Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.2, April 2026, pp. 1568~1576
e-ISSN: 2548-6861

Analysis of the Impact of Image Brightness Normalization on the Accuracy of Apple Leaf Disease Classification

Nurlina Ambon[1], Yufis Azhar [2]*
Program Studi Informatika, Universitas Muhammadiyah Malang
nurlinaambon@webmail.umm.ac.id [1], yufis@umm.ac.id [2]

Article Info

Article history:

Received 2026-02-20
Revised 2026-03-12
Accepted 2026-04-10

Keyword:

Apple Leaf Disease,
Brightness Normalization,
Deep Learning,
Image Classification,
MobileNetV2.

ABSTRACT

Apple crops represent an economically important horticultural sector, yet it faces vulnerabilities from various leaf diseases, including Apple Scab, Cedar Rust, and Black Rot. Translate . Variations in illumination during image acquisition can lead to inconsistencies in color and contrast, adversely affecting the performance of Convolutional Neural Network (CNN)–based classification systems. This study aims to explore how brightness normalization techniques can enhance the accuracy of apple leaf disease classification. The dataset utilized in this research consists of 480 images of apple leaves, categorized into three disease classifications. To facilitate an objective evaluation of the model, the dataset was partitioned into 70% training set, 15% validation set, and 15% test set. Eight different brightness normalization methods were applied during preprocessing, including no normalization, min-max scaling, z-score normalization, gamma correction, histogram equalization, CLAHE, logarithmic transformation, and square root transformation. For this study, the MobileNetV2 architecture was selected as the primary model due to its efficiency in parameters and strong performance in image recognition tasks. Experimental results reveal that the Min-Max normalization technique yielded the highest accuracy at 95.83%, followed by Histogram Equalization at 94.44% and Gamma Correction at 83.33%. In contrast, the baseline model without any normalization only achieved an accuracy of 38.89%. These findings underscore the significant role that brightness normalization plays in enhancing the resilience of CNN models against variations in illumination, thereby improving the automated classification of apple leaf diseases.

I. INTRODUCTION

Apple cultivation is a vital part of the horticultural sector and is extensively grown in many region. However, the productivity of apple trees is frequently compromised by leaf diseases such as Apple Scab, Cedar Rust, and Black Rot [1]. These diseases not only diminish crop yields but also impact fruit quality, leading to substantial economic losses for farmers and the broader agricultural sector. Consequently, the early detection of apple leaf diseases is crucial for facilitating prompt, accurate, and effective control measures

In recent years, deep learning technology, particularly Convolutional Neural Networks (CNNs), has become increasingly prominent in the classification of plant diseases based on digital images [2] Convolutional Neural Networks effectively learn to identify distinctive visual features from images, minimizing the reliance on manual feature engineering, resulting in enhanced classification performance. Research studies, including MCFFA-Net [3], LCAMNet [4], and Lightweight-CNN [5] , demonstrate that modifications to CNN architecture through the integration of feature fusion and attention mechanisms can achieve apple leaf disease classification accuracy exceeding 95%. Recent studies have also investigated the integration of advance CNN architectures to enhance the accuracy and interpretability of plant disease detection, these findings demonstrate that deep learning models can achieve impressive performance when paired with suitable image preprocessing techniques[6].

One of the primary challenges in implementing image classification systems in the field is the variation in lighting during image capture [7]. Images captured at different times, under varying weather conditions, or using different devices may exhibit inconsistent brightness levels. This variation in lighting directly impacts the color distribution, contrast, and texture of leaves, ultimately diminishing the classification accuracy of CNN models [8]. For instance, images that are excessively dark may obscure disease spots, while those that are overly bright can mask important color differences on the leaf surface. Consequently, the model is prone to misclassification, mistakenly identifying disease types based on lighting discrepancies rather than actual disease patterns.

Numerous studies have sought to tackle this issue by employing image brightness normalization techniques during the pre-processing stage. The application of Midpoint Normalization (MPN) for balancing light intensity has demonstrated a notable enhancement in classification accuracy, achieving up to 93% [9]. Additionally, other methods such as Contrast Limited Adaptive Histogram Equalization (CLAHE) [10] and Gamma Correction have shown effectiveness in normalizing pixel intensity distribution, thereby increasing the model's resilience to variations in lighting conditions. For example, algorithms based on CNNs for image enhancement have been introduced to improve brightness distribution in low-light images[11]. Additionally, Retinex-based CNN methods have proven effective in enhancing image illumination while preserving structural details[12]. Furthermore, preprocessing techniques like CLAHE have been successfully utilized in CNN-based classification tasks to enhance image contrast and improve model accuracy, particularly in medical imaging applications [13].

In the context of leaf disease image classification, variations in brightness levels significantly influence the perception of color and texture features [9], [14]. When images are captured in low light conditions, two distinct diseases may appear quite similar; conversely, images depicting the same disease can look markedly different under high light conditions. This phenomenon may lead CNN models to prioritize lighting artifacts rather than the actual morphological characteristics of the diseases. Consequently, employing brightness normalization techniques becomes essential, enabling deep learning models to focus on features pertinent to disease symptoms rather than variations in light intensity. The integration of Contrast Limited Adaptive Histogram Equalization (CLAHE) with deep residual CNN models has demonstrated enhanced classification accuracy in crop image analysis [15]. Additionally, the application of hybrid brightness preserving bi-histogram equalization (BBHE) in conjunction with CNN has proven effective in improving plant disease detection performance [16]. As a result, the use of brightness normalization techniques is crucial, allowing deep learning models to concentrate on features relevant to disease symptoms, rather than being affected by variations in light intensity.

This study aims to analyze the impact of eight image brightness normalization methods on the accuracy of apple leaf disease classification using a transfer learning-based MobileNetV2 model. The eight methods examined contains: no normalization (None), Min–Max Scaling, Z-score Normalization, Gamma Correction, Histogram Equalization (hist_eq), Contrast Limited Adaptive Histogram Equalization (CLAHE), Log Transformation, and Square Root Transformation (SQRT). This research is intended to provide scientific insights into the role of image normalization in enhancing the stability of CNN model performance, as well as offering practical contributions to the development of plant disease detection systems that can adapt to varying lighting conditions in the field.

The primary contributions of this research are: (i) a comparative analysis of eight brightness normalization techniques, (ii) an examination of their impact on the performance of transfer learning-based MobileNetV2, and (iii) recommendations for optimal methods under varying lighting conditions.

II. METHOD

This study employs a quantitative experimental approach aimed at analyzing the impact of eight brightness normalization techniques on the accuracy of apple leaf disease classification utilizing the MobileNetV2 model. The experiment involved comparing the model's performance across each normalization method, which included: no normalization (None), Min-Max Scaling, Z-score Normalization, Gamma Correction, Histogram Equalization (hist_eq), Contrast Limited Adaptive Histogram Equalization (CLAHE), Log Transformation, and Square Root Transformation (SQRT). Each method was evaluated using an apple leaf image dataset that was divided into training, validation, and testing sets with a ratio of 70% training, 15% validation, and 15% testing.

A. Dataset Collection (Plant Pathology 2020)

The image dataset utilized in this study was sourced from the Plant Pathology 2020 challenge available on Kaggle, It comprises labeled images of apple leaves demonstrating symptoms of various common diseases [17]. This research specifically targets the classification of three disease categories: Apple Black Rot, Apple Cedar Rust, and Apple Scab. In total, it contains 480 images of apple leaves, distributed as follows: Apple Black Rot: 170 images, Apple Cedar Rust: 160 images, and Apple Scab: 150 images. This relatively balanced class distribution enhances the model training process by minimizing the risk of bias toward any particular class. The dataset served as the foundation for a series of processes involving preprocessing, model training, and the evaluation of the performance of the apple leaf disease classification system.

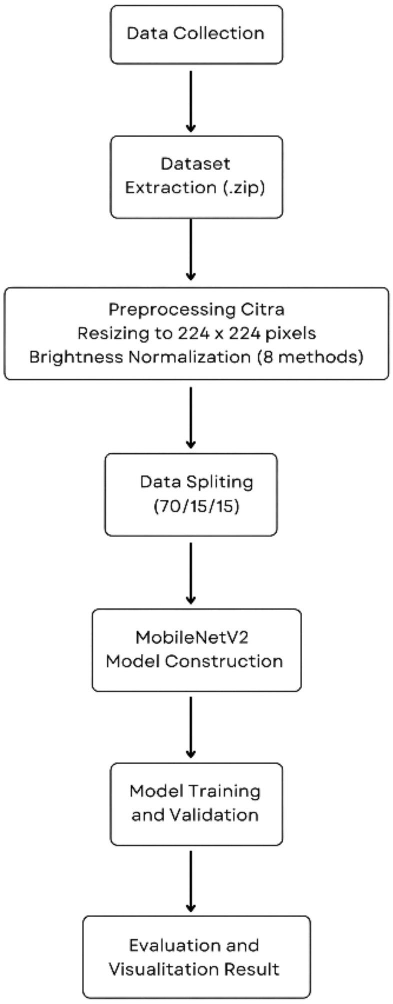


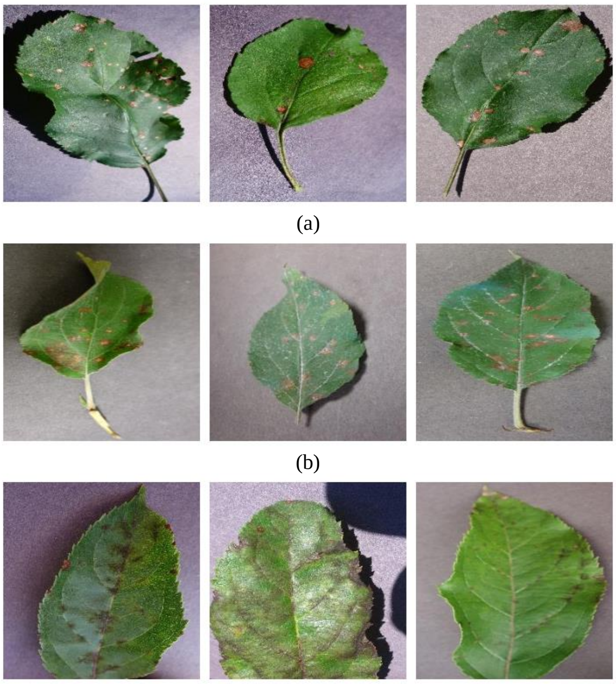
Figure 1. Flowchart of Research on Image Brightness Normalization for Apple Leaf Disease Classification

The images in the Plant Pathology 2020 dataset were captured under authentic field conditions using a variety of cameras and mobile devices. As a result, the dataset features variations in illumination, background, leaf orientation, and image quality. These differences make the dataset particularly well-suited for assessing the efficacy of brightness normalization techniques in enhancing classification robustness.

B. *Dataset Extraction (.zip)*

The retrieved dataset files are generally saved in .zip compressed format. At this stage, the extraction process is conducted to unpack and reorganize all image files into the appropriate directory structure. This procedure is essential to ensure that each image is easily accessible by the system and prepared for use in the subsequent image preprocessing phase.

Each class exhibits distinct visual characteristics, including variations in spot patterns, leaf color, and disease severity. Additionally, differences in lighting intensity during image capture play a significant role in influencing the complexity of the deep learning-based image classification process.


(a)

(b)

(c)

Figure 2. Examples of apple leaf disease classes: (a) Apple Black Rot, (b) Apple Cedar Rust, and (c) Apple Scab.

The representations of each class of apple leaf disease are displayed in Figure 2, which highlights the differences in leaf surface texture and color caused by various pathogen infections.

C. *Image Preprocessing*

The image preprocessing stage is essential for processing images to maintain consistency and adequate quality before they are used for model training. (a) Image Resizing (224 × 224): All apple leaf images are scaled to an image size of 224 × 224 pixels, which aligns with the standard input size required by the MobileNetV2 architecture. This resizing process guarantees uniform input dimensions across the dataset, allowing for optimal and consistent feature extraction. (b) Brightness Normalization: To tackle variations in illumination conditions present in the apple leaf dataset, several brightness normalization techniques are applied and evaluated individually. These techniques include using the original images without normalization as a baseline, intensity rescaling via Min–Max Scaling, standardization through Z-score normalization, and contrast enhancement methods such as Gamma Correction and Histogram Equalization (hist_eq). Additionally, local contrast enhancement is addressed through Contrast Limited Adaptive Histogram Equalization (CLAHE), along with Log Transformation, and Square Root Transformation (SQRT). The purpose of these normalization techniques is to mitigate illumination variability and enhance the visibility of disease-related features, such as color changes, texture, and lesion patterns. Each normalization technique's performance is assessed to determine its impact on classification accuracy.

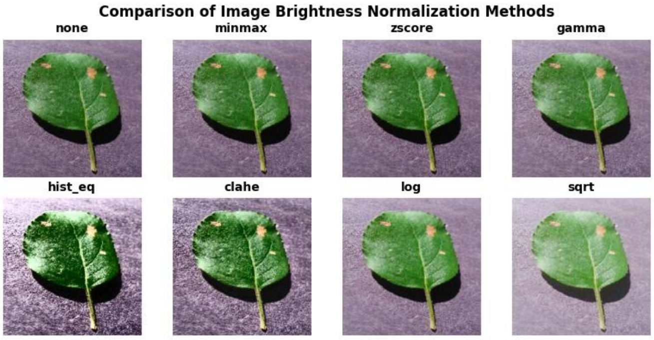


Figure 3. Visual comparison of brightness normalization methods applied to an apple leaf image.

Figure 3 presents a visual comparison of the brightness normalization techniques applied to a sample apple leaf image. The original image without normalization shows uneven illumination, where certain regions appear darker due to lighting conditions. After applying Min–Max Scaling and Z-score normalization, the overall brightness distribution becomes more balanced while preserving the natural color characteristics of the leaf. Gamma Correction slightly enhances contrast, making disease spots more noticeable. In contrast, Histogram Equalization and CLAHE significantly increase image contrast and highlight texture details on the leaf surface. However, these methods may also introduce stronger brightness variations in some regions. Meanwhile, Log Transformation and Square Root Transformation produce softer brightness adjustments that tend to reduce contrast, making some disease symptoms less visually distinct. These visual differences demonstrate how each normalization technique modifies the pixel intensity distribution, which may influence the ability of the deep learning model to extract relevant disease features.

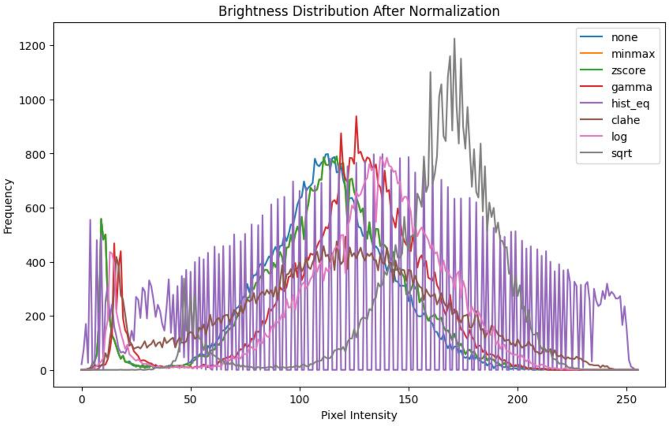


Figure 4. Histogram of pixel intensity distribution for different brightness normalization methods.

Figure 4 shows the distribution of pixel intensities following the application of various brightness normalization techniques. It is evident that each method yields a distinct intensity distribution pattern. Techniques such as Min–Max Scaling, Z-score Normalization, and Gamma Correction generally preserve the intensity distribution around the central range, albeit with slight variations in the shape of the curve. In contrast, Histogram Equalization and CLAHE extend intensity values over a broader spectrum, enhancing image contrast as reflected in their wider distributions. Meanwhile, Log and Square Root transformations gradually shift the intensity distribution, with the Square Root method typically resulting in higher intensity values compared to the other approaches. These variations in brightness distribution suggest that each normalization technique modifies image characteristics in unique ways, which may ultimately impact the deep learning model's capability to extract crucial features for apple leaf disease classification. (c) Data Augmentation Consideration: Given that the dataset comprises only 480 images, data augmentation is typically employed to enhance data diversity and mitigate overfitting. However, in this study, data augmentation was deliberately excluded.

This choice was made to ensure that the experimental results accurately reflect the effects of brightness normalization techniques on classification performance. Implementing augmentation transformations like rotation, flipping, zooming, or shifting could introduce additional variations that might complicate the evaluation of brightness normalization methods. Thus, the original dataset was utilized directly to preserve the integrity of the experimental comparison.

D. Dataset Division (70/15/15)

The dataset was divided into three subsets consisting of 70% training set, 15% validation set, and 15% test set. The training set was employed to train the MobileNetV2 model, while the validation set served to monitor model performance during training and to prevent overfitting by tuning hyperparameters. The testing set was reserved exclusively for final performance evaluation to assess the generalization capability of the trained model on previously unseen data. This data partitioning strategy ensures that the model is trained with sufficient data while maintaining an objective and unbiased evaluation of its classification performance.

To ensure consistent class distribution across all subsets, the dataset was split using a stratified sampling strategy. This method preserves the proportion of each disease class (Apple Black Rot, Apple Cedar Rust, and Apple Scab) in the training, validation, and testing sets. By implementing stratified division, we can minimize the risk of class imbalance during model training and evaluation, leading to a more reliable and representative performance assessment of the classification model.

E. MobileNetV2 Model

The model was developed based on the MobileNetV2 architecture, employing a transfer learning approach. MobileNetV2 was selected due to its relatively small number of parameters while still demonstrating effectiveness in extracting visual features in depth [18]. In this research, the MobileNetV2 model was initialized with pretrained weights obtained from the ImageNet dataset. Rather than being utilized solely as a fixed feature extractor, the pretrained model underwent a retraining process involving fine-tuning of all convolutional layers to tailor the model for the specific task of classifying apple leaf diseases.

To enhance the classification process, additional layers were incorporated on top of the base model. These layers included a global average pooling layer, a dropout layer, and a fully connected layer equipped with a softmax activation function. The model was trained independently for each brightness normalization method, employing the same training, validation, and testing data split across all scenarios. During the training phase, a validation process was implemented to monitor the model’s performance, ensuring stability and consistency while also preventing overfitting in the different normalization contexts.

MobileNetV2 was chosen for its lightweight architecture and impressive computational efficiency, which render it ideal for use on mobile and edge devices in agricultural settings. The main focus of this research is not to compare various CNN architectures but to examine how different brightness normalization methods affect classification performance within a consistent model framework.

F. Model Training and Validation

Model training was conducted using data generated from eight brightness normalization methods. During this phase, the MobileNetV2 model, which has been modified at its final layer, was run through several epochs to learn the visual patterns associated with each class of apple leaf disease. Validation data was utilized concurrently at each epoch to monitor the model's performance throughout the training process. This use of validation data is crucial for detecting signs of overfitting, a situation where the model simply memorizes the training data and fails to generalize to new, unseen data.

Throughout the training process, the loss and accuracy metrics for both the training and validation subsets were carefully monitored to assess the stability of the model. The initial weights, pre-trained on the ImageNet dataset, were leveraged as feature extractors to enhance the speed and accuracy of the training. The model architecture was adjusted by incorporating Global Average Pooling, a Dropout rate of 0.4 to mitigate overfitting, and a Dense (Softmax) output layer for multi-class classification. The model was trained for a total of 20 epochs utilizing the Adam optimizer, with a learning rate set to 0.0001 and a batch size of 32. To address the multi-class classification challenge, the categorical cross-entropy loss function was employed. Additionally, several callback mechanisms were incorporated to enhance training performance. Early stopping was implemented with a patience of 5 epochs to mitigate the risk of overfitting, while ReduceLROnPlateau was employed to decrease the learning rate when the validation loss ceased to improve.

Upon completing the training process, the final performance of each normalization method was documented. The model exhibiting the best performance was subsequently evaluated using the testing dataset, allowing for an assessment of its generalization capabilities on previously unseen data.

G. Evaluation and Visualization of Results

During the evaluation phase, the performance of the model is assessed using accuracy metrics based on test data from each applied brightness normalization method. The objective of this evaluation is to determine how effectively normalization enhances the MobileNetV2 model's capacity to identify disease patterns on apple leaves. Accuracy has been selected as the primary metric, aligning with the research goal of evaluating the model's proficiency in classifying images into three distinct disease categories. In addition to calculating accuracy, the evaluation process incorporates an analysis of classification errors through the use of a confusion matrix. This matrix illustrates the distribution of model predictions for each disease class, allowing for the identification of classes that are frequently misclassified. Understanding this information is crucial for recognizing the limitations of the model, particularly with regard to classes that exhibit similar visual characteristics.

To clarify the evaluation results, this study employs two forms of visualization. First, a comparative graph illustrates the model's performance across the eight normalization methods tested. This graph effectively highlights the methods that yield the most significant improvements in accuracy, enabling readers to easily observe performance trends among the various techniques. Second, a confusion matrix is presented for the best-performing methods. This visualization depicts the relationship between predictions and actual labels for each class, allowing for a more thorough analysis of error patterns.

III. RESULT AND DISCUSSION

A. Model Testing Result

The predictive performance of the MobileNetV2 architecture was assessed using classification accuracy as the main metric on the test dataset to determine its effectiveness in identifying apple leaf diseases. This evaluation incorporated eight different methods for image brightness normalization: None, Min–Max Scaling, Z-score Normalization, Gamma Correction, Histogram Equalization (hist_eq), Contrast Limited Adaptive Histogram Equalization (CLAHE), Log Transformation, and Square Root Transformation (SQRT).

The comparative performance of the MobileNetV2 model utilizing various brightness normalization techniques is summarized in Table I. This study employs four evaluation metrics: accuracy, precision, recall, and F1-score. Together, these metrics offer a thorough assessment of the model's classification performance on the testing dataset.

TABLE I
PERFORMANCE COMPARISON OF MOBILENETV2 USING DIFFERENT BRIGHTNESS NORMALIZATION METHODS

Normalization Method	**Result**			
	Accuracy (%)	**Precision**	**Recall**	**F1-score**
None	38.89	78.43	38.89	27.43
Min-Max Scaling	95.83	96.25	95.83	95.86
Z-Score Normalization	80.56	86.95	80.56	80.16
Gamma Correction	83.33	86.52	83.33	83.08
Histogram Equalization	94.44	94.72	94.44	94.47
CLAHE	69.44	84.38	69.44	68.23
Log Transformation	58.33	80.45	58.33	60.29
Square Root Transformation	42.06	80.71	43.06	37.77

Accuracy measures the overall classification performance of the trained model when evaluated on the testing dataset, which comprises previously unseen data and highlights the model's generalization capability. Precision quantifies the proportion of correctly predicted positive instances among all predicted positives, thereby indicating the reliability of the model's positive predictions. Recall reflects the proportion of accurately identified positive instances among all actual positives, showcasing the model's ability to detect relevant samples in each class. The F1-score, serving as the harmonic mean of precision and recall, offers a balanced evaluation of the model's classification performance. Collectively, these metrics provide a comprehensive assessment of the effectiveness of various brightness normalization techniques when applied to the MobileNetV2 model.

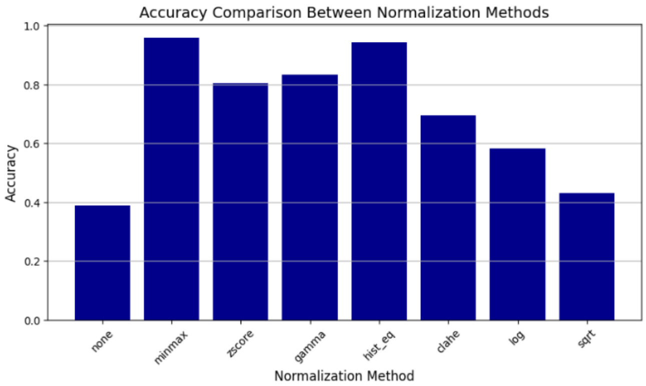


Figure 6. Accuracy comparison of classification performance across different brightness normalization methods.

Figure 6 presents a comparison of classification accuracy achieved through various brightness normalization methods. The findings reveal that Min–Max Scaling yields the highest accuracy at 95.83%, followed by Histogram Equalization at 94.44%, and Gamma Correction at 83.33%. In contrast, the model without any normalization demonstrates the lowest accuracy, reaching only 38.89%. These results underscore the significant role of brightness normalization in enhancing the MobileNetV2 model's robustness when addressing illumination variations in apple leaf images.

B. *Analysis of Result*

Based on the test results presented in Table I, the Min–Max Scaling normalization method exhibited the highest performance, achieving an accuracy of 95.83%. It also demonstrated impressive precision (96.25%), recall (95.83%), and F1-score (95.86%). This suggests that Min–Max Scaling effectively normalizes pixel intensity values into a consistent range, enabling the MobileNetV2 model to learn discriminative visual features with greater efficacy. The Histogram Equalization method also delivered strong results, attaining an accuracy of 94.44%. It was followed by Gamma Correction, which achieved 83.33% accuracy, and Z-Score Normalization, with an accuracy of 80.56%. These methods enhance image contrast and enhance the distribution of pixel intensities, thereby aiding the in capturing crucial visual characteristics such as color variations, spot patterns, and the texture of infected leaf regions. CLAHE demonstrated moderate performance, achieving an accuracy of 69.44%. In comparison, Log Transformation and Square Root Transformation yielded lower accuracies of 58.33% and 43.06%, respectively. While these techniques adjust brightness distribution, they may inadvertently alter important visual details, potentially impairing the model's effectiveness in identifying disease patterns.

In contrast, the model trained without any brightness normalization (None) exhibited the lowest performance, with an accuracy of just 38.89%. This outcome underscores the significance of lighting variation as a critical factor influencing classification performance. Uneven illumination can obscure disease patterns and diminish the model's ability to reliably recognize features of infected leaves. The marked difference in accuracy between the model without normalization (38.89%) and the highest-performing method, Min–Max Scaling (95.83%), underscores the importance of brightness normalization techniques in enhancing classification performance. This significant disparity likely arises from the apple leaf image dataset's substantial variations in illumination conditions, including differences in light intensity, shadows, and environmental factors during image capture. In the absence of a normalization process, these lighting discrepancies lead to inconsistencies in pixel value distribution, which complicates the model's ability to extract relevant visual features such as spot patterns, color variations, and disease textures on the leaves.

The application of Min–Max Scaling facilitates the normalization of pixel value ranges into a consistent scale, effectively diminishing the impact of lighting variations and enhancing the uniformity of feature representation in images. This consistency allows the MobileNetV2 model to extract features more reliably during the learning process, as the input data distribution becomes more standardized. Consequently, the model can concentrate more effectively on learning the genuine characteristics of leaf diseases rather than being affected by irrelevant lighting differences. Min–Max Scaling normalizes pixel intensity values to a consistent range between 0 and 1, which stabilizes the numerical distribution of input data fed into the CNN model. This normalization retains the relative differences between pixel intensities while mitigating the impact of extreme brightness variations caused by differing illumination conditions. Given that convolutional neural networks depend on stable numerical input distributions during convolution and feature extraction, Min–Max Scaling enhances the model's capacity to learn meaningful visual representations. In contrast to contrast enhancement techniques such as Histogram Equalization or CLAHE, which can significantly alter image contrast and possibly introduce artificial patterns, Min–Max Scaling preserves the natural visual structure of the leaf while effectively reducing illumination variation. This characteristic allows the MobileNetV2 model to concentrate more efficiently on disease-related visual features, including lesion texture, color variation, and spot distribution.

It is essential to evaluate whether the observed improvements in model performance are statistically reliable. In this study, the Min–Max Scaling method achieved the highest classification accuracy compared to the other normalization techniques. To enhance the reliability of these findings, conducting a statistical significance analysis can prove beneficial when comparing performance differences among normalization methods. Commonly employed statistical techniques, such as the paired t-test or the Wilcoxon signed-rank test, can be utilized in machine learning studies to ascertain whether the observed improvements in classification accuracy are statistically significant or merely a result of chance. While this study primarily centers on a comparative evaluation of normalization methods based on performance metrics, integrating statistical significance testing could offer additional validation of the performance differences observed. Consequently, future research should consider incorporating statistical testing procedures to further confirm the robustness and reliability of the enhancements achieved by specific normalization techniques.

It is crucial to consider the potential risks associated with over-enhancement when employing certain contrast enhancement techniques. Methods such as Histogram Equalization and CLAHE can significantly boost image contrast, thereby aiding in the identification of disease spots and textures. However, excessive contrast enhancement may also intensify image noise or background artifacts, particularly in areas with subtle intensity variations. This effect can generate irrelevant visual patterns that may disrupt the feature extraction process of deep learning models. The relatively lower performance of CLAHE compared to Histogram Equalization and Min-Max Scaling may be partially due to this issue.

In addition to classification accuracy, the computational cost of preprocessing techniques is a crucial factor for practical implementation. Methods such as Min-Max Scaling and Z-score normalization involve relatively straightforward mathematical operations that require only basic arithmetic transformations of pixel values, resulting in minimal computational overhead. In contrast, techniques like Histogram Equalization and Contrast Limited Adaptive Histogram Equalization (CLAHE) necessitate histogram computations and local contrast adjustments, which can increase preprocessing time when applied to large image datasets. Therefore, from the perspective of computational efficiency, Min-Max Scaling not only achieves high classification accuracy but also maintains a low level of preprocessing complexity, making it a practical normalization method for real-world CNN-based systems in plant disease classification.

To enhance the performance analysis, an evaluation was conducted using a Confusion Matrix for the method that demonstrated the best performance, specifically Min-Max Scaling. This matrix illustrates the distribution of classifications across the three classes of apple leaf diseases that were tested, as depicted in Figure 5.

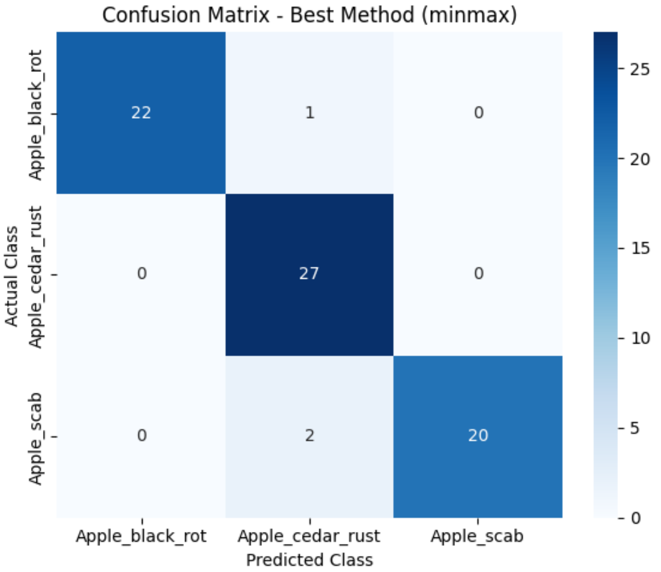


Figure 5. Confusion Matrix – Min-Max Scaling Method.

The results depicted in Figure 5 showcase the classification performance of the MobileNetV2 model utilizing the Min–Max Scaling normalization method. The confusion matrix reveals that the model successfully recognized the majority of samples across all classes. In the Apple Black Rot class, the model achieved impressive classification performance, accurately identifying 22 images, with only one sample misclassified as Apple Cedar Rust. This suggests that the visual characteristics of Apple Black Rot, such as dark lesions and irregular spot patterns, are sufficiently distinctive for the model to recognize.

For the Apple Cedar Rust class, the model exhibited outstanding performance by correctly classifying all 27 samples without any misclassifications. This implies that the visual features of Apple Cedar Rust, including the characteristic orange or rust-colored spots, present clear patterns that the MobileNetV2 model can effectively capture. In the Apple Scab class, the model accurately classified 20 samples, while two images were misclassified as Apple Cedar Rust. This misclassification may stem from the similarities in visual characteristics between the two diseases, particularly regarding spot distribution and color intensity on the leaf surface.

Overall, the confusion matrix indicates that the MobileNetV2 model, when combined with the Min–Max Scaling normalization method, demonstrates robust classification capabilities across the three apple leaf disease classes. Although minor misclassifications occur between Apple Scab and Apple Cedar Rust, the model maintains stable performance overall, suggesting that Min–Max Scaling effectively enhances the visibility and consistency of disease-related features in the images.

While the experimental results indicate that Min-Max Scaling yields the best performance on the Plant Pathology 2020 dataset, the broader applicability of this finding to other plant disease datasets is still uncertain. Variations in lighting conditions, image resolution, background complexity, and disease characteristics across different datasets may influence outcomes. Consequently, further experiments utilizing additional plant disease datasets are essential to determine whether the effectiveness of Min-Max Scaling is consistent across diverse image sources and environmental contexts.

IV. CONCLUSION

This study illustrates that the implementation of image brightness normalization significantly enhances the performance of deep learning-based classification of apple leaf diseases when using the MobileNetV2 model. Among the eight evaluated normalization methods—None, Min–Max Scaling, Z-Score Normalization, Gamma Correction, Histogram Equalization, Contrast Limited Adaptive Histogram Equalization (CLAHE), Log Transformation, and Square Root Transformation—Min–Max Scaling yielded the most impressive results, achieving an accuracy of 95.83%, with a precision of 96.25%, recall of 95.83%, and an F1-score of 95.86%. In contrast, the model without normalization reached only 38.89% accuracy, underscoring the critical role that brightness normalization plays in enhancing the model's classification capabilities.

Overall, the two top-performing normalization methods, Min–Max Scaling and Histogram Equalization, attained an average accuracy of 95.14%, translating to an improvement of approximately 56.25% over the model that did not incorporate normalization. This outcome emphasizes the significance of brightness normalization in bolstering the model’s ability to reliably identify patterns of apple leaf diseases across varying illumination conditions.

The analysis of the confusion matrix reveals that the majority of classification errors occur between the Apple Scab and Apple Cedar Rust classes. This is likely due to their similar visual characteristics, including spot distribution, color intensity, and texture patterns. Nevertheless, the combination of MobileNetV2 and Min–Max Scaling demonstrates stable and reliable classification performance across all disease classes.

These findings confirm that brightness normalization plays a crucial role in stabilizing the distribution of pixel intensity values and mitigating the effects of lighting variations in image datasets. By standardizing the pixel value range and enhancing image contrast, the preprocessing stage allows the MobileNetV2 model to concentrate more effectively on key visual features, such as color variations, texture patterns, and characteristics of disease spots on apple leaves. As a result, the model is able to learn more representative features, leading to improved classification performance.

The findings of this study have practical implications for developing mobile or edge-based plant disease detection systems. MobileNetV2 is designed as a lightweight convolutional neural network architecture with relatively low computational requirements, making it well-suited for deployment on mobile devices or embedded systems. Additionally, Min–Max Scaling involves only straightforward mathematical operations with minimal computational overhead. Together, MobileNetV2 and Min–Max normalization create an efficient processing pipeline that could be implemented in smartphone-based agricultural monitoring applications, allowing farmers to perform early disease detection directly in the field without needing high-performance computing resources.

The primary contribution of this research lies in its thorough comparative evaluation of eight brightness normalization techniques and their impact on the performance of MobileNetV2 for apple leaf disease classification. In contrast to many prior studies that typically apply a single preprocessing technique without systematic comparisons, this research evaluates multiple normalization methods and demonstrates that the choice of brightness normalization technique significantly affects the performance of deep learning models. The results identify Min–Max Scaling as the most effective approach for enhancing classification robustness and accuracy under varying illumination conditions.

REFERENCES

[1] M. U. Ali, M. Khalid, M. Farrash, H. F. M. Lahza, A. Zafar, and S. H. Kim, “AppleLeafNet: a lightweight and efficient deep learning framework for diagnosing apple leaf diseases,” *Front. Plant Sci.*, vol. 15, 2024, doi: 10.3389/fpls.2024.1502314.

[2] A. S. Abade, P. A. Ferreira, and F. de B. Vidal, “Plant Diseases recognition on images using Convolutional Neural Networks: A Systematic Review,” Sep. 2020, [Online]. Available: http://arxiv.org/abs/2009.04365

[3] Md. R. Ahmed, A. F. Ashrafi, R. U. Ahmed, and T. Ahmed, “MCFFA-Net: Multi-Contextual Feature Fusion and Attention Guided Network for Apple Foliar Disease Classification,” Nov. 2022, doi: 10.1109/ICCIT57492.2022.10055790.

[4] Y. Jiao *et al.*, “LCAMNet: a lightweight model for apple leaf disease classification in natural environments,” *Front. Plant Sci.*, vol. 16, 2025, doi: 10.3389/fpls.2025.1626569.

[5] L. Fu, S. Li, Y. Sun, Y. Mu, T. Hu, and H. Gong, “Lightweight-Convolutional Neural Network for Apple Leaf Disease Identification,” *Front. Plant Sci.*, vol. 13, May 2022, doi: 10.3389/fpls.2022.831219.

[6] B. M. Usman, S. K. Johl, and P. A. Khan, “Enhancing energy sector sustainability through robust green governance mechanisms for carbon neutrality,” Dec. 01, 2025, *Springer Nature*. doi: 10.1007/s43621-025-01064-5.

[7] C. O. Artizzu, G. Allibert, and C. Demonceaux, “OMNI-CONV: Generalization of the Omnidirectional Distortion-Aware Convolutions,” *J. Imaging*, vol. 9, no. 2, Feb. 2023, doi: 10.3390/jimaging9020029.

[8] J. P. Schwarz Schuler, S. Romani, M. Abdel-Nasser, H. Rashwan, and D. Puig, “Reliable Deep Learning Plant Leaf Disease Classification Based on Light-Chroma Separated Branches,” in *Frontiers in Artificial Intelligence and Applications*, IOS Press BV, Oct. 2021, pp. 375–382. doi: 10.3233/FAIA210157.

[9] E. A. Taufik, A. F. Parsa, and S. A. M. Mostafa, “Efficient Leaf Disease Classification and Segmentation using Midpoint Normalization Technique and Attention Mechanism,” May 2025, [Online]. Available: http://arxiv.org/abs/2505.21316

[10] N. Sulaiya, “Enhanced Plant Leaf Disease Detection Using CLAHE-Processed Images and Custom CNN Deep Learning Architecture,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 13, no. 9, pp. 906–915, Sep. 2025, doi: 10.22214/ijraset.2025.74149.

[11] X. Niu and G. Zheng, “Research on Low Illumination Image Enhancement Algorithm Based on Convolutional Neural Network,” *Academic Journal of Computing & Information Science*, vol. 6, no. 9, pp. 115–121, 2023, doi: 10.25236/ajcis.2023.060917.

[12] H. Mao, W. Peng, Y. Tian, and X. Zhu, “A Novel Low-Illumination Image Enhancement Method Based on Convolutional Neural Network with Retinex Theory,” *Applied Sciences (Switzerland)*, vol. 15, no. 22, pp. 1–14, 2025, doi: 10.3390/app152212324.

[13] G. Restu Baihaqi, S. Reditha Shalsadilla, A. Maya Nur Maulidiya, and L. Muflikhah, “Enhancing DenseNet Accuracy in Retinal Disease Classification with Contrast Limited Adaptive Histogram Equalization,” *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*, vol. 10, no. 4, pp. 858–869, 2024, doi: 10.26555/jiteki.v10i4.30327.

[14] Z. Bi *et al.*, “Apple leaf disease severity grading based on deep learning and the DRL-Watershed algorithm,” *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-15246-8.

[15] V. S. Babu and N. V. Ram, “Deep Residual CNN with Contrast Limited Adaptive Histogram Equalization for Weed Detection in Soybean Crops,” *Traitement du Signal*, vol. 39, no. 2, pp. 717–722, 2022, doi: 10.18280/ts.390236.

[16] V. Rajpoot *et al.*, “Mango Plant Disease Detection System Using Hybrid BBHE and CNN Approach,” *Traitement du Signal*, vol. 39, no. 3, pp. 1071–1078, 2022, doi: 10.18280/ts.390334.

[17] C. Kaeser-Chen, F. Pathology, - Maggie, and S. Dane, “Plant Pathology 2020 - FGVC7,” Kaggle. Accessed: Dec. 09, 2025. [Online]. Available: https://www.kaggle.com/competitions/plant-pathology-2020-fgvc7

[18] M. T. Ahad, Y. Li, B. Song, and T. Bhuiyan, “Comparison of CNN-based deep learning architectures for rice diseases classification,” *Artificial Intelligence in Agriculture*, vol. 9, pp. 22–35, Sep. 2023, doi: 10.1016/j.aiia.2023.07.001.