

A Controlled Comparative Study of EfficientNetB0, ResNet50, and DenseNet121 for Multi-Class Skin Disease Classification

Nasiru Bashir Umar¹, B. L. Pal², Ahmad Aminu³, Ajay Singh Dhabariya⁴

¹MSc. (AI) Research Scholar, Mewar University Gangrar, Chittorgarh, Rajasthan, India

²Assistant Professor Mewar University Gangrar, Chittorgarh, Rajasthan, India

³MSc. (AI) Research Scholar, Mewar University Gangrar, Chittorgarh, Rajasthan, India

⁴Assistant Professor, Mewar University Gangrar, Chittorgarh, Rajasthan, India

Abstract

Skin disease classification from images is difficult because different disease types share visual characteristics, and image appearance can be altered by image quality and acquisition conditions. A controlled comparison of EfficientNetB0, ResNet50, and DenseNet121 was conducted under a common experimental configuration. The Skin Disease Dataset was obtained from Kaggle, consisting of 4,357 images of skin diseases divided into five classes: Acne, Eczema, Psoriasis, Tinea, and Vitiligo. Common pre-processing, training-only data augmentation, class weighting, and two-phase transfer learning from ImageNet-pretrained models were applied to all models. The accuracy, precision, recall, F1 score, confusion matrices, and results for three training seeds were used for the evaluation. Classical machine-learning baselines, including a support vector machine (SVM) and a random forest, were also evaluated. Results: ResNet50 achieved the highest mean accuracy ($80.5\% \pm 0.9\%$) and macro-F1 ($81.8\% \pm 0.9\%$). DenseNet121 achieved $74.6\% \pm 0.5\%$ accuracy and $76.4\% \pm 0.5\%$ macro-F1, while EfficientNetB0 achieved $72.2\% \pm 2.8\%$ accuracy and $73.8\% \pm 2.5\%$ macro-F1. Random Forest achieved 57.8% accuracy and 58.3% macro-F1, while SVM achieved 53.8% accuracy and 53.9% macro-F1. All deep-learning models appeared to classify images of Acne and Vitiligo more successfully, whereas images of Eczema, Psoriasis, and Tinea were more frequently misclassified. ResNet50 appears to have achieved the best overall classification accuracy under the experimental conditions. The results, however, are only specific to the chosen evaluation setup and dataset. More generalizability testing is required to determine if observed performance applies across populations, skin tones, and image-acquisition conditions in larger, more diverse, independent datasets.

Keywords: Deep learning, convolutional neural networks, transfer learning, EfficientNetB0, ResNet50, DenseNet121, multi-class classification.

1. Introduction

The visual properties of skin diseases are diverse, with differences in the appearance, texture, color, and distribution of the skin. Automated image classification remains challenging because diseases may exhibit similar visual features, while image quality and illumination may also

Published: 17 October 2026

DOI: <https://doi.org/10.70558/IJMRS.2026.v2.i4.301233>

Copyright © 2026 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

affect their appearance. Skin disease classification (Allugunti, 2022; Fraiwan & Faouri, 2022) using a convolutional neural network (CNN) and transfer learning has been investigated. (Saifan & Jubair, 2022) explored multi-class classification of CNN on a small dataset, and (Sadik et al., 2023) proposed a comparison of six pretrained CNN models on five skin disease classes. (AlSuwaidan, 2023) also investigated the impact of architecture selection on the classification accuracy. Additional research further shows that results are dependent on the composition of the dataset, the categories of diseases, and experimental settings (Abbas et al., 2022; Deepak et al., 2025). The findings suggest that deep learning may support skin disease classification and that controlled evaluation and comparison of CNN architectures may be warranted.

Transfer learning may facilitate the fine-tuning of pretrained convolutional neural networks (CNNs) for skin disease classification when only limited medical-image data are available. Different architectures and performance differences have been noted on datasets of dermatological images (AlSuwaidan, 2023; Sadik et al., 2023). (Abbas et al., 2022) compared DenseNet121 and ResNet50 with a custom CNN trained on HAM10000, and (Gulzar et al., 2025) examined hybrid transfer-learning techniques for multi-class skin disease prediction. (Alruwaili & Mohamed, 2025) tested EfficientNet-B0, EfficientNet-B2, and ResNet50 on a set of 10 skin disease classes with augmentation and class weighting. (Deepak et al., 2025) also explored the EfficientNet-B0 and ResNet models with three disease classes. The differences among the datasets, class selection, dataset preprocessing, and training configurations emphasize the importance of comparing the performance of EfficientNet-B0, ResNet50, and DenseNet121 under the same experimental settings (Alruwaili & Mohamed, 2025; Deepak et al., 2025; Gulzar et al., 2025).

One of the major difficulties in skin disease classification is the imbalance in the number of images in each disease class. All but one of the samples were minority classes, which may increase the risk of under-representation during training, as pointed out by (Fraiwan & Faouri, 2022). In the comparison between DenseNet121 and ResNet50, (Abbas et al., 2022) employed a data augmentation technique; (Naeem & Anees, 2024) studied how to deal with the class imbalance problem in multi-class skin-cancer classification using an oversampling technique. Preprocessing, augmentation, and class-balancing methods were used in other studies for the classification of psoriasis and lichen planus, further demonstrating the importance of appropriate data preparation. There have been recent studies investigating augmentation to diversify training data for dermatology (Kim et al., 2025). Because of variation in datasets, disease classes, preprocessing, data splitting, augmentation, and evaluation metrics, cross-study results may not be directly comparable. (Alruwaili & Mohamed, 2025) have trained EfficientNet-based and ResNet-based models on a set of ten skin disease classes, while Gulzar et al. (2025) have investigated DenseNet121 and EfficientNetB0 on another multi-class skin disease dataset. Because these studies use different databases and settings, a controlled evaluation using the same database, preprocessing, training, and evaluation may be warranted.

In this study, we compared three models, EfficientNetB0, ResNet50, and DenseNet121, to classify five skin diseases: Acne, Eczema, Psoriasis, Tinea, and Vitiligo using 4,357 images taken from a public Kaggle dataset. Previous research has compared the convolutional neural

network (CNN) architectures under different experimental conditions (Abbas et al., 2022; Sadik et al., 2023). The same pre-processing, augmentation, class-weighting, and two-phase transfer-learning methods were applied to all models. The accuracy, precision, recall, macro-F1 score, and confusion matrices were used to evaluate the performance. The controlled design may mitigate variation in datasets and configurations reported in previous studies, such as (Alruwaili & Mohamed, 2025; Gulzar et al., 2025).

2. Materials and Methods

2.1 Dataset Description

The analysis drew on a publicly available Skin Disease Dataset from Kaggle (Alruwaili & Mohamed, 2025; Gulzar et al., 2025). Recently, transfer learning CNNs have also been used for multi-class skin disease classification (Abdulqader & Abraham, 2026; Sarı & Keser, 2025). Five classes were chosen: Acne, Eczema, Psoriasis, Tinea, and Vitiligo. Images were tested for integrity, RGB format, image size, and class distribution (Abdulqader & Abraham, 2026; Eskandari & Sharbatdar, 2024). Although the file-level audit detected 4,358 images, Keras loaded 4,357 for model development and evaluation because one image failed to load.

Class	Train	Validation	Test	Total
Acne	466	118	65	649
Eczema	738	196	112	1,046
Psoriasis	630	161	88	879
Tinea	712	181	101	994
Vitiligo	567	141	82	790
Total	3,113	797	448	4,358

Table 1. File-level distribution of the selected skin-disease images

To prevent unrealistic performance estimates, duplicate images were identified, and data leakage was assessed (Abdulqader & Abraham, 2026; Abhishek et al., 2025).

2.2.Data Preprocessing

Readability and RGB format assessments were performed on images, followed by checks for duplicate images and cross-partition leakage to avoid inaccurate performance estimates (Abhishek et al., 2025; Choy et al., 2023). For EfficientNetB0, ResNet50, and DenseNet121, images were resized from 100×100 to 224×224 and kept as RGB. Prior studies of deep learning in skin and health appear to have employed standardized images with these dimensions (S. Abbas et al., 2025; Sun et al., 2024).

The main preprocessing procedures are briefly summarized in Table 2.

Stage	Procedure	Purpose
Image Verification	Readability and RGB checking	Ensure valid image files

Duplicate Detection	MD5 hashing	Identify duplicate images
Leakage Checking	Cross-partition comparison	Prevent data overlap
Resizing	224 × 224 pixels	Standardize model input
Model Preparation	Architecture-specific preprocessing	Prepare images for pretrained CNNs

Table 2. Data preprocessing procedures

The same set of preprocessed images was supplied to all three CNN architectures, thereby establishing a common basis for comparing their results.

2.3 Data Augmentation

Training images underwent augmentation through rotation, shifting, shearing, zooming, horizontal flipping, and brightness adjustment (85–115%) to increase diversity and reduce overfitting (Abbas et al., 2024; Gulzar et al., 2025; Mir et al., 2024). These methods may also help mitigate class imbalance (Anand et al., 2025).

Class weighting accounted for variation in class size (Anand et al., 2025): Acne 1.3356, Eczema 0.8434, Psoriasis 0.9879, Tinea 0.8754, and Vitiligo 1.0977. EfficientNetB0, ResNet50, and DenseNet121 were applied to the identical methods, with the same validation and test images reserved for independent evaluation (Gulzar et al., 2025).

2.4 Model Development using Transfer Learning

ImageNet-pretrained EfficientNetB0, ResNet50, and DenseNet121 were adapted for skin disease classification via transfer learning, thereby reusing learned features and potentially reducing training requirements (S. Abbas et al., 2025; G. Yang et al., 2025; Y.-C. Yang et al., 2025). EfficientNetB0 employs compound scaling, ResNet50 incorporates residual connections, and DenseNet121 enables feature reuse. These architectures have been demonstrated in the field of skin disease classification (Alruwaili & Mohamed, 2025; Gulzar et al., 2025a; Venkatesh et al., 2026; Wang & Chen, 2026).

Model	Pretrained weight	Pooling layer	Dense layer	Dropout layer	Output Layer
EfficientNetB0	ImageNet	Average Pooling	256 neurons, ReLU	0.4	5-class Softmax
ResNet50	ImageNet	Average Pooling	256 neurons, ReLU	0.4	5-class Softmax
DenseNet121	ImageNet	Average Pooling	256 neurons, ReLU	0.4	5-class Softmax

Table 3. Configuration of the Transfer-Learning Models

Average pooling was used instead of the original classification layers; a dense layer with 256 ReLU neurons replaced the original dense layer, a dropout layer with a rate of 0.4 replaced the original dropout layer, and a five-class Softmax output replaced the original classification layer. To make the models suitable for the target classes, the models were fine-tuned (Ahmmed et al., 2025; Khullar et al., 2025; Sharma et al., 2026; G. Yang et al., 2025).

2.5 Two-Phase Training Strategy

A two-phase transfer-learning strategy was implemented for EfficientNetB0, ResNet50, and DenseNet121. The classification head was initially frozen and trained for 10 epochs with an Adam learning rate of 1×10^{-3} , following the standard procedure of training the model with both layers (Bijoy et al., 2025; Song et al., 2025).

Training component	Phase 1: Frozen Base	Phase 2: Fine-Tuning
Base Layers	Frozen	Top 35% unfrozen
Epochs	10	10
Optimizer	Adam	Adam
Learning rate	1×10^{-3}	1×10^{-5}
Class weights	Applied	Applied
Early stopping	Patience = 5	Patience = 5
Learning-rate reduction	On plateau	On plateau
Checkpoint	Best validation accuracy	Best validation accuracy

Table 4. Two-Phase Training Configuration

The top 35% of layers were unfrozen and fine-tuned on the same machine for 10 epochs using the same learning rate of 1×10^{-5} , as has been done in dermatology (Turuta et al., 2026). In both phases, class weights, early stopping (patience = 5), and learning-rate reduction on plateau were applied. The checkpoint trained with the best validation accuracy was chosen for the final testing (Venkatachalam et al., 2025).

2.6 Model Evaluation

Accuracy, precision, recall, and F1-score (Aquil et al., 2025) were used to assess performance on the independent test set.

Metric	Formula	Purpose
Accuracy	$(TP + TN) / (TP + TN + FP + FN)$	Overall classification performance
Precision	$TP / (TP + FP)$	Correctness of prediction
Recall	$TP / (TP + FN)$	Ability to identify each class

F1-Score	$2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$	Balance between precision and recall
-----------------	---	--------------------------------------

Table 5. Performance Evaluation Metrics

These metrics were applied to all five classes: Acne, Eczema, Psoriasis, Tinea, and Vitiligo; confusion matrices revealed correct and incorrect predictions (Chaitan et al., 2026; Mousa et al., 2025). The study compared the models using Macro-F1; the architecture achieving the maximum score represents the best-performing model (Sonuç et al., 2026).

3. Results and Discussion

This section compares EfficientNetB0, ResNet50, and DenseNet121 for classifying five skin diseases using accuracy, precision, recall, F1-score, and confusion matrices (Dahiya et al., 2026; Mehboob et al., 2026; Noaman et al., 2025). Class-level analysis may identify challenging diseases and misclassification patterns (Wahyuni et al., 2026), while comparative studies may highlight performance differences across architectures and training approaches (Mehboob et al., 2026; Ozdemir & Pacal, 2025). SVM attained 53.8% accuracy, whereas Random Forest achieved 57.8% accuracy. Skin-image features (Sarı & Keser, 2025) have been extracted, and SVM has been used with these features, while classical classifiers have been investigated and compared with CNNs (Prayesy, 2025).

Model	Accuracy	Macro F1	Weighted F1
SVM (RBF, scaled features)	53.8%	53.9%	53.2%
Random Forest	57.8%	58.3%	57.7%

Table 6. Classical Machine-Learning Baseline Performance

Random Forest may have outperformed SVM by 4.0 percentage points. These methods, alongside deep-learning approaches, have been applied to skin-disease classification (K N & V, 2026; Prayesy, 2025).

4. Overall Model Comparison

ResNet50 attained the highest mean accuracy ($80.5\% \pm 0.9\%$) and macro-F1 ($81.8\% \pm 0.9\%$), followed by DenseNet121 and EfficientNetB0 across three training seeds. The convolutional neural network (CNN) architectures for skin-disease classification were compared similarly (N et al., 2026).

Model	Accuracy	Macro F1	Weighted F1
EfficientNetB0	$72.2\% \pm 2.8\%$	$73.8\% \pm 2.5\%$	$72.4\% \pm 2.7\%$
ResNet50	$80.5\% \pm 0.9\%$	$81.8\% \pm 0.9\%$	$80.7\% \pm 0.8\%$
DenseNet121	$74.6\% \pm 0.5\%$	$76.4\% \pm 0.5\%$	$75.0\% \pm 0.4\%$

Table 7. Overall Test-Set Performance across the Three Models

To compare pretrained CNN architectures, we used (Wahyuni et al., 2026; Firdous et al., 2026) several performance metrics. Table 3.2 summarizes overall test-set performance, while Figure 3.1 provides a comparative visualization of the three models.

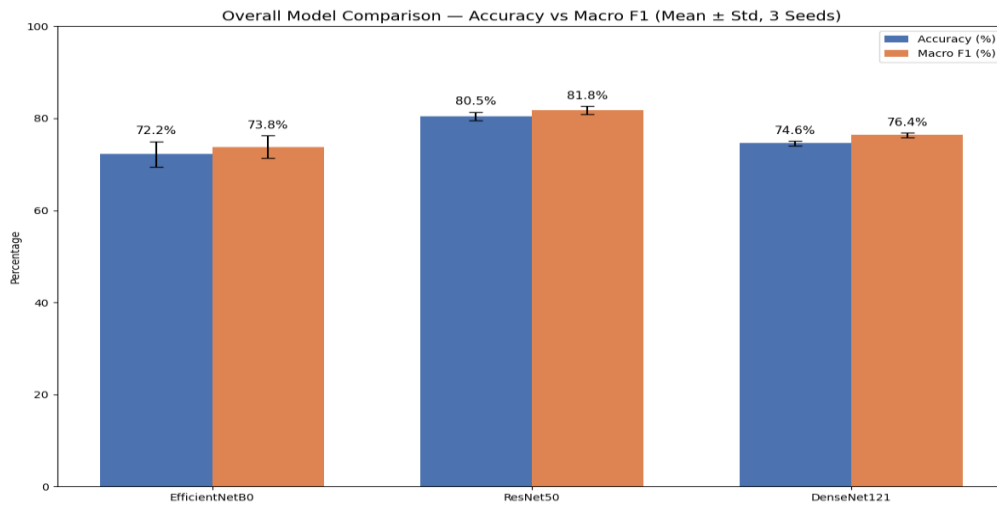


Figure 1. Overall Model Performance Comparison

Compared to EfficientNetB0 and DenseNet121, ResNet50 achieved the highest macro-F1 score with a small accuracy variance of $\pm 0.9\%$ across the three seeds, while DenseNet121 was within $\pm 2.8\%$ and EfficientNetB0 within $\pm 2.8\%$. The performance and stability are also considered, and these results are indicative of the importance of both (Wang & Chen, 2026)

5. Per-Class Performance Analysis

Performance appears to have differed across the three models and five skin-disease classes. The results are presented by class-wise precision, recall, and F1-scores to highlight the differences that may not be apparent from the overall accuracy (Badr et al., 2025; Noaman et al., 2025; G. Yang et al., 2025). Similarly, (Bijoy et al., 2025) reported evidence suggesting statistically significant differences among class-wise F1 scores.

Class	EfficientNetB0	ResNet50	DenseNet121
Acne	81.9%	90.6%	89.9%
Eczema	68.8%	75.2%	66.4%
Psoriasis	58.5%	71.8%	65.4%
Tinea	64.9%	76.1%	62.3%
Vitiligo	90.9%	93.6%	92.8%

Table 8. Per-Class F1-Score Comparison across the Three Transfer-Learning Models

As seen in Figure 3.2, the class-wise F1-score differences appear consistent with findings from other studies that analyzed the performance of the models on a set of classes that have similarities in terms of their image appearance and overlap (Bijoy et al., 2025; Noaman et al., 2025; G. Yang et al., 2025).

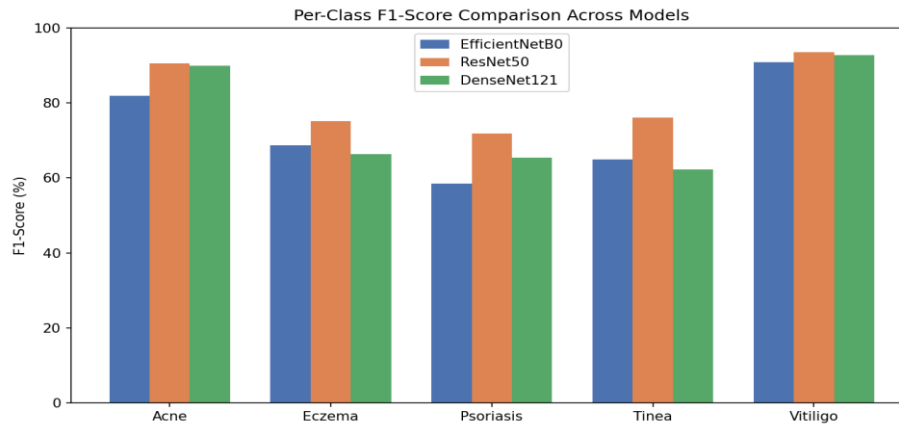


Figure 2. Per-Class F1-Score Comparison across Models

Class-level results show strengths and weaknesses in the model's performance other than the general accuracy. (Badr et al., 2025; Song et al., 2025) and reported on issue variations related to feature representation and transfer learning, respectively, and (Elhadidy et al., 2025) discussed the importance of misclassification analysis.

6. Confusion Matrix Analysis

The confusion-matrix analysis suggests that misclassification rates for Eczema, Psoriasis, and Tinea may have exceeded those for Acne and Vitiligo. The challenges with visually related inflammatory diseases were reported by (Eskandari & Sharbatdar, 2024; Hammad et al., 2023; Yodrabum et al. 2026).

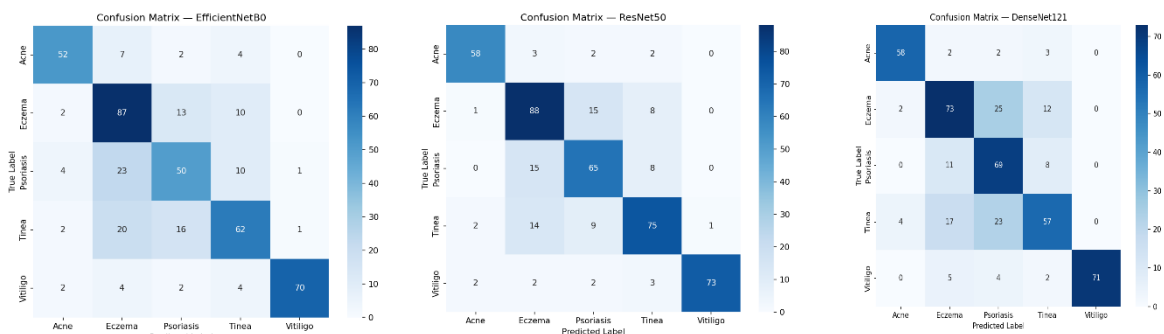


Figure 3(a): Confusion Matrix of EfficientNetB0

Figure 3(b): Confusion Matrix of ResNet50

Figure 3(c): Confusion Matrix of DenseNet121

EfficientNetB0 appears to have incorrectly classified 23 Psoriasis images as Eczema and 20 Tinea images as Eczema. Such errors may reflect visual similarity or limitations of transfer learning (Alwakid et al., 2026; Kharisma et al., 2025). ResNet50 appears to have made fewer errors, misclassifying 15 Psoriasis and 14 Tinea images as Eczema. Confusion-matrix analysis has been employed in the assessment of ResNet50 and the comparison of CNN architectures (Alruwaili & Mohamed, 2025; Kapoor et al., 2025; Mahmud et al., 2024). DenseNet121 also appears to have misclassified Eczema, Psoriasis, and Tinea, although it may have identified Acne and Vitiligo more accurately. (Raditya et al., 2025; Ravi, 2022), and (Gulzar et al., 2025)

also found similar differences in errors between classes. In this experiment, ResNet50 appears to have provided the strongest class-wise discrimination.

7 Training and Validation Behavior

ResNet50 appeared to exhibit the greatest training stability, as validation accuracy increased while loss declined and subsequently stabilized. The relevance of fine-tuning and validation monitoring has also been noted by (Das & Addya, 2025; Davila et al., 2024; Eliwa, 2025).

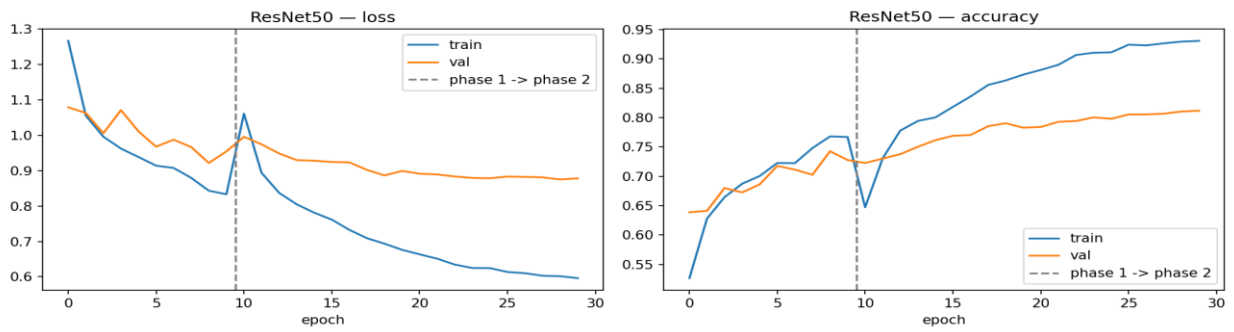


Figure 4. Training and validation curves for ResNet50

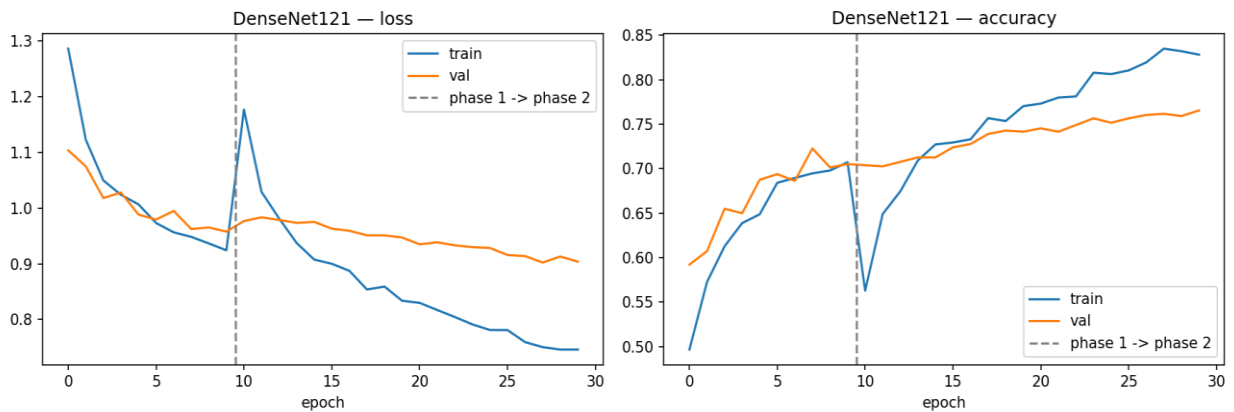


Figure 5. Training and validation curves — DenseNet121

DenseNet121 experienced a modest post-fine-tuning decline, followed by recovery, and EfficientNetB0 saw a more significant drop from ~76% in Phase 1 to 67% in Phase 2. (Shaik et al., 2025) noted the same type of differences between architectures.

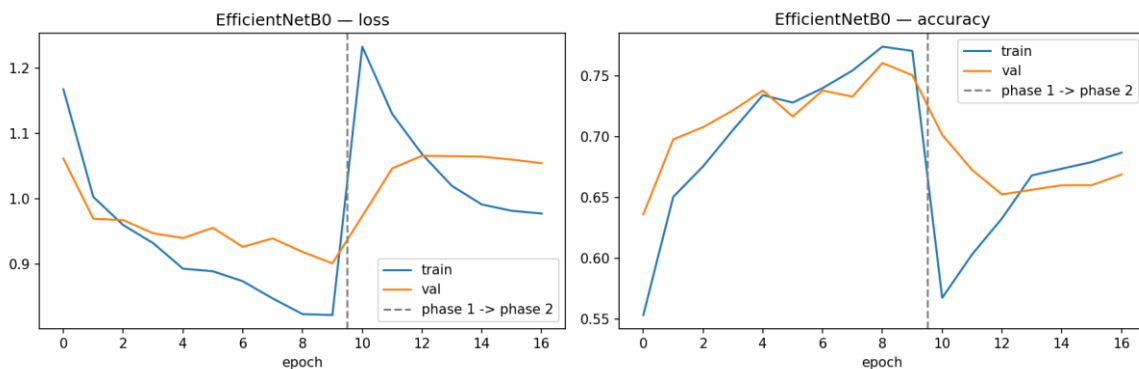


Figure 6. Training and validation curves for EfficientNetB0

The performance of a model may vary with training configuration, and validation monitoring is important (Gabani et al., 2026; Rajeshkumar & Chowdhary, 2026). Overall, ResNet50 achieved the best performance and stability, followed by DenseNet121, while EfficientNetB0 was less stable. This result aligns with ResNet50 for this dataset, though the results vary with the experimental setup (A et al., 2024; Mustafa et al., 2025).

8. Discussion of Findings

ResNet50 performed best with the highest accuracy (80.5%) and macro-F1 (81.8%) scores, thus discriminating better between five classes of skin diseases. Similarly, the performance of ResNet has been seen to be satisfactory in previous studies, and the difference in performance among CNN architectures has also been reported based on different datasets and training setups (AlSuwaidan, 2023; Kharisma et al., 2025; Oztel et al., 2023; Wang & Chen, 2026).

The classification of Eczema, Psoriasis, and Tinea was more challenging than the classification of Acne and Vitiligo due to their visual similarity and image representation (which is consistent with other studies reporting disease class differences based on visual similarity and image representation (Güler, 2025; Velasco et al., 2023)).

Although EfficientNetB0 is a very efficient network, it is not the most accurate, which verifies that lightweight networks are not always the optimal models to be used in CNN models (Oztel et al., 2023).

9. Conclusion

This study compared EfficientNetB0, ResNet50, and DenseNet121 using 4,357 images. Overall, ResNet50 appears to have performed best under the experimental conditions; however, further validation may be required to establish an F1 of $81.8\% \pm 0.9\%$. Acne and Vitiligo were classified more successfully than Eczema, Psoriasis, and Tinea. For the overall performance, ResNet50 performed best under the experimental conditions.

10. Recommendations and future work

more validation is External validation may also be required to assess whether these findings generalize to other datasets of different populations, skin tones, disease presentation, and image-acquisition conditions should be used in future studies. Testing outside of the confines of this study is also required to assess whether the findings apply to other data.

The use of additional training seeds in future comparisons may provide a more reliable estimate of performance variability, given that this study included only three, and Tinea model's size, memory usage, and inference speed may also warrant consideration alongside training seeds should be used in future comparisons to get a more reliable estimate of performance likely be required before their use in seeds. The size of the model from this selected image set should not be construed as establishing clinical reliability or readiness performance.

Validation of these models, followed by appropriate clinical evaluation, is required before they can be used for healthcare. The findings in this selected image set are not intended to demonstrate clinical reliability or preparedness for independent medical application.

Authors' Contribution

All authors contributed to the interpretation of the reviewed literature, critically revised the paper, approved the final version for publication and agreed to be held responsible for the entire contents.

Acknowledgements

This publication was the result of research conducted without external funding.

Conflicts of Interest

There are no conflicts of interest for the authors to disclose

References

- Allugunti, V. R. (2022). A machine learning model for skin disease classification using convolution neural network. *International Journal of Computing, Programming and Database Management*, 3(1), 141–147. <https://doi.org/10.33545/27076636.2022.v3.i1b.53>
- Alruwaili, M., & Mohamed, M. (2025). An Integrated Deep Learning Model with EfficientNet and ResNet for Accurate Multi-Class Skin Disease Classification. *Diagnostics*, 15(5). <https://doi.org/10.3390/diagnostics15050551>
- AlSuwaidan, L. (2023). Deep Learning Based Classification of Dermatological Disorders. *Biomedical Engineering and Computational Biology*, 14. <https://doi.org/10.1177/11795972221138470>
- Alwakid, G. N., Tehsin, S., Humayun, M., Farooq, A., Alrashdi, I., & Alsirhani, A. (2026). Graph Attention Networks for Skin Lesion Classification with CNN-Driven Node Features. *Computers, Materials and Continua*, 86(1). <https://doi.org/10.32604/cmc.2025.069162>
- Das, S., & Addya, R. K. (2025). Skin Cancer Detection and Classification Through Medical Image Analysis Using EfficientNet. *NDT*, 3(4), 23. <https://doi.org/10.3390/ndt3040023>
- Davila, A., Colan, J., & Hasegawa, Y. (2024). Comparison of fine-tuning strategies for transfer learning in medical image classification. *Image and Vision Computing*, 146, 105012. <https://doi.org/10.1016/j.imavis.2024.105012>
- Eliwa, E. H. I. (2025). Enhancing Skin Cancer Diagnosis Through Fine-Tuning of Pretrained Models: A Two-Phase Transfer Learning Approach. *International Journal of Breast Cancer*, 2025(1). <https://doi.org/10.1155/ijbc/4362941>
- Fraiwan, M., & Faouri, E. (2022). On the Automatic Detection and Classification of Skin Cancer Using Deep Transfer Learning. *Sensors*, 22(13). <https://doi.org/10.3390/s22134963>

- Gabani, V., Navamani, T. M., Shyamala, K., & Vaswani Rajpal, V. K. (2026). Multimodal skin lesion classification for early cancer diagnosis using deep learning. *Frontiers in Physiology*, 17. <https://doi.org/10.3389/fphys.2026.1717517>
- Güler, F. (2025). Investigation of Binary and Multiclass Classification Performance of Skin Cancer Images Using Transfer Learning Methods. *Journal of Clinical Practice and Research*, 47(3), 235–245. <https://doi.org/10.14744/cpr.2025.27623>
- Gulzar, Y., Agarwal, S., Soomro, S., Kandpal, M., Turaev, S., Onn, C. W., Saini, S., & Bounsiar, A. (2025). Next-generation approach to skin disorder prediction employing hybrid deep transfer learning. *Frontiers in Big Data*, 8. <https://doi.org/10.3389/fdata.2025.1503883>
- Kapoor, D., Gupta, D., Gulzar, Y., Alkanan, M., Mir, M. S., Soomro, A. B., Uppal, M., & Reshma, R. (2025). Convolutional Neural Network-based Multi-Classification of Skin Disease with Fine-Tuned ResNet50 and VGG16. *The Open Bioinformatics Journal*, 18(1). <https://doi.org/10.2174/0118750362387304250716055022>
- Kharisma, I. L., Munandar, R. A., Maulana, D. I., Rizky, M. N., & Kamdan. (2025). A Comparative Evaluation of CNN and Transfer Learning Models for Multi-Class Skin Disease Classification †. *Engineering Proceedings*, 107(1). <https://doi.org/10.3390/engproc2025107052>
- Mahmud, M. Z., Reza, M. S., Alve, S. R., Islam, S., & Fahmid, N. (2024). *Advance Transfer Learning Approach for Identification of Multiclass Skin Disease with LIME Explainable AI Technique*. <https://doi.org/10.1101/2024.12.02.24318311>
- Mustafa, S., Jaffar, A., Rashid, M., Akram, S., & Bhatti, S. M. (2025). Deep learning-based skin lesion analysis using hybrid ResUNet++ and modified AlexNet-Random Forest for enhanced segmentation and classification. *PLoS ONE*, 20(1). <https://doi.org/10.1371/journal.pone.0315120>
- Oztel, I., Yolcu Oztel, G., & Sahin, V. H. (2023). Deep Learning-Based Skin Diseases Classification using Smartphones. *Advanced Intelligent Systems*, 5(12). <https://doi.org/10.1002/aisy.202300211>
- Raditya, D. I., Slamet, I., Sulandari, W., Susanto, I., Zukhronah, E., & Isnaini, B. (2025). A Comparative Study of Densenet121 And Resnet50 With Grad-Cam For Skin Disease Classification for Improving Good Health. *Journal of Lifestyle and SDGs Review*, 5(7), e06526. <https://doi.org/10.47172/2965-730x.sdgsreview.v5.n07.pe06526>
- Rajeshkumar, S., & Chowdhary, C. L. (2026). Improved CNN with BiLSTM model for early melanoma and skin lesion classification. *Frontiers in Artificial Intelligence*, 9. <https://doi.org/10.3389/frai.2026.1808770>
- Ravi, V. (2022). Attention Cost-Sensitive Deep Learning-Based Approach for Skin Cancer Detection and Classification. *Cancers*, 14(23). <https://doi.org/10.3390/cancers14235872>

- Sadik, R., Majumder, A., Biswas, A. A., Ahammad, B., & Rahman, M. M. (2023). An in-depth analysis of Convolutional Neural Network architectures with transfer learning for skin disease diagnosis. *Healthcare Analytics*, 3. <https://doi.org/10.1016/j.health.2023.100143>
- Saifan, R., & Jubair, F. (2022). Six skin diseases classification using deep convolutional neural network. *International Journal of Electrical and Computer Engineering*, 12(3), 3072–3082. <https://doi.org/10.11591/ijece.v12i3.pp3072-3082>
- Shaik, A., Dutta, S. S., Sawant, I. M., Kumar, S., Balasundaram, A., & De, K. (2025). An attention based hybrid approach using CNN and BiLSTM for improved skin lesion classification. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-00025-2>
- Velasco, J. S., Catipon, J. V., Monilar, E. G., Amon, V. M., Virrey, G. C., & Tolentino, L. K. S. (2023). Classification of Skin Disease Using Transfer Learning in Convolutional Neural Networks. *International Journal of Emerging Technology and Advanced Engineering*, 13(4), 1–7. https://doi.org/10.46338/ijetae0423_01
- Wang, D., & Chen, Y. (2026). Comparative Study of Transfer Learning Strategies for Multi-Class Skin Lesion Classification: Architectures, Fine-Tuning, and Data Augmentation. *ICCK Journal of Image Analysis and Processing*, 2(3), 153–167. <https://doi.org/10.62762/jiap.2026.390206>