JOIN (Jurnal Online Informatika)
p-ISSN: 2528-1682, e-ISSN: 2527-9165
Volume 11 Number 1 | June 2026: 98-112
DOI: 10.15575/join.v11i1.1758

Comparative Analysis of Pre-Trained Deep Learning Models for Classifying Tropical Fungal Skin Infections

Suhardi Aras[1], Muhammad Rizwan Darwis[2], Muhammad Zhaky Arkan[3]
[1,2,3]Department of Informatics, Faculty of Engineering, Universitas Muhammadiyah Sorong, Indonesia

Article Info

Article history:

Received July 15, 2025
Revised September 29, 2025
Accepted October 22, 2025
Published April 25, 2026

Keywords:

Deep Learning
Fungal Skin Infections
Medical Image Classification
MobileNetV3
Tropical Skin Infections

ABSTRACT

Tropical fungal skin infections, including Tinea corporis, Tinea versicolor, Tinea pedis, and Tinea nigra, are common health problems in tropical countries such as Indonesia. Although not life-threatening, these diseases can cause discomfort, reduce self-confidence, and interfere with daily activities. Conventional diagnostic methods still rely on subjective visual observation, which is often inaccurate—especially in regions with limited infrastructure and scarce access to specialists. Moreover, existing studies rarely provide a comparative evaluation of deep learning architectures for tropical fungal infections using small and diverse datasets. Therefore, this study aims to address these challenges by conducting a systematic comparative evaluation of three pre-trained models—MobileNetV3, EfficientNet-B2, and SE-ResNet101—to determine the most accurate and computationally efficient architecture for multi-class classification of tropical fungal skin diseases. In this study, a dataset of 660 clinical skin images sourced from the Kaggle repository was used, covering four tropical fungal infection classes. The dataset consisted of 165 images per class, which were divided into training, validation, and testing subsets. Experimental results demonstrated that MobileNetV3 achieved the best performance, with a validation accuracy of 95.08%, a test accuracy of 97.34%, and the shortest training time of 15 minutes, compared to EfficientNet-B2 (16 minutes) and SE-ResNet101 (22 minutes). The main contribution of this study is to provide a systematic comparative evaluation of deep learning models for the classification of tropical fungal skin infections, while recommending MobileNetV3 as the most suitable model for practical implementation of automated image-based diagnosis in primary healthcare services with limited resources.

Corresponding Author:

Muhammad Rizwan Darwis,
Department of Informatics, Faculty of Engineering, Universitas Muhammadiyah Sorong
Jalan Pendidikan Km 8, Kota Sorong Papua Barat Daya, Indonesia.
Email: m,rizwan.darwiss@gmail.com

1. INTRODUCTION

Indonesia is a tropical country with consistently high temperatures and humidity throughout the year [1]. These conditions strongly support the growth of various pathogenic microorganisms, including fungi that cause skin infections. Fungal skin infections represent one of the most significant public health issues in tropical regions, with approximately 1.73 billion cases reported globally in 2021 [2]. This condition is further supported by WHO data, which estimates that around 1.8 billion people experience skin-related problems, including fungal infections, at any given time [3]. In Indonesia, the 2018 Basic Health Research (Riskesdas) reported more than 10 million cases of skin diseases annually, most of which are caused by superficial fungal infections [4]. Common types of infections include Tinea corporis (ringworm), Tinea versicolor, Tinea pedis (athlete's foot), and Tinea nigra. These diseases are

contagious and can affect anyone regardless of social status or physical condition. Although not classified as life-threatening, fungal skin infections can cause discomfort, reduce self-confidence, interfere with daily activities, and tend to recur if not properly treated. Therefore, this issue deserves further attention in the context of public health in tropical regions, particularly to support early detection efforts and the development of more effective and sustainable disease management strategies

In many tropical regions, fungal skin infections are among the most commonly encountered cases at primary healthcare facilities [5]. However, limited medical infrastructure and restricted access to dermatology specialists often result in diagnoses being conducted by general healthcare workers or even self-diagnosed by patients. In such circumstances, disease identification relies heavily on visual observation, which is not always accurate—especially when symptoms appear atypical or resemble other skin conditions [6]. These challenges highlight the need for technological support that can facilitate the classification of fungal skin infections in a more objective, rapid, and reliable manner across various healthcare settings.

Although deep learning technology has demonstrated great potential in medical image classification, the classification of fungal skin infections in tropical regions still presents distinct challenges. Several types of infections, such as Tinea corporis, Tinea versicolor, Tinea pedis, and Tinea nigra, share certain visual similarities that may confuse automated classification systems if not properly trained. In addition, the use of small-scale datasets compiled from various image sources amplifies these challenges, as the data may not fully represent the visual characteristics of tropical skin conditions. Therefore, a comprehensive evaluation of several pretrained deep learning models is necessary to assess their effectiveness and efficiency in performing multiclass classification on limited and diverse image data from tropical environments [7], [8].

Deep learning-based research for the classification of fungal skin infections has been explored in various studies. Alruwaili et al. proposed a hybrid model combining EfficientNet and ResNet for multi-class skin disease classification, achieving a high accuracy of 98.67% [9]. Thohari et al. conducted a comparison of five pre-trained CNN models—VGGNet16, MobileNetV2, InceptionResNetV2, ResNet152V2, and DenseNet201—for the classification of three common types of skin infections, with ResNet152V2 achieving the highest accuracy of 95.84% [10]. Alyas et al. compared five pre-trained CNN models—VGGNet16, MobileNetV2, InceptionResNetV2, ResNet152V2, and DenseNet201—for classifying three types of common skin infections, where ResNet152V2 achieved the best performance with an accuracy of 95.84% [11], However, the study only classified two types of fungal infections and did not implement a multi-class classification approach, which is essential for addressing the broader variety of tropical skin diseases. Tsedenya et al. employed a CNN model that achieved an accuracy of 83% [12], However, it was limited to a single model and did not evaluate broader visual variability, which may restrict the model's generalization capability in the context of highly diverse tropical skin diseases. Similarly, Debebe et al. developed a CNN model for skin infection classification [13], but it was limited to a single model type and did not cover various types of tropical fungal skin infections. Furthermore, the model was not evaluated in the context of broader visual diversity. Therefore, a study is needed that systematically compares the performance of several pre-trained deep learning architectures, particularly in the context of multi-class classification of tropical fungal skin diseases that exhibit highly diverse visual characteristics.

To address the challenges in classifying tropical fungal skin infections, this study focuses on the multi-class classification of four infection types—Tinea corporis, Tinea versicolor, Tinea pedis, and Tinea nigra—while simultaneously performing a comparative evaluation of three pre-trained deep learning architectures: MobileNetV3, EfficientNet-B2, and SE-ResNet101. Unlike previous studies that typically considered only two or three classes or overlooked computational efficiency, this research adopts a more comprehensive approach by evaluating both classification accuracy and training efficiency on diverse tropical image data. The main contribution of this work is to identify the most accurate yet lightweight and efficient model, making it suitable for deployment in primary healthcare facilities with limited infrastructure**.** The evaluation results are expected to provide valuable insights for the development of adaptive, automated, image-based diagnostic systems tailored to the needs of tropical regions.

2. METHOD

This study applies a deep learning approach to develop an image classification system, focusing on the performance analysis of three pre-trained CNN architectures: MobileNetV3, EfficientNet-B2, and

p-ISSN: 2528-1682
e-ISSN: 2527-9165

SE-ResNet101. The methodological steps used in this study are illustrated in the flowchart below. The process begins with importing the image dataset, which is then divided into three subsets: training data, validation data, and testing data. Data augmentation is then performed to enrich image variation, followed by image preprocessing before being used for model training. The deep learning model architectures are then constructed and trained using the prepared data. The training process is visualized to monitor model performance, after which the trained model is tested using the test data. Evaluation is conducted to quantitatively assess the model's performance [14], and in the final stage, a comparative analysis of the model architectures is conducted to identify the most effective and optimal model for the classification task.

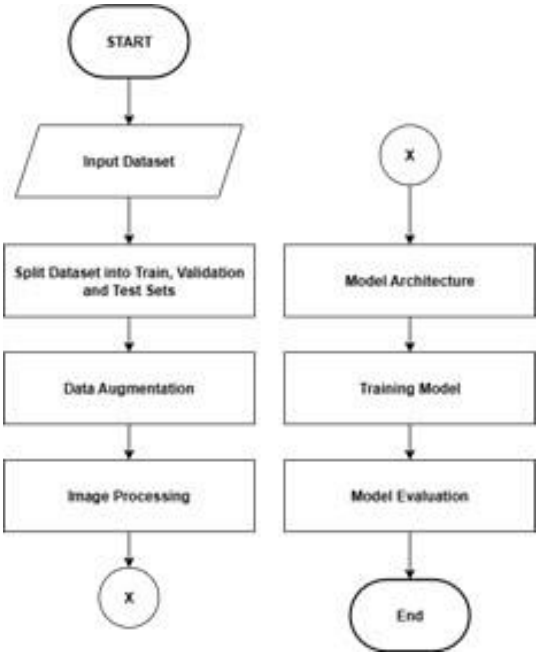
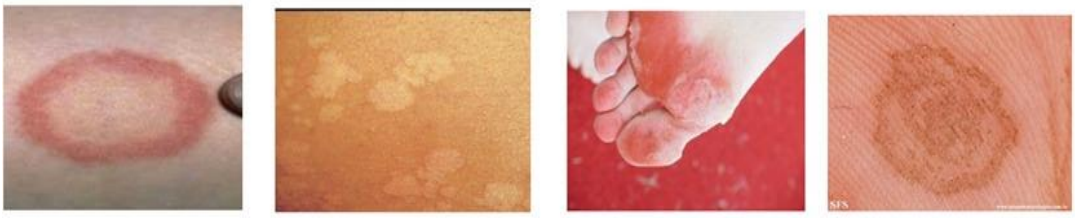


Figure 1. Research Flowchart

2.1. Dataset

The dataset used in this study was obtained from the Kaggle platform and consists of 660 skin images that have been categorized into four classes of tropical fungal skin infections: Tinea corporis (ringworm), Tinea versicolor, Tinea pedis (athlete's foot), and Tinea nigra. The classification was performed based on the distinctive visual characteristics of each disease.



(a) (b) (c) (d)

Figure 2. Sample Dataset ; (a) Tinea corporis, (b) Tinea versicolor, (c) Tinea pedis, (d) Tinea nigra.

Figure 2 presents sample images from each class. Tinea corporis (ringworm) is characterized by circular, reddish lesions [15]; Tinea versicolor appears as white patches that spread across the skin [16] ; Tinea pedis (athlete's foot) commonly occurs between the toes, presenting as cracked or blistered

skin [17]; whereas Tinea nigra manifests as dark patches typically found on the palms or soles [18]. All images were then divided into three subsets: 70% for training, 15% for validation, and 15% for testing, in order to ensure optimal model training while maintaining class balance across the dataset. Since the dataset size is relatively small for deep learning training, the model tends to memorize patterns in the training data excessively (*memorization*), which reduces its generalization capability and leads to *overfitting*. Therefore, data augmentation techniques were applied in the next stage to increase both the quantity and diversity of the training sample.

2.2. Augmentation

Data augmentation is a technique used to artificially expand a dataset by generating new variations of existing images [19]. This approach reduces overfitting and enhances the generalization ability of pretrained deep learning models. The initial dataset contained 660 images (165 per class) for four classes: Tinea corporis, Tinea versicolor, Tinea pedis, and Tinea nigra. Data augmentation was applied because the initial dataset was relatively limited in size. To overcome this limitation and improve generalizationTo increase both quantity and diversity, several augmentation techniques were applied. Horizontal flipping (probability 0.5) made the model invariant to image orientation, rotation up to 15° improved robustness against rotated objects, and color adjustments (brightness ±15%, contrast 0.2, saturation 0.2, hue 0.2) using ColorJitter enhanced generalization to varying lighting conditions. Additionally, resizing to 224 × 224 pixels standardized input dimensions, and RGB normalization (mean = [0.485, 0.456, 0.406]; std = [0.229, 0.224, 0.225]) ensured pixel values were suitable for pretrained networks. After augmentation, the dataset expanded to 1,980 images, providing more diverse training samples for improved model performance. The detailed augmentation settings are shown in Table 1

Table 1. Augmentation and Preprocessing Settings for Model Training

Augmentation	Value
Horizontal Flip	Probability 0.5 (random)
Brightness Adjustment	Maximum ±15°
Contrast Adjustment	0.2
Saturation Adjustment	0.2
Hue Adjustment	0.2
Resize	224 × 224
Rotation	0.2
Normalization (RGB)	mean = [0.485, 0.456, 0.406]; std = [0.229, 0.224, 0.225]

2.3. Image Processing

The image processing stage was carried out after the augmentation process to prepare the images in accordance with the standard input format required by deep learning models. This process included resizing the images, converting them to RGB color format, and normalizing the pixel values [20]. Specifically, the images were resized to 224×224 pixels, converted to tensor format using transforms.ToTensor(), and normalized based on the statistical distribution of the ImageNet dataset, with mean = [0.485, 0.456, 0.406] and standard deviation = [0.229, 0.224, 0.225]. This step aims to ensure data compatibility with the architecture of the pre-trained models used, as well as to maintain the stability of input distribution during the training process. Resizing ensures uniform input dimensions across the dataset, conversion to RGB preserves color information critical for skin lesion recognition, while normalization aligns input statistics with those of pretrained models (ImageNet), thus stabilizing training and accelerating convergence.

2.4. Model Architecture

After the data preprocessing stage was completed, the next step in this study was to build and train deep learning models based on Convolutional Neural Networks (CNNs). Three pre-trained architectures were selected for performance comparison: EfficientNet-B2, SE-ResNet101, and MobileNetV3. These models have been widely used in image classification tasks due to their ability to achieve high accuracy with efficient computational complexity. The selection of these models also considered their suitability for deployment in systems with limited resources, such as those commonly found in primary healthcare settings in tropical regions. A comparative evaluation was conducted to determine which model performs most optimally in handling multi-class classification of tropical fungal skin infection.

Table 2 provides a summary of the three pre-trained architectures employed in this study. EfficientNet-B2 achieves a balance between accuracy and efficiency through compound scaling of depth, width, and resolution. SE-ResNet101 enhances deep residual learning with squeeze-and-excitation modules, allowing better feature selection at the cost of higher complexity. MobileNetV3 integrates inverted residual blocks and h-swish activation, resulting in a lightweight architecture suitable for resource-constrained environments. These complementary characteristics justify their selection for fair performance comparison in tropical fungal skin disease classification. In this study, all three models were fine-tuned on the same dataset, consisting of 1,980 images across four classes of tropical fungal skin infections, under identical training configurations. This ensures a fair comparison of their performance in terms of classification accuracy, efficiency, and generalization capability.

Table 2. Summary of Pre-trained Deep Learning Models

Model	Input Size	Number of Parameters	Core Architecture / Building Block	Key Characteristics
EfficientNet-B2	260×260	~9.2M	MBConv blocks, compound scaling (depth, width, resolution) [21]	High accuracy with efficient parameter utilization, balanced scaling
SE-ResNet101	224×224	~44.5M	Residual blocks + Squeeze-and-Excitation modules	Strongfeature representation, but computationally heavy
MobileNetV3	224×224	~5.4M	Inverted residual blocks, depthwise separable convolution, h-swish activation [22]	Lightweight and fast, optimized for resource-constrained devices

2.4.1. EfficientNet-B2

EfficientNet-B2 is a pre-trained model composed of multiple convolutional blocks optimized using the compound scaling technique, which proportionally balances network depth, width, and input resolution[23]. In general, the EfficientNet-B2 architecture is divided into two main components: the feature extraction layers and the final fully connected classification layer. The feature extraction layers are responsible for identifying visual patterns from input images and transforming them into one-dimensional vector representations, while the fully connected layer learns from these representations to perform the final classification.

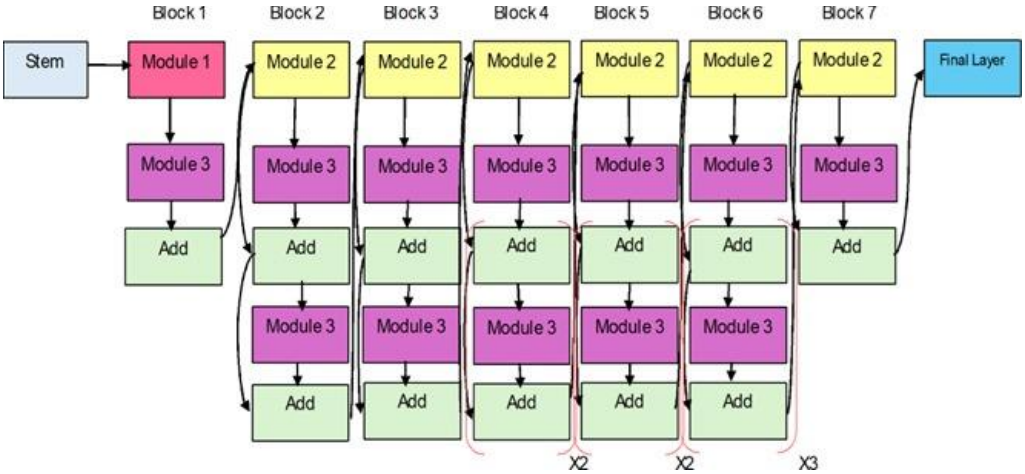


Figure 3.EfficientNet-B2 Architectur

By default, EfficientNet-B2 employs an input resolution of 260×260 pixels. However, in this study the input resolution was standardized to 224×224 pixels to ensure consistency with the other architectures (SE-ResNet101 and MobileNetV3) and to maintain fairness in performance comparison.

With its high parameter efficiency, EfficientNet-B2 is capable of achieving strong accuracy while maintaining relatively low computational cost compared to conventional models.

2.4.2. SE-ResNet101

SE-ResNet101 is an extension of the ResNet (Residual Neural Network) architecture, enhanced with a Squeeze-and-Excitation (SE) module. This model belongs to the category of pre-trained CNNs and has been widely applied in various computer vision tasks, such as image classification, object detection, and semantic segmentation[24]. The main advantage of ResNet lies in its ability to train very deep networks through residual connections, which help mitigate the vanishing gradient problem [25]. SE-ResNet101 improves this architecture by incorporating an attention mechanism through the SE module, enabling the network to assign greater importance to informative features within each channel.

In this study, SE-ResNet101 was implemented with an input size of 224×224 pixels to maintain consistency with the other architectures (EfficientNet-B2 and MobileNetV3). Although relatively complex and containing a large number of parameters, the inclusion of the SE module enhanced classification performance and helped reduce the risk of overfitting when applied to a limited dataset, as in this study.

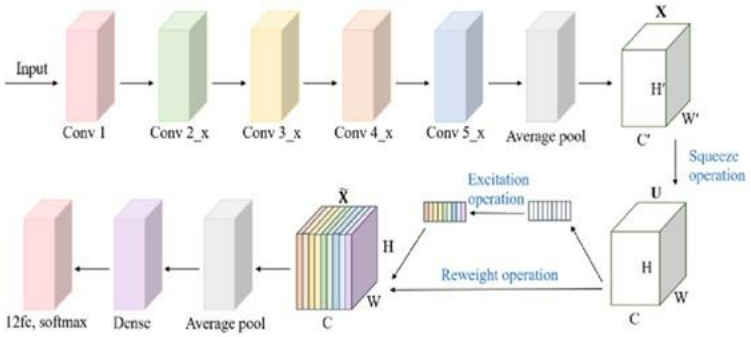


Figure 4. Architecture of the proposed SE- ResNet101. [26]

2.4.3. MobileNet V3

MobileNetV3 is an improved version of its predecessors, MobileNetV1 and V2, combining manual design approaches with Neural Architecture Search (NAS) techniques. This architecture introduces several new components, such as the h-swish activation function, Squeeze-and-Excitation (SE) modules, and more efficient inverted residual blocks. MobileNetV3 is specifically designed to deliver high performance on devices with limited computational resources, such as smartphones or embedded systems. Due to its architectural efficiency, this model is highly suitable for fast and lightweight image classification tasks. Compared to its predecessors, MobileNetV3 offers a better balance between accuracy and training time, making it well-suited for implementation in resource-constrained environments. In this study, MobileNetV3 was also implemented with an input size of 224×224 pixels to ensure consistency with the other models.

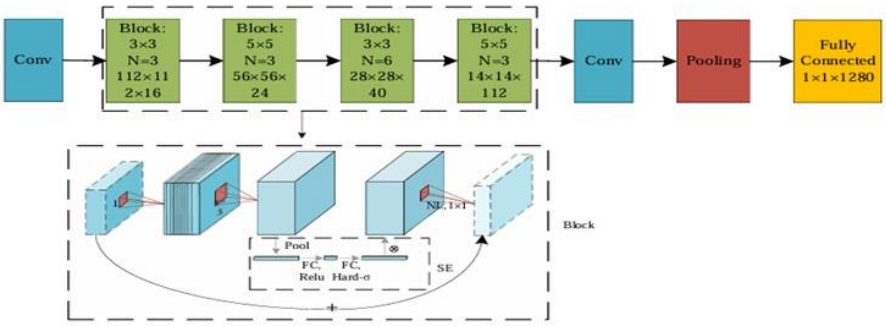


Figure 5. MobileNet V3 Architecture [27]

p-ISSN: 2528-1682
e-ISSN: 2527-9165

2.5 Training model

After the three pre-trained deep learning models, namely EfficientNet-B2, SE-ResNet101, and MobileNetV3, were constructed, the next stage was the training process using the preprocessed training data. All models were fine-tuned under the same configuration to ensure a fair comparison. Training was conducted for 50 epochs using a GPU device, which significantly accelerated computation and enabled large-scale image processing. The training procedure included the forward pass to generate predictions, the computation of the loss value using CrossEntropyLoss, and backpropagation to update model weights based on error gradients. The Adam optimizer was employed for its ability to accelerate and stabilize convergence.

The training parameters were determined based on preliminary experiments. The learning rate was set to 0.001 as it provided stable convergence. Smaller values (e.g., 0.0001) slowed down convergence, whereas larger values (e.g., 0.01) caused instability in the training curves. The batch size was set to 32, as this configuration offered the best balance between GPU memory efficiency and gradient stability. Smaller batch sizes resulted in noisier gradients, while larger batch sizes were computationally inefficient. The number of epochs was set to 50 after testing variations of 30, 50, and 100 epochs. The results showed that models suffered from underfitting at 30 epochs, while extending training to 100 epochs provided only marginal improvements compared to 50 epochs and increased the risk of overfitting. Therefore, 50 epochs were selected as the most optimal configuration, consistent with the accuracy and loss curves that demonstrated stable convergence at this point.

During training, accuracy and loss were periodically evaluated on both the training and validation datasets to monitor performance and detect possible overfitting. After training was completed, the results were visualized in the form of accuracy and loss curves across epochs. These curves confirmed that the chosen hyperparameter configuration produced a stable training process: accuracy increased significantly until around the 50th epoch before plateauing, while the loss consistently decreased without divergence. Hence, the training configuration was proven to support optimal performance, which is further analyzed in the Results *and* Discussion section.

Table 3. Training Value

Parameter	Value
Optimizer	Adam
Learning Rate	0.001
Loss Function	CrossEntropyLoss
Batch Size	32
Epochs	50
Input Size	224 × 224
Device	GPU

2.6. Model Evaluation

The trained model was then evaluated using the test set to measure its actual performance on previously unseen data. The evaluation involved calculating several classification performance metrics, including accuracy, precision, recall, and F1-score, as well as constructing a confusion matrix to assess the model's prediction accuracy for each class in greater detail. The confusion matrix consists of four main components: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), which serve as the basis for calculating the evaluation metrics. The formulas used in this evaluation are as follows:

$$\text{Accuracy} = \frac{(TP+TN}{(TN+FP+FN+TN} \quad (1)$$

$$\text{Precission} = \frac{TP}{(TN+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{(TP+FN} \quad (3)$$

$$\text{F1 Score} = 2x\frac{Recall\ x\ precision}{Recall+\ precision} \quad (4)$$

In this study, a performance comparison was conducted between three pre-trained deep learning architectures: EfficientNet-B2, SE-ResNet101, and MobileNetV3. The evaluation focused on multi-class classification accuracy, training time efficiency, and classification stability across classes based on metrics such as the F1-score distribution. This stage aimed to identify a model that not only achieves high accuracy but also offers computational efficiency and reliability in handling the diverse visual characteristics of tropical fungal skin infections.

3. RESULT AND DISCUSSION

All models used in this study—EfficientNet-B2, SE-ResNet101, and MobileNetV3—were trained using uniform parameters, with 50 epochs, a batch size of 32, and an image resolution of 224×224 pixels. Based on the training history, all three models exhibited relatively low accuracy during the initial epochs. However, starting around the 15th epoch, the performance of each model began to show significant improvement. MobileNetV3 and EfficientNet-B2 tended to achieve faster and more stable accuracy gains compared to SE-ResNet101, which showed a slightly slower progression. Nevertheless, none of the models exhibited signs of overfitting, as accuracy continued to improve gradually until the end of training without any drastic drop in validation performance.

Figure 6 below presents the loss and accuracy curves of the EfficientNet-B2 model over 50 training epochs. At the beginning of training, the model's accuracy was relatively low and the loss value was quite high; however, performance improved rapidly within the first 10–15 epochs. The accuracy curves show a stable upward trend and gradually converged toward the end of training, with validation accuracy remaining close to the training accuracy throughout most epochs. The small gap between the two curves indicates that the model achieved good generalization. Starting around epoch 35, the validation loss began to plateau while the training loss continued to decrease slightly, but this had no significant impact on the model's final performance. This is supported by the consistently high validation accuracy, which peaked at around 96% and remained stable at approximately 94% by the end of training. Overall, EfficientNet-B2 demonstrated stable and reliable classification performance, confirming its suitability for practical diagnostic applications.

Figure 7 below presents the loss and accuracy curves of the SE-ResNet101 model over 50 training epochs. During the early stages of training, the model's performance was relatively low; however, a significant improvement was observed within the first 10 to 15 epochs. The accuracy curves indicate that both training and validation accuracy increased progressively and began to stabilize toward the end of training, with a relatively small difference between the two. However, the loss curve shows fluctuations in validation loss starting from around epoch 25 until the end of training. Despite this, the performance gap was not substantial, and the validation accuracy peaked at approximately 96% and remained stable at 93.19% by the end of training. Overall, SE-ResNet101 demonstrated strong classification capability, although further attention may be required regarding validation stability during the final stages of training.

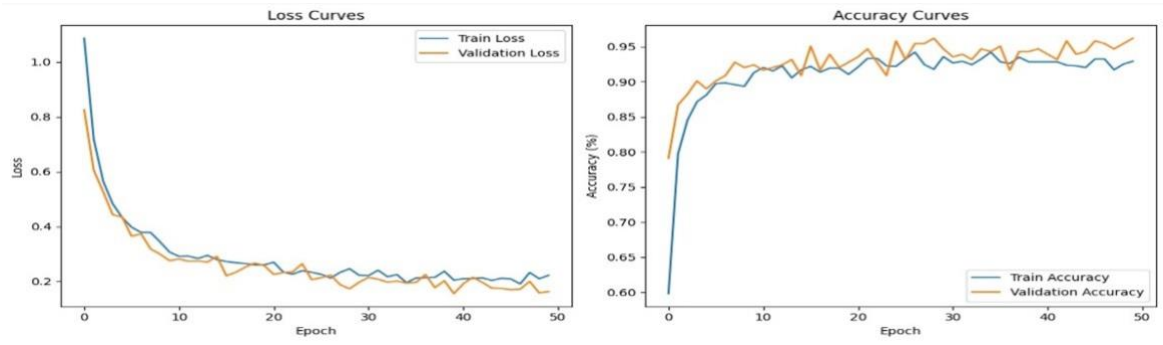


Figure 6. Training performance graph of EfficientNet-B2 based on accuracy and loss.

p-ISSN: 2528-1682
e-ISSN: 2527-9165

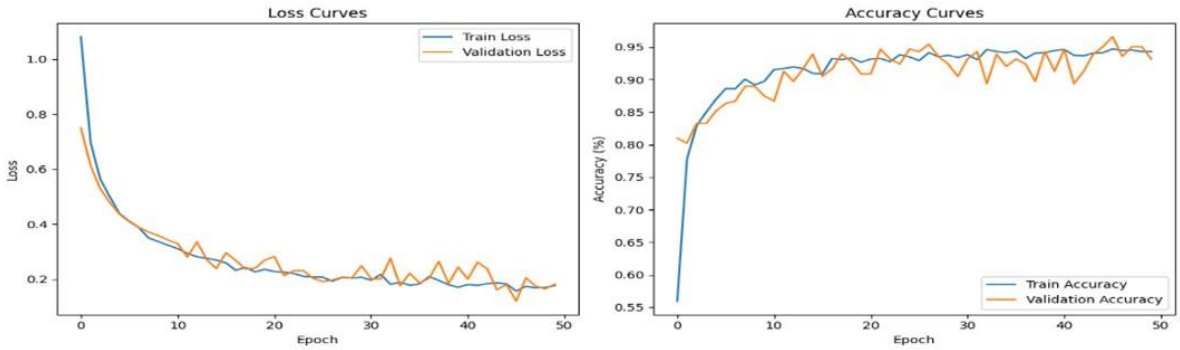


Figure 7. Training performance graph of SE Resnet 101 based on accuracy and loss

Figure 8 below shows the loss and accuracy curves during the training process of the MobileNetV3 model over 50 epochs. At the beginning of training, the accuracy was still below 70%, and the loss value was relatively high; however, a rapid performance improvement occurred within the first 10 epochs. The accuracy curve demonstrates a stable and consistent trend from epoch 15 until the end of training, where validation accuracy continued to increase with minimal fluctuations and consistently remained above the training accuracy. This indicates that the model did not experience overfitting, but instead was able to generalize very well to the validation data. The loss patterns further support this, as both training and validation loss showed a parallel decline with no significant divergence. With a stable trend and a very small performance gap between training and validation, it can be concluded that MobileNetV3 is an efficient and reliable model for classifying tropical skin disease images, as evidenced by its high validation accuracy of 95.08%.

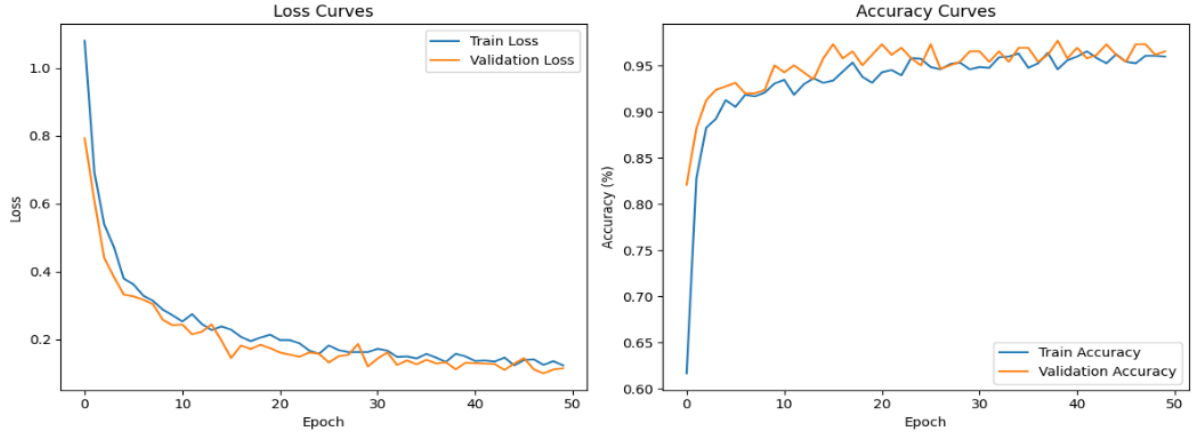


Figure 8. Training performance graph of MobileNet V3 based on accuracy and loss

Based on the overall evaluation using the model.evaluate function, EfficientNet-B2 achieved a test accuracy of 94.32%. Furthermore, the classification report presented in *Table 3* shows an overall accuracy of 89%, indicating balanced and consistent performance across all classes. The best performance was achieved in the *Tinea pedis (Athlete's Foot)* class, with an F1-score of 0.93, followed by *Tinea corporis (Ringworm)* with 0.89. The other two classes, *Tinea versicolor* and *Tinea nigra*, exhibited slightly lower but still acceptable performance, with F1-scores of 0.85 and 0.86, respectively—likely due to visual similarities between these classes. Overall, these findings demonstrate that EfficientNet-B2 is capable of consistently classifying the four types of tropical fungal skin infections with high accuracy, making it a reliable and efficient model for image-based diagnostic applications in primary healthcare settings.

Table 4. EfficientNet-B2 classification report

	precision	recall	f1-score	support
Tinea Corporis	0.90	0.88	0.89	82
Tinea Pedis	0.94	0.93	0.93	69
Tinea Versicolor	0.86	0.84	0.85	58
Tinea Nigra	0.83	0.89	0.86	55
Accuracy			0.89	264
Macro Avg	0.88	0.89	0.88	264
Weighted Avg	0.89	0.89	0.89	264

Based on the classification report presented in Table 4 below, the model achieved an overall accuracy of 94% on the test set, which consisted of 264 images. The best performance was recorded in the *Tinea pedis* class, with a precision of 0.99, recall of 1.00, and an F1-score of 0.99, indicating the model's ability to detect all cases in this class with minimal error. The *Tinea corporis* and *Tinea versicolor* classes also demonstrated strong performance, each with an F1-score of 0.93. In the *Tinea corporis* class, the high precision (0.96) combined with a recall of 0.91 suggests that a small number of cases remained undetected. Meanwhile, the *Tinea versicolor* class had the second-highest recall (0.97), despite a lower precision (0.90), indicating the potential for misclassification with other classes. The lowest performance was observed in the *Tinea nigra* class, with an F1-score of 0.89. Although the precision was fairly high (0.91), the slightly lower recall (0.88) implies that the model had some difficulty capturing the full visual variability of this class. Overall, both the macro and weighted average F1-scores reached 0.94, reinforcing the model's ability to perform consistent and accurate multi-class classification across the evaluated tropical fungal skin infections.

Table 5. SeResnet 101 classification report

	precision	recall	f1-score	support
Tinea Corporis	0.96.	0.91	0.93	75
Tinea Pedis	0.99	1.00	0.99	75
Tinea Versicolor	0.90	0.97	0.93	66
Tinea Nigra	0.91	0.88	0.89	48
Accuracy			0.94	264
Macro Avg	0.94	0.94	0.94	264
Weighted Avg	0.94	0.94	0.94	264

Based on the classification report shown in Table 5 below, the MobileNetV3 model achieved an overall accuracy of 97% on the test dataset consisting of 264 images. The best performance was observed in the *Tinea pedis* and *Tinea nigra* classes, with F1-scores of 0.99 and 0.98, respectively, reflecting highly accurate and stable detection capabilities. The *Tinea versicolor* and *Tinea corporis* classes also demonstrated excellent performance, with F1-scores of 0.96 and 0.95, respectively. The consistently high precision and recall across all classes indicate that the model is capable of performing multi-class classification effectively without exhibiting bias toward any particular class. Both the macro and weighted average F1-scores reached 0.97, confirming that MobileNetV3 is the most reliable and efficient model for classifying tropical fungal skin infections in this study.

Table 6. MobileNet V3 classification report

	precision	recall	f1-score	support
Tinea Corporis	0.95	0.96	0.95	77
Tinea Pedis	0.99	0.99	0.99	80
Tinea Versicolor	0.97	0.95	0.96	62
Tinea Nigra	0.98	0.98	0.98	45
Accuracy			0.97	264
Macro Avg	0.97	0.97	0.97	264
Weighted Avg	0.97	0.97	0.97	264

Based on Figure 9, the EfficientNet-B2 model correctly classified most of the test samples, as indicated by the dominance of values along the main diagonal of the confusion matrix. The *Tinea pedis* class exhibited the highest accuracy, with 64 out of 69 images correctly predicted. Meanwhile, the *Tinea corporis* and *Tinea nigra* classes also showed good classification results, although a number of misclassifications occurred, particularly into visually similar categories such as *Tinea versicolor* and

p-ISSN: 2528-1682
e-ISSN: 2527-9165

Tinea corporis. The *Tinea versicolor* class experienced the highest rate of misclassification, frequently being predicted as *Tinea corporis* or *Tinea nigra*, indicating potential overlap in visual features among these classes. Despite these confusions, the matrix reflects relatively stable overall performance, with high true positive rates across all classes.

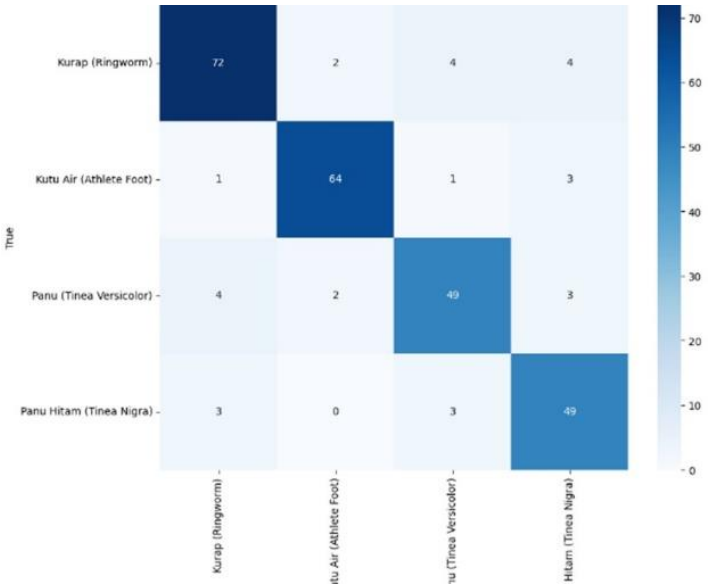


Figure 9. EfficientNet-B2 Confution Matrix

Figure 10 demonstrates that the SE-ResNet101 model was able to classify the test data with high accuracy. The *Tinea pedis* class was classified perfectly with no errors (75/75), highlighting the model's strong ability to recognize the distinctive features of this class. The *Tinea corporis* and *Tinea versicolor* classes also showed excellent performance, with correct classification rates exceeding 90%. However, some misclassifications were observed in the *Tinea nigra* class, particularly with samples being incorrectly classified as *Tinea versicolor*. This suggests visual similarity between the two classes, which may have affected classification precision. Overall, this confusion matrix reinforces the findings of the classification report, confirming that SE-ResNet101 delivers strong multi-class classification performance with a low error rate.

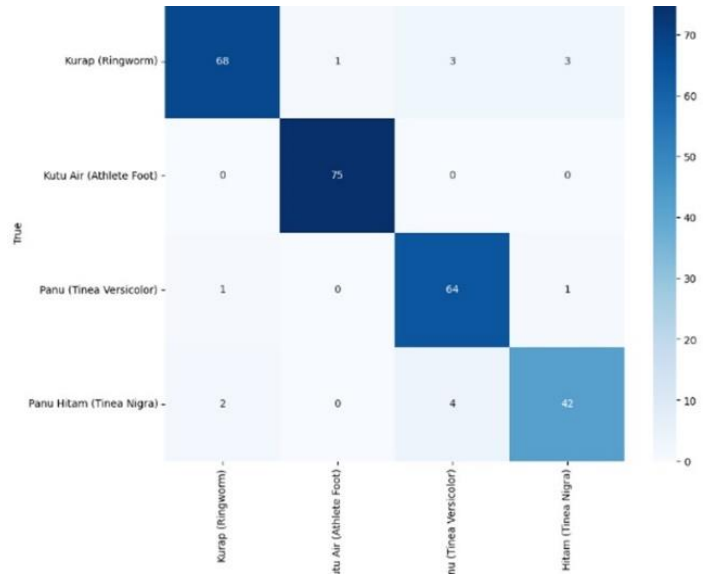


Figure 10. SE-ResNet101 Confution Matrix

Figure 11 shows that the MobileNetV3 model was able to classify the test data with high accuracy and minimal errors. The *Tinea pedis* class was classified with the highest precision, with only one misclassification out of 80 samples. The *Tinea corporis* and *Tinea versicolor* classes each had only 2–3 misclassifications, most of which occurred between classes with similar visual characteristics. The *Tinea nigra* class also demonstrated excellent performance, with just a single misclassified instance. The dominance of values along the main diagonal indicates that the model has a strong generalization capability across all classes, which aligns with the high F1-scores reported in the classification report. These results further reinforce MobileNetV3's position as the most efficient and accurate model for multi-class classification of tropical fungal skin infections.

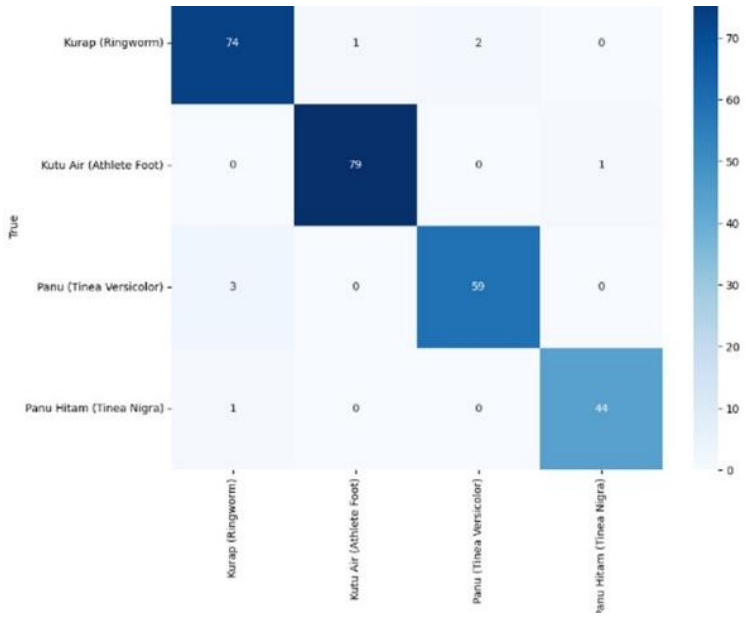


Figure 11. MobileNet V3 Confution Matrix

A visualization was conducted to compare the Test accuracy of the three pre-trained models used in this study. The results, as shown in Figure 12, indicate that the MobileNetV3 architecture achieved the highest Test accuracy of 97.34%. Therefore, the MobileNetV3-based model can be considered the most optimal for classifying tropical fungal skin disease images, as it produced more accurate predictions compared to the other architectures.

The entire training process was carried out using Google Colab with GPU support as the hardware accelerator, with each model trained for 50 epochs. The training time of each architecture is presented in Figure 13. The results demonstrate that MobileNetV3 not only outperformed the other models in terms of accuracy but also achieved the shortest training time. This advantage is attributed to its lightweight architecture, relatively fewer layers, and its design, which was specifically optimized for efficiency on devices with limited computational resources.

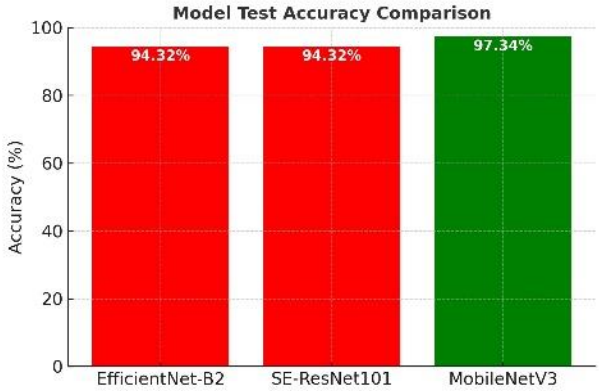


Figure 12. Model Test Accuracy Comparison

p-ISSN: 2528-1682
e-ISSN: 2527-9165

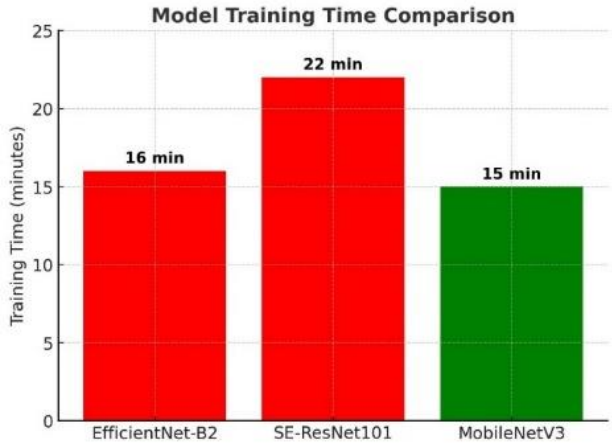


Figure 13. Training Time Comparison

In addition to these visualizations, Table 7 presents a summary of the maximum validation accuracy and the final test accuracy of the three models. MobileNetV3 again achieved the best performance, with a validation accuracy of 95.08% and a test accuracy of 97.34%. Meanwhile, EfficientNet-B2 and SE-ResNet101 obtained nearly comparable test accuracies of 94.32%. The small differences (approximately 1–2%) between validation and test results across all models suggest that they generalize well to unseen data without showing significant signs of overfitting.

Table 7. Validation and Test Accuracy of the Pre-trained Deep Learning Models

Model	Validation Accuracy (Max)	Test Accuracy
EfficientNet-B2	96.20%	94.32%
SE-ResNet101	96.58%	94.32%
MobileNetV3	95.08%	97.34%

Table 8. Comparison Results

Model	Year	Reference	Dataset/Collections	Test Accuracy	Key Difference
MobileNet V3	2025	Alruwaili et al. [9]	Kaggle Skin Diseases Dataset (10 classes, ~27k images)	85%	Used only as a baseline model; reported test accuracy without detailed analysis; no specific focus on tropical fungal infections; applied only basic augmentation
MobileNet V3	-	This Study	Kaggle Skin Diseases Dataset (focused on 4 tropical fungal classes)	97,34%	Conducted detailed evaluation; achieved 97.34% validation accuracy and 95.08% test accuracy; supported by per-class classification reports; applied data augmentation (rotation, flipping, normalization) to improve generalization.
EfficientNet-B2	2023	Mohan et al.[28]	Atlas Dermatology + ISIC (31 classes, 4,910 images → ~49k after augmentation)	87,15%	Trained on diverse dataset (31 general skin disease classes). Applied intensive augmentation (10 techniques). Reported test accuracy of 87.15%.
EfficientNet-B2	.	This Study	Kaggle Skin Diseases Dataset (focused on 4 tropical fungal classes)	94,32%	Focused on specific tropical fungal infections (4 classes). Used moderate augmentation. Achieved higher validation accuracy (96.20%) and test accuracy (94.32%) compared to Mohan et al., indicating strong performance in a specialized domain.
Resnet 101	2021	Kahn et al. [29]	HAM10000 Dataset (10,015 dermoscopic images, 7 general skin lesion classes)	86.78	Utilized standard ResNet101 architecture for classifying general skin lesions (e.g., melanoma, nevus). Focused on non-tropical conditions with limited augmentation (rotation, flipping). Reported only test accuracy without detailed

Model	Year	Reference	Dataset/Collections	Test Accuracy	Key Difference
					per-class evaluation or computational efficiency analysis.
SEResnet 101	-	This Study	Kaggle Skin Diseases Dataset (focused on 4 tropical fungal classes	94,32	Introduced Squeeze-and-Excitation (SE) modules to enhance channel-wise attention. Focused on tropical fungal infections with advanced augmentation (rotation ±15°, flipping, color jitter, normalization). Achieved higher accuracy (validation 96.58%, test 94.32%) and provided detailed evaluation (precision, recall, F1-score), demonstrating strong generalization and computational efficiency suitable for low-resource healthcare deployment.

4. CONCLUSION

This study was conducted to address the challenges in diagnosing tropical fungal skin infections, which have long relied on subjective visual observation and the availability of medical specialists. By comparing three pre-trained deep learning architectures—EfficientNet-B2, SE-ResNet101, and MobileNetV3—this research contributes a systematic comparative evaluation to identify the most optimal model for multi-class classification of tropical fungal skin diseases.

The evaluation results demonstrate that MobileNetV3 is the most superior model, achieving a validation accuracy of 95.08% and a test accuracy of 97.34%, along with the fastest training efficiency (15 minutes), compared to EfficientNet-B2 (16 minutes) and SE-ResNet101 (22 minutes). With consistent performance across all evaluation metrics (accuracy, precision, recall, and F1-score), MobileNetV3 has proven to be not only accurate but also computationally efficient. The main contribution of this study is to establish MobileNetV3 as a practical candidate model for automated classification systems of tropical fungal skin diseases, particularly in primary healthcare facilities with limited infrastructure.

The main contribution of this study lies in providing an empirical benchmark that highlights the balance between accuracy and computational efficiency in deep learning–based classification of tropical fungal skin infections. These findings offer a strong scientific foundation for the development of practical and adaptive automated diagnostic tools, particularly for deployment in healthcare facilities with limited resources. Future research can focus on expanding the dataset by incorporating a wider variety of skin conditions and exploring real-time implementation on mobile or embedded medical systems to enhance clinical applicability and accessibility.

REFERENCE

[1] P. N. Srinivasu, J. G. Sivasai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, "Classification of skin disease using deep learning neural networks with mobilenet v2 and lstm," *Sensors*, vol. 21, no. 8, Apr. 2021, doi: 10.3390/s21082852.

[2] D. Li *et al.*, "Worldwide trends and future projections of fungal skin disease burden: a comprehensive analysis from the Global Burden of Diseases study 2021," *Front. Public Health*, vol. 13, 2025, doi: 10.3389/fpubh.2025.1580221.

[3] World Health Organization (WHO), "Report of the first WHO global meeting on skin-related neglected tropical diseases," World Health Organization (WHO), 2023. Accessed: Oct. 06, 2025. [Online]. Available: https://www.who.int/publications/i/item/9789240081201

[4] Ministry of Health and Republic of Indonesia (Kemenkes RI), "Laporan Nasional Riskesdas," Ministry of Health, Republic of Indonesia (Kemenkes RI), Jakarta, 2019. Accessed: Jul. 06, 2025. [Online]. Available: https://www.litbang.kemkes.go.id/laporan-riskesdas-2019/

[5] S. Azzahra, E. Ervianti, and R. Setiabudi, "Clinical Patterns and Demographic Characteristics of Dermatophytosis in Surabaya," *Indonesian Journal of Tropical and Infectious Disease*, vol. 12, no. 3, pp. 252–267, Dec. 2024, doi: 10.20473/ijtid.v12i3.66511.

[6] M. F. Petrucelli *et al.*, "Epidemiology and diagnostic perspectives of dermatophytoses," *Journal of Fungi*, vol. 6, no. 4, pp. 1–15, Dec. 2020, doi: 10.3390/jof6040310.

[7] V. Ravi, "Attention Cost-Sensitive Deep Learning-Based Approach for Skin Cancer Detection and Classification," *Cancers (Basel).*, vol. 14, no. 23, Dec. 2022, doi: 10.3390/cancers14235872.

[8] M. K. Khalifaturohman, "Systematic Literature Review: Deep Learning Models in Arabic Script Classification," *Khazanah Journal of Religion and Technology*, vol. 3, no. 1, pp. 18–22, Jul. 2025, doi: 10.15575/kjrt.v3i1.1619.

[9] M. Alruwaili and M. Mohamed, "AEfficientNetn Integrated Deep Learning Model with and ResNet for Accurate Multi-Class Skin Disease Classification," *Diagnostics*, vol. 15, no. 5, Mar. 2025, doi: 10.3390/diagnostics15050551.

[10] 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: Jurnal Informatika*, vol. 10, no. 1, pp. 9–18, 2022, Accessed: Oct. 06, 2025. [Online]. Available: https://jurnal.stmiktime.ac.id/index.php/juita/article/view/802

[11] T. Alyas, K. Alissa, A. S. Mohammad, S. Asif, T. Faiz, and G. Ahmed, "Innovative Fungal Disease Diagnosis System Using Convolutional Neural Network," *Computers, Materials and Continua*, vol. 73, no. 3, pp. 4869–4883, 2022, doi: 10.32604/cmc.2022.031376.

[12] T. D. Nigat, T. M. Sitote, and B. M. Gedefaw, "Fungal Skin Disease Classification Using the Convolutional Neural Network," *J. Healthc. Eng.*, vol. 2023, 2023, doi: 10.1155/2023/6370416.

[13] T. Debebe and B. Molla, "Human Skin Fungal Diseases Classification Using Deep Learning Technique," *Haramaya Journal of Engineering and Technology (HJET)*, vol. 1, no. 2, pp. 16–37, 2022, [Online]. Available: https://journals.ddu.edu.et/index.php/HJET

[14] E. da S. Puls, M. V. Todescato, and J. L. Carbonera, "An evaluation of pre-trained models for feature extraction in image classification," Oct. 2023, [Online]. Available: http://arxiv.org/abs/2310.02037

[15] P. Chanyachailert, C. Leeyaphan, and S. Bunyaratavej, "Cutaneous Fungal Infections Caused by Dermatophytes and Non-Dermatophytes: An Updated Comprehensive Review of Epidemiology, Clinical Presentations, and Diagnostic Testing," *Journal of Fungi*, vol. 9, no. 6, Jun. 2023, doi: 10.3390/jof9060669.

[16] N. Łabędź, C. Navarrete-Dechent, H. Kubisiak-Rzepczyk, M. Bowszyc-Dmochowska, A. Pogorzelska-Antkowiak, and P. Pietkiewicz, "Pityriasis Versicolor—A Narrative Review on the Diagnosis and Management," Oct. 01, 2023, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/life13102097.

[17] Q. Yu *et al.*, "Clinical Analysis of 57 Patients With Interdigital Infection in Shanghai, China: A Cross-Sectional Study," *Int. J. Dermatol. Venereol.*, vol. 8, no. 2, pp. 100–105, Jun. 2025, doi: 10.1097/JD9.0000000000000233.

[18] A. K. C. Leung, J. M. Lam, K. F. Leong, and K. L. Hon, "Tinea corporis: An updated review," Jul. 20, 2020, *Bioexcel Publishing LTD*. doi: 10.7573/dic.2020-5-6.

[19] M. Trigka and E. Dritsas, "A Comprehensive Survey of Deep Learning Approaches in Image Processing," *Sensors (Basel)*, vol. 25, no. 2, Jan. 2025, doi: 10.3390/s25020531.

[20] D. Murcia-Gómez, I. Rojas-Valenzuela, and O. Valenzuela, "Impact of Image Preprocessing Methods and Deep Learning Models for Classifying Histopathological Breast Cancer Images," *Applied Sciences (Switzerland)*, vol. 12, no. 22, Nov. 2022, doi: 10.3390/app122211375.

[21] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *arXiv preprint arXiv:1905.11946*, no. 6, Sep. 2020, doi: 10.48550/arXiv.1905.11946.

[22] A. Howard *et al.*, "Searching for MobileNetV3," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, 2019. doi: 10.48550/arXiv.1905.02244.

[23] I. Manole, A. I. Butacu, R. N. Bejan, and G. S. Tiplica, "Enhancing Dermatological Diagnostics with EfficientNet: A Deep Learning Approach," *Bioengineering*, vol. 11, no. 8, Aug. 2024, doi: 10.3390/bioengineering11080810.

[24] S. Thirumalaisamy *et al.*, "Breast Cancer Classification Using Synthesized Deep Learning Model with Metaheuristic Optimization Algorithm," *Diagnostics*, vol. 13, no. 18, Sep. 2023, doi: 10.3390/diagnostics13182925.

[25] B. Ashwini, M. Kaur, D. Singh, S. Roy, and M. Amoon, "Efficient Skip Connections-Based Residual Network (ESRNet) for Brain Tumor Classification," *Diagnostics*, vol. 13, no. 20, Oct. 2023, doi: 10.3390/diagnostics13203234.

[26] M. Huang and Y. Xu, "Image classification of Chinese medicinal flowers based on convolutional neural network," *Mathematical Biosciences and Engineering*, vol. 20, no. 8, pp. 14978–14994, 2023, doi: 10.3934/mbe.2023671.

[27] S. zhenlin, "Lightweight Face Anti-spoofing for Improved MobileNetV3," *Journal of Image Processing Theory and Applications*, vol. 7, no. 1, 2024, doi: 10.23977/jipta.2024.070117.

[28] J. Mohan, A. Sivasubramanian, V. Sowmya, and R. Vinayakumar, "Enhancing Skin Disease Classification Leveraging Transformer-based Deep Learning Architectures and Explainable AI," *Comput. Biol. Med.*, Jul. 2024, doi: 10.1016/j.compbiomed.2025.110007.

[29] M. A. Khan, M. Sharif, T. Akram, R. Damaševičius, and R. Maskeliūnas, "Skin lesion segmentation and multiclass classification using deep learning features and improved moth flame optimization," *Diagnostics*, vol. 11, no. 5, 2021, doi: 10.3390/diagnostics11050811.