Skin Disease Detection and Classification Based on Deep Learning Techniques

Rose aljehani 2111206 , Reem Alattas 2110378
Sjoud Alqhtani 2112034, Samaher Alharthi 2110632
Afnan Aljohani 2008745
Supervisor prof. Araek tashkndi

Abstract

skin disorders pose a significant worldwide health issue and timing is important in diagnosing, this study introduces a deep learning framework for the automated classification of skin diseases, utilizing a dataset of 30,000 dermoscopic images that includes seven categories: acne, eczema, psoriasis, melanoma, basal cell carcinoma, seborrheic keratosis, and normal skin. Three advanced architectures, MobileNet, a Hybrid EfficientNet-ResNet model and a Vision Transformer (ViT) were utilized with uniform preprocessing and augmentation techniques to guarantee a superior comparison. Among these models the Vision Transformer got the highest overall performance by achieving an accuracy of 97.38% showing its capability to effectively capture complex visual representations compared to traditional convolutional neural networks. this top performing model was implemented through a Streamlit based web interface, enabling users to upload images and promptly obtain classification results along with confidence scores. This research illustrates how deep learning can improve skin disease diagnosis and acts as a basis for clinical decision assistance especially in remote or resource poor areas

1. Introduction

Skin diseases are and have been a major concern of public health as they are a major worldwide affect an estimated 1.9 billion people globally representing the fourth-largest cause of non deadly illness, apart from physical symptoms dermatological conditions affect one's psychological health and may result in depression. Studies have revealed that these effects can greatly affect a person's quality of life [1] and achieving the best diagnosis is crucial. We focused on the six most common skin diseases: melanoma, eczema, psoriasis, basal cell carcinoma (BCC), acne, and seborrheic keratosis. These conditions vary in severity and are presented in detail in the following discussions. Firstly, Melanoma known as malignant melanoma, being the deadliest of skin cancer types with high mortality rates [2] .Secondly, Eczema makes daily clothing a hard task, given that it's a chronic skin condition that causes dry, itchy, and inflamed patches of skin. It commonly affects areas such as the elbows and knees and may be triggered by environmental factors or allergies. Another disease is psoriasis, which may cause patients to face psychological and social dilemmas as it's a growth of skin cells forming thick, red, and scaly patches on the skin. And can be associated with itching, pain, and inflammation. Another important condition included in this study is basal cell carcinoma (BCC), one of the most common cancers in the world originating in the outer layer of the skin [3].Also, acne is the most common skin condition among them all.This condition can psychologically devastate individuals who are affected by it, particularly teenagers. That said, 95% of patients experienced significant emotional and psychological distress When they experienced acne. Finally, we will also include seborrheic keratosis. This condition is frequently misdiagnosed due to its superficial similarity to either melanoma or moles which can cause enormous disturbance to such patients and potentially harmful medical procedures [4] .

Despite the availability of effective treatments, traditional diagnosis relies on doctor consultations, where dermatologists examine the skin and prescribe appropriate treatments. However, access to medical professionals is not always possible, especially in remote areas. Patients in these areas frequently experience long waits in receiving medical consultations, which can lead to worsening symptoms, higher healthcare costs, and treatment complexity [5]. Manual diagnosis relies on the doctor's expertise which plays the most important role in the diagnosis of patients, not every physician has the same level of expertise or specialization. In some cases, certain skin conditions are misdiagnosed or diagnosed inconsistently because of limited experience or the complexity and rarity of their conditions. Such variation in expertise may lead to miss diagnoses, making it prone to human error and resulting in inaccurate diagnoses [6] .

To address these challenges, this research will work toward meeting the following objectives:

• Bulid a deep learning model for detecting and classifying six skin diseases using real-world images. The aim is to harness the power of computer vision to detect and identify visual patterns suggestive of early skin disease signs for quicker and more reliable detection of skin conditions.

•Design a user-friendly application capable of allowing individuals, to upload images and obtain quick diagnose , this way our system will be meaningful and usable by non-specialists.

•improving classification accuracy; we are looking to reduce bias with respect to each disease category through careful preprocessing steps, data augmentation, and model optimization. By doing this; we will improve the quality of input to the model as well as the fine-tuning of model parameters to improve the reliability of predictions..

Building on these objectives, we will present several contributions that improve the efficiency and reliability of automated skin disease diagnosis, we will start by:

• Generalizing Disease Classification: Unlike previous studies, we will develop a deep-learning system capable of classifying and detecting the six most common skin diseases based on advanced deep-learned techniques.

• Improving Healthcare Accessibility: The system facilitates early diagnosis, particularly in remote areas, ensuring better healthcare access and timely medical intervention.

• Enhancing model accuracy: Some researchers address the low accuracy (as low as 60.5%) observed in some previous models [7] by implementing advanced deep-learning techniques to improve performance. After conducting a comprehensive comparison of several deep-learning models for skin disease classification, results showed significant variations in accuracy and performance. Based on this analysis, MobileNet emerged as the most efficient model, demonstrating the highest accuracy and balanced performance across all disease categories it outperformed other models in terms of lightweight architecture and speed, making it the optimal choice for real-world applications, particularly in scenarios requiring high efficiency and fast processing. However, despite its superior performance, further improvements in accuracy are possible. This can be achieved by fine-tuning hyperparameters and expanding the dataset.

• implementing Data Augmentation: techniques to increase the diversity of training data and improve the model's generalization.Moreover, Ensemble Learning can be explored, where MobileNet is integrated with other models to leverage their strengths, ultimately boosting prediction accuracy and system performance.

• Treatment Recommendations Beyond diagnosis, the system provides treatment suggestions, offering users an initial step in managing their condition before consulting a specialist.

These contributions enhance the efficiency, accuracy, and accessibility of skin disease diagnosis, ultimately improving patient outcomes and healthcare solutions.

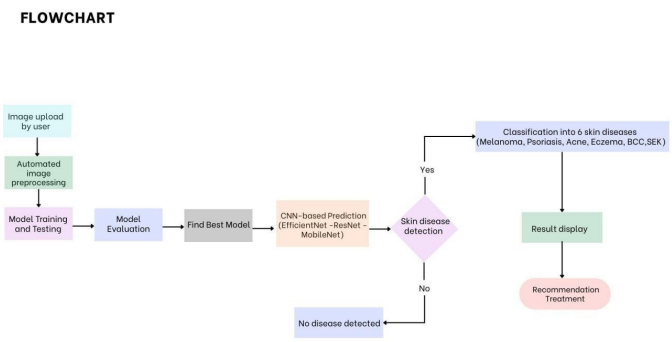

Figure 1: **Workflow of the proposed skin disease classification system** - This figure shows the main steps the system takes to help users check their skin health. After an image is uploaded, it's automatically processed and analyzed using deep learning. If a condition is found, the system identifies the type of skin disease and suggests a possible treatment.

2. Related Work

Deep learning has opened new possibilities for classifying skin diseases, providing much-needed assistance to dermatologists who often face challenges while diagnosing diseases that look alike. This section presents some previous important deep learning studies for skin disease classification, where we analyze the methodologies used, the datasets used..

[7]Aziz and Saputri developed this efficient approach for skin disease detection with the YOLOv9 model. The study aimed to classify six commonly known skin diseases: acne, atopic dermatitis, keratosis pilaris, leprosy, psoriasis, and warts. It is characterized by the integrating deep learning methods adopted in PGI and GELAN to improve performance and efficiency. The training was conducted on a dataset with 2,721 images, 288 images for validation, and 145 images for testing, attaining an MAP of 81.4%. However, a few issues were found, such as difficulty in detecting certain diseases, such as atopic dermatitis and psoriasis, they also used a small dataset size, which influenced generalization; class imbalance, where some diseases received less attention than others; and data quality issues, such as labeling errors and image noise. Our project addresses this by adopting MobileNet as an efficient, lightweight model as compared to YOLOv9, larger and more diversified datasets for better prediction outcomes, and employing data-augmentation techniques to balance the classes.

[6]Rangaswamy et al. used deep learning to classify skin diseases. Two well-known CNN architectures InceptionV3 and VGG16-formed the backbone structures of the classification of 13 skin diseases including acne, eczema, melanoma, psoriasis, urticaria, and vascular disorders. The performance of InceptionV3 was higher than VGG16 in terms of classification, with InceptionV3 rendering an accuracy of 80.88% compared to 74.17% for support vector machines. Some major challenges were with image processing, since problems related to image quality would significantly affect the

model's performance and accuracy. Many images had inconsistent lighting, noise, and blurry details, making it difficult for the model to detect significant features with precision.Our project overcomes these challenges with Normalization techniques, for instance, we will use pixel values within a specific range (e.g., [0,1]). By normalizing the data in that way, the process achieves enhanced stability during training. Other measures include some notable noise reduction methods like Gaussian Blur and Median Filtering that were adopted to achieve the reduction of noise, ultimately increasing feature extraction accuracy.

[1] Hemavathi and Velmurugan used CNN model for diagnosing 3 skin diseases, melanoma, eczema, and impetigo, achieving 91% accuracy. PC-based system was developed in this study and proved very effective, particularly in remote areas. The model will use image processing techniques and deep learning techniques to analyze images and give a proper diagnosis. Research showed certain limitations, most importantly, the classification of three diseases only, and no mobile application was provided for easy access. In our research, we will be expanding the classification approaches, furthermore, it will be integrated with a mobile-based system that enhances accessibility and usability.

[5]Raju et al. proposed a CNN-based classification model for the identification of five classes of skin diseases-acne, eczema, melanoma, psoriasis, and urticaria-training a model using a training dataset of 5,633 images. The research derived an efficiency measure of 82% classification accuracy with their model inwards, emphasizing the application power of deep learning in skin diagnostics. This research based the classification of skin diseases on convolutional neural networks (CNN), which enables the model to learn and detect patterns in images very efficiently. The study had some major limitations as it excluded major skin diseases such as basal cell carcinoma and seborrheic keratosis, while the classification accuracy of 82% was acknowledged to be lesser than some models available now, which depicts a slot for improvement. This research attempts to fill these gaps by including basal cell carcinoma and seborrheic keratosis to enhance the classification range while improving accuracy with better diversity of datasets and advanced preprocessing techniques.

[10] Kalaivani A. and Karpagavalli S. classified various forms of skin diseases using the latest Artificial Intelligence techniques. They developed a hybrid model called RF-DCNN that merges Random Forest with Deep Convolutional Neural Networks. This hybrid model achieved a classifying accuracy of 96.1% on the ISIC 2019 dataset. They further tested a traditional CNN, which achieved a much lower accuracy of 88%.Despite these promising results, their study was limited to the classification of only cancerous diseases. We are extending our research by including six other forms of diseases into our model. Three of these diseases are immune-related, namely, eczema, psoriasis, and acne; the rest, as mentioned above, are cancerous. Therefore, we can address all these diseases as a common model with high efficiency and performance using the lightweight MobileNet architecture.

[9] Spyridonos et al. and Ibraheem et al. aimed at skin cancer detection and established the efficacy of merging deep learning along with better extraction techniques to enhance skin disease classification. Spyridonos focused on melanoma and SK-like lesions and achieved, with the use of VGG16, ResNet50, and SVM classifiers, an 80.7% accuracy in the diagnoses of 978 images, whereas Ibraheem focused on melanoma and basal cell carcinoma (BCC) with GBT and an accuracy of 97.5using 150 ISIC images. They both had limited disease and small datasets, affecting generalization. Our project expands classification to six diseases, utilizes a larger dataset, and applies data augmentation to enhance model diversity and generalization.

[11]Kuldeep Viadandio et al. integrated CNN, KNN, and SVM models and achieved 97% accuracy in classifying various diseases like eczema, psoriasis, melanoma, and basal cell carcinoma. Others like GMM and TGMM have their accuracy reach 93% and 97%, CNN DenseNet had 95%. Additionally, Dragonfly Optimization Algorithm was used for selecting important features. Although the accuracy was still high, the lack of real-world applications limited usability. Our project addresses this by integrating six diseases into a MobileNet model while providing treatment recommendations for practical applications.

[12] Hammad et al. utilized the same dataset we will be using in our research. However, we will enhance the dataset with additional sources of data to further improve the performance of our model. Their study proposed a deep learning-based model called Derma Care, through which they achieved a diagnosis accuracy of 96.2% on eczema and psoriasis. One of the key strengths of their research is the model's ability to detect multiple skin diseases at once, backed by a huge amount of data, which heightened its scope for accuracy. The model was also intended to run on smartphones to make it more accessible and provide treatment ata time. The study, however, had some limitations. It dealt only with two skin diseases, eczema and psoriasis, and also they did not include treatment recommendations after diagnosis. Our research, however, intends to expand this and include four more categories: acne, melanoma, basal cell carcinoma, and normal skin. We pursue to enhance the system through the augmentation of treatment recommendations together with the diagnosis to provide broad support to our users.

[13]Krishnan et al. developed a Deep Forest based Classification model that highlights the superiority of the proposed model over previous models such as CNN and GAN, which demonstrates its efficiency in classifying skin cancer with an accuracy of 96.03% on ISIC-2019 and 96.36% on HAM10000. However, they targeted skin cancers only and they also didn't use data augmentation or deep learning improvements, which can help with low-quality images and rare cases. Our project

Citation	Purpose of this research	Model used	Best Models	Accuracy	Dataset
[7]	predicted six diseases: acne, atopic dermatitis, keratosis pilaris, leprosy, psoriasis, and warts.	YOLOv9	YOLOv9	71%	private
[6]	predicted thirteen skin diseases: acne, eczema, melanoma, psoriasis, urticaria, and vascular disorders.	VGG16, InceptionV3	InceptionV3	80.88%	public [14] [15]
[5]	predicted five types of skin diseases: acne, eczema, melanoma, psoriasis, and urticaria.	CNN	CNN	82%	private
[1]	predicted three skin diseases: melanoma, eczema, and impetigo.	CNN	CNN	91%	public [16]
[8]	predicted two diseases: Melanoma (MM), Basal Cell Carcinoma (BCC)	Gradient Boosting Trees (GBT)	GBT	97.5%	public [17]
[9]	predicted two diseases: Melanoma (MM), SK-like lesions (BKL)	ResNet50, SVM, VGG16	VGG16	80.7%	public [17]
[10]	predicted skin cancer.	RF-DCNN	RF-DCNN	96%	public [14]
[11]	Predicted eczema, psoriasis, melanoma, and basal cell carcinoma (BCC)	TGMM, RL, CNN, Meta-Learning	Meta-Learning	97%	public [14] [16] [17]
[12]	predicted eczema and psoriasis	Derma Care	Derma Care	96.20%	public [18]
[13]	predicted skin cancer	Deep forest classification	Deep forest classification	96%	public [14] [17]
[22]	predicted three skin diseases Psoriasis, Eczema, and Atopic Dermatitis	MobileNet	MobileNet	95.7%	Public [16] [26]
[23]	predicted five skin diseases: Acne, Blister, Cold Sore, Psoriasis, and Vitiligo	MobileNet	MobileNet	90%	private
[24]	predicted Psoriasis	Vision Transformer	Vision Transformer	97.53%	Public [14]
[25]	predicted skin cancer (benign vs malignant)	Vision Transformer	Vision Transformer	96.15%	Public [18]

Table 1: Summary of Research on Skin Disease Detection using AI Models

overcomes these challenges through a wider scope of other conditions, and we will also add data augmentation and use better deep learning methods to improve accuracy.

[22] AlSuwaidan (2023) explored the effectiveness of various convolutional neural network models for dermatological image classification, focusing on three skin conditions: eczema, psoriasis, and atopic dermatitis. By applying the BM3D algorithm to enhance image clarity before classification, the study demonstrated how image pre processing can significantly improve diagnostic accuracy. Among the tested models VGG16, EfficientNet, InceptionV3, NasNet, and ResNet50 But MobileNet achieved the highest accuracy at 95.7% , highlighting its strength as a lightweight yet high-performing architecture.

[23] Similarly,Nath et al. (2023) investigated deep learning models for classifying five different skin diseases, including acne, blisters, cold sores, psoriasis, and vitiligo. Their use of data augmentation to address class imbalance allowed the models to generalize better, resulting in a validation accuracy close to 99% and a testing accuracy of around 90%. MobileNet again emerged as one of the most effective models, demonstrating an exceptional balance between accuracy, speed, and computational efficiency. While these two studies have shown the strong potential of MobileNet in dermatological diagnosis, they were limited in scope to a small number of skin conditions. Our research builds upon these findings by expanding the diagnostic capability beyond three or five conditions to include six distinct skin diseases as well as normal skin. This broader classification system provides a more realistic and clinically relevant solution, reflecting the complexity of real world dermatology practice and making our model more practical.. Advancements in recent years have begun moving beyond traditional CNNs toward transformerbased models.

[24]Himel et al. (2024) adopted a Vision Transformer architecture integrated with a segmentation model (SAM) to classify skin cancer lesions from the HAM10000 dataset. The ability of ViT to capture long range spatial dependencies enabled it to achieve an accuracy of 96.15%, outperforming many conventional CNN based models.

[25] Mashal et al. (2025) applied the transformer based architecture to the classification of psoriasis, achieving an accuracy of 97.53% and an F1-score of 0.98. These studies demonstrate how transformer models are redefining dermatological image analysis by providing deeper feature extraction and improved classification performance. Inspired by this emerging direction, our project extends the application of Vision Transformers beyond single-disease classification by using this architecture to classify six skin diseases along with normal skin. By doing so, our research bridges the gap between accuracy and clinical applicability, demonstrating that ViT is not only powerful in theory but also scalable and effective when applied to diverse dermatological conditions in a single, unified diagnostic system.

Based on a comparison with ten previous studies, we established that most of the previous work was either on cancerous skin diseases only or on fewer than six common skin diseases. In our research, we aim to predict six skin conditions, three of

which are immune-related eczema, psoriasis, and acne, and are most common among young people, and three cancerous conditions, melanoma, basal cell carcinoma, and seborrheic keratosis that are often difficult to distinguish. What is unique to our research is its focus on these six target diseases to produce a more equitable and broader coverage of immune-related as well as cancerous disease conditions that commonly occur in everyday practice. Our contribution also extends beyond classification. We will design an easy-to-use mobile application that allows one to quickly read and use the system easily. In addition to offering accurate diagnoses, the application will give users initial treatment advice to help them before seeing a medical professional.

3. Methodology

3.1. Dataset understanding & exploration

Before we get into training models, we recognized that preparing the data set properly was going to be one of the most important parts of the project. With a project involving skin disease images, there were many different factors to consider image size, image quality, the lighting that accompanied the photos, and the color distribution in the image itself. We spent time researching and cleaning the dataset, so we could be sure that the data that we were training on was as consistent and clear as possible. And now we will mention important steps we took.

1.Image Size Inspection

The first data point that we looked at was image size. Skin disease images can vary in size, and we knew image size mattered when image dimensions and pixel density were inconsistent because that would potentially lower model performance. Smaller image sizes can miss important visual characteristics, and larger image sizes can affect training speed across epochs and reduce our digital accessibility in a study. To understand the set of images that we were working with, we used os.path.getsize(), which gives you the original dimensions and file size, to document the size of each image. This gave us the information we needed to standardize all of our inputs to a set size later, which provided a level of consistency across the dataset.

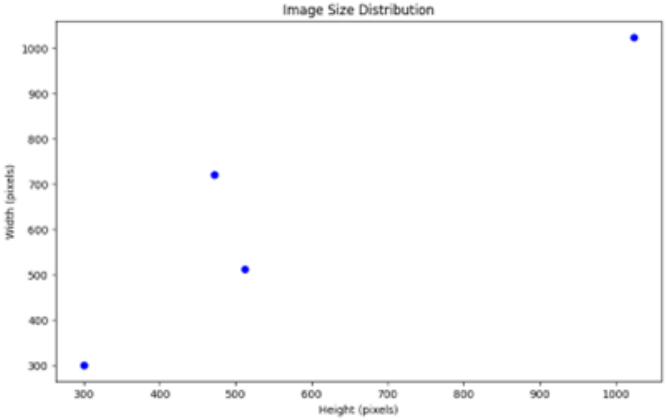

Figure 2: **Scatter plot showing the distribution of image sizes (height vs width) in the skin disease dataset.** .

2. Image Count per Category

One of the important first steps before analysis or model training was to determine how many images we had for each category of skin disease. This knowledge was critical to reveal, for the dataset, whether the data was balanced among the different skin diseases, or if one or more classes were overrepresented. We accomplished this with the use of os.listdir() which allowed us to count the number of images in each folder or directory. This step allowed us to provide a clearer picture of the structure of the dataset and to demonstrate where data augmentation should likely be applied to provide balance.

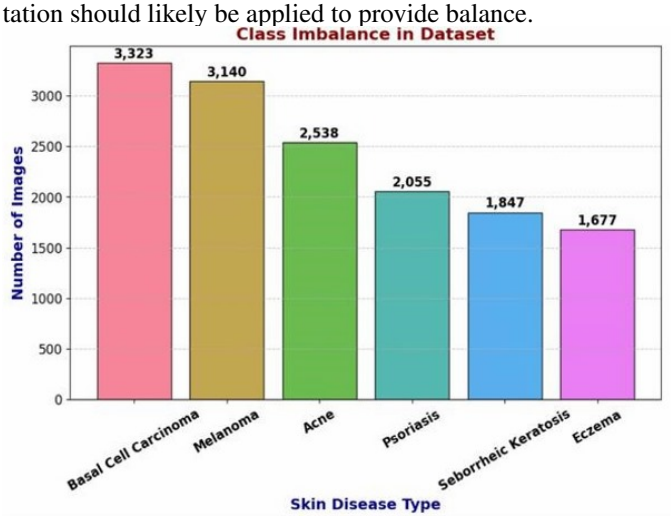

Figure 3: **Class imbalance among different skin disease categories in the dataset** .

3. Color Distribution Analysis

We were interested in visualizing the color distribution in our dataset because several skin conditions have unique color characteristics. For each image we computed the average values for the red, green, and blue (RGB) channels, and presented those averages with plt.hist(). While exploring the color distributions helped us see what possible color patterns existed across different disease types, it also helped us determine what the model might learn from the color patterns.

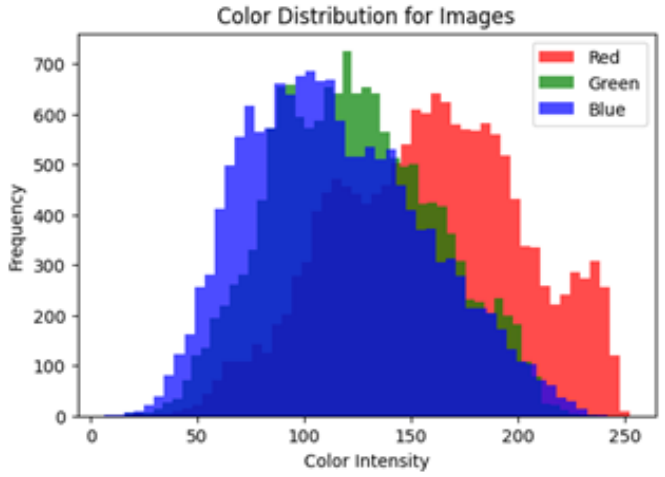

Figure 4: **Color intensity distribution across the Red, Green, and Blue channels in the skin disease images dataset** .

4. Image Clarity (Sharpness and Blur Detection)

Since we are working with medical pictures this means that one of the most important things is how clear the images in our dataset are. Sharp images are essential for enabling the model to detect fine details such as edges and textures. To evaluate this, we used Laplacian variance[figure5] to measure sharpness and identify blurry images[figure6]. As we reviewed the results, we noticed that a large portion of the images had low sharpness scores, and several appeared noticeably blurry. This raised concerns about the model's ability to learn effectively from these samples. Based on this, we decided to flag all low-sharpness images for either enhancement or removal, ensuring that the model would be trained on data with sufficient visual clarity.

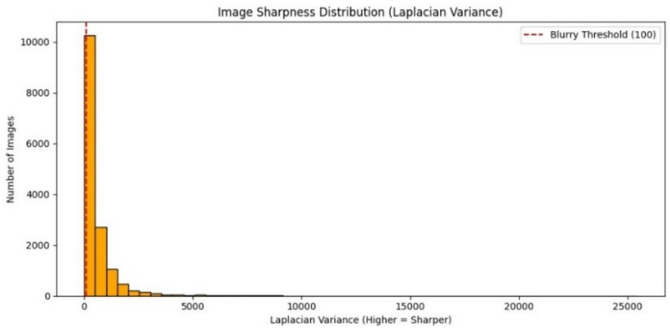

Figure 5: **Distribution of image sharpness based on Laplacian variance, with the blurry threshold indicated at 100** .

Figure 6: **Correlation matrix showing the relationships between different image quality metrics, including sharpness, brightness, contrast, noise, illumination variability, and texture** .

5.Destribution of image Sharpness and Noise

Once we examined sharpness, we dedicated time to examining the overall image quality place we spent time looking into noise, which could occur as pixel fluctuations that could reduce clarity and mislead the model. For our purpose, we took the standard deviation of grayscale values to estimate noise levels for each image. Most images were approximately moderate in noise (15 to 50) and deemed acceptable, but few images were higher than moderate/noise with values above 60. We then flagged these images for further investigation and possible denoising. The denoising pathway or step made us aware of a quality issue that we likely would not have flagged on our own, but nonetheless could potentially impact training performance.

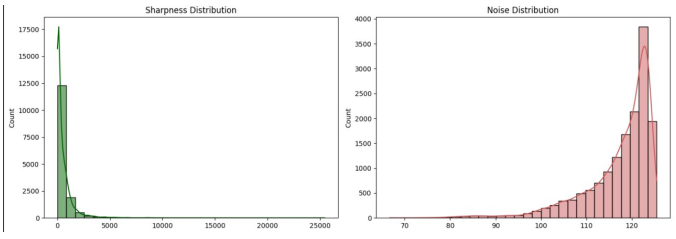

Figure 7: **Distribution of sharpness (left) and noise levels (right) across the skin disease images** .

6.Background and Lighting Consistency

During our dataset exploration, we noticed an imperfect lighting configuration or a distracting background in some images. These factors can pull the model's attention from the skin itself and decrease classification reliability.

As a first step, we ran a check for illumination variability across image regions and an analysis of background noise using

cv2.Laplacian for visual noise. Our suspicions were correct: some of the images contained shadows, glare, or were rather busy in the background. In response, we tagged those images for normalization or cropping to prioritize the model's attention towards the provided features alone.

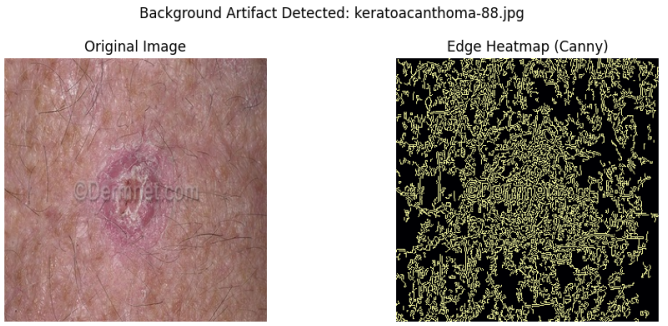

Figure 8: **Detection of background artifacts using Canny edge detection. The original image (left) and the corresponding edge heatmap (right) highlight the presence of unwanted background elements** .

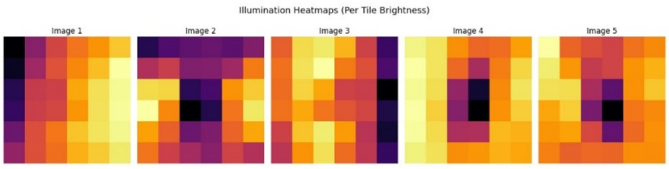

Figure 9: **Illumination heatmaps showing per-tile brightness distribution for sample skin disease images** .

7.Destribution of image Brightness and contrast

To gain greater insights into the clarity, and visual balance of the data set, we explored both brightness (average pixel intensity) and contrast (standard deviations of intensity values)[figure10]. The histogram of brightness was relatively normal, suggesting that many of the images had good lighting. However, there were a number of images that were either too bright or too dark and this may hinder the model's ability to learn relevant features.

In terms of contrast, many images fell into a low to medium contrast range which means there are some dark images[figure11]. Although this is somewhat expected in skin imagery, it also suggested to us that images may not present the color or texture differences sharply enough to support the model's learning..

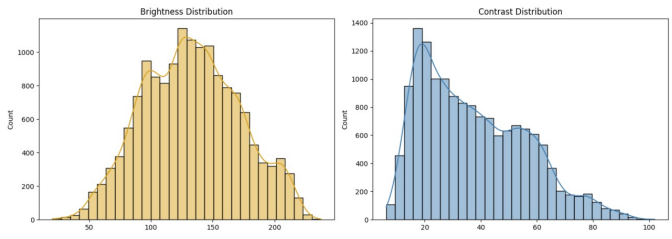

Figure 10: **Distribution of brightness (left) and contrast (right) values across the skin disease image dataset** .

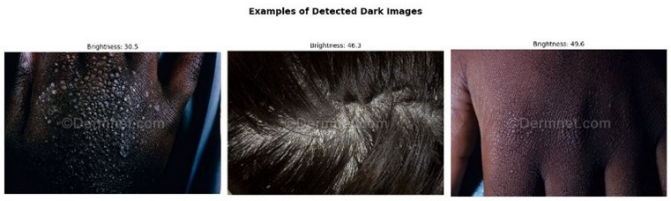

Figure 11: **Examples of detected dark images in the dataset, with brightness levels indicated for each image** .

8.texture

As skin conditions generally present a texture on the surface, we wanted to make sure our dataset included enough texture variability. We looked at edge density through a Sobel filter to estimate texture richness. In general most of the images contained sufficient texture, some however had a very low presence of edge which may suggest problems such as blur or flat lighting, and were considered lower importance for training, or in higher quality patterns considered for enhancement.

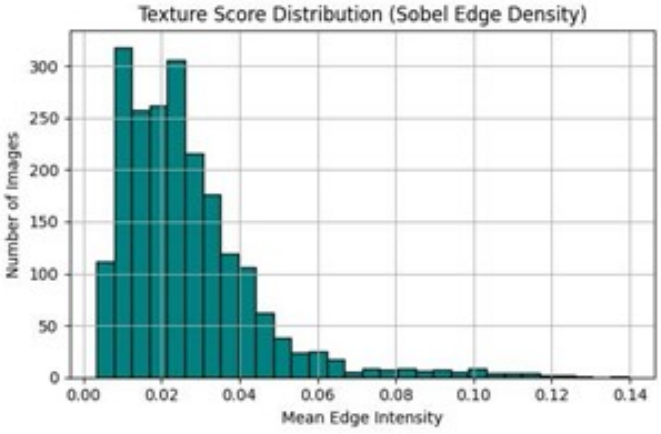

Figure 12: **Distribution of texture scores based on Sobel edge density, showing the variation of mean edge intensity across the dataset**.

Final Thoughts

After fully examining all the visual quality of our dataset, we found many issues like low levels of sharpness, disparate lighting, high levels of noise, etc., which we thought could have affected model performance. For that reason, we did not treat image quality simply as an inconvenience, but we took it into consideration to guide how we ultimately shaped our dataset and what informative, consistent, clean samples we trained on. Ultimately, this process provided us with a better foundation to build the model on and increased our confidence in the reliability of the learning process moving forward.

3.2. Preprocessing

Before training any deep learning model, we knew that data quality is everything. Our raw dataset of skin disease images came with many imperfections: inconsistent brightness, blurry

details, dominant backgrounds, varying image sizes, and visible noise. These issues could confuse even the most advanced CNN models. So, we began preprocessing. We applied a total of 15 detailed steps to prepare the images for training. Each step was carefully selected based on the best practices for medical images that were supported by research.

We started by removing duplicate images using imagehash, then resized all images to 300×300 pixels, a standard size compatible with ResNet, MobileNet, and EfficientNet. The decision to resize the images 300×300 was not spontaneous, rather it was a difficult, but intentional decision based on the characteristics of the skin diseased images and good results in previous studies.

When building our model pipeline, we examined the unique nature of dermatological data. Skin diseases often have small differences in tiny visual cues, such as slight variations in texture or small shifts in pigmentation. These details are important for accurate classification. Resizing down to 224×224 pixels, even though its standard, runs the risk of losing some of these small features. Since our task is clinically sensitive we needed the tiny features and we wanted to keep every bit.

Our decision became stronger after recent work on skin disease classification that ignored some assumptions about size. Srinivasu et al. [20]experimented with MH and LSTM using a MobileNetV2 style from the HAM10000 dataset. Rather than using 224×224 input size, they resized all images to 300×300 pixels, believing that higher input resolution would offer better feature extraction. Their results confirmed this assumption: high classification accuracy was achieved without negatively impacting performance. The key was fine tuning and they also froze the lower convolutional layers of MobileNetV2 (which learn basic visual patterns like edges and colors) and retrained the higher-level layers and classifier head, adapting the model to skin lesion-specific features.

Velasco et al. [21]examined MobileNet and ResNet50 using larger image input sizes of 300×300 specifically to classify seven types of skin diseases. They used transfer learning and fine tuned the upper layers of MobileNet and ResNet50 obtaining an accuracy rating of 94.1%. Their study also confirmed that smaller input original trained models (like 224×224) could perform well and in some cases even better than utilizing larger high resolution images.

Building on these insights, we confidently standardized all our input images to 300×300 pixels. This allowed us to implement a consistent and standardized preprocessing pipeline around EfficientNet-B3, our baseline model, as this model is optimized for that input size. Additionally, this input size maintained compatibility with other architectures such as MobileNet and ResNet. We were able to maintain the key detail in the images needed to classify skin disease while also standardizing the data flow across all of the models, which enabled us to balance accuracy, efficiency, and scalability. After we resized the images,

the next step in improving image quality was manipulate brightness, contrast, and sharpness through CLAHE on the L channel of the LAB color space to improve contrast, reveal delicate skin patterns, and improve the visibility of the disease in the image. We eliminated images that were too dark, blurry, or low in texture to ensure the model only learned from meaningful and clean data. Next, we moved into noise reduction, where images were reviewed for clarity, as well as signal strength. Images with high noise or poor focus were filtered out using a sharpness score and a texture score To enhance the dataset, we applied considerable data augmentation using the Albumentations library to produce varied images until it reached 4,000 images per disease class. In addition to balancing out the dataset, the models were also able to generalize better to unseen patterns.

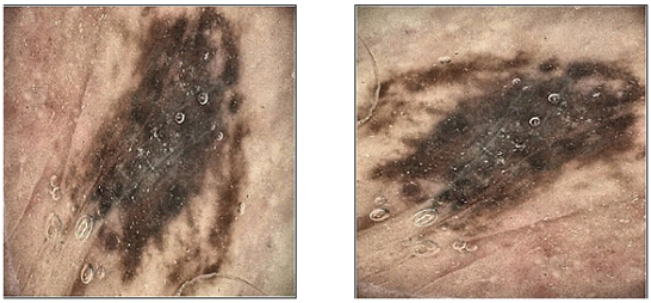

Figure 13: **Illustration of Data Augmentation. This figure illustrates the data augmentation process by showing an original image (before) and its corresponding augmented version (after) for visualization purposes.** .

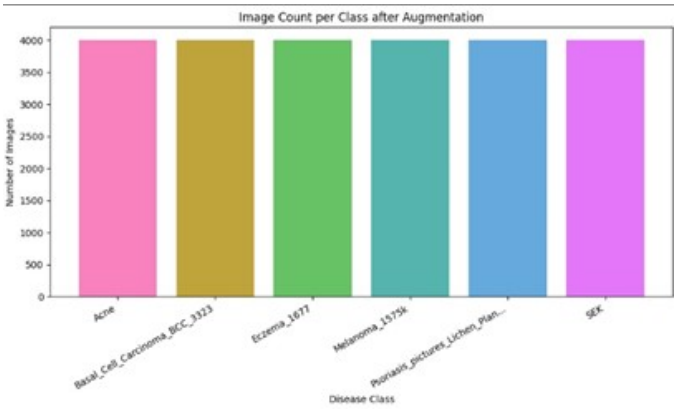

Figure 14: **Image count after augmentation All disease classes were balanced to exactly 4000 images each to ensure data uniformity and improve model performance.** .

At the end of our process, we put together a before and after grid. Each disease class has one representative image that conveys a clear representative example of the remedies described above. The results were very impressive We now have clean borders, sharper lesion areas, less background noise, and a more uniform visual structure across the full dataset.

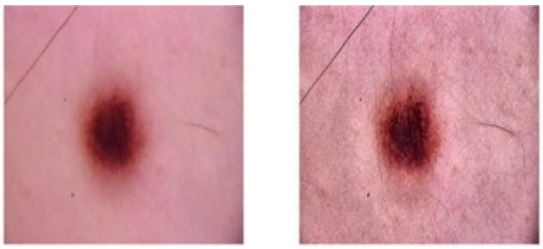

Figure 15: **Melanoma** - Before vs After
• Before: The lesion is slightly blurry with low contrast, mak ing it hard to distinguish edges.
• After: The lesion is sharper, the skin texture is clearer, and the contrast is improved using CLAHE .

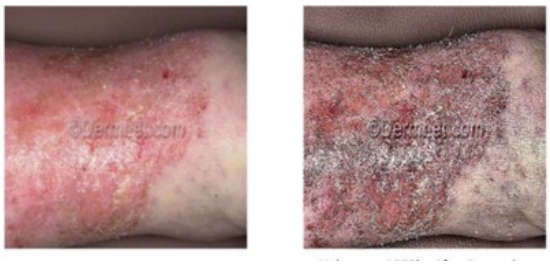

Figure 16: **Eczema** - Before vs After
• Before: High brightness and reflections obscure key skin features.
• After: Brightness normalized and noise removed, al lowing better visibility of texture and inflammation. .

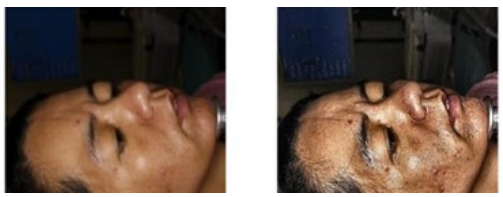

Figure 17: **Acne** - Before vs After
• Before: Uneven lighting and skin gloss obscure acne regions.
• After: Noise reduced and skin smoothed to highlight acne severity for better analysis.

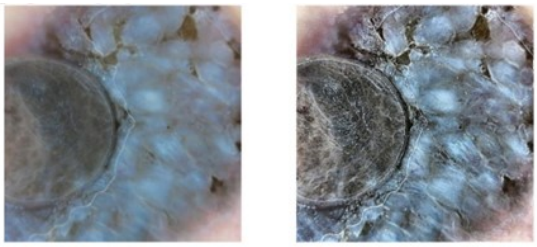

Figure 18: **Basal Cell Carcinoma** - Before vs After
• Before: The lesion appears low in contrast with a hazy tex ture. The edges between the lesion and surrounding skin are not clearly defined, which could limit the model's ability to learn discriminative features.
• After: Contrast and sharpness were significantly enhanced. Fine details of the lesion structure are now visible, and the tex ture is more defined, allowing better feature extraction during training

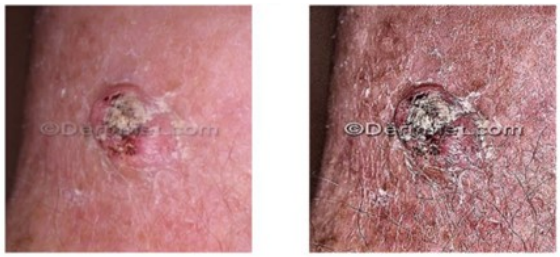

Figure 19: **SEK(Seborrheic Keratosis)** - Before vs After
• Before: The image has mild lighting and visible color soft ness, with moderate skin detail. However, the presence of back ground noise and lack of clear contrast reduces clarity.
• After: Noise was reduced and contrast improved. The lesion becomes more prominent and the surrounding skin texture is better defined, which helps highlight irregularities and sharp ens the lesion boundary

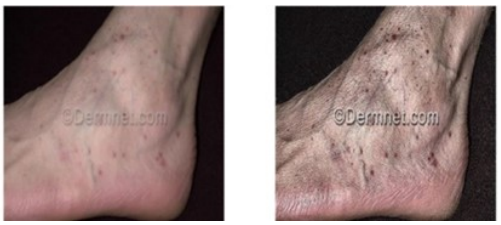

Figure 20: **Psoriasis / Lichen Planus and Related Diseases** - Before vs After
• Before: The image shows a generally low-contrast skin sur face with faint lesion visibility. The red patches are visible but subtle, blending with the skin tone and background.
• After: Color correction and edge sharpening techniques en hanced lesion clarity. The red patches and skin texture became more defined, supporting clearer identification of affected ar eas by the model

With this carefully constructed and detailed preprocessing pipeline, along with researched image size, we were able to set up strong building blocks for our deep learning models to perform at an optimal level, in terms of both accuracy and reliability.

4. Results and Discussion

4.1. Model Architectures

We developed three models To evaluate the effectiveness of various deep learning approaches for multiclass skin disease classification and fine tuned using the same dataset which included seven categories : acne, eczema, psoriasis, melanoma, basal cell carcinoma, seborrheic keratosis, and normal skin. These models went through standardized training , validation and testing procedures to ensure an effictive evaluation. In this section we will present the detailed design, training methodology, and performance outcomes of each model.

4.1.1. MobileNet Model

Mobilenetv2 was employed primarily due to its compactness and real time as well as web-based applicability. Initial total test accuracy of 71.63%. Although the model showed robust

performance across multiple classes its accuracy for psoriasis was low at 38% highlighting a lack of adequate feature representation for this less represented category. To overcome this we applied fine tuning where deeper convolutional layers were unfrozen to improve extraction of disease specific characteristics. Extra optimization methods like the AdamW optimizer with weight decay, cosine learning rate scheduling, and Exponential Moving Average were added to enhance model generalization and stabilize training dynamics. The accuracy of psoriasis classification went from 38% to 55% and the overall test accuracy increased from 71.63% to 81.84%, and the model reached a training accuracy of 76.10% showing effectiveness without signs of overfitting. These enhancements demonstrate that MobilenetV2, when combined with focused fine tuning.

4.1.2. Hybrid EfficientNet–ResNet Model

In order to enhance the performance of classification above that could be achieved through MobileNet, a combination of EfficientNet-B3 and ResNet-50 was developed.The approach employs EfficientNet's scaling efficiency and deep residual learning capacity of ResNet such that the model is capable of learning high-level semantic representations as well as high-resolution dermatological textures. Features from the pre trained backbones were extracted with Global Average Pooling and concatenated subsequently. A custom classification head consisting of fully connected layers with ReLU activation and dropout regularization was employed and followed by a Softmax layer at the end to generate the final predictions. Training employed the Adam optimizer at a learning rate of 1e-5 and early stopping in order to prevent overfitting. This integrated architecture achieved a training accuracy level of approximately 93.57% and a test accuracy of 88.92%. The model was found to be highly accurate in the detection of melanoma and eczema both at 98.20% making sure that it could generalize on varying types of lesions. Overall, the hybrid model was more balanced with greater accuracy and computation rate, achieving higher performance than single CNN models in most classes.

4.1.3. Vision Transformer (ViT) Model

The Vision Transformer (ViT) model, a transformer model to process long range spatial dependencies was employed as the third model. In contrast to the conventional CNNs, ViT avoids using convolutional operations to acquire complex inter-pixel relations and applies self attention mechanisms instead. The ViT-Tiny Patch16-224 configuration was selected since it attempted to strike a balance between computation efficiency and classification performance. Fine tuning consisted of three incremental stages:

1. **Head-only training:** Only the classifier head was trained by keeping the transformer backbone frozen.
2. **Partial unfreezing:** The remaining four transformer blocks also normalization layers and embeddings were not frozen, which allowed the model to learn its representations from the medical domain.
3. **fine tuning :** All the layers were not frozen and were trained with a reduced learning rate to achieve stable convergence.

Model was optimized using the AdamW optimizer in combination with label smoothing and data augmentation techniques to enable generalization and prevent overfitting. ViT achieved a decisive test accuracy of 97.38%, and precision, recall, and F1-scores over 97% across all seven classes. The model excelled by performing excellent ability for discriminating visually related conditions such as eczema and psoriasis, showcasing the superior performance of self attention compared to the standard convolutional approach.

4.2. Comparative Discussion of Model Performance

Comparative analysis between the three models revealed a clear pattern of performance improvement with architectural complexity. MobileNet, while effective as a lean baseline, was limited in capturing complex dermatological features, particularly for rare classes. The Hybrid EfficientNet–ResNet model made a great leap by achieving complementary characteristics of two robust architectures and enjoying overall balanced performances across disease classes. But the Vision Transformer outperformed both the CNN-based models by achieving a test accuracy of 97.38% since it is capable of learning fine-grained and contextual information through self attention. Briefly, although the hybrid model offers the best trade-off between computation and accuracy for real-time clinical deployment, the Vision Transformer made the most accurate diagnosis and thus is the best architecture for machine-learning-based skin disease classification in this work.

4.3. Model Deployment

Following model testing, the Vision Transformer was selected for release due to increased performance. The model was hosted as a web app with Streamlit, a Python library enabling rapid interactive machine learning interface development. The release involved publishing the trained ViT model and Streamlit script to a project repository and linking it with Streamlit Cloud through GitHub, enabling the model to be deployed automatically and continuously integrated. The final application allows the upload of dermoscopic images, which are processed by the ViT model to generate real time classification predictions with confidence scores. The interface was designed particularly user friendly and accessible so that the system not only suits dermatologists but can also be used in remote or less developed areas.

Figure 4, Figure 5, and Figure 6 present the confusion matrices for the three models — MobileNet, Hybrid EfficientNet–ResNet, and Vision Transformer (ViT). These matrices illustrate the class-wise performance of each model. ViT achieved the highest accuracy (97.38%), followed by the hybrid model (88.92%) and MobileNet (81.84%)

Disease	Accuracy	CI low	CI high
Acne	79.80%	76.28%	83.32%
Eczema	82.80%	79.49%	86.11%
Melanoma	98.60%	97.57%	99.63%
Normal skin	90.30%	87.63%	92.96%
Psoriasis	55.00%	50.64%	59.36%
Seborrheic keratosis	84.80%	81.65%	87.95%
Basal cell carcinoma	79.44%	75.90%	82.98%
Total Accuracy (MobileNet)	**81.84%**		

Table 2: Classification report of MobileNet model showing per-class accuracy, precision, recall, and F1-score. (Overall Accuracy: 81.84%)

Disease	Precision	Recall	F1-Score	Accuracy
Acne	91.51%	88.40%	89.93%	88.40%
Basal cell carcinoma	97.41%	82.44%	89.30%	82.44%
Melanoma	89.11%	98.20%	93.43%	98.20%
Eczema	84.51%	98.20%	90.84%	98.20%
Normal skin	97.36%	93.46%	95.37%	93.46%
Psoriasis	85.65%	76.40%	80.76%	76.40%
Seborrheic keratosis	80.00%	85.60%	82.71%	85.60%
Total Accuracy of Hybrid Model (EfficientNet–ResNet)	**88.92%**			

Table 3: Classification report of the Hybrid EfficientNet–ResNet model showing improved overall accuracy (88.92%) and balanced precision across all classes.

Disease	Precision	Recall	F1-Score	Accuracy
Acne	94.69%	99.80%	97.18%	99.80%
Eczema	99.20%	99.80%	99.50%	99.80%
Basal cell carcinoma	99.20%	99.40%	99.30%	99.40%
Melanoma	100.00%	98.40%	99.19%	98.40%
Normal Skin	99.77%	92.83%	96.17%	92.83%
Psoriasis	91.83%	96.60%	94.15%	96.60%
Seborrheic keratosis	97.73%	94.60%	96.14%	94.60%
Micro avg	97.49%	97.35%	97.38%	97.38%
Weighted avg	97.47%	97.38%	97.39%	97.38%
Total Accuracy of Vision Transformer Model	**97.38%**			

Table 4: Classification report of the Vision Transformer (ViT) model achieving the highest accuracy (97.38%) with superior precision and recall for all seven classes

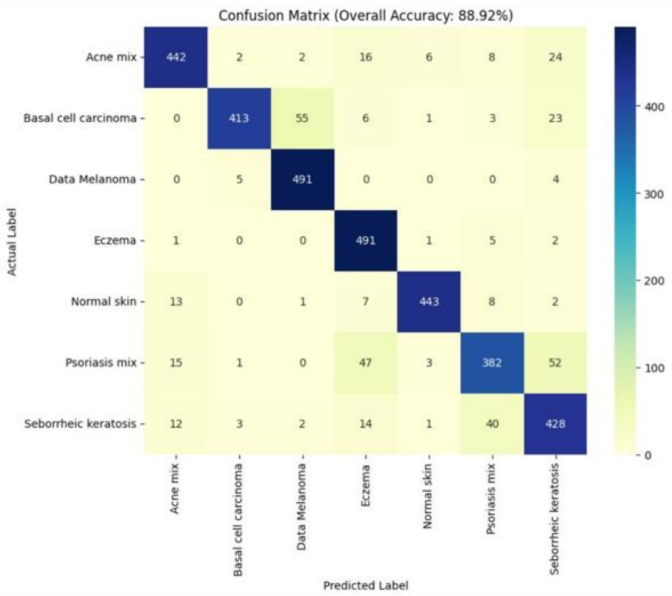

Figure 21: **Confusion Matrix –** MobileNet

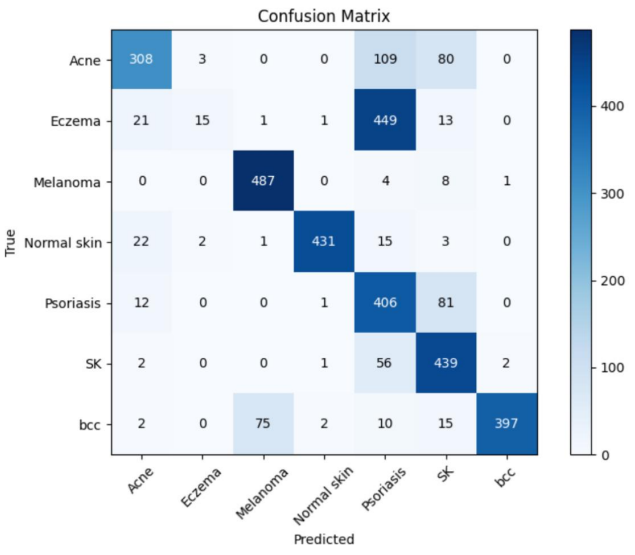

Figure 22: **Confusion Matrix –** Hybrid EfficientNet–ResNet

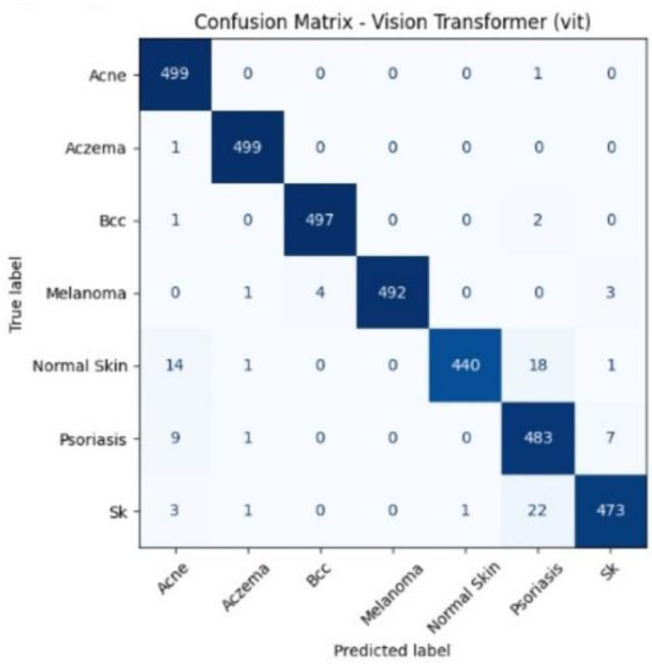

Figure 23: **Confusion Matrix –** Vision Transformer (ViT)

4.4. Conclusion and Future Work

This project presented the use and evaluation of a deep learning based diagnosis platform for machine skin disease classification using three models: MobileNet, Hybrid EfficientNet–ResNet model, and the Vision Transformer (ViT). Of them ViT achieved the best accuracy of 97.38% showing excellent skill in distinguishing among seven dermatological classes with high accuracy. The best performing model was subsequently used as a functional web application through Streamlit, providing users with real time diagnostic predictions following input of uploaded skin photos. This transition from research model to functional diagnostic machine shows the real world potential of deep learning on medical imaging. Automated medical report generation such as treatment recommendations, and scaling the system to other classes of disease at high levels of accuracy are possible avenues for future research. The long term aspiration is to secure clinical validation and regulatory approval, ultimately offering an inexpensive, accurate, and accessible method for the early detection of skin disease, particularly for remote or underserved individuals. Moreover, this technology could be used in hospital systems to assist dermatologists in clinical decision making and improve outcomes for patients. Overall, this research highlights the promise of artificial intelligence to be an agent of transformation in healthcare and closing the technology vs access gap in medicine.

References

[1] Hemavathi, S., & Velmurugan, K. J. (2021). *Skin Disease Prediction and Provision of Medical Advice Using Deep Learning. Journal of Physics: Conference Series, 1724(1), 012048.

[2] American Academy of Dermatology Association. Melanoma. J Med Inform Technol. 2025;2(1):11-15.

[3] Mayo Clinic. Basal cell carcinoma. J Med Inform Technol. 2025;2(1):11-15.

[4] Mayo Clinic. Seborrheic keratosis. J Med Inform Technol. 2025;2(1):11-15.

[5] Raju KS, Sharma A, Reddy NCS, Latha GS, Kumar GP, Vidyasagar G. Prediction and classification of skin diseases using convolution neural network techniques. In: Devi BR, editor. Proceedings of Fifth International Conference on Computer and Communication Technologies. Singapore: Springer; 2024. p. 403-411.

[6] Rangaswamy S, Tantry SS, Lal TS. Skin disease classification using deep learning. Natl Acad Sci Lett. 2024;2(3):1-10.

[7] Aziz F, Saputri DUE. Efficient skin lesion detection using YOLOv9 network. J Med Inform Technol. 2024;2(1):11-15.

[8] Ibraheem, M. R., & Elmogy, M. (2020). A non-invasive automatic skin cancer detection system for characterizing malignant melanoma from seborrheic keratosis. Proceedings of IEEE (pp. 1-8).

[9] Spyridonos, P., Gaitanis, G., Likas, A.,& Bassukas, I. (2021). Characterizing malignant melanoma clinically resembling seborrheic keratosis using deep knowledge transfer. Cancers, 13(24), 6300.

[10] Kalaivani, A., & Karpagavalli, S. (2022). Detection and classification of skin diseases with ensembles of deep learning networks in medical imaging. International Journal of Health Sciences, 6(S1), 13624–13637.

[11] Vayadande, K., Bhosle, A. A., Pawar, R. G., Joshi, D. J., Bailke, P. A., & Lohade, O. (2024). Innovative approaches for skin disease identification in machine learning: A comprehensive study. Oral Oncology Reports, 10, 100365.

[12] Hammad, M., Pławiak, P., ElAffendi, M., Abd El-Latif, A. A., & Abdel Latif, A. A. (2023). Enhanced deep learning approach for accurate eczema and psoriasis skin detection. Sensors, 23(16), 7295.

[13] Krishnan, V. G., Murthy, K. S., Kongara, S., Chowdary, K. U., Somaskandan, M., & Simla, A. J. (2022). A prediction of skin cancer using mean-shift algorithm with deep forest classifier. Iraqi Journal of Science, 63(7), 3200-3211.

[14] Kaggle. (2024). Skin Cancer MNIST: HAM10000. Kaggle.Retrieved from https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000

[15] Roboflow. (n.d.). Dataset source. Retrieved from https://www.roboflow.com

[16] DermNet NZ. (n.d.). DermNet NZ Image Library. Retrieved from https://dermnetnz.org

[17] International Skin Imaging Collaboration (ISIC). (n.d.). ISIC Archive. Retrieved from https://www.isic-archive.com

[18] IsmailPromus. (2021). Skin diseases image dataset [Dataset]. Kaggle. Retrieved from https://www.kaggle.com/datasets/ismailpromus/skin-diseases-image-dataset

[19] Billal Manzoor. (2023). FYP Skin Disease Dataset. Kaggle. [Online]. Available: https://www.kaggle.com/datasets/billalmanzoor/fyp-skin-disease-dataset/suggestions?status=pending

[20] Srinivasu, P. N., et al. (2023). Deep Learning-Based Skin Diseases Classification using MobileNet V2 and LSTM. Advanced Intelligent Systems, 5(3), 2300211. https://doi.org/10.1002/aisy.202300211

[21] Velasco, J. S., et al. (2023). Classification of Skin Disease Using Transfer Learning in Convolutional Neural Networks. *arXiv:2304.02852*. Retrieved from https://arxiv.org/abs/2304.02852

[22] AlSuwaidan, L. (2023). Deep Learning Based Classification of Dermatological Disorders. Biomedical Engineering and Computational Biology,14,1–9.

[23] Nath, S., Das Gupta, S., & Saha, S. (2023). Deep Learning-Based Common Skin Disease Image Classification. Journal of Intelligent & Fuzzy Systems, 44(6), 7483–7499

[24] Himel, G. M. S., Islam, M. M., Al-Aff, K. A., Karim, S. I., & Sikder, M. K. U. (2024). Skin Cancer Segmentation and Classification Using Vision Transformer for Automatic Analysis in Dermatoscopy-Based Noninvasive Digital System. International Journal of Biomedical Imaging, 2024, Article ID 3022192, 18 pages.

[25] Mashal, M., Ahmad, H. I., Ali, M., Ashraf, E., Alharbi, F., Javed, I. T., & Alkayal, E. (2025). Classification of Psoriasis Using Vision Transformer: A Deep Learning-Based Approach. IEEE Access, 13, 163885–163895.

[26] University of Iowa, Department of Dermatology. (n.d.). Dermatology Online Atlas. Retrieved from https://dermatology.medicine.uiowa.edu/