jsret.knpub.com
E-ISSN: 2962-6110
P-ISSN: 2964-7843
JSRET
Journal of Scientific Research, Education, and Technology
Interdisciplinary science, open access, and peer-reviewed journal that disseminates research findings from lecturers, researchers, teachers, and scientists in a variety of scientific and technological fields. 

# Enhancing Convolutional Neural Network Accuracy for Herbal Leaf Classification Using Squeeze and Excitation Attention

**Ragil Gigih Utomo, Rodhiyah Mardhiyyah**
Department of Informatics, University of Technology Yogyakarta, Yogyakarta, Indonesia

**ABSTRACT**

Accurate identification of herbal plant species is crucial for human health, but remains challenging, even for experts. To address this critical challenge with an advanced, robust solution, this study introduces a novel approach: the integration of a Convolutional Neural Network (CNN) with the Squeeze-and-Excitation (SE) attention mechanism for reliable herbal leaf classification, leveraging the efficient MobileNetV2 architecture. The dataset consisted of 1,500 images across 10 classes of herbal leaves, split into 80% for training, 10% for validation, and 10% for testing. Both the native MobileNetV2 architecture and the enhanced CNN with Attention Mechanism (CNN-AM) were trained and evaluated using standard metrics (accuracy, precision, recall, and F1-score). The comparison results decisively demonstrated the effectiveness of the attention mechanism. Crucially, the integration of Squeeze-and-Excitation yielded a significant performance improvement: the model's average accuracy increased substantially from 68% to 72% across all classes, and the average loss decreased from 1.03 to 1.02. Furthermore, the best-performing CNN-AM model achieved a high classification accuracy of 86% with a remarkably low loss of 0.53. These findings confirm that the proposed SE attention mechanism effectively and significantly enhances herbal leaf classification performance, establishing a promising and highly reliable foundation for developing efficient automated identification systems in the field.

***Keywords:*** *Attention Mechanism, Convolutional Neural Network, Deep Learning, Image Classification*

***Corresponding author***
***Name: Ragil***
***Email: dionisiusragil48@gmail.com***

## INTRODUCTION

Plants have a vital role in human life, for example as an ingredient for herbal medicine. Since the past, herbs have been used as alternative traditional medicine to prevent and naturally cure diseases. According to WHO, herbal plant usage is very high, reaching 65% in developed countries and 80% in developing countries (Basri et al., 2023). From 40.000 herbal plant species in the world, around 30.000 of them are exist in Indonesia. From that number only 1.200 species are being used by Indonesian people as an alternative medicine (Meiriyama et al., 2022). One of the most used parts of herbal plants are leaf, because it contains beneficial compounds including vitamin, mineral, and antioxidant that can help to prevent diseases and maintain health (Darmawati et al., 2024).

It's really hard for most people to identify traditional plant species, and even experts who work with plants every day, like conservationists, farmers, foresters, and landscape architects, find it quite difficult. Sometimes, even botanists, who are specialists in plants, struggle to correctly identify different species (Goeau et al., 2025). So, there is a need for a simple and dependable way to identify herbs. Using computers along with statistical methods could be a strong solution for recognizing herbs. This approach is non-invasive and can quickly identify different herbs, especially for people who can't afford costly equipment (Azlah et al., 2019). One of the solutions is utilizing machine learning, it’s a skill to get information by finding patterns in raw data. Machine learning lets computers handle problems that require understanding the real world and make choices that seem personal (Mirtehari et al., 2022).

Various study on herbal leaf classification using different algorithm have been conducted previously. In a study by L. P. R. Noviana comparing Logistic Regression and Decision Tree for herbal leaf classification, feature extracted for this classification was the color, Grey Level Co-Occurrence Matrix (GLCM) and shape. resulting test accuracy 60.24% for Logistic Regression and 78.31% for Decision Tree (Noviana et al., 2023). Roki Hidayat et al. Implementing backpropagation using segmentation process to separate the foreground and the background then, extract it’s feature with statistic method for classify herbal leaf out of 10 types resulting in average of 75% accuracy (Hidayat et al., 2025). A study by R. I. Borman using first-order feature extraction and the Multiclass Support Vector Machine (Multiclass SVM) algorithm. Resulting test identification of an average of 79% (Borman et al., 2021). These result show that older methods require features to be extracted manually mean they are less robust to real world image. A study by Alfitriana Riska et al. Compared K-Nearest Neighbour (KNN) with feature extraction Geometric Moment Invariant (GMI) and Hue Saturation Value (HSV) to CNN without feature extraction. It was concluded that the CNN had a better performance by having a higher accuracy, precision, recall, and an F1 score (Khatib et al., 2023). Another study by Rosida Pujiati and Naim Rochmawati using only CNN method with custom model including 3 layers of convolution and pooling, 3 dense layers, resulting in test accuracy 84% and 0.28 loss (Pujiati et al., 2022). Despite these advances, simple custom models often do not have high parameter efficiency or the ability to dig into image features as deeply as advanced architectures such as MobileNetV2.

There are several problems that hinder the use of CNN when relying on basic architecture. Such as overfitting, study by P. E. Mujahid et al. resulting 92.66% model accuracy however the models frequently suffer from overfitting, where they memorize the training data rather than generalizing well to new, unseen data. This limitation means that while a simple model might perform well on its training set, its real-world effectiveness is significantly compromised, highlighting the need for more robust and advanced CNN architectures (Mujahid et al., 2024). Furthermore, the use of native CNNs often results in suboptimal accuracy, indicating that there is still substantial room for improvement beyond what basic models can achieve (Purnama et al., 2020). This gap necessitates the development of more sophisticated CNN designs and innovative training techniques to overcome the inherent limitations of standard architectures and unlock the full potential of deep learning for a given task.

In light of the identified challenges with standard CNNs, specifically their susceptibility to overfitting and their limited accuracy, this study introduces a strategic solution, the integration of an attention mechanism. Attention mechanisms have proven to be a very useful way to improve the performance of convolutional neural networks in image super-resolution. As research in deep learning and image upscaling continues, many new methods have been introduced (Zhu et al., 2021). This targeted focus not only serves as a powerful regularization method to mitigate overfitting but also directly boosts the model's performance. Empirical evidence from various studies, including a notable example where an attention mechanism improved a deep learning algorithm accuracy by 3.5% (Kardakis et al., 2021).

This study is expected to make a significant contribution to the field of herbal leaf classification, offering a new and more effective system built on deep learning enhanced with an attention mechanism. The findings of this research are expected to be highly beneficial to a diverse range of professionals. Botanists can utilize this system to more accurately identify plant species, aiding in biodiversity research and conservation efforts. Pharmacists and herbalists may leverage this technology for quality control and authentication of medicinal plants, ensuring the correct herbs are used. Furthermore, the methodology and results of this research can serve as a valuable resource for other researchers, providing a robust foundation for future studies in deep learning and computer vision applications for botany and beyond.

## METHOD

### 1. Research Stages

In this research several stages were caried out. The first stage is acquisition through the same device with 48MP smartphone camera and average 10cm of range this done to avoiding outlier on the dataset, the second stage is dataset creation, the third stage is image preprocessing, the fourth stage is dataset split where the dataset splited into 3 part including training set, validation set, and test set, the fifth stage is building the deep learning models and training it using the training set and validation set, the sixth stage is evaluating the model using test set, and the seventh stage is testing the result. For a more detailed explanation, you can see the proposed research method in Fig 1.

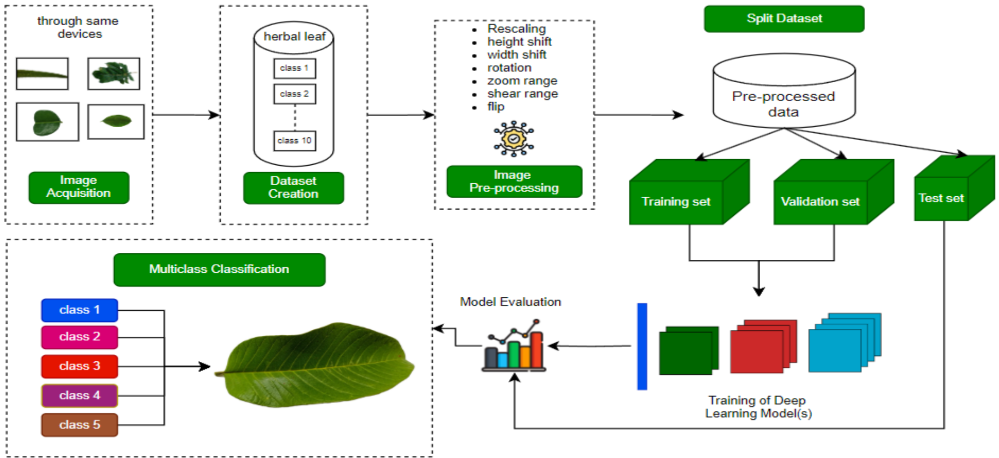


***Figure 1.*** *Proposed research method, starting from image acquisition, dataset creation, image pre-processing, split dataset, training model, model evaluation, classification.*

## 2. Dataset

Images are collected from Merapi Herbal Farma for the primary data, and supported with opensource data from Indonesian Herbal Leaf 3500 (Minarno et al., 2022), because processing of herbal leaf classification requires many images. The dataset consists of 10 type tropical leaves especially those that can grow in Indonesia. The final dataset is 20 images from Merapi Herbal Farma and 130 images from Indonesian Herbal Leaf, each image has dimension of 1600 x 1200. In addition, the example of each leaf is shown in the Table 1.

***Table 1: leaf type used in this study.***

| Image | Annotation |
|---|---|
|  | Averrhoa bilimbi leaf (Belimbing wuluh) |
|  | Psidium guajava leaf (Jambu biji) |
|  | Citrus aurantiifolia leaf (Jeruk nipis) |
|  | Ocimum citratum leaf (Kemangi) |

| | |
|---|---|
|  | Aloe vera leaf (Lidah buaya) |
|  | Artocarpus heterophyllus leaf (Nangka) |
|  | Pandanus amaryllifolius leaf (Pandan) |
|  | Carica papaya leaf (Pepaya) |
|  | Apium graveolens leaf (Seledri) |
|  | Piper betle leaf (Sirih) |

**3. Image Preprocessing**

In this process images not only being preprocessed but also augmented, first image will be resized. This resizing process is done to adjust the image size in terms of pixels. In this research, the image size is made smaller than the original size from 1600 x 1200 into 224 x 224 this size will have relation with the deep learning model later. The goal is to make the image processing faster without losing any of the information in the image (Saifullah et al., 2021). Next to prevent the limited data in this study we applied data augmentation, Data augmentation involves expanding the training dataset through creating synthetic or modified data to increase data diversity and bridge the gap between datasets and real-world scenarios. It is widely adopted and proven to be both effective and efficient. The basic concept is to create modified versions of existing data (Wang et al., 2024). For the example of augmentation result can be seen in Fig. 2 below.

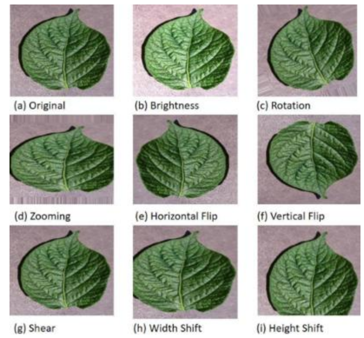


***Figure 2.*** *Augmentation sample result. Khan et al. (20).*

**4. Convolutional Neural Network**

Convolutional Neural Networks are the most widely used deep learning technique, a subtype of artificial neural network (Taye et al., 2023). Convolutional neural networks (CNNs) have been widely used successfully in various image-based applications such as classification, object detection, pattern recognition, and segmentation. Throughout their development, many different CNN architectures have been proposed with the aim of achieving superior results (Nguyen et al., 2022).

In this study the model will use transfer learning with MobileNetV2 architecture as the backbone, this architecture consists of input layer, initial convolution layer, batch normalization, 7 layers of bottleneck including of 3 x 3 depth wise layer, fully connected layer, final convolutional neural network layer, average pooling layer, and the output layer or fully connected (Swathi et al., 2025). For the detail of the MobileNetV2 architecture can be seen on this Fig. 3 below.

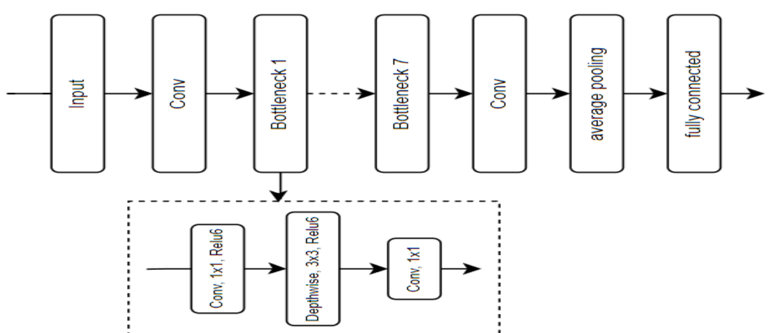


***Figure 3.*** *MobileNetV2 architecture in detail, showing the layer consist in the architecture starting from the initialize convolution until the output layer.*

a. Input Layer

The proposed CNN-attention mechanism model will start by retrieving a size of 224 x 224 x 3 input image.

b. Convolution Layer

The convolution layer is the main part and most important part of a CNN. It creates feature maps that show both simple and complex features in images. Simple features are basic things you can see in an image, like color, brightness, texture, and edges (Rathore et al., 2020). The mathematical formula is equation (1).

$$S(i,j) = \left( \sum_{m} \sum_{n} I(i-m, j-n) K(m,n) \right) + b \quad (1)$$

Where S is the output, I refer to the tensor input, K is filters convolution, m and n are the index.

c. Fully Connected Layer

The model suggests adding a fully connected layer after the convolutional layers, which combines features from several images. This layer helps identify the most important patterns used to classify images.

d. Average Pooling Layer

The pooling layer processes each group of feature maps and makes them smaller. It this study applies a 3x3 filter in two separate steps and works on each part of the input data. The mathematical formula is equation (2).

$$Y_{i,j,k} = \frac{1}{|R|} \sum_{(m,n) \in R} X_{m,n,k} \quad (2)$$

Where Y is the output, X is the input from position (m, n) in canal (k), R is the region, and |R| is the sum of the pixel element in region R. for the example of average pooling layer can be seen in the Fig. 4 below.

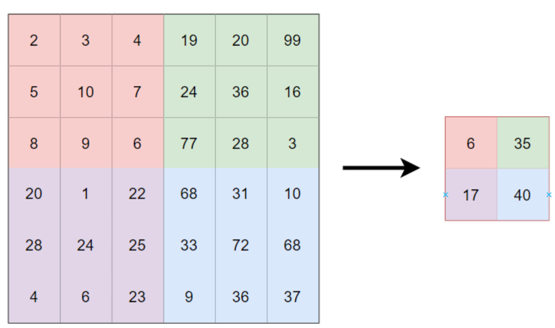


***Figure 4**. Sample of Average Pooling, it's grouping in size of 3x3 and take the average of it to make a new value.*

d. Output/Softmax Layer

The activation function is used when you need to classify things into more than one category. It takes the output from the fully connected layers, calculates the probability for each possible category, and gives results that are numbers between 0 and 1. The total of all these probabilities adds up to 1 (Azmi et al., 2023). The mathematical formula is equation (3).

$$P(y = j|z) = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \tag{3}$$

Where P is the probability of the input is a class of j, $e^{zj}$ is the exponentiation of the logit value for class j. The exponential function ensures that all resulting values are positive. And the last part of the equation is the sum of the exponentials of all logits for all classes (K).

**5. Attention Mechanism**

Attention Mechanism is a technique in deep learning that makes the model dynamically assign different weights or levels of importance to various parts of the input data, these weights are assigned based on how relevant those parts are to the task being performed, for example classification (Ruan et al., 2025). In this study type of attention mechanism will be used is squeeze-and-excitation attention. The squeeze-and-excitation (SE) block enhances representation quality by gathering global information and understanding channel relationships. This block consists of two parts: the squeeze module, which collects global spatial information through global average pooling, and the excitation module, which uses fully-connected layers and activation functions (ReLU and sigmoid) to generate attention vectors and capture channel-wise relationships (Guo et al., 2021). The mathematical formula for squeeze is equation (4) and excitation is equation (5).

$$z_c = F_{sq}(x_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_c(i,j) \tag{4}$$

$$s = F_{ex}(z, W) = \sigma\big(W_2 \delta(W_1 z)\big) \tag{5}$$

Where $x_c$ is the feature map input, H x W is the feature map dimension, $z_c$ is the scalar value of average computation for canal going to c that's for squeeze equation. Then δ is ReLU activation, σ is sigmoid activation, and $W_1$, $W_2$ is the weight matrix of two connected layer.

**6. Matrix Evaluation**

In this study confusion will be applied as matrix evaluation. The confusion matrix shows the prediction results either correct or incorrect in binary classification. It consists of four elements: false negative (FN, incorrectly predicted negative values), true negative (TN, correctly predicted negative values), false positive (FP, incorrectly predicted positive values), and true positive (TP, correctly predicted positive values) (Lu et al., 2021). Then the accuracy can be calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$F1 = \frac{2}{\frac{1}{P} + \frac{1}{R}} = \frac{2 \times P \times R}{P + R} \quad (9)$$

In research about classifying especially in leaves, accuracy is the most widely used measure (Atha et al., 2022), (Pratama et al., 2025), (Suhardin et al., 2021). Higher values of accuracy, precision, and recall are better. Within a certain range, a lower F1 score often means the model generalizes better. Once the training and testing are done, the model gains a new ability, which is then used on new data

## FINDING AND DISCUSSION

In this section, the results of the experiments are presented, followed by a comprehensive discussion of the findings. The results section will detail the performance metrics of both the baseline CNN model and the proposed model enhanced with the Squeeze and Excitation (SE) Attention mechanism. Subsequently, the discussion will analyze and interpret these results, exploring the effectiveness of the SE attention module in improving classification accuracy for herbal leaves, and will also contextualize the findings within the broader field of deep learning for plant classification.

The dataset for this study comprises 1,500 images of leaves, representing 10 distinct types or classes. To ensure a robust and unbiased evaluation of the model, the dataset was systematically partitioned into three separate sets: 80% for the training set, 10% for the test set, and 10% for the validation set. The specific distribution of images for each class is detailed in Table 2.

***Table 2: Split data distribution.***

| Class | Training set | Test set | Validation set |
|---|---|---|---|
| Averrhoa | 120 | 15 | 15 |
| Guajava | 120 | 15 | 15 |
| Citrus | 120 | 15 | 15 |
| Ocimum | 120 | 15 | 15 |
| Aloe Vera | 120 | 15 | 15 |
| Artocarpus | 120 | 15 | 15 |
| Pandanus | 120 | 15 | 15 |
| Papaya | 120 | 15 | 15 |
| Graveolens | 120 | 15 | 15 |
| Betle | 120 | 15 | 15 |

Once the dataset is prepared, the proposed CNN model will be implemented. This model utilizes a MobileNetV2 architecture, which is enhanced with a Squeeze and Excitation

(SE) attention mechanism. This strategic upgrade is anticipated to yield superior results compared to the native MobileNetV2 architecture. The detailed structure and specifications of our proposed model can be found in Table 3.

***Table 3: Proposed model summary.***

| Layer | Output Shape | Params |
|---|---|---|
| Input layer | (None, 224, 224, 3) | 0 |
| MobileNetV2 | (None, 7, 7, 1280) | 2,257,984 |
| Average pooling | (None, 1280) | 0 |
| Reshape | (None, 1, 1, 1280) | 0 |
| Dense | (None, 1, 1, 80) | 102,400 |
| Dense | (None, 1, 1, 1280) | 102,400 |
| Multiply | (None, 7, 7, 1280) | 0 |
| Average pooling | (None, 1280) | 0 |
| Dropout | (None, 1280) | 0 |
| Dense/output | (None, 10) | 12,810 |
| Total params: 2,475,594 | | |
| Trainable params: 217,610 | | |
| Non-trainable params: 2,257,984 | | |

The model's architecture begins with an input layer accepting 224x224 pixel RGB images. The core feature extractor is the pre-trained MobileNetV2 backbone, which processes the input and outputs a (7, 7, 1280) feature tensor. This output then passes through several subsequent layers, including average pooling, reshape, dense, multiply, and dropout.The model contains a total of 2,475,594 parameters. Critically, only 217,610 are trainable (optimized during training), while the majority (2,257,984) are non-trainable.

The training strategy aims to comprehensively compare the native CNN and the enhanced CNN models. This is achieved by systematically testing various hyperparameter combinations. The parameters varied and tested are the Learning Rate (0.01, 0.001, 0.0001) and the Optimizer (Adam, RMSprop, SGD). Other parameters (epoch, batch size, and dropout) are kept constant. This systematic evaluation determines the configuration that yields the highest accuracy and the most stable training process. The resulting performance is summarized using the test data in Table 4.

***Table 4: training result comparison between native cnn and cnn enhanced with attention mechanism.***

| | CNN | | | CNN/AM | | |
|---|---|---|---|---|---|---|
| **Parameter** | **Accuracy** | **loss** | **F1-score** | **accuracy** | **loss** | **F1-score** |
| Adam/0.01 | 0.79 | 0.83 | 0.78 | 0.82 | 0.88 | 0.80 |
| Adam/0.001 | 0.80 | 0.52 | 0.79 | 0.79 | 0.98 | 0.77 |
| Adam/0.0001 | 0.65 | 1.18 | 0.64 | 0.78 | 1.09 | 0.77 |
| RMSprop/0.01 | 0.76 | 1.03 | 0.76 | 0.86 | 0.53 | 0.86 |

| | | | | | | |
|---|---|---|---|---|---|---|
| RMSprop/0.001 | 0.78 | 0.61 | 0.78 | 0.82 | 0.62 | 0.81 |
| RMSprop/0.0001 | 0.72 | 0.96 | 0.72 | 0.76 | 0.95 | 0.74 |
| SGD/0.01 | 0.78 | 0.64 | 0.78 | 0.80 | 0.79 | 0.80 |
| SGD/0.001 | 0.66 | 1.22 | 0.65 | 0.68 | 1.21 | 0.66 |
| SGD/0.0001 | 0.20 | 2.30 | 0.16 | 0.17 | 2.12 | 0.15 |
| Average | 0.68 | 1.03 | 0.67 | 0.72 | 1.02 | 0.71 |

The experimental results show that adding an attention mechanism to a Convolutional Neural Network (CNN) greatly improves the performance of classifying herbal leaves. This mechanism helps the network focus on the most important and useful features in the leaf images, which increases accuracy. In nine experiments, the attention-enhanced CNN (CNN/AM) performed better than the standard CNN each time, achieving an average accuracy of 0.72 compared to 0.68. The improvement is most clear in the F1-score, where CNN/AM scored 0.71 versus 0.67 for the standard CNN, showing a better balance between precision and recall. Even though the loss difference is small (1.02 vs. 1.03), the lower loss still means the model makes fewer mistakes.

The biggest improvement was achieved with the RMSprop/0.01 setup. In this configuration, adding attention raised both accuracy and F1-score from 0.76 to 0.86, an increase of 0.10, while the loss dropped significantly from 1.03 to 0.53. This peak performance of 86% accuracy marks a clear advancement compared to previous deep learning approaches in the field, surpassing the 84% accuracy by (Pujiati et al., 2022) using a custom CNN model. Furthermore, the significant loss reduction confirms the theoretical implication that the Squeeze-and-Excitation (SE) effectively recalibrates channel-wise features, leading to a much more robust and confident learning process.

The best model's top performance is clearly shown not just by the final test results but also by how well it behaved during training. This adds more proof that the CNN with attention mechanism model, with the right setup, is both reliable and efficient. The training result can be seen in Fig. 5.

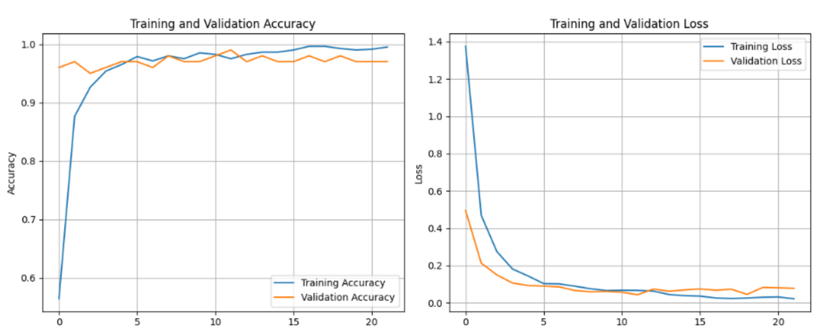


***Figure 5.*** *highest training result with proposed model.*

During training, the model's accuracy chart showed a smooth and steady increase. It didn't jump up and down but gradually went up over time, showing that the model was learning consistently and effectively with each training round. This steady rise suggests the model didn't overfit and was able to learn the important features from the training data. At the same time, the loss chart showed the opposite trend and was just as good. The loss value kept going down steadily, meaning the model was constantly improving its predictions and reducing mistakes. This consistent drop shows that the training process worked well and the model found the best settings to classify the data correctly. The fact that the accuracy went up and the loss went down together is a common sign of a well-trained model.

Later on the best model were also tested using confusion matrix to see it is performance on data the model have not see. The data used for testing are 10% from the dataset or 150 images with 15 images for each class. For the result of confusion matrix can be seen in Fig. 6.

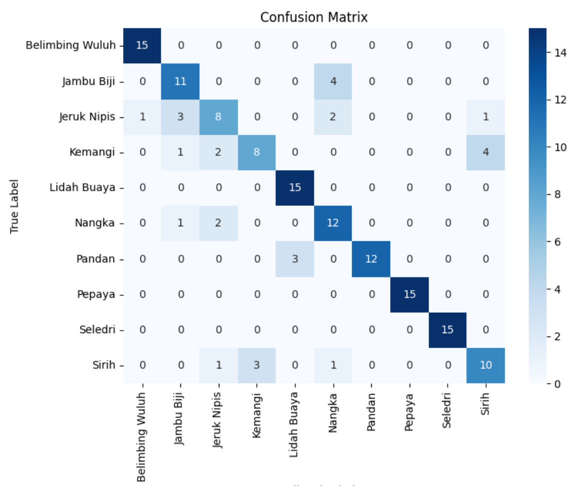


***Figure 6.*** *confusion matrix result with proposed model.*

As shown in the matrix, the model exhibits a very high level of accuracy across most classes, which is clearly indicated by the large number of correct predictions along the main diagonal. For instance, classes such as Avverhoa, Aloe vera, and Papaya were classified with near-perfect accuracy, with all 15 samples correctly identified. Similarly, the model performed exceptionally well on Guajava and Pandanus leaf classifications. However, the matrix also highlights specific areas where the model encountered some challenges, which are represented by the non-zero values outside the main diagonal. The primary confusion appears to be among a few classes, such as Citrus, Ocimum, and Betle.

## CONCLUSION

### 1. Summary

This study proposed a new approach to improve the accuracy of classifying Indonesian herbal leaves by adding the Squeeze-and-Excitation (SE) attention mechanism to a MobileNetV2 model. The results were very promising: the CNN with the attention mechanism (CNN-AM) raised the average accuracy across nine different training situations from 68% to 72%. This shows that the SE attention helps the model pay more attention to the most important features, making the attention-based design a better choice. However, the study also has some drawbacks. One issue is that the dataset wasn't very varied, so it didn't include different environmental conditions, lighting situations, or various stages of leaf growth. This lack of diversity might make it harder for the model to work well in real-world settings. Another limitation is that the study only tested one type of model backbone (MobileNetV2) and one specific attention method (Squeeze-and-Excitation).

**2. Future Work**

To further advance the model's performance and generalization, we recommend that future research focus on three key areas. First, diversifying the dataset by incorporating a wider range of data points will improve the model's robustness to real-world variations. Second, investigating alternative backbone architectures, such as Vision Transformers (ViT) or more advanced CNNs like EfficientNet, could lead to more effective feature extraction. Finally, exploring novel types and combinations of attention mechanisms beyond a single implementation will allow the model to better identify and utilize the most important features, potentially yielding more accurate and nuanced results.

## REFERENCES

Al-Zoghby, A. M., Al-Awadly, E. M. K., Moawad, A., Yehia, N., & Ebada, A. I. (2023). Dual Deep CNN for Tumor Brain Classification. *Diagnostics 2023, Vol. 13, Page 2050*, *13*(12), 2050. https://doi.org/10.3390/DIAGNOSTICS13122050

Atha, A. M., & Zuliarso, E. (2022). Herbal Plants Detection Specifically For Skin And Hair Diseases Using The Convolutional Neural Network (CNN) and Tensorflow. *JUPITER: Jurnal Penelitian Ilmu Dan Teknologi Komputer*, *14*(2-a), 01–10. https://doi.org/10.5281./4736/5.jupiter.2022.10

Azlah, M. A. F., Chua, L. S., Rahmad, F. R., Abdullah, F. I., & Alwi, S. R. W. (2019). Review on Techniques for Plant Leaf Classification and Recognition. *Computers 2019, Vol. 8, Page 77*, *8*(4), 77. https://doi.org/10.3390/COMPUTERS8040077

Azmi, K., Defit, S., & Putra Indonesia YPTK Padang Jl Raya Lubuk Begalung-Padang-Sumatera Barat, U. (2023). Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat. *JURNAL UNITEK*, *16*(1), 28–40. https://doi.org/10.52072/unitek.v16i1.504

Basri, H., Purnawansyah, P., Darwis, H., & Umar, F. (2023). Klasifikasi Daun Herbal Menggunakan K-Nearest Neighbor dan Convolutional Neural Network dengan Ekstraksi Fourier Descriptor. *Jurnal Teknologi Dan Manajemen Informatika*, *9*(2), 79–90. https://doi.org/10.26905/JTMI.V9I2.10350

Borman, R. I., Rossi, F., Jusman, Y., Rahni, A. A. A., Putra, S. D., & Herdiansah, A. (2021). Identification of Herbal Leaf Types Based on Their Image Using First Order Feature Extraction and Multiclass SVM Algorithm. *2021 1st International Conference on*

*Electronic and Electrical Engineering and Intelligent System, ICE3IS 2021*, 12–17. https://doi.org/10.1109/ICE3IS54102.2021.9649677

Darmawati, A. R. R., Purnawansyah, Darwis, H., & Ilmawan, L. B. (2024). Digital Image Classification of Herbal Leaves Using Support Vector Machine and Convolutional Neural Network with Fourier Descriptor Features. *CSRID (Computer Science Research and Its Development Journal)*, *16*(1), 01–12. https://doi.org/10.22303/CSRID-.16.1.2024.01-12

Daun, K., Berdasarkan, H., Bentuk, F., Menggunakan, D. T., Meiriyama, K., Devella, S., & Adelfi, S. M. (2022). Klasifikasi Daun Herbal Berdasarkan Fitur Bentuk dan Tekstur Menggunakan KNN. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, *9*(3), 2573–2584. https://doi.org/10.35957/JATISI.V9I3.2974

Goeau, H., Bonnet, P., & Joly, A. (2025). *Plant identification based on noisy web data: the amazing performance of deep learning (LifeCLEF 2017)*. https://arxiv.org/pdf/2509.20856

Guo, M.-H., Xu, T.-X., Liu, J.-J., Liu, Z.-N., Jiang, P.-T., Mu, T.-J., Zhang, S.-H., Martin, R. R., Cheng, M.-M., & Hu, S.-M. (2021). Attention Mechanisms in Computer Vision: A Survey. *Computational Visual Media*, *8*(3), 331–368. https://doi.org/10.1007/s41095-022-0271-y

Hidayat, R., Toyib, R., Apridiansyah, Y., Reswan, Y., Bali, J., Bali, K., Segara, T., & Bengkulu, K. (2025). KLASIFIKASI CITRA DAUN HERBAL DENGAN KHASIATNYA UNTUK PENGOBATAN MENGGUNAKAN JARINGAN SYARAF TIRUAN (BACKPROPAGATION). *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(4), 5571–5577. https://doi.org/10.36040/JATI.V9I4.13860

Kardakis, S., Perikos, I., Grivokostopoulou, F., & Hatzilygeroudis, I. (2021). Examining Attention Mechanisms in Deep Learning Models for Sentiment Analysis. *Applied Sciences 2021, Vol. 11, Page 3883*, *11*(9), 3883. https://doi.org/10.3390/APP11093883

Khatib Sulaiman, J., Perbandingan Kombinasi GMI, S., CNN pada Klasifikasi Daun Herbal Alfitriana Riska, dan, Darwis, H., & Astuti, W. (2023). Studi Perbandingan Kombinasi GMI, HSV, KNN, dan CNN pada Klasifikasi Daun Herbal. *The Indonesian Journal of Computer Science*, *12*(3). https://doi.org/10.33022/IJCS.V12I3.3210

Lu, J., Tan, L., & Jiang, H. (2021). Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification. *Agriculture 2021, Vol. 11, Page 707*, *11*(8), 707. https://doi.org/10.3390/AGRICULTURE11080707

Marpaung, N. L., Butar, R. J. H. B., & Hutabarat, S. (2023). Implementasi Deep learning untuk Identifikasi Daun Tanaman Obat Menggunakan Metode Transfer learning. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*, *9*(3), 348–354. https://jurnal.untan.ac.id/index.php/jepin/article/view/63895

Minarno, A. E., Wicaksono, G. W., Azhar, Y., & Hasanuddin, M. Y. (2022). *Indonesian Herb Leaf Dataset 3500*. *1*. https://doi.org/10.17632/S82J8DH4RR.1

Mirtaheri, S. L., & Shahbazian, R. (2022). Machine Learning : Theory to Applications. *Machine Learning Theory to Applications*. https://doi.org/10.1201/9781003119258
Mujahid, P. E., Manik, R., Simbolon, J. S., Sinaga, M. R. R. S., Aisyah, S., Nababan, M., & Harmaja, O. J. (2024). Herbal Leaf Image Classification Using Convolutional Neural Network (CNN). *Jurnal Sistem Informasi Dan Ilmu Komputer*, *8*(1), 52–68. https://doi.org/10.34012/JURNALSISTEMINFORMASIDANILMUKOMPUTER.V8I1.5145
Nguyen, H. T., Li, S., & Cheah, C. C. (2022). A Layer-Wise Theoretical Framework for Deep Learning of Convolutional Neural Networks. *IEEE Access*, *10*, 14270–14287. https://doi.org/10.1109/ACCESS.2022.3147869
Noviana, L. P. R., & Nugraha, I. N. B. S. (2023). Perbandingan Klasifikasi Citra Daun Herbal Menggunakan Metode Logistic Regression dan Decision Tree Classifier Berdasarkan Fitur (Warna, GLCM, Bentuk). *JITU : Journal Informatic Technology And Communication*, *7*(2), 126–133. https://doi.org/10.36596/JITU.V7I2.1241
*Perbandingan Klasifikasi Citra Daun Herbal Menggunakan Metode Logistic Regression dan Decision Tree Classifier Berdasarkan Fitur (Warna, GLCM, Bentuk) | JITU : Journal Informatic Technology And Communication*. (n.d.). Retrieved September 6, 2025, from https://ejournal.uby.ac.id/index.php/jitu/article/view/1241
Pratama, A., Sugiharto, T., & Novantara, P. (2025). Classification of Avocado Plant Varieties Based on Leaf Shape Using CNN Algorithm. *Jurnal CoSciTech (Computer Science and Information Technology)*, *6*(2), 120–128. https://doi.org/10.37859/COSCITECH.V6I2.9474
Pujiati, R., & Rochmawati, N. (2022). Identifikasi Citra Daun Tanaman Herbal Menggunakan Metode Convolutional Neural Network (CNN). *Journal of Informatics and Computer Science (JINACS)*, *3*(03), 351–357. https://doi.org/10.26740/JINACS.V3N03.P351-357
Purnama, I. N. (2020). HERBAL PLANT DETECTION BASED ON LEAVES IMAGE USING CONVOLUTIONAL NEURAL NETWORK WITH MOBILE NET ARCHITECTURE. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, *6*(1), 27–32. https://doi.org/10.33480/JITK.V6I1.1400
Rathore, S., Niazi, T., Iftikhar, M. A., & Chaddad, A. (2020). Glioma Grading via Analysis of Digital Pathology Images Using Machine Learning. *Cancers 2020, Vol. 12, Page 578*, *12*(3), 578. https://doi.org/10.3390/CANCERS12030578
Ruan, T., & Zhang, S. (2024). *Towards understanding how attention mechanism works in deep learning*. https://arxiv.org/pdf/2412.18288
Saifullah, S., Suryotomo, A. P., & Yuwono, B. (2021). Fish Detection Using Morphological Approach Based-on K-Means Segmentation. *Compiler*, *10*(1). https://doi.org/10.28989/compiler.v10i1.946
Suhardin, I., Patombongi, A., Muhammad Islah, A., & Catur Sakti Kendari Jl Abdullah, S. H. (2021). MENGIDENTIFIKASI JENIS TANAMAN BERDASARKAN CITRA DAUN MENGGUNAKAN AlGORITMA CONVOLUTIONAL NEURAL NETWORK. *Simtek :*

*Jurnal Sistem Informasi Dan Teknik Komputer*, *6*(2), 100–108. https://doi.org/10.51876/SIMTEK.V6I2.101

SWATHI, N., HARI, DR. N. S., & VEDANTAM, DR. R. (2025). SATELLITE IMAGE CLASSIFICATION: A DEEP LEARNING APPROACH USING MOBILENETV2. *International Journal of Innovative Research in Technology*, *12*(2), 3856–3862. https://ijirt.org/article?manuscript=182947

Taye, M. M. (2023). Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. *Computation 2023, Vol. 11, Page 52*, *11*(3), 52. https://doi.org/10.3390/COMPUTATION11030052

Wang, Z., Wang, P., Liu, K., Wang, P., Fu, Y., Lu, C.-T., Aggarwal, C. C., Pei, J., & Zhou, Y. (2024). *A Comprehensive Survey on Data Augmentation*. https://arxiv.org/pdf/2405.09591

Zhu, H., Xie, C., Fei, Y., & Tao, H. (2021). Attention Mechanisms in CNN-Based Single Image Super-Resolution: A Brief Review and a New Perspective. *Electronics 2021, Vol. 10, Page 1187*, *10*(10), 1187. https://doi.org/10.3390/ELECTRONICS10101187