

# 基于MobileNetV3的陈皮年份鉴别

陈康辉, 陈开权, 郑成勇\*, 梁桓绵

五邑大学数学与计算科学学院, 广东 江门

收稿日期: 2025年5月14日; 录用日期: 2025年6月13日; 发布日期: 2025年6月20日

## 摘要

陈皮(学名柑橘皮)的贮藏年限与其品质特征及市场价值呈现显著正相关性。然而, 传统的陈皮年份鉴别方法主要依赖经验判断或化学分析, 存在主观性强、操作复杂、耗时耗力等诸多局限。为此, 本文提出一种基于轻量化深度学习框架MobileNetV3的陈皮年份智能鉴别方法。通过构建涵盖不同贮藏年限的陈皮图像数据集, 采用区域裁剪对原始图像样本进行标准化处理, 生成分辨率为 $224 \times 224$ 像素的黑色背景规范图像。在数据预处理与增强阶段, 应用像素值归一化、随机旋转等策略, 以提升模型的泛化能力与鲁棒性。实验环节对比了MobileNetV3、ResNet、DenseNet、FasterNet、AlexNet、VGG16以及EfficientFormerV2共七种经典与轻量级卷积神经网络模型。结果表明, MobileNetV3-large在模型参数量仅为16.3 MB的前提下, 实现了86.96%的分类准确率与87.55%的精确率, 单张图像推理时间仅为5.36毫秒, 综合性能显著优于其他对比模型。总之, 本文提出了一种更为简易且准确的陈皮年份识别方法, 为在边缘计算设备的部署提供了高效可靠的技术支持。

## 关键词

陈皮年份鉴别, 轻量化, 区域裁剪, MobileNetV3

# Tangerine Peel Age Identification Based on MobileNetV3

Kanghui Chen, Kaiquan Chen, Chengyong Zheng\*, Huanmian Liang

School of Mathematics and Computational Sciences, Wuyi University, Jiangmen Guangdong

Received: May 14<sup>th</sup>, 2025; accepted: Jun. 13<sup>th</sup>, 2025; published: Jun. 20<sup>th</sup>, 2025

## Abstract

The storage age of dried tangerine peel (*Citrus reticulata* peel) is significantly positively correlated with its quality characteristics and market value. However, traditional methods for identifying the

\*通讯作者。

age of dried tangerine peel mainly rely on expert judgment or chemical analysis, which are subjective, complex, time-consuming, and labor-intensive. To address these limitations, this paper proposes an intelligent method for identifying the age of dried tangerine peel based on the lightweight deep learning framework MobileNetV3. A dataset of tangerine peel images with varying storage ages was constructed, and regional cropping was used to standardize the original image samples, generating standardized images with a resolution of  $224 \times 224$  pixels and a black background. During the data preprocessing and augmentation phase, strategies such as pixel value normalization and random rotation were applied to enhance the model's generalization ability and robustness. The experimental section compares seven classic and lightweight convolutional neural network models, including MobileNetV3, ResNet, DenseNet, FasterNet, AlexNet, VGG16, and EfficientFormerV2. The results show that MobileNetV3-large achieves a classification accuracy of 86.96% and a precision of 87.55%, with a single image inference time of only 5.36 milliseconds, while having only 16.3 MB of model parameters. Its overall performance significantly outperforms the other models. In conclusion, this paper presents a simpler and more accurate method for identifying the age of dried tangerine peel, providing efficient and reliable technical support for deployment on edge computing devices.

## Keywords

Tangerine Peel Age Identification, Lightweight, Regional Cropping, MobileNetV3

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

陈皮是芸香科植物橘(*Citrus reticulata* Blanco)及其栽培变种的成熟果皮,经干燥后所得的中药材。根据产地及加工工艺的不同,主要分为“陈皮”和“广陈皮”两类[1]。传统的陈皮加工方法通常包括果实的适时采摘、果皮剥离以及采用自然晾晒或低温干燥等方式进行脱水处理。陈皮因其理气健脾、燥湿化痰等显著药理作用,在中医临床中被广泛应用。同时,凭借其独特的香气与风味,陈皮也在食品加工、调味品制造及日常烹饪等领域发挥着重要作用。近年来,随着中医药保健理念的普及和陈皮多元化应用场景的拓展,其市场需求持续增长。尤以贮藏时间较长、具有年份沉淀的陈皮最为珍贵,因其药效成分更为丰富、香气更为浓郁,受到消费者及相关产业的高度关注。目前,陈皮贮藏年份的主流鉴别方法主要包括外观特征分析与化学成分检测两类。前者依据果皮的颜色、纹理与质感等感官指标进行初步判断,但该方法主观性较强,易受操作者经验和环境条件的影响,导致准确性和重复性较低;后者则通过检测黄酮类、挥发油等关键成分的含量变化,来推测其贮藏时间[2][3]。尽管化学检测法在提升鉴别准确性方面具有一定优势,但由于其依赖高精度仪器与专业操作人员,检测成本较高,难以在基层单位及商品流通环节实现广泛应用。

针对上述年份鉴别方法存在的局限性,本文提出了一种基于 MobileNetV3 深度学习模型的陈皮年份识别方法。该方法旨在充分发挥深度卷积神经网络在图像特征提取方面的优势,有效提升了年份分类的准确性与处理效率。MobileNetV3 作为一种优化后的轻量级神经网络结构,在保持良好分类性能的同时大幅压缩了模型参数规模与计算开销,特别适用于移动端部署或计算资源受限的实际应用场景,例如在 Android 平台上的边缘设备运行[4]。为全面验证所提出方法的有效性与应用前景,本文以 MobileNetV3 作为核心模型,并选取 ResNet、DenseNet 和 FasterNet 等典型深度学习架构作为对比组,对其在陈皮年份

识别任务中的性能进行系统性评估。评估体系涵盖多项关键性能指标，包括损失函数值、分类精度、准确率(Accuracy)、召回率(Recall)、F1 分数，以及浮点运算次数(FLOPs)等，以从模型分类准确性与计算效率两个维度进行多角度综合分析。

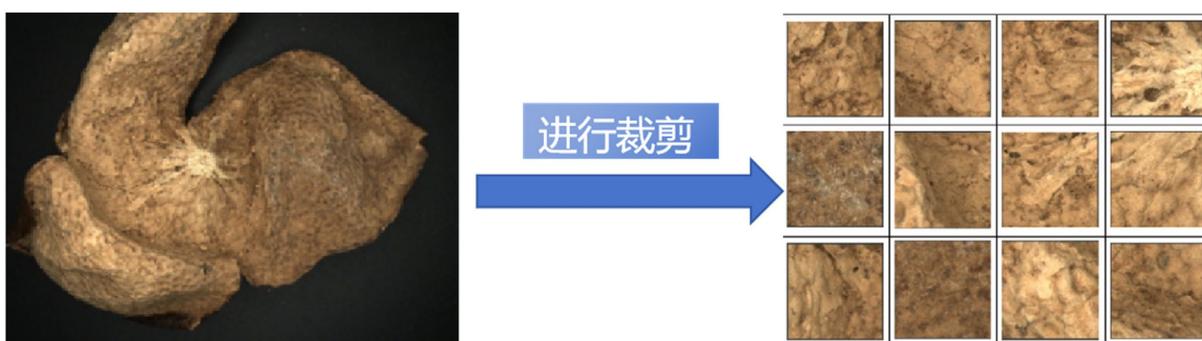
## 2. 材料与方法

### 2.1. 样品获取及数据集制备

本文以广东省江门市新会区生产的陈皮为研究对象，采集了 2014 年至 2022 年共 9 个不同年份的陈皮样品，其中 2014、2015、2016 三个年份的陈皮数分别为 80 片、100 片和 130 片，其他年份的陈皮数各为 180 片。为保证数据的代表性与实验的可重复性，每个年份的样品均在统一的实验环境下进行标准化图像采集。拍摄过程中严格控制光照条件、拍摄角度及背景布置，所有样品均置于统一的黑色背景上进行拍摄，以最大限度减少背景复杂度对图像特征提取的干扰，从源头上避免了后期图像分割可能引入的不确定性。

在图像采集完成后，进一步对原始图像进行裁剪预处理，以构建符合深度学习输入需求的标准化数据集。裁剪操作结合图像清晰度评估(clearDegree = 6)与背景比例控制(background percentage = 5%)策略，确保所提取图像区域具备良好的可辨识性和目标集中性。具体而言，首先以设定的步长遍历原始图像，对每个候选区域计算清晰度得分，并仅保留得分高于设定阈值的区域，以剔除模糊或信息不足的部分。随后，根据背景像素占比进行筛选，排除背景占比过高的区域，进一步提升图像内容的有效性。所有保留区域统一裁剪为尺寸为  $224 \times 224$  像素的图像片段，图像裁剪效果如图 1 所示。该裁剪过程不仅满足主流卷积神经网络对输入尺寸一致性的要求，从而提升模型训练的稳定性与效率，也有助于模型更专注于学习不同年份陈皮图像中的关键纹理特征，进而增强对年份差异的判别能力。

经过上述标准化采集与预处理流程，最终构建的数据集共包含 7808 张高质量陈皮图像，其中 2014 年至 2022 年各年份的图片数分别为：1029、1047、1046、1024、999、1028、692、532、411。为进一步提升模型训练效果并保证评估结果的科学性与可靠性，本文按照 7:3 的比例对数据集进行随机划分，其中 70%用于训练集构建，30%用于测试集评估，从而实现模型在训练与验证阶段的有效分离，确保模型性能评估的客观性。



**Figure 1.** Example of image cropping effects  
**图 1.** 图像裁剪效果示例图

### 2.2. 数据预处理与增强

为了提升模型的训练效果与泛化能力，本文在训练前对图像数据进行了规范化处理，并在训练阶段引入数据增强策略，以扩展样本多样性、降低过拟合风险。具体操作如下：

在数据预处理阶段，首先对图像进行像素值归一化处理。将图像从整型格式转换为浮点型张量，并将像素值从[0,255]映射至[0,1]区间，以满足神经网络输入格式要求。转换公式如下：

$$I'(x,y) = \frac{I(x,y)}{255} \quad (1)$$

其中， $I(x,y)$ 表示原始像素值， $I'(x,y)$ 表示归一化后的像素值。

为进一步规范数据分布、加速模型收敛过程，我们对输入图像进行标准化处理。标准化计算公式如下：

$$I'' = \frac{I - \mu}{\sigma} \quad (2)$$

其中， $I$ 为输入张量， $\mu$ 和 $\sigma$ 分别表示图像各通道的均值与标准差， $I''$ 为归一化后的张量。

在数据增强阶段，引入了随机旋转(RandomRotation)策略，以提高模型对图像姿态变化的鲁棒性。具体操作为：在训练过程中对图像随机施加 $\pm 30^\circ$ 的旋转，模拟现实采集条件中样本方向变化的情况。旋转操作的数学表达如下：

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3)$$

其中， $(x,y)$ 为旋转前的坐标， $\theta$ 为旋转角度， $(x',y')$ 为旋转后的坐标。

### 2.3. 所提方法

在本文中，采用基于轻量级卷积神经网络(CNN)架构的 MobileNetV3 模型对陈皮年份进行图像识别与分类。MobileNetV3 是一种专为移动端和资源受限设备优化设计的深度学习网络，其核心构建基于深度可分离卷积(Depthwise Separable Convolution)，在显著降低模型参数数量和计算复杂度的同时，仍能保持较高的图像分类精度[4]。

研究中，我们将经过前期标准化处理的陈皮图像数据作为输入，载入至基于迁移学习的 MobileNetV3 网络中进行训练与优化。通过引入在 ImageNet 等大型数据集上预训练的模型参数，迁移学习不仅显著加快了模型的收敛速度，也有效提升了图像识别的准确率[5]。

在特征提取阶段，MobileNetV3 主要采用深度可分离卷积结构，通过分别处理空间信息与通道信息，实现高效且具有判别力的特征表达。同时，引入 Hard-Swish 激活函数替代传统 ReLU，进一步增强网络的非线性表达能力与性能表现。网络中堆叠多个倒残差结构(Inverted Residual Blocks)，结合瓶颈结构与跳跃连接(Skip Connection)，在提升网络深度的同时缓解梯度消失问题，从而增强模型的表达能力与训练稳定性[6]。

最终，模型通过全局平均池化(Global Average Pooling)将高维特征压缩为定长向量，并经全连接层(Fully Connected Layer)输出对应年份的分类结果。整体架构具备高效、精准及良好适配性的特点，验证了 MobileNetV3 在陈皮年份识别任务中的应用可行性与有效性。

## 3. 实验与讨论

### 3.1. 实验过程

为全面评估不同卷积神经网络架构在陈皮贮藏年份图像识别任务中的性能表现，本文选取了七种具有代表性模型开展对比实验，涵盖从传统结构到新型轻量化模型。所选模型包括：早期经典架构 AlexNet 和 VGG，深层网络 ResNet 与 DenseNet，以及近年来广泛应用于移动端与边缘计算场景的轻量化模型

MobileNetV3、EfficientFormerV2 和 FasterNet。所有模型均在相同的标准化图像数据集上进行训练，并在统一测试集上评估其性能，以确保实验的可比性与客观性。

AlexNet 作为早期深度学习图像分类的里程碑模型，采用五层卷积层与三层全连接层，激活函数使用 ReLU，并通过重叠池化与局部响应归一化(LRN)提升模型泛化能力[7]。

VGG 网络结构简洁，使用多个  $3 \times 3$  卷积核堆叠构建深层网络，以增加感受野并提升特征表达能力。本文采用 VGG16 版本，其在多个图像分类任务中表现稳定[8]。

ResNet 网络通过引入残差连接有效缓解深层网络中的梯度消失问题，提升了网络的可训练性。本文选取 ResNet101 版本进行评估，验证其在深度识别任务中的优势[9]。

DenseNet 采用密集连接机制，即每一层接收所有前面层的特征图输入，显著增强了特征复用效率并改善信息流传递。本文采用 DenseNet121 版本，其高效的参数利用率在图像识别任务中表现出色[10]。

MobileNetV3 属于轻量级神经网络结构，结合深度可分离卷积(Depthwise Separable Convolution)和倒残差模块(Inverted Residual Block)，引入 SE 注意力机制与 Hard-Swish 激活函数，兼顾高精度与低计算成本，适用于嵌入式场景[4]。

EfficientFormerV2 为高效视觉 Transformer-CNN 混合模型，结合局部卷积与多头注意力机制(MHSA)，在实现轻量化的同时具备较强的语义建模能力，尤其适用于精度与推理效率并重的实际应用环境[11]。

FasterNet 则通过重新设计瓶颈模块，优化卷积计算路径，并集成轻量级通道注意力机制，实现推理速度与精度的有效平衡，是新一代高性能轻量模型的代表之一[12]。

通过对上述模型的分分类准确率、F1 分数、推理时间(FLOPs)等多个关键性能指标进行综合评估，旨在选出在陈皮年份识别任务中性能最优、效率最优的网络架构。

### 3.2. 实验设置

为确保不同模型在统一条件下进行公平对比，所有模型的训练轮数(Epochs)统一设定为 500，批次大小(Batch Size)设置为 16，以在模型收敛效率与显存资源消耗之间取得平衡。训练过程中引入早停策略(Early Stopping)，当准确率在连续 30 个训练周期内无显著提升时，自动中止训练过程，以防止模型过拟合。在优化器设置方面，采用 Adam 自适应梯度算法，并针对不同结构将模型划分为特征提取模块与分类器模块，分别设置学习率为 0.0001 和 0.0002。同时引入 StepLR 学习率调度策略，每 5 个训练周期后将学习率按 0.9 的比例进行衰减，从而提升训练稳定性与收敛效率。所有模型均使用多分类交叉熵损失函数(Cross-Entropy Loss)作为目标函数，以提升模型对不同年份陈皮图像的区分能力和分类性能。

### 3.3. 评价指标

为全面评估所提模型在陈皮年份识别任务中的性能，本文采用了交叉熵损失(Loss)、准确率(Acc)、精确率(P)、召回率(R)、F1 分数(F1)和浮点运算量(FLOPs)等常见评价指标。相关公式具体如下：

$$Loss = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(p_{i,c}) \quad (4)$$

其中， $N$  为样本总数， $C$  为年份类别数， $y_{i,c} \in \{0,1\}$  表示第  $i$  个样本的真实标签， $p_{i,c}$  为模型预测的第  $i$  个样本属于类别  $c$  的概率。

$$Acc = \frac{1}{N} \sum_{i=1}^N I(\hat{y}_i = y_i) \quad (5)$$

其中， $\hat{y}_i$  为第  $i$  个样本的预测标签， $y_i$  为真实标签， $I(\cdot)$  为指示函数，表示预测与真实标签一致时取值为 1，否则为 0。

$$P = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FP_c} \quad (6)$$

$$R = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FN_c} \quad (7)$$

其中,  $TP_c$ 、 $FP_c$ 、 $FN_c$  分别表示类别  $c$  的真阳性、假阳性和假阴性样本数。

$$F1 = \frac{2PR}{P+R} \quad (8)$$

$$FLOPs = 2 \cdot H_{out} \cdot W_{out} \cdot C_{in} \cdot K_h \cdot K_w \cdot C_{out} \quad (9)$$

其中,  $H_{out}$ 、 $W_{out}$  为输出特征图的尺寸,  $C_{in}$ 、 $C_{out}$  为输入和输出通道数,  $K_h$ 、 $K_w$  为卷积核的尺寸。

**Table 1.** Evaluation metrics results of different models

**表 1.** 各模型评价指标结果数据

模型名称	损失值(loss)	推理时间 (毫秒/张)	准确率(%)	精确率(%)	召回率(%)	F1 分数(%)	模型大小(MB)
AlexNet	0.7439	15.12	79.79	80.02	78.68	78.87	233
VGG	0.7679	13.47	82.05	82.86	82.48	82.24	527.8
ResNet	0.4597	10.16	85.60	85.68	86.07	85.83	162
DenseNet	0.5149	13.55	79.79	79.78	80.59	79.87	28.5
MobileNetV3	0.4532	<b>5.36</b>	86.96	<b>87.55</b>	87.16	87.28	16.3
EfficientFormerV2	0.3950	10.60	<b>87.61</b>	86.97	<b>87.81</b>	87.23	<b>13.5</b>
FasterNet	<b>0.3896</b>	9.03	87.22	87.31	87.74	<b>87.32</b>	114

### 3.4. 实验结果

为验证所提模型在陈皮年份识别任务中的有效性, 本文对 MobileNetV3-Large 与 ResNet101、DenseNet121、EfficientFormerV2-S0 及 FasterNet-T0 等主流模型进行了系统对比。实验结果如表 1 所示, MobileNetV3-Large 在分类性能、推理效率与模型轻量化方面实现了最佳平衡。

具体来看, MobileNetV3-Large 实现了 86.96% 的准确率、87.16% 的召回率与 87.28% 的 F1 分数, 在精度上与最优模型 EfficientFormerV2-S0 (Acc: 87.61%) 相差不足 0.8%, 但推理速度提升近一倍 (5.36 ms vs. 10.60 ms), FLOPs 减少约 49%。相比 ResNet101 (Acc: 85.60%) 与 DenseNet121 (Acc: 79.79%), MobileNetV3-Large 在精度与速度上均明显占优。

此外, MobileNetV3-Large 模型体积仅为 16.3 MB, 远小于 FasterNet-T0 (114 MB) 与 VGG (527.8 MB), 在边缘部署中更具优势。虽然 FasterNet-T0 损失值最低 (0.3896) 且推理较快 (9.03 ms), 但其模型规模大、计算效率逊色。DenseNet121 尽管结构紧凑 (28.5 MB), 推理延迟 (13.55 ms) 和损失值 (0.5149) 仍表现不佳。

MobileNetV3-Large 在陈皮年份识别任务中表现良好, 潜在的原因可能是: 它采用混合卷积核策略, 结合深度可分离卷积与普通卷积, 有效降低计算复杂度且保持高精度; 通过引入注意力机制 (如 SE 模块), 使模型能聚焦于陈皮纹理、色泽等关键年份特征区域, 提升特征表达能力。

综合分析各项指标, MobileNetV3-Large 凭借高精度、低延迟与小体积, 展现出出色的综合性能与边缘部署潜力, 是陈皮年份识别任务中最具实用价值的模型选择。

## 4. 结论

本文构建了一种基于 MobileNetV3 的陈皮年份识别模型，成功实现了在保证高分类精度的同时显著降低计算成本，展示了该模型在中药材图像识别领域的应用潜力。实验结果表明，该方法具备良好的泛化能力与实际部署价值，特别适用于资源受限的场景。未来工作将扩大样本年份范围，探索多模态数据(近红外光谱)融合，并开发移动端部署应用，助力中药材产业数字化转型。

## 基金项目

本文获五邑大学大学生创新创业训练计划项目(202311349014)资助，部分获五邑大学港澳联合研发基金(2022WGALH16)资助。

## 参考文献

- [1] 国家药典委员会. 中华人民共和国药典(一部) [M]. 北京: 中国医药科技出版社, 2020.
- [2] Yin, J., Wang, Y., He, L., Jiang, W. and Liang, Y. (2022) Chemical Variation of Chenpi (Citrus Peels) and Corresponding Antioxidant Activities during the Aging Process. *Frontiers in Nutrition*, **9**, Article ID: 825381.
- [3] Chen, Y., Sun, J., He, X., Yu, L. and Lin, Y. (2021) UPLC-Q-TOF-MS with Chemometrics Approach Analysis of Citri Reticulatae Pericarpium. *Phytochemical Analysis*, **32**, 793-802.
- [4] Howard, A., Sandler, M., Chen, B., Wang, W., Chen, L., Tan, M., *et al.* (2019) Searching for MobileNetV3. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 1314-1324. <https://doi.org/10.1109/iccv.2019.00140>
- [5] Pan, S.J. and Yang, Q. (2010) A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, **22**, 1345-1359. <https://doi.org/10.1109/TKDE.2009.191>
- [6] Sandler, M., Howard, A., Zhu, M., *et al.* (2018) MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, 18-22 June 2018, 4510-4520. <https://doi.org/10.1109/CVPR.2018.00474>
- [7] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 25, 1097-1105.
- [8] Simonyan, K. and Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. *International Conference on Learning Representations (ICLR)*, San Diego, 7-9 May 2015, 1-14. <https://arxiv.org/abs/1409.1556>
- [9] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/cvpr.2016.90>
- [10] Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. (2017) Densely Connected Convolutional Networks. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 4700-4708. <https://doi.org/10.1109/cvpr.2017.243>
- [11] Li, Y., Du, J., Song, G. and Qian, C. (2022) EfficientFormerV2: Smaller, Faster and Stronger. <https://arxiv.org/abs/2212.08059>
- [12] Chen, Y., Dai, X., Liu, M., *et al.* (2023) FasterNet: Rethinking Bottleneck Structure for Real-Time Efficient Image Recognition. <https://arxiv.org/abs/2206.03492>