Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 4, Juni 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i4.9818
Hal: 1583-1592

Analysis of Image Preprocessing on EfficientNet-B5 Performance in Acne Severity Classification

Dita Kurnia Rachmasari[1], Risqy Siwi Pradini[*], Ahsanun Naseh Khudori

Faculty of Science and Technology, Informatics Study Program, Institut Teknologi, Sains, dan Kesehatan RS.DR. Soepraoen Kesdam V/BRW, Malang, Indonesia
Email: [1]ditakurnia857@gmail.com, [2,*]risqypradini@itsk-soepraoen.ac.id, [3]ahsanunnaseh@itsk-soepraoen.ac.id
Correspondence Author Email: risqypradini@itsk-soepraoen.ac.id[*]
Submitted: **19/05/2026**; Accepted: **25/06/2026**; Published: **30/06/2026**

Abstract−Deep learning based acne severity classification requires consistent color distribution and image illumination for stable feature extraction. Color imbalance, noise, and lighting variations can affect the accuracy and generalization ability of the model, making image preprocessing optimization a crucial aspect in dermatological classification. This study analyzes the impact of image preprocessing on the performance of EfficientNet-B5 in classifying three levels of acne severity using the Kaggle Acne Grading dataset (999 images; 80% training, 20% testing). The experiment compares the default preprocessing (resize, normalization) with the proposed preprocessing: gray-world white balance for color stabilization, bilateral filtering for edge preservation, and adaptive gamma correction for adaptive illumination. The evaluation uses accuracy and loss curves, confusion matrices, and classification reports, focusing on the macro F1-score to assess the balance between precision and recall. The results show a slight increase in accuracy from 77% to 78%, a macro F1 score of 75%, and more controlled overfitting with smaller differences in accuracy and loss between training and validation. Improving image quality before feature extraction contributes to feature representation and balance in multi-class classification.

Keywords: Acne Classification; Acne Severity; Efficientnet-B5; Image Preprocessing; Macro F1-Score

1. INTRODUCTION

Acne is a skin problem that is often experienced by all groups, both teenagers and adults [1]. Acne occurs when pores become clogged with dust, dirt, oil, or dead skin cells, causing inflammation [2]. Based on Global Burden of Disease (GBD) data, the global prevalence of acne sufferers is estimated at 9.4%, and around 85% of these sufferers are young people aged 12 to 25 years [1]. This shows that acne is a skin health issue that requires attention, as it affects many young people in various countries. Acne not only affects physical appearance but also psychological well-being [3], such as decreased self-confidence along with the severity of the acne experienced [4]. By classifying the severity of acne it can help to choose the right treatment method [5]. Diagnosing acne severity often requires the assistance of a dermatologist, which is time-consuming and expensive. Therefore, an early health screening system for automatically and accurately classifying acne severity is necessary.

As artificial intelligence technology develops, there is research that uses deep learning to classify acne. [6]. High-quality images are required for deep learning implementation. Using images of varying quality can make it difficult for the model to find consistent patterns [7]. Examples of varying image quality include differences in lighting, varying skin tones, noise or visual disturbances, varying resolutions, and shooting angles. Uneven lighting can result in areas of the face being too bright or dark [8]. In addition, the skin color of each individual can affect the contrast between lesions and normal skin [9]. The presence of noise interference that covers the area around the face will trigger errors in detection [10].

To address varying image quality, image preprocessing is necessary. Image preprocessing can improve classification performance. [11]. Several previous studies have used image preprocessing to classify acne severity, such as the study by Suriani et al. [12] which image preprocessing techniques such as resizing and automatic cropping using the MobileNetV2 model to classify acne severity, achieving an accuracy of 92%. Another study conducted by Watanabe et al. [13] applied image preprocessing techniques such as ROI extraction, resizing, and normalization, then processed using the EfficientNet-B2 model to classify acne severity, achieving 90% accuracy. Meanwhile, research conducted by Ramadhani et al. [14] applied image preprocessing, namely resizing and rescaling, to the MobileNetV2 model to classify acne severity, achieving an accuracy of 87.29%. Therefore, selecting the right image preprocessing algorithm and technique can yield high accuracy in classifying acne severity.

In this study, three image preprocessing techniques will be used gray world white balance, bilateral filter, and adaptive gamma correction. These three image preprocessing techniques were chosen because they can address three fundamental problems in acne images color non-uniformity, skin noise, and low contrast. Furthermore, the combination of these image preprocessing techniques has the potential to provide greater accuracy. The gray world white balance technique is used to balance the image color, ensuring that the average of the three RGB channels is the same under the light source [15]. The bilateral filter technique is used to smooth the image while maintaining the edges of the acne so that they do not disappear [16]. Meanwhile, the adaptive gamma correction technique is used to brighten an image by calculating each pixel based on the luminance factor and average color [17].

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 4, Juni 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i4.9818

This study used the EfficientNet-B5 model, a pretrained Convolutional Neural Network (CNN) model with a compound scaling concept. This model was chosen because it balances depth, width, and resolution, resulting in good image computation and generalization [18]. In addition, the EfficientNet model can produce good accuracy with fewer parameters [19]. Several previous studies conducted by Shams et al. [20] using the EfficientNet-B0 model with 99% accuracy compared to Resnet50 and MobileNet. Meanwhile, Pramono et al. [21] using EfficientNet-V for multi-class acne classification achieved an accuracy of 96.66%.

Based on the previous studies mentioned above, there is a gap in the research the lack of specific image preprocessing techniques for classifying acne severity. Previous studies generally only used default image preprocessing methods, such as resizing and normalization [22], [23]. On the other hand, the use of more varied image preprocessing techniques can improve the quality of the dataset [24]. Most studies primarily focus on model architecture and classification performance, while the influence of image preprocessing on feature representation and model performance has received less attention. Therefore, this study proposes a more diverse image preprocessing stage by adding Gray World White Balance, Bilateral Filter, and Adaptive Gamma Correction. Furthermore, a comparison of the model performance between the default preprocessing and the proposed preprocessing is conducted, and the results are analyzed to assess the effect of using image preprocessing.

Thus, the purpose of this study is to apply the EfficientNet-B5 architecture to classify acne severity in facial images. The performance of the EfficientNet-B5 model will then be compared with the default image preprocessing and the proposed image preprocessing in the acne severity classification process. The comparison results will then be used to analyze the effect of the proposed image preprocessing stage on improving the model's accuracy in performing global grading of acne severity. It is hoped that the combination of gray world white balance, bilateral filter, and adaptive gamma correction methods in the EfficientNet-B5 architecture will be able to produce accurate acne severity classification and improve model performance.

2. RESEARCH METHODOLOGY

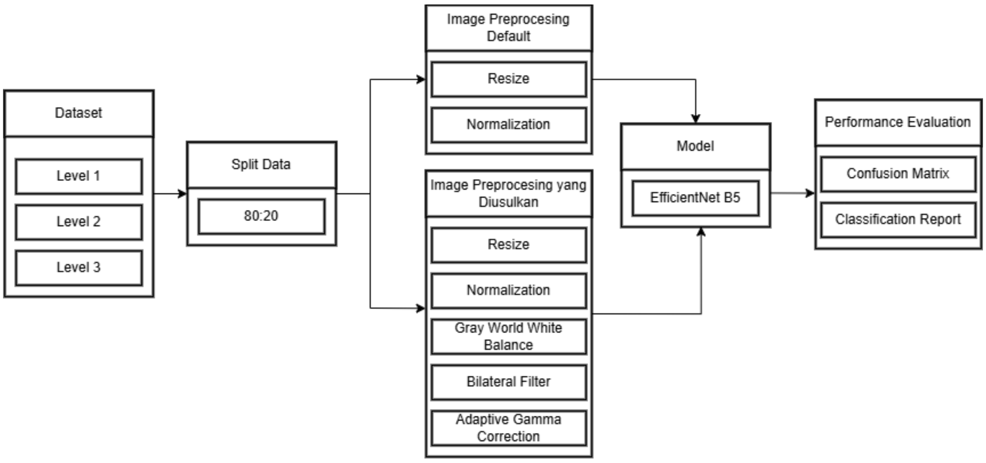


Figure 1. Research Stages

This study uses a quantitative experimental approach to analyze the effect of image preprocessing on the performance of the EfficientNet-B5 model in acne severity classification. The effect is determined by comparing the results of the model's performance before and after the application of certain image preprocessing methods. The research flow is carried out systematically, starting from data collection to model performance evaluation, with each stage being interconnected. The stages consist of dataset collection, data splitting, application of default and proposed image preprocessing methods, model development, and performance evaluation. As shown in figure 1, the complete flow of the research stages.

2.1 Dataset

Figure 2. Example of acne grading dataset with severity levels (a) level 0, (b) level 1, (c) level 2. Adapted from [25]

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 4, Juni 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i4.9818

The dataset used in this study is a secondary dataset obtained from the public repository on Kaggle [25]. It is used exclusively for academic research purposes in accordance with the terms and conditions of the dataset provider. The dataset concerns the classification of acne severity with a total of 999 images divided into three classes level 0 with 387, level 1 with 473, and level 2 with 139, as shown in Figure 2 Although this dataset has been widely used in acne classification studies, it has several limitations. First, the total number of images is relatively limited for deep learning applications. Second, the class distribution is unbalanced, especially for level 2, which contains significantly fewer samples than level 0 and level 1. This can impact the model's generalization ability and classification performance, particularly for minority class predictions. Due to the imbalance in distribution between classes, class weighting will be applied so that the model is not too biased towards the majority class. Each class is a label category that indicates differences in the severity of the skin condition. Level 0 or mild indicates a skin condition with very few lesions and a predominance of blackheads. Level 1 or moderate describes the presence of a greater number of papules and pustules with a more pronounced level of inflammation. Level 2 or severe indicates a significant increase in the number and size of lesions, accompanied by extensive inflammation that is characteristic of severe acne.

2.2 Split Data

The dataset will be split into two subsets training and testing, with 80% for training and 20% for testing. This proportion was chosen to ensure the model receives sufficient learning and maintains the quality of previously unseen data. The data split will be performed before image preprocessing to avoid data leakage [26]. This is done to prevent the model from simply memorizing patterns and to ensure that model evaluation can be carried out objectively.

2.3 Image Preprocessing

The image preprocessing stage aims to enhance image quality, reduce the effects of uneven lighting, and preserve important details in the skin area. The proposed image preprocessing pipeline is applied consistently to the training and testing datasets. In this study, the first stage of the proposed image preprocessing technique that will be implemented before model training includes resizing, which adjusts the entire image to maintain consistency [27]. At this stage, the adjusted image will have the same standard dimensions of 456x456. In the second stage, the image is normalized to ensure that the pixel values are within a consistent range [27]. This step balances the distribution of pixel values with a range of [0,1] so that there is no difference in scale between images. The next step in applying the Gray World White Balance technique is to normalize skin color so that the distribution of RGB channel intensity is balanced [15]. This technique assumes the image is close to neutral gray, thus eliminating bias due to differences in lighting conditions. The next step is the Bilateral Filter to reduce noise in the image without removing edges, so that the resulting image is more representative [16]. This Bilateral filter is applied with a diameter (d) of 7, sigma color 50, and sigma space 50. In the final stage, the Adaptive Gamma Correction technique is applied to adjust the contrast and brightness levels of the image [17]. This technique uses a gamma value calculated as $\gamma = 0.8 + 0.5 \times (\text{mean intensity} - 0.5)$, with the gamma value limited to the range 0.5–1.5. Images that have gone through the image preprocessing stage are then applied to training and final evaluation. The choice of image preprocessing in this study influences classification accuracy, primarily through color stability, noise, and contrast levels. This approach ensures that each image entered into the model has uniform and optimal quality.

2.4 Model

The model used in this study is EfficientNet-B5, which is a Convolutional Neural Network architecture shown in **Error! Reference source not found.**.

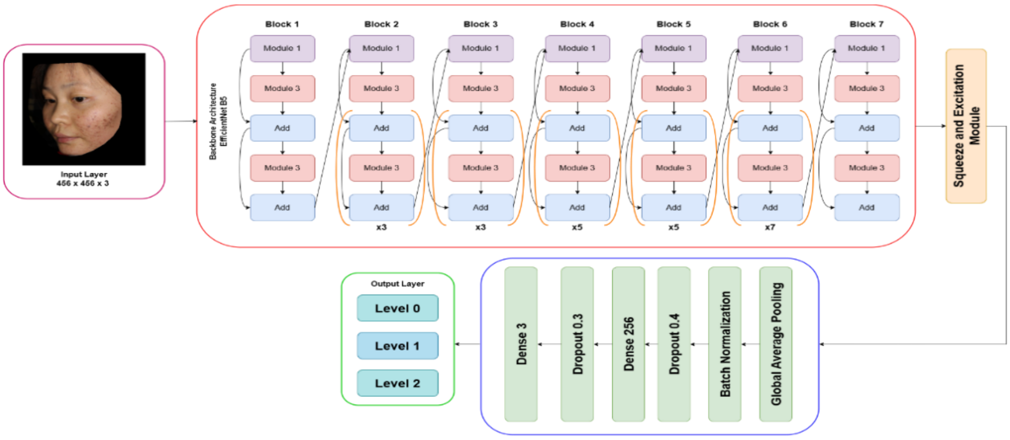


Figure 3. EfficientNet-B5 Model Architecture

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 4, Juni 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i4.9818

EfficientNet-B5 was chosen because of its compound scaling concept, which can balance depth, width, and resolution [28]. Furthermore, this architecture supports transfer learning on limited datasets. Input images of size 456 x 456 will be extracted by EfficientNet-B5, which has been pre-trained on ImageNet [29]. In the backbone, there are 7 blocks consisting of Mobile Inverted Bottleneck Convolutional (MBConv), skip connection through Add operation, and Squeeze and Excitation Module, which are used to improve the model in emphasizing the most informative feature channels [30]. The features are then processed by a global average pooling layer, producing a 2048-dimensional feature vector to reduce dimensionality. The results are then normalized using batch normalization to stabilize the distribution, followed by a 0.4 dropout layer to reduce overfitting. This is followed by a dense layer of 256 neurons as a fully connected layer, added with a dropout layer of 0.3, ending with a dense layer of 3 neurons [5]. The final result of the neuron is a prediction of three classes of acne severity, namely level 0, level 1, and level 2. After constructing the EfficientNet-B5 architecture, the model was trained using the configuration described as follows. The EfficientNet-B5 model was trained using the Adam optimizer with a batch size of 16 and an input image size of 456 × 456 pixels. Transfer learning was implemented in two stages. In the first stage, the EfficientNet-B5 backbone was frozen and only the classification head was trained for 15 epochs using a learning rate of 1e-3. In the second stage, the last 80 layers of the backbone were unfrozen and fine-tuned for an additional 10 epochs using a reduced learning rate of 1e-4. Class imbalance was addressed by applying weights of 0.86, 0.70, and 2.39 to Level 0, Level 1, and Level 2, respectively. Early stopping and learning rate reduction strategies were also employed to improve training stability and reduce the risk of overfitting.

2.5 Model Evaluation

The evaluation of model performance was conducted utilizing a confusion matrix and a classification report. The dataset was divided into training and testing sets using an 80:20 ratio. The training set was used for model learning, while the testing set was utilized as validation data during the training process to monitor model performance through accuracy and loss curves. After training was completed, the same testing set was used for final evaluation using confusion matrix, precision, recall, F1-score, and macro F1-score metrics. The confusion matrix was used to measure the correct predictions and incorrectly classified instances per class. Referring to the confusion matrix, the True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) values were obtained [31]. While the classification report displays performance per class with metrics such as accuracy, precision, recall, and F1-score [32]. However, the primary metric used in this study is the macro F1-score. This metric was chosen based on the characteristics of multi-class classification with an uneven data distribution across acne severity classes. These evaluation results serve as the basis for understanding the effect of image preprocessing and comparing model performance between the default and proposed image preprocessing methods.

3. RESULT AND DISCUSSION

3.1 Image Preprocessing

This study applies two image preprocessing scenarios, namely the default image preprocessing and the proposed image preprocessing. In the default image preprocessing, resize is applied to change the image size to a uniform size of 456 x 456 pixels. This size is in accordance with the EfficientNet-B5 model and ensures model consistency. After the resizing process, normalization is performed to stabilize the range of intensity values to suit the model's needs.

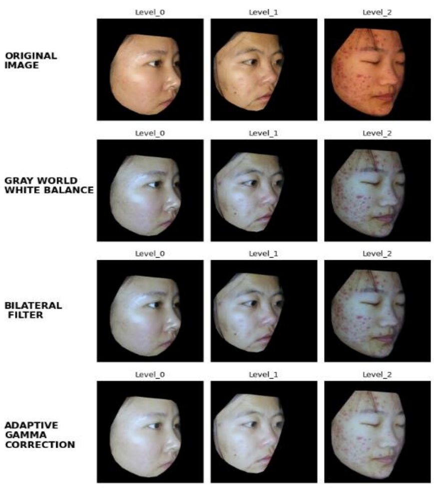
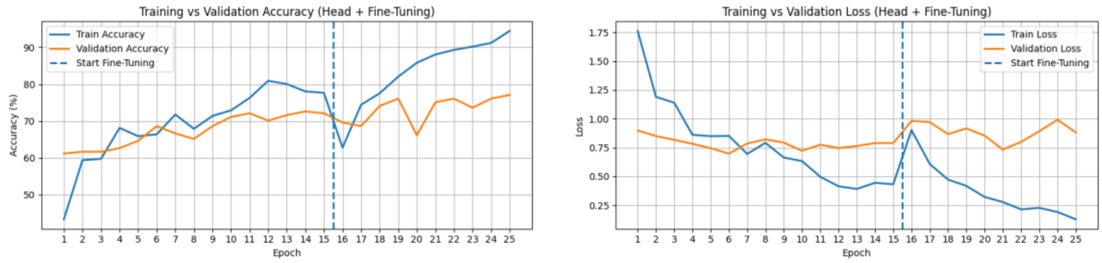


Figure 4. Results of Image Preprocessing Gray World White Balance, Bilateral Filter, and Adaptive Gamma Correction

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 4, Juni 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i4.9818

This study applies two image preprocessing scenarios, namely the default image preprocessing and the proposed image preprocessing. In the default image preprocessing, resize is applied to change the image size to a uniform size of 456 x 456 pixels. This size is in accordance with the EfficientNet-B5 model and ensures model consistency. After the resizing process, normalization is performed to stabilize the range of intensity values to suit the model's needs. The proposed Image Preprocessing applies several additional techniques, namely Gray World White Balance, Bilateral Filter, and Adaptive Gamma Correction, as shown in figure 4. The application of Gray World White Balance helps normalize the varying conditions due to lighting. Furthermore, the bilateral filter is able to reduce noise in the image but still maintains the edges from being lost. Meanwhile, adaptive gamma correction corrects areas that are too dark or too bright.

3.2 Default Image Preprocessing Performance Results



(a) (b)

Figure 5. Model performance results from default image preprocessing (a) Training and Validation Accuracy Curves, (b) Training and Validation Loss Curves

The performance evaluation of the EfficientNet-B5 model with default image preprocessing is shown in Figure 5(a) through training and validation accuracy curves. In the initial training stage, only training the classifier head portion in epochs 1-15 consistently increased from around 43-45% in the first epoch to ± 78-80% in the 15th epoch. The improvement is seen in the validation accuracy, which increased from 61% to around ± 72-73%. In the fine-tuning stage in the 16th epoch, there was a temporary decrease to around 63-65%, accompanied by an increase in training loss. This indicates a re-adjustment process on the EfficientNet-B5 backbone to better suit the specific characteristics of acne images. After the fine-tuning stage was applied, the model achieved additional performance improvement during epochs 17–25. Training accuracy continued to increase to around 94-95% at the end of training, while validation accuracy was around 75-77%. In line with the accuracy curve, the training and validation loss graphs shown in Figure 5(b) show a decline in training loss from approximately 1.75 to ± 0.12 at the end of the epoch. In contrast, the validation loss is in the range of 0.80–1.00 and tends to stagnate after the fine-tuning phase. In addition to the default image preprocessing, the training process also applies class weights to address data imbalance between classes. Assigning higher weights to level 2 classes helps the model continue to learn the characteristics of the minority class even though it has fewer samples than the other classes.

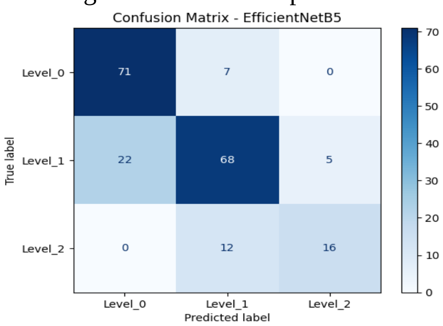


Figure 6. Confusion Matrix

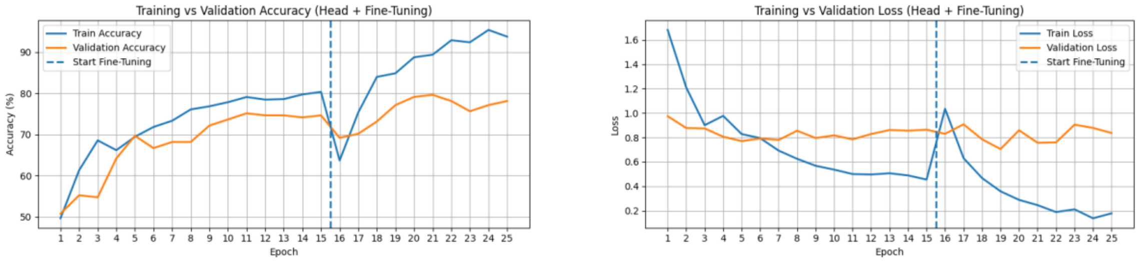
Severity Level	Precision	Recall	F1-Score	Accuracy
Level 0	0.76	0.91	0.83	
Level 1	0.78	0.72	0.75	
Level 2	0.76	0.57	0.65	77%
Macro avg	0.77	0.73	0.74	
Weighted avg	0.77	0.77	0.77	

Table 1. Classification Report

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 4, Juni 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i4.9818

The evaluation results using the confusion matrix are shown in Figure 6, and the classification report in Table 1. shows an accuracy of 77% on the test data. The Level 0 class has a precision of 0.76, a recall of 0.91, and an F1-score of 0.83, represents a fairly good model for recognizing mild acne. Level 1 obtained a precision of 0.78, a recall of 0.72, and an F1-score of 0.75, represents a similarity in characteristics between mild and moderate acne. Meanwhile, Level 2 had a precision of 0.76, a recall of 0.57, and an F1-score of 0.65, indicating that many severe acne images were still misclassified into lower classes. Overall, the model obtained a macro average F1-score of 0.74 and a weighted average F1-score of 0.77. The lower macro average value compared to the weighted average indicates an imbalance in performance between classes, especially in the minority class.

3.3 Performance Results of the Proposed Image Preprocessing



(a) (b)

Figure 7. Model performance results from the proposed image preprocessing (a) Training and Validation Accuracy Curves, (b) Training and Validation Loss Curves

The performance evaluation of the EfficientNet-B5 model with the proposed image preprocessing is shown in Figure 7(a) through the training and validation accuracy curves. In the initial training stage, only training the classifier head part in epochs 1-15 consistently increases by around 50% to nearly ± 80% in the 15th epoch. Meanwhile, the validation accuracy increases by around 51% to around ± 74-75%. In the fine-tuning stage in the 16th epoch, there is a temporary decrease of around 64%, accompanied by an increase in the training loss. In epochs 17-25, the training accuracy reaches around 94-95%, while the validation accuracy is around 76-79%. In line with the accuracy curves, the training and validation loss graphs shown in Figure 7(b) show a decline in loss from around 1.65 to ± 0.15 at the end of training. Meanwhile, the validation loss tends to be relatively stable at around 0.75-0.90 despite experiencing slight fluctuations during the fine-tuning process. Similar to the previous scenario, model training applies class weights to mitigate the effects of data imbalance. However, when combined with the proposed image preprocessing, the model is able to utilize better feature information, increasing sensitivity to minority classes. This is reflected in the slight increase in Level 2 class recall and a higher macro F1-score compared to the default image preprocessing scenario.

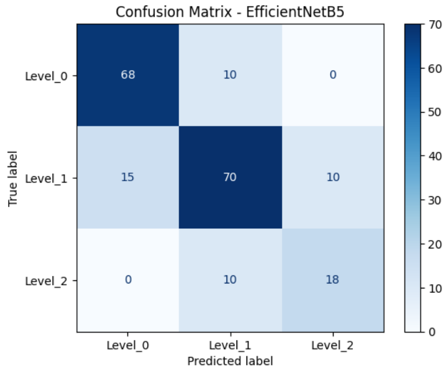


Figure 8. Confusion Matrix

Table 2. Classification Report

Severity Level	Precision	Recall	F1-Score	Accuracy
Level 0	0.76	0.91	0.83	
Level 1	0.78	0.72	0.75	
Level 2	0.76	0.57	0.65	
				77%
Macro avg	0.77	0.73	0.74	
Weighted avg	0.77	0.77	0.77	

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 4, Juni 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i4.9818

The evaluation results using the confusion matrix are shown in Figure 8, and the classification report in Table 2 shows an accuracy of 78% on the test data. The level 0 class obtained a precision value of 0.82, a recall of 0.87, and an F1-score of 0.84, indicating that the model is very good at recognizing mild acne. Level 1 obtained a precision value of 0.78, a recall of 0.74, and an F1-score of 0.76, represents a balanced classification in recognizing moderate acne images. Meanwhile, the level 2 class had a precision value of 0.64, a recall of 0.64, and an F1-score of 0.64, indicating that some severe acne images were misclassified to a lower class. Overall, the macro average F1-score was 0.75, indicating that performance between classes was more balanced.

3.4 Comparison with Previous Research

Based on the comparison table shown in **Error! Reference source not found.**, this study is within the context of an acne classification study using the Acne Grading Classification dataset. However, direct comparisons require caution due to differences in model architecture, preprocessing, data splitting strategies, and evaluation metrics. Therefore, these results are contextual and not a claim of absolute superiority. The findings suggest that image preprocessing improves the performance of EfficientNet-B5 by strengthening feature representations before classification. In previous studies, the majority used accuracy as the primary metric in evaluation, such as the results of Singhapathiranage et al. with an accuracy of 75.7% [33], Paluri et al. with 96% accuracy [34], Shetty et al. with 80% accuracy [35], and Ali et al. with an accuracy of 99.47% [36]. Our research yielded an accuracy of 78%, but this was not the primary metric, using the macro F1-score, which yielded 75%. Although this is still below Paluri et al. 93% result [34], these differences must be viewed proportionally, taking into account the methodology and evaluation metrics. The macro F1-score was chosen as the primary metric based on an imbalanced dataset, so the performance evaluation considers the balance between precision and recall for each class. Unlike accuracy, which tends to be biased toward the majority class, the macro F1-score can demonstrate the model's performance on the classification task of all levels of acne severity without bias toward a particular class. Although the macro F1-score is still low compared to the study by Paluri et al. [34] in studies with complex architectures or aggressive augmentation, these results show that optimizing image preprocessing can produce stable and competitive performance.

Compared to previous research, this study uses EfficientNet-B5 with an image preprocessing approach that focuses on image stabilization. The proposed strategy focuses on global illumination correction, noise reduction while preserving texture details, and adaptive lighting adjustment. Overall, this study confirms that improving image quality through structured image preprocessing significantly contributes to the stability of feature representations. This approach also improves the performance of acne severity classification using EfficientNet B5, especially when evaluated with metrics sensitive to class imbalance.

Table 3. Comparison with Previous Research Based on the Dataset

References (Author & Year)	Dataset	Input	Architectural Model	Image Preprocessing	Results
Singhapathiranage et al., 2025 [33]	Kaggle + DermNet NZ	Facial image (3 levels + European and South Asian skin)	ResNet-152	OpenCV (Facial Landmark), Rolling Augmentation	Accuracy: 75,7% Macro F1-score: -
Paluri et al., 2024 [34]	Kaggle Acne Grading Classification	Facial image (3 levels)	Ensamble Model (ETLoVIT)	Resize, Normalization, Augmentation (horizontal flip, random rotation, and zoom)	Accuracy: 96% Macro F1-score: 93%
Shetty et al., 2025 [35]	Kaggle Acne Grading Classification	Facial image (color)	EfficientNet-B0	Normalization, Clustering (YCrCb and HSV)	Accuracy: 80% Macro F1-score: -
Ali et al., 2025 [36]	Kaggle Acne Grading Classification	Facial image (3 levels)	DenseNet 121	Resize, Rotation, Horizontal Flipping, Random Brightness/Contrast, Gamma Correction, CLAHE, RGB Shifting, Bluring	Accuracy: 99.47% Macro F1-score: -
Our Research	Kaggle Acne Grading Classification	Facial image (3 levels)	EfficientNet-B5	Resize, Normalisasi, Gray World White	Accuracy: 78%

References (Author & Year)	Dataset	Input	Architectural Model	Image Preprocessing	Results
				Balance, Bilateral Filter, Adaptive Gamma Correction	Macro F1-Score: 75%

3.5 Discussion

3.5.1 Analysis of the Image Preprocessing on Model Performance

The results show that the proposed image preprocessing provides improvements compared to the default image preprocessing. The accuracy increased from 77% to 78%, and the macro F1-score of 75% indicate that the combination of gray world white balance, bilateral filter, and adaptive gamma correction contributes to the feature extraction process. Specifically, gray world white balance helps balance colors between images, minimizing variations in lighting. This color normalization allows the EfficientNet-B5 backbone to extract texture and pattern features from acne more consistently. The bilateral filter plays a role in reducing noise without removing edges, thus maintaining structural information about acne. Meanwhile, adaptive gamma correction enhances both dark and light areas, making details of acne severity clearer. This indicates that the combination of these three techniques can produce more informative features and support the model's extraction process. The impact is also seen in the improvement in level 2 class performance, where recall values increase compared to the default image preprocessing scenario. These results indicate that the proposed image preprocessing can help the model recognize more severe acne cases than previously missed. However, the recall value is lower compared to other classes due to the similarity of characteristics between moderate and severe acne. This is influenced by the limited amount of data in the level 2 class and the diversity of acne inflammation patterns. Although the proposed image preprocessing technique can improve performance and balance the error distribution, the macro F1-score value of 75% indicates that the model's performance is not yet fully optimal. These results indicate that the applied image preprocessing is proven to help improve feature representation, but does not fully separate the severity characteristics optimally. Thus, there is still room for further development, particularly through additional image preprocessing techniques or in combination with other image quality improvement strategies to strengthen the discrimination of inflammatory features.

3.5.2 Model Performance using Macro F1-Score

The model performance analysis in this study used the macro F1-score as the primary metric. This metric was chosen based on the characteristics of the dataset, which exhibits an imbalanced class distribution, particularly for high-severity classes with a smaller sample size. Under these conditions, accuracy can potentially bias performance estimates toward the majority class. This occurs because accuracy only represents the overall proportion of correct predictions without considering the distribution of errors within each class. In contrast, the macro F1-score calculates the average of the F1-scores for each class independently, thus giving equal weight to all acne severity levels. This method allows for a fairer evaluation of the model's ability to detect minority classes. Furthermore, this metric also assesses the equilibrium between precision and recall across all categories. Therefore, the use of the macro F1-score in this study is more relevant for evaluating multi-class classification performance, which is sensitive to data imbalance. This metric also provides a more representative interpretation of the model's ability to recognize all acne severity classes proportionally.

4. CONCLUSION

The application of image preprocessing with a combination of gray-world white balance, bilateral filtering, and adaptive gamma correction has been shown to provide a slight improvement in feature representation quality. This technique also supports the performance of acne severity classification with EfficientNet-B5. Analysis shows that this approach results in more stable training, a more balanced distribution of errors between classes, and better sensitivity to minority classes with a macro F1-score of 75%. This research contribution demonstrates that improving image quality plays a direct role in strengthening model generalization without increasing architectural complexity. Therefore, image preprocessing plays a strategic role in developing an optimal and stable deep learning-based dermatological image classification system. Future research can be directed by exploring more adaptive image preprocessing techniques, integrating lesion-characteristic augmentation strategies, and testing on more diverse datasets. These efforts are expected to improve discriminatory capabilities and strengthen the stability of model generalization comprehensively.

REFERENCES

[1] D. L. Guguluș *et al.*, “The Epidemiology of Acne in the Current Era: Trends and Clinical Implications,” *Cosmetics*, vol. 12, no. 3, Art. no. 106, 2025, doi: 10.3390/cosmetics12030106.
[2] N. U. R. Sifatullah, “Jerawat (Acne vulgaris): Review Penyakit Infeksi Pada Kulit,” *Pros. Semin. Nas. Biol.*, vol. 7, no.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 4, Juni 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i4.9818

1, pp. 19–23, 2021, doi: 10.24252/psb.v7i1.22212.
[3] A. M. Layton, V. Bettoli, V. Delore, E. Puentes, and J. K. L. Tan, "The Burden of Acne Vulgaris on Health-Related Quality of Life and Psychosocial Well-Being Domains: A Systematic Review," *Am. J. Clin. Dermatol.*, vol. 27, no. 1, pp. 17–47, 2025, doi: 10.1007/s40257-025-00983-3.
[4] R. F. Autrilia and R. H. Ninin, "Eksplorasi Dampak Psikologis pada Remaja yang Memiliki Masalah Penampilan dengan Jerawat," *J. Psikol. Udayana*, vol. 9, no. 2, pp. 194–205, 2022, doi: 10.24843/jpu.2022.v09.i02.p09.
[5] V. D. Br Sebayang and I. G. N. L. W. Kusuma, "Klasifikasi Jenis Jerawat Berdasarkan Citra Menggunakan Convolutional Neural Network dengan Arsitektur MobileNetV2," *J. FASILKOM, vol. 14, no. 3, pp. 766–774, 2024,* doi: 10.37859/jf.v14i3.8202.
[6] D. O. Traini, G. Palmisano, C. Guerriero, and K. Peris, "Artificial Intelligence in the Assessment and Grading of Acne Vulgaris: A Systematic Review," *J. Pers. Med.*, vol. 15, no. 6, Art. no. 238, 2025, doi: 10.3390/jpm15060238.
[7] D. Hussain and Y. H. Gu, "Exploring the Impact of Noise and Image Quality on Deep Learning Performance in DXA Images," *Diagnostics*, vol. 14, no. 13, Art. no. 1328, 2024, doi: 10.3390/diagnostics14131328.
[8] Z. Younis, M. S. M. Rahim, and F. Bin Mohamed, "Enhancement of Unevenly Illuminated Images: An Experiment Based Review," *PROtek J. Ilm. Tek. Elektro*, vol. 12, no. 2, pp. 112–120, 2025, doi: 10.33387/protk.v12i2.8533.
[9] M. Benčević, M. Habijan, I. Galić, D. Babin, and A. Pižurica, "Understanding skin color bias in deep learning-based skin lesion segmentation," Comput. Methods Programs Biomed., vol. 245, Art. no. 108044, 2024, doi: 10.1016/j.cmpb.2024.108044.
[10] J. A. Rodríguez-Rodríguez, E. López-Rubio, J. A. Ángel-Ruiz, and M. A. Molina-Cabello, "The Impact of Noise and Brightness on Object Detection Methods," *Sensors*, vol. 24, no. 3, Art. no. 821, 2024, doi: 10.3390/s24030821.
[11] M. Resa, A. Yudianto, P. Sukmasetya, R. A. Hasani, and D. Sasongko, "Pengaruh Data Preprocessing terhadap Imbalanced Dataset pada Klasifikasi Citra Sampah menggunakan Algoritma Convolutional Neural Network," *Build. Informatics, Technol. Sci.*, vol. 4, no. 3, pp. 1367–1375, 2022, doi: 10.47065/bits.v4i3.2575.
[12] N. S. Suriani, S. S. A. Tarmizi, M. N. H. Mohd, and S. M. Shah, "Acne Severity Classification on Mobile Devices using Lightweight Deep Learning Approach," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 6, pp. 680–687, 2024, doi: 10.14569/IJACSA.2024.0150668.
[13] K. Watanabe, K. Iinuma, C. Nakashima, H. Yamamoto, N. Oiso, and A. Otsuka, "Deep Learning-Based Acne Severity Classification Using Standardized Facial Images of Japanese Patients," *Cureus*, vol. 17, no. 9, Art. no. e91944, 2025, doi: 10.7759/cureus.91944.
[14] F. Ramadhani, S. Rahardiantoro, and M. Masjkur, "Acne Severity Classification Study Using Convolutional Neural Network Algorithm with MobileNetV2 Architecture," *Indones. J. Stat. Its Appl.*, vol. 8, no. 2, pp. 112–128, 2024, doi: 10.29244/ijsa.v8i2p112-128.
[15] S. J. Kuryanti *et al.*, "Peningkatan Kualitas Citra Bawah Air Berbasis Algoritma Fusion Dengan Keseimbangan Warna, Optimalisasi Kontras," *INTI Nusa Mandiri*, vol. 16, no. 1, pp. 1–8, 2021, doi: 10.33480/inti.v16i1.2286.
[16] J. Maus *et al.*, "Deep Learning Based Bilateral Filtering for Edge-Preserving Denoising of Respiratory-Gated PET," *EJNMMI Phys.*, vol. 11, Art. no. 58, 2024, doi: 10.1186/s40658-024-00661-z.
[17] G. A. Alhamzawi, A. S. Alfoudi, and A. H. Alsaeedi, "Fusion Deep Learning with Adaptive Gamma Correction (DLAGC) to Enhance Images in Low Light," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 36s, pp. 882–893, 2025, doi: 10.52783/jisem.v10i36s.6610.
[18] M. A. Ghofur and A. B. Ulum, "Analisis Perbandingan Kinerja Model ResNet50 Dan EfficientNet-B0 Dalam Klasifikasi Kerusakan Jalan," *J. Media Inform.*, vol. 6, no. 5, pp. 2504–2511, 2025, doi: 10.55338/jumin.v6i5.6843.
[19] I. de Zarza, J. de Curto, and C. T. Calafate, "Detection of glaucoma using three-stage training with EfficientNet," *Intell. Syst. Appl.*, vol. 16, Art. no. 200140, 2022, doi: 10.1016/j.iswa.2022.200140.
[20] M. Y. Shams *et al.*, "Skin Disease Classification: A Comparison of ResNet50, MobileNet, and Efficient-B0," *J. Curr. Multidiscip. Res.*, vol. 1, no. 1, pp. 1–7, 2025, doi: 10.21608/jcmr.2025.327880.1002.
[21] A. Y. Pramono and K. Kusnawi, "Multi-Class Facial Acne Classification using the EfficientNetV2-S Deep Learning Model," *Aviat. Electron. Inf. Technol. Telecommun. Electr. Control.*, vol. 7, no. 3, pp. 299–309, 2025, doi: 10.28989/avitec.v7i3.3157.
[22] M. I. Anshori, R. A. Yaqin, M. Ali, Z. Sidiq, and W. P. Agung, "Klasifikasi Jenis Jerawat Secara Otomatis Dengan Convolutional Neural Network Menggunakan Arsitektur ResNet-50," *J. Manaj. Inform.*, vol. 15, no. 1, pp. 73–84, 2025, doi: 10.34010/jamika.v15i1.13712.
[23] S. N. Hasan, "Accurate Deep Learning Algorithms for Skin Lesion Classification," *Ing. Syst. Inf.*, vol. 29, no. 4, pp. 1529–1539, 2024, doi: 10.18280/isi.290426.
[24] A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging," *Cancers*, vol. 15, no. 16, Art. no. 4172, 2023, doi: 10.3390/cancers15164172.
[25] R. Lathiya Teksun, "Acne Grading Classification Dataset," Kaggle, 2020. [Online]. Available: https://www.kaggle.com/datasets/rutviklathiyateksun/acne-grading-classificationdataset. [Accessed: Nov. 2025].
[26] A. Apicella, F. Isgrò, and R. Prevete, "Don't push the button! Exploring data leakage risks in machine learning and transfer learning," *Artif. Intell. Rev.*, vol. 58, Art. no. 339, 2025, doi: 10.1007/s10462-025-11326-3.
[27] M. I. Burhanuddin, A. Syaifullah, S. A. P. Jaya, and M. G. Somoal, "Analisis Komparatif Model MobilenetV1 Dan EfficientnetB0 Dalam Klasifikasi Citra Empat Musim Menggunakan Transfer Learning," *J. Tek. Inform.*, vol. 5, no. 2, pp. 508–521, 2025, doi: 10.58794/jekin.v5i2.1378.
[28] A. L. Rukuna *et al.*, "Enhancing Monkeypox Detection with Efficientnet-B5 And Image Augmentation Fusion Technique," *Int. J. Sci. Res. Sci. Technol.*, vol. 11, no. 6, pp. 646–661, 2024, doi: 10.32628/IJSRST241161119.
[29] K. Kanchana, S. Kavitha, K. J. Anoop, and B. Chinthamani, "Enhancing Skin Cancer Classification Using Efficient Net B0-B7 through Convolutional Neural Networks and Transfer Learning with Patient-Specific Data," Asian Pac. J. Cancer Prev., vol. 25, no. 5, pp. 1795–1802, 2024, doi: 10.31557/APJCP.2024.25.5.1795.
[30] R. I. Sumon, H. Ali, S. Akter, S. Muhammad, and I. Uddin, "A Deep Learning-Based Approach for Precise Emotion

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 4, Juni 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i4.9818

Recognition in Domestic Animals Using EfficientNetB5 Architecture," *Engineering*, vol. 6, no. 1, Art. no. 9, 2025, doi: 10.3390/eng6010009.

[31] A. Nur, A. Thohari, L. Triyono, I. Hestiningsih, B. Suyanto, and A. Yobioktobera, "Performance Evaluation of Pre-Trained Convolutional Neural Network Model for Skin Disease Classification," *JUITA J. Inform.*, vol. 10, no. 1, pp. 9–18, 2022, doi: 10.30595/juita.v10i1.12041.

[32] R. S. Pradini, W. T. Kusuma, and A. S. Budi, "Convolutional neural network model for fingerprint-based gender classification using original and degraded images," *Int. J. Adv. Appl. Sci.*, vol. 14, no. 4, pp. 1350–1358, 2025, doi: 10.11591/ijaas.v14.i4.pp1350-1358.

[33] A. P. Singhapathiranage, D. D. K. Kelaart, A. A. M. Thathsarani, C. D. L. Liyanage, and B. N. S. Lankasena, "Optimizing Acne Severity Detection: A Deep Learning Approach with Electronic Medical Record System Integration and Diverse Image Data," *Adv. Technol.*, vol. 5, no. 1, pp. 52–66, 2025, doi: 10.31357/ait.v5i01.8414.

[34] K. V. Paluri and A. Gupta, "ETLoViT: An acne diagnosis approach using vision transformers and model ensembling," Mach. Learn.: Sci. Technol., vol. 6, no. 3, Art. no. 035028, 2024, doi: 10.1088/2631-8695/ad7ad9.

[35] B. Shetty, N. Shetty, S. Suvarna, and P. R. Kolte, "Dermavision: Transforming Skincare, One Recommendation at a Time," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 13, no. 5, pp. 2890–2897, 2025, doi: 10.22214/ijraset.2025.70322.

[36] D. Ali, S. Azzaben, and R. Bajaman, "Optimized Acne Classification Using Transfer Learning with DenseNet121 and ResNet152," *in Proc. 2025 16th Student Res. Conf. Appl. Comput. (SRC), 2025, pp. 1–5,* doi: 10.1109/SRC65875.2025.11263689.