Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th February 2026. Vol. 5. No. 2; e-ISSN: 2808-4519

Optimization of Convolutional Neural Networks Using Resizing Techniques for Banana Leaf Disease Classification

Aldiyansyah Kurniawan[1]*, Ade Irma Purnamasari[2], Denni Pratama[3], Edi Tohidi[4], Edi Wahyudin[5]

[1,2]Prodi Teknik Informatika, STMIK IKMI Cirebon
[3]Prodi Manajemen Informatika, STMIK IKMI Cirebon
[4,5]Prodi Komputerisasi Akuntansi, STMIK IKMI Cirebon
aldiyansyahkurniawann@gmail.com[1], irma2974@yahoo.com[2], pratamadenni@gmail.com[3], tohidiikmi@gmail.com[4], serdos.ikmi@gmail.com[5]*

Abstract

Early and accurate identification of banana leaf diseases is essential for supporting digital agriculture, as visual symptoms often require rapid and reliable analysis. This study investigates the impact of three image resizing techniques squashing, letterboxing, and random resized crop on the performance of the MobileNetV2 architecture in classifying four categories of banana leaf images using the Banana Leaf Disease Dataset v4 consisting of 4,675 samples. The experiments were conducted using a transfer learning approach with an 80:10:10 data split, standardized normalization, and data augmentation. The results show that all resizing techniques achieved test accuracies above 92%. Squashing produced the highest accuracy and fastest training time, letterboxing demonstrated the most stable performance with the lowest validation loss, and random resized crop improved generalization to variations in object position. These findings confirm that resizing strategies significantly influence the stability and effectiveness of CNN models. Overall, MobileNetV2 proves capable of delivering accurate and efficient classification of banana leaf diseases when supported by an appropriate preprocessing pipeline. This study provides empirical evidence for developing image-based plant disease diagnosis systems within smart agriculture.

Keywords: *MobileNetV2, image resizing, banana leaf disease, deep learning, transfer learning.*

1. Introduction

Advancements in information technology and the adoption of smart agriculture have transformed approaches to monitoring plant health, including the detection of banana leaf diseases that significantly affect productivity. Manual methods based on visual observation are considered inefficient, making computer vision technologies powered by deep learning a viable alternative capable of rapidly and accurately identifying visual patterns such as color changes, texture variations, and leaf structural abnormalities [1], [2]. The integration of image-based detection systems into field devices, including edge computing platforms, further strengthens the urgency of using efficient and adaptive models for plant diagnosis under diverse environmental conditions [3].

Convolutional Neural Network (CNN) architectures have proven effective for classifying leaf diseases across various agricultural commodities due to their ability to hierarchically extract visual features [4], [5]. However, CNN performance is highly influenced by the quality of input images, particularly in the preprocessing stage, which includes variations in lighting, background, image size, and aspect ratio [6]. Image resizing techniques play a critical role as they determine the spatial consistency of images before being processed by the model. Studies such as [7], [8], and [9] demonstrate that resizing methods can affect feature representation, training speed, and model stability, yet in-depth comparative evaluations in the context of banana leaf disease remain limited.

These limitations indicate a research gap, particularly regarding the differential impact of three resizing techniques—squashing, letterboxing, and random resized crop—when applied to lightweight architectures such as MobileNetV2. Each technique produces distinct input characteristics: squashing may distort shape, letterboxing introduces non-informative padding areas, while random resized crop may cut out important features. To date, few studies have systematically examined these three techniques in banana leaf disease classification, even though variations in aspect ratio and image size are typically high in agricultural datasets and may influence a model's generalization ability [10], [11].

This study aims to evaluate the impact of the three resizing techniques using a transfer learning approach on the MobileNetV2 architecture with the Banana Leaf Disease Dataset v4. The evaluation is performed based on accuracy, loss values, and training time efficiency. Theoretically, this research enriches the literature on the relationship between image preprocessing and CNN performance in the agricultural domain [12], [13]. Practically, it offers strategic recommendations for selecting the most effective resizing technique to develop accurate, efficient, and field-ready image-based disease detection systems, applicable both at the farmer level and within modern precision agriculture systems.

2. Research Method

2.1. Research Design

This study employs a quantitative experimental design to evaluate the effect of three image resizing techniques on the performance of a Convolutional Neural Network (CNN) model in classifying banana leaf diseases. The resizing techniques—squashing, letterboxing, and random resized crop serve as independent variables, while the classification performance of the MobileNetV2 model functions as the dependent variable. This approach follows the principles of controlled experimental design as described by [14] and [15], which emphasize the importance of controlling variables to ensure that changes in model performance are influenced solely by the applied treatments. The entire research process is organized in a structured workflow, beginning with dataset exploration, followed by preprocessing, transfer-learning-based model training, and concluding with final evaluation. This workflow is visualized in the research methodology diagram shown in Fig 1.

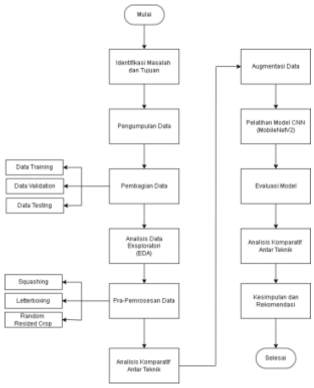

Fig. 1. Research Design

The study adopts a between-group design, where each resizing technique is treated as a separate experimental group. Each experiment is conducted using identical training parameters to maintain internal validity, and all treatments are replicated three times to ensure experimental stability, as recommended by [16].

2.2. Dataset and Image Description

The study adopts a between-group design, where each resizing technique is treated as a separate experimental group. Each experiment is conducted using identical training parameters to maintain internal validity, and all treatments are replicated three times to ensure experimental stability, as recommended by [16].

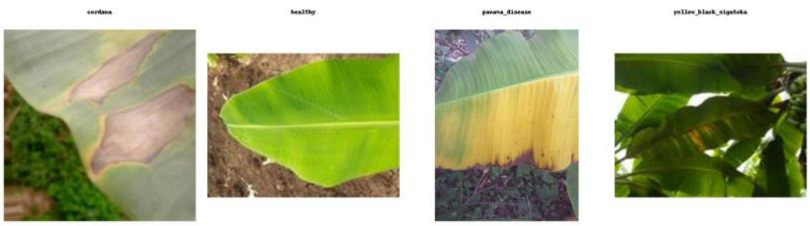

Fig. 2. Sample Images of Each Class

Preliminary exploration shows that the image sizes vary widely, typically ranging from 256×256 to 1024×1024 pixels. Color visualizations also reveal distinct patterns across color channels between healthy and infected leaves. Sample images and color histograms may be included to illustrate the dataset's characteristics. The dataset is split using stratified sampling with proportions of 80% training, 10% validation, and 10% testing. This distribution ensures that class proportions remain balanced across each subset to maintain the model's generalization quality. Previous studies such as [17] and [18] have adopted similar approaches in leaf image analysis, but have not compared resizing techniques, making this research contribution a fill to that gap.

2.3. Preprocessing and Resizing Techniques

The preprocessing stage is applied to standardize images before they are processed by the model. This study evaluates three resizing techniques. The *squashing* technique resizes images to 224×224 without preserving the aspect ratio, which may distort the shape of the objects. The *letterboxing* technique maintains the original aspect ratio by adding padding before resizing. Meanwhile, *random resized crop* introduces spatial variation by randomly cropping portions of the image before scaling them to the standard size.

Fig. 3. Banana Leaf Images After Preprocessing

All images are normalized to the range [0,1] and augmented through rotation, shifting, zooming, and flipping to increase data variability and reduce the risk of overfitting. These steps follow the recommendations of [19], which highlight the importance of augmentation in image classification tasks.

2.4. MobileNetV2 Model

The primary model used in this study is MobileNetV2, a lightweight CNN architecture based on *depthwise separable convolution* and *linear bottlenecks*, known for its efficiency in edge computing applications [20], [21]. The model is applied through a transfer learning approach with pretrained ImageNet weights. The initial layers are frozen, while the final layers are modified to produce four outputs corresponding to the number of classes.

Model: "functional"

Layer (type)	Output Shape	Param #
input_layer_1 (InputLayer)	(None, 224, 224, 3)	0
mobilenetv2_1.00_224 (Functional)	(None, 7, 7, 1280)	2,257,984
global_average_pooling2d (GlobalAveragePooling2D)	(None, 1280)	0
dropout (Dropout)	(None, 1280)	0
dense (Dense)	(None, 4)	5,124

Total params: 2,263,108 (8.63 MB)
Trainable params: 5,124 (20.02 KB)
Non-trainable params: 2,257,984 (8.61 MB)

Fig. 4. MobileNetV2 Architecture

The model is trained using the Adam optimizer with a learning rate of 0.0001, a batch size of 32, and callbacks including EarlyStopping and ReduceLROnPlateau. Each resizing technique is executed in three training replications to ensure consistency.

2.5. Research Implementation Workflow

The implementation workflow begins with dataset exploration, data splitting, preprocessing, and the application of resizing techniques. Afterward, the processed images undergo augmentation to expand sample variability. The next stage is training the MobileNetV2 model using transfer learning, followed by validation during training. Once the model reaches convergence, evaluation is performed on the test data to measure its performance. This workflow is applied consistently across all resizing techniques so that performance differences arise solely from the resizing treatments.

where NNN is the number of samples, yiy_iyi is the true label, and y^i\hat{y}_iy^i is the predicted label. This formulation is more appropriate for multi-class settings compared to the TP–TN representation, which is primarily suitable for binary classification.

In addition to accuracy, loss is computed using the categorical cross-entropy function, which is standard for multi-class classification tasks. The loss function is defined as:

$$Loss = -\left(\frac{1}{N}\right) sum_{i=1}^{N} sum_{c=1}^{C} y_{ic} * \log(\hat{y}_{ic})$$

with CCC representing the number of classes, yic=1y_{ic} = 1yic=1 if the iii-th sample belongs to class ccc, and y^ic\hat{y}_{ic}y^ic the predicted probability for that class. This loss value reflects how well the model captures the distribution of visual patterns in the training data and maintains them on the test data. To analyze classification errors more specifically, this study uses a confusion matrix, formulated as:

$$CM_{i,j} = |\{ x : y = i, \hat{y} = j \}|$$

which represents the number of samples with actual class i that are predicted as class j. This matrix helps identify which classes are most frequently misclassified and the visual characteristics that may lead to errors.

All evaluations are conducted on the previously separated test data, using the Scikit-learn library to ensure consistent computation. Meanwhile, the accuracy and loss curves will be discussed in more detail in the results section. This evaluation approach follows the practices used in prior studies by [13] and [12].

2.6. Model Evaluation

Model performance in this study is evaluated using three main metrics: accuracy, loss values, and the confusion matrix. These metrics collectively describe the model's prediction quality on the test dataset. Accuracy is calculated based on the proportion of correct predictions relative to all test samples. The accuracy formula used follows the standard definition in classification evaluation:

$$Accuracy = \frac{\left(sum_{i=1}^{N} 1(\hat{y}_i = y_i)\right)}{N}$$

where N is the number of samples, y_i is the true label, and $\hat{y}_i$ is the predicted label. This formulation is more appropriate for multi-class settings compared to the TP–TN representation, which is primarily suitable for binary classification.

In addition to accuracy, loss is computed using the categorical cross-entropy function, which is standard for multi-class classification tasks. The loss function is defined as:

$$Loss = -\left(\frac{1}{N}\right) sum_{i=1}^{N} sum_{c=1}^{C} y_{ic} * \log(\hat{y}_{ic})$$

with C representing the number of classes, y_{ic}=1 if the i -th sample belongs to class c, and $\hat{y}_{ic}$ the predicted probability for that class. This loss value reflects how well the model captures the distribution of visual patterns in the training data and maintains them on the test data. To analyze classification errors more specifically, this study uses a confusion matrix, formulated as:

$$CM_{i,j} = |\{x : y = i, \hat{y} = j\}|$$

which represents the number of samples with actual class i that are predicted as class j. This matrix helps identify which classes are most frequently misclassified and the visual characteristics that may lead to errors.

All evaluations are conducted on the previously separated test data, using the Scikit-learn library to ensure consistent computation. Meanwhile, the accuracy and loss curves will be discussed in more detail in the results section. This evaluation approach follows the practices used in prior studies by [13] and [12].

3. Results And Discussion

3.1. Data Analysis

Initial analysis of the Banana Leaf Disease Dataset v4 was carried out to understand class distribution, variations in image dimensions, and color characteristics relevant to the classification process. The dataset consists of 4,675 images divided into four main classes: Cordana, Panama Disease, Yellow and Black Sigatoka, and Healthy. Identification results show a noticeable imbalance in distribution, where the Yellow and Black Sigatoka class accounts for more than half of the total data, while Cordana is the least represented class. This condition is important to consider because it may influence the model's sensitivity to visual patterns in minority classes. A summary of the number of images in each class is presented in Table 1.

Table 1. Dataset Distribution per Class

Kelas	Train	Val	Test	Total
Cordana	239	52	51	342
Healthy	701	150	150	1001
Panama Disease	584	125	126	835
Yellow & Black Sigatoka	1748	374	375	2497
Total	**3272**	**701**	**702**	**4675**

In addition to class distribution, image dimensions also vary widely, ranging from low-resolution images of 150×113 pixels to as large as 1024×1024 pixels. This heterogeneity reinforces the need for appropriate resizing techniques before feeding images into MobileNetV2. A summary of the statistical characteristics of image sizes is presented in Table 2.

Table 2. Image Dimension Statistics

Statistik	Lebar	Tinggi	Rasio Aspek
Minimum	150	113	0.45
Maksimum	1024	1024	2.22
Rata-rata	491.73	507.66	1.05
Simpangan Baku	208.19	214.46	0.43

Color characteristics show consistent patterns: healthy leaves tend to have high intensity in the green channel, whereas diseased classes exhibit increased red and blue channel intensities due to pigmentation changes. Although brightness levels differ across classes, their ranges still overlap, indicating that color alone is insufficient as a standalone feature for accurate classification.

3.2. Experimental Results

This experiment compares three resizing techniques—squashing, letterboxing, and random resized crop—using the MobileNetV2 architecture. All methods were evaluated under identical training parameters to ensure that the performance differences genuinely reflect the impact of each resizing technique. Each method achieved a test accuracy above 92%, although their stability, convergence patterns, and sensitivity to image shape variations exhibited notable differences.

With the squashing technique, images are directly resized to 224×224 without preserving the aspect ratio. The resulting distortion does not significantly hinder learning; in fact, the model converges more quickly and achieves the highest accuracy because the texture patterns associated with leaf disease remain intact and are easily captured by the CNN. In the letterboxing approach, the original aspect ratio is preserved through padding, resulting in the most stable training process with the lowest validation loss. The leaf structure remains intact, enabling the model to more easily recognize contours and spot patterns, while the padding areas do not meaningfully interfere with MobileNetV2's feature extraction. Meanwhile, random resized crop forces the model to learn from random portions of the images, improving adaptability to data variation. However, the risk of cropping out important regions leads to lower accuracy and higher loss compared to the other two methods, even though the model becomes more robust to changes in framing and orientation. The quantitative results of all three techniques are presented in Table 3.

Table 3. Summary of Resizing Performance

Teknik	Train Acc	Val Acc	Test Acc	Val Loss	Waktu (menit)
Squashing	0.978	0.961	**0.963**	0.12	**18.2**
Letterboxing	0.981	**0.965**	0.956	**0.10**	19.1
RandomResizedCrop	0.951	0.940	0.927	0.18	20.5

3.3. Performance Analysis

Analysis of the confusion matrix shows that the classification error patterns differ across the three resizing techniques. Squashing tends to misclassify the Cordana class, which is often predicted as Panama Disease due to changes in spot shape caused by distortion. In contrast, letterboxing produces a more consistent prediction distribution because the leaf shape remains intact. For random resized crop, errors are more widely distributed, particularly when disease regions are partially cropped, resulting in incomplete features.

The accuracy and loss curves provide additional insights into the learning dynamics. Squashing demonstrates the fastest convergence; letterboxing shows the greatest stability; whereas random resized crop exhibits fluctuations that reflect the high spatial variation introduced by the training data.

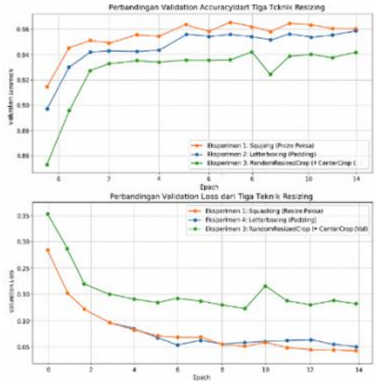

Fig. 5. Comparison of Validation Accuracy and Loss Across Techniques

On the test set, squashing continues to yield the highest accuracy, while letterboxing maintains the lowest prediction errors. Meanwhile, random resized crop offers the best generalization capability for non-uniform field images.

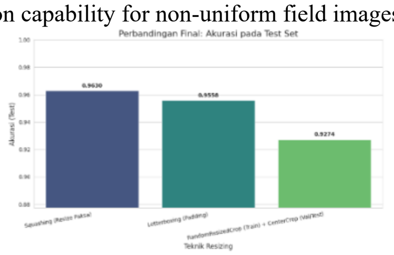

Fig. 6. Comparison of Test Set Accuracy

3.4. Comparative Analysis and Discussion

The performance of the three resizing techniques indicates that preprocessing quality has a direct impact on the effectiveness of feature extraction in MobileNetV2. Techniques that alter image proportions, such as squashing, improve input consistency and accelerate learning, but the resulting distortion may interfere with geometric features. Techniques that preserve proportions, such as letterboxing, produce the most stable learning behavior. Meanwhile, randomized techniques like random resized crop enhance generalization but sacrifice classification precision. These findings align with, yet also differ from, several previous studies. A summary is presented in Table 4.

Table 4. Comparison of Previous Studies

Penelitian	Arsitektur	Fokus	Resizing	Hasil Utama	Gap Penelitian
Althuniyan et al. (2024)	DeepLeaf CNN	Daun umum	Tunggal	Akurasi meningkat oleh augmentasi	Tidak menguji resizing
Helmawati & Utami (2025)	MobileNetV2	Daun pisang	224×224 tetap	Akurasi tinggi	Tidak membandingkan resizing
Auni & Sugiharti (2025)	MobileNetV2 + DenseNet	Mangga	Standar	Preprocessing membantu	Tidak mengevaluasi tiga teknik
Batool et al. (2024)	T-Net	Tomat	Dasar	Model efisien	Tidak fokus pada resizing
Penelitian ini (2025)	**MobileNetV2**	**Daun pisang**	**3 teknik resizing**	**Akurasi 92–96% dengan pola berbeda**	**Mengisi gap evaluasi resizing**

The final results confirm that letterboxing is the most balanced technique, squashing is the fastest and most accurate, and random resized crop is the most robust to field-level data variations. Therefore, the choice of resizing technique should be aligned with the intended implementation needs of real-world classification systems.

4. Conclusion

This study evaluated the impact of three image resizing techniques—squashing, letterboxing, and random resized crop on the performance of MobileNetV2 in classifying banana leaf diseases. Overall, the model demonstrated high performance across all methods; however, each technique presents distinct characteristics and implications for model stability and generalization. Based on the experimental results, the main conclusions are as follows:

1. All resizing techniques achieved test accuracies above 92%, indicating that MobileNetV2 can adapt well to variations in image dimensions as long as preprocessing is applied consistently.
2. Squashing provided the best accuracy and efficiency, yielding the fastest training time and highest test accuracy, although it may distort leaf shape under extreme aspect ratios.
3. Letterboxing produced the most stable performance, characterized by the lowest validation loss and preservation of the original aspect ratio, making it suitable for long-term deployment.

4. Random resized crop improved model robustness, as the variability of cropped regions during training enhanced generalization to real-world conditions, despite yielding lower accuracy due to the risk of losing important features.
5. The choice of resizing technique is a crucial component of the CNN pipeline, with squashing excelling in efficiency, letterboxing excelling in stability, and random resized crop excelling in adaptiveness to real-world data.

5. Recommendations

To broaden the scope of research and further improve model performance, several directions for future work can be considered:

1. Implementing alternative CNN architectures such as EfficientNet, ResNet, or MobileNetV3 to evaluate whether the influence of resizing techniques remains consistent across modern models.
2. Incorporating advanced evaluation metrics, including precision, recall, F1-score, and explainability analysis based on Grad-CAM to better understand dominant visual features.
3. Conducting generalization tests on real-field images, including those with uneven lighting, complex backgrounds, and random capture angles.
4. Performing cross-dataset evaluations to assess the reliability of resizing techniques on banana leaf datasets from other sources or on different plant disease datasets.
5. Testing the model on mobile or edge devices to evaluate inference speed, memory consumption, and deployment readiness for digital agriculture disease-detection systems.

Acknowledgement

The author expresses sincere appreciation to the academic supervisor for providing guidance throughout the research and manuscript preparation process. Gratitude is also extended to the higher education institution for providing academic facilities, as well as to the contributors of the Banana Leaf Disease Dataset v4, which served as the foundation of this study's experiments. The moral, technical, and scientific support from fellow researchers and the computational laboratory environment greatly contributed to the completion of this work.

References

[1] M. Bagga and et al., "Image based detection and classification of plant diseases: A review," *[Journal name]*, 2024, doi: 10.1002/uar2.20053.
[2] P. Sajitha, A. Andrushia, N. Mostafa, S. Y. Shdefat, and N. Anand, "A review of machine learning and deep learning image-based plant disease classification for industrial farming systems," *Sustain. Oper. Comput.*, 2024, doi: 10.1016/j.susoc.2023.03.001.
[3] A. T. Khan and et al., "Plant disease detection model for edge computing devices," *Front. Plant Sci.*, vol. 14, 2023, doi: 10.3389/fpls.2023.1308528.
[4] R. Archana and S. Jeevaraj, "Deep learning models for digital image processing: a review," *Artif. Intell. Rev.*, 2024, doi: 10.1007/s10462-023-10631-z.
[5] J. Lu, L. Tan, and H. Jiang, "Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification," *Agriculture*, vol. 11, no. 8, p. 707, 2021, doi: 10.3390/agriculture11080707.
[6] A. Joshi, D. Guevara, and M. Earles, "Standardizing and centralizing datasets for efficient training of agricultural deep learning models," *Appl. Sci.*, vol. 15, no. 16, p. 9206, 2023, doi: 10.3390/app15169206.
[7] S. Saponara and A. Elhanashi, "Impact of image resizing on deep learning detectors," in *Advances in Sensors and Sensing Technology*, Springer, 2022, pp. 15–32. doi: 10.1007/978-3-030-95498-7_2.
[8] O. Rukundo, "Effects of image size on deep learning," *Electronics*, vol. 12, no. 4, p. 985, 2023, doi: 10.3390/electronics12040985.
[9] H. Talebi and P. Milanfar, "Learning to resize images for vision tasks," *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 2021. doi: 10.48550/arXiv.2103.09950.
[10] H. J. D. Koresh, A. H. Al-Kubati, and R. Kumar, "Impact of the preprocessing steps in deep learning based image classifications," *NAS Lett.*, vol. 47, no. 6, pp. 645–647, 2024, doi: 10.1007/s40009-023-01372-2.
[11] Y. P. Liang, X. Wu, and Z. Chen, "Implementation of lightweight convolutional neural network for embedded systems: Impact of image resizing interpolation methods," *Sensors*, vol. 24, no. 7, p. 2265, 2024, doi: 10.3390/s24072265.
[12] A. R. Auni and E. Sugiharti, "Optimization of Mango Plant Leaf Disease Classification Using Concatenation Method of MobileNetV2 and DenseNet201 CNN Architectures," *Sci. J. Informatics*, vol. 11, no. 4, pp. 1023–1034, 2025, doi: 10.15294/sji.v11i4.15169.
[13] N. Helmawati and E. Utami, "Analysis for Detecting Banana Leaf Disease Using the CNN Method," *JUITA J. Inform.*, vol. 13, no. 1, pp. 29–36, 2025, doi: 10.30595/juita.v13i1.24514.
[14] R. Arboretti, T. T. Aye, and D. Tondini, "Combining design of experiments and machine learning for process modeling: A systematic approach," *Comput. Ind. Eng.*, vol. 160, p. 107590, 2021, doi: 10.1016/j.cie.2021.107590.
[15] N. Al Kharusi and E. López, "Design of experiments and machine learning integration in tissue engineering: A systematic review," *Heliyon*, vol. 8, no. 11, p. e11624, 2022, doi: 10.1016/j.heliyon.2022.e11624.
[16] H. Xu, J. Zhao, and L. Zhang, "Benchmarking AutoML frameworks for image classification: An experimental evaluation," *Neural Comput. Appl.*, vol. 36, no. 9, pp. 5317–5332, 2024, doi: 10.1007/s00521-023-08793-6.
[17] V. Schiano di Cola *et al.*, "Exploratory data analysis and supervised learning in plant phenotyping studies," *Commun. Appl. Ind. Math.*, vol. 15, no. 2, pp. 69–90, 2024, doi: 10.2478/caim-2024-0014.
[18] N. Jiménez, J. López, P. García, and A. Martín, "Detection of leaf diseases in banana crops using deep learning techniques," *AI*, vol. 6, no. 3, p. 61, 2025, doi: 10.3390/ai6030061.
[19] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre processing and data augmentation techniques," *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 91–99, 2022, doi: 10.1016/j.gltp.2022.04.020.
[20] Y. Gulzar, F. Ahmad, A. Shafique, and S. Malik, "Fruit image classification model based on MobileNetV2 with deep transfer learning technique," *Sustainability*, vol. 15, no. 3, p. 1906, 2023, doi: 10.3390/su15031906.
[21] A. Batool, J. Kim, S.-J. Lee, J.-H. Yang, and Y.-C. Byun, "An Enhanced Lightweight T-Net Architecture Based on CNN for Tomato Plant Leaf Disease Classification," *PeerJ Comput. Sci.*, vol. 10, p. e2495, 2024, doi: 10.7717/peerj-cs.2495.