Orange Fruit Images Classification using Convolutional Neural Networks

Dhiya Mahdi Asriny, Septia Rani, Ahmad Fathan Hidayatullah
Department of Informatics, Universitas Islam Indonesia, Yogyakarta, Indonesia
Email: 15523033@students.uii.ac.id, septia.rani@uii.ac.id, fathan@uii.ac.id

Abstract. The quality of fruit is important to increase sales in the market. Right now, the quality selection of orange fruit is mostly still complete by humans. Several drawbacks such as inaccuracy and inconsistency in results happened due to the limitation of human perceptions. The development of computer vision, making it possible to train the computer to classify images based on specified characteristics. This paper proposes the classification model to classify orange images using Convolutional Neural Network (CNN). Five classes of orange namely good-orange-grade-1, good-orange-grade-2, immature-orange, rotten-orange, and damaged-orange are classified using deep learning CNN. Total of 1000 orange images is collected using the smartphone camera. Each class consists of 200 images which are divided into 60% as the training data, 20% as the validation data, and 20% as the testing data. K-Fold Cross-Validation method is used to validate the model. In this paper, the hidden layer of CNN consists of 256 nodes. Two activation functions, ReLU and Tanh, are employed for comparing the accuracy of classification with the Softmax classifier. The result shows that the accuracy of ReLU activation function is 96%, which is better than the Tanh activation function that gives only 93.8%.

Keywords: classification, convolutional neural network, deep learning, orange fruit

1. Introduction
The productivity of orange fruits in Indonesia is high. Badan Pusat Statistik (Central Bureau of Statistics) Indonesia [1] reported that the commodity of orange fruits ranked in the third position with a high productivity rate after banana and mango. Therefore, orange fruits have become the main commodity because of their great contribution to the national economy. In 2017, the productivity of orange fruits in Indonesia increased by 150,978 ton or about 7.50% compared with the previous year. In line with that, orange fruits have given high profit for Indonesian farmers [2].

In order to increase sales and market competitiveness of orange fruits (oranges), it is important to select the oranges with high quality. Several indicators such as size, shape, and color are used to determine the quality during the selection process. However, the selection process of the oranges is still conducted manually by humans, which is time-consuming. Some other problems such as visual limitation, fatigue, slow speed, and low precision have also happened which can make the selection’s results are not in accordance with the requirements of the market. Therefore, we need to overcome those problems by proposing a model that is able to recognize the quality of the oranges automatically.
This study aims to build a classification model for orange fruit images. We propose Convolutional Neural Network (CNN) as the method to classify orange fruit images into 5 classes, namely good-orange-grade-1, good-orange-grade-2, damaged-orange, rotten-orange, and immature-orange. It has been proven that CNN shows great performance in some tasks, including object detection and image recognition [3]. Moreover, we also investigate the performance of our model using several scenarios and metrics to obtain the best model for classification.

The rest of this paper is organized as follows. Section 2 describes the related work. Section 3 describes the methodology. Section 4 describes the results and discussion. Finally, the concluding remarks are described in section 5.

2. Related Work

Research on the quality of oranges has previously been done by [4] using Learning Vector Quantization (LVQ) method and get 76% accuracy. In their research, they focused on classifying one type of orange, which is lime. Subsequent research was conducted by [5] using Naive Bayes on orange images and obtained an accuracy result of 91.6%. In 2016, research conducted by [6] on orange images obtained an accuracy result of 88%. All of those research still use traditional machine learning approach for classifying orange images. A recent study carried out by [7] also conducted orange images classification. However, they were not only utilized traditional machine learning but also utilized a deep learning approach. The study revealed that Hybrid Artificial Neural Network-Harmony Search (ANN-HS) outperformed k-Nearest Neighbors (kNN) with accuracies of 94.28% and 70.88% respectively.

Therefore, we will discuss more about the application of deep learning approach for image classification. One of the most popular deep learning methods that are suitable for image classification is CNN. CNN seeks to emulate the image recognition system in visual cortex such as humans, thus by applying CNN, machine (computer) can be able to process image information [8]. Other research stated that CNN can reduce the number of parameters and able to handle the input image deformation such as translation, rotation, and scale [9].

Research using CNN was conducted by [10] to classify advertising images. According to their experiments, the accuracy result is 86%. Other research [11] also using CNN on different images, which is chili images. Their model achieved 80% accuracy on the test datasets. Recent research by [12] conducted a classification of fruit images (apples, bananas, blueberries, kiwi, and raspberry) using CNN. The result of their study showed that the accuracy achieved 94%.

3. Methodology

In this study, we classified images of orange using Convolutional Neural Network (CNN). CNN is suitable for processing data with grid-like topologies, such as time-series data (1D grid) and image data (2D grid of pixels). CNN used a mathematical operation called a convolution, which is a type of special linear operation. The convolution network is simply a neural network that uses convolution as a substitute for common matrix multiplication in at least one layer [13]. The model proposed in this research has several stages, consisting of collecting data image, image pre-processing, training the model using CNN, tuning the parameters to produce the best model, and testing the model using new data.

3.1. Dataset

The orange images were taken using a camera phone with 16 Megapixels resolution under the room lighting environment. The orange fruit is placed on a white surface. The five classes of images, namely good-orange-grade-1, good-orange-grade-2, damaged-orange, rotten-orange, and immature-orange are used (shown in figure 1). The pre-processing steps including cropping and resizing the object into 32 × 32 pixel.

We have collected 1000 images consists of 200 images for each class. The k-Fold Cross-Validation is used to distribute the dataset and validate the model. We use $k = 5$, which means that for each class the data will be divided into five folds (F1, F2, F3, F4, and F5). Each fold in each class will consist of 40 data. The distribution of these datasets is as follows: data number 1-40 will be categorized as group
F1, data number 41-80 as group F2, data number 81-120 as group F3, data number 121-160 as group F4, and data number 161-200 as group F5. In addition, the illustration of k-Fold Cross-Validation with \( k = 5 \) is shown in figure 2. There are five data sharing scenarios: D1, D2, D3, D4, and D5. For each scenario, 60% of data will be used as the training data, 20% will be used as the validation data, and 20% will be used as the testing data.

3.2. Model Building

Figure 3 shows our proposed CNN architecture with several different layers, including convolution layer, pooling layer, dropout layer, flatten layer, and layer dense.

Pooling layer aims to reduce the dimensions of the feature map (downsampling) and overcome overfitting. Dropout layer is a regularization technique in a neural network where some neurons will be selected randomly and not applied during the training process. Flatten layer is used to transform the feature map in a multidimensional array into a vector form that will be used as an input for the fully-connected layer.

The difference between traditional neural network and CNN is the convolution layers. The input of the convolution layer is an image which is used for the classification process and the output is the feature map. The convolution layer has three main processes: convolution, activation, and pooling [15]. So, in this research, the convolution process was conducted 4 times to train the model and measure the performance of the model. We used two activation functions: Rectified Linear Unit (ReLU) and Tanh to compare the performance of the model. Based on the study of Krizhevsky, ReLU is preferable, work faster than Tanh without making a significant difference to the generalization of the accuracy [16]. The size of kernel/filter used for each convolution layer is \( 3 \times 3 \) and the size of pooling is \( 2 \times 2 \). To keep the size of the input, the pooling is applied to every output from the convolution filter, which makes the input does not reduce drastically on every process.

The number of filter/kernel used on layer 1 and 2 are 32 filters, while the number of filter on layer 3 and 4 are 64 filters. Layer 3 and 4 need more filters because these layers need to extract the information
from the image which is smaller after the process in the first and second convolution layers. The optimal neurons in the hidden layer used in this research are 256 nodes. The Softmax classifier is applied in the classification to provide more intuitive results and easy to classify probabilistic interpretations for all labels produced. Finally, the model will be implemented using the Keras package with Rstudio.

**Figure 3.** The convolutional neural network model.

4. Results and Discussion

In our experiments, we employed ReLU and Tanh activation function and compared their performance. We use two metrics: loss and accuracy to measure the performance. The loss function is used to penalize the output of the model. It indicates the magnitude of error of the model made on its prediction. We used categorical cross-entropy loss, which is the common setting for classification problems with multi classes. Meanwhile, the accuracy indicates how close a measurement is to the correct value. The accuracies of both activation function for each fold are shown in Table 1-3.

| Data Training | Number of Data | Loss Value | Accuracy Value |
|---------------|----------------|------------|----------------|
|               |                | ReLU       | Tanh           | ReLU | Tanh |
| D1            | 600            | 0.040389   | 0.002499       | 99%  | 100% |
| D2            |                | 0.010151   | 0.001387       | 99%  | 100% |
| D3            |                | 0.004420   | 0.002139       | 100% | 100% |
| D4            |                | 0.088673   | 0.006887       | 97%  | 100% |
| D5            |                | 0.020158   | 0.001592       | 98%  | 100% |
| Average       |                | 0.032758   | 0.002901       | 98.6%| 100% |

The loss value on the training, validation and testing data using the ReLU activation function are 0.032758, 0.239608, and 0.147128. While, Tanh activation function created loss value 0.002901 on training data, 0.198132 on validation data, and 0.204804 on testing data. Both the loss value from each activation function is relatively low. This implies that the results of the proposed models are great.

The accuracy value represents the correct classification rate. It is reasonable while getting a low level of loss and get high accuracy. Table 1 shows 98.6% and 100% accuracy on the training data for both activation functions. On the validation data, the accuracies are 92.8% and 93.2% for both activation functions (see Table 2). Meanwhile on the testing data (see Table 3), the accuracies are 96% and 93.8%
by using ReLU activation function and Tanh activation function, respectively. The more detailed performance of the model along with the increase in the number of the epoch can be seen in figure 4.

Table 2. The accuracy of validation data for each fold.

| Data Validation | Number of Data | Loss Value ReLU | Loss Value Tanh | Accuracy Value ReLU | Accuracy Value Tanh |
|-----------------|---------------|-----------------|-----------------|---------------------|---------------------|
| D1              |               | 0.418156        | 0.159932        | 91%                 | 95%                 |
| D2              |               | 0.032446        | 0.241223        | 98%                 | 90%                 |
| D3              | 200           | 0.371931        | 0.407679        | 90%                 | 90%                 |
| D4              |               | 0.183812        | 0.086802        | 93%                 | 95%                 |
| D5              |               | 0.191694        | 0.095025        | 92%                 | 96%                 |
| Average         |               | 0.239608        | 0.198132        | 92.8%               | 93.2%               |

Table 3. The accuracy of testing data for each fold.

| Data Testing | Number of Data | Loss Value ReLU | Loss Value Tanh | Accuracy Value ReLU | Accuracy Value Tanh |
|--------------|---------------|-----------------|-----------------|---------------------|---------------------|
| D1           |               | 0.098135        | 0.047535        | 98%                 | 98%                 |
| D2           |               | 0.274564        | 0.460625        | 93%                 | 86%                 |
| D3           | 200           | 0.207179        | 0.149645        | 96%                 | 95%                 |
| D4           |               | 0.052459        | 0.048771        | 98%                 | 98%                 |
| D5           |               | 0.103303        | 0.317443        | 95%                 | 92%                 |
| Average      |               | 0.147128        | 0.204804        | 96%                 | 93.8%               |

Figure 4. Loss and accuracy of data training and validation with (a) ReLU Activation Function and (b) Tanh Activation Function

In line with the increasing number of the epoch, generally, the value of the loss is approaching zero. While the value of accuracy is increasing to 1.00 (maximum value) for both data training and data validation. From this, we can conclude that the number of epoch used in the training process has a positive correlation with accuracy. Otherwise, the correlation between the number of epoch and the value of the loss is a negative correlation. Based on this, we can minimize the value of loss by expanding the number of epoch in the training process. Thus we will get a CNN model that produces high accuracy.

However, when we look more closely at figure 4(a), the performance of the training process using ReLU activation function is decreased when the number of epoch approaches 40. It is shown by the
graphic trend of loss value on data validation that increases after the 40th epoch. This phenomenon implied that it is not always necessary to use a large number for the epoch. In our case, the model will give the best performance if we choose 40 as the number of the epoch.

5. Concluding Remarks
In this study, we applied Convolutional Neural Networks to classify the orange images. The CNN model consists of four layers: two pooling layers using $3 \times 3$ kernel size, a softmax layer, and a hidden layer consists of 256 nodes. In the hidden layers, we compared the use of both ReLU and Tanh activation function to obtain the best model performance. According to the experiments, the performance of ReLU is better than Tanh. Tanh produces lower accuracy when we use a new dataset. The classification results of data training by using both activation function are 98.6% and 100%. The accuracies of data validation are 92.8% for ReLU activation function and 93.2% for Tanh activation function. Furthermore, the accuracy of data testing is 96% by using ReLU activation function and 93.8% by using Tanh activation function.

References
[1] Badan Pusat Statistik 2017 Statistik tanaman buah-buahan dan sayuran tahunan Indonesia Badan Pusat Statistik 05120.1807 p 46
[2] Nainggolan C I, Tarigan K and Salmiah 2013 Analisis usaha tani jeruk dan faktor-faktor yang mempengaruhi penerimaan petani J. Agriculture and Agribusiness Socioeconomics 2
[3] LeCun Y, Bengio Y and Hinton G 2015 Deep learning Nature Int. J. of Science 521 pp 436-44
[4] Romadhon A S and Widyaningrum V T 2015 Klasifikasi mutu jeruk nipis dengan metode learning vector quantization (LVQ) Jurnal Ilmiah Rekayasa 8 pp 121–8
[5] Agustian W, Setyaningsih S and Qur’ania A 2015 Klasifikasi buah jeruk menggunakan metode naive bayes berdasarkan analisis tekstur dan normalisasi warna Program Studi Ilmu Komputer FMIPA - Univ. Pakuan 2
[6] Capizzi G, Sciuto G L, Napoli C, Tramontana E and Wozniak M 2016 A novel neural networks-based texture image processing algorithm for orange defects classification Int. J. Computer and Science Applications 13 pp 45–60
[7] Sabzi S, Abbaspour-gilandeh Y and Garcia-Mateos G 2018 A New approach for visual identification of orange varieties using neural networks and metaheuristic algorithms Information Processing in Agriculture 5 pp 162–72
[8] Wayan S E P I, Arya Y W and Soelaiman R 2016 Klasifikasi citra menggunakan convolutional neural network (CNN) pada caltech 101 Jurnal Teknik ITS 5 pp 65–9
[9] LeCun Y, Bottou L, Bengio Y and Haffner P 1998 Gradient-based Learning Applied to Document Recognition vol 86 pp 2278-324 (USA: IEEE)
[10] An T V, Tran H S and Le T H 2017 Advertisement image classification using convolutional neural network The 9th Int. Conf on Knowledge and Systems Engineering 1 pp 197–202
[11] Purwaningsih T, Anjani I A and Utami P B 2018 Convolutional neural networks implementation for chili classification 2018 Int. Symp. on Advanced Intelligent Informatics (SAIN) pp 190-4
[12] Khaing Z M, Naung Y and Htut P H 2018 Development of control system for fruit classification based on convolutional neural network 2018 IEEE Conf. Russ. Young Res. Electr. Electron. Eng. EIConRus.2018.8317456 pp 1805–7
[13] Goodfellow I, Bengio Y and Courville A 2016 Deep Learning Book (United States: MIT Press) p 326
[14] Suyanto 2018 Machine Learning: Tingkat Dasar dan Lanjut (Bandung: Informatika) p 350
[15] Rismiyati and SN A 2016 Convolutional neural network implementation for image-based salak sortation 2nd Int. Conf. Sci. Technol. (ICST) ICSTC.2016.7877351 pp 77-82
[16] Krizhevsky A, Sutskever A and Hinton G E 2012 ImageNet classification with deep convolutional neural networks Adv. Neural Inf. Process. Syst. 8 pp 713–72