基于时序多源信息融合的果蔬新鲜度在线检测系统

黄先果，朱启兵*，黄 敏

（江南大学 物联网工程学院，江苏无锡 214122，中国）

摘 要：［**目的/意义**］为实现果蔬在储运过程中的实时、精准品质监控，提出并实现了一种基于时序多源信息融合的在线检测系统。该系统旨在克服传统检测方法存在的离线性、单一模态及无法捕捉动态演变等缺陷，为冷链供应链的精细化管理与早期腐败预警提供新的技术途径。［**方法**］该系统通过自主设计的便携式采集节点，同步采集时序的环境传感数据与视觉指示标签图像。为深度挖掘两种异构时序数据间的复杂关联，本研究提出了一种新颖的协同注意力卷积循环网络（Co-Attention-Based Convolutional Recurrent Network, Co-ACRN）深度学习模型。该模型独创性地采用“协同注意力+自注意力”双重机制，前端通过协同注意力对时序多模态数据进行智能对齐，后端利用自注意力对时序信息进行全局上下文复盘。最终将模型高效部署于Qt端，实现了对果蔬新鲜度的边缘侧在线检测。［**结果和讨论**］Co-ACRN模型在芒果新鲜度检测任务中准确率达到了96.93%，显著优于多种先进的时序多模态融合模型；消融实验进一步证实了研究提出的“时序信息+多模态信息”输入的必要性，以及“双重注意力”架构的优越性，其中模型的各类别召回率最高可达99.16%。部署于客户端后，系统单次诊断耗时小于2 s，验证了该方案的高精度与实时性。［**结论**］开发的系统为分布式果蔬品质的在线、智能监控提供了一套兼具创新性与实用性的完整解决方案。

关键词：果蔬；冷链物流；时序多模态融合；双重注意力机制；在线无损检测

中图分类号：TP29 **文献标识码：**A **文章编号：**SA202505037

引用格式：黄先果, 朱启兵, 黄敏. 基于时序多源信息融合的果蔬新鲜度在线检测系统[J]. 智慧农业(中英文), 2026, 8(1): 203-212. DOI：10.12133/j.smartag.SA202505037

HUANG Xianguo, ZHU Qibing, HUANG Min. Online Detection System for Freshness of Fruits and Vegetables Based on Temporal Multi-source Information Fusion[J]. Smart Agriculture, 2026, 8(1): 203-212. DOI: 10.12133/j. smartag. SA202505037 (in Chinese with English abstract)

0 引 言

冷链储运技术是维持生鲜农产品品质并提高流通效率的重要途径[1]。近年来，随着人们生活水平的不断提高，对生鲜农产品的质量需求越来越高，冷链储运技术得以快速发展。据统计，2022年中国食品冷链物流需求总量为3.3亿t，其中果蔬占比为50.30%[2]。中国果蔬的冷链流通率为22%，冷藏运输率为35%，腐损率控制在15%以下。尽管如此，与发达国家5%以下的腐损率相比，仍存在明显差距[3]。因此，开发一种能够对果蔬新鲜度进行实时、精准、无损检测的在线监控系统，已成为保障生鲜供应链质量、减少经济损失的迫切需求。

存储环境会对果蔬质量产生重要影响[4]。ZHANG等[5]在运输和供应链过程中使用传感设备对温度、相对湿度和呼吸（O_2和CO_2浓度）指标进行采集，利用人工智能模型预测了黄瓜和草莓的品质。傅泽田等[6]在贮藏微环境中测量O_2、CO_2和乙烯的含量，利用反向传播（Back Propagation, BP）神经网络建立了蓝莓的货架期预测模型。尽管这些研究验证了环境数据在一定程度上可以反映果蔬新鲜度，但其核心缺陷在于无法直接量化果蔬内部生理状态（如细胞膜完整性、酶活性变化），导致模型普适性较差且难以应对突变情况[7]。果蔬新鲜度的指示标签是一种通过物理或化学变化直接反映果

收稿日期：2025-05-30

基金项目：国家重点研发计划项目（2022YFD2100601）

作者简介：黄先果，硕士研究生，研究方向为传感器与检测技术、智能检测。E-mail：18236021521@163.com

*通信作者：朱启兵，博士，教授，研究方向为信息感知与智能处理、物联网系统集成。E-mail：zhuqibing@jiangnan.edu.cn

copyright©2026 by the authors （登录www.smartag.net.cn免费获取电子版全文）

蔬新鲜状态的智能标签技术，其核心原理是通过监测与腐败相关的关键指标（如气体释放、pH变化等），以直观形式（如颜色变化）提示新鲜度[8]。近年来，基于指示剂的果蔬品质检测得到了广泛的研究[9]。FENG等[10]制备了基于花青素染料，甲基红和溴百里酚蓝混合染料的pH敏感指示标签，利用指示剂标签的颜色信息建立了青椒的新鲜度检测模型。黄霜霜等[11]以海藻酸钠和聚乙烯醇为基材，复配甲基红和溴百里酚蓝为指示剂，制备pH显色水凝胶用于鲜切哈密瓜的新鲜度监测，发现哈密瓜从新鲜转变为次新鲜时水凝胶从黄绿色变为橙色。然而，现有的研究大多依赖于人工视觉观测，难以满足果蔬品质检测的实时性及自动化的要求，且其精度容易受光照等环境因素干扰。综上所述，现有研究存在两大核心挑战：一是如何有效融合能够反映“内因”的环境数据与能够表征“外因”的视觉数据，以实现更全面的状态感知；二是如何将复杂的融合模型进行自动化、在线化部署，以满足实际应用的需求。

为应对上述挑战，本研究提出了一种基于时序多源信息融合的果蔬新鲜度在线检测系统。该系统的核心创新在于设计并构建了一种新颖的协同注意力卷积循环网络（Co-attention-based Convolutional Recurrent Network, Co-ACRN）。该网络采用“协同注意力+自注意力”的双重机制，旨在深度挖掘环境传感器时序数据与指示标签视觉时序数据之间的动态关联与长距离依赖。本研究将该模型成功部署于Qt客户端，通过与云平台和下位机采集节点的联动，最终实现了对果蔬新鲜度的实时、无损的在线检测与预警。本研究为利用先进人工智能技术解决复杂的生鲜供应链品质监控问题提供了有效的理论和实践范例。

1 果蔬新鲜度检测系统设计

1.1 检测系统构成

为实现对果蔬品质的实时、在线检测，设计并构建了1套由下位机（数据采集端）和上位机（数据处理与应用端）组成的物联网系统，其总体架构，如图1所示。下位机由内部移植了uC/OS-III实时操作系统的STM32F407ZGT6微处理器[12, 13]为核心，集成了用于采集冷链环境信息的传感器模块（温湿度、CO_2浓度、乙烯浓度）和用于获取指示标签信息的图像采集模块（OV5640摄像头）。所有采集到的数据通过ESP8266无线通信模块，以传输控制协议（Transmission Control Protocol，TCP）进行封装。上位机由OneNet云平台与部署了协同注意力卷积循环网络（Co-attention-based Convolutional Recurrent Network, Co-ACRN）神经网络模型的Qt客户端构成。下位机通过增强设备协议（Enhanced Device Protocol, EDP）将数据安全、高效地上传至OneNet云平台；Qt客户端则通过超文本传输协议（Hypertext Transfer Protocol, HTTP）向云平台发送数据请求指令，获取指定时间段内的时序多模态数据；最终在Qt客户端，这些数据被输入到Co-ACRN模型中进行新鲜度等级的实时在线检测，并将结果进行可视化展示。

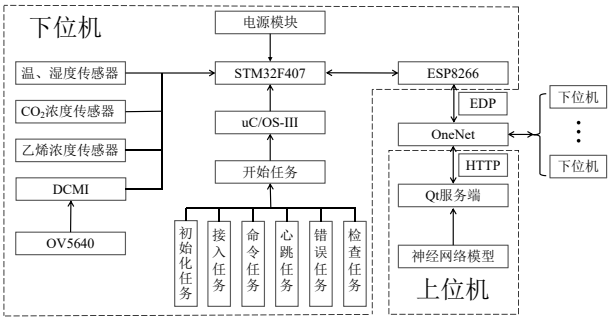


图1 果蔬新鲜度在线检测系统总体框图

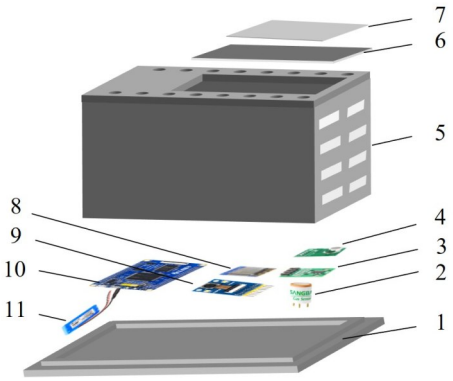
Fig. 1 Overall block diagram of online freshness detection system for fruits and vegetables

1.2 硬件设计

下位机装置核心组件如图2所示，由盒体底板、乙烯传感器、温湿度传感器、CO_2传感器、装置盒体、透明玻璃、无线通信模块、摄像头模块、微处理器和电源模块组成。装置外形尺寸为124 mm×91 mm×62 mm。盒体上方及其侧边设计了很多通风小孔以便传感器对外部环境的测量，透明玻璃承载指示剂标签配合摄像头模块完成图像的采集。

装置的数据采集与传输功能由STM32F407ZGT6微处理器控制并与多个专用模块协同完成[14]。传感器数据由SHT30（温湿度传感器）、MG-812（CO_2传感器）和HH-4C2H4-100（乙烯传感器）负责检测，它们分别通过集成电路总线（Inter-Integrated Circuit, I^2C）、模/数转换器（Analog-to-Digital Converter, ADC）和推荐标准485总线（Recommended Standard 485, RS485）与主控芯片通信；指示剂标签信息由1颗具有120º广角镜头的ATK-OV5640 500万像素互补金属氧化物半导体（Complementary Metal-Oxide-Semiconductor,

CMOS）摄像头模块[15]负责采集，并通过数字摄像头接口（Digital Camera Interface, DCMI）进行高速数据传输，其自带的两颗LED补光灯确保了图像质量的稳定性；所有采集到的数据，最终通过ESP8266（无线通信模块）上传至云平台；为保证装置的便携性与工作的可靠性，整个装置由5 V可充电锂电池供电。



注：1. 盒体底板；2. 乙烯传感器；3. 温湿度传感器；4.CO_2传感器；5. 装置盒体；6. 透明玻璃；7. 指示剂标签；8. 无线通信模块；9. 摄像头模块；10. 微处理器；11. 电源模块。

图2　下位机装置核心组件示意图

Fig. 2　Schematic diagram of core components of lower computer device

1.3 软件设计

1.3.1　下位机软件设计

为保证数据采集的实时性与多任务并发的稳定性，下位机软件基于uC/OS-III实时操作系统[16]进行开发。uC/OS-III通过基于优先级的抢占式调度[17]，确保高优先级的图像采集任务不会被低优先级的传感器轮询任务阻塞，从而保证了多模态数据在时间戳上的精准同步。系统启动后，主任务动态创建并管理4个核心子任务：温湿度传感器任务、气体传感器（CO_2与乙烯）任务、图像采集任务，以及网络通信任务，并通过信号量与消息队列机制[18]实现任务间的安全通信与资源共享。

1.3.2　上位机软件架构与模型部署

上位机软件采用“云-端”协同架构，由OneNet物联网云平台[19]作为数据中转与存储后端，Qt客户端作为用户交互与智能检测前端。在数据链路层面，本研究选择了EDP协议用于下位机到云端的数据上传，其在物联网场景下具有低开销、高效率的优势；而客户端与云端之间则采用通用的HTTP进行指令下发与数据请求。这种异构协议的组合，兼顾了嵌入式端的资源限制与应用端的开发便利性。

上位机软件架构的核心任务在于将训练好的Co-ACRN模型高效地部署于Qt客户端。为实现实时、低延迟的在线检测以克服传统云端推理存在的网络延迟与带宽依赖等缺陷，本研究采用了一种模型边缘侧部署（Edge-side Deployment）的优化方法，即采用ONNX（Open Neural Network Exchange）作为连接Python训练环境与C++应用环境的关键桥梁，通过ONNX Runtime推理引擎在客户端实现高性能的本地化推理。这种部署策略通过将AI计算任务前移至本地客户端，极大地提升了检测的实时性与鲁棒性，不仅实现了模型的跨平台应用，也保证了在客户端进行实时、高效的本地化新鲜度检测与预警，是实现在线检测系统的关键技术。

1.3.3　系统显示界面

在Qt端按下post按钮执行命令下发任务，按下get按钮进行数据上传，按下classify按钮进行果蔬品质分级并伴随预警功能：识别果蔬品质为新鲜时绿灯亮起，次新鲜时黄灯亮起，腐败时红灯亮起。按下auto按钮系统默认按照每1 h频率自动下发命令、上传数据与果蔬品质识别，其频率可通过左侧选项栏进行调整。界面左栏可以切换为其他子设备进行该子设备数据显示与果蔬品质分级。图3为Qt端该系统自动检测芒果新鲜度时的1个子设备界面，经测试，从发送数据采集命令到新鲜度识别时间间隔小于2 s，相比于裸机开发，实时操作系统的优势得以体现。

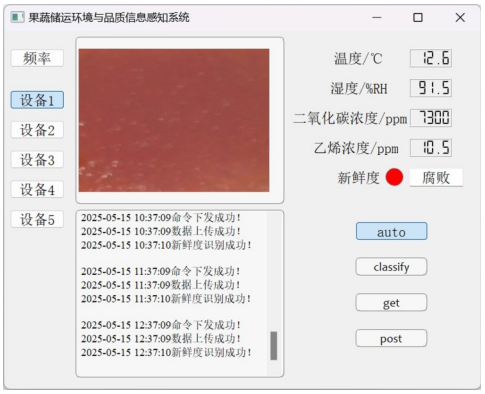


图3　Qt显示界面

Fig. 3　Qt display interface

2　果蔬新鲜度在线检测模型

为实现对果蔬在冷链储运中新鲜度状态的精准、动态在线检测，并解决传统方法在信息融合与时序分析上的局限性，本研究提出了一种Co-

ACRN模型。将得到的传感器数据与指示标签信息进行预处理操作后，对Co-ACRN架构设计、核心模块原理进行详细阐述，通过消融实验系统性地验证模型输入与模型架构各组件的贡献，最后将本研究模型与多种主流基线及前沿模型进行对比，以全面评估其性能。

2.1 数据预处理

为保证模型的训练效率与最终性能，原始数据在输入模型前必须进行标准化的预处理。

1）图像标签数据预处理。原始指示剂标签图像以320×240像素的JPEG格式存储，为适应卷积神经网络（Convolutional Neural Network, CNN）的输入要求并提升计算效率，选择对图像数据进行标准化预处理。首先，将所有原始图像采用双立方插值算法[20]统一降采样并重塑为64×64像素，在保留颜色、斑点等关键视觉特征的同时，显著降低模型的计算负担；确保所有图像均为RGB三通道格式，并将其像素值从原始的［0，255］整数范围线性归一化到［0，1］的浮点数区间。经过此流程，每张图像最终被转换为1个64×64×3的浮点数张量，作为后续模型的标准输入。

2）传感器数据预处理。在1 h的采样间隔内，温湿度、CO_2和乙烯浓度数据通常不会发生剧烈的阶跃式变化，因此主要采用前向填充策略处理潜在的缺失值，即使用每个缺失值之前最后1个有效的观测值来进行填充。另外两种极端情况：数据序列开头的连续缺失采用后向填充，即使用序列中第1个出现的有效观测值进行回填；序列中间出现连续时间较长的缺失则采用线性插值，在缺失段的起始和结束有效值之间创建线性趋势来进行填充，以保证数据在长时段内的平滑过渡。在所有缺失值通过上述组合策略被妥善处理后，采用最小-最大缩放方法消除各特征间的量纲差异，使所有特征值线性地缩放到［0，1］。

2.2 Co-ACRN模型

针对现有果蔬新鲜度检测方法存在的离线检测、信息融合肤浅、无法捕捉动态演变等缺陷，本研究提出了一种用于在线检测的新型深度学习模型—Co-ACRN，其详细架构如图4所示。该模型的核心创新在于构建了1个“跨模态早期对齐”与“时序内全局复盘”相结合的双重注意力体系，旨在从数学上模拟专家进行综合诊断的思维过程。模型主要包括4个阶段：CNN视觉特征提取、协同注意力融合、深度特征提取和分类决策。

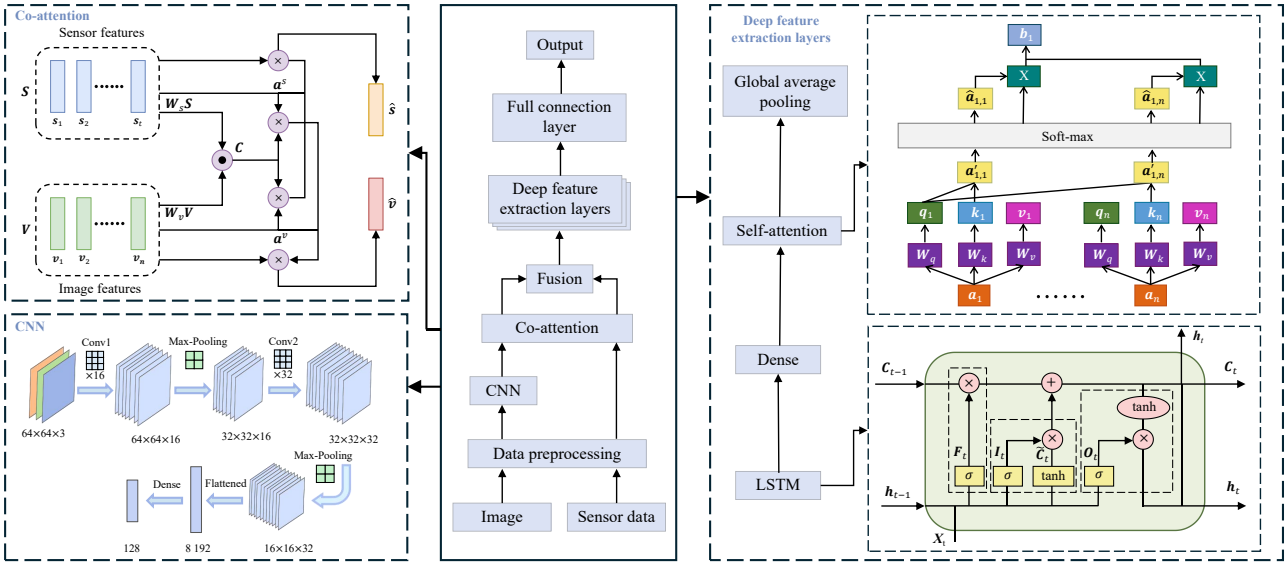


图4 Co-ACRN模型架构

Fig. 4 Co-ACRN model architecture

2.2.1 CNN视觉特征提取

为从图像序列中提取时序视觉特征，采用TimeDistributed层包裹1个轻量级CNN模型。该CNN旨在自动学习图像中的颜色、斑点及纹理等关键视觉信息，其结构包含两轮“卷积-池化”操作，并将最终的特征图通过全连接层映射为128维的视觉特征向量。该模块为输入的每1帧图像都生成1个对应的特征向量，形成有时序结构的视觉特征序列。

2.2.2　协同注意力融合

在多模态学习领域，传统的数据融合策略主要分为早期融合（Early Fusion）与晚期融合（Late Fusion）。在处理视觉特征序列与传感器特征序列这类异构时序数据时，传统的数据融合策略存在固有局限性。早期融合，即在时序分析前进行简单的特征拼接，其“生硬”的融合方式无法显式地建模两种模态在不同时间点上的动态关联，难以模拟专家根据一个信息源（如气体浓度）动态调整对另一个信息源（如颜色变化）关注度的诊断过程。晚期融合，即在各自独立完成时序分析后再进行决策融合，这种融合方式完全忽略了模态间在特征层面的底层交互，与果蔬腐败过程中内外因相互作用的物理现实相悖。

为克服上述传统方法的局限性，本研究引入了一种先进的协同注意力机制（Co-attention）[21] 进行智能融合。该机制的优越性在于其并非简单的特征拼接，而是通过构建跨模态的亲和力矩阵，来显式地计算视觉特征序列与传感器特征序列在每1个时间步之间的相互关联度，从而在特征层面就实现深度、智能的交互。其具体实现原理如下：假设输入的传感器特征序列为 $\boldsymbol{S}=\{\boldsymbol{s}_1,\ \boldsymbol{s}_2,\ \cdots,\ \boldsymbol{s}_t\}$，视觉特征序列为 $\boldsymbol{V}=\{\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ \cdots,\ \boldsymbol{v}_n\}$，其中 $\boldsymbol{S}\in\boldsymbol{R}^{d\times t}$，$\boldsymbol{V}\in\boldsymbol{R}^{d\times n}$，$d$ 为时间步长，t、n 为传感器特征维度与视觉特征维度。利用这两个特征序列计算它们的亲和力矩阵 $\boldsymbol{C}\in\boldsymbol{R}^{t\times n}$，如公式（1）所示。

$$\boldsymbol{C}=\tanh\left(\boldsymbol{S}^{T}\boldsymbol{W}_b\boldsymbol{V}\right) \tag{1}$$

式中：$\boldsymbol{W}_b\in\boldsymbol{R}^{d\times d}$ 为权重参数。得到亲和力矩阵后，一种简单的计算方式是直接取不同模态特征间的相似程度最大值作为该模态特征的注意力权重，即：$\boldsymbol{a}^{v}[n]=\max_i\left(\boldsymbol{C}_{i,\ n}\right)$ 和 $\boldsymbol{a}^{s}[t]=\max_j\left(\boldsymbol{C}_{t,\ j}\right)$。而使注意力机制表现出更优性能的计算方式是将所得亲和力矩阵视为一个特征，并通过学习来获得注意力权重，如公式（2）～公式（5）所示。

$$\boldsymbol{H}^{v}=\tanh\left(\boldsymbol{W}_v\boldsymbol{V}+\left(\boldsymbol{W}_s\boldsymbol{S}\right)\boldsymbol{C}\right) \tag{2}$$

$$\boldsymbol{a}^{v}=\mathrm{softmax}\left(\boldsymbol{w}_{hv}^{T}\boldsymbol{H}^{v}\right) \tag{3}$$

$$\boldsymbol{H}^{s}=\tanh\left(\boldsymbol{W}_s\boldsymbol{S}+\left(\boldsymbol{W}_v\boldsymbol{V}\right)\boldsymbol{C}^{T}\right) \tag{4}$$

$$\boldsymbol{a}^{s}=\mathrm{softmax}\left(\boldsymbol{w}_{hs}^{T}\boldsymbol{H}^{s}\right) \tag{5}$$

式中：$\boldsymbol{W}_v,\ \boldsymbol{W}_s\in\boldsymbol{R}^{k\times d}$；$\boldsymbol{w}_{hv},\ \boldsymbol{w}_{hs}\in\boldsymbol{R}^{k}$ 为权重参数；$\boldsymbol{H}^{v}$，$\boldsymbol{H}^{s}$ 为融合了跨模态上下文信息的中间特征表示；$\boldsymbol{a}^{v}\in\boldsymbol{R}^{n}$ 和 $\boldsymbol{a}^{s}\in\boldsymbol{R}^{t}$ 分别对应图像特征 $\boldsymbol{v}_n$ 和传感器特征 $\boldsymbol{s}_t$ 的注意力权重矩阵。基于前面计算所得的注意力权重，图像和传感器注意力向量可以分别通过其注意力权重与对应的原输入特征值进行加权结合得到最终输出，如公式（6）所示。

$$\hat{\boldsymbol{v}}=\sum_{i=1}^{n}\boldsymbol{a}_i^{v}\boldsymbol{v}_i,\hat{\boldsymbol{s}}=\sum_{j=1}^{t}\boldsymbol{a}_j^{s}\boldsymbol{v}_j \tag{6}$$

最后，在Fusion步骤中，将这两个被注意力增强过的序列进行拼接，形成一个统一的、信息更丰富的时序特征序列 $\boldsymbol{F}=\mathrm{concatenate}\left([\,\hat{\boldsymbol{v}},\ \hat{\boldsymbol{s}}\,]\right)$，作为后续深度特征提取模块的输入。

2.2.3　深度特征提取

1）时序动态编码：长短期记忆网络（Long Short-Term Memory, LSTM）

果蔬的腐败是一个累积性的、有时间依赖的过程。为有效捕捉这一特性，研究采用LSTM对经过协同注意力融合后的特征序列进行时序动态编码。LSTM凭借其独特的门控机制 [22] 能够对状态的“渐变”过程进行建模，并有效记忆如早期温度骤降等关键事件对后续状态的持续影响，从而形成初步的隐藏状态序列 $\boldsymbol{h}_t$，为后续的深层分析提供高质量的时序表示。

2）全局上下文建模：自注意力机制

为进一步挖掘序列内部的深层联系，克服传统循环神经网络（Recurrent Neural Network, RNN）对长距离依赖捕捉不足的缺陷，LSTM输出的隐藏状态序列（经过线性投影层以匹配模型维度后）被送入1个自注意力（Self-attention）模块。该模块通过计算查询、键和值矩阵之间的关系，捕捉序列中任意两个时间步之间的依赖关系，从而生成富含全局上下文信息的特征表示 [23]。其计算步骤如公式（7）～公式（9）所示。

$$\boldsymbol{q}_i=\boldsymbol{W}_q\boldsymbol{a}_i \tag{7}$$

$$\boldsymbol{k}_i=\boldsymbol{W}_k\boldsymbol{a}_i \tag{8}$$

$$\boldsymbol{v}_i=\boldsymbol{W}_v\boldsymbol{a}_i \tag{9}$$

式中：$\boldsymbol{a}_i$ 为输入向量；$\boldsymbol{W}_q$、$\boldsymbol{W}_k$、$\boldsymbol{W}_v$ 为可学习参数；$\boldsymbol{q}_i$ 为查询向量；$\boldsymbol{k}_i$ 为键向量；$\boldsymbol{v}_i$ 为值向量。

自注意力机制的核心思想是根据相似性来辨别输入之间的关系，使用内积来计算两个向量的相似性，如公式（10）所示。

$$\boldsymbol{a}'_{ij}=\frac{\boldsymbol{q}_i\times\boldsymbol{k}_j}{\sqrt{d_k}} \tag{10}$$

式中：$\boldsymbol{a}'_{ij}$ 为 $\boldsymbol{q}_i$ 与另一个输入 $\boldsymbol{k}_j$ 之间的相似性；d_k 为缩放参数。

在获得所有的 $\boldsymbol{a}'_{ij}$ 后，将softmax应用于 $\boldsymbol{a}'_{ij}$ 得到归一化权重 $\hat{\boldsymbol{a}}_{ij}$，这样 $\sum_{j=1}^{d}\hat{\boldsymbol{a}}_{ij}=1$。最后，通过确定有

关权重的值与所有输入的值的加权和，计算出输出向量，如公式（11）所示。

$$\boldsymbol{b}_i = \hat{\boldsymbol{a}}_{i1}\boldsymbol{v}_1 + \hat{\boldsymbol{a}}_{i2}\boldsymbol{v}_2 + \cdots + \hat{\boldsymbol{a}}_{id}\boldsymbol{v}_d \tag{11}$$

自注意力机制打破了循环网络对局部邻近信息的依赖，能够直接计算序列中任意两个时间点的相互影响。这使其能够发现腐败过程中跨越数小时的长距离因果关系，如识别出早期的1个温度骤降是导致后期气体浓度飙升的关键诱因，从而捕捉到状态的突变节点。

由于自注意力模块的输出仍为1个序列，为得出最终的单一诊断结论，本研究采用全局平均池化（Global Average Pooling）层对该序列进行信息汇总，通过时间维度上的平均操作将时间序列压缩为固定长度的特征向量，该操作在保留重要特征信息的同时也减少了参数量和计算复杂度。

2.2.4 分类决策

最终的分类决策由一个前馈神经网络完成。该网络接收池化后的特征向量，首先通过1个使用ReLU激活函数的隐藏全连接层进行非线性变换。为防止过拟合，加入Dropout层进行正则化；最后1层利用Softmax函数将高维特征映射为3种新鲜度的概率分布，由概率最高的类别决定最终的类别预测结果。

3 实验及结果分析

3.1 数据集制作

1）指示标签制作。JIA等[24]用甲基红与溴甲酚绿以3∶2的组合比例制作了一种混合指示标签，通过实验验证此比例对pH变化具有最佳的颜色表征性能，并有效地检测了芒果、猕猴桃与葡萄的新鲜度。

为了验证Co-ACRN模型对果蔬新鲜度分级的可行性，实验以芒果为例制作数据集进行测试。首先以甲基红与溴甲酚绿3∶2制作指示剂标签，该标签处于不同pH下的颜色，如图5所示。

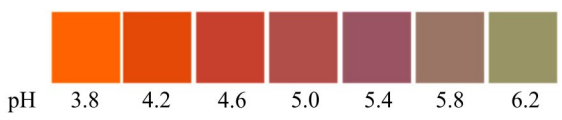


图5 不同pH下指示剂标签颜色变化

Fig. 5 Color changes of indicator labels at different pH levels

2）样本制作。实验使用的芒果在无锡市场新鲜采购，确保其处于成熟之前的阶段并无任何损坏。将芒果每5个为1组，分为10组，将各组的芒果与放有指示剂标签的下位机设备一起放进密封透明的存储箱中，然后将存储箱放入12 ℃的冷链环境中10 d，设置下位机设备每隔1 h自动对传感器数据与对标签数据进行1次采集。为构建数据集，采用每份芒果从第12时至第239时的有效数据，以确保所有模型的比较都在1个共同的时间区间内。对于时序模型（本研究模型）：使用步长为12的滑动窗口，每个样本包含连续12 h的数据，用于检测第12时末的状态；对于非时序模型（用于后续消融实验）：每个时间点的单帧数据被视为1个独立样本，用于检测其当前时刻的状态。基于这两种方法，每份芒果数据均可生成228个独立的、时间上对齐的样本。最终，10份芒果数据构成了2 280个样本的数据集。

3）新鲜度评价指标。本研究采用9分制专家感官评估法构建新鲜度标签。在数据采集期间，与下位机数据同步记录芒果的多角度高清图像，最后由15名经过标准培训的学生依据统一的感官标准（色泽、斑点、萎缩程度）对芒果新鲜程度进行打分，评分标准详见表1。为确保评分的客观性和一致性，将结果取平均值作为新鲜度依据[25]。为进行3分类模型训练，本研究依据分数区间将样本划分为3个等级：7～9分为“新鲜”；4～6分为“次新鲜”；1～3分为“腐败”。

表1 芒果感官评分标准

Table 1 Sensory evaluation criteria for mango

评分	感官评价
9	色泽鲜亮,颜色均匀;无任何瑕疵;果形饱满坚实
7～8	色泽良好,颜色基本均匀;存在极少量自然斑点;无萎缩迹象
5～6	光泽度开始下降,局部出现轻微暗沉;少量斑点;轻微皱缩
4	果皮普遍暗沉,颜色不均;斑点增多但未连成片;出现可见皱缩
3	果皮大面积暗沉,发黑;斑点扩大并有连片趋势;明显皱缩
1～2	大面积软烂,发黑;大面积连片斑点;果实严重皱缩

3.2 实验环境与网络训练

实验使用的硬件环境：AMD Ryzen 97945 HX（CPU），NVIDIA GeForce RTX 4060 （GPU），显存为8 GB；软件环境：操作系统为Windows 11，CUDA版本11.8，深度学习框架Pytorch 2.3.0，编程语言Python 3.9；设置模型网络参数：优化器采用Adam，初始学习率为0.001，批量样本个数为64，迭代次数为100，LSTM层的隐藏单元数128，步长为12。

将样本数据按数据集以7∶1∶2划分为训练集、验证集与测试集放入模型训练。

3.3 模型性能评价指标

本研究采用基于测试集混淆矩阵的标准化评价指标对模型的性能进行评估。考虑到在果蔬新鲜度监控中，及时发现状态变质具有重要的预警价值，研究重点关注召回率（Recall）与准确率（Accuracy）。其中召回率表示实际为正类的样本中，被正确预测为正类的比例；准确率表示所有样本中被正确分类的样本所占的总体比例，如公式（12）～公式（13）所示。

$$Recall = \frac{TP}{(TP + FN)} \tag{12}$$

$$Accuracy = \frac{TP + TN}{(TP + FP + TN + FN)} \tag{13}$$

式中：TP指把正类预测为正类的样本数；FP指把负类预测为正类的样本数；TN指把负类预测为负类的样本数；FN指把正类预测为负类的样本数。

3.4 实验结果与分析

3.4.1 芒果新鲜度检测结果

Co-ACRN模型基于芒果测试集新鲜度分类结果的混淆矩阵如图6所示。模型在测试集上的总体准确率为96.93%，对新鲜、次新鲜、腐败的召回率分别达到99.16%、96.84%和95.53%。模型不仅对“新鲜”与“腐败”两种极端类别具有较强的识别力，而且将极少数的分类错误严格限制在了物理状态最相近的相邻类别之间，从未出现跨类别的严重误判。

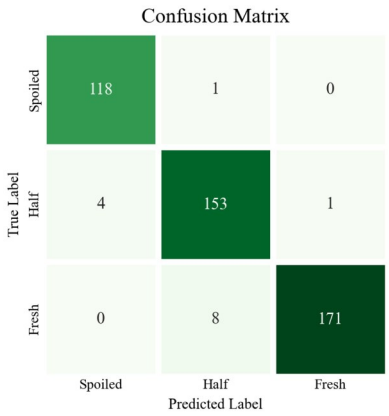


图6 Co-ACRN模型测试集混淆矩阵

Fig. 6 Confusion matrix of the Co-ACRN model on the test set

3.4.2 消融实验

为系统、定量地评估各项设计决策的有效性，本研究从模型输入和模型架构两个维度，设计了两组独立的消融实验。所有实验均在统一的数据集和实验框架下进行，以确保对比的公平性。

1）基于模型输入的消融实验。本组实验旨在深入探究不同信息源（传感器/图像）与信息模式（时序/非时序）对芒果新鲜度检测性能的独立及协同贡献。实验共设置了5个对比模型，以本研究时序多模态输入作为性能参照组。实验结果如表2所示。

表2 基于模型输入的消融实验结果

Table 2 Ablation experimental results based on model input

网络模型	模型输入			新鲜召回率/%	次新鲜召回率/%	腐败召回率/%	准确率/%
	时序	传感器数据	指示标签图像				
MLP	×	√	×	93.30	84.28	95.76	90.79
CNN	×	×	√	94.41	84.81	96.64	91.67
CNN-MLP	×	√	√	87.71	93.04	98.32	92.32
LSTM-Attn	√	√	×	91.62	88.24	94.96	91.23
CNN-LSTM-Attn	√	×	√	90.50	98.73	96.64	94.96
Co-ACRN	√	√	√	95.53	96.84	99.16	96.93

注：×表示不使用该项输入，√表示使用该项输入。

实验结果充分证明了时序动态与多模态信息协同的必要性。首先，无论是否引入时序，融合了视觉与传感器信息的模型性能均超越了任何单一模态模型，证实了多模态信息的互补性；其次，任何引入了时序信息的模型均大幅优于其对应的静态模型，彰显了捕捉动态演变趋势的核心价值。这为本研究提出的一个专门为此类数据而设计的先进模型架构（Co-ACRN）提供了坚实的立论基础。

2）基于模型架构的消融实验。在确立了采用“时序的多模态数据”作为最佳输入后，本组实验

旨在聚焦于模型内部架构，验证本研究核心创新模块的优越性。实验共设置了两个简化模型，同样以完整 Co-ACRN 模型作为参照。实验结果如表 3 所示。

表 3 基于模型架构的消融实验结果

Table 3 Ablation experimental results based on model architecture

模型架构		新鲜召回率/%	次新鲜召回率/%	腐败召回率/%	准确率/%
Co-attention	Self-attention				
×	×	89.50	91.03	94.96	91.45
√	×	93.30	94.94	96.64	94.73
×	√	92.27	94.23	95.80	93.86
√	√	95.53	96.84	99.16	96.93

注：×表示不使用该模块，√表示使用该模块。

基于模型架构的消融实验结果，清晰地量化了本研究双重注意力体系中每一个核心组件的独特贡献。以不包含高级注意力的骨干网络（91.45%）为基准，单独引入协同注意力或自注意力均能带来显著的性能提升，证明了“跨模态早期对齐”和“时序内全局复盘”两个策略本身的有效性。最终，集成了两个模块的完整 Co-ACRN 模型在所有指标上均达到最优表现（准确率 96.93%）。这一结果有效地证明 Co-ACRN 模型的优越并非源于单一技术的堆砌，而是其“先对齐、后复盘”的系统性架构设计，使得模型能够更全面、更深入地理解复杂的时序多模态数据。

3.4.3 多种检测模型对比

为进一步验证本研究提出的 Co-ACRN 模型在处理异构时序数据上的先进性，选择了几种在时序多模态学习领域具有代表性的前沿模型进行比较，这些模型为解决复杂的时序多模态问题提供了不同的思路，结果如表 4 所示。

MISA 通过将每个时间步的特征解耦为模态共享和模态独有的表示，为从复杂多模态信号中提炼纯净特征提供了新的视角。RAVEN 利用多跳记忆网络，让视觉模态能够反复查询传感器的时序记忆库，为捕捉由浅入深的跨模态动态交互提供了新的方法。MulT 采用多层交叉注意力 Transformer 直接对齐两个原始时序序列，为未对齐的多模态时序数据进行端到端融合提供了强大的解决方案。HHMPN 通过 1 个分层指针网络，同时建模模态内部（如气体浓度随时间的变化）和模态之间（如气体变化如何指向关键的视觉特征）的依赖关系，为异构信息的层次化融合提供了一种有效策略。

表 4 时序多模态模型对比结果

Table 4 Comparison results of temporal multimodal models

模型	新鲜召回率/%	次新鲜召回率/%	腐败召回率/%	准确率/%
MISA[26]	93.41	93.63	97.44	94.52
RAVEN[27]	94.41	94.38	98.31	95.39
MuIT[28]	92.82	94.27	96.61	94.30
HHMPN[29]	94.48	93.67	98.29	95.18
Co-ACRN	95.53	96.84	99.16	96.93

由实验结果可以看出，在时序多模态方法的对比实验中，本研究提出的 Co-ACRN 模型在所有评估指标上，包括各类别召回率和总体准确率，都明显超过了所有对比的先进方法。相较于致力于特征解耦的 MISA 和采用迭代式融合的 RAVEN，Co-ACRN 通过更直接、高效的协同注意力机制取得了显著的性能优势。即使与同样采用强大注意力机制的 MulT 和 HHMPN 相比，本研究的模型依然在所有关键指标上取得了最优表现，准确率分别领先 2.63% 和 1.75%，模型在避免了 MulT 庞大参数量的同时，其对称、双向的融合策略也比 HHMPN 的非对称融合更契合果蔬腐败中内外因互为因果的物理本质。综上所述，实验结果有力地证明了本研究模型在保持高性能的同时，为果蔬新鲜度在线检测实现了一种更加灵活、高效且鲁棒的解决方案。

4 结 论

本研究针对传统果蔬品质检测方法存在的离线性、单一模态和无法捕捉动态演变等核心缺陷，成功设计并实现了一种基于时序多源信息融合的在线检测系统。通过对系统各组件的设计与实现，以及对核心算法 Co-ACRN 的深入验证，本研究得出以下主要结论：

1）实验结果决定性地证明了精准的果蔬新鲜度检测必须依赖于时序信息与多模态信息的协同，任何单一维度的信息都存在固有局限性。

2）本研究提出的“协同注意力+自注意力”双重机制通过“先对齐、后复盘”的策略在处理异构时序数据上表现出显著的优越性，其检测准确率（96.93%）优于多种主流及先进模型。这证实了该架构在模拟复杂物理过程（如腐败）的动态演变上具有强大的潜力。

3）通过将高性能模型以 ONNX 格式高效部署于客户端，研究验证了边缘侧在线检测方案的可行

性，证明了该系统兼具高精度与实时响应能力。同时，系统具有广泛的适用性，针对任何代谢过程中挥发出酸性物质使指示剂颜色变化的果蔬，若要实现果蔬品质的分级，根据自制的指示剂标签制作数据集，然后将数据集放入模型训练即可。

综上所述，本工作为冷链供应链的精细化、智能化品质管控提供了一套行之有效的解决方案，并在理论和实践上为该领域的发展开辟了新的路径。展望未来，将当前的检测模型扩展为预测模型，实现对果蔬剩余货架期的精准预测为本工作后续重要的研究方向。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] 杨天阳, 田长青, 刘树森 . 生鲜农产品冷链储运技术装备发展研究[J]. 中国工程科学, 2021, 23(4): 37-44.
YANG T Y, TIAN C Q, LIU S S. Technology and equipment for cold-chain storage and transportation of fresh agricultural products[J]. Strategic Study of CAE, 2021, 23(4): 37-44.

[2] 徐静, 朱玉, 戴盼倩, 等 . 生鲜农产品冷链物流过程优化研究[J]. 中国农机化学报, 2024, 45(2): 115-121.
XU J, ZHU Y, DAI P Q, et al. Optimization research on cold chain logistics process of fresh agricultural foods[J]. Journal of Chinese Agricultural Mechanization, 2024, 45(2): 115-121.

[3] 白舒婕 . 补齐冷链短板,助农产品“鲜”行[N]. 国际商报, 2021-12-20(2).

[4] 曲劲亮 . 生鲜果蔬农产品物流模式研究: 评《果蔬贮藏及物流保鲜实用技术》[J]. 中国瓜菜, 2020, 33(5): 94.
QU J L. A review of "Practical techniques for storage, logistics, and fresh-keeping of fruits and vegetables": A study on logistics models for fresh fruits and vegetables[J]. China Cucurbits and Vegetables, 2020, 33(5): 94.

[5] ZHANG L H, ZHANG M, MUJUMDAR A S, et al. Advanced model predictive control strategies for nondestructive monitoring quality of fruit and vegetables during supply chain processes[J]. Computers and Electronics in Agriculture, 2024, 225: 109262.

[6] 傅泽田, 高乾钟, 李新武, 等 . 基于气体传感信息的蓝莓贮藏货架期预测方法[J]. 农业机械学报, 2018, 49(8): 308-315.
FU Z T, GAO Q Z, LI X W, et al. Blueberry shelf life prediction method based on sensor information stored gas[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(8): 308-315.

[7] HUANG X Y, YU S S, XU H X, et al. Rapid and nondestructive detection of freshness quality of postharvest spinaches based on machine vision and electronic nose[J]. Journal of Food Safety, 2019, 39(6): e12708.

[8] SHAO P, LIU L M, YU J H, et al. An overview of intelligent freshness indicator packaging for food quality and safety monitoring[J]. Trends in Food Science & Technology, 2021, 118: 285-296.

[9] HAILU F W, FANTA S W, TSIGE A A, et al. Current status and challenges of colorimetric intelligent packaging for fruit and vegetables: A review[J]. Journal of Food Quality, 2025, 2025(1): 6669828.

[10] FENG T L, CHEN H Z, ZHANG M. Applicability and freshness control of pH-sensitive intelligent label in cool chain transportation of vegetables[J]. Foods, 2023, 12(18): 3489.

[11] 黄霜霜, 王涛, 周倩怡, 等 . 海藻酸钠/聚乙烯醇pH指示水凝胶的制备及其对鲜切哈密瓜的新鲜度指示[J]. 食品与发酵工业, 2024, 50(16): 185-192.
HUANG S S, WANG T, ZHOU Q Y, et al. Preparation of pH indicating hydrogel based on sodium alginate/poly vinyl alcohol and application for freshness monitoring of fresh-cut cantaloupe[J]. Food and Fermentation Industries, 2024, 50(16): 185-192.

[12] ZENG J S, WANG Y, WU H T, et al. Research on the positioning method of steel belt anchor holes applied in coal mine underground[J]. Applied Sciences, 2024, 14(11): 4360.

[13] LIU Y S, CAO H K, LIU Y. Design of an online COD detection system for water quality[J]. Engineering Advances, 2024, 4(1): 1-10.

[14] DENG C, BIAN E H, GE Z S. Design and realization of stepping motor drive system controlled by single-chip microcomputer[J]. Wireless Personal Communications, 2022, 124(4): 3703-3724.

[15] YANG X T, ZHANG Y H, CHEN X Y, et al. Design and experimental Study of Plasma Device for Accurate Contour Scanning [J]. Vacuum, 2022, 205: 111442.

[16] REN Z Q, YU H Z, WANG S F, et al. Research on multi-functional intelligent ventilator based on UC/OS-III operating system for gas concentration detection[J]. Journal of Physics: Conference Series, 2022, 2246(1): 012028.

[17] UNGUREAN I. Timing comparison of the real-time operating systems for small microcontrollers[J]. Symmetry, 2020, 12 (4): 592.

[18] GAITAN N C, UNGUREAN I. Software vs hardware implementations for real-time operating systems[J]. International Journal of Advanced Computer Science and Applications, 2018, 9(12): 42-45.

[19] JIA S W, ZOU N Y, XU S H, et al. Applied research of the UAV illumination measurement system in sports stadiums[J]. Applied Sciences, 2023, 13(11): 6774.

[20] GOR A, BHENSDADIA C K. Self-supervised image denoiser design using multiscale bicubic image interpolation and U-Net network[J]. SN Computer Science, 2025, 6(3): 192.

[21] ZHOU X G, CHEN C, ZUO E G, et al. Cross branch co-attention network multimodal models based on Raman and FTIR spectroscopy for diagnosis of multiple selected cancers[J]. Applied Soft Computing, 2024, 166: 112204.

[22] XU Q, YANG G Y, YIN X B, et al. Reconstruction of sea surface chlorophyll-a concentration in the Bohai and Yellow Seas using LSTM neural network[J]. Remote Sensing, 2025, 17(1): 174.

[23] PALEKAR V, KUMAR L S. An effective image annotation using self-attention based stacked bidirectional capsule network[J]. Computer Standards & Interfaces, 2025, 93: 103973.

[24] JIA H J, WU C L, HUANG M, et al. An intelligent fruit freshness monitoring system using hydrophobic indicator labels based on methylcellulose, k-carrageenan, and sodium tripolyphosphate, combined with deep learning[J]. International Journal of Biological Macromolecules, 2025, 291: 140001.

[25] 李鑫, 朱磊, 张媛, 等 . 基于注意力时间卷积网络的香蕉新鲜度识别与剩余货架期预测[J]. 食品与机械, 2024, 40(11): 153-159.
LI X, ZHU L, ZHANG Y, et al. Freshness recognition and remaining shelf life prediction of banana based on attention temporal convolutional network[J]. Food & Machinery, 2024, 40(11): 153-159.

[26] HAZARIKA D, ZIMMERMANN R, PORIA S. MISA: Modality-invariant and-specific representations for multimodal sentiment analysis[C]// Proceedings of the 28th ACM International Conference on Multimedia. Stroudsburg, PA, USA: ACM, 2020: 1122-1131.

[27] WANG Y S, SHEN Y, LIU Z, et al. Words can shift: Dynamically adjusting word representations using nonverbal behaviors[J]. Proceedings of The Aaai Conference on Artificial Intelligence Aaai Conference on Artificial Intelligence, 2019, 33(1): 7216-7223.

[28] TSAI Y H, BAI S J, LIANG P P, et al. Multimodal transformer for unaligned multimodal language sequences[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA, USA: ACL, 2019: 6558-6569.

[29] ZHANG D, JU X C, ZHANG W, et al. Multi-modal multi-label emotion recognition with heterogeneous hierarchical message

passing[J]. Proceedings of The Aaai Conference On Artificial Intelligence, 2021, 35(16): 14338-14346.

Online Detection System for Freshness of Fruits and Vegetables Based on Temporal Multi-source Information Fusion

HUANG Xianguo, ZHU Qibing*, HUANG Min

(*School of Internet of Things Engineering*, *Jiangnan University*, *Wuxi* 214122, *China*)

Abstract: [Objective] Real-time and accurate quality monitoring of fruits and vegetables during cold chain logistics is of great importance for ensuring supply chain quality and reducing economic losses. However, traditional detection methods generally suffer from several core deficiencies, such as being offline, relying on unimodal information, and being unable to capture dynamic evolution. To overcome these challenges, an online freshness detection system is proposed and implemented for fruits and vegetables based on temporal multi-source information fusion. The system was designed to achieve precise online detection of fruit and vegetable freshness, providing an effective technical solution for the refined management and early spoilage warning within the cold chain supply chain, thereby significantly reducing economic losses. **[Methods]** A complete system was constructed, consisting of a lower-computer data acquisition node, an IoT cloud platform, and an upper-computer Qt client. The lower-computer synchronously collected environmental temporal sensing data (temperature, humidity, CO_2, ethylene) and visual temporal images of indicator tags via a self-designed portable acquisition node. A novel co-attention-based convolutional recurrent network (Co-ACRN) deep learning model was proposed for deeply mining the complex correlations between the two heterogeneous time-series data streams. This model innovatively employed a "co-attention + self-attention" dual mechanism. Firstly, in the early fusion stage, a co-attention module intelligently aligned and deeply integrated visual and sensor feature sequences by constructing a cross-modal affinity matrix. Subsequently, the fused sequence was fed into a long short-term memory (LSTM) network to encode temporal cumulative effects. Finally, a self-attention module performed a global contextual review on the LSTM output to capture long-range temporal dependencies. In the specific implementation, visual features were extracted by a lightweight convolutional neural network (CNN) with two convolutional-pooling layers; the co-attention calculated weights by generating context-aware intermediate features; and the self-attention adopted the standard scaled dot-product attention mechanism. For application deployment, the model was efficiently deployed to the Qt client in the open neural network exchange (ONNX) format, achieving real-time, edge-side inference. **[Results and Discussions]** Experimental results showed that the proposed Co-ACRN model achieved an overall accuracy of 96.93% on the test set in the three-class mango freshness detection task, with its performance significantly surpassing that of various mainstream baselines and advanced temporal multimodal fusion models, such as modality-invariant and specific-representations for multimodal sentiment analysis (MISA), recurrent attended variation embedding network (RAVEN), multimodal transformer (MulT), and heterogeneous hierarchical message passing network (HHMPN). To verify the rationale of the model design, two sets of ablation experiments were conducted. The input-based ablation study decisively proved that the combination of "time-series information + multimodal information" is a necessary prerequisite for accurate detection, as any model relying on unimodal or static information exhibited significant performance bottlenecks. The architecture-based ablation study further confirmed the superiority of the proposed "dual-attention" system; compared to a backbone network without any attention mechanism, its accuracy was improved by more than five percentage points, and the recall rate for the critical "spoiled" category was as high as 99.16%. An in-depth analysis of the confusion matrix revealed that the vast majority of the model's errors occurred between adjacent categories with the most similar physical states, with no serious cross-category misclassifications, demonstrating its strong robustness. After being deployed on the client side, the system's single diagnosis time was less than 2 s, verifying the solution's combination of high accuracy and real-time performance. **[Conclusions]** The developed online detection system and Co-ACRN model successfully enabled the real-time, accurate, and non-destructive intelligent detection of fruit and vegetable freshness. The research findings indicate that by combining advanced co-attention and self-attention mechanisms, the fusion challenges of complex multimodal temporal data can be effectively solved. In summary, this study provides a complete solution that combines theoretical innovation with engineering practicality for the online and intelligent detection of distributed fruit and vegetable freshness, and paves new paths for the development of this field in both theory and practice.

Key words: fruits and vegetables; cold chain logistics; temporal multimodal fusion; dual attention mechanism; online non-destructive testing

Foundation items: National Key R&D Program of China (2022YFD2100601)

First author: HUANG Xianguo, E-mail:18236021521@163.com

***Corresponding author:** ZHU Qibing, E-mail: zhuqibing@jiangnan.edu.cn