中国光学 (中英文) Chinese Optics 核心期刊

激光多普勒测振技术无损检测果冻橙粒化病

刘智 赖庆荣 张天禹 李斌 宋云峰 陈楠

Non-destruction detection of jelly orange granulation disease using laser Doppler vibrometry

LIU Zhi, LAI Qing-rong, ZHANG Tian-yu, LI Bin, SONG Yun-feng, CHEN Nan

引用本文:
刘智, 赖庆荣, 张天禹, 李斌, 宋云峰, 陈楠. 激光多普勒测振技术无损检测果冻橙粒化病[J]. 中国光学, 2025, 18(1): 160–169. doi: 10.37188/CO.2024–0115
LIU Zhi, LAI Qing-rong, ZHANG Tian-yu, LI Bin, SONG Yun-feng, CHEN Nan. Non-destruction detection of jelly orange granulation disease using laser Doppler vibrometry[J]. *Chinese Optics*, 2025, 18(1): 160-169. doi: 10.37188/CO.2024-0115

在线阅读 View online: https://doi.org/10.37188/CO.2024–0115

文章编号　2097-1842(2025)01-0160-10

激光多普勒测振技术无损检测果冻橙粒化病

刘　智[1]，赖庆荣[1]，张天禹[1]，李　斌[1,2]，宋云峰[3]，陈　楠[1,2*]
（1. 华东交通大学 机电与车辆工程学院，江西 南昌 330000；
2. 华东交通大学 水果智能光电检测技术与装备国家地方联合工程研究中心，江西 南昌 330000；
3. 宁波舜宇仪器有限公司，浙江 余姚 315400）

摘要：粒化是柑橘类水果常见的一种内部病害，患有该病害的水果外部特征并不明显，难以从外观上直接识别出来。本文使用微型激光多普勒测振仪(micro-LDV)和共振喇叭搭建了一套声学振动实验装置，将其用于采集"爱媛 38 号"果冻橙的振动响应信号。然后，将一维的振动响应信号转换为振动多域图像，并构建了一个 Resnet-Transformer(ResT)网络，用于提取振动多域图像中的深层特征，以识别果冻橙粒化病。本文中，使用振动多域图像分别训练 ResT、Resnet50 和 Vision Transformer(ViT)模型，并将 ResT 的性能与 Resnet50 和 ViT 进行比较。最后，使用振动多域图像纹理特征或振动频谱特征训练偏最小二乘判别分析(PLS-DA)和支持向量机(SVM)模型，并与 ResT 模型进行性能对比。结果表明，使用振动多域图像训练的 ResT 模型可以精准识别果冻橙粒化病并且检测准确率为 98.61%，模型的 F1 为 0.986、精确率为 0.986、召回率为 0.986。由上述结果可知，提出的方法可在简单、快速、低成本的前提下准确识别粒化果冻橙。

关　键　词：激光多普勒测振；声学振动；柑橘粒化病；无损检测；振动多域图像
中图分类号：TN247　　**文献标志码**：A　　**doi**: 10.37188/CO.2024-0115　　**CSTR**: 32171.14.CO.2024-0115

Non-destruction detection of jelly orange granulation disease using laser Doppler vibrometry

LIU Zhi[1], LAI Qing-rong[1], ZHANG Tian-yu[1], LI Bin[1,2], SONG Yun-feng[3], CHEN Nan[1,2*]
(1. *School of Mechatronics & Vehicle Engineering*, *East China Jiaotong University*, *Nanchang* 330000, *China*;
2. *National and Local Joint Engineering Research Center of Fruit Intelligent Photoelectric Detection Technology and Equipment*, *East China Jiaotong University*, *Nanchang* 330000, *China*;
3. *Ningbo Sunny Instrument Co.*, *Ltd*, *Yuyao* 315400, *China*)
* *Corresponding author*, *E-mail*: *chennan@enjtu.edu.cn*

Abstract: Granulation is a common internal disease of citrus fruits, and it is difficult to identify the fruits with this disease from their external features. In this study, an acoustic vibration experimental setup was constructed using a micro-laser Doppler vibrometer (micro-LDV) and a resonance speaker. This was used to collect vibration response signals of 'Aiyuan 38' jelly orange. The one-dimensional vibration response signals

收稿日期：2024-06-21；**修订日期**：2024-07-23
基金项目：国家重点研发计划(No. 2022YFD2001804, No. 2023YFD2001301)；国家自然科学基金(No. 12304447)
Supported by National Key Research and Development Program of China (No. 2022YFD2001804, No. 2023YFD2001301); National Natural Science Foundation of China (No. 12304447)

were converted into vibration multi-domain images, and a Resnet-Transformer network (ResT) was constructed to extract deeper features from the vibration multi-domain images for identifying granulation disease in jelly oranges. In this paper, the ResT, Resnet50, and Vision Transformer (ViT) models were trained using vibration multi-domain images, and their performances were compared. Then, partial least squares discriminant analysis (PLS-DA) and support vector machine (SVM) models were trained using vibration multi-domain image texture features or vibration spectrum features, and the performance was compared with the ResT model. The results show that the ResT model trained using vibration multi-domain images can achieve accurate identification of jelly orange granulation disease with detection accuracy of 98.61%, model F1 of 0.986, precision of 0.986, and recall of 0.986. The proposed method can accurately identify granulated jelly oranges with simplicity, fast speed, and low cost.

Key words: laser Doppler vibrometry; acoustic vibration; citrus granulation disease; non-destructive detection; vibration multi-domain image

1　引　言

“爱媛 38 号”果冻橙是一种由日本培育的优良杂交柑橘品种,具有皮薄、果肉嫩、口感好、早熟和丰产的特点。目前,这一品种已在中国南方地区广泛种植,并深受消费者喜爱。然而,柑橘粒化问题严重影响了果实的品质和食用价值。粒化病会导致柑橘果实的可溶性固形物、总糖和柠檬酸含量降低,颗粒状部分的细胞壁变硬、变厚,汁水含量减少。粒化在柑橘果实生长的早期和晚期容易发生,但从外观上难以区分正常水果和粒化水果。因此,为了提升消费者满意度,防止粒化的果冻橙流入市场,对粒化果进行无损检测显得尤为重要。目前,研究者们已开发了多种检测方法,包括近红外光谱技术[1-4]、高光谱成像技术[5]、X 射线成像技术[6]。尽管这些方法在一定程度上可以检测粒化问题,但成本和操作复杂性方面存在一定的局限性。声学振动技术由于其简单、经济和快速的特点,越来越受到关注。在已有的研究中,声学振动技术已经成功应用于梨的黑心病检测[7-8]、苹果的霉心病检测[9]、蛋壳裂纹检测[10-11]和水果硬度的检测[12-13]。然而,针对柑橘类水果粒化病的研究仍然较少。对于将声学振动技术用于检测柑橘粒化病展开研究,有助于提升柑橘果实的品质控制和市场竞争力。

据报道,声振水果无损检测的效果与振动特征参数选择和建模方法密切相关。过去的研究中,研究者们常使用共振频率等参数[14-15]来预测水果品质或识别缺陷,但这些参数往往受到水果大小、重量和形状的影响,导致信息不足、结果不够准确。为确保声振法水果质量检测的精度,选择合适的声振激励方式和声振特征提取方法显得至关重要。通过调研可知一维振动信号可以通过各种信号处理方法灵活地转换为时频多域图像[16]。时频多域图像可以充分反映声振信号的能量分布和时频特征[17]。在声振信号的多域图像中挖掘潜在的有用信息,并充分利用深度学习算法来识别水果内部缺陷,是非常有潜力和有价值的。在图像识别深度学习算法架构中,卷积神经网络(CNN)和 Transformer 作为两大主流架构,在各个领域中都有广泛应用。CNN 可以自动挖掘隐藏在图像中的局部深度特征[18],而 Transformer 可以自动挖掘图像中的全局深度特征[19],两者结合往往可以获得更好的性能。Liu 等人[20]结合 CNN 和 Transformer 对刀具磨损进行估计,并取得了较单一使用 CNN 或 Transformer 更好的结果。Alshammari 等人[21]使用 Vision Transformer 和 CNN 相结合,用于检测和分类橄榄叶疾病,并证实了两者相结合的结果优于单一使用 CNN 或 Vision Transformer 的结果。

通常水果声振检测装置主要包含 3 个部分:激励装置、检测传感器和信号分析系统。为了实现快速且无损的检测,研究者们尝试多种不同的振动激励装置和信号采集设备。例如,Zhao 等人[22]使用压电换能器激励并采集苹果的振动响应信号。Kadowaki 等人[7]使用振动器激励梨产生振动,以检测梨的黑心病。Zhang 等人[23]使用激振器激励梨,让其产生振动并用激光多普勒测振仪(LDV)采集振动信号。Ding 等人[24]使用扬声器正弦扫频信号激励西瓜振动,并用 LDV 采集西

瓜的响应信号。LDV 可以实现非接触式的振动测量，并在灵敏度、精度、响应速度方面具有显著优势，但其体积大和成本高，从而限制了其使用范围[25]。值得一提的是，最近的一篇综述[26]指出，随着集成光学芯片的发展，LDV 将实现微型化，这会使其成本大幅降低，在可预见的未来，基于 micro-LDV 的水果质量检测设备有望被开发出来。

本文采用一种基于 micro-LDV 和共振喇叭的声学振动装置用于采集果冻橙的振动响应信号。此外，分析了正常果冻橙和粒化果冻橙的振动频谱的差异，并将振动响应信号转换为振动多域图像。在此基础上，构建了一个 Resnet-Transformer 网络（ResT），用于提取振动多域图像中的深层特征并识别果冻橙粒化病。

2 材料与方法

2.1 果冻橙样本

本研究使用的果冻橙品种为“爱媛 38 号”，于 2023 年 10 月 29 号在中国江西省吉安市泰和县的果园自行采摘。将采摘的 720 个果冻橙放置于实验环境中，每天采集 360 个果冻橙样本的振动信号。每个样本的振动信号采集完后，立即沿样本赤道切开并观察样本的粒化情况，记录果冻橙样本是否发生粒化。果冻橙内部出现粒化病变即为粒化果实，未出现则为正常果实。发生粒化的果冻橙样品会出现果皮与果肉分离、节段变厚、变硬、变白、果汁减少等症状，具有明显的病害特征，如图 1 所示。

图 1 正常和粒化的果冻橙横切面

Fig. 1 Cross sections of normal and granulated jelly orange

2.2 实验装置与振动信号的采集

使用自行设计的声振检测装置采集果冻橙的振动信号，如图 2 所示。该装置主要包括一个共振喇叭、一个 micro-LDV(LV-M200, Sunnyinnova, China)。该装置将共振喇叭作为激励源，并使用缓冲泡沫和吸音泡沫材料有效地支撑和固定共振喇叭，以避免外界振动的影响。当水果与共振喇叭的共振盘接触时，水果将直接作为信号产生的介质，发出包含水果物理特性的声音和振动信号。使用 Pearson 相关性分析和测量变异性分析对不同程度的输出增益和输出时间下的振动频谱进行分析，以确定最佳输出增益和输出时间。装置经过优化后的输出增益为 0.12，输出扫频信号时间为 1 s。采集水果振动响应信号时，将果冻橙放置于共振喇叭的共振盘上，由电脑输出 100~1 500 Hz 的正弦扫频信号发送至共振喇叭，激励果冻橙产生振动。随后，使用 micro-LDV 检测水果表面的振动响应信号 X_V，振动信号由采集卡(INV3060V, COINV, China) 传输至电脑进行记录，采集卡的采样频率设置为 8 192 Hz。最终采集了 477 个正常果冻橙和 240 个粒化果冻橙的振动响应信号。将这 717 个样本按 3:1:1 的比例随机划分为训练集、验证集和测试集。

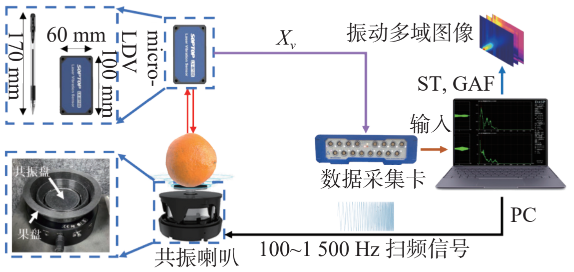


图 2 声振检测装置示意图

Fig. 2 Schematic diagram of acoustic vibration detection device

2.3 振动多域图像的生成

由于 Stockwell 变换 (ST) 可以从时域信号中获得时间与频域的关系，并在时域和频域上都有较好的分辨率。本文使用 ST 生成振动多域图像，如图 3（彩图见期刊电子版）所示。为了从振动响应信号 X_V 中充分提取时频信息，使用 ST 将振动响应信号 (图 3(b)) 转换为时频域图像。计算公式如下：

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t)\, w(t-\tau, f) \exp(-\mathrm{i}2\pi f t)\,\mathrm{d}t \quad , \tag{1}$$

$$w(t-\tau,f)=\frac{|f|}{\sqrt{2\pi}}\exp\left(\frac{-(t-\tau)^2f^2}{2}\right)\quad,\qquad (2)$$

其中, $w(t-\tau,f)$为高斯窗函数, $h(t)$为输入信号, τ为平移因子, 用于控制高斯窗函数在时间轴t上的位置, f为频率, i为虚数单位。

在生成振动时频域图像的过程中 (图 3(c)), 首先使用 S 变换将振动响应信号转换成大小为 224×8 192 的中间时频数据矩阵, 然后使用卷积块 Conv1 对中间时频数据矩阵降维, 得到 224×224 大小的振动 ST 时频域图像 (V-ST), 振动 ST 时频域图像中的数据是包含幅度和相位信息的复矩阵, 行是时间, 列是频率, 其可视化结果如图 3(d) 所示。

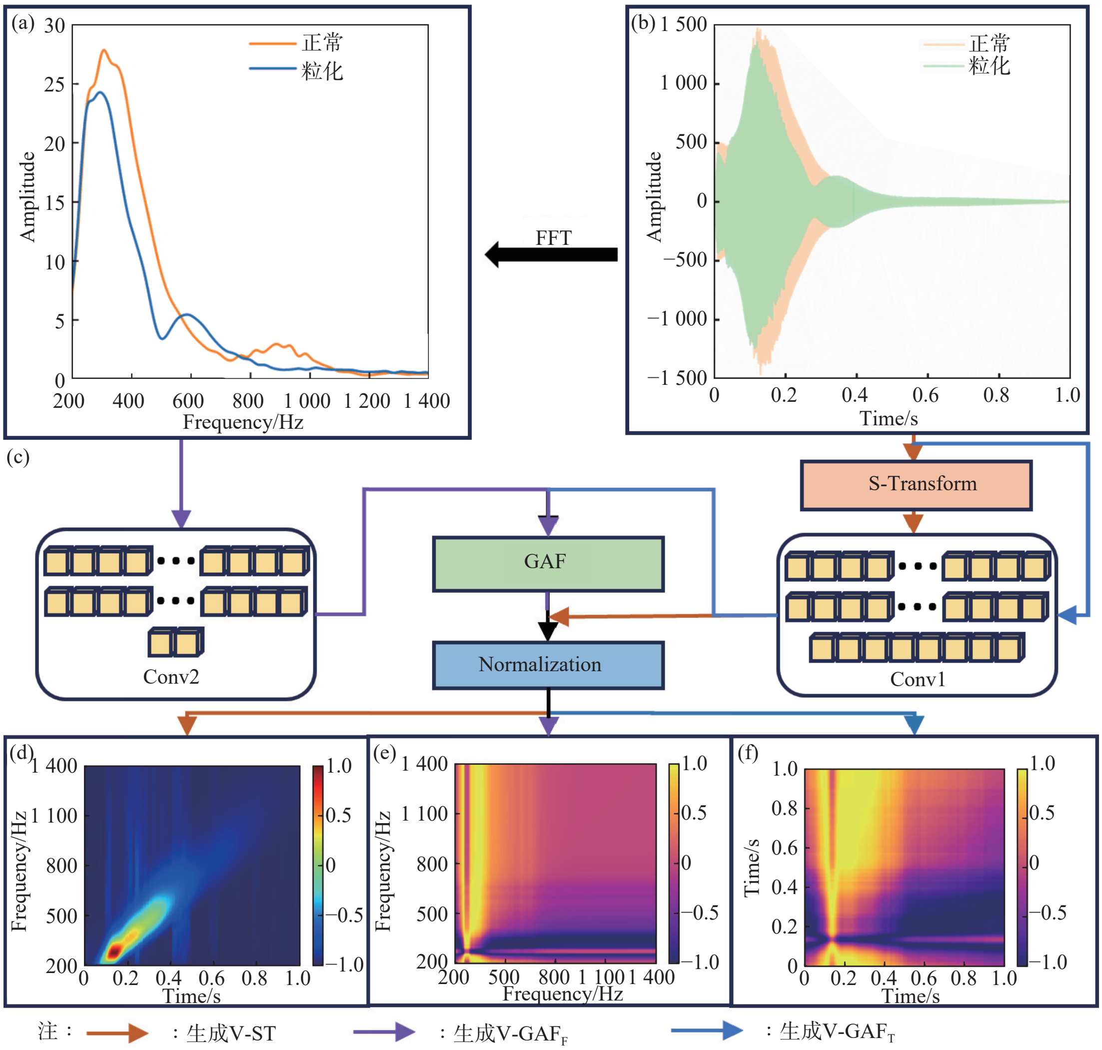


图 3　振动多域图像生成流程图。(a) 经过快速傅立叶变换得到的果冻橙振动频谱曲线; (b) 果冻橙振动响应信号; (c) 振动多域图像生成模块, 其包含对时频信号降维的卷积块、生成时频图像的 Stockwell transform(ST)块、将时频域信号生成时域或频域图像的格拉姆角场 (GAF) 块以及归一化块; (d) 振动 ST 时频域图像可视化结果; (e) 振动 GAF 频域图像可视化结果; (f) 振动 GAF 时域图像可视化结果

Fig. 3　Flow chart of vibration multi-domain image generation. (a) Vibration spectrum curves of jelly orange obtained by fast Fourier transform; (b) vibration response signal of jelly orange; (c) vibration multi-domain image generation module which contains the convolution block for downscaling the time-frequency signal, the Stockwell transform (ST) block for generating the time-frequency image, the Gramian Angle Field (GAF) block for generating from the time-frequency signal into the time-domain or frequency-domain images, and the normalization block; (d) the result of visualizing the vibration ST time-frequency domain image; (e) the visualizing vibration GAF frequency domain image; (f) the visualizing vibration GAF time-domain image

为了进一步利用果冻橙的时域和频域信号中的时频信息, 使用格拉姆角场 (GAF) 对一维时频信号进行编码, 生成 GAF 时域和频域图像。格拉姆角差场 (GADF) 计算公式如下:

$$GADF = \begin{bmatrix} \sin(\varphi_1-\varphi_1) & \cdots & \sin(\varphi_1-\varphi_n) \\ \sin(\varphi_2-\varphi_1) & \cdots & \sin(\varphi_2-\varphi_n) \\ \vdots & \ddots & \vdots \\ \sin(\varphi_n-\varphi_1) & \cdots & \sin(\varphi_n-\varphi_n) \end{bmatrix}, \tag{3}$$

其中，$\varphi_i=\arccos(\tilde{x}_i), -1\leqslant\tilde{x}_i\leqslant1, \tilde{x}_i\in X$。$\varphi_i$是时间序列中第$i$个角度值，$\tilde{x}_i$是时间序列$X$中的第$i$个归一化后的数据点。

果冻橙的振动时域和频域图像生成过程如图 3(c) 所示。使用卷积块 Conv1 将振动响应信号 (图 3(b)) 降维到大小为 1×224 的时域信号，使用 GADF 将时域信号编码成 224×224 振动 GAF 时域图像 (V-GAF$_T$)，可视化结果如图 3(f) 所示。振动响应信号经过快速傅立叶变换转换成振动频域信号 (图 3(a))，再经过卷积块 Conv2 降维成 1×224 的频域信号后，使用 GADF 将频域信号编码成 224×224 振动 GAF 频域图像 (V-GAF$_F$)，可视化结果如图 3(e) 所示。生成的振动多域图像的大小为 224×224×3，每组图像都包含 ST 时频图像、GAF 时域图像和 GAF 频域图像。

2.4 一维数据处理

2.4.1 振动多域图像纹理特征提取

得到振动多域图像后，使用灰度梯度共生矩阵 (GLGCM) 统计振动多域图像中的不同梯度值之间的空间分布特征，进而提取图像的梯度纹理特征。每张图像计算 15 个纹理特征，从振动多域图像中提取了 45 个纹理特征。

2.4.2 振动频谱特征的选择

振动频谱中的大部分频率分量与果冻橙是否粒化的相关性不高，并且无效的频率分量会增加模型的计算量。为了加快模型的预测速度并减少无效频率的输入，需要对有效频率进行选择。在本研究中，采用竞争自适应重加权采样 (CARS) 提取有效频率。CARS 算法首次采用全部变量进行建模，在迭代过程中逐次减少确定的变量。根据设置的循环迭代次数进行循环计算，选择 PLS 模型最小的 RMSECV 子集为最佳的变量集合。在本研究中，通过 5 折交叉验证来确定变量的数量，蒙特卡罗采样次数设置为 50 次。

2.5 识别果冻橙粒化病的建模方法

2.5.1 Resnet-Transformer 分类模型的构建

在这项研究中，建立了一个 Resnet-Transformer 分类模型 (ResT) 来识别果冻橙是否发生粒化，如图 4 所示。CNN 的优势主要表现在局部特征提取，而在全局特征提取上具有局限性。Transformer 则可以关注全局信息，将 CNN 与 Transformer 相结合可以充分提取振动多域图像中的全局时频信息，从而提高模型的识别准确率。

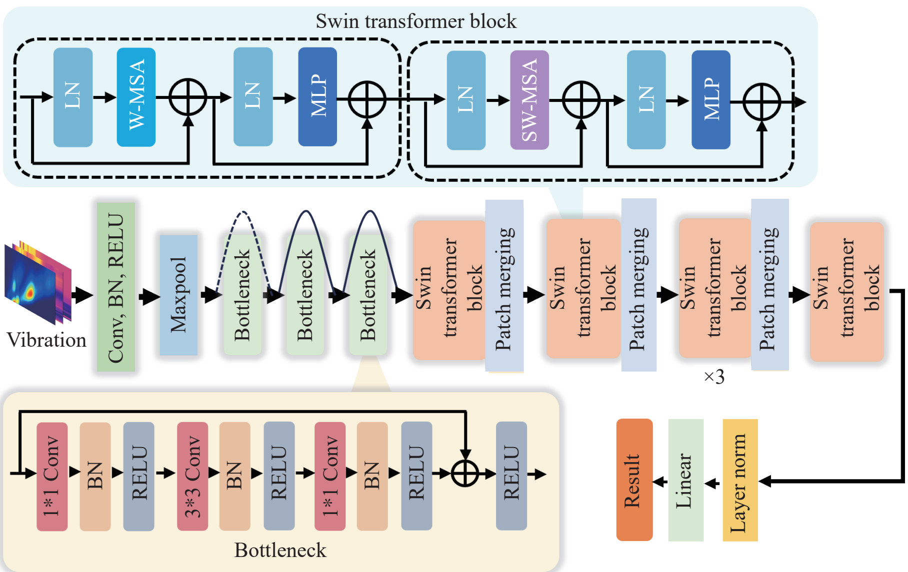


图 4 ResT 模型结构示意图。ResT 主要由多个基于 CNN 的 Bottleneck 块和基于 Transformer 的 Swin Transformer 块组成，分别提取振动多域图像的局部和全局的深层信息

Fig. 4 Schematic diagram of ResT model structure. ResT mainly consists of multiple CNN-based Bottleneck blocks and Transformer-based Swin Transformer blocks, respectively, which extract local and global deep information from vibration multi-domain images

首先, 将振动多域图像作为 ResT 的输入, 经过 Conv、Maxpool 以及 Bottleneck 层进行初步的特征提取, 将大小为 224×224×3 的输入图像转换成大小为 56×56×96 的中间特征矩阵。其中 Bottleneck 为 Resnet 网络中带有残差连接的特征提取模块。残差连接可以有效解决网络性能的退化问题。随后, 将特征输入到基于 Transformer 的 Swin Transformer Block 中进行全局特征提取, 经过 1 层 3 头、1 层 6 头、3 层 12 头以及 1 层 24 头的 Swin Transformer Block 后, 特征矩阵大小变为 49×768。随后经过归一化层后, 将特征矩阵沿列方向上取均值, 变成大小为 1×768 的特征矩阵, 最后经过神经元个数为 2 的全连接层输出结果。在训练 ResT 模型、Resnet50 模型和 Vision Transformer 模型 (ViT) 时, 使用自适应矩估计 (Adam) 优化器进行训练, 学习率为 0.0001。

2.5.2　PLS-DA 和 SVM 分类模型的构建

为了更客观地评价 ResT 的性能, 还将其与 PLS-DA 以及 SVM 模型进行比较。由于振动多域图像不属于一维数据, 不能作为 PLS-DA 以及 SVM 的输入, 因此使用 GLGCM 从振动多域图像中提取 45 个图像纹理特征作为 PLS-DA 和 SVM 的输入。此外, 还使用 CARS 从振动频谱中选择有效频率变量训练 PLS-DA 和 SVM 模型。在训练 PLS-DA 模型的过程中, 使用 5 折交叉验证方法, 并根据验证集中最大的准确率来确定最佳成分数量。在训练 SVM 模型的过程中, 使用网格搜索来确定最佳的核函数、惩罚系数、多项式次数以及 gamma 值。

2.6　分类模型的评价指标

使用准确率、精确率、F1 和召回率这四个指标评价 ResT、Resnet50、Vision Transformer、SVM 和 PLS-DA 的分类性能。这些指标的定义如下:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad , \tag{4}$$

$$Precision = \frac{TP}{TP+FP} \quad , \tag{5}$$

$$Recall = \frac{TP}{TP+FN} \quad , \tag{6}$$

$$F1 = 2\times\frac{Precision\times Recall}{Precision+Recall} \quad . \tag{7}$$

上述指标中, TP 是正确分类的粒化果冻橙的数量。TN 是正确分类的健康果冻橙的数量。FP 是健康果冻橙被错误地归类为粒化果冻橙的数量。FN 是被错误地归类为健康的粒化果冻橙的数量。本研究的所有算法都是在 Python3.7 环境下运行的。计算机的配置为 Intel(R) Xeon (R) Platinum 8338C CPU @ 2.60 GHz RTX 3090 (24 GB)。

3　结果与讨论

3.1　振动频谱曲线分析

图 3(a) 显示了粒化果冻橙和正常果冻橙的振动频谱响应曲线, 可以看出这两种样本之间的振动频谱曲线存在明显的差异。正常果冻橙的第二共振频率在 500~700 Hz 之间, 而粒化果冻橙的第二共振频率在 800~1000 Hz 之间。相较于正常果冻橙的振动频谱曲线, 粒化果冻橙的振动频谱曲线在 500~1000 Hz 范围内的能量会偏左。考虑到共振频率的峰宽较大, 且共振频率受水果大小、重量和形状影响, 基于共振频率参数的建模方法无法进行有效区分。同一水果的不同摆放位置下的振动频谱如图 5(彩图见期刊电子版)所示, 可以看到不同摆放位置的振动频谱有较大的差异, 因此不同的摆放位置对基于共振频率的检测方式会有较大的影响。这也表明仅仅通过频谱共振频率这类简单的识别方式还不足以对粒化病进行精确识别。图 6(彩图见期刊电子版)是振动频谱中 200~1400 Hz 频率成分相关性分析的可视化结果。如图 6 所示, 振动频谱的 400~1400 Hz 范围的频率成分相关性较高, 特别是 400~800 Hz 和 1000~1400 Hz 之间的频率相关性大于 0.75。这说明振动频谱中存在大量相关性较高的频率成分, 因此需要使用 CARS 算法从振动频谱中挑选有效频率, 去除冗余信息, 以提高 PLS-DA 和 SVM 的分类性能。

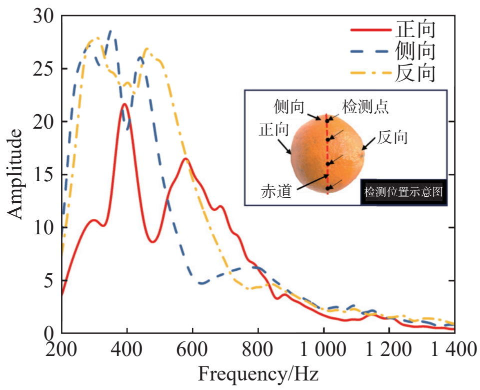


图 5　果冻橙不同摆放位置的振动频谱

Fig. 5　Vibration spectra of jelly orange at different placement positions

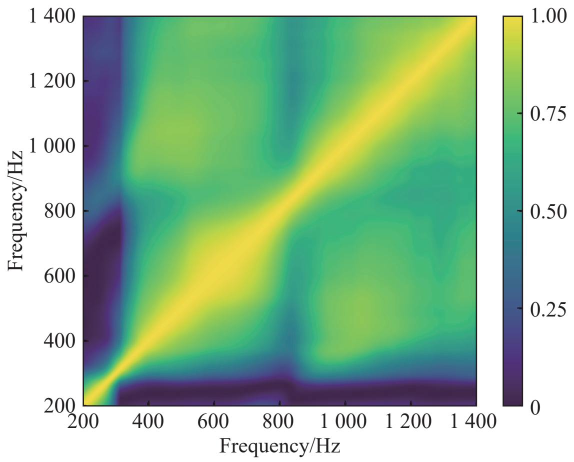


图 6　717 条振动频谱曲线的相关性分析

Fig. 6　Correlation analysis of 717 vibration spectrum curves

3.2　基于一维特征的 SVM 和 PLS-DA 分类模型

3.2.1　基于振动多域图像纹理特征建模分析

为了衡量基于振动多域图像的 ResT 的性能，建立了基于振动多域图像纹理特征 (VMIT) 的 SVM 和 PLS-DA 模型。对于 SVM 模型，使用网格搜索算法对 SVM 的惩罚系数‘*c*’, Gamma 值‘*g*’以及‘*p*’进行优化。当模型参数确定为 c=0.1, p=7, 以及核函数为‘Poly’时具有相对较高的识别性能。使用振动多域图像纹理特征训练 SVM 和 PLS-DA 模型，结果如表 1 所示。从表中可以看出，使用振动多域图像纹理特征训练的 SVM 模型的总体识别准确率比 PLS-DA 模型更高，其中 PLS-DA 对正常果冻橙的识别准确率为 92.08%，对粒化果冻橙的识别准确率为 74.42%。SVM 对正常果冻橙的识别准确率为 95.04%，对粒化果冻橙的识别准确率为 95.35%。在总体准确率上，SVM 的结果比 PLS-DA 高 8.32%，并且只将 2 个粒化果冻橙被识别错误。结果表明，基于振动多域图像纹理特征的 SVM 模型在果冻橙粒化识别方面性能良好。

表 1　基于振动多域图像纹理特征的 PLS-DA 和 SVM 模型对果冻橙粒化病的分类结果

Tab. 1　Classification results of jelly orange granulation disease by PLS-DA and SVM models based on vibrational multi-domain image texture features

模型	实际类别	预测类别		类别准确率 (%)	总体准确率(%)
		正常	粒化		
SVM	正常	96	5	95.04%	95.13%
	粒化	2	41	95.35%	
PLS-DA	正常	93	8	92.08%	86.81%
	粒化	11	32	74.42%	

3.2.2　基于振动频谱特征建模分析

由 3.1 节可知，振动频谱中存在大量相关性较高的频率成分，因此使用 CRAS 算法从振动频谱中提取有效频率成分。如图 7 所示，从振动频谱中选取了 26 个有效频率进行果冻橙粒化病识别。

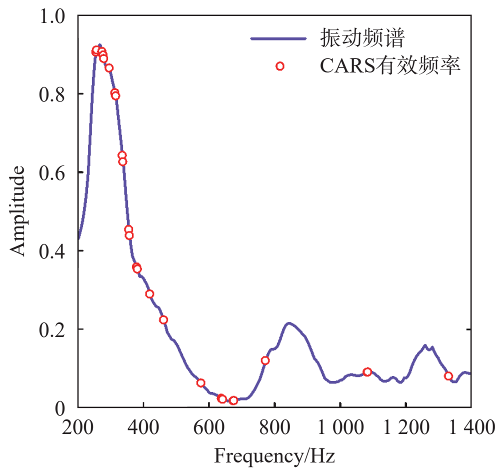


图 7　使用 CARS 算法从振动频谱中选择有效频率

Fig. 7　Selection of effective frequencies from the vibration spectrum using the CARS algorithm

使用网格搜索算法对 SVM 的惩罚系数‘*c*’, Gamma 值‘*g*’进行优化。当 SVM 模型参数确定为 c=0.1, g=0.4, 以及核函数为‘RBF’时具有最高的识别性能。使用振动频谱特征 (VST) 训练 PLS-DA 和 SVM 模型，结果如表 2 所示。从表中可以看出利用振动频谱特征训练的 SVM 和 PLS-DA

模型对果冻橙粒化病的识别效果均比使用振动多域图像纹理特征训练的 SVM 模型和 PLS-DA 模型的总体识别准确率更低。其中振动频谱特征训练的 PLS-DA 模型相较于 SVM 模型具有最高的识别准确率, 为 90.97%。但有 4 个正常果冻橙和 9 个粒化果冻橙识别错误。

表 2　基于振动频谱特征的 PLS-DA 和 SVM 模型对果冻橙粒化病的分类结果

Tab. 2　Classification results of jelly orange granulation disease by PLS-DA and SVM models based on vibration spectral features

模型	实际类别	预测类别		类别准确率	总体准确率
		正常	粒化		
SVM	正常	96	5	95.04%	89.58%
	粒化	10	33	76.74%	
PLS-DA	正常	97	4	92.08%	90.97%
	粒化	9	34	74.42%	

结果表明, 结合基于振动频谱特征的 PLS-DA 和 SVM 对果冻橙粒化病的识别性能较好, 但是还是有较多的正常果冻橙和粒化果冻橙识别错误。

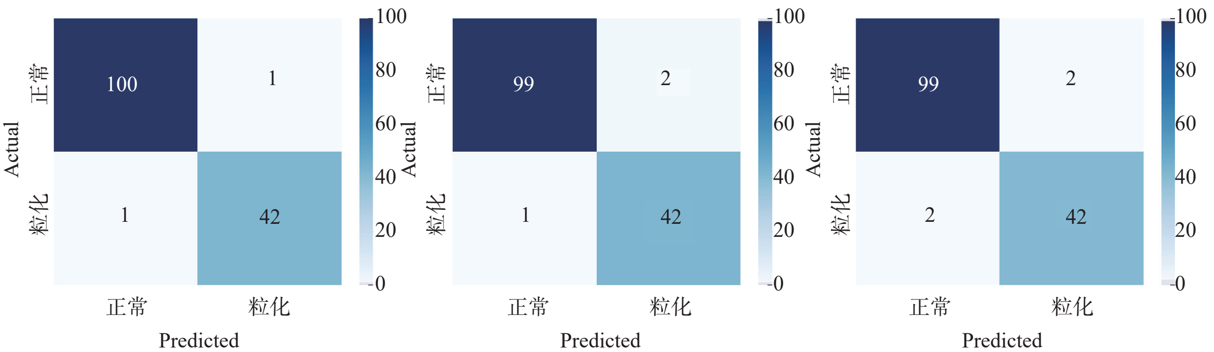
3.3　基于振动多域图像的 ResT 分类模型

构建了一个 Resnet-Transformer 分类模型 (ResT), 用于识别正常果冻橙和粒化果冻橙。同时, 将其性能与典型的 Resnet50 模型和 ViT 进行对比。使用振动多域图像训练 ResT、Resnet50 和 ViT 模型, 在模型训练过程中, 设置的 batch-size 大小为 64, 训练轮数为 100。表 3 给出了这 3 个模型的训练集和测试集结果。可以看出: ResT 和 Resnet50 在训练集上的准确率都为 100%; ViT 模型在训练集上的准确率为 99.53%。在预测集上, ResT 模型对正常果冻橙的分类准确率为 99.01%, 粒化果冻橙的分类准确率为 97.67%, 总体分类准确率为 98.61%。而 Resnet50模型对正常果冻橙的准确率为 98.02%, 粒化果冻橙的准确率为 97.67%, 总体准确率为 97.92%, ViT 模型的总体准确率为 97.22%, ResT 的准确率最高。由图 8 可知, ResT 可以识别出全部粒化果冻橙, 仅将 1 个正常果冻橙识别成粒化果冻橙。相比之下, Resnet50 将 2 个正常果冻橙识别成了粒化果冻橙, 将 1 个粒化果冻橙识别成了正常果冻橙。ViT 将 2 个正常果冻橙识别成了粒化果冻橙, 将 2 个粒化果冻橙识别成了正常果冻橙。

表 3　ResT、Resnet50 和 ViT 的训练集和预测集结果

Tab. 3　Training and prediction set results by ResT, Resnet50 and ViT models　(%)

模型	训练集准确率			预测集准确率		
	正常	粒化	总体	正常	粒化	总体
ResT	100.00	100.00	100.00	99.01	97.67	98.61
Resnet50	100.00	100.00	100.00	98.02	97.67	97.92
ViT	100.00	98.61	99.53	98.02	95.35	97.22



(a) ResT的结果
(a) Results for ResT
(b) Resnet50的结果
(b) Results for Resnet50
(c) ViT的结果
(c) Results for ViT

图 8　预测集的结果混淆矩阵

Fig. 8　Confusion matrix of the prediction set results

基于 VMIT 的 SVM (VMIT-SVM) 和 PLS-DA (VMIT-PLS-DA) 以及基于 VST 的 SVM (VST-SVM) 和 PLS-DA (VST-PLS-DA) 有较好的分类性能。为了进一步分析模型的分类性能, 采用 F1、精度和召回率评价指标对模型进行评价, 结果如图 9(彩图见期刊电子版)所示。可见, 与基于一维数据的 VMIT-SVM、VST-SVM、AVIT-PLS-DA 和 AVS-PLS-DA 模型相比, 基于振动多

域图像的 ResT 模型的 F1、Precision 和 Recall 均达到了 0.98 以上，显示出了最好的分类能力。与基于振动多域图像的 Resnet50 模型和 ViT 模型相比，ResT 模型的性能同样更好。这表明基于振动多域图像的 ResT 模型可以很好地对果冻橙粒化病进行检测。虽然 LDV 测振技术显示出较高的准确性，但设备成本较高，微型化的 LDV 设备还处于发展阶段，未来的适用性和推广还需考虑设备的可得性和经济性。此外，果实的大小、重量和形状会影响振动频谱，导致共振频率发生变化，这可能会限制该方法在不同水果中的普适性，尤其是对于形状、结构差异较大的水果。

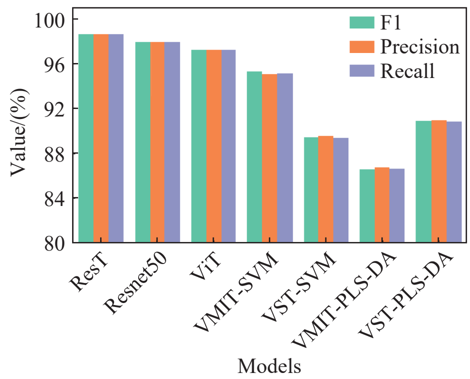


图 9 ResT、Resnet50、ViT、VMIT-SVM、VST-SVM、VMIT-PLS-DA 和 VST-PLS-DA 模型对果冻橙粒化病的识别性能比较

Fig. 9 Comparison of the jelly orange granulation disease identification performances obtained by ResT, Resnet50, ViT, VMIT-SVM, VST-SVM, VMIT-PLS-DA, and VST-PLS-DA models

4 结 论

本研究使用自行设计的声学振动检测装置采集果冻橙的振动响应信号，并将振动响应信号转换成振动多域图像。同时构建了 Resnet-Transformer 分类模型 (ResT)，以便从振动多域图像中提取全局和局部的振动特性和时频信息，实现对果冻橙粒化病的精准检测。此外，为了验证 ResT 模型的分类性能，将其与 Resnet50 和 ViT 模型进行对比。还将基于振动多域图像的 ResT 模型与基于振动多域图像纹理特征或振动频谱特征的 PLS-DA 和 SVM 模型进行比较。结果表明，相较于 Resnet50 和 ViT 模型，ResT 模型有更好的性能。此外，相较于使用从振动多域图像中提取的振动多域图像纹理特征或者振动频谱特征建立的 PLS-DA 和 SVM 模型，使用振动多域图像建立的 ResT 模型具有更好的性能。实验结果表明，基于振动多域图像的 ResT 模型对果冻橙粒化病的检测准确率为 98.61%，模型的 F1 为 0.986、精确率为 0.986、召回率为 0.986。由此说明，构建的基于振动多域图像的 ResT 模型可以用于果冻橙粒化检测任务。然而，虽然实现了果冻橙粒化病的精确检测，但还未将其应用到在线分选设备中。此外，声振法已经广泛应用于苹果、梨、西瓜、桃等水果的品质检测中。因此，开发在线分选设备以及检测不同水果品质将是下一步的工作重点。

参考文献:

[1] ZHENG Y J, TIAN SH J, XIE L J. Improving the identification accuracy of sugar orange suffering from granulation through diameter correction and stepwise variable selection[J]. *Postharvest Biology and Technology*, 2023, 200: 112313.

[2] THEANJUMPOL P, WONGZEEWASAKUN K, MUENMANEE N, *et al*.. Non-destructive identification and estimation of granulation in ‘Sai Num Pung’ tangerine fruit using near infrared spectroscopy and chemometrics[J]. *Postharvest Biology and Technology*, 2019, 153: 13-20.

[3] 陈玥瑶, 夏静静, 韦芸, 等. 近红外光谱法无损检测平谷产大桃品质方法研究[J]. 分析化学, 2023, 51(3): 454-462.
CHEN Y Y, XIA J J, WEI Y, *et al*.. Research on nondestructive quality test of Pinggu peach by near-infrared spectroscopy[J]. *Chinese Journal of Analytical Chemistry*, 2023, 51(3): 454-462. (in Chinese).

[4] 于水, 宦克为, 王磊, 等. 基于卷积神经网络的近红外光谱多组分定量分析模型研究[J]. 分析化学, 2024, 52(5): 695-705.
YU SH, HUAN K W, WANG L, *et al*.. Multicomponent quantitative analysis model of near infrared spectroscopy based on convolution neural network[J]. *Chinese Journal of Analytical Chemistry*, 2024, 52(5): 695-705. (in Chinese).

[5] JIE D F, WU SH, WANG P, *et al*.. Research on *Citrus grandis* granulation determination based on hyperspectral imaging through deep learning[J]. *Food Analytical Methods*, 2021, 14(2): 280-289.

[6] NAYAK S L, SETHI S, SAHOO R N, *et al*.. Potential of X-ray imaging to detect citrus granulation in different cultivars with progress in harvesting time[J]. *Indian Journal of Experimental Biology*, 2022, 60(4): 263-268.

[7] KADOWAKI M, NAGASHIMA S, AKIMOTO H, *et al*.. Detection of core rot symptom of Japanese pear (*Pyrus pyrifolia* cv. Kosui) by a nondestructive resonant method[J]. *Journal of the Japanese Society for Horticultural Science*, 2012, 81(4): 327-331.
[8] ZHANG H, ZHA ZH H, KULASIRI D, *et al*.. Detection of early core browning in pears based on statistical features in vibro-acoustic signals[J]. *Food and Bioprocess Technology*, 2021, 14(5): 887-897.
[9] ZHAO K, LI H, ZHA ZH H, *et al*.. Detection of sub-healthy apples with moldy core using deep-shallow learning for vibro-acoustic multi-domain features[J]. *Measurement: Food*, 2022, 8: 100068.
[10] JIN CH, XIE L J, YING Y B. Eggshell crack detection based on the time-domain acoustic signal of rolling eggs on a step-plate[J]. *Journal of Food Engineering*, 2015, 153: 53-62.
[11] SUN L, BI X K, LIN H, *et al*.. On-line detection of eggshell crack based on acoustic resonance analysis[J]. *Journal of Food Engineering*, 2013, 116(1): 240-245.
[12] WANG D CH, FENG ZH, JI SH Y, *et al*.. Simultaneous prediction of peach firmness and weight using vibration spectra combined with one-dimensional convolutional neural network[J]. *Computers and Electronics in Agriculture*, 2022, 201: 107341.
[13] 崔笛, 张文, 应义斌. 激光多普勒测振技术在农产品品质检测中的应用[J]. 农业机械学报, 2013, 44(7): 160-164.
CUI D, ZHANG W, YING Y B. Applications of laser doppler vibrometer technology in nondestructive detection of agro-product quality[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2013, 44(7): 160-164. (in Chinese).
[14] FATHIZADEH Z, ABOONAJMI M, BEYGI S R H. Nondestructive firmness prediction of apple fruit using acoustic vibration response[J]. *Scientia Horticulturae*, 2020, 262: 109073.
[15] ABBASZADEH R, RAJABIPOUR A, DELSHAD M, *et al*.. Application of vibration response for the nondestructive ripeness evaluation of watermelons[J]. *Australian Journal of Crop Science*, 2011, 5(7): 920-925.
[16] AKAN A, CURA O K. Time-frequency signal processing: Today and future[J]. *Digital Signal Processing*, 2021, 119: 103216.
[17] FU W L, JIANG X H, LI B L, *et al*.. Rolling bearing fault diagnosis based on 2D time-frequency images and data augmentation technique[J]. *Measurement Science and Technology*, 2022, 34(4): 045005.
[18] HUANG X F, LEI Q, XIE T L, *et al*.. Deep transfer convolutional neural network and extreme learning machine for lung nodule diagnosis on CT images[J]. *Knowledge-Based Systems*, 2020, 204: 106230.
[19] TANG X Y, XU Z B, WANG ZH G. A novel fault diagnosis method of rolling bearing based on integrated vision transformer model[J]. *Sensors (Basel, Switzerland)*, 2022, 22(10): 3878.
[20] LIU H, LIU ZH Y, JIA W Q, *et al*.. Tool wear estimation using a CNN-transformer model with semi-supervised learning[J]. *Measurement Science and Technology*, 2021, 32(12): 125010.
[21] ALSHAMMARI H, GASMI K, BEN LTAIFA I, *et al*.. Olive disease classification based on vision transformer and CNN models[J]. *Computational Intelligence and Neuroscience*, 2022, 2022: 3998193.
[22] ZHAO K, ZHA Z H, LI H, *et al*.. Early detection of moldy apple core based on time-frequency images of vibro-acoustic signals[J]. *Postharvest Biology and Technology*, 2021, 179: 111589.
[23] ZHANG W, CUI D, YING Y B. The impulse response method for pear quality evaluation using a laser Doppler vibrometer[J]. *Journal of Food Engineering*, 2015, 159: 9-15.
[24] DING CH Q, WANG D CH, FENG ZH, *et al*.. Integration of vibration and optical techniques for watermelon firmness assessment[J]. *Computers and Electronics in Agriculture*, 2021, 187: 106307.
[25] 张弛, 王顺, 关向雨, 等. 激光多普勒测振技术应用的新进展[J]. 激光与光电子学进展, 2022, 59(19): 1900006.
ZHANG CH, WANG SH, GUAN X Y, *et al*.. New progress in application of laser doppler vibration measurement technology[J]. *Laser & Optoelectronics Progress*, 2022, 59(19): 1900006. (in Chinese).
[26] WANG D CH, DING CH Q, FENG ZH, *et al*.. Recent advances in portable devices for fruit firmness assessment[J]. *Critical Reviews in Food Science and Nutrition*, 2023, 63(8): 1143-1154.

作者简介：

陈楠(1992—), 男, 江西兴国人, 博士, 讲师, 硕士生导师, 2022 年于中国科学院微电子研究所获得博士学位, 主要从事光电智能传感方面的研究。E-mail: chennan@ecjtu.edu.cn