Research on External Quality Automatic Detection and Classification Method of Navel Orange Based on Simple Dense Network

Yumei Tan¹, Peng Li¹, Xiaoxin Li¹, Longwen Yang²*

¹ College of Artificial Intelligence, Hezhou University, Hezhou, Guangxi, 542899, China
² Hezhou Electric Power Supply Bureau, Guangxi Southern Power Grid Company Limited, Hezhou, Guangxi, 542899, Country

Abstract. We propose a simple dense convolutional neural network model (SDenseNet) to solve the problem of low accuracy and poor performance in classification of navel orange with large external differences. Specifically, AlexNet is adopted as the backbone network, through the introduction of batch normalization (BN) and reduction initial size of the convolution kernel to accelerate convergence of the network model. Besides, we design a new feature reuse structure to promote the connection between layers, and modify the full connection layer by using global pooling to reduce training parameters. The experimental results demonstrate that our proposed SDenseNet significantly improves the performance with 99.33% accuracy on self-built navel orange dataset, outperforming the classic models of LeNet, AlexNet, SqueezeNet and ResNet with significant improvements of 5.33%, 1.55% and 3.55% respectively. The research results provide a new solution for external quality automatic detection and classification of navel orange.

1. Introduction
With the development of modern agricultural informatization and the improvement of people’s consumption, people’s requirement for the quality of fruits increases. Moreover, it will bring more profits for fruit industry by external quality automatic detection and classification of fruits. The existing fruit detection and classification methods can mainly be divided into two categories: traditional methods and machine vision methods based on deep learning. The traditional methods mainly rely on artificial classification and mechanical classification[1][2][3], among them, the artificial classification method will inevitably cause time consumption, low efficiency and uneven quality, mechanical classification method will cause damage to the appearance of fruit, even though its fast detection speed[4], while the machine vision methods based on deep learning can be very fast, accurate and non-destructive for classification.

Recently, deep learning has led to enormous progress on a variety of tasks, such as the convolutional neural network (CNN) performs in image classification and recognition. Many works on recognition, classification and defect detection of various fruits based on convolutional neural network have progressed enormously. Peng et al.[5] proposed an improved SSD network model for identifying apple, litchi, navel orange and imperial orange, achieving average detection accuracy by 89.53%. Cheng[2] designed an improved residual network model for navel orange automatic detection and
classification with accuracy of 92%. Fu et al.[6] proposed a kiwi fruit image recognition system based on the optimized LeNet network model, which obtained 94.78% recognition accuracy for separated fruit. Li[4] designed a simple orange classification recognition model based on convolution neural network, achieving accuracy of 94.34% on self-built orange dataset. Gene-Mola et al.[7] adopted the convolutional neural network model based on VGG16 to recognize Red Fuji Apple in RGB-D images, obtaining an AP of 94.8%. Nevertheless, special works on the quality classification of navel orange are numbered, and recognition accuracy fails to meet the actual demand, what’s more, there is no public dataset available. We do research on Fuchuan navel orange and build navel orange dataset, and train several convolutional neural network models. The experimental results demonstrate the effectiveness and feasibility of our proposed model, which provides a new idea for automatic quality classification of navel orange.

2. Materials and Methods

2.1. Navel orange data acquisition
Firstly, we need to build navel orange dataset for quality classification. It mainly includes three steps: data acquisition, data labeling and data segmentation.

**Data acquisition:** We capture a series of navel orange fruit images by image acquisition devices (such as digital cameras and smart phones) in Fuchuan navel orange planting bases and fruit supermarkets. It is important to note that these images are taken under various lighting conditions, different scene backgrounds, and different camera angles. After collecting all images, we make preprocessings, such as resizing all images with 256x256 pixels.

![Table 1. Samples of navel oranges of different qualities.](image)

| Label | Training set | Test set |
|-------|--------------|----------|
| 0     | 550          | 137      |
| 1     | 546          | 136      |
| 2     | 568          | 142      |
| 3     | 575          | 143      |

**Data labeling:** As presented in Table 1, navel oranges are divided into four categories, namely, unqualified, qualified, good and high quality. Its classification standards are mainly based on navel orange diameter, color ratio at maturity, brightness, surface defects, surface scars and so on. Finally, the collected datas are annotated manually according to the aboved standards.

**Data segmentation:** We leverage image data augmentation method including cropping, flipping, mirroring, etc., to obtain 3000 navel orange images, and divide these images into 80% training set and 20% test set, as shown in Table 2.
2.2. Proposed SDenseNet model

The original AlexNet model[8] has been widely used in various recognition tasks in transfer learning manner due to its superior performance. However, the parameters of AlexNet model are huge, and it is prone to be over-fitting with inadequate data[9]. To address the problem above, we design a new SDenseNet model based on AlexNet by modifying the convolution layers, and leveraging a new designed feature reuse structure to redesign the convolution structure in this paper, an overview of the architecture is shown in Fig. 1.

- **1.** BN layer replace LRN layer
  
  Since the batch normalization (BN) algorithm is beneficial to speed up model convergence and alleviate the over-fitting problem in deep network model training, a large number of networks with superior performance all adopt BN instead of local response normalization(LRN), such as GoogleNet[10], ResNet[11], ShuffleNet[12], MobileNet[13] and so on, those models add BN layers to accelerate model convergence. Inspired by this, we replace the LRN layer with BN layer in AlexNet model to improve performance.

- **2.** Redesign convolution structure
  
  Traditional standard convolution structure, as shown in Fig. 2 (a), extract features layer by layer without information interaction between different layers, which is not conducive to extract more discriminative features for classification. Enlightened by the idea of feature reuse structure in DenseNet[14], we redesign the traditional standard convolution structure to obtain a new feature reuse convolution structure used for feature extraction in our task, as shown in Fig. 2 (b). By doing this, information interaction across different layers can be improved by concatenating the front N layers on channel, which will be an input of later layer, so as to fuse more effective features. Simultaneously, we introduce 1x1 small size convolution kernels for dimension reduction (in Layer4), thus, the whole structure will reduce model parameters and increase the utilization of feature maps to get better classification performance.

- **3.** Simplify full connection layer
  
  The AlexNet model contains three full connection layers, leading to dense feature weights, which makes it easy to occur over-fitting. In order to relieve the problem, we use the global pooling to simplify the full connection layer to reduce training parameters and prevent network over-fitting.
3. Results and analysis

3.1. Experimental settings
We trained our SDenseNet on the Caffe platform with Python 3.6, and computer configuration is shown in Table 3. We adopted Stochastic Gradient Descent (SGD) optimizer, set the initial learning rate to 0.005, the maximum iterations of model can be 10000 times.

Table 3. Computer configuration.

| Type            | Parameter(M)                  |
|-----------------|-------------------------------|
| CPU             | Intel® Core™ i7-9700@3.0GHz   |
| GPU             | NVIDIA GeForce GTX 1080 Ti   |
| RAM             | 32GB                          |
| Operating system| Ubuntu 16.04 LTS 64           |

3.2. Training results and analysis
Fig.3 shows change curves of accuracy and loss value for different models, (a) shows the accuracies change of different models along training iterations, (b) shows the losses change of different models along training iterations. As seen from Fig.3, compared with other models, our proposed SDenseNet model has the fastest convergence rate. Meanwhile, test accuracies and losses of all models tend to be stable after 10000 iterations, and all accuracies are kept above 94%.

3.3. Performance comparisons of different models
The comparison results of different models are shown in Table 4. It is clear that there are great differences in recognition accuracy, training time and model size of different CNN models. Among all the results, shallow model, namely Improved LeNet, costs least training time but its accuracy is only 94.00%, while those deep models, SqueezeNet, AlexNet and ResNet, achieve 95.78%, 97.78% and 98.23 respectively, but they either bring too much parameters or take much training time. On the contrary, our proposed SDenseNet achieves accuracy of 99.33%, and its parameter is only 0.63M compared, which greatly reduce calculations. Experimental results show that our proposed model has better performance than the traditional ones, which is more accurately and effectively for navel orange classification.

Table 4. Performance comparison of different models.

| Model       | Parameter(M) | Training time(s) | Accuracy(%) |
|-------------|--------------|------------------|-------------|
| Improved LeNet | 3.63         | 130              | 94.00       |
| Model     | Accuracy | Parameters | Training Time |
|-----------|----------|------------|---------------|
| AlexNet   | 56.88    | 360        | 97.78         |
| SqueezeNet| 0.72     | 305        | 95.78         |
| ResNet-18 | 11.17    | 1076       | 98.23         |
| SDenseNet | 0.63     | 218        | 99.33         |

4. Conclusions
In this paper, we proposed a simple dense network for navel orange detection and classification, and built navel orange dataset according to navel orange classification tasks. The experimental results show that our SDenseNet outperforms those classical deep learning models, less parameters, less training time and the highest accuracy of 99.33%, which is appropriate for practical application.

Acknowledgments
This work was financially supported by the project of improving the basic scientific research ability of young and middle-aged teachers in Guangxi Universities (2020KY18016).

References
[1] He W.B, Wei A.Y, Ming W.Y, Jia H.J. (2020) Survey of fruit quality detection based on machine vision[J]. Computer Engineering and Applications, 56 (11):10-16.
[2] Cheng C.Y. (2018) Research on The Technology Of Navel orange Quality Automatic Detection and Classification Based on Machine Vision[D]. Jiangxi University of Science and Technology.
[3] Fu L.S, Song Z.Z, Zhang X, Li R, Wang D, Cui Y.J. (2020) Application and research progress of deep learning in agriculture[J]. Journal of China Agriculture University, 25(02):105-120.
[4] Li L.F, Wang J.J, Wei Z, Feng X.G, An J.L. (2020) Research on orange classification and recognition based on convolutional neural network[J]. Journal of Henan Institute of Science and Technology (Natural Science Edition), 48(03):68-73.
[5] Peng H.X, Huang B, Shao Y.Y, Li Z.S. (2018) General improved SSD model for picking object recognition of multiple fruits in natural environment[J]. Transactions of the Chinese Society of Agricultural Engineering, 34(16):155-162.
[6] Fu L.S, Feng Y.L, Elkamil T, Liu Z.H, Li R, Cui Y.J. (2018) Image recognition method of multi-cluster kiwifruit in field based on convolutional neural networks[J]. Transactions of the Chinese Society of Agricultural Engineering, 34(02):205-211.
[7] Jordi G.M, Verónica V, Rosell Polo J.R. (2019) Multi-modal deep learning for Fuji apple detection using RGB-D cameras and their radiometric capabilities[J]. Computers and Electronics in Agriculture, 162:689-698.
[8] Krizhevsky A, Sutskever I, Hinton G. (2012) ImageNet Classification with Deep Convolutional Neural Networks[C]. In: NIPS. Curran Associates Inc. 2012.
[9] Guo X.Q, Fan T.J, Shu X. (2019) Tomato leaf diseases recognition based on improved Multi-Scale AlexNet[J]. Transactions of the Chinese Society of Agricultural Engineering, 35(13):162-169.
[10] Szegedy C, Vanhoucke V, Ioffe S, et al. (2016) Rethinking the inception architecture for computer vision[C]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition., pp.2818-2826.
[11] He K, Zhang X, Ren S, et al. (2016) Deep residual learning for image recognition[C]. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770-778.
[12] Zhang X, Zhou X, Lin M, et al. (2018) ShuffleNet: An extremely efficient convolutional neural network for mobile devices [C]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 6848-6856.
[13] Sandler M, Howard A, Zhu M, et al. (2018) MobileNetV2: inverted residuals and linear bottlenecks[C]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4510-4520.

[14] Huang G, Liu Z, Van Der Maaten L, et al. (2017) Densely connected convolutional networks[C]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4700-4708.