

结合三维卷积神经网络和 Haar 小波滤波器的高光谱猕猴桃品质分类

金科¹, 郭志强^{1*}, 曾云流^{2**}, 丁港²

¹武汉理工大学信息工程学院, 湖北 武汉 430070;

²华中农业大学园艺林学学院园艺植物生物学教育部重点实验室, 国家柑橘保鲜技术研发专业中心, 湖北 武汉 430070

摘要 针对猕猴桃硬度品质无损检测分类困难的问题, 提出了结合高光谱成像技术和卷积神经网络的分类模型。该模型融合 Haar 小波核提取的空间特征信息和三维卷积核提取的空谱联合信息, 采用分解数据通道连接的方式确保所有特征能够流到模型末尾, 提升了网络特征提取的能力。通过自制的猕猴桃硬度品质 Kiwi_seed 数据集上的实验表明, Haar 小波变换模块可以显著提升网络的特征提取能力; 通过消融实验表明, 在增加 Haar 小波变换模块后模型分类准确率提升了 7.4%, 最优可达 97.3%, 优于经典的图像分类网络, 可以很好地解决猕猴桃硬度品质的无损检测分类问题。

关键词 高光谱图像; 猕猴桃; 图像分类; 卷积神经网络; Haar 小波变换

中图分类号 TP183

文献标志码 A

DOI: 10.3788/LOP223142

Hyperspectral Imaging-Based Quality Classification for Kiwifruit by Incorporating Three-Dimensional Convolution Neural Network and Haar Wavelet Filter

Jin Ke¹, Guo Zhiqiang^{1*}, Zeng Yunliu^{2**}, Ding Gang²

¹College of Information Engineering, Wuhan University of Technology, Wuhan 430070, Hubei, China;

²Key Laboratory of Horticultural Plant Biology, Ministry of Education, National R&D Center for Citrus Preservation, College of Horticulture and Forestry Sciences, Huazhong Agricultural University, Wuhan 430070, Hubei, China

Abstract To address challenges in the non-destructive inspection and classification of kiwifruit hardness quality, we propose a classification model that incorporates hyperspectral imaging technology and a convolution neural network. This network combines the spatial feature information extracted by the Haar wavelet and the space-spectrum joint information extracted by the three-dimensional (3D) convolution kernel. In this network, the data decomposition of channel connections is executed to ensure that all features can be utilized by the model, which improves the ability of network feature learning. Experiments on the acquired hyperspectral image-based, self-made kiwifruit hardness quality dataset (named Kiwi_seed) demonstrate that the Haar wavelet transform module can significantly improve the feature extraction ability of the network. Ablation experiments reveal that the classification accuracy of the model incorporating the Haar wavelet transform module is increased by 7.4% and reaches the optimum level at 97.3%, which is better than the classical image classification network. The proposed classification model can be effectively used for the non-destructive inspection and classification of kiwifruit quality.

Key words hyperspectral image; kiwifruit; image classification; convolutional neural network; Haar wavelet transform

1 引言

猕猴桃是重要的园艺作物, 富含多种对人体有益

的微量元素, 硬度是猕猴桃重要的品质指标之一, 到手即可食是国产猕猴桃发展的主要方向, 过硬的果实不宜食用, 过软的果实不方便运输销售。经过催熟的猕

收稿日期: 2022-11-23; 修回日期: 2022-12-08; 录用日期: 2022-12-15; 网络首发日期: 2023-01-05

基金项目: 国家重点研发计划(2021YFD1200202-08)、国家自然科学基金(32272779)、猕猴桃质量安全与加工保鲜岗位项目(CARS26)、湖北省重点研发项目(2021BBA090)

通信作者: *guozhiqiang@whut.edu.cn; **zengyl@mail.hzau.edu.cn

猕猴桃,一般要求硬度为 3 kg/cm^2 时停止催熟并出库, $1 \sim 3 \text{ kg/cm}^2$ 时货架销售,并在 $0.3 \sim 1 \text{ kg/cm}^2$ 时为最佳可食用窗口期,因此硬度对猕猴桃的营销和食用质量非常重要^[1]。与一些可通过外观判断成熟的果品(如柑橘、苹果等)不同,猕猴桃等果实的硬度指标很难用肉眼区分,产业中通常采用手指按压法判断硬度,不仅耗时且判断不准确。实验室中, Kiwifirm 是一种较为常用的便携式设备,但该设备为接触式装备,无法批量检测且软果容易破损^[2]。因此寻找一种有效、可靠、非接触式的无损检测猕猴桃硬度的方法对猕猴桃采后储存、运输、分级具有十分重要的意义。

随着高光谱遥感技术的快速发展,传感器能够获得分辨率更高、谱信息更丰富的高光谱图像(HSI)^[3],该技术已在农业^[4]、食品^[5]、医疗^[6]等领域得到广泛的应用,高光谱图像丰富的谱信息和空间信息给农业无损检测和农产品分类带来了机遇。目前,高光谱在农业工程上的应用研究主要包括传统的机器学习方法和兴起的神经网络分类技术研究。

近年来,高光谱图像在农业监测,农产品无损检测方面的应用受到广泛关注,已有很多利用高光谱进行农业监测和无损检测的研究。尚静等^[7]利用高光谱成像技术和机器学习方法进行贵长猕猴桃的硬度检测,实验获取猕猴桃感兴趣区域内的光谱信息,并进行波段选择和建立多元线性回归模型,结果表明模型具有非常好的检测性能。Benelli等^[8]采用高光谱成像技术来评估“海沃德”猕猴桃的成熟度,结合高光谱成像技术和偏最小二乘(PLS)回归模型来估计猕猴桃可溶性固体物含量(SSC)和果肉硬度(FF),进而确定猕猴桃的成熟度,SSC和FF的分类准确率分别可达97%和93%,表明高光谱成像技术适用于猕猴桃品质属性的预测和分类。Wang^[9]指出深度学习技术在农业工程上的应用将是趋势,结合高光谱成像技术和深度学习神经网络进行蓝莓的机械损伤分类,采用残差网络(ResNet)模型对实验数据的验证准确率可达88.44%,验证了结合高光谱成像技术和深度学习技术进行蓝莓损伤鉴定的可行性。以上研究说明,结合神经网络分类技术和高光谱成像技术进行农业无损检测具有很好的应用价值。

传统的机器学习方法只关注到了高光谱图像的平

均光谱信息,并没有充分利用到高光谱图像的空间信息^[10],此外,传统的机器学习方法提取特征依赖于人工经验,提取的特征通常无法保证后续分类器的识别精度。基于高光谱成像技术和深度学习图像分类算法对猕猴桃硬度的品质进行分类,为了增强对HSI空谱特征的提取,提出了一种融合Haar小波变换的3DCNN_SpectralNet网络。该网络利用前置三维卷积神经网络(3DCNN)卷积核提取空谱联合信息,然后利用Haar小波变换增强对空间特征信息的提取,并采用分解数据通道连接的方式确保所有特征流到模型末尾,增强分类的准确率。本文以采集到的猕猴桃高光谱图像作为数据样本,通过特征区域提取和数据块裁剪获得猕猴桃高光谱数据集,最后将模型与其他主流网络进行对比分析并进行消融实验,证明网络模型的有效性。

2 分类模型

本方法针对的是高光谱图像三维数据,首先使用3DCNN方法提取特征,得到包含纹理-光谱联合特征的三维立方体数据;然后将三维数据展开成二维数据,使用Haar小波核和2D卷积核分别进行特征提取,并采用Concat方法拼接Haar小波核分支和2D卷积核分支提取的特征;最后将提取到的特征展开接入全连接层进行猕猴桃品质的分类。本节介绍3DCNN_SpectralNet网络的高光谱图像分类思想,包括3DCNN的基本原理和小波核函数进行特征提取的原理,最后,介绍3DCNN_SpectralNet分类网络的整体设计。

2.1 3DCNN

卷积神经网络(CNN)包括输入层、卷积层、池化层、全连接层和输出层,具有局部连接和权值共享的特点,目前在图像分类、目标检测、目标跟踪等领域应用广泛^[11]。在图像分类和目标检测领域,图像的通道数量较少,应用更多的仍是2DCNN,2DCNN的卷积核是二阶张量,二维卷积过程如图1(a)所示,可以有效的提取图像数据中的空间特征。由于高光谱图像具有丰富的光谱信息,利用二维卷积核提取特征将会破坏高光谱的三维结构信息,导致光谱信息的丢失,而3DCNN在2DCNN的基础上增加了一个维度,适合直接处理三维数据,可以有效提取高光谱图像的空间信

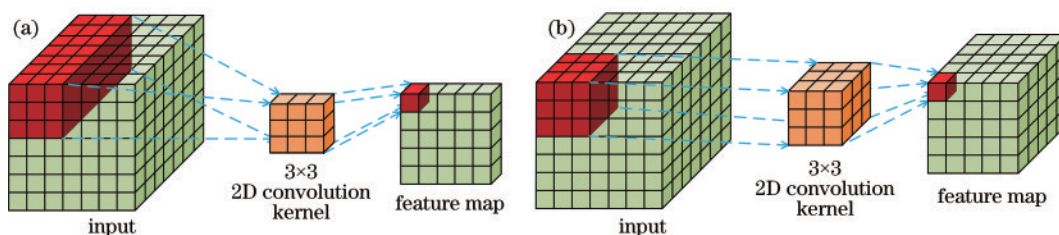


图1 卷积过程。(a)二维卷积;(b)三维卷积

Fig. 1 Convolution process. (a) 2D Convolution; (b) 3D Convolution

息和谱信息,三维卷积过程如图 1(b)所示。

3DCNN 模型采用三维卷积算子,其卷积核可以同时从 3 个方向进行卷积,得到的特征图仍是一个三阶张量。3D-CNN 卷积层的第 i 层、第 j 个特征图的 (x, y, z) 位置的输出表达式为

$$v_{ij}^{xyz} = f \left(\sum_{k=1}^K \sum_{w=1}^{W_i} \sum_{h=1}^{H_i} \sum_{c=1}^{C_i} w_{ijk}^{whc} v_{(i-1)k}^{(x+w)(y+h)(z+c)} + b_{ij} \right), \quad (1)$$

式中: $f(\cdot)$ 为激活函数,本文中使用了 ReLU 激活函数,ReLU 的主要优点为对所有大于零的输入都有固定的导数,加快了网络训练; K 为上一层的特征图数量; W_i, H_i, C_i 分别为卷积核的宽度、高度和光谱维深度; w_{ijk}^{whc} 为第 i 层第 j 个卷积核与上一层第 k 个特征图像连接的卷积核在空间位置 (w, h, c) 处的值; $v_{(i-1)k}^{(x+w)(y+h)(z+c)}$ 为第 $i-1$ 层输出的第 k 个特征图在空间位置 $(x+w, y+h, z+c)$ 处的值; c_{ij} 为第 i 层第 j 个特征图所对应的偏置项。

目前在高光谱遥感领域,已有采用 3DCNN 模型和混合 3D-2DCNN^[12-14] 模型进行高光谱图像分类的研究,在公开的 HSI 数据集上的实验表明 3DCNN 可以对高光谱图像进行有效的空谱信息提取。但简单堆叠 3DCNN 卷积层也存在参数量过多,模型难以收敛,容易过拟合等问题^[15]。

2.2 二维小波特征变换

小波变换是数字图像处理中非常重要的工具,常用于图像的压缩和编码等应用领域^[16]。小波分解是多分辨率分析的有力工具,已被广泛用于提取多分辨率的子带图像,小波变换在时域、频域都能很好地表征信号的局部特征,与人类视觉机理相似^[17]。二维离散小波变换(2D-DWT)的表达式为

$$\begin{cases} \Phi(a, b) = \phi(a)\phi(b) \\ \Psi^H(a, b) = \psi(a)\phi(b) \\ \Psi^V(a, b) = \phi(a)\psi(b) \\ \Psi^D(a, b) = \psi(a)\psi(b) \end{cases}, \quad (2)$$

式中: a 为图像横坐标; b 为图像纵坐标; Ψ^H 为对应列方向的变化,反映了图像的水平边缘信息; Ψ^V 为对应行方向的变化,反映了图像的垂直边缘信息; Ψ^D 为对应角线方向的变化,反映了图像的对角线方向的信息^[16]。式(2)的等式均为 2 个一维函数的乘积,基于 1 个二维尺度函数和 3 个二维小波函数。

由于高光谱图像三维数据的特殊性,以往的研究通常更注重谱维度特征,而忽略了空间特征,二维 Haar 滤波器能够很好的提取图像的空间纹理信息。刘鹏举^[18]在 2D-DWT 中,使用 4 个滤波器用于与特征图卷积实现下采样,以获得 4 个子带特征图 $\{X_1, X_2, X_3, X_4\}$ 。王欣^[19]在运动窗口中嵌入双 Haar 小波变换,进行了信号的局部多尺度分析,发现二维 Haar 小波变换非常适合于图像去噪、边缘检测和边缘增强。计算

小波变换比计算 CNN 卷积更简单,Chakraborty^[20]在 CNN 中融入小波变换对 HSI 公开数据集进行分类证实了 CNN 中融入小波变换的有效性。

在 3DCNN 的基础上增加了一个单独 Haar 滤波分支,通过分解数据通道连接的方式将初始提取的 Haar 特征传输到后续连接层和输出层,增强网络模型的特征提取能力,采用以下 4 个核 $(f_{L,L}, f_{L,H}, f_{H,L}, f_{H,H})$ 进行小波变换:

$$\begin{aligned} f_{L,L} &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, f_{L,H} = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \\ f_{H,L} &= \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, f_{H,H} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \end{aligned} \quad (3)$$

式(3)的 4 个核代表 4 个滤波器,经过 Haar 小波变换的特征输出可表示为

$$\begin{aligned} \text{Haar}(i, j) &= x(2i-1, 2j-1) + x(2i-1, 2j) + \\ & x(2i, 2j-1) + x(2i, 2j), \end{aligned} \quad (4)$$

式中: $\text{Haar}(i, j)$ 为小波变换提取到的特征; $x(i, j)$ 为输入的特征。

2.3 整体网络结构

所提方法包括卷积神经网络、2D-DWT 以及分类器 3 个部分,网络结构如图 2 所示。采集到的 HSI 数据维度为 $800 \times 800 \times 150$,提取感兴趣区域并剪切,得到大小为 $N \times N \times 150$ 的 3D 数据块, $N \times N$ 为剪切窗的大小,本文中窗口选取 30×30 ,小的特征数据块的尺寸,可以减小模型的大小,缩减训练时间。然后,将数据块输入到设计的模型进行训练,类别标签由测定的猕猴桃硬度值进行标定。

模型通过 CNN 从高光谱图像数据块中提取深层空间信息和谱信息,输入数据块的大小为 $30 \times 30 \times 150$,首先经过 2 个 3D 卷积核,由于高光谱数据存在数据维度高且冗余较大的特点,所以第 1 个 3D 卷积核在光谱维度上采用卷积核宽和采样步距相同的方法来去除相邻波段的冗余。第 1 个卷积核的大小为 $(3, 3, 5)$,步距为 $(1, 1, 5)$,这样可以将波段维度由 150 个波段降至 30 个维度。第 2 个卷积核的大小为 $(3, 3, 3)$,步距均为 1,可以进一步提取光谱数据的空间信息和谱信息。

经过 2 层 3D 卷积核后将特征图重组得到新的特征图,对重组的特征图进行 2D-DWT 操作。将 2D-DWT 提取得到的特征和重组后得到的特征图分别经过 2 个 2D 卷积核,得到 $2\text{DCov}_1_x, 2\text{DCov}_3_x$ 和 2DCov_2_x ,将 2DCov_1_x 和 2DCov_2_x 进行拼接得到 Concat_1 , Concat_1 融合了 2D-DWT 提取到的图像空间信息和 2D 卷积核提取到的空谱联合信息; Concat_1 再经过 2 个 2D 卷积核得到 2DCov_4_x ,将 2DCov_3_x 和 2DCov_4_x 再次进行拼接得到 Concat_2 ,通过这种跨层连接的方式将 2D-DWT 得到的图像空间特征向后续特征提取层传递; Concat_2 经过 3 个 2D 卷积核后得到 2DCov_5_x ,进一步实现了特征提取。

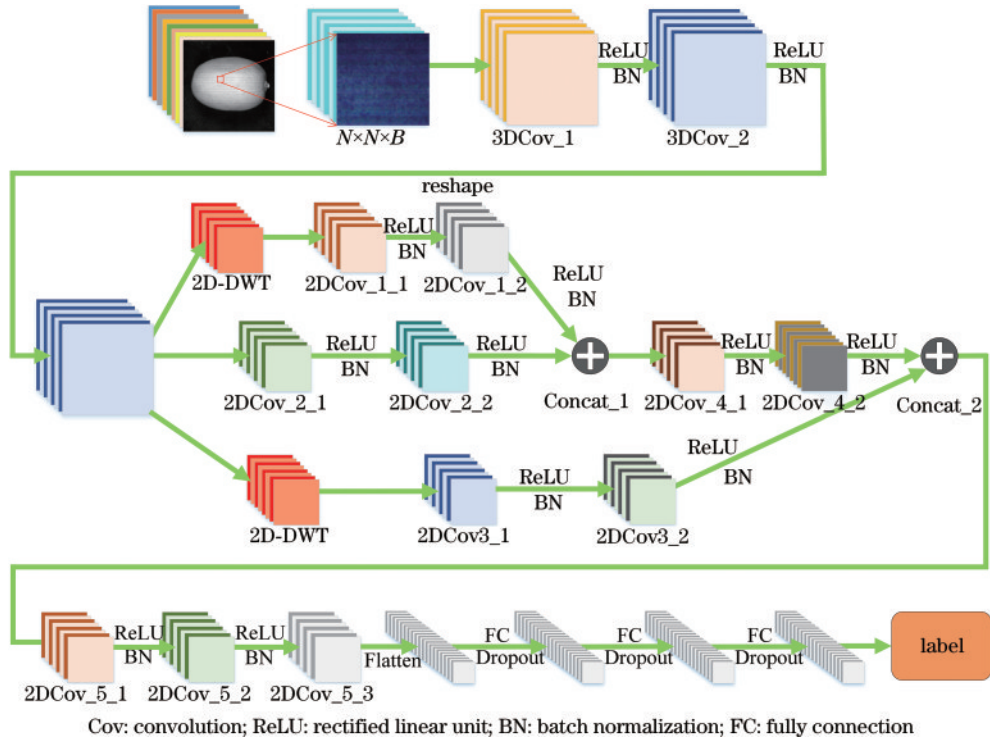


图 2 整体网络结构

Fig. 2 Overall network structure

在卷积层之后将得到的特征图展开并输入到全连接层,然后将全连接层输出给 softmax 激活函数得到最终的分标签。本文中使用 ReLU 作为激活函数,该函数具有收敛速度快的优点。为了防止过拟合,在每次激活函数之前进行批量归一化操作,并在全连接层

采用 Dropout 正则化方式,在训练过程中随机去除隐藏层。

模型中各层的输入尺寸、输出尺寸、卷积核尺寸、步距以及使用的 Padding 方式,如表 1 所示,其中 s 为步长。整个模型的参数量为 15.21 MB。

表 1 各模块参数

Table 1 Parameters of each module

Layer name	Input size	Output size	Kernal_size	Padding
3DCov_1	30×30×150×1	28×28×30×8	3×3×5, 8, s=1×1×5	
3DCov_2	28×28×30×8	26×26×28×1	3×3×3, 1, s=1	
2D-DWT	26×26×28	13×13×112		
2DCov_1_x	13×13×112	13×13×128	$\begin{bmatrix} 3 \times 3, 128, s=1 \\ 3 \times 3, 128, s=1 \end{bmatrix}$	'same'
2DCov_2_x	26×26×28	13×13×128	$\begin{bmatrix} 3 \times 3, 64, s=1 \\ 3 \times 3, 128, s=2 \end{bmatrix}$	'same'
2DCov_3_x	13×13×112	13×13×128	$\begin{bmatrix} 3 \times 3, 128, s=1 \\ 3 \times 3, 128, s=1 \end{bmatrix}$	'same'
2DCov_4_x	13×13×256	13×13×128	$\begin{bmatrix} 3 \times 3, 256, s=1 \\ 3 \times 3, 128, s=1 \end{bmatrix}$	'same'
2DCov_5_x	13×13×256	9×9×64	$\begin{bmatrix} 3 \times 3, 256, s=1 \\ 3 \times 3, 128, s=1 \\ 3 \times 3, 64, s=1 \end{bmatrix}$	$\begin{bmatrix} \text{'same'} \\ - \\ - \end{bmatrix}$

3 实验与分析

3.1 实验数据

本文使用自制的猕猴桃高光谱图像数据集,实验材料为武汉市某奇异果园新采摘的“云海一号”品种猕猴桃,实验选取了 500 个大小相近,表面无损伤的猕猴桃样本,采摘后猕猴桃需在室内静置 12 h 以去除田间热并用二氧化氯药剂处理以避免软腐病的影响^[21]。高光谱图像采集设备为比利时 IMEC 公司制造的便携式高光谱图像仪,工作波段范围为 470~900 nm,最大分辨率为 3650 pixel×2048 pixel,采集到的光谱波段数为 150 个,猕猴桃硬度采集设备为 GY-4 数显水果硬度计(托普仪器)。数据采集时选取 800 pixel×800 pixel 的像素分辨率,先进行校正、调焦,保证图像清晰,修改相机参数积分时间为 4.0 ms,模拟增益为 1.6,高光谱图像采集前,先将高光谱成像仪和光源预热 10 min 左右,且在拍摄温度超过 45 °C 时需关闭成像仪等待降温。数据采集分 9 d 进行,每天选取 50~60 个猕猴桃,提前标好序号,按序号取猕猴桃正放在镜头下进行扫描,扫描完成后标记好拍摄面以方便后续测定拍摄面的硬度数据。

深度学习模型需要大量的数据作为训练集,但高光谱图像的采集较为费时费力,无法快速大量地采集高光谱图像,所以需要进行数据集的创建和扩充。高光谱图像数据占用存储空间大,包含的谱信息丰富,基于高光谱图像数据量丰富的特点,本文数据集扩充的思路是

进行数据块的裁切。

为了去除背景信息,得到有效信息,需要提取感兴趣区域。获取感兴趣区域掩模图像的步骤如图 3 所示。

首先,手动裁取猕猴桃果肉和背景区域的平均反射信息,绘制出平均反射光谱曲线,绘制出背景区域、边缘区域、中心区域的平均反射率,如图 4 所示。由于猕猴桃果实的结构呈椭圆形,由图 4 可知,猕猴桃边缘区域和中心区域的平均反射率存在明显的差异,选取感兴趣区域掩模时,需要集中选取中心反射光谱强的区域,略去边缘区域反射率小的区域。同时,由图 4 可知,背景区域的反射率要远小于果实区域的反射率,选取果肉与背景平均反射率差值最大的波长作为特征波段(868 nm),得到特征波段下的灰度图像。

然后,将特征波长下的灰度图依次进行中值滤波、腐蚀运算、阈值分割、开运算及再次腐蚀运算,获取猕猴桃中心区域的感兴趣区掩模图像。在获取掩模图像的过程中采用中值滤波的目的是去除背景噪点,需要注意的是,由于猕猴桃椭圆形的特点,果实中心区域和边缘区域的反射率存在差异,采用两次腐蚀运算(采用椭圆形的结构元)的方法缩小感兴趣区域,以减小对猕猴桃边缘的选取,集中选取中心区域。

在裁切得到的掩模图像的基础上进行图像分割,在掩模区域内选取合适的窗大小对原始 HSI 数据进行不重叠切块,使用的切割窗宽大小为 30 pixel×30 pixel。实验共采集到 472 张高光谱图像数据,经过

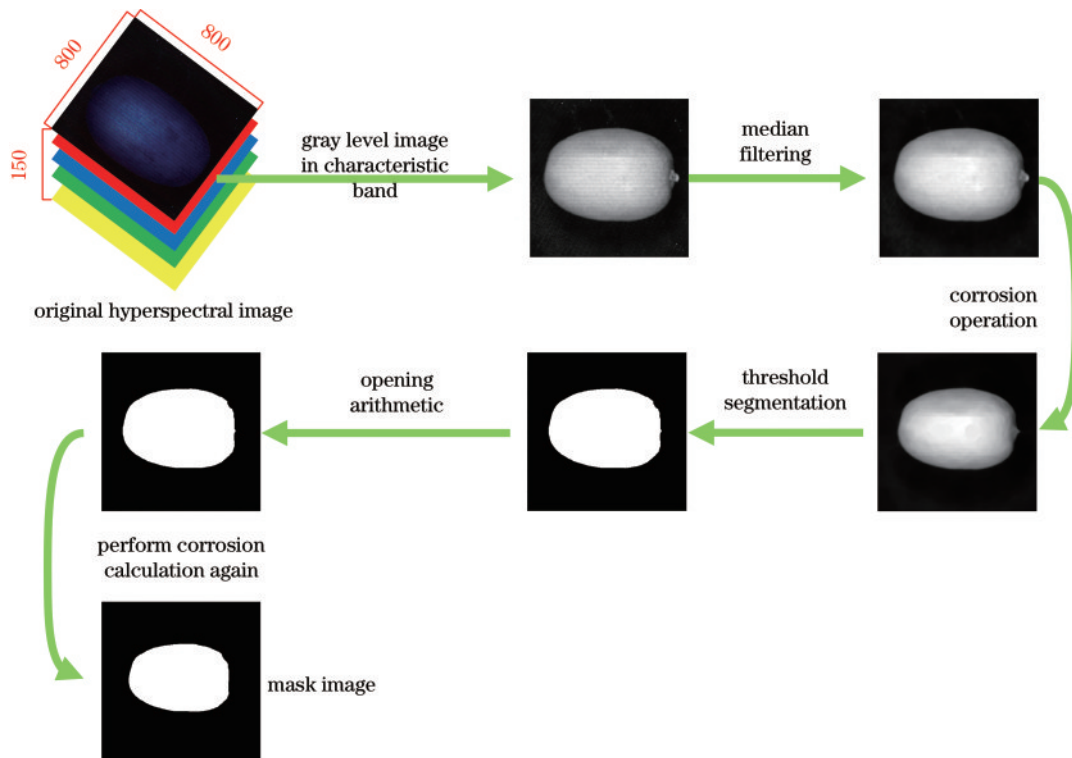


图 3 获取感兴趣区域掩模的步骤

Fig. 3 Steps of acquiring region of interest mask

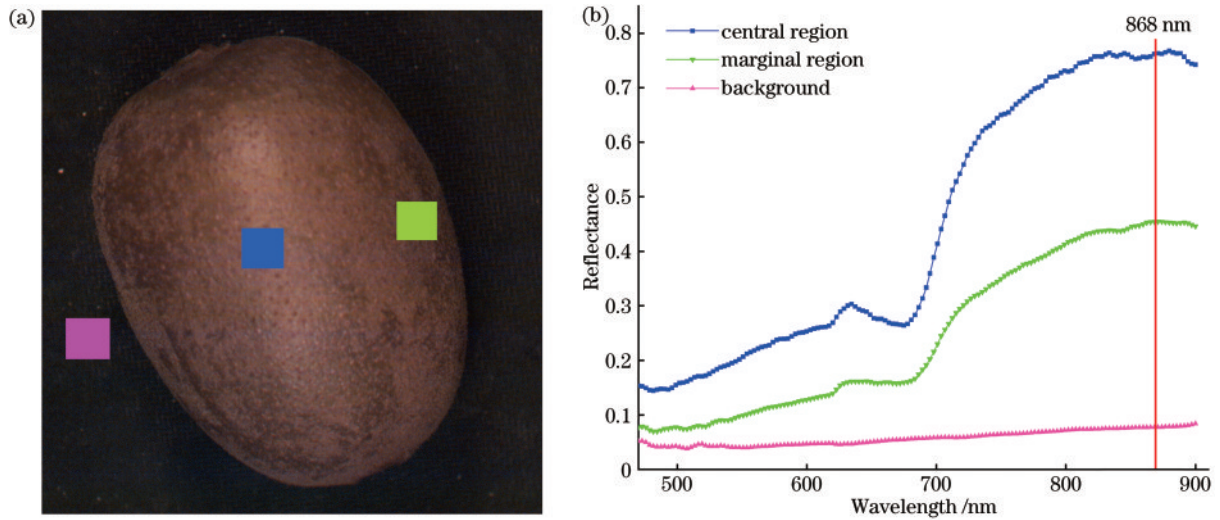


图 4 中心、边缘和背景区域的平均反射率。(a)采样区域;(b)平均反射光谱曲线

Fig. 4 Average reflectivity of center, edge and background area. (a) Sampling area; (b) average reflectance spectra curve

分割共得到 70416 个数据块,达到了数据扩充的目的。在制作数据集过程中与农业领域的专家共同确定成熟果,半熟果和未熟果的分类标准。将硬度值大于 3 kg/cm^2 的猕猴桃定为未熟果,这部分果实硬度值较大,对应猕猴桃的运输期;将硬度值介于 $1 \sim 3 \text{ kg/cm}^2$ 的猕猴桃定为半熟果,这部分果实的硬度值适中,对应猕猴桃的货架期;将硬度值小于 1 kg/cm^2 的猕猴桃定为成熟果,对应猕猴桃的可食用期。将制作的数据集命名为 Kiwi_seed,采集到的硬度值统计信息及各类别的数据块个数,如表 2 所示。

表 2 数据集统计信息
Table 2 Dataset Statistics

Category	Mature group	Semimature group	Immature group
Number of fruit	228	164	101
Number of data block	31540	23746	15130

3.2 实验参数设置

实验在 Inter Xeon CPU E5-2650 v3@2.3 GHz, 128 GB 内存, NVIDIA Tesla K40c GPU, 64 位 Windows 10 操作系统, Python 3.6 和 TensorFlow 2.0 框架下进行。使用 Kiwi_seed 数据集对模型进行训练和验证,设定训练集、验证集和测试集的划分比例为 7:2:1。选择模型训练的轮数为 200 轮,批次大小设置为 512,使用 Adam 优化器进行优化计算。为了解决模型收敛慢的问题,实验中采用动态学习率,在前 30 轮训练中,学习率设置为 0.001,保证了模型快速收敛;在训练轮数达到 30 轮后,每轮的学习率减少为上一轮的 0.95,这样保证了模型能够接近最优训练效果。

3.3 实验结果与分析

为了验证本文方法的分类精度,使用 Kiwi_seed 数据集对模型进行训练和验证,分别与图像分类领域经典的网络模型 ResNet^[22], GoogleNet^[23] 和近三年表现

优异的 EfficientNet^[24] 以及 HSI 地物分类领域广泛关注的 HybridSN^[12] 模型进行对比实验。在实验过程中,由于本文的数据尺寸与经典网络的输入尺寸并不匹配,需要修改经典网络的输入尺寸以匹配 Kiwi_seed 数据集的数据尺寸。在 ResNet34、GoogLeNet、EfficientNet 这些网络中,使用了大量的降采样步骤降低特征图的尺寸,本文的对比试验中去掉了这些池化层降采样部分,并在 ResNet34、GoogLeNet、EfficientNet 模型前端添加 3DCNN 卷积部分进行输入数据尺寸的匹配。

为了更好的比较模型的性能,绘制出了实验训练过程中每个网络在每个训练轮次下的准确率(Acc)和损失(Loss)的变化曲线来进行比较。不同模型在训练集和验证集上的 Acc 曲线和 Loss 曲线,如图 5 所示。由图 5 可知,提出的 3DCNN_SpectralNet 网络在训练集上能够得到更高的准确率,且训练过程中的收敛速度更快,这说明模型能够更有效地学习到目标特征,对空间信息和谱信息的提取更加充分,加快了网络的训练速度。

为了更准确地衡量网络模型分类的准确率,实验选择分类 Acc、精确率(PPV)、召回率(R)和 F1 分数(F1)对测试集进行评价。各个模型的 PPV、R、F1 实验结果如表 3 所示。由表 3 可知,所提网络模型在测试集上准确率可达 97.3%,为所有模型中最佳的准确率,每个类别的 PPV、R、F1 表明:所提模型对每个类别都可以进行较为准确的区分,采用该模型可以有效实现猕猴桃果实的硬度品质分类。

为了验证神经网络端到端的特征提取方式的效果,设置了另一组对比实验,比较传统的特征波段选择方式进行特征波段的选择和神经网络端到端方式特征提取的效果。在对比实验中,丢弃了前置的 3DCNN 特征提取部分,而改用传统的特征波段选择的方式,将波段选择后的数据输入到后续网络进行训练。对照组

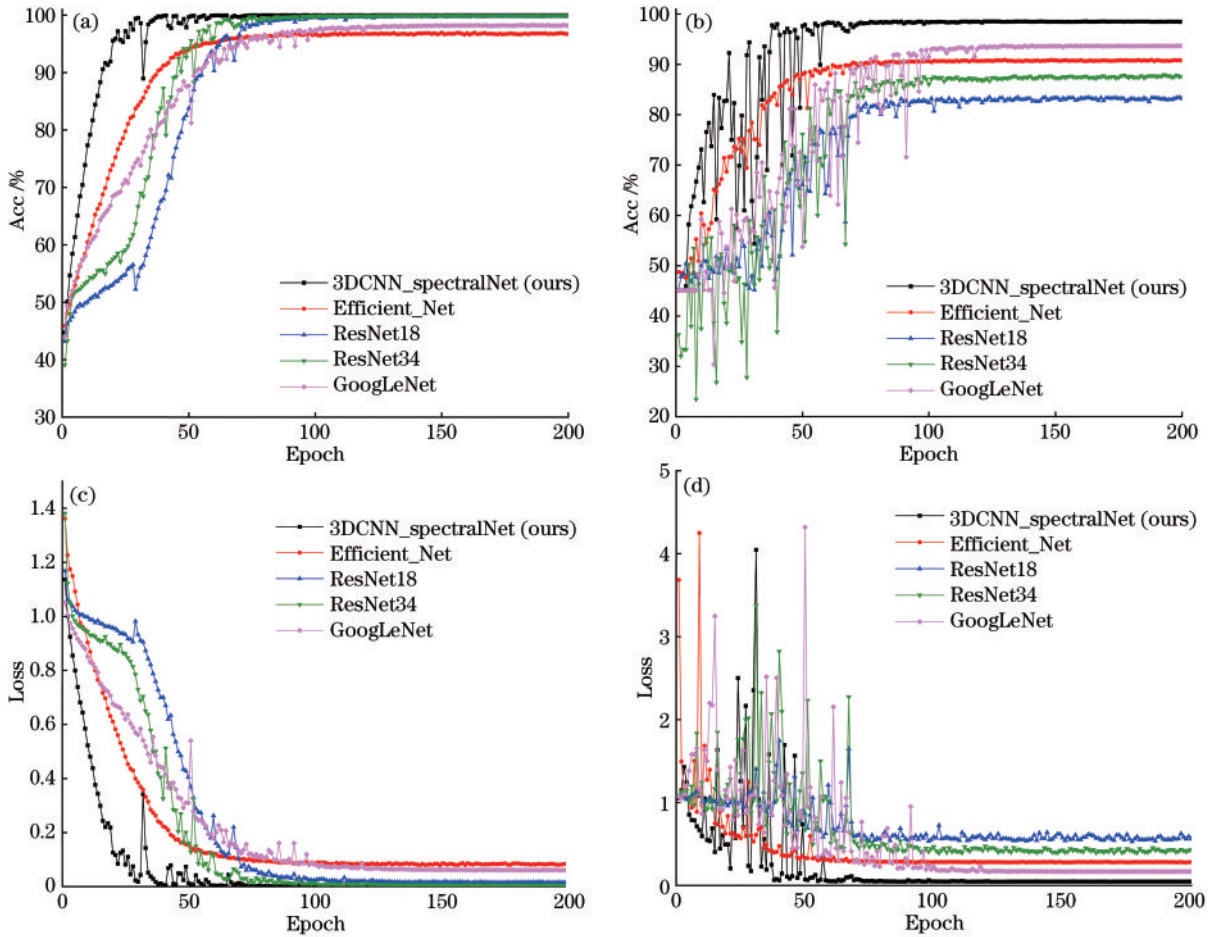


图 5 不同模型的准确率和损失曲线。(a)训练集上的准确率曲线;(b)验证集上的准确率曲线;(c)训练集上的损失曲线;(d)验证集上的损失曲线

Fig. 5 Accuracy and loss curves of different models. (a) Accuracy curves on training set; (b) accuracy curves on validation set; (c) loss curves on training set; (d) loss curves on validation set

表 3 不同模型在测试集上的分类结果

Table 3 Classification results of different models on test sets

Model	Acc	Mature kiwi			Half_Mature kiwi			Unripe kiwi		
		PPV	R	F1	PPV	R	F1	PPV	R	F1
EfficientNet_B0	0.909	0.932	0.942	0.937	0.879	0.904	0.892	0.909	0.847	0.877
GoogLeNet	0.932	0.948	0.953	0.951	0.919	0.921	0.920	0.916	0.900	0.907
ResNet18	0.824	0.857	0.895	0.876	0.793	0.787	0.790	0.797	0.731	0.763
ResNet34	0.879	0.906	0.918	0.912	0.858	0.852	0.855	0.854	0.840	0.847
HybridSN	0.750	0.820	0.844	0.831	0.702	0.664	0.682	0.676	0.691	0.684
3DCNN_SpectralNet	0.973	0.977	0.980	0.979	0.968	0.969	0.969	0.970	0.961	0.965

实验利用竞争性自适应重加权采样法(CARS)^[25]进行波段选择,将选择后的图像数据进行切块并输入分类器得到分类结果。CARS是一种采用“适者生存”原则的波段选择方法,常用于光谱波段的选择,由于CARS选择波段具有随机性和不确定性,这会造成特征波段的漏选或冗余问题,因此实验中进行多次CARS,并将结果进行统计筛选,本实验采用执行50次CARS,统计挑选出被选择次数排名前30的波段输入后续网络进行训练。

对照组模型(CARS_SpectralNet)和所提模型在测试集上的混淆矩阵,如图6所示。2个模型的Acc、PPV、R、F1指标结果,如表4所示。由图6、表4可知,直接采用神经网络进行端到端特征提取并分类的方式效果比采用波段选择后输入神经网络分类器的方式效果好,采用端到端的特征提取方式比采用CARS提取后再分类的方式总体准确率提升了5.6%,体现了采用3DCNN进行光谱信息和空间信息提取的良好效果。

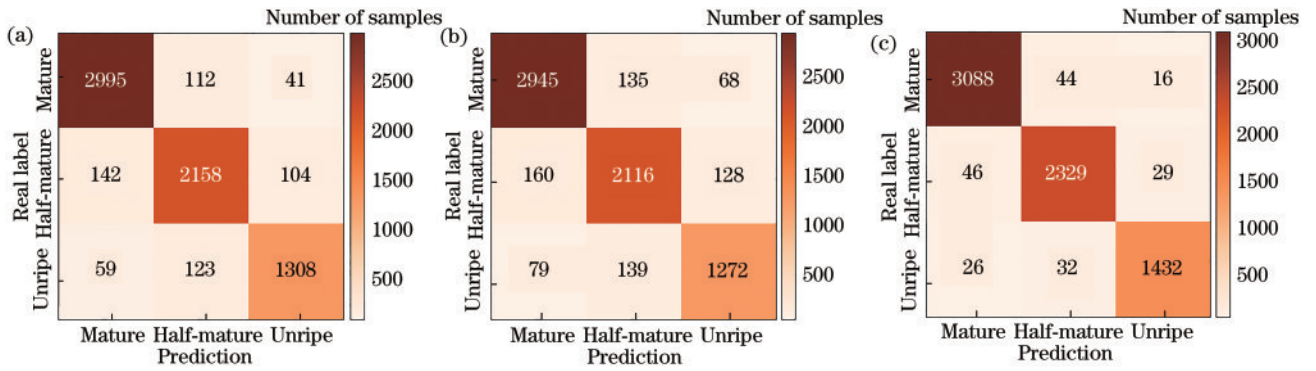


图 6 对照组模型的混淆矩阵。(a) CARS_SpectralNet模型的混淆矩阵;(b) Without_2D-DWT模型的混淆矩阵;(c) 3DCNN_SpectralNet模型的混淆矩阵

Fig. 6 Confusion matrix of control group. (a) Confusion matrix of CARS_SpectralNet; (b) confusion matrix of Without_2D-DWT; (c) confusion matrix of 3DCNN_SpectralNet

表 4 对照组实验结果

Table 4 Experimental results of control group

Model	Acc	Mature kiwi			Half_Mature kiwi			Unripe kiwi		
		PPV	R	F1	PPV	R	F1	PPV	R	F1
CARS_SpectralNet	0.917	0.937	0.951	0.944	0.902	0.898	0.900	0.900	0.878	0.889
Without_2D-DWT	0.899	0.936	0.925	0.930	0.880	0.885	0.883	0.854	0.866	0.860
3DCNN_SpectralNet	0.973	0.977	0.980	0.979	0.968	0.969	0.969	0.970	0.961	0.965

为了验证 Haar 小波变换提取空间特征的效果,在 2.3 节网络结构中去除了 2D-DWT 结构并进行实验,图 6 绘制出了网络模型去除 2D-DWT 结构后 (without_2D-DWT) 在验证集上的混淆矩阵,表 4 给出了 without_2D-DWT 模型在验证集上的实验结果。可以发现增加了 2D-DWT 结构后效果提升明显,准确率由 89.9% 提升至 97.3%。

利用 t-SNE 数据可视化的方法在测试集上对

3DCNN_SpectralNet 提取到的特征和 without_2D-DWT 提取到的特征进行降维可视化的结果,如图 7 所示。由图 7 可知,Without_2d-DWT 模型提取到的特征之间有聚类效果但存在着混叠,可分性并不强。增加了 Haar 小波特征变换模块后的 3DCNN_SpectralNet 模型提取特征的能力有了明显提升,不同类别提取到的特征可分性明显增强,体现了 Haar 小波变换提取空间信息的良好效果。

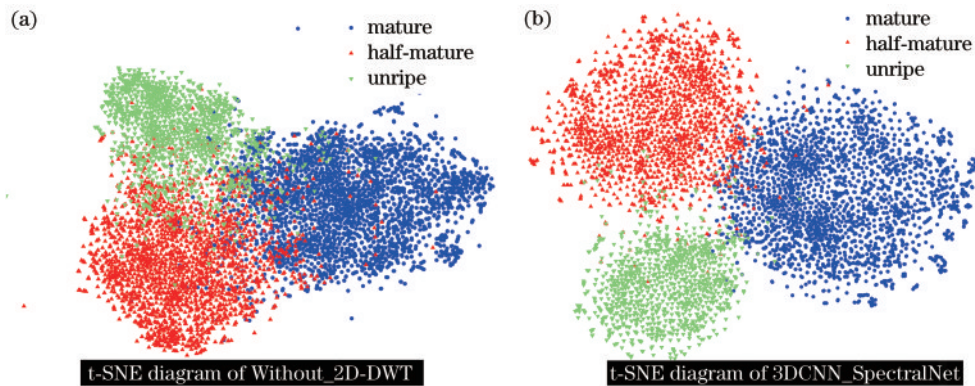


图 7 测试集上不同模型的 t-SNE 示意图。(a) Without_2D-DWT 模型提取特征的 t-SNE 示意图;(b) 3DCNN_SpectralNet 模型提取特征的 t-SNE 示意图

Fig. 7 t-SNE diagram of different models on the test set. (a) t-SNE diagram of features extracted from Without_2D-DWT model; (b) t-SNE diagram of features extracted from 3DCNN_SpectralNet model

4 结 论

结合高光谱成像技术和卷积神经网络分类算法解决猕猴桃硬度品质分类的问题。针对高光谱图像光谱

分辨率高、谱信息丰富的特点,提出了一种融合 Haar 小波特征和 3DCNN 的 3DCNN_SpectralNet 网络。首先用三维卷积层提取高光谱图像数据块的空谱联合特征;然后利用 Haar 小波变换增强对空间特征信息的提

取,并分解数据的通道连接确保所有特征流到模型末尾,经全连接层对特征进行展平重组;最后通过输出层完成分类。

本研究提出的网络模型的收敛速度更快,相较于传统图像分类网络和高光谱分类领域广受关注的 HybridSN 模型,提出的方法在准确率、召回率和 F1 分数指标上更加优秀,这是由于 Haar 小波滤波器能够提取到更有效的空间特征,分解数据通道连接的方式也能够更有效地保证所有特征接入模型末尾。采用 3DCNN 卷积核可以有效提取高光谱图像的空谱信息,且端到端的特征提取方式要优于传统的波段选择方式进行特征选择后再分类的方式,采用 CARS 方法进行波段选择后再分类的方法准确率为 91.7%,而直接使用 CNN 进行特征提取并分类的 SpectralNet 方式准确率可以达到 97.3%,提升效果明显;增加 Haar 小波变换部分可以有效的增加模型的特征提取能力,消融实验表明增加 Haar 小波变换模块后提取到的特征可分性明显变好,证明了利用 Haar 小波核进行空间特征提取的良好效果。

在研究高光谱图像分类算法的基础上,对猕猴桃硬度品质分类进行了实际考察和研究,对现有的高光谱图像分类算法进行总结和改进,提出利用神经网络技术进行猕猴桃的硬度品质分类。为了增强模型的特征提取能力,将 Haar 小波特征变换融合到卷积神经网络中,并通过实验证明了模型提取特征的能力。实验结果表明了利用卷积神经网络技术和高光谱成像技术可以较好地完成猕猴桃硬度品质的分类问题,该方法可为实现猕猴桃硬度的无损检测提供有效指导,具有一定的理论和实践价值。

参 考 文 献

- [1] Burdon J, Lallu N. Kiwifruit (*actinidia* spp.)[M]//Yahia E M. Postharvest biology and technology of tropical and subtropical fruits. Amsterdam: Elsevier, 2011: 326-362.
- [2] Li H J, PidakalaPaul, BillingDavid, et al. Kiwifruit firmness: measurement by penetrometer and non-destructive devices[J]. Postharvest Biology and Technology, 2016, 120: 127-137.
- [3] 李非燕, 霍宏涛, 李静, 等. 基于多特征和改进稀疏表示的高光谱图像分类[J]. 光学学报, 2019, 39(5): 0528004.
Li F Y, Huo H T, Li J, et al. Hyperspectral image classification via multiple-feature-based improved sparse representation[J]. Acta Optica Sinica, 2019, 39(5): 0528004.
- [4] 吴永清, 李明, 张波, 等. 高光谱成像技术在谷物品质检测中的应用进展[J]. 中国粮油学报, 2021, 36(5): 165-173.
Wu Y Q, Li M, Zhang B, et al. Application of hyperspectral imaging technology in grain quality detection[J]. Journal of the Chinese Cereals and Oils Association, 2021, 36(5): 165-173.
- [5] 赵婕秀, 董庆利, 陈培琴, 等. 高光谱成像技术在动物源性食品微生物污染快速检测中的应用进展[J]. 食品工业科技, 2022, 43(7): 467-473.
Zhao J X, Dong Q L, Chen P Q, et al. Application progress of hyperspectral imaging technology in rapid detection of microbial contamination in animal derived food[J]. Science and Technology of Food Industry, 2022, 43(7): 467-473.
- [6] 李伟, 吕蒙, 陈天虹, 等. 高光谱图像在生物医学中的应用[J]. 中国图象图形学报, 2021, 26(8): 1764-1785.
Li W, Lü M, Chen T H, et al. Application of a hyperspectral image in medical field: a review[J]. Journal of Image and Graphics, 2021, 26(8): 1764-1785.
- [7] 尚静, 冯树南, 谭涛, 等. 基于高光谱成像的贵长猕猴桃硬度快速预测[J/OL]. 食品工业科技: 1-10[2022-10-20]. DOI:10.13386/j.issn1002-0306.2022060168.
Shang J, Feng S N, Tan T, et al. Rapid prediction for the firmness of Guichang Kiwifruit by hyperspectral imaging[J/OL]. Science and Technology of Food Industry: 1-10[2022-10-20]. DOI: 10.13386/j.issn1002-0306.2022060168.
- [8] Benelli A, Cevoli C, Fabbri A, et al. Ripeness evaluation of kiwifruit by hyperspectral imaging[J]. Biosystems Engineering, 2022, 223: 42-52.
- [9] Wang Z D, Hu M H, Zhai G T. Application of deep learning architectures for accurate and rapid detection of internal mechanical damage of blueberry using hyperspectral transmittance data[J]. Sensors, 2018, 18(4): 1126.
- [10] 邵园园, 王永贤, 玄冠涛, 等. 高光谱成像的猕猴桃货架期快速预测[J]. 光谱学与光谱分析, 2020, 40(6): 1940-1946.
Shao Y Y, Wang Y X, Xuan G T, et al. Hyperspectral imaging technique for estimating the shelf-life of kiwifruits [J]. Spectroscopy and Spectral Analysis, 2020, 40(6): 1940-1946.
- [11] 张号逵, 李映, 姜晔楠. 深度学习在高光谱图像分类领域的研究现状与展望[J]. 自动化学报, 2018, 44(6): 961-977.
Zhang H K, Li Y, Jiang Y N. Deep learning for hyperspectral imagery classification: the state of the art and prospects[J]. Acta Automatica Sinica, 2018, 44(6): 961-977.
- [12] Ghaderizadeh S, Abbasi-Moghadam D, Sharifi A, et al. Hyperspectral image classification using a hybrid 3D-2D convolutional neural networks[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 14: 7570-7588.
- [13] Xu Q, Xiao Y, Wang D Y, et al. CSA-MSO3DCNN: multiscale octave 3D CNN with channel and spatial attention for hyperspectral image classification[J]. Remote Sensing, 2020, 12(1): 188.
- [14] Zheng J W, Feng Y C, Bai C, et al. Hyperspectral image classification using mixed convolutions and covariance pooling[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(1): 522-534.
- [15] 李冠东, 张春菊, 高飞, 等. 双卷积池化结构的 3D-CNN 高光谱遥感影像分类方法[J]. 中国图象图形学报, 2019, 24(4): 639-654.

- Li G D, Zhang C J, Gao F, et al. Doubleconvpool-structured 3D-CNN for hyperspectral remote sensing image classification[J]. *Journal of Image and Graphics*, 2019, 24(4): 639-654.
- [16] 李文娟. 基于局部特征提取的人脸识别方法研究[D]. 天津: 天津大学, 2017.
- Li W J. A study of face recognition based on local feature extraction[D]. Tianjin: Tianjin University, 2017.
- [17] 刘望华, 刘光帅, 陈晓文, 等. 结合微分特征和 Haar 小波分解的鲁棒纹理表达[J]. *计算机应用*, 2020, 40(9): 2728-2736.
- Liu W H, Liu G S, Chen X W, et al. Robust texture representation by combining differential feature and Haar wavelet decomposition[J]. *Journal of Computer Applications*, 2020, 40(9): 2728-2736.
- [18] Liu P J, Zhang H Z, Zhang K, et al. Multi-level wavelet-CNN for image restoration[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 18-22, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 886-88609.
- [19] Wang X. Moving window-based double haar wavelet transform for image processing[J]. *IEEE Transactions on Image Processing*, 2006, 15(9): 2771-2779.
- [20] Chakraborty T, Trehan U. SpectralNET: exploring spatial-spectral WaveletCNN for hyperspectral image classification[EB/OL]. (2021-04-01)[2022-08-09]. <https://arxiv.org/abs/2104.00341>.
- [21] 王亚萍, 郭叶, 费学谦. 二氧化氯处理对“徐香”猕猴桃贮藏品质的影响[J]. *西北林学院学报*, 2014, 29(3): 151-154.
- Wang Y P, Guo Y, Fei X Q. Impacts of chlorine dioxide treatment on preservation quality of kiwifruit[J]. *Journal of Northwest Forestry University*, 2014, 29(3): 151-154.
- [22] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [23] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1800-1807.
- [24] Tan M X, Le Q V. EfficientNet: rethinking model scaling for convolutional neural networks[C]//International Conference on Machine Learning, June 9-15, 2019, Long Beach, California, USA. Copenhagen: MLR Press, 2019: 6105-6114.
- [25] 许丽佳, 陈铭, 王玉超, 等. 高光谱成像的猕猴桃糖度无损检测方法[J]. *光谱学与光谱分析*, 2021, 41(7): 2188-2195.
- Xu L J, Chen M, Wang Y C, et al. Study on non-destructive detection method of kiwifruit sugar content based on hyperspectral imaging technology[J]. *Spectroscopy and Spectral Analysis*, 2021, 41(7): 2188-2195.