Maturity Detection of Crystal Guava Fruit Using Convolutional Neural Network Algorithm

[1*]Toriq Fatur Rohman, [2]Suhartono
[1,2]Magister Informatika, UIN Maulana Malik Ibrahim Malang, Malang 65144 - Indonesia
[*]*230605210005@student.uin-malang.ac.id*

ABSTRACT
Guava fruit in Latin is called Psidium Guajava L. is a tropical plant originating from Brazil, and distributed to Indonesia. One of the fruit commodities found in Indonesia which is a leading commodity and continues to increase in production is guava. Guava is a climacteric fruit due to chemical changes, namely the activity of the enzyme pyruvate which causes an increase in the amount of acetaldehyde and ethanol so that CO2 production increases and ethylene produced during fruit ripening will increase the respiration process. This research focuses on crystal guava fruit which is very commonly found in Indonesia, especially in East Java Province. Where at the time of harvesting guava fruit of course have a different maturity. From these problems, researchers used the Convolutional Neural Network (CNN) Algorithm to identify ripe guava fruit (Psidium Guajava L.) with image classification. From this study, the CNN model using the VGG16 architecture that has been created in this study has an accuracy of 96% which is approximately the same as the comparison of the accuracy of other CNN models and gets good performance on the testing model with an accuracy rate of 83%.

Keywords: Convolutional Neural Network, identification, VGG-16

1. INTRODUCTION

Guava fruit in Latin called Psidium Guajava L. is a tropical plant originating from Brazil and spread to Indonesia through Thailand. Guava fruit is round or oval in shape with light green skin. This fruit is one of the fruits that has good benefits for the body because of the many nutrients contained in the fruit. Guava fruit contains high phenolic compounds and flavonoids. These compounds are effective as antidotes to DPPH radicals [1]. The type of fruit that we know so far is divided into two groups, the first climatic group and the second non-climatic group, where the climatic fruit will only hold a respiration reaction when ethylene is given at the pre-climatic level and is no longer sensitive to ethylene after the increase in respiration begins [2]. Guava is a climatic fruit due to chemical changes, namely the activity of the enzyme pyruvate which causes an increase in the amount of acetaldehyde and ethanol so that CO2 production increases and ethylene produced in ripening fruit will increase the respiration process. Therefore, the fruit must be harvested in a mature condition and until now there are still many farmers harvesting guava fruit simultaneously without seeing the condition of the maturity of the fruit. Whereas at the time of harvesting guava fruit certainly has a different maturity.

Where at the time of harvesting guava fruit certainly has different maturities. The results of the researcher's observation to guava fruit farmers, stated that this plant was able to bear fruit when it was 2-3 months old even though it was planted from seeds which caused the maturity of guava fruit to vary and the green fruit skin when the fruit was young and old would make it difficult for farmers to identify the level of fruit maturity. Therefore, technology is needed that can facilitate farmers to identify the maturity of guava fruit. Talking about the technology used to identify guava fruit, researchers focus on image processing, which later from a photo can produce identifications about image processing. Several methods can be used to process an image, namely using the Convolutional Neural Network (CNN) algorithm. Researchers chose this method because this method is widely used by various researchers to classify images and is very accurate in identifying two-dimensional images [3] [4] [5] [6].

This research is related to the implementation of CNN (Convolutional Neural Network) to identify the maturity of crystal guava fruit. Some research uses the Convolutional Neural Network (CNN) algorithm to process an image and get a good accuracy value, namely the Design of Deep Learning Algorithm Convolutional Neural Network (CNN) for Image Classification of Arabica and Robusta Coffee Types [7][8]. The research was made by Agricultural Engineering Students of Jember University which has similarities in one of the characteristics [3], namely Convolutional Neural Network. In the study, the CNN model used has the provisions of input data measuring 224x224 pixels, convolution layers four times, kernel size 3x3, train data totalling 15,984, testing data totalling 3,996 and producing parameters 12,944,034 neurons in the training model. From the results of these provisions, the accuracy rate is 100% of the training data and 98.8% of the testing data. This accuracy is obtained when the learning rate is 0.00001 and epoch 100. The dataset used in this study was taken in the crystal guava fruit plantation area in Tembokrejo Village, Gumuk Mas District, Jember Regency with coordinates - 8.2709613 LS, 113.4369311 BT. The CNN architecture used is VGG16. In this research, the CNN model used is VGG-16 by using 224×224 for the input image size. It is expected that the VGG16 architecture can identify the type of maturity of crystal guava fruit by getting high accuracy, as well as with a fast model training process compared to previous researchers.

2. RESEARCH METHOD

Convolutional Neural Network (CNN) is an algorithm developed from Multilayer Perceptron (MLP) and is one of the deep learning algorithms that can train large datasets with millions of parameters on two-dimensional data, such as images and sounds [9]. LeCun was the first to introduce CNNs, inspired by previously established neurocognition models. LeNet is a form of CNN that was first applied by LeCun in the problem of handwritten number recognition. Convolutional Neural Network (CNN) has a layer with a three-dimensional arrangement of neurons (width, height, depth). Width and height are measurements based on the layer, while depth itself is set at the number of layers. The layer then sequentially goes through a series of processes. Processing can include the convolution layer, pooling layer, normalization layer, fully connected layer, loss layer, and so on [10].

In general, the layers in CNN are divided into two types, namely the feature extraction layer and the classification layer. In the first layer (feature extraction layer) is located at the beginning of the architecture over several layers and each layer is composed of neurons connected to the local region of the previous layer. The convolutional layer is the first type of layer and the pooling layer is the second layer. An activation function is applied in each layer. In this feature extraction layer receives input in the form of an image directly by processing it so that the output is a vector to be processed in the next layer vector to be processed in the next layer. This research aims to design a modified classification of guava fruit maturity automatically using the Convolutional Neural Network (CNN) algorithm and find out how the accuracy level is obtained from the classification using the Convolutional Neural Network (CNN) algorithm [11]. This research has a flow chart framework process to achieve these objectives, as shown in Figure 1.

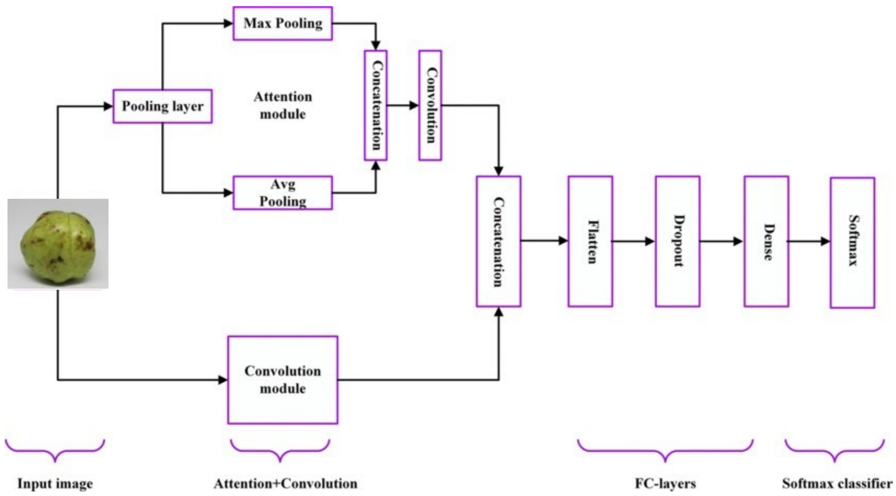

Figure 1. Flow of Research Method

2.1. Data Collecting

The author conducts direct observation with crystal guava farmers on their crystal guava plantations to collect datasets in the form of crystal guava image images. The first step is to determine the size of the crystal guava fruit that will be used for the CNN system process with a provision of 30 days. Data acquisition was conducted using a Canon EOS 4000D camera. Previously, guavas were picked from the tree and collected, then placed in a studio box. The dataset will be grouped into two groups with a predetermined arrangement. From the direct imaging of guava crystal objects, 184 images were obtained for the dataset.

2.2. Pre-processing

At this stage is a process to prepare the crystal guava dataset before other processes are carried out. Image data from the crystal guava dataset is processed until it is ready to be used by the CNN architecture built. Data pre-processing receives input from image data, then the data is divided into two parts, namely testing data and training data. Next, the image is transformed by resizing the image pixels.

2.3. Modelling

This dataset was created by collecting all the data obtained from live portraits of individual crystal guava fruits. Two groups of datasets were used: ripe and unripe. From all the live shots of the object, it was divided into two parts of 80: 20. Of these, 80% are training data and 20% are test data. The coefficient distribution in the dataset is because the training data should be larger than the test data to improve the performance of the CNN model, but the amount of test data is too small to be larger than the test data. This is so as not to reduce the quality of accuracy.

2.4. Model Evaluation

In the model training stage, this dataset is created by collecting all the data obtained from the portrait of the crystal guava fruit object directly. The dataset used has two groups, namely ripe and unripe. All the results of the object portrait are directly divided into two parts 70:30, where 70% for training data and 30% for testing data. Therefore, each dataset consists of mature and raw types that will be used to train CNN. Where this training data will later be used to train the CNN model algorithm for identification and test data is used to test the performance of the trained model. The material used in this study consists of 180 ripe crystal guava fruit datasets and 180 unripe crystal guava fruit datasets. The division of the dataset ratio is because the training data must be more than the testing data to improve the performance of the CNN model, but the amount of test data is not too small so as not to reduce the quality of testing.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$Sensitivity\ (Recall) = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

Where TP is True Positive (True detected positive data), TN is True Negative (Number of correctly detected negative data), FP is False Positive (Negative data but detected as positive data), and FN is False Negative (Positive data but detected as negative data).

3. RESULT AND ANALYSIS

3.1. Data Collecting

The Convolutional Neural Network algorithm in the network architecture can affect the accuracy of the model created. Then the architecture can produce different accuracy or performance. In the image identification stage using the CNN algorithm, training data that has high accuracy will be included first. The data that will be included in this training process is data that is ready to be processed for the trial stage. Previously, it was necessary to identify 3 learning parameters. The parameters that must be identified to train the architecture in this research are learning rate, batch size, and epochs.

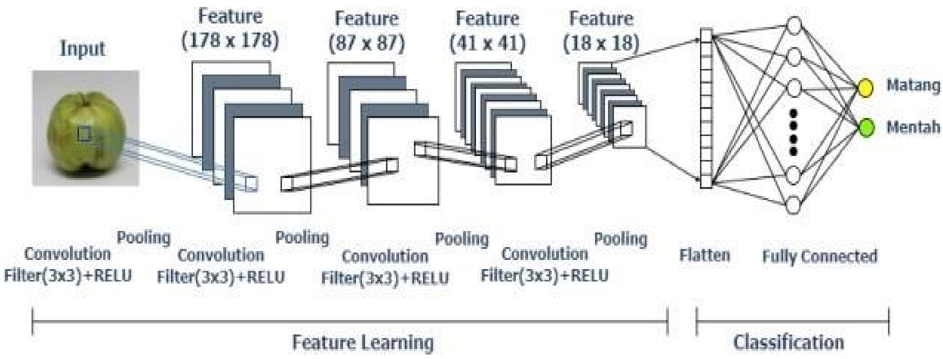

Figure 2. Convolutional Neural Network Architecture Design

Shown in Figure 2 is the CNN architecture design in the training process to identify crystal guava fruit. The architecture involves convolution, pooling, flattening or fully connecting, and sigmoid activation processes. This research uses an input data size of 800x600x3. The length is 800 pixels, the height is 600 pixels and the thickness or number of three indicates the number of channels in the image (Red, Green, Blue). In the first convolution process using a 3x3 kernel with the number of filters 16. Then the ReLu activation function is added, which aims to convert negative values to zero (eliminate negative values in a convolution matrix). The convolution output then enters the pooling process, which is used to reduce the matrix size. This research uses max pooling to get the new pooled matrix value. The result is that the new matrix from the first convolution process is 178x178x16. Next, the feature map will be activated using ReLu.

The second convolution process continues the first convolution result with an input matrix of 87x87x32. Just like before, this second convolution process also uses the ReLu activation function. The third convolution process uses a 41x41 matrix input, a total of 64 filters and a 3x3 kernel size. Just like the previous process, it uses the ReLu activation function and uses max pooling. In the fourth convolution process, the input matrix is 18x18, using 128 filters with a 3x3 kernel size. Then add the ReLu activation function and proceed to the max pooling process. The following is a table of model structures formed from the training process.

Table 1. Test Results Using Data Testing

No	Name	Size	Parameter
0	Input	(800, 600, 3)	0
1	max_pooling2d_2	(None, 28, 28, 256)	0
2	conv2d_7	(None, 28, 28, 512)	1180160
3	batch_normalization_7	(None, 28, 28, 512)	2048

4	conv2d_8	(None, 28, 28, 512)	2359808
5	batch_normalization_8	(None, 28, 28, 512)	2048
6	conv2d_9	(None, 28, 28, 512)	2359808
7	batch_normalization_9	(None, 28, 28, 512)	2048
8	max_pooling2d_3	(None, 14, 14, 512)	0
9	conv2d_10	(None, 14, 14, 512)	2359808
10	batch_normalization_10	(None, 14, 14, 512)	2048
11	conv2d_11	(None, 14, 14, 512)	2359808
12	batch_normalization_11	(None, 14, 14, 512)	2048
13	conv2d_12	(None, 14, 14, 512)	2359808
14	batch_normalization_12	(None, 14, 14, 512)	2048
15	max_pooling2d_4	(None, 7, 7, 512)	0
16	Flatten	(None, 25088)	0
17	Dense	(None, 4096)	102764544
18	Dropout	(None, 4096)	0
19	dense_1	(None, 4096)	16781312
20	dropout_1	(None, 4096)	0
21	dense_2	(None, 1)	4097
	Total		134,281,537

Based on Table 2 above, this research forms the structure of the CNN model. To calculate the convolution input value using the formula: $input\ size + (2 \times padding) - (filter\ size - 1)$. The total parameters formed from the model are 134,281,5374 neurons.

3.2. Pre-processing

The validation process uses 360 images of test data, which are divided into 2 classes, namely the raw class as many as 180 images and the ripe class as many as 180 images, in the process of obtaining accuracy is done using a confusion matrix calculation. In identifying the level of maturity of the crystal guava fruit contained in the image of the object portrait directly. This can be seen in the table 2.

Table 2. Prediction Results of The Validation Process

Confusion Matrix		Predict Class	
		Mature	Raw
	Mature	151	29
Actual Class	Raw	29	151

Based on the prediction table above, the model prediction results show quite good results compared to the verification data. Among the 180 tested maturity images, 29 data were wrongly predicted, while in the 180 tested original images, only 29 data had wrong predictions. From the table above, the accuracy can be calculated to be 83%. The training accuracy value is the value of calculating the accuracy of the training dataset and the prediction of the model. Accuracy shows the ability of the model to classify the dataset correctly [12]. [Loss indicates the number of errors that occur for each example in the dataset [12]. The results of data processing using the CNN model design show different accuracy values for each treatment. Different treatments are learning rate and epoch. The difference in treatment is done to compare which model is the best by paying attention to the parameter value. There are two parameter values in testing model training data, namely accuracy train and loss train. A good model has a high accuracy value with a low loss value [12].

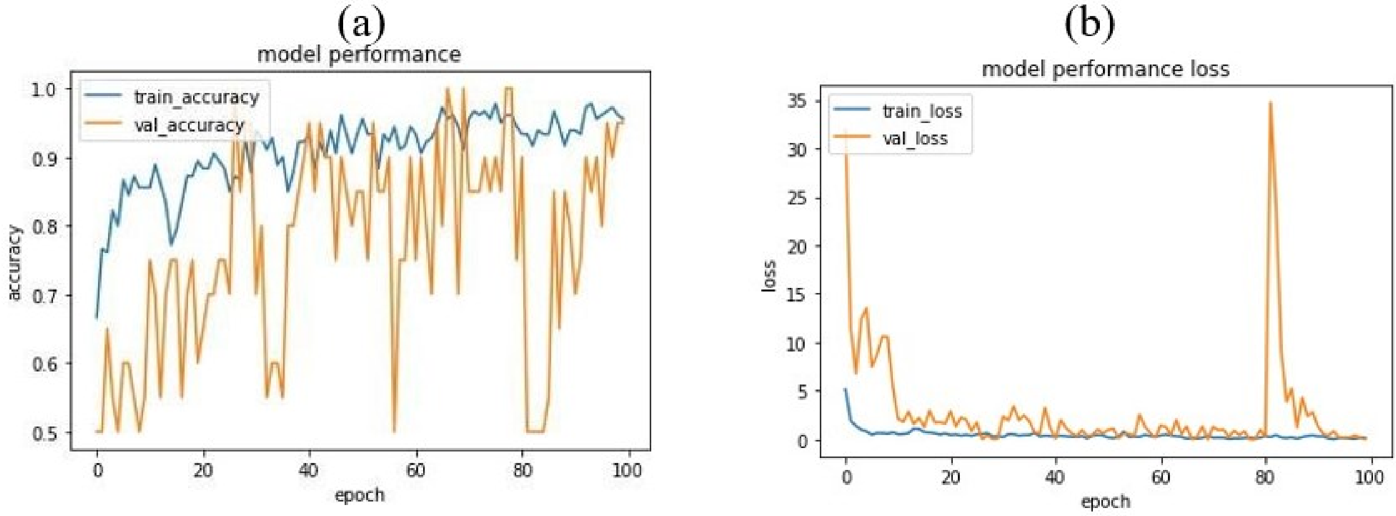

Figure 3. Training graph epoch 100, (a) training accuracy, (b) training loss

It is shown in Figure 3 that the training accuracy value at epoch = 100 graph is relatively constant. While the training loss graph fluctuates, but not too large. The training accuracy value of epoch = 100 occurs at a learning rate of 0.0001 which is 96% with a training loss of 0.140827. In the learning process the dataset will be rotated until it returns to its initial position in one round which is done repeatedly, the process is called epoch. No research can claim the best epoch range in the learning process [13]. So that each study has a different best epoch value. The learning process also requires optimization methods such as Stochastic Gradient Descent (SGD). This research uses Adam's optimization. Optimization is carried out by learning and updating weights on neurons in the fully connected layer. With the existence of a learning rate value parameter, the weight update process can be more optimal. The training accuracy value is obtained when the learning rate is 0.0001, which is 96% with a training loss of 0.140827.

4. CONCLUSION

Based on the results of an analysis during the design, implementation, and testing of the system, the following conclusions can be drawn The CNN model in this study uses input data with a size of 224x224 pixels, four convolution layers, 3x3 kernel size, 300 training data while 60 testing data and produces neuron parameters with a total of 134,281,537 in the training model. From several learning rate and epoch parameters, the best accuracy performance is obtained during training which is 96% with a training loss of 0.14. This accuracy is obtained when the learning rate is 0.0001 and epoch 100. Parameter determination is done by trial and error to get optimal results, because each classification has certain parameters and is different from one another.

REFERENCES

[1] R. A. D. Widyastuti et al., “Fruit Quality of Guava (Psidium guajava ‘Kristal’) under Different Fruit Bagging Treatments and Altitudes of Growing Location,” JTCS, vol. 9, no. 01, pp. 8–14, Feb. 2022, doi: 10.29244/jtcs.9.01.8-14.

[2] Pantastico, E.R.B. 1993. Fisiologi Pasca Panen, Penanganan dan Pemanfaatan Buah Buahan dan Sayuran Tropika dan Subtropika. Terjemahan Komeriyani.UGM Press, Yogyakarta https://balaiyanpus.jogjaprov.go.id/opac/detail-opac?id=87312

[3] Ilahiyah, S., & Nilogiri, A. 2018. "Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network". JUSTINDO (Jurnal Sistem Dan Teknologi Informasi Indonesia), 3(2), 49-56.D. Pramadhana, http://jurnal.unmuhjember.ac.id/index.php/JST/article/view/7384

[4] Sarraf, A. (2020). Binary Image Classification Through an Optimal Topology for Convolutional Neural Networks. American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 68(1), 181–192.H. https://asrjetsjournal.org/index.php/American_Scientific_Journal/article/view/5938.

[5] N. R. Y. Putri, I. Khoiriyah, and Suhartono, Bibliometrics Using VOS viewer with Publish or Perish: Neural Network Research in Plant Disease Classification, Journal of Informatics and Science Media, vol. 1, no. 1, pp. 38–51, Jul. 2024, [Online]. Available: https://jismandiri.com/index.php/jisman/article/view/8/10.

[6] A. Damayanti, R. Murtiani, and Suhartono, “Mapping Soft Computing Concepts and Research Trends in Computer Vision through Bibliometric Analysis Using VOS Viewer and Publish or Perish,” Journal of Informatics and Science Media, vol. 1, no. 1, pp. 52–63, Jul. 2024, [Online]. Available: https://jismandiri.com/index.php/jisman/article/view/9/9.

[7] Arrofiqoh, E. N., & Harintaka, H. (2018). Implementasi Metode Convolutional Neural Network Untuk Klasifikasi Tanaman Pada Citra Resolusi Tinggi. Geomatika,24(2),6168. https://doi.org/10. 24895/jig.2018.24-2.810

[8] Setiawan, H. H. 2018. "Klasifikasi Jenis Buah Pisang Dengan Image Processing Menggunakan Metode Backpropagation". Skripsi S1, Universitas Sanata Dharma https://ojs.jurnaltechne.org/index.php/techne/article/view/231

[9] Chauhan, R., Ghanshala, K. K., & Joshi, R. C. 2018. Convolutional neural network (CNN) for image detection and recognition. In 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC) (pp. 278- 282). https://ieeexplore.ieee.org/document/8703316

[10] Albelwi, S., & Mahmood, A. (2017). A framework for designing the architectures of deep Convolutional Neural Networks. Entropy,19(6). https://doi.org/10.3390/e1 9060242

[11] Shi, J., Wu, J., Li, Y., Zhang, Q., & Ying, S. (2017). Histopathological Image Classification with Colour Pattern Random Binary Hashing-Based PCANet and Matrix-Form Classifier. IEEE Journal of Biomedical and Health Informatics,21(5), 1327–1337. https://ieeexplore.ieee.org/document/8037745

[12] Wikarta, A., Pramono, A. S., & Ariatedja, J. B. (2020). Analisa Bermacam Optimizer Pada Convolutional Neural Network Untuk Deteksi Pemakaian Masker. Seminar Nasional Informatika 2020 (SEMNASIF 2020), 2020 (Semnasif), 69–72. http://jurnal.upnyk.ac.id/index.php/semnasif/article/view/4087.

[13] Öztürk, Ş., & Akdemir, B. (2018). Effects of Histopathological Image Pre-processing on Convolutional Neural Networks. *Procedia Computer Science*, *132*(Iccids), 396–403. https://dx.doi.org/10.1016/j.procs.2018.05.166