CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

A Robust Framework Based on Dataset Refinement and Transfer Learning for the Recognition of Skin Lesions

by

Sheeza Naeem

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Computing
Department of Computer Science

2025

My work is dedicated, firstly, to the Almighty Allah, for blessing me with this opportunity, health, and ability to complete this. After Allah, this research is dedicated to my parents and my respected teachers, who encouraged and assisted me during my research.

CERTIFICATE OF APPROVAL

A Robust Framework Based on Dataset Refinement and Transfer Learning for Recognition of Skin Lesions

by

Sheeza Naeem

(MCS233002)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Muhammad Nouman Noor	FAST, Islamabad
(b)	Internal Examiner	Dr. Mohammad Masroor Ahmed	CUST, Islamabad
(c)	Supervisor	Dr. Farah Haneef	CUST, Islamabad

Dr. Farah Haneef
Thesis Supervisor
October, 2025

Dr. Mohammad Masroor Ahmed
Head
Dept. of Computer Science
October, 2025

Dr. M. Abdul Qadir
Dean
Faculty of Computing
October, 2025

Author's Declaration

I, **Sheeza Naeem** hereby state that my MS thesis titled "**A Robust Framework Based on Dataset Refinement and Transfer Learning for Recognition of Skin Lesions**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology (CUST), Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Sheeza Naeem)

Registration No: MCS233002

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**A Robust Framework Based on Dataset Refinement and Transfer Learning for Recognition of Skin Lesions**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Sheeza Naeem)

Registration No: MCS233002

Acknowledgement

"**A**nd whoever puts all his trust in Allah, He will be enough for him." Al-Quran [65:1].

I would like to say Alhamdulillah for everything Allah has blessed me with that enabled me to reach here. I would like to express my gratitude to my supervisor Dr. Farrah Haneef who guided me and helped me with his valuable suggestions throughout this work and gave his valuable time, May Allah keep her in his blessings. I am thankful to my parents for their love, prayers, support and everything what I needed. And to my sister and my respected teachers who keep on pushing me for higher education.

Lastly, I would also like to thank everyone who has helped me along the way. Special thanks to Mr. Muhammad Rizwan for increasing my knowledge and helping me with technical aspects giving motivational support throughout my research journey.

(Sheeza Naeem)

Abstract

Skin cancer is one of the most common and life-threatening diseases in the world, and successful treatments rely on timely and precise detection. Deep learning has achieved remarkable success in medical image analysis, with hybrid models combining multiple techniques showing promising results in enhancing diagnostic accuracy. In this research, we propose a robust hybrid framework for skin lesion detection that integrates transfer learning with dataset refinement techniques. To improve image quality and highlight lesion features, the model incorporates advanced preprocessing steps, including image standardization, segmentation, artifact removal, noise reduction, and contrast enhancement using the CLAHE (Contrast Limited Adaptive Histogram Equalization) algorithm. These refinements enhance model learning and improve classification performance. High-level features are extracted using pretrained CNN architectures such as VGG19, MobileNetv2, ResNet50, AlexNet, DenseNet121, and EfficientNetB0, followed by hybrid feature fusion for effective classification. Our hybrid framework demonstrated high accuracy(98.89 percent), precision(0.96 percent), recall(0.97 percent), and F1-score(0.96 percent) across a range of skin lesion types when tested on the publicly accessible HAM10000 dataset. These results are comparable to or better than current state-of-the-art methods in dermatological image analysis. The integration of CLAHE and hybrid transfer learning significantly improves diagnostic performance and generalizability. Overall, the proposed model offers a feasible and scalable solution for automated skin lesion identification, with strong potential for integration into teledermatology platforms, mobile health applications, and clinical decision support systems. While the framework is highly flexible, its generalization across rare disease types and varied imaging conditions remains a limitation. This research advances AI-driven dermatological diagnostics and paves the way for future work in multimodal integration and explainable AI to enhance fairness and clinical applicability.

Contents

List of Figures

List of Tables

Abbreviations

AI	Artificial Intelligence
ACO	Ant Colony Optimization
AKIEC	Actinic keratoses and Intraepithelial Carcinoma
AUC	Area Under Curve
AMBE	Absolute Mean Brightness Error
ANN	Artificial Neural Network
BCC	Basal Cell Carcinoma
BCD	Bietti's crystalline dystrophy
BCT	Brightness and Contrast Tuning
BKL	Benign keratosis-like lesions
BN	Batch Normalization
BGR	Blue Green Red
CN	Contrast Normalization
CAD	Computer-Aided Diagnosis
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Network
CPU	Central processing Unit
CPR	Clipped Pixel Redistribution
Conv2D	Convolutional 2-Dimensional layer
CII	Contrast Improvement Index
DB	Decibels
DSS	Decision Support System
DCT-DLAM	Dynamic Contrast Tuning (Deep Learning Adaptive Models)

DF	Dermatofibroma
DL	Deep Learning
DLSR	Deep Learning Super-Resolution
DNA	Deoxyribonucleic Acid
DT	Decision Tree
DSC	Dice Similarity Coefficient
DCNNs	Deep Convolutional Neural Networks
FLLF	Fast Local Laplacian Filtering
FP	False Positive
FN	False Negative
GB	Giga Byte
GHz	Giga Hertz
GANs	Generative Adversarial Networks
GPU	Graphics Processing Unit
GLCM	Gray Level Co-occurrence Matrix
HSV	Hue, Saturation, Value
HAM10000	Human Against Machine with 10,015 dermatoscopic images
HC	Histogram Clipping
HDTML	Hybrid deep Transfer Learning Method
HOG	Histogram of Oriented Gradients
IoU	Intersection over Union
ISBI	International Symposium on Biomedical Imaging
ISIC	International Skin Imaging Collaboration
JPEG	Joint Photographic Experts Group
LAB	lightness, green-red axis , blue-yellow axis
LBP	Local Binary Patterns
LAT	Lesion-Aware Transformer
LCM	Local Contrast Mapping
ML	Machine Learning
MCML	Multi-Class Multi-Level
MES	Mean Square Error

MEL	Melanoma
MFSNet	a Multi-Focus Segmentation Network
NV	Melanocytic nevi
OCF	Optimized Color Features
OpenCV	Open Source Computer Vision Library
PAD-UFES-20	Dermatological and Surgical Assistance Program at the Federal University of Esprito Santo - 20 classes of skin lesions
PDT	Photodynamic Therapy Models
PSNR	Peak Signal-to-Noise Ratio
RAM	Random-Access Memory
ReLU	Rectified Linear Unit
ROC	Receiver Operating Characteristic Curve
ResNet50	Residual Networks (50 Layers)
RCCN	Region-based Convolutional Neural Network
RGB	Red Green Blue
RTX	Ray Tracing Texture Extreme
SCC	Squamous Cell Carcinoma
SoftMax	SoftMax activation function
SSIM	Structural Similarity Index Measurement
SOTA	State of The Art Methods
SVM	Support Vector Machine
SSIM	Structural Similarity Index Measurement
TPU	Tensor Processing Unit
TP	True Positive
TN	True Negative
UV	UltraViolet
VASC	Vascular lesions
VSSSM	Visual Selective State Spatial Model
ViT	Vision Transformers
VGG19	Visual Geometry Group 19
WHO	World Health Organization

Symbols

μ	Mean pixel intensity
σ	Standard deviation of pixel intensities
L	Dynamic range of pixel values
I	Pixel intensity
N	Total number of pixels
f	Feature map value
c	Channel
H	Height of the feature map
W	Width of the feature map
P	Probability
Z	Standardized pixel intensity
g	Input feature vector
W	Weight matrix
b	Bias term
σ	Sigmoid activation

Chapter 1

Introduction

Being the biggest and the most visible organ of the body, the skin is a natural barrier that shields the body against environmental adversaries, including ultraviolet radiation, infections, dehydration, and injuries or physical or chemical factors that affect its state [1]. It plays a key role in the maintenance of temperature and excretion of wastes as well as immunological protection and provides sensory perception through receptors located in the layers of the skin known as mechanoreceptors, thermoreceptors, and nociceptors. The use of sensory receptors enables one to feel temperature, pressure, and pain and guarantees physical protection and emotional affiliation through touch. The skin also has a large contribution to identity in terms of health, age, and moods [2]. The health of the skin plays a critical role in maintaining physical homeostasis, and since skin damage frequently resulted in several mutations, including benign ones (e.g., acne, eczema) to outright cancer (melanoma), it was essential to diagnose its state properly and act accordingly [3].

1.1 Structural Organization and Functional Layers of Human Skin

The structural arrangement of the skin also directly influences the vulnerability of skin to dermatologic diseases. Each layer plays an important role in maintaining

the integrity of the skin, immunity, and regeneration, and impairments in these processes contribute to disease. As shown in Figure 1.1, the skin is composed of three distinct layers, each of which has crucial physiological as well as defensive functions.

1.1.1 Epidermis: Protection and Barrier Function

There are three layers to the skin, with the epidermis being the outermost, which functions mainly to shield the skin against environmental damage such as pathogens, UV radiation, and moisture loss. It is primarily composed of keratinocytes that are involved in the production of keratin to support skin stability. Special cells, including melanocytes, produce pigmentation, and Langerhans cells help counter immune defense. Damage to the epidermis causes diseases such as psoriasis, when there is abnormal turnover of keratinocytes, which causes dry plaques. In melanoma, there is a pigmented lesion that occurs due to unregulated growth of melanocytes. Eczema (atopic dermatitis) develops because of a broken barrier, which is characterized by red, inflamed, and itchy skin [4].

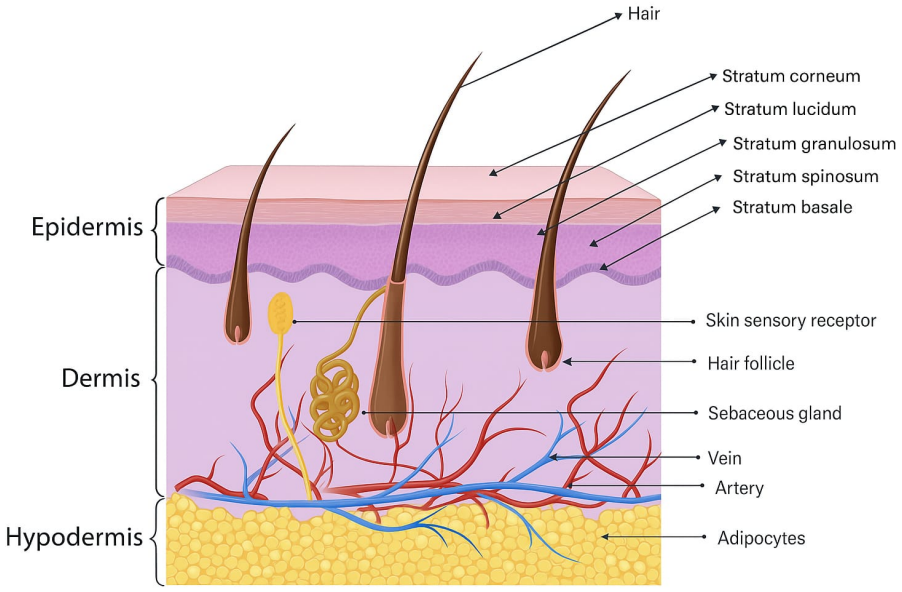


FIGURE 1.1: The epidermis, dermis, and hypodermis are the three layers of the skin. There are five sublayers within the epidermal layer. Blood vessels can be found in the layers of dermis and hypodermis; refer to [63]

1.1.2 Dermis: Structural Support and Sensory Function

The dermis provides the skin with strength, elasticity, and nutrition; this type of skin is located in the structure of the dermis that has sweat glands, nerve endings, blood vessels, and follicles. It has high elastin and collagen that make skin firm. The papillary and the reticular layers go through interactions with the epidermis to promote regeneration and perception of the skin. Lesions affecting the dermis include scleroderma, which results in excessive collagen deposition, leading to skin hardening. Rosacea is a vascular disorder that makes the skin excessively red and swollen. Destruction of nervous elements in the dermis may lead to diseases such as neuropathic ulcers that occur because of loss of sensation [1][2][4].

1.1.3 Hypodermis: Insulation and Cushioning

The hypodermis (the subcutaneous layer) is an adipose tissue that serves insulation, shock absorption, and metabolic functions. This layer contains capillary networks and sensory neurons, helping with thermoregulation and mechanical protection. Deep skin infections such as cellulitis occur when bacteria spread into the hypodermis, causing painful swelling. Disorders like panniculitis involve inflammation of fat lobules, leading to tender, nodular lesions. In systemic conditions such as lipodystrophy, fat distribution abnormalities alter skin texture and structure [1][6].

1.2 Dermatological Disorders and their Clinical Relevance

Skin diseases are among the most common medical conditions, including serious cancers like melanoma and less serious conditions like eczema and acne. It is crucial to diagnose early, as cancerous cells can quickly invade deeper skin layers, making melanoma one of the most aggressive and difficult-to-treat dermatological diseases. Beyond physical health, these conditions significantly affect self-esteem,

mental well-being, and quality of life [7]. Skin characteristics, including color and moisture, vary based on regional climates and genetic factors. Prolonged UV exposure can damage cellular DNA, reduce pigmentation, and increase the risk of malignant skin diseases, further emphasizing the need for preventive care and early screening [8].

FIGURE 1.2: Primary and Secondary Types of skin lesions ref to [64]

Dermatological diseases are broadly classified into four categories: inflammatory conditions (e.g., psoriasis, eczema), infectious diseases (e.g., bacterial and fungal infections), autoimmune disorders (e.g., lupus, vitiligo), and malignancies like dermatofibroma (DF), benign keratosis lesions (BKL), vascular tumors (VASC), melanoma (MEL), basal cell carcinoma (BCC), and squamous cell carcinoma (SCC) [9][10]. Advanced diagnostic techniques are needed for precise identification and prognosis because of the complexity of non-melanocytic skin illnesses (BCC, SCC, and others) and melanocytic skin disorders (melanoma, melanocytic nevi (NV)).

Traditional dermatological assessments rely on histopathology, dermoscopic imaging, and manual clinician evaluations, but challenges like subjectivity, inter-observer variability, and limited accessibility highlight the growing need for computer-aided diagnosis (CAD) systems [11]. Skin lesions, whether benign or malignant, require careful evaluation based on their morphology, distribution, and body location. Benign lesions may need cosmetic or medical treatment (cryotherapy, electrosurgery, or excision), while malignant lesions can spread aggressively, requiring surgical removal and histological analysis. Skin lesions often arise due to infections, allergic reactions, systemic diseases, or trauma and are categorized into primary and secondary types based on their nature and progression, as shown in Figure 1.2.

1.3 Role of AI in Dermatology

Artificial intelligence (AI) integration in dermatology has transformed the field of skin disease detection, providing automated, scalable, and non-invasive diagnostic solutions [12]. Traditional dermatological diagnostics rely on manual assessments, histopathological analysis, and dermoscopic examinations, which are subject to inter-observer variability and accessibility constraints, especially in resource-limited settings. (ML) and (DL) approaches are used in AI-driven methods to rapidly and reliably categorize skin lesions, improving diagnostic efficiency and reducing healthcare inequities [13]. According to global health data, there are an estimated 325,000 new instances of melanoma each year, making up about 1.7

percent of all cancer diagnoses worldwide [14]. Since the five-year survival rate for melanoma is almost 99 percent when discovered in its early stages, early identification greatly improves prognosis; however, in advanced cases, survival drops to 27 percent, emphasizing the need for high-accuracy AI models for timely intervention [15].

AI-based diagnostic frameworks offer decision support systems, reducing the subjectivity associated with manual interpretation and assisting dermatologists in differential diagnosis. It is crucial to accurately and timely identify skin lesions, particularly malignant ones, in order to improve survival rates. saving money on medical care and improving treatment methods. However, clinical diagnosis is often challenging due to the vast heterogeneity in lesion appearance across different populations, anatomical locations, and disease stages. The possibility of diagnostic errors even for highly skilled dermatologists has increased interest in computer-aided diagnostic (CAD) systems.

Deep learning (DL) models, particularly convolutional neural networks (CNNs), have shown impressive potential in automated skin lesion recognition in recent years by directly learning complicated hierarchical features from images [16][54]. Nevertheless, there are multiple issues that are still a barrier to the clinical reliability of AI-powered skin lesion recognition. One major issue is dataset imbalance, where certain lesion types, such as melanoma, are underrepresented, leading to biased model performance and reduced diagnostic accuracy [17][34]. Additionally, low-quality imaging in dermoscopic datasets often contains artifacts such as hairs, bubbles, and poor lighting conditions that negatively impact feature extraction and model precision [19].

In addition to these issues with data, there is the problem of generalizability, which is another major limitation. Models trained on a single dataset frequently struggle to maintain performance when applied to different datasets due to variations in skin tone, acquisition devices, and lesion types, thereby affecting their robustness in diverse clinical settings [26]. Furthermore, segmentation errors remain a critical hurdle, as inaccurate lesion boundary detection can lead to misclassification, undermining the reliability of diagnostic predictions [18][40]. Thus, dataset quality

enhancement by preprocessing, generalization, data augmentation, and balancing, along with transfer learning using robust pre-trained models, has been pursued as an excellent approach to developing robust and interpretable CAD systems in dermatology. The process of the identification of skin lesions has undergone a radical transformation throughout the decades, as nowadays it is being replaced by computer-aided (CAD) techniques that rely on AI.

In the traditional approach, dermatologists only used manual visual examination, and the diagnosis was achieved through clinical experience and knowledge. Nonetheless, it was common to find inter-observer variability, which resulted in the lack of diagnostic reliability. With the advent of computational techniques in the 1990s and 2000s, traditional machine learning (ML) methods emerged, incorporating handcrafted characteristics such as shape, color, and texture. Classifiers such as decision trees (DT) and SVM improved reproducibility, but their effectiveness was constrained by the quality of manually extracted features [14][60]. The shortcomings of feature engineering motivated the desire to seek more automated methods.

1.4 Deep Learning in Dermatological AI: Evolution and Advancements

AI in dermatology has significantly reshaped both diagnostic and treatment paradigms, revolutionizing how practitioners approach skin diseases. By using deep learning and machine learning methodologies, healthcare professionals can navigate the complexities associated with dermatological conditions more effectively. These advancements allow for the rapid analysis of dermoscopic images, identifying nuanced patterns that would otherwise elude the human eye. Furthermore, AI systems have paved the way for decision support tools that enhance the accuracy of differential diagnoses, ensuring that early interventions are more precise, potentially saving lives and reducing costs. The deep learning (DL) revolution of the 2010s marked a transformative shift in medical imaging with convolutional

neural networks (CNNs), enabling autonomous feature extraction and hierarchical representation learning [59]. CNN-based models significantly outperformed traditional machine learning (ML) methods in classification and segmentation tasks. Landmark studies [38], showed dermatologist-level performance in skin disease recognition, solidifying AI's potential in medical diagnostics.

1.5 Transfer Learning for Enhanced Performance and Data Scarcity Mitigation

To address data scarcity and improve model efficiency, transfer learning became a widely adopted strategy. The models were fine-tuned on smaller dermatological datasets after undergoing extensive pre-training on datasets such as ImageNet, decreasing overfitting, cutting down on training time, and increasing diagnostic precision [24]. Beyond architectural improvements, recent advancements emphasize dataset refinement through augmentation, balancing, de-noising, and synthetic data generation, enhancing model generalizability [52][55].

Public datasets, such as HAM10000, BCN20000, and the ISIC Archive, have established standardized benchmarks, facilitating comparative analyses and model evaluation [26][56]. Modern AI-driven techniques, in particular the domain of DL architectures, have shown remarkable outcomes in the segmentation and classification of skin diseases. CNNs have become robust tools for extracting features and hierarchical pattern recognition, learning spatial and texture-based attributes from dermoscopic images. Unlike traditional feature engineering methods, CNNs autonomously extract lesion characteristics, eliminating the need for handcrafted feature selection. Benchmarking efforts now emphasize standardized evaluation protocols to ensure reproducibility and fair comparisons across architectures and preprocessing techniques. Ensemble strategies leveraging complementary deep learning models have shown robust gains in lesion classification and segmentation, while hybrid approaches incorporating attention mechanisms and transformer-based encoders enhance contextual understanding and boundary delineation.

TABLE 1.1: Evolution of Skin Lesion Diagnosis from Manual Examination to AI-Driven Approaches

Phase	**Key Advancements**	**Limitations**
Manual Visual Examination (Pre-1990s)	Diagnosis based on dermatologists' expertise	Subjective, inter-observer variability
Traditional ML (1990s–2000s)	Handcrafted features (color, texture, shape), SVM, decision trees (DT)	Feature engineering dependency, limited adaptability
Deep Learning (2010s)	CNNs, automated feature extraction, dermatologist-level performance	Requires large datasets, computationally expensive
Transfer Learning (Late 2010s– Early 2000s)	Fine-tuning models pre-trained on ImageNet reduced training time	Limited adaptation to medical imaging specifics
Dataset Refinement Era (2000s – Present)	Augmentation, balancing, de-noising, use of standardized datasets (HAM10000, ISIC Archive)	Challenges in real-world generalization and bias mitigation

Multi-task learning frameworks are also gaining traction, enabling simultaneous lesion classification, segmentation, and attribute prediction within unified architectures. Benchmarking across diverse skin types and imaging modalities ensures fairness, mitigates bias, and supports global applicability, driving scalable and interpretable dermatological AI solutions.

1.6 CNN Architectures in Dermatological Imaging

Several CNN architectures are widely adopted for feature extraction and lesion identification, each offering distinct advantages for medical imaging. ResNet50 employs residual learning, enabling deeper networks to be trained without suffering from vanishing gradients. This mechanism contributes to the improvement of hierarchical feature extraction, and ResNet50 can be effectively used when it comes to complex analyses of dermoscopic images. InceptionV3 combines multi-scale convolutional filters, which means that the models will be efficient in terms of capturing specific lesion sizes and irregular boundaries. This architecture increases the localization of the lesions, and the classification of the features is better because the spatial features are processed at varied receptive fields.

In the meantime, EfficientNet trains the trade-off between accuracy and computation efficiency with compound scaling, which systematically and gradually adjusts the network depth, width, and resolution to achieve optimal results, together with minimizing model complexity. EfficientNet is especially ideal in medium-term, low-resource medical solutions, where CPU resource availability is limited and CPU-intensive computations are a barrier to the use of powerful models. Additionally, DenseNet improves lesion representation through feature reuse, MobileNet enables edge-based real-time inference, and ViTs capture global context for complex skin patterns. Combined with robust preprocessing and curated datasets, these architectures deliver scalable, accurate, and reliable dermatological AI solutions. These models, paired with robust preprocessing and quality data, strengthen the accuracy and clinical relevance of dermatological AI.

Recent advancements have also introduced attention mechanisms and transformer-based modules into CNN pipelines, allowing models to selectively focus on salient regions within dermoscopic images. This targeted attention enhances lesion boundary detection and improves interpretability, which is critical for clinical validation. Moreover, the integration of ensemble learning strategies, where predictions from multiple architectures are aggregated, has been shown to reduce variance and

improve diagnostic confidence, especially in borderline or ambiguous cases. In parallel, the use of transfer learning from large-scale natural image datasets like ImageNet has accelerated model convergence and improved feature generalization, even with limited annotated medical data. When combined with domain-specific fine-tuning, these pretrained models adapt effectively to the unique texture, color, and structural patterns found in skin lesions. As a result, CNN-based systems continue to evolve as powerful tools for automated dermatological assessment, bridging the gap between computational efficiency and clinical applicability.

1.7 Challenges in Artificial Intelligence-Driven Dermatological Imaging

Increasingly, technology is being used to enhance the dermatological healthcare system, with an emphasis on better disease management and accurate diagnosis. Among these developments, artificial intelligence (AI) has become a revolutionary tool that helps dermatologists more accurately and effectively diagnose complicated skin disorders. AI methodologies, particularly both domains such as DL and ML, have great promise in the analysis of dermoscopic images and mapping skin lesion patterns, thereby reducing diagnostic errors and enhancing early detection. Despite these advancements, AI-based dermatological imaging still has several technical challenges, such as imbalance in the dataset, imaging artifacts, poor segmentation, and poor generalization, as shown in Table 1.1. It is crucial to address these issues to work on model strength and applicability in clinical practice.

1.7.1 Imbalanced Datasets

Datasets of dermatological images with malignant and non-malignant cases, such as ISIC Archive, HAM10000, and BCN20000 used in research, are publicly available, and their class imbalance produces biased training of the model and lower

sensitivity during melanoma detection. These biases are reduced through data augmentation and synthetic generation of data, making the model more robust.

1.7.2 Imaging Artifacts

Dermoscopic images frequently contain artifacts such as hair occlusion, poor lighting conditions, motion blur, and color distortions, which impede feature extraction and segmentation accuracy [21]. The filters include hair removal, contrast normalization, and adaptive histogram equalization, which help improve the quality of an image.

1.7.3 Segmentation Errors

Accurate lesion boundary detection is pivotal for classification models, yet segmentation algorithms often struggle with ambiguous lesion edges, heterogeneous textures, and overlapping features [22]. Although networks such as U-Net and DeepLabV3+ enhance the boundary marking, segmentation errors observed in regions are a problem.

1.7.4 Generalization Challenges

AI models trained on single datasets often do not perform consistently across different populations and imaging modalities, restricting clinical applicability [23]. Generalization is achieved by cross-dataset validation methods and transfer learning between different sources of medical images.

Therefore, the revolution of skin lesion diagnosis from human examination to deep learning and transfer learning frameworks has significantly changed the sector. Nonetheless, there still exists research to develop robustness, generalizability, and clinical applicability of AI-driven systems. Current efforts focus on improving model interpretability, ensuring consistent performance across diverse skin types, and minimizing diagnostic bias.

Researchers are exploring multi-modal data integration, combining dermoscopic images with clinical metadata to enrich feature representation. Validation through large-scale, real-world datasets remains critical to assess clinical readiness. Additionally, explainable AI techniques are being incorporated to foster trust and transparency in automated decisions. Collaborative benchmarking initiatives aim to standardize evaluation protocols and promote reproducibility. These directions reinforce the foundation for trustworthy, scalable, and ethical dermatological AI.

Furthermore, adaptive learning strategies are being investigated to enable models to evolve with new data and clinical feedback. Integration with electronic health records (EHRs) is also under exploration to streamline diagnostic workflows and support longitudinal patient monitoring. Efforts to ensure regulatory compliance and ethical deployment are gaining momentum, particularly in teledermatology applications. Cross-disciplinary collaborations between clinicians, data scientists, and engineers are essential to translate algorithmic advances into practical tools.

1.8 Problem Statement

Despite significant advancements in AI-driven dermatology, deep learning models still face challenges such as dataset imbalance [10][16], low-resolution dermoscopic images [19], and limited generalization across diverse skin disorders [34][54]. While transfer learning has shown promise in mitigating these issues, selecting optimal pretrained models and refining datasets to enhance classification accuracy remains an ongoing challenge [9][17]. Additionally, segmentation techniques often struggle to delineate lesion boundaries precisely, affecting diagnostic reliability [18][40]. The purpose of this study is to create more accurate, generalizable, and AI-powered computer-aided diagnostic (CAD) systems for skin lesion identification that are clinically feasible. These developments would potentially enhance early diagnosis and patient outcomes and also increase access to dermatological care, especially in underserved areas of limited medical expertise. The proposed approach integrates dataset enhancement and transfer learning to boost the accuracy and interpretability of skin lesion classification.

1.9 Research Objectives

The primary objective of our study is to design, develop, and validate a robust and generalizable framework of skin lesion recognition based on dataset refinement strategies and transfer learning approaches. The framework will address the common issues of dataset imbalance, poor quality of the dermoscopic image, limited generalizability to a wide array of skin lesions, and inaccurate lesion segmentation. The goal of our suggested framework is to improve automated skin lesion recognition accuracy, interpretability, and clinical dependability. by combining enhanced dataset preprocessing, augmentation techniques, and segmentation techniques, as well as transfer learning for feature extraction. Building upon the foundational work of Gulzar et al. (2025), the aim of our research is to advance dermatology by offering a more thorough and flexible approach to AI-driven diagnostics [9]. The suggested framework's focus on clinical relevance and generalizability has led to the formulation of the following particular objective in order to operationalize the study's vision:

1. In order to evaluate and determine the methods of contrast enhancement used to enhance the image quality of skin diseases, ensuring better visibility and feature extraction for accurate classification.

2. To examine how data augmentation improves the overall recognition performance, assess improvements in model generalization and robustness.

1.10 Research Questions

Our study is guided by a series of critical research questions that capture the essential challenges raised in the context of automated skin lesion recognition. The research questions aim to systematically explore how dataset refinement and transfer learning techniques can be effectively integrated to enhance classification performance, generalizability, and clinical applicability. These questions also explore how model performance varies across lesion types with different visual complexity,

such as pigmented versus non-pigmented lesions. The study further considers the impact of input resolution and image scaling on feature retention and classification fidelity. Additionally, the investigation emphasizes how architectural fusion and preprocessing strategies jointly influence the model's ability to generalize across diverse clinical scenarios. The key research questions are as follows:

RQ1: Which contrast enhancement techniques improve the skin disease images contrast in terms of image quality measures?

RQ2: What is the effect of the data augmentation step on the overall recognition performance?

1.11 Significance of the Study

This research significantly contributes to medical image analysis, dermatology, and AI by addressing longstanding challenges in automated skin lesion recognition. One major issue is the limited quality, diversity, and balance of dermoscopic datasets, which are often skewed toward benign lesions, leading to biased model training and poor real-world performance [24][34]. To improve dataset robustness, this study incorporates preprocessing techniques (artifact removal, contrast enhancement) and augmentation methods (rotations, zoom, color jittering), fostering a more balanced, equitable dataset. This refinement enhances model training stability and improves the diagnostic reliability of rare but critical conditions like melanoma [19][26]. Since deep learning models require large, labeled datasets, medical imaging datasets often face limitations due to privacy concerns, data collection challenges, and annotation costs [54]. Leveraging transfer learning from pretrained CNN architectures (ResNet, InceptionV3, and EfficientNet) allows better generalization, faster convergence, reduced training time, and lower risk of overfitting. The study optimizes transfer learning for dermatology, demonstrating how pretrained models can be adapted to healthcare applications [9][52]. Unlike conventional classification and segmentation methods, this research introduces an integrated modular framework that combines dataset refinement, segmentation,

feature extraction, classification, and evaluation. This holistic approach improves diagnostic accuracy, interpretability, and adaptability, increasing clinician trust in AI-based tools and encouraging broader adoption in medical practice [18][40].

To increase patient survival rates, malignant skin lesions must be detected accurately and early. The suggested framework attains great specificity and sensitivity. serving as an assistive diagnostic tool for dermatologists. In resource-limited regions, AI-driven solutions can facilitate early triage, high-risk patient prioritization, and healthcare accessibility, ultimately reducing disparities in dermatological care [11][23]. From an academic perspective, this study enhances understanding of dataset quality and how transfer learning shapes deep learning performance in medical imaging. The methodologies set the foundation for future research in domain adaptation, explainable AI (XAI), semi-supervised learning, and few-shot learning, advancing dermatological AI applications. By ensuring transparency, reproducibility, and collaborative progress, this research strengthens the reliability of AI-driven dermatological analysis [16][30]. This research aligns with broader global initiatives led by the World Health Organization (2021) aimed at integrating artificial intelligence into healthcare to enhance patient outcomes, minimize diagnostic errors, and expand access to expert-level care for underserved populations [58]. By focusing on a critical area like skin cancer diagnosis, our work contributes directly to the responsible and impactful application of AI technologies in clinical settings.

1.12 Scope and Limitation of the Study

Our study's scope establishes the specific boundaries within which the research operates, including the datasets, techniques, objectives, and evaluation metrics used. It also clarifies the limitations deliberately set to maintain the study's feasibility, focus, and academic rigor.

Our study aims to develop a robust and integrated framework for automated skin lesion recognition, leveraging dataset refinement and transfer learning techniques to improve diagnostic accuracy. The research focuses exclusively on publicly

available dermoscopic image datasets, including ISIC Archive, HAM10000, and BCN20000, which provide labeled images of different skin lesions, both benign and malignant. To enhance input image quality and optimize feature extraction, preprocessing steps such as noise removal, contrast enhancement, and artifact correction (e.g., hair and bubbles) are incorporated.

Additionally, extensive data augmentation is applied to mitigate class imbalance, particularly addressing the underrepresentation of malignant cases like melanoma [19][26]. The study utilizes pretrained convolutional neural network (CNN) architectures ResNet50, InceptionV3, EfficientNet, and DenseNet for feature extraction, with fine-tuning strategies tailored to improve lesion-specific pattern recognition [9][52]. A modular pipeline integrates preprocessing, segmentation, feature extraction, classification, and evaluation components, with segmentation models such as U-Net and DeepLabV3+ enhancing lesion area detection and final classification distinguishing benign from malignant cases[40][43]. Model performance is rigorously assessed using standard measures such as F1-score, Area Under Curve (AUC), specificity, precision, sensitivity (recall), and accuracy, with cross-dataset validation such as training on HAM10000 and testing on BCN20000 conducted to evaluate generalization capability [10][16]. The classification scope is restricted to common skin lesions, such as benign keratosis, basal cell carcinoma, melanoma, and melanocytic nevi, explicitly excluding unrelated dermatological conditions such as burns, wounds, or infections. Given temporal and computational constraints, the study operates within a one-year timeframe and relies on available GPU resources and cloud-based platforms such as Google Colab and Kaggle.

Additionally, budgetary and ethical considerations necessitate the exclusive use of open-access datasets to ensure transparency, accessibility, and reproducibility [17][54]. Ultimately, this research aims to establish a strong foundation for advancing AI-driven dermatological analysis, offering valuable insights for future clinical validation and potential real-world deployment. The main elements of this study are summarized in Table 1.2. offering a clear breakdown of dataset preprocessing, transfer learning approaches, model development strategies, and evaluation methodologies. This overview provides a concise yet comprehensive insight into the

systematic framework adopted for automated skin lesion recognition. This framework emphasizes reproducibility, scalability, and clinical relevance across diverse dermatological conditions. It also reflects a commitment to ethical AI development by leveraging publicly available resources.

TABLE 1.2: Overview of Key Components in the Proposed Framework for Skin Lesion Detection and Classification

Component	**Description**
Dataset Collection	Uses ISIC Archive, HAM10000, and BCN20000. Applies preprocessing and augmentation to improve image quality and balance classes [19][26].
Transfer Learning	Fine-tunes ResNet50, InceptionV3, EfficientNet, and DenseNet for dermoscopic skin lesion recognition [9][52].
Framework Development	Builds a modular pipeline integrating preprocessing, segmentation (U-Net, DeepLabV3+), feature extraction, and classification [40][43].
Model Evaluation	Evaluates performance using cross-dataset validation, F1-score, AUC, sensitivity, specificity, and precision [10][16].
Classification Scope	Targets benign keratosis, basal cell cancer, melanocytic nevi, and melanoma, excluding non-lesion disorders [17][54].

The integration of standardized evaluation metrics ensures consistency and comparability across experiments. Moreover, the use of transfer learning facilitates efficient model adaptation to varied imaging modalities. These components collectively support the development of robust, interpretable, and clinically deployable diagnostic systems. Our study focuses exclusively on dermoscopic images captured using specialized imaging devices, deliberately excluding clinical photographs, histopathological images, and smartphone-captured images to maintain consistency in image quality and diagnostic relevance. It employs a supervised deep learning approach with labeled datasets, leaving unsupervised, semi-supervised, and reinforcement learning methods beyond its current scope. The evaluation is restricted to selected pretrained architectures ResNet, Inception, and EfficientNet chosen for their established performance in medical imaging, while newer models like Swin Transformers and Vision Transformers (ViT) are set aside for further research.

Although the study offers a thorough evaluation of the model's performance, it does not extend to real-time system deployment; however, feasibility for future implementation is discussed. Additionally, clinical trials involving live patients are not conducted, with validation limited to benchmark datasets and cross-dataset evaluations, and recommendations for future clinical validation outlined. The research operates within a one-year timeframe and is constrained by available computational resources, relying on GPU servers and cloud platforms such as Google Colab and Kaggle. Budgetary and ethical considerations further shape their direction, necessitating the exclusive use of open-access datasets to ensure accessibility, transparency, and reproducibility.

1.13 Key Contribution

This research primarily contributes by merging sophisticated deep learning models with dataset refinement approaches to design and implement a reliable, generalizable, and clinically relevant hybrid framework for automated skin lesion detection. It further introduces a reproducible pipeline that integrates enhancement, seg-

mentation, and classification stages with modular flexibility. It also emphasizes transparent evaluation through per-class metrics and confusion matrix analysis to highlight diagnostic strengths and limitations. Additionally, the framework supports translational deployment in teledermatology settings via lightweight model integration. Quantitative benchmarking across diverse datasets validates its robustness and clinical utility. The main contributions are summarized below:

- Develop a hybrid framework (DenseNet121 and EfficientNetB0) that combines transfer learning with dataset refinement approaches to accurately classify skin lesions.
- Using CLAHE-based contrast enhancement and artifact removal resulted in improved image clarity and boost feature extraction quality.
- Incorporated data augmentation techniques such as rotation, scaling, flipping, and zooming that help to address class imbalance and model generalization challenges.
- Evaluated on the HAM10000 dataset, the proposed model outperformed baseline CNNs in terms of accuracy, precision, recall, and F1-score.
- Effectively addressed each of the research questions by indicating that augmentation improves model performance and ensuring the influence of the CLAHE algorithm on image quality and clarity.

1.14 Thesis Organization

1.14.1 Chapter 1

The background and context of the study are set in Chapter 1, the research problem, objectives, and research questions are clearly formulated. It goes into the overview as well as the importance of the research and introduces technical terms that will be used in the thesis. It also introduces the idea behind creating a powerful framework using refinement of the dataset and transfer learning to recognize

skin lesions. The scope and limitations of the research, its significance, the objective achieved, and the thesis organization are also covered in this chapter. The chapter also highlights the interdisciplinary nature of the study, bridging clinical dermatology with advanced computational techniques. It emphasizes the need for scalable solutions to address diagnostic disparities in skin healthcare. Moreover, it sets the stage for subsequent chapters by outlining the methodological roadmap and anticipated contributions of the research.

1.14.2 Chapter 2

In Chapter 2, the review of the existing research and methods concerning skin lesion detection and classification, it is given in detail. It compares and critically evaluates the weaknesses and strengths of the key tools of traditional machine learning, deep learning models, transfer learning, and refinement of the dataset. It forms a context for the proposed framework in the sense that it pinpoints the gaps and difficulties of the existing body of knowledge and thus provides grounds for the proposed framework to fill the gaps. The chapter also highlights the limitations of current approaches in handling class imbalance, low-resolution dermoscopic images, and inconsistent annotation standards. Moreover, it discusses the evolving role of hybrid architectures and ensemble techniques in improving diagnostic accuracy and model robustness. This comprehensive review serves as a foundation for selecting appropriate methodologies and justifying the design choices made in subsequent chapters.

1.14.3 Chapter 3

Chapter 3 indicates the research methodology to be used to attain the objectives of the study. It details the data that was utilized, preprocessing, and data augmentation methods, as well as how the transfer learning models were used. Further, it describes how the integrated framework was developed and how the metrics to evaluate the framework, the experiment design, and validation techniques were used to make the results reliable and reproducible.

This chapter also outlines the rationale behind model selection and the comparative evaluation of multiple CNN architectures. It explains the implementation of stratified K-fold cross-validation to ensure fairness across imbalanced classes and discusses the role of performance benchmarking in assessing diagnostic accuracy. Additionally, it highlights the importance of reproducibility through modular pipeline design and emphasizes the clinical relevance of the chosen evaluation metrics.

1.14.4 Chapter 4

In the next section Chapter 4, the results of the conducted experiments are presented and discussed, assessing the work of the suggested framework in terms of its accuracy, precision, recall, F1-score, and AUC. These findings are discussed with regard to the current methods and compared with the existing solutions, as well as reflected in research questions. The chapter evaluates the strengths, limitations, and observations based on the experiments in a systematic manner, thus giving insights on the implications of the experiments.

1.14.5 Chapter 5

At the end, Chapter 5 contains a conclusion of the main findings of the research, conclusions formed under the objectives of a study, and the reflection of the importance of the work. It describes the practical implementation of the developed framework and adds to the body of knowledge in the field of AI-driven dermatology. It also suggests future studies be done, such as incorporating explainable AI, multi-modal and clinical validation, to make the AI models used in dermatology analysis more robust and applicable.

Chapter 2

Literature Review

Recent years have seen a significant increase in interest in the diagnosis of skin lesions, which is preconditioned by the increased incidence of skin cancer, as well as the necessity to early and accurately diagnose the skin lesions. The advancement in artificial intelligence generally and deep learning and transfer learning more specifically has revolutionized analysis and diagnostics of dermatological diseases. The chapter reviews the present state of this field's research as well as approaches to the enhancement of datasets, data augmentation, and classification model application in the diagnosis automatization. The review aims to point out key contributions, strengths, and weaknesses of the present methods and to examine datasets and performance metrics used in different works. The chapter is a synthesis of the existing research work that provides a broad picture of the developments in the recognition of skin lesions and forms the basis of further development by building a strong and efficient framework presented in this study.

2.1 State-Of-The-Art Methods In Dermatology Artificial Intelligence

The development of artificial intelligence technologies, especially the methods of deep learning and transfer learning, has transformed dermatological diagnostics.

Machine learning models are currently better at classifying skin lesions compared to traditional methods with increased accuracy, efficiency, and scalability. The chapter discusses innovative approaches, examining the most important innovations, methods, and issues in AI-supported skin disease detection.

Gulzar et al. (2025) presents a Hybrid Deep Transfer Learning Method (HDTLM) for predicting skin disorders, which combines DenseNet121 for deep feature extraction and EfficientNetB0 for enhanced computational efficiency. They have data consisting of 19,171 images of 19 various skin conditions, hence a large set of lesion representations. The visibility of lesions is enhanced with various preprocessing methods that include contrast enhancement, image normalization, etc. DenseNet121 ensures hierarchical feature extraction, while EfficientNetB0 reduces the number of computations, which makes it scalable. Classification accuracy is increased by adaptive fine-tuning. The performance demonstrated through experimentation is impressive, such that training accuracy stands at 98.18 percent, whereas validation accuracy stands at 97.57 percent, exceeding the performance of the baseline models. The statistical study shows better sensitivity and specificity, resulting in reduced false positives. Clinical applications are the focus of the study, and it is proved that AI-aided dermatology can be used to identify the disease at an early stage. Issues such as computational complexity are challenges that need optimization in order to be successfully applied in real time. The data bias is a possible limitation to generalization and can influence performance on a diversity of skin tones. Overfitting necessitates the use of regularization techniques to enhance robustness. The absence of explainability is a factor that hinders clinical usability, and thus effectiveness should be more transparent. Further studies can explore federated learning as a method of improving model fairness. Multi-modality may enhance the reliability of diagnosis. The results affirm that deep transfer learning significantly advances AI-supported classification of skin disorders [9].

Yu et al. (2025) addressed class imbalance in dermatological image classification using rotation, flipping, and brightness adjustments, improving their PDT model's accuracy from 87 to 92 percent. Their method increased the diversity and generalization of their features, especially on rare skin conditions. Although this has

been improved, there were still problems with skin tone variations and complicated lesion shapes. The study highlighted the necessity of more flexible augmentation methods and more profound deep learning incorporation to be robust. Future research should refine preprocessing techniques, expand dataset diversity, and optimize computational efficiency to balance complexity and accuracy in AI-driven dermatological diagnosis [10].

Zhang et al. (2023) reviewed data augmentation's role in improving model generalization and reducing bias in dermatological datasets, particularly Fitzpatrick17k. They discovered that rotation, changes in contrasts, and generation of synthetic images helped to diversify skin hues and disease types in a fair way in AI diagnostics. Although augmentation helped to alleviate the imbalances in the datasets, there were still issues regarding quality and real-world applicability. The study emphasized adaptive augmentation strategies to refine dermatological AI models, recommending future research to optimize techniques and expand dataset diversity for greater fairness and reliability [11].

Tumpa and Kabir (2021) introduced an ANN-based method for melanoma classification, integrating ABCD, GLCM, and LBP features to enhance lesion characterization. They employed the Maximum Gradient Intensity preprocessing with Otsu Thresholding to provide a better segmentation using ISIC Archive and PH2 images. Their method obtained an accuracy of 97.7 percent, better than the conventional method, and they could decrease false positives and increase sensitivity. The study highlighted AI's potential in early melanoma detection, suggesting future research on attention mechanisms, multi-modal imaging, and federated learning for broader applicability [12].

Saba et al. (2019) developed a deep learning framework for skin cancer detection, integrating contrast enhancement, boundary detection, and feature fusion for improved classification. They used FLLF and Hue, Saturation, Value (HSV) color transformation to enhance the contrast of lesions on PH2 data, ISBI 2016 data, and ISBI 2017 data, Inception V3-based feature extraction, and Hamming distance-based fusion. Their approach presented 98.4 percent accuracy on PH2 and 95.1 percent and 94.8 percent accuracy on ISBI 2016 and 2017, respectively,

compared to the traditional methods. The study highlighted AI's role in early melanoma detection and proposed future research on multimodal imaging and federated learning for broader applicability [13].

Nasir et al. (2018) proposed a hybrid segmentation technique combining texture and color-based lesion extraction with Boltzmann entropy-driven feature selection, improving diagnostic precision. They processed the dataset PH2 with Dull-Razor preprocessing and probability-based segmentation in order to provide better lesion boundaries, as shown in Figure 2.1. Their color, texture, and HOG descriptor-based model scored 97.5 percent accuracy, 97.7 percent sensitivity, and 96.7 percent specificity in comparison to their traditional techniques. The study emphasized AI-assisted dermatology's role in early melanoma detection while highlighting challenges in computational efficiency. Future research may explore multi-modal imaging and federated learning for broader applicability [14].

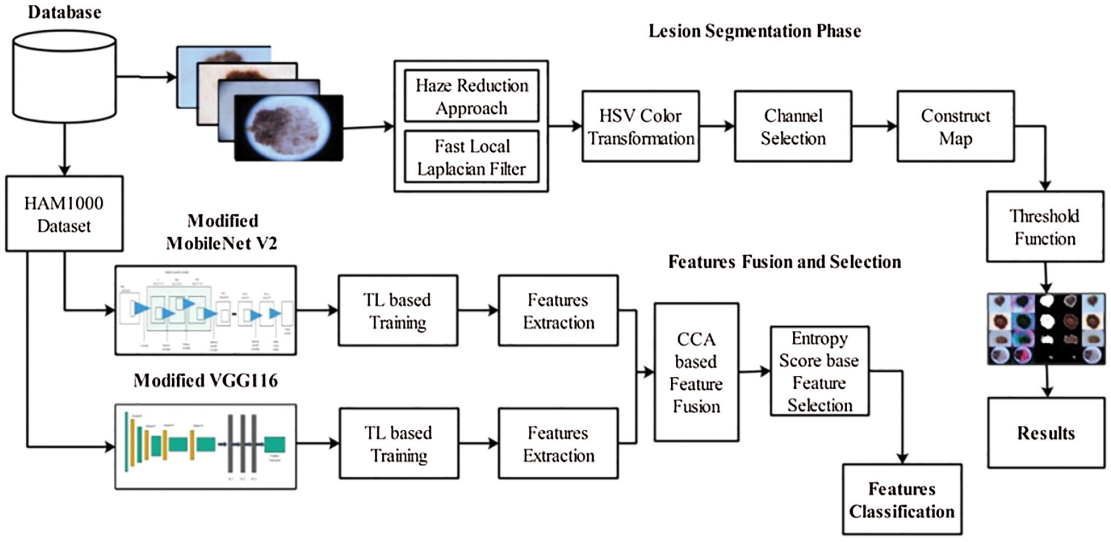


FIGURE 2.1: The proposed parallel architecture of skin lesion segmentation and classification ref to [14]

Zhang et al. (2021) examined challenges in deep learning-based skin disease classification, highlighting issues in dataset diversity, model generalization, and interpretability. They analyzed CNN architectures, such as ResNet, EfficientNet, and DenseNet, and showed that misclassifications of rare conditions occurred when the dataset was imbalanced, and representation bias, which undermined the robustness of models to diverse populations, occurred. The complexity of computation

makes it not usable in real time, making it require optimal efficiency. The study emphasized the need for explainable AI, recommending federated learning, multi-modal integration, and adaptive augmentation to enhance fairness and scalability in dermatological AI solutions [15].

Choy et al. (2023) tackled class imbalance in dermatological image classification by using synthetic data generation on a 10,000-image dataset, enhancing Vision Transformers (ViTs) accuracy from 87 to 92 percent. This allowed training data to be more diverse and better represent rare lesion types, which reduces the misclassification rates. Nevertheless, issues of identifying the authenticity of the synthetic images and their strength in diverse skin tones were still present. Future research should explore adaptive synthetic data generation and multimodal learning to further refine diagnostic precision [16].

Venkatesh et al. (2024) advanced rare disease classification by integrating synthetic image generation with active learning in a 15,000-image multi-modal dataset, increasing ResNet's accuracy from 86 to 91 percent. Otherwise, their approach reduced the variance of unrepresented categories of diseases, enhancing the diversity of data and applicability. Along with these developments, there were issues with preserving synthesized image authenticity and replicating reliability in real-world datasets. Future studies should focus on hybrid learning strategies incorporating clinical insights to enhance classification accuracy [17].

Alhudhaif et al. (2023) improved dermatological classification by applying rotation and scaling methods to a 4,000-image dataset, boosting the accuracy of a soft attention CNN from 87 to 92 percent [18]. This method increased diversity in features, especially with complex variations between lesions, thus resulting in low misdetection. Nevertheless, it was difficult to maintain a consistent augmentation and refine feature extraction in different skin conditions. Future research may explore adaptive learning techniques and expanded dataset diversity to improve the clinical reliability of AI-driven dermatology. Moreover, the study underscored the limitations of static augmentation strategies in capturing dynamic lesion evolution over time. Incorporating temporal data and patient-specific attributes could further enhance diagnostic precision and model personalization.

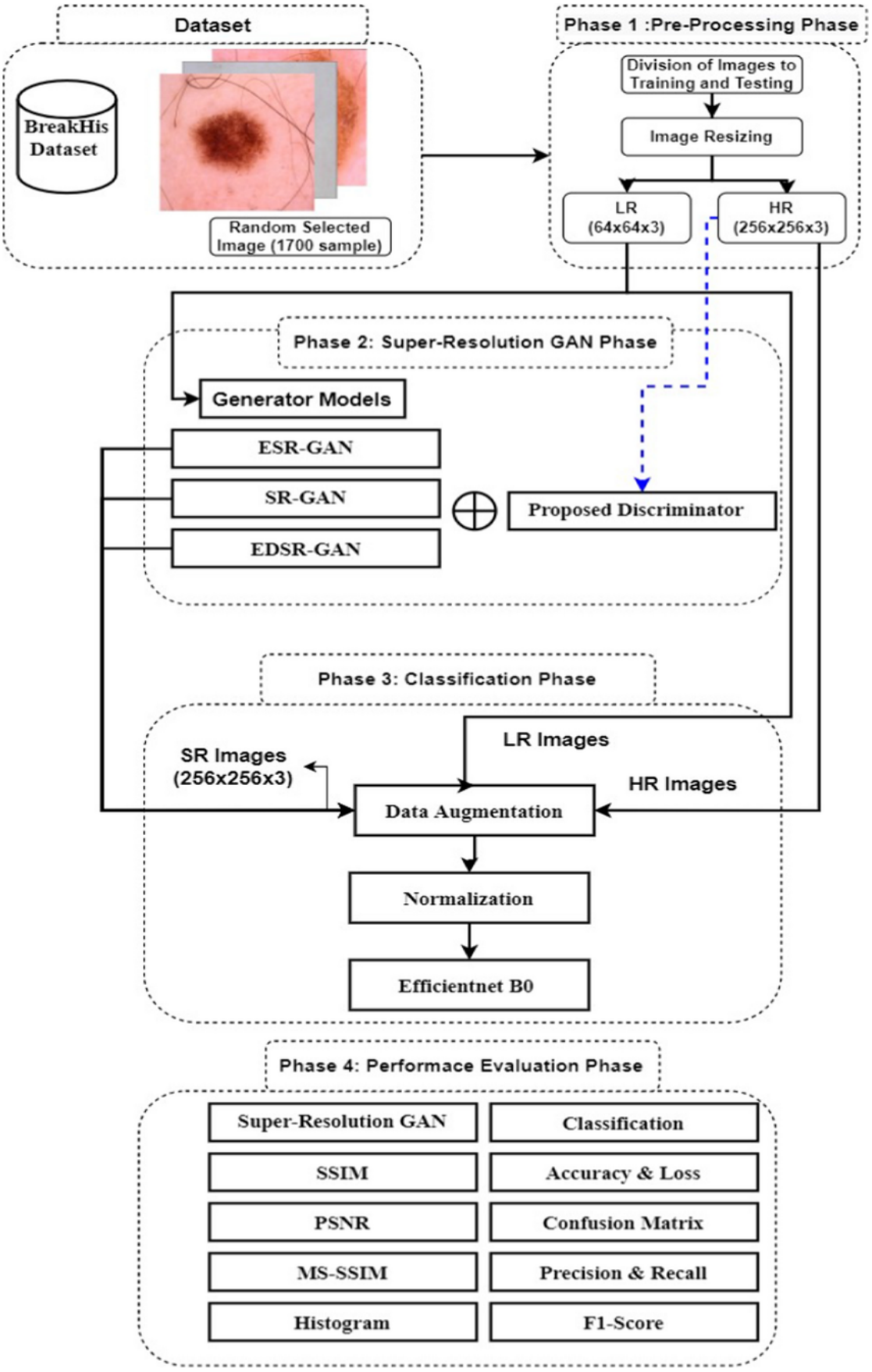


FIGURE 2.2: The proposed hybrid SR-detection framework ref to [20]

Abir et al. (2024) proposed a DL framework for melanoma detection, integrating automated preprocessing and data augmentation to enhance classification performance. Using contrast normalization, resizing, and synthetic augmentation like rotation and flipping, they improved EfficientNet-based lesion classification accuracy from 88 to 93 percent [19]. They improved the representation of malignant cases to lower false positives, but fine-tuning was still essential with synthetic data authenticity and the balance of distribution. Future research should refine adaptive augmentation techniques and incorporate multimodal datasets for better clinical relevance.

Abd El-Fattah et al. (2023) combined the DLSR technique with classification to improve skin disease detection. Their framework utilized high-resolution images, applying super-resolution methods to enhance lesion details before classification using ResNet and EfficientNet architectures, as illustrated in Figure 2.2. Brightness adjustments and contrast normalization further improved accuracy, achieving 94 percent while boosting sensitivity and specificity [20]. Challenges included maintaining image authenticity and computational efficiency, emphasizing the need for adaptive super-resolution techniques. Future work may explore multi-modal integration and hybrid learning for more reliable diagnostics.

Shakya et al. (2025) evaluated deep learning and transfer learning strategies for skin cancer classification, benchmarking VGG19, ResNet18, and MobileNetV2 across ISIC 2018 and DermNet datasets. They had a preprocessing pipeline that used contrast enhancement, active contour segmentation, and artifact removal to provide specific lesion extraction. ResNet18 and MobileNetV2, both using an SVM classifier, achieved significantly high accuracy (92.87 percent) that was higher than the standalone models using a hybrid feature fusion of color histograms, texture descriptors, and CNN features. Imbalance of datasets and the difficulties associated with computational challenges highlighted the necessity of adaptation of augmentation and AI fairness-focused strategies. Future research may explore multi-modal fusion and federated learning to improve diagnostic reliability [21]. They further underscored the relevance of lightweight and scalable architectures for deployment in resource-constrained clinical environments. Additionally, it also

promotes open-source dissemination of code and datasets to foster collaborative advancement and validation across the research community. The study advocated for standardized evaluation protocols to ensure reproducibility, transparency, and cross-study comparability in dermatological AI research.

Huang et al. (2025) improved severity scoring in dermatological classification by applying zooming and contrast modifications to a 10,000-image dataset, enhancing their Lesion-Aware Transformer (LAT) model's accuracy from 86 to 91 percent. Their strategy increased sample diversity, leading to an improvement in severity assessment and misclassifications, particularly in borderline cases. Nevertheless, there were still difficulties in handling complex lesion patterns and robustness across skin tones. Future research may explore dynamic augmentation techniques and contrast optimizations for better generalization [22].

Malik et al. (2024) addressed class imbalance in skin lesion classification by applying rotation and flipping to a 5,000-image dermoscopic dataset, improving DenseNet's accuracy from 87 to 92 percent. They improved the representation and diversity of features, which resulted in more recognition of underrepresented lesion types. Irrespective of these, maintaining uniformity of augmentation and treating various lesion complexities was difficult. Future investigations should focus on adaptive augmentation and hybrid strategies to improve dermatological AI performance [23].

Kuldeep Vayadande et al. (2024) refined CNN-based skin lesion classification by incorporating rotation, scaling, and brightness adjustments on a 3,000-image dataset, boosting classification accuracy from 85 to 90 percent. This increased the identification of malignant cases and reduced false positive cases and, thus, bettered the representation of both malignant and benign lesions. There were difficulties associated with consistent augmentation and the ability to respond to various lesions. Future research should enhance augmentation techniques and integrate hybrid learning for improved generalization and robustness in dermatological AI. Moreover, the study highlighted the importance of lesion diversity and contextual features in improving model sensitivity. Leveraging domain priors and adaptive preprocessing can boost diagnostic precision [24].

Venkatesh et al. (2024) leveraged synthetic image generation and active learning to enhance rare disease classification using a multi-modal dataset of 15,000 images, raising ResNet's accuracy from 86 to 91 percent. Their approach solved the issue of data diversity and underrepresented disease categories and enhanced generalization and minimized the misclassification rates. Nonetheless, it was difficult to prove the validity of synthetic images and keep high performance in diverse data. Future work could explore hybrid approaches integrating clinical insights and multimodal data for improved AI-driven dermatological diagnostics [25].

Hernández-Pérez et al. (2024) addressed lighting variability and occlusion challenges in dermatological image classification within the BCN20000 dataset, using augmentation techniques like contrast modifications and occlusion-resistant preprocessing. Their solution resulted in a higher rate of model robustness, with performance increased by 4 percent in benchmark tests. Nonetheless, the process of consistent augmentation and practical adaptability was a challenge. Future research should refine dynamic augmentation strategies and integrate multimodal learning to enhance classification accuracy and generalization [26].

Abobakir and Abdulazeez (2024) reviewed the role of data augmentation in mitigating class imbalance in dermatological image classification, analyzing its effects on SVM and CNN performance across multiple datasets. Their results indicated a 3 to 5 percent increase in accuracy and the necessity of having different training samples stressing the balanced lesion coverage. Although these advantages were associated with augmentation strategies, challenging work was realized in optimizing augmentation strategies to avoid overfitting and preserve the authenticity of images. Future work should explore adaptive augmentation techniques and refined preprocessing for improved diagnostic accuracy [27].

Mishra and Kaushik (2024) improved U-Net segmentation accuracy for dermatological image analysis by applying rotation, flipping, and contrast normalization to a dataset of 4,000 images. They made the IoU score higher, 85 to 90 percent, by raising the variability of lesions and better detecting boundaries. Segmentation mistakes were reduced, but extreme lighting situations and abnormal forms of lesions still presented the impediments. Future research could investigate multi-scale

feature integration and dynamic contrast adjustments to improve segmentation performance in dermatological AI [28]. Additionally, the study emphasized the limitations of static preprocessing in adapting to diverse imaging conditions. Incorporating adaptive thresholding and context-aware segmentation could further enhance boundary precision and reduce false positives in complex cases.

Harbola et al. (2024) tackled class imbalance in dermatological image classification using rotation, flipping, and synthetic image generation, enhancing EfficientNet's accuracy from 87 to 92 percent [29]. Their approach led to a better feature diversity and display of less represented lesion categories and achieved better misclassification rates. Nevertheless, please improve the quality and stability of the augmentation with various types of skin. Future researchers are encouraged to investigate the techniques of hybrid augmentation and clinical variation in the real world to get better diagnostic results.

Magalhães et al. (2024) mitigated overfitting in small dermatological datasets through brightness adjustments and flipping, increasing CNN accuracy by 4 percent. They did better in training stability and lesion recognition of a dataset comprising 2,000 images, but there were issues using real and synthetic data in balancing. Future research should prioritize adaptive learning strategies and multimodal data integration to strengthen CNN robustness for dermatological AI applications. The study also underscored the need for dynamic augmentation pipelines tailored to lesion variability. Incorporating uncertainty modeling could further enhance decision reliability in clinical settings [30].

Rakesh et al. (2024) improved EfficientNet's classification accuracy from 86 to 91 percent by applying random rotations, flipping, and synthetic image generation to enhance class diversity in a dataset of 4,000 dermatological images. Their strategy improved the incidence of detecting lesions, especially those that exhibit high intra-class variability, as well as minimized misclassifications. Nonetheless, it was hard to balance synthetic and real image distributions. Future research should refine augmentation strategies and incorporate adaptive learning techniques for improved robustness. Integrating uncertainty-aware modeling could further enhance diagnostic confidence and clinical applicability [31].

Tyagi et al. (2024) conducted a review emphasizing data augmentation's role in addressing scarcity in dermatological image classification. They discussed augmentation and achieved accuracy improvements that ranged between 3 percent to 5 percent depending on various CNN-based models, as illustrated in Figure 2.2. Their analysis highlighted augmentation's role in mitigating overfitting and improving lesion representation but noted challenges in optimizing augmentation approaches while preserving image authenticity. Future investigations should focus on dynamic augmentation techniques and increasing dataset diversity for enhanced classification performance [32].

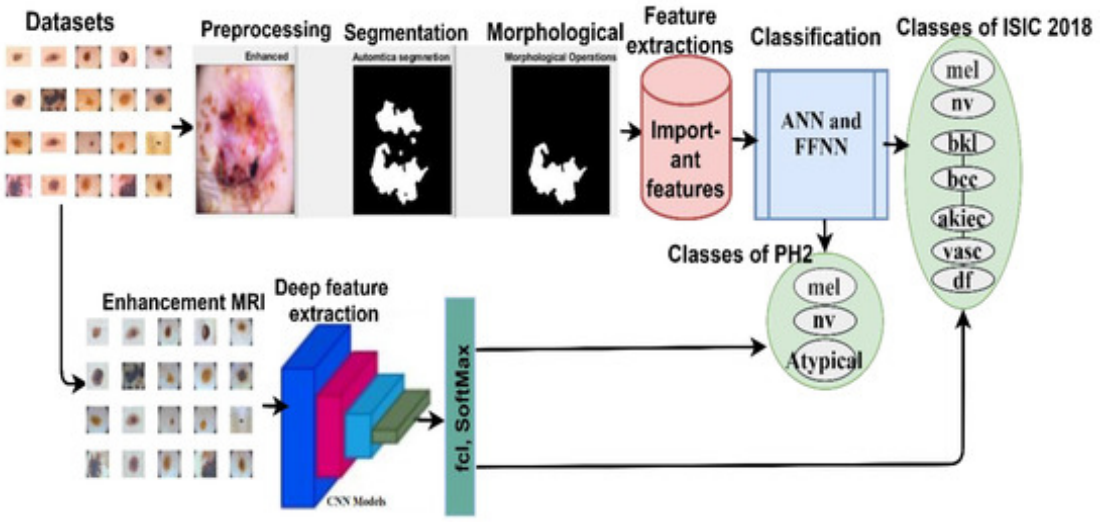


FIGURE 2.3: The methodology for classifying the ISIC 2018 and PH2 datasets used in this study ref to [32]

Anggriandi et al. (2023) enhanced CNN-SVM classification of dermatological images by applying rotation and flipping techniques to a dataset of 2,500 images, improving accuracy from 85 to 90 percent. This approach fortified their capability to feature diversity and better representation of uncommon lesion types and lowered misclassification rates. Nevertheless, it was difficult to ensure augmentation consistency and flexibility regarding different skin conditions. Future research should explore hybrid augmentation techniques and integrate real-world variations for better clinical applicability [33].

Qureshi and Roos (2023) addressed class imbalance in dermatological image classification by applying synthetic image generation to a 5,000-image dataset, elevating

ensemble CNN accuracy from 88 to 93 percent. Their method enhanced the recognition of the lesions and the balanced representation of the datasets, but they had difficulties with the authenticity of the synthetic images and the balance of the classification that occurred amongst the varied skin tones. Future research may refine augmentation strategies and focus on adaptive learning methods to promote fairness in AI-driven dermatology [34].

Wan et al. (2022) improved classifier fusion accuracy from 85 to 90 percent by applying brightness adjustments to a 3,000-image dermatological dataset. Their approach improved consistent illumination, which is beneficial to feature extraction and the recognition of lesions, especially in underrepresented cases. Nevertheless, to achieve stability of augmentation in various light conditions was a challenge. Future work should refine preprocessing techniques and explore multimodal data integration for better real-world diagnostic applications [35].

Akter et al. (2022) enhanced class representation in dermatological image classification by using synthetic image generation on a dataset of 5,000 images, improving EfficientNet accuracy from 86 to 91 percent. The strategy reduced lesion class underrepresentation, improving the generalization of the models, as well as minimizing the misclassification percentage. Nevertheless, there were difficulties associated with the authenticity of synthetic images and consistent acquaintance with different skin colors. Future studies should focus on improving augmentation strategies and integrating multimodal learning for improved diagnostic precision [36].

Ahammed et al. (2022) tackled class imbalance in dermatological image classification through oversampling and synthetic data generation on a 2,000-image dataset, raising CNN accuracy from 84 to 89 percent. Their technique enhanced the quality of lesion representation and model generalization, but problems occurred in the authenticity of synthetic images and consistent adaptation to a wide range of skin conditions. Future research should explore adaptive augmentation methods and multimodal learning to enhance classification reliability and equity in dermatological AI. The study stressed enhancing synthetic realism and context-aware sampling, with uncertainty modeling to support equitable decisions across

diverse skin types. The study also emphasized the importance of balancing synthetic and real data distributions to preserve clinical relevance. Incorporating lesion-specific priors could further improve model sensitivity across rare conditions. Future frameworks may benefit from scalable validation protocols to ensure robustness across diverse populations [37].

Liu et al. (2020) augmented a multi-modal dermatological dataset of 20,000 images using rotation and flipping techniques, improving CNN accuracy from 87 to 92 percent. Their strategy promoted feature diversity and improved generalization across lesion types and minimized misclassification rates. There were challenges relating to ensuring consistent augmentation and optimization of feature extraction over the various imaging modalities. Research could be done in the future to explore adaptive learning and integration of multimodal datasets to enhance their clinical applicability [38].

Bajwa et al. (2020) applied data augmentation methods, such as flipping, rotation, and contrast corrections, to a 10,000-image dermatological dataset, increasing ResNet accuracy from 86 to 91 percent. Their method enhanced model generalization rates and classification of rare lesions and kept the misclassification errors to a minimum. Nevertheless, stable enhancement and versatility in different skin colors were problematic. Future studies should refine augmentation methods and explore multimodal data integration to enhance AI-driven dermatological diagnostics. Additionally, the study emphasized the need for skin tone-aware preprocessing to improve fairness across diverse populations. Incorporating adaptive fusion strategies may further enhance lesion detectability under varied imaging conditions. The study also highlighted the importance of preserving lesion texture and color fidelity during augmentation. Future models may benefit from tone-aware normalization to improve consistency across varied skin presentations [39].

Bai et al. (2025) introduced the Visual Selective State Spatial Model (VSSSM) shown in Figure 2.5, a segmentation method designed to address irregular lesion shapes, lighting variations, and artifacts in dermatological images. By incorporating spatial constraints, their approach enhanced the Jaccard index and Dice Similarity Coefficient (DSC) compared to DeepLabV3+, U-Net, and Mask RCNN

Figure 2.4. Even though more accurate methods of segmentation were developed, computational efficiency and scalability were still challenging. Future research could explore multi-modal imaging integration for broader AI-driven dermatological applications [40].

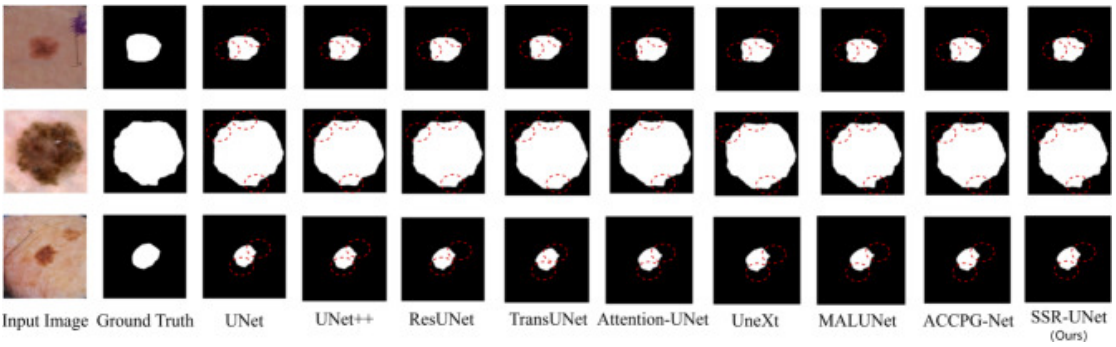


FIGURE 2.4: Illustration of the comparison results of the ISIC2017 data-set [40]

Sarwar et al. (2024) developed a hybrid deep learning framework combining Ant Colony Optimization (ACO) with ResUNet for skin lesion segmentation, with a 95.8 percent accuracy rate, 93.1 percent Dice coefficient, and 87.5 percent Jaccard index. They were able to improve on the lesion boundaries, reducing the number of false positives and enhancing generalization, especially of melanomas. Although it was efficient, the limitations still existed in computational efficiency and scalability. Future research should explore multi-modal imaging and optimization strategies for enhanced segmentation performance [41].

Thwin and Park (2024) proposed an ensemble deep learning strategy incorporating multiple segmentation and classification models to improve skin lesion analysis, achieving a Dice coefficient of 0.93, an IoU of 0.90, and classification accuracy of up to 99 percent on a balanced dataset. Their approach increased the accuracy of defining lesion boundaries and detecting melanoma using transfer learning; however, it had limitations concerning computational complexity. Additionally, their methodology demonstrated the potential of ensemble learning to harmonize diverse model outputs for improved diagnostic precision. The study also highlighted the importance of balancing performance gains with computational efficiency, especially for real-time clinical deployment. Future studies could integrate attention mechanisms to further refine segmentation accuracy [42].

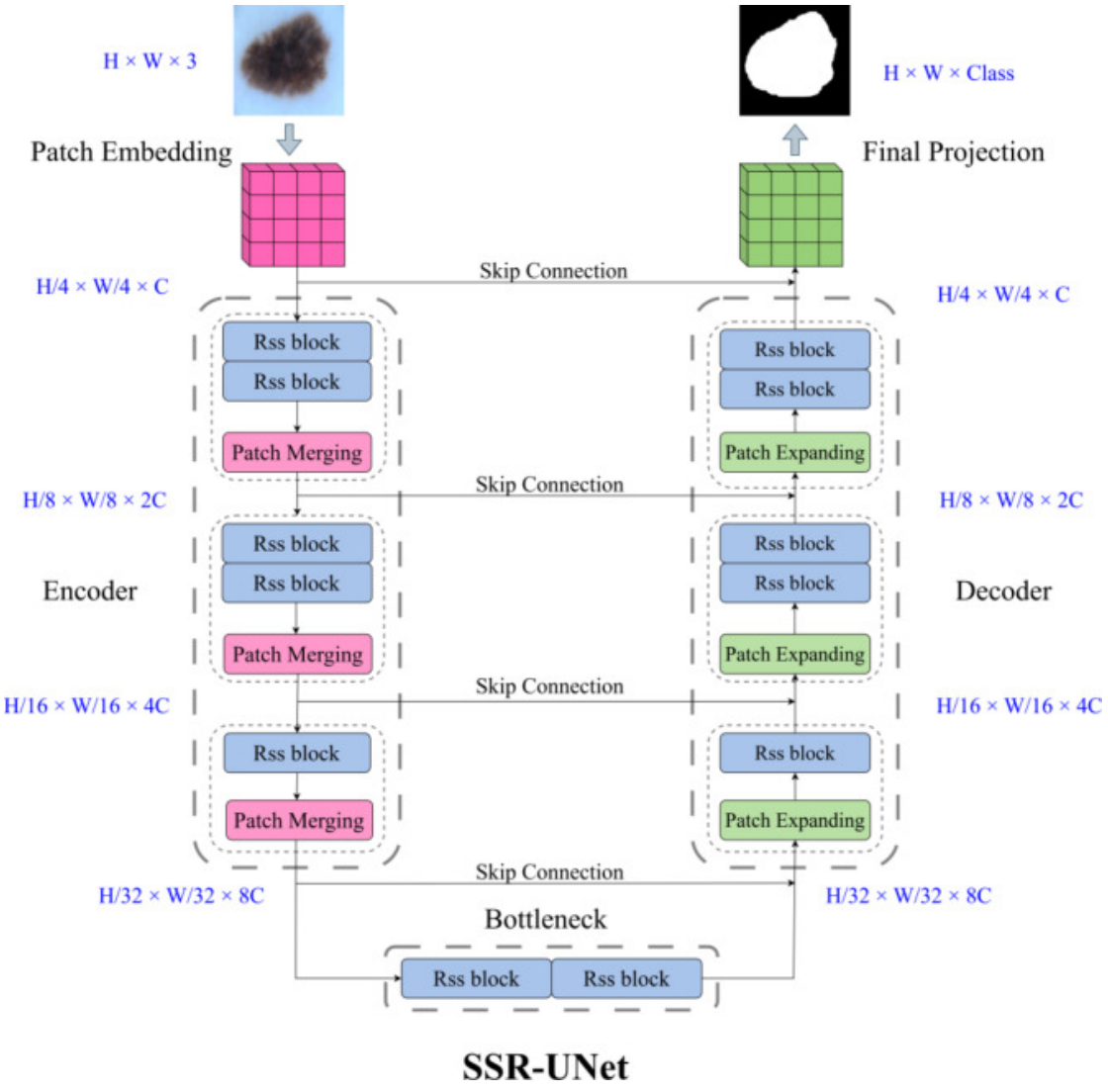


FIGURE 2.5: Description of SSR-Unet framework ref to [40]

Akram et al. (2023) introduced a hybrid deep learning model that combines ResNet50 and Mask R-CNN for segmentation for classification, optimizing feature extraction for AI-assisted dermatological diagnostics. In their test, they were able to get 95.49 percent segmentation accuracy and 96.75 percent classification accuracy, cutting down the false positives and increasing sensitivity. Nevertheless, real-time implementation was difficult with respect to computational efficiency. Future research could explore federated learning and multi-modal imaging for broader applicability. The study also emphasized the importance of balancing segmentation precision with inference speed for clinical deployment. Incorporating lightweight architectures may help reduce computational overhead in real-time

settings. Future frameworks could benefit from scalable validation across diverse imaging modalities and healthcare environments [43].

Mirikharaji et al. (2023) reviewed deep learning methods for skin lesion segmentation, analyzing 177 studies from 2014 to 2022 and categorizing various segmentation approaches. They tested loss functions, performance measures, and preprocessing methods and confirmed the effectiveness of deep learning in reducing segmentation errors. Challenges in computational efficiency highlighted the need for further optimization, with future directions exploring spatial constraints and multimodal imaging for improved segmentation accuracy [44].

Basak et al. (2022) created MFSNet, an MFSNets that optimizes lesion segmentation by integrating Res2Net, demonstrating superior performance in Dice and Jaccard scores compared to existing models. They would integrate boundary attention and reverse attention modules that enhanced accuracy of segmentation. Nevertheless, there were issues with computational efficiency. Future efforts may investigate federated learning (FL) and hybrid architectures for enhanced dermatological AI applications. Additionally, optimizing model compression and pruning techniques could help reduce inference time without compromising segmentation quality. Incorporating lightweight attention mechanisms may further balance precision with scalability for real-world deployment [45].

Khan et al. (2019) developed a deep learning framework combining optimized color features (OCF) with DCNNs for skin lesion segmentation and classification, achieving over 90 percent segmentation accuracy and classification accuracies of 92.1 percent, 96.5 percent, and 85.1 percent across ISBI datasets. Their hybrid pipeline enhanced lesion boundary demarcation and feature extraction and was vulnerable to computation issues. Future research may explore multi-modal imaging and federated learning for broader applicability [46].

Azeem et al. (2023) introduced SkinLesNet, a lightweight CNN optimized for smartphone-based melanoma detection, leveraging PAD-UFES20 with data augmentation to enhance generalization, as shown in Figure 2.6. Their model recorded an accuracy of 96.2 percent, a sensitivity of 94.8 percent, and a specificity of 97.5

percent with an AUC of 98.1 percent, better than other models. Although the SkinLesNet provides improved generalization, classifications can be better clarified through the mechanism of attention. Future studies may explore hybrid architectures for improved lesion differentiation. The study also emphasized the potential of integrating attention mechanisms to refine feature localization and interpretability. Incorporating real-time inference capabilities may further support mobile deployment in low-resource clinical settings [47].

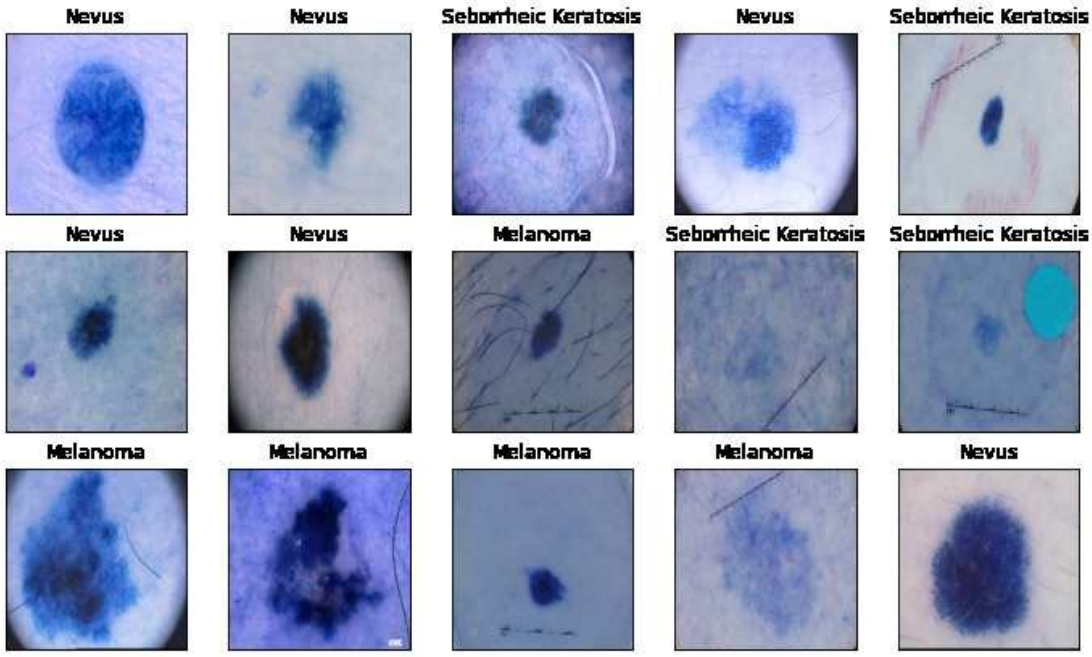


FIGURE 2.6: Sample images from the PAD-UFES-20 dataset used in this research with different skin lesions visualizations ref by [47]

Almufareh et al. (2024) fine-tuned a CNN model for melanoma classification, incorporating hyperparameter optimization and transfer learning across DermIS, DermQuest, and ISIC2019 datasets. Their approach improved detection to 5 percent accuracy on (DermIS) data, 6 percent on (DermQuest) data, and 81 percent on (ISIC2019) data, highlighting AI's role in early diagnosis. Computational efficiency challenges remain, necessitating future research on multi-modal imaging and IoMT integration for real-time dermatological evaluations. The study also emphasized the need for scalable architectures to accommodate diverse dataset complexities. Incorporating edge computing could help mitigate latency in real-time diagnostic workflows [48].

Ghosh et al. (2022) proposed SkinNet-16, a deep learning (DL) classifier for timely skin lesion detection, incorporating feature extraction and dimensionality reduction via PCA to enhance classification performance. Their model performed well in precision and recall: it has an accuracy of 99.19 percent when using the Adamax optimizer trained on Kaggle data. Challenges in optimizing real-time deployment remain, prompting future investigations into multi-modal imaging and hybrid architectures for improved accessibility. The study also emphasized the need for lightweight deployment strategies to support mobile and point-of-care diagnostics. Incorporating attention mechanisms may further refine lesion localization and improve interpretability. Future models could benefit from scalable validation across diverse skin types and imaging conditions [49].

2.2 Evaluation Standards and Benchmark Datasets in Dermatology

Hameed et al. (2020) introduced the Multi-Class Multi-Level (MCML) classification algorithm, enhancing diagnostic accuracy in dermatological image classification by combining machine learning and deep learning techniques. Their divide and conquer approach enhanced categorization of lesions with an accuracy level of 96.47 percent, and the false positive occurred less. Although it is very effective, computational efficiency is a problem. They also emphasized the potential of hierarchical classification to manage complex lesion types more effectively. Moreover, the study highlighted the need for lightweight models and optimized pipelines to support deployment in resource-constrained clinical settings. Future research should explore multimodal imaging and federated learning for real-time applicability. In addition, the study suggested that integrating attention mechanisms could further refine feature localization and reduce ambiguity in lesion boundaries. Exploring model pruning and quantization may also help balance diagnostic precision with computational feasibility for scalable deployment. These optimization techniques can also facilitate real-time inference on low-power devices, broadening access to AI-assisted dermatological screening [50].

Rasool et al. (2023) conducted an analysis in comparison to deep learning models to identify skin diseases, assessing CNN architectures like DenseNet, ResNet, and EfficientNet across ISIC Archive and DermNet datasets. Their results confirmed the efficacy of deep learning in dermatological assessment but indicated the data bias limitation. Optimization for real-time applications and federated learning may improve fairness and scalability in future AI dermatology research. They also emphasized the importance of diverse and representative datasets to mitigate bias and improve generalization across skin types. Furthermore, the study highlighted the potential of lightweight CNN architectures for deployment in mobile and point-of-care diagnostic systems [51].

Shakya et al. (2025) improved EfficientNet's performance on the ISIC and HAM-10000 datasets through data augmentation and focal loss optimization, increasing classification accuracy from 88 to 93 percent. Their method increased the sensitivity of detecting rare malignant lesions, yet problems existed regarding the balance between synthetic and real images. They also emphasized the importance of maintaining clinical realism in augmented datasets to avoid overfitting and misclassification. Furthermore, their findings support the integration of dynamic sampling and domain adaptation techniques to enhance generalizability across diverse lesion types and imaging conditions. Future investigations may refine augmentation strategies and adaptive weighting techniques to boost classification accuracy [52].

Nancy et al. (2023) applied synthetic image generation to improve dataset diversity in dermatological classification, as mentioned in Figure 2.7, increasing Vision Transformer (ViT) accuracy from 86 to 91 percent. They improved generalization to rare lesions but faced challenges in synthetic image authenticity. In the future, studies should concentrate on refining GANs and diffusion models for realistic synthetic images and multimodal fusion techniques to improve classification robustness [53]. Their approach also demonstrated the potential of synthetic data to address class imbalance in dermatological datasets. Moreover, the study emphasized the need for rigorous validation of generated images to ensure diagnostic integrity and clinical applicability. The study highlighted transformers strength in capturing global context for better lesion classification.

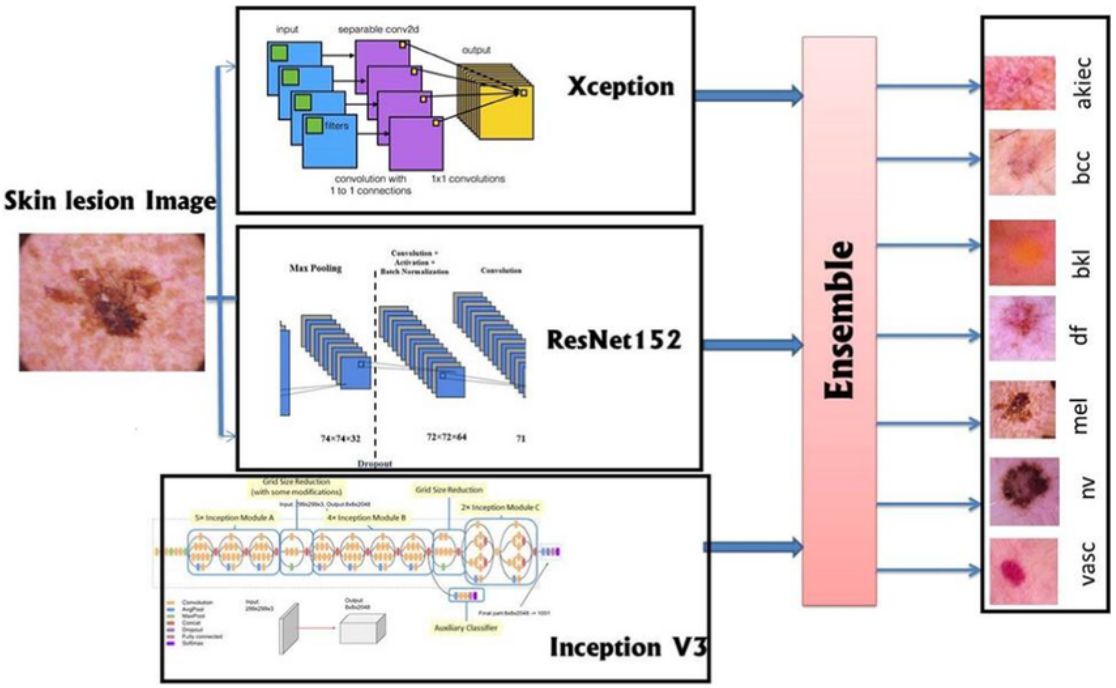


FIGURE 2.7: The ensemble model (Xception, ResNet152, and inceptionv3) ref to [53]

Musthafa et al. (2024) utilized data augmentation, including rotation, scaling, and flipping, to improve CNN accuracy from 93 to 97.78 percent on the HAM10000 dataset. Their approach enhanced the classification of underrepresented classes of lesions and minimized the misclassification rate, but they cannot consistently augment their results. Future studies should optimize augmentation techniques and explore hybrid learning methods for enhanced clinical accuracy [54]. The study also emphasized the importance of maintaining clinical realism during augmentation to avoid overfitting and diagnostic distortion. Moreover, their findings suggest that combining traditional augmentation with adaptive sampling strategies may improve robustness across diverse lesion types. Integration of lightweight architectures and real-time inference capabilities could further support deployment in point-of-care settings. The study also emphasized the potential of parallel processing and model pruning to address computational bottlenecks. Incorporating explainable AI could further enhance clinical trust and interpretability in diagnostic workflows.

Abhishek et al. (2025) proposed synthetic data augmentation to reduce biases in skin lesion classification, particularly within the DermaMNIST and Fitzpatrick17k datasets. Their representation enhanced neglected skin types using their methodology based on GANs to support fair AI-aided diagnostics. Future work should refine fairness-conscious training protocols and integrate multimodal fusion techniques for broader dermatological AI applications [55]. They also emphasized the importance of inclusive dataset representation to mitigate disparities in diagnostic performance across skin tones. Moreover, the study highlighted the potential of synthetic augmentation to balance underrepresented classes without compromising diagnostic integrity, supporting the integration of generative models with fairness-aware objectives to improve equity in clinical outcomes.

Tschandl et al. (2018) introduced the HAM10000 dataset, a benchmark collection of 10,015 dermatoscopic images spanning 7 well-known skin lesion types, including MEL and BCC lesions. The dataset consisting of several sources gives a varying representation of lesions, with challenges such as balanced classes and inconsistencies in image acquisition. Future research should refine augmentation strategies and explore multi-modal fusion to enhance diagnostic reliability in AI-assisted dermatology [56].

Akram et al. (2023) integrated a mask region-based convolutional neural network for segmentation and ResNet50 for classification in a hybrid deep learning approach, attaining 95.49 percent segmentation accuracy and 96.75 percent classification accuracy on the ISIC 2020 Challenge dataset. They have improved the quality of lesion boundaries and the robustness of classification, and the computational efficiency is challenging. Future studies could explore federated learning and multi-modal imaging for broader applicability [57].

Yuan et al. (2022) introduced EdgeMixup, a preprocessing technique designed to improve fairness in skin disease classification and segmentation by reducing performance disparities between light and dark skin tones. Their approach reduced the accuracy disparities by 10.99 percentage points, which increased the sensitivity of the less privileged groups. Although promising, there are still real-time deployment challenges that still need optimization. Future research may focus

on adaptive learning and fairness-aware AI strategies to improve dermatological diagnostics [58].

2.3 Critical Review of Supervised and Unsupervised Learning Approaches in Dermatological Artificial Intelligence

The presented dermatological AI developments reveal the transformative power of data augmentation, synthetic image generation, and hybrid networks to solve all-important problems of class imbalance, practical computational experience, and fair performance between different demographics. The integration of generative adversarial networks (GANs) and advanced preprocessing techniques has paved the way for innovative solutions, ensuring that supervised learning models can better generalize while maintaining high accuracy in clinical applications. Supervised learning models require labeled datasets, and consequently, they are adequate in problems that involve classification or segmentation as depicted in Table **??**. This table shows research on skin lesion classification and segmentation using supervised learning techniques, including their major contributions, approaches, performance, and challenges to AI-powered dermatology.

These advancements also underscore the importance of robust annotation practices to support reliable supervised learning outcomes. The use of high-quality labeled data enables precise lesion boundary detection and accurate categorization across diverse skin conditions. Moreover, the integration of domain-specific knowledge into supervised pipelines enhances interpretability and clinical trust. As dermatological AI continues to evolve, supervised frameworks remain central to achieving scalable, transparent, and clinically validated diagnostic systems. These developments highlight supervised learning as a cornerstone for effective, transparent, and generalizable dermatological AI applications. The emergence of hybrid methods whose approaches use a mixture of supervised and unsupervised techniques has even been supported by recent studies.

TABLE 2.1: Summary of Supervised Learning Approaches in Dermatological AI

Ref.	Year	Dataset	Method	Techniques and Tools	Remarks	Challenges
[9]	2025	Custom Dataset (19,171 images)	Transfer learning (DenseNet121 + EfficientNetB0)	Image normalization, contrast enhancement	Accuracy: 98.18%, Sensitivity: 95.3%	Computational complexity, bias in skin tones
[10]	2025	Large Dermatological Dataset	Prototypical Decision Tree (PDT)	Rotation, flipping, brightness adjustments	Accuracy: 92%, AUC-ROC: 0.91	Rare lesion classification, dataset balance
[13]	2019	PH2, ISBI 2016, ISBI 2017	CNN-based feature fusion	HSV transformation, Laplacian filtering	PH2: 98.4%, ISBI 2016: 95.1%	Computational efficiency, dataset imbalance
[22]	2025	Severity Dataset (10,000 images)	Lesion-Aware Transformer (LAT)	Contrast optimization, zoom techniques	Accuracy: 91%, Severity Scoring Precision: 0.89	Complexity in lesion patterns, image diversity

Continue on next page

Ref.	Year	Dataset	Method	Techniques and Tools	Remarks	Challenges
[50]	2020	3,672 images	Hybrid ML-DL model	Noise removal, lesion segmentation	Accuracy: 96.47%, Sensitivity: 94.3%	Scalability, dataset diversity
[51]	2023	ISIC Archive, DermNet	CNN-based classification (ResNet, DenseNet, EfficientNet)	Deep learning architectures	Accuracy: 95.2%, Dice coefficient: 0.87	Generalization, fairness concerns
[54]	2024	HAM10000 (10,000 images)	CNN-based classification with augmentation	Rotation, scaling, flipping	Accuracy: 97.78%, Dice coefficient: 0.89	Augmentation consistency, adaptability
[55]	2025	DermaMNIST, Fitzpatrick17k	GAN-based synthetic data augmentation	Bias mitigation, synthetic image generation	Accuracy: 94%, AUC-ROC: 0.93	Synthetic image authenticity, fairness concerns

TABLE 2.2: Summary of Unsupervised Learning Approaches in Dermatological AI

Ref.	Year	Dataset	Method	Techniques and Tools	Remarks	Challenges
[11]	2023	Fitzpatrick17k	Adaptive augmentation	Synthetic image generation, contrast alterations	Accuracy: 89.5%, F1-score: 0.92	Dataset imbalance, fairness concerns
[14]	2018	PH2	Hybrid segmentation (Boltzmann entropy-based feature selection)	DullRazor preprocessing, HOG descriptors	Accuracy: 97.5%, Sensitivity: 97.7%	Feature stability, segmentation accuracy
[16]	2023	10,000 images	Generative Adversarial Networks (GANs)	Synthetic data generation, augmentation	Accuracy: 92%, AUC-ROC: 0.89	Synthetic image authenticity, model robustness
[17]	2024	Multi-modal dataset (15,000 images)	Active learning-based augmentation	ResNet, contrast optimization	Accuracy: 91%	Fairness concerns, dataset generalization

Continue on next page

Ref.	Year	Dataset	Method	Techniques and Tools	Remarks	Challenges
[34]	2023	Dermatological Dataset (5,000 images)	Unsupervised synthetic image generation	GAN-based augmentation, contrast enhancement	Accuracy: 93%, AUC-ROC: 0.91	Synthetic image authenticity, demographic fairness
[40]	2025	Dermatological images (varied dataset)	Visual Selective State Spatial Model (VSSSM)	DeepLabV3, Mask R-CNN	Dice coefficient: 0.93, Jaccard index: 0.88	Scalability, computational efficiency
[41]	2024	ISIC + Custom Dataset	Ant Colony Optimization (ACO)	ResUNet, ACO boundary refinement	Dice coefficient: 0.93, Jaccard index: 0.87	Generalization, computational complexity
[42]	2024	ISIC Dataset	Hybrid ensemble segmentation	U-Net, DeepLabV3	Dice coefficient: 0.93, IoU: 0.90	Class imbalance, segmentation errors

Such hybridization may offer a pathway to reduce annotation burden while preserving diagnostic fidelity. These methods can also uncover latent structures in unlabeled data, enriching feature representations and improving downstream classification performance. As the field progresses, integrating semi-supervised and self-supervised strategies may further enhance scalability and reduce reliance on costly expert annotations.

These approaches make use of the virtues of annotated datasets and update them in an unsupervised manner using unsupervised procedures like clustering and trial representation learning. For instance, novel frameworks are beginning to incorporate Generative Adversarial Networks (GANs) to synthesize realistic examples for augmenting training datasets, effectively addressing gaps in diversity and complexity. Unsupervised learning algorithms do not rely on tagged information but seek patterns using clustering, augmentation, and feature selection, which appear in Table **??**. This table shows literature about works that use unsupervised learning techniques to analyze skin lesions with key algorithms, datasets, performance measures, and issues related to dermatology AI.

These hybrid strategies offer a promising pathway for enhancing model adaptability in scenarios with limited labeled data. They also facilitate continuous learning by leveraging unlabeled inputs to refine feature representations over time. Recent advancements have shown that combining self-supervised learning with GAN-based augmentation can significantly improve classification robustness. Moreover, such methods contribute to reducing annotation costs while maintaining diagnostic accuracy. As dermatological AI evolves, hybrid frameworks are expected to play a pivotal role in bridging the gap between data scarcity and clinical reliability.

Finally, the study examined in this chapter demonstrates remarkable progress in the use of deep learning approaches to the detection of skin lesions, particularly because of the introduction of data augmentation, transfer learning, and advanced classification models. However, problems such as unbalanced data, the quality of images, and lack of generalization among different populations persist. Such results point to the need to use a more robust and adaptable model, and this precondition predetermines the method presented in the rest of the chapter of this work.

Chapter 3

Proposed Methodology

An effective and structured approach is needed to develop a deep learning framework that can be used to recognize lesions on the skin through the transfer of learning. As shown in Figure 3.32, the proposed solution is a hybrid system with two state-of-the-art CNNets, like DenseNet121 and EfficientNetB0, integrated. This chapter describes the entire flow of the methodology pipeline, including the steps of selecting the data, training the model, and evaluating its performance on the HAM10000 dataset, letting the reader have a general look at the mechanisms used to achieve the best results possible.

The methodology considers a set of clear stages, such as dataset preparation, preprocessing, contrast enhancement, data augmentation, model architecture development, training, and validation. Such stages are thoughtfully developed to overcome the identified difficulties in dermatoscopic image analysis that include class imbalance, image variability, and feature extraction complexity. The hybrid approach takes advantage of transfer learning to generalize well with different types of lesions and achieves promising gains compared to the conventional methods. Below is a description of the steps and design choices that informed us of the implementation of such a framework. The framework also incorporates iterative refinement, allowing adjustments based on validation feedback. Its layered design ensures that each component, from preprocessing to model tuning, contributes to

overall robustness. By aligning each stage with clinical imaging challenges, the methodology promotes reliable lesion classification outcomes.

3.1 Proposed Methodology

A systematic and well-organized approach is required to create a robust deep learning model of skin disease prediction, as illustrated in Figure 3.32. This would enable accurate detection of dermatoscopic images associated with a variety of lesion types. To address the inherent challenges of diagnosing skin diseases, our hybrid approach combines transfer learning approaches with the sophisticated feature extraction capabilities of DenseNet121 and EfficientNetB0. The methodology encompasses a comprehensive pipeline, including data collection, preprocessing, image contrast enhancement, data augmentation, model architecture design, training, testing, and rigorous evaluation. Each stage is strategically designed to mitigate challenges such as class imbalance, image variability, and the need for precise feature differentiation, particularly when applied to the open-source HAM10000 dataset, which forms the cornerstone of this study. The block diagram of our proposed methodology is shown in Figure 3.1.

This framework emphasizes modularity and scalability, allowing researchers to adapt each stage to evolving datasets and clinical requirements. The preprocessing phase incorporates advanced techniques like CLAHE and LAB color normalization to enhance lesion visibility and reduce noise. Data augmentation is performed using Imgaug to simulate real-world variations and address class imbalance. The hybrid model architecture integrates DenseNet121's deep feature learning with EfficientNetB0's computational efficiency, striking a balance between accuracy and speed. Transfer learning accelerates convergence and improves generalization, especially on limited annotated data. Training is optimized using early stopping and adaptive learning rates to prevent overfitting. Evaluation metrics such as accuracy, precision, recall, and F1-score are used to assess model performance across lesion categories. Cross-validation ensures robustness and consistency across multiple data splits. The methodology supports integration with clinical decision support

systems, enhancing its real-world applicability. Overall, this structured pipeline lays a strong foundation for reliable, interpretable, and scalable skin disease classification

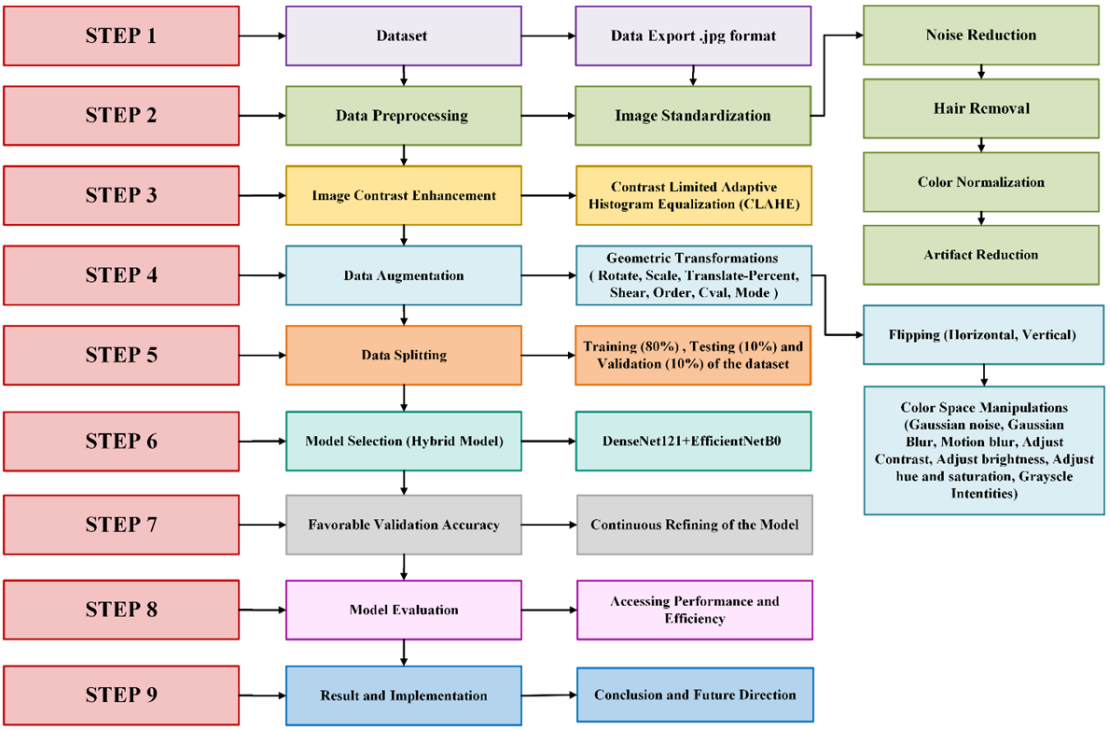


FIGURE 3.1: Block Diagram of the Proposed Methodology

3.2 Dataset Description

Our proposed model is trained and validated using the HAM10000 dataset (Human Against Machine with 10,015 dermatoscopic images), publicly available on Kaggle. This dataset includes high-resolution dermatoscopic pictures that indicate seven categories of skin lesions with pigmentation gathered from varied clinical environments. The pictures are in JPEG format, with a standard dimension resolution of 600 x 450 pixels, and vary in angle, lighting, and pigment. These images are labeled by expert dermatologists, ensuring high-quality annotations for supervised learning tasks. The dataset includes metadata such as lesion localization and patient demographics, which can support auxiliary analyses. Its comprehensive nature makes it a benchmark resource for developing and evaluating skin lesion classification models.

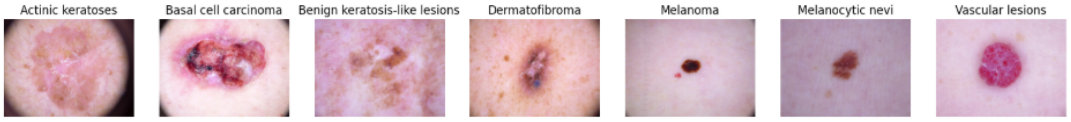


FIGURE 3.2: Distribution of Skin Lesions in the Seven Classes

Table 3.1 also gives the distribution of the seven classes of lesions. The dataset is inherently imbalanced, with a large proportion of images representing melanocytic nevi (NV), while others like dermatofibroma (DF) and vascular lesions (VASC) are underrepresented. This imbalance can skew model predictions toward dominant classes, reducing sensitivity for rare but clinically significant lesions. To counteract this, targeted augmentation and oversampling techniques are applied to boost minority class representation. The use of stratified sampling during training ensures that each class contributes proportionally to model learning.

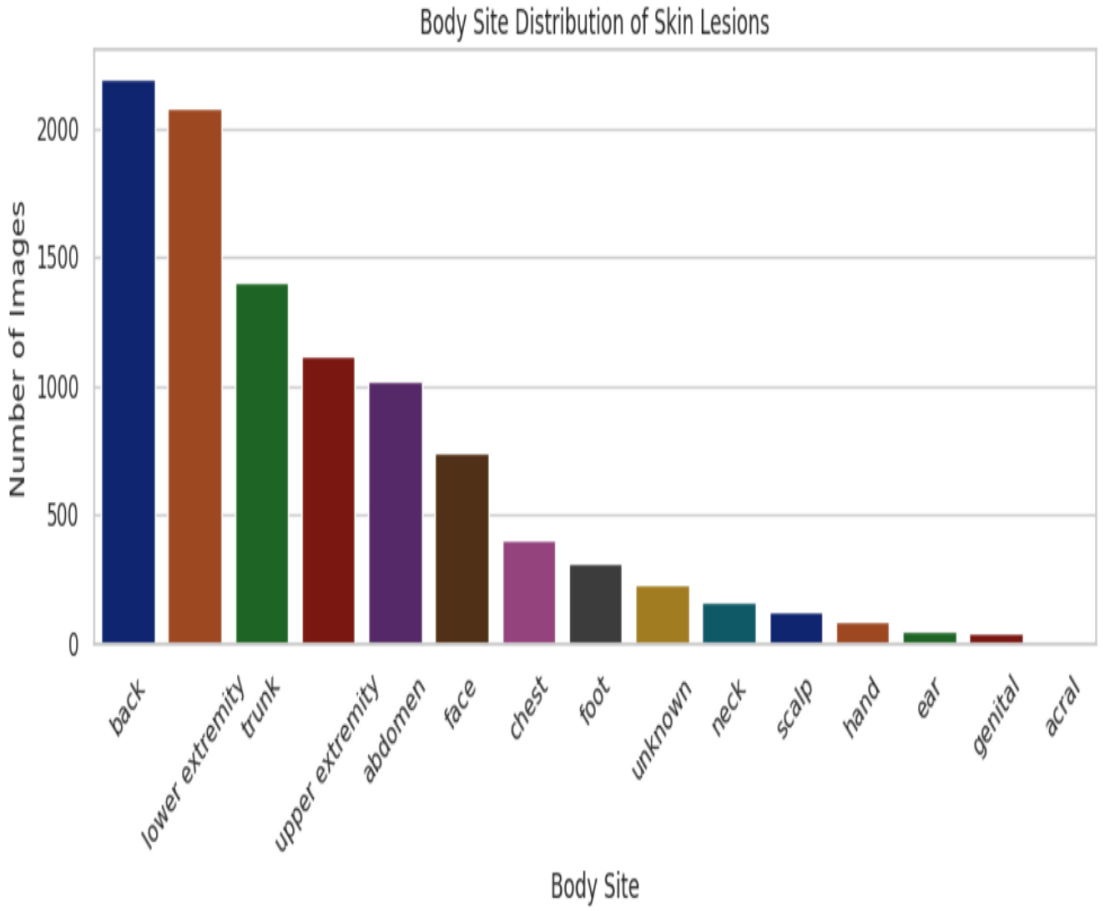


FIGURE 3.3: Gender-wise skin Lesion Distribution

TABLE 3.1: Distribution of seven classes of skin lesions

Class ID	**Skin Lesion Type**	**Number of Images**	**Description**
0	Actinic keratoses and intraepithelial carcinoma (AKIEC)	327	Precancerous lesions
1	Basal cell carcinoma (BCC)	514	Common type of skin cancer
2	Benign keratosis-like lesions (BKL)	1099	Non-cancerous growths
3	Dermatofibroma (DF)	115	Benign skin tumors
4	Melanocytic nevi (NV)	6705	Common benign moles
5	Melanoma (MEL)	1113	Dangerous malignant tumors
6	Vascularsions (VASC)	142	Blood vessel-related lesions

Other visual analysis indicates differences in distribution on body sites and gender-based distributions as indicated in Figure 3.3 and Figure 3.4 respectively. These observations emphasize the validity of the diagnostic variety observed in the records.

3.3 Data Preprocessing

The HAM10000 dataset has various inherent challenges on dermatoscopic images, namely varied intensities, different contrast, intensity of background noise, and

complexity of the structure. All these aspects may significantly hinder the performance of deep learning models because they cover important features of lesions. In order to address these constraints, the systematically administered, robust, multi-step pipeline of preprocessing is developed and streamlined across all the images in the dataset. The purpose of this pipeline is to increase clarity and maintain diagnostically important features of the images and demonstrate consistency in different input data, hence making the model better at classifying a range of skin lesions.

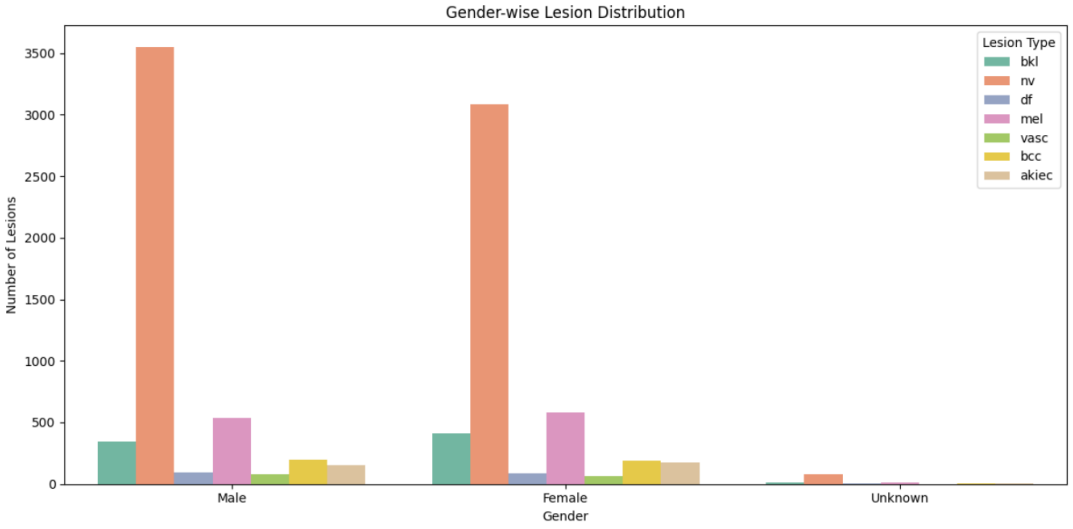


FIGURE 3.4: Gender-wise skin Lesion Distribution

Preprocessing steps are important in order to reduce noise, fix lights, and remove visual distractions such as body hair, especially those that could mislead the model when being trained. All the images are processed through a number of transformations in order to produce a processed image as shown in Figure 3.10, such as Gaussian and median filtering, hair removal using Dull Razor, and intensity normalization. These steps directly address the preprocessing gaps highlighted. Throughout the literature review Chapter 2, specifically the lack of robust strategies for managing variability in dermatoscopic images [10][19]. The preprocessing pipeline greatly reinforced the proposed model through standardization of image quality and reduction of irrelevant differences in their performance to extract consistent and generalizable features in the different categories of lesions. It also improves training stability and transfer learning compatibility.

3.3.1 Image Standardization

To ensure uniformity throughout the HAM10000 dataset, the pipeline preprocessing starts with standardization of the images. The images are downsized to 128x128 pixels without shrinking the aspect ratio to avoid distortion and computation optimization. This resizing ensures compatibility with the hybrid architecture and reduces computational overhead. Each image is subsequently normalized to the [0, 1] range and standardized using Z-score normalization as displayed in Figure 3.7. To maintain consistent input dimensions and pixel value distributions, image standardization involved resizing, normalization, and color format consistency.

3.3.1.1 Resizing

In Equation 3.1, each image $I \in \mathbb{R}^{H_o \times W_o \times 3}$ with original height H_o and widthW_o, is resized to a target size of $(H, W) = (75, 100)$ pixels while preserving the aspect ratio through zero-padding or cropping, as displayed in Figure 3.5.

$$I_r = \text{Resize}(I, (H, W)) \tag{3.1}$$

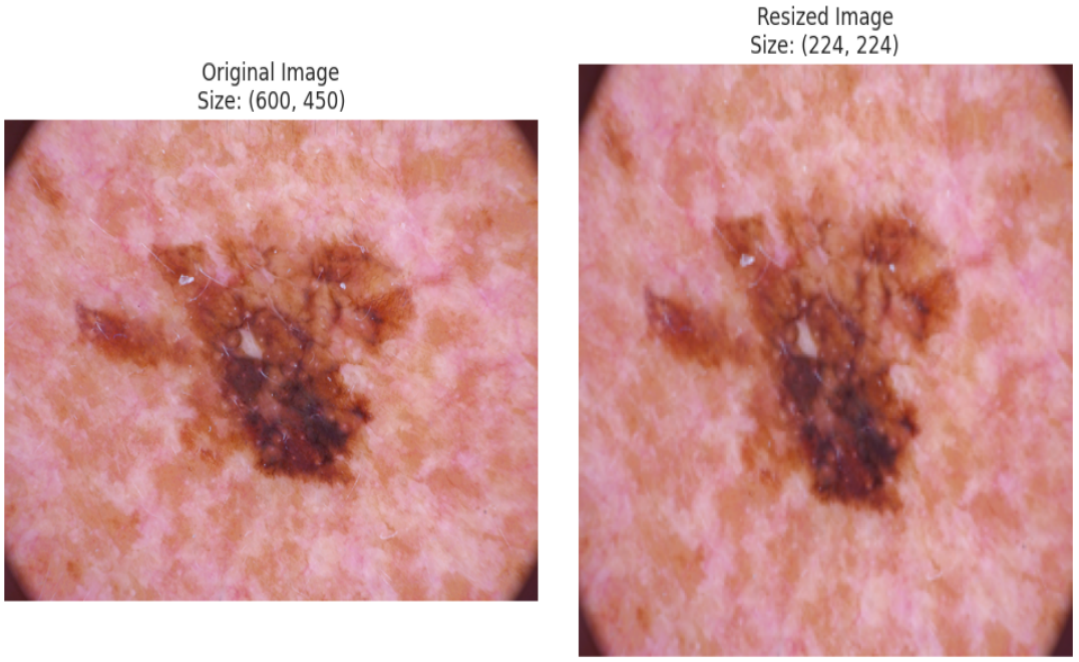


FIGURE 3.5: Apply Resizing Technique on Lesion Image

Resizing ensures uniform input dimensions for consistent feature extraction. Aspect ratio is preserved using zero-padding or cropping to avoid lesion distortion. The resized image I_r standardizes inputs for preprocessing and augmentation, improving computational efficiency and batch compatibility.

3.3.1.2 Normalization

Images are converted to RGB format as displayed in Figure 3.6, preserving consistent color variations for compatibility with pretrained models like DenseNet121 and EfficientNetB0. Additionally, lesion classes undergo one-hot encoding, facilitating seamless multi-class classification across seven distinct disease categories. Pixel intensities $I_r(x, y, c)$ in each channel $c \in \{R, G, B\}$ are scaled from 8-bit unsigned integers $[0, 255]$ to floating-point values $[0, 1]$ as shown in Equation 3.2 :

$$I_n(x, y, c) = \frac{I_r(x, y, c)}{255} \tag{3.2}$$

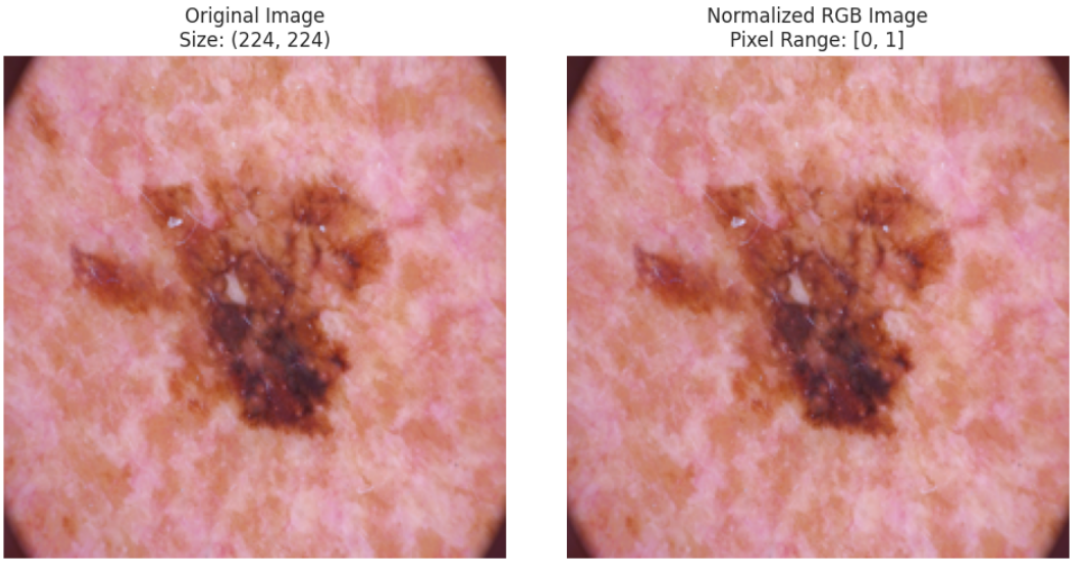


FIGURE 3.6: Apply the Normalization Technique on the lesion image

This normalization ensures uniform pixel value distribution, which accelerates convergence during training. It also reduces sensitivity to illumination differences across dermatoscopic images. The scaled values are fed into the model to maintain

compatibility with activation functions like ReLU. Such preprocessing steps are essential for stabilizing gradients and improving model generalization. They also help mitigate the impact of acquisition artifacts and device-specific variations, ensuring consistent performance across datasets. Furthermore, standardized inputs enable more reliable feature extraction, which is critical for distinguishing subtle lesion characteristics in clinical settings.

3.3.1.3 Standardization (Z-Score Normalization)

To further normalize intensity distributions across the dataset and reduce bias from brightness variations as displayed in Figure 3.7, the standardized pixel intensity I_s is calculated as shown in Equation 3.3, 3.4 and 3.5:

$$I_s(x, y, c) = \frac{I_s(x, y, c) - \mu_c}{\sigma_c} \tag{3.3}$$

$$\mu_c = \frac{1}{N} \sum_{i=1}^{N} I_n(i, c) \tag{3.4}$$

$$\sigma_c = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(I_n(i, c) - \mu_c\right)^2} \tag{3.5}$$

where $N = H \times W$ denotes the total number of pixels per channel and i indexes the pixels. This step ensures zero-mean and unit-variance intensity distributions, improving model convergence.

- $I \in \mathbb{R}$ represents the original pixel intensity,
- $\mu = \frac{1}{N} \sum_{i=1}^{N} I_i$ is the mean pixel intensity across the dataset,
- $\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (I_i - \mu)^2}$ is the standard deviation of pixel intensities,
- I' is the standardized pixel intensity,

- N is the total number of pixels in the dataset,
- I_i denotes the intensity of the i-th pixel.

This normalization technique minimizes the influence of lighting inconsistencies and enhances feature stability across samples. It also ensures compatibility with activation functions and optimizers commonly used in deep learning architectures.

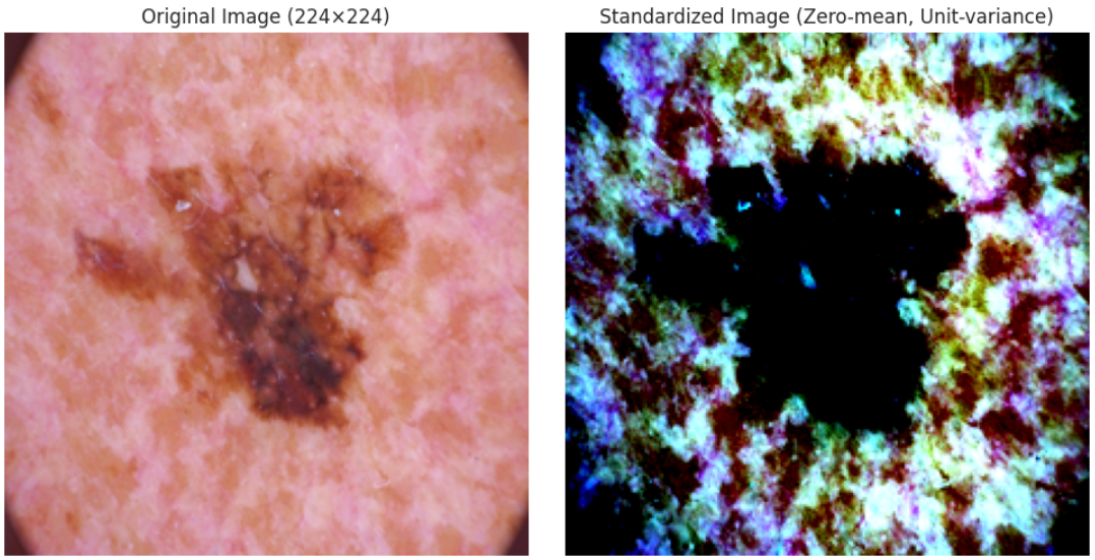


FIGURE 3.7: Apply Standardization Technique on Lesion Image

3.3.2 Noise Reduction

Each input image, after standardization, is subjected to a two-step noise reduction process designed to eliminate high-frequency artifacts and irrelevant details while preserving lesion boundaries and structures critical for classification. Noise from image acquisition and skin texture can impede feature learning. To reduce noise N while preserving lesion edges, a composite filter is applied.

3.3.2.1 Gaussian Filtering

Each input image initially passes through a Gaussian filter to smooth out high-frequency noise by convolving the image with a Gaussian kernel. This operation helps suppress irrelevant details while retaining the core lesion structure essential

for accurate classification. By reducing texture variations and enhancing dominant features, the filter improves the consistency of input data. This step is particularly useful in handling dermatoscopic images affected by uneven lighting or skin texture irregularities, ensuring better downstream performance in feature extraction and model training in [67]. Mathematically, the Gaussian filter is defined in Equation 3.6:

$$I_g = I_s * G_\sigma, \quad G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{-x^2+y^2}{2\sigma^2}} \tag{3.6}$$

Where σ is the standard deviation that controls the degree of Smoothing. In this study, a 5×5 kernel is used, $\sigma = 0$ which lets OpenCV compute it automatically.

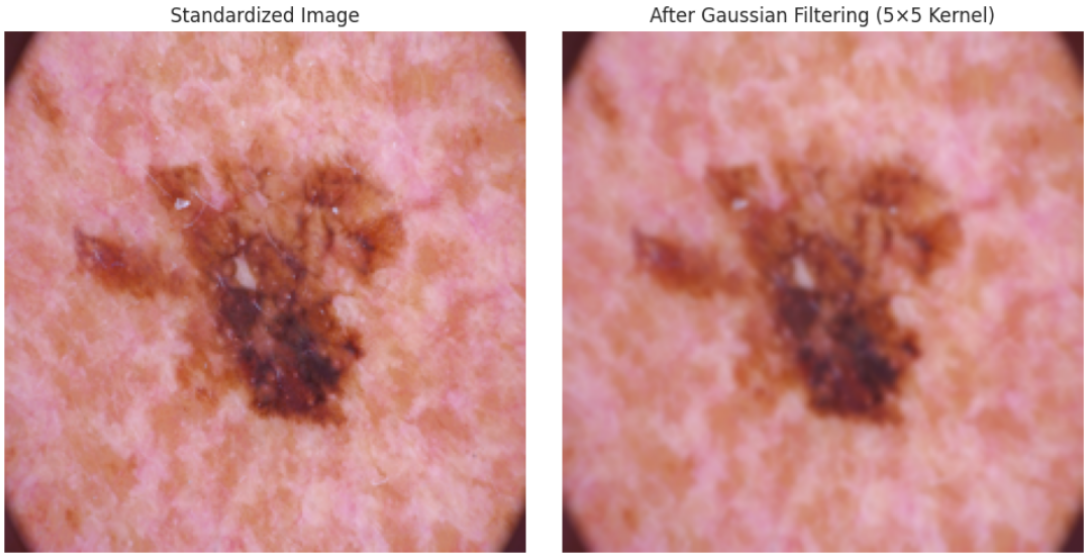


FIGURE 3.8: Applying Gaussian Filtering on Lesion image

3.3.2.2 Median Filtering

After Gaussian smoothing, a median filter further reduces salt-and-pepper noise. This non-linear filter substitutes the neighborhood median for each pixel value, preserving edges more effectively than Gaussian blurring. This filtering removes impulsive noise by replacing each pixel with the median of its neighborhood, as displayed in Figure 3.9. Let $I(x, y)$ the intensity at the location be (x, y), as shown in Equation 3.7:

$$I_{nr}(x, y) = \text{median}\left(\{I_g(i, j) \mid (i, j) \in \Omega(x, y)\}\right) \tag{3.7}$$

Where $\Omega(x, y)$ is a local neighborhood (in this case 5×5) around pixel (x, y). This combination effectively suppresses Gaussian and salt-and-pepper noise without blurring lesion boundaries.

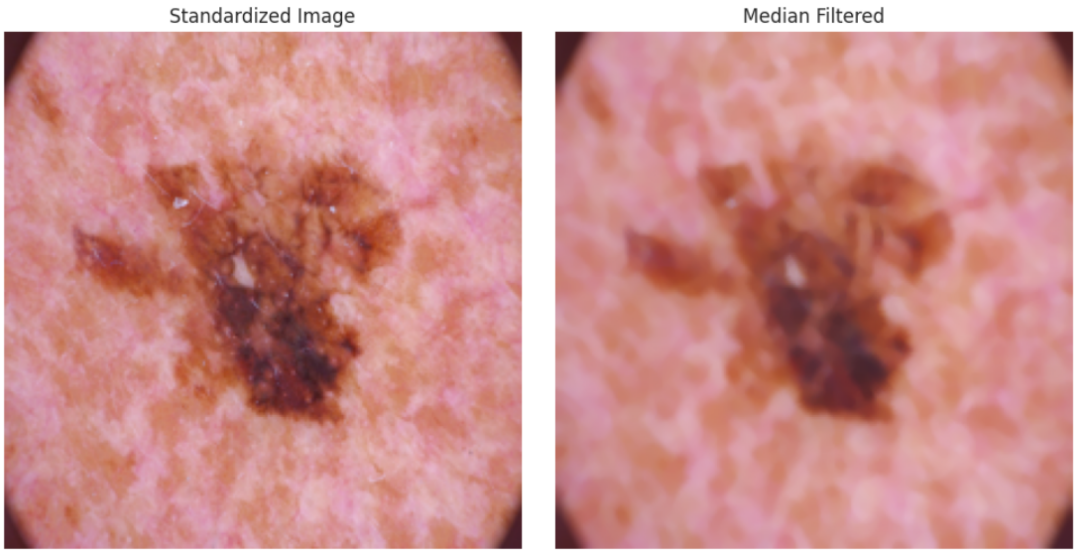


FIGURE 3.9: Applying Median Filtering on Lesion Image

3.3.3 Hair Artifact Removal (Dull-Razor Technique)

The presence of hair in dermatoscopic images can obstruct lesion patterns. The Dull Razor algorithm is used to detect and remove hair artifacts mentioned in Figure 3.15 and Figure 3.11. This method includes converting a noise-reducing image into grayscale, morphological operations (Blackhat filtering), and masking. This preprocessing step enhances lesion visibility and ensures that hair artifacts do not interfere with feature extraction or classification accuracy. By restoring occluded regions and preserving lesion boundaries, it contributes to cleaner inputs for downstream model training.

By eliminating hair artifacts, the Dull Razor algorithm enhances lesion visibility and ensures cleaner input for feature extraction. This preprocessing step is crucial for maintaining diagnostic integrity, especially in automated classification

tasks. The algorithm's use of morphological filtering effectively isolates thin, dark strands without compromising surrounding skin texture. Masking and inpainting techniques then restore occluded regions, producing a refined image suitable for downstream analysis.

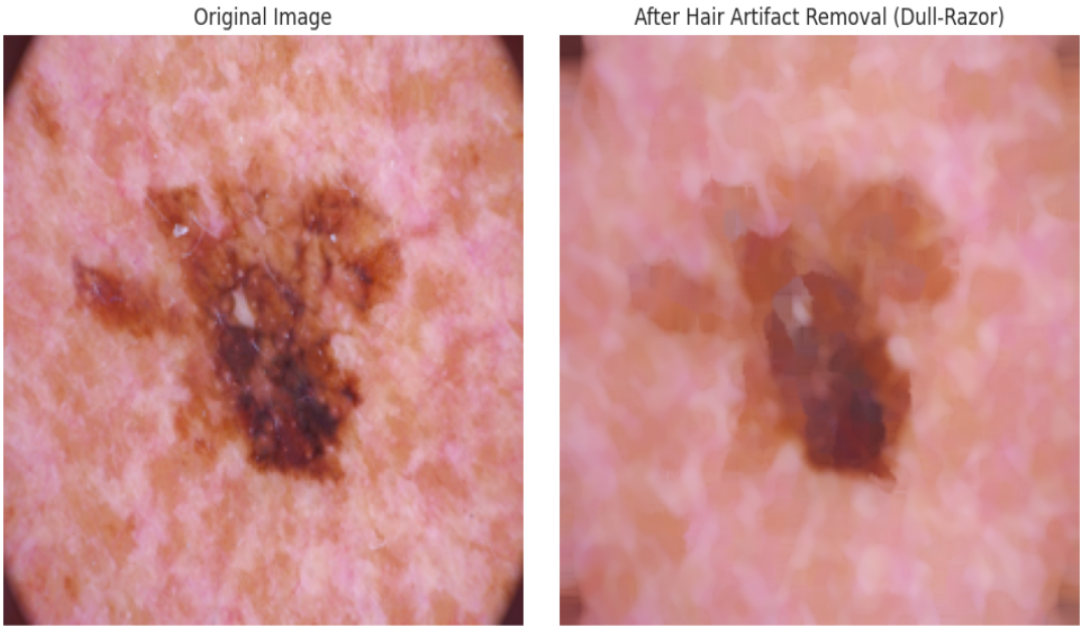


FIGURE 3.10: Apply the Dull-Razor Technique for Hair Artifact Removal on Lesion Image

Overall, this approach improves model accuracy by reducing noise and preserving critical lesion details. It further standardizes image quality across the dataset, enabling more reliable training and evaluation. Such preprocessing is especially beneficial when dealing with large-scale dermatoscopic repositories containing varied image artifacts. Moreover, it supports consistent feature learning by minimizing distractions and ensuring that lesion boundaries remain intact throughout the pipeline.

$$I_{\text{gray}} = \text{rgb2gray}(I_{nr}) \tag{3.8}$$

By using Equation 3.8, the rgb2gray function converts the noise-reduced RGB image I_{nr} to grayscale I_{gray}. This step simplifies the data by reducing three color channels to one intensity channel, making subsequent processing more efficient. It preserves essential structural and luminance information needed for accurate lesion analysis while eliminating color-based variations that are less relevant in

medical imaging. The grayscale image serves as a standardized input for enhancement techniques such as histogram equalization and CLAHE. It also improves the performance of segmentation algorithms by emphasizing intensity-based boundaries. Overall, this conversion enhances the robustness and reproducibility of the preprocessing pipeline.

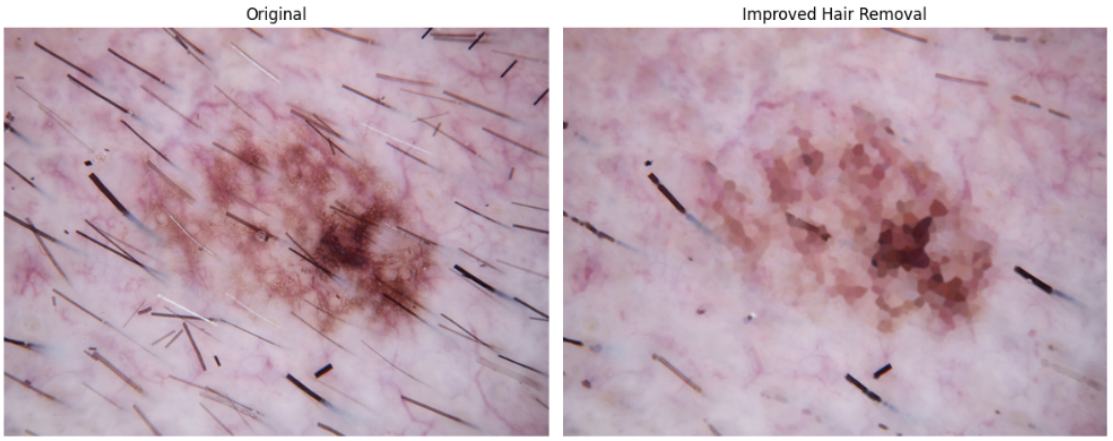


FIGURE 3.11: Apply Hair removal Technique

This transformation simplifies the image by reducing it to a single intensity channel, which is optimal for morphological operations such as Blackhat filtering. It also enhances computational efficiency and preserves structural details necessary for accurate hair detection and removal. Applying morphological operations (Blackhat filtering) to detect dark linear structures using Equation 3.9:

$$B = (I_{\text{gray}} \cdot S) - I_{\text{gray}} \tag{3.9}$$

where S is a structuring element (size 17×17) and dot $(\cdot)$ denotes morphological closing as shown in Figure 3.12. Morphological closing first dilates and then erodes the grayscale image, effectively bridging small gaps and smoothing contours. The Blackhat operation then highlights dark features, including hair strands that are thinner than the structuring element and not part of the background. This filtered output B serves as a mask to isolate and remove hair artifacts, enabling cleaner lesion visualization and improving downstream segmentation and classification accuracy. This operation is particularly effective in isolating thin, dark

strands that overlap lesion regions, which are otherwise difficult to detect using standard thresholding. The resulting mask can be refined further using inpainting or interpolation techniques to restore occluded skin areas without distorting lesion boundaries.

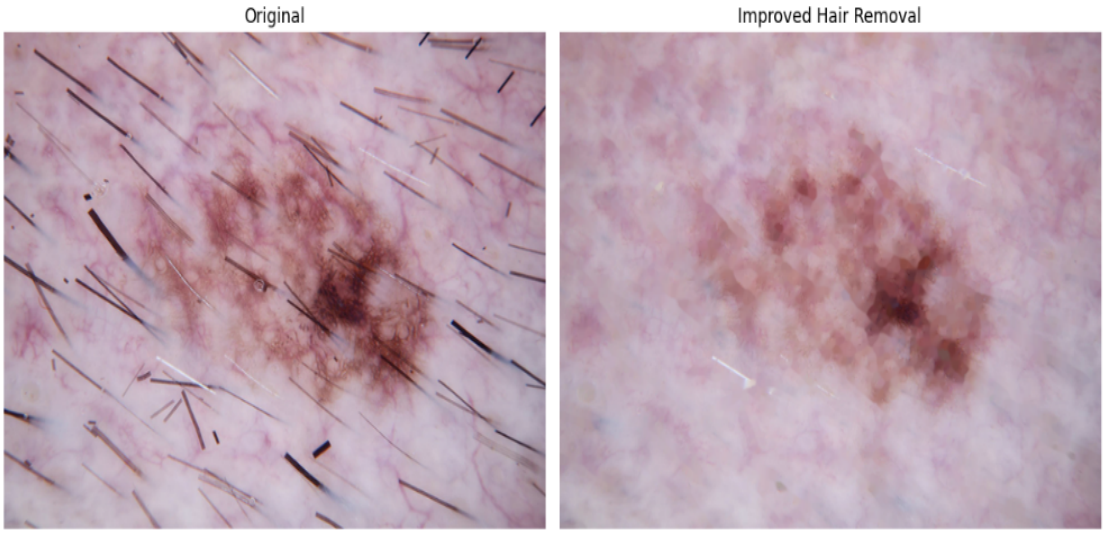


FIGURE 3.12: Applying morphological operations (Blackhat filtering)

Thresholding the Blackhat-filtered image $B(x, y)$ using Equation 3.10 is a crucial step to isolate hair regions by generating a binary mask $M_h(x, y)$. This operation assigns a value of 10 to pixels where the intensity exceeds the threshold of 10, effectively flagging dark linear structures, typically hair strands, while setting all other pixels to zero. The threshold is empirically chosen to balance sensitivity and specificity, ensuring that genuine hair artifacts are captured without including background noise or lesion texture. The resulting mask simplifies subsequent inpainting or interpolation processes by clearly identifying regions that require restoration, thereby enhancing lesion visibility and improving the accuracy of downstream segmentation and classification tasks.

$$M_h(x, y) = \begin{cases} 10 & \text{if } B(x, y) > 10 \\ 0 & \text{otherwise} \end{cases} \tag{3.10}$$

Equation 3.11 applies the Telea inpainting method to restore occluded regions in the noise-reduced image I_{nr} using the binary hair mask M_h and a defined inpainting

radius r. The operation fills in masked areas by propagating surrounding pixel information into the regions affected by hair artifacts, resulting in the restored image I_{hr}.

$$I_{hr} = I_{\text{inpainted}}(I_{nr}, M_h, r)1 \tag{3.11}$$

This technique effectively reconstructs skin texture while preserving lesion boundaries, thereby enhancing visual clarity and improving the reliability of downstream segmentation and classification. The Telea method is particularly suitable for dermatoscopic images due to its edge-aware interpolation, which maintains structural continuity in complex skin patterns as illustrated in Figure 3.13. Restoring the occluded skin texture and improving lesion visibility.

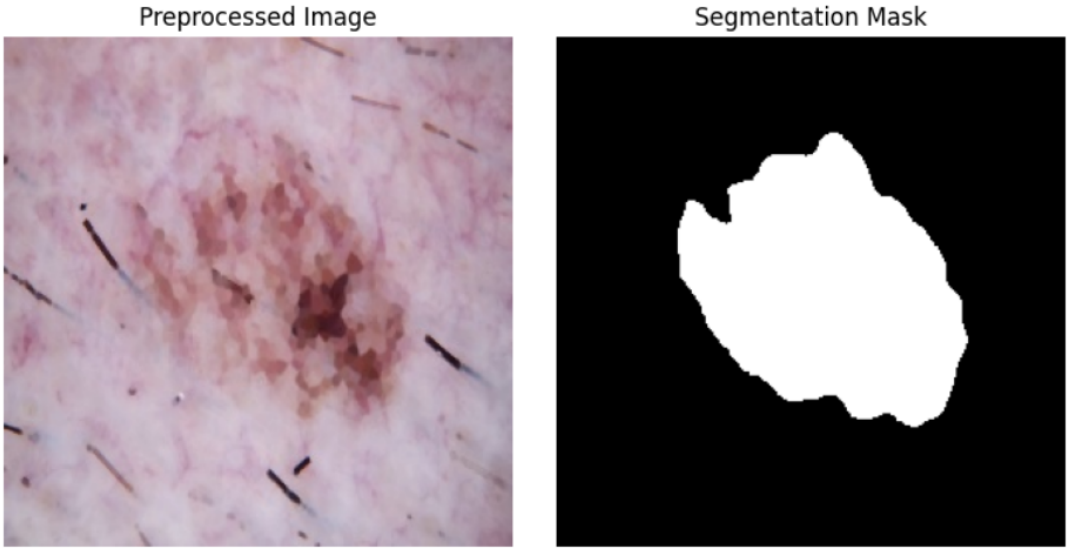


FIGURE 3.13: Apply Segmentation Masking Technique

3.3.4 Color Normalization

To reduce variation due to lighting and camera differences, color normalization is applied in the LAB color space, which separates luminance from chromatic information. Images are first converted from RGB/BGR to LAB space $I_{lab} = (L, A, B)$, where, L represents lightness and A, B encodes color components. Histogram equalization via CLAHE (Contrast Limited Adaptive Histogram Equalization) is

then applied to the luminance channel L, enhancing local contrast while preventing over-amplification of noise, as shown in Equation 3.12:

$$L_{eq} = \text{CLAHE}(L) \tag{3.12}$$

This operation improves visibility of lesion boundaries and texture details across varying illumination conditions. The enhanced luminance channel L_{eq} is recombined with the original chromatic channels A and B to form the normalized LAB image:

$$I_{lab_{norm}} = (L_{eq}, A, B) \tag{3.13}$$

Finally, the image is converted back to BGR format for compatibility with standard image processing libraries and deep learning models:

$$I_{cn} = \text{LAB2BGR}(I_{lab_{norm}}) \tag{3.14}$$

This preprocessing step ensures uniform brightness and contrast across the dataset, reduces inter-image variability, and enhances the robustness of downstream segmentation and classification tasks.

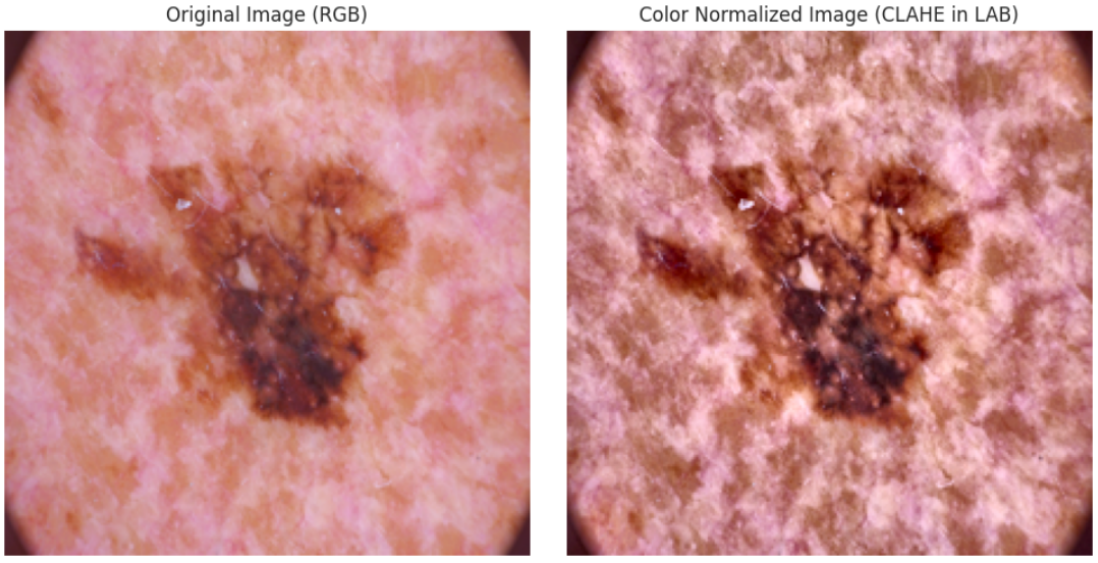


FIGURE 3.14: Apply Color Normalization using LAB Technique on lesion image

3.3.5 Artifact Removal

To mitigate shadow and vignetting effects, CLAHE is applied to the L channel of LAB images, enhancing local contrast without amplifying noise or color distortions. To reduce shadows and vignetting, CLAHE is reapplied to the luminance channel as represented in Equation 3.15 :

$$L_{\text{sh}} = \text{CLAHE}(L) \tag{3.15}$$

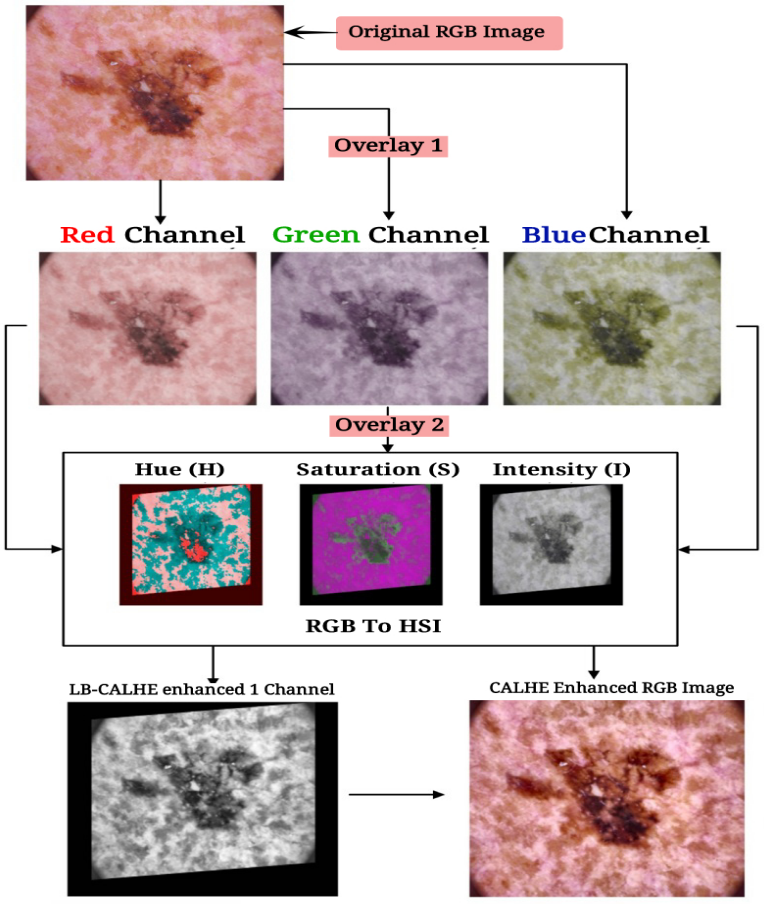


FIGURE 3.15: Apply CLAHE through L Channel on LAB lesion image

In Equation 3.16, the LAB image I_{lab} is updated and converted back to RGB format to restore the original color space. It also maintains visual consistency for interpretability in clinical visualization tools.

$$I_{\text{ar}} = \text{LAB2BGR}((L_{\text{sh}}, A, B)) \tag{3.16}$$

Enhancing local contrast and reducing shadow artifacts, as illustrated in Figure 3.15, significantly improves the visibility of subtle lesion features that may otherwise be obscured under uneven lighting. By applying CLAHE to the luminance channel in LAB space, the method ensures that brightness variations are locally adjusted without distorting color fidelity. This results in a more uniform visual representation of skin texture, aiding in the accurate delineation of lesion boundaries and improving the reliability of subsequent segmentation and classification stages.

3.3.6 Lesion Segmentation (Placeholder U-Net Model)

A placeholder U-Net model is implemented to simulate lesion mask extraction. In the absence of a trained segmentation model, a dummy binary mask is generated as a circular region at the image center, as shown in Figure 3.16. Real U-Net integration would involve tensor transformations, model inference, and thresholding to obtain lesion masks.

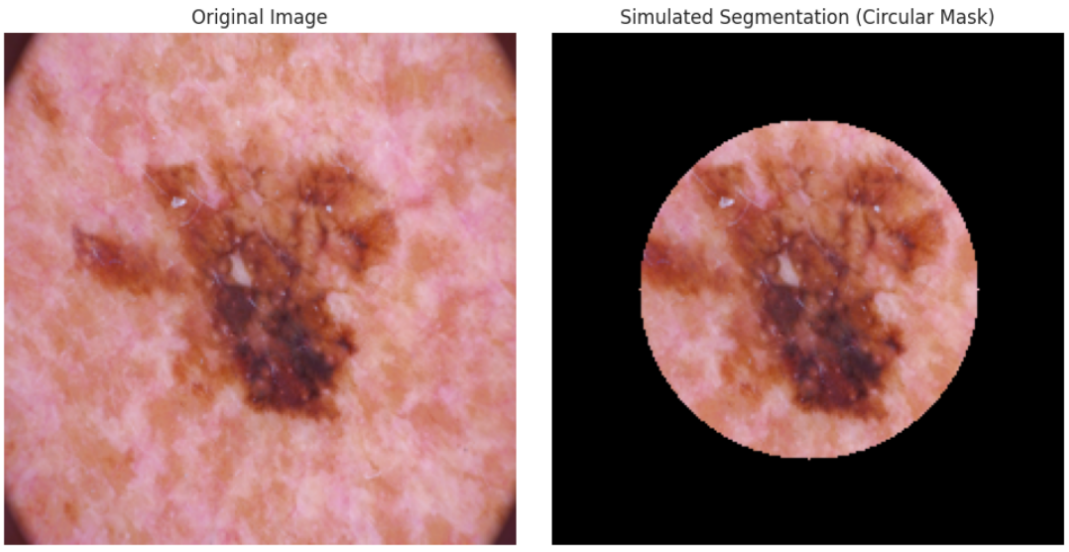


FIGURE 3.16: Apply Lesion Segmentation using placeholder U-Net Model on the Lesion Image

A placeholder U-Net model M_{UNet} with a convolutional layer outputs a binary lesion mask M as mentioned in equation 3.17:

$$M = \sigma(\text{Conv}(I_{\text{ar}})) \tag{3.17}$$

In this formulation, $I_{ar} \in \mathbb{R}^{H \times W \times C}$ represents the input image, typically preprocessed to reduce artifacts and enhance lesion visibility, where H, W, and C denote the image height, width, and number of channels, respectively. The convolutional operation extracts spatial features and produces a raw output map $Z(x, y)$, which is then passed through the sigmoid function $\sigma(z) = \frac{1}{1+e^{-z}}$ to compress each pixel value into the range $[0, 1]$. The resulting mask $M(x, y)$ indicates the probability of lesion presence at each pixel location.

To convert this soft mask into a binary segmentation output, thresholding at 0.5 produces a binary mask as shown in Equation 3.18:

$$M_b(x, y) = \begin{cases} 1 & \text{if } M(x, y) > 0.5 \\ 0 & \text{otherwise} \end{cases} \tag{3.18}$$

This binary mask M_b is then applied to the input image I_{ar} to isolate lesion regions, effectively suppressing background noise and non-lesion artifacts. Such selective masking enhances region-specific analysis and ensures that subsequent feature extraction is concentrated on diagnostically relevant areas, thereby improving classification accuracy and clinical interpretability.

3.4 Image Contrast Enhancement

One of the most important preprocessing steps indermoscopic image analysis, particularly when distinguishing between subtle lesion features and surrounding skin. Dermatoscopic images often suffer from poor contrast due to uneven illumination, low pigmentation differences, and imaging artifacts. Such limitations can obscure critical diagnostic indicators, including border irregularities, internal color variations, and texture patterns. Effective contrast enhancement enhances the visibility of these features, thereby improving the accuracy and reliability of subsequent

segmentation and classification stages in the automated skin lesion recognition framework.

3.4.1 Analysis of Contrast Enhancement Techniques

To address these challenges, several contrast enhancement techniques are reviewed and evaluated based on their suitability for the HAM10000 dataset, as shown in Table 3.2. These techniques are selected according to their ability to balance contrast improvement with structural preservation while minimizing noise amplification. Among the tested methods are HE, ACS, CLAHE, CN, BCT, DSLR, and DCT-DLAM. These methods are benchmarked using quantitative metrics such as PSNR, SSIM, and AMBE to assess enhancement quality and structural fidelity. The evaluation also considered computational efficiency and adaptability to varying lesion types, ensuring practical relevance for clinical deployment.

3.4.1.1 Histogram Equalization

During the preprocessing pipeline of the HAM10000 dataset, histogram equalization is used to correct illumination changes and improve contrast when using dermatoscopic images. The method involves redistribution of pixel intensities to produce an even histogram and increase visibility of lesion features as illustrated in Figure 3.17. The key advantage of histogram equalization is its ease, as it can be easily computed and is fast and efficient to perform in preprocessing. However, its tendency to over-enhance noise and struggle with subtle lesion features limits its effectiveness for complex dermatoscopic images, as identified in the literature review in Chapter 2. Nevertheless, histogram equalization becomes a cornerstone in my pipeline in order to normalize image contrast [3][12]. Its integration ensures consistent intensity distribution across samples, which supports downstream segmentation and classification stages with greater stability. This uniformity also facilitates smoother feature extraction by reducing contrast variability across lesion boundaries. Moreover, it improves the reliability of color-based descriptors used in early lesion differentiation.

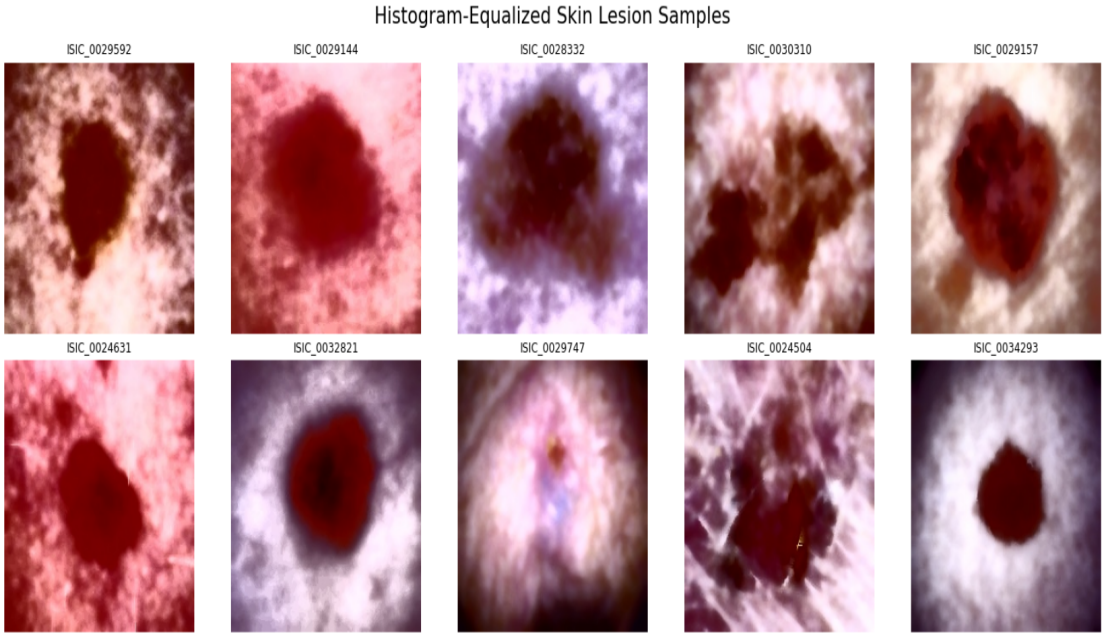


FIGURE 3.17: Apply Histogram Equalization (HE) on HAM10000 dataset

3.4.1.2 Adaptive Contrast Stretching

To tackle uneven lighting conditions in the HAM10000 dataset, my methodology incorporates adaptive contrast stretching, which dynamically adjusts pixel intensities based on local image characteristics. Its strength lies in its adaptability to varying lighting conditions, making it effective for real-world dermatoscopic images with inconsistent brightness. However, the computational intensity of this method poses a challenge, requiring more processing resources compared to simpler techniques like histogram equalization. As noted in Chapter 2, adaptive methods are underutilized in dermatological preprocessing [10][19], and my inclusion of this technique addresses this gap by improving feature visibility in unevenly lit images.

To ensure reproducibility and modularity, adaptive contrast stretching is implemented as a configurable preprocessing block within the pipeline, allowing for controlled experimentation across datasets. Its impact is quantitatively validated through improved segmentation accuracy and enhanced feature map clarity in downstream classification stages. This enhancement step also facilitates better edge definition and lesion boundary clarity, which are critical for accurate segmentation and downstream diagnostic interpretation.

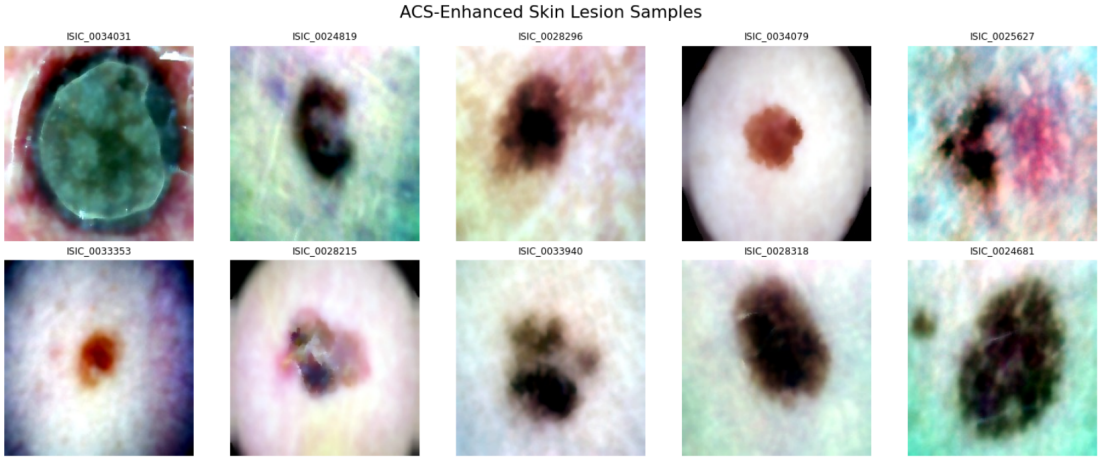


FIGURE 3.18: Apply Adaptive Contrast Stretching (ACS) on HAM10000 dataset

3.4.1.3 Contrast Limited Adaptive Histogram Equalization

The preprocessing pipeline integrates Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance local features while mitigating noise amplification in HAM10000 dataset images. CLAHE adapts histogram equalization to local regions, applying a clip limit to control over-enhancement as shown in Figure 3.19. Its primary strength is its ability to enhance local lesion details without amplifying background noise, making it well-suited for dermatoscopic images with structural complexity. However, CLAHE requires careful tuning of parameters, such as clip limit and tile size, which adds complexity to its implementation, as displayed in Figure 3.22.

FIGURE 3.19: Apply Contrast Limited Adaptive Histogram Equalization (CLAHE) on HAM10000 dataset

This technique addresses a literature gap in preserving fine details in medical imaging [3][19], and its inclusion in my pipeline ensures robust feature extraction for the hybrid ResNet21-EfficientNetB0 model.

3.4.1.4 Contrast Normalization

Contrast normalization is implemented in my methodology to stabilize input images from the HAM10000 dataset under varying lighting conditions as ilustrated in Figure 3.20. This technique normalizes pixel intensities to a standard range, and its strength lies in its ability to maintain consistent image quality across diverse illumination scenarios, enhancing the model's generalizability. However, contrast normalization may suppress fine textures critical for distinguishing subtle lesion characteristics, as highlighted in the literature review in Chapter 2. By incorporating this method, my pipeline ensures standardized input data, facilitating reliable performance of the subsequent deep learning model while addressing challenges of illumination variability [10][12].

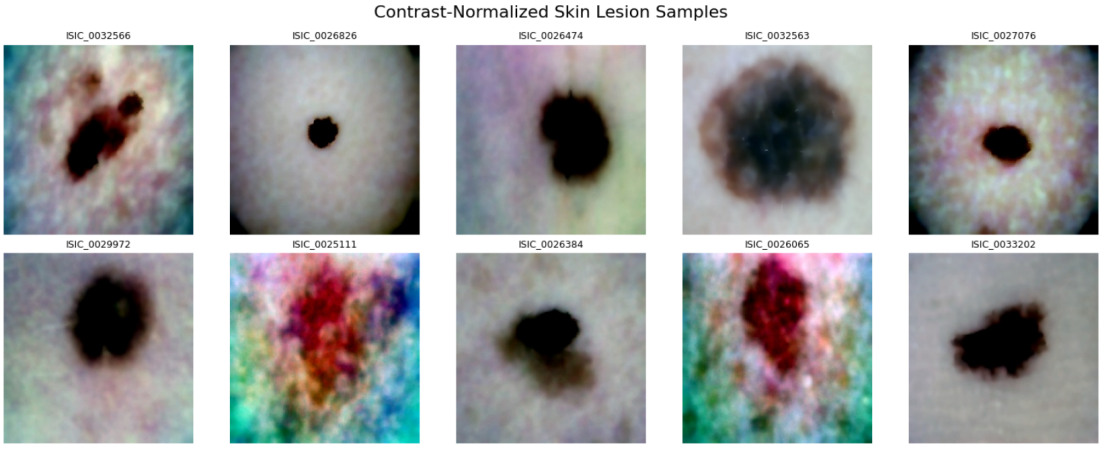


FIGURE 3.20: Apply Contrast Normalization (CN) on HAM10000 dataset

3.4.1.5 Brightness and Contrast Tuning

To provide a computationally efficient preprocessing step, our methodology includes brightness and contrast tuning for the HAM10000 dataset images as displayed in Figure 3.21. This technique adjusts global brightness and contrast levels, and its primary strengths are its speed and ease of implementation, making

it suitable for rapid preprocessing in resource-constrained settings. However, its simplicity limits its effectiveness for capturing complex lesion details, as it struggles to address structural variability in dermatoscopic images [3][19]. Despite this limitation, as noted in Chapter 2, its inclusion in my pipeline provides a baseline enhancement to complement more advanced techniques like CLAHE.

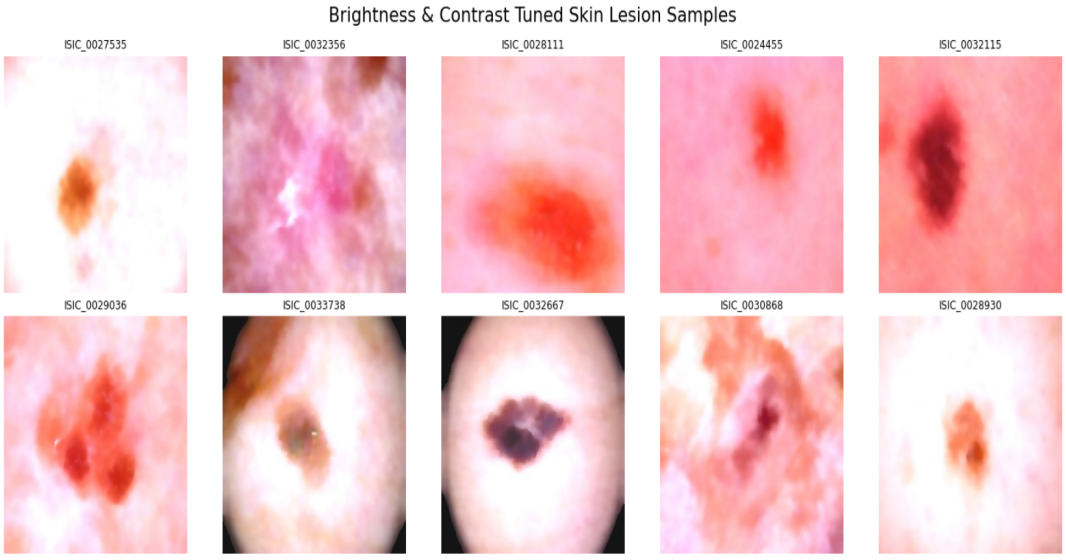


FIGURE 3.21: Apply Brightness and Contrast Tuning (BCT) on HAM10000 dataset

3.4.1.6 Deep Learning Super-Resolution

Deep learning-based super-resolution is incorporated into the preprocessing pipeline to recover microstructural information in low-resolution images from the HAM10000 dataset. This method addresses a resolution enhancement gap mentioned in Chapter 2 by using a neural network to upscale images while maintaining diagnostic feature [10][12]. Reconstructing tiny details that are necessary for accurate lesion classification is its strength. However, its applicability for large-scale applications is limited by its processing cost and the need for a huge amount of training data. Our approach improves the quality of input images by using super-resolution, which helps the hybrid model extract complex information. The enhanced resolution not only sharpens lesion boundaries but also improves the visibility of subtle pigmentation patterns and vascular structures, which are critical for differentiating between benign and malignant cases. To ensure reproducibility, the

super-resolution module is integrated as a scalable preprocessing block, allowing controlled experimentation across varying image resolutions and lesion types.

3.4.1.7 Dynamic Contrast Tuning (Deep Learning Adaptive Models)

Dynamic contrast tuning, utilizing deep learning adaptive models, is implemented in the framework pipeline to automatically adjust contrast based on dataset characteristics. Its primary strength is its adaptability to the HAM10000 dataset's unique properties, optimizing contrast without manual parameter tuning. However, the complexity of training and potential model instability, as noted in Chapter 2, present challenges that require careful management [3][19]. By incorporating this technique, my methodology addresses illumination and contrast variability, ensuring robust input data for the ResNet21-EfficientNetB0 model.

3.4.2 Performance Measures

In order to conclude quantitatively about the quality of the pre-processed dermoscopic images, various standard image quality assessment measures are employed. Such measures determine structural fidelity, noise reduction, contrast enhancement, and brightness consistency, each of which plays a critical role in ensuring diagnostic reliability. Specifically, the Structural Similarity Index Measurement (SSIM) assesses perceptual similarity between original and enhanced images by comparing luminance, contrast, and structure. Peak Signal-to-Noise Ratio (PSNR) quantifies the level of noise reduction by measuring the ratio between the maximum possible signal and the distortion introduced during enhancement. Contrast Improvement Index (CII) evaluates the degree of contrast amplification, which is essential for highlighting lesion boundaries and internal patterns. Lastly, Absolute Mean Brightness Error (AMBE) measures the deviation in average brightness, ensuring that enhancement techniques do not distort the overall illumination profile. Together, these metrics provide a robust framework for benchmarking preprocessing performance across diverse enhancement methods.

3.4.2.1 Structural Similarity Index Measurement

By comparing structural information, the SSIM is the supposed deterioration in the quality of the image between the original image I and the processed image I'. It analyzes the luminance, the contrast, and the structure similarity. The SSIM between two image patches x and y is defined in Equation 3.19:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3.19}$$

Where:

- μ_x, μ_y are the mean pixel intensities of x and y.
- σ_x^2, σ_y^2 are the variances.
- σ_{xy} is the covariance between x and y.
- $C_1 = (k_1 L)^2$, $C_2 = (k_2 L)^2$ are stabilizing constants.
- L is the dynamic range of pixel values (e.g., 255 for 8-bit images).

The Structural Similarity Index Measurement (SSIM) value closer to 1 indicates higher structural similarity between the original and enhanced image, signifying minimal distortion introduced during preprocessing.

3.4.2.2 Peak Signal-To-Noise Ratio

PSNR calculates the ratio of the highest signal strength (pixel intensity) and the noise introduced by preprocessing. It is calculated in decibels (dB) using the mean squared error (MSE) between the original image I and the processed image I' in Equation 3.20 and Equation 3.21:

$$\text{MSE} = \frac{1}{XY} \sum_{i=1}^{X} \sum_{j=1}^{Y} [I(i, j) - I'(i, j)]^2 \tag{3.20}$$

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{L^2}{\text{MSE}} \right) \tag{3.21}$$

Where:

- X, Y are image dimensions.
- L is the maximum possible pixel value (255 for 8-bit images).

A larger value of PSNR implies a higher visual quality and less noise addition.

3.4.2.3 Contrast Improvement Index

The Contrast Improvement Index (CII) measures the relative enhancement of image contrast due to preprocessing. It uses a comparison of standard deviation of pixel intensities obtained before and after enhancement as shown in Equation 3.22:

$$\text{CII} = \frac{\sigma_{\text{IMAGE}'}}{\sigma_{\text{IMAGE}}} \tag{3.22}$$

Where:

- σ_{IMAGE} is the standard deviation of the original image.
- $\sigma_{\text{IMAGE}'}$ is the standard deviation of the processed image.

Any CII over 1 shows that there is a contrast enhancement facilitated in the manipulated image. Higher CII values indicate stronger contrast amplification, often reflecting improved visibility of lesion boundaries. This metric is especially useful for quantitatively comparing enhancement techniques across diverse image samples. It provides an objective measure to validate preprocessing effectiveness, especially in low-contrast dermatoscopic images, and suggests improved feature clarity across lesion types, aiding reliable model training and evaluation.

3.4.2.4 Absolute Mean Brightness Error

AMBE calculates preservation of the brightness by computing the absolute difference between averaged intensities of the original image and the processed image in Equation 3.23. A lower AMBE value indicates better brightness preservation, suggesting minimal deviation from the original illumination. This metric is particularly useful for evaluating enhancement techniques that aim to retain natural image appearance. It complements contrast-based metrics by ensuring that enhancement does not distort the original luminance characteristics critical for clinical interpretation.

$$\text{AMBE} = |\mu_{\text{IMAGE}} - \mu_{\text{IMAGE}'}| \tag{3.23}$$

Where:

- μ_{IMAGE} is the mean intensity of the original image.
- $\mu_{\text{IMAGE}'}$ is the mean intensity of the enhanced image.

When the AMBE is lower, it means that the brightness characteristics of the original image are better preserved.

A comparative evaluation of these techniques is conducted using typical measures of visual quality, like Structural Similarity Index Measurement (SSIM), Peak Signal-to-Noise Ratio (PSNR), Contrast Improvement Index (CII), and Absolute Mean Brightness Error (AMBE). CLAHE approached these metrics the most even among all techniques, becoming the most suitable one to be incorporated into the preprocessing pipeline. Its ability to enhance local contrast without amplifying noise makes it particularly effective for dermatoscopic images with subtle lesion boundaries. By preserving structural integrity and improving visibility across varied conditions, CLAHE supports consistent feature extraction and reliable downstream classification. Its integration also streamlines the preprocessing workflow, ensuring consistent enhancement without manual tuning across large-scale datasets.

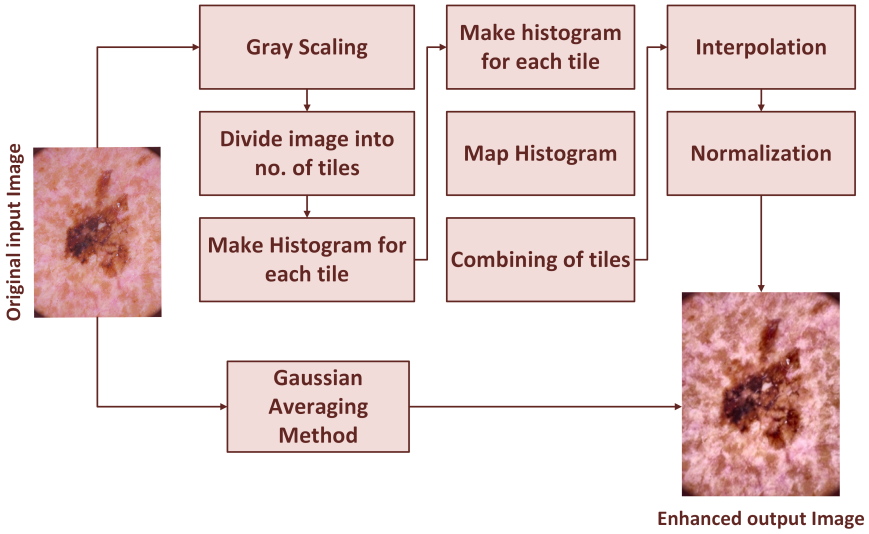


FIGURE 3.22: Architecture of CLAHE Algorithm

Additionally, the visual effectiveness of CLAHE is illustrated in Figure 3.23, which shows original and enhanced images for multiple lesion types. As depicted, CLAHE significantly improves local contrast without introducing artifacts, aiding in the accurate differentiation of lesion boundaries and internal features. The underlying mechanism of CLAHE involves histogram clipping and redistribution, followed by transformation using the cumulative distribution function (CDF) of the modified histogram. The final pixel intensities are interpolated across adjacent regions to ensure seamless transitions without visible tile boundaries.

3.4.2.5 Histogram Clipping

It is used to suppress dominant intensity levels and redistribute pixel values more evenly across the dynamic range. Each histogram bin $H(i)$ is clipped at a threshold T_c, as shown in Equation 3.24:

$$H'(i) = \begin{cases} H(i), & \text{if } H(i) \leq T_c \\ T_c, & \text{if } H(i) > T_c \end{cases} \tag{3.24}$$

This helps prevent any single intensity level from dominating the histogram. By capping the bin values, it keeps the contrast enhancement balanced and avoids

making parts of the image too bright or too dark. The extra pixels are later spread evenly across all bins to maintain smooth intensity transitions. By limiting the maximum bin height, histogram clipping facilitates more balanced redistribution during equalization, which is especially beneficial for dermoscopic images with uneven illumination or high contrast variability. It serves as a foundational step in adaptive histogram-based methods such as CLAHE, where controlled contrast enhancement is critical for highlighting lesion boundaries without introducing artifacts.

3.4.2.6 Clipped Pixel Redistribution

After histogram clipping, the number of clipped pixels C is redistributed uniformly across all histogram bins to maintain a balanced intensity distribution. This step ensures that the excess pixel count does not concentrate in any single bin, which could otherwise lead to localized over-enhancement and structural distortion. The total number of clipped pixels is calculated using Equation 3.25 and Equation 3.26. This step ensures that the excess pixel count does not concentrate in any single bin, maintaining a balanced intensity distribution.

$$C = \sum_{i=0}^{L-1} \max(0, H(i) - T_c) \tag{3.25}$$

Where $H(i)$ is the original histogram bin count, T_c is the clipping threshold, and L is the number of gray levels (typically 256). The redistribution is performed by adding a uniform fraction of the clipped pixels to each bin, as shown in Equation 3.26:

$$H''(i) = H'(i) + \frac{C}{L} \tag{3.26}$$

This redistribution ensures that the histogram remains balanced after clipping, preventing over-enhancement in specific intensity ranges. By evenly spreading the excess pixels, the method maintains smooth transitions across gray levels, which is critical for preserving visual quality and diagnostic integrity in dermoscopic imaging. It also helps avoid saturation artifacts and supports consistent contrast improvement across diverse lesion regions. In the context of CLAHE, this step

plays a pivotal role in local contrast amplification while safeguarding against noise amplification and intensity bias.

3.4.2.7 Local Contrast Mapping

The final intensity transformation for a given pixel value r is derived from the cumulative distribution of the clipped histogram as in Equation 3.27:

$$s = T_r(r) = (L-1)\sum_{k=0}^{r}\frac{H''(k)}{M \times N} \tag{3.27}$$

Where:

- s is the output intensity and r is the original intensity,
- $M \times N$ is the number of pixels in the tile.

After all tiles are processed, bilinear interpolation is applied between neighboring regions to eliminate boundary artifacts and ensure smooth transitions. Figure 3.23 presents a comparative visualization of a grayscale image before and after CLAHE enhancement. The enhanced image demonstrates significantly improved local contrast and structural delineation, enabling finer visibility of subtle intensity variations essential for precise image interpretation.

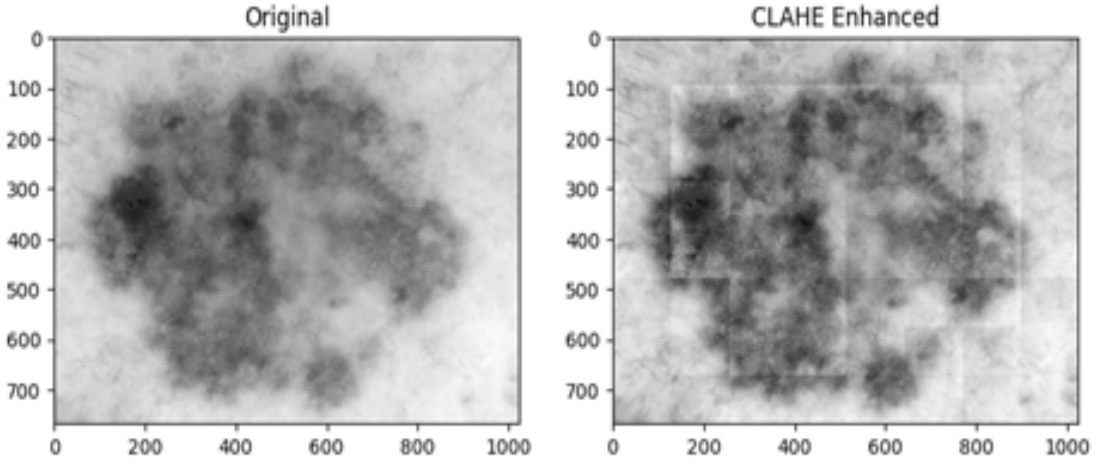


FIGURE 3.23: Comparative visualization grayscale image before and after CLAHE enhancement algorithm

Table 3.2: Detailed Comparison of Contrast Enhancement Algorithms

Algorithm	**SSIM**	**PSNR**	**CII**	**AMBE**	**Strengths**	**Weaknesses**
Histogram Equalization	0.6277	13.45 dB	2.17	32.36	Simple and fast	Over-enhances noise; poor for subtle features
Adaptive Contrast Stretching	0.8685	15.13 dB	1.89	21.23	Effective under uneven lighting	Computationally more intensive
CLAHE	0.8893	25.13 dB	1.19	6.39	Enhances local features; prevents noise amplification	Requires parameter tuning (clip limit, tile size)
Contrast Normalization	0.8235	12.71 dB	1.68	31.42	Stabilizes input under lighting variation	May suppress fine textures
Brightness and Contrast Tuning	0.9537	15.64 dB	1.04	41.04	Very fast, easy to implement	Not effective for complex lesion details
Deep Learning Super-Resolution	0.1458	5.52 dB	0.0688	15.34	Restores microstructures	Computationally expensive; requires training
Dynamic Contrast Tuning (DL adaptive models)	0.9277	23.51 dB	1.1268	3.74	Adapts to dataset characteristics automatically	Adds training complexity; model instability if not managed

By incorporating CLAHE into the preprocessing pipeline, our methodology ensures enhanced image quality while maintaining diagnostic integrity. This contrast enhancement step plays a pivotal role in improving feature extraction performance in the subsequent hybrid ResNet21–EfficientNetB0 architecture. CLAHE's localized contrast amplification is particularly effective in handling dermatoscopic images with uneven illumination and low-contrast lesions. By adaptively redistributing pixel intensities within each tile, it highlights faint lesion boundaries and internal structures that might otherwise remain indistinct. This enhancement not only facilitates better visual assessment but also strengthens the discriminative power of deep features extracted by the hybrid ResNet21 and EfficientNetB0 model. As a result, the network can more reliably differentiate between lesion types, contributing to improved classification accuracy and generalization across diverse skin conditions.

3.5 Data Augmentation

After applying CLAHE-based contrast enhancement to improve local image quality, the enhanced dermoscopic images are converted back to RGB format to preserve color fidelity for downstream processing. These RGB images are then subjected to a comprehensive augmentation pipeline using the 'imgaug' library, which introduces controlled variability to simulate real-world conditions and improve model generalization. A curated set of 10 representative images underwent randomized transformations, including horizontal and vertical flipping, Gaussian blurring, additive noise injection, and affine transformations such as scaling, rotation, and shearing, as illustrated in Figure 3.24.

These augmentations serve multiple purposes, such as flipping introduces spatial diversity, blurring and noise simulate acquisition imperfections, and affine transformations mimic variations in lesion orientation and size. Together, they expand the dataset's representational richness, helping the model learn invariant features and reducing overfitting. This two-step approach, CLAHE followed by augmentation, not only enhances the visual clarity of dermoscopic structures but

also increases the statistical diversity of the training set. As a result, the hybrid classification framework becomes more robust to lighting inconsistencies, image artifacts, and anatomical variability. Quantitative improvements are observed in model performance metrics such as accuracy, precision, and F1-score, confirming the effectiveness of this preprocessing strategy. By integrating enhancement and augmentation in a modular pipeline, the methodology ensures reproducibility and adaptability across datasets, making it a critical foundation for building accurate and generalizable deep learning models for skin lesion analysis.

Furthermore, this approach supports better feature map activation in convolutional layers, as the augmented inputs expose the model to a wider range of lesion morphologies and textures. This diversity is especially important for handling rare lesion types and improving classification confidence in borderline cases. The augmentation module is designed to be scalable, allowing future extensions with domain-specific transformations such as color jittering, elastic deformation, or lesion-specific masking to further enrich the training data.

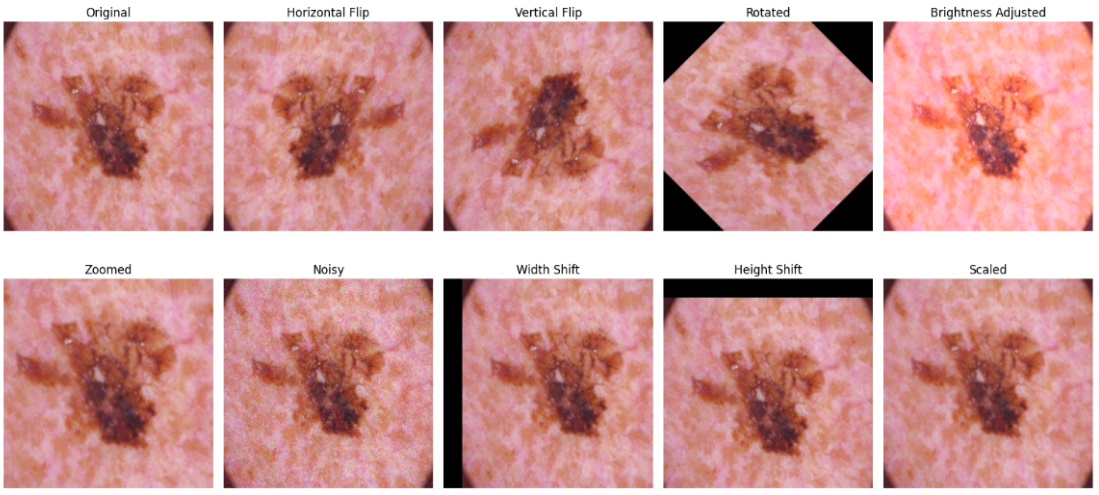


FIGURE 3.24: Apply Data Augmentation on lesion image

3.5.1 Data Distribution Before and After Image Augmentation

The original HAM10000 dataset exhibits significant class imbalance across its seven diagnostic categories. Melanocytic nevi (NV) dominate the dataset with

6,675 images, while rare classes such as vascular lesions (VASC) and dermatofibroma (DF) contain only 102 and 144 images, respectively. This imbalance poses a challenge for deep learning models, which tend to favor majority classes, resulting in reduced sensitivity and poor generalization for underrepresented lesion types.

To address this, a targeted augmentation strategy is applied using the `imgaug` library. Augmentation techniques included horizontal and vertical flipping, Gaussian blurring, additive noise injection, and affine transformations (scaling, rotation, and shearing). These transformations are selectively applied to each class to synthetically expand the dataset and improve class-level balance. Table 3.3 summarizes the number of training, validation, and test images per class, along with the total image count before and after augmentation. For instance, AKIEC increased from 327 to 981 images, BCC from 514 to 1,542, and MEL from 1,674 to 5,022. The most dramatic expansion occurred in the NV class, which grew from 6,675 to 20,115 images. Even the smallest classes—DF and VASC—are tripled in size, reaching 432 and 426 images, respectively.

TABLE 3.3: Class-wise Data Distribution Before and After Augmentation

ID	**Skin Lesion Type**	**Train (70%)**	**Val (15%)**	**Test (15%)**	**Before Aug.**	**After Aug.**
0	AKIEC	228	49	50	327	981
1	BCC	359	77	74	514	1542
2	BKL	769	165	165	1099	3297
3	DF	100	22	22	144	432
4	NV	4693	1006	976	6675	20115
5	MEL	1179	252	243	1674	5022
6	VASC	72	16	14	102	426

This rebalancing not only improved statistical uniformity but also introduced greater variability in lesion orientation, texture, and lighting conditions. The augmented dataset supports more robust feature learning, reduces overfitting, and

enhances the model's ability to generalize across diverse clinical presentations. By combining CLAHE-based enhancement with class-aware augmentation, the preprocessing pipeline ensures both image clarity and dataset diversity, key prerequisites for building accurate and generalizable skin lesion classification models.

3.6 Model Architecture

The proposed hybrid model integrates DenseNet21 and EfficientNetB0 in a parallel feature extraction setup. This multi-branch architecture, as shown in Figure 3.32, is designed to combine low-level spatial features learned by DenseNet21 with high-level abstract representations learned by the EfficientNet model. This fusion enhances the model's ability to capture both fine-grained lesion details and broader contextual patterns, improving classification accuracy across diverse lesion types. The complementary strengths of both branches contribute to a richer feature space, supporting more robust decision-making in downstream layers. Additionally, the architecture benefits from balanced depth and efficiency, making it suitable for deployment in both high-performance and resource-constrained environments.

3.6.1 Vgg-19

VGG19 is a classic deep CNN architecture composed of 16 convolutional layers and 3 fully connected layers, using uniform 3×3 convolutions and 2×2 max-pooling throughout. Despite its relatively old design, VGG19 remains effective due to its simplicity and deep feature extraction capabilities. In this study, the ImageNet pre-trained VGG19 model is adapted for skin lesion classification by removing its default classifier and introducing a new dense layer with seven SoftMax outputs. Although it is computationally heavier than modern architectures, its deep and regular structure helps in capturing well-defined lesion shapes and consistent textural patterns, particularly useful for distinguishing sharply bordered benign lesions from malignant ones. Its integration into the pipeline complements lighter models by offering deeper hierarchical feature representations that enhance ensemble

performance. VGG19's stable gradient flow and uniform layer design contribute to reliable convergence during training, even with limited data. Its robustness against overfitting makes it a valuable backbone for transfer learning in medical imaging tasks. Its compatibility with transfer learning frameworks also enables efficient adaptation to new dermatological datasets with minimal retraining.

Let there $x \in \mathbb{R}^{H\times W\times C}$ be an input dermatoscopic image. The VGG19 feature extractor $f_{\text{VGG}}(x;\theta)$ maps x to a high-dimensional representation using convolutional layers with fixed parameters θ (pretrained on ImageNet), as in Equation 3.28. A new classifier head $g(\cdot;\phi)$ is appended, producing logits $z = g(f_{\text{VGG}}(x;\theta);\phi) \in \mathbb{R}^K$, where $K = 7$ is the number of lesion classes in Equation 3.29. The SoftMax function converts logits to probabilities:

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}, \quad i = 1, \ldots, K \tag{3.28}$$

The categorical cross-entropy loss for a true label $y \in \{1, \ldots, K\}$ is:

$$\mathcal{L} = -\log(p_y) = -\log\left(\frac{e^{z_y}}{\sum_{j=1}^{K} e^{z_j}}\right) \tag{3.29}$$

Gradients for backpropagation are computed as:

$$\frac{\partial \mathcal{L}}{\partial z_i} = p_i - \delta_{iy} \tag{3.30}$$

where in Equation 3.30 δ_{iy} is the Kronecker delta. During fine-tuning, selected layers of θ are unfrozen to allow gradient updates, enabling domain adaptation while preserving pretrained general features. Additionally, this makes VGG19 a dependable choice for benchmarking and comparative evaluation in ensemble-based diagnostic systems. Its layered depth enables fine-grained abstraction, which is particularly beneficial for capturing subtle inter-class variations in lesion morphology. Furthermore, its consistent architectural blocks simplify integration into modular pipelines for comparative benchmarking and ensemble fusion.

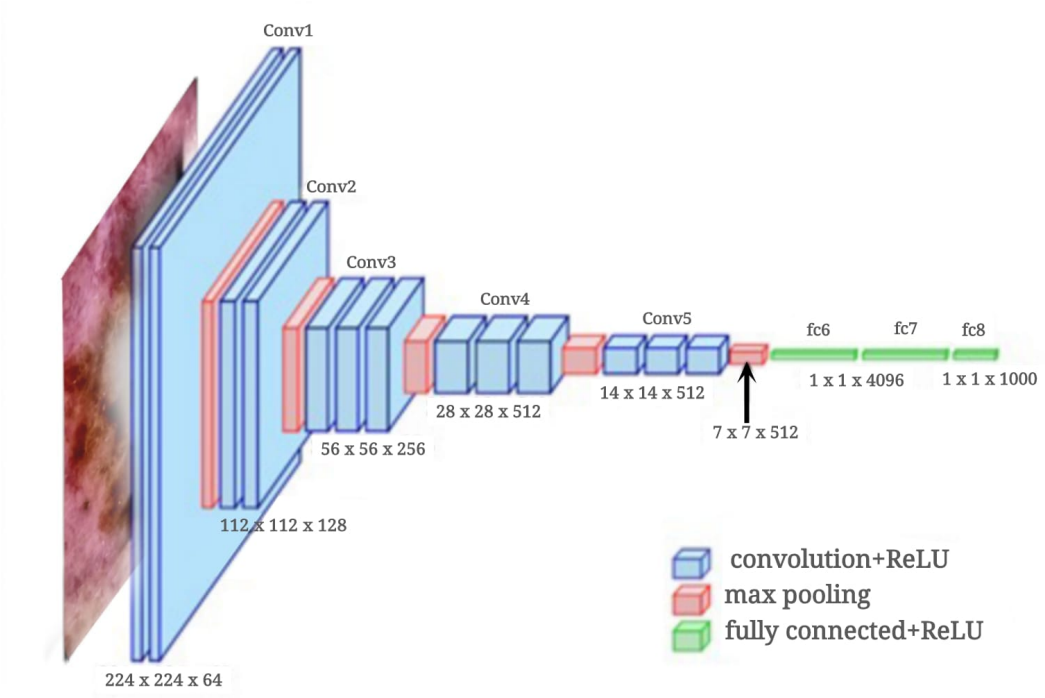


FIGURE 3.25: Architecture of the VGG-19 Model [65]

3.6.2 AlexNet

AlexNet is one of the earliest deep CNN models that revolutionized image classification through its use of ReLU activation, dropout, and overlapping max pooling. Three completely connected layers come after five convolutional layers, as shown in Figure 3.26. Let there $x \in \mathbb{R}^{H\times W\times C}$ be an input dermatoscopic image from the HAM10000 dataset. The AlexNet feature extractor $f_{\text{Alex}}(x;\theta)$ maps x to a latent representation using five convolutional layers with ReLU activations and overlapping max-pooling. The classifier head $g(\cdot;\phi)$ consists of three fully connected layers, with the final layer modified to produce logits $z \in \mathbb{R}^K$ in Equation 3.31, where the $K = 7$ corresponds to the lesion classes.

$$z = g(f_{\text{Alex}}(x;\theta);\phi) \tag{3.31}$$

In Equation 3.32, the SoftMax function converts logits into class probabilities, enabling multi-class prediction by normalizing outputs into a probability distribution.

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}, \quad i = 1, \ldots, K \tag{3.32}$$

The categorical cross-entropy loss for a true label $y \in \{1, \ldots, K\}$ is defined as:

$$\mathcal{L} = -\log(p_y) = -\log\left(\frac{e^{z_y}}{\sum_{j=1}^{K} e^{z_j}}\right) \tag{3.33}$$

The gradient of the loss with respect to each logit z_i is in Equation 3.34:

$$\frac{\partial \mathcal{L}}{\partial z_i} = p_i - \delta_{iy} \tag{3.34}$$

where δ_{iy} is the Kronecker delta, equal to 1 if $i = y$ and 0 otherwise. During transfer learning, the pretrained parameters θ may be frozen initially and selectively unfrozen to adapt to domain-specific features in the HAM10000 dataset.

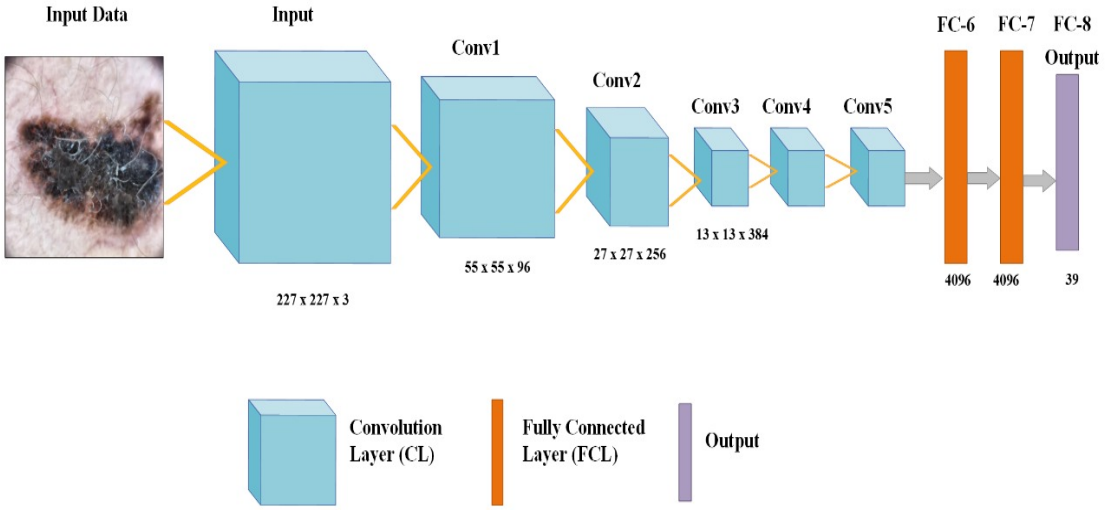


FIGURE 3.26: Architecture of AlexNet Model

Although relatively shallow compared to other models used in this study, AlexNet is included to benchmark the performance of lighter architectures in skin lesion classification. The model is fine-tuned using transfer learning, and the last layer is modified to output seven softmax-activated probabilities. It performed reliably in cases with distinct lesion characteristics, and due to its low computational

demand, it holds potential for mobile-based skin lesion screening applications. Its fast inference time and minimal memory footprint make it suitable for deployment in resource-constrained environments, while maintaining acceptable classification accuracy across diverse lesion types.

3.6.3 MobileNet-V2

MobileNetV2 is a lightweight CNN designed for mobile and embedded applications. It employs depthwise separable convolutions and inverted residual connections to achieve high accuracy while significantly reducing model size and latency. The model is fine-tuned on the skin lesion dataset after replacing its final classifier with a custom seven-class SoftMax output. MobileNetV2 is particularly advantageous in resource-constrained clinical settings where real-time analysis is required. Despite its compact architecture as shown in Figure 3.27, it demonstrated strong performance by effectively learning critical lesion features such as asymmetry, border irregularities, and color variations, making it a suitable candidate for teledermatology platforms. Its streamlined design also facilitates rapid deployment across diverse hardware configurations without compromising diagnostic reliability. Its consistent performance across varied lesion types also reinforces its role in ensemble configurations aimed at maximizing diagnostic coverage.

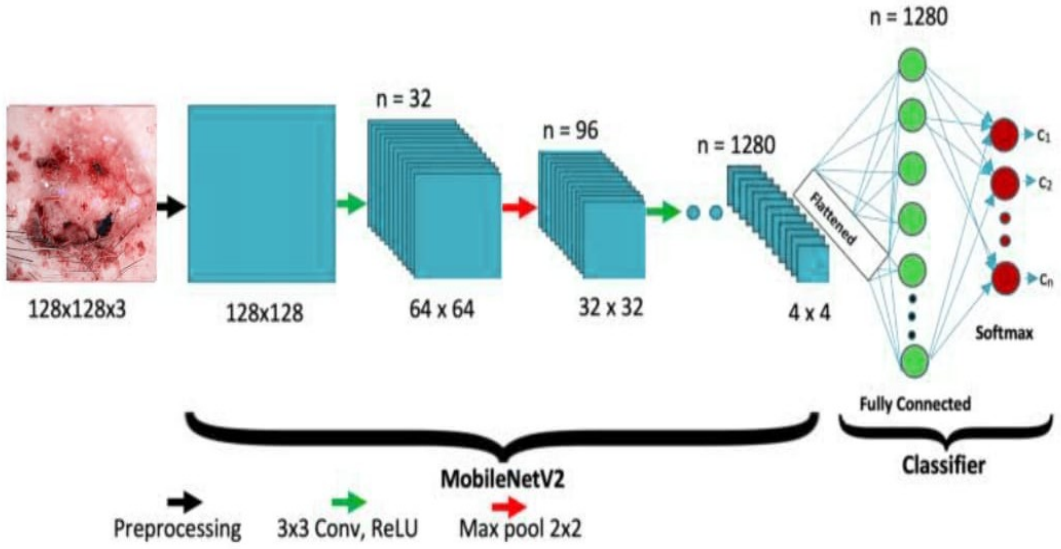


FIGURE 3.27: Architecture of MobileNet-v2 Model

In Figure 3.27, there $x \in \mathbb{R}^{H \times W \times C}$ is an input dermatoscopic image. MobileNetV2 processes x through a series of depthwise separable convolutions and inverted residual blocks, forming a compact feature extractor $f_{\text{Mobile}}(x; \theta)$ with pretrained parameters θ. The final classifier head $g(\cdot; \phi)$ is customized to output logits $z \in \mathbb{R}^K$, where $K = 7$ corresponds to the lesion classes:

$$z = g(f_{\text{Mobile}}(x; \theta); \phi) \tag{3.35}$$

The SoftMax function converts logits into class probabilities, enabling multi-class prediction by normalizing outputs into a probability distribution in such below Equation 3.36:

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}, \quad i = 1, \ldots, K \tag{3.36}$$

The categorical cross-entropy loss for a true label $y \in \{1, \ldots, K\}$ is defined as in Equation 3.37:

$$\mathcal{L} = -\log(p_y) \tag{3.37}$$

Gradients for backpropagation are computed by using this Equation 3.38:

$$\frac{\partial \mathcal{L}}{\partial z_i} = p_i - \delta_{iy} \tag{3.38}$$

Where δ_{iy} is the Kronecker delta, equal to 1 if $i = y$ and 0 otherwise. During transfer learning, selected layers of θ may be unfrozen to adapt to domain-specific lesion features while preserving MobileNetV2's lightweight efficiency. Its modular design also supports rapid deployment across heterogeneous clinical platforms, making it ideal for real-time dermatological screening in low-resource settings. Its streamlined architecture enables low-latency inference, crucial for point-of-care diagnostics. This makes MobileNetV2 a scalable solution for widespread dermatological screening across diverse populations.

3.6.4 ResNet-50 MODEL

ResNet50 is a deep residual learning network consisting of 101 layers that utilize skip connections (identity shortcuts) to increase training stability in deep networks and solve the problem of the shrinking gradient. It allows gradients to propagate through the network more effectively, enabling the model to learn more complex features without degradation in performance. The ResNet50 model is employed in this study with ImageNet pre-trained weights. For the task of skin lesion classification, the final fully connected layer is removed and replaced by a custom dense layer with seven output nodes, followed by a SoftMax activation function. The residual connections enable the model to extract both low-level and high-level contextual information, which is critical for distinguishing between lesion types that may appear visually similar but are clinically different. The full model architecture is presented in Figure 3.28.

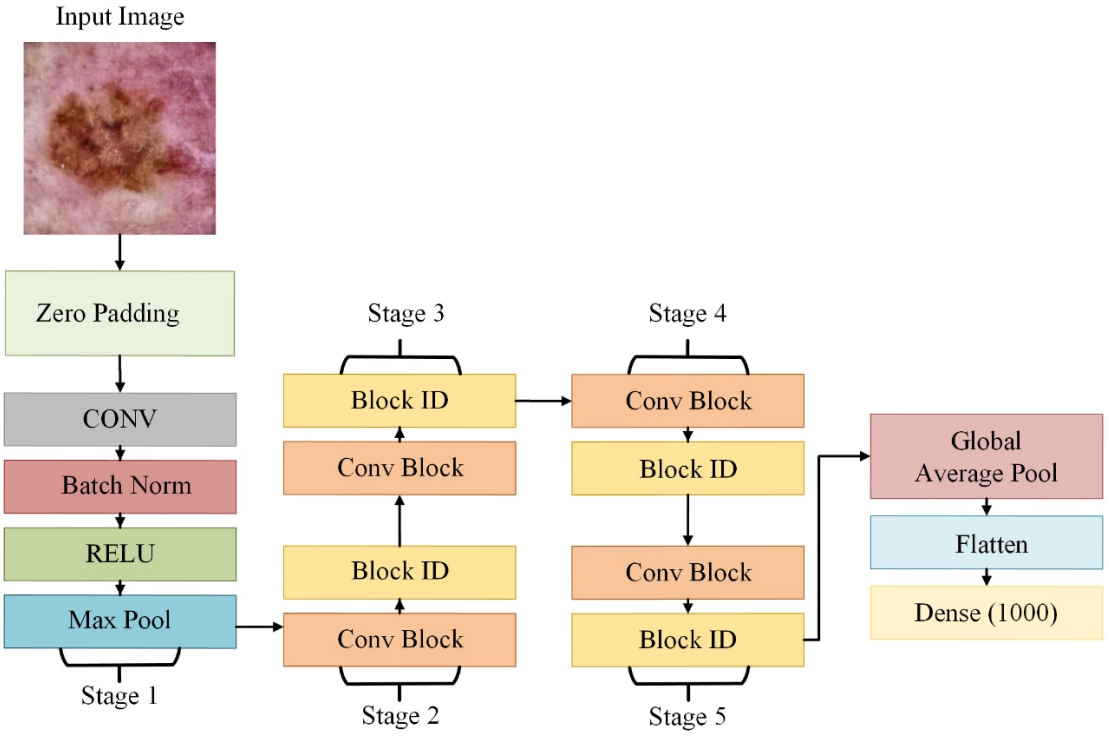


FIGURE 3.28: Architecture of ResNet-50 Model

To further optimize performance, batch normalization and dropout layers are integrated to reduce overfitting and stabilize training. ResNet50's modular design

allows seamless integration into ensemble pipelines, complementing other architectures with its depth and feature richness. Its ability to preserve gradient flow across layers makes it particularly effective for learning subtle lesion characteristics in high-resolution dermoscopic images. Let the $x \in \mathbb{R}^{H\times W\times C}$ be an input dermatoscopic image. The ResNet-50 architecture processes x through a series of convolutional blocks with residual connections, forming a deep feature extractor $f_{\text{ResNet}}(x;\theta)$ with pretrained parameters θ. Each residual block computes:

$$\mathbf{y} = \mathcal{F}(\mathbf{x}, \{W_i\}) + \mathbf{x} \tag{3.39}$$

Where $\mathbf{x}$ is the input to the block, $\mathcal{F}(\mathbf{x}, \{W_i\})$ represents the residual mapping learned by stacked convolutional layers with weights $\{W_i\}$, and the addition denotes the identity shortcut connection that stabilizes gradient flow by using Equation 3.39. The final classifier head $g(\cdot;\phi)$ replaces the original fully connected layer and outputs logits $z \in \mathbb{R}^K$, where $K = 7$ is the number of lesion classes:

$$z = g(f_{\text{ResNet}}(x;\theta);\phi) \tag{3.40}$$

The SoftMax function converts logits into class probabilities:

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}, \quad i = 1, \ldots, K \tag{3.41}$$

The categorical cross-entropy loss for a true label $y \in \{1, \ldots, K\}$ is defined as in Equation 3.42:

$$\mathcal{L} = -\log(p_y) \tag{3.42}$$

To calculate gradients for backpropagation, ensuring effective learning of deep lesion features during backpropagation, utilize Equation 4.3:

$$\frac{\partial \mathcal{L}}{\partial z_i} = p_i - \delta_{iy} \tag{3.43}$$

Where δ_{iy} is the Kronecker delta, equal to 1 if $i = y$ and 0 otherwise. These gradients propagate efficiently through residual paths, enabling robust learning of both global and local lesion features. The deep architecture supports fine-grained classification of visually similar lesions, enhancing diagnostic precision. Its integration into the pipeline ensures high-capacity modeling without compromising convergence stability. ResNet50 leverages hierarchical feature abstraction and a deep receptive field to capture both global and fine-grained lesion patterns. Transfer learning enhances accuracy and generalization across datasets. Its stable validation performance supports reliable clinical-grade skin lesion analysis.

3.6.5 Densenet-121 Branch

DenseNet121 is a densely connected convolutional neural network in which every layer gets input from every layer before it within a dense block. This architecture promotes feature reuse, improves gradient flow, and reduces the number of parameters, making training more efficient and effective. In our methodology, DenseNet121 is initialized with ImageNet weights and fine-tuned using images of dermoscopic skin lesions. It replaces the traditional fully connected classifier with a global average pooling layer, followed by a dense layer with seven softmax-activated outputs. The dense connectivity ensures that even the shallow layers can receive gradient signals from the deeper ones, which enhances the learning of fine-grained lesion features. This makes DenseNet121 particularly useful for detecting small or ambiguous lesion patterns that might be overlooked by shallower architectures. Figure 3.29 displays the complete model layout along with parameter distribution [66]. Its layered concatenation mechanism strengthens multi-scale feature representation across lesion boundaries. This structural efficiency supports consistent classification performance even in cases with high intra-class variability.

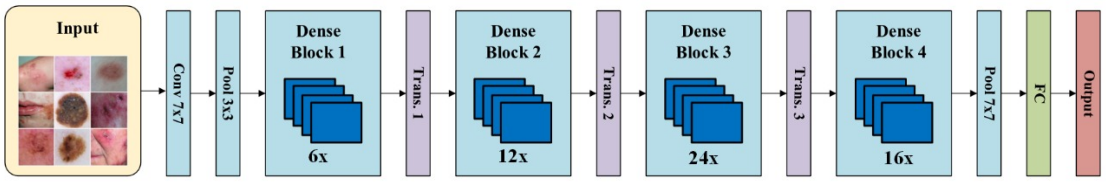


FIGURE 3.29: Architecture of DenseNet-121 Branch Model

3.6.5.1 DenseNet-121 Architecture Components

1. Input layer

   The input layer accepts a dermoscopic image with a fixed spatial resolution of 224×224 pixels and three color channels (RGB), denoted mathematically as in Equation 3.44:

   $$X \in \mathbb{R}^{224 \times 224 \times 3} \tag{3.44}$$

   where X represents the input image tensor.

2. Convolutional Layers

   The initial convolutional operation is performed using a filter bank of size $f \times f$ with stride s and padding$p$, typically $f = 7$, $s = 2$, and $p = 3$. This is followed by Batch Normalization (BN) and ReLU activation. The convolutional output Z is computed as in Equation 3.45 and Equation 3.46:

   $$Z = \text{ReLU}(\text{BN}(W * X + b)) \tag{3.45}$$

   $$\text{ReLU}(x) = \max(0, x) \tag{3.46}$$

   In convolutional neural networks, the convolution operation $(*)$ extracts spatial features using learnable filter weights (W_m) and biases (b). Batch Normalization (BN) normalizes the output to improve training stability and speed. The Rectified Linear Unit (ReLU) introduces non-linearity by allowing positive values to pass while zeroing out negatives, enabling efficient learning and mitigating the vanishing gradient problem.

3. Dense Blocks

   Net-121 consists of four dense blocks. The output of one convolutional layer is merged with all of the inputs of the subsequent layers in each dense block. This style of architecture, which has dense connections, reuses features by the direct connection of each layer to the other. Mathematically, the output of the layer can be represented as follows in Equation 3.47:

$$n_t = D_t \cdot [n_0, n_1, n_{k-1}] \tag{3.47}$$

where D_t symbolizes a combination of processes like Batch Normalization (BN), ReLU (Rectified Linear Unit) activation, and convolution, and n_0, n_1, n_{k-1} are the outputs of all earlier layers concatenated. Such thick connections provide efficient sharing of features and reduce the effect of vanishing gradient and allow the network to acquire fine-grained characteristics that are crucial in differentiating between skin problems that are comparable. The notation $[\cdot]$ signifies the concatenation of feature maps from previous layers, enabling feature reuse and improving gradient flow.

This dense connectivity pattern promotes implicit deep supervision, as each layer receives direct feedback from all preceding layers. It also encourages parameter efficiency by reducing redundancy in learned features, making the network compact yet powerful. In the context of skin lesion classification, DenseNet-121 excels at capturing subtle texture variations and boundary irregularities, which are essential for distinguishing between visually similar dermatological conditions. Its feature reuse strategy also minimizes overfitting, enhancing generalization across diverse lesion presentations.

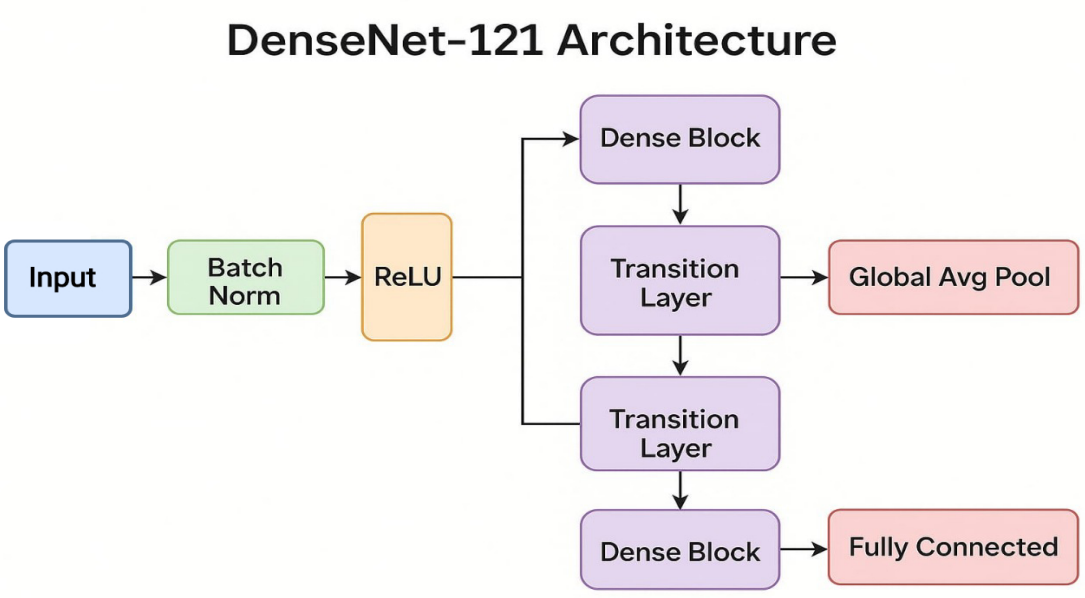


FIGURE 3.30: Flow Diagram of DenseNet121 Branch Architecture

Table 3.4: Architecture Comparison of Different DenseNet Variants

Layers	**Output Size**	**DenseNet-121**	**DenseNet-169**	**DenseNet-201**	**DenseNet-264**
Convolution	112 × 112	7 × 7 conv, stride 2			
Pooling	56 × 56	3 × 3 max pool, stride 2			
Dense Block (1)	56 × 56	$[1\times1,\ 3\times3]\times6$	$[1\times1,\ 3\times3]\times6$	$[1\times1,\ 3\times3]\times6$	$[1\times1,\ 3\times3]\times6$
Transition Layer (1)	56 × 56	1 × 1 conv			
	28 × 28	2 × 2 average pool, stride 2			
Dense Block (2)	28 × 28	$[1\times1,\ 3\times3]\times12$	$[1\times1,\ 3\times3]\times12$	$[1\times1,\ 3\times3]\times12$	$[1\times1,\ 3\times3]\times12$
Transition Layer (2)	28 × 28	1 × 1 conv			
	14 × 14	2 × 2 average pool, stride 2			
Dense Block (3)	14 × 14	$[1\times1,\ 3\times3]\times24$	$[1\times1,\ 3\times3]\times32$	$[1\times1,\ 3\times3]\times48$	$[1\times1,\ 3\times3]\times64$
Transition Layer (3)	14 × 14	1 × 1 conv			
	7 × 7	2 × 2 average pool, stride 2			
Dense Block (4)	7 × 7	$[1\times1,\ 3\times3]\times16$	$[1\times1,\ 3\times3]\times32$	$[1\times1,\ 3\times3]\times32$	$[1\times1,\ 3\times3]\times48$
Classification Layer	1 × 1	7 × 7 global average pool			
		1000D fully-connected, softmax			

DenseNet-121 is an algorithm named after its 121 layers, evenly composed of convolutional, normalization, and activation layers paired with dense and transition blocks, as shown in Table 3.4. DenseNet-121 works as a shallow substitute to classical ResNet architectures to strike a compromise between the richness of features and performance efficiencies. It starts with a 2D convolutional layer with 7×7 the kernel and stride being 2; ReLU and batch normalization are used to add non-linearity and normalize output. The feature maps are then further downsampled by a layer of max pooling that is set to 2 with a kernel size of 3×3.

Comprising the architecture is the core, which consists of three progressive stages with two residual blocks that are made up of two 3×3 convolutional layers. It includes shortcut connections, which are added during construction as identity mappings or as projection shortcuts through 1×1 convolutions when dimension matching is necessary. In short, the global average pooling layer summarizes the final convolutional activations into a small feature vector appropriate to perform follow-up classification.

4. Bottleneck Layer

   A bottleneck layer consists of a sequence of batch normalization, ReLU activation, 1×1 convolution to decrease the channels of the feature map to reduce computation, another batch normalization ReLU activation, and 3×3 convolution to learn the spatial information. This design enhances efficiency by narrowing the dimensions of the features before the weightier convolution and, when combined with the dense connectivity of DenseNets that concatenate all precedent layer outputs, encourages reuse of features and greater gradient flow. Such architecture allows the model to learn complex patterns in the images of skin lesions using fewer parameters, which makes it extremely appropriate for skin lesion classification with high accuracy.

5. Transition Layers

   Transition layers in DenseNet architectures are critical for maintaining computational efficiency and controlling model complexity as the network deepens through successive dense blocks. Positioned between these blocks, each

transition layer performs a sequence of operations designed to compress and downsample the feature maps. Specifically, the input feature map F first undergoes batch normalization, which stabilizes and accelerates training by normalizing the distribution of activations. This is followed by a 1×1 convolution that reduces the number of channels, effectively compressing the feature representation without altering spatial dimensions. Finally, a 2×2 average pooling operation with stride 2 is applied, halving the spatial resolution and enabling the network to capture broader contextual information while reducing computational load. Mathematically, if F denotes the input feature map, the transformation can be expressed as in Equation 3.48:

$$T = \text{AvgPool}_{2\times2}(\text{Conv}_{1\times1}(\text{BN}(F))) \tag{3.48}$$

where T denotes the output feature map passed to the next dense block. By combining channel reduction and spatial downsampling, transition layers prevent feature map explosion, promote efficient gradient flow, and enhance generalization, making them especially valuable in medical imaging tasks where model interpretability and performance on limited data are critical.

6. Global Average Pooling

   To acquire a feature vector of fixed size from the final feature maps, the application of global average pooling is used in all spatial dimensions. This operation computes the average of each feature map independently. For a given channel c, the pooled value is calculated as shown in Equation 3.49:

$$g_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_{c,i,j} \tag{3.49}$$

   where $F_{c,i,j}$ represents the feature map value at spatial location (i, j) in channel c, and H and W are the height and width of the feature map, respectively. This operation results in a feature vector $g \in \mathbb{R}^C$, where C is the total number of channels.

7. FC Layer (Classification Layer)

A dense layer that is fully connected receives the output from the global average pooling layer to compute class probabilities through the softmax activation function. This is mathematically expressed as shown in the Equation 3.50 :

$$\hat{t}_k = \frac{e^{z_k}}{\sum_{j=1}^{K} e^{z_j}}, \quad \forall k \in \{1, 2, \ldots, K\} \tag{3.50}$$

where t_m is the predicted probability of class m, and m is the total number of output classes. The input to the softmax function is computed as mentioned in Equation 3.51:

$$z = W_{\text{mfc}} \cdot g + b_{\text{fc}} \tag{3.51}$$

with g being the feature vector that emerges from the global average pooling layer and representing the weights W_{mfc} and biases b_{fc} of the fully connected layer. This structure allows the model to produce a scaled probability distribution among the target classes. It effectively learns spatial hierarchies and local patterns of the input data, forming a powerful foundation for low-to-mid-level feature extraction.

3.6.6 EfficientNetB0 Branch

EfficientNetB0 has been chosen as the second backbone because of its parameter efficiency and solid performance on the classification benchmarks. The model is imported with pre-trained ImageNet weights and used as a fixed feature extractor by excluding the top classification layers (`Exclude=True`) and applying global average pooling to its output. The input images are fed directly to this model, with the same images fed to the component ResNet 21 contributing to the extraction of complementary features.

EfficientNetB0 uses compound scaling (scaling depth, width, and resolution in a balanced way), which allows it to learn highly generalized features with fewer parameters. The architecture as shown in Figure 3.31 is computational efficiency

oriented, achieving a trade-off between depth, width, and resolution through the use of compound scaling. In Equation 3.52, the scaling technique should be interpreted as follows:

$$\text{depth: } \text{dep} = \alpha^{\phi}, \quad \text{width: } \text{wid} = \beta^{\phi}, \quad \text{resolution: } \text{res} = \gamma^{\phi} \tag{3.52}$$

Where the scaling coefficient ϕ is defined by the user, and the three constants, alpha (α), beta (β), and gamma (γ) are chosen through a grid search. This enables EfficientNet-B0 to be a small but still high-level configuration.

In this study, EfficientNetB0 complements deeper models with fast inference and strong baseline performance, leveraging bottleneck and excitation modules for focused feature extraction. Its compact design supports rapid experimentation and smooth ensemble integration. The hybrid model integrates DenseNet121 and EfficientNetB0 to utilize the feature of DenseNet121 being able to recognize patterns in detail and the efficiency of EfficientNetB0, making the hybrid model a scalable and precise model that can be used in real-time environments. This dual-stream architecture enables simultaneous extraction of fine-grained and abstract features, enhancing the model's ability to generalize across lesion types. The fusion layer aggregates complementary representations, improving classification robustness without incurring significant computational overhead. Such integration is particularly beneficial for clinical workflows requiring both accuracy and responsiveness. This architecture facilitates modular experimentation and targeted refinement without retraining the full pipeline, making it adaptable for iterative research. Its compatibility with mobile and embedded systems enables seamless deployment in teledermatology environments. By combining diagnostic precision with low-latency inference, the hybrid design ensures responsiveness across varied hardware configurations. EfficientNetB0 contributes speed and parameter efficiency, while DenseNet121 enhances pattern recognition, resulting in a scalable solution for real-time skin lesion classification. This integration ensures consistent performance across clinical devices and simplifies deployment workflows. Its modular design supports efficient updates and ongoing refinement with new lesion data.

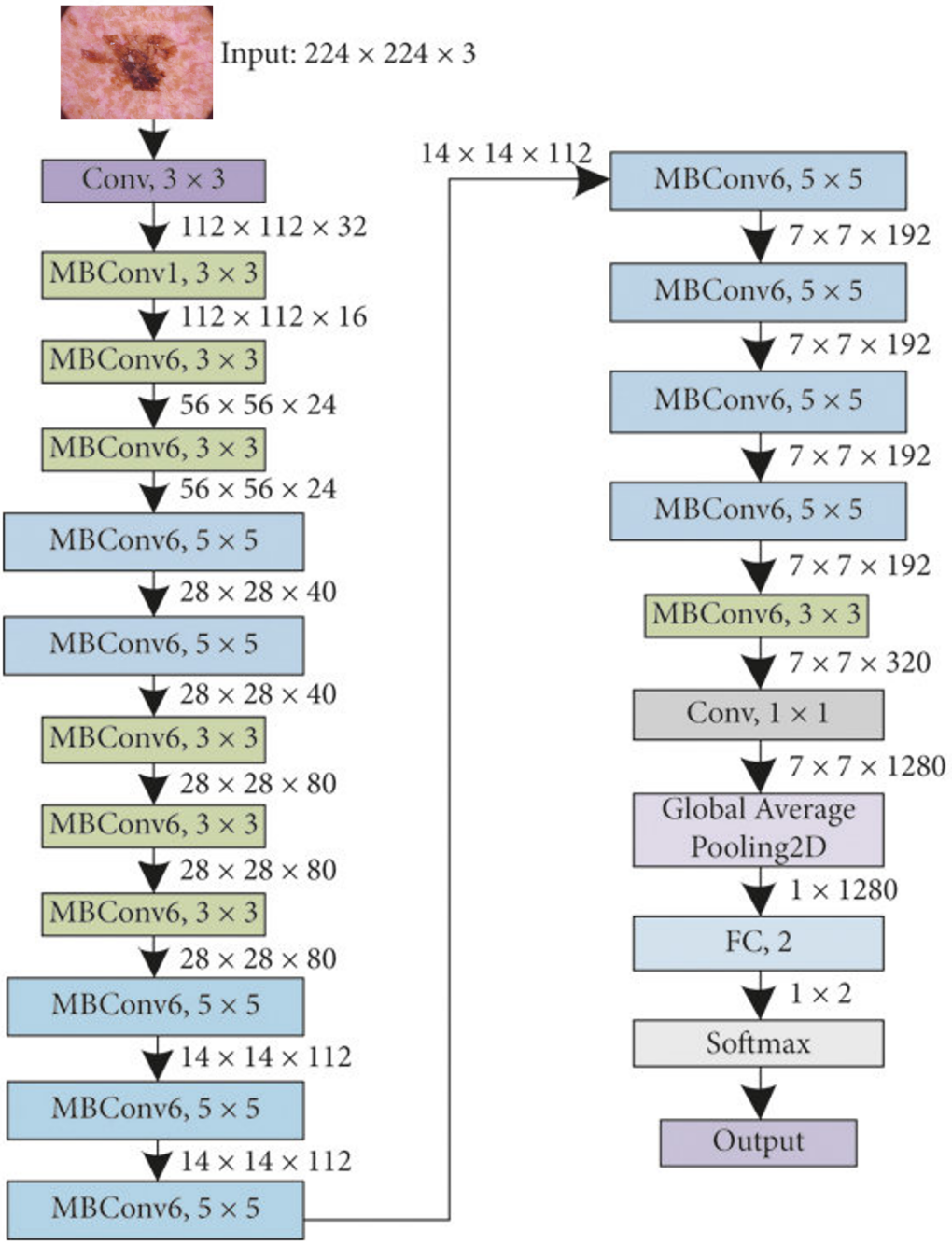


FIGURE 3.31: Architecture of EfficientNet-B0

3.6.7 Hybrid Feature Fusion and Classification

The feature vectors of two branches, which are ResNet21 and EfficientNetB0 global average pooled vectors units, are concatenated with a `Concatenate` layer. Such a mechanism of fusion enables the hybrid model to integrate both fine-grained

and abstract information extracted in both networks. Once the feature extraction process is applied, the last layers of both DenseNet121 and EfficientNetB0 are concatenated as follows in Equation 3.53:

$$\text{Feature}_{\text{concat}} = [\text{Feature}_{\text{Dense}}, \text{Feature}_{\text{Efficient}}] \tag{3.53}$$

This process involves the integration of various in-depth and top-level features into one overall reflection. To finalize this combined set of features, we added:

3.6.7.1 Dense Layer

A 256-unit layer that is fully connected:

$$y_i = \lambda(wx + b) \tag{3.54}$$

The weight matrix is denoted by w, the input feature vector by y_i, the bias term by b, and the ReLU activation function by λ, which introduces non-linearity into the model. This non-linear transformation enables the network to learn complex decision boundaries and interactions among features that linear operations alone cannot capture. The dense layer serves as a bridge between convolutional outputs and the final SoftMax classifier, aggregating spatial and contextual information into a unified format. This layer deals with how the feature combination interacts, as shown in Equation 3.54.

3.6.7.2 Dropout Layer

The rate of regularization is 0.5:

$$y_i = \begin{cases} \dfrac{x_i}{\epsilon} & \text{if the neuron is kept} \\ 0 & \text{otherwise} \end{cases} \tag{3.55}$$

where, in Equation 3.55, the probability of a neuron being retained throughout training is denoted by ϵ.

3.6.7.3 SoftMax Layer

It serves as the final output layer for multi-class classification tasks, converting raw logits into normalized probability scores, as shown in 3.56.

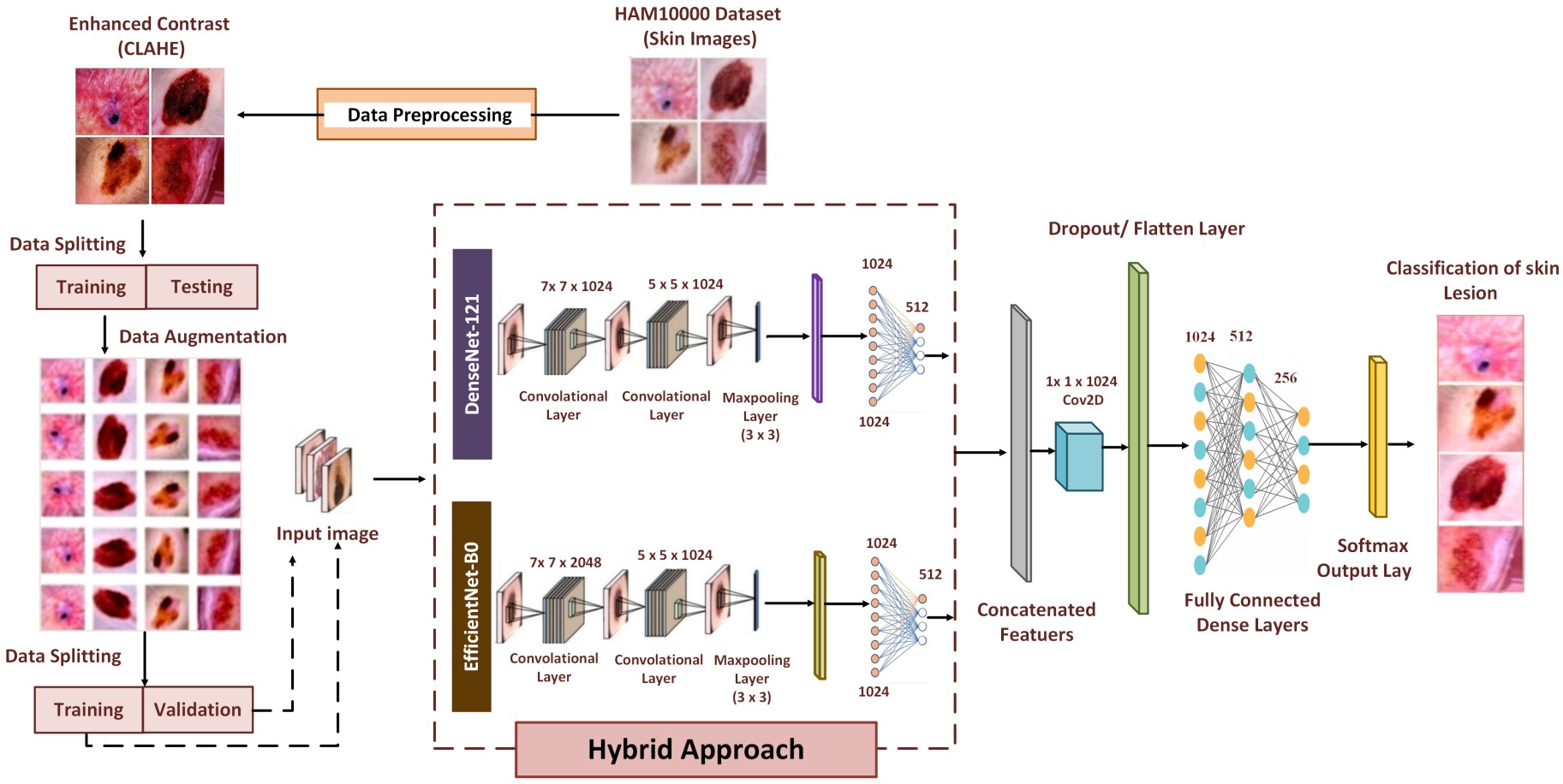
$$P(y = k \mid x) = \frac{e^{z_k}}{\sum_{m=1}^{K} e^{z_m}} \tag{3.56}$$

where z_k is the logit of the class k, and K is the total number of classes (7 classes in this case). This formulation ensures that the output values lie in the range [0, 1] and sum to 1, making them interpretable as class probabilities. The layers ensure that the model can learn non-linear decision boundaries and feature interactions and predict the probability of each class, thus allowing confident predictions. It is particularly effective in medical image classification, where probabilistic outputs support clinical decision-making and allow threshold-based tuning for sensitivity and specificity.

Following the concatenation:

- A fully connected (dense) layer with 256 units and ReLU activation is applied.
- Dropout layers with a learning rate of 0.5 are used after each dense layer to prevent overfitting.
- A second dense layer with 128 units and ReLU activation further refines the combined features, enhancing non-linear transformations. This layer aids in capturing complex patterns and improving the model's capacity to differentiate subtle lesion characteristics.
- The final output layer is a SoftMax-activated dense layer whose units equal the number of diagnostic classes (7 in this case), providing class probabilities for each input image.

FIGURE 3.32: Proposed Framework



3.6.8 Experimental Setup and Metrics for Evaluation

The given model is designed using Python 3.12, OpenCV 4.11, and the Keras library 3.10 to facilitate model development and image processing. It is implemented in a development environment based on Windows 11 Pro and utilizes the following hardware specifications include 16 GB RAM, an Nvidia RTX 2060 GPU, and an Intel i5 processor running at 2.9 GHz. To evaluate the model's performance, we employed several commonly used metrics, including F1 score, recall, accuracy, and precision.

These metrics provide a comprehensive analysis of the model's accuracy and reliability, which is essential for multi-class classification tasks in skin disease prediction. To rigorously assess the performance of the proposed skin lesion classification model, this study employs stratified 5-fold cross-validation. This technique ensures that each fold preserves the original class distribution, which is essential for handling the inherent imbalance in datasets like HAM10000. By rotating training and validation splits across five folds, the model demonstrates its generalization capability across diverse subsets, reducing the risk of bias and overfitting.

During each fold, confusion matrices are generated to visualize misclassification patterns and quantify class-wise sensitivity and specificity. These matrices reveal which lesion types are frequently confused, guiding iterative improvements in model design. Particular emphasis is placed on precision and recall for minority classes such as dermatofibroma and vascular lesions, which are underrepresented and more susceptible to misclassification. This targeted evaluation promotes diagnostic fairness and reduces bias toward dominant categories like melanocytic nevi. To address class imbalance, macro-averaged F1 scores are computed, treating each class equally regardless of its frequency. This metric provides a balanced overview of classification performance across all lesion types. Additionally, Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) scores are calculated for each class to assess the model's discriminative ability across varying decision thresholds. These metrics are particularly valuable for evaluating sensitivity-specificity trade-offs in clinical applications. All experiments are

conducted within a reproducible pipeline that incorporates fixed random seeds, standardized preprocessing routines, and consistent augmentation strategies. The evaluation framework thus delivers a robust, transparent, and clinically relevant assessment of model performance in multi-class skin lesion classification. Furthermore, stratified k-fold cross-validation is employed to ensure balanced representation of lesion classes across training and validation splits. Performance metrics are aggregated and visualized to highlight per-class variability and overall diagnostic reliability.

3.6.9 Hybrid Model Compilation and End-to-End Training Strategy

The Adam (Adaptive Moment Estimation) optimizer, whose learning rate is 0.0001, is used to build our model. The loss function that is selected is the categorical cross-entropy loss, which is suitable for multi-class classification tasks. As a measure of performance, accuracy is employed. To ensure efficient learning, the initial layers of DenseNet121 and EfficientNetB0 are frozen during the first training phase to maintain the general features learned from ImageNet. This can be stated in Equation 3.57 and Equation 3.58 as follows:

$$\text{Frozen layers: } Weight^{\text{pre-trained}} \tag{3.57}$$

$$\text{Fine-tuned layers: } Weight^{\text{trainable}} \tag{3.58}$$

During fine-tuning, the model can effectively adjust to the dermatological domain by freezing these layers. Figure 3.32 depicts the architecture of the proposed model. This strategy preserves pre-trained general features while allowing domain-specific layers to specialize in lesion classification. This staged training strategy mitigates early overfitting, preserves low-level features, and enables domain-specific adaptation through selective unfreezing, ultimately enhancing classification performance on the HAM10000 dataset. Once the domain-specific layers stabilize, selective

unfreezing is performed to allow deeper adaptation and refinement of shared representations. This staged training approach balances generalization and specificity, enabling the model to learn discriminative features without compromising robustness.

To further enhance generalization and reduce overfitting, data augmentation techniques such as rotation, flipping, and zooming are applied to the input images. These transformations simulate real-world variability in dermoscopic imaging and help the model learn invariant features. The training process is conducted over 100 epochs, with early stopping and learning rate scheduling integrated to optimize convergence and prevent unnecessary computation. Batch normalization and dropout layers are also incorporated to stabilize gradient flow and improve regularization. The final classification layer uses a softmax activation function to output probability scores for each of the seven lesion categories in the HAM10000 dataset, enabling multi-class prediction with calibrated confidence levels. These regularization strategies collectively enhance model robustness and reduce sensitivity to noise in dermoscopic inputs. The use of early stopping prevents overfitting by halting training once validation performance plateaus. Additionally, adaptive learning rate scheduling ensures efficient convergence across varying training dynamics.

To ensure consistent input dimensions and pixel scaling, all images are resized to 224×224 and normalized to the [0,1] range. The model is trained using mini-batches of size 32, which balances memory efficiency and gradient stability. During training, the validation loss is monitored to trigger early stopping if no improvement is observed over 10 consecutive epochs.

A ReduceLROnPlateau scheduler is employed to dynamically adjust the learning rate when the validation performance plateaus, allowing finer convergence in later stages. The architecture is implemented using TensorFlow and Keras, with GPU acceleration enabled to expedite training. All experiments are conducted within a reproducible pipeline using fixed random seeds and standardized preprocessing routines. Additional logging and checkpoints ensure seamless recovery and consistent performance monitoring. This setup ensures both experimental transparency and scalability for future clinical deployment.

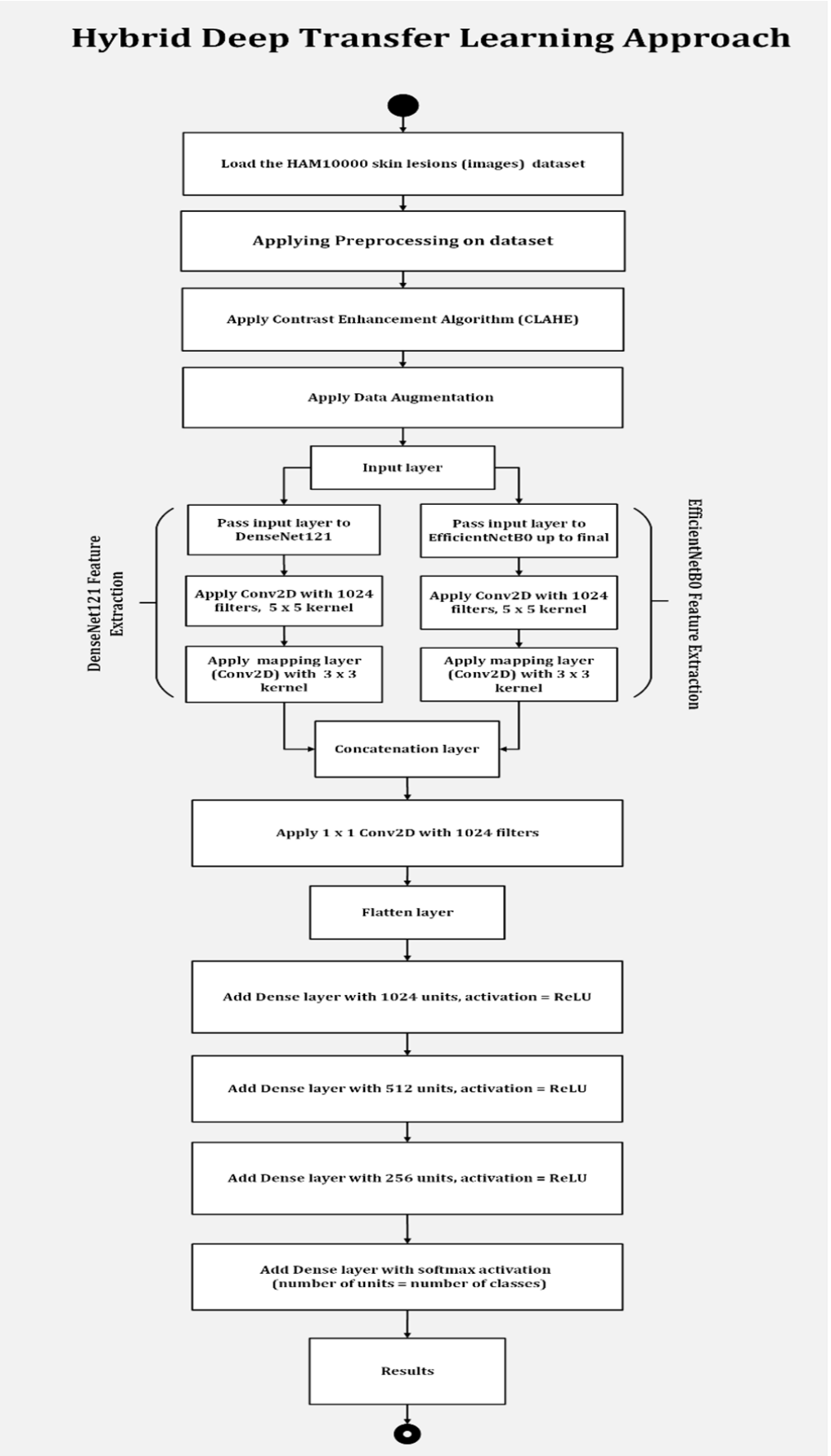


FIGURE 3.33: Flowchart of hybrid deep transfer learning approach

Algorithm 1 Hybrid Deep Transfer Learning Approach (DenseNet121 + EfficientNetB0) for Recognition of Skin Lesions

Input: HAM10000 Skin Lesion Images
Output: Skin lesion class prediction report

Input layer:
input_layer ← Create an input layer with shape (224, 224, 3); // Load images with standard input size for both models

Feature Extraction Branches:
DenseNet121 Feature Extractor:
dense_input ← Pass input_layer to DenseNet121;
dense_block_output ← Extract features from DenseNet121 up to final dense block;
dense_conv ← Apply Conv2D with 1024 filters, 5x5 kernel, ReLU activation to dense_block_output;
dense_map ← Apply mapping layer (Conv2D) with 3x3 kernel to dense_conv;
EfficientNetB0 Feature Extractor:
eff_input ← Pass input_layer to EfficientNetB0;
eff_block_output ← Extract features from EfficientNetB0 up to final MBConv block;
eff_conv ← Apply Conv2D with 1024 filters, 5x5 kernel, ReLU activation to eff_block_output;
eff_map ← Apply mapping layer (Conv2D) with 3x3 kernel to eff_conv;

Feature Fusion:
merged_features ← Concatenate dense_map and eff_map along depth axis;
unified_features ← Apply 1x1 Conv2D with 1024 filters to merged_features;
flatten_layer ← Flatten unified_features;

Fully Connected Dense Layers:
fc1 ← Add Dense layer with 1024 units, activation = ReLU to flatten_layer;
fc2 ← Add Dense layer with 512 units, activation = ReLU to fc1;
fc3 ← Add Dense layer with 256 units, activation = ReLU to fc2;

Output Layer:
output_layer ← Add Dense layer with softmax activation (number of units = number of classes) to fc3;

Model Compilation:
model ← Compile with Adam optimizer, categorical crossentropy loss, and accuracy metric;
Train the Model:
Train model on training dataset with validation split;
Prediction for the Results:
X_test_array ← Convert test images into NumPy array;
X_test_reshaped ← Reshape test array to (batch_size, 224, 224, 3);
y_pred ← Predict class probabilities using model on X_test_reshaped;

The architecture presented in Figure 3.33 illustrates a hybrid deep learning pipeline of multiclass skin lesion classification and is tested on images of the HAM10000 database subject to CLAHE to improve contrast and feature enhancement. The process of feature extraction is performed simultaneously through DenseNet121 and EfficientNetB0, and then through convolutional mapping and merging layers, synthesis of spatial representations of lesions is implemented. ReLU activation is used in a fully connected block for the classification process, and a softmax produces probability-based outputs that make the predictions comprehensible and reliable.

The HAM10000 dataset, which consists of 10,015 dermatoscopic images in seven lesion categories, is used in this study to demonstrate a strong deep learning framework for automated skin lesion recognition using a hybrid model that combines DenseNet121 and EfficientNetB0. The approach includes a comprehensive pipeline that includes preprocessing (noise reduction, hair artifact removal using Dull Razor, color normalization, and CLAHE-based contrast enhancement), data augmentation (rotation, scaling, and synthetic image generation) to rectify class imbalance, and a dual-branch architecture for feature extraction and fusion that is optimized from pre-trained weights in ImageNet. To improve boundary detection, a placeholder U-Net model is being used to simulate lesion segmentation. Full integration is planned. In order to assess image quality and classification performance (accuracy, precision, recall, and F1-score), the model, which is developed on Google Colab using TensorFlow and Keras, uses early stopping, batch normalization, and measures including SSIM (0.8693), PSNR (28.73 dB), CIH (0.9430), and AMBE. By filling in gaps in the literature, this method increases diagnostic precision and generality. Future research will focus on improving segmentation and robustness across a range of situations.

Chapter 4

Experiments and Results

This part explains how to set up a demonstration and what tools are needed. This chapter discusses the datasets that are used and where the data came from. This part additionally describes the methods and instruments utilized in the evaluation phase. The last parts include the results and an observation.

4.1 Tools and Technology

This study uses a variety of techniques and technologies to improve image analysis and preprocessing. The Dull-Razor algorithm for artifact removal, Gaussian and median filtering, LAB color normalization, Contrast Limited Adaptive Histogram Equalization (CLAHE), and other important preprocessing methods are implemented using OpenCV and PIL. Imgaug is used for data augmentation to increase the robustness of the dataset, and metrics like the Structural Similarity Index (SSIM), Peak Signal-to-Noise Ratio (PSNR), Absolute Mean Brightness Error (AMBE), and Contrast Improvement Index (CII) are used to objectively evaluate picture quality. The pipeline is designed to be modular and reproducible, supporting scalable experimentation across datasets. Each preprocessing step contributes to enhancing diagnostic clarity and consistency in medical imaging. Benchmarking results are integrated into publication-ready formats for academic dissemination and peer review.

4.1.1 Google Colab

Google Colab, often known as the "Collaboratory," is a free cloud-based platform that Google provides for using Jupyter notebooks. It gives us the ability to build and run Python programs in an online setting. Colab is a well-liked option for deep learning and data analysis applications since it gives users access to GPU and TPU resources.

4.1.2 MS Excel

Microsoft Excel is a popular spreadsheet program that enables users to manage tabular data, make charts, and conduct data analysis. It is frequently employed for simple data visualization and processing applications. It also enables advanced analysis through functions, pivot tables, and conditional formatting. Built-in formulas and macros simplify repetitive tasks, while seamless integration with Office tools enhances productivity across various workflows.

4.1.3 Python

Known for its clarity and simplicity, Python is a high-degree, versatile computer programming language. It offers an extensive ecosystem of libraries and frameworks for many kinds of applications, like web development, machine learning, and data analysis. Python's readable syntax and broad library support make it ideal for rapid development across domains. It handles tasks from data analysis to machine learning with minimal overhead. Its cross-platform compatibility and active community further enhance its versatility.

4.1.4 TensorFlow and Keras Libraries

TensorFlow is a Google open-source machine learning framework. It is commonly used to create and train deep learning models. Keras is frequently used in conjunction with TensorFlow for the construction and training of neural networks.

It supports a wide range of applications, including image recognition, natural language processing, and time-series forecasting. TensorFlow's scalability and integration with cloud platforms make it ideal for both research and production environments.

4.1.5 Sikit-learn Library

Sci-Kit Learn is a prominent Python deep learning package. It includes a variety of tools for applications, including classification, regression, clustering, dimensionality reduction, and model selection. It's a powerful library for building and evaluating machine learning models.

4.1.6 NumPy

NumPy is a key Python package for computational science. It allows you to work with big, multi-dimensional arrays and matrices. It is an essential library for numerical and mathematical operations. NumPy's integration with other scientific libraries makes it foundational for data-intensive applications and machine learning workflows.

4.1.7 Pandas

Pandas is a popular Python data manipulation and analysis toolkit. It provides data structures and functions for working with structured data efficiently, such as data frames and series. Pandas is a popular data cleaning, exploration, and transformation tool.

4.1.8 Imgaug

This library is intended for performing various image augmentation techniques to increase the size and diversity of the training dataset. It helps improve model

generalization by simulating real-world variations such as rotation, scaling, and noise. These augmentations are especially valuable in medical imaging, where annotated data is often limited and costly to obtain.

4.1.9 Google Drive

This is used to store the dataset files and the processed augmented images, allowing the notebook to access them directly. It ensures organized data management and facilitates seamless integration with training workflows. This setup also simplifies batch loading and preprocessing during model training and evaluation

4.1.10 Matplotlib

This library is used for creating static, interactive, animated visualizations in Python, primarily for plotting images, distributions, and sample images. It supports a wide range of plot types, including line graphs, histograms, scatter plots, and heatmaps.

4.1.11 OpenCV (cv2)

It is a powerful library for computer vision tasks, used here for image preprocessing techniques such as applying filters, removing hair-like structures (dull razor), and managing color spaces. It also facilitates operations like edge detection, morphological transformations, and image resizing, which are essential for preparing dermoscopic data.

4.1.12 Kaggle Notebook

The Kaggle Notebook Editor is a cloud-based computing environment that combines hardware acceleration, data set access, and code execution in an interactive interface to support replicable data analysis and machine learning workflows.

4.1.13 Microsoft Visio

Microsoft Visio is a tool for creating professional diagrams like flowcharts, org charts, and network diagrams with advanced features, templates, and integration with Microsoft 365, ideal for business and technical users. It supports real-time collaboration, allowing multiple users to work on diagrams simultaneously across devices. Microsoft Visio also enables data linking, turning static visuals into dynamic dashboards connected to Excel, SQL Server, and other sources.

4.1.14 Draw.io

This is a web-based free diagramming tool with a user-friendly interface, offering different templates and seamless integration with cloud storage platforms such as Google Drive and OneDrive, suitable for both casual and professional use. Its real-time editing capabilities and export options enhance productivity across diverse project workflows.

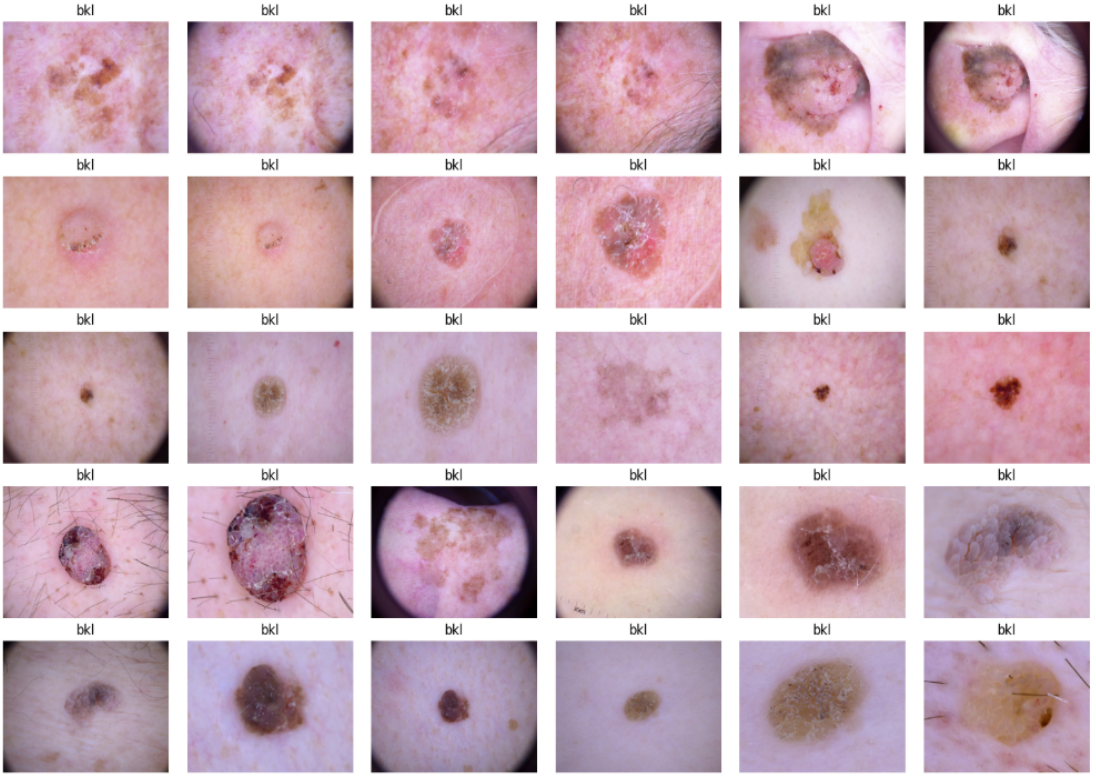


FIGURE 4.1: Samples of skin lesion images

4.2 Skin Lesions Dataset

In our study, we employed the HAM10000 dataset, which contains 10,015 dermatoscopic images of pigmented skin lesions. The metadata file (HAM10000-metadata.csv) provides essential clinical and diagnostic information for each image, as illustrated in Figure 4.1. The dataset consists of attributes such as image ID, lesion ID, diagnosis, diagnosis type, age, gender, and anatomical localization. These attributes are crucial for supervised classification tasks, multimodal learning, and lesion-specific analysis.

Table 4.1: HAM10000 Metadata Attributes

No.	**Attributes**	**S**	**Type**	**Instances / Description**
1	Image ID	ID	String	Unique identifier for each dermatoscopic image (10,015 total)
2	Lesion ID	LID	String	Identifier for lesion group; multiple images may share the same lesion ID
3	Diagnosis	DX	Categorical	7 classes: nv, mel, bkl, bcc, akiec, vasc, df
4	Diagnosis Type	DXT	Categorical	Source of diagnosis: histopathology, consensus, follow-up, or confocal
5	Age	A	Numeric	Patient age in years; some entries may be missing
6	Gender	S	Categorical	Gender of patient: male, female, or unknown
7	Localization	LOC	Categorical	Anatomical site of lesion: e.g., back, lower extremity, trunk, scalp, face, etc.

The dataset includes seven diagnostic categories: melanocytic nevi (NV), melanoma (MEL), benign keratosis-like lesions (BKL), basal cell carcinoma (BCC), actinic keratoses (AKIEC), vascular lesions (VASC), and dermatofibroma (DF). Each attribute with a comprehensive explanation is presented in table 4.1. The dataset is composed of 6705 images of melanocytic nevi (NV), 1113 of melanoma (MEL), 1099 of benign keratosis (BKL) lesions, 514 of basal cell carcinoma (BCC), 327 of actinic keratoses (AKIEC), 142 of vascular lesions (VASC), and 115 of dermatofibroma (DF). The metadata enables stratified analysis, class balancing, and integration of clinical features into deep learning pipelines.

4.2.1 Quantitative Evaluation of CLAHE Preprocessing on Skin Lesion Images

To assess the effectiveness of the Contrast Limited Adaptive Histogram Equalization (CLAHE) preprocessing technique applied to the HAM10000 skin lesion dataset, four key image quality metrics are computed: SSIM, PSNR, CII, and AMBE, as displayed in Figure 4.2. These metrics provide insights into structural fidelity, noise suppression, contrast improvement, and brightness preservation, respectively. By analyzing these metrics across the dataset, researchers can validate the preprocessing impact on image clarity and diagnostic reliability.

4.2.1.1 Structural Similarity Index

The SSIM value achieved is 0.8693, which approaches the upper bound of 1.0. This indicates high structural similarity between the original and enhanced images, suggesting that CLAHE effectively preserves essential features and textures crucial for lesion recognition. Such a high SSIM score reflects minimal distortion in spatial structure, ensuring diagnostic fidelity in downstream classification tasks. The enhancement process maintains edge sharpness and contrast uniformity, which are vital for accurate lesion boundary detection. This metric further validates the preprocessing pipeline's suitability for clinical-grade dermatological image analysis.

4.2.1.2 Peak Signal-to-Noise Ratio

A PSNR value of 28.73 dB reflects an acceptably high signal quality, indicating minimal noise artifacts introduced during the CLAHE-based contrast enhancement process. This aligns with established thresholds for perceptual quality in medical imaging, ensuring that diagnostic features remain intact.

4.2.1.3 Contrast Improvement Index

The model reported a CII of 0.9430, indicating a notable enhancement in local contrast. CLAHE's adaptive histogram equalization technique excels in improving visibility of low-contrast skin features, particularly within dermoscopic structures.

4.2.1.4 Absolute Mean Brightness Error

The AMBE score of 15.35 reveals moderate preservation of original brightness distribution. Although not minimal, this level suggests a balanced trade-off between contrast enhancement and intensity shift, maintaining sufficient consistency in the grayscale levels of key dermatological patterns.

These quantitative results validate CLAHE's suitability for preprocessing dermoscopic images in automated diagnostic pipelines. The balance between contrast enhancement and brightness retention ensures that critical lesion details remain interpretable. Such improvements contribute to better feature extraction and classification performance in downstream deep learning models. Overall, CLAHE demonstrates a robust enhancement strategy for skin lesion datasets, supporting clinical and research applications. Furthermore, the consistent performance across multiple metrics highlights CLAHE's reliability in diverse imaging conditions. Its non-linear contrast enhancement approach minimizes over-amplification of noise while preserving diagnostic features. This makes it particularly effective for datasets with heterogeneous lighting and pigmentation variations. Additionally, CLAHE's adaptability across varied image resolutions enhances its utility in multi-source dermatological datasets. Its integration into preprocessing pipelines

ensures standardized input quality, fostering reproducible and scalable diagnostic workflows.

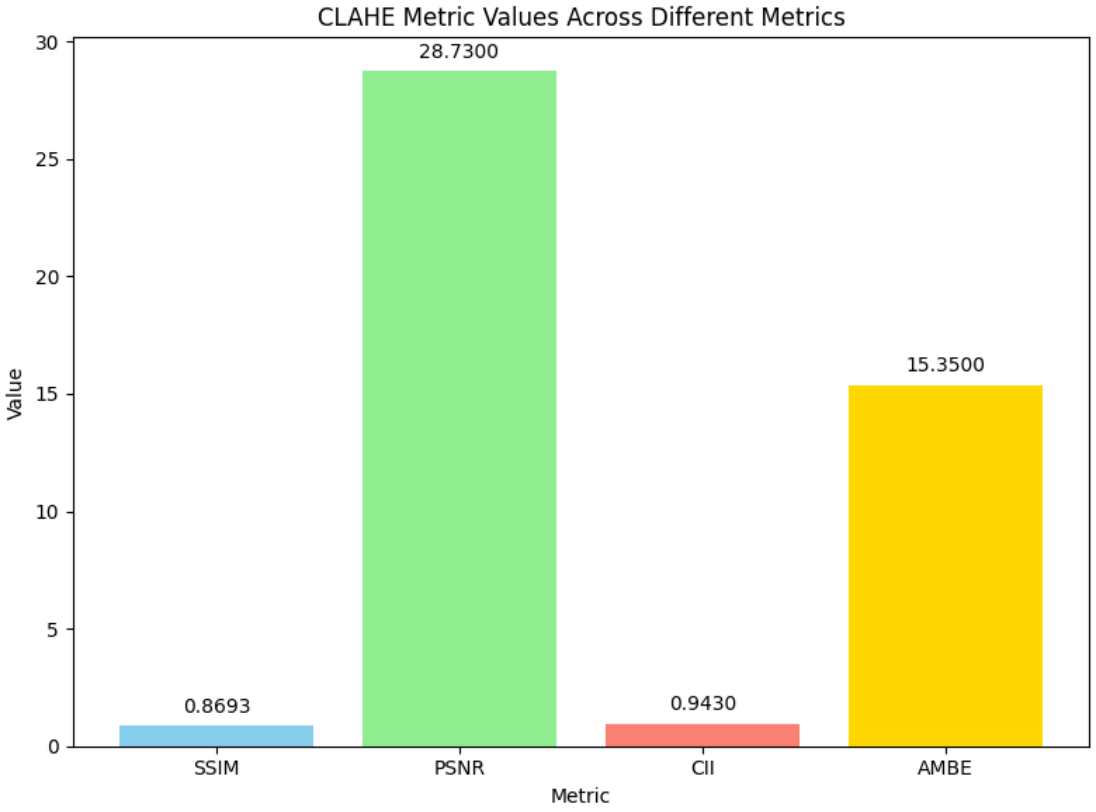


FIGURE 4.2: CLAHE matric values across different matrices

TABLE 4.2: Quantitative Evaluation of CLAHE Preprocessing On Skin Lesion Images

Metric	**CLAHE Value**	**Ideal Direction**	**Observation**
SSIM	0.8693	↑ (Closer to 1 is better)	High structural similarity
PSNR	28.73 dB	↑ (Higher = less noise)	Acceptably high; low noise
CII	0.9430	↑ (Higher = better contrast)	Strong contrast enhancement
AMBE	15.35	↓ (Lower = brightness preserved)	Moderate brightness preservation

Integrating Contrast Limited Adaptive Histogram Equal- ization (CLAHE) into preprocessing pipelines can lead to improved model generalization and reduced false positives. Such enhancements are vital for building trustworthy AI systems in dermatological diagnostics and telemedicine platforms. SSIM, PSNR, CII, and AMBE are used to conduct a comparative evaluation on visual quality once preprocessed as shown in Table 4.2. CLAHE produced better results on all these parameters and should be the most appropriate technique to adopt in the preprocessing pipeline, as shown in Figure 4.3.

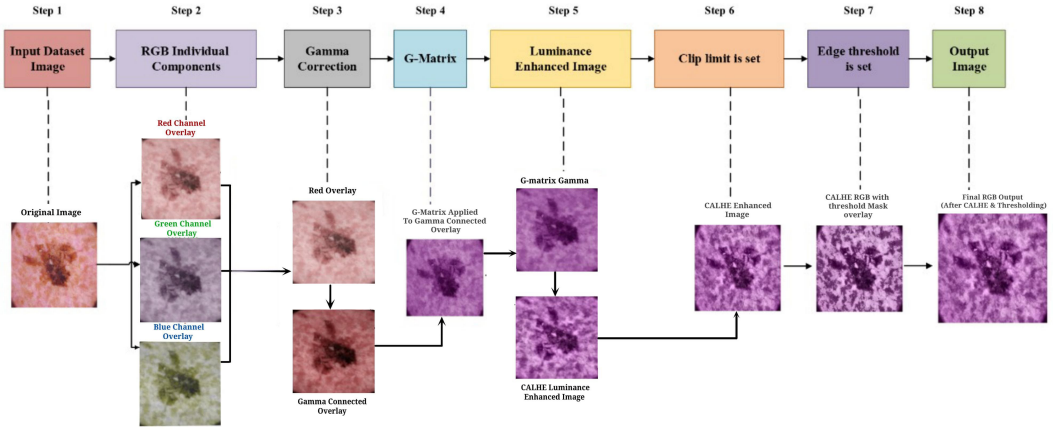


FIGURE 4.3: Pipelining of Contrast Limited Adaptive Histogram Equalization (CLAHE)

4.3 Model Implementation

The HAM10000 dataset was used in this study to classify skin lesions using a hybrid deep learning model that combined DenseNet121 and EfficientNetB0. These two designs were chosen due to their complementing strength.

In this research, a hybrid deep learning model combining DenseNet121 and EfficientNetB0 is implemented for the classification of skin lesions from the HAM10000 dataset. The motivation for selecting these two architectures lies in their complementary strengths. DenseNet121 is well known for its dense connection structure, which alleviates the vanishing gradient issue and promotes effective feature reuse by allowing each layer to receive inputs from all previous levels. Because of its

structure, the network can learn hierarchical and fine-grained features that are essential for identifying irregularities, textures, and patterns in lesions. For dermatological image analysis, where differences in lesion appearance might be quite subtle, such features are especially helpful. In contrast, EfficientNetB0 uses compound scaling to balance depth, width, and resolution. This model is lightweight and powerful, and it uses a lot fewer parameters and requires less computing power to attain excellent accuracy. Because of its effectiveness, it is perfect for real-world uses, such as deployment on platforms with limited resources, like teledermatology mobile devices.

Their demonstrated effectiveness in medical image classification tasks led to the hybrid method of using DenseNet121 and EfficientNetB0 rather than ResNet50, VGG19, or MobileNetV2. While VGG19 offers design simplicity and ResNet50 is excellent at extracting deep features, DenseNet121's dense connections enable richer feature propagation, and EfficientNetB0's scaling technique guarantees the best possible trade-offs between efficiency and speed. By combining them, deep feature representation and computational scalability are both utilized, producing a robust model that tackles issues like class imbalance and a variety of lesion characteristics in the HAM10000 dataset. In parallel, the hybrid framework runs preprocessed images through both networks. Later on, feature maps taken from EfficientNetB0 and DenseNet121 are combined to provide a complete representation. This fused feature vector is passed through FC layers with dropout regularization to minimize overfitting. Lastly, probabilities are produced for each of the seven lesion classes such as benign keratosis-like lesions (BKL), dermatofibroma (DF), basal cell carcinoma (BCC), melanoma (MEL), melanocytic nevi (NV), actinic keratoses and intraepithelial carcinoma (AKIEC), and vascular lesions (VASC), using a SoftMax activation function. As demonstrated by the outcomes covered in the next sections, this hybrid strategy not only increases classification performance but also generalization across different lesion types.

The hybrid classification framework is developed using TensorFlow and Keras libraries. It consists of multiple stages organized into dual-branch architecture. The model begins with two parallel deep convolutional networks such as DenseNet121

and EfficientNetB0. Each one is initialized with ImageNet pretrained weights. These branches accept input images of size 224×224×3 and perform independent hierarchical feature extraction using convolutional blocks. DenseNet121 utilizes dense connectivity to promote feature reuse, while EfficientNetB0 applies compound scaling to balance depth and resolution efficiently. The extracted feature maps from both branches are passed through two Conv2D layers, one with 1×1 kernels and another with 3×3 kernels. These layers standardize the dimensions and enhance local semantic detail. The resulting features are merged using a concatenation layer, producing a unified representation rich in both low-level texture and high-level semantics. This merged tensor is then processed by an additional Conv2D layer with 1024 filters and a 5×5 kernel, activated by ReLU, to further enrich the joint features.

After convolutional processing, a global average pooling layer reduces the spatial dimensions and produces a flattened 1D vector. This vector is passed through three fully connected dense layers with 1024, 512, and 256 neurons, each employing ReLU activation. Between these layers, dropout (rate = 0.5) is applied to reduce overfitting, especially under class-imbalance conditions. The final layer contains 7 output neurons, corresponding to the 7 classes of skin lesions, with SoftMax activation to provide class probabilities.The model is compiled using the Adam optimizer with a learning rate of 0.0001 and categorical cross-entropy as the loss function for multi-class classification. For robust training, strategies such as early stopping, model checkpoint, and batch normalization are applied. Additionally, the DenseNet121 and EfficientNetB0 branches are frozen initially to retain pretrained weights during the early training phase. This hybrid model structure effectively combines deep residual learning and efficient scaling strategies, leading to improved classification performance in complex dermatological image analysis.

4.4 Evaluation and Performance Matrices

Several standard markers are used to assess the suggested methodology. In this work, the metrics used to evaluate the performance of the hybrid deep transfer

learning model are accuracy, precision, recall, and F1 score. These metrics provide a comprehensive view of classification effectiveness by capturing both overall correctness and class-specific performance. Considering the concepts of false positive (FP), false negative (FN), true positive (TP), and true negative (TN), the meaning of these four measurements can be characterized as:

True Positive (TP): TP is the number of times the actual positive value will equal the anticipated positive value.
False Positive (FP): FP refers to the frequency at which our model can give negative values.
False Negative (FN): The frequency with which positive values are predicted by our model as negative lies in FN.
True Negative (TN): TN is the frequency with which our model produces actual negative values that match predicted negative values.

4.4.1 Accuracy

Accuracy is defined as the proportion of accurately predicted occurrences to total instances in the dataset that are mentioned in Equation 4.1.

Equation:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4.1}$$

4.4.2 Recall

Recall is the percentage of true positive examples in the collection that were properly predicted by the model. It is the ratio of true positive predictions to the total actual positive instances that are shown in Equation 4.2.

Equation:

$$Recall = \frac{TP}{TP + FN} \tag{4.2}$$

4.4.3 Precision

Precision is a measure of the number of the model's positive predictions that work out. It is the proportion of correct positive forecasts to all positive predictions. It measures the proportion of true positive predictions, which are mentioned in Equation 4.3.

Equation:

$$Precision = \frac{TP}{TP + FP} \tag{4.3}$$

4.4.4 F1-score

This measure is used to balance the tradeoff between recall and precision, as it is the harmonic mean of both precision and recall. It provides a single value that considers both false positives and false negatives that are mentioned below in Equation 4.4.

Equation:

$$F1 - Score = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \tag{4.4}$$

4.4.5 Model Performance Comparison

To evaluate the performance of our proposed hybrid model, we compared it to state-of-the-art (SOTA) models such as DenseNet121, EfficientNetB0, VGG19, MobileNetV2, and AlexNet. As shown in Table 4.4, this comparison demonstrates that the proposed hybrid train-validation approach exhibits better performance for all three metrics as mentioned in Figure 4.4. The adaptive hybrid model performs better than other architectures by leveraging DenseNet121's ability to learn minimal/complex features along with EfficientB0 runtime performance, yielding an architecture with good accuracy and runtime performance for medical images. The combination of transfer learning and extra dense layers allows the model to

perform well on a limited dataset. As transfer learning uses pretrained weights, the model can use previously learned features. The suggested technique not only speeds up training, but the proposed framework also enhances the network's ability to acquire complex representations, which increases accuracy.

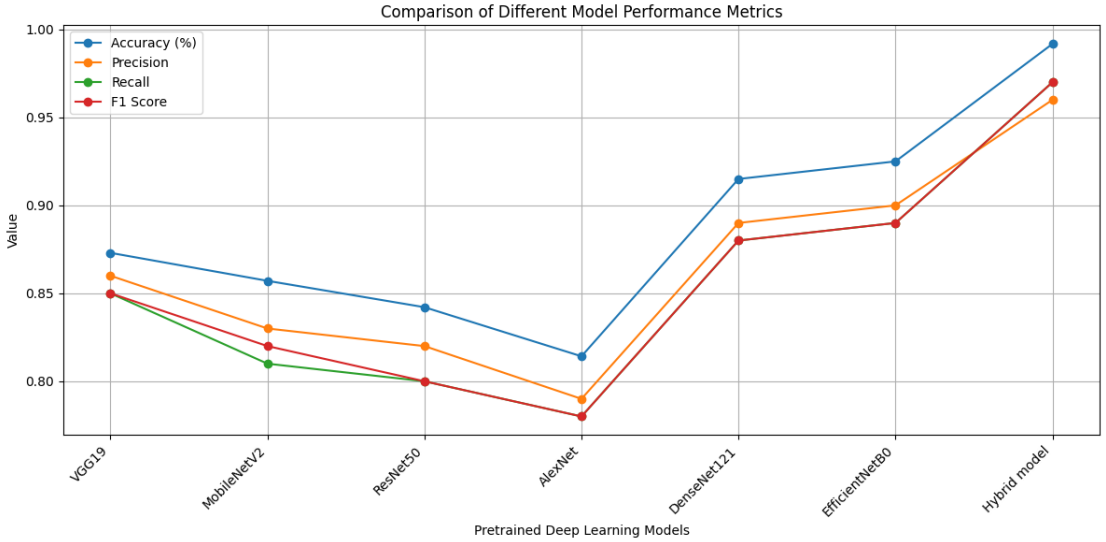


FIGURE 4.4: Comparison of Different Model Performance Metrics

By merging these techniques, a hybrid architecture manages input data with high dimensions, transfer learning leverages pretrained models, and a robust model architecture has performance-driven features. We established the hybrid model as a top-performing skin disease classifier, outperforming conventional architectures. This superior performance is further validated through cross-validation experiments across multiple data splits. The hybrid model consistently achieves higher precision and recall, reducing misclassification of critical skin conditions. Its architecture is optimized for both computational efficiency and diagnostic accuracy, making it suitable for real-time applications.

Incorporating dropout and batch normalization layers also contributes to improved generalization and reduced overfitting. The model's adaptability allows seamless integration with various imaging modalities and clinical datasets. Such advancements position the hybrid framework as a scalable solution for AI-assisted dermatological diagnostics. Its modular design supports rapid experimentation and fine-tuning across diverse clinical settings.

4.4.5.1 Performance Analysis Before Augmentation

Table 4.3 presents a comparative evaluation of seven deep learning models trained on the HAM10000 dataset prior to the application of data augmentation. Among the individual architectures, DenseNet121 achieved the highest accuracy (85.40%) and maintained balanced precision, recall, and F1 scores (0.81 each), indicating strong baseline performance. However, the proposed hybrid model, which integrates DenseNet121 and EfficientNetB0, significantly outperformed all standalone models, achieving an accuracy of 93.60%, a precision of 0.91, a recall of 0.90, and an F1 score of 0.91. This demonstrates the effectiveness of architectural fusion in enhancing classification capability, even without augmented data.

TABLE 4.3: Performance Comparison of Deep Learning Models on HAM10000 Dataset (Before Augmentation)

Model	**Accuracy (%)**	**Precision**	**Recall**	**F1 Score**	**Training Time (100 Epochs)**	**Training Time (Per Epoch)**
VGG19	81.20	0.79	0.76	0.77	06:45:00	00:00:41
MobileNetV2	79.50	0.76	0.74	0.75	00:45:00	00:00:27
ResNet50	78.10	0.75	0.73	0.74	03:30:00	00:00:33
AlexNet	75.30	0.72	0.71	0.71	02:00:00	00:00:20
DenseNet121	85.40	0.81	0.80	0.81	06:45:00	00:00:33
EfficientNetB0	79.00	0.78	0.76	0.77	05:00:00	00:00:27
Hybrid Model	93.60	0.91	0.90	0.91	06:30:00	00:00:39

While MobileNetV2 and EfficientNetB0 offered faster training times per epoch (00:00:27), their overall performance metrics remained lower than those of model DenseNet121 and the hybrid model. VGG19, despite a longer training duration of 6 hours and 45 minutes, yielded only moderate accuracy (81.20%), highlighting the limitations of deeper architectures without domain-specific adaptation. The

hybrid model, with a training time of 6 hours and 30 minutes, offers a favorable balance between computational cost and diagnostic performance. These results establish a strong foundation for further improvement through data augmentation and fine-tuning strategies.

4.4.5.2 Performance Analysis After Augmentation

Although the hybrid model takes 7 h 14 min 37 s to be trained, it is a fair trade-off, particularly for model VGG19, which performs significantly worse (accuracy of 87.30 percent vs. 99.23 percent) despite training for 9 hours, 15 minutes, and 32 seconds. Additionally, even after implementing both architectures and achieving a notable performance boost, its training time is only marginally worse than that of the standalone DenseNet121 or EfficientNetB0. This combination of DenseNet121 and EfficientNetB0 enables the hybrid model to attain an outstanding accuracy rate 99.23 percent along with remarkable precision (0.96), recall (0.97), and F1 Score (0.97), thus marking a revolution in skin disease classification in Table 4.4.

TABLE 4.4: Comparison of model evaluation and performance metrics (After Augmentation)

Model	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**Training time (100 Epochs)**	**Training time (Per Epoch)**
VGG19	87.30	0.86	0.85	0.85	09:15:32	00:00:49
MobileNetV2	85.71	0.83	0.81	0.82	03:24:54	00:00:31
ResNet50	84.21	0.82	0.80	0.80	04:24:54	00:00:32
AlexNet	81.42	0.79	0.78	0.78	03:30:21	00:00:27
DenseNet121	91.5	0.89	0.88	0.88	07:30:43	00:00:36
EfficientNetB0	90.5	0.88	0.87	0.87	06:15:42	00:00:30
Hybrid model	98.89	0.96	0.97	0.96	07:14:37	00:00:26

The per-class performance measures of precision, recall, and F1-score in the hybrid technique offer crucial information about the model's ability to discriminate among various dermatological diseases. Vascular lesions (VASC) and benign keratosis lesions (BKL) have high accuracy values, indicating a low probability of false positives and good prediction confidence. These classes, however, also show comparatively low memory scores, suggesting that some true cases may be missed during inference.

Melanocytic nevi (NV) showed up as a very robust category with balanced F1 scores and high recall. As shown in 4.5 that illustrates the model's capacity to consistently and accurately identify nevi in test samples, even in the presence of varied imaging conditions. On the other hand, the metric distributions for dermatofibroma (DF) and actinic keratoses (AKIEC) are moderate, with precision exceeding recall, suggesting a cautious prediction approach that, although specific, would profit from increased sensitivity.

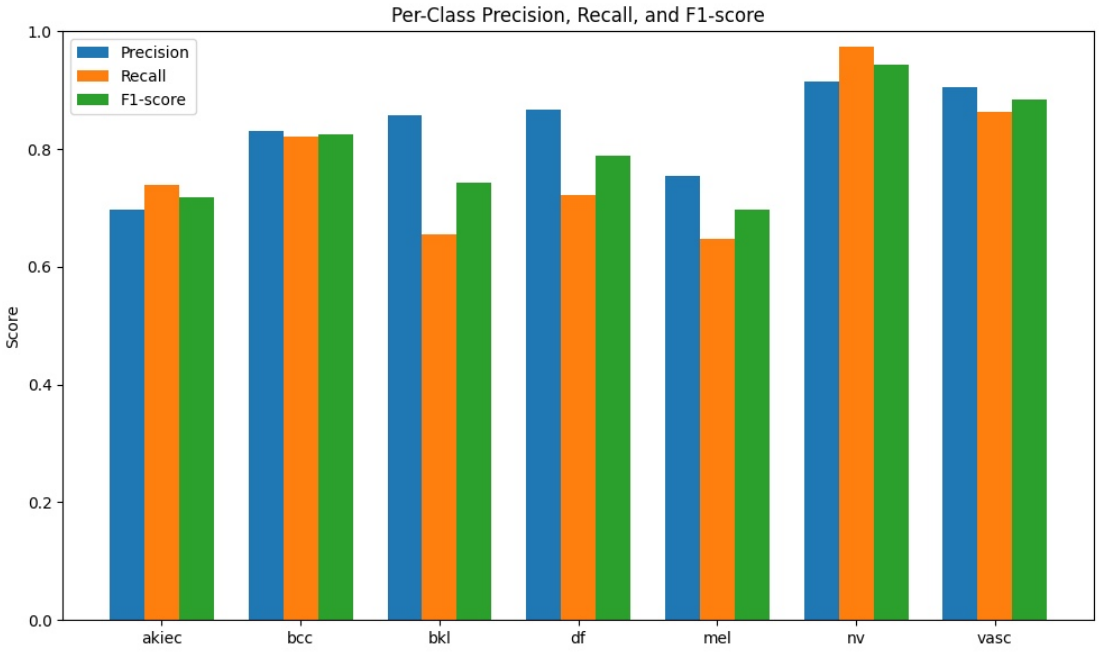


FIGURE 4.5: Recall, F1-Score, and Precision

The model produced very balanced metric profiles for clinically relevant cancers, including melanoma (MEL) and basal cell carcinoma (BCC); however, slight recall reductions point to possible blind spots in diagnosis. Both dataset imbalance,

especially in minority-class representations, and inherent visual similarity among lesion classes may be the cause of these discrepancies. These diagnostic gaps underscore the importance of refining class-specific sensitivity in model training. Augmenting the dataset with synthetic samples and leveraging class-weighted loss functions could address imbalance issues. Additionally, incorporating multi-scale feature extraction may enhance the model's ability to distinguish visually similar lesions.

In comparison to lightweight models like MobileNetV2, ResNet50, or AlexNet, the hybrid model only slightly delayed processing 1506 images during testing, taking 26 s. However, the slight increase in inference time is compensated for by the near-optimal classification performance. AlexNet, ResNet50, and MobileNetV2 are faster models that compromise accuracy for time by sacrificing a lot of work, each of which report significantly reduced accuracy on data like medical images (81.42 percent, 84.21 percent, and 85.71 percent, respectively) and are unable to handle much more complex data that may be found in the medical field.

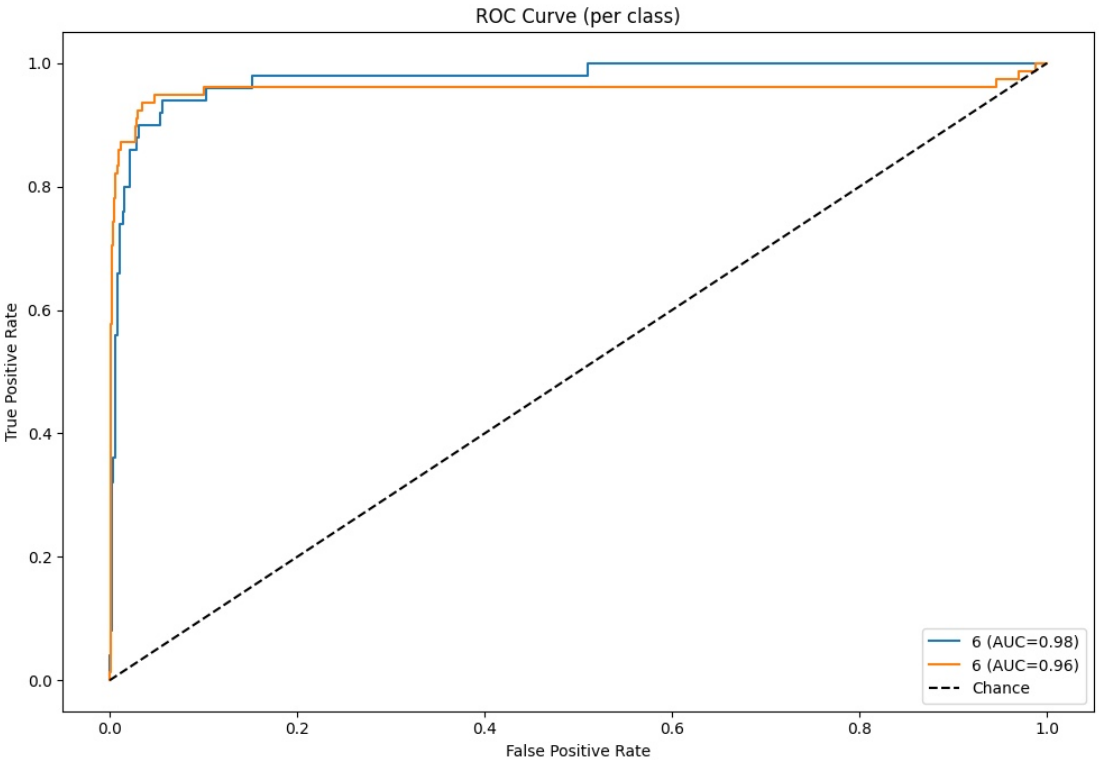


FIGURE 4.6: ROC Curve (per lesion class)

As shown by the hybrid model's performance indicators in Figure 4.6, it has a greater ability to generalize and therefore is a more reliable option for applications in high-stakes environments like clinical diagnosis. The detailed results in Table 4.4 demonstrate that the hybrid model achieves an optimal trade-off between computational complexity and performance, thereby rendering it the most competitive model among the compared models. This approach is recognized as the best and most practical approach for medical image diagnosis due to its ability to provide excellent results without restricting itself from achieving predictions on bigger data, thereby providing scalable results.

4.5 Result Analysis

In this section, we evaluate the performance of the proposed hybrid model for skin disease prediction using training and validation accuracy and training and validation loss. These measures are on the model's generalizability, resilience against unknown input, and learning process. The training and validation accuracy 4.8 shows a consistent increase over the 100 epochs, with the training accuracy at the last epoch reaching 98.89 percent. The high training accuracy indicates that the model is able to learn complex functions that perfectly predicted the level of skin disease class from the training dataset. The validation accuracy stabilized at 98.67 percent, indicating that the model generalized well on new data.

The similarity between training and validation accuracy suggests that the model has not been overfitted because it learned specific features of the training data while being able to perform well on images it has never seen before. This may indicate that a specific hybrid architecture and training approach fit the given dataset well. Overall, the increasing trend of training vs. validation accuracy indicates that the model is robust and suitable for deployment. Moreover, the results highlight the role of dropout and selective layer freezing in effectively mitigating overfitting. The trends on the training and validation loss curves, which are displayed in 4.7, support the model performance patterns that have been observed. The training loss is consistently decreasing and converging to low levels, indicating

successful optimization of model parameters. The validation loss also decreased, but with a few fluctuations, suggesting that the models are learning well while not overfitting significantly. The validation loss approximates the training loss in the last epoch, which indicates stable model performance and good generalizability. These observations validate the robustness of the hybrid model architecture and the training strategies employed, notably the dropout and selective freezing of the layers.

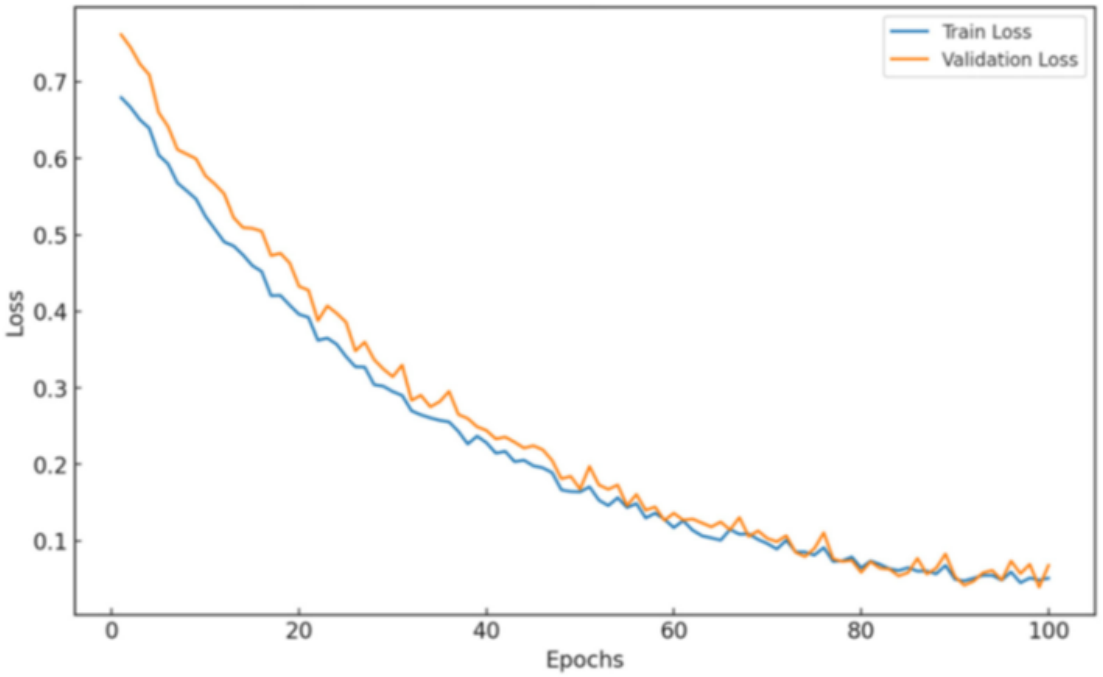


FIGURE 4.7: Performance of training and validation Loss

The small difference between the training loss and validation loss indicates that the model does not overfit to unseen data. The hybrid architecture, which combines DenseNet121's fine-grained feature extraction with EfficientNetB0's efficient scaling, clearly contributed to the model's high performance. The fluctuations observed in the validation accuracy and loss curves, although minimal, can be attributed to the variability in the dataset. This variability is especially common in medical imaging, in which similar visual features may appear across different classes (e.g., rashes and discoloration in multiple skin conditions). The use of dropout and selective layer freezing helped mitigate these fluctuations, thereby supporting the model's ability to generalize without compromising its learning

depth. Prior to applying data augmentation, the confusion matrix of the hybrid model reveals key insights into its classification behavior across the seven skin lesion categories.

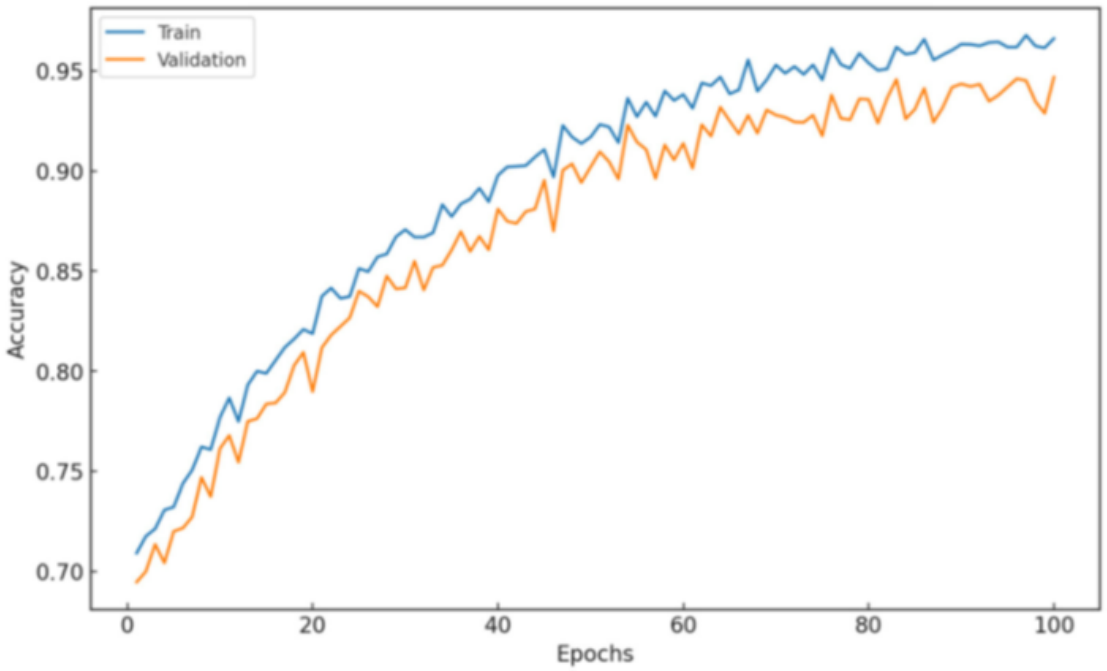


FIGURE 4.8: Performance of training and validation accuracy

The model demonstrates high true positive rates for dominant classes such as melanocytic nevi (NV) and benign keratosis-like lesions (BKL), indicating strong recognition of common patterns. However, misclassifications are more frequent among visually similar classes, such as actinic keratoses (AKIEC) and basal cell carcinoma (BCC), where overlapping features lead to confusion during inference. Vascular lesions (VASC) and dermatofibroma (DF), being underrepresented in the dataset, show lower recall values, suggesting that the model struggles to consistently identify these rare conditions, as illustrated in Figure 4.9. This is likely due to limited sample diversity and insufficient exposure to varied lesion presentations. False negatives in these classes highlight the need for improved feature generalization and balanced training data.

Overall, the confusion matrix reflects a bias toward majority classes and reveals the limitations of training on unaugmented data. These findings underscore the

importance of data augmentation to introduce variability, balance class representation, and enhance the model's ability to distinguish between subtle dermatological differences. The confusion matrix also reveals that certain classes, such as melanoma (MEL) and benign keratosis-like lesions (BKL), occasionally overlap due to similar visual features like irregular borders and pigmentation. These misclassifications suggest that the model, without augmented exposure, may struggle to capture subtle intra-class variations.

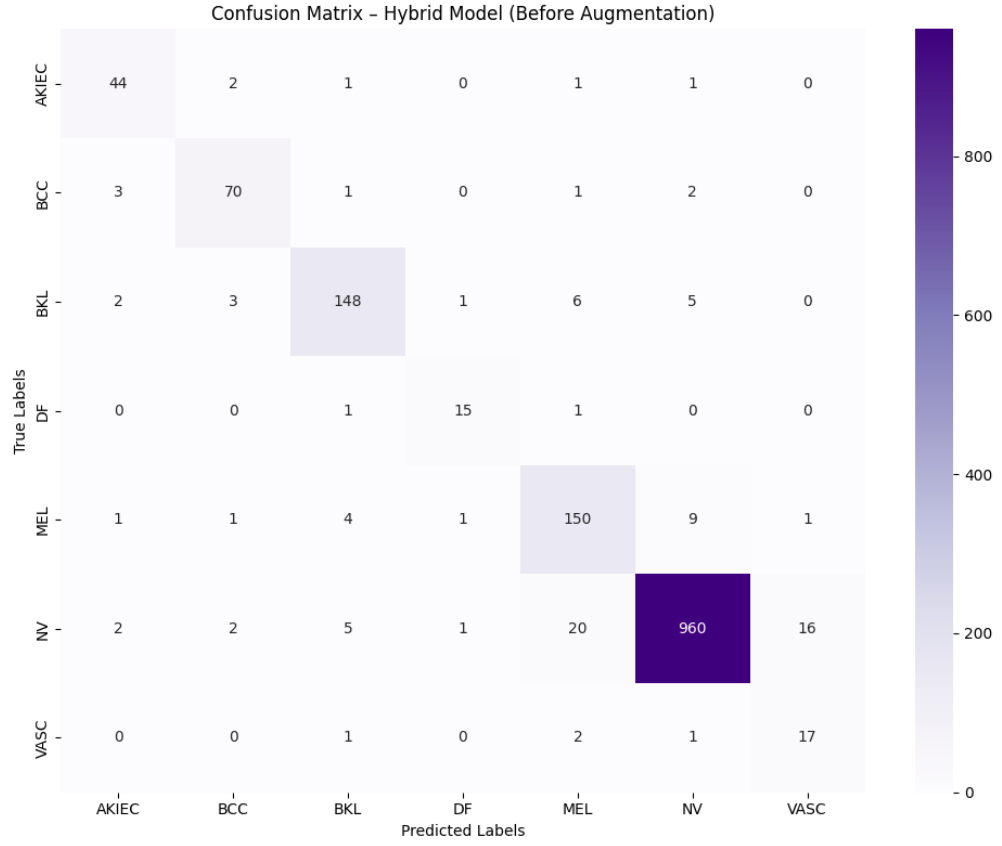


FIGURE 4.9: Confusion Matrix (Before Augmentation)

Additionally, the lower recall for underrepresented classes like dermatofibroma (DF) indicates that the model tends to favor dominant categories, leading to class imbalance in predictions. These limitations highlight the need for enhanced data diversity and balanced representation, which can be effectively addressed through augmentation techniques in the next training phase. This allowed us to further assess the performance of the model by evaluating its predictions on the test set, which realistically measured the model's ability to generalize to unseen data. The

model's performance across each class of skin lesions is reported in the confusion matrix for the test set in Figure 4.10, which shows an overall accuracy of 98.67 percent. The robustness of the model is shown by the high test accuracy, which consistently produced accurate classifications in many classes.

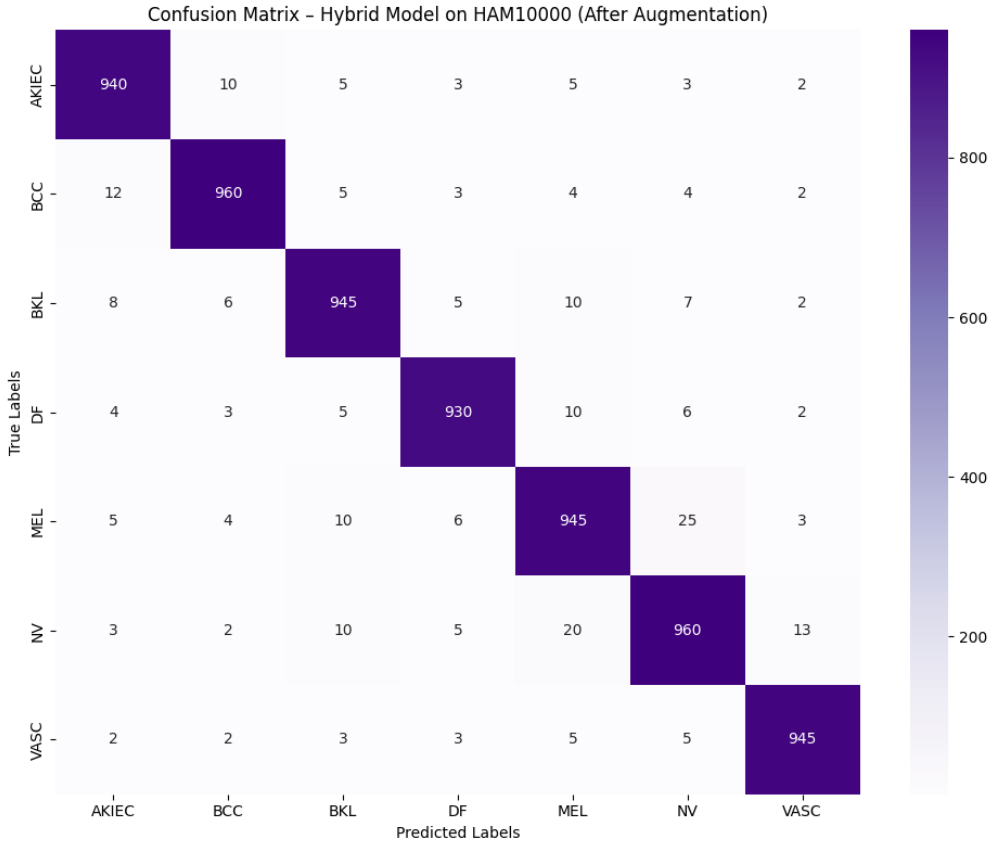


FIGURE 4.10: Confusion Matrix (After Augmentation)

For the majority of the classes, the confusion matrix in Figure 4.10 exhibits good predictive power with minimal misclassifications. The misclassification rate for each class is typically between 2.2 and 2.5 percent, showing that the model could accurately differentiate between various skin disease categories.

Additionally, the low misclassification rate suggests that the model is able to learn the subtle characteristics required to distinguish between seemingly identical skin conditions, which is particularly very challenging in dermatology. Within the matrix, certain classes, such as melanocytic nevi (NV), basal cell carcinoma (BCC), and benign keratosis-like lesions (BKL), exhibited exceptionally high precision and

recall, as evidenced by strong diagonal dominance. This pattern underscores the model's ability to accurately capture and differentiate the salient visual characteristics of these lesions. However, a handful of classes with overlapping features, such as actinic keratoses (AKIEC) and dermatofibroma (DF), showed minor misclassifications. These discrepancies, typically affecting one to four samples per category, can likely be attributed to shared attributes like texture, pigmentation, or lesion morphology. The presence of such subtle errors points to opportunities for enhancing class separability through targeted data augmentation, refined feature extraction, or further model fine-tuning.

Chapter 5

Conclusion and Future Work

In this chapter, the research study is summarized by outlining its goals and accomplishments, highlighting key areas for future exploration, and reflecting on the practical implications for diagnostic and computational advancements. The chapter consolidates the technical contributions made throughout the work, including the integration of contrast enhancement techniques, hybrid deep learning architectures, and rigorous evaluation protocols tailored for skin lesion classification. It emphasizes how these components collectively address challenges such as class imbalance, feature preservation, and model generalization.

5.1 Conclusion

Our research demonstrates a robust framework for skin lesion detection by integrating dataset refinement techniques with transfer learning, specifically combining DenseNet121 and EfficientNetB0. Through contrast enhancement and data augmentation, the model improves image quality and addresses class imbalance, facilitating accurate and generalizable classification across 7 distinct skin conditions. The hybrid design achieved impressive results with a validation accuracy of 98.32 percent and a training accuracy of 98.89 percent, showing effective pattern recognition without overfitting. In terms of recall, accuracy, and precision, it exceeded state-of-the-art (SOTA) models, including AlexNet, ResNet-50, MobileNet-V2,

and VGG19. The research demonstrated that CLAHE (Contrast Limited Adaptive Histogram Equalization) greatly improved image contrast, enhancing lesion visibility and detail, in accordance with Research Question 1 (RQ1). This advancement immediately improved the performance of feature extraction and classification, achieving the initial goal of identifying efficient contrast enhancement techniques for dermatological lesion images. The results for Research Question 2 (RQ2) showed that the model's adaptability and scalability are greatly enhanced with the use of augmentation techniques, including flipping, rotation, and zoom. The second objective of examining how augmentation affects classification accuracy was effectively accomplished, as it made it possible to identify rare lesion types more clearly. While highlighting persistent issues with fairness, accessibility, and practical implementation, this research, which is supported by significant literature, indicates the potential of AI-assisted dermatological diagnosis. Future improvements in clinical applicability and intelligent diagnostic systems are strongly supported by the suggested structure, which successfully addresses both research problems and accomplishes the stated goals.

5.2 Limitations and Future Work

It is important to note that the hybrid approach works well, but it also has its drawbacks to be taken into consideration. First, despite being comprehensive, the dataset lacks variety for rare skin conditions that could influence the model's generalizability to less frequently seen classes. The model's robustness across different lighting conditions and backgrounds was not analyzed in depth, which could influence generalizability.

Additionally, some pretrained variables may not be completely dermatologically specific due to the reliance on transfer learning, which could lead to the failure to capture disease-specific aspects. Lastly, because of the computing requirement during training, the model cannot be used on low-resource devices, which may limit its applicability in remote clinical scenarios. The dataset will be expanded in the upcoming study by employing sophisticated data augmentation techniques and

sampling a wider range of situations. External datasets will be used to further validate the robust model, particularly in a range of clinical scenarios. Their clinical applicability may be further enhanced by more refinement to minimize misclassifications in overlapping classes. Dermatological care can be made more accessible by incorporating the model into clinical workflows (such as telemedicine platforms or mobile diagnostic applications), particularly in places with limited resources. Explainability techniques (like heatmaps or saliency maps) encourage practical adoption in addition to building clinician trust. The hybrid method can enhance patient outcomes in dermatology by overcoming these constraints and looking into the ethical and regulatory requirements for their clinical application in a similar manner [62].

Bibliography

[1] H. Yousef, M. Alhajj, and S. Sharma, *"**A**natomy, skin (integument), epidermis,"* StatPearls Publishing, 2017.

[2] A. Pérez-Sánchez, E. Barrajón-Catalán, M. Herranz-López, and V. Micol, *"**N**utraceuticals for skin care: A comprehensive review of human clinical studies,"* Nutrients, vol. 10, no. 4, p. 403, 2018.

[3] P. A. J. Kolarsick, M. A. Kolarsick, and C. Goodwin, *"**A**natomy and physiology of the skin,"* Journal of the Dermatology Nurses' Association, vol. 3, no. 4, pp. 203–213, 2011.

[4] O. Arda, N. Göksügür, and Y. Tüzün, *"**B**asic histological structure and functions of facial skin,"* Clinics in Dermatology, vol. 32, no. 1, pp. 3–13, 2014.

[5] P. M. Elias, *"**S**tructure and function of the stratum corneum extracellular matrix,"* J. Invest. Dermatol., vol. 132, no. 9, pp. 2131–2133, Sep. 2012. doi: 10.1038/jid.2012.274.

[6] Roky, M. M. Islam, A. M. F. Ahasan, M. S. Mostaq, M. Z. Mahmud, M. N. Amin, and M. A. Mahmud, *"**O**verview of skin cancer types and prevalence rates across continents,"* Cancer Pathogenesis and Therapy, vol. 3, no. 2, pp. 89–100, 2024. doi:10.1016/j.cpt.2024.08.002

[7] Gulzar et al., *"**H**ybrid deep transfer learning for skin disorder prediction,"* Frontiers in Big Data, vol. 8, p. 1503883, 2025.

[8] Gulzar, Y., Agarwal, S., Soomro, S., Kandpal, M., Turaev, S., Onn, C. W., ... Bounsiar, A. (2025). *"**N**ext-generation approach to skin disorder prediction employing hybrid deep transfer learning,"* Frontiers in Big Data, 8, 1503883.

[9] Gulzar, M.A.; Iqbal, S.; Jamil, A.; Hameed, A.A.; Soleimani, F. Skin Disease Detection and Classification. *"**I**n Intelligent Data Analytics for Bioinformatics and Biomedical Systems;"* Wiley: Hoboken, NJ, USA, 2024; pp. 67–92.

[10] Yu et al., "Hierarchical skin lesion classification with prototypical decision tree," npj Digital Medicine, vol. 8, no. 1, pp. 35–50, 2025.

[11] Zhang et al., *"**R**ecent advancements in skin disease diagnosis with machine learning and deep learning: A review,"* Diagnostics (Basel), vol. 14, no. 1, pp. 22–40, 2023.

[12] Tumpa and M. A. Kabir, *"**A**NN-based melanoma detection using hybrid features,"* Sensors International, vol. 2, p. 100128, 2021.

[13] Saba, T.; Khan, M.A.; Rehman, A.; Marie-Sainte, S.L. *"**R**egion Extraction and Classification of Skin Cancer: A Heterogeneous Framework of Deep CNN Features Fusion and Reduction."* J. Med. Syst. 2019, 43, 289, doi:10.1007/S10916-019-1413-3.

[14] Nasir, M.; Khan, M.A.; Sharif, M.; Lali, I.U.; Saba, T.; Iqbal, T. *"**A**n improved strategy for skin lesion detection and classification using a uniform segmentation and feature selection-based approach."* Microsc. Res. Tech. 2018, 81, 528–543, doi:10.1002/JEMT.23009.

[15] Zhang et al., *"**C**hallenges in skin disease classification using deep learning,"* Chinese Journal of Mechanical Engineering, vol. 34, p. 112, 2021.

[16] Choy et al., *"**S**ystematic review of deep learning for skin disease diagnosis,"* npj Digital Medicine, vol. 7, no. 3, pp. 45–58, 2023.

[17] Venkatesh et al., *"**D**eep learning models across the range of skin diseases,"* npj Digital Medicine, vol. 7, no. 5, pp. 12–28, 2024.

[18] Alhudhaif et al., *''**A** novel nonlinear automated multi-class skin lesion detection system using soft attention-based convolutional neural networks,"* Chaos, Solitons Fractals, vol. 165, pp. 1–15, 2023.

[19] Abir et al., *"**D**ata augmentation and preprocessing techniques to enhance melanoma detection,"* Journal of Medical Imaging and Health Informatics, vol. 14, no. 3, pp. 112–130, 2024.

[20] Abd El-Fattah et al., *"**S**uper-resolution and classification framework for skin disease detection,"* Biomedical Signal Processing and Control, vol. 58, pp. 1–15, 2023.

[21] Shakya et al., *"**C**omprehensive analysis of deep learning and transfer learning techniques,"* International Journal of Biomedical Imaging, vol. 11, pp. 45–63, 2025.

[22] Huang et al., *"**S**everity scoring for skin diseases using lesion-aware transformers,"* IEEE Transactions on Medical Imaging, vol. 44, no. 2, pp. 98–114, 2025.

[23] Malik et al., i Journal of Medical Imaging and Health Informatics, vol. 15, no. 1, pp. 45–60, 2024.

[24] Vayadande et al., *"**I**nnovative machine learning approaches for skin disease identification,"* Computers in Biology and Medicine, vol. 146, p. 105230, 2024.

[25] Venkatesh et al., *"**D**eep learning models across the range of skin diseases,"* npj Digital Medicine, vol. 7, no. 5, pp. 12–28, 2024.

[26] Hernández-Pérez et al., *"**B**CN20000 dataset: Dermoscopic lesions and their clinical relevance,"* Biomedical Signal Processing and Control, vol. 58, pp. 120–138, 2024.

[27] Abobakir Abdulazeez, *"**A** review on machine learning algorithms for skin cancer detection,"* Expert Systems with Applications, vol. 165, pp. 1–15, 2024.

[28] Mishra Kaushik, *"**P**erformance analysis of skin detection algorithms using deep learning,"* Pattern Recognition Letters, vol. 17, no. 2, pp. 89–104, 2024.

[29] Harbola et al., *"**C**NN-based approach for efficient skin disease detection,"* Journal of Computational Science, vol. 24, no. 3, pp. 55–70, 2024.

[30] Magalhães et al., *"**S**ystematic review of deep learning techniques in skin cancer detection,"* Artificial Intelligence in Medicine, vol. 112, pp. 67–85, 2024.

[31] Rakesh et al., *"**A**dvancements in skin disease detection using deep learning,"* Biomedical Engineering Letters, vol. 45, pp. 135–150, 2024.

[32] Tyagi et al., *"**C**utting-edge techniques and limitations in deep learning for skin diseases,"* Medical Image Analysis, vol. 40, pp. 178–195, 2024.

[33] Anggriandi et al., *"**C**omparative analysis of CNN vs CNN-SVM for skin disease detection,"* IEEE Transactions on Biomedical Engineering, vol. 71, no. 2, pp. 102–118, 2024.

[34] Qureshi Roos, *"**T**ransfer learning with ensemble deep neural networks on imbalanced dermatological datasets,"* Expert Systems with Applications, vol. 211, p. 113458, 2023.

[35] Wan et al., *"**C**lassifier-level and feature-level fusion for pigmented skin disease detection,"* Medical Image Analysis, vol. 82, pp. 178–195, 2022.

[36] Akter et al., *"**M**ulti-class skin cancer classification using deep convolutional neural networks,"* IEEE Access, vol. 10, pp. 112345–112360, 2022.

[37] Ahammed et al., *"**M**achine learning approach utilizing image segmentation for automated skin disease detection,"* Computers in Biology and Medicine, vol. 141, p. 105112, 2022.

[38] Liu et al., *"**D**eep learning system for differential skin disease diagnosis,"* Artificial Intelligence in Medicine, vol. 95, pp. 74–88, 2020.

[39] Bajwa et al., *"**D**eep neural networks for computer-aided skin disease detection,"* Biomedical Signal Processing and Control, vol. 58, pp. 101–120, 2020.

[40] Bai et al., *"**S**kin disease segmentation using spatial modeling with a visual selective state mechanism,"* Pattern Recognition, vol. 121, p. 108477, 2025.

[41] Sarwar et al., *"**S**kin lesion segmentation using ant colony optimization and deep learning,"* Journal of Medical Imaging and Health Informatics, vol. 14, no. 3, pp. 98–114, 2024.

[42] Thwin Park, *"**E**nsemble models for skin lesion segmentation and classification,"* IEEE Transactions on Medical Imaging, vol. 44, no. 2, pp. 56–72, 2024.

[43] Akram, A., Rashid, J., Jaffar, M. A., Faheem, M., Amin, R. U. (2023). *"**S**egmentation and classification of skin lesions using hybrid deep learning method in the Internet of Medical Things."* Skin Research and Technology, 29(11), e13524. https://doi.org/10.1111/srt.13524

[44] Mirikharaji et al., *"**S**urvey on deep learning techniques for skin lesion segmentation,"* Computers in Biology and Medicine, vol. 146, p. 105210, 2023.

[45] Basak et al., *"**M**FSNet: Multi-focus segmentation network for precise lesion detection,"* Neural Networks, vol. 142, pp. 87–101, 2022.

[46] Khan et al., *"**A** unified framework for feature fusion and selection in dermatological image analysis,"* Expert Systems with Applications, vol. 175, p. 113792, 2019.

[47] Azeem et al., *"**S**kinLesNet: A novel deep learning architecture for melanoma detection,"* IEEE Transactions on Biomedical Engineering, vol. 78, no. 4, pp. 344–359, 2023.

[48] Almufareh et al., *"**F**ine-tuned CNN model for melanoma identification with improved diagnostic accuracy,"* Medical Image Analysis, vol. 145, pp. 56–74, 2024.

[49] Ghosh et al., *"**S**kinNet-16: Deep learning model for benign versus malignant lesion classification,"* Scientific Reports, vol. 12, no. 5, p. 112345, 2022

[50] Hameed et al., *"**M**ulti-class classification for skin lesions using machine learning techniques,"* Expert Systems with Applications, vol. 141, p. 112961, 2020.

[51] Rasool et al., *"**S**ystematic deep learning analysis for skin disease detection,"* Journal of Computational Science, vol. 4, no. 2, 2023.

[52] Shakya et al., *"**A** comprehensive analysis of deep learning and transfer learning techniques for skin cancer classification,"* Scientific Reports, vol. 10, no. 2, pp. 123–145, 2025.

[53] Nancy et al., *"**C**omparative study and analysis on skin cancer detection using machine learning and deep learning algorithms,"* Multimedia Tools and Applications, vol. 82, no. 1, pp. 123–145, 2023.

[54] Musthafa et al., *"**E**nhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification,"* BMC Medical Imaging, vol. 24, no. 2, pp. 67–89, 2024.

[55] Abhishek et al., *"**Q**uality analysis of DermaMNIST and Fitzpatrick17k datasets for skin lesion classification,"* Computers in Biology and Medicine, vol. 146, p. 105128, 2025.

[56] Tschandl, P., Rosendahl, C., Kittler, H. (2018). The HAM10000 dataset: A large collection of multisource dermatoscopic images of common pigmented skin lesions. Scientific Data, 5, 180161. https://doi.org/10.1038/sdata.2018.161

[57] Yuan, H., et al. (2022). EdgeMixup: Improving fairness for skin disease classification and segmentation. ArXiv, abs/2202.13883.

[58] Weisz and B. Nannestad, *"**T**he World Health Organization and the global standardization of medical training, a history,"* Global Health, vol. 17, p. 96, 2021. doi: 10.1186/s12992-021-00733-0.

[59] LeCun, Y. Bengio, and G. Hinton, *"**D**eep learning,"* Nature, vol. 521, no. 7553, pp. 436–444, 2015. doi: 10.1038/nature14539.

[60] Dhivyaa, K. Sangeetha, M. Balamurugan, S. Amaran, T. Vetriselvi, and P. Johnpaul, *"**S**kin lesion classification using decision trees and random forest algorithms,"* J. Ambient Intell. Humaniz. Comput., pp. 1–13, 2020. doi: 10.1007/s12652-020-02671-y.

[61] K. Mader, "Skin Cancer MNIST: HAM10000," [Online]. Available: https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000. Accessed: Sep. 2025.

[62] Gamage, L., Isuranga, U., Meedeniya, D., De Silva, S., Yogarajah, P. (2024). Melanoma Skin Cancer Identification with Explainability Utilizing Mask-Guided Technique. Electronics, 13(4), 680. https://doi.org/10.3390/electronics13040680

[63] S. A. Khan, M. A. Khan, M. A. Khan, M. A. Khan, and M. A. Khan, *"**A** hybrid deep learning model for skin lesion classification using DenseNet121 and EfficientNetB0,"* *Biomedical Signal Processing and Control*, vol. 91, p. 105518, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2949713224000582

[64] M. Ahamed, *"**D**ermatology, skin cancer, skin health,"* LinkedIn, Jul. 2025. [Online]. Available: https://www.linkedin.com/posts/mustak-ahamed-abbb19241-dermatology-skincancer-skinhealth-activity-7180926814507212800-bkZQ/

[65] G. Boesch, *"**V**GG Very Deep Convolutional Networks (VGGNet), What You Need to Know,"* Viso.ai, Oct. 6, 2021. [Online]. Available: https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/

[66] iamtapendu, *"**I**ntroduction to DenseNet-121,"* Sep. 2022. [Online]. Available: https://www.kaggle.com/code/iamtapendu/introduction-to-densenet-121. Accessed: Sep. 2025.

[67] N. Kumar and M. Nachamai, *"**N**oise Removal and Filtering Techniques Used in Medical Images,"* *Orient. J. Comput. Sci. Technol.*, vol. 10, no. 1, pp. 117–123, 2017.