

基于CNN框架的LSTM融合优化模型用于芒果干物质的近红外光谱分析

林雪梅¹, 蔡肯², 黄家立¹, 蒙芳秀¹, 林钦永², 陈华舟^{1,3*}

(1. 桂林理工大学 数学与统计学院, 广西 桂林 541004; 2. 仲恺农业工程学院 自动化学院, 广东 广州 510225; 3. 广西高校应用统计重点实验室, 广西 桂林 541004)

摘要: 芒果中的干物质(DM)含量是评判芒果品质的重要指标之一。该文利用近红外光谱法(NIR)检验和预测芒果的干物质含量。主要基于卷积神经网络(CNN)框架, 研究其结构参数网格数值化筛选方案, 融入长短期记忆网络(LSTM)完成参数协同优化, 构建CNN-LSTM融合优化模型。实验过程中, 通过构建浅层CNN建模框架, 针对CNN-LSTM模型的核心参数进行局部规模的超参数联合调试。模型训练和模型测试结果显示, CNN模型和CNN-LSTM模型的最优化预测结果均明显优于常规的线性或非线性模型。该研究除了确定最优模型以外, 还提供了更多可选的模型优化参数组合, 有望在芒果的生产和培育过程中得到应用。浅层CNN框架融合LSTM优化模型及其参数网格数值化筛选方案能够为快速检测芒果果实中的干物质含量提供化学计量学技术支持。

关键词: 近红外(NIR); 芒果干物质; 卷积神经网络(CNN); 长短期记忆网络(LSTM); 参数优选; 网格数值化

中图分类号: O657.33; TB9 **文献标识码:** A **文章编号:** 1004-4957(2025)06-1176-07

Fusion Model of LSTM Optimization Based on CNN Framework and Its Application to NIR Spectroscopic Analysis of Mango Dry Matter

LIN Xue-mei¹, CAI Ken², HUANG Jia-li¹, MENG Fang-xiu¹, LIN Qin-yong², CHEN Hua-zhou^{1,3*}

(1. School of Mathematics and Statistics, Guilin University of Technology, Guilin 541004, China; 2. College of Automation, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China; 3. Guangxi Colleges and Universities Key Laboratory of Applied Statistics, Guilin 541004, China)

Abstract: The content of dry matter(DM) is one of the important indices to determine the quality of mango. In this paper, near-infrared spectroscopy(NIR) is used to predict the dry matter content of mango, so as to achieve rapid evaluation of mango quality. The study launched to propose the grid numericalization scheme for screening structural parameters based on the convolutional neural network(CNN) framework. The parameter optimization strategy was improved by the fusion of long short-term memory(LSTM) network, to propose the CNN-LSTM combined optimization model. In data experiment, a shallow CNN modeling architecture was constructed. The hyperparameters were for refine tuning by testing some local-scale values of the core parameters of CNN-LSTM model. Results showed that the optimal CNN model and CNN-LSTM models were obviously better than the conventional linear or nonlinear models in both the model training and model testing stages. In addition to identifying the most optimal models, we also provided some other appreciating less-optional models as well as their available parameter combinations. These findings are expected to be helpful in the production line of mango cultivation. The modeling framework of a shallow CNN architecture in fusion with the LSTM optimization provides chemometrics technical support for rapid detection of dry matter content in mango fruit.

Key words: near-infrared(NIR) spectroscopy; dry matter of mango fruit; convolutional neural network(CNN); long short-term memory(LSTM); parameter optimization; grid numericalization

芒果是最受欢迎、消费量最大的即食水果之一。芒果果实的品质好坏会影响当季销量, 间接影响

收稿日期: 2025-02-18; 修回日期: 2025-03-19

基金项目: 国家自然科学基金(62365008); 广西自然科学基金(2022GXNSFAA035499)

* 通讯作者: 陈华舟, 教授, 研究方向: 生物信息与化学计量学, E-mail: hzchengut@foxmail.com

区域水果经济。芒果果实中的干物质(DM)是与采摘时的果肉颜色相对应的判断芒果品质的一个重要指标^[1]。在芒果逐渐成熟的过程中,大部分的淀粉被分解成可溶性糖,最终决定芒果果实的甜度和口感。已有文献证实,成熟芒果果实中的干物质含量与可溶性固形物(SSC)含量高度相关^[2]。因此,对于干物质的预测是定量判断芒果果实品质水平的重要手段。

近红外(NIR)光谱技术利用光的辐射能量与水果成分的分子结构进行相互作用(吸收和散射),通过对光强衰减量的计算来测算样本的物质含量^[3],已成功应用于梨、橄榄等多种水果的成分预测,且能够有效测定新鲜水果中的干物质含量^[4-5]。NIR光谱技术对新鲜水果的成分分析需要建立定标模型,并进行模型优化。常规NIR光谱定标分析可以利用线性建模方法实现,例如主成分回归(PCA)^[6]、偏最小二乘法(PLS)^[7]以及它们的改进方法^[8-9]。然而,线性定标模型对于指定训练样本的预测能力表现良好,而对新一批未知样本的预测效果并不能得到保证,原因是线性模型对部分非线性光谱特征的提取不充分,导致模型的预测能力未能达到预期效果^[10]。为了解决这类问题,可以考虑通过不断积累新的样本参与模型训练,使NIR可检测的参考值范围动态扩大^[11],或者采用自适应学习等更高级的数据建模方法^[12]。

利用NIR光谱技术实现数据建模的精准分析仍面临着诸多挑战,开发能够自动提取特征、自适应处理复杂非线性关系、高效处理高维数据的新型化学计量学算法,已成为当前NIR光谱分析领域研究的核心发展方向^[13]。卷积神经网络(CNN)作为目前流行的深度学习框架模型,能够将特征提取、数据建模、迭代优化以及稳健性检验等环节融合一体,并成功应用于食物品种识别和中药药性鉴别^[14],以及农作物的营养成分预测^[15]。而且,浅层网络模式更能提供快速便利的分析决策手段^[16]。然而,常用的浅层CNN是一种前馈神经网络,通常一次性完成特征提取,但对于光谱快速检测数据的分析,一次性作业无法保证特征信号提取的高精确性和稳健性。长短期记忆网络(LSTM)作为一种特殊的循环神经网络,适用于处理具有序列性质的NIR光谱数据,例如,LSTM在烟草成分分析、药物来源识别、香料分类等应用中的模型预测能力均优于传统方法^[17-18]。在此基础上,研究基于CNN框架的LSTM优化模型,能够充分利用LSTM模型的序列特征和CNN处理图像信号的优势,提升模型的数据特征识别能力^[19-20]。

本文将基于CNN框架的LSTM融合优化模型,应用于芒果干物质含量的近红外定量预测,通过卷积操作捕捉光谱数据中的特征信息,并按照LSTM序列分析方式将卷积特征信号进行循环优化处理,旨在增强针对NIR波长空间的特征提取能力。通过构建卷积神经网络结合长短期记忆训练系统(简称CNN-LSTM系统),在多层CNN的结构单元中融入LSTM优化模块,将芒果果实的NIR光谱特征进行多次循环式序列化增强,以达到快速提高模型预测结果的目的,同时获取高信噪比的光谱特征建模变量,简化NIR预测模型,以期为芒果干物质含量的定量分析提供强有力的化学计量学建模技术支持。

1 实验部分

1.1 实验数据

数据来自Mendeley开源数据集(Mango DMC and spectra);该数据集所涉及的芒果样本是来源于澳大利亚境内跨季节、跨地域的主要芒果品种,覆盖面较广^[21]。数据包含了芒果完整果实的果皮表面NIR吸收光谱值及干物质含量值,所提供的光谱数据经F750 Produce Quality Meter光谱仪(Felix Instruments, Camas, USA)配合积分球附件进行测量,测量时的环境温度 $25\text{ }^{\circ}\text{C}$ 恒温,湿度保持在65%RH恒湿。每个样本进行4次重复扫描计算平均数值。原始光谱数据的采集范围为309~1 149 nm;共收集到10个不同品种总计6 050个芒果样本的完整光谱数据。

在上述数据集中,本文关注4个经典的芒果品种:Calypso(Caly),Honey Gold(HG),Kensington Pride(KP)和R2E2,选取814个芒果样本的NIR光谱数据(如图1所示),对应的干物质含量最大值、最小值、平均值、标准偏差分别为23.85%、9.87%、16.32%、2.55%。由图可见,在测量区间两端边缘位置的数据受到的干扰较大,存在较多高频噪声;而且,短波的边缘已经跨越到紫外的频率范围,近红外结合可见光区域数据进行分析较为常见,但结合紫外信号的分析并不常见。综合考虑,选择去除测量区间两端的部分光谱数据,保留480~1 080 nm频率区域的数据进行后续化学计量学建模分析。

1.2 浅层 CNN 模型框架

浅层 CNN 框架通常只包含一个卷积层和一个池化层，随后连接到一个平铺层和一个全连接层进行运算并输出^[22]。模型结构如图 2 所示。

卷积层(CONV)：将备用光谱数据输入卷积层，读取数据特征进行卷积计算。设计采用多个卷积通道(假设为 K 个)进行运算，每个卷积通道的输出结果采用激活函数进行数据变换，随后传输至池化层。

池化层(Pool)：池化计算过程主要关注提取局部信息，通常使用最大池化(Max_pooling)的方式，经过卷积激活的特征数据在池化层中进行二次特征提取。池化操作针对每一个卷积通道单独完成。

平铺层(FLAT)：将 K 个通道的卷积池化过后得到的二次提取特征图谱进行汇总，平铺到同一个数据水平，并同时去除重复的特征值。

全连接层(FC)：将卷积-池化-平铺后得到的特征集合作为输入变量进行模型训练，在 Softmax 中可以选择嵌入线性回归或线性判别模块进行定量、定性计算，随后输出预测结果。

基于上述浅层 CNN 模型框架，融合 LSTM 序列变量训练模块，进一步构建 CNN-LSTM 特征提取及优化模型。

1.3 CNN-LSTM 模型设计

在浅层 CNN 模型框架的基础上引入 LSTM 模块，根据 CNN 学习原理，生成 K 个卷积核计算通道，在每一个通道中完成卷积计算和池化计算；然后使用 LSTM 单元对每一个卷积通道中经过卷积-池化之后的特征变量群进行修正；进而将修正后的特征变量传递至平铺层，随后进入全连接网络层进行训练完成卷积网络的最终预测和输出。CNN-LSTM 融合模型的多通道联合优化方案如图 3 所示。

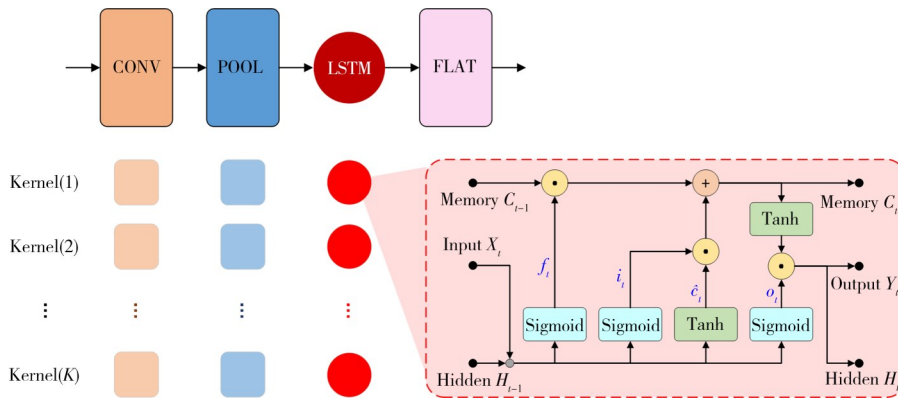


图 3 基于 CNN-LSTM 融合模型的多卷积通道联合优化设计

Fig. 3 Design for combined optimization of the multi-convolutional channels based on CNN-LSTM fusion model

LSTM 循环计算模块包括多个 LSTM 单元(记为 T 个)，每个单元完成一个步次的计算。整体 LSTM 模块通过相邻单元的连接和数据传递来完成循环计算。每个 LSTM 单元的处理过程主要是将数据分解为遗忘门(f_t)、输入门(i_t)、增补单元(\hat{c}_t)和输出门(o_t)的运算，实现对记忆数据(C_t)和隐藏数据(H_t)的更新和循环传递。LSTM 单元接收输入数据，在循环计算开始时(记为 $t = 0$)， X_t 接收的是卷积-池化处理后的特征变量群，当 LSTM 循环处理到第 $t (\neq 0)$ 次时，输入数据 X_t 实际为上一个步骤(第 $t - 1$ 步)处理得到的修正特征变量群。

在每一次循环计算中，LSTM 单元会逐步计算 f_t 、 i_t 、 \hat{c}_t 和 o_t ，即

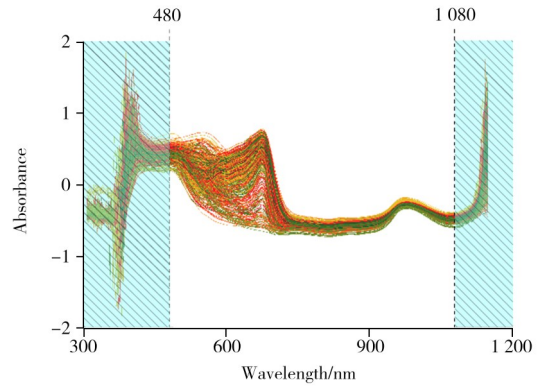


图 1 814 个芒果果实样本的近红外光谱

Fig. 1 Near-infrared spectra of 814 mango fruit samples

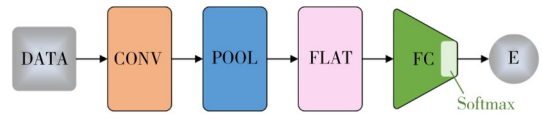


图 2 浅层 CNN 网络结构

Fig. 2 The structure of shallow CNN architecture

$$\begin{cases} f_t = \text{Sigmoid}(X_t \cdot W_f + H_{t-1} \cdot U_f + b_f) \\ i_t = \text{Sigmoid}(X_t \cdot W_i + H_{t-1} \cdot U_i) + b_i \\ \hat{c}_t = \text{Tanh}(X_t \cdot W_c + H_{t-1} \cdot U_c + b_c) \\ o_t = \text{Sigmoid}(X_t \cdot W_o + H_{t-1} \cdot U_o) + b_o \end{cases}$$

其中, W_f 、 U_f 、 W_i 、 U_i 、 W_c 、 U_c 、 W_o 、 U_o 分别为遗忘门、输入门、增补单元、输出门的权重变量, b_f 、 b_i 、 b_c 、 b_o 为相应的偏置向量。随后, 利用增补单元的特征信号来更新记忆数据 C_t , 结合输出门的特征信号进行计算, 获得当前第 t 步计算的输出结果 Y_t 。

$$\begin{aligned} C_t &= f_t * C_{t-1} + i_t * \hat{c}_t \\ Y_t &= o_t * \text{Tanh}(C_t) \end{aligned}$$

同时, 将 Y_t 记录为更新后的隐藏数据(即 $H_t = Y_t$), 以提供给下一次的 LSTM 循环计算进行使用。

当 LSTM 完成最后一次循环计算(当 $t = T$ 时), 最终的输出值 Y_T 将进一步传输至 CNN 网络结构中的平铺层, 完成后续网络计算, 输出模型预测结果。

1.4 模型训练机制和预测指标

按照近红外光谱数据建模及化学计量学分析的基本原理, 模型的训练、优化和测试需要将全体备用样本划分为定标集(Calibration)、检验集(Validation)和测试集(Testing)。定标集样本用于建立并训练模型; 检验集样本用于对模型的结构参数和计算参数进行调试并完成优化设置; 测试集样本通常不参与模型的训练和优化过程, 仅用于对优化后的模型及其参数进行测试, 以确定模型的有效性和预测能力。本研究中, 对全部 814 个的芒果样本进行 CVT 样本集划分, 采用随机抽样的方法选取测试集, 定标集和检验集的划分采用 SPXY 方法^[23]进行。按照常规 2 : 1 : 1 的数量比例, 将定标集、检验集和测试集的样本数量分别确定为 407、203、204。

在模型优化和模型测试的过程中, 采用均方根偏差(RMSE)和 Pearson 相关系数(R)作为模型的量化评价指标, 计算公式如下:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2}, R = \frac{\sum_{i=1}^n (y_i - \bar{y})(y'_i - \bar{y}')}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (y'_i - \bar{y}')^2}}$$

其中, y_i 表示芒果干物质含量的参考值, y'_i 表示 NIR 光谱分析模型计算得到的关于芒果干物质含量的预测值, $i = 1, 2, \dots, n$, n 表示指定样本集中的样本总数量; \bar{y} 和 \bar{y}' 分别为 $\{y_i | i = 1, 2, \dots, n\}$ 和 $\{y'_i | i = 1, 2, \dots, n\}$ 的平均值。特别的, 检验集的预测结果记为 RMSE_V 和 R_V , 测试集的结果记为 RMSE_T 和 R_T 。

2 结果与讨论

利用 CNN-LSTM 模型对芒果的干物质含量进行近红外定量分析。由于 CNN 模型计算过程包含了数据信息转换和特征提取过程, 模型优化主要取决于模型参数的选择。而 CNN 模型参数众多, 即便是浅层模型架构, 仍然涉及多个卷积通道的并行计算。

首先, 对 CNN 模型的结构参数进行设置, 分别对卷积通道数量(K)、卷积核长度(L)、池化元长度(M)3 个参数设置多个不同的取值(如表 1 所示), 构建网格化数值筛选方案, 通过对比不同的网络结构组成, 完成网络架构优化。

针对上述参数设置, 对 CNN 模型进行定标和检验, 考察参数(K, L, M)联合调试的协同模型优化结果(如图 4 所示)。考察了对比模型检验的偏差值, 在芒果干物质含量的近红外定量分析中, CNN 模型的最优参数组合为卷积通道 $K=64$, 卷积长度 $L=10$, 池化元长度 $M=8$, 对应的最优定量预测偏差为 $\text{RMSE}_V=1.057$ 。此外, 由于实际芒果种植和产出过程中受气候、环境、温度等自然条件以及人类生产活动的影响, 最优预测模型在实时检测过程中未必能够获得理想的预测结果。本研究通过模型参数局部大范围筛选, 同时寻找多种其他不同的参数组合(例如图 4 中的深蓝色和紫色区域, 均指向预测偏差 RMSE_V 值较小的模型参数组合), 以确定更多不同结构的 CNN 模型, 为实际生产监测过程的模型在线调试提供更多可能。

表 1 浅层 CNN 网络结构参数调试

Table 1 Settings for parameter tuning of the shallow CNN structure

卷积层(CONV)	卷积通道数量 $K = \{8, 16, 32, 64, 128\}$ 卷积核长度 $L = \{5, 10, 15, 20\}$ 卷积移动步长 $S = 1$
池化层(POOL)	池化元长度 $M = \{2, 4, 6, 8, 10\}$ 池化计算方式选用 Max-pooling
平铺层(FLAT)	采用 DROPOUT 方式去除重复特征变量
全链接层(FC)	输入节点数量等于特征数量 隐藏节点通过自适应筛选 Softmax 单元调用多元线性回归模型

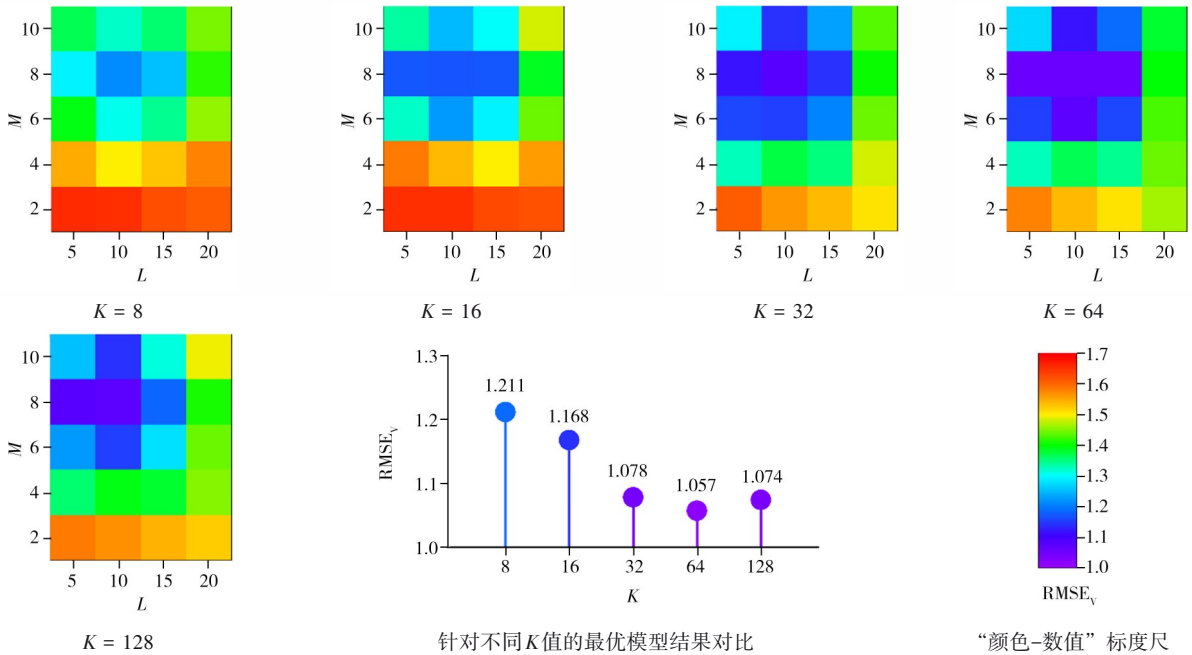


图 4 基于 CNN 框架的模型参数 (K, L, M) 联合调试结果

Fig. 4 Prediction results for tuning of model parameters (K, L, M) based on the CNN framework

其次，对 CNN 框架完成参数联合调试之后，尝试将 LSTM 循环优化模块融合到 CNN 模型架构的每个卷积通道中，建立 CNN-LSTM 模型的参数联合调试机制，实现数据特征信息的精细修正。设置 LSTM 循环计算总步次的可取值 $T = \{5, 10, 15\}$ ，将其联合卷积通道数量 K 进行融合优化，其余 CNN 参数遵循上述网络结构寻优规则；由此获得的 (K, T) 组合参数共有 15 种不同的选择，对应的模型预测结果如图 5 所示，作为对比，未融合 LSTM 模块的浅层 CNN 网络模型的预测结果也在图中显示。结果显示，对于芒果干物质的预测，当 LSTM 的循环总次数选择为 $T = 10$ 时，针对卷积核数量 K 的几种不同取值，均获得该卷积模式下的最佳预测结果，当 $K = 64$ 时，该模型获得最小的预测偏差值 $RMSE_v = 0.769$ ，从而确定最优的 CNN-LSTM 模型参数为 $(K, T) = (64, 10)$ 。该模型对于检验样本和测试样本的预测散点分布如图 6 所示，图中显示，数据结果没有明显异常值，说明模型拟合效果良好，预测结果具有一定的稳定性。此外，从模型效果对比中还可以发现，当 $K = 32$ 和 $K = 64$ 时，CNN-LSTM 模型的预测效果优于浅层 CNN 模型；而当 $K = 16$ 和 $K = 128$ 时，虽然 $T = 15$ 时的结果略逊于 CNN，但选择较少的 LSTM 循环总轮次数可获得更好的结果。模型训练和检验结果表明，适当调整参数之间的联合协同效应，可获得较好的 CNN-LSTM 模型。

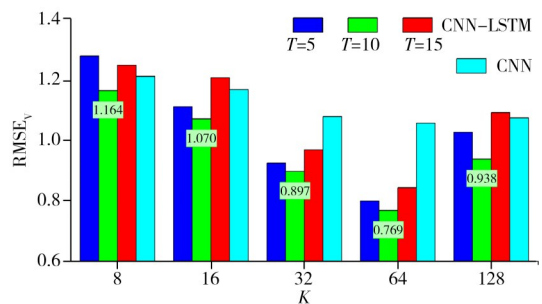


图 5 基于 CNN-LSTM 核心模型参数 (K, T) 的融合优化结果对比

Fig. 5 Comparison of the results by the fusion optimization of the key parameters (K, T) based on the CNN-LSTM model

为了进一步验证 CNN-LSTM 模型的应用有效性, 将 CNN-LSTM 模型、浅层 CNN 模型的预测结果与常规的偏最小二乘回归(PLSR)模型、支持向量机回归(SVR)模型的预测结果进行汇总对比, 其中 PLSR 模型的可调参数 LV 设置为从 1 到 30 的所有整数; SVR 模型采用径向基核函数, 正则化参数 γ 和核变换参数 σ^2 的取值范围均设置为 $\{2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8\}$ 。针对检验集样本和测试集样本分别计算 $RMSE_V$ 、 R_V 和 $RMSE_T$ 、 R_T , 各个模型的最优预测结果如表 2 所示, 与之对应的最优参数也在表中列出。值得注意的是, 测试集样本完全不参与定标和模型检验的训练过程, 其结果可验证模型具有一定的鲁棒性。由表 2 结果可见, 无论是模型训练和检验的结果, 还是模型测试结果, CNN 和 CNN-LSTM 模型的预测结果均明显优于常规的 PLSR 和 SVR 模型, 且 CNN 和 CNN-LSTM 模型将测试相关系数显著提升至 0.9 以上。结果表明, 选择采用浅层 CNN 模型框架对模型的结构参数进行网格数值化调试, 有利于在模型优化方面取得良好的效果; 而融合了 LSTM 模块的 CNN 模型能够进一步提升模型的预测能力。

表 2 CNN 和 CNN-LSTM 模型与传统 PLSR 和 SVR 模型的预测结果对比

Table 2 Comparison on the prediction results from the CNN and CNN-LSTM models against the conventional PLSR and SVR models

Model	Parameter	$RMSE_V$	R_V	$RMSE_T$	R_T
CNN	$K=64, L=10, M=4$	1.057	0.924	1.352	0.907
CNN-LSTM	$K=64, L=15, M=6, T=10$	0.769	0.940	1.084	0.916
PLSR	$LV=14$	2.059	0.871	2.525	0.860
SVR	$\gamma=2^3, \sigma^2=2^5$	1.584	0.913	2.079	0.889

3 结 论

本文构建了基于 CNN 框架的 LSTM 融合优化模型, 应用于芒果干物质的近红外定量预测。首先基于浅层 CNN 网络结构, 进行卷积网络参数网格数值化筛选, 设置卷积网络通道数量(K)可调、卷积核长度(L)可调、池化元长度(M)可调, 通过局部规模数值调整方式, 选择卷积网络模型优化参数组合。其次在每个卷积通道中融入 LSTM 循环计算模块, 同时设置每个模块中的 LSTM 单元数量(即 LSTM 总循环轮次数, T)可调, 并与 CNN 模型参数进行协同调试, 实现参数组合优化。芒果的近红外样本数据来自 Mendeley 开源数据集, 分别建立 CNN 网格化参数筛选以及 CNN-LSTM 融合模型参数寻优。模型训练和测试结果显示, CNN-LSTM 融合模型的预测结果均明显优于常规 PLSR 和 SVR 模型。同时, 网格数值化参数筛选方案还提供了更多可应用的模型参数选择的可能性。结果表明, 所提出的浅层 CNN 框架融合 LSTM 优化模型及其参数网格数值化筛选方案, 能够有效定量分析芒果的干物质含量, 为快速检测芒果果实中的干物质含量提供了化学计量学技术支持; 后续工作可以尝试将该模型优化模式推广到其他水果或农产品的成分检测。

参考文献:

- [1] Ratprakhon K, Neubauer W, Riehn K, Fritsche J, Rohn S. *Foods*, **2020**, 9: 1709.
- [2] de Freitas S T, Guimaraes I T, Vilvert J C, do Amaral M H P, Brecht J K, Marques A T B. *Postharvest Biol. Technol.*, **2022**, 189: 111917.
- [3] Chu X L, Chen P, Li J Y, Liu D, Xu Y P. *J. Instrum. Anal.* (褚小立, 陈瀑, 李敬岩, 刘丹, 许育鹏. 分析测试学报), **2020**, 39(10): 1181-1188.
- [4] Liu Y D, Liao J, Li B, Jiang X G, Zhu M W, Yao J L, Wang Q. *Spectrosc. Spectral Anal.* (刘燕德, 廖军, 李斌, 姜小刚, 朱明旺, 姚金良, 王秋. 光谱学与光谱分析), **2022**, 42(9): 2781-2787.
- [5] Mishra P, Paillart M, Meesters L, Woltering E, Chauhan A. *Postharvest Biol. Technol.*, **2022**, 183: 111739.
- [6] Duarte M S, Pontes M J C, Ramos C S. *J. Appl. Spectrosc.*, **2016**, 82: 1042-1045.
- [7] Ma H L, Wang J W, Chen Y J, Cheng J L, Lai Z T. *Food Chem.*, **2017**, 215: 108-115.

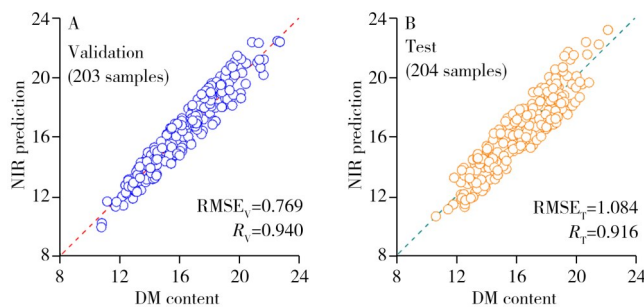


图 6 最优的 CNN-LSTM 模型的预测散点图

Fig. 6 The prediction scatter plots of the optimal CNN-LSTM model

- [8] Zhang H, Tan H N, Lin B R, Yang X C, Sun Z Y, Zhong L, Gao L L, Li L, Dong Q, Nie L, Zang H C. *Molecules*, **2023**, 28: 406.
- [9] Xu Y, Li M, Feng T, Jiao L, Wu F, Zhang T, Tang H, Li H. *Chem. Res. Chin. Univ.*, **2022**, 38: 1057–1064.
- [10] Wang J J, Yang J H, Shao X G. *J. Instrum. Anal.* (王家俊, 杨家红, 邵学广. 分析测试学报), **2020**, 39(10): 1218–1224, 1230.
- [11] Anderson N T, Walsh K B, Flynn J R, Walsh J P. *Postharvest Biol. Technol.*, **2021**, 171: 111358.
- [12] Tosin R, Martins R, Pocas I, Cunha M. *Biosyst. Eng.*, **2022**, 219: 235–258.
- [13] Yang J, Liu Y, Li J Y, Chen P, Xu Y P, Liu D, Chu X L. *Prog. Chem.* (杨健, 刘宇, 李敬岩, 陈瀑, 许育鹏, 刘丹, 褚小立. 化学进展), **2024**, 36(12): 1874–1892.
- [14] Zheng Y, Yang S Y, Wang T, Deng Z W, Lan W J, Yun Y H, Pan L Q. *Chin. J. Anal. Chem.* (郑运, 杨思雨, 王涛, 邓焯文, 兰维杰, 云永欢, 潘磊庆. 分析化学), **2024**, 52(9): 1266–1276.
- [15] Cai K, Chen H, Ai W, Miao X, Lin Q, Feng Q. *IEEE T. Ind. Inform.*, **2022**, 18(2): 1200–1209.
- [16] Acquarelli J, van Laarhoven T, Gerretzen J, Tran T N, Buydens L M C, Marchiori E. *Anal. Chim. Acta*, **2017**, 954: 22–31.
- [17] Zhu Z, Qi G, Lei Y, Jiang D, Mazur N, Liu Y, Wang D, Zhu W. *Chemosensors*, **2022**, 10(5): 164.
- [18] Chen C, Yang B, Si R M, Chen C, Chen F F, Gao R, Li Y Z, Tang J, Lv X Y. *Optik*, **2021**, 242: 167080.
- [19] Shafik W, Tufail A, Liyanage C D S, Apong R A A H M. *J. Sci. Food Agric.*, **2023**, 103: 5849–5861.
- [20] Liao F, Feng X, Li Z, Wang D, Xu C, Chu G, Ma H, Yao Q, Chen S. *J. Integr. Agric.*, **2024**, 23: 711–723.
- [21] Anderson N, Walsh K, Subedi P. *Mango DMC and Spectra, Mendeley Data, V1*, **2020**. doi: 10.17632/46htwnp833.1.
- [22] Chen H, Chen A, Xu L, Xie H, Qiao H, Lin Q, Cai K. *Agric. Water Manag.*, **2020**, 240: 106303.
- [23] Zhang L. *Anal. Methods*, **2019**, 11: 3936–3942.

(责任编辑: 盛文彦)