Classifying four maturity categories of coffee cherry using CNN-VGG19

Dominic Olango Cagadas[1], Dwi Sudarno Putra[2*], Kristine Mae Paboreal Dunque[1] and Meri Azmi[3]

[1] Department of Electronics Technology, University of Science and Technology of Southern Philippines, **Philippines**
[2] Department of Automotive Engineering, Faculty of Engineering, Universitas Negeri Padang, **Indonesia**
[3] Department of Information Technology, Politeknik Negeri Padang, **Indonesia**

* Corresponding Author: dwisudarnoputra@ft.unp.ac.id
Received August 08th 2024; Revised December 23rd 2024; Accepted December 25th 2024

Cite this https://doi.org/10.24036/teknomekanik.v7i2.31072

Abstract: The local coffee farmers employ manual inspection to identify the maturity of coffee cherries that are inefficient in labor and time. Thus, the objective of this study is to develop a CNN-VGG19 algorithm model that can accurately detect the maturity image of coffee cherry samples, and classify them into: unripe, semi-ripe, ripe, and overripe categories. The proposed solution will provide local coffee farmers with an automated and more accurate classification of the quality of coffee cherries. The visual geometry group-19 was employed to increase the object recognition model performance of the proposed algorithm while maintaining higher accuracy and quicker throughput, thus increasing revenues. The images are utilized as training and test set data. They were then processed by using the feature extraction of CNN-VGG19 deep learning model, and got four coffee cherry maturity classes. The model architecture attained a 90.00 % accuracy. Furthermore, the increase in both the validation and training accuracy graph with a corresponding decrease in both the validation and training loss graph propounds that the model performance has improved.

Keywords: coffee fruit; convolutional neural network; image processing; object detection; visual geometry group

1. Introduction

Coffee is the most valued and profitable beverage commodity traded on the global market with over 2 billion cups of coffee consumed every day around the world. It increases the consumption of coffee and the demand for high-quality coffee beans year by year [1]. According to Statistica on Global Coffee Consumption, around 166.63 million 60-kilogram bags of coffee were consumed in 2021, which is an increase from the 164.00 million bags marketed in 2020. Coffee is grown in 50 countries along the equatorial "The Bean Belt" zone. Interestingly, the Philippines lies within this favorable climatic condition location with a current 37,000 metric tons of average annual coffee cherries harvest. Referred to as cherries, the fruit is initially green and turns bright and deep red when it is ready for harvesting. Traditionally, the coffee cherries are harvested by hand through one of two strategic ways: either strip picking or selective picking [2].

Strip picking is usually done either by hand or machinery off the branches at one time. The harvested produce may not achieve the desired quality taste of coffee. It is because of the mixture of unripe, semi-ripe, ripe, and overripe cherries that require further sorting or will result in ununiform drying. This is a preferred method because it is quick and convenient, but it needs an expensive investment for the mechanical stripping. In addition to the high cost, the following important factors are considered to use the machinery: the topography of the land and the physiognomy of the coffee plants. Whereas, selective picking is the process of handpicking only the perfectly mature and ripe coffee cherries from the whole tree, leaving the unripe, semi-ripe, or overripe ones unharvested. The harvested coffee meets the quality standard due to the consistent maturity of the cherries. However, this strategy is costly because requires several labor-intensive picking rounds. They rotate among the trees 5 or even 10 times to choose the cherries which are at the peak of their ripeness.

The harvest of immature cherries in the production of coffee is associated with caustic, monotonous, or astringent sensory attributes [3]. When cherries are fresh and ripe-harvested, they tend to produce a cup of coffee with a clean floral tea-like clarity such as jasmine or hibiscus, and mild sweetness, that resonates like raw honey or similar to sugar cane juice. Moreover, beans from coffee berries at the green-cane maturity presented more dry matter and lower yield compared to the fully ripe coffee cherry [4], [5]. The most common approach used for detecting coffee cherries' quality is the Convolutional Neural Network (CNN), which showed excellent results in the area of object detection and classification [6], [7]. A previous study verified its findings through an intricate colorimetric analysis using the physicochemical features associated with the coffee cherry maturity classification [8]. Another study implemented an algorithm model based on computer vision technique to detect, classify, and map coffee cherries into three catagories only: unripe, ripe, or overripe coffee cherries [9]. However, the results required higher accuracy [10].

Therefore, the objective of this study was to develop and implement a CNN algorithm model that detects and classifies the percentage recognition based on the image maturity of the 4 categories: unripe, semi-ripe, ripe, and overripe cherries with higher recognition accuracy. Specifically, it aims to determine the accuracy percentage of coffee cherry among unripe, semi-ripe, ripe, and overripe classifications by the use of image processing. Thus, it assists the local farmers in the Philippines to automate the inspection of the coffee cherries.

2. Material and methods

This section condenses the detailed process and conditions of obtaining the data images. It also outlines the deep learning model architecture, the classification structure, and the evaluation criteria and metrics used to assess the performance of the proposed coffee cherry maturity assessment framework. This includes accuracy, precision, recall, F1 – score, and some domain analysis calculations.

2.1 Image acquisition

The 356 sample images were collected and classified by the expert coffee farmers into 4 different categories. These images were adjusted in alignment with the system's requirements. The coffee farmers sorted the fruit by observing the color of the epidermis of the coffee cherries (Figure 1).

Figure 1. Various coffee cherries color harvested

2.2 Proposed model for classifying coffee cherry maturity

The proposed model consists of collected coffee cherry images both from the internet and in a real-time coffee farm, image pre-processing, convolutional neural network – visual geometry group – 19 convolutional weight layers CNN – VGG - 19 algorithm model, and confusion matrix assessment [11]. Where the cherry maturity evaluation is divided into training and classification (Figure 2).

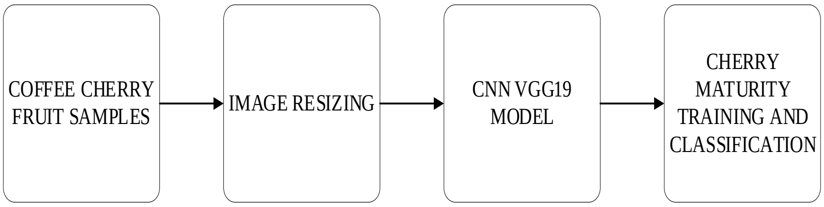

Figure 2. Proposed model

2.3 Maturity detection and classification of coffee cherry

The VGG was employed to increase the object recognition model performance of the proposed algorithm while maintaining higher accuracy and faster throughput turnover (Figure 3). From the input sampled coffee cherries, the convolutional layer generated a feature map by applying a filter that scanned the image. Then, the pooling layer scaled down the data recreated by the convolutional layer and effectively stored it. The sample images were resized to the preset proportion to increase the detection accuracy. This down-sampling stage improved the anti-distortion ability, retained the main features of the sample images, and reduced the number of characteristic parameters. The fully connected input layers flattened the outputs into a one-dimensional data vector and applied weights over the inputs produced by feature analysis. Accordingly, the fully connected output layer fabricated the final probabilities in determining the output image classes as Y_1= unripe, Y_2= semi-ripe, Y_3= ripe, and Y_4= overripe [12]. The convolution method integrates the three signals of interest: input, impulse response, and output. Whereas, a rectified linear unit as a piecewise linear function allowed the model to learn faster. Therefore, it performs better. Further, the pooling operation entailed an aggregating feature on the function map within the filter's area and utilized a flattening process to build a single long resulting function vector.

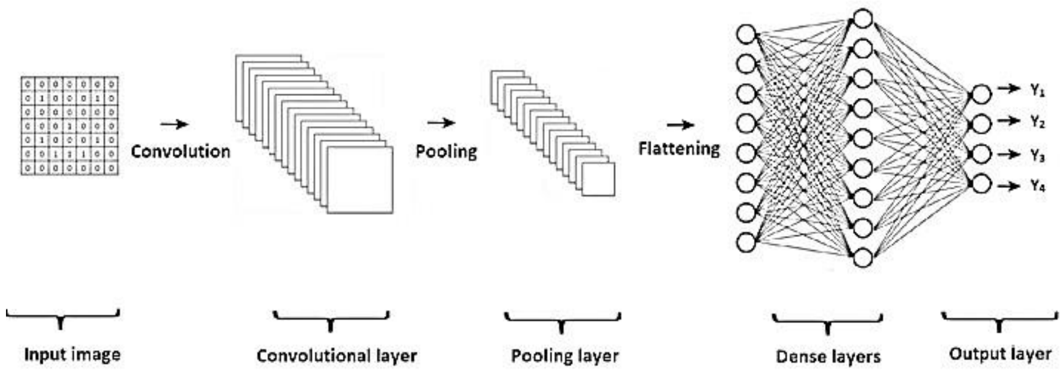

Figure 3. CNN-VGG 19

The implementation of the CNN-VGG 19 system model follows the algorithms:

a. The system model and date set size were loaded, wherein, the images of the different maturities of coffee cherries for training, testing, and validation were arbitrarily reorganized.

b. A function was set up to freeze the initial layers of VGG-19, which extracted the features and labels of the coffee images. It allowed the model to apply transfer learning.
c. The right data shape was configured to resonate the dimensions of the selected dataset.
d. An image binary category was concluded from the extracted features and labels of the coffee cherries.
e. The image classifier was built that visualized the model learned using the accuracy and loss metrics.

2.4 CNN-VGG 19 model

The VGG model consists of 19 weighted layers, with 16 convolutional layers, 5 max-pooling layers, and 3 fully connected layers. The 16 convolutional layers learned hierarchical visual feature representations while the 5 max-pooling layers reduced the spatial dimensionality of the featured image data maps. The 3 fully connected layers along with the transfer learning performed the 4 prediction tasks. The coffee cherry data images were resized into 224x224 pixels and were labelled into 4 groups for the CNN-VGG 19 model.

Sample images were then fed to the system model after being processed. Then, the VGG19 framework was trained. The fully connected input layers were replaced with a value of 128 for smooth and refined tuning. The Softmax classifier was substituted with four maturity classes as needed in the data analysis. The proposed system model had almost 26 layers, each with a different number of convolution and max pooling layers and provided more training parameters. Consequently, it gives a huge pool of training parameters. This deep learning model produced high-accuracy [6], [12], [13].

2.5 Performance evaluation criteria

The classification accuracy is the ratio between the total number of predictions produced and the number of right predictions made. The precision rate evaluates the accuracy with which a class is determined. Recall or rate of sensitivity reflects the classifier's capacity to predict the class correctly. The F1 score defines the harmonic sensitivity and accuracy value of the model. Whereas, the specificity rate demonstrates the classifier's separation abilities. The confusion matrix shows the ratio between the predicted and the actual classes. All the assessment indicators to measure the success of the classification model apply the basic terms from the confusion matrix, true positive (TP), false positive (FP), true negative (TN), and false negative (FN), with the following statistical equations [14],[15]:

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

$$Precision: \frac{TP}{(FP+TP)} \quad (2)$$

$$Recall: \frac{TP}{(FN+TP)} \quad (3)$$

$$F1\,Score: \frac{2(Precision*Recall)}{Precision+Recall} \quad (4)$$

$$Specify: \frac{TN}{FP+TN} \quad (5)$$

3. Results and discussion

The proposed classification solution is simple and can learn robust visual features. Whereas, the classification prediction encapsulated the evaluation of the study's performance on the held-out test set.

3.1 Accuracy and loss of CNN-VGG-19 model

After the test image simulations were done, it returned an overall accuracy of 89.71 %. The system loss, along with the model accuracy, between train, and validation were laid out in Table 1.

Table 1. CNN-VGG19 model classification report

Class	Precision	Recall	F1 Score	Support
Class0 (Unripe)	0.93	0.88	0.90	16
Class1 (Semi-Ripe)	0.89	0.94	0.82	18
Class2 (Ripe)	0.88	**0.92**	**0.91**	22
Class3 (Over-Ripe)	0.93	0.82	0.87	17
Average	**0.91**	**0.90**	**0.90**	**73**

Table 1 outlines the classification report for the CNN VGG19 with the four maturity catagories: Class_0 = unripe, Class_1 = semi-ripe, Class_2 = ripe, and Class_3 = overripe evaluated categories. The model architecture attained a 90.00 % average accuracy, this value represented the percentage of the number of coffee cherry images that were correctly predicted from the total number of predictions. Table 1 shows that Class_2 = ripe has a recall rate of 92.00 %, wherein, it resonated a high percentage of ripe coffee cherry images classified as true ripe cherries from the total number of samples. The proposed CNN-VGG19 has an F1 Score of 0.91 that returned both high Precision and Recall rates. The classification ability of the system can accurately capture and evaluate true, ripe matured coffee cherries. The performance analysis of the system in accuracy (Figure 4) and loss graphs (Figure 5).

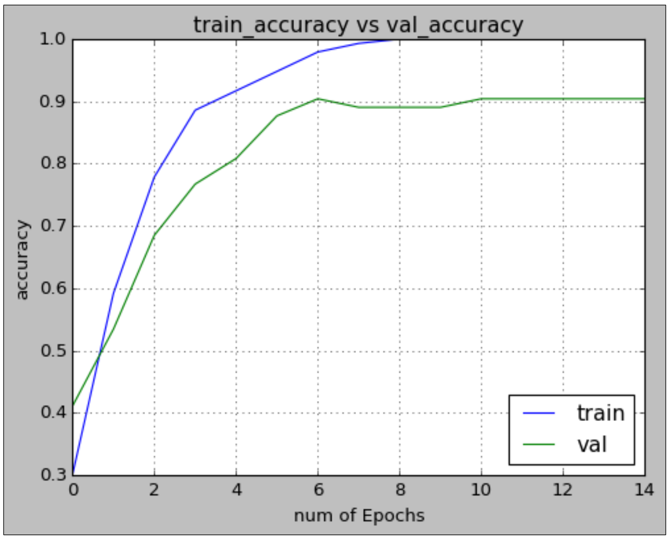

Figure 4. Training and validation accuracy

The increase in both the validation and training accuracy graphs with a corresponding decrease in both validation and training loss graphs, respectively, propounded that the model performance had improved.

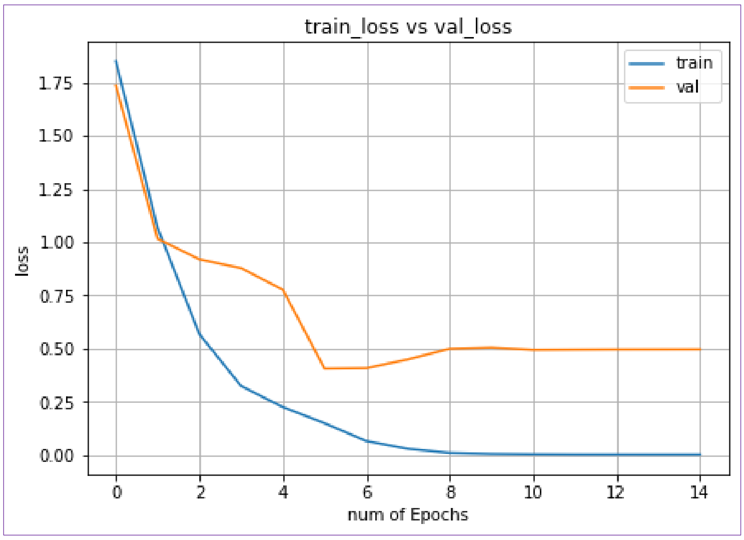

Figure 5. Training and validation loss

The confusion matrix of the proposed model where classes 0-3, designated as unripe, semi-ripe, ripe, and over-ripe coffee cherries (Figure 6). Thus, validated that the fruit classification was performed precisely among the four maturity categories, with an 87.14 % entire accuracy. It is shown that in Class_0 = unripe, 14 samples were predicted out of 16 images, and 2 samples were identified as Class_1 = semi-ripe. In the Class_1 = semi-ripe column, 17 samples were predicted out of 18 images and 1 sample was identified as Class_0 = unripe. While in the Class_2 = ripe group, 21 samples were predicted out of 22 images and 1 sample was identified as Class_3 = over-ripe. Lastly for the Class_3 = overripe section, 14 were predicted out of 17 images and 3 samples were identified as Class_2 = ripe.

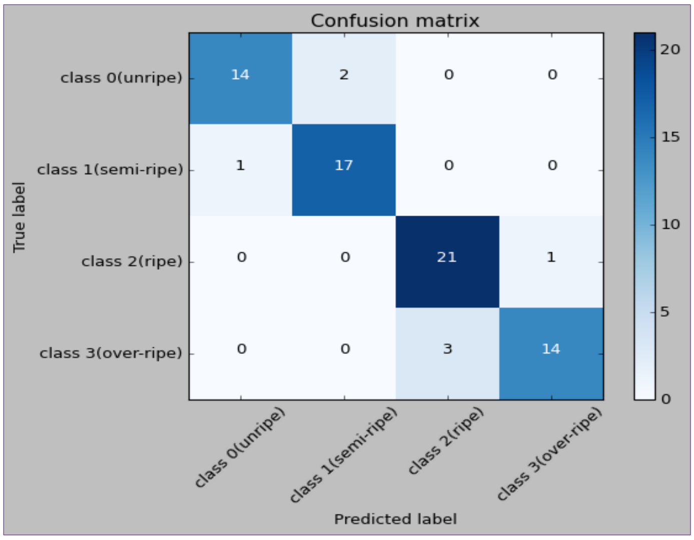

Figure 6. Confusion matrix

3.2 Classification prediction

Table 2 summarises the different maturity classification findings. It tested the system to each maturity class whenever an input coffee cherry image was fed. The system will either return P or NP as a classification evaluation status.

Table 2. Classification predictions

	Image Under Test	Class0 (Unripe)	Class1 (Semi-Ripe)	Class2 (Ripe)	Class3 (Over-Ripe)
		**P**	NP	NP	NP
Unripe		**P**	NP	NP	NP
		**P**	NP	NP	NP
		NP	**P**	NP	NP
Semi-ripe		NP	**P**	NP	NP
		NP	**P**	NP	NP
		NP	NP	**P**	NP
Ripe		NP	NP	**P**	NP
		NP	NP	**P**	NP
Over-ripe		NP	NP	**P**	NP
		NP	NP	NP	NP

Legend:
P – Correct Prediction
NP – Incorrect Prediction

4. Conclusion

This study used the CNN-VGG19 model in detecting the coffee cherries' maturity and replacing the two completely connected layers and SoftMax classifier with 128 fine-tuning frames and four maturity classes. The developed CNN-VGG19 model was used to compare the efficient and accurate results. The test image simulations were done and returned an overall accuracy of 89.71 %. The ongoing work attempted to improve the model accuracy to 90.00 % and increased the classification into 4 categories while maintaining a decrease in the training time. It successfully implemented the coffee cherry quality identification system by image processing. Thus, the proposed model was effective in using a single deep-learning CNN-VGG19 model for classification with reduced decision time.

Author's declaration

Author contribution

Dominic Olango Cagadas: conceived and designed the study and methodology, supervised the creation of datasets, and contributed to the development of code and visualizations. **Dwi Sudarno Putra**: collected and analyzed the primary data, synthesized the findings into a coherent narrative, interpreted the results, managed the submission process, and revised the manuscript based on feedback. **Kristine Mae Paboreal Dunque**: co-developed the research framework, oversaw dataset preparation, and refined the code and visualizations. **Meri Azmi**: assisted in data collection and analysis, integrated findings into the manuscript, and supported revisions based on reviewer feedback.

Funding statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgements

The authors would like to express their sincere gratitude to the University of Science and Technology of Southern Philippines (USTP), Universitas Negeri Padang (UNP), and Politeknik Negeri Padang (PNP) for their invaluable support in facilitating this collaborative research.

Competing interest

The authors declare no competing interest.

Ethical clearance

This research did not involve human or animal subjects.

AI statement

This article is the author's original work, written from original research and no sections or figures are generated by AI. English is checked using Grammarly and has been verified by the authors.

Publisher's and Journal's note

Universitas Negeri Padang as the publisher and Editor Teknomekanik state that there is no conflict of interest towards this article publication.

References

[1] "The Caffeinated History of Coffee | Stories | PBS Food." [Online]. Available: https://www.pbs.org/food/stories/history-coffee

[2] M. Haile and W. Hee Kang, "The Harvest and Post-Harvest Management Practices' Impact on Coffee Quality," in *Coffee - Production and Research*, IntechOpen, 2020. https://doi.org/10.5772/intechopen.89224

[3] V. Osorio Pérez, L. G. Matallana Pérez, M. R. Fernandez-Alduenda, C. I. Alvarez Barreto, C. P. Gallego Agudelo, and E. C. Montoya Restrepo, "Chemical Composition and Sensory Quality of Coffee Fruits at Different Stages of Maturity," *Agronomy*, vol. 13, no. 2, p. 341, Jan. 2023, https://doi.org/10.3390/agronomy13020341

[4] P. D. Leandro, S. S. Ney, S. A. Gilberto, R. C. Paulo, A. P. da S. Fernando, and dos S. G. O. Lidiane, "Coffee production through wet process: Ripeness and quality," *African Journal of Agricultural Research*, vol. 12, no. 36, pp. 2783–2787, Sep. 2017, https://doi.org/10.5897/AJAR2017.12485

[5] E. Yusibani, R. I. Putra, A. Rahwanto, M. S. Surbakti, Rajibussalim, and Rahmi, "Physical properties of Sidikalang robusta coffee beans medium roasted from various colors of coffee cherries," *Journal of Physics: Conference Series*, vol. 2243, no. 1, p. 012046, Jun. 2022, https://doi.org/10.1088/1742-6596/2243/1/012046

[6] M. A. Tamayo-Monsalve *et al.*, "Coffee Maturity Classification Using Convolutional Neural Networks and Transfer Learning," *IEEE Access*, vol. 10, pp. 42971–42982, 2022, https://doi.org/10.1109/ACCESS.2022.3166515

[7] J. Jayanth, M. Mahadevaswamy, and M. Shivakumar, "Fruit Quality Identification and Classification by Convolutional Neural Network," *SN Computer Science*, vol. 4, no. 3, p. 220, Feb. 2023, https://doi.org/10.1007/s42979-022-01527-w

[8] S. Velasquez, A. Patricia Franco, N. Peña, J. Carlos Bohorquez, and N. Gutierrez, "Classification of the maturity stage of coffee cherries using comparative feature and machine learning," *Coffee Science*, vol. 16, pp. 1–11, 2021, https://doi.org/10.25186/.v16i.1710

[9] H. C. Bazame, J. P. Molin, D. Althoff, and M. Martello, "Detection, classification, and mapping of coffee fruits during harvest with computer vision," *Computers and Electronics in Agriculture*, vol. 183, p. 106066, Apr. 2021, https://doi.org/10.1016/j.compag.2021.106066

[10] J. Xiao, J. Wang, S. Cao, and B. Li, "Application of a Novel and Improved VGG-19 Network in the Detection of Workers Wearing Masks," *Journal of Physics: Conference Series*, vol. 1518, no. 1, p. 012041, Apr. 2020, https://doi.org/10.1088/1742-6596/1518/1/012041

[11] T.-H. Nguyen, T.-N. Nguyen, and B.-V. Ngo, "A VGG-19 Model with Transfer Learning and Image Segmentation for Classification of Tomato Leaf Disease," *AgriEngineering*, vol. 4, no. 4, pp. 871–887, Oct. 2022, https://doi.org/10.3390/agriengineering4040056

[12] S. Raveena and R. Surendran, "ResNet50-based Classification of Coffee Cherry Maturity using Deep-CNN," in *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, IEEE, Jan. 2023, pp. 1275–1281. https://doi.org/10.1109/ICSSIT55814.2023.10061006

[13] C. Chan Enriquez, J. Marcelo, D. Rae Verula, and N. Joy Casildo, "Leveraging deep learning for coffee bean grading: A comparative analysis of convolutional neural network models," *2024. Transactions on Science and Technology*, vol. 11, no. 1, pp. 1–6, 2024, [Online]. Available: http://tost.unise.org/

[14] N. Ullah *et al.*, "Diabetic Retinopathy Detection Using Genetic Algorithm-Based CNN Features and Error Correction Output Code SVM Framework Classification Model," *Wireless Communications and Mobile Computing*, vol. 2022, 2022, https://doi.org/10.1155/2022/7095528

[15] Y. S. Taspinar, M. Koklu, and M. Altin, "Fire Detection in Images Using Framework Based on Image Processing, Motion Detection and Convolutional Neural Network," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 9, pp. 171–177, Dec. 2021, https://doi.org/10.18201/IJISAE.2021473636