由图 2 可知, 猪肉组织中观察到 H 656.2 nm, O 777.1 nm 等非金属元素谱线, 由于实验在空气环境中进行, 空气等离子体发射光谱可能会对猪肉组织 LIBS 光谱中的 H 和 O 非金属元素造成一定的干扰, 因此 H 和 O 元素不能作为分析谱线。由于猪肉组织有大量的碳水化合物脂肪组织, 因此实验中观察到了 C—N 谱线如 385.4 和 387.1 nm。猪肉组织 LIBS 光谱中还存在 Mg II 279.8 nm, Ca II (393.3, 396.8,

422.7, 430.2 和 445.4 nm 等), Na I 588.9 nm, Na I 589.0 nm, K II 766.4 nm 和 K I 769.9 nm 等金属元素谱线。一般 LIBS 的特征谱线的选取, 要选择干扰较少, 信背比较高, 峰值明显, 易激发的波段, 如图 3 所示。通过实验所得的光谱信息, 参照美国 NIST 原子光谱数据库, 选取 Ca II 393.3 和 422.7 nm, Na I 588.9 和 589.0 nm, K II 766.4 nm, K I 769.9 nm 作为猪肉组织分类和识别的 LIBS 特征分析光谱线。

脂肪含量较多的五花肉 LIBS 光谱中存在大量 C—N 及 C 元素谱线峰, 且谱线强度明显比其他三种部位组织样品高。由于实验选择的前腿肉、梅花肉、里脊肉、五花肉四类组织都来自同一种生物个体, 获得的 LIBS 特征谱线都较为相似。若从 LIBS 光谱图中分辨猪肉部位, 脂肪含量较高的五花肉和含钙较高的前腿肉可以直接区分, 而梅花肉、里脊肉两种猪肉组织元素 LIBS 光谱强度差异较小, 不能进行直接分辨。因此仅通过 4 种部位组织的 LIBS 光谱无法对猪肉部位组织进行分类鉴别, 需结合机器学习算法对猪肉部位进行分类识别。

### 2.1 PCA 猪肉部位聚类分析

分析猪肉组织的 LIBS 光谱, 元素线状光谱和分子光谱有十多条之多, 因此对猪肉部位进行分类识别的过程中, 需要减少 LIBS 谱线的分析数据量。主成分分析(PCA)是一种

常用的线性变换技术，能够降低噪声和维度<sup>[12]</sup>，使用 PCA 对猪肉部位 LIBS 光谱数据进行降维处理，同时能够保留大部分原始光谱数据信息。将选取的 Ca II 393.3 和 422.7 nm, Na I 588.9 和 589.0 nm, K II 766.4 nm, K I 769.9 nm 六个特征谱线强度作为输入变量，得到 PC1, PC2, PC3, PC4, PC5 和 PC6 的主成分贡献率及累计贡献率如图 3 所示。

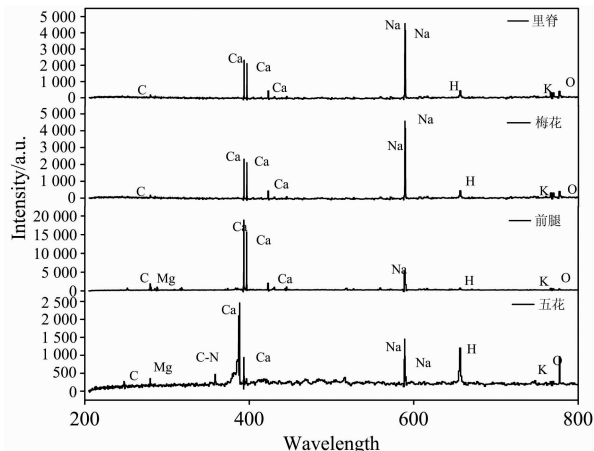


图 2 四种猪肉部位 LIBS 光谱 (里脊、梅花、前腿、五花)  
Fig. 2 LIBS spectrograms of four pork parts

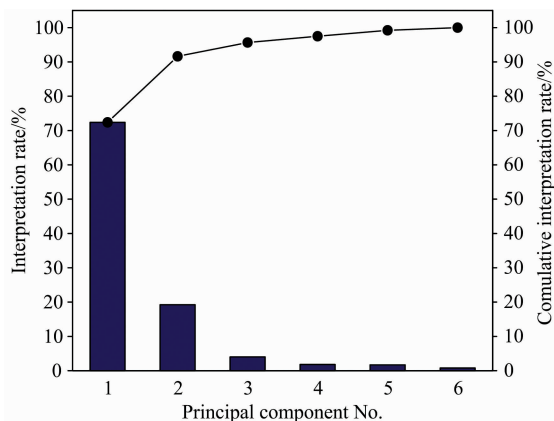


图 3 主成分贡献率及累计贡献率  
Fig. 3 Principal component contribution rate and cumulative contribution rate

由图 3 可知，PC1, PC2 和 PC3 的累计贡献率达到了 95% 以上，说明由 PC1, PC2 和 PC3 三个主成分基本能够表征猪肉部位的 LIBS 光谱全部信息，PC1, PC2 和 PC3 三个主成分的三维散点图如图 4 所示。

由图 4 可知，经 PCA 降维后发现，里脊、梅花、五花肉数据点呈现出明显的聚类，但里脊和梅花肉部位信息存在部分重叠。前腿的光谱数据点较为分散，没有呈现出聚类，表明前腿猪肉组织部位包括了里脊、梅花和五花肉类组织的生物特性，因此前腿部位的光谱数据和里脊、梅花、五花等部位光谱数据重叠区域较大。

2.2 PCA-SVM 猪肉部位识别

由于生物组织光谱数据所固有的高维性，需结合一种使用大数据统计的分类算法，支持向量机 (support vector machine, SVM)<sup>[13]</sup> 是一种基于统计学习理论的新型学习机。本文将采用 SVM 结合 PCA 算法对 4 种猪肉部位进行分类识别，以 PC1, PC2 和 PC3 主成分代替猪肉的特征光谱，构建猪肉样品的特征空间向量，特征向量构成的数据矩阵作为 SVM 识别算法的输入量。为保证猪肉部位识别模型的稳定性，采用十倍交叉验证的方式对模型进行评估，得到分类识别结果混淆矩阵如图 5 所示。

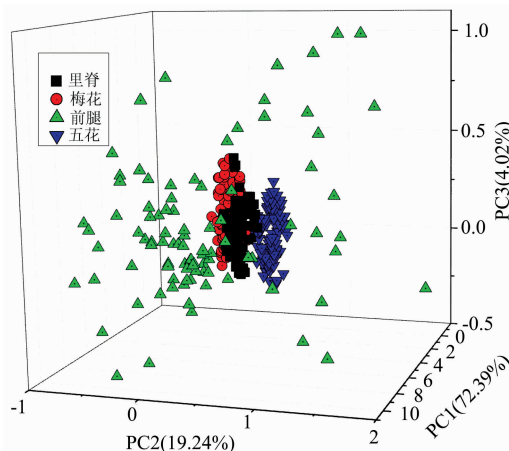


图 4 四种猪肉组织散点图

Fig. 4 Scatter plots of four pork tissues

	里脊	梅花	前腿	五花	真实结果
里脊	96	4			
梅花	2	98			
前腿	2	1	97		
				100	
	里脊	梅花	前腿	五花	预测结果

图 5 猪肉部位 PCA-SVM 分类识别混淆矩阵  
Fig. 5 PCA-SVM classification recognition confusion matrix for pork parts

由图 5 的猪肉部位 PCA-SVM 分类识别混淆矩阵可知，里脊组织、梅花组织、前腿组织、五花组织的分类识别率分别为 96%，98%，97% 和 100%，说明基于 LIBS 技术结合 PCA-SVM 分类模型可以有效地实现猪肉部位的快速识别。猪肉组织相邻较近的组织成分接近，因此 LIBS 光谱数据的信息冗余性较高且噪声干扰较强，通过 PCA 实现 LIBS 光谱分析变量减少，克服了生物组织光谱数据信息重叠性及噪声干扰。在 PCA 基础上，构造核函数的方式将两条相近不可分的谱线映射到高维空间中，再从高维空间得到线性区分，采

用 SVM 算法克服了猪肉组织光谱数据维度较高的缺点,因此,PCA-SVM 机器学习算法可有效提高了猪肉部位分类准确率。

### 3 结 论

通过 LIBS 技术结合 PCA-SVM 机器学习对猪肉部位识别进行了研究。通过对比分析同一头肉猪的前腿肉、梅花肉、里脊肉、五花肉的 4 个部位组织的 LIBS 特征光谱,发现

4 个部位组织的 LIBS 特征光谱具有较高的相似性。通过 PCA 算法对猪肉部位 LIBS 光谱数据进行降维处理,里脊、梅花、五花呈现出明显的聚类,但里脊和梅花部位信息存在部分重叠,前腿的光谱数据点较为分散,没有呈现出聚类。通过十倍交叉验证的方式,采用 PCA-SVM 分类模型得到里脊组织、梅花组织、前腿组织、五花组织的分类识别率分别为 96%, 98%, 97% 和 100%,表明 LIBS 与机器学习相结合的方法可以作为一种绿色、快速的生物组织检测技术手段。

### References

- [ 1 ] YUAN Wei-li, LIU Chao, WANG Yong-song, et al(袁伟丽, 刘超, 王永松, 等). *Acta Laser Biology Sinica*(激光生物学报), 2018, 27(4): 320.
- [ 2 ] Varricchio E, Russolillo M G, Maruccio L, et al. *European Journal of Histochemistry*, 2013, 57(1): 10.
- [ 3 ] LUO Jia-qin, WANG Jia-qi, BU Deng-pan, et al(罗家琴, 王加启, 卜登攀, 等). *Scientia Agricultura Sinica*(中国农业科学), 2008, (7): 2112.
- [ 4 ] TANG Yu-long, GUO Zhou-yi(唐玉龙, 郭周义). *Acta Laser Biology Sinica*(激光生物学报), 2004, 13(5): 386.
- [ 5 ] Tognoni E, Palleschi V, Corsi M, et al. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 2002, 57(7): 1115.
- [ 6 ] Yao M, Huang L, Zheng J, et al. *Optics & Laser Technology*, 2013, 52: 70.
- [ 7 ] Leon Radziemski, David Cremers. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 2013, 87: 3.
- [ 8 ] LI Jia-ming, CHU Ying-bo, ZHAO Nan, et al(李嘉铭, 褚应波, 赵楠, 等). *Chinese Journal of Analytical Chemistry*(分析化学), 2016, 44(7): 1042.
- [ 9 ] WANG Cai-hong, HUANG Lin, LIU Mu-hua, et al(王彩虹, 黄林, 刘木华, 等). *Chinese Journal of Analysis Laboratory*(分析实验室), 2017, 36(1): 32.
- [10] Su Yuanhang, Lin Ruiyuan, Jay Kuo C C. *Expert Systems with Applications*, 2019, 118: 355.
- [11] WANG Hai-yan, LI Jian-hui, YANG Feng-lei(汪海燕, 黎建辉, 杨风雷). *Application Research of Computers*(计算机应用研究), 2014, 31(5): 1281.
- [12] Cocianu C. *Economic Computation and Economic Cybernetics Studies and Research*, 2013, 47(1): 41.
- [13] ZHU Yi-ning, YANG Ping, YANG Xin-yan, et al(朱毅宁, 杨平, 杨新艳, 等). *Chinese Journal of Analytical Chemistry*(分析化学), 2017, 45(3): 336.

# Identification of Pork Parts Based on LIBS Technology Combined With PCA-SVM Machine Learning

XU Yu-ting<sup>1</sup>, SUN Hao-ran<sup>2</sup>, GAO Xun<sup>1\*</sup>, GUO Kai-min<sup>3\*</sup>, LIN Jing-quan<sup>1</sup>

1. College of Science, Changchun University of Science and Technology, Changchun 130022, China

2. College of Mechanical and Electrical Engineering, Changchun University of Technology, Changchun 130012, China

3. College of Physical Science and Technology, Baotou Normal University, Baotou 014030, China

**Abstract** In recent years, laser-induced breakdown spectroscopy (LIBS) is gradually emerging to classify and identify biological organizations by combining them with algorithms. Due to each part of the pork similar spectral characteristics, it is difficult to achieve accurate identifications only through the analysis of the effect of spectral information, so in this paper studied pork from four different parts of the same individual and sliced and planished them, then applied LIBS technology on four parts, i. e., the organization, fillet, plum flower, and front legs. 100 specimens of each sample were collected and the spectrum analysis was conducted. A preliminary analysis of the spectrum was performed on Ca, Na, K and 6 lines. It was found that other tissues were difficult to distinguish except for the C—N tissue of plum flower with more fat content and higher C content than other tissues, so the Principal Component Analysis (PCA) on these 6 principal components was carried out. The cumulative contribution rate of PC1, PC2 and PC3 reached 95%. The Support Vector Machines (SVM) classification model was established by employing feature scores as the input source of SVM model, and the confusion matrix diagram of these samples got obtained. Through observation of the confusion matrix, the classification accuracy of each type of samples could be clearly distinguished. The results showed that the accuracy of the four samples was 96%, 98%, 97% and 100%, respectively, with an average accuracy surpassing 97%. The study proved that LIBS combined with PCA-SVM can be used as a fast identification method for different parts of pork tissues.

**Keywords** Laser-induced breakdown spectroscopy; Principal component analysis; Support Vector Machines; Pork tissue

(Received Oct. 23, 2020; accepted Feb. 4, 2021)

\* Corresponding authors