

高光谱透射成像联合特征区间筛选的鸽蛋受精无损检测方法

郭文艳^{1,2} 李晓玉^{1,2} 徐善金³ 冯春刚⁴ 刘龙申^{1,2}

(1. 南京农业大学智慧农业学院(人工智能学院), 南京 210031; 2. 农业农村部养殖装备重点实验室, 南京 210031; 3. 南京东晨鸽业有限公司, 南京 211500; 4. 南京农业大学动物科技学院, 南京 210095)

摘要: 为提升鸽蛋育种与孵化管理的无损检测效率与准确性,构建了一套面向孵化期种鸽蛋受精状态快速判别的高光谱透射成像识别流程。在高光谱透射式装置下采集原始高光谱数据,进行黑白校正后,将整个蛋体作为感兴趣区域获取其平均光谱信息。采用了标准正态变量校正(Standard normal variate, SNV)、小波变换(Wavelet transform, WT)、多元散射校正(Multiplicative scatter correction, MSC)3种预处理算法;通过KS样本选取策略按照比例7:3划分训练集与测试集;结合BiPLS、iPLS与SiPLS对光谱区间进行特征波段区间筛选;在此基础上分别训练线性模型(LDA、LR)与非线性分类模型(XGBoost),并分析最优预处理—特征选择—建模组合与关键判别波段。采取一阶导数处理方法,对受精/非受精的光谱差异进行详细分析。结果表明,在孵化前分类效果最好的是BiPLS-LDA-WT算法,对受精种蛋的召回率达到0.98,准确率为0.87,F1分数为0.89;在孵化第1天、第2天和第3天,分类准确率分别为0.92、0.95和1.00,召回率分别为0.98、1.00、1.00。本研究所提出的集成式高光谱特征区间筛选与多模型对比框架具备良好的准确性与可解释性,可为种鸽蛋孵化早期无损受精检测及孵化管理提供技术支撑。

关键词: 种鸽蛋; 高光谱透射成像; 受精信息; 波段选择; 无损检测

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2026)15-0094-09

OSID:



Non-destructive Fertility Detection of Pigeon Eggs Using Hyperspectral Transmission Imaging Combined with Feature Interval Selection

Guo Wenyan^{1,2} Li Xiaoyu^{1,2} Xu Shanjin³ Feng Chungang⁴ Liu Longshen^{1,2}

(1. College of Smart Agriculture (College of Artificial Intelligence), Nanjing Agricultural University, Nanjing 210031, China
2. Key Laboratory of Breeding Equipment, Ministry of Agriculture and Rural Affairs, Nanjing 210031, China
3. Nanjing Dongchen Pigeon Industry Co., Ltd., Nanjing 211500, China
4. College of Animal Science and Technology, Nanjing Agricultural University, Nanjing 210095, China)

Abstract: Aiming to improve the efficiency and accuracy of non-destructive testing in pigeon egg breeding and incubation management, a hyperspectral transmission imaging recognition process was established for the rapid determination of the fertilization status of breeding pigeon eggs in the early stages of incubation. Original hyperspectral data were collected by using a hyperspectral transmission device, and after black-and-white correction, the entire egg was taken as the region of interest to obtain its average spectral information. Three preprocessing algorithms were applied: standard normal variate (SNV), wavelet transform (WT), and multiplicative scatter correction (MSC). Using a KS sample selection strategy, the data were split into training and test sets at a ratio of 7:3. Feature wavelength intervals were selected by BiPLS, iPLS, and SiPLS for spectral interval screening. Based on this, linear models (LDA, LR) and nonlinear classification models (XGBoost) were trained, and the optimal preprocessing-feature selection-modeling combination and key discriminative wavelengths were analyzed. First derivative processing was used for a detailed analysis of the spectral differences between fertilized

收稿日期: 2026-03-15 修回日期: 2026-04-15

基金项目: 江苏省重点研发计划项目(BE2022315)

作者简介: 郭文艳(2000—),女,硕士生,主要从事农业智能检测技术研究,E-mail: 2023119014@stu.njau.edu.cn

通信作者: 刘龙申(1983—),男,副教授,博士,主要从事畜禽精准养殖信息感知关键技术研究,E-mail: liulongshen@njau.edu.cn

and unfertilized eggs. The results showed that, before incubation, the BiPLS - LDA - WT algorithm achieved the best classification performance, with a recall rate of 0.98 for fertilized eggs, an accuracy of 0.87, and an F1-score of 0.89. On the 1st, 2nd, and 3rd days of incubation, the classification accuracies were 0.92, 0.95, and 1.00, respectively, with recall rates of 0.98, 1.00, and 1.00. The proposed integrated hyperspectral feature interval selection and multi-model comparison framework demonstrated good accuracy and interpretability, providing technical support for early non-destructive fertilization detection and incubation management of breeding pigeon eggs.

Key words: breeding pigeon eggs; hyperspectral transmission imaging; fertilization information; band selection; non-destructive testing

0 引言

受精蛋的早期精准识别是鸽业生产中的核心环节,对于提升孵化率、优化种蛋管理以及降低资源消耗具有关键意义^[1]。受精蛋与未受精蛋在外观上难以区分,传统的判别方式难以满足现代规模化养殖对精准、高效识别的需求。因此,探索科学高效的受精蛋早期识别技术,成为推动鸽业可持续发展的重要课题。

目前,种鸽蛋受精识别方法主要包括人工观察、光照法等传统手段。人工方法依赖经验,主观性强且效率低下;光照法虽然简单,但对蛋壳厚度和光源条件敏感,准确率有限。上述方法往往需要在鸽蛋孵化第 5 天进行操作,然而未受精蛋在孵化过程中容易变质腐败,滋生细菌和霉菌,成为孵化环境中的潜在污染源。早期剔除未受精蛋有助于控制孵化场的生物安全,降低疾病传播风险。实现禽蛋孵化早期无损检测的主要方法有光电检测技术、机器视觉法、光谱技术等。Khaliduzzaman 等^[2]用输入电流、平均电压来记录种鸡蛋的发育过程,以种蛋不透明度的变化过程作为判断种蛋胚胎受精的依据。但是光电法存在受蛋壳干扰大、早期检测灵敏度不足的问题。

机器视觉具有成本低、效率高和客观等优点,也是常用的无损检测方法。机器视觉在禽蛋无损检测中具有高效率、高准确率和自动化的优势,但对光照和硬件依赖性强、早期判别难、模型数据依赖大、泛化能力不足,是其推广应用的主要瓶颈。

近年来,随着高光谱成像技术的发展,其在农业^[3]及禽蛋^[4]领域的应用日益广泛。高光谱成像结合了光谱信息与空间分布特征,能够实现对蛋体内部结构及成分进行无损、快速检测。在鸡蛋受精检测方面,已有大量研究验证了方法的有效性与可行性。研究者引入并优化了 SSA - SVM^[5]、朴素贝叶斯^[6]及深度神经网络(DNN)^[7]等机器学习模型,并结合有效波段筛选技术^[8];在采集策略上,明确了最佳检测部位(如赤道侧)、孵化时间^[9]以及多维

度(X 、 Y 、 Z 轴^[10])对精度的积极影响;此外,祝志慧等^[11]将单一光谱拓展为光谱与高分辨率图像的多模态融合。与此同时,深度学习框架也被引入高光谱分类。Wang 等^[12]提出了一种基于透射光谱的新型混合深度学习检测框架 CVAE - DF,在鸡蛋测试数据集上达到了 95.94% 的整体准确率。关于鸭蛋与鸽蛋的研究亦有进展,陈灼廷等^[13]提出的 CAE - 1DCNN 模型能较好地提取鸭蛋光谱数据中的有效信息,判别准确率为 95.06%。李庆旭等^[14]在 Alexnet 神经网络基础上进行改进,将孵化第 3 天的种鸭蛋透射图像直接输入到深度卷积神经网络,测试集分类准确率高达 98.87%。Li 等^[15]提出的 MobileOne - YOLO 模型在未受精鸭蛋、死胚蛋和健康胚蛋的检测中表现出色,准确率达到 98.10%,精确率和召回率均为 98.55%。Lin 等^[16]在孵化的第 3 天采集每个鸽蛋的光谱,采用 MWPLS - DA 模型结合 STD - MC 方法,预测受精状态的准确率可达到 98.60%。

已有研究表明,高光谱技术在鸡蛋^[17-19]、鸭蛋^[20-21]等禽蛋的受精状态识别中取得了较好效果,但在鸽蛋的受精识别领域相关研究尚处于起步阶段,现有的方法体系和应用模式仍不成熟,难以满足种鸽蛋早期受精状态识别的实际需求。鸽蛋与鸡蛋、鸭蛋在生物学特性、蛋壳结构及光谱响应上存在显著差异,难以直接套用现有研究的方法体系。首先,在物理结构与光学响应上,鸽蛋体积显著偏小、曲率更大且蛋壳极薄^[22],与常规家禽蛋的光谱采集参数不匹配;其次,在内部组分与发育节律上,鸽蛋具有独特的蛋清蛋黄比例,且其孵化周期较短(约 18 d),早期胚胎发育进程更为迅速。为此,本研究面向种鸽蛋孵化早期受精状态精准判别的实际需求,系统整合高光谱透射成像技术、光谱数据预处理方法与有效光谱特征区间筛选算法,深入挖掘表征受精状态的关键光谱特征区间,并在此基础上构建适配的判别模型,旨在建立一套专用于种鸽蛋受精情况的智能、无损识别方法,为鸽蛋精准化、智能化育种与孵化管理提供技术支撑。

1 材料与方法

1.1 实验材料

实验样品是产自南京东晨鸽业公司的种鸽蛋，所用鸽种为灰羽王，种鸽蛋大小接近，蛋壳颜色均为白色，如图 1 所示。实验共采集鸽蛋 325 枚，分 3 个批次采集：第 1 批次采集 58 枚，开始时间为 2025 年 5 月 7 日；第 2 批次采集 118 枚，开始时间为 2025 年 6 月 12 日；第 3 批次采集 149 枚，开始时间为 2025 年 10 月 14 日。



图 1 孵化机内鸽蛋
Fig. 1 Pigeon eggs in incubator

在进孵化箱前，用体积分数为 75% 的乙醇溶液擦拭种鸽蛋的表面脏污，随后自然风干。进行称量与测量大小，挑选质量较为均匀的种蛋（ (20 ± 3) g）。所有种鸽蛋在孵化前均编号。孵化采用济南四只熊网络科技有限公司的孵化机，初始孵化温度设定为 38.2℃，相对湿度为 60%，每 2 h 翻蛋一次。

种鸽蛋在孵化箱孵化 5 d 后取出，进行照蛋观察，若能在蛋内发现血丝，则该枚鸽蛋为受精蛋，无法分辨的种蛋采取破壳方式进行确认。统计后共得到 173 枚受精蛋，152 枚无精蛋，共 325 枚。剔除孵化前(d0)阶段 2 个采集质量不合格的受精样本，最终该阶段保留有效数据 323 条。

1.2 光谱数据采集与预处理

采用推扫式高光谱仪采集鸽蛋的高光谱透射图像，种鸽蛋横向放置，卤素光源位于正下方，光谱相机位于上方 40 cm 处。光线穿过鸽蛋，最终被相机收集到。具体地，高光谱采集系统的物理布局如图 2 所示。成像过程中，相机扫描范围覆盖整个蛋体，采集图像空间分辨率设为 361 像素 × 1 632 像素，光谱范围为 358.5 ~ 1 021.7 nm，共 616 个波段，覆盖紫外-可见-近红外区域。剔除波谱头尾信噪比低的波段（如 358.5 ~ 400 nm 与 1 000 ~ 1 021 nm），实际有效分析波段限定在 400 ~ 1 000 nm 区间。

采集所得高光谱数据保存为 ENVI 标准格式，包括原始数据文件(. raw)和头文件(. hdr)。以入孵前采集的光谱数据作为第 0 天(d0)，并在入孵后 3 d 内每天采集一次光谱数据，即一共采集 4 d 的光

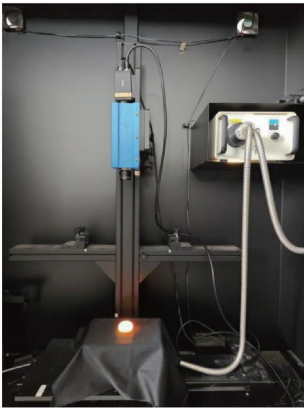


图 2 高光谱采集系统
Fig. 2 Hyperspectral acquisition system

谱数据，最终采集高光谱数据 1 298 条。
为保证光谱数据的准确性，采集后进行黑白校正，如图 3 所示。

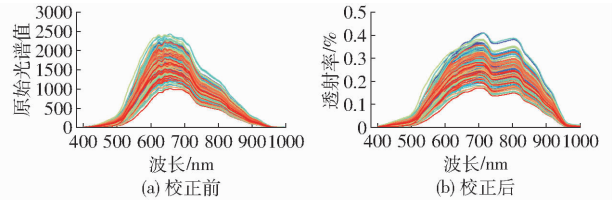


图 3 黑白校正前后 d0 光谱
Fig. 3 The d0 spectrum before and after black-and-white correction

具体方法为：先采集黑板（将镜头盖住，使其不受光）和白板图像（让线性光源完全无遮蔽对准镜头入光）。分别获得系统的零参考和最大参考值。随后将样品图像进行归一化校正，以消除设备噪声和环境光影响。

$$I_{calibrated} = \frac{I_{sample} - I_{black}}{I_{white} - I_{black}} \tag{1}$$

式中 I_{sample} ——样品图像的原始光谱值
 I_{black} ——黑板图像的光谱值
 I_{white} ——白板图像的光谱值
 $I_{calibrated}$ ——归一化校正后的光谱值

1.3 提取感兴趣区域

在每个样品图像中，手动划定蛋体区域，呈椭圆形，图 4 为孵化第 3 天的鸽蛋图像。并对该区域内所有像素的光谱数据进行均值提取，作为后续分析的代表光谱。所有样品均采用统一感兴趣区域 ROI (Region of Interest) 选取方法，保证数据处理的一致性与可比性，选取的 ROI 如图 4 所示。

1.4 数据集划分

为保证模型的训练与验证具有良好的代表性和泛化能力，本文采用 KS (Kennard-stone) 算法对原始数据集进行样本划分。具体方法为按照比例 7:3 将全部样本合理分为训练集和测试集。其中，KS 算法

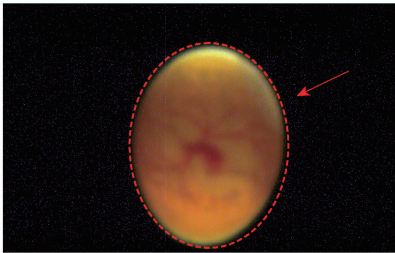


图 4 选取的 ROI
Fig. 4 Selected ROI

通过最大化样本间的空间分布距离,优先选取分布均匀的样本作为训练集,能够有效覆盖特征空间,提高模型鲁棒性。KS 算法旨在从 N 个样本中选择一个具有代表性的子集。为了确保该子集沿着 x (仪器响应)数据空间均匀分布,KS 在远离已选择样本的空间区域中进行新的选择。为此,该算法采用每对样本 (p, q) 特征向量 $\mathbf{x}$ 之间的欧几里得距离 $d_{x(p,q)}$,计算公式为

$$d_{x(p,q)} = \sqrt{\sum_{j=1}^J (x_p(j) - x_q(j))^2} \quad (p, q \in [1, N]) \quad (2)$$

对于光谱数据, $x_p(j)$ 和 $x_q(j)$ 分别是样品 p 和 q 在第 j 个波长处的仪器响应。 J 表示光谱中波长的数量。首先取距离最大的样本对为初始样本对,在每个后续迭代中,算法会计算剩余样本到已选样本的最小距离,并从中挑选出该距离最大值对应的样本。重复这样的程序,直到达到指定的样本数量。

1.5 光谱数据预处理

在本研究中,为提升高光谱数据分析的准确性与鲁棒性,系统性地对原始数据采用多种常用预处理方法,包括:标准正态变量校正 (Standard normal variate, SNV)、小波变换 (Wavelet transform, WT)、多元散射校正 (Multiplicative scatter correction, MSC)。上述方法有效消除了噪声、基线漂移、散射效应和幅值差异等非分析性干扰^[23],为后续特征提取与建模提供了更加可靠的数据基础。图 5 为 d0 中的两个样本经过预处理算法处理后的光谱曲线。

1.6 光谱特征区间筛选

为有效筛选与受精状态判别相关的光谱特征,本文在训练集内采用多种特征选择与降维策略对全波段数据进行处理。

偏最小二乘回归 (PLS) 通过对光谱矩阵 $\mathbf{X}$ 和响应矩阵 $\mathbf{Y}$ 同时进行分解,找到 $\mathbf{X}$ 中最能预测潜在变量的隐藏变量,可使 $\mathbf{X}$ 和 $\mathbf{Y}$ 之间的协方差最大化,并和矩阵 $\mathbf{Y}$ 进行建模,并避免了多重共线性,是应用最广泛的多元校正方法。而样品光谱中通常包含不相关或无用的变量信息,需要进行筛选和剔除,从

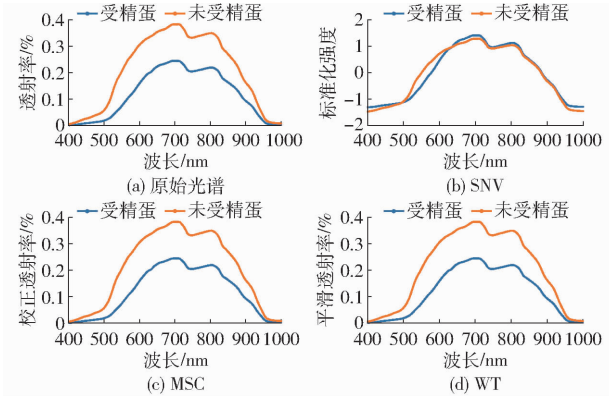


图 5 各种方法预处理后的光谱曲线

Fig. 5 Spectral curves after preprocessing by various methods

而提高模型的预测性。间隔偏最小二乘法 (Interval partial least-squares regression, iPLS)^[24] 是将模型区间划分为多个具有相同宽度的子区间,并在每个子区间上分别建立 PLS 模型,然后选择 5 折交叉验证 (5-fold cross-validation, 5 - CV) 准确率最高的子区间作为特征波段。联合区间偏最小二乘法 (Synergy iPLS, SiPLS) 是在 iPLS 的基础上,考虑多个区间的协同作用。首先将建模区间分成若干等份,利用自由组合 C ,从几个建模子区间中随机选择 m (通常 $m \leq 5$) 个子区间建立有机质含量的 PLS 定量模型,然后以 5 - CV 准确率最高的区间组合作为特征波段建立 PLS 回归模型。

后向间隔偏最小二乘线性判别分析 (Backward interval partial least squares linear discriminant analysis, BiPLS - LDA)^[25] 引入了“反向间隔”的概念。完整算法流程如图 6 所示。

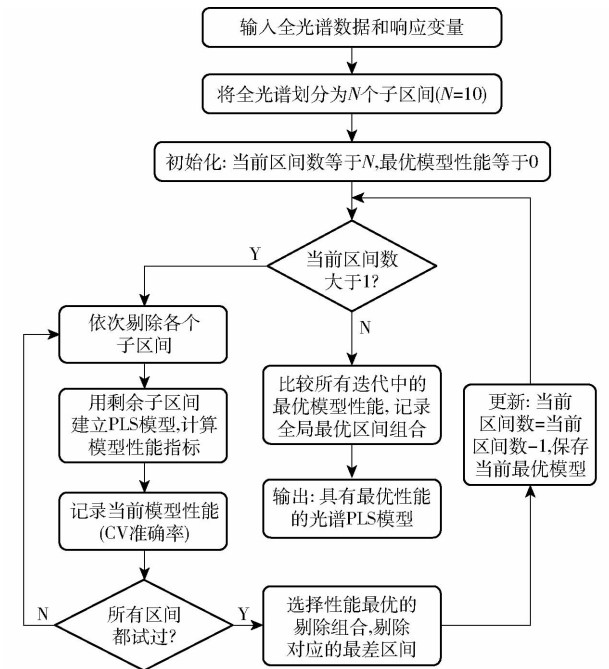


图 6 BiPLS 流程图

Fig. 6 BiPLS algorithm flowchart

具体而言,首先将全波段光谱等分为 N 个连续区间,基于全部区间构建 PLS 模型,并计算模型性能指标。随后,算法进入“反向区间剔除”过程:在每一次迭代中,依次对每个候选区间执行“暂时剔除”操作,重建 PLS 模型并评估性能,剔除对模型贡献最小(剔除后模型性能最优)的区间。此过程重复直至所有区间均被剔除,整个过程中记录每一轮剔除后得到的模型性能,最终选取 CV 准确率最高的一组区间作为最优特征区间组合。得到最优区间后,使用该区间内的光谱数据重新建立 PLS 模型,并将其输出变量或特征得分作为输入送入分类器,从而构建分类模型。

1.7 分类模型建立

在完成预处理和特征提取后,针对受精状态判别任务,采用 ROS(Random over sampling)随机过采样策略以缓解类别不平衡问题。基于训练集,分别训练线性判别分析(Linear discriminant analysis, LDA)、逻辑回归(Logistic regression, LR)和极端梯度提升(Extreme gradient boosting, XGBoost)3 种分类模型。LDA 是一种经典的线性分类方法,通过在特征空间中寻找能最大化类间方差、最小化类内方差的投影方向,实现样本的线性可分。其计算效率高、模型解释性强。LR 通过学习特征与类别之间的线性关系,使用 Sigmoid 函数将输出映射为类别概率,是一种稳健的二分类算法。该方法能够提供概率意义的分类结果,并可通过正则化项防止过拟合。XGBoost 是一种基于梯度提升框架的集成学习算法,通过多棵弱分类树的迭代加权组合实现非线性特征学习。其具有高精度、良好的泛化能力和处理非线性复杂关系的优势。通过分层 k 折交叉验证($k=5$)对模型进行超参数优化与选择,最后在独立测试集上报告各模型的泛化性能,同时结合不同特征方案的结果进行对比,选择性能与可解释性兼顾的最优模型应用于受精状态自动判别。

1.8 性能评价方法

为评估所构建分类模型的性能,本文采用准确度(A)、精确率(P)、召回率(R)与 F1 分数(F1)作为评价指标。

2 结果与分析

2.1 光谱特征对比分析

一阶导数在光谱数据处理中具有重要作用,它能够突出光谱曲线中的细微变化,使吸收峰位置更加明显。通过对原始透射率光谱求一阶导数,可以揭示出隐藏在整体趋势下的局部差异,放大受精蛋与未受精蛋等样本之间的特征差异,能够更明显地

观察受精蛋与非受精蛋的差异。

随机选取一个受精蛋样品与非受精蛋样品在鸽蛋入孵前(d0)的光谱数据,绘制其一阶导数对比图,如图 7 所示。

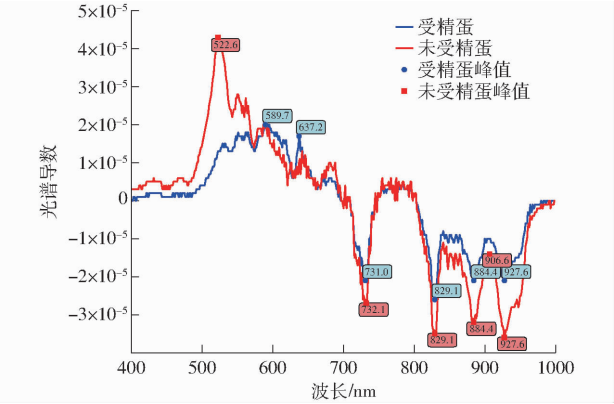


图 7 d0 受精蛋与未受精蛋平均光谱曲线
Fig. 7 Average spectral curves of d0 fertilized and unfertilized eggs

受精蛋与非受精蛋明显的光谱差异峰值有: 522.6、731.0、829.1、884.4、927.6 nm。522.6 nm 附近的光谱峰差异可能与类胡萝卜素水平的变化有关,由于类胡萝卜素在保护发育中的胚胎免受氧化应激方面的作用,类胡萝卜素水平在受精蛋中较高^[26]。731.0 nm 附近的波段与水的 O—H 第 3 次倍频和鸽蛋中的碳水化合物有关。胚胎发育过程消耗碳水化合物、蛋白质和脂肪。因此,两者在 731.0 nm 附近的差异可能是由于与非受精样品相比,受精样品中的碳水化合物、蛋白质和脂肪含量相对较低。829.1 nm 附近的吸收是由 O—H 伸缩振动的第 3 次谐波引起的,与鸽蛋内部的水分含量相关,927.6 nm 附近的强吸收可以与 C—H 伸缩吸收的第 3 次谐波相关联,这可能与此鸽蛋中的碳水化合物含量有关^[27]。一项研究说明,830 ~ 880 nm 范围与近红外水分倍频带和蛋白质振动相关,能够捕捉到与受精率相关的蛋成分差异,包括蛋白质和脂质含量^[26]。

图 8 展示了同一受精蛋 d0 ~ d3 的光谱变化,在 526.8、594.0、729.9、732.1、829.1、884.4、927.6 nm 处有明显差异。在孵化 0 ~ 4 d,胚胎主要处于发育准备阶段。发育前期,胚胎会发生一系列理化指标变化,如含水率、pH 值、蛋白质质量分数、脂肪质量分数等,526 nm 附近的光谱峰差异可能与蛋白质质量分数的变化有关^[27]。Gielen 等^[28]研究表明 585 ~ 610 nm 之间的范围可用作计算血液值的参考值。730 nm 附近的波段与水的 O—H 第 3 次倍频和鸽蛋中的碳水化合物有关。830 nm 附近的吸收可能与 O—H 伸缩吸收的第 3 谐波相关,根据鸽蛋成分,该处的吸收可归因于鸽蛋中的碳水化合物和水

含量。由于胚胎在发育过程中消耗碳水化合物、蛋白质和脂肪,因此随着发育阶段的变化,光谱透射值会发生变化。927 nm 附近的强吸收可能与 C—H 伸缩吸收的第 3 次谐波相关联,这可能与此鸽蛋中的碳水化合物含量有关^[8]。一阶导数预处理光谱揭示了原始光谱中模糊的峰,并确定了受精蛋和非受精蛋之间明显的光谱峰差异。这些观察结果表明,与非受精蛋相比,受精蛋的代谢活动和生化变化导致了不同的光谱特征。

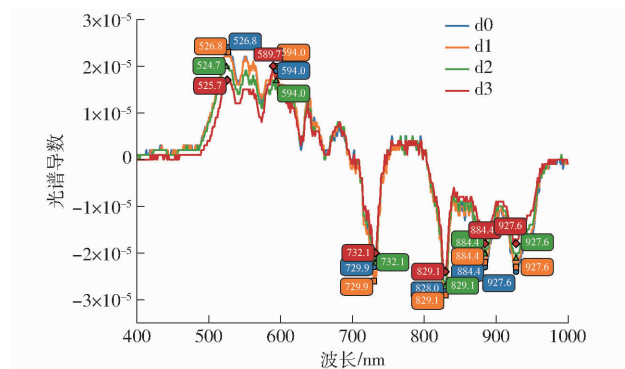


图8 受精蛋平均光谱曲线

Fig.8 Average spectral curves of fertilized eggs

2.2 波段选择分析

在孵化 d0 阶段的对比中,3 种区间选择方法 (BiPLS、iPLS、SiPLS) 均与 LDA 分类算法结合进行了训练和测试,性能参数如表 1 (RAW 表示原始光谱) 所示。结果表明,BiPLS 算法可以得到最佳的受精蛋分类效果,高达 0.98 的召回率。在分析预处理方法对于分类效果的影响时,小波变换 (WT) 带来最优的综合表现:BiPLS-LDA-WT 达到召回率为 0.98、F1 分数为 0.89,明显优于未预处理的分类结果。iPLS-LDA 和 SiPLS 整体弱于 BiPLS。从机理上看,WT 的去噪和平滑提升了光谱差异的可变性,有利于 LDA 的线性判别;MSC/SNV 主要校正散射

表 1 不同特征区间选择模型对比

Tab.1 Comparison of different feature interval selection models

模型	预处理算法	准确率	精确率	召回率	F1 分数
BiPLS - LDA	RAW	0.83	0.81	0.88	0.84
	SNV	0.79	0.79	0.81	0.80
	WT	0.87	0.81	0.98	0.89
	MSC	0.74	0.75	0.77	0.76
SiPLS - LDA	RAW	0.84	0.88	0.76	0.81
	SNV	0.89	1.00	0.76	0.86
	WT	0.87	0.97	0.74	0.84
	MSC	0.89	1.00	0.76	0.86
iPLS - LDA	RAW	0.84	0.81	0.90	0.85
	SNV	0.85	0.81	0.92	0.86
	WT	0.85	0.79	0.96	0.87
	MSC	0.84	0.81	0.90	0.85

与光程差,可以提升精度但可能降低召回率。在孵化早期(d0),受精蛋的准确识别至关重要,因为漏检(受精蛋误判为非受精蛋)会导致孵化率下降,因此后续选用BiPLS算法进行特征波段选择。

特征选择结果显示, BiPLS 算法选取了 401 ~ 458 nm、517 ~ 576 nm 和 637 ~ 696 nm 等多个敏感区段, 光谱信息更为丰富。且 BiPLS 和 iPLS 算法一致地选择了 517 ~ 576 nm 作为受精识别的最优波段区间, 如图 9 所示。

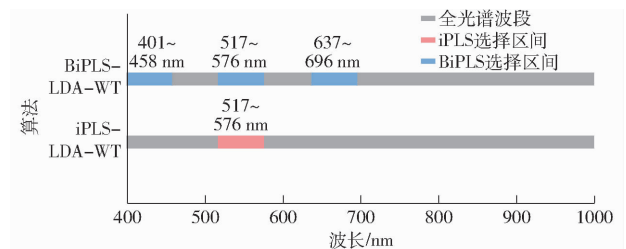


图 9 模型选择光谱区域

Fig. 9 Model selection spectral region

这一结果具有重要的生物学意义:该波段覆盖了文献所报道的 500 ~ 599 nm 关键区域,而该波段被证实在受精蛋血液形成过程中起着决定性作用^[26]。算法选择与生物学机制的高度一致性验证了本研究特征选择策略的科学性和可靠性。

2.3 分类模型识别结果与分析

为验证基于 BiPLS 特征选择的不同分类模型在孵化早期受精蛋判别中的效果,比较了线性判别分析 LDA、LR 和 XGBoost 3 种分类器在不同光谱预处理条件下的表现,结果如表 2 所示。

总体来看,分类性能随孵化时间增加显著提升。在孵化前(d0),受精蛋与未受精蛋光谱差异尚不明显,整体分类准确率在 0.73 ~ 0.87 之间。其中,BiPLS-LDA 结合 WT 预处理表现最佳,准确率为 0.87,召回率为 0.98,F1 分数为 0.89,说明在早期阶段合理的光谱预处理有助于放大微弱差异、提高分类性能。由于孵化前鸽蛋内部的物质变化不明显,因此采取小波去噪(WT)能有效平滑信号并增强特征。

在模型对比方面, BiPLS - LDA 以 F1 分数 0.89 位居首位, BiPLS - LR 以 0.80 次之, 而 BiPLS - XGBoost 为 0.79, 表明线性模型在光谱分类任务中具有独特优势。预处理方法的选择对模型性能影响显著, WT 在 BiPLS - LDA 模型中表现最佳, SNV 在 BiPLS - LR 模型中效果最优, 特别是在 d3 孵化阶段, BiPLS - LDA 结合 WT 预处理实现了完美的分类性能。值得注意的是, 在孵化早期 (d0 ~ d1), 模型普遍呈现高召回率、低精确率的特点, 这有利于避免受精蛋的漏检, 符合实际应用需求。

表 2 不同分类器不同孵化时间的分类效果对比

Tab. 2 Comparison of classification performance of different classifiers with different incubation times																	
模型	预处理 方法	d0				d1				d2				d3			
		A	P	R	F1	A	P	R	F1	A	P	R	F1	A	P	R	F1
BiPLS – LDA	RAW	0.83	0.81	0.88	0.84	0.87	0.84	0.92	0.88	0.83	0.87	0.79	0.83	0.95	0.91	1.00	0.95
	SNV	0.79	0.79	0.81	0.80	0.88	0.84	0.94	0.89	0.84	0.88	0.81	0.84	0.98	0.96	1.00	0.98
	WT	0.87	0.81	0.98	0.89	0.88	0.82	0.98	0.89	0.89	0.84	0.98	0.90	1.00	1.00	1.00	1.00
	MSC	0.74	0.75	0.77	0.76	0.78	0.81	0.75	0.78	0.91	0.92	0.90	0.91	0.97	0.95	1.00	0.97
BiPLS – LR	RAW	0.77	0.74	0.87	0.80	0.73	0.71	0.85	0.77	0.95	0.93	0.98	0.95	1.00	1.00	1.00	1.00
	SNV	0.85	0.80	0.94	0.87	0.89	0.85	0.96	0.90	0.92	0.88	0.98	0.93	0.96	0.93	1.00	0.96
	WT	0.77	0.74	0.87	0.80	0.73	0.71	0.85	0.77	0.95	0.93	0.98	0.95	1.00	1.00	1.00	1.00
	MSC	0.85	0.80	0.94	0.87	0.86	0.82	0.94	0.88	0.93	0.88	1.00	0.94	0.96	0.94	0.98	0.96
BiPLS – XGBoost	RAW	0.78	0.77	0.83	0.80	0.91	0.88	0.96	0.92	0.82	0.78	0.90	0.84	0.98	0.96	1.00	0.98
	SNV	0.73	0.75	0.75	0.75	0.83	0.78	0.94	0.85	0.91	0.89	0.94	0.92	0.96	0.93	1.00	0.96
	WT	0.77	0.75	0.85	0.79	0.92	0.89	0.96	0.93	0.85	0.82	0.90	0.86	0.97	0.95	1.00	0.97
	MSC	0.81	0.81	0.83	0.82	0.82	0.84	0.81	0.82	0.90	0.88	0.94	0.91	0.98	0.96	1.00	0.98

为了更直观地分析模型的分类性能,对图 10 所示的混淆矩阵进行进一步分析。

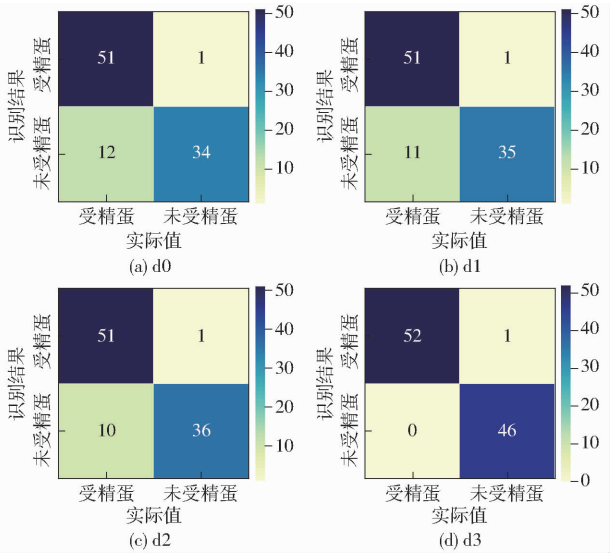


图 10 BiPLS - LDA - WT 混淆矩阵

Fig. 10 BiPLS - LDA - WT confusion matrix

在 d0 效果最好的 BiPLS - LDA - WT 模型中,对受精蛋表现出较高的识别准确性(识别率为 51/52,98.08%),仅有 1 个未受精蛋样本被误分为受精蛋。随着胚胎发育的进行,受精蛋的生理与化学特征变化更加明显,模型的分类准确率进一步提高。未受精蛋的误判数量逐渐减少,表明模型对类别间特征差异的捕捉能力增强,分类边界更加清晰。图 10d 模型实现了完美分类,所有受精蛋与未受精蛋均被正确识别。这表明到第 3 天时,受精蛋与未受精蛋在光谱特征上的差异已非常显著,模型能够充分提取这些差异。

综上所述,BiPLS - LDA 在多数情况下能够较好地区分受精蛋与未受精蛋。但需要注意的是,部分模型仍存在“未受精蛋被预测为受精蛋”的偏差,

这可能与未受精蛋样本量较少、类别分布不均衡有关。在后续研究中,可通过增加样本数据、引入数据平衡策略(如类别权重调整、数据增强等技术)等进一步改善对未受精蛋的识别效果。与现有无损检测方法相比,本研究方法也具有显著优势。传统可见光成像技术主要依赖外观、纹理或色差特征,在胚胎早期由于外观变化不明显,容易出现识别准确率偏低的情况;近红外光谱方法虽然具有较高的化学指示能力,但其单点光谱或少量波段难以充分刻画胚胎发育过程中的复杂光学变化,且模型容易受噪声或局部光照波动影响。相比之下,高光谱具有波段密集、光谱分辨率高、重复性好等特点,能够捕捉受精蛋与未受精蛋在早期阶段形成的微弱透射光学差异。结合特征区间筛选算法(BiPLS),可进一步提取最具判别力的特征波段,避免了全波段冗余信息对模型的干扰,提高了模型稳健性与泛化能力。

2.4 误判样本分析

如图 10 中混淆矩阵所示,孵化前(d0)批次测试共产生误判样本 13 个,误判率为 13.27%。为了深入探究模型产生误判的原因,随机抽取 BiPLS - LDA - WT 模型的 1 个误检(未受精蛋→受精蛋)与 3 个漏检(受精蛋→未受精蛋)样本进行了案例分析。随机抽取基于小波变换(WT)系数曲线提取了一组核心光谱特征。在图 11 的雷达图分析中,将误判样本的特征值与标准正确样本的平均特征值进行对比,从而快速定位导致错判的关键特征偏差。

提取了 5 种核心 WT 系数指标。平均系数(Mean):代表信号的整体强度尺度,如果误判样本的均值大幅偏离其真实类别,意味着该样本的整体化学成分或背景信号异常;峰值系数(Peak):捕捉特征波段最强的信号,峰值偏差往往指向核心目标

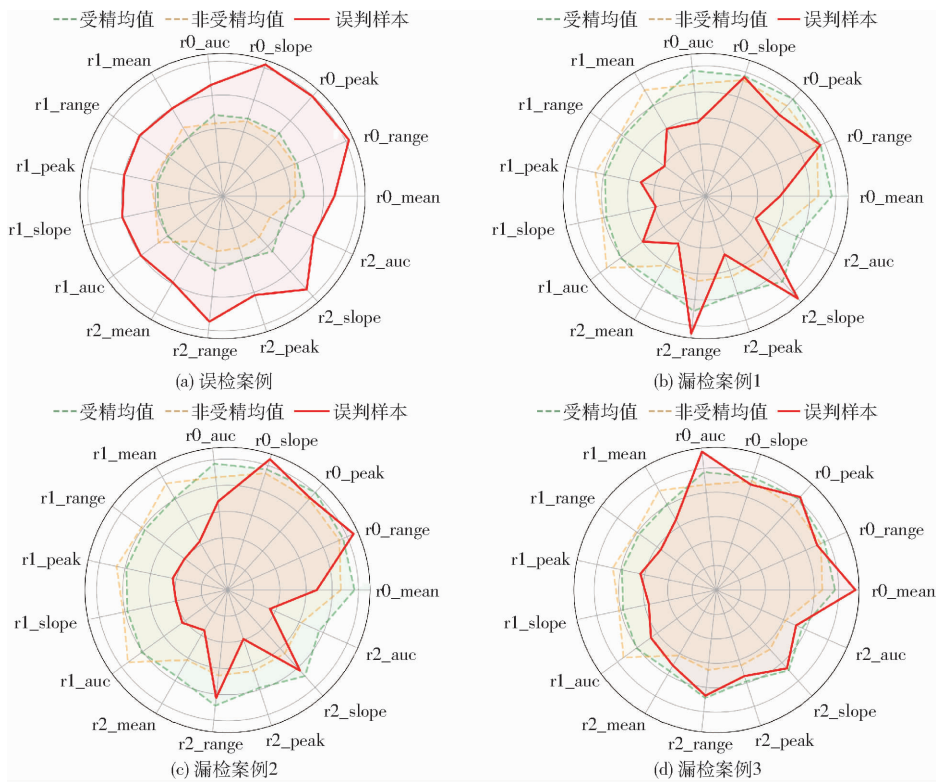


图 11 误判样本雷达图

Fig. 11 Misclassified sample radar chart

物质含量或特征的异常;系数范围(Range):量化信号的波动幅度,范围过小可能意味着该特征不明显或信噪比较低,范围过大则可能意味着异常干扰;WT 系数曲线下面积(AUC):综合强度和宽度,AUC 偏差反映了该波段信号的整体能量规模异常;平均绝对斜率(Slope):描述系数曲线的几何形状变化,斜率偏差通常指示了信号的基线趋势、背景漂移或曲线边缘的异常。r0、r1 和 r2 分别代表选中的 3 个光谱区间(401 ~ 458 nm, 517 ~ 576 nm, 637 ~ 696 nm)。

图 11a 中的受精误判为非受精的样本特征值在 3 个区域显著高于正常受精蛋的平均水平。这表明该受精蛋的信号特征因个体差异、发育不良或者测量偏差而导致其特征向量错误地趋向于非受精蛋的信号。图 11b ~ 11d 中,非受精误判为受精的样本在 r1 区间的特征偏低,而在 r0 和 r2 区域存在不同程度的偏高或偏低。这些样本可能因蛋内物质分布与受精蛋相似,或受表面污染物、蛋壳结构异常等因素影响,使其光学特性错误地模拟了受精蛋的光谱轮廓。

3 结束语

融合高光谱透射成像技术与光谱预处理算法、特征区间选择算法,构建了鸽蛋孵化早期受精情况判别模型,实现了鸽蛋受精情况的快速无损分类。实验结果表明,BiPLS 特征区间筛选算法在孵化前(d0)阶段,表现出优异的特征选择能力。通过对比多种预处理方法,明确小波变换(WT)能够有效消除噪声干扰、提升光谱数据质量。BiPLS-LDA-WT 分类模型在 401 ~ 458 nm、517 ~ 576 nm 和 637 ~ 696 nm 组合波段区间内取得了最佳判别性能(准确率为 0.87,召回率为 0.98, F1 分数为 0.89),验证了早期筛分的可行性。随着孵化时间的增加,受精蛋与未受精蛋的光谱判别能力显著提升,孵化第 3 天(d3)分类效果最优,LDA 与 LR 模型的分类指标均达到 100%。有效的区间筛选为后续窄带化与低成本光谱传感器设计提供依据,并为探索生产线在线检测提供可行性基础。该方法在孵化场早期分选中具备良好应用前景。

参 考 文 献

[1] 刘成康,刘勇,王巧华,等.种蛋受精与性别信息无损检测技术与装备研究进展[J].农业机械学报,2024,55(5):1-20.
Liu Chengkang, Liu Yong, Wang Qiaohua, et al. Progress and prospect in non-destructive assessment technologies and equipment of fertilization and gender determination in hatching eggs[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(5): 1-20. (in Chinese)

[2] Khaliduzzaman A, Kashimori A, Suzuki T, et al. Chick embryo growth modeling using near-infrared sensor and non-linear least

- square fitting of Egg opacity values[J]. *Sensors*,2020,20(20):5888.
- [3] 贾志鑫,杨霖,史策,等.农产品品质在线感知技术应用研究进展[J]. *农业机械学报*,2025,56(6):17–32.
Jia Zhixin, Yang Lin, Shi Ce, et al. Research progress on application of online perception technology for agricultural product quality[J]. *Transactions of the Chinese Society for Agricultural Machinery*,2025,56(6):17–32. (in Chinese)
- [4] 张晓雨,籍颖,锡建中,等.基于光谱技术的种蛋受精早期检测研究[J]. *中国家禽*,2023,45(3):106–112.
Zhang Xiaoyu, Ji Ying, Xi Jianzhong, et al. Study on early fertilization detection of breeding eggs based on spectroscopic technology[J]. *China Poultry*,2023,45(3):106–112. (in Chinese)
- [5] 刘云飞,张晓雨,籍颖,等.基于透射光谱技术的种蛋受精无损检测研究[J]. *中国家禽*,2025,47(6):153–161.
Liu Yunfei, Zhang Xiaoyu, Ji Ying, et al. Non-destructive detection of fertilized chicken eggs based on transmission spectroscopy technology[J]. *China Poultry*,2025,47(6):153–161. (in Chinese)
- [6] Dong J, Dong X, Li Y, et al. Prediction of infertile chicken eggs before hatching by the Naïve-Bayes method combined with visible near infrared transmission spectroscopy[J]. *Spectroscopy Letters*,2020,53(5):327–336.
- [7] Ghaderi M, Mireei S A, Masoumi A, et al. Chick embryo development assessment and fertility detection using pixel-wise hyperspectral image analysis and deep learning[J]. *Poultry Science*,2025,104(12):106064.
- [8] Ghaderi M, Mireei S A, Masoumi A, et al. Fertility detection of unincubated chicken eggs by hyperspectral transmission imaging in the Vis – SWNIR region[J]. *Scientific Reports*,2024,14(1):1289.
- [9] 张伏,崔夏华,张亚坤,等.多位置可见/近红外光谱检测与种鸡蛋受精信息的关系研究[J]. *光谱学与光谱分析*,2021,41(10):3064–3068.
Zhang Fu, Cui Xiahua, Zhang Yakun, et al. Relationship between visible/near infrared spectral data and fertilization information at different positions of hatching eggs[J]. *Spectroscopy and Spectral Analysis*,2021,41(10):3064–3068. (in Chinese)
- [10] Hosseini S A, Banakar A, Minaei S, et al. Determination of brown egg fertility by multi-directional visible/near-infrared spectroscopy and machine learning[J]. *Journal of Applied Poultry Research*,2025,34(2):100534.
- [11] 祝志慧,刘婷,马美湖.基于高光谱信息融合和相关向量机的种蛋无损检测[J]. *农业工程学报*,2015,31(15):285–292.
Zhu Zhihui, Liu Ting, Ma Meihu. Hatching eggs nondestructive detection based on hyperspectral-imaging information and RVM[J]. *Transactions of the CSAE*,2015,31(15):285–292. (in Chinese)
- [12] Wang D, Wang Q, Chen Z, et al. CVAE – DF: a hybrid deep learning framework for fertilization status detection of pre-incubation duck eggs based on VIS/NIR spectroscopy [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*,2024,320:124569.
- [13] 陈灼廷,王巧华,王东桥,等.可见-近红外光谱与联合优化策略的孵前种鸭蛋受精信息无损检测[J]. *光谱学与光谱分析*,2025,45(5):1469–1475.
Chen Zhuoting, Wang Qiaohua, Wang Dongqiao, et al. Non-destructive detection of pre-incubation breeding duck egg fertilization information based on visible/near infrared spectroscopy and joint optimization strategy[J]. *Spectroscopy and Spectral Analysis*,2025,45(5):1469–1475. (in Chinese)
- [14] 李庆旭,王巧华,顾伟,等.基于深度学习的种鸭蛋孵化早期受精信息无损检测[J]. *农业机械学报*,2020,51(1):188–194.
Li Qingxu, Wang Qiaohua, Gu Wei, et al. Non-destructive testing of early fertilization information in duck egg laying based on deep learning[J]. *Transactions of the Chinese Society for Agricultural Machinery*,2020,51(1):188–194. (in Chinese)
- [15] Li Q, Shao Z, Zhou W, et al. MobileOne – YOLO: improving the YOLO v7 network for the detection of unfertilized duck eggs and early duck embryo development-a novel approach[J]. *Computers and Electronics in Agriculture*,2023,214:108316.
- [16] Lin Q, Fang Q, Feng Y, et al. An optimization strategy for detection of fertile pigeon egg based on NIR spectroscopy analysis [J]. *Infrared Physics & Technology*,2023,132:104733.
- [17] Zhang W, Pan L, Tu K, et al. Comparison of spectral and image morphological analysis for egg early hatching property detection based on hyperspectral imaging[J]. *PLoS One*,2014,9(2):e88659.
- [18] Adegbenjo A O, Liu L, Ngadi M O. An adaptive partial least-squares regression approach for classifying chicken egg fertility by hyperspectral imaging[J]. *Sensors*,2024,24(5):1485.
- [19] 祝志慧,叶子凡,杨凯,等.种蛋孵化信息的无损检测研究进展[J]. *华中农业大学学报*,2021,40(4):94–101.
Zhu Zhihui, Ye Zifan, Yang Kai, et al. Progress of nondestructive detection of hatching information of breeding eggs[J]. *Journal of Huazhong Agricultural University*,2021,40(4):94–101. (in Chinese)
- [20] Dong J, Dong X, Li Y, et al. Identification of unfertilized duck eggs before hatching using visible/near infrared transmittance spectroscopy[J]. *Computers and Electronics in Agriculture*,2019,157:471–478.
- [21] 李庆旭,王巧华,顾伟,等.基于卷积神经网络和光谱特征的孵前种鸭蛋受精信息无损检测[J]. *光谱学与光谱分析*,2020,40(12):3847–3853.
Li Qingxu, Wang Qiaohua, Gu Wei, et al. Non-destructive testing of fertilization information of pre-incubation duck eggs based on convolutional neural network and spectral features[J]. *Spectroscopy and Spectral Analysis*,2020,40(12):3847–3853. (in Chinese)

commercial farming environments[J]. Artificial Intelligence in Agriculture, 2026, 16(1): 266 – 283.

[17] Man, Zhang X, Zheng H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C] // Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 116 – 131.

[18] Han K, Wang Y, Tian Q, et al. GhostNet: more features from cheap operations[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020: 1577 – 1586.

[19] Wang Q, Wu B, Zhu P, et al. ECA – Net: efficient channel attention for deep convolutional neural networks[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020: 11531 – 11539.

[20] Khanam R, Hussain M. YOLO 11: an overview of the key architectural enhancements[EB/OL]. (2024 – 10 – 24) [2026 – 03 – 08]. <https://arxiv.org/abs/2410.17725>.

[21] Li C, Li L, Jiang H, et al. YOLO v6: a single-stage object detection framework for industrial applications[EB/OL]. (2022 – 09 – 06) [2026 – 03 – 08]. <https://arxiv.org/abs/2209.02976>.

[22] Tu P, Xie X, Ai G, et al. FemtoDet: an object detection baseline for energy versus performance tradeoffs[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, 2023: 13272 – 13281.

[23] Geffen O, Yitzhaky Y, Barchilon N, et al. A machine vision system to detect and count laying hens in battery cages[J]. Animal, 2020, 14(12): 2628 – 2634.

[24] 赵春江,梁雪文,于合龙,等. 基于改进 YOLO v7 的笼养鸡/蛋自动识别与计数方法[J]. 农业机械学报, 2023,54(7): 300 – 312.

Zhao Chunjiang, Liang Xuewen, Yu Helong, et al. Automatic identification and counting method of caged hens and eggs based on improved YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(7):300 – 312. (in Chinese)

[25] Wu Z, Yang J, Zhang H, et al. Enhanced methodology and experimental research for caged chicken counting based on YOLO v8[J]. Animals, 2025, 15(6): 853.

(上接第 102 页)

[22] 唐徐华,林君,杨海明,等. 不同家禽蛋壳结构及其成分研究[J]. 福建农业学报,2022,37(2):156 – 163.

Tang Xuhua,Lin Jun,Yang Haiming,et al. Shell structure and composition of poultry eggs[J]. Fujian Journal of Agricultural Sciences,2022,37(2):156 – 163. (in Chinese)

[23] 张云伟,张健涛,张海,等. 基于近红外光谱与 Transformer 的烟叶感官指标预测方法[J]. 农业机械学报,2026,57(1): 386 – 396.

Zhang Yunwei, Zhang Jiantao, Zhang Hai, et al. Prediction method of tobacco sensory indicators based on near infrared spectroscopy and transformer[J]. Transactions of the Chinese Society for Agricultural Machinery,2026,57(1):386 – 396. (in Chinese)

[24] Norgaard L,Saudland A,Wagner J,et al. Interval partial least-squares regression(iPLS): a comparative chemometric study with an example from near-infrared spectroscopy[J]. Applied Spectroscopy,2000,54(3):413 – 419.

[25] Leardi R,Nørgaard L. Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions[J]. Journal of Chemometrics,2004,18(11):486 – 497.

[26] Ahmed M W, Sprigler A, Emmert J L, et al. Non-destructive detection of pre-incubated chicken egg fertility using hyperspectral imaging and machine learning[J]. Smart Agricultural Technology,2025,10:100857.

[27] 宋科,潘磊庆,杨崇龙,等. 受精种蛋理化指标与光谱信息的响应关系[J]. 南京农业大学学报,2023,46(6):1187 – 1195.

Song Ke,Pan Leiqing,Yang Chonglong,et al. The response of physical and chemical indicators of fertilized breeder eggs based on spectral information[J]. Journal of Nanjing Agricultural University,2023,46(6):1187 – 1195. (in Chinese)

[28] Gielen R M A M,De Jong L P,Kerkvliet H M M. Electrooptical blood-spot detection in intact eggs[J]. IEEE Transactions on Instrumentation and Measurement,1979,28(3):177 – 183.