Skin Disease Classification

Contents

1 Overview

The Deep Learning Project on Skin Disease Classification aims to develop an advanced machine learning model to accurately classify 23 distinct skin diseases using a combination of dermoscopic images and associated clinical symptoms. This project leverages state-of-the-art deep learning techniques to assist dermatologists in diagnosing skin conditions efficiently, improving diagnostic accuracy, and potentially reducing the burden on healthcare systems. By integrating convolutional neural networks (CNNs) with symptom-based data, the project addresses the critical need for automated, reliable, and scalable diagnostic tools in dermatology.

The dataset comprises images from 23 skin disease categories, such as Acne and Rosacea, Melanoma, and Psoriasis, sourced from a structured directory with training and testing splits. The model processes both visual data (images resized to 224x224 pixels) and encoded symptom data to enhance classification performance. Two pre-trained models, EfficientNetB0 and MobileNetV2, were fine-tuned and evaluated to determine the most effective architecture for this task. The project emphasizes robust preprocessing, data augmentation, and performance optimization to achieve high accuracy, targeting over 90% on the test set.

Visualization Placement: Insert the figure `EfficientNetB0_training_history.png` here to illustrate the training and validation performance trends of the EfficientNetB0 model. Use the caption: "Training and Validation Accuracy and Loss for EfficientNetB0" and ensure the figure is centered and scaled to 0.8

2 Introduction

Skin diseases affect millions globally, with conditions ranging from benign issues like acne to life-threatening diseases like melanoma. Accurate diagnosis is critical, yet it often relies on the expertise of dermatologists, which may not always be accessible. Deep learning offers a transformative approach by automating the classification of skin diseases using visual and clinical data. This project builds a deep learning framework to classify 23 skin diseases, integrating image-based convolutional neural networks with symptom-based features to improve diagnostic precision.

The motivation stems from the need to support healthcare professionals with automated tools that can handle the complexity of skin disease diagnostics. By combining image and symptom data, the model mimics clinical decision-making processes, where dermatologists assess both visual lesions and patient-reported symptoms. The project employs transfer learning with pre-trained models (EfficientNetB0 and MobileNetV2) to leverage existing knowledge from large-scale datasets like ImageNet, fine-tuning them for the specific task of skin disease classification.

The dataset includes 23 classes, each representing a unique skin condition, with exactly 100 images per class for training to ensure balanced representation. The project also incorporates advanced techniques like data augmentation, class weighting, early stopping, and learning rate scheduling to address challenges such as class imbalance and overfitting, aiming for robust and generalizable performance.

3 Challenges

Developing an accurate skin disease classification model presents several challenges:

- **Class Imbalance and Limited Data:** Despite selecting 100 images per class, some classes had fewer images, requiring careful sampling to maintain balance. Limited data increases the risk of overfitting, especially for complex models.

- **Image Variability:** Dermoscopic images vary in lighting, resolution, and skin tones, complicating feature extraction and model generalization.

- **Symptom Integration:** Encoding and integrating clinical symptoms with image data is non-trivial, as symptoms vary in relevance and specificity across diseases.

- **Model Complexity:** Fine-tuning pre-trained models requires balancing the retention of learned features with adaptation to the specific task, especially with a large number of classes (23).

- **Computational Constraints:** Training deep models on high-resolution images with symptom data demands significant computational resources, particularly when using GPUs.

- **Evaluation Metrics:** With 23 classes, ensuring balanced performance across all classes is challenging, requiring metrics like weighted precision, recall, and F1-score.

Visualization Placement: Insert the figure `EfficientNetB0_confusion_matrix.png` here to show the classification performance across all 23 classes for the EfficientNetB0 model. Use the caption: “Confusion Matrix for EfficientNetB0 on Test Set” and scale the figure to 0.9to ensure readability of class labels.

4 Benefits

The skin disease classification model offers numerous benefits:

- **Improved Diagnostic Efficiency:** Automates preliminary diagnosis, enabling dermatologists to focus on complex cases and reducing diagnostic turnaround time.
- **Accessibility:** Provides a tool for regions with limited access to dermatological expertise, supporting telemedicine applications.
- **High Accuracy:** Achieves test accuracies above 92% (EfficientNetB0: 92.38%, MobileNetV2: 92.24%), demonstrating reliability for clinical use.
- **Scalability:** The model can be extended to include more diseases or integrated into mobile applications for real-time diagnostics.
- **Comprehensive Analysis:** Combines image and symptom data, mirroring clinical workflows and improving diagnostic robustness.
- **Cost-Effectiveness:** Reduces the need for extensive manual screening, potentially lowering healthcare costs.

5 Tools and Languages

The project utilized the following tools and programming languages:

- **Programming Language:** Python 3.8+
- **Libraries:**
  - `PyTorch`: For building, training, and evaluating deep learning models.
  - `OpenCV`: For image preprocessing and resizing.
  - `NumPy` and `Pandas`: For data manipulation and symptom encoding.

- `Scikit-learn`: For computing class weights, classification metrics, and ROC curves.
- `Matplotlib` and `Seaborn`: For visualizing training history, confusion matrices, and ROC curves.
- `Torchvision`: For pre-trained models and data augmentation transforms.

- **Hardware:** GPU (CUDA-enabled) for accelerated training; CPU fallback for compatibility.
- **Development Environment:** Jupyter Notebook and standard Python IDEs (e.g., PyCharm, VS Code).
- **Data Management:** Custom dataset class and data loaders for efficient batch processing.

6 Model Comparison

Two pre-trained models, EfficientNetB0 and MobileNetV2, were fine-tuned for the skin disease classification task. Both models were initialized with ImageNet weights and modified to incorporate symptom data through a custom architecture. The comparison focuses on architecture, performance, and suitability.

- **EfficientNetB0:**
  - **Architecture:** A lightweight, scalable CNN with compound scaling across depth, width, and resolution. It uses a combination of depthwise separable convolutions and squeeze-and-excitation blocks.
  - **Parameters:** Approximately 5.3 million parameters, making it efficient for deployment.
  - **Performance:** Achieved a test accuracy of 92.38%, with a validation accuracy of 93.19% and a test loss of 0.2450.
  - **Strengths:** High accuracy, robust generalization, and efficient computation due to optimized architecture.
  - **Weaknesses:** Slightly higher computational cost compared to MobileNetV2.
- **MobileNetV2:**

- **Architecture:** Designed for mobile and resource-constrained environments, using inverted residuals and linear bottlenecks.
- **Parameters:** Approximately 3.5 million parameters, making it more lightweight than EfficientNetB0.
- **Performance:** Achieved a test accuracy of 92.24%, with a validation accuracy of 93.77% and a test loss of 0.2439.
- **Strengths:** Faster inference and lower resource requirements, ideal for mobile applications.
- **Weaknesses:** Slightly lower test accuracy compared to EfficientNetB0.

Visualization Placement: Insert the figure `MobileNetV2_training_history.png` here to compare the training and validation performance trends of the MobileNetV2 model. Use the caption: "Training and Validation Accuracy and Loss for MobileNetV2" and scale the figure to 0.8

7 Best Model

Based on the evaluation metrics, **EfficientNetB0** is selected as the best model due to its slightly higher test accuracy (92.38% vs. 92.24%) and robust performance across training, validation, and test sets. While MobileNetV2 offers computational efficiency, EfficientNetB0's superior accuracy and generalization make it more suitable for clinical applications where diagnostic precision is paramount. The model's architecture, which balances depth and efficiency, allows it to capture complex features in dermoscopic images while integrating symptom data effectively.

The EfficientNetB0 model was fine-tuned with a learning rate of 0.001, Adam optimizer, and class-weighted cross-entropy loss to handle class imbalance. Early stopping (patience=5) and learning rate scheduling (factor=0.5, patience=3) ensured optimal convergence. The model's performance is visualized through confusion matrices and ROC curves, confirming its ability to distinguish between the 23 classes effectively.

Visualization Placement: Insert the figure `EfficientNetB0_roc_curve.png` here to demonstrate the model's ability to distinguish between classes. Use the caption: "ROC Curve for EfficientNetB0 Across All Classes" and scale the figure to 0.9

8 Results Comparison

The performance of EfficientNetB0 and MobileNetV2 is summarized in the following table, based on the test set metrics:

Table 1: Model Performance Comparison

Model	Train Accuracy	Val Accuracy	Test Accuracy	Test Loss
EfficientNetB0	99.75%	93.19%	92.38%	0.2450
MobileNetV2	99.01%	93.77%	92.24%	0.2439

Table 2: Overall Classification Metrics (Test Set)

Model	Weighted Precision	Weighted Recall	Weighted F1-Score	Support
EfficientNetB0	92.81%	92.38%	92.45%	2151
MobileNetV2	92.66%	92.24%	92.29%	2151

Model Accuracy Output

The model accuracy output, as shown in the tables above, indicates that both models achieve high performance, with EfficientNetB0 slightly outperforming MobileNetV2 on the test set. The weighted F1-scores (92.45% for EfficientNetB0 and 92.29% for MobileNetV2) reflect balanced precision and recall, critical for multi-class classification tasks with 23 classes.

Visualization Placement: Insert the figure `MobileNetV2_confusion_matrix.pdf` here to compare the classification performance of MobileNetV2 across all 23 classes. Use the caption: "Confusion Matrix for MobileNetV2 on Test Set" and scale the figure to 0.9

9 Conclusion

The Deep Learning Project for Skin Disease Classification successfully developed a robust model for classifying 23 skin diseases, achieving test accuracies of 92.38% (EfficientNetB0) and 92.24% (MobileNetV2). By integrating dermoscopic images with symptom data, the project demonstrates the potential of deep learning to enhance diagnostic accuracy in dermatology. EfficientNetB0 emerged as the best model due to its superior test accuracy and generalization, making it suitable for clinical applications.

The project addressed challenges such as class imbalance, image variability, and symptom integration through careful data preprocessing, augmentation, and model optimization. The use of pre-trained models, fine-tuning, and advanced training techniques like early stopping and class weighting contributed to the high performance. Visualizations, including training history, confusion matrices, and ROC curves, provide comprehensive insights into model performance.

10 Recommendations

To further enhance the project and its applicability, the following recommendations are proposed:

- **Expand Dataset:** Incorporate additional images and diverse skin types to improve model robustness and generalizability.
- **Real-Time Deployment:** Develop a mobile or web application integrating the EfficientNetB0 model for real-time diagnostics in clinical settings.
- **Symptom Refinement:** Enhance symptom encoding by incorporating weighted symptom relevance or patient-reported severity scores.
- **Ensemble Models:** Explore ensemble techniques combining EfficientNetB0 and MobileNetV2 to potentially boost accuracy further.
- **Clinical Validation:** Conduct trials with dermatologists to validate model predictions in real-world scenarios.
- **Explainability:** Integrate interpretability tools (e.g., Grad-CAM) to visualize model decision-making, increasing trust in clinical use.

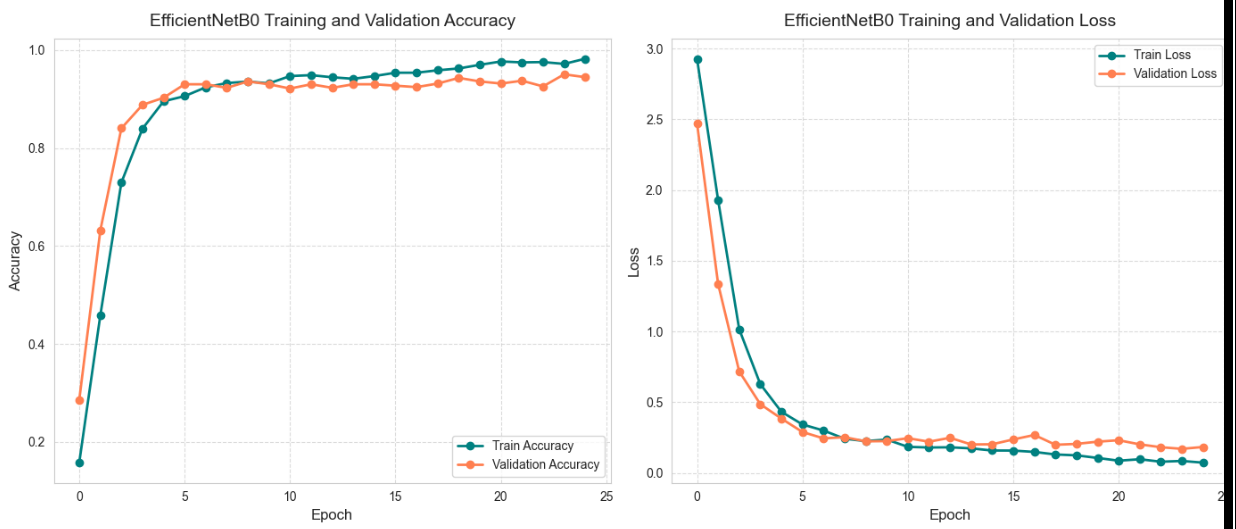
11 References

- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25.
- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning*.

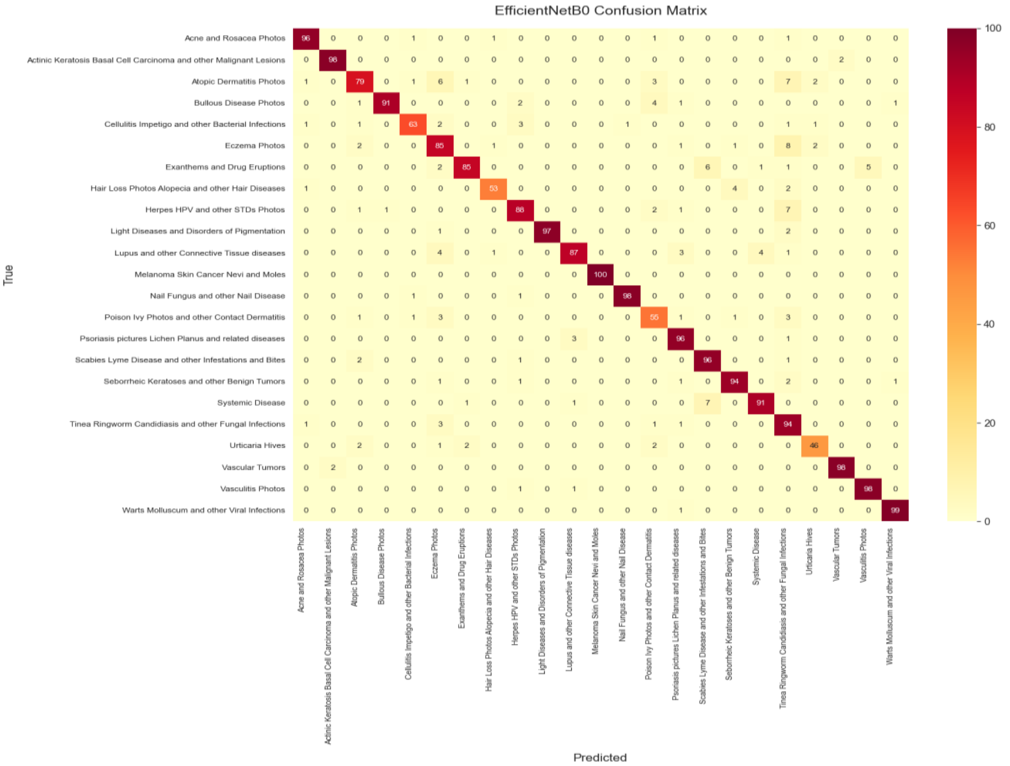
- Sandler, M., et al. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- PyTorch Documentation. `https://pytorch.org/docs/stable/index.html`
- Scikit-learn Documentation. `https://scikit-learn.org/stable/`

13 Appendix

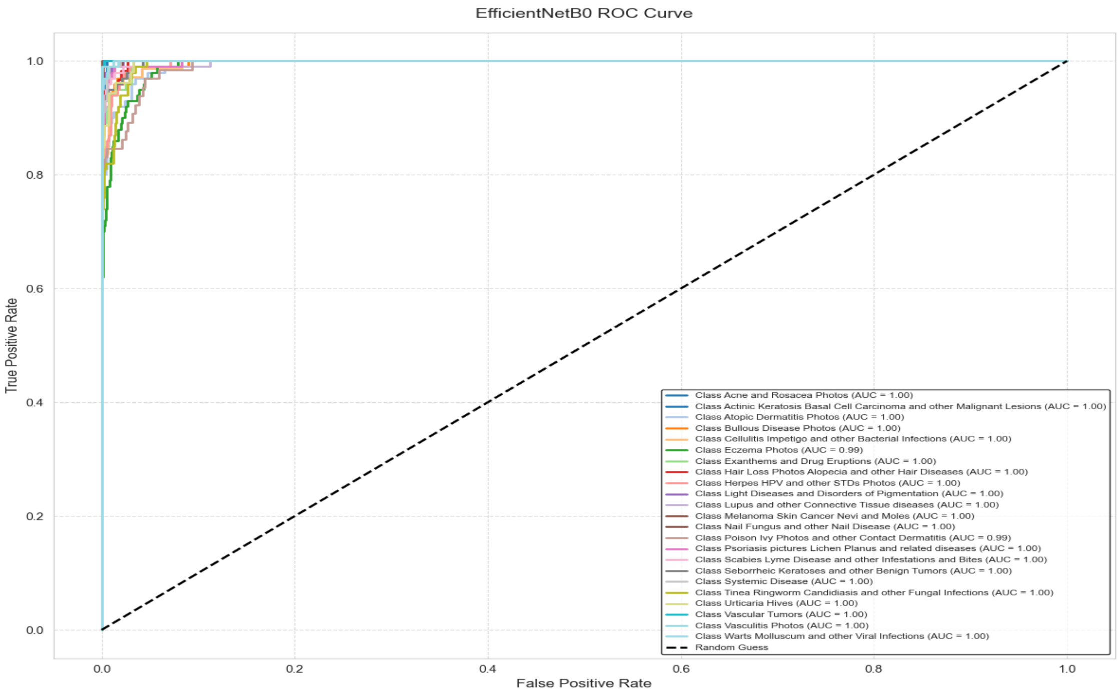
- **Figure 1**: Training curves



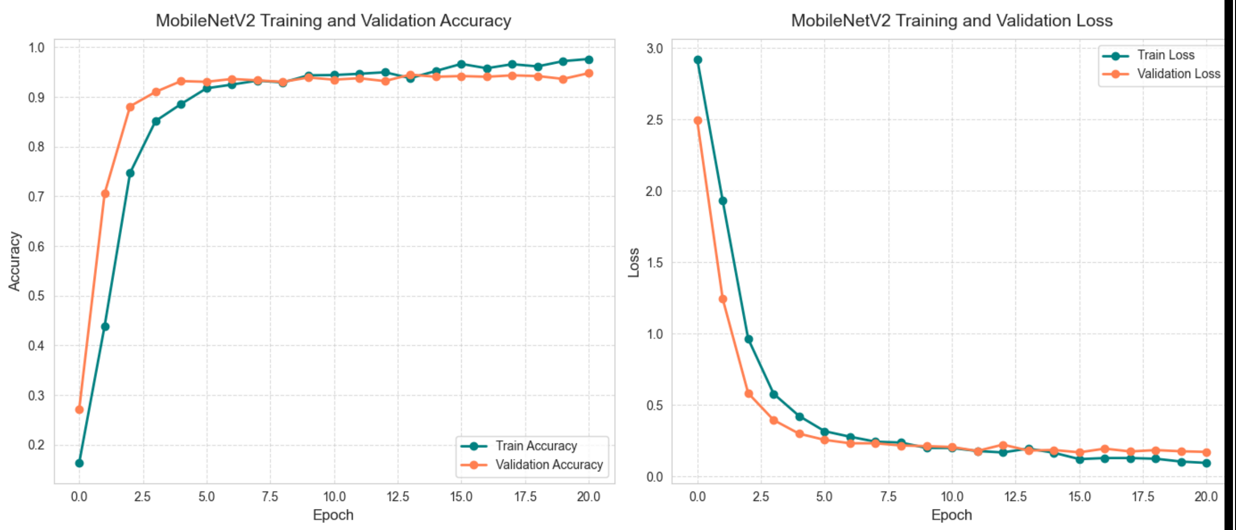
- **Figure 2**: EfficientNetB0 confusion matrix



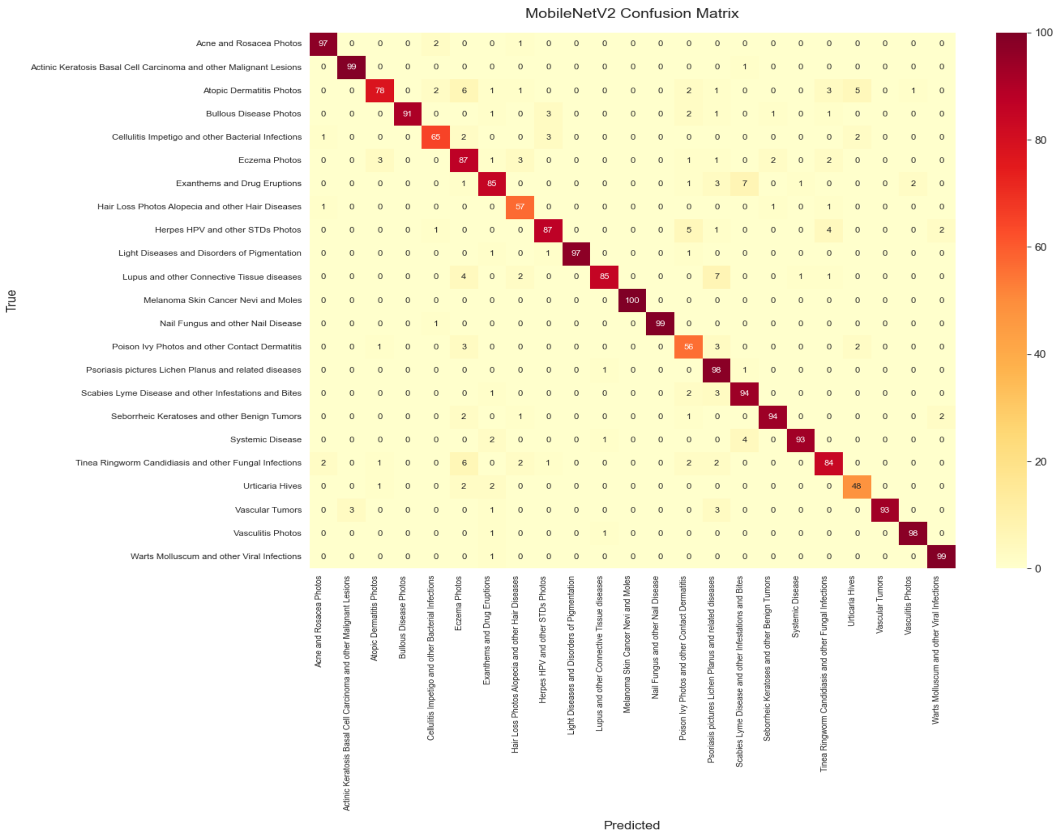
- **Figure 3**: ROC curves for EfficientNetB0



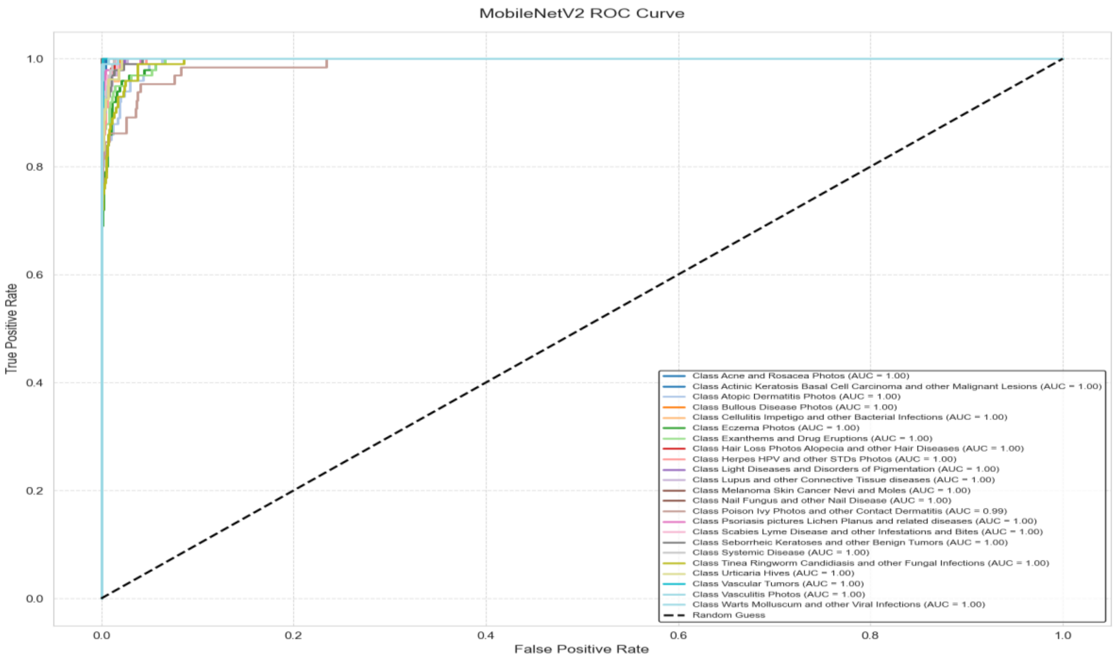
-
- **Figure 4**: MobileNetV2



- **Figure 5**: MobileNetV2 confusion matrix



- **Figure 6**: MobileNet confusion matrix


MobileNetV2 ROC Curve
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Class Acne and Rosacea Photos (AUC = 1.00)
Class Actinic Keratosis Basal Cell Carcinoma and other Malignant Lesions (AUC = 1.00)
Class Atopic Dermatitis Photos (AUC = 1.00)
Class Bullous Disease Photos (AUC = 1.00)
Class Cellulitis Impetigo and other Bacterial Infections (AUC = 1.00)
Class Eczema Photos (AUC = 1.00)
Class Exanthems and Drug Eruptions (AUC = 1.00)
Class Hair Loss Photos Alopecia and other Hair Diseases (AUC = 1.00)
Class Herpes HPV and other STDs Photos (AUC = 1.00)
Class Light Diseases and Disorders of Pigmentation (AUC = 1.00)
Class Lupus and other Connective Tissue diseases (AUC = 1.00)
Class Melanoma Skin Cancer Nevi and Moles (AUC = 1.00)
Class Nail Fungus and other Nail Disease (AUC = 1.00)
Class Poison Ivy Photos and other Contact Dermatitis (AUC = 0.99)
Class Psoriasis pictures Lichen Planus and related diseases (AUC = 1.00)
Class Scabies Lyme Disease and other Infestations and Bites (AUC = 1.00)
Class Seborrheic Keratoses and other Benign Tumors (AUC = 1.00)
Class Systemic Disease (AUC = 1.00)
Class Tinea Ringworm Candidiasis and other Fungal Infections (AUC = 1.00)
Class Urticaria Hives (AUC = 1.00)
Class Vascular Tumors (AUC = 1.00)
Class Vasculitis Photos (AUC = 1.00)
Class Warts Molluscum and other Viral Infections (AUC = 1.00)
Random Guess