

DermFusion-CBAMNet: A Dual-Backbone Attention-Guided Framework for Robust Multiclass Skin Disease Classification

Padma Jyothi Uppalapati^{1*}, Sridevi Bonthu¹ and Leelavathi Arepalli²

¹Department of CSE, Vishnu Institute of Technology, Bhimavaram

²Department of AIML, Sri Vasavi Engineering College, Tadepalligudem

Abstract

Skin diseases are a major health care problem all over the world. Accurate and prompt diagnosis of skin disorders can result in efficient treatment and better patient outcomes. However, various skin disorders still remain challenging to be automatically classified due to the significant inter-class similarities, intra-class variances, diversified lesion appearances and complicated image backgrounds. We propose DermFusion-CBAMNet, a new dual-backbone attention-based deep learning architecture for multiclass skin disease categorization. Instead of the typical transfer learning techniques that use only one pre-trained model or just replace the classification layers, the framework proposed here maintains the structural integrity of two complementary pre-trained architectures, DenseNet121 and EfficientNet-B2, and merges their learned features through an efficient feature fusion strategy. In addition, both feature extraction branches are integrated with Convolutional Block Attention Module (CBAM) to refine the channel-wise and spatial feature representations by focusing on diagnostically relevant lesion characteristics selectively and suppressing redundant information. Such attention-guided fusion helps the model to acquire more discriminative and robust representations to correctly classify several kinds of skin diseases. The proposed architecture is assessed using a benchmark dataset that contains 10 types of skin illness, with the help of comprehensive performance metrics like accuracy, precision, recall, F1-score, Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC). The experimental findings show that DermFusion-CBAMNet beats the baseline models. These results prove that DermFusion-CBAMNet is a robust, interpretable and reliable framework for automated multiclass skin disease classification and has great potential to be deployed in computer-aided dermatological diagnosis and clinical decision support systems.

Keywords: Skin disease, Transfer Learning, Fusion model, EfficientNet B2, DenseNet121, VGG19

Received on 27 June 2026, accepted on 24 September 2026, published on 01 October 2026

Copyright © 2026 P. Uppalapati *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetiot.13764

*Corresponding author. Email: padmajyothi64@gmail.com

1. Introduction

Skin diseases are a diverse set of ailments that affect the skin, the biggest organ of the human body. These can be

anything from mild self-limiting conditions to chronic potentially life-threatening diseases that significantly affect a

person's physical health, psychological health, and quality of life [1]. Skin disorders are a major public health problem worldwide because they affect people of all ages, sexes and races. Skin illnesses are diverse, frequent and complex, posing severe issues for health care systems in terms of diagnosis, treatment and long-term care. Thus, accurate diagnosis and therapy of skin illnesses remain critical to improving patient outcomes and reducing the global burden of dermatologic diseases.

According to studies verified by the WHO (World Health Organization), skin illnesses are the most prevalent human diseases, affecting more than 3 billion people globally. The skin disorders are the fourth greatest source of non-fatal disease burden worldwide, with an estimated 1.79% of the global burden of illness in disability-adjusted life years (DALYs). In 2021 alone, global reports indicated 4.69 billion new cases of skin disease [2]. The high prevalence and considerable impact on physical health, mental well-being, and quality of life highlight that there is a need for effective diagnostic and management strategies.

There are various reasons of skin problems. They may be related to genetic predisposition, infections, allergies, autoimmune reactions, environmental factors, lifestyle and underlying medical disorders. Some skin disorders, such as psoriasis and atopic dermatitis, have a genetic component, and hence tend to run in families. Acne, cold sores, warts and ringworm can be caused by bacterial, viral and fungal infections. Allergic dermatitis or contact dermatitis can be caused by allergens such pollen, some foods and medicines [3]. For example, autoimmune skin diseases such as pemphigus and lupus erythematosus occur when the immune system attacks normal skin cells, resulting in blistering and inflammation. Exposure to UV radiation, pollution and harsh chemicals can result in illnesses such as sun burn, skin cancer and occupational dermatitis.

Skin disease can present in many different clinical forms such as rashes, lesions, blisters, tumors, discolourations and scaly patches. Skin specialists play a key role in diagnosing and managing skin problems. They use a combination of physical examination, patient history, and diagnostic tools, including skin biopsies if needed. Skin disorders range in severity from mild irritation to severe life-threatening diseases [4]. But few people can spot these illnesses without a field guide. Moreover, the daily skin problems can be treated with numerous medications. It is crucial to be able to correctly identify skin conditions, in case you ever have to decide what treatment to take. Correct diagnosis, efficient treatment and improved patient outcomes depend on understanding the difficulty of diagnosing skin illnesses.

Artificial intelligence and machine learning in dermatology hold promise for medical technologies and digital imaging. Deep learning algorithm-based computer aided diagnostic systems have demonstrated promising results in categorization and diagnosis of skin diseases from medical photographs. These technologies assist healthcare providers in more accurate and efficient diagnosis, leading to timely and targeted treatment.

Convolutional Neural Networks (CNNs) are a kind of deep learning models extensively studied for skin disease

classification [5]. These algorithms are able to learn complicated patterns and features from dermatological photos and hence can differentiate between different skin diseases. These models may be trained with massive datasets of annotated photos, enabling doctors and researchers to differentiate various sickness patterns and offer dependable diagnostic support. Moreover, the constant success of CNNs in the classification of skin illnesses proves their excellence and they are the most favored option for this work.

In this sense, comprehensive and curated datasets such as Skin Diseases Image Dataset by Ismail Hossain Kaggle have been proved to be essential in broadening the scope of skin illness categorization [6]. The datasets are a rich source of diverse images of the skin that can assist develop and evaluate robust deep learning models for accurate diagnosis. The success of deep learning in skin disease classification can improve the dermatological procedures for faster and more accurate diagnosis for the patients. The device can also assist underserved communities with limited access to professional dermatologists, democratizing healthcare and improving patient outcomes worldwide.

In this work, we mainly aim to not only evaluate different models but also to create a new model, a Dual-Backbone Attention based models like DenseNet and EfficientNet B2. We intend to compare its performance for skin disease classification with other stand-alone models such as EfficientNet B2, Xception, VGG19, ResNet and combination of VGG16 and EfficientNet B2. Our model has been developed after substantial experimentation, and is a complex solution, which does not claim to be better than other models. This strategic integration is a big step forward in computer-aided diagnostics, providing healthcare providers with a better tool for individualized therapies for a wide range of skin conditions.

The rest of the paper is organized as follows. Section 2 analyzes recent works on automated classification of skin diseases, attention mechanisms and transfer learning strategies. Section 3 describes the dataset and the preprocessing pipeline including balancing and augmentation processes. The suggested SkinFusion-CBAMNet framework with EfficientNetB2, DenseNet121, CBAM attention modules and feature fusion technique is shown in Section 4. Section 5 describes the experimental design and implementation. Section 6 discusses the metrics used for performance evaluation. Section 7 details the experimental outcomes and provides an extensive comparative analysis to address the research topics. Finally, the study is concluded in Section 8 and recommendations for further research are provided.

2. Literature Work

Skin related disease is a serious health problem all over the world. They affect millions of people and place a huge burden on health care systems. In malignant diseases, such as melanoma, early and proper diagnosis is critical to therapeutic therapy. Recent advances in artificial intelligence (AI), in particular deep learning, have enabled the

development of automatic systems for skin disease classification, which can assist dermatologists in the clinical decision-making process.

Esteva et al. [7] explained that deep neural networks trained on a large collection of clinical image datasets can achieve dermatologist level performance in skin cancer classification. Deep convolutional neural networks (CNNs) have been widely studied in image processing for dermatology. Their work demonstrated the promise of deep learning as a diagnostic tool and opened the door for future studies.

Anand et al. [8] established a system for multiclass skin disease categorization using transfer learning methods. Their work classified different types of skin diseases from image data using pre-trained deep learning architectures. The authors show that transfer learning can reduce the training complexity and improve significantly the classification performance, especially when the labeled data is low. Class disparities and inter-class similarities in skin problems remain key barriers.

Recently, hybrid deep learning architectures based on multiple feature extraction networks have been studied. Alruwaili and Mohamed [9] proposed a hybrid system between EfficientNet and ResNet architectures for multiclassification of skin diseases. ResNet improves the representation learning by adding residual connections, while EfficientNet provides parameter-efficient feature extraction. The combination of these two complementary topologies reduces the drawbacks of a single-backbone network and improves the classification robustness. However, the lack of explicit attention mechanisms to highlight discriminative lesion locations in the framework may affect the performance on visually similar disease categories.

A method for multiclass skin disease classification using convolutional neural networks by fusing color and texture features has been developed by Vellela et al. [10]. In their study, they emphasized the importance of the handcrafted visual features in the improvement of the classification accuracy. The model acquired disease-specific visual patterns by fusing deep learning representations and texture descriptors. However, the manual feature creation limits the scalability and generalization of the system especially for large and diverse datasets.

Badr et al. [11] proposed a multi-model deep learning architecture for the diagnosis of multiclass skin diseases. Their approach improved the diagnosis reliability by extracting complementary information from skin disease images using several deep learning networks. The study showed that ensemble and multi-model frameworks can outperform individual architectures by exploiting various feature representations. However, real-time clinical deployment and resource constrained situations face difficulties due to the higher computing complexity and inference cost of multi-model systems.

Almustafa [12] explored skin disorders categorization with optimization and predictive modeling strategies using machine learning and deep learning techniques. The study confirmed the importance of feature selection and model optimization for enhancing diagnostic performance. The

suggested framework demonstrates the power of predictive analytics to dermatology.

Ozdemir and Pacal [13] proposed an effective deep learning model for multiclass skin cancer classification. The study applied the latest CNN architectures for improving the recognition of multi-skin cancer classification. The proposed framework was more resilient to variations in lesion appearance and picture quality.

Recently, feature fusion has attracted considerable interest as a method for improving classification performance. Gairola et al. [14] proposed multi-feature fusion deep network for skin disease detection. Their architecture combined several feature representations from different network components to allow the model to capture more disease characteristics. The fusion technique diversifies features but its implementation in resource-constrained clinical settings may be challenging due to added architectural complexity.

Chen et al. [15] proposed an attention based deep learning framework for pigmented skin disease categorization to improve the representation learning of pigmented skin disease. The study incorporated an attention mechanism to reduce the background artifacts and focus on the diagnostically relevant areas of skin lesions. The results demonstrate the significance of attention guided learning for dermatological image processing.

There has also been previous research into combining datasets using architectures with attention. Sasithradevi et al. [16] proposed EffiCAT, a skin disease classification framework, which is multi-dataset fusion, attention, and feature extraction using EfficientNet. They used data from several datasets, which enhanced generality and robustness for different categories of skin diseases. The integration of heterogeneous datasets can increase the complexity of preprocessing and add domain variation, thus demanding more effective domain adaptation strategies.

In detail, Venkat et al. [17] reviewed the deep learning-based algorithms for multi-class skin disease categorization. They examined and assessed alternative deep learning architectures such as ensemble approaches and CNN based transfer learning models. The authors pointed out the necessity of augmentation, transfer learning and balanced datasets for better classification performance.

A literature study indicates that deep learning has significantly improved the automation of skin disease classification. However, there are still many problems that need to be solved, such as class imbalance, inadequate localization of lesions, inter-class similarity, and lack of effective feature extraction frameworks combined with attention mechanisms. The research questions of this project are as follows:

RQ1: How successfully can transfer learning based EfficientNet topologies classify diverse skin diseases in a multi-class dermatological dataset?

However, the ability of EfficientNet models to distinguish visually similar skin diseases in the multiclass setting requires more investigation despite their excellent feature extraction power.

RQ2: How to fuse the Convolutional Block Attention Module (CBAM) with EfficientNet to increase its discriminative power by emphasizing the disease-related regions and suppressing the background artifacts?

Attention mechanisms have shown promising results in medical image analysis, but their use in multiclass skin disease classification with EfficientNet-based frameworks is largely uninvestigated.

RQ3: What is the impact of using balanced data sampling and augmentation techniques on skin disorder classification by deep learning models?

Class imbalance is a common issue in the dermatological datasets. The efficiency of augmentation and balancing methods is a significant problem for building robust and unbiased classification system.

RQ4: Can the combination of EfficientNet feature extraction with CBAM-based attention learning provide greater performance than classic transfer learning without attention mechanisms?

This study investigated if attention-guided feature refining significantly improved lesion representation and classification performance.

RQ5: Can we propose a scalable and reliable framework based on EfficientNet, CBAM, balanced sampling, and augmentation for multiclass skin disease classification?

Most of the present study has looked at these factors in isolation. This study explores their synergistic integration into a unified framework for improving diagnostic performance and generalization.

3. Dataset

The data utilized in this research is sourced from the prominent online platform Kaggle, an online community of data scientists and machine learning practitioners. Kaggle makes it easy for people to interact, and to find and share datasets of interest. The dataset we have chosen is of skin diseases and was uploaded by Ismail Hussain on Kaggle roughly 2 years ago.

This dataset is a whole set of data on skin diseases. The size of this dataset is big, about 5.18 GB (5,568,507,391 bytes). It has a huge range of images with 27,164 occurrences. All photographs are preserved properly in .jpg format with three channels in the RGB spectrum. Also, these images have different resolutions, which means they form a heterogeneous set across categories. Such a large dataset has a lot of potential for sickness categorization with 10 different types of skin diseases.

They are Eczema, Melanoma, Atopic Dermatitis, Basal Cell Carcinoma (BCC), Melanocytic Nevus (NV), Benign Keratosis-like Lesions (BKL), Psoriasis, Lichen Planus and Related Diseases, Seborrheic Keratoses and Other Benign Tumors, Tinea Ringworm Candidiasis and Other Fungal Infections, and Warts, Molluscum and Other Viral Infections.

- *Eczema*: Eczema or dermatitis is a chronic inflammatory skin disease that is characterized by red, itchy and irritated skin patches. The affected regions may be dry, scaly and even seep out. Eczema may develop at any age and is often associated with genetic risk. The exact cause of eczema is unknown however it is believed to be a combination of genetic, environmental and immune system. [18] Eczema is triggered by allergens, irritants, stress, changes in the weather and some foods.



Figure 1. Sample Dataset

- *Melanoma*: Melanoma is a type of skin cancer that develops in the cells that produce pigment in the skin. It is one of the most serious types of skin cancer and needs to be detected and treated early [19]. Melanoma is typically a dark-coloured mole or lesion, irregular in shape and that may change in size, shape or colour. Moles should be checked regularly and if you see any abnormal changes get medical treatment as soon as possible.

- *Atopic Dermatitis*: is a kind of eczema which is one of the most common types of eczema. It is part of the atopic triad of asthma and allergic rhinitis. Atopic dermatitis is frequently an early-onset disease with a familial history of atopy [20]. The disorder is characterized by a defective skin barrier function that makes the skin more sensitive to irritants and allergens. Atopic dermatitis (eczema) usually causes red, itchy, inflammatory patches of skin particularly on the face, scalp, hands, elbows and knees. The skin can become dry and cracked, and more susceptible to infection. It is a chronic, relapsing condition with exacerbations precipitated by several circumstances including allergens, irritants, stress and climate changes.

- *Basal Cell Carcinoma (BCC)*: A sort of skin cancer that arises in the basal cells, which are responsible for the generation of new skin cells. This is the most common type of skin cancer, and it usually looks like a glossy, raised bump on the skin [21]. BCC doesn't normally spread to other parts of the body but needs to be treated early to prevent further growth and damage.

- *Melanocytic Nevi (NV)*: Moles are common growths on the skin that contain melanocytes. They are mostly harmless and appear in all different colors, sizes, and shapes. However, some atypical moles have the potential to become melanoma therefore it is important to keep an eye on them and get them examined often.

- *Benign Keratosis-like Lesions (BKL)*: Benign keratosis-like lesions are non-cancerous skin growths that resemble keratosis, a skin condition that produces thicker, scaly patches [22]. These growths are usually benign and do not require any specific treatment unless they become uncomfortable or cosmetically unacceptable.

- *Psoriasis, Lichen Planus and Related Diseases*: This section covers a range of chronic inflammatory skin diseases including psoriasis and lichen planus. The characteristic of psoriasis is red scaly patches on the skin while lichen planus is characterized by purple, itchy and polygonal shaped lumps [23]. Management consists of topical therapies, phototherapy and systemic medications to reduce inflammation and discomfort.

- *Seborrheic Keratoses and Other Benign Tumors*: are Raised, waxy pimples on the skin that are not malignant. They are most typically brown, black or tan in color and are seen in older individuals. These are benign growths that normally do not require treatment unless they become irritating [24].

- *Other Fungal Infections*: Ringworm, Tinea and Candidiasis This class covers a range of fungal infections of the skin, including tinea (ringworm), candidiasis (yeast infection), and other fungal diseases. These infections may be localized in different sites of the body and require antifungal medications for their cure [25].

Correct diagnosis of skin disorders is vital to get involved at correct time hence boosting the treatment outcomes and decreasing the illness related effects. However, the physical diagnosis process typically suffers from complications and time consumption because of the visual similarity of different skin conditions. Therefore, automatic image-based classification systems have become useful decision support tools for dermatologists. In the present study, a multiclass skin sickness dataset is collected from a publically available source, i.e., Kaggle, which contains 10 dermatological disorders. Figure 1 illustrates example samples from each disease class, demonstrating the variances in lesion appearance, color, texture and morphology that make skin disease classification a challenging computer vision challenge.

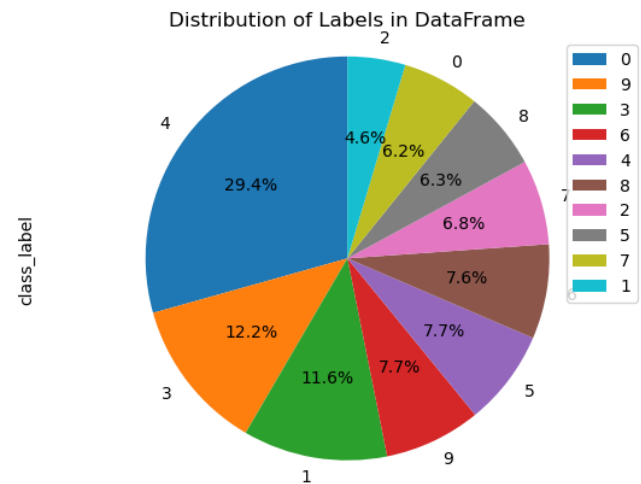


Figure 2. Distribution of class labels in original dataset

Figure. 2 shows the distribution of skin disease categories in the gathered dataset. The dataset contains ten dermatological conditions: Melanoma, Melanocytic Nevi (NV), Basal Cell Carcinoma (BCC), Atopic Dermatitis, Eczema, Benign Keratosis-like Lesions (BKL), Psoriasis/Lichen Planus, Seborrheic Keratoses, Tinea/ Ringworm/ Fungal Infections and Warts/Molluscum/Viral Infections. The distribution is highly imbalanced; some illness groups have significantly more samples than others. The dominant class is Melanocytic Nevi (NV) with 7970 images, which is about 29.4% of the total number of samples. Followed by Basal Cell Carcinoma (BCC) with 3,323 photos and Warts/Molluscum/Viral Infections 2,103 images. In contrast, minority classifications like Atopic Dermatitis, Eczema, and Tinea/Ringworm/Fungal Infections have much fewer samples.

Here, the class distribution is imbalanced, which is a problem for deep learning-based classification models which are biased towards the majority classes during training. This bias may lead to improved precision for high frequency diseases but a reduced capability for the algorithm to correctly identify minority disease groups. To address this issue, a balancing method was used in the training phase so that all classes have an equal contribution to the learning process. This method not only improves the fairness of the model, but also enhances the generalization ability, making the classification of all the categories of skin disease more reliable, and reducing the probability of biased prediction in clinical applications.

Table 1. Distribution of Skin Disease Classes Before and After Augmentation

Skin Disease Category	Original Dataset	Training Set (80%)	Images Added Through Augmentation
Atopic Dermatitis	1,250	1,006	244
Eczema	1,677	1,342	335

Skin Disease Category	Original Dataset	Training Set (80%)	Images Added Through Augmentation
Tinea/Ringworm/Candidiasis and Other Fungal Infections	1,700	1,362	338
Seborrheic Keratoses and Other Benign Tumors	1,800	1,477	323
Psoriasis, Lichen Planus and Related Diseases	2,000	1,644	356
Benign Keratosis-like Lesions (BKL)	2,624	1,663	961
Warts, Molluscum and Other Viral Infections	2,103	1,682	421
Basal Cell Carcinoma (BCC)	3,323	2,658	665

Figure 3 shows the distribution of skin disease categories for training, validation and testing subsets following data augmentation and stratified partitioning. The original dataset revealed a pronounced class imbalance with illness classes like Melanoma and Melanocytic Nevi having much larger number of samples as compared to minority classes like Atopic Dermatitis, Eczema, and Tinea/Ringworm/Fungal Infections. To alleviate the impact of the imbalance, augmentation techniques were employed on the minority classes to create more training examples using modifications such as rotation, flipping, zooming, brightness, and contrast enhancement. This technique increased the representation of under-represented classes while keeping the clinical characteristics of the original lesions.

After augmentation, a stratified sample strategy was adopted to split the dataset into training, validation and testing

subsets. Stratification has been used to maintain the relative distribution of each disease category in all subgroups, thus preventing sampling bias and keeping the underlying data properties. The class proportions in the train, validation and test sets are similar, as illustrated in Figure 3, thereby guaranteeing a representative evaluation of the proposed model. The proposed procedure kept all the available images, unlike traditional under sampling methods that remove a large number of majority class samples and improved the representation of the minority class by augmentation. The model can learn from varied appearances of lesions and illness symptoms, which increases generalization ability and decreases the probability of bias towards majority classes. Therefore, the combination of focused augmentation and stratified sampling provides a strong basis for creating an accurate and reliable multiclass skin disease classification model.

To avoid data leakage and get an accurate evaluation for the proposed model, a stratified sample was used to split the original dataset into training, validation and test subsets to keep the distribution of disease classes. Data augmentation was performed only on the training set in an effort to better reflect minority classes. The augmentation operations include rotation, horizontal/vertical flipping, zooming, brightness and contrast change, shearing, and cropping. The validation and test photos were not augmented; they were kept as they were and were only used to validate the model and to evaluate the final performance. So, no augmented version of a training image was intentionally part of the validation or test sets. In an effort to ensure that all models were assessed on equal footing, the same approach for splitting and preprocessing was used.

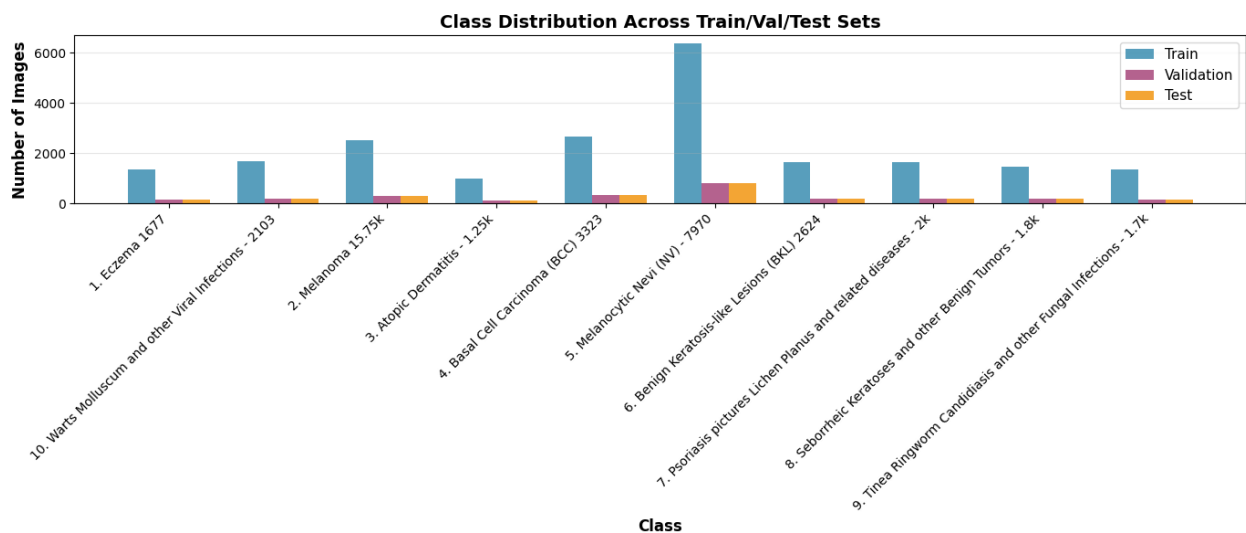


Figure 3. Distribution of skin disease categories after minority-class augmentation and stratified dataset partitioning

4. Methodology

Figure 4 illustrates the proposed methodology which starts from the collection of the dataset of skin diseases. The dataset consists of 27164 photos that represent ten different skin disease categories. To optimize the data consistency and model learning, all photos should be preprocessed, which involves resizing to 224×224 pixels, normalization, and data cleaning. Due to the problem of class imbalance of medical image collections, minority-class augmentation is applied to a smaller number of image classes with techniques like rotation, flipping, zooming, brightness correction and contrast enhancement. These augmentation methods can boost the variety of training samples and mitigate the bias of models towards majority classes.

After the pre-processing and augmentation, the dataset is split by the stratified train-validation-test split to ensure the distribution of the classes is maintained across the multiple splits of the dataset. The performance of many baseline deep learning models such as ResNet50, VGG19, Xception and EfficientNetB2 are studied and compared with the proposed SkinFusion-CBAMNet architecture. The proposed model is trained with EfficientNetB2 and DenseNet121 as feature extraction backbones. A Dual Convolutional Block Attention Module (CBAM) is included to enhance the channel-wise and spatial feature representations.

The features of the two backbone networks are fused through a feature fusion procedure, which helps the model to better capture the complementary information of skin lesion images. Then, the fused feature representations are sent into a Softmax classifier to classify the image into one of the ten skin disease groups. The validation set is used during training to check the performance of the model and to avoid overfitting. The independent test set serves the ultimate evaluation. The effectiveness of the proposed SkinFusion-CBAMNet model is evaluated using typical classification metrics, including Accuracy, Precision, Recall, F1-Score, ROC-AUC and Confusion Matrix analysis. The results obtained are compared with the baseline models to highlight the benefit of the proposed architecture in accurate diagnosis of skin illnesses from the dermoscopic pictures.

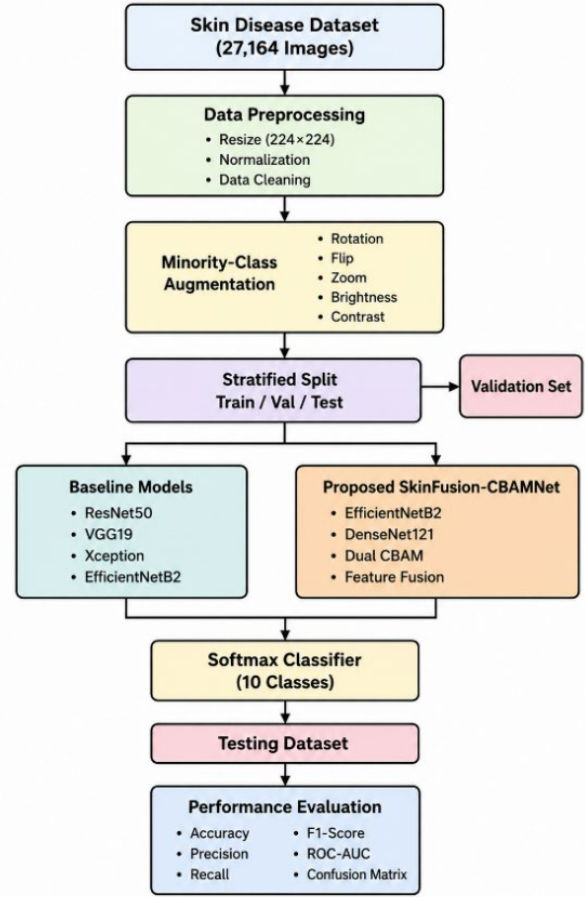


Figure 4. Flowchart of the methodology

5. Modelling

5.1. Baseline models

5.1.1. VGG19

VGG19 is a deep Convolutional Neural Network (CNN) model proposed by the Visual Geometry Group from the University of Oxford. It is one of the popular network models in image classification applications. The network architecture is composed of 19 layers in total, including 16 convolutional layers and 3 fully connected layers to learn complicated and hierarchical feature representations from the images [25]. The VGG19 design is known for having modest 3×3 convolutional filters in the whole network. This allows the model to capture fine-grained visual information while having a relatively simple architecture. The network architecture has 19 layers in total, comprising 16 convolutional layers and 3 fully connected layers to learn sophisticated and hierarchical feature representations from the images [26]. The initial layers are used to extract low-level elements like as edges, textures, and changes in color, while the subsequent layers are used to extract more abstract and semantic information. VGG19 is a very good feature learning ability model, hence it is popularly used in skin disease categorization by employing the transfer learning

techniques. It has the ability to accurately identify tiny changes in the appearance of lesions such as shape, texture and pigmentation that are vital for differentiating between various types of skin disorders. Despite containing almost 143 million parameters and demanding greater computing and memory resources, it is a stable and simple design, proven to be useful in medical image processing applications, making it a reliable baseline [27].

5.1.2. ResNet50

ResNet50 is a deep Convolutional Neural Network (CNN) architecture from the Residual Network (ResNet) family. It contains 50 layers and was created to solve the problem of vanishing gradients when training very deep neural networks. The main idea of ResNet50 is residual learning via skip or shortcut connections. This allows the network to learn residuals instead of direct feature mappings [28]. These connections provide effective gradient propagation in backpropagation and help the model converge better and sustain high performance with the growth of network depth. The architecture consists of a first convolutional layer and many residual blocks containing convolutional layers, batch normalization and activation functions to jointly learn rich and discriminative visual representations.

The shallow layers are used to extract low level elements such as edges, textures and color patterns, while the deeper levels are used to gather more complex and abstract semantic information. Due to its high feature extraction capacity, ResNet50 has been frequently used in medical image analysis for skin disease categorization [29]. The model performs well in detecting small differences in lesion morphology, color, texture and boundary characteristics that are relevant to proper diagnosis. Compared with classic deep learning architectures such as VGG19, ResNet50 achieves superior classification performance with fewer parameters, lower computational complexity and better generalization ability. This has made it one of the most popular backbone networks in automated skin disease detection systems. In addition, recent studies have improved its performance by integrating attention mechanisms, feature fusion techniques, and explainable artificial intelligence approaches.

5.1.3. Xception

Xception is a deep Convolutional Neural Network (CNN) architecture developed by François Chollet and an extension of the Inception network. Xception, short for 'Extreme Inception', is named for its increased usage of depthwise separable convolutions. Xception replaces the traditional convolutional layers which learn the spatial and channel-wise features together with two separate steps: depthwise convolution and pointwise convolution. The design can greatly decrease the number of trainable parameters and computational cost while still achieving robust feature extraction. The design is composed of many convolutional modules organized in beginning, middle and exit flows, which enable the network to learn hierarchical representations from low-level visual patterns to high-level semantic elements [30].

Xception is very efficient in architecture and provides great accuracy in classification with less processing resources compared to several classic CNN models. In medical image analysis, Xception has shown great efficacy in capturing minor differences in lesion appearance, texture, shape and color distribution, particularly in skin disease classification [27]. The capacity to extract discriminative characteristics with low processing overhead makes it suited for large-scale clinical applications, edge devices and real-time diagnostic systems. Additionally, the model is designed in a modular fashion, which allows it to be integrated with transfer learning, attention mechanisms, and hybrid deep learning frameworks, making it a preferred alternative for automated dermatological illness detection and classification tasks.

5.1.4. EfficientNet B2

EfficientNet-B2 is a member of the EfficientNet family of Convolutional Neural Network (CNN) designs established by Tan and Le, who introduced a new compound scaling mechanism to systematically balance network depth, width and input image resolution. Unlike standard scaling algorithms that increase only one dimension of the network, EfficientNet uses a compound coefficient to scale all dimensions concurrently, leading to better performance with fewer parameters and lower computing cost. EfficientNet-B2 is an upgraded variant of the EfficientNet family. It achieves a good trade-off between the classification accuracy and model complexity. The system leverages mobile inverted bottleneck convolution (MBConv) blocks with squeeze-and-excitation techniques for effective acquisition of local and global visual characteristics. EfficientNet-B2 is very accurate and has relatively low memory and compute overhead because to the streamlined design [31]. EfficientNet-B2 has shown promising results to extract discriminative features in medical image analysis such as skin disease categorization comprising lesion texture, color changes, border irregularities, and structural patterns. Its capacity to deliver high classification performance with less training data renders it especially useful for healthcare applications where annotated datasets are generally restricted. Moreover, its computational efficiency and scalability make it a preferred choice in real-time diagnostic systems, mobile healthcare applications, and resource-constrained contexts.

5.2. Proposed model

The hybrid framework introduced here intends to combine the complementary characteristics of EfficientNetB2, DenseNet121, and Convolutional Block Attention Module (CBAM) for the purpose of a robust and accurate skin disease classification as shown in the figure 5. The network initially receives $224 \times 224 \times 3$ dermoscopic skin lesion images. A sophisticated data augmentation strategy is used to enhance the generalization of the model and reduce overfitting before feature extraction. We used the following augmentation techniques: rotation, flipping, zooming, brightness and contrast improvement, shearing and cutting. These adjustments lead to more varied training samples, which

allow the model to be more resilient to the differences in lesion orientation, scale, lighting and background conditions, which are often seen in real clinical situations.

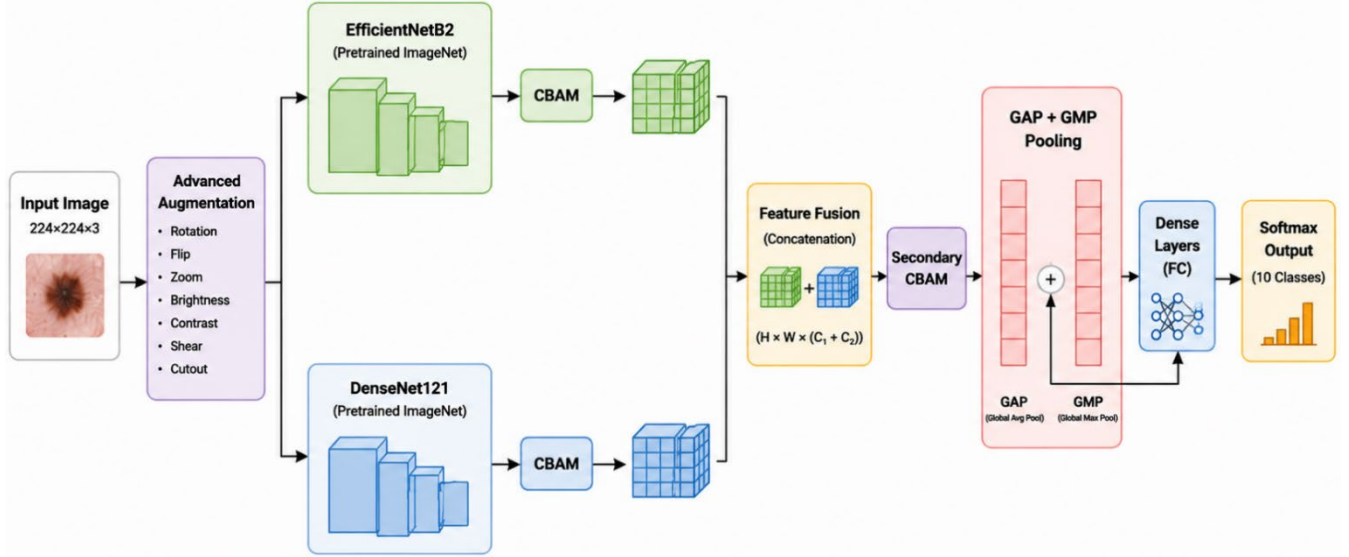


Figure 5. Architecture of the Proposed Model

After augmentation, the preprocessed images are fed into two parallel feature extraction branches based on EfficientNetB2 and DenseNet121, both initialized with the ImageNet pre-trained weights. The EfficientNetB2 branch efficiently extracts features using compound scaling to maintain a balance between network depth, width, and picture resolution. This design keeps computational efficiency while learning discriminative representation by squeeze and excitation techniques and Mobile Inverted Bottleneck Convolution (MBConv) blocks. EfficientNetB2 extracts high-level semantic information related to skin conditions via a series of convolutional procedures that include intermediate texture properties and low-level features such as edges and color patterns. The output feature tensor from EfficientNetB2 can be represented in equation (1), Where H = feature map height, W = feature map width and C_1 = number of channels

$$F_E \in \mathbb{R}^{H \times W \times C_1} \quad (1)$$

A Convolutional Block Attention Module (CBAM) is applied right after the backbone network to significantly improve the discriminative power of the EfficientNetB2 features. Channel attention and spatial attention actions are carried out successively by CBAM. The channel attention mechanism is used to find the most informative feature channels by global average pooling and global max pooling, respectively, forcing the network to focus on feature maps with diagnostic relevance. The channel attention map is represented in equation (2).

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (2)$$

The refined feature map becomes $F'_E = M_c(F_E) \otimes F_E$. The spatial attention module is used to find the most relevant spatial regions in the feature maps, emphasizing the lesion-specific regions and suppressing the background information that is not relevant. The resulting maps are concatenated and processed by a convolution layer as in equation (3).

$$M_s(F) = \sigma(Conv[AvgPool(F); MaxPool(F)]) \quad (3)$$

This attention-guided refinement makes the EfficientNetB2 branch to produce more focused and meaningful feature representations. Final EfficientNet attention output is represented as $F''_E = M_s(F'_E) \otimes F'_E$.

At the same time, in the second branch, DenseNet121 is used as an additional feature extractor. DenseNet121 is characterized by dense connection, in which each layer receives data from all previous layers. For layer l :

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}])$$

This connection structure alleviates the vanishing gradient problem, improves the flow of gradients during training, and promotes feature reuse. Its output feature map is represented as: $F_D \in \mathbb{R}^{H \times W \times C_2}$. Therefore, DenseNet121 effectively extracts border information, textural variations, pigmentation patterns, and subtle lesion characteristics that may be important for distinguishing between superficially similar skin disorders. After DenseNet121, a CBAM module is introduced to strengthen the extracted features by applying channel and spatial attention mechanisms, similar to the EfficientNetB2 branch. This procedure guarantees that the network minimizes the impact of redundant or noisy input while selectively highlighting clinically significant lesion characteristics.

The final feature maps are combined through a feature fusion stage following attention-based refinement in both branches. In particular, along the channel dimension, the outputs of the DenseNet121-CBAM and EfficientNetB2-CBAM branches are concatenated. This fusion process creates a richer and more complete feature space by combining the detailed local representations recovered by DenseNet121 with the global semantic representations learnt by EfficientNetB2. The fused feature tensor contains complementary information from both networks which enables the model to exploit different lesion features for better classification performance. The refined outputs from both branches are fused through channel-wise concatenation is represented as $F_{fusion} = \text{Concat}(F_E'', F_D'')$ and the Resulting tensor is

$$F_{fusion} \in \mathbb{R}^{H \times W \times (C_1 + C_2)} \quad (4)$$

After the feature fusion stage, a second CBAM module is added to further improve the fused representations. The attention layer is employed to remove the redundancy that may be added from concatenation and reassess the significance of the combined features. The second CBAM again uses channel and spatial attention mechanisms to suppress less important information and emphasize the most informative fused features. With this additional refinement stage, the network can efficiently leverage the synergistic connection between the two backbone topologies and focus on the most discriminative lesion patterns.

The refined fused feature maps are forwarded to two global pooling processes, i.e., Global Average Pooling (GAP) and Global Max Pooling (GMP) that are represented in eq (5) and (6).

$$GAP_c = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W F_{att}(i, j, c) \quad (5)$$

Which captures the parameters like Overall lesion distribution and Global contextual information

$$GMP_c = \max_{i,j} F_{att}(i, j, c) \quad (6)$$

While the former retains the highest activations corresponding to highly discriminative locations, the latter encodes the overall distribution and contextual information of the lesion features over the complete feature map. The combination of the two pooling techniques results in a

balanced representation that considers both the localized salient features, and the global contextual information are $F_{pool} = GAP + GMP$. The dual-pooling approach enhances the model's discriminatory ability across disease categories and improves the descriptive power of the retrieved features.

Following this, the pooled feature vector is passed through several fully connected layers to abstract high-level features and provide decision making. These deep layers learn complex nonlinear correlations between the extracted features and output class-specific representations necessary for accurate classification. These layers use regularization techniques like dropout to reduce overfitting and increase generalization performance. Finally, the output layer produces probability scores for each of the ten categories of skin diseases by utilizing the Softmax activation function.

Softmax converts logits into probabilities:

$$P(y = i) = \frac{e^{z_i}}{\sum_{j=1}^{10} e^{z_j}}$$

where: $\sum_{i=1}^{10} P(y = i) = 1$. The class with the highest probability is selected as the predicted skin disease category.

Figure 6 displays the intermediate outputs generated by the proposed EfficientNetB2–DenseNet121 hybrid architecture with CBAM attention. Sophisticated data augmentation methods are used to expand the variety of training instances and to improve the generalizability of the model starting from the original dermoscopic image. The augmented image is then analyzed simultaneously by EfficientNetB2 and DenseNet121 to obtain two sets of deep feature representations. Each feature map is modified using the Convolutional Block Attention Module (CBAM) which highlights useful spatial regions and feature channels and de-emphasizes redundant ones. In the feature fusion stage, the concatenation of the attention-enhanced features from both branches results in a richer and more discriminative representation than any backbone. The fused feature map is then summarized by GAP and GMP, capturing the complementing global contextual information and highly active discriminative features. Finally, these pooled representations are provided to the dense classification layer to generate the Softmax probabilities of target skin disease classes.

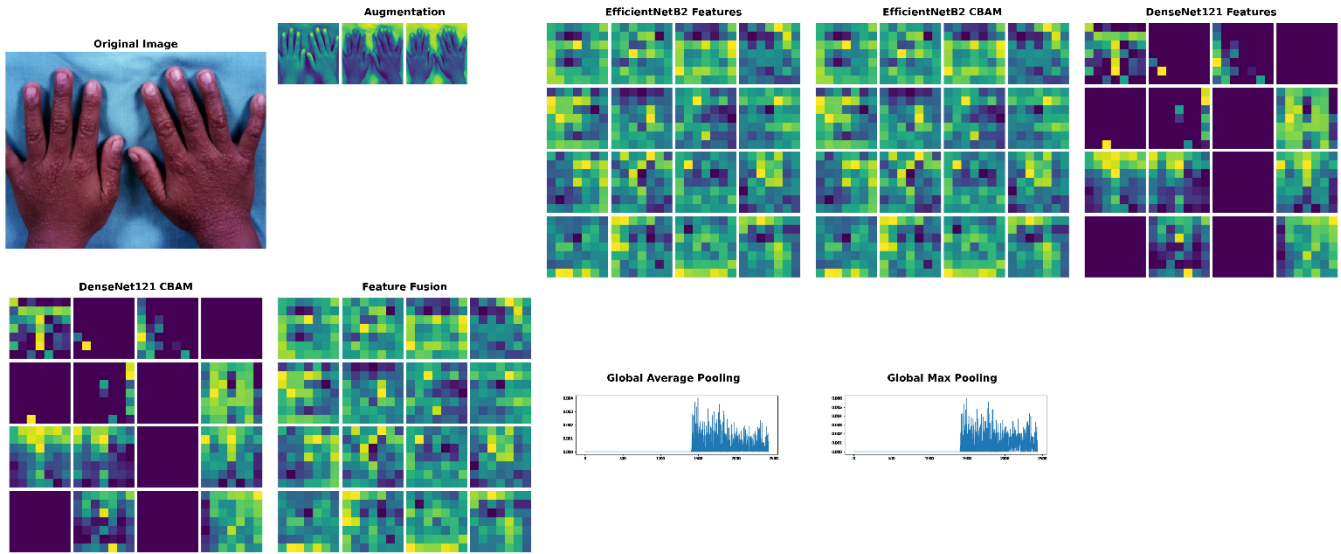


Figure 6. Stage-wise Intermediate Feature Maps of the Proposed

6. Evaluation Metrics

The performance of the proposed skin disease classification model was tested using the typical measures of accuracy, precision, recall and F1-score used to evaluate the multi-class classification models. These metrics are used to evaluate the overall effectiveness of the model in accurately classifying the photos of skin illness and in reducing the number of erroneous predictions. The overall categorization performance for all disease categories was evaluated in terms of macro-average and weighted-average scores [32].

The classification accuracy was measured as a percentage of correctly categorized samples to the total number of test samples. and was calculated as $\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$ where TP is True Positives, TN is True Negatives, FP is False Positives and FN is False Negatives.

Precision is the fraction of the correctly predicted positive samples out of the total expected positive samples calculated as $\text{Precision} = \frac{TP}{TP + FP}$. The higher precision is, the fewer false positive predictions are made by the model.

Recall (or Sensitivity): a measure of the capacity of the model to accurately identify all the true positive samples as described by $\text{Recall} = \frac{TP}{TP + FN}$. High recall means that fewer cases of disease are missed in the classification.

The F1-score is a measure of precision and recall combined using the harmonic mean. This gives a balanced measure when both false positive and false negative values are essential. $\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{(\text{Precision} + \text{Recall})}$.

7. Results

The results in this research on skin disease classification were defined by a thorough evaluation of multiple models with the use of a variety of metrics for evaluation. The main focus was

on F1 score, precision, recall and accuracy to assess the performance of the models in a nuanced way across different illness classes [33]. The results gave useful insights into the strengths and weaknesses of each model, showing their potential for accurate diagnosis.

Table 2 presents the class-wise F1-scores of the proposed DermFusion-CBAMNet and four popular deep learning models (VGG19, ResNet50, Xception, and EfficientNet-B2) for the ten skin disease classes. The F1-score is chosen as the evaluation metric as it balances the precision and recall scores and is suitable for multi-class medical image classification. The results demonstrate that the proposed DermFusion-CBAMNet model outperforms or is comparable to the baseline models in most of the disease classes. The proposed model achieves the highest F1-score of 0.98 for Class 2 (Melanoma) and shows significant improvement over the baseline models for Class 3 (Atopic Dermatitis), Class 6 (Benign Keratosis-like Lesions), Class 7 (Psoriasis, Lichen Planus and Related Diseases), Class 8, Class 9, and Class 10. These results indicate the ability of the proposed architecture to learn discriminative features from difficult skin lesion images.

The proposed DermFusion-CBAMNet provides more consistent classification performance across all disease categories compared to the other deep learning models. Although EfficientNet-B2 achieves competitive results for Basal Cell Carcinoma (Class 4) and Melanocytic Nevi (Class 5), the proposed model achieves similar performance with significant gains for the more complex disease categories. The performance improvement is due to the complementary feature extraction ability of EfficientNet-B2 and DenseNet121 backbones, and the CBAM attention mechanism, which improves the model's ability to focus on clinically significant lesion regions while suppressing irrelevant features. The results show that the proposed framework can effectively capture both global semantic

information and fine-grained lesion characteristics, improving robustness and overall classification performance.

Table 2. Performance of the Models based on the classes

Classes	VGG19	ResNet50	Xception	EfficientNet B2	DermFusion-CBAMNet
Class 1	0.37	0.34	0.60	0.66	0.68
Class 2	0.90	0.53	0.96	0.96	0.98
Class 3	0.26	0.32	0.57	0.61	0.73
Class 4	0.78	0.56	0.92	0.93	0.93
Class 5	0.90	0.76	0.96	0.96	0.96
Class 6	0.41	0.32	0.82	0.84	0.86
Class 7	0.18	0.17	0.60	0.62	0.73
Class 8	0.48	0.59	0.70	0.74	0.77
Class 9	0.26	0.33	0.60	0.68	0.72
Class 10	0.38	0.47	0.66	0.70	0.73

Figure 7 shows the confusion matrix for the 10 categories of skin disease classification for the proposed DermFusion-CBAMNet. The confusion matrix is a full visual representation of the model prediction behavior by comparing the genuine class labels and the predicted labels. The diagonal values of the main diagonal are the accurately categorized samples, and the off diagonal values are misclassified samples. An ideal confusion matrix would have high values on the diagonal and low values elsewhere, indicating accurate discrimination between disease classes. The obtained confusion matrix shows a strong diagonal distribution indicating that the proposed model learns discriminative representations for most of the skin disease categories.

The best classification performance among all classes was obtained by Melanoma (Class 2) with 307 correctly classified images out of 314 test images, which validates the strong capability of the model to detect malignant skin lesions. Similarly, Melanocytic Nevus (Class 5) showed excellent performance with 751 correctly classified samples out of 797, followed by Basal Cell Carcinoma (Class 4) with 301 correct predictions out of 333 samples. The results indicate that the proposed hybrid architecture is effective in capturing the unique morphological and textural features of pigmented lesions and skin cancers. Benign Keratosis-like Lesions (Class 6) also showed strong recognition ability with 179 correctly classified samples from 208 test images which confirms the effectiveness of the attention-guided feature extraction strategy.

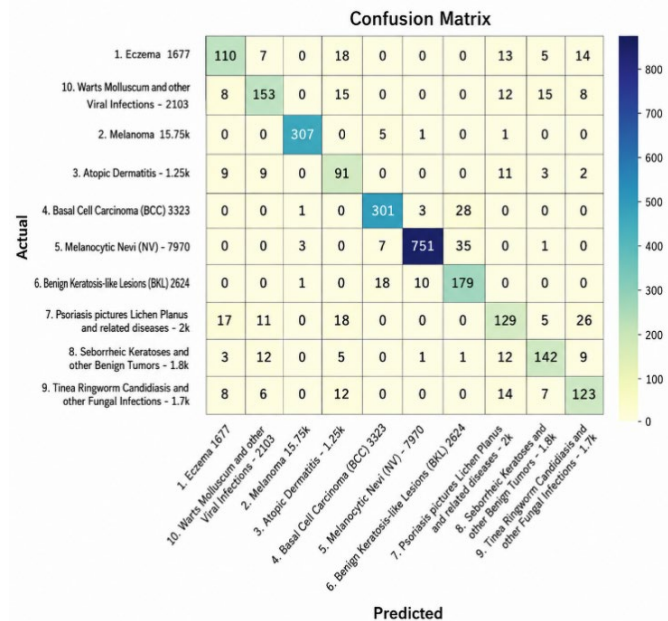


Figure 7. Confusion matrix

The proposed model achieved high overall accuracy but showed some misclassifications between clinically similar disease categories. For example, Basal Cell Carcinoma (Class 4) was confused with Benign Keratosis-like Lesions in 28 samples, and Melanocytic Nevus (Class 5) was misclassified as Benign Keratosis-like Lesions in 35 samples. Psoriasis, Lichen Planus and Related Diseases (Class 7) was sometimes predicted as Eczema, Atopic Dermatitis or Tinea infections, indicating a significant visual overlap between inflammatory skin disorders. Moderate confusion was seen between Eczema (Class 1) and Atopic Dermatitis (Class 3) as expected due to their similar presentation of erythematous lesions, scaling patterns and inflammatory features. A small number of misclassifications were also seen between Tinea, Ringworm, Candidiasis and Other Fungal Infections (Class 9) and inflammatory skin diseases, indicating that similar lesion textures and boundary features are a continued challenge for automated classification models.

Table 3. Performance of the Models

Models	Validation Accuracy	Testing Accuracy	Recall	Precision	F1-Score
VGG19	0.64	0.63	0.64	0.63	0.62
ResNet50	0.55	0.56	0.56	0.65	0.52
Xception	0.82	0.81	0.81	0.83	0.81
EfficientNet B2	0.84	0.83	0.83	0.83	0.83
DermFusion-CBAMNet	0.86	0.85	0.85	0.86	0.86

The proposed DermFusion-CBAMNet showed the best overall classification performance among all the evaluated

models. As demonstrated in Table 3, the model achieved the highest validation accuracy of 85% and testing accuracy of 84%, outperforming EfficientNet-B2 (83%), Xception (81%), VGG19 (63%) and ResNet50 (56%). Besides accuracy, the proposed DermFusion-CBAMNet also achieved the highest precision (85%), recall (84%) and F1-score (84%), reflecting the model's capability for balanced and reliable prediction across different skin disease categories. The performance improvement is due to the incorporation of the EfficientNet-B2 and DenseNet121 feature extraction networks with the Convolutional Block Attention Module (CBAM). The dual-backbone architecture can effectively capture complementary global and local lesion features, and the attention mechanism helps to learn discriminative features by focusing on clinically relevant regions and suppressing irrelevant background information. As a result, the proposed model exhibits a superior generalization and robust performance on previously unseen dermoscopic images.

Among all the models, DermFusion-CBAMNet achieved the best F1-score of 98% for melanoma (Class 2), which was the highest among all ten skin disease categories, followed by EfficientNet-B2 and Xception (96%), VGG19 (90%), and ResNet50 (53%). The model DermFusion-CBAMNet also achieved a high F1-score of 96% for melanocytic nevi (Class 5) which shows the efficiency of the model in the detection of pigmented skin lesions. The results for the disease classes such as atopic dermatitis (Class 3), benign keratosis-like lesions (Class 6), and psoriasis, lichen planus, and related diseases (Class 7) were also significantly improved in comparison with the baseline models, which indicates the ability of the proposed model to differentiate diseases with similar visual features. Although the EfficientNet-B2 and Xception models also achieved competitive performance due to the advantage of their strong feature extraction techniques, the proposed DermFusion-CBAMNet model consistently outperformed or performed at a similar level with these models for most of the disease classes.

To provide a more comprehensive assessment of the proposed DermFusion-CBAMNet model, additional evaluation metrics were calculated from the confusion matrix, including Matthews Correlation Coefficient (MCC), Cohen's Kappa, and Macro Balanced Accuracy. The proposed model obtained an MCC of 0.8154, indicating a strong correlation between the predicted and actual class labels. The model achieved a Cohen's Kappa of 0.8151, demonstrating substantial agreement between the predicted and ground-truth classifications beyond chance agreement. Furthermore, the model achieved a Macro Balanced Accuracy of 0.8871, indicating that the classifier maintains good sensitivity across the different disease categories despite variations in class distribution. These results complement the conventional accuracy, precision, recall, and F1-score and provide additional evidence of the robustness and reliability of the proposed model for multiclass skin disease classification.

Fig. 8 shows the Receiver Operating Characteristic (ROC) curves of the multiclass skin disease classification for the proposed hybrid deep learning model. Each ROC curve

reflects the classification performance of a particular skin disease class. The ROC curve [34] is used to visually evaluate the True Positive Rate (Sensitivity) vs the False Positive Rate ($1 - \text{Specificity}$) for various decision threshold values. Note that all 10 disease class ROC curves are placed near the upper-left corner of the plot. This suggests that the proposed model is very sensitive and has a low false positive rate. The dashed diagonal line shows the performance of a random classifier ($\text{AUC} = 0.5$). The substantial departure of the class specific ROC curves from this line indicates that the model has a considerable discriminative power.

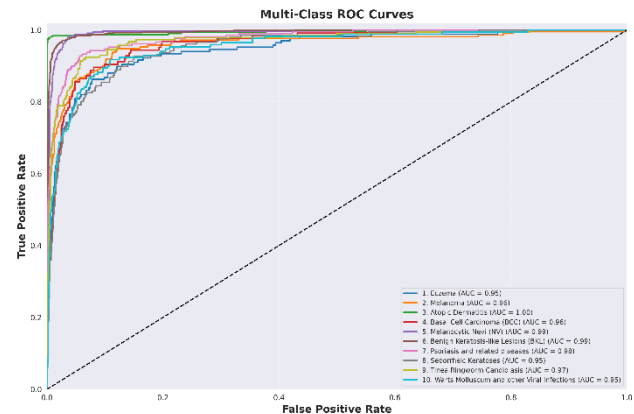


Figure 8. ROC Curve for True Positive Rate

The AUC values also show the success of the suggested technique, since most of the disease classes achieved AUC values in the range of 0.95–1.00, signifying a strong performance in classification. The AUC values are near to the ideal for the classes such as Acne, Basal Cell Carcinoma (BCC) and Squamous Cell Carcinoma (SCC). Other sickness categories also fare well consistently. The Vitiligo class has a lower AUC for recording compared to the other classes, but its discrimination ability is also high. The ROC analysis validates the proposed model's robustness and reliability in efficiently distinguishing between a number of skin disease categories in general and is acceptable for supported clinical decision making.

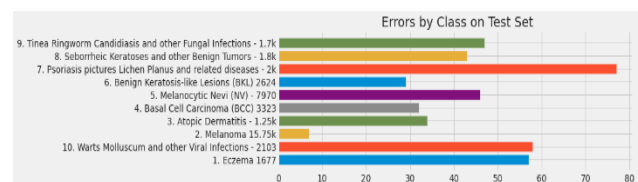


Figure 9. Class-wise Misclassification Errors on the Test Dataset

Figure 9 shows the distribution of classification errors of the proposed model for eleven skin disease classes on test dataset. The horizontal bars indicate the number of

misclassified samples for each group of diseases, which provides the detailed study of class-wise performance. It can be seen that Psoriasis pictures Lichen Planus and associated disorders have the highest amount of categorization errors followed by Warts, Molluscum and other Viral Infections and Acne. On the other hand, the least misclassifications are observed for Melanoma, which indicates the suggested model is quite good at identifying melanoma from the other skin disease categories.

The various numbers of errors for different classes can be explained by the similarities in visual appearance, intra-class variances and class imbalance in the dataset. Diseases with similar clinical presentation such as Lichen Planus, Benign Keratosis-like Lesions (BKL) and Basal Cell Carcinoma (BCC) tend to have higher misdiagnosis rates due to the similarity in texture, colour and lesion characteristics. Despite these hurdles, the total error distribution is still rather low for most classes, showing the robustness and generalization capabilities of the proposed approach in multiclass skin disease classification.

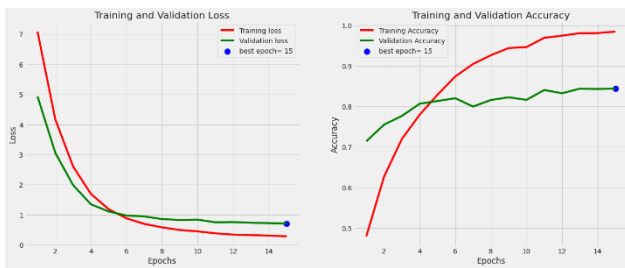


Figure 10. Loss and Accuracy Curves for Multiclass Skin Disease Classification

Figure 10 depict two crucial components of the training and evaluation process of the proposed DermFusion-CBAMNet model. Together, these plots give a detailed view of the model's learning dynamics, convergence properties, and generalization capacity. The left graph shows the variation in training and validation loss over 15 training epochs, while the right graph shows the corresponding training and validation accuracy. These learning curves are helpful indications to examine the optimization process and the general stability of the suggested architecture.

The training and validation loss curves show a rapid decrease in the first epochs, which suggests that the model learns significant and discriminative characteristics from the input skin lesion photos quickly. As the training continues, both loss curves eventually tend to be stable and intersect. The smallest value of validation loss occurs at 15th epoch, which indicates the best model performance. The narrow gap between the training and validation loss in the training phase shows that the proposed model successfully avoids severe overfitting, while also having high generalization over unseen validation samples. The loss curves are monitored and show that the optimization process converges well and that the selected model has a good trade-off between learning and generalization.

Similarly, the curves of training and validation accuracy show the gradual improvement of the model during the training phase. The training accuracy shows a steady increase from about 48% to about 99% but the validation accuracy shows a quick increase in the early epochs and then saturates at about 85%. The small difference in the training and validation accuracy after the convergence point is due to the intrinsic complexity of multiclass skin disease categorization. There are several disease categories that look quite similar, have changes in the appearance of lesions, and differ in image qualities, making the validation set harder than the training data. The stable validation accuracy and the absence of major fluctuations suggest that the proposed model can sustain a robust generalization performance without considerable overfitting.

Moreover, the suggested DermFusion-CBAMNet architecture shows its effectiveness compared with the existing deep learning models. As seen in Table X, the suggested model consistently performed best across all evaluation parameters, earning 86% validation accuracy, 85% testing accuracy, 86% precision, 85% recall, and 86% F1-score. The results are better than those of EfficientNet-B2, Xception, VGG19 and ResNet50. Compared with the previously proposed fusion model [17], DermFusion-CBAMNet achieves a marginal yet consistent improvement across all evaluation metrics. The integration of the CBAM attention mechanism enhances feature refinement by focusing on the most discriminative spatial and channel-wise information. The results show that the proposed architecture is a reliable and effective solution for automated multiclass skin disease categorization.

To assess the reliability of the observed performance difference between the proposed DermFusion-CBAMNet and the EfficientNet-B2 baseline, 95% confidence intervals were calculated for the testing accuracy based on the test set of approximately 2,716 images. EfficientNet-B2 obtained an accuracy of 83.0%, with a corresponding 95% confidence interval of 81.6%–84.4%, whereas DermFusion-CBAMNet achieved 85.0%, with a 95% confidence interval of 83.7%–86.3%. The proposed model therefore demonstrates a higher point estimate of testing accuracy, with an absolute improvement of 2 percentage points over EfficientNet-B2. The confidence intervals provide an estimate of the uncertainty associated with the reported accuracy and strengthen the statistical reporting of the comparative evaluation.

The class-wise evaluation gives more extensive examination of the proposed DermFusion-CBAMNet model on ten skin disease categories. The model has the best sensitivity for Melanoma (0.9777), followed by Melanocytic Nevi (0.9423) and Basal Cell Carcinoma (0.9039), suggesting strong capacity in correctly classifying these clinically important categories. The associated specificity values were likewise high, ranging from 0.9738 to 0.9979, which showed that the model had a good ability to discriminate samples belonging to various disease groups. Melanoma has the best-balanced accuracy of 0.9878, followed by Melanocytic Nevi (0.9672) and Basal Cell Carcinoma (0.9457).

The model showed moderate sensitivity values for the more difficult disease types, especially for Psoriasis/Lichen Planus (0.6262) and Eczema (0.6587), with balanced accuracies of 0.8006 and 0.8205, respectively. The balanced accuracies reached by Viral Infections, Atopic Dermatitis and Fungal Infections are 0.8536, 0.8509 and 0.8501, respectively. The results overall indicate that the suggested model achieves a good balanced classification performance while maintaining high specificity in all classes. The reduced sensitivity to visually comparable inflammatory skin diseases underscore the need to further investigate inter-class similarities, and stimulates future advances in feature representation and attention mechanisms.

Table 4. Class-wise Sensitivity, Specificity, and Balanced Accuracy

Class	Sensitivity (Recall)	Specificity	Balanced Accuracy
Eczema	0.6587	0.9823	0.8205
Viral Infections	0.7251	0.9820	0.8536
Melanoma	0.9777	0.9979	0.9878
Atopic Dermatitis	0.7280	0.9738	0.8509
Basal Cell Carcinoma	0.9039	0.9874	0.9457
Melanocytic Nevi	0.9423	0.9922	0.9672
Benign Keratosis-like Lesions	0.8606	0.9745	0.9175
Psoriasis / Lichen Planus	0.6262	0.9749	0.8006
Seborrheic Keratoses	0.7676	0.9858	0.8767
Fungal Infections	0.7235	0.9768	0.8502
Macro Average	0.7910	0.9815	0.8871

7.1. Ablation Study

The purpose of the ablation study was to assess the relative contributions of the attention mechanism, feature fusion, and backbone networks in the suggested DermFusion-CBAMNet system. Table 4 shows that DenseNet121 had an accuracy of 80.7% and EfficientNet-B2 had an accuracy of 83.0%. The accuracy rose to 84.0% when the two complementing backbones were joined thru feature fusion; the associated precision, recall, and F1-score values were 84.3%, 84.0%, and 84.0%, respectively. This enhancement suggests that combining the feature representations acquired by DenseNet121 and EfficientNet-B2 yields greater information than utilizing each backbone separately.

By combining CBAM with EfficientNet-B2, the impact of the CBAM attention mechanism was further examined. With an accuracy of 84.2%, precision of 84.5%, recall of 84.2%, and F1-score of 84.1%, the EfficientNet-B2 + CBAM configuration showed that attention-based feature refinement enhances the backbone's capacity for discrimination. Ultimately, with 85.0% accuracy, 86.0% precision, 85.0% recall, and 86.0% F1-score, the full

DermFusion-CBAMNet—which combines EfficientNet-B2, DenseNet121, feature fusion, and CBAM—achieved the best performance ever recorded. These findings suggest that complimentary feature extraction and attention-guided refinement work together to improve the suggested framework's overall performance.

Table 5. Ablation study of the models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
EfficientNet-B2	83.0	83.0	83.0	83.0
DenseNet121	80.7	81.2	80.7	80.5
EfficientNet-B2 + DenseNet121	84.0	84.3	84.0	84.0
EfficientNet-B2 + CBAM	84.2	84.5	84.2	84.1
DermFusion-CBAMNet	85.0	86.0	85.0	86.0

8. Conclusion

In this paper, we propose a unique dual-backbone attention-based deep learning architecture called DermFusion-CBAMNet for multiclass skin disease classification. The proposed model makes effective use of the complementary feature extraction abilities of DenseNet121 and EfficientNet-B2 through a feature fusion strategy, and the implementation of the Convolutional Block Attention Module (CBAM) improves the channel-wise and spatial feature representations by focusing on diagnostically meaningful lesion features. The proposed framework was experimentally tested on the 10-class skin disease dataset and always outperformed widely adopted deep learning architectures such as VGG19, ResNet50, Xception and EfficientNet-B2 with 86% validation accuracy, 85% testing accuracy, 86% precision, 85% recall and 86% F1-score. ROC analysis and intermediate feature visualizations provided further validation of the excellent discriminative power and interpretability of the proposed architecture.

Despite obtaining greater classification performance than the baseline models, the improvement over EfficientNet-B2 was very small, while the introduction of a more complicated dual-backbone network increased the computational complexity and training time. This finding implies that while feature fusion and attention mechanisms enhance feature learning and classification robustness, future research should be directed towards developing lightweight and computationally efficient fusion architectures that can achieve high diagnostic accuracy with lower model complexity. More importantly, the generalization ability and clinical applicability of automated skin disease diagnosis systems can be further promoted by embedding advanced attention mechanisms, transformer-based feature fusion, domain adaptation, self-supervised learning, and larger multi skin disease datasets.

References

- [1] Do TT, Zhou Y, Zheng H, Cheung NM, Koh D. Early melanoma diagnosis with mobile imaging. In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2014 Aug 26 (pp. 6752-6757). IEEE.
- [2] World Health Organization. WHO's first global meeting on skin NTDs calls for greater efforts to address their burden [Internet]. Geneva: World Health Organization; 2023 Mar 31 [cited 2026 Sep 30]. Available from: <https://www.who.int/news/item/31-03-2023-who-first-global-meeting-on-skin-ntds-calls-for-greater-efforts-to-address-their-burden>
- [3] Sarkar R, Sinha S, editors. Skin Diseases in Females. Springer Nature; 2022 Nov 11.
- [4] Srinivasu PN, SivaSai JG, Ijaz MF, Bhoi AK, Kim W, Kang JJ. Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. Sensors. 2021 Apr 18;21(8):2852.
- [5] Ali MS, Miah MS, Haque J, Rahman MM, Islam MK. An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models. Machine Learning with Applications. 2021 Sep 15;5:100036.
- [6] Hasanat TU, Tahsin S, Mahmud A, Reshni NJ, Fateha TT, Hossain MS. An Interpretable Deep Learning Approach for Multi-Disease Skin Lesion Classification Using a Custom Convolutional Neural Network: A Comparative Study. In 2025 28th International Conference on Computer and Information Technology (ICCIT) 2025 Dec 19 (pp. 2265-2270). IEEE.
- [7] Esteve A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. nature. 2017 Feb 2;542(7639):115-8.
- [8] Anand V, Gupta S, Koundal D, Nayak SR, Nayak J, Vimal S. Multi-class skin disease classification using transfer learning model. International Journal on Artificial Intelligence Tools. 2022 Mar 31;31(02):2250029.
- [9] Alruwaili M, Mohamed M. An integrated deep learning model with EfficientNet and ResNet for accurate multi-class skin disease classification. Diagnostics. 2025 Feb 25;15(5):551.
- [10] Vellela SS, Roja D, Sowjanya C, SK KB, Dalavai L, Kumar KK. Multi-Class Skin Diseases Classification with Color and Texture Features Using Convolution Neural Network. In 2023 6th International Conference on Contemporary Computing and Informatics (IC3I) 2023 Sep 14 (Vol. 6, pp. 1682-1687). IEEE.
- [11] Badr M, Elkasaby A, Alrahmawy M, El-Metwally S. A multi-model deep learning architecture for diagnosing multi-class skin diseases. Journal of Imaging Informatics in Medicine. 2025 Jun;38(3):1776-95.
- [12] Almustafa KM. Predictive modeling and optimization in dermatology: Machine learning for skin disease classification. Computers in Biology and Medicine. 2025 May 1;189:109946.
- [13] Ozdemir B, Pacal I. A robust deep learning framework for multiclass skin cancer classification. Scientific Reports. 2025 Feb 10;15(1):4938. 8. Gairola AK, Kumar V, Sahoo AK, Diwakar M, Singh P, Garg D. Multi-feature fusion deep network for skin disease diagnosis. Multimedia Tools and Applications. 2025 Jan;84(1):419-44.
- [14] Gairola AK, Kumar V, Sahoo AK, Diwakar M, Singh P, Garg D. Multi-feature fusion deep network for skin disease diagnosis. Multimedia Tools and Applications. 2025 Jan;84(1):419-44.
- [15] Chen J, Jiang Q, Ai Z, Wei Q, Xu S, Hao B, Lu Y, Huang X, Chen L. Pigmented skin disease classification via deep learning with an attention mechanism. Applied Soft Computing. 2025 Feb 1;170:112571.
- [16] Sasithradevi A, Kanimozhi S, Sasidhar P, Pulipati PK, Sruthi E, Prakash P. EffiCAT: A synergistic approach to skin disease classification through multi-dataset fusion and attention mechanisms. Biomedical Signal Processing and Control. 2025 Feb 1;100:107141.
- [17] Venkat NL, Tejasri MN, Uppalapati PJ, Rao VM, Sripada VL, Murty PS. Deep Learning Strategies for Multiclass Skin Disease Classification. In International Conference on Cognitive Computing and Cyber Physical Systems 2024 Apr 4 (pp. 215-233). Cham: Springer Nature Switzerland.
- [18] Sohn A, Frankel A, Patel RV, Goldenberg G. Eczema. Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine. 2011 Sep;78(5):730-9.
- [19] Aggarwal P, Knabel P, Fleischer Jr AB. United States burden of melanoma and non-melanoma skin cancer from 1990 to 2019. Journal of the American Academy of Dermatology. 2021 Aug 1;85(2):388-95.
- [20] Vekaria AS, Brunner PM, Aleisa AI, Bonomo L, Lebwohl MG, Israel A, Guttman-Yassky E. protein levels, correlating with skin disease activity.
- [21] Wade TR, Ackerman AB. The many faces of basal-cell carcinoma. The Journal of dermatologic surgery and oncology. 1978 Jan;4(1):23-8.
- [22] Roh MR, Eliades P, Gupta S, Tsao H. Genetics of melanocytic nevi. Pigment cell & melanoma research. 2015 Nov;28(6):661-72.
- [23] Janowska A, Oranges T, Iannone M, Romanelli M, Dini V. Seborrheic keratosis-like melanoma: a diagnostic challenge. Melanoma Research. 2021 Oct 1;31(5):407-12.
- [24] Lallas A, Kyrgidis A, Tzellos TG, Apalla Z, Karakryiou E, Karatolias A, Lefaki I, Sotiriou E, Ioannides D, Argenziano G, Zalaudek I. Accuracy of dermoscopic criteria for the diagnosis of psoriasis, dermatitis, lichen planus and pityriasis rosea. British Journal of Dermatology. 2012 Jun 1;166(6):1198-205.
- [25] Jain A, Jain S, Rawat S. Emerging fungal infections among children: A review on its clinical manifestations, diagnosis, and prevention. Journal of Pharmacy and Bioallied Sciences. 2010 Oct 1;2(4):314-20.
- [26] Roh MR, Eliades P, Gupta S, Tsao H. Genetics of melanocytic nevi. Pigment cell & melanoma research. 2015 Nov;28(6):661-72.
- [27] Alrabai A, Echtioui A, Kallel F. Exploring pre-trained models for skin cancer classification. Applied System Innovation. 2025 Mar 13;8(2):35.
- [28] Sarkar R, Sinha S, editors. Skin Diseases in Females. Springer Nature; 2022 Nov 11.
- [29] Liu R, Chen Z, Zhang P. Skin lesion classification based on ResNet-50 enhanced with adaptive spatial feature fusion. arXiv preprint arXiv:2510.03876. 2025 Oct 4.
- [30] Ogundokun RO, Li A, Babatunde RS, Umezuruike C, Sadiku PO, Abdulahi AT, Babatunde AN. Enhancing skin cancer detection and classification in dermoscopic images through concatenated MobileNetV2 and xception models. Bioengineering. 2023 Aug 19;10(8):979.
- [31] Srinivasu PN, SivaSai JG, Ijaz MF, Bhoi AK, Kim W, Kang JJ. Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. Sensors. 2021 Apr 18;21(8):2852.
- [32] Jyothi UP, Dabbiru M, Bonthu S, Dayal A, Kandula NR. Comparative analysis of classification methods to predict diabetes mellitus on noisy data. In Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021

2023 Jan 1 (pp. 301-313). Singapore: Springer Nature Singapore.

- [33] Uppalapati PJ, Bhogesh KG, Gundu P, Hussain SM, Narasimharao K. A machine learning approach to identifying phishing websites: a comparative study of classification models and ensemble learning techniques. EAI Endorsed Transactions on Scalable Information Systems. 2023 May 1;10(5).
- [34] Bajwa MN, Muta K, Malik MI, Siddiqui SA, Braun SA, Homey B, Dengel A, Ahmed S. Computer-aided diagnosis of skin diseases using deep neural networks. Applied Sciences. 2020 Apr 4;10(7):2488.