doi:10.6041/j.issn.1000-1298.2025.01.010

基于多模态信息融合的皮蛋溏心沙心分类方法

汤文权[1]　王巧华[1,2]　张　浩[1]　杨　烝[1]　范　维[1,3]

(1.华中农业大学工学院, 武汉 430070; 2.农业农村部长江中下游农业装备重点实验室, 武汉 430070;
3.国家蛋品加工技术研发分中心, 武汉 430070)

摘要: 溏心皮蛋与沙心皮蛋有着各自的口感和味道,均有各自受众,目前只能根据腌制时间来判断是溏心皮蛋还是沙心皮蛋,而这种方法不仅需要丰富的经验且误判比例较高。为了解决这一问题,本文设计了皮蛋红外图像和可见/近红外光谱采集装置,以及配套的溏心皮蛋和沙心皮蛋的分类模型。根据采集到的红外图像数据,在 ResNet18 网络添加 MLCA(Mixed local channel attention)模块,得到的改进模型 ResNet_MLCA 实现了溏心皮蛋和沙心皮蛋的分类,准确率为 95.0%。根据采集到的可见/近红外光谱数据,基于一维卷积设计了一维残差模块用于可见/近红外光谱数据的特征提取和分类,其对溏心皮蛋和沙心皮蛋分类准确率也达到 95.0%。为了进一步提高模型检测准确率,将 ResNet_MLCA 模型所提取的红外图像特征和 1D_ResNet 所提取的可见/近红外光谱特征进行融合,得到的融合模型 ResNet_OP 对溏心皮蛋和沙心皮蛋分类准确率达到 98.3%。研究成果提供了一种更低计算成本、更高准确率的溏心皮蛋和沙心皮蛋分类模型,对于指导皮蛋生产和提升皮蛋品质具有重要意义。

关键词: 皮蛋; 多模态信息融合; ResNet; 红外图像; 可见/近红外光谱; 溏心沙心

中图分类号: TS253.7　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)01-0092-10　　**OSID**:

Classification Methods for Soft-yolk and Hard-yolk Preserved Eggs Based on Multimodal Information Fusion

TANG Wenquan[1]　WANG Qiaohua[1,2]　ZHANG Hao[1]　YANG Zheng[1]　FAN Wei[1,3]

(1. *College of Engineering*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*
2. *Key Laboratory of Agricultural Equipment in Mid-Lower Yangtze River*, *Ministry of Agriculture and Rural Affairs*, *Wuhan* 430070, *China*
3. *National Egg Processing Technology Research and Development Branch Center*, *Wuhan* 430070, *China*)

Abstract: The soft-yolk preserved eggs (SYP eggs) and hard-yolk preserved eggs (HYP eggs) each possess distinct textures and flavors, captivating their respective discerning consumers. Presently, artisans can only discern whether an egg is a soft-yolk or hard-yolk preserved egg based on the duration of the brining process, a method that not only demands their extensive expertise but also entails a high rate of misjudgment. To address this issue, the design of infrared imaging and visible/near-infrared spectroscopy acquisition devices was introduced, alongside a classification model for SYP eggs and HYP eggs. Utilizing gathered infrared image data, an enhanced model, ResNet_MLCA, was crafted by incorporating a mixed local channel attention (MLCA) module into the ResNet18 framework, achieving a noteworthy classification accuracy of 95.0% in distinguishing SYP eggs from HYP eggs. Furthermore, leveraging visible/near-infrared spectroscopy data, a one-dimensional residual module was designed, and through its stacking, the 1D_ResNet model for feature extraction and classification of visible/near-infrared spectroscopy data was developed, yielding an identical accuracy of 95.0% in discriminating SYP eggs from HYP eggs. In a bid to further augment detection accuracy, the infrared image features extracted by the ResNet_MLCA model and the visible/near-infrared spectroscopy features extracted by the 1D_ResNet were amalgamated. The resultant fusion model, ResNet_OP, achieved an outstanding classification

收稿日期: 2024-07-07　修回日期: 2024-08-20

基金项目: 国家自然科学基金面上项目(32072302)、湖北省重点研发计划项目(20230611)和重庆市技术创新与应用发展专项乡村振兴(对口帮扶)项目(CSTB2023TIAD-ZXX0011)

作者简介: 汤文权(1997—),男,博士生,主要从事图像处理和深度学习研究,E-mail: 349346946@qq.com

通信作者: 王巧华(1970—),女,教授,博士生导师,主要从事农畜产品无损检测技术及装备研究,E-mail: wqh@mail.hzau.edu.cn

accuracy of 98.3% in distinguishing SYP eggs from HYP eggs. In summary, this research can offer a novel, cost-effective, and high-precision classification model for SYP eggs and HYP eggs, which held significant implications for guiding preserved egg production and enhancing its quality. Additionally, the proposed method offered a theoretical reference for enhancing the performance of classification models for other agricultural products, aiming to further increase their accuracy and reduce the number of parameters in the fusion model.

Key words: preserved egg; multimodal information fusion; ResNet; infrared imaging; visible/near-infrared spectroscopy; soft-yolk and hard-yolk

0　引言

皮蛋作为中国传统美食,不仅具有消炎、抗癌、抗肿瘤、降血压的功效[1-3],也因其独特的风味一直受到消费者喜爱[4]。市场上销售的皮蛋主要分为溏心皮蛋和沙心皮蛋两类,其主要区别为:溏心皮蛋的蛋黄为半液体,口感细腻浓郁;沙心皮蛋的蛋黄为固体,口感紧实绵密,不同消费者对这两类皮蛋有各自的偏爱。目前工厂中主要通过腌制时间来判断是溏心皮蛋还是沙心皮蛋,沙心皮蛋往往需要更久的腌制时间,但皮蛋腌制时间太久,不仅浪费资源,还会导致皮蛋蛋清变得松散,影响蛋清口感[5]。但实际上,由于皮蛋个体差异,同一缸腌制液的浓度分布也不完全均匀,腌制时的温度也在随着时间发生改变,导致溏心皮蛋转变为沙心皮蛋的时间有所差异,因而根据时间来判断是溏心皮蛋或沙心皮蛋是不完全准确的。因此需要开发一种装置来实现溏心皮蛋和沙心皮蛋分类,帮助工人判断溏心皮蛋转变为沙心皮蛋的时间,避免皮蛋腌制时间过长。

机器视觉技术因其包含丰富的形状、纹理、空间等信息,不仅被大量应用于蛋品外部品质的检测中[6-8],在蛋品内部品质无损检测领域也有大量应用[9-12],证明了机器视觉技术在皮蛋检测领域的有效性。由于皮蛋透光性差,而红外光具有更强的穿透能力,相较于 RGB 图像,红外图像拥有更加丰富的内部信息;并且红外图像因不受自然光影响,其图像中的干扰信息更少。鉴于此,本文选择红外图像作为机器视觉技术的数据。

光谱能够表征物质化学成分和结构信息,包括分子振动、转动以及电子能级等,因其丰富的细节信息在蛋品品质无损检测领域应用非常广泛[13-17],并已经实现了皮蛋凝胶弹性、成熟度、含水量等指标的预测[18-19],证明了光谱检测皮蛋内部品质的有效性。

多模态信息融合技术目前主要应用于图像描述生成[20]、视觉问答[21]、跨模态检索[22]和情感分析[23]等任务中,而在农产品品质检测任务中的应用很少。合适的多模态信息融合模型的检测效果通常远高于单一模态信息模型的检测效果[24],但现有的多模态信息融合模型往往需要很大的计算量和大量的训练数据,然而在农产品品质检测任务中实验数据往往很难获取,且实际应用的检测设备通常是计算能力有限的嵌入式设备,为了实现应用,要求所建立的模型不能太大。为此本文分别采用红外相机和便携式可见/近红外光谱仪采集皮蛋内容物的红外图像和可见/近红外光谱;将 ResNet18 作为基础网络进行改进,建立皮蛋红外图像的分类模型;利用一维卷积设计一维残差模块,建立皮蛋可见/近红外光谱分类模型;将所建立皮蛋红外图像分类模型和可见/近红外光谱分类模型所提取的特征进行融合,建立溏心皮蛋和沙心皮蛋多模态信息融合分类模型。

1　材料和设备

1.1　实验设备和环境

模型运行的硬件设备为:32 GB 内存,搭载 AMD Rezen 3600 CPU,GPU 为 NVIDIA RTX1070;运行环境为:Windows 10(64 位)操作系统,Python 3.8.13 编程语言版本,Pytorch 1.13 深度学习框架,CUDA 版本为 11.3,运行软件为 Pycharm 2020.1。

1.2　数据获取

采用如图 1 所示的装置采集皮蛋红外图像和可见/近红外光谱,将皮蛋水平放置于托板上,皮蛋位置刚好在光源正上方,托板中间开有小孔,可以使光源光线通过,皮蛋完全覆盖住小孔。

准直透镜位于皮蛋右侧,正对着皮蛋的一端,采集漫透射皮蛋的可见/近红外光谱数据,由于皮蛋透光性差,此方法相较于从皮蛋的正上方采集,其光强度衰减更少,数据特征也更加明显。采集到的两类皮蛋平均光谱曲线和误差带如图 2 所示,其中两类皮蛋光谱曲线在 750 nm 附近和 900 nm 附近的波峰相对强度存在差异,但是总体上看每一类皮蛋光谱曲线个体之间差异巨大,给模型识别造成一定困难。

红外相机采集的图像波段范围为 750 ~ 900 nm,位于皮蛋正上方,可以更加完整地采集到整个皮蛋的内部信息,所采集到的红外图像如图 3 所示。溏心

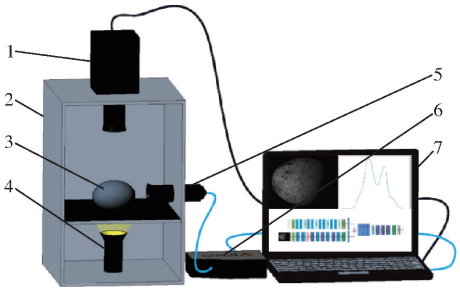


图 1 红外图像和可见/近红外光谱数据采集装置

Fig. 1 Device for acquiring infrared images and visible/near-infrared spectra data

1. 红外相机 2. 暗箱 3. 皮蛋 4. 光源 5. 准直透镜 6. 光谱仪 7. 计算机

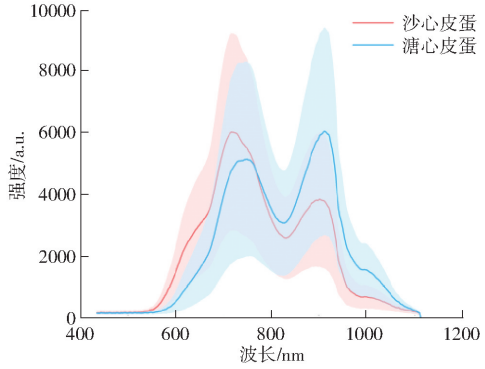
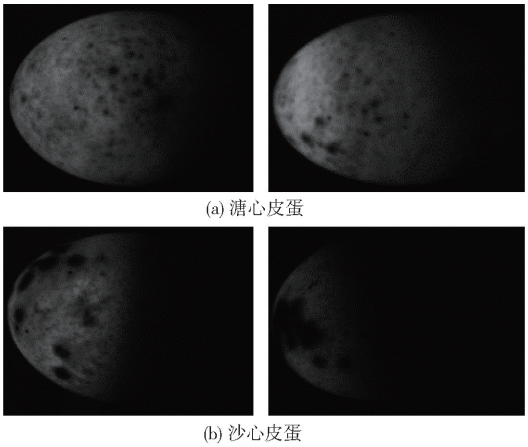


图 2 皮蛋的可见/近红外光谱图

Fig. 2 Visible/near-infrared spectra of preserved egg



(a) 溏心皮蛋

(b) 沙心皮蛋

图 3 皮蛋红外图像

Fig. 3 Infrared images of preserved egg

皮蛋腌制时间更短,其透光性更好,因此透光区域面积更大、整体亮度也更高;沙心皮蛋腌制时间更久、蛋壳内侧更易形成大块斑点,因此能看到内部图像大块斑点更多。

沙心皮蛋样本和溏心皮蛋样本各 296 枚,总样本为 592 枚。将采集到的红外图像和可见/近红外光谱按照相同方式随机划分为 7∶2∶1(在 0 ~ 296 之间生成一组 7∶2∶1 随机序列,取出序列对应编号的皮蛋数据划分进相应的数据集中,为保证红外图像和可见/近红外光谱的数据在各数据集中能够一一对应,本实验仅生成一次随机序列,溏心皮蛋和沙心皮蛋共用一套随机序列,同时红外图像数据和可见/近红外光谱数据也共用这一套序列)作为训练集、验证集和测试集来建立模型,其样本分布情况如表 1 所示。

表 1 数据集划分

Tab. 1 Dataset division

类别		训练集	验证集	测试集	总计
红外图像/幅	溏心皮蛋	207	59	30	296
	沙心皮蛋	207	59	30	296
可见/近红外光谱/条	溏心皮蛋	207	59	30	296
	沙心皮蛋	207	59	30	296

2 分类模型构建

2.1 红外图像分类模型

残差神经网络(Residual network,ResNet)是一种深度卷积神经网络架构,通过引入残差学习的概念,在网络中建立 Shortcut Connections,从而使得信息可以更容易地传播,解决了深度神经网络训练过程中出现的梯度消失和梯度爆炸问题。ResNet 以其优秀的分类性能[25],其原始模型及其变体模型被广泛应用于食品无损检测领域[26-31]。因此本文以 ResNet18 作为基础模型,对其进行改进,来实现皮蛋红外图像分类。

MLCA(Mixed local channel attention)模块是一种混合局部空间信息和全局空间信息的轻量级注意力机制模块,实现检测效果、速度和模型参数量之间的平衡,并使该注意力机制同时包含通道信息、空间信息、局部通道信息和全局通道信息[32]。如图 4 所示,该模块首先将输入特征向量进行一次局部池化,将输入转换为 $C \times k_s \times k_s$,以提取局部空间信息。初始阶段使用两个分支将输入转换为一维向量,第 1 个分支包含全局信息,第 2 个分支包含局部空间信息;经过一维卷积后,通过反池化恢复两个向量原始分辨率,然后将信息融合,实现了混合注意力的目标。

图 5 描述了 GAP、LAP 和 UNAP 的具体过程。GAP(Global average pooling)为全局平均池化,通过矩阵整体执行平均池化将特征图变为 1 × 1;UNAP 为反平均池化,主要作用是将缩小后的特征图扩展回原较大的特征图,通过将池化后的值平均分配到对应的上采样区域来实现;LAP(Local average pooling)为局部平均池化,与 GAP 通道的总特征图不同,LAP 将整个特征图分成 $k \times k$ 个矩阵,然后对每个矩阵执行平均池化。

基于 ResNet18 模型进行改进,所设计的红外图

像特征提取分类模型如图 6 所示，其主要改进点为 Layer1 和 Layer2 之间引入 MLCA 模块，大幅提升了模型特征提取能力，且模型参数量和计算量均无明显提升。

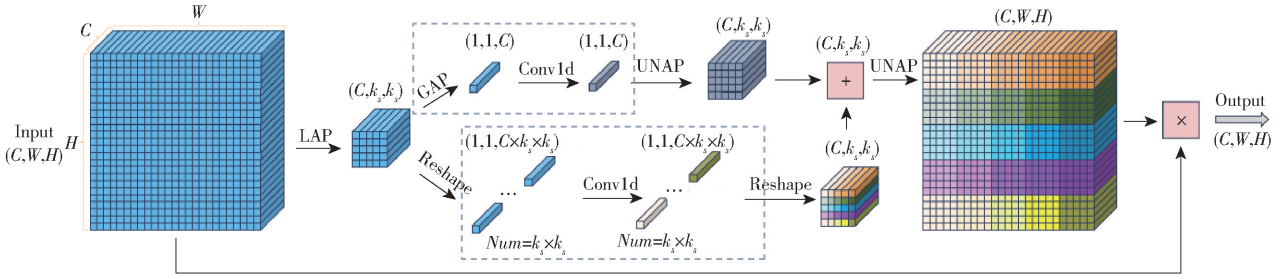


图 4　MLCA 原理图

Fig. 4　Schematic of MLCA

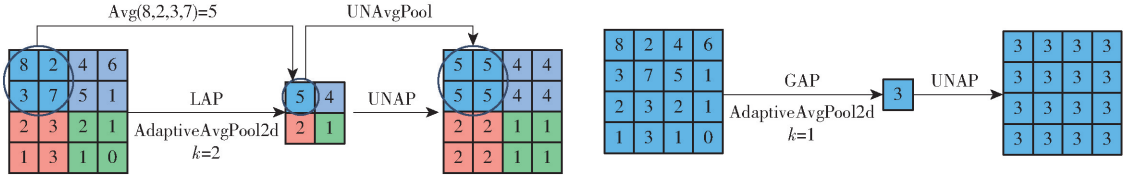


图 5　GAP、LAP 和 UNAP 计算原理图

Fig. 5　Principles of computation for GAP, LAP and UNAP

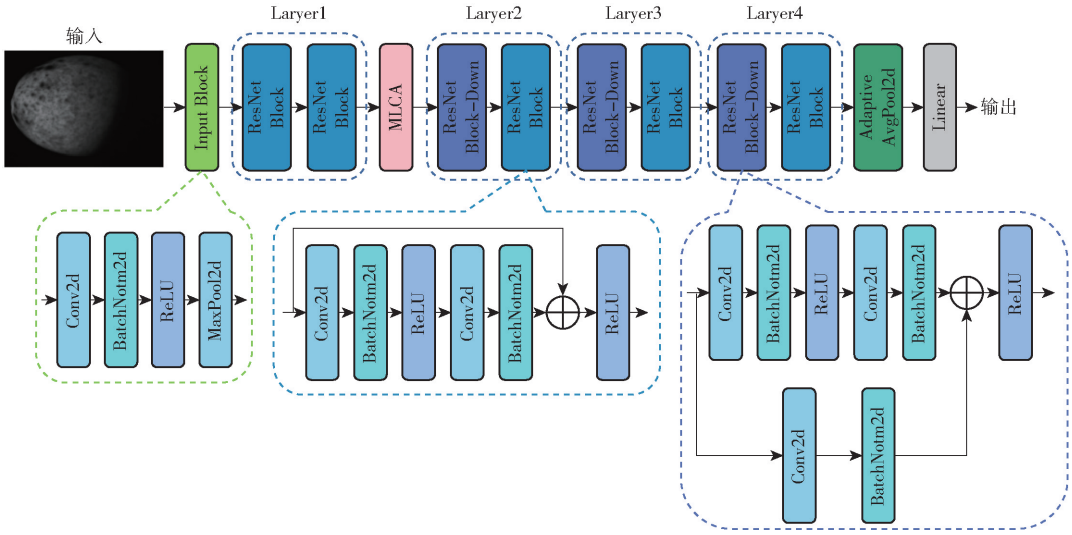


图 6　红外图像分类模型结构图

Fig. 6　Structure diagram of infrared image classification model

2.2　可见/近红外光谱分类模型

近年来，卷积神经网络（Convolutional neural network，CNN）以其强大特征提取能力被大量应用于光谱数据分析中[33-35]，BENMOUNA 等[36]所提出的 CNN 模型在苹果成熟度的分类任务中，表现好于人工神经网络（Artificial neural network，ANN）、支持向量机（Support vector machine，SVM）和 k 近邻（K-nearest neighbors，KNN）这些传统分类模型。CNN 仅需提供输入和输出数据，即可在模型内部完成特征提取和分类过程，其优秀的学习能力在可见/近红外光谱数据分析中存在巨大潜力[37]。

为了消除散射、噪声和光程等对数据的影响，在输入模型之前添加 SG 模块来平滑光谱曲线，并设计了如图 7 所示的 1D_ResNet 网络来对光谱信息进行分类。采用一维卷积来建立一维残差模块，通过 Res1D Block 模块和升维一维卷积的堆叠，实现光谱特征提取和分类。相较于 ResNet 网络，1D_ResNet 网络在输入模块和 Res1D Block 模块中未使用 batchnorm 模块，同时没有构建升维残差模块，而是通过升维一维卷积实现特征层升维。模型输入为 1 × 1 568 的一维光谱数据，经过 Input Block 中的一维卷积升维到 32 × 784，再经 Layer1 最后面的一维卷积形状变为 64 × 392，经过 Layer2 形状为 128 × 196，经 Layer3 形状为 256 × 98，经 Layer4 形状为 512 × 49，经过 Adaptive AvgPool1d 形状为 512 × 1，最后经一个线性层实现分类，经检验，此设计有较强的分类和特征提取能力。

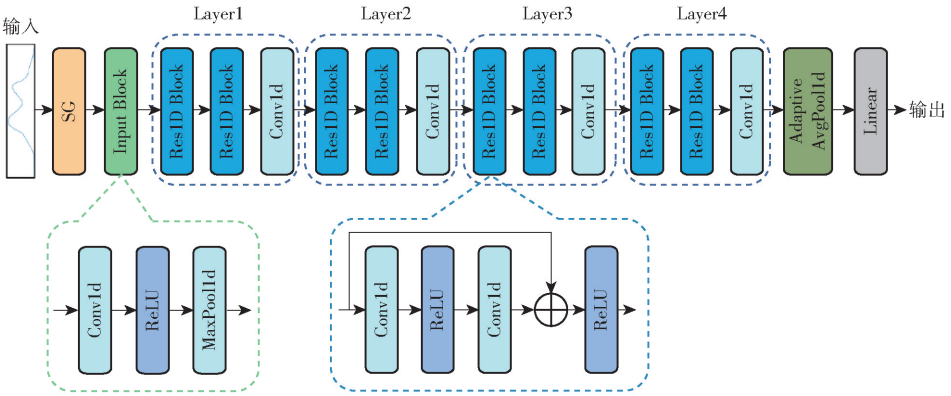


图 7 可见/近红外光谱分类模型结构图

Fig. 7 Structure diagram of visible/near-infrared spectral classification model

2.3 多模态信息融合分类模型

目前常见的两个一维特征矩阵的融合方式为拼接(Concatenate, Cat)、相加(Addition, Add)、外积(Outer product, OP)。Cat 的融合方式直接将第 2 个矩阵拼在第 1 个矩阵的后面形成一个更长的一维矩阵,计算方式如图 8a 所示,其优势在于可以保留原始特征的所有信息,增加了特征的丰富性,它在融合时也不用要求被融合特征的长度相等,但没有耦合两类特征之间的关系;Add 的融合方式为直接将 2 个矩阵的对应位置相加,得到一个与之前维度相同的特征矩阵,计算方式如图 8b 所示,其优势在于可以加强特征之间的协同作用,减少了干扰信息,但其特征量较少,会丢失部分特征信息,同时要求被融合的矩阵维度相同;OP 的融合方式为先将第 2 个一维矩阵进行转置操作,再将 2 个一维矩阵相乘形成一个二维特征矩阵,计算方式如图 9c 所示,其优势在于能够捕捉到特征之间的交互作用和高阶关系,可提供更加丰富的特征表示,但会增加计算成本,由于本文仅采用卷积结构来搭建融合模型且模型均选取参数量较小的模型,因此外积所增加的计算量是可以接受的。

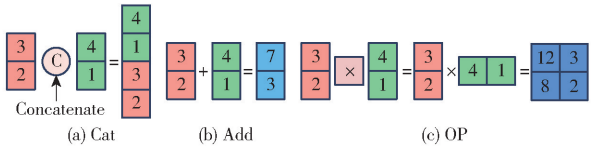


图 8 特征融合方式

Fig. 8 Method of feature fusion

经对比,OP 的融合方式最优,因此采用 OP 的融合方式设计了如图 9 所示的模型来实现可见/近红外光谱数据和红外图像数据的融合,将所建立的红外图像分类模型和可见/近红外光谱分类模型的最后一个线性层和分类层去掉,直接获取两个模型的特征提取部分,并将各自提取到的两组一维特征矩阵计算外积得到一个二维特征矩阵,再经过融合模块进一步提取融合后的特征,将所提取的关键特征经过分类模块最终实现溏心皮蛋和沙心皮蛋的分类。

所采用的融合方式为:将特征提取层得到两组 512 个特征数据组成的一维向量计算外积得到一个

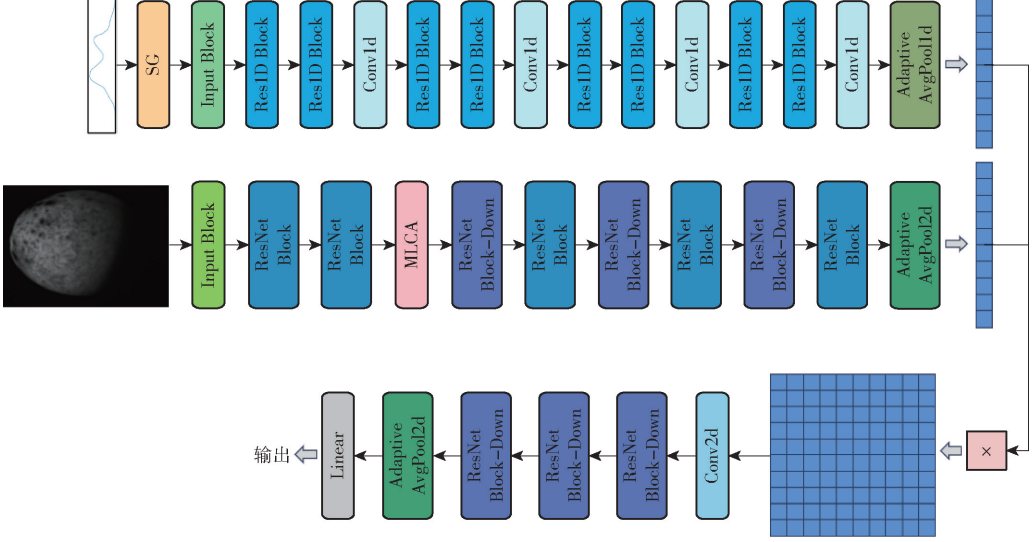


图 9 特征融合模型结构图

Fig. 9 Structure diagram of feature fusion model

1×512×512 的特征矩阵，特征矩阵通过一个升维卷积将矩阵形状变为 32×256×256，再经过 3 个具有下采样功能的残差模块进一步提取关键特征，其形状转变为 256×32×32，最后经过全局平均池化提取出融合后的 256 个关键特征，再经过一个线性层实现分类。

3　模型评价指标

为了评估分类模型性能，采用准确率(Accuracy)来评价模型整体性能，召回率(Recall)、精确率(Precision)、F1 值(F1 score)来评价模型对两类样本的分类性能。

4　结果与分析

4.1　红外图像分类结果对比

4.1.1　经典模型

为了找到最适合本文数据集的基础模型，选择经典分类模型 DenseNet[38]、ResNet[25]、EfficientNet[39]、MobileNetV3[40]在该数据集下进行分类效果对比。为了便于本文模型实现应用，降低模型对硬件的依赖，需要选择参数量小的模型，因此除 MobileNetV3 外，本文选取的模型均为参数量最小的版本，其检测结果如表 2 所示。

表 2　经典模型检测结果

Tab. 2　Detection results of classical models　%

模型	召回率		精确率		F1 值		准确率
	溏心皮蛋	沙心皮蛋	溏心皮蛋	沙心皮蛋	溏心皮蛋	沙心皮蛋	
MobileNetV3_large	89.7	87.1	86.7	90.0	88.1	88.5	88.3
EfficientNet_b0	96.3	87.9	86.7	96.7	91.2	92.0	91.7
DenseNet121	100	85.7	83.3	100	90.9	92.3	91.7
ResNet18	100	88.2	86.7	100	92.9	93.8	93.3

由表 2 可知，在上述经典模型中，ResNet18 对溏心皮蛋和沙心皮蛋的分选能力最强，因此选择 ResNet18 作为基础模型进行改进。从总体分类准确率上看，MobileNetV3_large 的分选能力最弱，其准确率仅为 88.3%，EfficientNet_b0 和 DenseNet121 准确率相同，均为 91.7%，而 ResNet18 准确率最高，可达 93.3%。从各类召回率、精确率和 F1 值看，MobileNetV3_large 对两类模型分选能力最均衡，但其对各类的分选效果和整体准确率都不理想，其他 3 个模型都表现为更加倾向于将沙心皮蛋识别为溏心皮蛋，只有 EfficientNet_b0 会将少量的溏心皮蛋识别为沙心皮蛋。

4.1.2　注意力机制对模型的提升效果

为了找到合适的注意力机制来提升模型检测效果，在基础模型 ResNet18 中分别添加注意力模块 SE[41]、CBAM[42]、GAM[43]、ECA[44]、CA[45]、MLCA[32]进行试验，添加注意力模块后得到的模型分别为 ResNet_SE、ResNet_CBAM、ResNet_GAM、ResNet_ECA、ResNet_CA、ResNet_MLCA，这些模型在本文数据集中的分类效果如表 3 所示。

表 3　注意力机制对模型的提升效果

Tab. 3　Enhancement effect of attention mechanisms on model　%

模型	召回率		精确率		F1 值		准确率
	溏心皮蛋	沙心皮蛋	溏心皮蛋	沙心皮蛋	溏心皮蛋	沙心皮蛋	
ResNet_SE	93.1	90.3	90.0	93.3	91.5	91.8	91.7
ResNet_CBAM	96.3	87.9	96.7	96.7	91.2	92.1	91.7
ResNet_GAM	93.3	93.3	93.3	93.3	93.3	93.3	93.3
ResNet_ECA	100	88.2	86.7	100	92.9	93.8	93.3
ResNet_CA	96.4	90.6	90.0	96.7	93.1	93.5	93.3
ResNet_MLCA	100	90.9	90.0	100	94.7	95.2	95.0

由表 3 可知，在初始的 ResNet18 模型中添加的所有注意力模块中，仅有 MLCA 模块可以在本文的数据集上提升模型的分类准确率，因此选择 MLCA 来提升模型检测效果。从分类准确率上看，SE 和 CBAM 模块降低了模型检测准确率，其准确率均由 93.3%下降至 91.7%，GAM、ECA 和 CA 模块对模型分选准确率没有提升效果，分类准确率与原模型相同，均为 93.3%，仅有 MLCA 使模型准确率由 93.3% 提升至 95.0%。从各类的召回率、精确率和 F1 值上看，SE 和 CBAM 降低了模型分类能力，ECA 模块的添加对模型检测效果几乎没有改变，GAM 模块和 CA 模块均可使模型对两类的识别能力变更均衡，其中 GAM 模块的添加使模型对每一类的识别能力相近，MLCA 模块在原模型的基础上提升了模型对两类的检测效果。

4.2　可见/近红外光谱

4.2.1　特征波长选择算法对比

为了验证特征波长选择算法对本文数据集是否有效，采用有竞争自适应重加权算法(Competitive adaptive reweighted sampling，CARS)、无信息变量消除(Uninformative variable elimination，UVE)和连续投影算法(Successive projections algorithm，SPA)这 3 种最常见的特征波长选择算法进行对比，并建立 SVM 分类模型，特征波长选择及分类结果如表 4 所示。

由表 4 可知，总体上特征波长选择算法虽然可以使特征波长点数大幅减少，但也会使模型分类效果下降，其中 CARS 算法作为最优的特征波长选

表 4 特征波长选择算法分类结果

Tab. 4 Classification results of feature wavelength selection algorithm %

特征选择算法	特征波长点数	召回率		精确率		F1 值		准确率
		溏心皮蛋	沙心皮蛋	溏心皮蛋	沙心皮蛋	溏心皮蛋	沙心皮蛋	
无	1 558	100	90.9	90.0	100	94.7	95.2	95.0
CARS	26	93.3	93.3	93.3	93.3	93.3	93.3	93.3
UVE	309	96.7	86.7	87.9	96.3	92.1	91.2	91.7
SPA	80	96.7	90.0	90.6	96.4	93.5	93.1	93.3

表 5 可见/近红外光谱分类模型结果对比

Tab. 5 Detection results of visible/near-infrared spectroscopy classification models %

模型	召回率		精确率		F1 值		准确率
	溏心皮蛋	沙心皮蛋	溏心皮蛋	沙心皮蛋	溏心皮蛋	沙心皮蛋	
SVM	100	90.9	90.0	100	94.7	95.2	95.0
ANN	93.1	90.3	90.0	93.3	91.5	91.8	91.7
1DCNN	100	85.7	83.3	100	90.9	92.3	91.7
1D_ResNet	96.6	93.4	93.3	96.7	94.9	95.1	95.0

择算法,所提取的特征波长点数量最少,模型的检测准确率下降幅度较小。常见的特征波长选择算法无法提升模型的分类准确率,为此只能寻找更为有效的皮蛋溏心沙心分类算法。

4.2.2 分类模型对比

选取可见/近红外光谱中常用的分类模型——支持向量机(SVM)、人工神经网络(ANN)、一维卷积神经网络(1DCNN)与本文所设计的一维残差神经网络(1D_ResNet)在本文数据集上的分类效果进行对比,结果如表 5 所示。

从表 5 可知,1D_ResNet 和 SVM 分类准确率均为 95.0%,是 4 个模型中准确率最高的,ANN 和 1DCNN 分类准确率仅为 91.7%,说明在本文的可见/近红外光谱数据的分类中 1D_ResNet 和 SVM 效果更好。从各类召回率、精确率和 F1 值来看,1D_ResNet 比 SAM 模型对每一类的识别能力更加均衡,且 1D_ResNet 中包含有特征提取和分类的功能,SVM 仅能实现分类,因此 1D_ResNet 将特征提取部分取出,实现光谱特征和图像特征的融合,有更大的改进空间。

4.3 特征融合模型

将 Cat、Add、OP 这 3 种融合方式所得到结果进行对比,3 种融合方式得到的模型分别为 ResNet_Cat、ResNet_Add 和 ResNet_OP。为了节省计算资源,使融合模型快速收敛,采用分布式训练的方式来训练融合模型,即将已经训练好的红外图像分类模型和可见/近红外光谱分类模型的特征提取部分的权重直接加载到融合模型的特征提取部分,并在训练时冻结融合模型的红外图像和可见/近红外光谱的特征提取部分权重,仅训练特征融合和分类部分权重,融合模型分类效果如表 6 所示。

由表 6 可知,Cat 和 Add 这两种特征融合方式对模型分类准确率几乎没有提升,其对各类检测效果也都与未融合之前的红外图像分类模型相同,而 OP 融合方式则有效提升了模型分类准确率,使得准确率提升为 98.3%,其对各类检测效果也都有明显提升,沙心皮蛋召回率由 90.9% 提升至 96.8%,溏心皮蛋精确率由 90.0% 提升至 96.7%,两类 F1 值也都有明显提升。从单枚皮蛋运算时间上看,OP 融合方式虽然相较于其他两种方式在检测时间上增加 30% 以上,但运算时间仍远小于实时检测所需时间,因此 OP 为本文中最合适的融合方式。

表 6 特征融合模型分类结果

Tab. 6 Classification results of feature fusion model

特征选择算法	单枚皮蛋运算时间/ms	召回率/%		精确率/%		F1 值/%		准确率/%
		溏心皮蛋	沙心皮蛋	溏心皮蛋	沙心皮蛋	溏心皮蛋	沙心皮蛋	
ResNet_Cat	61.1	100	90.9	90.0	100	94.7	95.2	95.0
ResNet_Add	60.1	100	90.9	90.0	100	94.7	95.2	95.0
ResNet_OP	80.1	100	96.8	96.7	100	98.3	98.4	98.3

4.4 讨论

从整体上看,溏心皮蛋召回率大于沙心皮蛋,沙心皮蛋精确率大于溏心皮蛋,所有模型均倾向于将溏心皮蛋分类为沙心皮蛋,其可能的原因是在腌制沙心皮蛋过程中,大部分沙心皮蛋由于腌制时间较长,其特征与溏心皮蛋相差较大,此部分皮蛋在分选过程中可被模型较好地识别,而在腌制溏心皮蛋的过程中,由于部分皮蛋因本身孔隙度较大,或处于腌制液浓度更高的位置上,这些皮蛋更快转变为沙心皮蛋,而此部分沙心皮蛋特征与溏心皮蛋更为接近,从而导致了模型的误判,能否将此部分皮蛋成功地识别是本文模型识别能力的关键。

传统的光谱特征波长选择算法在本文的数据中表现不佳,其可能原因是皮蛋个体差异较大,导致采集到的皮蛋光谱数据相差较大,经过特征选择之后可能会放大个体之间的差异,导致模型分类效果下

降。本文在设计 1D_ResNet 模型时，没有直接套用二维图像分类模型中的 ResNet 结构，这是因为在试验中发现，设计 Res1D Block 时添加 BatchNormal1d 模块反而会使模型失去识别能力，所以在设计该模块时仅采用一维卷积和 ReLU 作为激活函数。其中可能的原因是原光谱曲线中原始数据不同波段的实际强度是两类数据的主要差异所在，而经过 BatchNormal1d 模块处理之后数据损失了实际值的大小差异，仅保留了曲线的形状，由此丢失的重要信息导致模型在光谱数据中失去了识别能力。同时本文所构建的模型中进行了 5 次下采样操作，试验结果表明该方法优于 4 次下采样操作，其可能原因是光谱数据长度为 1 558，比图像长、宽（224 ×224）大的多，多次升维操作所提取的特征可能关注到更大范围的数据，因此效果更好。

为了确保深度学习模型低参数量和低计算量，本文在红外图像分类任务中仅对比了不同模型的最小版本，但是相对而言 ResNet18 仍是这 4 个模型中参数量最大的模型，因此有可能是模型参数量导致 ResNet18 的识别能力强于其他 3 个模型，为了证明这一点后期将在更好的运行设备上去尝试模型参数量对红外图像分类效果的影响。

5　结束语

设计了可同时采集红外图像和可见/近红外光谱的装置，成功采集了溏心皮蛋和沙心皮蛋的红外图像和可见/近红外光谱，并以 ResNet18 作为基础网络，通过添加 MLCA 注意力机制，得到的 ResNet_MLCA 模型实现了红外图像分类，仅使用红外图像数据对溏心皮蛋和沙心皮蛋的分类准确率可达 95.0%。同时基于一维卷积设计的 1D_ResNet 模型在可见/近红外光谱数据的特征提取和分类中均有较好的表现，使用该模型在仅有可见/近红外光谱数据的情况下其对溏心皮蛋和沙心皮蛋的分类准确率可达 95.0%。将 ResNet_MLCA 模型和 1D_ResNet 模型的特征提取部分所提取的特征进行融合，采用分布式训练将最佳的 ResNet_MLCA 模型权重和1D_ResNet 模型权重加载到融合模型的特征提取部分，将所提取的特征采用计算外积的方式得到特征矩阵，采用下采样残差模块进一步融合并提取特征，最终得到了溏心皮蛋和沙心皮蛋多模态信息融合识别模型，其准确率可达 98.3%，满足实际应用中对准确率的要求，同时模型参数量也在可应用的范围内。

参　考　文　献

[1] BATOOL Z, HU G, XINYUE H, et al. A comprehensive review on functional properties of preserved eggs as an excellent food ingredient with anti-inflammatory and anti-cancer aspects[J]. Food Bioscience, 2021, 44: 101347.

[2] MAO C, YU Z, LI C, et al. The functional properties of preserved eggs: from anti-cancer and anti-inflammatory aspects[J]. Korean J. Food Sci. Anim. Resour., 2018, 38(3): 615 - 628.

[3] ZHANG M, ZHAO Y, WU N, et al. The anti-inflammatory activity of peptides from simulated gastrointestinal digestion of preserved egg white in DSS-induced mouse colitis[J]. Food & Function, 2018, 9(12): 6444 - 6454.

[4] GANASEN P, BENJAKUL S. Physical properties and microstructure of pidan yolk as affected by different divalent and monovalent cations[J]. LWT—Food Science and Technology, 2010, 43(1): 77 - 85.

[5] XUE H, HAN T, XU M, et al. Processing technology, principle, and nutritional characteristics of preserved eggs: a review [J]. Trends in Food Science & Technology, 2022, 128: 265 - 277.

[6] 郭建军, 杨霖, 张恩威, 等. 基于机器视觉的裂纹鸡蛋分拣系统设计[J]. 现代农业装备, 2024, 45(1): 45 - 52.
GUO Jianjun, YANG Lin, ZHANG Enwei, et al. Design of cracked egg sorting system based on machine vision[J]. Modern Agricultural Equipment, 2024, 45(1): 45 - 52. (in Chinese)

[7] 梁旭, 王玲, 赵书涵. 基于残差神经网络的鸡蛋分类识别研究[J]. 河南农业大学学报, 2024, 58(3): 456 - 466.
LIANG Xu, WANG Ling, ZHAO Shuhan. Research on egg classification and recognition based on residual neural network[J]. Journal of Henan Agricultural University, 2024, 58(3): 456 - 466. (in Chinese)

[8] TANG W, HU J, WANG Q. High-throughput online visual detection method of cracked preserved eggs based on deep learning [J]. Applied Sciences, 2022, 12(3):952.

[9] HORKAEW P, KUPITTAYANANT S, KUPITTAYANANT P. Noninvasive in ovo sexing in Korat chicken by pattern recognition of its embryologic vasculature[J]. Journal of Applied Poultry Research, 2024, 33(3): 100424.

[10] SAH İN H A, ONDER H. Use of deep learning to determine the freshness of egg[J]. Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 2024, 14(1): 493 - 500.

[11] 白竣文, 靳家欣, 孙力, 等. 基于透射图像的高压脉动腌制咸鸭蛋含盐量检测[J]. 农业工程学报, 2024, 40(8): 245 - 253.
BAI Junwen, JIN Jiaxin, SUN Li, et al. Detection of salt content in high-pressure pulsed salted duck eggs based on transmission images[J]. Transactions of the CSAE, 2024, 40(8): 245 - 253. (in Chinese)

[12] TANG W, ZHANG H, CHEN H, et al. A non-destructive detection and grading method of the internal quality of preserved

eggs based on an improved ConvNext[J]. Foods, 2024, 13(6):925.
[13] 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)
[14] 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.
[15] CAI K, FANG Q, LIN Q, et al. Fast discrimination of female and male pigeon eggs using internet of things in combined with Vis-NIR spectroscopy and chemometrics[J]. Microchemical Journal, 2024, 203: 110883.
[16] COZZOLINO D, SANAL P, SCHREUDER J, et al. Predicting egg storage time with a portable near-infrared instrument: effects of temperature and production system[J]. Foods, 2024, 13(2): 212.
[17] YAO K, SUN J, ZHANG B, et al. On-line monitoring of egg freshness using a portable NIR spectrometer combined with deep learning algorithm[J]. Infrared Physics & Technology, 2024, 138: 105207.
[18] 陈远哲, 王巧华, 高升, 等. 皮蛋凝胶品质含水率和弹性的高光谱预测及其可视化[J]. 食品科学, 2022, 43(2): 324-331.
CHEN Yuanzhe, WANG Qiaohua, GAO Sheng, et al. Hyperspectral imaging for prediction and visualization of water content and springiness as indicators of the gel quality of preserved eggs[J]. Food Science, 2022, 43(2): 324-331. (in Chinese)
[19] CHEN Y, WANG Q, FAN W, et al. Non-destructive determination and visualization of gel springiness of preserved eggs during pickling through hyperspectral imaging[J]. Food Bioscience, 2023, 53: 102605.
[20] XU L, TANG Q, LV J, et al. Deep image captioning: a review of methods, trends and future challenges [J]. Neurocomputing, 2023, 546: 126287.
[21] XI Y, ZHANG Y, DING S, et al. Visual question answering model based on visual relationship detection [J]. Signal Processing: Image Communication, 2020, 80: 115648.
[22] BELTRÁN L, CAICEDO J, JOURNET N, et al. Deep multimodal learning for cross-modal retrieval: one model for all tasks [J]. Pattern Recognition Letters, 2021, 146: 38-45.
[23] BIRJALI M, KASRI M, BENI-HSSANE A. A comprehensive survey on sentiment analysis: approaches, challenges and trends [J]. Knowledge-Based Systems, 2021, 226: 107134.
[24] 李善军, 宋竹平, 梁千月, 等. 基于 X-ray 和 RGB 图像融合的实蝇侵染柑橘无损检测[J]. 农业机械学报, 2023, 54(1): 385-392.
LI Shanjun, SONG Zhuping, LIANG Qianyue, et al. Nondestructive detection of citrus infested by bactrocera dorsalis based on X-ray and RGB image data fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 385-392. (in Chinese)
[25] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
[26] BU Y, HU J, CHEN C, et al. ResNet incorporating the fusion data of RGB & hyperspectral images improves classification accuracy of vegetable soybean freshness[J]. Scientific Reports, 2024, 14(1): 2568.
[27] 姜红花, 杨祥海, 丁睿柔, 等. 基于改进 ResNet18 的苹果叶部病害多分类算法研究[J]. 农业机械学报, 2023, 54(4): 295-303.
JIANG Honghua, YANG Xianghai, DING Ruirou, et al. Identification of apple leaf diseases based on improved ResNet18[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(4): 295-303. (in Chinese)
[28] MOMENY M, JAHANBAKHSHI A, NESHAT A, et al. Detection of citrus black spot disease and ripeness level in orange fruit using learning-to-augment incorporated deep networks[J]. Ecological Informatics, 2022, 71: 101829.
[29] ZHU H, YANG L, FEI J, et al. Recognition of carrot appearance quality based on deep feature and support vector machine [J]. Computers and Electronics in Agriculture, 2021, 186: 106185.
[30] XUE G, LI S, HOU P, et al. Research on lightweight YOLO coal gangue detection algorithm based on ResNet18 backbone feature network[J]. Internet of Things, 2023, 22: 100762.
[31] SUN X, GU J, HUANG R. A modified SSD method for electronic components fast recognition [J]. Optik, 2020, 205: 163767.
[32] WAN D, LU R, SHEN S, et al. Mixed local channel attention for object detection[J]. Engineering Applications of Artificial Intelligence, 2023, 123: 106442.
[33] NG W, MINASNY B, MONTAZEROLGHAEM M, et al. Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra[J]. Geoderma, 2019, 352: 251-267.
[34] HAO Y, LI X, ZHANG C, et al. Online inspection of browning in Yali Pears using visible-near infrared spectroscopy and interpretable spectrogram-based CNN modeling[J]. Biosensors, 2023, 13(2): 203.
[35] WANG D, ZHAO F, WANG R, et al. A lightweight convolutional neural network for nicotine prediction in tobacco by near-infrared spectroscopy[J]. Frontiers in Plant Science, 2023, 14: 1138693.

[36] BENMOUNA B, GARCÍA-MATEOS G, SABZI S, et al. Convolutional neural networks for estimating the ripening state of fuji apples using visible and near-infrared spectroscopy[J]. Food and Bioprocess Technology, 2022, 15(10): 2226 – 2236.
[37] ZENG J, GUO Y, HAN Y, et al. A review of the discriminant analysis methods for food quality based on near-infrared spectroscopy and pattern recognition[J]. Molecules, 2021, 26(3): 749.
[38] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700 – 4708.
[39] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks[C]//International Conference on Machine Learning. PMLR, 2019: 6105 – 6114.
[40] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetv3[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314 – 1324.
[41] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132 – 7141.
[42] WOO S, PARK J, LEE J, et al. Cbam: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3 – 19.
[43] LIU Y, SHAO Z, HOFFMANN N. Global attention mechanism: retain information to enhance channel-spatial interactions[J]. arXiv Preprint, arXiv:211205561, 2021.
[44] 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, 2020: 11534 – 11542.
[45] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713 – 13722.

(上接第 91 页)

[14] 蒋易宇, 王硕, 张丽娜, 等. 基于水培生菜力学特征的成熟度分类方法[J]. 农业工程学报, 2023, 39(1): 179 – 187. JIANG Yiyu, WANG Shuo, ZHANG Li'na, et al. Maturity classification method based on the mechanical characteristics of hydroponic lettuce[J]. Transactions of the CSAE, 2023, 39(1): 179 – 187. (in Chinese)
[15] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale[J]. arXiv Preprint, arXiv: 2010.11929, 2020.
[16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017,30:5998 – 6008.
[17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[18] HE T, ZHANG Z, ZHANG H, et al. Bag of tricks for image classification with convolutional neural networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
[19] ANSEL J, YANG E, HE H, et al. PyTorch 2: faster machine learning through dynamic python bytecode transformation and graph compilation[C]//Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, 2024, 5(2):929 – 947.
[20] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Computer Vision-ECCV 2018, Lecture Notes in Computer Science, 2018:3 – 19.
[21] ZHANG W, PANG J, CHEN K, et al. K – Net: towards unified image segmentation[J]. Advances in Neural Information Processing Systems, 2021, 34: 10326 – 10338.
[22] GIRSHICK R. Fast R – CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV), 2015.
[23] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
[24] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019.
[25] QIN D, LEICHNER C, DELAKIS M, et al. MobileNetV4 – universal models for the mobile ecosystem[J]. arXiv Preprint, arXiv:2404.10518, 2024.
[26] 胡松涛,翟瑞芳,王应华,等. 基于多源数据的马铃薯植株表型参数提取[J]. 智慧农业(中英文),2023,5(1):132 – 145. HU Songtao, ZHAI Ruifang, WANG Yinghua, et al. Extraction of potato plant phenotypic parameters based on multi-source data[J]. Smart Agriculture, 2023, 5(1): 132 – 145. (in Chinese)
[27] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018.