

# 基于高光谱技术的猪肉肌红蛋白含量无损检测

王立舒, 胡金耀, 房俊龙\*, 陈曦, 李闯

(东北农业大学电气与信息学院, 哈尔滨 150030)

**摘要:** 为充分利用猪肉光谱与图像信息, 实现猪肉肌红蛋白含量的在线检测, 该研究提出一种基于深度学习模型的猪肉肌红蛋白含量无损检测方法。采用高光谱设备采集冷藏过程中猪肉高光谱图像, 通过 ENVI5.3 选择图像感兴趣区域 (Region Of Interest, ROI), 分别提取 ROI 平均光谱信息与主成分图像信息。利用卷积自动编码器 (Convolutional Auto Encoder, CAE) 提取光谱与图像信息深度特征, 分别建立光谱特征、图像特征及图-谱融合特征与肌红蛋白含量之间关系的卷积神经网络 (Convolutional Neural Network, CNN) 预测模型。其中基于融合深度特征 CNN 预测模型准确度较高, 该模型对脱氧肌红蛋白 (DeoMb)、氧合肌红蛋白 (OxyMb)、高铁肌红蛋白 (MetMb) 含量预测集决定系数  $R_p^2$  分别为 0.964 5、0.973 2、0.958 5, 预测集均方根误差 RMSEP 分别为 0.015 8、0.226 6、0.381 6。为进一步验证图-谱融合特征与猪肉肌红蛋白存在对应关系, 分别建立偏最小二乘回归 (Partial Least Squares Regression, PLSR)、支持向量机回归 (Support Vector Regression, SVR) 预测模型。结果表明: CAE 能充分提取图像与光谱特征; 基于融合特征建立回归模型能提高肌红蛋白含量预测精度, 相比于光谱信息与图像信息, 以 MetMb 为例, 其  $R_p^2$  分别提高 5.42%、16.12%。该检测方法为肉类质量在线检测提供参考, 具有好的应用前景。

**关键词:** 无损检测; 光谱特征; 高光谱图片; 卷积神经网络; 卷积自编码器

doi: 10.11975/j.issn.1002-6819.2021.16.035

中图分类号: S126

文献标志码: A

文章编号: 1002-6819(2021)-16-0287-08

王立舒, 胡金耀, 房俊龙, 等. 基于高光谱技术的猪肉肌红蛋白含量无损检测[J]. 农业工程学报, 2021, 37(16): 287-294.

doi: 10.11975/j.issn.1002-6819.2021.16.035 <http://www.tcsae.org>

Wang Lishu, Hu Jinyao, Fang Junlong, et al. Non-destructive detection of pork myoglobin content based on hyperspectral technology[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(16): 287-294. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2021.16.035 <http://www.tcsae.org>

## 0 引言

中国是猪肉生产和消费第一大国<sup>[1]</sup>, 猪肉在市场销售过程中, 色泽是影响消费者购买行为的主要因素<sup>[2]</sup>。生鲜肉色泽主要由肉中肌红蛋白相对含量及存在状态决定。随着冷藏时间延长, 脱氧肌红蛋白与氧合肌红蛋白逐步氧化为高铁肌红蛋白, 使生鲜猪肉逐渐由鲜红色转变为红褐色<sup>[3]</sup>, 影响猪肉在市场上的销售。快速检测肌红蛋白相对含量, 及时调节影响肌红蛋白改变的因素, 在保证猪肉质量前提下, 使猪肉维持鲜红色对肉品销售至关重要。目前对肌红蛋白相对质量分数检测主要有分光光度法、电化学法等, 这几种方法准确度高, 但对样本破坏性大, 操作过程复杂<sup>[4]</sup>, 因此实现对猪肉肌红蛋白快速无损检测仍是值得研究的问题。

目前, 光谱技术可对样本内部特征快速检测, 已经广泛应用于土壤<sup>[5-7]</sup>、农产品<sup>[8-9]</sup>、食品<sup>[10]</sup>等领域。高光谱成像能同时获取样本图像与光谱信息, 基于图谱特征建立数学模型, 能实现对样本快速分类与提高预测精度<sup>[11]</sup>。

孙俊等<sup>[12]</sup>利用堆叠自动编码器 (Stacked Auto Encoder, SAE) 提取不同放置时间大米光谱与图像融合特征, 建立支持向量回归 (Support Vector Regression, SVR) 预测模型, 实现对大米蛋白质含量在线检测。王彩霞等<sup>[13]</sup>采用连续投影算法、变量组合聚类分析法提取特征光谱与第一主成分图像纹理特征建立偏最小二乘回归 (Partial Least Squares Regression, PLSR) 预测模型, 实现对羊肉中饱和脂肪酸含量预测。翁士状等<sup>[14]</sup>利用卷积神经网络 (Convolutional Neural Network, CNN) 模型融合大米的图谱特征, 实现对大米品质的无损检测。由于高光谱图像数据特征相关性强, 冗余度高, 采用线性方法处理高光谱数据, 将直接影响模型预测精度。SAE 具有非线性深层网络结构, 能实现对输入数据特征提取<sup>[15]</sup>, 使其在故障检测<sup>[16]</sup>、光谱特征提取<sup>[17]</sup>、图像分类<sup>[18]</sup>等领域广泛应用。对于大容量的数据样本, 因 SAE 层数增加导致特征提取时间增加。深度学习技术快速发展<sup>[19]</sup>, 把 SAE 中全连接网络结构替换成卷积神经网络, 使用卷积自动编码器 (Convolutional Auto Encoder, CAE) 提取高光谱图像数据特征, 能获得鲁棒性强、可判别性高的光谱与图像的深度特征。这种特征提取方式可解决线性方法提取特征能力不足与 SAE 计算速度过慢等问题。

目前 CNN 在机器视觉领域表现优异<sup>[20]</sup>, 基于特征建立 CNN 预测模型相比与传统机器学习预测模型如 SVR、PLSR 等, 能减少对数据预处理并提高模型预测精度<sup>[21]</sup>。

收稿日期: 2020-03-18 修订日期: 2021-08-13

基金项目: 黑龙江省教育厅科技课题 (12521038)

作者简介: 王立舒, 教授, 博士, 博导。研究方向: 农业电气化与自动化; 电力新能源开发与利用。Email: wanglishu@neau.edu.cn

\*通信作者: 房俊龙, 教授, 博士, 博导。研究方向: 电力系统自动化、信息处理与智能测控。Email: junlongfang@126.com

本文采用 CNN 对光谱特征与图像特征及图-谱融合特征分别建立预测模型,实现对猪肉脱氧肌红蛋白、氧合肌红蛋白、高铁肌红蛋白相对含量的无损检测,以期为生鲜肉类品质在线检测提供技术支持。

## 1 材料与方法

### 1.1 样本制备与处理

试验猪肉样本品种为东北农业大学三花猪肉。该猪已达到 7 个月的出栏期,屠宰后经过 24 h 排酸。获取总质量为 3 kg 的里脊肉样本并剔除周边脂肪与结缔组织,用保鲜袋密封包装迅速运回无损检测实验室。先把猪肉样本切成 10 份形状为 10 cm×10 cm×1 cm (长×宽×高)立方体,然后把每份样品切成 2 cm×2 cm×1 cm 的立方体,共计 240 个样本。将所用样品置于 4 ℃ 恒温恒湿箱中放置 0~5 d,每天取出 40 个样本送往高光谱实验室进行高光谱图像采集与肌红蛋白含量测量。针对不同部位的猪肉该检测方法仍有适用性,但对不同品种的猪肉,仍需要建立新的光谱图像数据库。

### 1.2 肌红蛋白测量

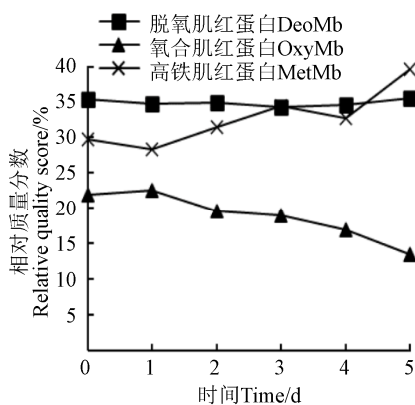
本次研究中对肌红蛋白含量测量参考 Krzywick<sup>[22]</sup>分光光度法,得出样本测量值。将磨碎猪肉样品 5 g 与 25 mL 磷酸钠缓冲液 (0.04 mol/L, pH 值为 6.8) 混合,然后用匀浆器以 10 000 r/min 均质 30 s。将均质液放置 4 ℃ 恒温恒湿箱保存 1 h 后取出。以 4 500 r/min 离心 20 min 后过滤上清液。用分光光度计分别在 525、545、565 和 572 nm 测定滤液吸光度值。由吸光度值计算脱氧肌红蛋白、氧合肌红蛋白、高铁肌红蛋白相对质量分数,计算公式如下:

$$P_{\text{DeoMb}} = 0.369R_1 + 1.140R_2 - 0.941R_3 + 0.015 \quad (1)$$

$$P_{\text{OxyMb}} = 0.882R_1 - 1.267R_2 + 0.809R_3 - 0.361 \quad (2)$$

$$P_{\text{MetMb}} = -2.514R_1 + 0.777R_2 + 0.800R_3 + 1.098 \quad (3)$$

式中  $R_1$ 、 $R_2$ 、 $R_3$  分别为 572 与 525 nm、565 与 525 nm、545 与 525 nm 吸光度比值。 $P_{\text{DeoMb}}$  为脱氧肌红蛋白质量分数 (%),  $P_{\text{OxyMb}}$  为氧合肌红蛋白相对质量分数 (%),  $P_{\text{MetMb}}$  为高铁肌红蛋白相对质量分数 (%)。贮藏期间生鲜猪肉肌红蛋白相对质量分数的变化趋势如图 1 所示。



注: 4 ℃ 恒温恒湿条件下保存。

Note: Store at 4 ℃ constant temperature and humidity conditions.

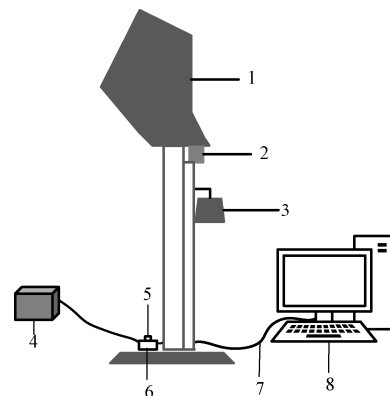
图 1 猪肉肌红蛋白相对质量分数变化曲线

Fig.1 Change curve of relative quality score of pork myoglobin

由图 1 可知,在 0~5 d 试验周期内脱氧肌红蛋白相对质量分数下降缓慢,氧和肌红蛋白相对质量分数明显下降,高铁肌红蛋白相对质量分数先下降再上升。随冷藏时间延长,生鲜猪肉发生褐变最终腐败变质。

### 1.3 高光谱图像采集

高光谱图像采集在东北农业大学电气与信息学院高光谱图像处理实验室进行,高光谱成像系统硬件部分如图 2 所示。该硬件系统主要由高光谱成像仪(HyperSpec® VNIR-A, Headwall Photonics Inc)、电控传输平台、卤素灯等组成。高光谱成像仪作为高光谱成像系统核心部件,其摄像机为图像传感器 (Charge Coupled Device, CDD)、光谱仪为可见/近红外光谱仪 (光谱范围 400~1 000 nm, 光谱采样间隔 0.74 nm, 光谱通道数 810, 光谱分辨率 2~3 nm)。



1.光谱仪 2.镜头 3.光源 4.控制器 5.样本 6.移动平台 7.数据线 8.计算机

1.Spectrometer 2.Lens 3.Light source 4.Controller 5.Sample 6.Mobile platform 7.Data cable 8.Computer

图 2 高光谱硬件系统示意图

Fig.2 Schematic diagram of hyperspectral hardware system

高光谱系统开机预热 30 min, 保证照射光源稳定。将样本平铺在移动平台,通过 Hyperspec 软件平台设置载物台移动速度为 5 mm/s。为消除暗电流及光源分布不均匀对高光谱成像造成影响,需要对样本图像进行黑白校正<sup>[23]</sup>,用以下公式可以获得校正后的反射强度  $R_j$ :

$$R_j = \frac{R_i - R_d}{R_w - R_d} \quad (4)$$

式中  $R_i$  为猪肉未经校正的高光谱图像,  $R_w$  为 100% 反射率条件下的白色标定图像,  $R_d$  为 0% 反射率条件下的全黑色标定图像,  $R_j$  为校正后的光谱反射强度。

### 1.4 光谱特征预处理

使用 ENVI5.3 软件,提取每个样本感兴趣区域,并计算该区域内猪肉像素平均反射率作为光谱特征。其中 240 个样本,每个样本采集 5 个点,共测得 1 200 试验点。每个样本点的光谱维度为 800,光谱信息矩阵存储格式为 1 200 行 800 列。由于仪器精准度与测量环境导致光谱数据产生偏差,为消除噪声提高光谱分辨率,采取卷积平滑 (Savitzky-Golay, SG) 对光谱信号去噪,图 3 反映预处理前后光谱特征变化,对比图 3a、3b 发现:经过 SG 预处理后高光谱曲线平滑度提高,毛刺减少。

### 1.5 高光谱图像信息提取

样本在每个波长有一张图像，共计 800 幅图像，相邻波长图像信息高度相关，不利于图像信息提取与储存<sup>[24]</sup>。使用主成分分析法（Principal Component Analysis, PCA）对高光谱图像数据进行降维（样本选取的矩阵为： $B \times H \times W$ ，其中  $B$  为光谱波段数， $B=800$ ， $H$  为二维图像的宽度， $H=50$  pixels， $W$  为二维图像的高度， $W=50$  pixels），提取方差贡献率大的主成分因子。使用 ENVI5.3 软件将高光谱图像经过线性组合后形成主成分图像。

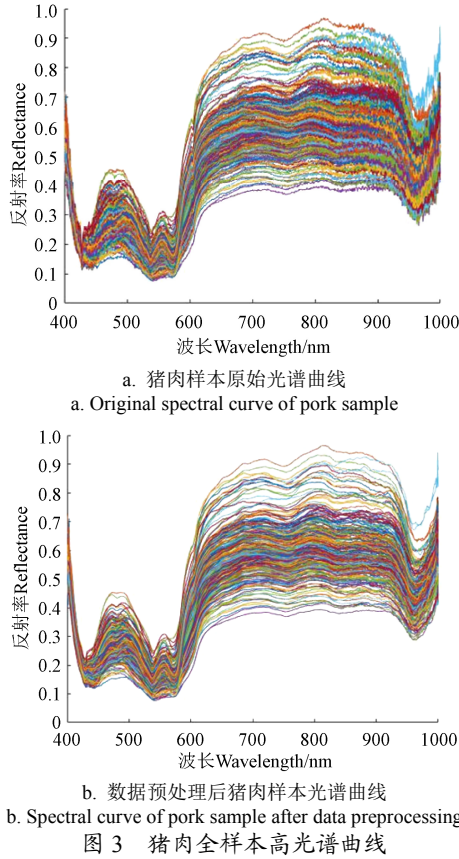


Fig.3 Hyperspectral curve of the whole sample of pork

前 3 个主成分图像累计贡献率达到 90.62%。其中第一主成分贡献率为 88.50%，表达信息量最多，选取第一主成分图像用于图像信息提取。将第一主成分图像尺寸统一为 16 pixel×16 pixel，并展平为一维向量，每幅图像中包含 768 个像素点。

### 1.6 卷积自编码器

Hinton 等<sup>[25]</sup>提出自编码器（Autoencoder, AE）用于特征提取，Chen 等<sup>[26]</sup>将多个 AE 采用级联堆叠构成 SAE 用于高光谱数据深层特征提取。CAE 为 SAE 改进形式，以端对端方式完成卷积与反卷积运算，实现光谱与图像信息深度特征提取。卷积自编码器利用 CNN 模型稀疏连接和权值共享特性，解决 SAE 因层数增加参数成指数增长问题，减少模型参数避免算法过拟合<sup>[27-28]</sup>，提高特征提取效率。卷积自编码器模型如图 4 所示。该模型分为编码器、解码器两部分，编码器由各种卷积层与池化层组成，对输入向量进行编码，提取向量深度特征，降低向量维度。解码器主要由反卷积层与上采样层构成，用

于特征数据重构。

编码器将输入映射到低维空间作为特征  $N$ ，解码器将低维特征映射到原始输入空间，生成解码矢量  $Y$ 。使用均方根误差作为模型损失函数  $L(\theta)$  衡量解码结果与真实值拟合程度。

$$N = F(X) \quad (5)$$

$$Y = G(N) \quad (6)$$

$$L(\theta) = \sqrt{\frac{1}{n} \sum_{i=1}^n (X - Y)^2} \quad (7)$$

式中  $X$  输入特征， $n$  为训练样本个数， $F(X)$  为编码器映射函数， $G(N)$  为解码函数。

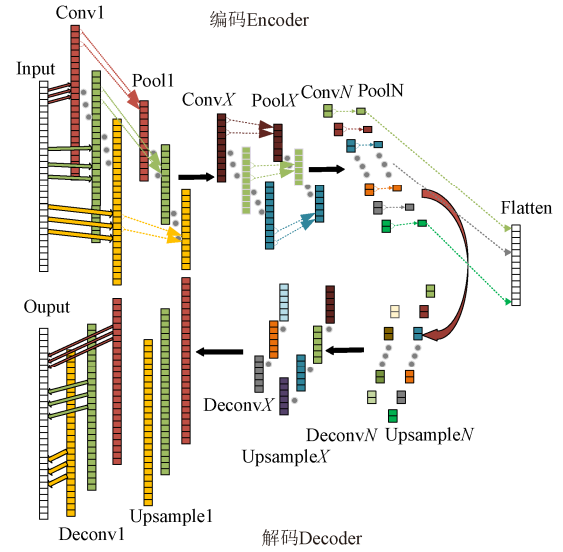


图 4 卷积编码器网络示意图

Fig.4 Schematic diagram of convolutional encoder network

### 1.7 CNN 回归模型

卷积神经网络回归模型是一种多层监督学习的神经网络，包括输入层、卷积层、池化层、输出层，其基本结构如图 5 所示。卷积层与池化层是实现卷积神经网络特征提取功能核心模块<sup>[29]</sup>，卷积层中通过卷积核对输入特征矢量进行卷积操作，再利用非线性激活函数构建输出特征矢量，其数学模型如式（8）所示：

$$x_j^l = \text{ReLU}(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l) \quad (8)$$

式中  $M_j$  为输入特征矢量； $k_{ij}^l$  为加权系数； $b_j^l$  为网络偏置； $x_j^l$  为第  $l$  层的第  $j$  个卷积核输出特征映射； $*$  代表卷积运算操作，ReLU 为激活函数，相比于传统的激活函数，其计算量少、收敛速度快，可缓解模型过拟合<sup>[30]</sup>。

最大池化层是对输入数据缩放映射，在输入中提取局部最大值，降低训练参数数量，提高特征鲁棒性，其数学模型如式（9）所示：

$$P_i^{l+1}(j) = \max_{(j-1)w+1 \leq l \leq jw} \{q_i^l(t)\} \quad (9)$$

式中  $q_i^l(t)$  表示第  $l$  层的第  $i$  个特征矢量中第  $t$  个神经元值； $w$  为池化区域宽度； $P_i^{l+1}(j)$  表示第  $l+1$  层神经元对应



值。为评估网络模型训练与测试过程准确度,使用决定系数  $R^2$  和均方根误差判断回归模型拟合程度。

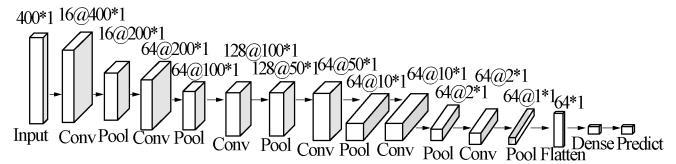
### 1.8 PLSR 与 SVR 回归模型

PLSR 集成主成分分析、多元线性回归分析等优点,在光谱信息存在多重相关性条件下建立回归模型,该模型潜变量通过自变量与目标变量间的协方差提高模型预测精度,是对光谱数据分析的一种多元统计分析方法。

SVR 模型能解决有限样本与数据非线性问题,对高光谱数据分析有较大优势。该算法将样本集从原始特征空间映射到高维特征空间,然后在高维空间中构造线性决策函数来实现线性回归,本文采用 PLSR、SVR 与 CNN 模型建立猪肉中肌红蛋白无损检测预测,通过对比 3 个模型决定系数与均方根误差选择较优模型。

### 1.9 试验平台

本文中模型的训练与测试所用电脑的主要配置为 PC Intel(R) Core(TM) i5-4200H CPU @ 2.80GHz 2.79 GHz、操作系统为 windows10。使用 Keras 深度学习框架,采用 python3.7 作为编程语言。



注: Conv 代表卷积操作。Pool 代表池化操作。a@b\*c: a 为矩形宽度,即卷积核个数。b 为矩形高度,代表数据长度。c 为矩形宽,代表输入数据样本个数。  
Note: Conv stands for convolution operation. Pool stands for pooling operation. a@b\*c: a is the width of the rectangle, that is, the number of convolution kernels. b is the height of the rectangle, representing the data length. c is the width of the rectangle, representing the number of input data samples.

图5 CNN网络预测模型

Fig.5 CNN network prediction model

## 2 结果与分析

### 2.1 光谱深度特征建模结果与分析

预处理后的光谱信息经过卷积编码器特征提取后,每个样本值的维度由 800 降到 64。图 6a 为某个样本的原始光谱信息,经过卷积编码器后提取的深度特征结果如图 6b 所示,深度特征经过解码器重构后光谱信息如图 6c 所示。经过对比发现,重构后光谱信息变化趋势与原始光谱信息大致相同。

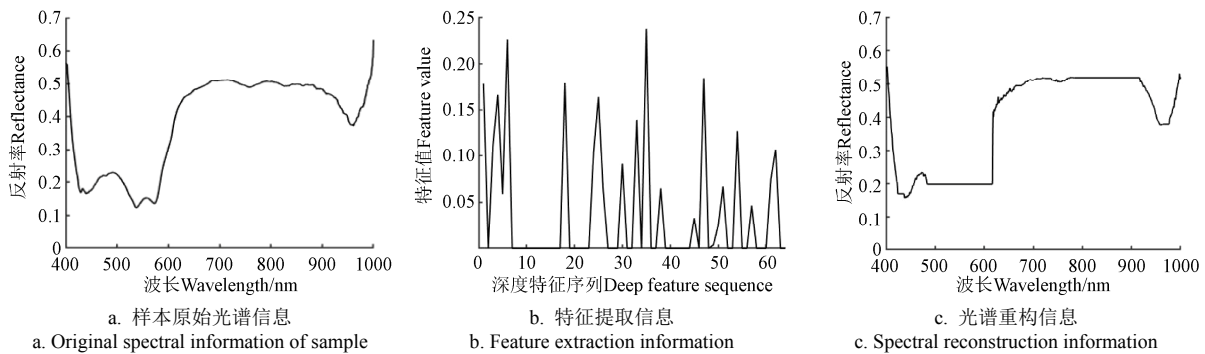


图6 光谱深度特征提取

Fig.6 Extraction of spectral depth features

针对不同放置时间的猪肉样本,共获得 1 200 个样本值(训练集 900 个,预测集 300 个),分别对全光谱波段与光谱深度特征建立 CNN 预测模型。800 个全光谱波段存储矩阵格式为 1 200×800 (行×列),深度光谱特征的存储矩阵格式为 1 200×64。基于全波段与深度光谱特征的肌红蛋白含量 CNN 预测模型评价结果如表 1 所示。

表 1 基于光谱特征的 CNN 模型预测结果

Table 1 CNN model prediction results based on spectral features

组别 Group	全波段 All bands				卷积自编码器 Convolutional Auto Encoder			
	$R_c^2$	RMSEC	$R_p^2$	RMSEP	$R_c^2$	RMSEC	$R_p^2$	RMSEP
脱氧肌红蛋白	0.872 9	0.064 3	0.855 1	0.074 4	0.936 7	0.027 6	0.923 8	0.033 4
氧合肌红蛋白	0.900 0	0.831 6	0.886 2	0.977 9	0.925 6	0.587 9	0.920 3	0.619 7
高铁肌红蛋白	0.894 7	0.853 3	0.861 8	1.174 4	0.938 0	0.597 7	0.909 2	0.809 1

注:  $R_c^2$  为测试集决定系数, RMSEC 为测试集均方根误差,  $R_p^2$  为预测集决定系数, RMSEP 为预测集均方根误差。下同。

Note:  $R_c^2$  is the determination coefficient of the test set, RMSEC is the root mean square error of the test set,  $R_p^2$  is the determination coefficient of the prediction set, and RMSEP is the root mean square error of the prediction set. The same below.

由表 1 知, CNN 预测模型对预处理后全波段高光谱信息建模,以 DeoMb 含量预测为例,其  $R_c^2$  为 0.872 9, RMSEC 为 0.064 3,  $R_p^2$  为 0.855 1, RMSEP 为 0.074 4。其结果表明全波段光谱信息维度较高,含有大量冗余数据,需要对光谱信息降维,提取深度特征信息。

全波段光谱信息经过卷积自编码器特征提取后,光谱数据维度由 800 降到 64。基于 64 个光谱信息深度特征值建立 CNN 回归模型,其  $R_p^2$  为 0.923 8, RMSEP 为 0.033 4,相比于全部光谱信息,预测精度有所提高。结果表明,卷积自编码器可用于光谱信息深度特征提取,降低光谱数据维度,提高模型预测精度。

### 2.2 图像特征信息建模结果与分析

通过 ENVI5.3 软件得到某样本猪肉第一主成分图像如图 7a,将第一主成分图像转换为 768 维列向量如图 7b 所示,将列向量作为 CAE 的输入,提取主成分图像的深度特征如图 7c,深度特征经过重构解码,图像信息如图 7d 所示。对比发现经过 CAE 重构的图像信息与原始信息变化趋势大致相同,可以得出卷积编码器可用于对高光谱主成分图像深度特征提取。

全图像信息经过卷积编码器特征提取后得到深度图像信息，维度由 768 降到 64。分别对全图像特征与深度图像特征建立 CNN 预测模型，训练集与预测集的划分方

法与光谱信息建模相同，按照测试集与预测集 3：1 的比例划分。基于全部图像特征与深度图像特征建立 CNN 肌红蛋白预测模型评价结果如表 2 所示。

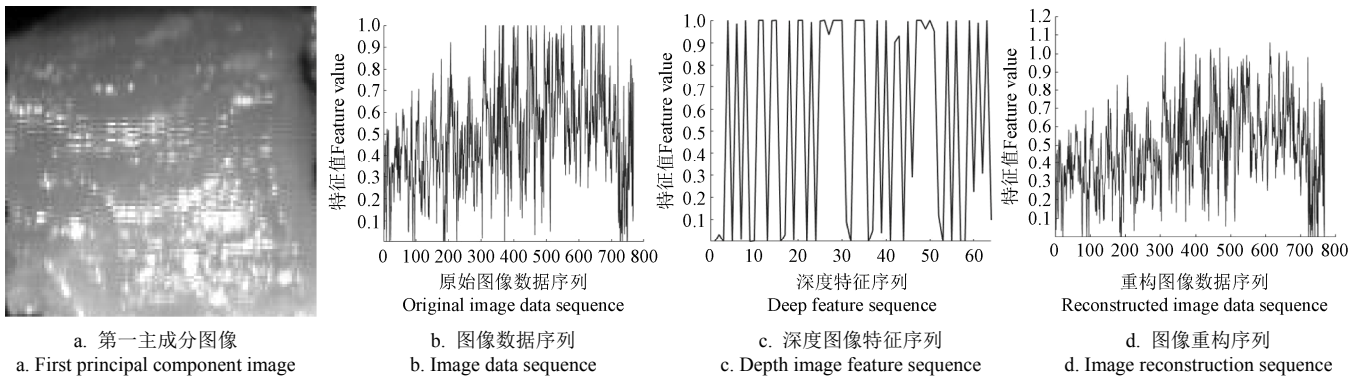


图 7 图像深度特征提取  
Fig.7 Image depth feature extraction

表 2 基于图像特征的 CNN 模型预测结果

Table 2 CNN model prediction results based on image features

组别 Group	全图像信息 Full image information				卷积自编码器 Convolutional Auto Encoder			
	$R_c^2$	RMSEC	$R_p^2$	RMSEP	$R_c^2$	RMSEC	$R_p^2$	RMSEP
脱氧肌红蛋白	0.737 4	0.114 9	0.725 0	0.146 7	0.837 0	0.082 4	0.772 1	0.105 8
氧合肌红蛋白	0.792 5	1.479 0	0.761 5	1.674 1	0.843 3	1.228 5	0.828 7	1.302 7
高铁肌红蛋白	0.791 9	1.914 5	0.743 9	2.276 3	0.887 8	1.074 3	0.825 4	1.566 7

由表 2 得到，对于第一主成分高光谱图片全部图像信息，以 OxyMb 为例， $R_c^2$  为 0.792 5，RMSEC 为 1.479 0， $R_p^2$  为 0.761 5，RMSEP 为 1.674 1，模型预测效果较差，原因在于 768 维列向量未经特征提取，包含大量冗余信息。经过 CAE 深度特征提取后，图像信息原始维度由 768 降到 64 维，预测模型  $R_c^2$  为 0.843 3，RMSEC 为 1.228 5， $R_p^2$  为 0.828 7，RMSEP 为 1.302 7，与全图像信息模型对比发现，基于深度图像特征建模效果提升，但预测精度达不到预测模型需求。

### 2.3 基于融合信息的建模结果与分析

参考文献[12]数据融合方法，将光谱信息与图像信息进行数据层的融合，800 维的光谱信息与 768 维的主成分图像数据得到 1 568 维列向量，并输入到 CAE 提取融合深度特征，样本数据集的划分与 2.1 节相同。为进一步验证基于高光谱图像信息预测猪肉肌红蛋白含量的有效性，设计 PLSR、SVR 与 CNN 模型的肌红蛋白对比试验，建模结果如表 3。

由表 3 得到基于图谱融合特征建立预测模型的  $R_p^2$  均大于 0.85，试验结果验证基于融合深度特征 CNN、PLSR、SVR 预测模型对肌红蛋白含量预测合理性。对比 3 个模型，以预测 DeoMb 含量为例，CNN 回归模型的  $R_p^2$  为 0.964 5，RMSEP 为 0.015 8；PLSR 预测模型的  $R_p^2$  为 0.867 3，RMSEP 为 0.069 2；SVR 预测模型  $R_p^2$  为 0.881 9，RMSEP 为 0.063 7；不同模型的决定性系数  $R^2$  接近于 1 表明模型回归模型拟合效果越好，CNN 模型通过卷积操作能抽取出比 PLSR、SVR 更好的非线性特征。通过非线性特征预测肌红蛋白含量的均方根误差较小，相比于线性模型预测更准确。

表 3 基于融合信息模型预测结果

Table 3 Fusion information modeling results

组别 Group	卷积神经网络 CNN				偏最小二乘回归 PLSR				支持向量回归 SVR			
	$R_c^2$	RMSEC	$R_p^2$	RMSEP	$R_c^2$	RMSEC	$R_p^2$	RMSEP	$R_c^2$	RMSEC	$R_p^2$	RMSEP
脱氧肌红蛋白	0.976 9	0.013 0	0.964 5	0.015 8	0.871 0	0.064 9	0.867 3	0.069 2	0.894 4	0.056 5	0.881 9	0.063 7
氧合肌红蛋白	0.978 3	0.173 9	0.973 2	0.226 6	0.905 0	0.733 9	0.855 7	1.068 2	0.918 7	0.621 5	0.907 2	0.721 5
高铁肌红蛋白	0.962 6	0.317 6	0.958 5	0.381 6	0.868 1	1.175 3	0.858 3	1.375 9	0.909 4	0.806 0	0.860 8	1.176 6

结合表 1、2、3 总结不同特征的 CNN 模型预测对比。以 MetMb 为例，基于光谱深度特征 CNN 回归模型其  $R_p^2$  为 0.909 2，RMSEP 为 0.809 1；基于图像深度特征 CNN 回归模型其  $R_p^2$  为 0.825 4，RMSEP 为 1.566 7；基于融合信息深度特征 CNN 回归模型其  $R_p^2$  为 0.958 5，RMSEP 为 0.381 6。对比 3 个模型效果，基于融合深度特征模型相

比于光谱特征与图像特征。其预测集决定系数分别提高 5.42%、16.12%。该结果表明图-谱融合特征能较好地反映出猪肉内部与外部信息，基于融合特征建立回归模型能提高预测精度。

基于以上结果分析在提取光谱特征时，采集区域集中在精瘦肉，采集的光谱信息与在采集较大光斑的图像反射率不同，导致两者的光谱曲线有差异，光谱信息包

含特征点不足。通过提取样本主成分图像的图像特征,来弥补不足。猪肉样本高光谱主成分图像包含样本颜色、纹理等特征,卷积神经网络较强的特征提取能力,提取图像深层次特征。采用图像特征与光谱特征的融合能获得更加全面的特征点。为进一步验证模型的可靠性,再次随机选取 50 个样本数据作为预测集,肌红蛋白含量预测值与实测值比较如图 8 所示,预测集决定系数均大于 0.85,进一步验证模型具有较好的预测能力。

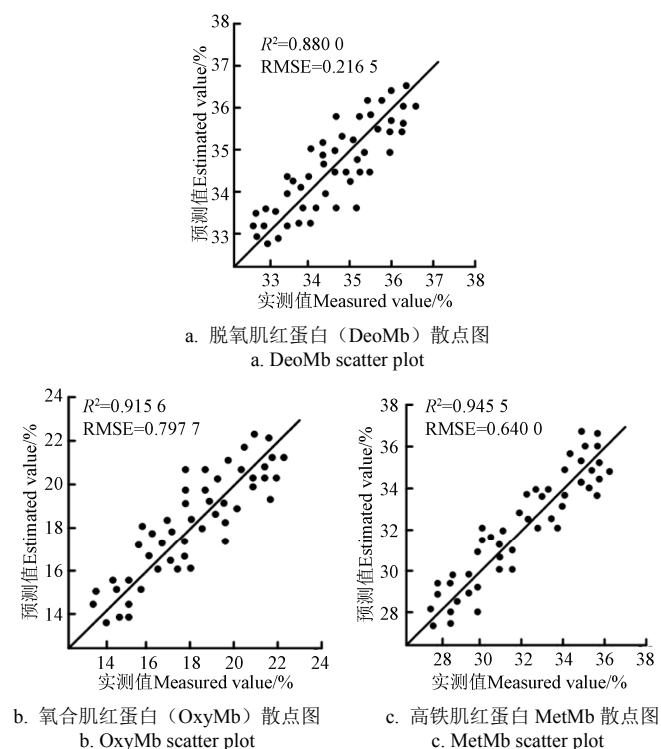


图 8 肌红蛋白含量的预测值与实测值

Fig.8 Predicted and measured values of myoglobin content

### 3 结论

本文采集冷藏 4℃ 的猪肉在 0~5 d 试验周期内猪肉高光谱的光谱与图像信息,采用卷积自编码器对光谱信息、图像信息及两者融合信息进行深度特征提取,并建立卷积神经网络 (Convolutional Neural Network, CNN)、偏最小二乘回归 (Partial Least Squares Regression, PLSR)、支持向量机回归 (Support Vector Regression, SVR) 猪肉肌红蛋白含量预测模型,得到以下结论:

1) 基于全波段与经过卷积自编码器提取深度特征建立 CNN 肌红蛋白预测模型,其中全波段光谱信息建立肌红蛋白模型,脱氧肌红蛋白 (Deoxymyoglobin, DeoMb)、氧合肌红蛋白 (Oxygenated myoglobin, OxyMb)、高铁肌红蛋白 (Metmyoglobin, MetMb) 的预测集决定系数分别为 0.855 1、0.886 2、0.861 8。基于深度特征建立回归模型 DeoMb, OxyMb, MetMb 的预测集决定系数分别 0.923 8、0.920 3、0.909 2,基于深度光谱特征建立模型决定系数均有提高。可以得出,卷积神经网络对于光谱数据有特征提取功能,可用于光谱数据研究与分析。

2) 基于光谱-图像深度融合特征建立卷积神经网络肌红蛋白回归模型,DeoMb, OxyMb, MetMb 的预测集决定系数分别 0.964 5、0.973 2、0.958 5,相比于建立的光谱、图像特征模型,其预测集决定系数较高,均方误差较低。说明融合特征包含更加全面的猪肉样本信息,基于融合特征建立回归模型能提高预测准确度。

3) 基于图谱融合特征建立 CNN、PLSR、SVR 3 个回归模型,对比三者决定系数可以得出:利用融合特征建立 CNN 预测模型准确度较高,有广阔应用场景,为高光谱图像处理提供新的方法。

### [参 考 文 献]

- [1] 汤晓艳, 赵小丽, 徐学万. 大宗猪肉质量分级现状与实施对策建议[J]. 农产品质量与安全, 2020(5): 36-40.  
Tang Xiaoyan, Zhao Xiaoli, Xu Xuewan. Current status of bulk pork quality classification and suggestions for implementation[J]. Quality and Safety of Agro-products, 2020(5):36-40. (in Chinese with English abstract).
- [2] Neethling N E, Suman S P, Sigge G O, et al. Exogenous and endogenous factors influencing color of fresh meat from ungulates[J]. Meat and Muscle Biology, 2017, 1(1):253-257.
- [3] 马超. 贮藏期内滩羊熟肉肌红蛋白含量变化的高光谱模型构建及空间分布[D]. 宁夏: 宁夏大学, 2019.  
Ma Chao. Hyperspectral Model and Spatial distribution of Myoglobin Content in Cooked Meat of Tan Sheep During Storage Period[D]. Ningxia: Ningxia University, 2019. (in Chinese with English abstract).
- [4] Yuan R, Liu G, He J, et al. Determination of metmyoglobin in cooked tan mutton using Vis/NIR hyperspectral imaging system[J]. Journal of Food Science, 2020, 85(22):1403-1410.
- [5] 唐海涛, 孟祥添, 苏循新, 等. 基于 CARS 算法的不同类型土壤有机质高光谱预测[J]. 农业工程学报, 2021, 37(2): 105-113.  
Tang Haitao, Meng Xiangtian, Su Xunxin, et al. Hyperspectral prediction on soil organic matter of different types using CARS algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(2): 105-113. (in Chinese with English abstract).
- [6] 杨越超, 赵英俊, 秦凯, 等. 黑土养分含量的航空高光谱遥感预测[J]. 农业工程学报, 2019, 35(20): 94-101.  
Yang Yuechao, Zhao Yingjun, Qin Kai, et al. Prediction of black soil nutrient content based on airborne hyperspectral remote sensing[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(20): 94-101. (in Chinese with English abstract).
- [7] Wang Y, Ma H B, Wang J Z, et al. Hyperspectral monitor of soil chromium contaminant based on deep learning network model in the Eastern Junggar coalfield[J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2021, 257: 119739.
- [8] 邵园园, 王永贤, 玄冠涛, 等. 高光谱成像快速检测壳聚糖涂膜草莓可溶性固形物[J]. 农业工程学报, 2019, 35(18): 245-254.  
Shao Yuanyuan, Wang Yongxian, Xuan Guantao, et al. Rapid detection of soluble solids content in strawberry coated with chitosan based on hyperspectral imaging[J]. Transactions of



- the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(18): 245-254. (in Chinese with English abstract).
- [9] 袁瑞瑞, 刘贵珊, 何建国, 等. 可见近红外高光谱成像对灵武长枣定量损伤等级判别[J]. 光谱学与光谱分析, 2021, 41(04): 1182-1187.  
Yuan Ruirui, Liu Guishan, He Jianguo, et al. Quantitative damage identification of lingwu long jujube based on visible near-infrared hyperspectral imaging, spectroscopy and spectral analysis, 2021, 41(04): 1182-1187. (in Chinese with English abstract)
- [10] Kucha C T, Liu L, Ngadi M, et al. Hyperspectral imaging and chemometrics as a non-invasive tool to discriminate and analyze iodine value of pork fat[J]. Food Control, 2021, 127: 108145.
- [11] Wan G, Liu G, He J, et al. Feature wavelength selection and model development for rapid determination of myoglobin content in nitrite-cured mutton using hyperspectral imaging[J]. Journal of Food Engineering, 2020, 287: 110090.
- [12] 孙俊, 靳海涛, 芦兵, 等. 基于高光谱图像及深度特征的大米蛋白质含量预测模型[J]. 农业工程学报, 2019, 35(15): 295-303.  
Sun Jun, Jin Haitao, Lu Bing, et al. Prediction model of rice protein content based on hyperspectral image and deep feature[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(15): 295-303. (in Chinese with English abstract).
- [13] 王彩霞, 王松磊, 贺晓光, 等. 高光谱图谱融合检测羊肉中饱和脂肪酸含量[J]. 光谱学与光谱分析, 2020, 40(02): 595-601.  
Wang Caixia, Wang Songlei, He Xiaoguang, et al. Detection of saturated fatty acid content in mutton by using the fusion of hyperspectral Spectrum and image information[J]. Spectroscopy and Spectral Analysis, 2020, 40(2): 595-601. (in Chinese with English abstract).
- [14] 翁士状, 唐佩佩, 张雪艳, 等. 高光谱成像的图谱特征与卷积神经网络的名优大米无损鉴别[J]. 光谱学与光谱分析, 2020, 40(9): 2826-2833.  
Wong Shizhuang, Tang Peipei, Zhang Xueyan, et al. Non-destructive identification method of famous rice based on image and spectral features of hyperspectral imaging with convolutional neural network[J]. Spectroscopy and Spectral Analysis, 2020, 40(9): 2826-2833.
- [15] 孙俊, 靳海涛, 武小红, 等. 基于低秩自动编码器及高光谱图像的茶叶品种鉴别[J]. 农业机械学报, 2018, 49(8): 316-323.  
Sun Jun, Jin Haitao, Wu Xiaohong, et al. Tea variety identification based on low-rank stacked auto-encoder and hyperspectral image[J]. Transaction of the Chinese Society for Agricultural Machinery, 2018, 49(8): 316-323. (in Chinese with English abstract).
- [16] 李晴晴, 侯瑞春, 丁香乾. 基于改进堆叠自编码器的滚动轴承故障诊断[J]. 计算机工程与设计, 2019, 40(7): 2064-2070.  
Li Qingqing, Hou Ruichun, Ding Xiangqian. Roller bearing fault diagnosis based on improved stacked auto-encoder[J]. Computer Engineering and Design, 2019, 40(7): 2064-2070.
- [17] 史杨, 王儒敬, 汪玉冰. 利用改进自动编码器光谱法预测土壤有机质[J]. 发光学报, 2018, 39(10): 1458-1465.
- Shi Yang, Wang Rujing, Wang Yubing. Prediction of soil organic matter by improved auto encoder based on Near-infrared spectroscopy[J]. Chinese Journal of Luminescence, 2018, 39(10): 1458-1465. (in Chinese with English abstract).
- [18] Feng J, Liu L, Cao X, et al. Marginal stacked autoencoder with adaptively-spatial regularization for hyperspectral image classification[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 6: 15224-15235.
- [19] 田启川, 王满丽. 深度学习算法研究进展[J]. 计算机工程与应用, 2019, 55(22): 25-33.  
Tian Qichuan, Wang Manli. Research progress on deep learning algorithms[J]. Computer Engineering and Applications, 2019, 55(22): 25-33. (in Chinese with English abstract).
- [20] 张思雨, 张秋菊, 李可. 采用机器视觉与自适应卷积神经网络检测花生仁品质[J]. 农业工程学报, 2020, 36(4): 269-277.  
Zhang Siyu, Zhang Qiuju, Li Ke. Detection of peanut kernel quality based on machine vision and adaptive convolution neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(4): 269-277. (in Chinese with English abstract).
- [21] Acquarelli J, Van Laarhoven T, Gerretzen J, et al. Convolutional neural networks for vibrational spectroscopic data analysis[J]. Analytica Chimica Acta, 2017, 954: 22-31.
- [22] Krzywicki K. The determination of haem pigments in meat[J]. Meat Science, 1982, 7(1): 29-36.
- [23] 姜微. 高光谱技术在马铃薯品种鉴别及品质无损检测中的应用研究[D]. 哈尔滨: 东北农业大学, 2017.  
Jang Wei. Study on Nondestructive Detection of Identification Varieties and Quality Potato Using Hyperspectral Technology[D]. Harbin: Northeast Agricultural University, 2017. (in Chinese with English abstract).
- [24] 王春雷, 陈婉芝, 卢彩云, 等. 基于高光谱成像的玉米收获后根茬行分割方法[J]. 农业机械学报, 2020, 51(S2): 421-426.  
Wang Chunlei, Chen Wanzhi, Lu Caiyun, et al. Segmentation method for maize stubble row based on hyperspectral Imaging[J]. Transaction of the Chinese Society for Agricultural Machinery, 2020, 51(S2): 421-426. (in Chinese with English abstract).
- [25] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [26] Chen Y, Lin Z, Xing Z, et al. Deep learning-based classification of hyperspectral data[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2017, 7(6): 2094-2107.
- [27] 曲建岭, 余路, 袁涛, 等. 基于一维卷积神经网络的滚动轴承自适应故障诊断算法[J]. 仪器仪表学报, 2018, 39(7): 134-143.  
Qu Jianling, Yu Lu, Yuan Tao, et al. Adaptive fault diagnosis algorithm for rolling bearings based on one-dimensional convolutional neural network[J]. Chinese Journal of Scientific Instrument, 2018, 39(7): 134-143. (in Chinese with English abstract).

- [28] 陈健, 刘明, 熊鹏, 等. 基于卷积自编码神经网络的心电信号降噪[J]. 计算机工程与应用, 2020, 56(16): 148-155. Chen Jian, Liu Ming, Xiong Peng, et al. ECG signal denoising based on convolutional auto-encoder neural network[J]. Computer Engineering and Applications, 2020, 56 (16): 148-155. (in Chinese with English abstract).
- [29] 周琪. 基于自编码器与卷积神经网络的遥感对象分类[D]. 武汉: 武汉理工大学, 2018. Zhou Qi. Classification of Remote Sensing Objects Based on Autoencoder and Convolutional Neural Network[D]. Wuhan: Wuhan University of Technology, 2018(in Chinese with English abstract).
- [30] 李小占, 马本学, 喻国威, 等. 基于深度学习与图像处理的哈密瓜表面缺陷检测[J]. 农业工程学报, 2021, 37(1): 223-232. Li Xiaozhan, Ma Benxue, Yu Guowei, et al. Surface defect detection of Hami melon using deep learning and image processing[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(1): 223-232. (in Chinese with English abstract).

## Non-destructive detection of pork myoglobin content based on hyperspectral technology

Wang Lishu, Hu Jinyao, Fang Junlong<sup>\*</sup>, Chen Xi, Li Chuang

(School of Electricity and Information, Northeast Agricultural University, Harbin 150030)

**Abstract:** Hyperspectral imaging system can widely be expected to acquire a set of sample images within certain spectral bands in each pixel at the same time. In this study, rapid detection was proposed for the myoglobin content in pork samples using spectral images and deep learning. The pork was placed under the cold storage conditions at 4°C, where a total of 250 pork samples were settled at different times (0-5 d). A hyperspectral imager was used to collect the pork hyperspectral images (400 to 1 000 nm). ENVI5.3 software was also selected to determine the region of interest (ROI) in the hyperspectral images, thereby extracting the full-band average spectrum and principal component image of ROI. Subsequently, a Savitzky-Golay (SG) filter was used to denoise the spectral information for the curve smoothness and spectral resolution. A convolutional auto encoder (CAE) was utilized to extract spectral depth features. A prediction model was finally established for the content of deoxymyoglobin (DeoMb), oxymyoglobin (OxyMb), and metmyoglobin (MetMb) in the pork samples. The results showed that the determination coefficients of test datasets were 0.923 8, 0.920 3, and 0.909 2, and the root mean square errors (RMSE) were 0.033 4, 0.619 7, and 0.809 1, respectively. Furthermore, the image information of adjacent wavelengths was highly correlated against the image extraction and storage. Principal Component Analysis (PCA) was utilized to reduce the dimension of hyperspectral images for better storage and processing. As such, the images under all bands were linearly combined to form a principal component image in the ENVI5.3 software. The first three principal component images represented 90.62% of the original hyperspectral image, where the contribution rate of the first principal component was 88.50%, indicating the most information. Therefore, the first principal component image was selected for the subsequent image extraction. The first principal component image was unified to the size of 16×16 pixels, and then converted into a 768-dimensional column vector for the extraction of image depth features using a convolutional encoder. DeoMb, OxyMb, and MetMb content prediction models were established using image depth features, in which the determination coefficients of test datasets were 0.772 1, 0.828 7, and 0.825 4, while the RMSE of prediction were 0.105 8, 1.302 7, and 1.566 7. The spectral and image features were fused at the data level, and then the fusion data was input into the CAE to extract the deep fusion features. The DeoMb, OxyMb, and MetMb content prediction models were also established using the fusion depth features. The determination coefficients of test datasets were 0.964 5, 0.973 2, and 0.958 5, while the RMSE of prediction were 0.015 8, 0.226 6, and 0.381 6. Obviously, the determination coefficients of the test dataset were improved, while the RMSE were reduced, compared with the individual image and spectrum information. Partial least square regression (PLSR) and support vector machine regression (SVR) prediction models were also established to further verify the relationship between the graph-spectrum fusion feature and pork myoglobin. It was found that the determination coefficients of the test dataset were greater than 0.85. Consequently, the convolutional autoencoder can be expected to extract the deep fusion features of image and spectral information. Moreover, the fusion features can better reflect the internal and external information of pork. The CNN regression model using the fusion features can also be used to improve the prediction accuracy. This finding can provide a new better way to detect the myoglobin content in pork using hyperspectral imaging.

**Keywords:** nondestructive detection; spectral feature; hyperspectral image; convolutional neural network; convolutional autoencoder