CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 11 (1) (2026), Pages 842-856
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Analysis of Color Space Transformation on MobileNetV2 Performance for Image Classification

Sherlyn Vironica*, Sugiyarto Surono, and Aris Thobirin

Program Studi Matematika, Fakultas Sains dan Teknologi Terapan, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

Abstract

This study analyzes the effect of color space transformation on the performance of MobileNetV2 for rice leaf disease classification using RGB, HSV, CIELab, and their combinations. The RGB color space is used as the baseline representation, while HSV and CIELab are applied to provide alternative representations of color information. In addition, a dual-stream architecture is employed to combine different color spaces for feature extraction. The results show that the choice of color space influences classification performance. In the single color-space scenario, RGB achieves the highest accuracy of 91.42%, while in the combined scenario, the RGB+CIELab model achieves the best performance with an accuracy of 97.00%. These findings suggest that the use of multiple color spaces can provide richer feature representations and may improve classification performance. Furthermore, the results indicate that optimizing input representation plays an important role in improving model performance, particularly when using lightweight architectures such as MobileNetV2. This study shows that color space transformation can improve classification performance in the rice leaf disease dataset used in this study.

Keywords: CIELab; HSV; MobileNetV2; RGB; Rice Leaf Disease

1. Introduction

Deep learning is an approach in machine learning that models data through stratified layers of neural networks, where each layer produces increasingly abstract representations [1]. This approach builds hierarchical representations to capture complex patterns that are difficult to recognize using conventional methods [2]. Its ability to automatically extract complex visual structures makes deep learning widely applied in fields such as text analysis, signal processing, and various modern prediction systems [3–5]. Among the various deep learning architectures, Convolutional Neural Networks (CNNs) are among the most widely used methods for processing grid-structured data, particularly images due to its ability to learn spatial patterns and visual structures [6], and they are widely applied in image classification [7], plant disease detection [8], medical image analysis [9], and industrial automation [10]. However, conventional CNN architectures such as VGG and ResNet require substantial computational resources, making them less suitable for deployment on devices with limited power and memory [11, 12]. To address these limitations, lightweight CNN architectures such as MobileNet, ShuffleNet, and EfficientNet were developed to improve computational efficiency [13].

*Corresponding author. E-mail: 2200015016@webmail.uad.ac.id

Submitted: February 26, 2026 Reviewed: March 14, 2026 Accepted: March 30, 2026
DOI: https://doi.org/10.18860/cauchy.v11i1.41353

MobileNetV2 is one of the lightweight architectures that uses inverted residual block and linear bottleneck to extract features efficiently with far fewer parameters, while maintaining important information representations in low dimensions and overcoming problems [14, 15]. The efficiency characteristic and generalization of MobileNetV2 make it not only suitable for implementation in real-time scenarios and data-limited devices, but also ideal for various smart agriculture applications [16].

Rice is the main food source for more than half of the world's population. Its production is greatly affected by the health of the plant, while leaf diseases can reduce yields if not detected early. The biggest challenge for farmers, especially in rural areas, is limited access to experts who can make accurate diagnoses. This condition has created a need for an image-based diagnostic system that is fast, consistent, and runs on mobile devices. The application of deep learning in plant disease detection is growing rapidly because disease symptoms can generally be identified through changes in color and texture in leaves [17, 18]. The integration of lightweight architectures, such as MobileNetV2, enables image-based diagnostic systems to be applied more efficiently in the identification of rice plant diseases, thereby increasing the speed and accuracy of decision-making in the field [19, 20].

Previous studies have shown that color representation plays an important role in improving the accuracy of CNN-based classification models. Previous studies have shown that the use of color spaces such as HSV, CIELab, and YCbCr can improve the performance of CNN-based image classification models and machine learning [21]. Another study that specifically evaluated the effectiveness of color features for corn leaf disease found that the combination of HSV and CIELab achieved the highest accuracy of 92.48% in classification [22]. However, these results cannot be generalized to rice plants, which have different visual characteristics and symptom patterns. In addition, no study has specifically assessed the effect of color-space transformation on lightweight deep learning architectures, such as MobileNetV2, for rice leaf disease classification.

Different color spaces represent visual information in different ways. The RGB color space directly stores red, green, and blue intensity values, which are often sensitive to illumination variations. In contrast, the HSV color space separates hue, saturation, and brightness information, allowing color characteristics of plant diseases to be represented more consistently under different lighting conditions. Meanwhile, the CIELab color space separates luminance and chromaticity components, which can highlight subtle color differences that frequently appear in diseased leaf tissues. These properties make HSV and CIELab potentially more effective for emphasizing disease characteristics compared to the standard RGB representation. Based on this, this study was designed in two scenarios. The first scenario uses a single color space, namely RGB, HSV, or CIELab. The second scenario uses a combination of two color spaces, namely RGB and HSV, RGB and CIELab, and HSV and CIELab. This approach aims to determine the extent to which differences and combinations of color representations affect the performance of MobileNetV2 in image classification.

The main contribution of this study lies in systematically evaluating the impact of both single and combined color-space representations on the performance of a lightweight deep learning model for rice leaf disease classification. Unlike previous studies that focus on general plant datasets or heavier architectures, this study specifically investigates the role of color-space transformation in improving classification performance using MobileNetV2, providing insights into efficient model design for resource-limited agricultural applications.

2. Methods

Rice leaf diseases are detected and classified using the MobileNetV2 architecture through three types of color space transformations, namely RGB, HSV, and CIELab. Color space transformations are used to provide image representation variations before the classification process, while MobileNetV2 is used as the main model due to its ability to extract image features efficiently and accurately. The methodological stages applied in this study are shown in Fig. 1

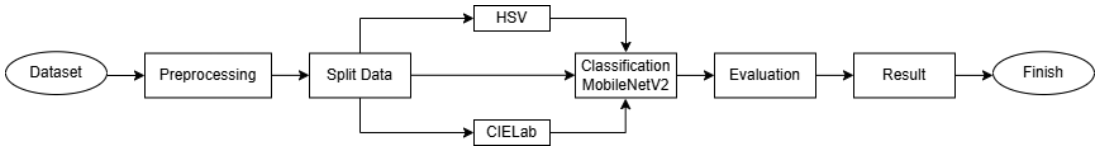


Fig. 1: Flowchart of the methodology

The stages listed in the research method flowchart will be explained in more detail in the next subsection.

2.1. Data Collection

The data used in this study came from the rice leaf disease dataset obtained through the Kaggle platform. This dataset consists of 4,658 images of rice leaves, which are divided into three disease categories, namely Bacterial Blight with 1,584 images, Brown Spot with 1,614 images, and Leaf Smut with 1,460 images. Each image contains variations in lighting conditions, shooting angles, textures, and severity of symptoms, providing sufficient visual diversity to support a representative model training process. All images were originally in RGB format and grouped based on their respective disease classes. This dataset was used as the main source for developing a rice leaf disease classification model based on color space transformation and the MobileNetV2 architecture. An example image from each category is shown in Fig. 2

(a) Bacterial Blight **(b)** Brown Spot **(c)** Leaf Smut

Fig. 2: Example data for each class category

2.2. Preprocessing and Data Splitting

In the preprocessing stage, all images were resized to 224 × 224 pixels to match the input size required by the MobileNetV2 architecture. Each image was normalized to the range [0, 1] to obtain a uniform pixel intensity scale and improve training stability [23]. Data augmentation was applied only to the training set to increase data diversity and reduce overfitting. The augmentation techniques included horizontal flip, vertical flip, shift, scale, rotation, and random brightness contrast. For the validation and test sets, only resizing and tensor conversion were applied without augmentation.

After preprocessing, the dataset was divided into three subsets: training, validation, and testing data. The split was performed using a stratified sampling strategy with a ratio of 80:10:10 to preserve the class distribution in each subset. The data partition was created once before the training process, and the same split was used for all experimental scenarios to ensure a fair comparison between different color space configurations.

To prevent data leakage, images from the same file were not shared between the training, validation, and testing sets. All models were trained and evaluated using the same dataset partition so that performance differences only reflect the effect of color space transformation and model architecture. The distribution of the number of images in each subset is shown in Table 1.

Table 1: Splitting data distribution

Data Splitting	Class	Amount
Train	Bacterialblight	1267
	Brownspot	1291
	Leafsmut	1168
	Total	3726
Validation	Bacterialblight	159
	Brownspot	161
	Leafsmut	146
	Total	466
Test	Bacterialblight	158
	Brownspot	162
	Leafsmut	146
	Total	466

2.3. Color Space Tranformation

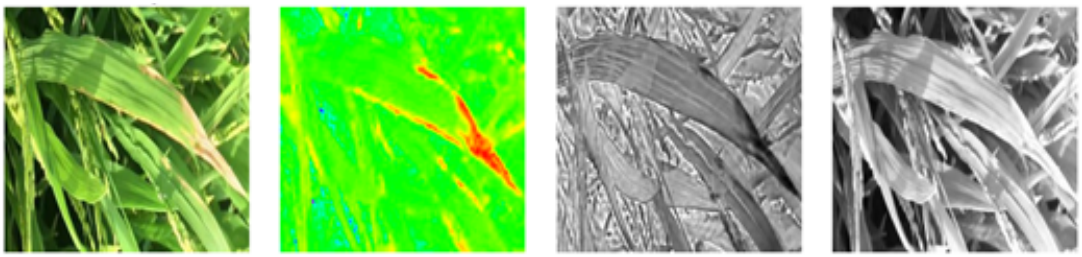
The color space transformation process in this study utilizes three representations of color space, namely RGB, HSV, and CIELab. The RGB color space is used as the original form of the image without additional color transformation, while HSV and CIELab are used to produce a more informative and stable color representation. These two color spaces were chosen because they are able to capture changes in the color of rice leaves more sensitively, especially changes due to disease symptoms. For HSV and CIELab representations, channel-wise normalization is applied to scale the values into a consistent range before being used as input to the model. An explanation of the HSV and CIELab color spaces is presented in the following subsection.

2.3.1. HSV

The HSV (Hue, Saturation, Value) color space represents color information using three components: hue, which indicates the type of color; saturation, which reflects the intensity of the color; and value, which represents brightness [24]. This representation allows color information to be separated from intensity, making it more robust to illumination variations.

In this study, the transformation from RGB to HSV is performed using the standard implementation provided by the OpenCV library. Therefore, the conversion follows the internal formulation defined by the library rather than a manually specified mathematical model. This approach ensures consistency between the implemented method and the reported results, avoiding discrepancies between theoretical formulations and practical implementation

The transformation from RGB to HSV results in different color representations for each component, as shown in Fig. 3.



RGB Image *Hue* *Saturation* *Value*

Fig. 3: The transformation from the RGB image to the HSV image

2.3.2. CIELab

The CIELab color space is a color model that is designed so that the color changes according to human visual perception. This color space is made up of three main components: L* which represents the brightness level, a* which represents the green gradation-red, and b* which depicts a blue-yellow gradation. Image conversion from RGB to CIELab is done through two stages: linear transformation to XYZ space and non-linear transformation to Lab space [25].

The first stage begins by converting the image from RGB to XYZ space using the following matrix (1):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

Then, the values X, Y, and Z are normalized against the white reference values (X_n, Y_n, Z_n), as defined in Eq. (2):

$$X' = \frac{X}{X_n} \tag{2}$$

$$Y' = \frac{Y}{Y_n} \tag{3}$$

$$Z' = \frac{Z}{Z_n} \tag{4}$$

The non-linear transformation to the Lab space is performed using the function defined in Eq. (5):

$$f(t) = \begin{cases} t^{1/3}, & t > 0.008856 \\ 7.787t + \frac{16}{116}, & t \leq 0.008856 \end{cases} \tag{5}$$

The lightness component is computed as defined in Eq. (6):

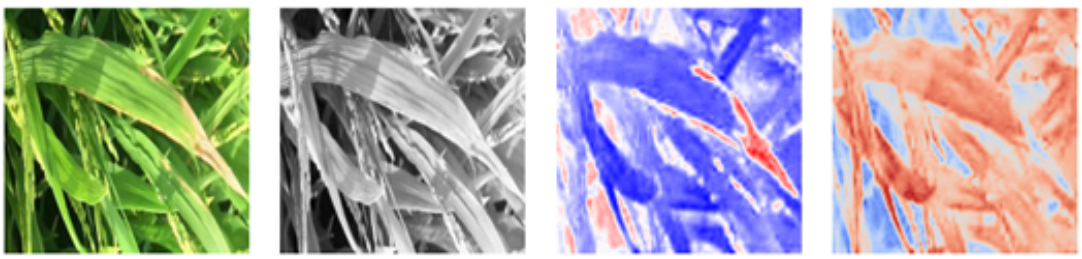
$$L^* = \begin{cases} 116f(Y') - 16, & Y' > 0.008856 \\ 903.3Y', & Y' \leq 0.008856 \end{cases} \tag{6}$$

The chromaticity components are calculated using Eq. (7) and Eq. (8):

$$a^* = 500\left[f(X') - f(Y')\right] \tag{7}$$

$$b^* = 200\left[f(Y') - f(Z')\right] \tag{8}$$

The CIELab representation is able to highlight the color differences more subtly, making it very effective in distinguishing between healthy rice leaves and those that have undergone color changes due to disease. The transformation from RGB to Lab results in different color representations for each component, as shown in Fig. 4.



RGB Image *L* *a** *b**

Fig. 4: The transformation from the RGB image to the CIELab image

2.3.3. Classification Using MobileNetV2

At the classification stage, this study uses the MobileNetV2 architecture as the main model to recognize three classes of rice leaf diseases, namely Bacterial Blight, Brown Spot, and Leaf Smut. This architecture was chosen because it is lightweight, efficient, and capable of producing representative features from image data while requiring relatively low computational resources. In the single color-space scenario, the model receives input images of size 224×224×3, where each image consists of three channels representing the selected color space (RGB, HSV, or CIELab). In the combined color-space scenario, two color representations are processed using a dual-stream architecture. Each color space is provided as a separate input to the model, and each input is processed by an independent MobileNetV2 backbone so that each representation can be learned independently while maintaining the original MobileNetV2 structure.

The features extracted from the two MobileNetV2 backbones are then combined using feature fusion before entering the classification layer. The fusion is performed by concatenating the feature vectors obtained from each backbone, allowing the model to utilize complementary information from different color spaces without modifying the internal convolutional structure. This feature-level fusion is chosen because it provides more stable learning compared to directly stacking channels from different color spaces.

The extracted features are summarized using global average pooling and passed to a fully connected layer with three output neurons using the softmax activation function to produce class predictions. The training process uses the AdamW optimizer with categorical cross-entropy loss to optimize the model parameters. All models in both experimental scenarios are trained using the same configuration with 10 epochs to ensure a fair comparison between single and combined color-space inputs, and all models are trained from scratch without transfer learning. The configuration of the layers used in the feature extraction process is shown in Table 2 as an example of the MobileNetV2 architecture. The same backbone structure is applied to all scenarios, and in the combined color-space scenario, two identical MobileNetV2 backbones are used in parallel before the fusion stage.

Table 2: Example of MobileNetV2 Layer Configuration

Layer	Type	Description
Input Layer	Input	Image input of size 224×224×3. In the combined scenario, two inputs are processed separately using parallel MobileNetV2 backbones.
Conv2D (Layer 1)	Convolution	Initial convolution layer for feature extraction
BatchNorm2D	Normalization	Normalization to stabilize training
ReLU6	Activation	Non-linear activation function
Inverted Residual Block×17	Depthwise + Pointwise Conv	Main feature extraction block of MobileNetV2
BatchNorm2D	Normalization	Feature normalization
Conv2D (1×1)	Pointwise Conv	Channel projection
ReLU6	Activation	Non-linear activation function
Global Average Pooling	Pooling	Reduces spatial dimensions
Dropout	Regularization	Prevents overfitting
Fully Connected (Dense)	Classification	Output layer with 3 classes
Softmax	Activation	Produces class probabilities

2.3.4. Model Evaluation

Model evaluation was carried out to assess the performance of MobileNetV2 in classifying rice leaf diseases in each color space scenario. The evaluation process includes monitoring the accuracy and loss curves during training to observe the model's learning progress and detect potential

overfitting. After the training is complete, the model's performance is tested using unseen test data. The assessment is conducted using four main metrics, namely accuracy, precision, recall, and F1-score, which are defined in Eqs. (9), (10), (11), and (12) [26–28]:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{10}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{11}$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{12}$$

3. Results

This section presents the experimental results obtained from the application of the MobileNetV2 model under different color space scenarios. The evaluation focuses on analyzing the model's performance in terms of training behavior and classification effectiveness across all scenarios. The results are organized into several subsections to provide a clearer and more structured discussion.

3.1. Training Performance Analysis

Training performance is evaluated based on the progression of accuracy and loss values during the training process. In this study, the loss function used is categorical cross-entropy, which is widely applied in multi-class classification problems [29]. The categorical cross-entropy loss is defined in Eq. (13):

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} y_{i,k} \log\left(\hat{y}_{i,k}\right) \tag{13}$$

where N denotes the total number of samples, K represents the number of classes, $y_{i,k}$ is the true label, and $\hat{y}_{i,k}$ is the predicted probability for class k of the i-th sample.

This function measures how well the model predicts the correct class in each training process. The training curve for each scenario is shown in Fig. 5–Fig. 10.

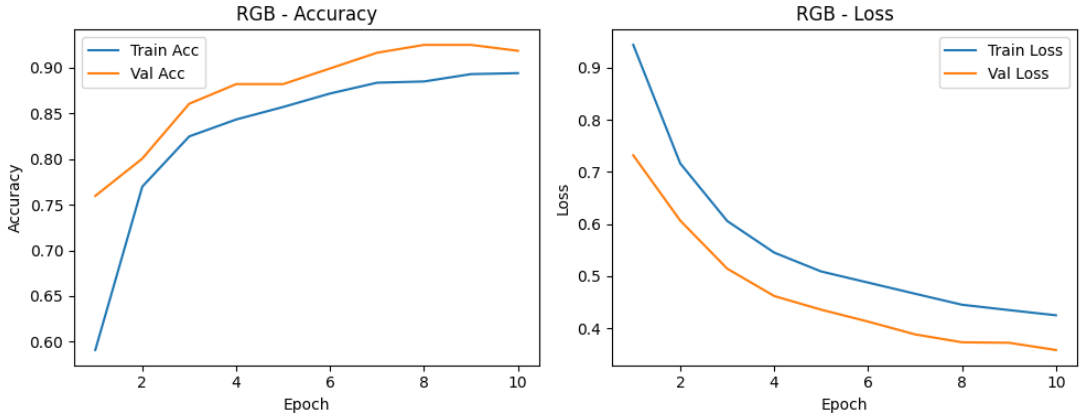


Fig. 5: Model accuracy and loss curves in RGB color space

In the RGB color space shown in Fig. 5, the validation accuracy increases steadily to close to 0.91 at the final epoch. The same training and validation loss curves decrease without significant distances, indicating that the model is not overfitting. The representation of original colors helps the model capture the texture and patterns of the disease more consistently.

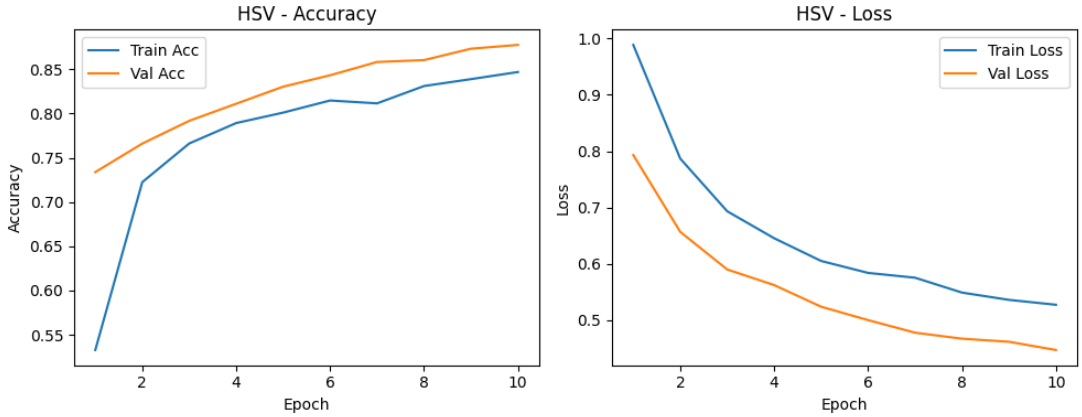


Fig. 6: Model accuracy and loss curves in HSV color space

Fig. 6 shows that the HSV color space has a gradually increasing accuracy curve with a maximum value of about 0.87. Validation loss decreases steadily, but more slowly than in the RGB color space. This happens because the intensity information in HSV is reduced, so it takes longer for the model to recognize differences between classes.

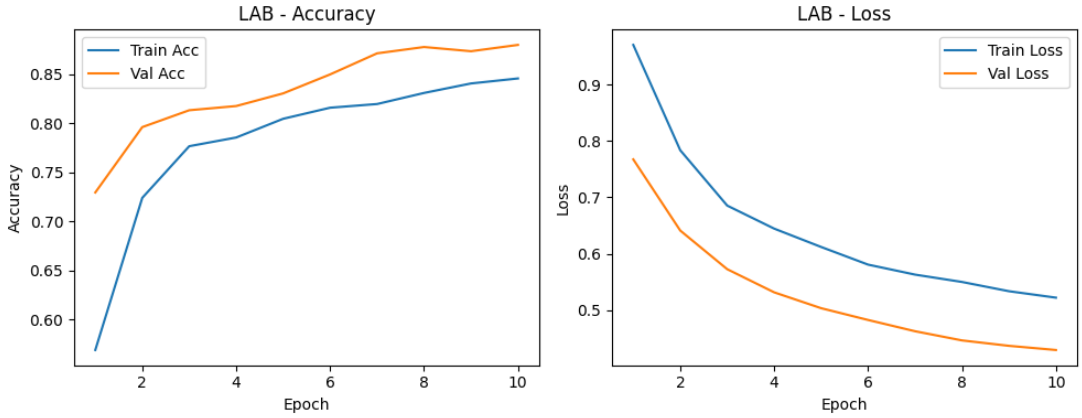


Fig. 7: Model accuracy and loss curves in CIELab color space

In the CIELab color space shown in Fig. 7, the validation accuracy is stable but slightly lower than that of RGB. The validation loss curve continues to decrease, but the difference between train and val is greater than that of the other two color spaces. CIELab transformation tends to smooth out micro-contrast that is important for the identification of disease spots, so that the learning process becomes less optimal.

Furthermore, the combination of two color spaces enriches visual information through feature fusion obtained from two parallel feature extraction paths. Each color space is processed independently by a separate MobileNetV2 backbone, producing feature representations f_1 and f_2. The final feature representation is obtained by concatenating the extracted features as defined in Eq. (14):

$$F = [\, f_1 \parallel f_2 \,] \tag{14}$$

where f_1 and f_2 denote the feature vectors extracted from each color space, and F represents the fused feature vector resulting from the concatenation operation. This approach allows the model to access complementary color representations simultaneously, resulting in a more diverse and informative feature space for the classification process.

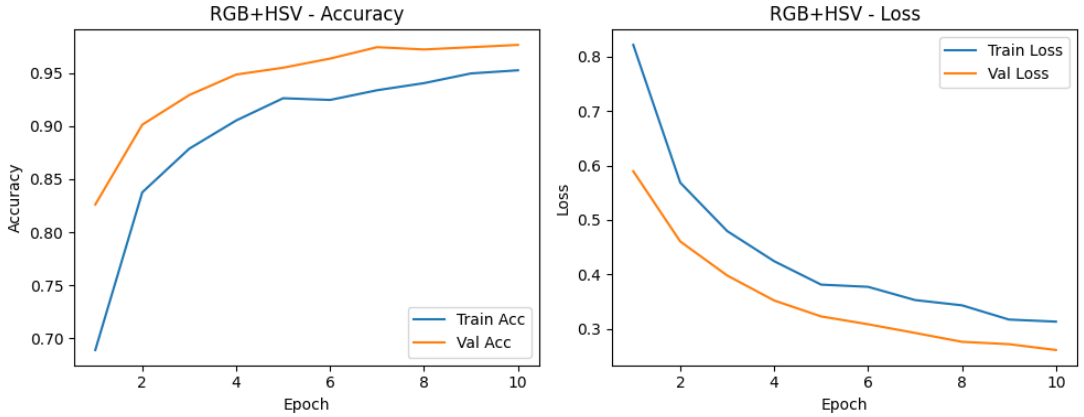


Fig. 8: Model accuracy and loss curves on RGB + HSV combination

Fig. 8 shows that the combination of RGB and HSV converges rapidly, even though it is trained for only 10 epochs. The validation accuracy increased sharply and remained stable throughout. Validation loss decreases consistently, suggesting that combining native colors with hue stability provides richer features and accelerates learning.

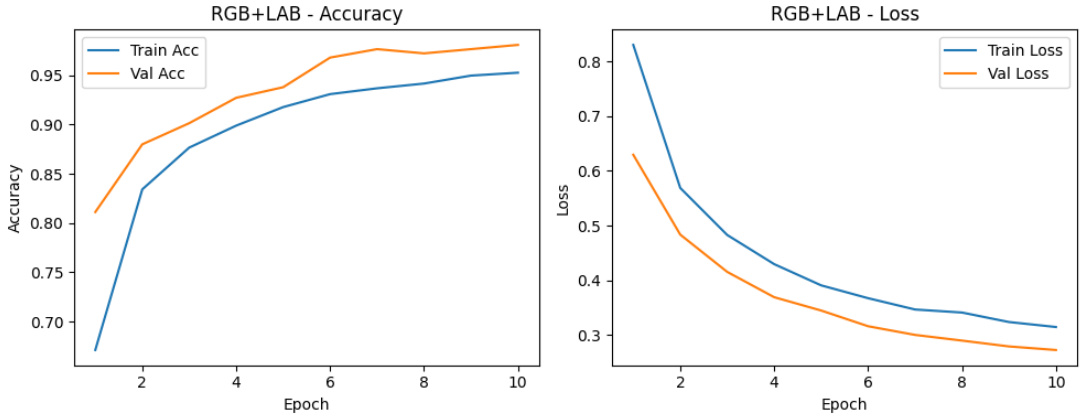


Fig. 9: Model accuracy and loss curves on RGB + CIELab combination

Fig. 9 shows that RGB+CIELab has excellent performance with high and stable validation accuracy. A consistent reduction in losses indicates an effective learning process. The luminance component of CIELab helps highlight differences in leaf texture, while RGB retains natural color, providing the model with a more discriminating feature.

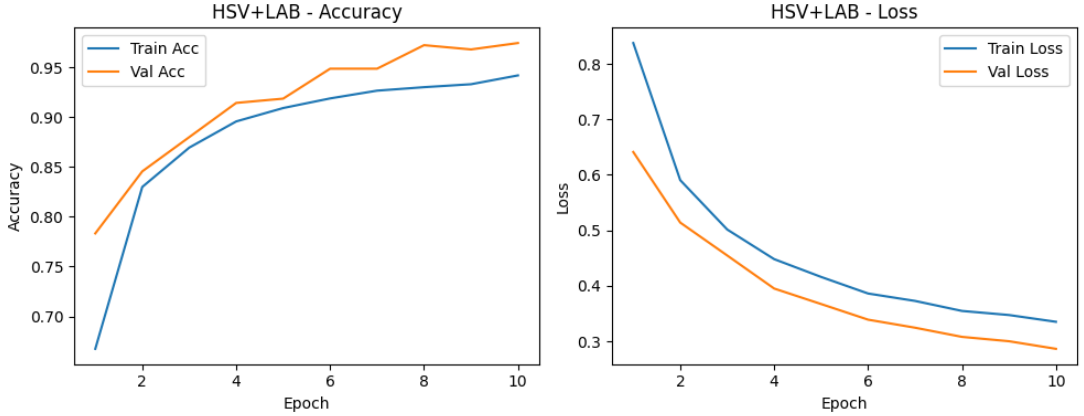


Fig. 10: Model accuracy and loss curves on HSV + CIELab combination

The combination of HSV+CIELab in Fig. 10 shows that the performance improvement is still noticeable but not as large as the other combinations because both color spaces reduce sensitivity to intensity, so the combined information is less diverse.

3.2. Evaluation on Test Data

Model performance evaluation is carried out using accuracy, precision, recall, and F1-Score metrics. The test results are shown in Table 3.

Table 3: Model Evaluation Results on Test Data

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
RGB	91.42	91.69	91.30	91.33
HSV	87.12	87.36	87.22	87.13
CIELab	86.05	86.03	86.03	85.97
RGB + HSV	96.35	96.32	96.27	96.29
RGB + CIELab	97.00	96.96	97.03	96.98
HSV + CIELab	96.14	96.12	96.04	96.07

Based on Table 3, models with RGB inputs perform best in single-color space scenarios. The HSV color space shows the lowest accuracy, while the CIELab provides medium performance. In the scenario of combining two color spaces, there is a consistent increase in performance. The combination of RGB+CIELab provides the best results with accuracy 97.00%, indicating that the incorporation of native color information (RGB) with luminance-chromaticity separation (CIELab) results in a richer representation of features for the classification process. Other combinations, such as RGB+HSV and HSV+CIELab, also show improved performance compared to the use of a single color space.

To further analyze the classification performance, confusion matrices were generated for the RGB model as the best result in the single color-space scenario and the RGB+CIELab model as the best result in the combined color-space scenario. The confusion matrices provide a more detailed view of the classification results for each disease class and help identify possible misclassification patterns.

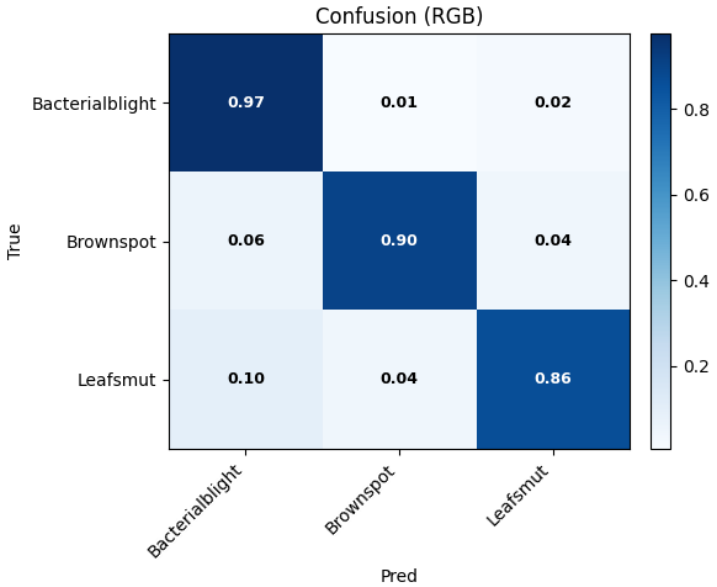


Fig. 11: Confusion matrix of the RGB model on the test dataset

Fig. 11 shows the confusion matrix for the RGB model. Most samples are correctly classified, especially in the Bacterial Blight class, which reaches the highest prediction accuracy. However,

several misclassifications occur between Brown Spot and Leaf Smut. These two classes have similar visual characteristics, such as dark spots and color changes on the leaf surface, which makes them more difficult to distinguish when only the RGB color space is used. This indicates that the RGB representation alone is sometimes insufficient to capture subtle differences in disease patterns.

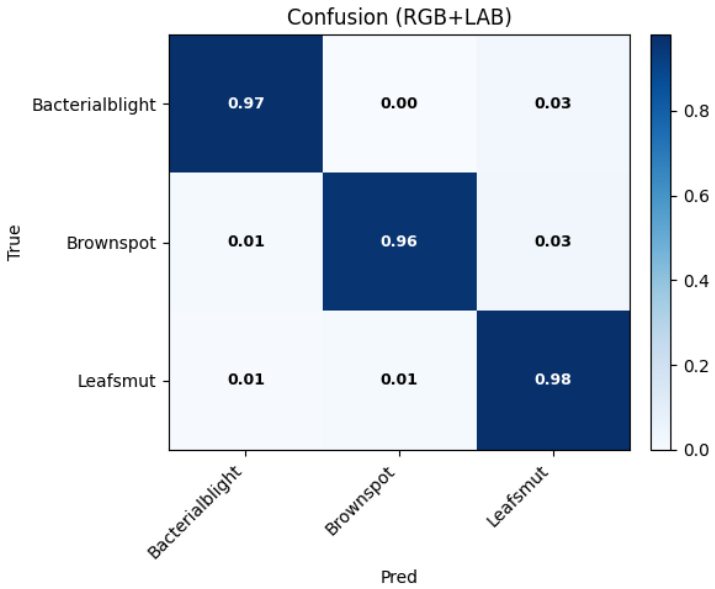


Fig. 12: Confusion matrix of the RGB+CIELAb model on the test dataset

Fig. 12 presents the confusion matrix for the RGB+CIELab model. Compared to the RGB model, the number of misclassifications decreases in all classes. The improvement is particularly noticeable in the Leaf Smut and Brown Spot classes, where the prediction accuracy becomes higher and more balanced. The combination of RGB and CIELab provides richer visual information, because RGB preserves natural color details while CIELab separates luminance and chromaticity components. This combination helps the model recognize subtle color variations and texture differences more effectively, resulting in better classification performance.

3.3. Comparison Between Color Spaces

All scenarios were trained for 10 epochs using the same training configuration to ensure consistent experimental conditions. In the single scenario, RGB performed best at 91.42%, CIELab was in the middle, and HSV had the lowest performance. In the combination scenario, all models showed improved performance under the same training configuration. The RGB+CIELab combination achieved the highest accuracy at 97.00%, followed by RGB+HSV and HSV+CIELab. These results indicate that combining two color spaces produces richer feature representations, allowing the model to learn more effectively. In addition, the baseline RGB-only model has limitations in capturing subtle color variations associated with disease symptoms. Transformations to HSV and CIELab provide additional information not available in RGB, and when combined, this information complements each other, resulting in a more stable learning process and faster convergence. Thus, the results suggest that color space transformation can improve classification performance in the experimental setting used in this study.

4. Discussion

This study analyzes the effect of color space transformation on the performance of MobileNetV2 for rice leaf disease classification using RGB, HSV, CIELab, and their combinations. The results show that the choice of color space affects classification performance. In the single

color-space scenario, RGB achieves the best accuracy of 91.42%, while in the combined scenario, the RGB+CIELab model produces the highest accuracy of 97.00%. These findings suggest that combining color spaces can provide richer feature information and improve classification performance. All scenarios were trained using the same number of epochs and identical training settings to maintain consistent experimental conditions. Therefore, the observed performance differences are likely related to the differences in input representation rather than variations in the training configuration.

These findings are consistent with previous studies showing that alternative color spaces such as HSV and CIELab can better capture color variations related to plant disease symptoms. In addition, combining multiple color representations can improve classification accuracy by providing complementary visual features. The results of this study indicate that optimizing input representation through color-space transformation and feature fusion can enhance model performance, particularly when using lightweight architectures such as MobileNetV2.

Compared to deeper architectures such as VGG, ResNet, and EfficientNet, lightweight models like MobileNetV2 are generally reported to provide a better balance between accuracy and computational efficiency, making them suitable for mobile-based or resource-limited applications. These results suggest that performance improvement may also be achieved by improving input representation, not only by increasing model complexity.

It is important to note that the combined color-space scenario utilizes a dual-stream MobileNetV2 architecture, where each color space is processed by an independent backbone. This design increases the number of parameters and overall model capacity compared to the single-stream configuration. Therefore, the observed performance improvement may not be solely attributed to the use of multiple color spaces, but may also be influenced by the increased representational capacity of the model. This limitation should be considered when interpreting the results.

This study has several limitations. The experiments were conducted using a single dataset and one deep learning architecture. Each scenario was trained under the same configuration, and repeated runs or cross-validation were not performed, which may introduce variability in the results due to the stochastic nature of the training process. Therefore, the reported performance differences should be interpreted with caution. In addition, the dataset was obtained from a public Kaggle source and may not fully represent real-world conditions with varying illumination, backgrounds, and acquisition devices. Future work may include larger and more diverse datasets, repeated experiments, cross-validation, and comparisons with multiple architectures to obtain more reliable and generalizable results.

5. Conclusion

This study indicates that color space transformation can influence the performance of the MobileNetV2 model in rice leaf disease classification. In the single color-space scenario, the RGB-based model achieved the highest accuracy of 91.42%, followed by CIELab and HSV. In the two-color space combination scenario, all models show improved performance while using the same training configuration, with the combination of RGB+CIELab resulting in the best accuracy of 97.00%. These results indicate that the incorporation of complementary color representations can enrich visual information and improve the performance of lightweight models such as MobileNetV2.

This study has several limitations. The experiments were conducted using a single dataset and one deep learning architecture, and the results were obtained from a single training run without cross-validation. These conditions may affect the generalization of the model to different datasets or real-world environments. Future research can use larger datasets, repeated experiments, cross-validation, and comparisons with other architectures to obtain more reliable and comprehensive results.

CRediT Authorship Contribution Statement

Sherlyn Vironica: Conceptualization, Methodology, Software development, Data curation, Formal analysis, Visualization, and Writing – Initial draft. **Sugiyarto Surono**: Supervision, Validation, Methodological Direction, and Writing – Review & Editing. **Aris Thobirin**: Supervision, Validation, and Writing – Review & Editing.

Declaration of Generative AI and AI-assisted technologies

In the preparation of this manuscript, the author used ChatGPT and Grammarly on a limited basis to help improve grammar, sentence structure, and writing clarity. The use of these tools is only for language editing purposes and is not involved in methodological design, experimental implementation, data analysis, or interpretation of research results. All scientific substance in this article is entirely the work of the author.

Declaration of Competing Interest

The author declares that there are no conflicts of interest, whether financial, personal, or professional, that could affect the research, analysis, or conclusions reported in this article.

Funding and Acknowledgments

This research did not receive special funding support from government agencies, commercial institutions, or non-profit organizations. The author expresses sincere gratitude to the supervisor for his valuable guidance, direction, and input during the implementation of the research and preparation of this manuscript. The author also expressed his appreciation to the Faculty of Applied Science and Technology and the Mathematics Laboratory of Ahmad Dahlan University for the support and facilities that have been provided so that this research can be carried out properly.

Data and Code Availability

The data used in this study comes from a dataset that is publicly available through the Kaggle platform and is publicly accessible through the original source. The code used in the model development, training, and evaluation process is not publicly available, but may be obtained by contacting the author of the correspondence upon reasonable request.

References

[1] Laith Alzubaidi, Jinglan Zhang, et al. “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions”. In: *Journal of Big Data* 8.1 (2021). DOI: 10.1186/s40537-021-00444-8.

[2] N. I. Putri and Z. Munawar. “Deep Learning and Big Data Technology for IoT Security”. In: *J. Inform. – Comput.* 7.1 (2020), pp. 48–73.

[3] F. Esmaeili et al. “Utilizing Deep Learning Algorithms for Signal Processing in Electrochemical Biosensors: From Data Augmentation to Detection and Quantification of Chemicals of Interest”. In: *Bioengineering* 10.12 (2023), p. 1348. DOI: 10.3390/bioengineering10121348.

[4] W. She. “A review of deep learning-based text sentiment analysis”. In: (2023). DOI: 10.54254/2755-2721/32/20230204.

[5] X. Kong et al. “Deep learning for time series forecasting: a survey”. In: *Springer* (2025). DOI: 10.1007/s13042-025-02560-w.

[6] A. Younesi et al. "A Comprehensive Survey of Convolutions in Deep Learning : Applications , Challenges , and Future Trends". In: *IEEE Access* (2024). DOI: 10.1109/ACCESS.2024.3376441.

[7] K. Azmi, S. Defit, and S. Sumijan. "Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat". In: *J. Unitek* (2023). DOI: 10.52072/unitek.v16i1.504.

[8] S. Suhendar et al. "Deteksi Penyakit Pada Daun Tanaman Ubi Jalar Menggunakan Metode Convolutional Neural Network". In: *J. Ilm. Inform. Glob.* (2023). DOI: 10.36982/jiig.v14i3.3478.

[9] Media Ali Ibrahim et al. "Deep Learning in Medical Image Analysis Article Review Media". In: *Indonesian Journal of Computer Science* 13.2 (2024), pp. 2293–2311.

[10] N. Hütten et al. "Deep Learning for Automated Visual Inspection in Manufacturing and Maintenance: A Survey of Open-Access Papers". In: *Applied System Innovation* 7.1 (2024), p. 11. DOI: 10.3390/asi7010011.

[11] C. Luo et al. *Comparison and Benchmarking of AI Models and Frameworks on Mobile Devices*. arXiv preprint arXiv:2005.05085. 2020. http://arxiv.org/abs/2005.05085.

[12] A. Younesi et al. "A Comprehensive Survey of Convolutions in Deep Learning: Applications, Challenges, and Future Trends". In: *IEEE Access* 12 (2024).

[13] T. Shahriar. *Comparative Analysis of Lightweight Deep Learning Models for Memory-Constrained Devices*. arXiv preprint arXiv:2505.03303. 2025. http://arxiv.org/abs/2505.03303.

[14] O. V. Putra, M. Z. Mustaqim, and D. Muriatmoko. "Transfer Learning untuk Klasifikasi Penyakit dan Hama Padi Menggunakan MobileNetV2". In: *Techno.Com* 22.3 (2023), pp. 562–575. DOI: 10.33633/tc.v22i3.8516.

[15] P. T. Huong et al. "Enhancing deep convolutional neural network models for orange quality classification using MobileNetV2 and data augmentation techniques". In: *Journal of Algorithms and Computational Technology* 19 (2025). DOI: 10.1177/17483026241309070.

[16] S. F. D. Wardhana and A. Nugroho. "Perbandingan Arsitektur MobileNetV2 dan MobileNetV3 Dalam Klasifikasi Jenis Jeruk". In: *Journal of Computing Science and Business* 16.1 (2025), pp. 25–34. DOI: 10.47927/jikb.v16i1.916.

[17] J. Lu, L. Tan, and H. Jiang. "Review on convolutional neural network (CNN) applied to plant leaf disease classification". In: *Agriculture* 11.8 (2021). DOI: 10.3390/agriculture11080707.

[18] Y. Toda and F. Okura. "How convolutional neural networks diagnose plant disease". In: *Plant Phenomics* 2019 (2019). DOI: 10.34133/2019/9237136.

[19] D. Verma, D. Bordoloi, and V. Tripathi. "Plant Leaf Disease Detection Using Mobilenetv2". In: *Webology* 18.5 (2021), pp. 3241–3246. DOI: 10.29121/WEB/V18I5/60.

[20] H. Chen et al. "Classification and identification of agricultural products based on improved MobileNetV2". In: *Scientific Reports* 14.1 (2024). DOI: 10.1038/s41598-024-53349-w.

[21] H. M. Hassan et al. *A Survey on Different Plant Diseases Detection Using Machine Learning Techniques*. 2022.

[22] P. F. Johari et al. "Corn Leaf Diseases Classification Using CNN with GLCM, HSV, and L*a*b* Features". In: *Jurnal Teknik Informatika* 6.2 (2025), pp. 709–722. DOI: 10.52436/1.jutif.2025.6.2.4345.

[23] M. Ikhsan and M. Rahardi. "Image-Based Classification of Indonesian Traditional Houses Using a Hybrid CNN-SVM Algorithm". In: *Journal of Computer Science* 9.5 (2025), pp. 2303–2309. DOI: 10.30871/jaic.v9i5.10864.

[24] D. Giuliani. "Metaheuristic Algorithms Applied to Color Image Segmentation on HSV Space". In: *Journal of Imaging* 8.1 (2022). DOI: 10.3390/jimaging8010006.

[25] I. A. P. F. Imawati et al. "A Study of Lab Color Space and Its Visualization". In: *ICAMSAC 2023*. Atlantis Press, 2024. DOI: 10.2991/978-94-6463-413-6_3.

[26] M. B. N et al. "Plant Leaf Disease Detection using MobileNetV2". In: (2025). DOI: 10.1051/itmconf/20257901021.

[27] M. S. Aryanta, C. A. Sari, and E. H. Rachmawanto. "A Banana Disease Detection Using MobileNetV2 Model Based on Adam Optimizer". In: 9.4 (2025), pp. 1207–1218. DOI: 10.30871/jaic.v9i4.9870.

[28] G. Morris et al. "Classification of Diabetic Wounds Using Transfer Learning Model: EfficientNetB1 and ResNet50". In: 11.1 (2025), pp. 207–217. DOI: 10.31154/cogito.v11i1.1001.207-217.

[29] A. R. Bushara. "Efficient net-based deep learning model for accurate plant disease classification and diagnosis". In: 14.1 (2025), pp. 1264–1270. DOI: 10.30574/ijsra.2025.14.1.0170.