JOSS:
Journal of Social Science

Volume 2 Number 8 August 2023
p-ISSN 2963-1866-e-ISSN 2963-8909

IMAGE CLASSIFICATION OF MANGOES USING CNN VGG16 AND ALEXNET

Aldo Leviko Marshal[1], Ahmad Nurul Fajar[2]
Bina Nusantara University Jakarta, Indonesia
Email: aldo.marshal@binus.ac.id[1], afajar@binus.edu[2]

KEYWORDS
Classification, CNN, VGG16, AlexNet, Deep Learning, Machine Learning, Mangoes

ABSTRACT
Mango is one of the fruits that is often consumed by Indonesian people and is the fruit with the third largest amount of production in Indonesia, but currently, there are obstacles in the export of mangoes in Indonesia where fruit conditions and regulations weaken the mango export process in Indonesia. This study focuses on comparisons and classifications in the use of the CNN VGG16 and AlexNet architectures for classifying arum manis mango images. In this research, 4 images of Arum Manis mangoes will be used with a dataset of 400. With the results of the classification carried out with the two architectural models, CNN VGG16 provides a high accuracy of 92.50% with the use of epoch 50 while the use of AlexNet gets an accuracy of 79.64%. This research was conducted to provide a solution to overcome the quality of mangoes that are not in accordance with the production standards of mangoes to be exported.

INTRODUCTION

Indonesia is a country that has high potential in agriculture, such as fruit, vegetable, and ornamental plant plantations (Vyas, Talati, & Naik, 2014). There are several agricultural sectors in Indonesia, one of which is plantations that focus on fruit production. One of the fruits in Indonesia which produces quite a lot is mango (Arya & Singh, 2019). This fruit is often consumed in people's daily lives because the fruit brings sweetness to the taste. Based on data from the Agricultural Data Center and Information System regarding mango consumption in the household, it is stated that the amount of consumption fluctuates. The average mango consumption rate in 2018 – 2022 is 8.25 kg/capita/week. In 2018 consumption reached 10.47 kg/capita/week, but decreased drastically in 2019, consumption per capita decreased and fluctuated again until 2022. Consumers certainly want fruit that is fresh and has benefits and durability in the fruit (Kazi & Panda, 2022).

Mango is a fruit that is easy to find in Indonesia. This fruit is usually found in the plains often consumed are the presence of some of the ingredients contained in these mangoes. Some of the ingredients contained in the fruit based on (Pham, Van Tran, & Dao, 2020) include a fiber content of 1.8%, there are Vitamin C, and Vitamin E. Mango fruit also has a source of minerals in the form of potassium, iron, copper, calcium, phosphorus, zinc, manganese, and selenium. Then the mango fruit also has a high- water source, so it is often used as juice by taking juice. Mango fruit is a fruit that has a high level of production in Indonesia. According to the Central Statistics Agency (BPS), there was a production of 2.89 million tons of mangoes in 2020. The number in that year has increased from the previous year when there was a total production of 2.80 million tons of mangoes in 2019 (Ashok & Vinod, 2016; Nagaraju, Sahana, Swetha, & Hegde, 2020).

Seeing the amount of production of these mangoes, mangoes still have a problem, namely the export process of the fruit (Kumar et al., 2021; Roomi, Priya, Bhumesh, & Monisha, 2012). According to researchers at the Center for Tropical Horticultural Studies, IPB, they said that the conditions of the mangoes and regulations were the reasons for Indonesia's weak mango exports (Aherwadi et al., 2022). There are mangoes that can compete, such as the Arum Manis mango which has a good taste compared to some other types of mangoes, however, the Arum Manis mango does not have a long shelf life making it difficult to store for a long period of time (Sahu & Dewangan, 2017; Singh, Chouhan, Jain, & Jain, 2019). This is due to the starch content of the arum manis mango which is quite small (Mahmood, Singh, & Tiwari, 2022). Then there are differences in the characteristics of each type of mango where the selection of gedong mangoes is more desirable when served but when presented in pieces, the sweet arum mangoes are more desirable (Garillos-Manliguez & Chiang, 2021). Meanwhile, the color of the fruit possessed by the arum manis mango is not attractive to people because the color of the mango is brownish green, which can be said to be an unripe or rotten fruit, while the other fruits have bright colors (Prabu & Chelliah, 2022). As for the use of vapor heat treatment (VHT) or hot steam technology which can help in killing pests or flies on mangoes as a condition for exporting mangoes, their use in Indonesia is still not optimal so it is still a big obstacle for mango exports (Ansah, Amo-Boateng, Siabi, & Bordoh, 2023).

Knowing the current conditions for selecting mangoes with more advanced technology, you should no longer make selections manually, but the use of machine learning will provide great assistance in determining the quality of these mangoes. The selection of quality mangoes can be determined by the color of the mangoes.

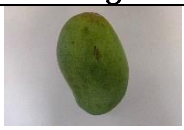
The purpose of this study is to build and train a CNN model based on VGG16 and AlexNet that is able to distinguish various types of mangoes based on visual characteristics in the image. The main goal is to achieve a high degree of accuracy in classifying mangoes.

METHOD

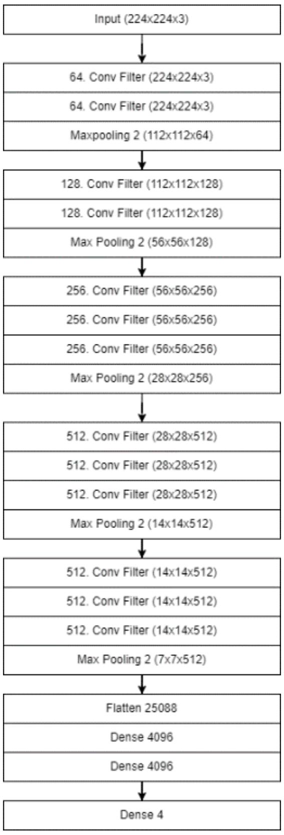
The classification that will be carried out will use tools from jupyter notebook (anaconda) to do data renaming and data resizing and google collab which will be used for classifying mango fruit images and modeling from CNN VGG16. There are stages that will be carried out in the classification of this mango. The first stage is data preprocessing, at this stage the use of mangoes used is Arum Manis mango. Taking pictures of mangoes will use the smartphone camera and then using white paper as a background for the mangoes.

Then make considerations regarding other factors that affect the results of capturing photos such as lighting on objects which is also one of the determining factors for the image of arum manis mangoes, then other factors such as placing a smartphone on a stable surface so that taking photos is higher quality and accurate. To overcome the lighting factor, the photos will be taken in a place with a high focus on lighting. The number of datasets taken is 400 data which has classifications including immature mangoes, grade 1 ripe mangoes, grade 2 ripe mangoes, and rotten mangoes. Each of the mango classifications has 100 datasets divided into 2 types of data, namely training data and testing data with a ratio of 70:30. The following are the variables of the sweet arum mangoes that will be used:

Table 1
Variables and descriptions of the 4 classifications of arum manis mango

Image	Variable	Description
	Not Ripe	Unripe mangoes with a dark green color. The condition of mangoes that are not yet ready to be consumed
	Ripe Grade 1	Ripe mangoes with a slightly yellowish-green color. The condition of mangoes suitable for consumption
	Ripe Grade 2	Mangoes that are still ripe with a long maturity level. It has a yellowish-green color with a yellow color which is quite dominant in mangoes. The condition of mangoes that are still suitable for consumption.
	Rotten	Mangoes with a past degree of ripeness. It Has a blackish brown color to the mango. The condition of mangoes that are unfit for consumption.

In classifying the mango image, the image size is 224x224 with a max. pool 2x2. Then the kernel size is 3 with convolution filter stages starting from 64,128,256,512,512. Then for the flattened layer, there are dense 25088,4096,4096, then 4. The number of epochs to be used is 10, 20, and 30. The following is an overview of the structure of the CNN VGG16 design:



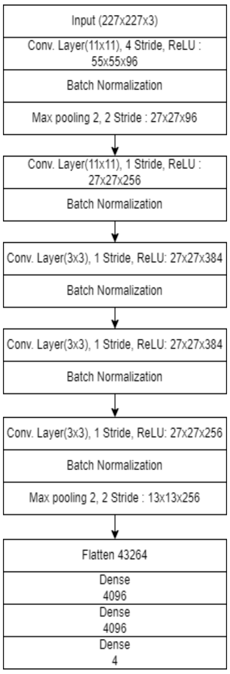
As for the design of the Convolutional Neural Network VGG-16 in this study, it has a reference based on the research that has been done by [11]. In the design of the Convolutional Neural Network in this study, there would be five times for the convolution process where each convolution is carried out with the aim of training the model and monitoring the performance of the model that has been designed. Then the design in the form of adding max pooling is carried out based on previous studies as a reference. Then each stage in the convolution layer will be in the form of a feature map.

The first convolution process uses 64 filters with a core size of 3x3. After that max pooling of size 2x2 will be used. The second and third convolution processes use 128 filters with a kernel size of 3x3. The fourth, fifth, and sixth convolution processes use 256 filters with the same kernel size as the previous convolution process. 512 filters are used in the convolution processing from the 7th layer to the 12th layer. The use of 512 filters is divided into two parts.

On the first part containing the 7th convolution process to the 9th convolution, a kernel of size 3x3 is used, followed by a max pooling process of size 2x2. was announced. After completing the convolution stage of the first part, we can proceed to the second part, the convolutions of the 10th to 12th convolution layers. Again, the use of 512 filters with the same conditions as in the previous section to extract the images. The convolution process in this work

In designing the AlexNet architecture using 5 stages of convolution starting with an image input with a size of 227x227. The first convolution uses filter 96, then the 11x11 kernel with stride 4 and the activation function used is ReLU. Then perform batch normalization to reduce the possibility of a significant decrease in accuracy and perform max pooling 2x2.

Then in the second convolution using filter 256, kernel with size 11x11, using stride 1, ReLU activation function and followed by batch normalization stages. The third and fourth convolutions use a 384 filter, then use a 3x3 kernel with stride 1, the ReLU activation function then performs batch normalization. The fifth convolution uses filter 256 with 3x3 kernel, stride 1, the activation function is used by ReLU, then batch normalization and max pooling 2x2 with stride 2. At the Flatten layer stage, neurons are used gradually starting from 43264, 4096, 4096. The following is an overview of the AlexNet design structure that will be used:



RESULT AND DISCUSSION

For the classification with the VGG16 model for the mango image, it is divided into 5 parts where classification is carried out with the number of epochs, which is 10, 20, 30, 40, and 50. For batch use in this experiment using size 32, with kernel size 3x3, image size 224x224. The results obtained in classifying mango images using the CNN VGG16 model can be seen in the following table:

Table 2.
VGG16 Accuracy Result

	Epoch	Test Accuracy	Loss	Train Accuracy
1	10	77.49%	0.5224	100%
2	20	98.33%	0.2926	100%
3	30	90.83%	0.1940	100%
4	40	90.83%	0.2024	100%
5	50	92.50%	0.1403	100%

Based on the results shown in table 1 the accuracy obtained in epoch 10 is 77.49% with a total loss of 0.5224. Then on epoch 20 experiments managed to get an increase in accuracy to 88.33% with the loss of a number is 0.2926. For epoch 30, an accuracy of 90.83% is obtained with a loss of 0.1940. At epoch 40 the accuracy is still the same as the experiment using epoch 30 which is 90.83% with the amount of loss being 0.2024. Then in the epoch 50 experiments, you get an increase in accuracy to 92.50% with a loss of 0.1403. These results illustrate the effect of epoch changes that have little impact on the accuracy of results obtained.

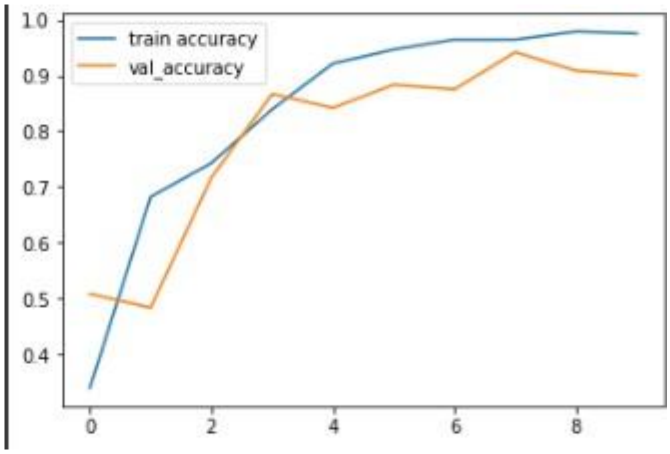


Figure 3
Graph of Accuracy Data train and Data test of VGG16

As for the classification with AlexNet model will also use 5 kind of epoch which is 10, 20, 30, 40, and 50. The results obtained in classifying mango images using the AlexNet model can be seen in the following table:

Table 3
AlexNet Accuracy Result

	Epoch	Test Accuracy	Loss	Train Accuracy
1	10	25.00%	256.63	25.00%
2	20	26.07%	70.57	22.49%
3	30	41.78%	26.36	18.33%
4	40	67.14%	3.50	75.00%
5	50	79.64%	0.49	79.16%

The accuracy results are shown in Table 3 where the first experiment used epoch 10 to get a low accuracy value of 25.00% with a loss of 256.63, then there was a slight increase in accuracy in the experiment the second uses epoch 20 which is equal to 26.07% with a loss of 70.57. On the third trial using epoch 30 experienced an increase in accuracy to 41.78% with losses of as much as 26.36. At epoch 40, the accuracy again increased, namely to 67.14% with a loss of 3.50. In the fifth experiment, the accuracy increased to 79.64% with a loss of 0.49.

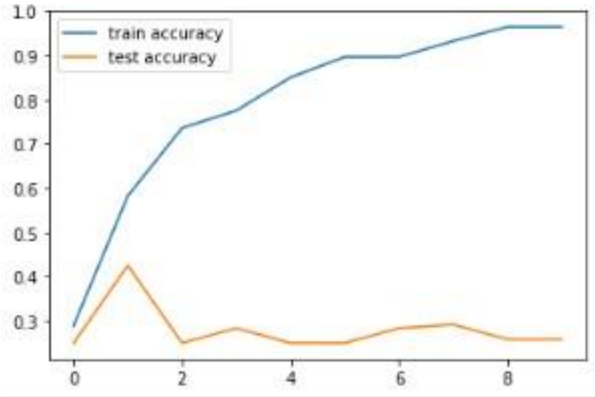


Figure 4
Graph of Accuracy Data train and Data test of AlexNet

The results obtained based on the use of the two architectures are shown using a confusion matrix. The following is a model of the confusion matrix based on the use of the two CNN architectures shown in Table 2 for the use of the VGG16 architecture and table 3 for the use of AlexNet.

Table 4
VGG16 Confusion Matrix

VGG16	NR	RG1	RG2	RO	TOTAL
NR	30	0	0	0	30
RG1	2	19	9	0	30
RG2	0	0	30	0	30
RO	0	0	0	30	30
TOTAL	32	19	39	30	120

Table 5
AlexNet Confusion Matrix

AlexNet	NR	RG1	RG2	RO	TOTAL
NR	25	0	5	0	30
RG1	3	13	14	0	30
RG2	0	0	0	30	30
RO	0	0	0	30	30
TOTAL	28	13	49	30	120

Note:
NR: Not Ripe
RG1: Ripe Grade 1
RG2: Ripe Grade 2

RO: Rotten

The following are the results of the prediction of Arum Manis mangoes based on their class. By using 120 datasets as test data, the results obtained from both architectures are shown in Table 6 for VGG16 and Table 7 for AlexNet.

Table 6
Mango Prediction Result of VGG16

VGG16	Match	Not Match
NR	30	0
RG1	19	11
RG2	30	0
RO	30	0
TOTAL	109	11

Table 7
Mango Prediction Result of AlexNet

AlexNet	Match	Not Match
NR	25	5
RG1	13	17
RG2	30	0
RO	30	0
TOTAL	98	22

CONCLUSION

The use of the CNN VGG16 architecture in classifying arum manis mangoes gives better results than the AlexNet architecture. After classifying the mango image using the CNN VGG16 and AlexNet architectural methods with the epochs used, namely 10, 20, 30, 40, and 50, it shows a good accuracy value, reaching 92.50% with a loss of 0.1403, while Alexnet get an accuracy of 79.64% with a loss of 0.49. The accuracy value obtained by the VGG16 architecture can be said to be quite good with little data loss, while the AlexNet architecture still has an accuracy that is not high enough. So it can be concluded that the use of CNN VGG16 in mango image classification provides a fairly good level of accuracy.

REFERENCES

Aherwadi, Nagnath, Mittal, Usha, Singla, Jimmy, Jhanjhi, N. Z., Yassine, Abdulsalam, & Hossain, M. Shamim. (2022). Prediction of fruit maturity, quality, and its life using deep learning algorithms. *Electronics*, *11*(24), 4100. Google Scholar

Ansah, Francisca Aba, Amo-Boateng, Mark, Siabi, Ebenezer K., & Bordoh, Paa Kwesi. (2023). Location of seed spoilage in mango fruit using X-ray imaging and convolutional neural networks. *Scientific African*, *20*, e01649. Google Scholar

Arya, Sunayana, & Singh, Rajeev. (2019). A Comparative Study of CNN and AlexNet for Detection of Disease in Potato and Mango leaf. *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, *1*, 1–6. IEEE. Google Scholar

Ashok, Vani, & Vinod, D. S. (2016). A comparative study of feature extraction methods in defect classification of mangoes using neural network. *2016 Second International Conference on Cognitive Computing and Information Processing (CCIP)*, 1–6. IEEE. Google Scholar

Garillos-Manliguez, Cinmayii A., & Chiang, John Y. (2021). Multimodal deep learning and visible-light and hyperspectral imaging for fruit maturity estimation. *Sensors*, *21*(4), 1288. Google Scholar

Kazi, Aafreen, & Panda, Siba Prasada. (2022). Determining the freshness of fruits in the food industry by image classification using transfer learning. *Multimedia Tools and Applications*, *81*(6), 7611–7624. Google Scholar

Kumar, Pankaj, Ashtekar, Sunidhi, Jayakrishna, S. S., Bharath, K. P., Vanathi, P. T., & Kumar, M. Rajesh. (2021). Classification of mango leaves infected by fungal disease anthracnose using deep learning. *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, 1723–1729. IEEE. Google Scholar

Mahmood, Atif, Singh, Sanjay Kumar, & Tiwari, Amod Kumar. (2022). Pre-trained deep learning-based classification of jujube fruits according to their maturity level. *Neural Computing and Applications*, *34*(16), 13925–13935. Google Scholar

Nagaraju, Y., Sahana, T. S., Swetha, S., & Hegde, Siddhanth U. (2020). Transfer learning based convolutional neural network model for classification of mango leaves infected by anthracnose. *2020 IEEE International Conference for Innovation in Technology (INOCON)*, 1–7. IEEE. Google Scholar

Pham, Tan Nhat, Van Tran, Ly, & Dao, Son Vu Truong. (2020). Early disease classification of mango leaves using feed-forward neural network and hybrid metaheuristic feature selection. *IEEE Access*, *8*, 189960–189973. Google Scholar

Prabu, M., & Chelliah, Balika J. (2022). Mango leaf disease identification and classification using a CNN architecture optimized by crossover-based levy flight distribution algorithm. *Neural Computing and Applications*, *34*(9), 7311–7324. Google Scholar

Roomi, S. Mohamed Mansoor, Priya, R. Jyothi, Bhumesh, S., & Monisha, P. (2012). Classification of mangoes by object features and contour modeling. *2012 International Conference on Machine Vision and Image Processing (MVIP)*, 165–168. IEEE. Google Scholar

Sahu, Dameshwari, & Dewangan, Chitesh. (2017). Identification and classification of mango fruits using image processing. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol*, *2*(2), 203–210. Google Scholar

Singh, Uday Pratap, Chouhan, Siddharth Singh, Jain, Sukirty, & Jain, Sanjeev. (2019). Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease. *IEEE Access*, *7*, 43721–43729. Google Scholar

Vyas, Ankur M., Talati, Bijal, & Naik, Sapan. (2014). Quality inspection and classification of mangoes using color and size features. *International Journal of Computer Applications*, *98*(1). Google Scholar