Herbal Leaf Classification Using Convolutional Neural Network (CNN) Method With VGG16 Architecture

Sisi Auliani[1], Dolly Indra[2], Fitriyani Umar[3]
Informatics Engineering, Faculty of Computer Science, Universitas Muslim Indonesia
sisiauliani09@gmail.com[1], dolly.indra@umi.ac.id[2], fitriyani.umar@umi.ac.id[3]

Article Info

Article history:

Received Aug 14, 2025
Revised Sep 19, 2025
Accepted Nov 25, 2025

Keyword:

Image Classification
Herbal Leaves
Convolutional Neural Network
VGG16
Fine-Tuning

ABSTRACT

Visual identification of various types of herbal leaves such as sweet leaf, moringa leaves, cat whiskers, bay leaf, and betel leaf is often difficult due to morphological similarities. This study presents a novel approach using a Convolutional Neural Network (CNN) model with the VGG16 architecture to automatically classify these five types of leaves. The main novelty of this study lies in the implementation of a two-stage fine-tuning strategy specifically tailored to the herbal leaf dataset. The first stage freezes the base layer and trains a new classification head, while the second stage fine-tunes several upper layers to improve model adaptability. The model was trained using 500 herbal leaf images and evaluated on 235 independent test images. The results demonstrated superior model performance with an overall accuracy rate of 91.06% and an average F1-score of 0.91. Qualitative analysis demonstrated the model's success in classifying leaves with unique features, such as cats whiskers and betel leaf. However, the model faced challenges in distinguishing leaves with high visual similarities, such as sweet leaf and moringa leaves. Practically, the developed model offers an effective and reliable solution for herbal leaf identification, reducing time and error rates compared to manual methods. Although this study is limited by the small dataset size, these results demonstrate the great potential of the optimized VGG16 architecture for applications in botany and traditional medicine, making it a valuable tool.

Corresponding Author:

Dolly Indra
Faculty of Computer Science
Universitas Muslim Indonesia
Jl. Urip Sumoharjo Km. 05, Makassar, Indonesia
Email : dolly.indra@umi.ac.id

1. INTRODUCTION

Indonesia is a country with extensive biodiversity and a wide variety of herbal plants[1]. This biodiversity makes it a highly respected country in the world in the field of herbal medicine[2]. The herbal plants that grow in this area are very beneficial, especially for families who have difficulty getting medical care[3]. Herbal plants are a reliable alternative treatment solution for alleviating the symptoms of common illnesses, treating minor wounds, and maintaining general health[4][5].

However, distinguishing between different types of herbal plants is often difficult for people, especially for most people[6]. Leaves such as Bay Leaf, Betel Leaf, Cat Whiskers, Moringa Leaves and Sweet Leaf are types of leaves that can be used for medicinal purposes. These five leaves show significant variations in shape, color, texture, and other characteristics, which can pose a challenge for plant leaf classification due to significant visual differences[7]. When leaf type classification is currently carried out, eye observation and individual assumptions are still used, so the process is time-consuming and tends to have a high error rate[8]. In addition, data augmentation including rotation, flip and brightness adjustment, is widely used to enrich the variety of training data, reducing bias in the dataset[9][10]. Therefore, a system is needed that can distinguish various types of herbal leaves based on their shape, color, and texture[11][12].

The author will use the Convolutional Neural Network (CNN) method with the VGG16 architecture in this context to recognize and classify various types of herbal plant leaves and use an input image size of 224 x 224 [13]. This method has a strong ability to learn important characteristics of images through convolution and pooling processes. This allows for more accurate and effective object recognition. VGG16 is a deep learning architecture with 16 layers[14]. Therefore, the author will use the CNN method to determine the type of leaf consisting of five types of leaves originating from herbal plants: Bay Leaf, Betel Leaf, Cat Whiskers, Moringa Leaves and Sweet Leaf[15].

Several previous studies have demonstrated the success of the CNN method in classifying herbal leaves. Research by Sri Adiningsi et al. (2022) yielded a test accuracy of 90.74% for identifying medicinal plant leaf types using the VGG16 model[4]. The medicinal plant leaf types consisted of guava, starfruit, lime, pandan, betel, papaya, basil, aloe vera, celery, and jackfruit. Meanwhile, research by Mahdarul Huda Ahmad et al. (2023) successfully achieved 96.2% accuracy using the VGG16 transfer learning model and 50% dropout for classifying four types of herbal leaves, namely curry leaves, betel leaves, mint leaves, and moringa leaves[11].

Based on these problems, this study aims to develop a classification model based on the VGG16 architecture that is able to recognize and classify 5 types of herbal leaves.

2. RESEARCH METHOD

The dataset used is a primary dataset collected manually by researchers using a 12 MP smartphone camera, so that image quality is maintained according to system requirements. There are a total of 500 images with five classes of herbal leaves, namely Bay Leaf, Betel Leaf, Cat Whiskers, Moringa Leaves and Sweet Leaf. The use of 500 images in this study is strategic because it uses the VGG16 fine-tuning method. This approach leverages a model already trained for a broader classification task, requiring only a small amount of data for adaptation. Furthermore, the data augmentation technique artificially enriches the dataset by creating image variations, effectively addressing data limitations and preventing overfitting. Details of the method stages can be seen in Figure 1.

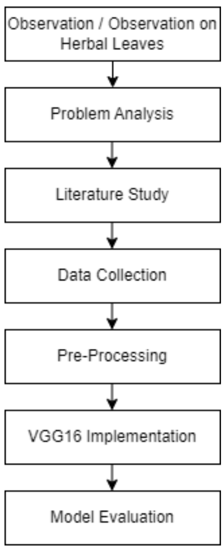

Figure 1. Research Stages

2.1. Observation of Herbal Leaves

This stage involves initial analysis to understand the problem to be solved. In the context of image classification, observation includes identifying the type of image data needed, such as images of herbal leaves to be classified.

2.2. Problem Analysis

The main problem in this research lies in developing a system capable of automatically classifying five types of herbal leaves: Bay Leaf, Betel Leaf, Cat Whiskers, Moringa Leaves and Sweet Leaf. The complexity of this problem stems from the variation within a class, where leaves of the same type can exhibit significant differences in size, shape, and color pattern.

2.3. Literature Study

In the literature review phase, an in-depth review of previous research relevant to herbal leaf image classification was conducted. This review aimed to understand the various approaches, methods, and architectures that have been successfully implemented, as well as to identify existing achievements and challenges.

2.4. Data Acquisition and Pre-Processing

Data collection was carried out through two main scenarios:

1. Individual Pick-up : Each leaf was photographed separately against a relatively clear background. Shots were taken from both the front and back of the leaf to capture the differences in texture and vein patterns.
2. Contextual Retrieval : Leaves were photographed while still attached to the tree, capturing complex variations in lighting, shadows, and background. This scenario helped the model recognize leaves in their natural state.

Some examples of data that were successfully collected from this process can be seen in Figure 2 :

Figure 2. Data Collection Results

After the image data has been successfully collected, the next step is pre-processing. This stage aims to prepare and standardize the raw data to conform to the format required by the model and to address challenges that could impact classification accuracy. The steps include:

1. Resize : Each image was resized so that its longest side matched the target dimensions (224 pixels), while the shorter sides were adjusted proportionally to preserve the original shape of the leaf[16].
2. Padding : After being resized, the image is not always perfectly square. To standardize the input dimensions, the empty space around the image is added (padding) with black pixels until it reaches a final size of 224x224 pixels. Adaptive padding techniques have been shown to improve classification accuracy compared to conventional methods [17].

With this method, the model can learn the shape and silhouette features of leaves without being affected by distortion. As a final step, the processed image is then normalized according to the requirements of the model architecture used. This transformation process is illustrated in Figure 3.

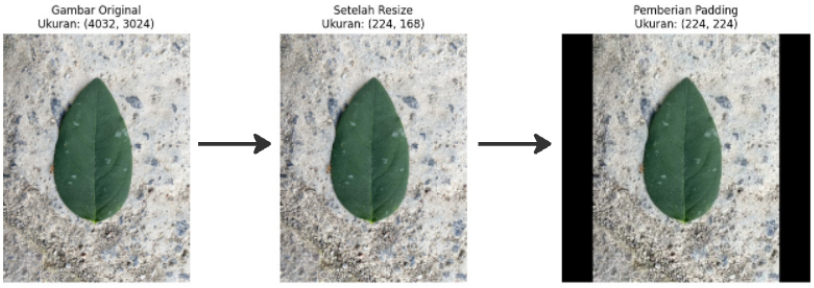

Figure 3. Data Resizing and Padding

Then, the data labeling process is carried out by organizing all the ready images into separate folders, where the name of each folder represents the respective leaf class (example: Bay Leaf, Betel Leaf, etc. folders) as seen in Figure 4.

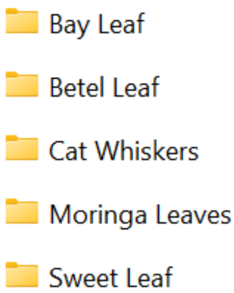

Figure 4. Data Folder Structure

To prepare the data before the training process, the dataset was divided into two parts: the training set and the validation set. This dataset was divided into 80% for training (400 images), 20% for validation (100 images). In addition, 235 additional independent test images were used for final evaluation. The following is the dataset division in Figure 5 below.

```
Found 500 files belonging to 5 classes.
Using 400 files for training.
Found 500 files belonging to 5 classes.
Using 100 files for validation.
Ditemukan 5 kelas: ['Katuk', 'Kelor', 'Kumis Kucing', 'Salam', 'Sirih']
```

Figure 5. Dataset Split

After the resizing and padding process, on-the-fly data augmentation techniques were applied to the training data. The goal of augmentation is to artificially increase the amount and diversity of data without the need to collect new images. This is crucial for preventing overfitting and improving the model's generalizability, especially when dealing with limited or imbalanced datasets. The results of the data augmentation are shown in Figure 6 below.

Figure 6. Data Augmentation Results

1.5. Architectural Models

At this stage, a model for classification will be developed. This study applies the CNN algorithm with the VGG16 architecture. In designing the VGG16 model, parameter adjustments are made where various parameters in the model are changed and adjusted to obtain the best configuration that can be used by the model[18]. The following is an illustration of the VGG16 architecture shown in Figure 7.

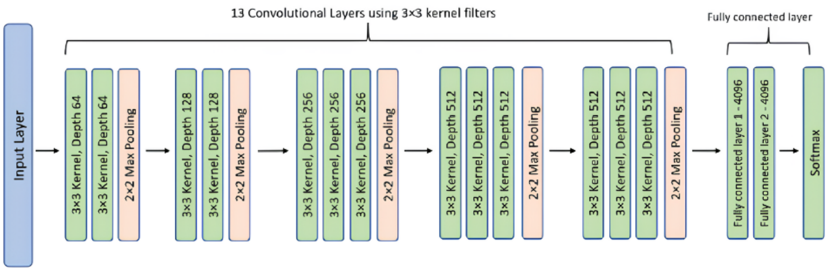

Figure 7. VGG16 Architecture

1.6. Training Models

The model is trained using training data to recognize patterns in the given image. In the training stage, the dataset has been processed using parameters such as learning rate, batch size, layer dropout value and other parameters that are adjusted to improve model performance[19], as shown in Table 1.

Table 1. List of Model Training Parameters

Parameters	Value
activation function	softmax, relu
optimizer	adam
learning rate	0.001
max epoch	400
batch size	32
layer dropout	0.6

1.7. Model Evaluation

At this stage, testing is carried out using separate test data to assess the performance of the trained model. Evaluation is carried out by calculating metrics using the Confusion Matrix, such as accuracy, precision, recall, and F1-score[20]. The following metric equations are shown in equations (1-4).

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

$$\text{Precision} = \frac{\mathrm{TP}}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{\mathrm{TP}}{TP+FN} \quad (3)$$

$$\text{F1} - \text{Score} = \frac{2 \text{ x precision x recall}}{precision+recall} \quad (4)$$

True Positive (TP) is the number of data from the positive class that are correctly classified as positive. True Negative (TN) refers to the number of data from the negative class that are correctly categorized as negative. False Positive (FP) is the number of data from the negative class that are incorrectly classified as positive. False Negative (FN) is the number of data from the positive class that are incorrectly identified as negative.

3. RESULTS AND ANALYSIS

This section presents the experimental results of classifying five types of herbal leaves using a Convolutional Neural Network (CNN) with a modified VGG16 architecture. The dataset consists of 500 training images and 235 independent test images. The training process was carried out using a two-stage fine-tuning strategy.

3.1. Model Training Results

This sub-chapter presents the results of the VGG16 model training process. Training was set to run for a maximum of 400 epochs, with an early stopping mechanism implemented. This mechanism aims to automatically stop the training process when performance on the validation data no longer shows improvement, preventing overfitting and maintaining the model at its best.

The development of model performance during the training process is illustrated through the accuracy and loss graphs presented in Figure 8.

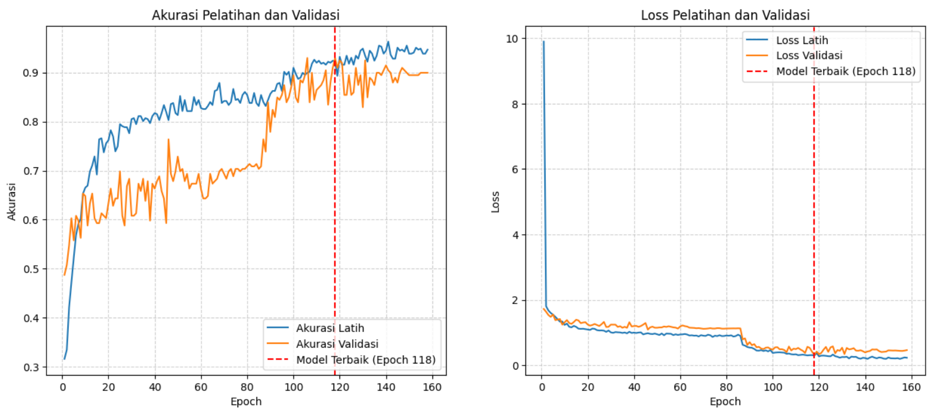

Figure 8. Model Training Results Graph

Figure 8 shows the training history for one of the experimental runs, showing the accuracy and loss curves for the training and validation data. The red vertical line at epoch 118 marks the point where the best-performing model on the validation data was retained. Table 2 summarizes the model performance before and after fine-tuning:

Table 2. Comparison of Model Performance Before and After Fine-Tuning

Process	Accuracy Training	Accuracy Validation	Loss Training	Loss Validation
before fine-tuning	0.8029	0.7638	0.9818	1.0919
after fine-tuning	0.9220	0.9296	0.3014	0.3229

From these results, a significant increase in validation accuracy (from 76.38% to 92.96%) and a decrease in validation loss (from 1.0919 to 0.3229) can be seen after applying the second stage of fine-tuning.

3.2. Model Testing Results

After the model is trained, the next step is to evaluate it to objectively measure its performance. Testing is conducted using a previously separated test set that has never been used in the training or validation process. The confusion matrix is shown in Figure 9.

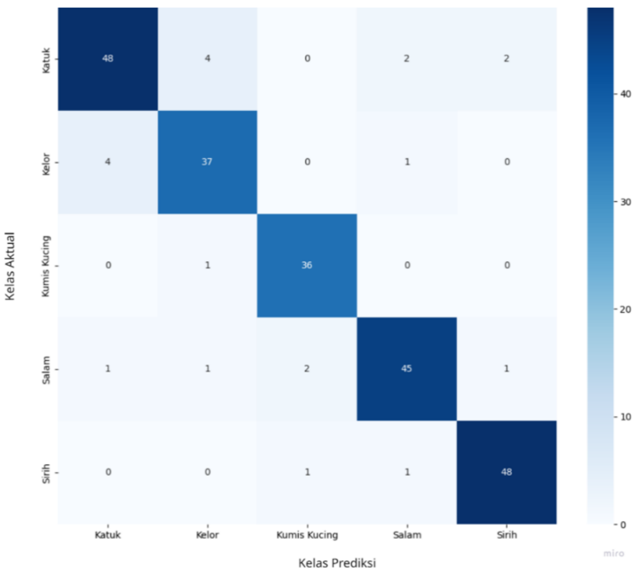

Figure 9. Confusion Matrix of Model Testing Results

Figure 7 shows that the values on the main diagonal, the row of cells stretching from the top left corner to the bottom right, in this example contain the values 48, 37, 36, 45, and 29. These values indicate all data that was correctly classified according to its original class. Meanwhile, values outside the diagonal indicate the number of data that were misclassified.

The results of the metric calculations for each class are summarized in Table 3 below.

Tabel 3. Summary of Model Testing Result Metrics

Leaf Class	Precision	Recall	F1-Score	Number of Samples
Sweet Leaf	0.91	0.86	0.88	56
Moringa Leaves	0.86	0.88	0.87	42
Cat Whiskers	0.92	0.97	0.95	37
Bay Leaf	0.92	0.90	0.91	50
Betel Leaf	0.94	0.96	0.95	50
Average	0.91	0.91	0.91	235

(Macro)				
Average (Weighted)	0.91	0.91	0.91	235

The table above presents a quantitative summary of the model testing results, measured using Precision, Recall, and F1-Score metrics. Overall, the model demonstrates very high and stable performance, as reflected in the two types of average calculations presented at the bottom of the table: Macro Average and Weighted Average.

In addition to these more detailed metrics, model performance is also measured using Accuracy to get a better idea. Accuracy essentially measures how many times the model correctly guesses across all the data tested.

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Data}$$

From the previous confusion matrix, we know that the number of correct guesses is 214 (the total number of numbers in the main diagonal) and the total number of data tested is 235. So, the calculation becomes:

$$Accuracy = \frac{214}{235} = 91.06\%$$

These results show that the model successfully guessed correctly 91.06% of the test data, which once again confirms that the overall performance of the model is very good.

4. CONCLUSION

The model demonstrated superior performance in classes with unique visual features. The Cat Whiskers and Betel Leaf classes performed best, achieving an F1-Score of 0.95. For the Cat Whiskers class, the model achieved a precision of 0.92 and a recall of 0.97, while for the Betel Leaf class, the precision reached 0.94 and a recall of 0.96. This success is attributed to the model's ability to recognize distinctive features, such as the serrated leaf edges of Cat Whiskers and the heart-shaped leaves of Betel Leaf.

ACKNOWLEDGEMENTS

With all humility, the author expresses her deepest gratitude to Ms. Haslinda, the owner of the herbal leaves, for her kindness and generosity in granting permission and full support to the author to photograph the herbal leaves. This assistance was invaluable and provided a crucial foundation for the success of this research. Herbal leaf data collection was carried out at BTP (Bukit Tamalanrea Permai), Makassar.

REFERENCES

[1] S. S. Tahoangako, D. Santosa, and N. Fakhrudin, “Study of the Utilization of Medicinal Plants by Traditional Healer of the Tolaki Ethnic Tribe, Southeast Sulawesi, Indonesia,” *Ethnobot. Res. Appl.*, vol. 28, 2024, doi: 10.32859/era.28.39.1-17.

[2] E. Singarimbun, E. Elfrida, and I. Indriaty, “Indigenous Knowledge and Herbal Medicine: Exploring the Ethnobotany of the Karo Tiganderket Tribe in Indonesia,” *Heca J. Appl. Sci.*, vol. 2, no. 2, pp. 74–86, 2024, doi: 10.60084/hjas.v2i2.208.

[3] F. Azzahra, A. D. Astuti, B. Arifin, and G. Alam, “Scoping Review: Study of Herbs Consumption for Self-Medication in Indonesia 2019-2022,” *Maj. Obat Tradis.*, vol. 29, no. 3, pp. 302–326, 2024, doi: 10.22146/mot.94091.

[4] S. A. E. ALBAKIA and R. A. Saputra, “Identifikasi Jenis Daun Tanaman Obat Menggunakan Metode Convolutional Neural Network (CNN) Dengan Model VGG16,” *J. Inform. Polinema*, vol. 9, no. 4, pp. 451–460, 2023, doi: 10.33795/jip.v9i4.1420.

[5] A. I. Sulaiman, S. Prastyanti, L. K. Sari, A. Rofik, and M. Yamin, "Community Empowerment in Developing Family Medicinal Plants to Realize a Green Economy Based on Local Wisdom," no. 9, pp. 54–65, 2025, doi: 10.53935/jomw.v2024i4.1158.

[6] K. Saputra S, D. Niska, I. Taufik, M. Hidayat, and D. Dharma, "Classification of Herbal Plants Based on Leaf Images using Convolutional Neural Network," 2022, doi: 10.4108/eai.11-10-2022.2325271.

[7] S. Salsabila, A. Suharso, and P. Purwantoro, "Comparison of Deep Learning Architectures in Identifying Types of Medicinal Plant Leaf Images," *J. Appl. Informatics Comput.*, vol. 8, no. 1, pp. 39–46, 2024, doi: 10.30871/jaic.v8i1.6289.

[8] N. L. Marpaung, R. J. H. Butar Butar, and S. Hutabarat, "Implementasi Deep learning untuk Identifikasi Daun Tanaman Obat Menggunakan Metode Transfer learning," *J. Edukasi dan Penelit. Inform.*, vol. 9, no. 3, p. 348, 2023, doi: 10.26418/jp.v9i3.63895.

[9] Z. Chen, "Diffusion Models-based Data Augmentation for the Cell Cycle Phase Classification," *J. Phys. Conf. Ser.*, vol. 2580, no. 1, 2023, doi: 10.1088/1742-6596/2580/1/012001.

[10] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, no. November, p. 100258, 2022, doi: 10.1016/j.array.2022.100258.

[11] M. H. Ahmad, M. Hana, T. Ghazi Pratama, and H. Aulida, "Klasifikasi Empat Jenis Daun Herbal Menggunakan Metode Convolutional Neural Network," *J. Ilmu Komput. dan Matemtika*, vol. 4, no. 2, pp. 69–76, 2023.

[12] O. E. Putra and R. Permana, "Hybrid Data Mining For Member Determination And Financing Prediction In Syariah Financing Saving And Loan Cooperatives," *J. RESTI*, vol. 8, no. 2, pp. 309–320, 2024, doi: 10.29207/resti.v8i2.5683.

[13] I. R. Ramadhani, A. Nilogiri, and A. Qurrota, "Klasifikasi jenis tumbuhan berdasarkan citra daun menggunakan metode convolutional neural network," *J. Smart Teknol.*, vol. 3, no. 3, pp. 249–260, 2022, [Online]. Available: http://jurnal.unmuhjember.ac.id/index.php/JST

[14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–14, 2015.

[15] N. N. Sinaga and A. Sembiring, "Klasifikasi Tumbuhan Obat Berdasarkan Citra Daun Menggunakan Algoritma CNN Classification of Medicinal Plants Based on Leaf Images Using CNN Algorithm," vol. 5, no. April, pp. 64–74, 2025, doi: 10.34007/incoding.v5i1.833.

[16] W. Liu *et al.*, "Is the aspect ratio of cells important in deep learning? A robust comparison of deep learning methods for multi-scale cytopathology cell image classification: From convolutional neural networks to visual transformers," *Comput. Biol. Med.*, vol. 141, 2022, doi: 10.1016/j.compbiomed.2021.105026.

[17] J. Kim, "PadChannel: Improving CNN Performance through Explicit Padding Encoding," 2023, [Online]. Available: http://arxiv.org/abs/2311.07623

[18] T. M. Arkan, A. Sugiharto, and H. A. Wibawa, "Optimizing VGG16 Architecture with Bayesian Hyperparameter Tuning for Tomato Leaf Disease Classification," *J. Masy. Inform.*, vol. 16, no. 2, pp. 162–173, 2025, doi: 10.14710/jmasif.16.2.73168.

[19] M. A. Rofiqul Zein, B. Rahmat, and E. Y. Puspaningrum, "Influence of Input Image Size Variations and Data Balancing on VGG-16 and VGG-16-ELM Models for Pneumonia Classification," *J. Artif. Intell. Eng. Appl.*, vol. 4, no. 3, pp. 2344–2352, 2025, doi: 10.59934/jaiea.v4i3.1166.

[20] S. Sathyanarayanan, "Confusion Matrix-Based Performance Evaluation Metrics," *African J. Biomed. Res.*, vol. 27, no. 4, pp. 4023–4031, 2024, doi: 10.53555/ajbr.v27i4s.4345.