CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Lesion Segmentation of Dermoscopic Images for Classification of Skin Cancer using Deep Neural Networks

by

Aqib Shahzaib

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

in the

Faculty of Computing
Department of Computer Science

2024

Copyright © 2024 by Aqib Shahzaib

All rights reserved. No part of this thesis may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, by any information storage and retrieval system without the prior written permission of the author.

Dedicated to all of the dermatologists and other healthcare workers.

CERTIFICATE OF APPROVAL

Lesion Segmentation of Dermoscopic Images for Classification of Skin Cancer using Deep Neural Networks

by

Aqib Shahzaib

(MCS221016)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Muhammad Nazir	HITEC University, Taxila
(b)	Internal Examiner	Dr. Imtiaz Ahmad Taj	CUST, Islamabad
(c)	Supervisor	Dr. Abdul Basit Siddiqui	CUST, Islamabad

Dr. Abdul Basit Siddiqui
Thesis Supervisor
September, 2024

Dr. Abdul Basit Siddiqui
Head
Dept. of Computer Science
September, 2024

Dr. M. Abdul Qadir
Dean
Faculty of Computing
September, 2024

Author's Declaration

I, **Aqib Shahzaib** hereby state that my MS thesis titled "**Lesion Segmentation of Dermoscopic Images for Classification of Skin Cancer using Deep Neural Networks**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, 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.

(Aqib Shahzaib)

Registration No: MCS221016

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Lesion Segmentation of Dermoscopic Images for Classification of Skin Cancer using Deep Neural Networks**" 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.

(Aqib Shahzaib)

Registration No: MCS221016

Acknowledgement

I want to thank **God** for giving me the ability to learn, stay strong, and be patient during my studies. I'm also very thankful to my supervisor, **Dr. Abdul Basit Siddiqui**. He closely watched my progress, gave me helpful ideas, and guided me in the right direction whenever I needed it.

(Aqib Shahzaib)

Abstract

Skin diseases are prevalent health issues worldwide, posing significant challenges for physicians in distinguishing between similar lesions. This often results in time-consuming examinations, highlighting the need for automated skin disease prediction to expedite treatment planning for both patients and dermatologists.

In this study, a comprehensive approach presented to enhance automated skin lesion classification. An Efficient Hair Removal (EHR) technique was introduced that combines a Deep Residual U-Net with the TELEA inpainting algorithm, effectively eliminating hair artifacts from dermoscopic images. For precise lesion delineation, the Deep Residual U-Net model was employed for skin lesion segmentation. To address dataset imbalance, data augmentation techniques were applied, significantly improving the robustness of model.

This approach progresses through five experimental stages, each building upon the previous. Starting with dataset balancing, which improved classification accuracy by 5%, then EHR framework was applied, further boosting accuracy by 2.53%. The integration of skin lesion segmentation contributed to an additional 1.5% improvement. In the last, a modified DenseNet169 architecture was used, which achieves a top accuracy of 97.72% on the ISIC2019 dataset, outperforming existing techniques. For segmentation, our model achieved state-of-the-art results on the ISIC2018 dataset, with an Intersection over Union (IoU) of 0.8981 and a Dice coefficient of 0.946.

Contents

List of Figures

List of Tables

Abbreviations

ABCD	Asymmetry, Border, Color, Diameter
AK	Actinic keratosis
BCC	Basal cell carcinoma
BKL	Benign keratosis lentigo
DCNN	Deep Convolutional Neural Networks
DF	Dermatofibroma
EHR	Efficient Hair Removal
FMM	Fast Marching Method
GPU	Graphics Processing Unit
IoU	Intersection over Union
MEL	Melanoma
NV	Melanocytic nevus
ReLU	Rectified Linear Unit
SCC	Squamous cell carcinoma
TELEA	Telea Inpainting Algorithm
TDS	Total Dermatoscopy Score
VASC	Vascular lesion

Chapter 1

Introduction

1.1 Background

The skin plays a vital role in protecting our internal organs from harmful microbes, regulating body temperature, and facilitating sensory perception [1]. The human skin is a remarkably complex organ composed of three distinct layers: the epidermis, dermis, and hypodermis [2]. The outermost layer, the epidermis, serves as a waterproof barrier, shielding the body from external elements. This layer is responsible for imparting our unique skin tone, as it hosts specialized cells called melanocytes that produce the pigment melanin. Under the epidermis lies the dermis, a strong layer containing tough connective tissues, hair follicles, and sweat glands. This layer's robust structure contributes to the skin's overall strength and resilience. The deepest layer, the hypodermis, is composed of fat and connective tissue, providing insulation and cushioning for the body. Any disease or disorder that compromises the integrity or function of these layers can be classified as a skin condition. The complex processes among these layers underscore the importance of maintaining skin health and promptly addressing any abnormalities or imbalances that arise.

Skin cancer is a type of cancer that begins in the skin cells, skin cancer stands as one of the most prevalent cancers globally and Its incidence rate is increasing

in many countries. [3]. The primary cause of skin cancer is exposure to ultraviolet (UV) radiation, which can damage the DNA of skin cells, paving the way for cancerous transformations [4]. This harmful radiation can originate from natural sources like sunlight or artificial ones such as tanning beds. Notably, an individual's lifetime accumulation of UV exposure, both in terms of amount and intensity, directly correlates with an elevated risk of developing skin cancer [5]. Other factors that can contribute to the development of skin cancer include:

Skin type: Individuals with fair skin, light hair, and freckles are more susceptible to UV damage and have a higher risk of developing skin cancer [6].

Family history: Having a family member with skin cancer increases an individual's risk of developing the disease [7].

Immunosuppression: Skin cancer is more common in people with compromised immune systems, such as organ transplant patients or those suffering from specific autoimmune diseases [8].

Skin cancer is divided into eight distinct diagnostic categories both benign and malignant, each representing a specific skin condition or type of lesion. These categories are:

1.1.1 Actinic Keratosis

Actinic Keratosis sometimes called solar keratosis, is a skin disorder that develops as a result of extended exposure to UV radiation from the sun. It is considered a precancerous condition. This condition appears as rough, scaly, or crusty patches on the skin, primarily affecting areas that receive substantial sunlight.

Although AK itself is not malignant cancerous, it has the potential to progress into squamous cell carcinoma, if left unaddressed. Early detection and appropriate treatment of Actinic Keratosis are crucial steps in preventing its advancement to a more severe condition and mitigating the risk of developing squamous cell carcinoma. Prompt intervention is recommended to manage this precancerous lesion effectively [9].

1.1.2 Basal Cell Carcinoma

BCC is the most prevalent form of skin cancer, originating from the basal cells located in the outermost layer of the skin, known as the epidermis. It commonly appears as a pearly, pink-colored bump, often exhibiting visible blood vessels. While BCC rarely spreads to other parts of the body, it can cause significant local tissue damage and destruction if left untreated. Prompt medical attention and treatment are crucial to prevent the progression of BCC and mitigate the risk of further complications [10]. Early detection and intervention are key to managing this common type of skin cancer effectively.

1.1.3 Benign Keratosis

BKL is a category of benign skin lesions. These lesions are characterized by well-defined borders, exhibiting colors ranging from brown to black, and often have a distinctive "stuck-on" appearance on the skin surface [11]. Despite their abnormal appearance, BKLs are benign and typically do not pose a significant health risk. However, it is advisable to monitor any changes or irregularities in these lesions and seek medical advice if necessary, particularly if they exhibit concerning features or cause discomfort.

1.1.4 Dermatofibroma

DF is a benign skin lesion that manifests as a firm, red-brown colored bump or nodule. These lesions frequently appear on the legs, although they can occur on other body parts as well. DFs are relatively common skin growths that originate from the dermis, the layer of skin beneath the epidermis. They are composed of fibroblasts, which are cells responsible for producing collagen fibers, and collagen itself [12]. While Dermatofibroma may have an unusual appearance, they are generally harmless and do not require treatment unless they become irritated or cause discomfort.

1.1.5 Melanocytic Nevus

NV commonly referred to as a mole, is a benign growth of melanocytes, which are the pigment-producing cells found in the skin. These nevi can either be present from birth (congenital) or develop later in life (acquired). Melanocytic nevi exhibit diverse morphological characteristics, including junctional nevi (located at the junction of the epidermis and dermis), compound nevi (extending into both the epidermis and dermis), and intradermal nevi (situated within the dermis) [13]. While nevi are generally harmless, it is essential to monitor any changes in their appearance, size, or color, as these alterations may indicate potential concerns that require medical evaluation.

1.1.6 Melanoma

MEL is a malignant form of skin cancer that originates from melanocytes, the pigment-producing cells in the skin. It is considered the most aggressive type of skin cancer and has the potential to spread rapidly if not detected and treated at an early stage. Melanomas often exhibit irregular borders, asymmetrical shapes, and color variations. Although melanoma accounts for only 4% of all skin cancer cases, it is responsible for 80% of skin cancer-related deaths. Patients diagnosed with metastatic melanoma have a five-year survival rate of just 15%. However, if melanoma is detected early, it has a remarkable 95% cure rate, highlighting the significance of early diagnosis in significantly improving survival prospects for patients [14].

1.1.7 Squamous Cell Carcinoma

SCC is the second most prevalent form of skin cancer, originating from the squamous cells located in the outermost layer of the skin, known as the epidermis. It typically appears as a scaly, reddish, or crusted patch or bump on the skin surface. While SCC is less aggressive than melanoma, it has the potential to spread to other

parts of the body if left untreated [15]. Early detection and prompt treatment are crucial in managing SCC effectively and preventing its potential metastasis, which can lead to more severe complications. Regular skin examinations and addressing any suspicious lesions promptly are recommended to mitigate the risks associated with this type of skin lesion.

1.1.8 Vascular Lesions

VASC refer to a group of benign skin lesions that involve abnormalities in the blood vessels. These lesions include conditions such as cherry angiomas, angiokeratomas, and pyogenic granulomas. The defining characteristic of vascular lesions is an increased number or dilation (widening) of blood vessels within the skin [16]. While these lesions may appear concerning due to their vascular nature, they are generally harmless and non-cancerous. However, it is advisable to monitor any changes or irregularities in these lesions and seek medical advice if necessary, particularly if they exhibit unusual symptoms or cause discomfort.

To illustrate the various types of skin lesions, example dermoscopic images are presented in 1.1.

Traditionally, dermatologists have relied on visual examination of skin lesions for diagnosis. However, dermoscopy, also known as dermatoscopy or epiluminescence microscopy, has emerged as an imaging technique that facilitates the elimination of skin surface reflections [17]. This technique employs a dermatoscope, a hand-held device equipped with a magnifying lens and an integrated light source, which enables enhanced visualization of the area of interest, Dermatoscope displayed in 1.2. https://www.alamy.com/. By eliminating surface reflections, dermoscopy enhances the visualization of deeper layers of the skin, providing valuable insights into the characteristics of skin lesions [18]. Numerous studies have demonstrated that when dermatologists utilize this technique, it can lead to higher diagnostic accuracy compared to standard photography [19]. To improve the diagnostic efficacy of skin diseases, dermoscopy has been introduced as a valuable tool to assist dermatologists in the diagnostic process [20]. The revealing characteristics of skin

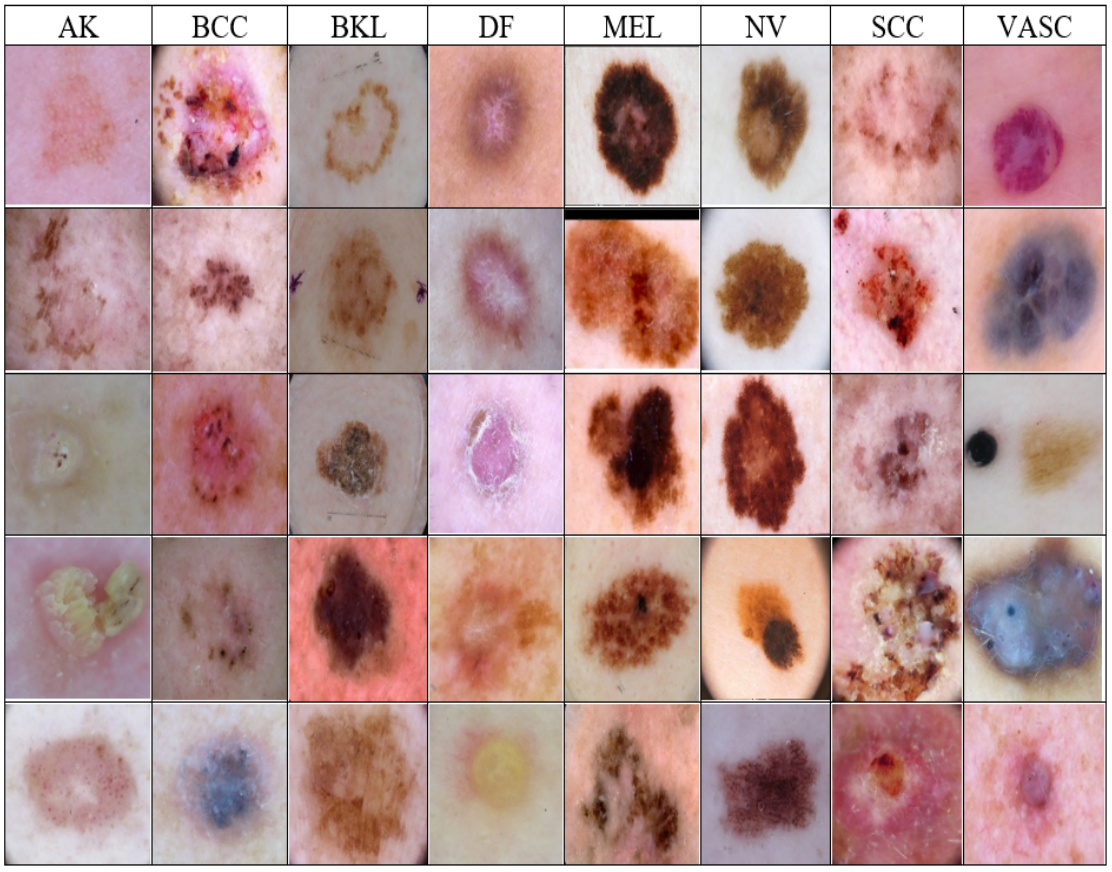


FIGURE 1.1: Types of Skin Lesions.

lesions that aid in diagnosis include the size, shape, color, edge, boundary, and location of the abnormality, as well as the presence or absence of other symptoms or signs. However, it is important to note that the experience of the dermatologist can influence the examination process.

In clinical medicine, there exists a certain degree of error in the naked-eye recognition rate of lesion images, particularly in the case of melanoma, when examined through dermoscopy [21]. Dermatologists are required to possess specific technical expertise and experience to accurately recognize skin cancer. The fault-tolerant rate of skin lesion recognition is relatively low, and false identification can have severe consequences for patients [22],[23]. Consequently, there is a need for advanced diagnostic methods that can enhance the accuracy and reliability of skin lesion recognition, minimizing the potential for misdiagnosis and its associated risks.

To aid practitioners in the diagnostic process, standard methods like the ABCD

FIGURE 1.2: Digital dermatoscope used for dermatoscopy.

rule (Asymmetry, Border, Color, Diameter) [24] or the 7-point checklist [25] offer guiding principles and algorithms.

The initial attempt to streamline the diagnosis procedure was made with the ABCD rule of dermatoscopy, which was the second algorithm created after pattern analysis. After a rigorous multivariate analysis of 31 dermoscopic criteria, four criteria were identified as significant co-factors for diagnosing melanoma: asymmetry (A), borders (B), colors (C), and different structural components (D). This semi-quantitative, mathematical approach assigns points for the identified criteria in a lesion and employs a formula to determine the total dermatoscopy score for each lesion. A TDS less than or equal to 4.75 typically indicates a benign lesion,

while a TDS between 4.80–5.45 is suggestive but not diagnostic of melanoma, warranting either excision or follow-up monitoring. With the emergence of digital systems, clinicians now create a database of lesions for follow-up, enabling side-by-side computer-monitor comparisons of baseline and follow-up images to detect significant dermoscopic changes over time. A TDS greater than or equal to 5.45 is highly suspicious but not conclusive for melanoma, as benign lesions can occasionally yield false-high TDS values [25]. 1.1 was formulated to calculate TDS.

$$TDS = (A \times 1.3) + (B \times 0.1) + (C \times 0.5) + (D \times 0.5). \tag{1.1}$$

Another form of pattern analysis that uses a point system is the 7-Point Checklist. It involves identifying and analyzing fewer criteria than pattern analysis, and the point system is less complicated than the ABCD rule of dermatoscopy. Major criteria receive 2 points each, while minor criteria receive 1 point. A score of 3 or greater has a 95% sensitivity for melanoma.

These methods can be used to differentiate between melanoma and non-melanoma. Dermatologists should be involved in classifying the 8 types of skin lesions. However, there exists an inherent risk that medical professionals may misinterpret an image as indicative of various forms of skin cancer. Consequently, there is a critical need for an automated computational system capable of analyzing datasets visually, to replicate and potentially surpass the expertise of human experts through a data-driven approach.

The manual interpretation of medical images by professionals is susceptible to human error and subjectivity, which can lead to incorrect diagnoses or wrong classifications. These errors can have severe consequences for patient outcomes, particularly in the context of skin cancer, where early and accurate detection is crucial for effective treatment and improved prognosis.

To mitigate this risk and enhance diagnostic accuracy, the development of an automated computational system is imperative. Such a system would leverage advanced computational techniques, such as machine learning and computer vision algorithms, to analyze visual datasets of skin lesions. By training these algorithms

on vast amounts of annotated data, the system can learn to recognize intricate patterns and features that are characteristic of various skin lesions, including malignant and benign conditions.

The data-driven approach used by this automated system will enable it to continuously improve its performance by incorporating new data and an ever-changing understanding of the characteristics of skin lesions. Ultimately, the goal is for the system to achieve, and possibly surpass, the diagnostic expertise of human experts, thereby reducing the risk of misinterpretation and misclassification.

The successful implementation of such an automated computational system would not only improve diagnostic accuracy but also facilitate more efficient and consistent screening processes, potentially leading to earlier detection and better patient outcomes in the management of skin cancer.

1.2 Motivation

Skin cancer has emerged as the most pervasive form of cancer in contemporary times, its reach extending to every corner of the globe. Each year, it is estimated that various manifestations of skin cancer, one of the most damaging strains of skin cancer, account for an astounding 3.5 million newly diagnosed cases worldwide. [26], [27]. However, when detected at an early stage, the five-year survival rate for melanoma patients rises to an impressive 92 percent. [28]. In contrast, delayed diagnosis of melanoma often translates into significant increases in treatment costs. [29].

The human factor in the visual evaluation of dermatoscopic images or manual dermoscopy by a dermatologist introduces an inherent limitation. The subjective nature of human decision-making, compounded by the considerable inter-class similarity among skin lesions and other confounding factors, renders this method susceptible to errors. General diagnostic procedures for identifying skin cancer, such as the ABCD rule [24] or the 7-point checklist [25], can only provide practitioners with broad guidelines. The potential for misinterpretation and misclassification

persists, as medical professionals may categorize the same dermatoscopic image sample as belonging to different types of skin cancer.

Artificial neural networks are complex computational models that inherit the architecture of biological neurons found in the human brain [30]. Convolutional neural networks (CNNs) represent a specialized variation of feedforward neural networks, specifically designed for image classification tasks [31]. With the rapid advancement in computational power, artificial intelligence, and deep learning techniques have been widely adopted [32], particularly in the realm of image processing, where they have achieved remarkable progress [33]. Deep convolutional neural networks (DCNNs) have demonstrated an exceptional ability to accurately detect, segment, and identify regions and objects within medical images [34]. The clinical medicine domain has increasingly embraced deep learning technology, yielding outstanding results in image segmentation and pattern recognition-based image classification [35].

1.3 Significance of the Proposed Solution

This study introduces an Innovative framework for classifying skin cancer into eight distinct categories. The proposed system integrates three key components: Hair removal from dermoscopic images which includes Deep Residual U-Net for hair mask image generation, TELEA (FMM) [36] for inpainting, Deep Residual U-Net for lesion segmentation and a modified DenseNet169 for classification. This comprehensive approach aims to maximize both accuracy and efficiency while fully utilizing the information present in dermoscopic images. Key Components are

1.3.1 Hair Removal

Step 1 : A Deep Residual U-Net is employed for precise hair segmentation.

Step 2 : The TELEA inpainting method is used to reconstruct areas where hair has been removed.

1.3.2 Skin Lesion Segmentation

A Deep Residual U-Net utilized for skin lesion segmentation.

1.3.3 Classification

An advanced deep learning model DenseNet169 adapted for skin lesion classification.

DenseNet169 is a convolutional neural network architecture introduced by Huang et al [37]. in their 2017 paper "Densely Connected Convolutional Networks". It is part of the DenseNet family, which was designed to address the vanishing gradient problem and feature reuse in very deep networks.

The proposed framework tackles the complex challenge of accurate skin cancer classification. By integrating the top layers, the modified DenseNet169 can adaptively refine its interpretation of lesion features, leading to more robust and precise classifications. This innovative approach has the potential to significantly advance the field of automated skin cancer diagnosis.

By combining hair removal techniques with this classification model, this framework offers a comprehensive solution to the challenges of skin cancer diagnosis from dermoscopic images.

1.4 Thesis Layout

The rest of the thesis is organized as follows:

1.4.1 Chapter 2

The second chapter of this study offers a comprehensive and critical examination of the existing state-of-the-art methodologies for classifying skin cancer into eight distinct categories. This in-depth review serves as a crucial foundation for

understanding the current landscape of skin cancer classification techniques and identifying areas for potential improvement.

1.4.2 Chapter 3

In this chapter hair removal framework is discussed.

1.4.3 Chapter 4

Chapter 4 of this study presents a detailed exposition of our methodological approach.

1.4.4 Chapter 5

Chapter 5, presents a comprehensive overview of our findings and engages in a nuanced discussion of the results. This chapter serves as the culmination of our research efforts, offering valuable insights into the performance and efficacy of our proposed skin cancer classification framework.

1.4.5 Chapter 6

The concluding chapter of this study serves as a comprehensive synthesis of our research findings, distilling key insights and drawing meaningful conclusions from the outcomes presented in previous chapters.

Chapter 2

Literature Review

This section provides an extensive overview of the most relevant and innovative methods used in the diagnosis of different forms of skin cancer, with an emphasis on the use of machine learning and deep learning approaches.

Our review aims to give an exhaustive understanding of the current state of automated skin cancer diagnosis research.

2.1 Survey of Existing Techniques

The fact that many forms of skin cancer have similar outward appearances makes it difficult to classify them, which is one of the main challenges in identifying the disease.

In the study conducted by [38], the researchers developed a comprehensive approach for skin lesion classification, incorporating several preprocessing techniques and machine learning algorithms. The researchers first implemented a digital hair removal technique, which utilized morphological filtering operations, specifically the Black-Hat transformation. This was followed by an inpainting algorithm to reconstruct the areas where hair had been removed. To further enhance image quality, they applied Gaussian filtering, which served to deblur and denoise the images. The researchers employed Grabcut segmentation and the Gray Level Co-occurrence Matrix (GLCM) technique for feature extraction. Three ML techniques

(SVM, DT, and KNN) are applied for classification. Initially, when applied to the imbalanced ISIC2019 dataset, the accuracy of their approach fell below 75%, to address this issue oversampling was used to balance the dataset and achieved the accuracy of 95% with SVM.

Hasan et al. [39] put forward a classification strategy tailored for skin disease categorization that leverages convolutional neural networks. Their approach facilitates multi-class skin lesion classification by employing the EfficientNetsB0-B7 architecture in conjunction with transfer learning techniques. The model's initial parameter values were derived from prior training on the ImageNet dataset, and the evaluation was conducted using the ISIC2019 dataset. The outcomes indicated that the proffered methodology could effectively classify multi-class skin lesions, achieving an accuracy of approximately 93.5%, a precision of 92.5%, a recall of 94.5%, and an F-Score of 93.5%. Additionally, digital hair removal was applied to the images. The study concluded that editing of top layers and fine-tuning of models significantly enhanced performance, the models of intermediate complexity, such as EfficientNetB5, outperformed others in classification performance.

In their study [40], Villa-Pulgarin et al. presented an approach that concentrates on skin lesion classification by employing optimized Convolutional Neural Network (CNN) models based on transfer learning. The architectures utilized were DenseNet201, Inception-V3, and Inception-ResNet-V2. The methodology involved conducting four experiments to analyze the classification performance, yielding accuracies of 98%, 97%, and 96% for the respective models when evaluated on the HAM10000 dataset. The best-performing model yielded 93% accuracy with the ISIC 2019 dataset. The research demonstrated how transfer learning and fine-tuning optimization can improve model performance for tasks involving the classification of skin lesions. Notably, the optimized DenseNet-201 model demonstrated high proficiency in classifying skin lesions on the ISIC 2019 dataset, with the incorporation of data augmentation techniques.

Rezaoana et al. [41] introduced an automated methodology for skin cancer classification that integrates image processing techniques with deep learning approaches. To bolster the dataset, they employed various image augmentation strategies to

enrich and diversify the available data. Furthermore, they leveraged transfer learning techniques to enhance the accuracy of their classification tasks. The proffered convolutional neural network (CNN) method yielded impressive results, attaining a weighted average precision of approximately 0.76, a 0.78 weighted average recall, a 0.76 weighted average f1-score, and a 79.45% overall accuracy.

In their work [42], Naveed et al. propose a progressive class-wise attention method for skin lesion classification tasks that is robust in its ability to generalize on unseen data and is end-to-end trainable. Their method addresses challenges such as inter-class similarity, significant intra-class variation, and unequal distribution of classes, which are commonly encountered in skin lesion classification problems. By focusing on discriminative areas of dermoscopic images, their progressive class-wise attention approach enhances the performance of skin lesion classification models. Furthermore, they employ Grad-CAM to acquire an understanding of the acquired solutions of their approach, offering significant assistance and interpretability to medical professionals. The proffered method achieved an impressive accuracy rate of 94.9% in their experiments.

Khan et al. [43] developed a technique that operates on two parallel streams. In the first stream, they present a contrast enhancement technique based on fusion, which enhances the images before feeding them into DenseNet201's pre-trained architecture. The features that were from this stream undergo further optimization using a moth-flame optimization technique with skewness control. Concurrently, in the subsequent stream, profound features are extracted from a fine-tuned and pre-trained MobileNetV2 network which are then down-sampled using their suggested feature selection methodology. After that, a unique parallel multimax coefficient correlation approach is employed to fuse the most unique features from both networks. Finally, the lesion images are classified using a multiclass extreme learning machine classifier. This comprehensive two-stream approach achieved an impressive accuracy rate of 89% on the ISIC2019 dataset in their experiments.

Li et al. [44] proposed an adaptive multi-modal fusion (AMF) method that integrates dermoscopy images with metadata information for skin lesion classification. The AMF technique dynamically adjusts the weight distribution between these

two modalities to achieve balanced optimization. However, the researchers noted limitations concerning the optimal ratio between the imaging modality and the metadata modality, suggesting that further exploration is needed in this area. They emphasized the significance of dermoscopy image information in the classification process, considering it more important than metadata. Despite these challenges, the proposed AMF method achieved an impressive accuracy rate of 89.09% and an AUC of 91.55% in their experiments.

Zhao et al. [45] introduced a novel technique for image augmentation in skin lesion classification. They developed a skin lesion image classification model using a GAN-based data augmentation method, aiming to enhance the performance of skin lesion classification. Specifically, they proposed a skin lesion augmentation SLA-StyleGAN model for medical image synthesis, based on the original architecture of style-based GANs. Additionally, they constructed a classifier using DenseNet201 based on transfer learning principles. Their results demonstrated that the DenseNet201 model performed effectively in the ISIC2019 classification task, with further improvements achieved through the proposed SLA-StyleGAN approach. The proposed method achieved an impressive accuracy rate of 93.64%.

In their study [46], Santos et al. explored class balancing through data augmentation techniques for skin lesion classification. They employed an Ensemble approach that combined three models: EfficientNetB2, DenseNet121, and InceptionResNetV2. This strategy aimed to leverage the strengths of multiple models to enhance classification performance. The Ensemble technique led to an accuracy of 89.1%, demonstrating the effectiveness of their class balancing and data augmentation methods in improving the overall classification accuracy for skin lesion analysis. Table 2.1 describes the summary of existing techniques.

2.2 Problem Statement

Accurate classification of skin cancer using dermoscopic images is crucial for better diagnosis. The irregular shapes, diversity of color, border, and size make this task

very challenging. The skin hair add another level of complexity to lesion detection. There is a significant need to enhance the ability of deep neural models for the better classification of dermoscopic images.

Table 2.1: Comparative analysis of existing techniques.

Ref.	**Technique**	**Dataset**	**Results**
Naveed et al. [42] (2024)	Progressive Class-wise attention mechanism and DenseNet-121	ISIC2019	Accuracy = 94.9%, Precision = 94.9%, F1-Score = 94.7%.
Hasan et al. [39] (2023)	Modified EfficientNets B0 - B7	ISIC2019	Accuracy = 93.5%, Precision = 92.5%, Recall = 94.5%, F1-Score = 93.5%.
Villa-Pulgarin et al. [40] (2023)	DenseNet-201, Inception-V3, and Inception-ResNet-V2	ISIC2019	Accuracy = 93.0%, Precision = 93.0%, Recall = 93.0%, F1-Score = 93.0%.
Li et al. [44] (2023)	Adaptive Multi-modal Fusion (AMF) method, also use meta-data	ISIC2019	Accuracy = 89.0%, Precision = 86.3%, AUC = 91.5% , F1-Score = 86.7%.
Ahammed et al. [38] (2022)	DHR, Grabcut segmentation, GLCM for Feature Extraction and SVM,KNN and DT as classifier	ISIC2019	Accuracy = 95.0%, Precision = 95.1%, Recall = 95.0%, F1-Score = 94.8%.
khan et al. [43] (2022)	Two-stream cnn-based framework	ISIC2019	Accuracy = 89.0%.
ZHAO et al. [45] (2021)	SLA-StyleGAN and Densenet201	ISIC2019	BMA = 93.6%.
Santos et al. [46] (2021)	Model Ensemble	ISIC2019	Accuracy = 89.1%, BMA = 84.6%.
Rezaoana et al. [41] (2020)	Proposed CNN	ISIC2019	Accuracy = 79.4%, Precision = 76.0%, Recall = 78.0%, F1-Score = 76.0%.

2.3 Research Questions

RQ1: How hair mask of a dermoscopic image can be generated?

RQ2: How hair removal from a dermoscopic image can be done for better lesion segmentation?

RQ3: What is the effect on the classification results of the preprocessing tasks done in RQ1, and RQ2?

Chapter 3

Hair Removal with Efficient Hair Removal Framework

This Figure 3.1 provides a visual representation of our workflow for the skin lesion classification task. This diagram offers a clear and concise overview of the various stages and processes involved in our research methodology. The figure illustrates the step-by-step progression of our work, from initial data preparation to the final classification output. It serves as a roadmap, guiding the reader through the key components of our approach.

The process began with the ISIC2019 dataset, which contains dermoscopic images. The first step in preprocessing is hair removal. Hair was removed to ensure the model focused on the actual lesion. Next, the lesion was segmented from the surrounding skin to highlight the lesion area. Data augmentation was then applied. Finally, the DenseNet169 architecture was used for skin cancer classification.

3.1 Efficient Hair Removal Framework

The appearance of hair on skin lesions produces a significant challenge in dermoscopic image analysis. These fine, often dark structures can obscure crucial

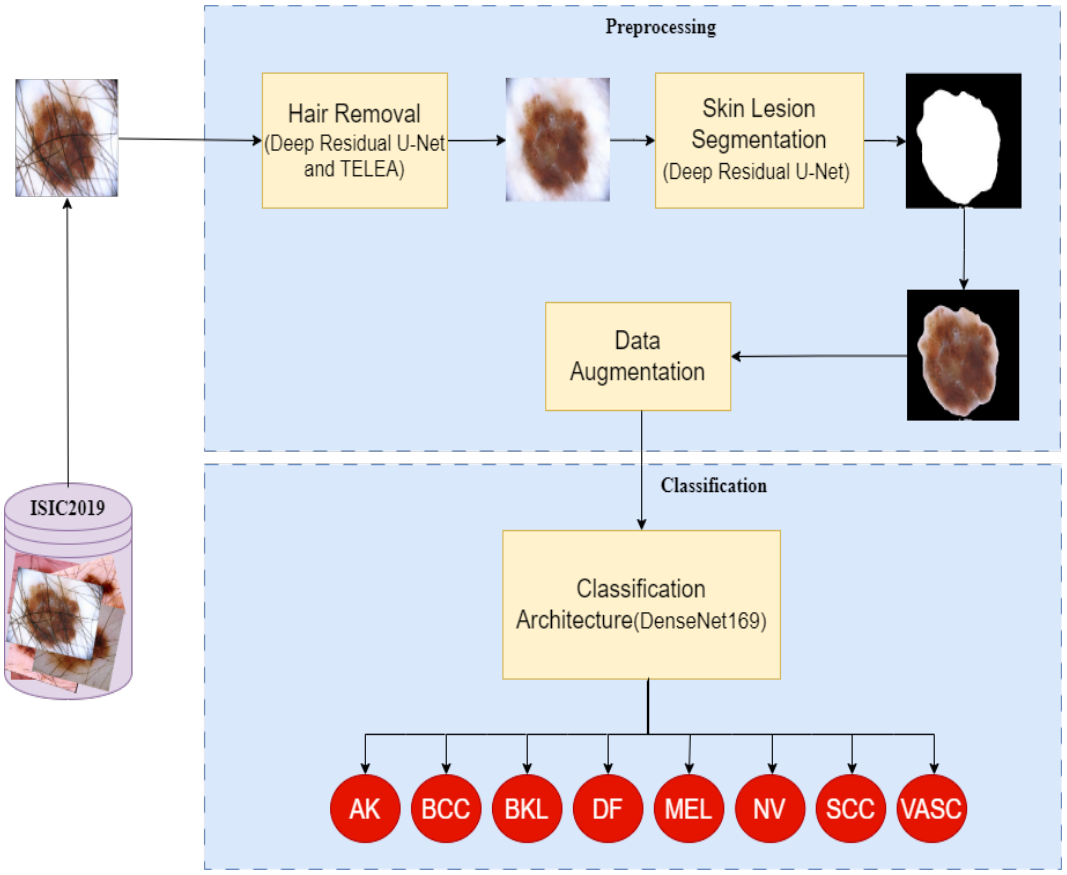


FIGURE 3.1: Flow of our Skin Lesion Classification task.

features of the lesion, potentially leading to misclassification or overlooked diagnoses. The variability in hair density, color, and distribution across images further complicates the classification task.

To address this issue, we have developed an innovative solution: The Efficient Hair Removal framework. This advanced system employs a two-step approach to mitigate the impact of hair on dermoscopic images. First, it utilizes sophisticated algorithms to accurately detect or segment hair structures within the images. After detection, the EHR framework used an intelligent Inpainting technique to fill in the gaps identified by the segmentation model, effectively reconstructing the underlying skin and lesion features.

This approach not only enhances the visibility of critical lesion characteristics but also standardizes the images, reducing the variability introduced by hair presence. By effectively removing this confounding factor, the EHR framework aims

TABLE 3.1: State of the art Hair Removal Algorithms

Year	**Ref.**	**Hair Segmentation**	**Inpainting Method**
1997	Lee et al. [47]	Grayscale closing	Bilinear interpolation
2009	Xie et al. [48]	Grayscale top-hat	Non-linear PDE
2011	Abbas et al. [49]	Derivative of Gaussians	Coherence transport
2013	Huang et al. [50]	Conventional matched filters	Region growing algorithms
2017	Bibiloni et al. [51]	Color top-hat	Morphological Inpainting

to improve the accuracy and reliability of subsequent lesion classification tasks, potentially leading to more precise diagnoses and improved patient outcomes.

The challenge of hair removal from dermoscopic images has been addressed by numerous researchers employing various computer vision techniques. We focus on several state-of-the-art algorithms that have gained prominence due to their availability and widespread adoption in the field. Notable contributions include the methods proposed by Lee et al. [47], Xie et al. [48], Abbas et al. [49], Huang et al. [50], and Bibiloni et al. [51].

These approaches represent a diverse range of strategies for tackling the hair removal problem in dermoscopic imagery. Each method brings its unique perspective and innovative techniques to address this critical preprocessing step in skin lesion analysis. To provide a comprehensive overview of these methodologies, we have compiled a summary table 3.1 that encapsulates the key approaches employed by each of these researchers. This table serves as a valuable reference point, highlighting the various techniques and their specific contributions to the field of hair removal in dermoscopic image analysis.

Our analysis revealed that the efficacy of hair removal is heavily dependent on the precision of hair segmentation. Recognizing this crucial relationship, we turned to a specialized dataset [52] to develop and validate our approach. This dataset, derived from the ISIC (International Skin Imaging Collaboration) database, comprises 500 dermoscopic images featuring visible hair.

The dataset's strength lies in its comprehensive nature, providing not only the original hair-containing images but also corresponding hair masks and overlaid images. This three-pronged approach - original images, hair masks, and overlays - offers a robust foundation for both training and evaluating hair detection algorithms.

By utilizing this well-curated dataset, we aimed to enhance the accuracy of our hair segmentation process, which in turn would lead to more effective hair removal. This strategic choice of training data underscores our commitment to addressing the hair removal challenge with precision and thoroughness, setting the stage for improved dermoscopic image analysis. We used a pair of **dermoscopic image** and **hair mask** as input for our model.

3.1.1 Hair Mask Image Generation

For our hair detection task, we employ Deep Residual U-net which is according to the original architecture of [53], an advanced semantic segmentation neural network that ingeniously combines the strengths of residual neural network and U-Net designs. This hybrid approach offers two key advantages:

3.1.1.1 Enhanced Training Efficiency

The incorporation of residual units significantly eases the training process of the network. This feature allows for more stable and efficient learning, particularly in deeper network configurations.

3.1.1.2 Improved Information Flow

The network's design includes skip connections within residual units and between low and high levels of the network. This architecture facilitates unimpeded information propagation throughout the network, mitigating the problem of information degradation commonly seen in deep neural networks.

Our implementation of Deep Residual Unet for skin hair detection or segmentation utilizes a sophisticated 7-level architecture, as illustrated in Figure 3.2. This network is structured into three primary components: encoding, bridge, and decoding.

The encoding section serves to compress the input image into compact, high-dimensional representations. This process distills the essential features of the image, creating a condensed form of the information. At the other end, the decoding section performs the inverse operation. It expands these compact representations back into a pixel-wise categorization, effectively producing a semantic segmentation of the original image. Connecting these two sections is the bridge component, which facilitates the flow of information between the encoding and decoding paths.

TABLE 3.2: Network structure of Deep Residual Unet.

	Unit level	**Conv layer**	**Filter**	**Stride**	**Output size**
Input					224 x 224 x 3
Encoding	Level 1	Conv 1	3 x 3/64	1	224 x 224 x 64
		Conv 2	3 x 3/64	1	224 x 224 x 64
	Level 2	Conv 3	3 x 3/128	2	112 x 112 x 128
		Conv 4	3 x 3/128	1	112 x 112 X 128
	Level 3	Conv 5	3 x 3/256	2	56 x 56 x 256
		Conv 6	3 x 3/256	1	56 x 56 x 256
Bridge	Level 4	Conv 7	3 x 3/512	2	28 x 28 X 512
		Conv 8	3 x 3/512	1	28 x 28 X 512
Decoding	Level 5	Conv 9	3 x 3/256	1	56 x 56 x 256
		Conv 10	3 x 3/256	1	56 x 56 x 256
	Level 6	Conv 11	3 x 3/128	1	112 x 112 x 128
		Conv 12	3 x 3/128	1	112 x 112 x 128
	Level 7	Conv 13	3 x 3/64	1	224 x 224 x 64
		Conv 14	3 x 3/64	1	224 x 224 x 64
Output		Conv 15	1 x 1	1	224 x 224 x 1

A key feature of this architecture is that all three components are constructed using residual units. Each unit comprises two 3x3 convolution blocks and an identity mapping. The convolution blocks are structured with a specific sequence: a batch normalization (BN) layer, followed by a ReLU activation layer, and finally

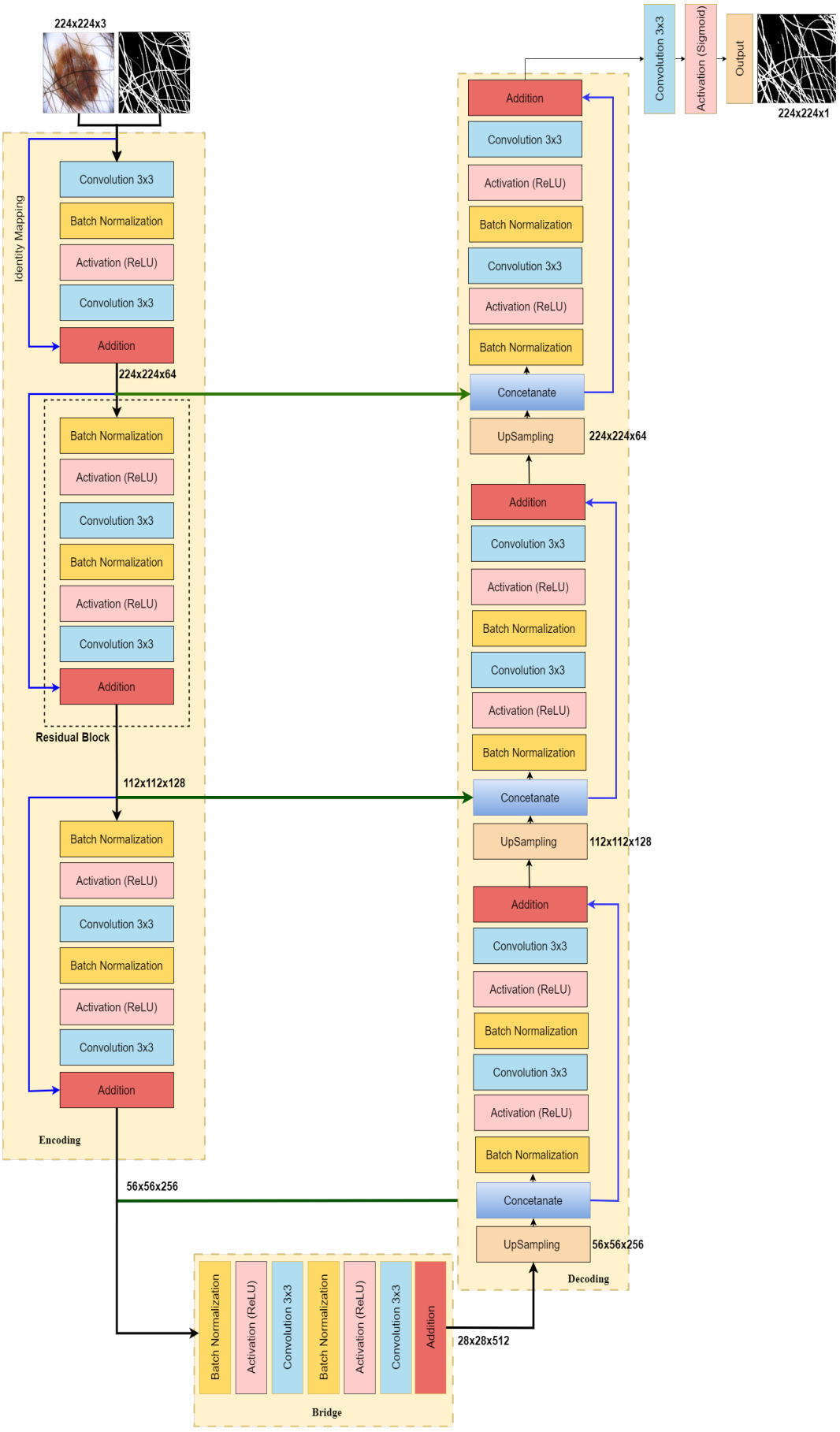


FIGURE 3.2: Architeture of Deep Residual Unet.

a convolutional layer. The identity mapping creates a direct connection between the input and output of each unit, allowing for unimpeded information flow. This carefully designed architecture enables the network to effectively capture and process the intricate details necessary for accurate hair detection and segmentation in dermoscopic images.

The encoding path of our Deep Residual Unet comprises three residual units, each designed to progressively compress the input information. Instead of traditional pooling operations for downsampling, we apply a stride of 2 with the first convolution block within each unit. This approach effectively halves the feature map size at each step, creating a more compact representation.

Mirroring this structure, the decoding path consists of three residual units. However, before each unit in the decoding path, we implement an up-sampling operation on the feature maps before residual blocks from the lower level. The matching feature maps from the encoding path are then concatenated with these up-sampled features, facilitating the reconstruction of spatial information.

A sigmoid activation layer and a 1x1 convolution are the last steps in the decoding pipeline. By projecting the multi-channel feature maps into the intended segmentation output, this combination successfully categorizes each pixel.

Our Deep Residual Unet architecture comprises a total of 15 convolutional layers, a notable reduction from the 23 layers found in the standard U-Net. For a comprehensive understanding of the network's structure, we have compiled 3.2, It describes the specifications and output sizes at every stage of the procedure. This table provides a clear overview of how information flows and is transformed throughout the network.

3.1.2 Hair Gap Inpainting

Inpainting, the process of reconstructing missing or degraded portions of images, can be achieved through various methods. These approaches can be broadly categorized into traditional computer vision techniques and more recent deep learning-based solutions. Among the traditional computer vision methods, two notable

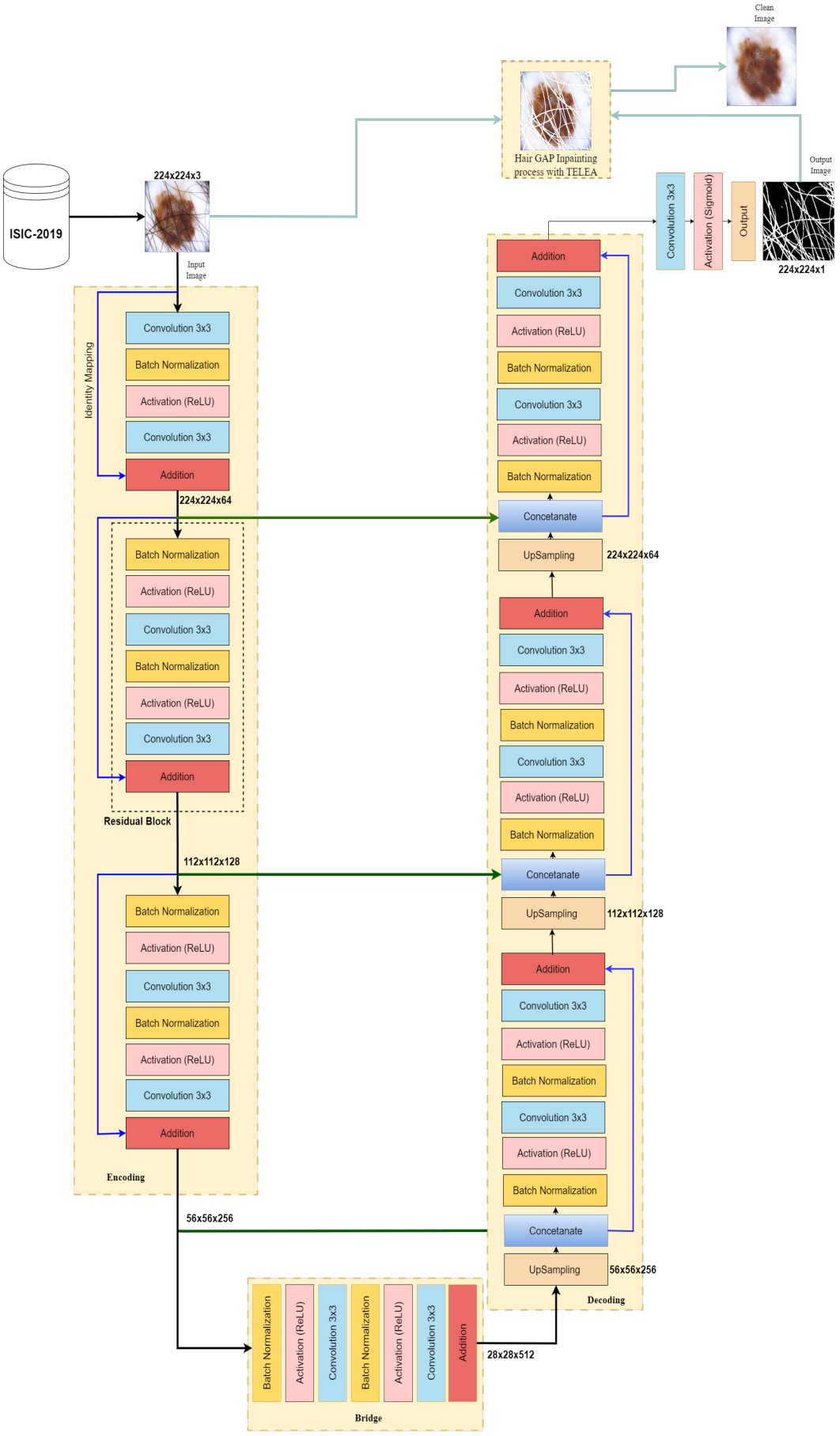


FIGURE 3.3: Efficient Hair Removal Framework.

approaches stand out: Navier-Stokes [54] and Telea [36]. These algorithms have been widely used and have proven effective in many inpainting scenarios.

In the realm of deep learning, two state-of-the-art techniques have gained prominence: SinGAN [55] and R-MNet [56]. These advanced methods leverage the power of neural networks to achieve high-quality inpainting results.

Interestingly, a comparative study [57] has shown that the performance of R-MNet is comparable to that of the Telea method. This finding suggests that despite the sophistication of deep learning approaches, traditional methods can still hold their own in certain inpainting tasks. Given this equivalence in performance and considering factors such as computational efficiency and ease of implementation, we have opted to use the Telea method as our inpainting algorithm of choice. This decision allows us to achieve high-quality results while potentially benefiting from the lower computational demands of traditional methods.

The Telea inpainting algorithm [36] employs an FMM, offering a more efficient and computationally less complex solution compared to other typical inpainting techniques [36]. This method ingeniously expands inpainted areas into target regions gradually using known regions, a process facilitated by the FMM. At its core, the FMM is a numerical method designed to solve boundary value problems [58]. In the context of inpainting, through the process, it maintains a narrow band, or boundary between known and unknown locations, by ensuring that pixels near known neighbors are processed before those with unknown neighbors.

This iterative process ensures a gradual, natural-looking reconstruction of the unknown areas, leveraging the information from known regions to inform the inpainting of target areas. The method's efficiency and effectiveness make it a valuable tool for hair removal in skin lesion images, where maintaining the integrity of the underlying skin features is crucial.

The comprehensive process of our EHR framework is visually represented in Figure 3.3. This illustration provides a detailed overview of the entire hair removal workflow, offering a clear and intuitive understanding of each step in our methodology.

Chapter 4

Data Augmentation, Skin Lesion Segmentation and Classification

4.1 Data Augmentation

Enhancing Dataset Volume and Balance, to optimize the training of our CNN model, we employed data augmentation techniques, which play a crucial role in maximizing data volume and balancing class distributions. This approach exposes the model to a variety of input data variations, improving the generalizability of the model as well as increasing the number of training samples[59].

Our augmentation strategy focused on offline techniques, applying a range of transformations including flips, zooms, width and height shifts, and rotations. The specific parameters used for these augmentations are detailed in TABLE 4.1. This carefully curated set of transformations was designed to preserve the original consistency and integrity of the input-output relationship while introducing meaningful variations. After conducting a thorough examination and executing multiple combinations of parameters, the augmentation settings detailed in 4.1 were ultimately determined to be optimal.

The augmentation process offers multiple benefits:

TABLE 4.1: Augmentation settings

Augmentation Setting	**Range**
rotation_range	90
width_shift_range	0.2
height_shift_range	0.2
zoom_range	0.2
horizontal_flip	True
fill_mode	reflect

4.1.1 Increased Data Volume

By generating additional images, we expanded our training set, providing more examples for the model to learn from.

4.1.2 Class Balancing

Augmentation helped address class imbalance issues, ensuring each category was adequately represented in the training data.

4.1.3 Improved Generalization

By exposing the model to various image transformations, we enhanced its ability to recognize key features across different orientations and scales.

4.1.4 Preservation of Essential Characteristics

Our chosen augmentation techniques were selected to maintain the critical features of skin lesions while introducing variability. This approach to data augmentation avoids distortions that could compromise the medical relevance of the images. By implementing these data augmentation techniques, we aim to build a more robust and versatile model capable of accurately detecting and classifying skin lesions across a wide range of presentations.
Figure 4.1 provides a clear visual representation of the impact of our data augmentation strategies. This visual representation provides a clear and comprehensive

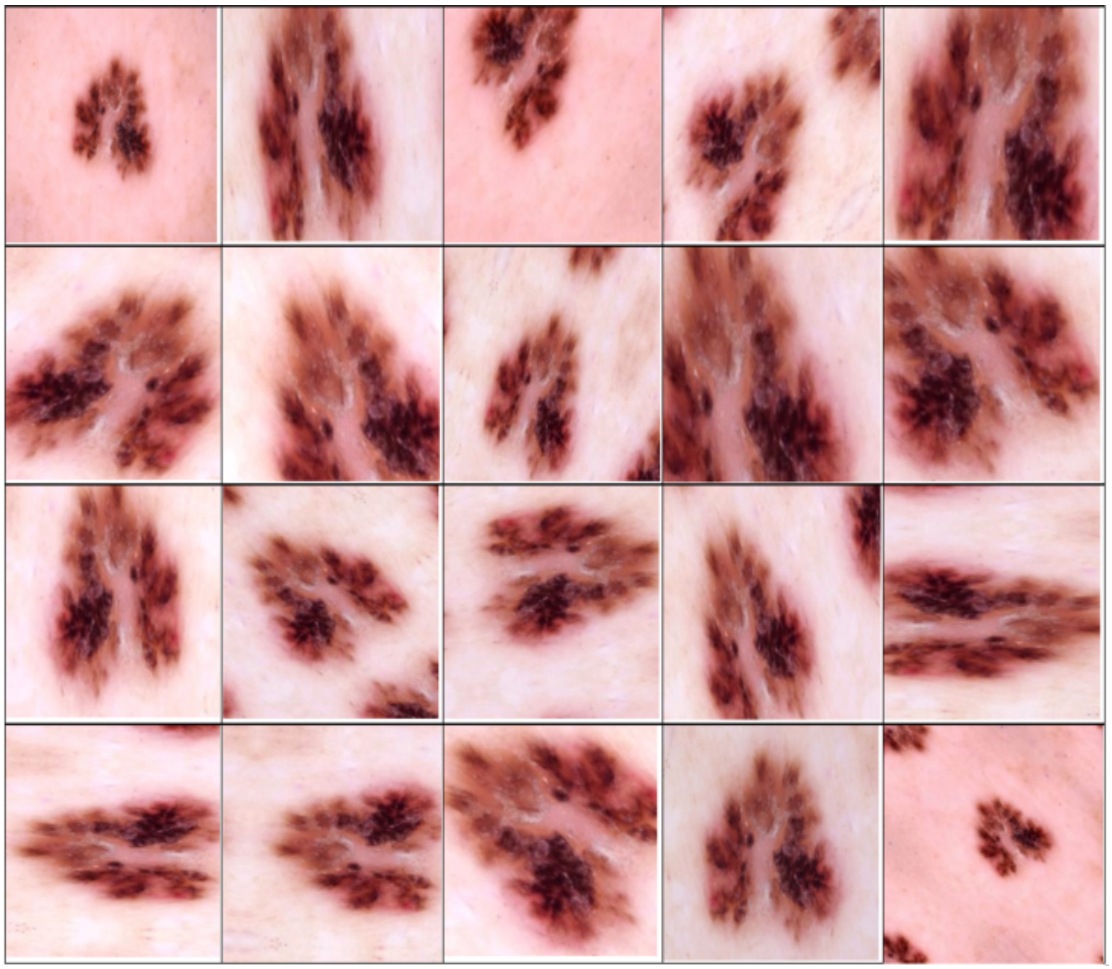


FIGURE 4.1: The Effect of Augmentation on Image.

showcase of how our augmentation methods transform and diversify the original dermoscopic images.

The visualization demonstrates how our augmentation methods maintain the integrity of essential skin lesion characteristics while introducing variability. It demonstrates how these techniques expand our dataset. This balance is crucial for creating a robust training dataset without compromising the medical relevance of the images.

4.2 Skin Lesion Segmentation

Accurately identifying and outlining skin lesions is a crucial yet challenging step in automated skin analysis. Many doctors use rule-based systems to diagnose skin

conditions, and these systems depend heavily on precise lesion boundaries. Without accurate outlines, it is difficult to assess key diagnostic factors like asymmetry, border irregularity, and size - all essential for applying common diagnostic tools such as the ABCD method [60] [24].

When examining large-area skin images that may contain multiple suspicious spots, being able to automatically detect and isolate individual lesions is incredibly helpful. Yet despite how important this task is, manually drawing lesion boundaries remains tedious and prone to inconsistency, even when done by the same person multiple times. Different experts may also produce varying results when outlining the same lesion [61]. Given these challenges, there's a clear need for a quick, dependable, and automated way to outline skin lesions. Such a tool would not only save time but could also provide more consistent results, potentially improving diagnostic accuracy.

The ISIC2019 dataset lacks lesion segmentation masks, which are crucial for training segmentation models. To overcome this limitation, researchers often turn to the ISIC2018 dataset as an alternative. This earlier dataset offers a substantial collection of 10,015 dermatological images, each paired with its corresponding lesion mask. The availability of these matched image-mask pairs makes the ISIC2018 dataset a valuable resource for training and developing skin lesion segmentation models.

4.2.1 Deep Residual U-Net for Skin Lesion Segmentation

Our implementation of Deep Residual Unet for skin lesion segmentation utilizes a sophisticated 7-level architecture, as illustrated in 4.2. This network is structured into three primary components: encoding, bridge, and decoding. The encoding section serves to compress the input image into compact, high-dimensional representations. This process distills the essential features of the image, creating a condensed form of the information. At the other end, the decoding section performs the inverse operation. It expands these compact representations back into a pixel-wise categorization, effectively producing a semantic segmentation of the

original image. Connecting these two sections is the bridge component, which facilitates the flow of information between the encoding and decoding paths.

A key feature of this architecture is that all three components are constructed using residual units. Each unit comprises two 3x3 convolution blocks and an identity mapping. The convolution blocks are structured with a specific sequence: a batch normalization (BN) layer, followed by a ReLU activation layer, and finally a convolutional layer. The identity mapping creates a direct connection between the input and output of each unit, allowing for unimpeded information flow. This carefully designed architecture enables the network to effectively capture and process the intricate details necessary for skin lesion segmentation of dermoscopic images.

The encoding path of our Deep Residual Unet comprises three residual units, each designed to progressively compress the input information. Instead of traditional pooling operations for downsampling, we apply a stride of 2 with the first convolution block within each unit. This approach effectively halves the feature map size at each step, creating a more compact representation.

Mirroring this structure, the decoding path consists of three residual units. However, before each unit in the decoding path, we implement an up-sampling operation on the feature maps before residual blocks from the lower level. The matching feature maps from the encoding path are then concatenated with these up-sampled features, facilitating the reconstruction of spatial information. A sigmoid activation layer and a 1x1 convolution are the last steps in the decoding pipeline. By projecting the multi-channel feature maps into the intended segmentation output, this combination successfully categorizes each pixel.

Our Deep Residual Unet architecture comprises a total of 15 convolutional layers, a notable reduction from the 23 layers found in the standard U-Net. For a comprehensive understanding of the network's structure, we have compiled 3.2, It describes the specifications and output sizes at every stage of the procedure. This table provides a clear overview of how information flows and is transformed throughout the network.

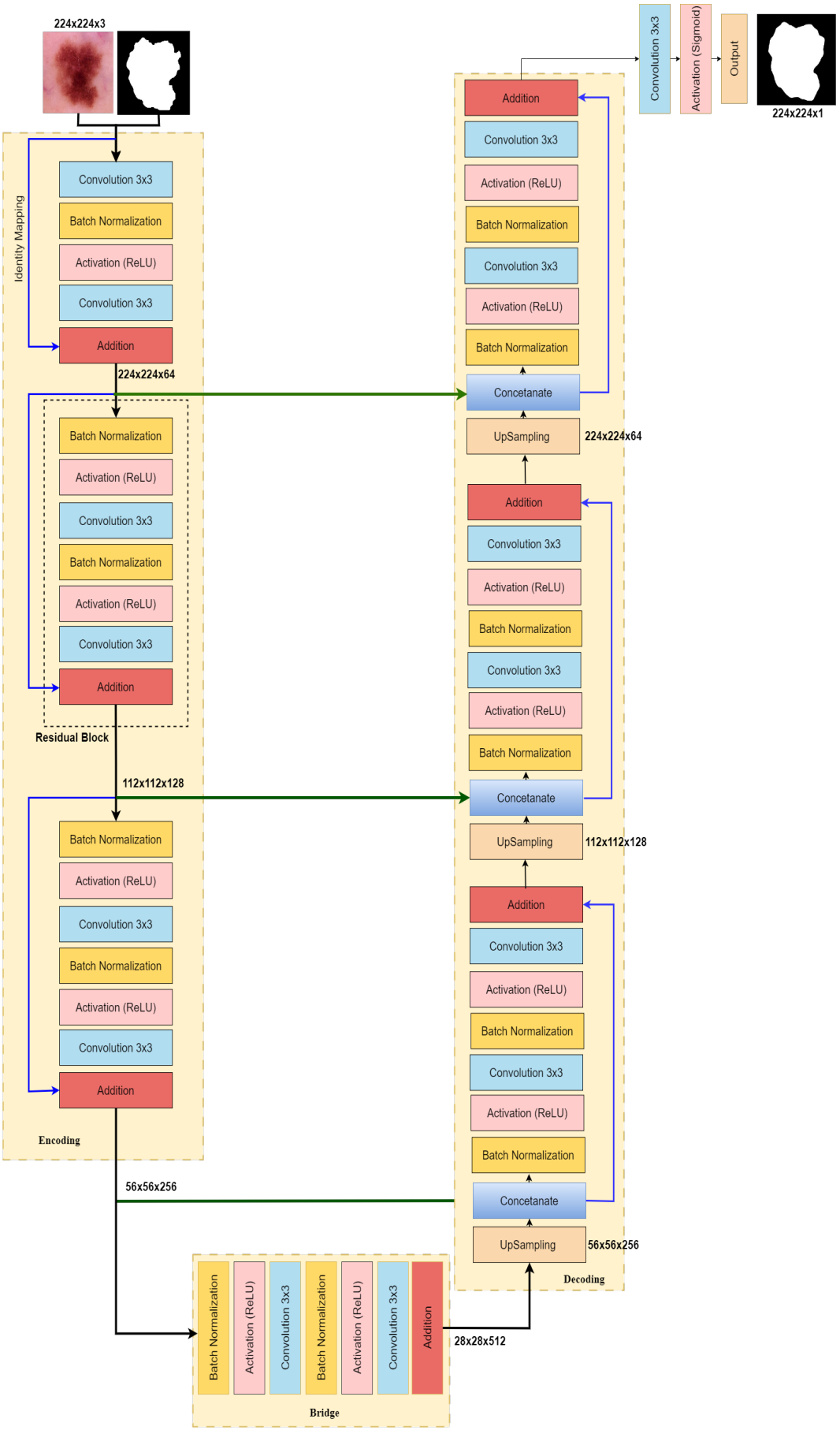


FIGURE 4.2: Architeture of Deep Residual Unet for Skin Lesion Segmentation.

4.3 Classification Architecture for Skin Cancer Diagnosis

To optimize our classification model of skin lesions for the ISIC2019 dataset, we employed transfer learning using the DenseNet169 architecture. This approach allowed us to leverage pre-trained neural network parameters, resulting in excellent accuracy for our specific task.

DenseNet-169 is a convolutional neural network architecture proposed by Huang et al. [37]. DenseNet169 is part of the DenseNet family, featuring 169 layers, which was created to deal with the problem of the vanishing gradient and feature reuse in very deep networks.

It's widely adopted for deep learning classification tasks due to its efficient design. Despite its depth, DenseNet169 boasts fewer trainable parameters compared to its shallower counterparts in the DenseNet family. This architecture mitigates the vanishing gradient problem, implements a robust feature propagation strategy, minimizes the number of trainable parameters, and encourages feature reuse.

4.3.1 Architectural Overview

The Modified structure of DenseNet169, as employed in our study, is illustrated in Figure 4.3.

DenseNet-169 consists of An initial convolution layer, Pooling Layer, 4 dense blocks, 3 transition layers between the dense blocks and final classification layers.

The key innovation of DenseNet is its dense connectivity pattern. All layers in a dense block are connected in a feed-forward fashion. In this structure, each layer doesn't just connect to the next one, it reaches out to every layer that comes after it. Every layer in the block acts like both a receiver and a broadcaster. It takes in all the feature maps from the layers that came before it, processing this wealth of

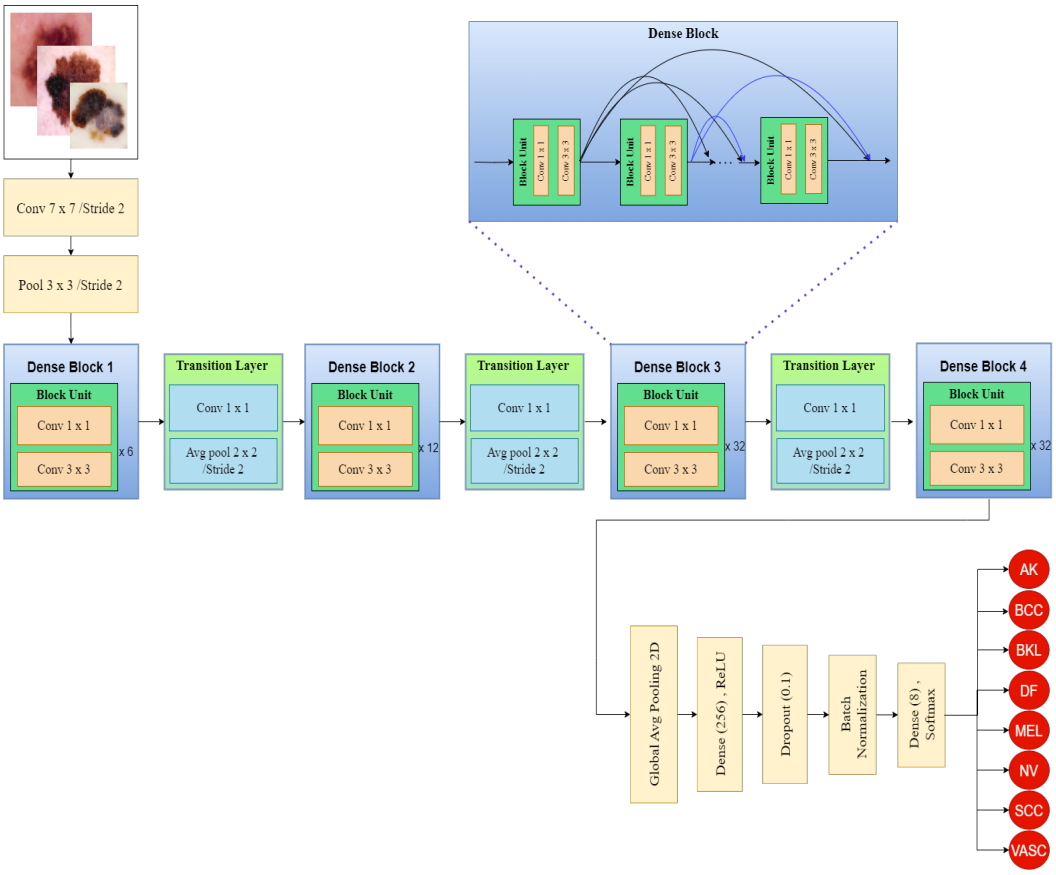


FIGURE 4.3: Modified Densenet169 Architecture for skin lesion classification.

information. Then, it generates its feature maps, which it shares with all the layers that follow. This dense connectivity pattern allows for maximum information flow throughout the network. It's as if each layer is having a conversation with every other layer, sharing what it has learned and building upon the collective knowledge of the entire block. This approach differs from traditional neural networks where information typically only flows to the immediately adjacent layer. In a dense block, the depth of connections creates a more robust and efficient learning environment.

The architecture of a dense block is built around a series of repeating elements called block units. Each of these units is composed of two key components: a 1x1 convolution followed by a 3x3 convolution. The output from each block unit moves forward to all subsequent units within the same dense block. The first dense block is relatively small, with 6 block units. The second block doubles this, containing 12 block units. The third and fourth blocks are the largest and most complex, each housing 32 block units.

In the DenseNet architecture, transition layers play a crucial role as bridges between neighboring dense blocks. These layers serve two main purposes: connecting blocks and streamlining the model. Think of transition layers as efficient intermediaries. They take the complex, information-rich output from one dense block and prepare it for the next. This preparation involves two key steps. First, the layer applies a 1x1 convolution. This step can be thought of as a feature distillation process, where the most important information is extracted and refined. Following this, a 2x2 average pooling operation is performed. This step effectively reduces the spatial dimensions of the feature maps, condensing the information into a more compact form.

These characteristics make DenseNet169 a reliable and efficient choice for our classification task. We utilized the Keras implementation of DenseNet169 for our study.

By adopting this sophisticated architecture and leveraging transfer learning, we aimed to create a highly effective model for classifying skin lesions in the ISIC2019 dataset, balancing the benefits of deep learning with computational efficiency.

4.3.2 Evaluation Metrics

4.3.2.1 For Classification

- Accuracy [68].
- Precision [68].
- Recall [68].
- F1-Score [68].

4.3.2.2 For Segmentation

- Intersection over Union [69].
- Dice similarity coefficient [69].

4.3.2.3 For Hair Removal

- Peak Signal-to-Noise Ratio [70].
- Structural Similarity Index [70].
- Universal Quality Image Index [70].
- Mean Square Error [70].

Chapter 5

Implementation, Experimental Results and Discussion

Building upon the methodological framework detailed in Chapters 3 and 4, this chapter delves into the practical application and empirical evaluation of our proposed approach for skin cancer classification. Here, we present a comprehensive account of our experimental process and the subsequent findings, offering insights into the efficacy and performance of our novel methodology.

5.1 Tools and Technology

Our experimental setup leveraged a combination of powerful tools and technologies, each chosen for its specific strengths in facilitating machine learning research and development:

5.1.1 Kaggle Jupyter Notebook

An online integrated development environment (IDE) called Kaggle Jupyter Notebook makes it easier to write and run Python code. This platform has gained significant popularity among data scientists and machine learning practitioners

due to its user-friendly interface and comprehensive suite of features tailored for data analysis and experimentation. The environment incorporates several key components that enhance the workflow of data scientists:

5.1.1.1 Code Editor

A robust interface for writing and editing Python scripts, with features such as syntax highlighting and auto-completion to improve coding efficiency.

5.1.1.2 Terminal

An integrated command-line interface that allows users to execute system commands and manage their development environment.

5.1.1.3 Debugger

A tool for identifying and resolving code issues, enabling users to step through their code execution and inspect variables.

These features, combined with Kaggle's cloud-based infrastructure, provide a powerful platform for conducting machine learning experiments. The environment allows users to seamlessly integrate data manipulation, model training, and result visualization within a single interface.

Moreover, Kaggle Jupyter Notebook's collaborative features and access to extensive datasets further enhance its utility for data science projects. This combination of accessibility, functionality, and resources has contributed to its widespread adoption in the field of machine learning and data analysis.

5.1.2 Kaggle GPU

Kaggle GPU is a specialized computational service that provides users with access to Graphics Processing Units optimized for machine learning applications. This service offers several key advantages: Accelerated Computation: GPUs are characterized by their superior processing speed compared to traditional Central

Processing Units, particularly in the context of parallel computations. This attribute makes them exceptionally well-suited for the computationally intensive tasks associated with deep learning model training.

Deep Learning Optimization: The architecture of GPUs is inherently conducive to the matrix operations that form the core of many deep learning algorithms. This alignment results in significant performance enhancements, especially when training complex neural network models.

Integrated Environment: The GPU service is seamlessly integrated within the Kaggle ecosystem, allowing users to leverage these computational resources in conjunction with Kaggle's other tools and datasets. This integration facilitates a streamlined workflow from data preparation to model deployment. The Kaggle GPU service thus serves as a critical infrastructure component for the machine learning community, enabling researchers and practitioners to conduct sophisticated experiments and analyses with enhanced computational efficiency. This capability accelerates the iterative process of model development and optimization.

5.1.3 Python Programming Language

Python is a highly versatile, general-purpose programming language that has gained significant prominence in the domain of machine learning. Several important elements have contributed to its global adoption:

5.1.3.1 Accessibility

Python's syntax is characterized by its readability and simplicity, making it accessible to both novice programmers and experienced developers. This ease of use facilitates rapid prototyping and implementation of machine learning algorithms.

5.1.3.2 Extensive Library Ecosystem

Python boasts a comprehensive suite of libraries and frameworks specifically tailored for scientific computing and machine learning applications. This robust

ecosystem significantly enhances productivity and enables efficient implementation of complex algorithms.

5.1.3.3 Machine Learning Frameworks

Python serves as the foundational language for several prominent machine learning frameworks:

- TensorFlow: An open-source framework renowned for its flexibility and scalability in the development and training of deep learning models. TensorFlow's architecture allows for efficient deployment across various platforms, from mobile devices to large-scale distributed systems.
- Keras: A high-level neural networks API that operates as an interface for TensorFlow. Keras simplifies the process of model construction and training by providing intuitive abstractions and pre-built components, thereby accelerating the development cycle of deep learning models.

The synergy between Python's inherent characteristics and its powerful machine-learning libraries has solidified its position as a preferred language for data scientists and machine-learning practitioners. This ecosystem enables researchers and developers to efficiently implement, test, and deploy sophisticated machine learning algorithms across a wide range of applications.

5.2 Dataset

The ISIC is a global database of dermatoscopic pictures. The ISIC2019 challenge [62] provides a comprehensive and diverse dataset for skin lesion classification research. This dataset consists of 25,331 dermatoscopic images and incorporates data from the previous year's (2018 and 2017) challenges.

The dataset is categorized into eight distinct classes of skin lesions. A thorough analysis of the number of dermoscopic images in each class in the ISIC2019 dataset

TABLE 5.1: Description of ISIC2019 Dataset.

Types of skin lesions	**Class abbreviations**	**Number of images**
Actinic Keratoses	AK	867
Basal cell Carcinoma	BCC	3323
Benign Keratosis	BKL	2624
Dermatofibroma	DF	239
Melanoma	MEL	4522
Melanocytic nevus	NV	12875
Squamous cell carcinoma	SCC	628
Vascular skin lesions	VASC	253

can be found in TABLE 5.1. The Distribution of the ISIC2019 Training images is shown graphically in Figure. 5.1.

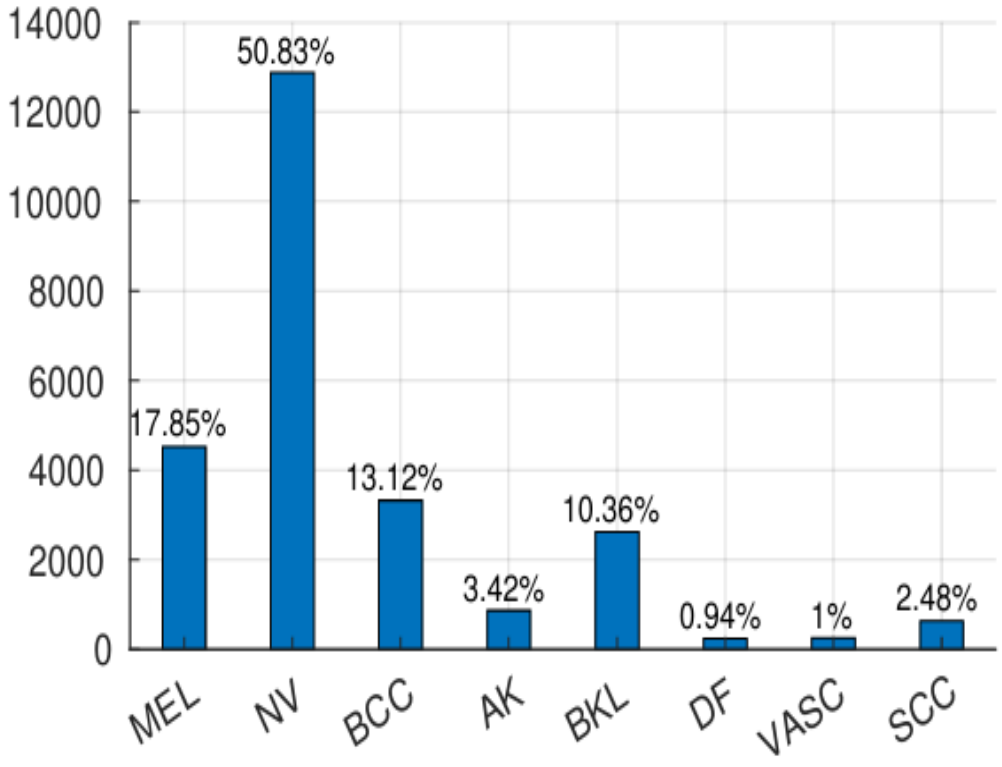


FIGURE 5.1: Frequency of Skin Lesion Categories in ISIC2019 dataset.

5.2.1 Challenges of Dataset

Challenges included in this dataset are as follows:

5.2.1.1 Imbalance Class Distribution

The dataset presents a significant hurdle in the form of class imbalance, which poses a formidable challenge for model training. Among the eight distinct categories of skin lesions, there exists a stark disparity in sample distribution. The majority class, Melanocytic nevus , dominates the dataset with a substantial 12,875 images. In stark contrast, the minority class, Dermatofibroma , is represented by a mere 239 images. This extreme imbalance in class representation complicates the training process, as models tend to be biased towards the overrepresented class, potentially leading to poor performance in underrepresented classes.

5.2.1.2 Hair Presence on Lesion

The presence of hair on skin lesions introduces a significant confounding factor. These fine, often dark structures can obscure important features of the lesion, potentially leading to misclassification or missed diagnoses. The variability in hair density, color, and distribution across images adds another layer of complexity to the classification task.

5.2.1.3 Inter-Class Similarity

The existence of visually similar images across different classes poses a substantial challenge for accurate differentiation.

This dataset's comprehensive nature, including its incorporation of data from previous years and various preprocessing methods, provides a robust foundation for developing and evaluating skin lesion classification models. The diverse range of

skin lesion types represented in the dataset allows for the development of models capable of distinguishing between multiple clinically relevant categories.

By utilizing this well-curated and extensive dataset, our research aims to develop a classification model that can effectively differentiate between these eight types of skin lesions, potentially contributing to more accurate and efficient dermatological diagnoses.

5.3 Implementation of Methodology

5.3.1 Skin Hair Mask Image Generation

The Residual U-Net model was implemented in TensorFlow and Keras libraries. The implementation of the Residual Unet model for skin hair detection and segmentation utilizes a sophisticated 7-level architecture. This network is structured into three primary components: encoding phase, bridge, and decoding phase.

The encoding section serves to compress the input image into compact, high-dimensional representations. This process distills the essential features of the image, creating a condensed form of the information. The Encoder consists of three residual blocks. Each residual block follows the sequence of Batch Normalization, ReLU, Convolution layer, Batch Normalization, ReLU, Convolution layer, and Addition. At the other end, the decoding section performs the inverse operation. It expands these compact representations back into a pixel-wise categorization, effectively producing a semantic segmentation of the original image. Up-sampling was performed before each residual block in the decoding section. Connecting these two sections is the bridge component, which facilitates the flow of information between the encoding and decoding paths.

For Deep Residual Unet training [52] dataset was used which contains 500 dermoscopic images, ground truth images and overlayed images. The pair of dermoscopic images and ground truth images were used for training.

5.3.1.1 Quantitative Analysis of Deep Residual U-Net Model Performance

Our implementation of the Res-Unet model for hair segmentation in dermoscopic images yielded promising results, demonstrating its effectiveness in accurately identifying and delineating hair structures. The performance of the model was evaluated using two key metrics **Intersection over Union** and **Dice Similarity Coefficient**.

The IoU score of 66% was achieved on the [52] dataset. This score indicates that there is a 66% overlap between the predicted hair segmentation and the ground truth. The DSC score of 77% was achieved and the Dice loss of 0.23 on the [52] dataset.

These results indicate that our Res-Unet model performs robustly in segmenting hair from dermoscopic images. The IoU of 0.66 and DSC of 0.77 suggest that the model correctly identifies and localizes hair structures in the majority of cases.

On the [52] dataset, we also trained three popular segmentation models, our proffered Deep Residual Unet outperforms the others. Details regarding the outcomes are provided in Table 5.3.

TABLE 5.2: Deep Residual Unet Model Parameters.

Parameters	**Value**
Batch Size	32
Initial Learning rate	0.0003
Epochs	100
Input Size	224x224x3
Augmentation	True
Optimizer	Adam
Loss	Dice loss
Metric	IoU, DSC

TABLE 5.3: Quantitative Comparison with other Segmentation models.

Segmentation Model	IoU (Jac)
DeepLabv3	0.48
Efficientnet-b5-Unet	0.62
Unet ++	0.61
Proffered Deep Residual U-Net	**0.66**

5.3.1.2 Qualitative Analysis of Residual U-Net Model Performance

The qualitative results presented in Figure 5.2 corroborate and enhance our und-

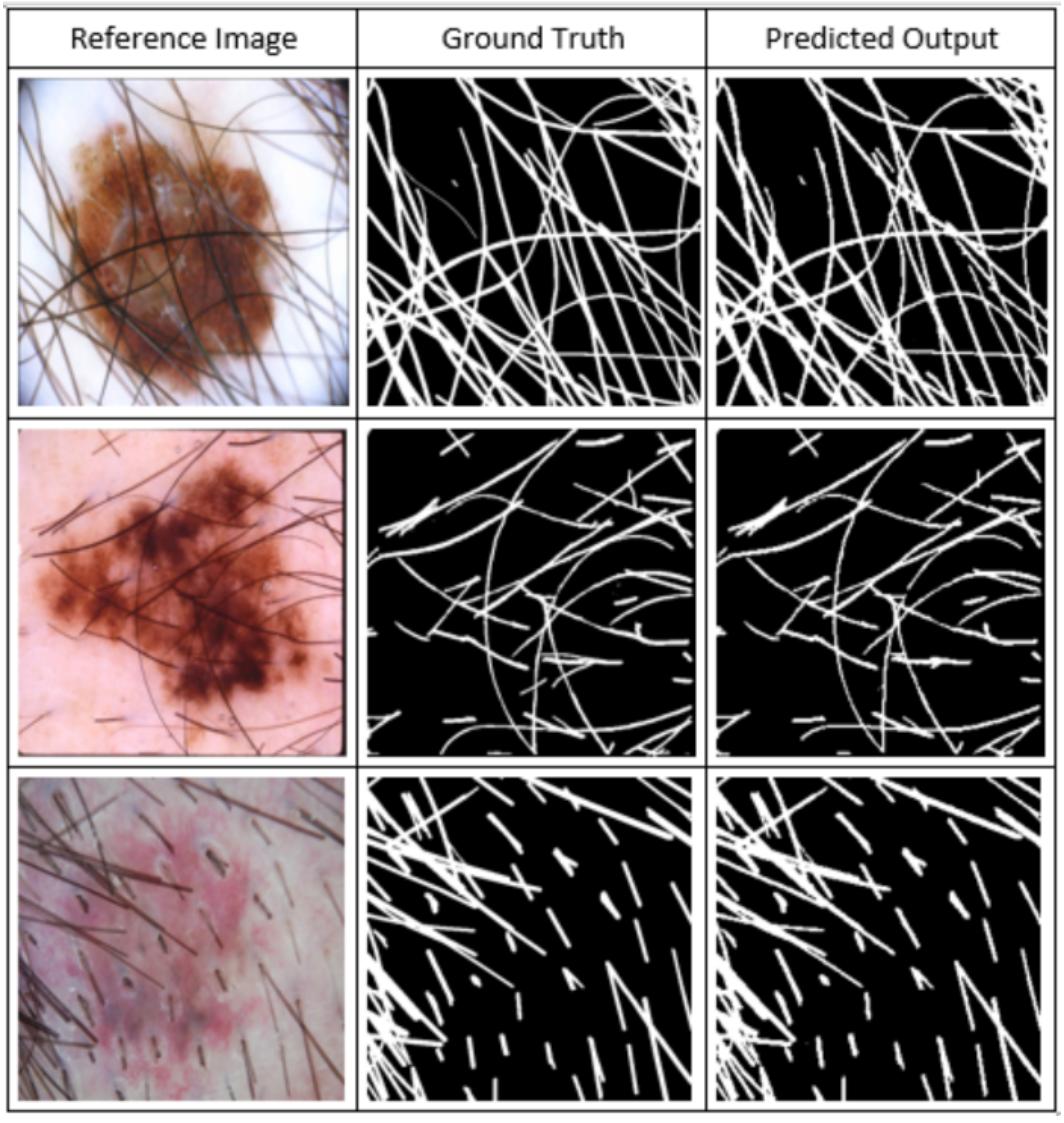


FIGURE 5.2: Qualitative Results of Deep Residual U-Net Model for Mask Images

erstanding of the model's performance beyond the quantitative metrics. The visual similarity between ground truth and predicted images across multiple samples demonstrates the robustness and reliability of Deep Residual U-Net model in the critical task of hair segmentation. This performance level provides a solid foundation for the subsequent stages of our skin cancer classification methodology, particularly in preparing clean, hair-free images for accurate lesion analysis.

5.3.2 Hair Removal

A quantitative and qualitative evaluation of the Hair Removal system's performance was conducted. These results are compared with those generated by five state-of-the-art methods: Lee et al. [47], Xie et al. [48], Abbas et al. [49], Huang et al. [50] and Bibiloni et al. [51]. Hair segmentation was performed on dermoscopy images. Segmented images can be seen in Figure 5.2. After the segmentation image is passed to the TELEA Inpainting algorithm, where hair gaps are filled morphologically.

5.3.2.1 Quantitative Results of Efficient Hair Removal

We analyzed the outcomes both qualitatively and quantitatively. We used Mean Squared Error (MSE), the Structural Similarity Index (SSIM), the Peak Signal-to-Noise Ratio (PSNR), and the Universal Quality Image Index (UQI) as quantitative evaluation metrics. In 5.4, we provided the mean of the results obtained for the images from the dataset [63] for each performance measure. Our method outperforms all compared methods in terms of MSE, PSNR and SSIM. For the

TABLE 5.4: Quantitative Comparison of EHR with Existing Techniques.

Year	Method	MSE	PSNR	SSIM	UQI
1997	Lee et al. [47]	62.76	33.91	0.94	0.9579
2009	Xie et al. [48]	16.75	36.96	0.97	0.998
2011	Abbas et al. [49]	130.7	27.74	0.90	0.995
2013	Huang et al. [50]	44.49	32.59	0.95	0.997
2017	Bibiloni et al. [51]	63.3	30.73	0.95	0.998
2024	Ours	**4.18**	**43.20**	**0.98**	0.9973

assessment task, we turned to a specialized dataset due to the lack of readily available paired images showing hair-covered lesions and their clean counterparts. This dataset was created by researchers who took a novel approach: they started with clean skin lesion images from the ISIC (International Skin Imaging Collaboration) database and artificially added hair structures to these images. This data allows us to accurately assess the performance of hair removal algorithms.

5.3.3 Qualitative Results of Efficient Hair Removal

In Figure 5.3, We may observe that the outcomes of our suggested approach are visually appealing. Furthermore, we provide an illustration contrasting our outcomes with those attained through alternative hair removal techniques presented.

Figure 5.3 illustrates that not all hair removal techniques are effective. The method proposed by Abbas and colleagues shows limitations in hair segmentation. It appears their algorithm struggles with accurately detecting hair structures in the images. This fundamental issue hinders the overall effectiveness of their hair removal process.

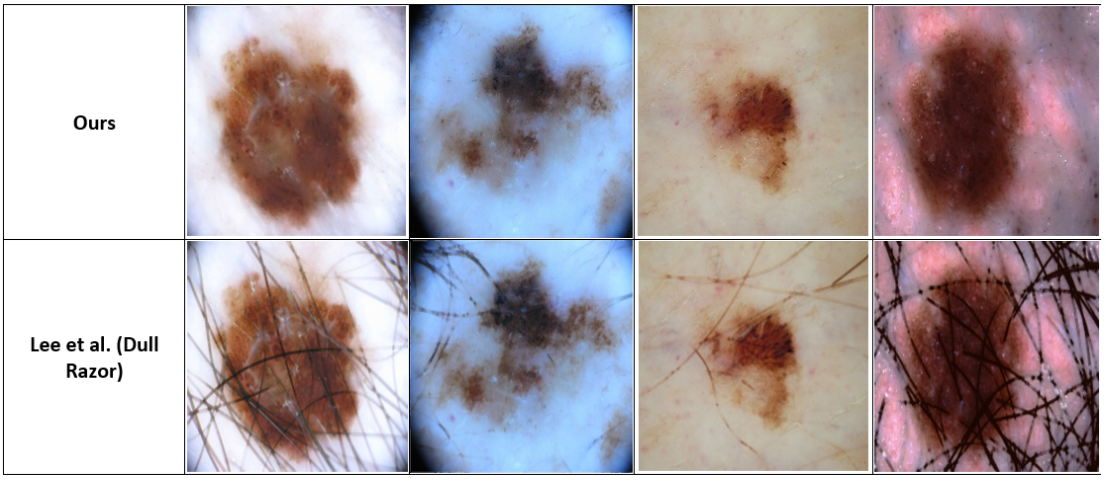


FIGURE 5.4: Performance on original hairs.

On the other hand, the approaches developed by Huang et al. and Xie et al. demonstrate a much stronger ability to detect hair in the images. However, these methods aren't without their flaws. While they successfully identify most of the

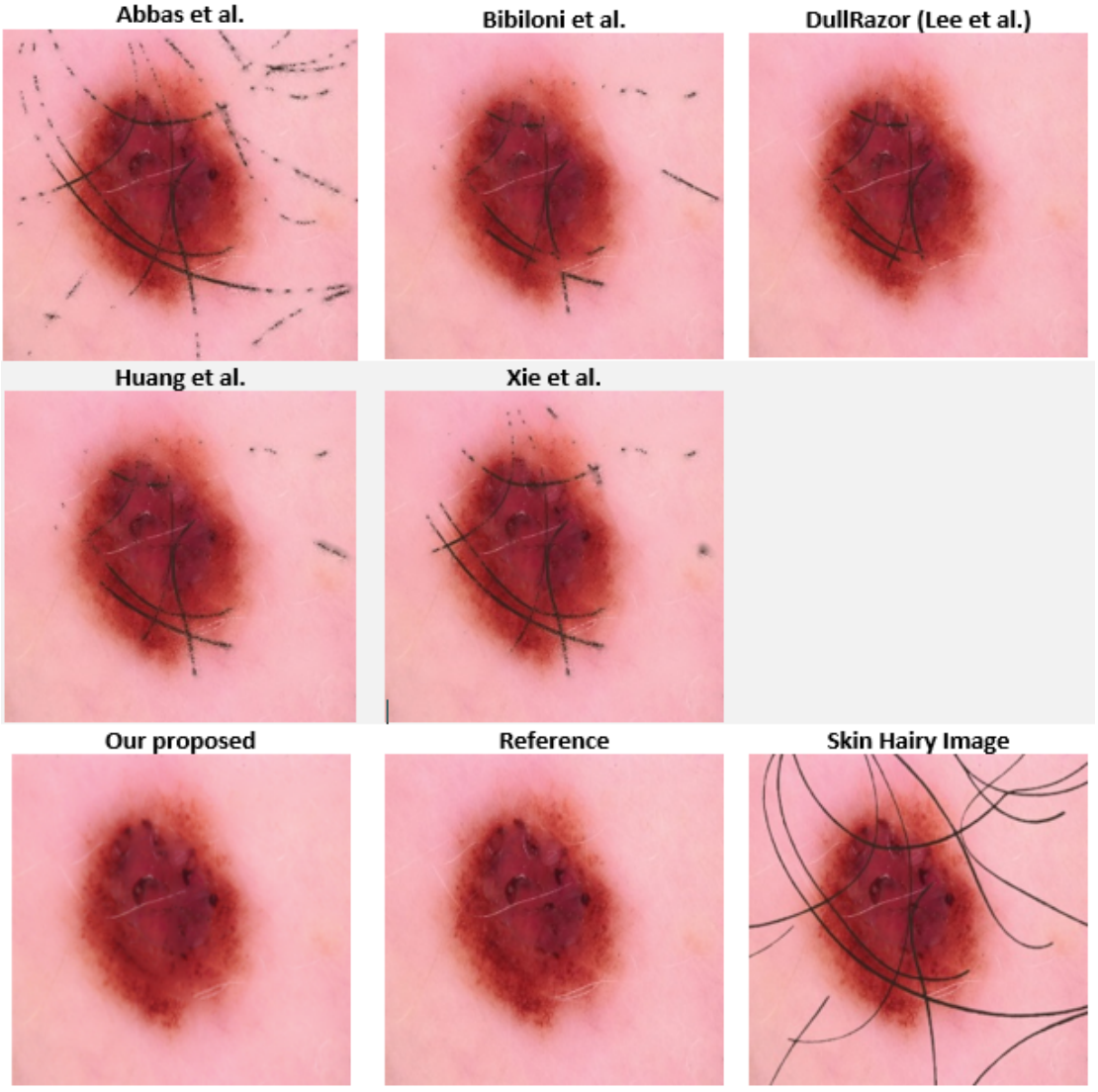


FIGURE 5.3: Qualitative Comparision of EHR with Existing Techniques.

hair, their inpainting process (the step where they fill in the areas where hair has been removed) leaves noticeable traces. Methods developed by Bibiloni et al. and Lee et al., appear to produce results more closer to reference images. Our proposed method is like a counterpart of the reference image.

Our initial evaluation was conducted using a dataset featuring artificially added hair structures. In this controlled environment, the method developed by Lee and colleagues demonstrated superior performance compared to other approaches. However, when we applied Lee et al.'s method to a dataset containing images with natural hair coverage, we observed some limitations. In several cases, the algorithm failed to accurately detect and remove hair from actual dermoscopic

images. Figure 5.4 contains the images obtained by the Lee et al. method and our method.

5.4 Skin Lesion Segmentation

Our implementation of the Deep Residual Unet model for skin lesion segmentation in dermoscopic images yielded promising results, demonstrating its effectiveness in accurately identifying the lesion areas. The performance of the model was evaluated using two key metrics IoU and Dice.

Our research has yielded impressive results in skin lesion segmentation, as evidenced by the data presented in Table 5.5. When compared to other segmentation models and previously published works using the ISIC2018 dataset, our method demonstrates superior performance. Specifically, our technique achieved an Intersection over Union (IoU) score of 0.8981 and a Dice coefficient of 0.946.

These metrics are widely used in the field of image segmentation to evaluate the accuracy of the predicted segmentation masks compared to the ground truth.

To provide a visual representation of our model's performance, we've included a Figure 5.5 in our study. This figure showcases a selection of original images alongside their corresponding ground truth segmentation masks and the masks generated by our model. This visual comparison allows for an intuitive understanding of how closely our model's predictions align with the expert-annotated ground truth.

TABLE 5.5: Comparison of skin lesion segmentaion

Ref.	**IoU**	**Dice**
U-Net	0.694	0.739
U-Net++	0.817	0.897
DeepLabV3	0.82	0.884
wang et al. [64]	0.843	0.912
Tang et al. [65]	0.867	0.923
Ben˘cevi´c et al. [66]	0.874	0.925
Dzieniszewska et al. [67]	0.88	0.937
Proffered Deep Residual U-Net	**0.8981**	**0.946**

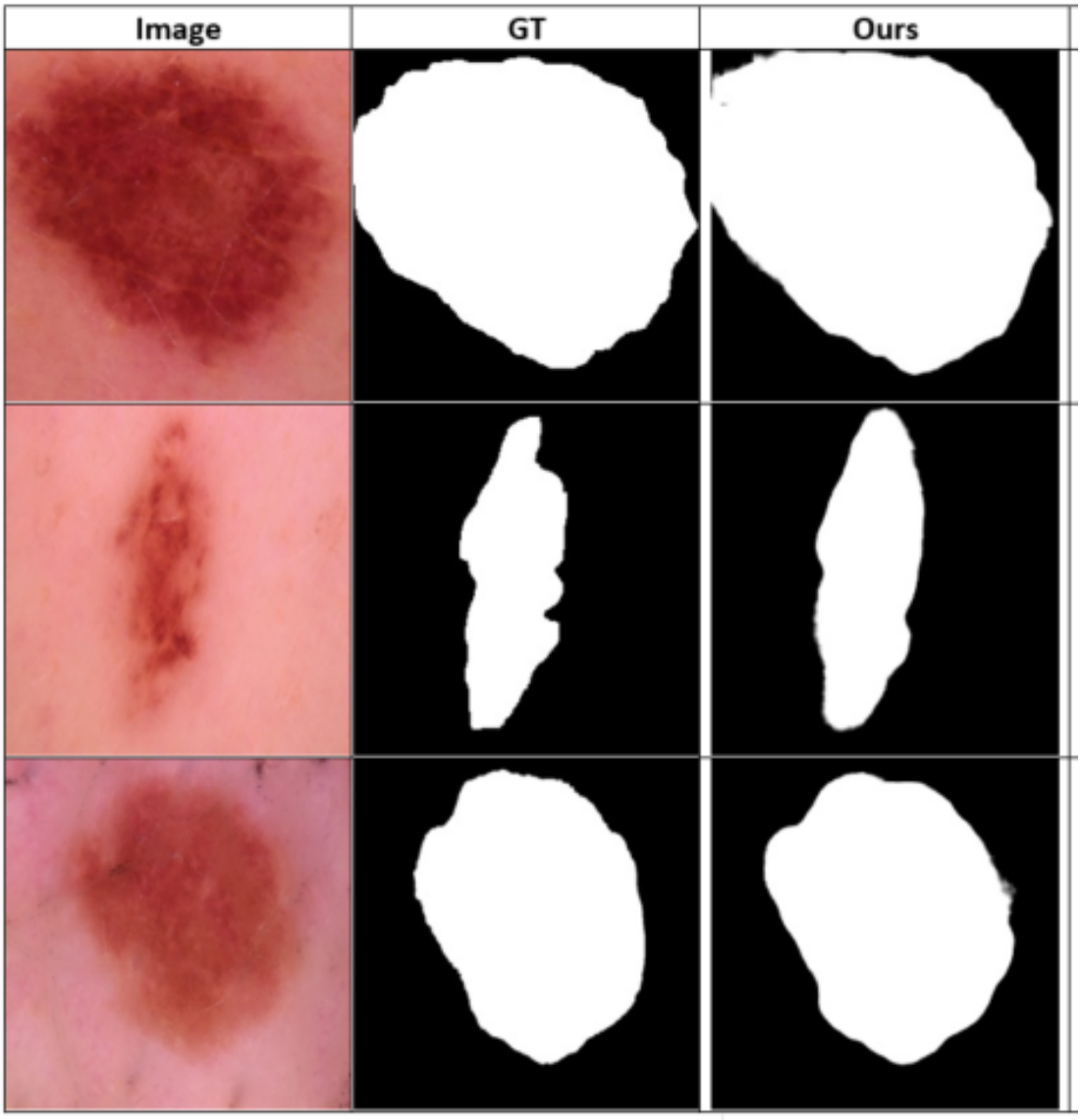


FIGURE 5.5: Comparison of skin lesion image, GT and Generated mask.

5.5 Classification Experiments and Results

To rigorously evaluate the efficacy of DenseNet169 in classifying various types of skin cancer, we conducted a series of five comprehensive experiments. These experiments were designed to provide a thorough assessment of the model's performance and reliability across different scenarios. There are five distinct experiments were performed. The 2019 ISIC dataset was utilized across all experiments. The evaluation of the experiments was assessed using Accuracy, Precision, Recall, F1-score, Confusion Matrix, and training curves.

The series of experiments demonstrates the robust capability of DenseNet169 in classifying different types of skin cancer. The achievement of 97% accuracy in Experiment 5, along with strong performance across various metrics, provides compelling evidence of the model's effectiveness. These results not only validate our approach but also open avenues for further research and potential clinical applications in the field of automated skin cancer diagnosis.

5.5.1 Experiment 1

The experiment was run on Python 3.11.7 on a Kaggle notebook with two T4 GPUs with 32 GB memory, and a CPU with 32 GB memory. The entire experiment was based on tensorflow 2 and keras.

In the preprocessing phase, the input images were transformed from their original dimensions of 1022x767x3 pixels to a standardized size of 224x224x3 pixels. This resizing operation was performed to ensure uniformity across the dataset and compatibility with the chosen neural network architecture. The dataset was divided into training and validation sets, with roughly 80% of the data going toward training and the remaining 20% being kept for validation.

The initial experiment was conducted using the imbalanced dataset, maintaining the original class distribution without applying any balancing techniques. The DenseNet169 architecture was employed in its original configuration, without any structural modifications. This decision was made to establish a baseline performance metric, facilitating future comparisons with potential architectural adaptations or alternative models. The experiment utilized the preprocessed images

Table 5.6: Hyperparameters maintained during training of Experiment 1.

Name	**Value**
Input size	224x224x3
Batch Size	32
Initial Learning Rate	0.0001
Epochs	50
Optimizer	Adam
Loss	Categorical Cross-entopy

of dimensions 224x224x3 as input, where the third dimension corresponds to the RGB color channels.

The hyperparameter configuration employed during the training phase was determined through iterative experimentation and optimization. The final set of hyperparameters utilized in this study is presented in 5.6. These parameters were carefully selected to balance model performance, computational efficiency, and generalization capability.

5.5.1.1 Results of Experiment 1

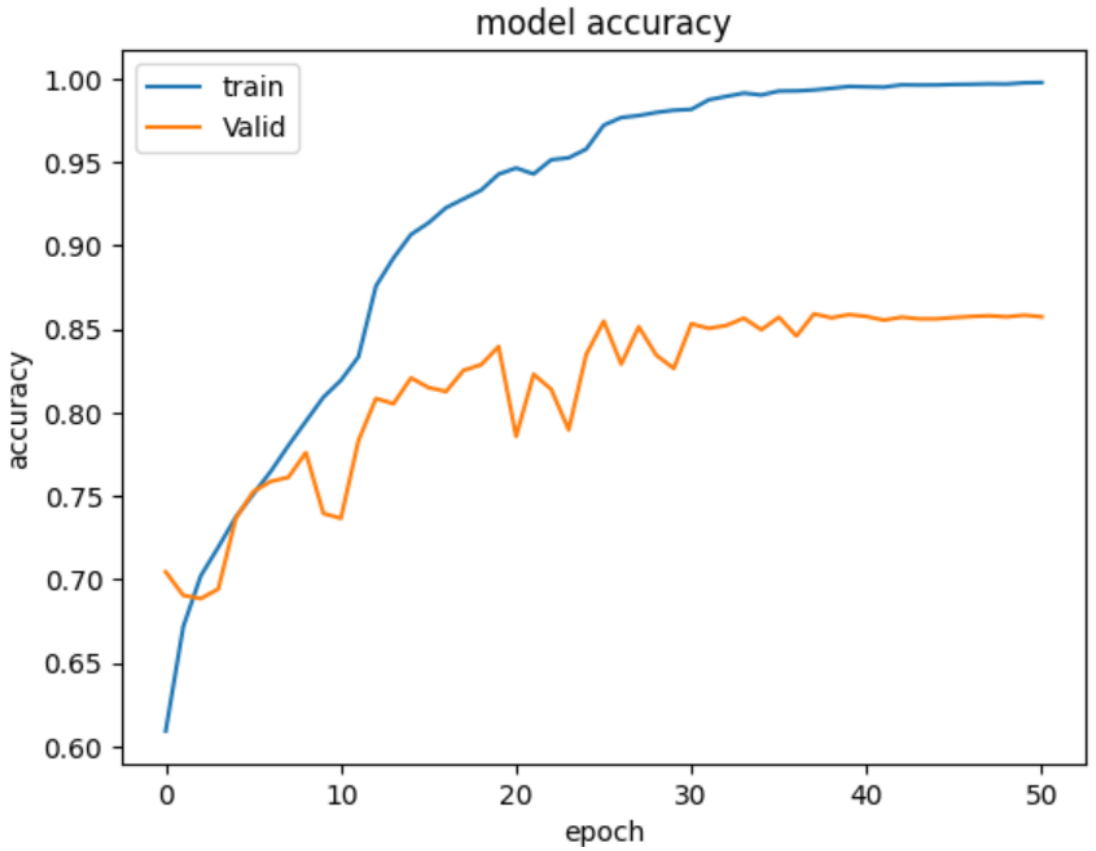
TABLE 5.7: Performance of the model with Original Architecture of Densenet169 and dataset imbalanced.

Class	**Precision**	**Recall**	**F1-Score**
AK	0.79	0.69	0.74
BCC	0.91	0.88	0.89
BKL	0.76	0.77	0.76
DF	0.88	0.75	0.81
MEL	0.83	0.72	0.77
NV	0.89	0.96	0.92
SCC	0.81	0.71	0.76
VASC	0.92	0.90	0.91
Average	0.87	0.87	0.86
Accuracy		0.87	

After 50 epochs of training, the model demonstrated promising performance. The test loss hit its low point, it continued to rise corresponding to a test accuracy of 0.87. This result suggests that the model has learned to effectively distinguish between different types of skin lesions with a high degree of reliability.

Table 5.7 presents a detailed breakdown of the classification metrics for each type of skin lesion, including Precision, Recall, and F1-score. Additionally, the table provides the average values for these metrics across all lesion types, along with the overall Accuracy. This comprehensive evaluation allows for a nuanced understanding of the model's performance across various lesion categories.

(a)



(b)

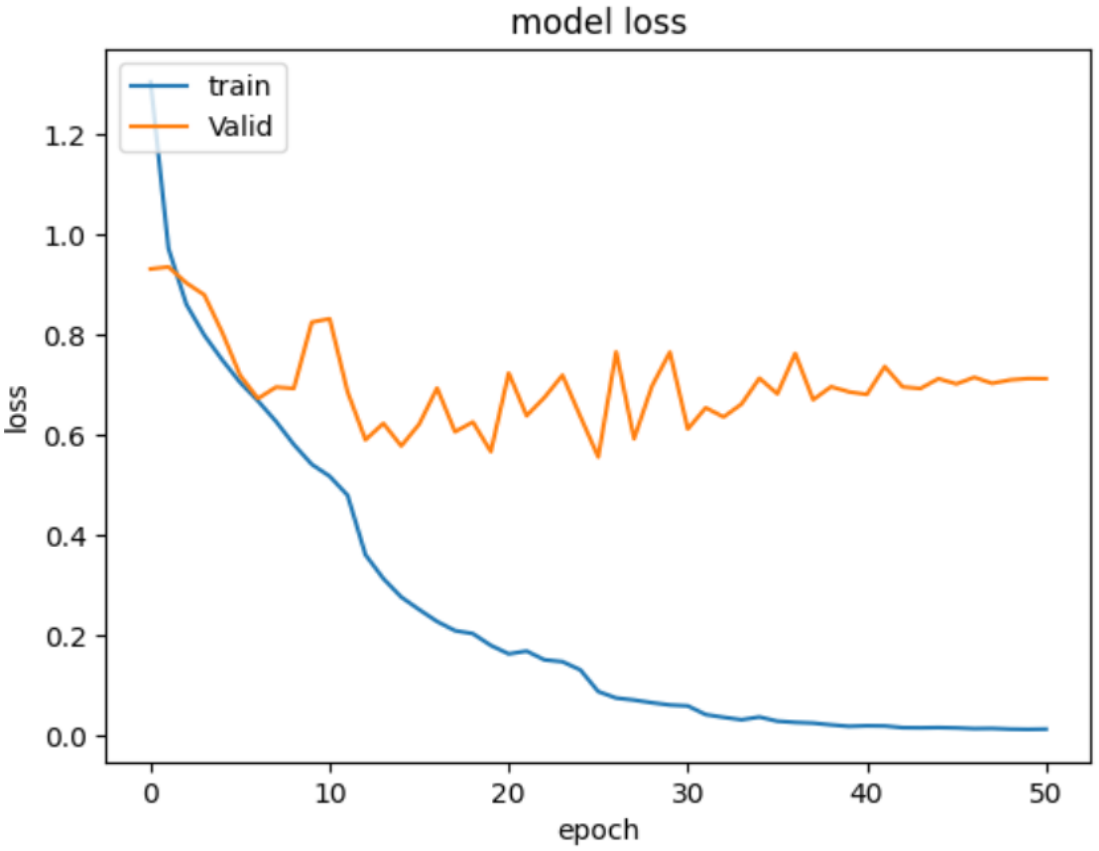


FIGURE 5.6: Experment 1: (a) Accuracy versus epochs, (b) Loss versus epochs.

TABLE 5.8: Hyperparameters maintained during training for Experiment 2, 3, 4 and 5.

Name	**Value**
Input size	224x224x3
Batch Size	32
Initial Learning Rate	0.0001
Epochs	30
Optimizer	Adam
Loss	Categorical Cross-entopy

Figure 5.6 illustrates the model's learning progression throughout the training process. Specifically, 5.6 (a) depicts the relationship between training and test Accuracy for the number of training epochs on the ISIC2019 test dataset.

Correspondingly, 5.6 (b) presents the evolution of training and test loss for the training epochs, also evaluated on the ISIC2019 test dataset. This graph offers valuable information about the model's convergence and optimization process, helping to identify any instances of divergence or instability during training.

These results collectively provide a thorough evaluation of the model's performance. Experiment 1 shows that the Original Architecture of DenseNet169 with an imbalanced ISIC2019 dataset using these hyperparameters achieved an overall Accuracy of 87%.

5.5.2 Experiment 2

The experiment was run on Python 3.11.7 on a Kaggle notebook with two T4 GPUs with 32 GB memory, and CPU with 32 GB memory. The entire experiment was based on TensorFlow 2 and Keras.

In the preprocessing phase, the input images were transformed from their original dimensions of 1022x767x3 pixels to a standardized size of 224x224x3 pixels. This resizing operation was performed to ensure uniformity across the dataset and compatibility with the chosen neural network architecture. The dataset was Augmented with settings shown in 4.1 so the dataset is almost balanced. Prepared

TABLE 5.9: Description of ISIC2019 Dataset after Augmentation.

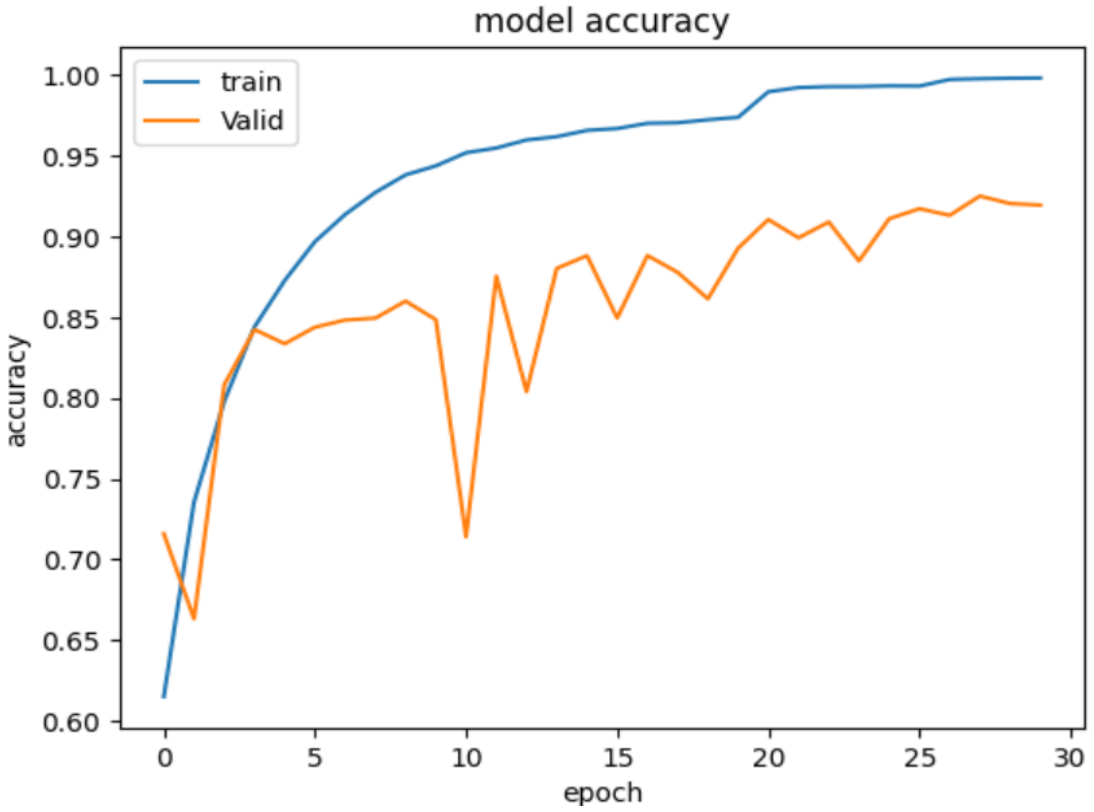
Classes	**Number of images**	**Data Augmentation**
AK	867	11941
BCC	3323	11951
BKL	2624	11972
DF	239	11092
MEL	4522	11972
NV	12875	12873
SCC	628	11726
VASC	253	11891

dataset for Experiments 2 and 3 detailed in 5.9. The dataset was split into training, validation, and test sets, with 70% of the data allocated for training, 15% of data allocated to the validation set, and the remaining 15% reserved for test purposes. To confirm there is no similar image in the training and test set we used the software dupeGuru https://dupeguru.voltaicideas.net/ with threshold 100.

The experiment was conducted using the balanced dataset. The DenseNet169 architecture was employed in its original configuration, without any structural modifications. The experiment utilized the preprocessed images of dimensions 224x224x3 as input, where the third dimension corresponds to the RGB color channels.

The hyperparameter configuration employed during the training phase was determined through iterative experimentation and optimization. The final set of hyperparameters utilized in this study is presented in 5.8. These parameters were carefully selected to balance model performance, computational efficiency, and generalization capability.

(a)



(b)

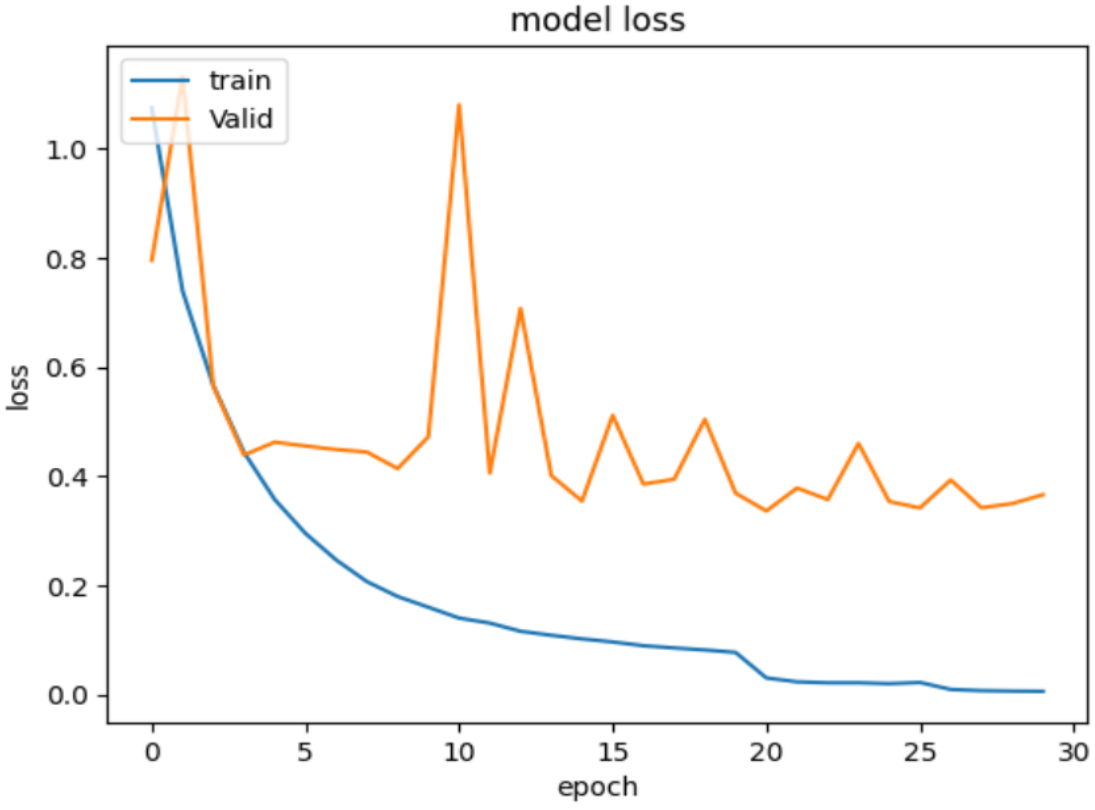


FIGURE 5.7: Experiment 2: (a) Accuracy versus epochs, (b) Loss versus the number of training epochs.

5.5.2.1 Results of Experiment 2

After 30 epochs of training, the model demonstrated promising performance. The test loss reached near to 0.4, corresponding to a test accuracy of 0.92. This result suggests that the model has learned to effectively distinguish between different types of skin lesions with a high degree of reliability.

Table 5.10 presents a detailed breakdown of the classification metrics for each type of skin lesion, including Precision, Recall, and F1-score. Additionally, the table provides the weighted average values for these metrics across all lesion types, along with the overall Accuracy. This comprehensive evaluation allows for a nuanced understanding of the model's performance across various lesion categories.

Figure 5.7 illustrates the model's learning progression throughout the training process. Specifically, 5.7 (a) depicts the relationship between training and test Accuracy for the number of training epochs on the ISIC2019 test dataset.

Correspondingly, 5.7 (b) presents the evolution of training and test loss for the training epochs, also evaluated on the ISIC2019 test dataset. This graph offers valuable information about the model's convergence and optimization process, helping to identify any instances of divergence or instability during training. These results collectively provide a thorough evaluation of the model's performance. Experiment 2 shows that the Original Architecture of DenseNet169 with a balanced

TABLE 5.10: Performance of the model with Original Architecture of Densenet169 and dataset Balanced.

Class	**Precision**	**Recall**	**F1-Score**
AK	0.92	0.95	0.94
BCC	0.89	0.92	0.90
BKL	0.86	0.85	0.85
DF	0.98	1.00	0.99
MEL	0.91	0.76	0.83
NV	0.89	0.94	0.91
SCC	0.95	0.98	0.96
VASC	1.00	1.00	1.00
Average	0.92	0.92	0.92
Accuracy		0.92	

ISIC2019 dataset using these hyperparameters achieved an overall accuracy of 92.28%.

5.5.3 Experiment 3

This experiment was conducted following the protocol established for experiment 2. The hyperparameters employed in this study were selected as per the specifications detailed in 5.8.

Before implementing data augmentation techniques, we applied our proposed Efficient Hair Removal Framework to the images. This preprocessing step was crucial in eliminating artifacts that could potentially interfere with the model's learning process. The application of this framework resulted in cleaned images with dimensions of 224x224x3 pixels, maintaining the standard input size for our neural network architecture.

Following the hair removal process, we proceeded with data augmentation on the cleaned dataset. The augmentation techniques and their corresponding parameters were applied as outlined in 4.1. Augmented dataset detail described in Table 5.9.

This augmentation step was designed to enhance the diversity of our training data, potentially improving the model's ability to generalize across various presentations of skin lesions. The experiment was conducted using the cleaned or hair removed balanced dataset. The DenseNet169 architecture was employed in its original configuration, without any structural modifications. The experiment utilized the preprocessed images of dimensions 224x224x3 as input.

5.5.3.1 Results of Experiment 3

After 30 epochs of training, the model demonstrated promising performance. The test loss reached near to 0.25, corresponding to a test accuracy of 0.9481. This result suggests that the model has learned to effectively distinguish between different types of skin lesions with a high degree of reliability. Table 5.11 presents a detailed breakdown of the classification metrics for each type of skin lesion, including Precision, Recall, and F1-score. Additionally, the table provides the weighted

TABLE 5.11: Performance of the model with Original Architecture of Densenet169 and Hair removed Balanced dataset.

Class	**Precision**	**Recall**	**F1-Score**
AK	0.95	0.98	0.97
BCC	0.95	0.93	0.94
BKL	0.93	0.90	0.91
DF	0.99	1.00	0.99
MEL	0.91	0.84	0.87
NV	0.91	0.95	0.93
SCC	0.96	0.99	0.98
VASC	1.00	1.00	1.00
Average	0.95	0.95	0.95
Accuracy		0.95	

average values for these metrics across all lesion types, along with the overall Accuracy. This comprehensive evaluation allows for a nuanced understanding of the model's performance across various lesion categories. Figure

Figure 5.8 illustrates the model's learning progression throughout the training process. Specifically, 5.8 (a) depicts the relationship between training and test Accuracy for the number of training epochs on the ISIC2019 test dataset. Correspondingly, 5.8 (b) presents the evolution of training and test loss versus the training epochs, also evaluated on the ISIC2019 test dataset.

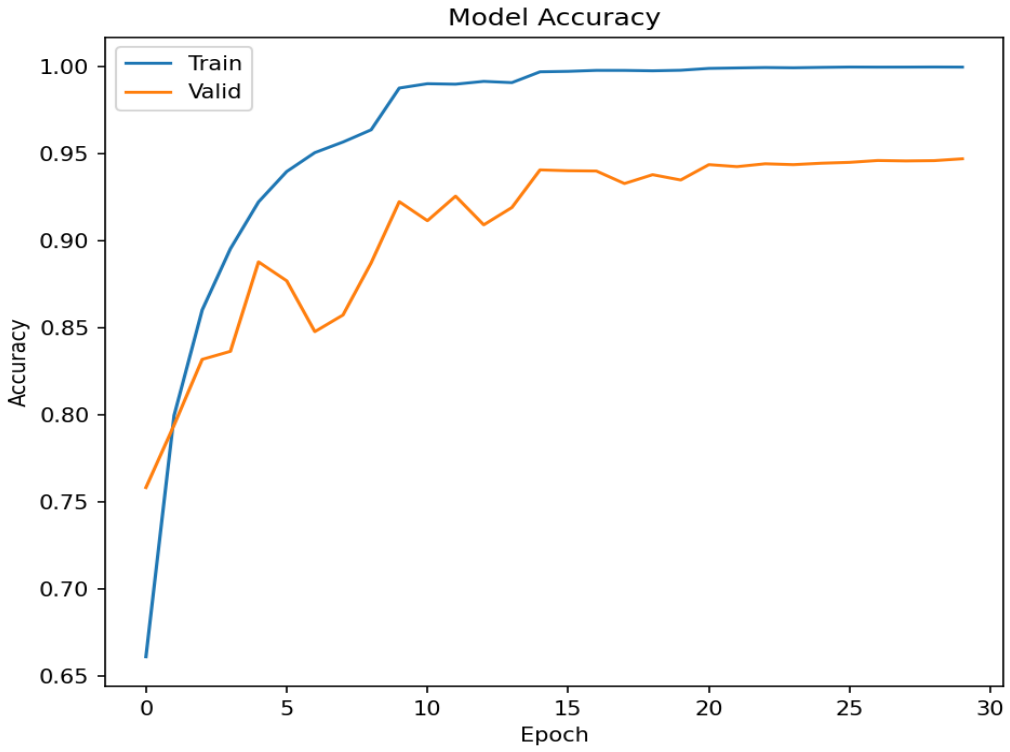
This graph offers valuable information about the model's convergence and optimization process, helping to identify any instances of divergence or instability during training.

These results collectively provide a comprehensive assessment of the model's performance. Experiment 3 shows that the Original Architecture of DenseNet169 with a hair removed balanced ISIC2019 dataset using these hyperparameters achieved an overall accuracy of 94.81%.

5.5.4 Experiment 4

This experiment was conducted following the protocol established for experiments 2, 3 and 4. The hyperparameters employed in this study were selected as per the

(a)



(b)

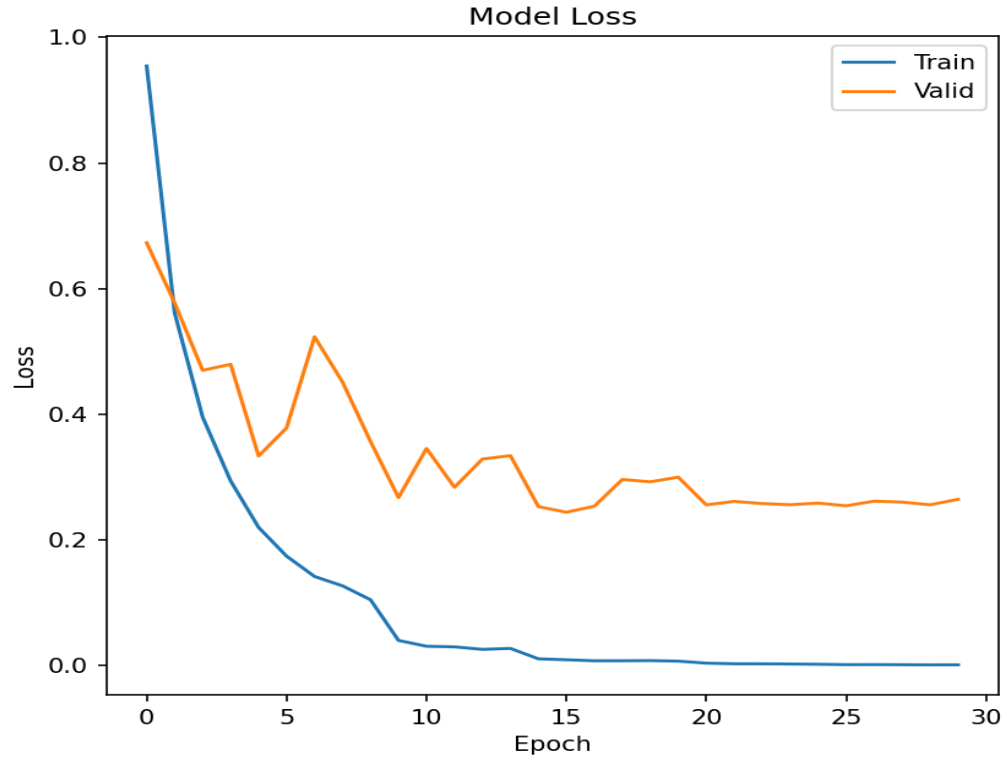


FIGURE 5.8: Experiment 3: (a) Accuracy versus epochs, (b) Loss versus the number of epochs.

specifications detailed in Table 5.8.

This experiment was conducted using a clean, balanced, and lesion-segmented dataset. For this experiment Skin lesion segmentation was performed using the Deep Residual Unet model and the whole dataset was prepared. The DenseNet169 architecture was employed in its original configuration, without any structural modifications. The experiment utilized the preprocessed images of dimensions 224x224x3 as input.

5.5.4.1 Results of Experiment 4

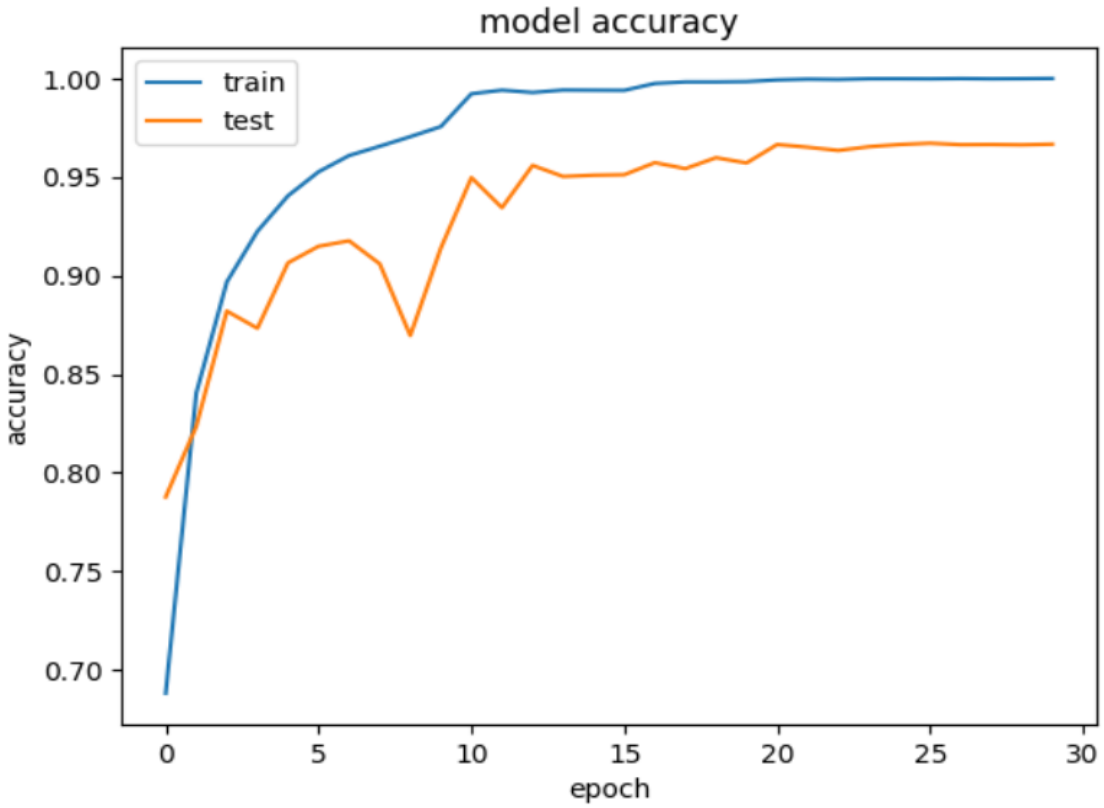
After 30 epochs of training, the model demonstrated promising performance. The test loss reached close to 0.2, corresponding to a test accuracy of **0.9631**. This result suggests that the model has learned to effectively distinguish between different types of skin lesions with a high degree of reliability. Table 5.12 presents a detailed breakdown of the classification metrics for each type of skin lesion, including Precision, Recall, and F1-score. Additionally, the table provides the average values for these metrics across all lesion types, along with the overall Accuracy.

The Figure 5.9 illustrates the model's learning progression throughout the training process. Specifically, 5.9 (a) depicts the relationship between training and test

TABLE 5.12: Performance of the model with Original Architecture of Densenet169, Hair removed and skin lesion segmented Balanced dataset.

Class	**Precision**	**Recall**	**F1-Score**
AK	0.99	1.00	0.99
BCC	0.96	0.96	0.96
BKL	0.93	0.96	0.95
DF	0.99	1.00	1.00
MEL	0.93	0.89	0.91
NV	0.92	0.93	0.92
SCC	0.99	1.00	0.99
VASC	1.00	1.00	1.00
Average	0.96	0.96	0.96
Accuracy		0.96	

(a)



(b)

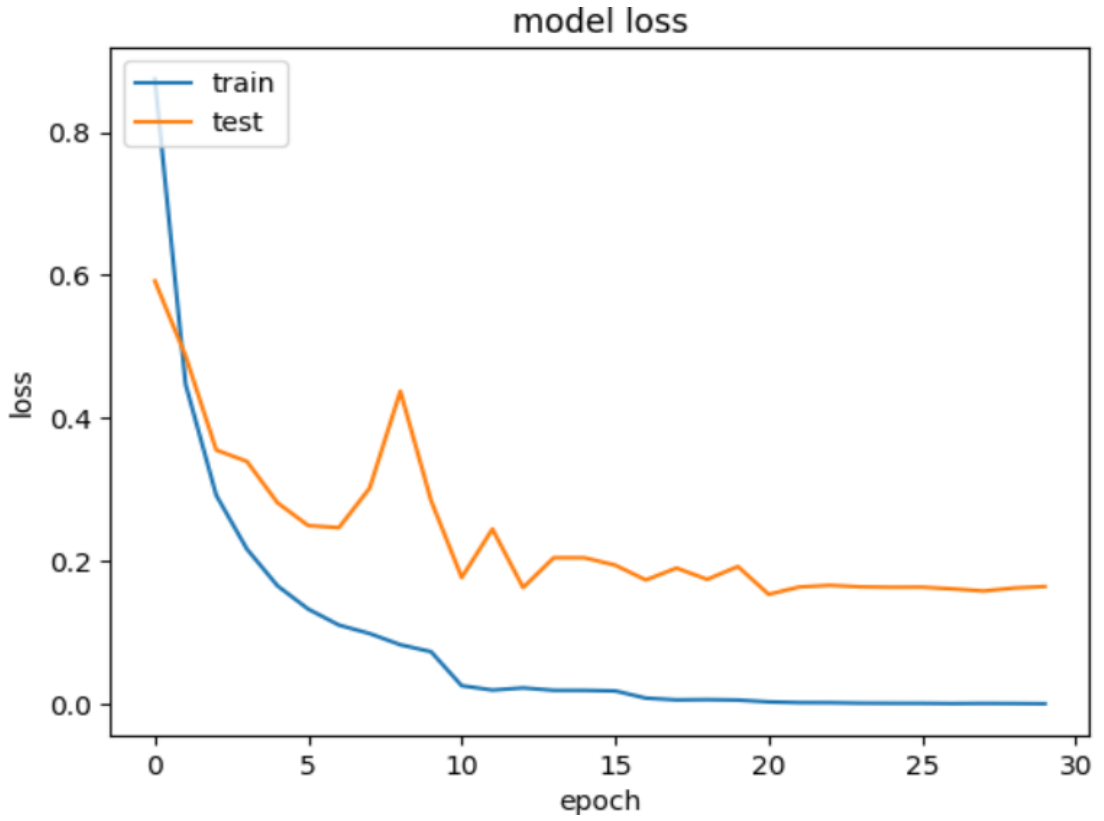


FIGURE 5.9: Experiment 4: (a) Accuracy versus epochs, (b) Loss versus the number of epochs.

accuracy for the number of training epochs on the ISIC2019 test dataset. Correspondingly, 5.9 (b) presents the evolution of training and test loss versus the training epochs, also evaluated on the ISIC2019 test dataset. These results collectively provide a comprehensive assessment of the model's performance. Experiment 4 shows that the original Architecture of DenseNet169 with a hair-removed, skin lesion segmented balanced ISIC2019 dataset using these hyperparameters achieved an overall accuracy of 96.31%.

5.5.5 Experiment 5

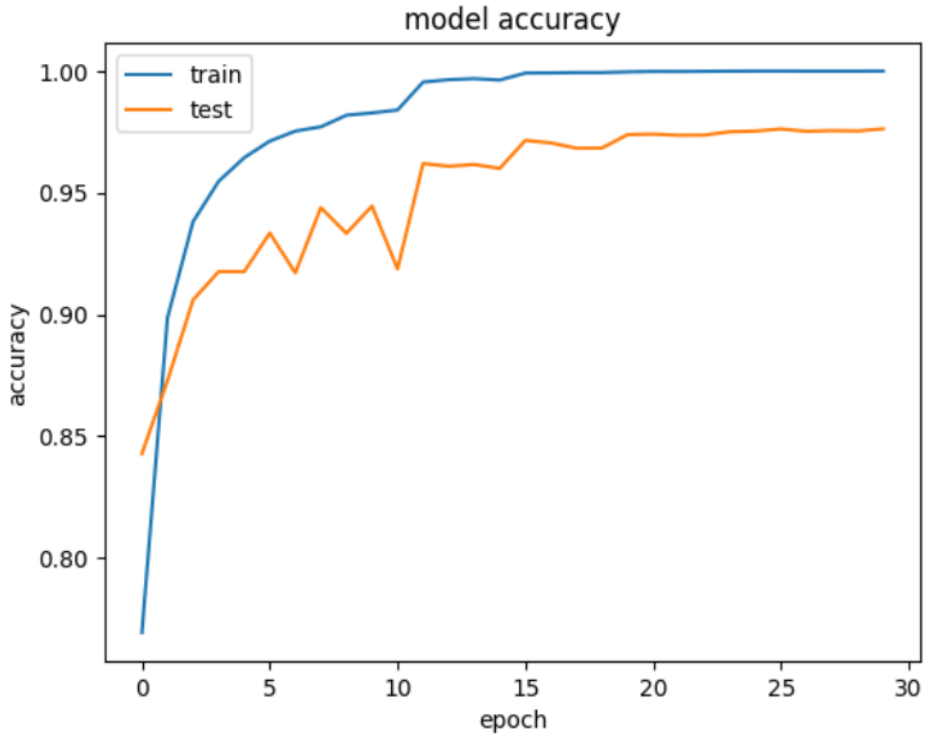
This experiment was conducted following the protocol established for experiments 2, 3 and 4. The hyperparameters employed in this study were selected as per the specifications detailed in 5.8.

The experiment was conducted utilizing a cleaned, balanced skin lesion-segmented dataset. This dataset was identical to the one employed in experiment 4. The primary distinction in this experiment lies in the architectural modifications applied to the DenseNet169 model. The DenseNet169 architecture was adapted according to the specifications detailed in Figure 4.3. These modifications were implemented to enhance the model's capacity to extract and process features specific to dermoscopy images. By tailoring the architecture to the unique characteristics of skin lesion images, we hypothesized that the model would achieve improved classification performance.

5.5.5.1 Results of Experiment 5

After 30 epochs of training, the model demonstrated promising performance. The test loss reached close to 0.15, corresponding to a test accuracy of **0.9772**. This result suggests that the model has learned to effectively distinguish between different types of skin lesions with a high degree of reliability. 5.13 presents a detailed breakdown of the classification metrics for each type of skin lesion, including Precision, Recall, and F1-score. Additionally, the table provides the weighted average values for these metrics across all lesion types, along with the overall accuracy.

(a)



(b)

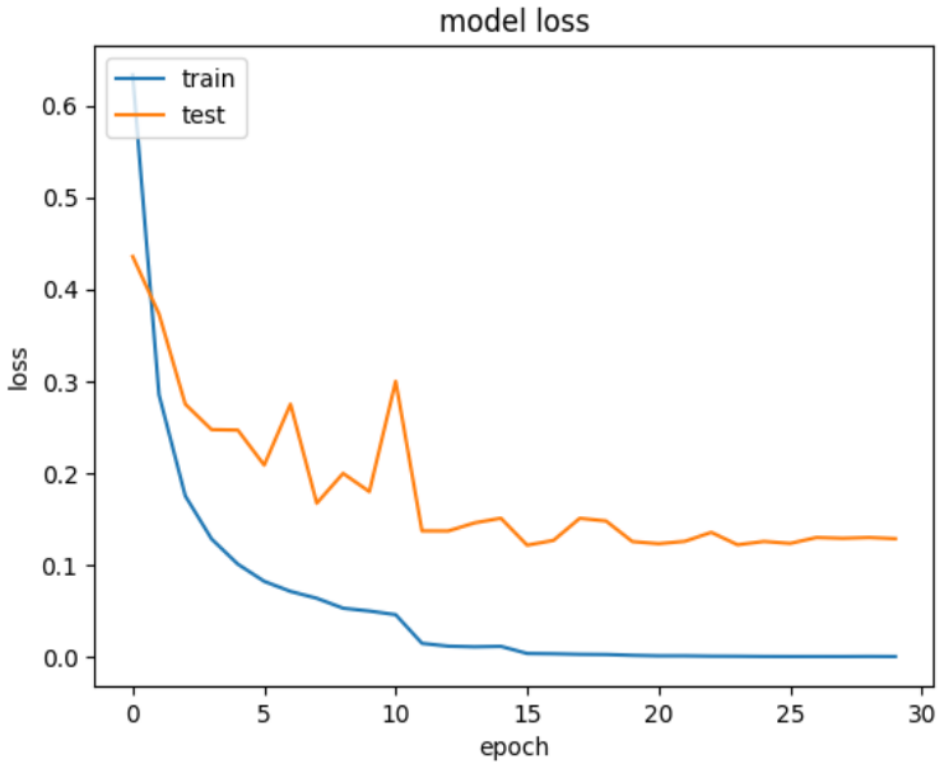


FIGURE 5.10: Experiment 5: (a) Accuracy versus epochs, (b) Loss versus the number of epochs.

TABLE 5.13: Performance of the model with Modified Architecture of Densenet169, Hair removed and skin lesion segmented Balanced dataset.

Class	**Precision**	**Recall**	**F1-Score**
AK	0.99	1.00	1.00
BCC	0.98	0.99	0.98
BKL	0.96	0.98	0.97
DF	0.99	1.00	1.00
MEL	0.95	0.93	0.94
NV	0.94	0.93	0.94
SCC	0.99	1.00	1.00
VASC	1.00	1.00	1.00
Average	0.97	0.97	0.97
Accuracy		0.97	

Figure 5.11 presents a visual representation of the confusion matrix, offering a comprehensive overview of the model's classification performance across different skin lesion categories. This matrix provides valuable insights into the model's strengths and limitations in distinguishing between various lesion types. The confusion matrix illustrates the distribution of correct classifications along the diagonal, while off-diagonal elements represent misclassifications. By examining these off-diagonal elements, we can discern the nature and frequency of classification errors.

5.10 illustrates the model's learning progression throughout the training process. Specifically, 5.10 (a) depicts the relationship between training and test Accuracy for the number of training epochs on the ISIC2019 test dataset. Correspondingly, 5.10 (b) presents the evolution of training and test loss versus the training epochs, also evaluated on the ISIC2019 test dataset. This graph offers valuable information about the model's convergence and optimization process, helping to identify any instances of divergence or instability during training.

These results collectively provide a comprehensive assessment of the model's performance. Experiment 5 shows that the Modified Architecture of DenseNet169 with a hair-removed and skin-lesion segmented balanced ISIC2019 dataset using these hyperparameters achieved an overall accuracy of 97.72%.

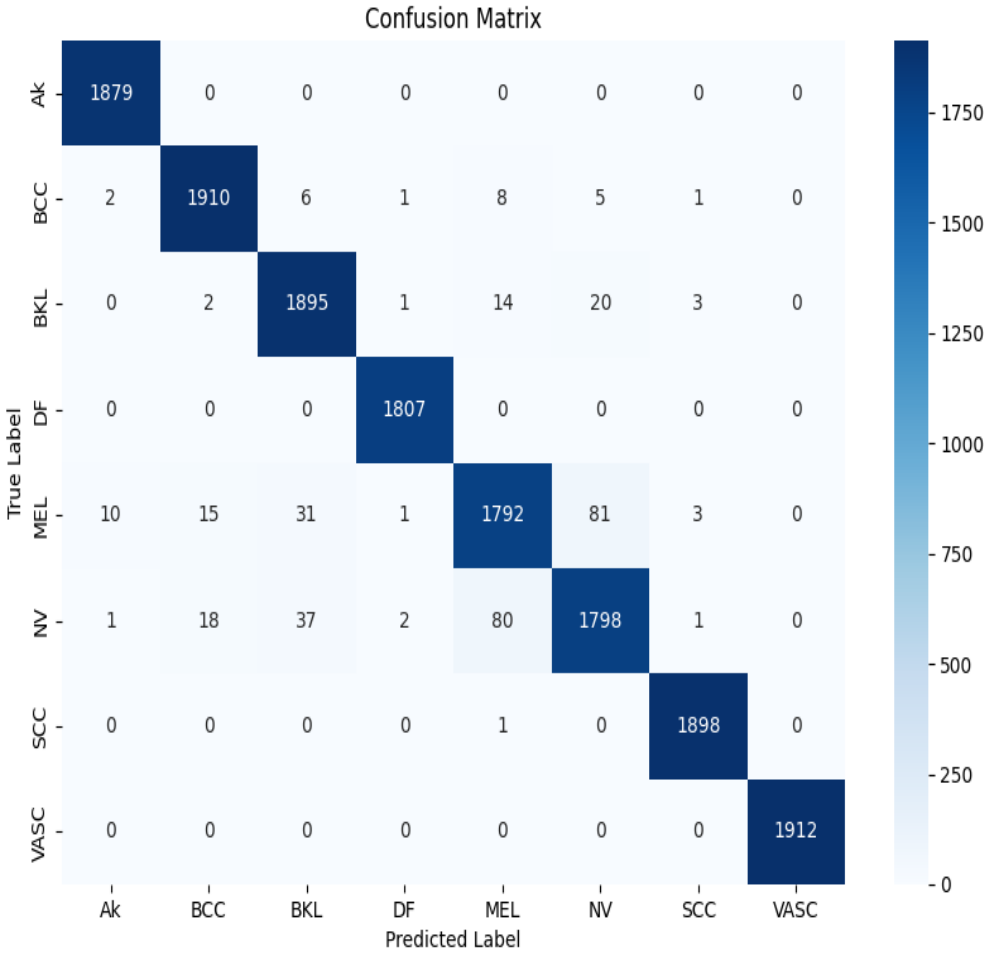


FIGURE 5.11: Confusion Matrix of Final Result.

5.5.6 Comparison and Discussion

Our research involved a series of five carefully designed experiments, each building upon the previous one to explore different aspects of skin lesion classification:

5.5.6.1 Baseline Experiment

We started with the original DenseNet169 architecture, applying it to the unmodified ISIC2019 dataset. This experiment served as our baseline, using the data in its raw form without any preprocessing or augmentation.

TABLE 5.14: Classification Accuracy of different combinations of various parts of the model.

Experiment	Model	Architecture	Data Augmentation	Hair Removed	Lesion Segmentation	Accuracy
Experiment 1	DenseNet169	Original	×	×	×	**87.06%**
Experiment 2	DenseNet169	Original	✓	×	×	**92.28%**
Experiment 3	DenseNet169	Original	✓	✓	×	**94.81%**
Experiment 4	DenseNet169	Original	✓	✓	✓	**96.31%**
Experiment 5	DenseNet169	Modified	✓	✓	✓	**97.72%**

5.5.6.2 Balanced Dataset Experiment

For our second experiment, we maintained the original DenseNet169 architecture but introduced dataset balancing. This step aimed to address any class imbalance issues in the original dataset.

5.5.6.3 Hair Removal Experiment

Building on the balanced dataset, our third experiment incorporated a hair removal process. We used the original DenseNet169 architecture on this cleaned, balanced dataset to assess the impact of hair artifacts on classification performance.

5.5.6.4 Segmentation Experiment

The fourth experiment combined multiple preprocessing steps. We used the original DenseNet169 architecture on a dataset that was balanced, had hair removed, and underwent skin lesion segmentation.

5.5.6.5 Modified Architecture Experiment

Our final experiment was the most comprehensive. We used the preprocessed dataset (balanced, hair-removed, and segmented) and introduced modifications to the DenseNet169 architecture itself. For a thorough analysis of each of the five trials, we have assembled Table 5.14 and Figure 5.12. Our series of experiments revealed a progressive improvement in skin lesion classification accuracy.

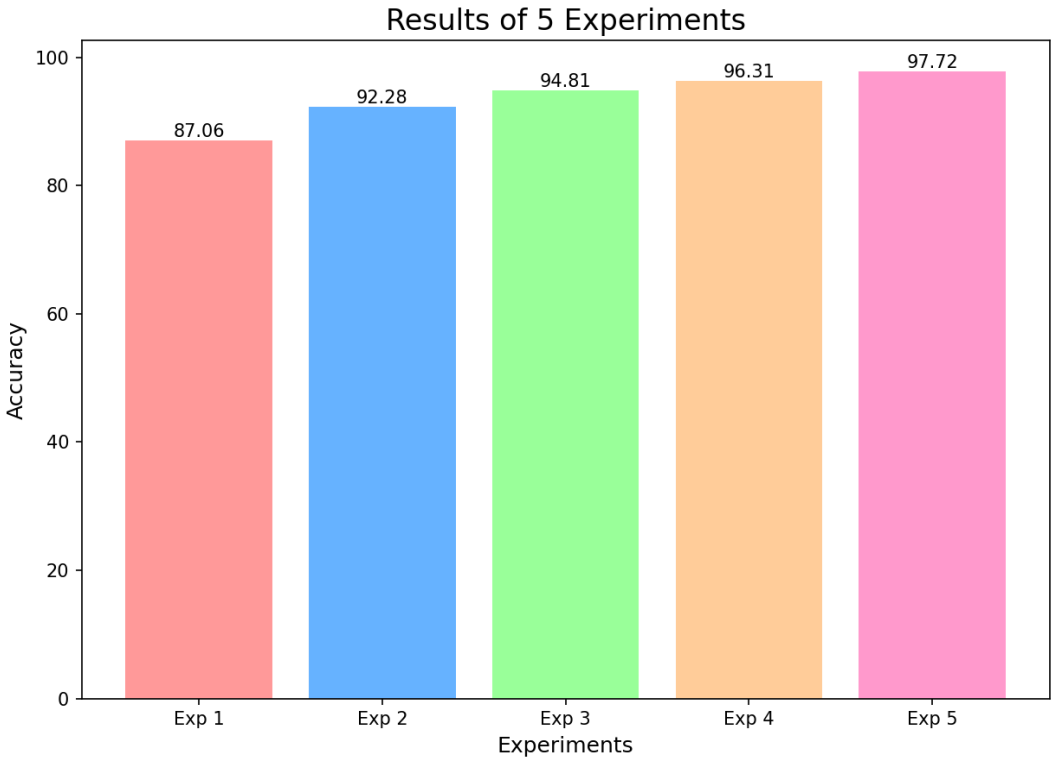


FIGURE 5.12: Experimental Results.

Using the original DenseNet169 with the imbalanced ISIC2019 dataset, we achieved an accuracy of 87%. This served as our starting point. By applying data augmentation techniques to balance the dataset, accuracy improved to 92.28%. This represents a significant 5% increase over the baseline, highlighting the importance of addressing class imbalance. Implementing our proposed EHR framework further boosted accuracy to 94.81%. This 2.53% improvement demonstrates the impact of removing hair artifacts from the images. By incorporating skin lesion segmentation using a deep residual U-Net, we reached an accuracy of 96.31%. This additional 1.5% gain shows the value of focusing the model's attention on the lesion area. In our final experiment, we modified the DenseNet169 architecture, pushing the accuracy to 97.72%. This represents a 1.41% improvement over the previous step.

These results showcase a cumulative improvement of 10.72% from our baseline to the final model. Each step in our process contributed to enhancing the model's performance, with the most substantial gains coming from dataset balancing and hair removal.

We conducted a comprehensive comparison with other existing methods in the

TABLE 5.15: Classification Accuracy comparison of existing techniques.

Method	**Dataset**	**Accuracy, %**
Naveed et al. [42]	ISIC2019	94.9
Hasan et al. [39]	ISIC2019	93.5
Villa-Pulgarin et al. [40]	ISIC2019	93.0
Li et al. [44]	ISIC2019	89.0
Ahammed et al. [38]	ISIC2019	95.0
Khan et al. [43]	ISIC2019	89.0
Zhao et al. [45]	ISIC2019	93.6
Santos et al. [46]	ISIC2019	89.1
Rezaoana et al. [41]	ISIC2019	79.4
Ours	ISIC2019	**97.72**

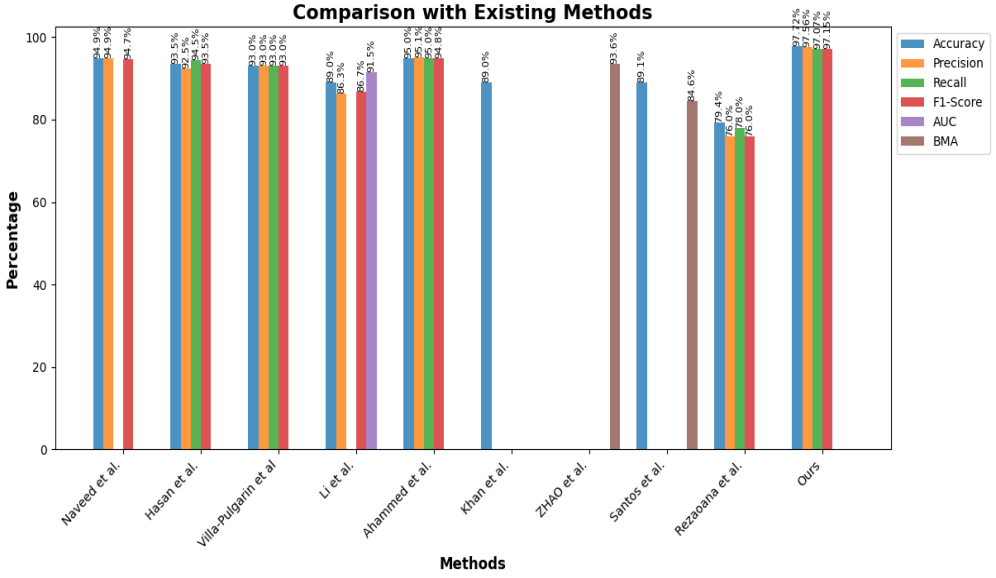


FIGURE 5.13: Comparison of Existing Methods.

field. This comparative analysis is visually represented in Figure 5.13 of our study. Our best-performing model, which achieved an accuracy of 97.72%, demonstrated superior performance when compared to all other existing techniques we evaluated.

The fact that our method outperforms existing techniques underscores the significance of our contributions. It suggests that our comprehensive approach addressing dataset imbalance, removing hair artifacts, performing lesion segmentation, and modifying the network architecture offers a promising direction for future research and potential clinical applications.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

Our comprehensive study on skin lesion classification and segmentation has yielded promising results that advance the field of automated dermatological diagnosis. Through a series of carefully designed experiments, we have demonstrated the effectiveness of a multi-faceted approach to improving classification accuracy. Starting with a baseline accuracy of 87% using the original DenseNet169 architecture on the imbalanced ISIC2019 dataset, we systematically addressed various challenges in skin lesion analysis. By implementing dataset balancing, we achieved a significant 5% improvement in accuracy. Our proposed EHR framework further boosted performance by 2.53%, highlighting the importance of artifact removal in image preprocessing. The integration of skin lesion segmentation using a deep residual U-Net contributed an additional 1.5% improvement. Finally, our modified DenseNet169 architecture pushed the accuracy to an impressive 97.72%, outperforming existing techniques in the field. In parallel, our segmentation model achieved state-of-the-art performance on the ISIC2018 dataset, with an IoU of 0.8981 and a Dice coefficient of 0.946. These results underscore the effectiveness of our approach in accurately delineating skin lesions, a crucial step in automated analysis.

While these results are encouraging and outperform existing techniques, there remains room for further improvement and exploration. A key direction for future research lies in addressing the challenge of inter-class similarity, which can be a significant hurdle in accurate skin lesion classification. To resolve this problem we can focus on the design of innovative loss functions specifically tailored to reduce inter-class similarity.

Bibliography

[1] Sadik, R., Majumder, A., Biswas, A.A., Ahammad, B. and Rahman, M.M., 2023. An in-depth analysis of Convolutional Neural Network architectures with transfer learning for skin disease diagnosis. Healthcare Analytics, 3, p.100143.

[2] Lawton, S.J.N.T., 2019. Skin 1: the structure and functions of the skin. Nurs. times, 115, pp.30-33.

[3] Apalla, Z., Nashan, D., Weller, R.B. and Castellsagué, X., 2017. Skin cancer: epidemiology, disease burden, pathophysiology, diagnosis, and therapeutic approaches. Dermatology and therapy, 7, pp.5-19.

[4] Armstrong, B.K. and Kricker, A., 2001. The epidemiology of UV induced skin cancer. Journal of photochemistry and photobiology B: Biology, 63(1-3), pp.8-18.

[5] Narayanan, D.L., Saladi, R.N. and Fox, J.L., 2010. Ultraviolet radiation and skin cancer. International journal of dermatology, 49(9), pp.978-986.

[6] Gandini, S., Sera, F., Cattaruzza, M.S., Pasquini, P., Zanetti, R., Masini, C., Boyle, P. and Melchi, C.F., 2005. Meta-analysis of risk factors for cutaneous melanoma: III. Family history, actinic damage and phenotypic factors. European journal of cancer, 41(14), pp.2040-2059.

[7] Florell, S. R., Boucher, K. M., Garibotti, G., Astle, J., Kerber, R., Mineau, G., ... Cannon-Albright, L. A. (2005). Population-based analysis of prognostic factors and survival in familial melanoma. Journal of Clinical Oncology, 23(28), 7168-7177.

[8] Euvrard, S., Kanitakis, J., Claudy, A. (2003). Skin cancers after organ transplantation. New England Journal of Medicine, 348(17), 1681-1691.

[9] Siegel, J.A., Korgavkar, K. and Weinstock, M.A., 2017. Current perspective on actinic keratosis: a review. British Journal of Dermatology, 177(2), pp.350-358.

[10] Cameron, M.C., Lee, E., Hibler, B.P., Barker, C.A., Mori, S., Cordova, M., Nehal, K.S. and Rossi, A.M., 2019. Basal cell carcinoma: Epidemiology; pathophysiology; clinical and histological subtypes; and disease associations. Journal of the American Academy of Dermatology, 80(2), pp.303-317.

[11] Hafner, C. and Vogt, T., 2008. Seborrheic keratosis. JDDG: Journal der Deutschen Dermatologischen Gesellschaft, 6(8), pp.664-677.

[12] Posso-De Los Rios, C.J., Lara-Corrales, I. and Ho, N., 2014. Dermatofibrosarcoma protuberans in pediatric patients: a report of 17 cases. Journal of Cutaneous Medicine and Surgery, 18(3), pp.180-185.

[13] Schaffer, J.V., 2015. Update on melanocytic nevi in children. Clinics in dermatology, 33(3), pp.368-386.

[14] Rastrelli, M., Tropea, S., Rossi, C.R. and Alaibac, M., 2014. Melanoma: epidemiology, risk factors, pathogenesis, diagnosis and classification. In vivo, 28(6), pp.1005-1011.

[15] Burton, K.A., Ashack, K.A. and Khachemoune, A., 2016. Cutaneous squamous cell carcinoma: a review of high-risk and metastatic disease. American journal of clinical dermatology, 17, pp.491-508.

[16] Hicks, M.J. and Flaitz, C.M., 2000. Oral mucosal melanoma: epidemiology and pathobiology. Oral oncology, 36(2), pp.152-169.

[17] Celebi, M.E., Iyatomi, H., Schaefer, G. and Stoecker, W.V., 2009. Lesion border detection in dermoscopy images. Computerized medical imaging and graphics, 33(2), pp.148-153.

[18] Conforti, C., Giuffrida, R., Vezzoni, R., Resende, F.S., di Meo, N. and Zalaudek, I., 2020. Dermoscopy and the experienced clinicians. International Journal of Dermatology, 59(1), pp.16-22.

[19] Yélamos, O., Braun, R.P., Liopyris, K., Wolner, Z.J., Kerl, K., Gerami, P. and Marghoob, A.A., 2019. Dermoscopy and dermatopathology correlates of cutaneous neoplasms. Journal of the American Academy of Dermatology, 80(2), pp.341-363.

[20] Zalaudek, I., Argenziano, G., Di Stefani, A., Ferrara, G., Marghoob, A.A., Hofmann-Wellenhof, R., Soyer, H.P., Braun, R. and Kerl, H., 2005. Dermoscopy in general dermatology. Dermatology, 212(1), pp.7-18.

[21] Lucas, R.M., Yazar, S., Young, A.R., Norval, M., De Gruijl, F.R., Takizawa, Y., Rhodes, L.E., Sinclair, C.A. and Neale, R.E., 2019. Human health in relation to exposure to solar ultraviolet radiation under changing stratospheric ozone and climate. Photochemical Photobiological Sciences, 18(3), pp.641-680.

[22] Sacchetto, L., Zanetti, R., Comber, H., Bouchardy, C., Brewster, D.H., Broganelli, P., Chirlaque, M.D., Coza, D., Galceran, J., Gavin, A. and Hackl, M., 2018. Trends in incidence of thick, thin and in situ melanoma in Europe. European journal of cancer, 92, pp.108-118.

[23] Duarte, A.F., Sousa-Pinto, B., Freitas, A., Delgado, L., Costa-Pereira, A. and Correia, O., 2018. Skin cancer healthcare impact: A nation-wide assessment of an administrative database. Cancer epidemiology, 56, pp.154-160.

[24] Stolz, W.J.E.J.D., 1994. ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. Eur J Dermatol, 4, pp.521-527.

[25] Argenziano, G., Fabbrocini, G., Carli, P., De Giorgi, V., Sammarco, E. and Delfino, M., 1998. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: comparison of the ABCD rule of dermatoscopy and

a new 7-point checklist based on pattern analysis. Archives of dermatology, 134(12), pp.1563-1570.

[26] Jamil, U., Akram, M.U., Khalid, S., Abbas, S. and Saleem, K., 2016. Computer based melanocytic and nevus image enhancement and segmentation. BioMed Research International, 2016(1), p.2082589.

[27] Mahbod, A., Schaefer, G., Wang, C., Ecker, R. and Ellinge, I., 2019, May. Skin lesion classification using hybrid deep neural networks. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1229-1233). IEEE.

[28] Ünver, H.M. and Ayan, E., 2019. Skin lesion segmentation in dermoscopic images with combination of YOLO and grabcut algorithm. Diagnostics, 9(3), p.72.

[29] McNoe, B.M., Morgaine, K.C. and Reeder, A.I., 2021. Effectiveness of sun protection interventions delivered to adolescents in a secondary school setting: a systematic review. Journal of skin cancer, 2021(1), p.6625761.

[30] Kasmi, R. and Mokrani, K., 2016. Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule. IET Image Processing, 10(6), pp.448-455.

[31] Celebi, M.E., Kingravi, H.A., Uddin, B., Iyatomi, H., Aslandogan, Y.A., Stoecker, W.V. and Moss, R.H., 2007. A methodological approach to the classification of dermoscopy images. Computerized Medical imaging and graphics, 31(6), pp.362-373.

[32] LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. nature, 521(7553), pp.436-444.

[33] Wang, L., Lee, C.Y., Tu, Z. and Lazebnik, S., 2015. Training deeper convolutional networks with deep supervision. arXiv preprint arXiv:1505.02496.

[34] Sharif Razavian, A., Azizpour, H., Sullivan, J. and Carlsson, S., 2014. CNN features off-the-shelf: an astounding baseline for recognition. In Proceedings

of the IEEE conference on computer vision and pattern recognition workshops (pp. 806-813).

[35] Zhu, L., Feng, S., Zhu, W. and Chen, X., 2020, February. ASNet: An adaptive scale network for skin lesion segmentation in dermoscopy images. In Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging (Vol. 11317, pp. 226-231). SPIE.

[36] Telea, A., 2004. An image inpainting technique based on the fast marching method. Journal of graphics tools, 9(1), pp.23-34.

[37] Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q., 2017. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[38] Ahammed, M., Al Mamun, M. and Uddin, M.S., 2022. A machine learning approach for skin disease detection and classification using image segmentation. Healthcare Analytics, 2, p.100122.

[39] Hasan, S.M., Mamun, A. and Srizon, A.Y., 2023, September. Enhancing Multi-Class Skin Lesion Classification with Modified EfficientNets: Advancing Early Detection of Skin Cancer. In 2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD) (pp. 94-98). IEEE.

[40] Villa-Pulgarin, J.P., Ruales-Torres, A.A., Arias-Garzon, D., Bravo-Ortiz, M.A., Arteaga-Arteaga, H.B., Mora-Rubio, A., Alzate-Grisales, J.A., Mercado-Ruiz, E., Hassaballah, M., Orozco-Arias, S. and Cardona-Morales, O., 2022. Optimized Convolutional Neural Network Models for Skin Lesion Classification. Computers, Materials Continua, 70(2).

[41] Rezaoana, N., Hossain, M.S. and Andersson, K., 2020, December. Detection and classification of skin cancer by using a parallel CNN model. In 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE) (pp. 380-386). IEEE.

[42] Naveed, A., Naqvi, S.S., Khan, T.M. and Razzak, I., 2024. PCA: progressive class-wise attention for skin lesions diagnosis. Engineering Applications of Artificial Intelligence, 127, p.107417.

[43] Attique Khan, M., Sharif, M., Akram, T., Kadry, S. and Hsu, C.H., 2022. A two-stream deep neural network-based intelligent system for complex skin cancer types classification. International Journal of Intelligent Systems, 37(12), pp.10621-10649.

[44] Li, Y., Yang, X., Chen, Y., Chen, S., Zeng, B. and Cai, H., 2023, December. Adaptive Multi modal Data Fusion Method for Dermatosis Diagnosis. In 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 2904-2911). IEEE.

[45] Zhao, C., Shuai, R., Ma, L., Liu, W., Hu, D. and Wu, M., 2021. Dermoscopy image classification based on StyleGAN and DenseNet201. Ieee Access, 9, pp.8659-8679.

[46] Santos, F., Silva, F. and Georgieva, P., 2021, August. Transfer learning for skin lesion classification using convolutional neural networks. In 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA) (pp. 1-6). IEEE.

[47] Lee, T., Ng, V., Gallagher, R., Coldman, A. and McLean, D., 1997. Dullrazor®: A software approach to hair removal from images. Computers in biology and medicine, 27(6), pp.533-543.

[48] F.-Y. Xie, S.-Y. Qin, Z.-G. Jiang, and R.-S. Meng, PDE-based unsupervised repair of hair-occluded information in dermoscopy images of melanoma, Comput. Med. Imag. Graph., vol. 33, no. 4, pp. 275282, Jun. 2009.

[49] Q. Abbas, M. E. Celebi, and I. F. Garcia, Hair removal methods: A comparative study for dermoscopy images, Biomed. Signal Process. Control, vol. 6, no. 4, pp. 395404, Oct. 2011.

[50] Huang, A., Kwan, S.Y., Chang, W.Y., Liu, M.Y., Chi, M.H. and Chen, G.S., 2013, July. A robust hair segmentation and removal approach for clinical

images of skin lesions. In 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC) (pp. 3315-3318). IEEE.

[51] Bibiloni, P., Gonz´alez-Hidalgo, M. and Massanet, S., 2017. Skin hair removal in dermoscopic images using soft color morphology. In Artificial Intelligence in Medicine: 16th Conference on Artificial Intelligence in Medicine, AIME 2017, Vienna, Austria, June 21-24, 2017, Proceedings 16 (pp. 322-326). Springer International Publishing.

[52] Hossain, S.I., Roy, S.S., De Herve, J.D.G., Mercer, R.E. and Nguifo, E.M., 2023. A skin lesion hair mask dataset with fine-grained annotations. Data in Brief, 48, p.109249.

[53] Zhang, Z., Liu, Q. and Wang, Y., 2018. Road extraction by deep residual u-net. IEEE Geoscience and Remote Sensing Letters, 15(5), pp.749-753.

[54] Bertalmio, M., Bertozzi, A.L. and Sapiro, G., 2001, December. Navier stokes, fluid dynamics, and image and video inpainting. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001 (Vol. 1, pp. II). IEEE.

[55] Shaham, T.R., Dekel, T. and Michaeli, T., 2019. Singan: Learning a generative model from a single natural image. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 4570-4580).

[56] Jam, J., Kendrick, C., Drouard, V., Walker, K., Hsu, G.S. and Yap, M.H., 2021. R-mnet: A perceptual adversarial network for image inpainting. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 2714-2723).

[57] Yap, H., Skin Hair dataset: Setting the benchmark for effective hair inpainting methods for improving the image quality of dermoscopic images.

[58] Sethian, J.A., 1996. A fast marching level set method for monotonically advancing fronts. proceedings of the National Academy of Sciences, 93(4), pp.1591-1595.

[59] Ahmed, T.U., Hossain, S., Hossain, M.S., ul Islam, R. and Andersson, K., 2019, May. Facial expression recognition using convolutional neural network with data augmentation. In 2019 Joint 8th International Conference on Informatics, Electronics Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision Pattern Recognition (icIVPR) (pp. 336-341). IEEE.

[60] Friedman, R.J., Rigel, D.S. and Kopf, A.W., 1985. Early detection of malignant melanoma: the role of physician examination and self-examination of the skin. CA: a cancer journal for clinicians, 35(3), pp.130-151.

[61] Birkenfeld, J.S., Tucker-Schwartz, J.M., Soenksen, L.R., Avilés-Izquierdo, J.A. and Marti-Fuster, B., 2020. Computer-aided classification of suspicious pigmented lesions using wide-field images. Computer methods and programs in biomedicine, 195, p.105631.

[62] "International skin imaging collaboration (isic) challenge 2019," 2019. [Online]. Available: https://challenge2019.isic-archive.com/

[63] Yap, H., Skin Hair dataset: Setting the benchmark for effective hair inpainting methods for improving the image quality of dermoscopic images.

[64] Wang, J., Wei, L., Wang, L., Zhou, Q., Zhu, L. and Qin, J., 2021. Boundary-aware transformers for skin lesion segmentation. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24 (pp. 206-216). Springer International Publishing.

[65] Tang, F., Xu, Z., Huang, Q., Wang, J., Hou, X., Su, J. and Liu, J., 2023, October. DuAT: Dual-aggregation transformer network for medical image segmentation. In Chinese Conference on Pattern Recognition and Computer Vision (PRCV) (pp. 343-356). Singapore: Springer Nature Singapore.

[66] Ben˘cevi´c, M., Gali´c, I., Habijan, M. and Babin, D., 2021. Training on polar image transformations improves biomedical image segmentation. IEEE access, 9, pp.133365-133375.

[67] Dzieniszewska, A., Garbat, P. and Piramidowicz, R., 2024. Improving skin lesion segmentation with self-training. Cancers, 16(6), p.1120.

[68] Vujović, Ž., 2021. Classification model evaluation metrics. International Journal of Advanced Computer Science and Applications, 12(6), pp.599-606.

[69] Wang, Z., Wang, E. and Zhu, Y., 2020. Image segmentation evaluation: a survey of methods. Artificial Intelligence Review, 53(8), pp.5637-5674.

[70] Samajdar, T. and Quraishi, M.I., 2015. Analysis and evaluation of image quality metrics. In Information Systems Design and Intelligent Applications: Proceedings of Second International Conference INDIA 2015, Volume 2 (pp. 369-378). Springer India.