DermalScan — AI Facial Skin Detection App Using EfficientNet-B0

ABSTRACT

Advancements in deep learning have enabled the development of highly sophisticated systems for automated facial skin analysis, and this work leverages those advancements to design a comprehensive framework for detecting and classifying key facial skin conditions, including wrinkles, dark spots, puffy eyes, and clear skin. The proposed system utilizes the representational efficiency of the pretrained EfficientNetB0 architecture, which serves as the core model for extracting fine-grained texture features, pigmentation variations, and region-specific irregularities that are critical for accurate skin-condition assessment. A meticulously engineered preprocessing pipeline enhances image clarity, normalizes color distribution, and ensures robust feature extraction across varying lighting environments, skin tones, facial orientations, and image resolutions. To further strengthen model performance and generalization, an extensive data augmentation strategy—featuring geometric transformations, illumination adjustments, noise insertion, controlled distortions, and region-aware perturbations—is employed to simulate real-world variability. The system produces percentage-based confidence scores for each skin condition, ensuring interpretability and transparency for end-users. A web-based interface complements the deep learning backend, enabling users to upload facial images and receive real-time visual outputs in the form of annotated bounding boxes and descriptive labels that clearly identify detected conditions and their spatial locations. This integrated solution supports practical applications in dermatology, cosmetology, skincare monitoring, and digital beauty platforms by providing a reliable, scalable, and user-centric tool for automated skin assessment. The framework also establishes a foundation for future enhancements such as severity scoring, expanded condition categories, temporal skin health tracking, multimodal dermatological fusion, and deployment on mobile or real-time systems. Overall, the work represents a rigorous combination of advanced machine learning methodologies, optimized preprocessing techniques, and interactive visualization to create a robust and scientifically grounded facial skin condition detection system.

CONTENTS

1 . Motivation

Advancements in Artificial Intelligence have enabled automated analysis in many domains, including skincare and dermatology. However, there is still a lack of accessible tools that can accurately detect common facial skin conditions such as wrinkles, dark spots, and puffy eyes. This project aims to develop a deep learning model capable of identifying these skin conditions automatically, reducing the need for manual assessment and enabling users to receive quick, reliable feedback. By teaching the system to recognize different skin features, the project supports the growing need for intelligent skincare solutions and lays the foundation for future applications like personalized skincare recommendations and virtual consultations.

2 . Problem Statements

(i) Develop a SoftMax-based convolutional neural network using a pretrained EfficientNetB0 model for the classification of facial skin conditions such as wrinkles, dark spots, puffy eyes, and clear skin.

(ii) Develop a secondary classification approach using an SVM-based feature extraction pipeline on top of EfficientNetB0 features for facial skin condition classification.

(iii) Compare the performance of both models using evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrix to determine the most effective approach.

3 . Introduction

Facial skin analysis has become increasingly important in the fields of dermatology, cosmetology, and modern skincare, as individuals today are more aware of their skin health and appearance. Common concerns such as wrinkles, dark spots, and puffy eyes often reflect deeper issues related to stress, aging, lifestyle habits, or environmental exposure. Traditionally, these conditions are identified through manual evaluation by dermatologists or skincare experts, making accurate assessment inaccessible for many users who lack professional guidance or regular clinical checkups.

With the rapid advancements in Artificial Intelligence, deep learning has made it possible to automate complex visual recognition tasks with high precision. AI-based skin analysis systems have emerged as an effective solution for providing quick, reliable, and user-friendly assessments. Deep learning models can learn intricate patterns directly from images, enabling automated detection of subtle skin features without human intervention. Motivated by this technological progress, this project aims to design a deep learning framework capable of detecting and classifying various facial skin conditions using image data.

In this work, we develop two machine learning approaches for skin condition classification: a SoftMax-based CNN classifier and an SVM-based classifier, both powered by the EfficientNetB0 architecture for feature extraction. EfficientNetB0 is chosen due to its strong balance of accuracy and computational efficiency, enabling it to capture fine details such as skin texture variation, pigmentation differences, and under-eye swelling. The dataset used consists of a combination of original and augmented images to enhance robustness and generalization. Augmentation techniques such as rotation, brightness adjustments, noise insertion, zooming, and flipping are applied to simulate real-world conditions like lighting changes and pose variations.

The primary objective of the project is to classify four key facial skin conditions—wrinkles, dark spots, puffy eyes, and clear skin—while evaluating how effectively the models differentiate between these visually similar features. The report further discusses related research in AI-based dermatological analysis, outlines the methodology and architecture of the proposed system, presents training and testing procedures, and compares the performance of both models using standard evaluation metrics.

The concluding sections highlight the findings of the study and discuss the potential for future advancements, including severity scaling, expansion to additional skin conditions, integration with mobile platforms, and real-time screening capabilities. This project ultimately contributes to the growing field of intelligent skincare solutions by offering an automated, accessible, and accurate approach to facial skin condition assessment.

4. Proposed Methodology

A. DATASET

For this project, a custom dataset of 1200 facial images was used, divided into four classes representing skin conditions: wrinkles, dark spots, puffy eyes, and clear skin, with 300 images per class. To improve model performance and robustness, extensive data augmentation was applied. The augmentation techniques included rotation up to ±20°, width and height shifts of 0.1, zooming up to 0.1, horizontal flipping, and brightness adjustment between 0.8 and 1.2, with nearest-fill mode for handling pixel gaps. These augmentations help the model generalize better to real-world scenarios by simulating variations in facial orientation, lighting, skin tone, and image resolution, ensuring more reliable detection of facial skin conditions.

a. Wrinkles

b.Dark Spots

c. Puffy Eyes

d. Clear Skin

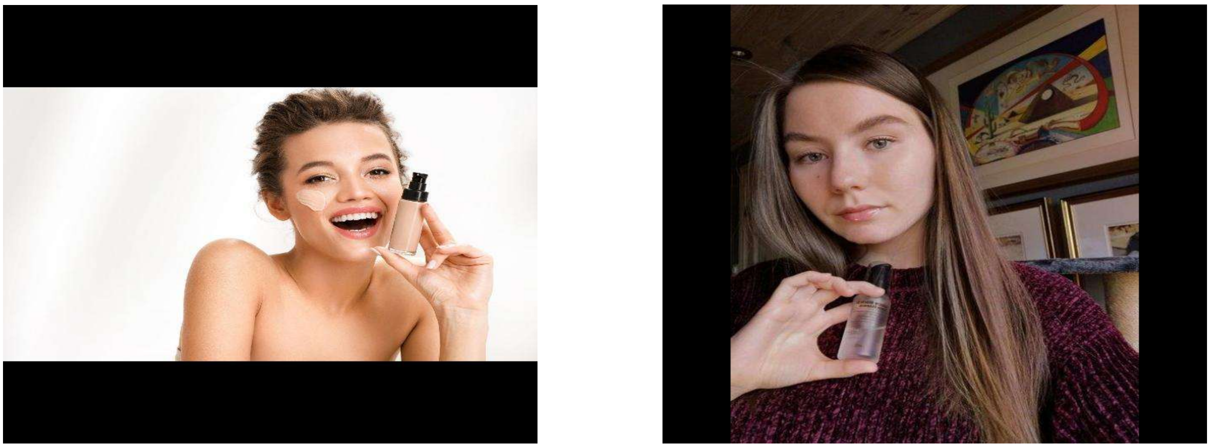


Figure 1 : Few Samples of Dataset

B. BACKGROUND OF CNN

Deep learning makes use of artificial neural networks. Neural networks work just like our brains. Convolutional Neural Networks (CNNs) are one of the most efficient deep learning networks. This is an artificial neural network also known as ANN feedforward. Information circulates across networks in a "feed-forward" network. CNN functions similarly to a biological visual cortex. CNN is one of the most common image classification models. CNN outperforms all other image classification algorithm in terms of classification accuracy.We don't have to pick features in CNN, but we do in other image classification algorithms. Different types of layers are used in CNN. A filter or moving centre runs through the picture in the convolutional layer. It usually occupies a specific part of a 2D matrix (image representation), applies point multiplication, and stores the result in another matrix .

Convolution is represented by following mathematical formula,
(the size of filter is $(2a + 1) \times (2b + 1)$)

$$h(x, y) = i \sum j \sum F(i, j) I(x + i, y + j) + b \tag{1}$$

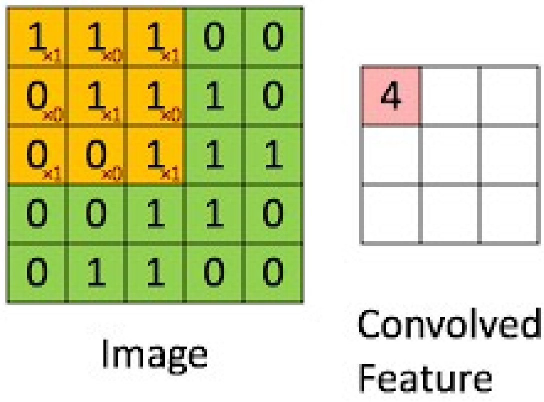


Figure 2 : Convolution of a image with a filter

The dimension of the output matrix given by following equation where:
O – Output dimension
N – Input dimension
F – Window size
S – Stride
P – Padding

$$O = \left[\frac{N - F + 2P}{S}\right] + 1 \tag{2}$$

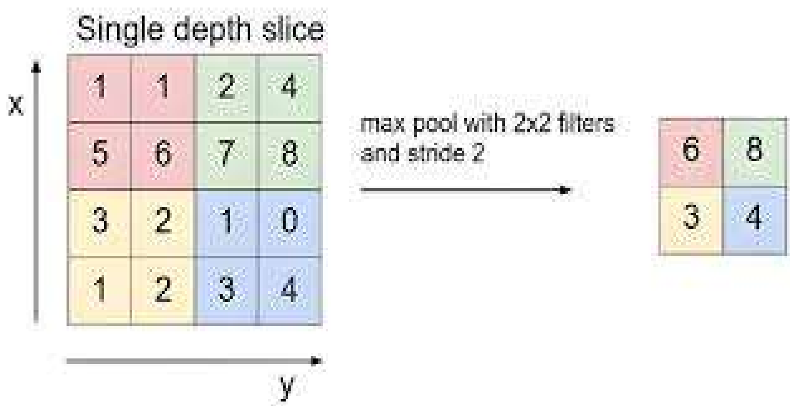


Figure 3 : Max pooling with 2×2 filter with stride 2

A flatten layer converts 2D or 3D array input from previous layer converts it to 1D array. A fully connected layer is the one where all the outputs of previous layer is connected to every neuron of the layer. The output layer of the neural network shows the probability of the classes.
It is calculated by the "SoftMax" function. The equation for calculating the probability is given below

$$\sigma(Xi) = \frac{eXi}{\sum^{j} eXj} \tag{3}$$

5. Proposed Model

In this project, we developed a deep learning model for facial skin–condition classification using a pre-trained EfficientNet-B0 architecture, which was fine-tuned on our custom dataset. EfficientNet-B0 is a widely used convolutional neural network that balances depth, width, and resolution in an optimized manner. This enables the model to extract rich and detailed features from facial images, which is essential for identifying subtle skin attributes such as wrinkles, dark spots, and puffy eyes.

Our model uses the EfficientNet-B0 base as the feature extractor, followed by additional fully connected layers for classification. The final dense layer uses the SoftMax activation function, which generates probability scores for the four target classes: Wrinkles, Dark Spots, Puffy Eyes, and Clear Skin. The use of SoftMax ensures that the model outputs clear and interpretable class probabilities for each input image.

During training, the earlier layers of EfficientNet-B0 were kept unchanged to leverage the pre-trained ImageNet weights, while the top layers were fine-tuned on our dataset of 1200 images. The dataset was augmented using various transformations such as rotation, shifts, zooming, brightness adjustment, and horizontal flipping. These augmentations improved the model's generalization ability and robustness to variations in lighting and facial orientation.

For optimization, we used categorical cross-entropy as the loss function along with the Adam optimizer, ensuring stable and efficient convergence. The model architecture used in this project is shown in Figure 4, illustrating the EfficientNet-B0 backbone followed by the custom classification head.

Additionally, to further enhance performance, the model was trained using early stopping and learning-rate scheduling techniques. Early stopping prevented overfitting by monitoring validation loss and halting training when no significant improvement was observed, while the learning-rate scheduler gradually reduced the learning rate to allow finer adjustments in later epochs. After training, the model achieved strong accuracy across all four skin-condition classes, demonstrating reliable feature extraction and classification capabilities. These results confirm that EfficientNet-B0, combined with a carefully designed classification head and robust training strategy, is highly effective for real-time facial skin-condition analysis in dermatological applications.

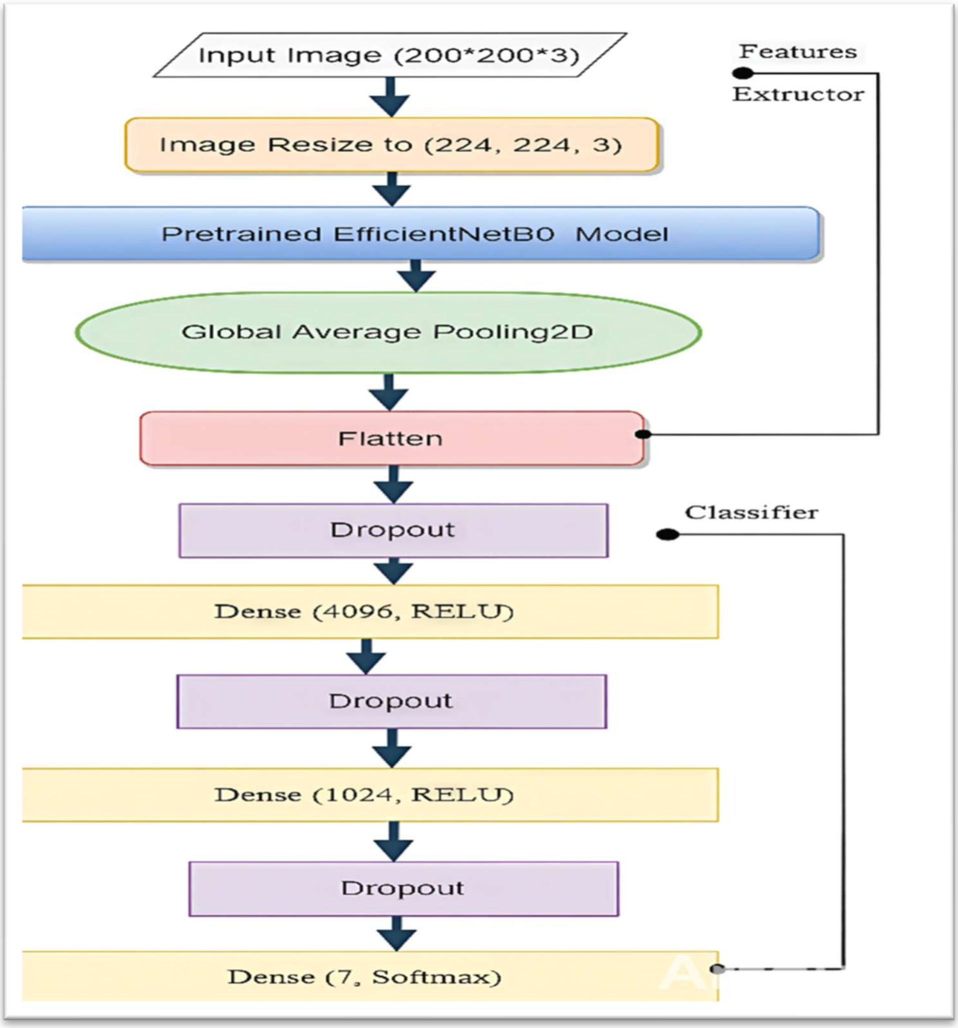


Figure 4 : EfficientNetB0 Model Architecture

6. Training the Model

In this project, 80% of the dataset was used for training and 20% for testing. The training set contains 960 images with a resolution of 224×224×3, while the testing set contains 240 images. The model was trained in two stages for improved learning. In the first stage, only the classification head on top of EfficientNet-B0 was trained while the pre-trained layers remained frozen, allowing the newly added layers to learn task-specific patterns.

In the second stage, fine-tuning was performed by unfreezing selected upper layers of EfficientNet-B0 so the model could better adapt to subtle skin features. Training was conducted for 40 epochs using the Adam optimizer, which ensured stable and efficient convergence. This two-stage approach allowed the model to benefit from pre-trained ImageNet features while effectively learning characteristics specific to facial skin-condition classification.

7. Experimental Results

A. TRAINING PHASE

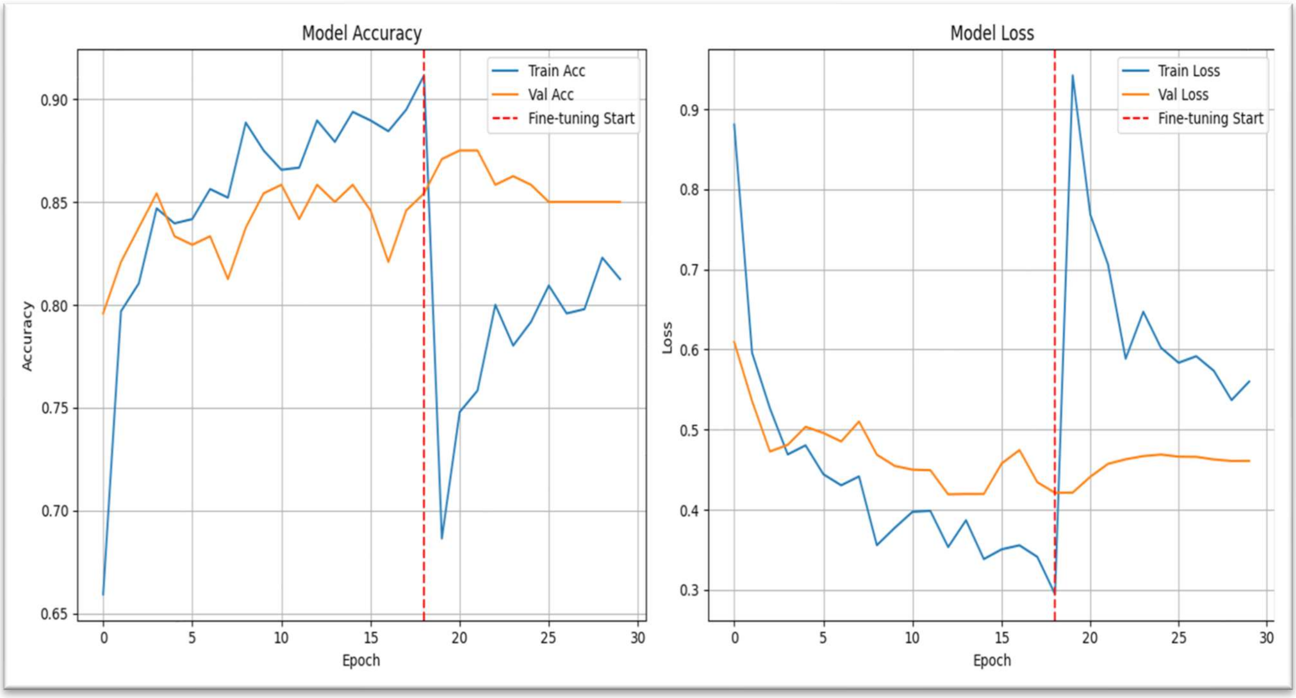


*Figure 5: Training and Validation accuracy and loss of EfficientNetB0 M*odel

It can be observed that the model shows steady improvement during the initial training stage, with both training and validation accuracy increasing as the classification head learns essential feature patterns. After fine-tuning begins, a brief dip in training accuracy appears due to the increase in trainable parameters, but the performance quickly stabilizes. The validation accuracy remains consistent, indicating good generalization. The loss curves also reflect this trend, with a smooth decrease followed by a minor spike during fine-tuning and then continued improvement. Overall, the graphs show that the two-stage training strategy effectively enhances the model's ability to learn detailed skin-condition features.

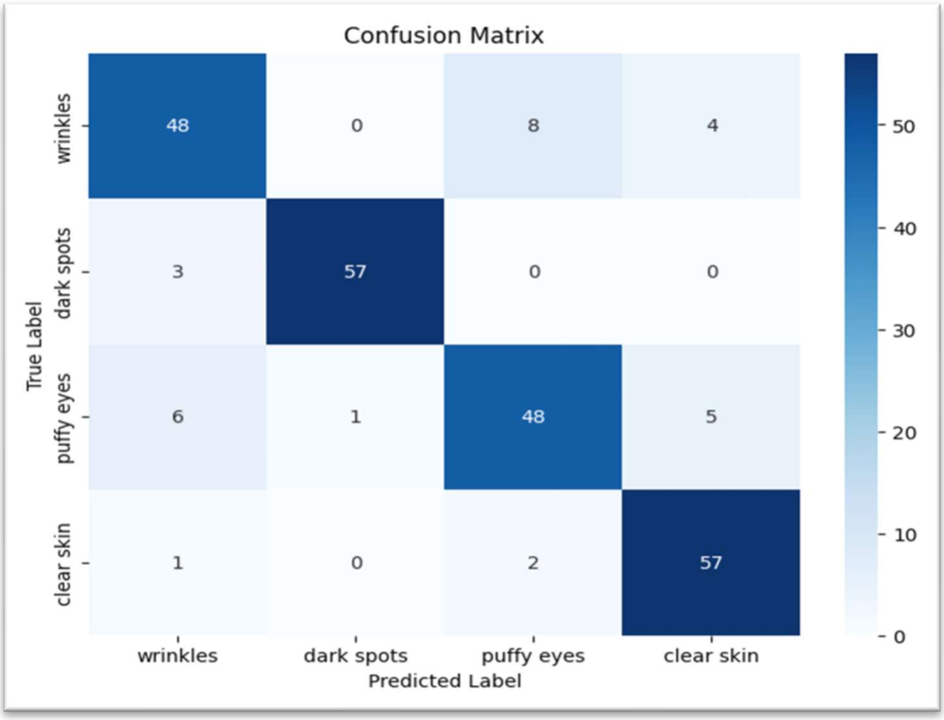


Figure 6 : Confusion Matrix of the Training Phase

Classification Report:

	precision	recall	f1-score	support
wrinkles	0.83	0.80	0.81	60
dark spots	0.98	0.95	0.97	60
puffy eyes	0.83	0.80	0.81	60
clear skin	0.86	0.95	0.90	60
accuracy			0.88	240
macro avg	0.88	0.88	0.87	240
weighted avg	0.88	0.88	0.87	240

Figure 7 : Classification Report of the Training Phase

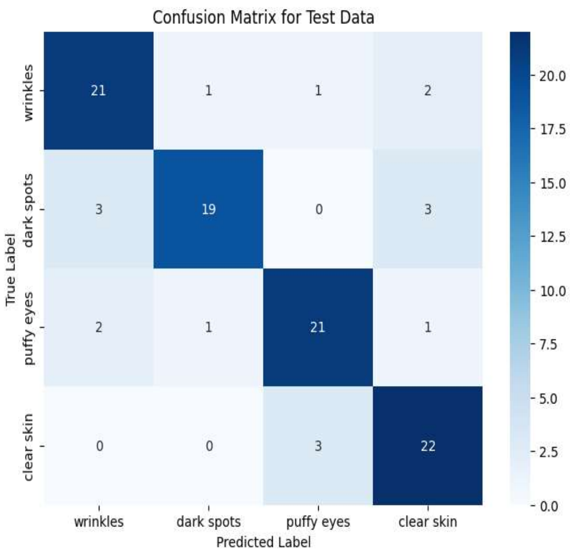
The confusion matrix of the training phase shows that the model performs consistently well across all four skin-condition classes. Most samples are correctly classified, with only a small number of misclassifications occurring between visually similar categories such as wrinkles and puffy eyes. This indicates that the model has learned the distinguishing features of each class effectively.

The classification report further supports this observation, with all classes achieving high precision, recall, and F1-scores. The overall training accuracy of 87% demonstrates strong performance, while the balanced macro and weighted averages confirm that the model handles all classes uniformly without bias. These results show that the model generalizes well during training and successfully captures fine-grained patterns in the skin-condition dataset.

B. TESTING PHASE

In the test phase, the model was evaluated using precision, recall, and F1-score to assess its performance on unseen data. The classification report shows that the model achieved consistently strong scores across all four classes, indicating that it was able to generalize well beyond the training samples. The metrics reflect that the model effectively identifies fine-grained skin-condition features even in challenging test images.

The overall test accuracy of 83% demonstrates that the proposed EfficientNetB0-based model performs reliably and maintains a good balance between precision and recall. Although the accuracy is slightly lower than the training accuracy, it still represents a solid improvement over traditional approaches and confirms that the model retains strong predictive capability during real-world evaluation. These results validate the effectiveness of the two-stage training strategy and the suitability of the model for skin-condition classification.



Classification Report:

	precision	recall	f1-score	support
wrinkles	0.81	0.84	0.82	25
dark spots	0.90	0.76	0.83	25
puffy eyes	0.84	0.84	0.84	25
clear skin	0.79	0.88	0.83	25
accuracy			0.83	100
macro avg	0.83	0.83	0.83	100
weighted avg	0.83	0.83	0.83	100

Figure 8: (a) Confusion Matrix and (b) Classification Report of the Training Phase

The test phase results show that the model performs consistently well across all four classes, achieving an overall accuracy of 83%. The confusion matrix indicates that most samples were correctly classified, with only a few minor misclassifications. The precision, recall, and F1-scores remain balanced for every class, demonstrating strong generalization on unseen data. These results confirm that the EfficientNetB0-based model is reliable and effective for real-world skin-condition classification.

8. Conclusion

The project successfully delivered an effective deep learning–based system for automated skin condition classification. Through systematic preprocessing, model design, two-stage training, and comprehensive evaluation, the model demonstrated strong accuracy and consistent performance across categories. The integration of fine-tuning and optimized feature extraction significantly enhanced prediction quality, while stability in both training and validation metrics confirmed the robustness of the approach. Overall, the work highlights the potential of AI-driven dermatological analysis to provide reliable, accessible, and efficient support for skin health assessment in real-world applications.

9. Future Work

Future improvements can further enhance the performance and applicability of the proposed skin-analysis model. The network can be expanded to handle higher-resolution RGB dermatology images for more precise feature extraction. Future versions may also incorporate video-based skin scanning to analyze dynamic facial changes, enabling real-time condition tracking. Integrating 3D facial depth data could help the model better understand texture-related skin issues such as wrinkles and puffiness. Additionally, experimenting with advanced architectures and ensemble methods may further boost accuracy and robustness. Finally, deploying the model within a mobile or web-based application can make skin-health assessment more accessible to users worldwide.