Journal of Electronics and Information Science (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/jeis.2025.100218
ISSN 2371-9524 Vol. 10 Num. 2

Design and Implementation of Apple Ripeness Grading System Based on Lightweight ResNet18

Jiajie Qiu[a], Xiaoying Su[b,*]

School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, China
[a]3887756843@qq.com, [b]25980669@qq.com
**Corresponding author*

Keywords: Machine Vision; ResNet18; Lightweight Improvement; Apple Ripeness; Image Classification

Abstract: To address the practical need for rapid ripeness detection of fruits after harvesting, an apple ripeness grading system based on a lightweight ResNet18 model is designed. Taking Red Fuji apples as the research object, images of apples at different ripening stages are collected by image acquisition equipment. After expanding the dataset through data augmentation, the ResNet18 model is optimized by channel pruning and activation function improvement to construct a lightweight classification model suitable for terminal deployment. Experimental results show that the improved model achieves an accuracy of 94.2% on the test set, with a 42% reduction in model parameter scale and a 35% increase in inference speed, which can meet the practical application requirements of rapid apple ripeness grading.

1. Introduction

As one of the major economic fruits in China, the ripeness of apples directly affects their market value and consumer experience. Traditional ripeness grading relies on manual observation of peel color, fruit shape and other features, which has problems such as low efficiency, strong subjectivity and inconsistent grading standards, making it difficult to adapt to the large-scale development needs of the modern fruit industry. With the rapid development of deep learning and machine vision technologies, image-based fruit quality detection technology, which benefits from the research foundation of target detection models like Fast R-CNN[2], has become a research hotspot in the agricultural field due to its advantages of non-contact, high efficiency and strong objectivity.

At present, some scholars have applied deep learning models to the research of fruit ripeness grading, such as using VGG16, ResNet50[1] and other models to realize the ripeness recognition of citrus, strawberries and other fruits. However, these models usually have large parameter scales and high computational complexity, making them difficult to deploy on resource-constrained devices such as embedded terminals. As a lightweight deep residual network, ResNet18 has the characteristics of simple structure and high training efficiency, but there are still problems of parameter redundancy and insufficient inference speed when directly applied to apple ripeness grading. Therefore, this paper takes Red Fuji apples as the research object, and conducts lightweight improvement on the ResNet18 model to construct an efficient and accurate apple ripeness grading system, providing technical support for the automatic grading of the fruit industry.

2. Related Technical Foundations

2.1. Principles of Machine Vision Image Acquisition

Machine vision technology converts the visual information of objects into digital images through image acquisition equipment, and then extracts features and completes detection tasks through image processing and pattern recognition algorithms. In apple ripeness grading, the core of image acquisition is to obtain image information that can accurately reflect ripeness characteristics. The ripeness of apples is mainly reflected by the color change of the peel. From unripe to ripe, the peel color of Red Fuji apples gradually transitions from light green to light red and dark red, and the fruit surface gloss also changes. Therefore, it is necessary to ensure stable lighting conditions during the image acquisition process to avoid color feature distortion caused by light changes.

2.2. ResNet18 Network Structure

ResNet18 is built based on the Deep Residual Network (Residual Neural Network), which solves the problem of gradient disappearance in deep network training by introducing residual connections. Its network structure includes 1 7×7 convolutional layer, 4 residual block groups and 1 fully connected layer. Each residual block group is composed of 2 3×3 convolutional layers, and the total number of network layers is 18. Compared with traditional CNN, ResNet18 simplifies the training process of deep networks through residual connections while maintaining high feature extraction capability, making it a commonly used model for lightweight image classification tasks.

2.3. Model Lightweight Technology

Model lightweight technologies mainly include model pruning, quantization, knowledge distillation and other methods[4]. This paper adopts channel pruning technology to reduce the model parameter scale. Its core idea is to remove the convolutional channels with low contribution in the network, so as to reduce the computational complexity while ensuring small loss of model performance. At the same time, the traditional ReLU activation function is replaced with Leaky ReLU[5] to solve the problem of neuron "death" caused by the zero gradient of the ReLU function in the negative semi-axis, and improve the feature extraction capability of the model.

3. Design of Apple Ripeness Grading System

3.1. Overall System Architecture

The apple ripeness grading system is mainly composed of an image acquisition module, a data preprocessing module, a lightweight ResNet18 classification module and a result output module. The image acquisition module is responsible for obtaining apple sample images; the data preprocessing module performs enhancement and standardization processing on the original images to improve the diversity of the dataset and the generalization ability of the model; the lightweight ResNet18 classification module is the core of the system, which realizes the extraction and grading of apple ripeness features; the result output module displays the classification results of the model in a visual form and outputs the grading report.

3.2. Image Acquisition and Dataset Construction

The image acquisition equipment adopts an industrial camera (model: Basler acA1920-40gc) with an 8mm fixed-focus lens. The acquisition environment is a closed light box to avoid interference from external light. Four uniformly distributed LED light sources are installed in the light box, and the color temperature is set to 5500K to ensure that the light intensity is stably maintained at 800-1000lux. The collection object is Red Fuji apples, which are divided into three grades according to agricultural grading standards: unripe (light green peel without red spots), semi-ripe (light red peel with red area accounting for 30%-60%), and ripe (dark red peel with red area exceeding 60%).

A total of 1500 apple samples were collected, with 500 samples in each grade. Each sample was photographed from different angles to obtain 3 images, and finally 4500 original images were obtained. To avoid model overfitting, data augmentation was performed on the original images, including random cropping (cropping size 224×224), horizontal flipping (flipping probability 0.5), brightness adjustment (brightness change range ±15%), and Gaussian noise addition (noise standard deviation 0.01). The expanded dataset contains 13500 images. The dataset is divided into training set (9450 images), validation set (2700 images) and test set (1350 images) in the ratio of 7:2:1[3].

3.3. Construction of Lightweight ResNet18 Model

Lightweight improvement is carried out based on the original ResNet18 model, and the improvement mainly includes two parts: channel pruning and activation function optimization. Channel pruning is realized by calculating the L1 norm of each convolutional layer channel. The smaller the L1 norm, the lower the feature contribution of the channel. The pruning threshold is set to 0.1, and the channels with L1 norm less than the threshold are removed. For each residual block group, the pruning ratios are set to 0.2, 0.3, 0.3, 0.2 in turn to avoid excessive pruning leading to a significant decline in model performance.

The ReLU activation function in the original ResNet18 is replaced with Leaky ReLU, and its slope parameter is set to 0.01, so that the input of the negative semi-axis can still generate a small gradient, improving the model's ability to extract weak features. In addition, a Batch Normalization layer is added before the fully connected layer to accelerate the convergence speed of model training and reduce the risk of overfitting. The improved model structure retains the core of residual connection of ResNet18, while the parameter scale and computational complexity are significantly reduced.

3.4. Model Training and Optimization

The model training environment is Python 3.8, the deep learning framework is PyTorch 1.12, and the hardware configuration is Intel Core i7-12700K CPU and NVIDIA RTX 3060 GPU. The training parameters are set as follows: Batch Size is 32, initial learning rate is 0.001, Adam optimizer is used to update the model parameters, the learning rate decay strategy is to decay to 0.5 of the original every 10 epochs, the number of training epochs is 50, and the cross-entropy loss function is used.

During the training process, the model parameters are dynamically adjusted through the accuracy of the validation set. When the accuracy of the validation set does not improve significantly for 5 consecutive epochs, the training is stopped and the optimal model is saved. To further improve the model performance, the Early Stopping strategy is adopted to avoid overfitting, and the gradient during training is clipped (gradient clipping threshold is 1.0) to prevent gradient explosion.

4. Experimental Results and Analysis

4.1. Evaluation Indicators

Accuracy, Precision, Recall and F1-Score are used as model performance evaluation indicators, and the parameter scale and inference speed of the model are also counted to comprehensively evaluate the effectiveness of the improved model. The definitions of each indicator are as follows: Accuracy is the proportion of correctly classified samples to the total number of samples; Precision is the proportion of samples predicted to be a certain grade and actually belonging to that grade to the number of samples predicted to be that grade; Recall is the proportion of samples predicted to be a certain grade and actually belonging to that grade to the number of samples actually belonging to that grade; F1-Score is the harmonic mean of Precision and Recall, which comprehensively reflects the classification performance of the model.

4.2. Comparative Experiment of Model Performance

To verify the superiority of the improved model, a comparative experiment is carried out with the original ResNet18 model and the lightweight VGG16 model. The experimental results are shown in Table 1.

Table 1 Accuracy Comparison between Transfer Learning and Training from Scratch

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Parameter Scale (MB)	Inference Speed (ms/image)
Lightweight VGG16	88.6	87.9	88.1	88.0	42.3	28.5
Original ResNet18	92.1	91.8	91.6	91.7	11.7	18.2
Improved ResNet18	94.2	93.9	93.7	93.8	6.8	11.8

It can be seen from Table 1 that the improved ResNet18 model is superior to the comparison models in all performance indicators. Compared with the original ResNet18, the accuracy of the improved model is increased by 2.1 percentage points, which is because the Leaky ReLU activation function enhances the model's ability to extract the gradient color features of apple peel; the parameter scale is reduced from 11.7MB to 6.8MB, with a reduction ratio of 42%, and the inference speed is increased from 18.2ms/image to 11.8ms/image, with an increase ratio of 35%, which reflects the lightweight effect of the channel pruning technology. Compared with the lightweight VGG16 model, the accuracy of the improved model is increased by 5.6 percentage points, and the parameter scale is reduced by 83.9%, which fully shows the dual advantages of the improved model in performance and lightweight.

4.3. Analysis of Classification Results for Different Ripeness Grades

To further analyze the classification performance of the model, the classification results of the improved ResNet18 model on different ripeness grades are counted. The accuracy, precision and recall of the unripe grade are 96.3%, 95.8% and 96.7% respectively; the accuracy, precision and recall of the semi-ripe grade are 92.1%, 91.5% and 92.8% respectively; the accuracy, precision and recall of the ripe grade are 94.2%, 94.5% and 93.8% respectively.

The classification performance of the semi-ripe grade is relatively low. The main reason is that the peel color of semi-ripe apples is in a transitional stage, and the red area of some samples is close to 30% or 60%, which has small feature differences from unripe or ripe samples, leading to easy misjudgment by the model. To solve this problem, the classification accuracy can be further improved by increasing the number of semi-ripe samples and optimizing the image acquisition angle.

5. Conclusions and Prospects

5.1. Research Conclusions

This paper designs an apple ripeness grading system based on lightweight ResNet18, which realizes the automatic grading of apple ripeness through image acquisition, data preprocessing, model improvement and training optimization. Experimental results show that through channel pruning and activation function optimization, the improved ResNet18 model significantly reduces the model parameter scale and computational complexity while ensuring classification accuracy. The test set accuracy reaches 94.2%, and the inference speed meets the requirements of real-time grading, providing a feasible solution for the automatic detection of apple ripeness.

5.2. Research Prospects

The research in this paper still has certain shortcomings, which can be further improved in the following aspects in the future: First, enrich the dataset by expanding its scope to include apple samples of different varieties and growth environments, thereby improving the model's generalization ability; second, introduce the attention mechanism [6] to enable the model to focus on key feature regions such as changes in apple peel color, further enhancing the classification accuracy; third, develop embedded terminal application programs, deploy the lightweight model on smartphones or special detection equipment, realize on-site rapid detection of apple ripeness, and promote the practical application of the technology.

References

[1] 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.

[2] Girshick R. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1440-1448.

[3] Deng J, Dong W, Socher R, et al. Imagenet: A large-scale hierarchical image database[C]//2009 IEEE conference on computer vision and pattern recognition. IEEE, 2009: 248-255.

[4] Han S, Pool J, Tran J, et al. Learning both weights and connections for efficient neural network[J]. Advances in neural information processing systems, 2015, 28.

[5] Maas A L, Hannun A Y, Ng A Y. Rectifier nonlinearities improve neural network acoustic models[C]//Proc. icml. 2013, 30(1): 3.

[6] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words[J]. arXiv preprint arXiv:2010.11929, 2020, 7: 5.