Faculty of Computer Science and Information Technology

Automatic Segmentation and Classification of Skin Lesions in Dermoscopic Images

Adil Humayun Khan

Doctor of Philosophy
2024

Automatic Segmentation and Classification of Skin Lesions in Dermoscopic Images

Adil Humayun Khan

A thesis submitted

In fulfillment of the requirements for the degree of Doctor of Philosophy

(Computer Science)

Faculty of Computer Science and Information Technology
UNIVERSITI MALAYSIA SARAWAK
2024

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature 

Name: Adil Humayun Khan

Matric No.: 17010172

Faculty of Computer Science and Information Technology

Universiti Malaysia Sarawak

Date : 12/31/2023

ACKNOWLEDGEMENT

I would like to acknowledge all the people who have supported me throughout my work but foremost I want to thank Almighty Allah the most merciful and gracious. I would like to express my deepest gratitude to my parents and wife as their support and encouragement was very crucial during all these years.

Then I would like to acknowledge the most important person, my mentor and advisor Associate Professor Dr. Dayang NurFatimah binti Awang Iskandar. Her continuous guidance and support during proposal, research publications and thesis write up was immense. This work would not have been possible without her remarkable input and supervision. She guided me at every single step of this study with her prompt and crucial support. I felt privileged and proud to work under her supervision.

My sincere gratitude to the Centre for Graduate Studies, and management of UNIMAS for the advice and support provided by them during my period of study and stay in Universiti Malaysia Sarawak.

Finally, I would like to thank my all colleagues and friends for their valuable feedback and assistance throughout the studies. Thank you all.

ABSTRACT

Skin cancer has a malignant type that is Melanoma and two benign types Nevus and Seborrheic Keratosis (SK). Diagnosis of actual skin cancer type is very challenging for physicians because there is no fixed colour, area, size, or boundary associated with a specific type. There is a remarkable resemblance between different types of skin cancer. Therefore, a computer-based process is required to help physicians in diagnosing actual cancer types timely. This process consists of four steps: pre-processing, segmentation, feature extraction, and classification. In this research, we evaluate proposed algorithms on two datasets, International Skin Imaging Collaboration (ISIC) and PH2 (Dermatology Service of Hospital Pedro Hispano, Matosinhos, Portugal). Statistical techniques such as Anisotropic Diffusion Filter (ADF), local contrast enhancement, and haze reduction in the CIELAB colour space, are incorporated in the proposed algorithms for the pre-processing step. Hair removal is implemented through black-hat morphological processing and total variation based inpainting. For segmentation, the first proposed algorithm is based on the boundary condition model, which is tested over the ISIC dataset and achieved 96% of accuracy. The second segmentation algorithm combines Delaunay triangulation clustering in the spatial domain and Particle Swarm Optimization (PSO). This proposed algorithm achieved segmentation accuracy of 96.8% and 92.1% for ISIC and PH2 datasets respectively. The third proposed segmentation algorithm involves two pipelines for feature extraction: split & merge methods and Contextual Encoding Network (EncNet) with an attention mechanism. This proposed algorithm achieved segmentation accuracy of 97.8% and 96.7% for ISIC and PH2 datasets respectively. The first proposed classification algorithm utilizes a Convolution Neural Network (CNN), in which the number of parameters and layers are reduced significantly, and 96% of classification accuracy is achieved on the ISIC dataset. The second proposed

classifier extracts statistical features, and a boost ensemble learning technique is implemented using Support Vector Machines (SVM) as initial classifiers and Artificial Neural Networks (ANN) as the final classifier. This proposed classifier achieved 97.9% classification accuracy on the ISIC dataset. In the third classification algorithm, hybrid features are extracted using AlexNet and VGG-16 through a transfer learning approach where parameter manipulation is implemented to simplify the network. This proposed classifier achieved 98.2% classification accuracy on the ISIC dataset. These algorithms are proposed while implying modifications to existing statistical, machine, and deep learning methods.

Keywords: Skin Lesion, Melanoma, Hair and artifact removal, Segmentation, Features extraction, Classification, Deep learning

Segmentasi Automatik dan Klasifikasi Lesi Kulit dalam Imej Dermoskopik

ABSTRAK

Terdapat dua jenis tumor, iaitu malignan dan benigna. Melanoma adalah tumor yang dihasilkan oleh transformasi malignan melanosit yang menyebabkan barah. Benigna adalah tumor jinak yang terdiri daripada Nevus dan Keratosis Seborrheic (SK). Bagi mengenalpasti jenis kanser kulit sebenar adalah sangat mencabar bagi pakar perubatan kerana tiada warna, kawasan, saiz atau sempadan tetap yang dikaitkan dengan jenis tertentu. Terdapat persamaan yang luar biasa antara pelbagai jenis kanser kulit. Oleh itu, proses berasaskan komputer diperlukan untuk membantu pakar perubatan dalam mendiagnosis jenis kanser sebenar. Proses ini terdiri daripada empat langkah: pra-pemprosesan, pembahagian, pengekstrakan ciri, dan pengelasan. Dalam penyelidikan ini, kami akan menguji algoritma yang dicadangkan pada dua set data, Kolaborasi Pengimejan Kulit Antarabangsa (ISIC) dan PH2 (Dermatologi Service Hospital Pedro Hispano, Matosinhos, Portugal). Teknik statistik seperti Penapis Resapan Anisotropik (ADF), peningkatan kontras tempatan dan pengurangan jerebu dalam ruang warna CIELAB, digabungkan dalam algoritma yang dicadangkan untuk peringkat pra-pemprosesan. Penyingkiran rambut dilaksanakan melalui pemprosesan morfologi topi hitam dan pengecatan berasaskan variasi total. Untuk pembahagian, algoritma pertama yang dicadangkan adalah berdasarkan model keadaan sempadan, yang diuji ke atas set data ISIC dan mencapai 96% ketepatan. Algoritma segmentasi kedua menggabungkan pengelompokan triangulasi Delaunay dalam domain spatial dan Particle Swarm Optimization (PSO). Algoritma yang dicadangkan ini mencapai ketepatan pembahagian masing-masing sebanyak 96.8% dan 92.1% untuk set data ISIC dan PH2. Algoritma segmentasi ketiga yang dicadangkan melibatkan dua saluran paip untuk pengekstrakan ciri: kaedah split & merge dan Rangkaian Pengekodan Kontekstual (EncNet)

dengan mekanisme perhatian. Algoritma yang dicadangkan ini memperoleh ketepatan pembahagian masing-masing sebanyak 97.8% dan 96.7% untuk set data ISIC dan PH2. Algoritma pengelasan pertama yang dicadangkan menggunakan Rangkaian Neural Convolution (CNN), di mana bilangan parameter dan lapisan dikurangkan dengan ketara, dan 96% ketepatan pengelasan dicapai pada dataset ISIC. Pengelas kedua yang dicadangkan mengekstrak ciri statistik dan teknik pembelajaran ensemble rangsangan dilaksanakan menggunakan Mesin Vektor Sokongan (SVM) sebagai pengelas awal dan Rangkaian Neural Buatan (ANN) sebagai pengelas akhir. Pengelas yang dicadangkan ini mencapai ketepatan pengelasan 97.9% pada set data ISIC. Dalam algoritma pengelasan ketiga, ciri hibrid diekstrak menggunakan AlexNet dan VGG-16 melalui pendekatan pembelajaran pemindahan di mana manipulasi parameter dilaksanakan untuk memudahkan rangkaian. Pengelas yang dicadangkan ini mencapai ketepatan pengelasan 98.2% pada set data ISIC. Algoritma ini dicadangkan sambil membayangkan pengubahsuaian kepada kaedah statistik, mesin dan pembelajaran mendalam yang sedia ada.

Kata kunci: *Lesi Kulit, Rambut dan penyingkiran artifak, Segmentasi, Pengekstrakan ciri, Pengelasan, Pembelajaran Mendalam*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Skin cancer also known as skin lesion is a disease that occurs when abnormal cells in the skin grow uncontrollably and form malignant tumors. These tumors can invade and damage surrounding tissues, in some cases, spread to other parts of the body. Not all types of skin cancer or lesion are malignant, some types are benign in nature. These benign types of skin lesions do not invade nearby tissues or spread to other parts of the body. While they may still grow, they typically remain localized and do not pose the same risks as a malignant type of skin lesion (Eedy, 2020). In this research, we will focus on one malignant and two benign types of skin lesion, which in this thesis we referred to as skin lesion classes. The malignant type of skin lesion is Melanoma. The two benign types are Nevus and Seborrheic Keratosis (SK). American Cancer Society has provided the statistics that almost 197,700 new cases of skin cancer diagnosed in 2022, out of which 97,920 were belonged to benign skin lesion and 99,780 were malignant melanoma. Skin cancer is expected to be the fifth most diagnosed cancer for both males and females in 2022 (Siegel et al., 2022). Fatality rates caused by melanoma skin lesions are the highest among skin cancer patients. This rate has been on the rise for the past few years, and it is expected that number of deaths due to melanoma will increase by 4.4% in 2023 (Siegel et al., 2023). It is crucial to remember that early detection, regular skin examinations, and protecting your skin from excessive sun exposure are key factors in preventing and effectively treating melanoma. If you notice any changes in your skin, such as new moles, changes in existing moles, or any suspicious growths, it is essential to consult a healthcare professional for evaluation and appropriate treatment.

Dermoscopy is a non-invasive process through which dermatologists diagnose the actual skin cancer type and cure accordingly. Dermatologists use dermatoscope to capture different skin images under varying conditions. This device uses a magnifier that is illuminated by a light input which helps in capturing the colour and microstructures of skin lesions more accurately. The images obtained from a dermatoscope are called dermoscopic images (Johr, 2002). Currently, most computer-automated systems use dermoscopic images for analyzing and diagnosing skin cancer patients. Melanoma can be cured easily with surgical procedures but due to the similarity among skin lesion types and the presence of artifacts accurate diagnosis is almost impossible even by experienced dermatologists therefore an automated computer-based diagnosis is required to overcome these problems (Mutlag et al., 2020).

A complete process for the detection of skin lesion consists of pre-processing, segmentation, feature extraction, and classification. In pre-processing stage artifacts that affect the lesion needs to be removed and skin hairs, in dermoscopic images are major artifacts that can make the detection process difficult. Human skin is covered with hairs of different colours, orientations, and textures. To classify skin lesion into either melanoma, nevus, or SK, the process of hair removal from dermoscopic images is very crucial (Sekniashvili, 2022; Zortea et al., 2017). After pre-processing, segmentation is applied to extract the lesion from dermoscopic images. This process is also very challenging due to the variable nature, texture, colour, and size of skin lesion (Barın & Güraksın, 2022; Yousef et al., 2022). Once a segmented lesion is secured, feature extraction is applied to understand the physical appearance of skin lesions with respect to all skin cancer types. Later these features are used to classify the skin lesion into different types, which are melanoma, nevus, and SK (Jeyakumar et al., 2022).

Human skin is divided into three main layers, but the topmost two layers are considered more important for skin cancer. The uppermost layer is called the Epidermis and the second layer is called the Dermis, usually skin cancer starts in the uppermost layer epidermis, more details on skin layers are provided in Chapter 2. Countries with Caucasian populations are affected more due to this disease (Wang et al., 2020). Ultraviolet (UV) radiations from the sun and artificial tanning beds are considered to be the major cause of skin cancer-type melanoma In fact as per research conducted by dermatologists nearly 86% of melanomas can be attributed to exposure to UV radiation from the sun and this number is believed to be on the increasing side (Eedy, 2020). Figure 1.1 shows different types of UV rays from the sun penetrates through different levels of human skin, image is courtesy from Adobe Stock, 2022. In earlier days, awareness of this disease was not very common that's why almost 54% of the population is affected by melanoma each year, and around 32% of the population could not benefit from dermatology care (Gouda et al., 2022). But more awareness regarding skin cancer, and an exponential increase in the rate of this disease in the last three decades, has changed the practice of dermatologists. Recent research showed that almost 50% of the workload of dermatologists is focused on this disease (Leigh, 2021).

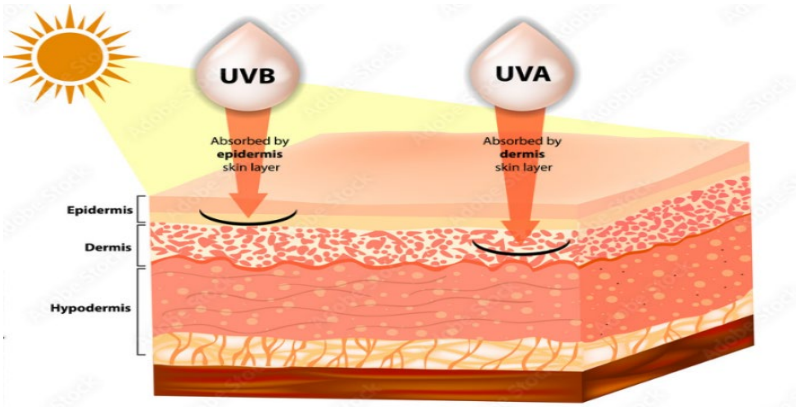


Figure 1.1: Effect of sun rays on skin

1.1 Motivation

Skin lesion segmentation and classification is a crucial task in the field of dermatology. It is used to identify and diagnose skin lesion classes, which are melanoma, nevus, and SK. Melanoma is the deadliest form of skin cancer. The latest statistics on skin cancer state that in the past decade, from 2013 to 2023, the number of diagnosed skin cancer cases increased by 27%. There is a prediction that almost 7,990 people will die of melanoma, of those 5,420 will be males and 2,570 will be females in 2023 (Siegel et al., 2023). Due to this high fatality rate, dermatologists are always keen to diagnose the correct skin cancer type timely to help patients in cure. Usually, dermatologists perform a manual investigation on dermoscopic images to diagnose the skin cancer type, but this diagnosis depends on the expertise and experience of the physician in which there is a chance of human judgmental error.

The diagnosis of skin cancer presents challenges due to various factors. Here are some of the problems that can arise in the diagnosis of different types of skin cancer. The first one is similar appearances; different types of skin cancer can have overlapping characteristics in terms of their appearance. These types have similar clinical and histological features along with a lack of typical pigmentation associated with a specific skin lesion type. Due to these challenges, this process of identification and diagnosis is very difficult (Bevan & Atapour-Abarghouei, 2022; Jeyakumar et al., 2022). Clinical heterogeneity poses another challenge, skin cancer can manifest in various ways, and its presentation can vary significantly from person to person. Lesions may appear differently depending on factors such as the location on the body, the individual's skin type, and previous sun exposure. This clinical heterogeneity can complicate diagnosis, as skin cancers may not always exhibit similar features. Also, melanoma can mimic other benign types (Anand et al., 2022). Some

skin cancers may exhibit atypical features that deviate from the usual presentation. For instance, certain melanomas can lack the typical ABCDE (Asymmetry, Border irregularity, Colour variation, Diameter greater than 6 mm, Evolution) signs used for melanoma identification. These atypical features can lead to confusion or oversight during diagnosis (Das et al., 2022). Another challenge in skin lesion detection is sampling errors and diagnostic expertise. Accurate diagnosis of skin cancer often relies on performing a biopsy, where a sample of the suspicious lesion is obtained for examination under a microscope. However, errors can occur during the biopsy process, such as inadequate sampling or misinterpretation of the biopsy results, leading to inaccurate diagnoses. Skill and experience in dermatology are crucial for accurate skin cancer diagnosis. Interpreting the microscopic features of skin biopsies requires expertise, and errors can occur, particularly in less experienced or less specialized healthcare settings (Gulzar & Khan, 2022). Given these challenges, it is important to introduce a computer-based diagnosis along with consultation from a healthcare professional with expertise in dermatology. The introduction of a computer-aided system in the process of diagnosis will surely help dermatologists for better examination of skin lesion. The accuracy of the diagnosis will improve, and dermatologists can detect actual skin lesion types timely and start respective procedures to cure this disease.

The motivation of this research work is to propose algorithms that can correctly segment the skin lesion from dermoscopic images after hair and other artifacts removal, then extract features of the skin lesion and identify the correct type of skin lesion through the classification process. Most of the existing research performs binary classification to identify whether a skin lesion is malignant (melanoma) or benign (Thomsen et al., 2022; Haenssle et al., 2023). Our proposed research work will perform multiclass classification for three types of skin lesions that are melanoma, nevus, and SK.

Most of the existing research articles on skin lesion segmentation and classification are validated over a limited number of images from one of the publicly available datasets. These datasets are International Skin Imaging Collaboration (ISIC) (ISIC Archive, 2016) and images provided by the Dermatology Service of Hospital Pedro Hispano, Matosinhos, Portugal also known as PH2 (Pennisi et al., 2016). Existing work mostly tested over selected images from one of these datasets (Marchetti et al., 2021; Murugan et al., 2022). Our proposed algorithms will be tested on all the images from both datasets which are publicly available. We are going to test our proposed algorithms on all 2000 dermoscopic images from the ISIC dataset and 200 dermoscopic images from the PH2 dataset. In this research, algorithms will be proposed to increase the overall performance of skin lesion detection as compared to existing methods in the literature. Proposed algorithms will incorporate statistical, machine learning, and deep learning methods for pre-processing, segmentation, feature extraction, and classification of skin lesion.

1.2 Research Problem

The first task in skin cancer detection is to separate the skin lesion from the normal skin which is done through segmentation. The process of segmentation becomes complicated due to the presence of noise, for example, illumination variation like shadows, reflections, low contrast, and artifacts especially the hairs in dermoscopic images. Human skin is covered with hairs of different colours, orientation, and texture. In dermoscopic images this problem becomes worse as dermis hairs are extremely thin, visually fade in images, mostly overlapping, and sometimes have similar contrast as underlying skin. Although significant research is done on pre-processing, the process of hair removal still needs to be integrated with a complete system of skin lesion detection (Choudhary et al., 2021; Jeyakumar et al., 2022; W. Li et al., 2021).

Besides hairs, the irregular shape and colour characteristics of skin lesion make the segmentation step very difficult. The boundaries of lesion are very diverse in structure and shape as compared to normal skin spots. Some sample images are shown in Figure 1.2 from ISIC and PH2 datasets (ISIC Archive, 2016; Pennisi et al., 2016), where some anomalies like dark spots, reflections, hair, and boundary variations such as blurry edges can be seen in images. Most of the existing segmentation methods used either statistical or deep learning techniques. Segmentation algorithms based on statistical techniques face the challenge of variations in the shape and size of skin lesion whereas deep learning-based segmentation algorithms require a very large number of images to train the model and avoid overfitting. In the existing literature, least work is done in which both techniques are incorporated to perform the segmentation (Hasan et al., 2023). Due to all these challenges, this research problem received immense attention in the last decade and still research is ongoing to address these issues (Dai et al., 2022; El-Wahab et al., 2022).

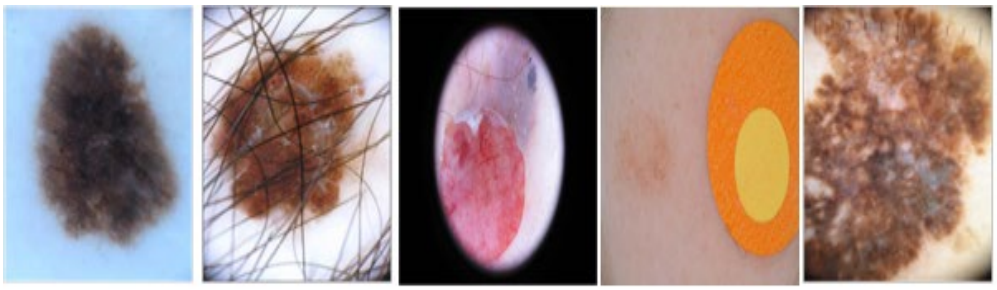


Figure 1.2: Sample images from the ISIC and PH2 datasets

For feature extraction and classification of skin lesion, similarity between skin cancer types complicates the process. Very little is known about the dermoscopic features of these skin cancer types as the biological behaviour of some benign skin lesions and subtypes, such as nevus and SK is similar to malignant melanoma (Carrera et al., 2017; Papageorgiou et al., 2018). Although existing methods have achieved better results for the classification of skin

lesion, they require a very large number of skin images interpreted by experienced doctors and this task of labelling is very time-consuming and expensive (Wang et al., 2021). Most of the methods presented in existing literature either use statistical methods which are validated over a limited number of images (Lv et al., 2022) or utilized resource-hungry deep learning-based methods for which computational cost is very high (Alzubaidi et al., 2021; Salma & Eltrass, 2022). For example, an algorithm based on hyper-connected Convolution Neural Network (CNN) is presented, in which an additional hyper branch is introduced (Bi et al., 2021). This additional hyper branch increased the number of parameters and layers of the network. Another CNN-based network is proposed, where additional filters, hyperparameters, and layers are introduced to tackle the raw pixels of the images (Shetty et al., 2022). A detailed review of different deep learning-based algorithms is presented, where comparative analysis concerning the number of filters, layers, parameters, etc among algorithms is provided (Hasan et al., 2023). Here the author also highlighted the pros and cons of statistical, machine, and deep learning methods. In this study, we will address the problem of resource-hungry deep learning-based methods. We will simplify CNN-based models, which are AlexNet (Krizhevsky et al., 2012) and VGG-16 (Simonyan & Zisserman, 2014) with respect to the number of layers, parameters, and filters while achieving better results for multi-class skin lesion classification.

1.3 Research Questions

The following are the research questions that are addressed in this research work:

i. How to remove the anomalies such as shadows, hairs, and artifacts from dermoscopic images so that skin lesion can be highlighted?

ii. How segmentation can be implemented using statistical, machine, and deep learning-based methods to extract the skin lesion from dermoscopic images?

iii. How features would be extracted from segmented skin lesion so that multi-class classification can be implemented through machine and deep learning-based methods?

1.4 Research Objectives

This research will target the above-mentioned problems and questions in an attempt to solve them by developing new algorithms and applying these algorithms to dermoscopic images. The following are the main research objectives:

i. To implement pre-processing through proposed algorithms that utilize existing methods to eliminate anomalies which are noise, artifacts, and hair from dermoscopic images to enhance the skin lesion.

ii. To implement segmentation for accurate boundary detection of skin lesion through proposed algorithms that utilize statistical methods such as Delaunay triangulation and Particle Swarm Optimization (PSO), machine learning methods such as split & merge, and a deep learning method such as Contextual Encoding Network (EncNet).

iii. To implement feature extraction and feed to the proposed classification algorithms that utilize machine learning and modified neural networks including AlexNet and VGG-16 with the reduced number of parameters and layers to classify the skin lesion as melanoma, nevus, or SK.

1.5 Methodology

The overall methodology consists of four steps for skin lesion detection. These steps are Pre-processing, Segmentation, Feature extraction, and Classification.

i. Propose algorithms for the pre-processing step which consists of filtering and contrast enhancing techniques to remove any unwanted structures, and background noise including hairs. In this step image quality will be improved to highlight the skin lesion from dermoscopic images.
ii. After pre-processing, segmentation will be implemented on dermoscopic images to obtain the foreground, which is skin lesion in this work. Existing statistical and deep learning techniques will be incorporated into the proposed algorithms for this step.
iii. In the third step, colour, texture, gradient, and geometry-based features of the segmented image will be extracted through the proposed algorithm. Also, deep features are extracted through the neural network, which is simplified in terms of network layers, so that adequate classification can be implemented.
iv. In the final step, based on the extracted features, skin lesion will be classified into melanoma, nevus, and SK. Proposed classification algorithms are the ensemble and hybrid of traditional machine learning and deep learning-based methods.

Proposed algorithms which are either based on statistical, machine learning or deep learning methods, for all four stages of skin lesion detection, will be compared with existing renowned benchmark techniques.

1.6 Contributions

Our main objective is to propose algorithms for each step of skin lesion detection that is pre-processing (including hair removal), segmentation, feature extraction and classification through statistical, machine and deep learning-based methods. Following are the contributions for each step of skin lesion detection.

i. In the first algorithm, pre-processing is implemented through Anisotropic Diffusion Filter (ADF) and other basic filtering. In the second proposed algorithm hair and artifact removal is implemented through the proposed six steps. In this algorithm, colour space transformation, weighted matrix, and black-hat morphological processing along with other techniques are utilized to remove the hairs. In the third proposed algorithm, pre-processing is implemented through local contrast enhancement, and haze reduction in the CIELAB colour space.

ii. For segmentation we proposed three algorithms. In the first algorithm, a statistical boundary condition-based model is used in which segmentation is performed through thresholding and two proposed equations. In the second algorithm, segmentation is implemented through Delaunay triangulation-based clustering with Particle Swarm Optimization (PSO). Here spatial clustering is introduced with Delaunay triangulation. In the third algorithm segmentation is implemented through the traditional split & merge method and deep learning-based Contextual Encoding Network (EncNet) with an attention mechanism. Since both techniques have origins in the region growing methods and combining them produced more favourable results as compared to existing methods.

iii. Feature extraction is implemented through statistical techniques such as the Hough transform, colour histogram, and Local Binary Patterns (LBP). Here fusion of statistical features, based on colour, pattern, and shape, is implemented through proposed values for the best division of skin lesion images. Deep learning-based features are also extracted using AlexNet and VGG-16 through a transfer learning approach. This proposed algorithm only selects the convolution features from the last convolution layer of ALexNet and VGG-16 while discarding the remaining layers,

which reduces the computational cost significantly while simplifying the overall architecture.

iv. In the first proposed algorithm, classification is implemented through a modified Convolution Neural Network (CNN) with reduced complexity. The number of layers, filters, and parameters are significantly reduced in this proposed algorithm as compared to existing architectures such as AlexNet, VGG-16, and VGG-19. A boost ensemble learning algorithm is proposed using Support Vector Machines (SVM) as initial classifiers and Artificial Neural Networks (ANN) as the final classifier in the second algorithm for classification. In the third proposed algorithm, hybrid features which were extracted through AlexNet and VGG-16, are used for classification. Here complexity in terms of layers is reduced by ignoring separate fully connected layers of AlexNet and VGG-16. This classifier is termed as a hybrid because features from both architectures are concatenated to perform the classification.

1.7 Thesis Organization

The rest of the thesis is organized as follows: in Chapter 2, the details of the literature review are provided. The materials and proposed algorithms are explained in Chapter 3 for all four steps of skin lesion detection. In Chapter 4, visual and numerical results secured through the proposed algorithms are discussed and evaluated on publicly available datasets. In Chapter 5, the conclusion and future directions are provided.

1.8 Summary

The study of the diagnosis of skin lesions is of the utmost importance. It allows early medical intervention that could cure this disease easily with reduced healthcare costs. Computer-based diagnosis can help dermatologists to detect actual skin lesion types. In this

chapter, the basics of skin lesion, their types, and the detection process are introduced. The motivation, research problem, research questions, research objectives, methodology, and thesis contributions are presented as well.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter presents the most commonly and frequently used methods for skin lesion detection. Existing methods for all four stages of skin lesion detection; pre-processing, segmentation, feature extraction, and classification are included in this chapter. These methods belong to multiple paradigms including statistical, machine learning and deep learning etc.

In the pre-processing stage, most of the existing literature is based on statistical methods in which traditional filtering techniques are used. These traditional filters commonly do not perform well because of the variable colour, size, dimensions, and illumination conditions of the skin lesion in dermoscopic images (Sau et al., 2018; Ashwini & Murugan, 2020; Bazulin, 2021). In this research, we will discuss existing methods for pre-processing from nonlinear filtering-based methods, total variation-based methods, techniques that use different colour spaces, and black & white hat transformation-based algorithms.

The existing literature uses both statistical and deep learning-based methods for segmentation, feature extraction, and classification. These methods have respective advantages and disadvantages. Although statistical methods can yield better results for a specific dataset, they may not be able to achieve similar results for another dataset of skin lesion images. This can happen because of the non-uniform nature of skin lesions in dermoscopic images (Anand et al., 2022; Hasan et al., 2020). In deep learning-based

methods, we may face the problem of limited images in publicly available datasets. Due to limited images, there would be partial learning of algorithms, and this may reduce the accuracy (Filali et al., 2018). Another issue with deep learning based-methods is that they can have multiple hidden layers, and this can increase the computational cost of overall procedure (Bi et al., 2022). In this chapter existing segmentation algorithms related to statistical, machine learning, and deep learning are presented. Our focus would be on algorithms that utilize contextual encoding, split & merge, adaptive thresholding, and Delaunay clustering-based methods.

Existing literature to extract shape, colour, and texture-based features, is presented in this chapter. A comprehensive literature review is presented to extract the features of skin lesion. This chapter also encompasses a review of the literature regarding machine and deep learning-based methods for the classification step. The most frequently utilized deep learning methods for classification, along with their corresponding pros and cons, are provided in detail. A brief introduction to human skin, dermoscopic images, and skin cancer types is also provided in this chapter.

2.1.1 Human Skin

Human skin is the largest organ of the body with an average surface area of 1.5-2.0 square meters. Our skin keeps the body's organs safe from UV and pathogens radiations, controls the body temperature and evaporation, and produces vitamin D. Human skin consists of three layers; epidermis, dermis, and hypodermis (Proksch et al., 2008; Eedy, 2020).

Epidermis is the topmost layer of the skin that is built of squamous and basal cells which are multi-layered. The flat cells of epidermis layer are termed squamous cells and

below these basal cells reside which are round-shaped. Interestingly topmost epidermis layer does not comprise any blood vessels and oxygen reaches deeper layers through a process called diffusion. Melanin pigment is the deepest part of the epidermis layer that determines a person's skin colour. (Stücker et al., 2002). Dermis is the second layer of human skin below the epidermis. This layer contains different types of cells that form blood vessels and sweat glands. This layer works like a cushion to protect the human body from stress and strain (Choudhary et al., 2021). The third layer is hypodermis which is below the dermis but usually, this does not consider as a proper layer of human skin. The main purpose of hypodermis is to connect skin with the muscles and bone. This layer consists of fat known as adipose tissue. Figure 2.1 depicts all the layers and basic cells of these layers (Yousef et al., 2022), this image is courtesy of Terese Winslow, (2008).

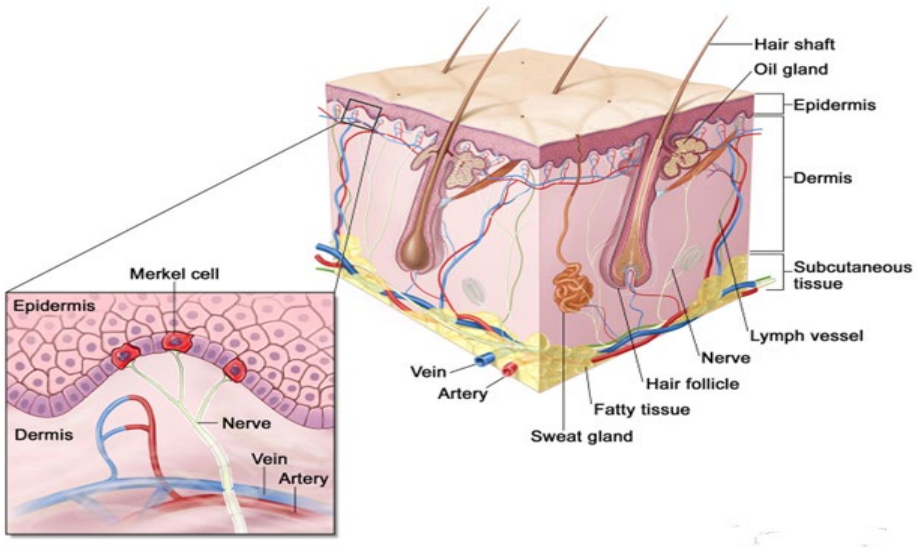


Figure 2.1: Layers structure in human skin

2.1.2 Dermoscopic Images

Dermoscopy is known to be a non-invasive technique to observe, diagnose and evaluate pigmented and non-pigmented skin lesions. With the help of this imaging technique

dermatologists can visualize better morphological structures of the skin layers or junctions where the chances of skin cancer development are higher. Skin cancer commonly develops at the epidermal and dermo junction. It is also found in the dermis papillary and other such skin portions which are not visible to the naked eye (Witkowski et al., 2017).

During the last decade, dermoscopy and related techniques have become a routine practice in dermatology which has made substantial contributions to exploring the morphology of malignant and benign skin lesions. It has been stated that with the help of dermoscopy, the diagnosis accuracy of the pigmented skin lesion improves due to the 10-27% increase in the sensitivity of dermoscopic images (Marghoob et al., 2019). These images are acquired through a simplified microscope which consists of a high-quality lens with magnification 20 to 30 times and a lighting system. There are three main and critical advantages of dermoscopy which can be explained in three-fold. The first one is surface magnification; magnified visualization of dermoscopic images allows the recognition of micro skin structures that are not visible to the naked eye. Second is subsurface visualization; dermoscopic images allow physicians to visualize the structures of the epidermis, dermo-epidermal junction, and superficial dermis that is located within 200µm depth of the skin surface. Third is simplicity and portability; this setup of dermoscopy is very easy and can be carried around for use in clinical settings and other applications (Papageorgiou et al., 2018). Dermoscopic assessment of skin lesions is generally based on the categories that are related to global and local features of the skin. A preliminary categorization of the skin lesion is provided by the global features by evaluating the overall morphology of the lesion before comprehensive assessment. Local features are mainly responsible for the detailed and final diagnosis. Precise extraction of local features can provide a detailed assessment of the lesion with focus on definite diagnostic clues.

In the last decade, smartphone cameras have improved health care and aided in the early diagnosis of skin blemishes. Various smartphone applications can use these images for dermatologist assessment or automated identification of skin lesions (Kassianos et al., 2015). Still, the process of skin lesion detection is very challenging because of different artifacts and similarities among skin lesion types. Even diagnosis through an expert dermatologist is not completely reliable and should be verified through other means. To address these issues, an automated computer-based diagnosis is required. Most of these computer-aided systems used dermoscopic images for evaluating and diagnosing skin cancer in patients (Massone et al., 2005). There are several existing algorithms used by these systems, the most popular are Menzis score-based technique (Nik-Zainal et al., 2012) a 7-point checklist, the CASH algorithm (Zaqout, 2019), and many other variations of these algorithms are available in the literature. However, these techniques are still in the early stages of development and should not be utilized as the only diagnosis source by dermatologists (Perrinaud et al., 2007; Salma & Eltrass, 2022).

2.1.3 Types of Skin Cancer

There are many cancer types, but skin cancer can be titled in the list of most common cancers. A prediction was made based on statistical data that 20% of the population will develop skin cancer during their lifetime. According to the American Cancer Society, over 97,920 males and females are diagnosed with skin cancer and there is a 27% increase in skin cancer cases (Siegel et al., 2022, 2023). As a disease, skin cancer is not always fatal if it is diagnosed and treated in an early stage then this can be cured very easily. To understand the early diagnosis of skin cancer, it is important to study and examine the types of skin cancers because these are very similar and can lead to incorrect diagnosis which may turn the skin cancer more fatal. Human skin is made of cells that grow, divide, and die then re-produce.

Cell division and reproduction is a continuous process in the human skin. However, the growth of abnormal cells can cause uncontrollable cell division which later can turn into cancer. These abnormal cells can penetrate the human body and spread through other organs. Skin cancer cases have been rising in recent years and dermatologists must thoroughly examine each patient for the signs of infection. Dermatologists can benefit from image processing technology by automatic screening of melanoma in affected people (Korotkov & Garcia, 2012; Xie et al., 2020).

There are multiple categories of skin cancer, as shown in Figure 2.2, image is courtesy of Adobe Stock, (2022). However, we can divide these under two general categories, malignant and benign. The malignant type of skin lesion is Melanoma. The two

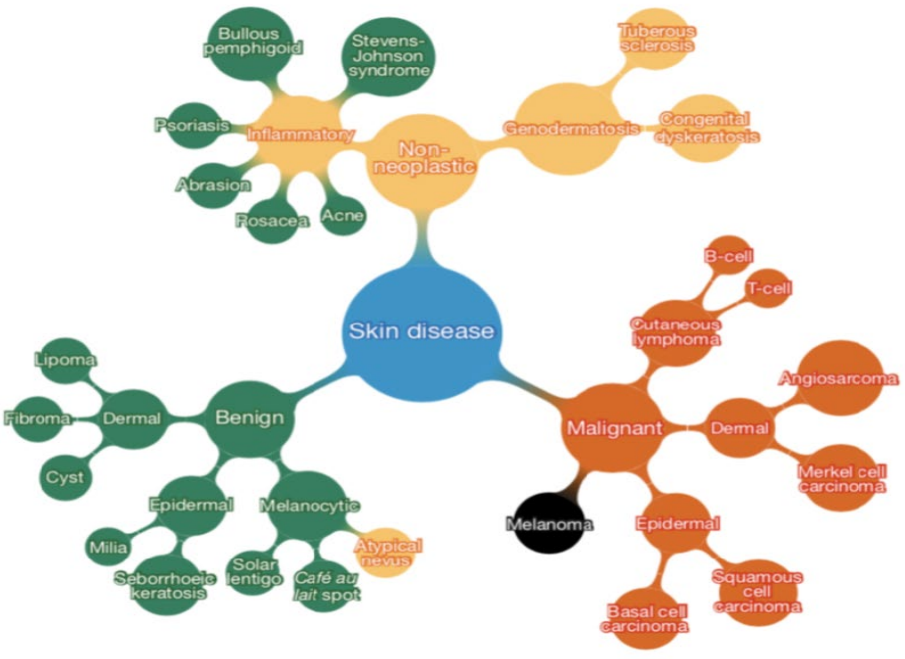


Figure 2.2: Skin Lesion Taxonomy

are Nevus and Seborrheic Keratosis (SK). Melanoma is the most dangerous and common skin cancer type with the highest fatality rate. Whereas, benign types of skin lesions do not invade nearby tissues or spread to other parts of the body usually, but in some cases, they

can turn into melanoma (Jenkins and Fisher, 2021). Different colours in Figure 2.2 indicate the degree of risk for various types of skin cancer. The benign types of skin cancer are those that are highlighted in green. Skin cancer types highlighted in yellow have a low likelihood of developing into melanoma, whereas types highlighted in orange have a higher risk. Black is used to represent melanoma because it is the deadliest type of skin cancer. As mentioned earlier that in this work we are going to classify skin lesions into melanoma, nevus, and SK because these are the most common skin cancer types (Sekniashvili, 2022).

Melanoma is one of the deadliest and fastest-growing cancer types in the world. In general skin cancer is not fatal except melanoma which is the sixth most common cancer type. People between the age of 25 to 29 are most affected by this type of skin cancer (Sabbaghi Mahmouei et al., 2019). A major cause of melanoma is UV light exposure due to sunlight and tanning beds which damage the DNA and can turn fatal. Other factor includes the genetic history of malignant melanoma in people having fair skin (Gordon, 2020). Melanoma is produced in the skin due to malignant melanocyte cells. These cells are on the outer side of the epidermis skin layer. As melanoma grows in the epidermis layer, therefore, it can be seen through the naked eye. Early diagnosis and treatment are very critical to prevent further harm. Once detected, melanoma can be cured easily through a simple procedure of excision. However, the main challenge in treatment is the high rate of false negative of malignant melanoma. Melanoma is commonly found on the back side of male patients whereas in the lower limbs of female patients, it can also be found on other body parts such as mouth, ears, head, and near eyes (Gouda et al., 2022). A sample melanoma image from ISIC (ISIC Archive, 2016) dataset is presented in Figure 2.3.

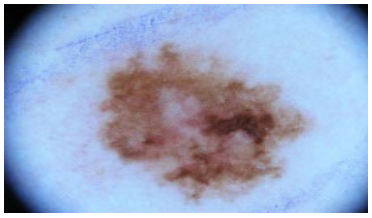


Figure 2.3: Melanoma sample image

The second type is nevus, which is a form of mole that looks different from a regular mole. The physical appearance of a nevus can be different from a common mole due to its size, colour, surface, and border. It is usually more than 5 millimeters wide (Goldstein and Tucker, 2013). This skin lesion type can have single or multiple colours, ranging from pink to dark brown colour. Usually, it has flat with a slightly scaly, smooth, or pebbly surface. Most of the time nevus has an irregular shape that may fade into the surrounding skin due to the edges. This type of skin lesion may occur anywhere on the human skin, but it is usually seen on the skin areas which are exposed to the sun, such as face, hands, and neck, etc. Nevus can also be formed in the areas which are not exposed to the sun, such as the areas below waist, breasts, and scalp (Diepgen and Mahler, 2002). It is common to have only a couple of nevi but people with fair and sensitive skin may have more than ten nevi at different skin parts and such people also have an increased number of common moles as well (Jain et al., 2022). A sample nevus image from ISIC (ISIC Archive, 2016) dataset is shown in Figure 2.4.

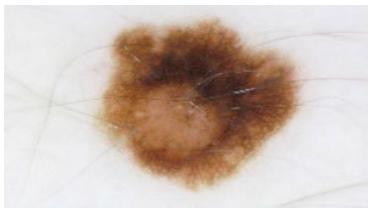


Figure 2.4: Nevus sample image

SK developed on the skin due to non-cancerous skin growth and this type of skin lesions are least dangerous. People with old age tend to get more such lesions. SK are usually light tan, black, or brown in colour. This skin lesion looks slightly raised and the growth of it is quite waxy or scaly. Normally, SK appears gradually on the face chest, neck, or back (Wollina, 2018). A sample of SK skin lesion from the ISIC (ISIC Archive, 2016) dataset is shown in Figure 2.5.

Figure 2.5: SK sample image

2.2 Medical Image Analysis

Medical imaging refers to the development of acquiring images of inner or outer body structures for scientific experiments and medical research. An image is a function that represents a measure of a perceived sight's qualities such as illumination or colour variations in terms of pixels. The main purpose of image analysis is to detect and cure illnesses. In this research, image analysis is implemented through four main steps: pre-processing, segmentation, feature extraction, and classification. Following is a brief description of these four steps.

- Pre-processing: The initial stage of image pre-processing is important as it enhances, cleans, and alters the image to improve its quality. This step makes it easier to extract information from the image and is critical to ensuring the success of further analysis or processing (Bazulin, 2021).

- Segmentation: Segmentation is essential in medical image analysis as it involves separating the lesion area from the surrounding skin region, enabling the extraction of specific features to aid in the accurate diagnosis of skin cancer (Joseph & Olugbara, 2022).

- Feature extraction: Feature extraction involves identifying and isolating significant details from an image. Skin lesions can vary in colour, shape, size, and texture. Therefore, it is important to extract distinct class dependent features which can help in deciding related skin lesion type. Sometimes in this step dimensionality reduction is also applied to keep only key information required for further examination and manipulation (Mutlag et al., 2020).

- Classification: Image classification involves selecting a label for an image from a list of predefined classes. The label should represent the main contents of the image. Image classification models take an image as input and return a prediction about which class the image belongs to. Features extracted at earlier stage, play a vital role in deciding image class (Afroz et al., 2022).

2.3 Pre-processing Techniques

Most of the available techniques for pre-processing use traditional filtering to remove noise and artifacts from images. But the most crucial artifact is hair in dermoscopic images. Dullrazor is a software-based approach proposed in 1997 that detects hair using morphological operations (Lee et al., 1997). It detects hairs from images through the grayscale morphological closing operation. The long and thin structures of hair pixels are verified and replaced through bilinear interpolation. In the end, smoothing is applied through

adaptive median filtering. This proposed method performed well in the hair removal process. Later different variations with respect to morphological operations, interpolation, and filtering are proposed in the Dullrazor-based method (Chi, 2010; Schmid-Saugeon et al., 2003). Zhou et al., (2008) proposed curve linear structure-based modeling to detect hair and artifacts then image inpainting is done through a feature-based guided method. This method preserves the morphological features of the image while implementing hair removal. A fast marching method was proposed to remove hair from dermoscopic images and improve segmentation accuracy (Hearn, 2008). In this method, the occurrence of bleeding across boundaries is minimized through automatic selection of starting point. This method is a combination of statistical, and user-influenced region expansion which support the competitive paradigm. Although reasonable results are presented in a variety of scenarios, poorer performance is observed in the presence of high intensity hair. Abbas et al., (2013) proposed an automatic hair detection and repairing algorithm, where modified mean filtering is used for hair detection. Later morphological operations and thresholding were incorporated to enhance the non-skin image parts. Inpainting of the image is done through a marching technique so that the texture pattern of skin lesion is preserved. Toossi et al., (2013), proposed a canny edge-based detection method that used morphological operation to detect hairs then coherent inpainting is used for image repairing. Kretzler, (2013), proposed a Generalized Radon Transformation (GRT) based method to detect hair through a binary mask of image followed by inpainting through pixel interpolation. Radon transform is utilized to find quadrative curve characteristics of hair. This algorithm performed well in removing hairs with curves. It also points out that better results can be achieved by replacing hair masking with interpolation algorithms. Contrast enhancement plays a key role in medical image analysis. A modified Adaptive Contrast Enhancement (ACE) is proposed that

used contrast gain to adjust frequency components of images (Chang & Wu, 1998). Here contrast gains are adjusted through a nonlinear function. Overall image contrast enhancement is carried out through Equation 2.1.

$$E(x,y) = \mu\,(x,y) + \varphi\,(x,y)\big(O\,(x,y) - \mu\,(x,y)\big) \qquad \text{Equation 2.1}$$

Where E is local contrast enhanced image, μ represents mean, O represents original image, and φ is the contrast gain which is varied through a nonlinear function as explained in (Chang and Wu, 1998). In another study, authors presented comprehensive review on enhancing contrast and image quality (Mustafa & Yazid, 2017). Here review of image enhancement techniques based on histogram equalization, wavelet transform, and homomorphic filtering is presented. Malik et al. (2022) proposed a new hybrid meta-heuristic preprocessor to optimize decision variables for contrast-enhancement transformation function. This algorithm enhances the efficiency of skin lesion segmentation and outperforms existing methods on publicly available datasets, as demonstrated using the Jaccard and dice coefficient performance metrics.

Li, L. F et al. (2020) discuss the significance of image preprocessing in skin lesion prediction using deep learning models, and it reviews the different image preprocessing techniques available for this purpose, including data cleaning and data conversion, while examining the pros and cons of each method. The use of deep learning, specifically Convolutional Neural Networks (CNN), in identifying malignant and benign skin cancer is demonstrated by (Gouda et al., 2022). This algorithm used Enhanced Super-Resolution Generative Adversarial Network (ESGRAN), image augmentation, and normalization in pre-processing step. This algorithm is tested on the ISIC2018 dataset, resulting in an accuracy of 83.2% and comparable results to other models. Even with all the classification

techniques and modifications, still results depend a lot on pre-processing of the images in which major artifact such as hair removal is required (Kretzler, 2013). There will be an immense effect on classification results due to the presence of hairs on skin lesion. In the following sub-sections, renowned methods and their variants are presented, which are used for pre-processing of the images.

2.3.1 Anisotropic Diffusion Filter (ADF) and Adaptive Thresholding

In non-linear filtering methods, all the pixels of the image are processed by a similar function which can merge different regions information. For the segmentation of skin lesion, precise detection of the cancer boundary is extremely important therefore these linear filtering techniques are not preferred. Anisotropic Diffusion Filter (ADF) solves this problem by increasing the spatial conduction in more homogeneous regions of the image and avoids major alteration of the signal along region boundaries (Bazulin, 2021). ADF has two main attributes, which include causality and piecewise smoothing. Causality means no fictitious information should be induced while passing from a coarser to a finer scale. Piecewise smoothing is a process in which intraregional smoothing should occur preferentially as compared to interregional. Both attributes can preserve the edges of skin lesion in smoothed image thus improving the process of segmentation (Perona and Malik, 1990). The paper (Sau et al., 2018), presents a novel method for enhancing skin lesion images to aid in the detection of skin cancer. This proposed method implements pre-processing by applying contrast enhancement and ADF-based filtering. The anisotropic diffusion step eliminates Gaussian noise and preserves image features, while the sigmoid function enhances the spatial domain. This proposed method has potential applications in diverse fields such as medical science and satellite imaging. Ashwini & Murugan (2020) presents a new methodology for handling Computer Tomography (CT) skin cancer images. Here image quality is improved through

filtering and contrast enhancement. In the pre-processing stage, ADF-based filtering followed by a Recursive Mean Separate Histogram Equalization (RMSHE) algorithm is introduced. Additionally, Public Contour Metric Based Segmentation (PCMBS), Mapping and Prioritized Patch Based Region Segmentation (PPBRS) algorithms are proposed for skin lesion segmentation, which can detect benign or malignant CT skin lesion from images with high accuracy.

A novel technique for automatically detecting melanoma skin cancer from dermoscopic images is presented by Rahman et al. (2022). Due to poor contrast and similarity between skin and lesions, distinguishing between the types can be challenging. The proposed approach involves using an anisotropic diffusion filter to eliminate speckle noise and the fast-bounding box method for segmentation. To extract the features from images, two methods are employed. A hybrid approach that combines histogram-oriented gradient and Local Binary Pattern (LBP) is implemented. In the second approach, robust features with convolutional neural network VGG19 are extracted. The resulting feature vectors are combined to design the classification model for discriminating between melanoma and non-melanoma skin cancer. Sunsuhi and Albin Jose (2022), addresses the difficulties in detecting brain tumors and suggests an automated approach to detect brain tumor through pre-processing, segmentation, and classification. The proposed method employs anisotropic diffusion filtering, binary-based boundary box detection, and a new segmentation technique named Adaptive Eroded Deep Convolutional neural network (AEDCNN) to recognize different regions of the brain, including meningioma, glioma, and pituitary. Additionally, the method employs Inception ResnetV2 for classification, which separates the tumor cell area and determines spatial membership without relying on multiple scales. A new method for the pre-processing and segmentation of skin lesion from digital

images is presented by Pereira et al. (2020). This method uses an adaptive threshold to ensure robust and invariant results, and statistical features such as mean and standard deviation to pre-process and segment the lesions. Image processing techniques such as average filtering and mathematical morphology are also employed to improve accuracy. The results show promising accuracy. Another study proposes the use of adaptive thresholding to separate objects of interest from the background in images with poor illumination (Gupta et al., 2017). The method uses a simpler and more accurate threshold surface and has been tested on various images with varying patterns and lighting, showing promising results compared to existing adaptive thresholding-based algorithms. The proposed method has been applied to both medical and document images.

2.3.2 Colour Space

There are multiple colour space transformations for effective pre-processing of the images and still there is no consensus on which colour space yields the best results as every colour space has its pros and cons. CIELAB colour space is preferred over RGB because it carries every single colour which can be perceived by human vision (Commission Internationale de l'Éclairage, 1978). Direct conversion from RGB to CIELAB is not possible and first, it needs to be transformed from RGB to XYZ tristimulus with the selection of a white point in the image and gamma function. Gamma function is method utilized to linearize simple RGB image components when transforming them into a linear space. The conversion from XYZ tristimulus to CIELAB is implemented through Equation 2.2 till 2.5 (Wold JH, 2001).

$$L = 116\ func(\hat{y}) - 16 \qquad \text{Equation 2.2}$$

$$a = 500\left(func(\hat{x}) - func(\hat{y})\right) \quad \text{Equation 2.3}$$

$$b = 200\,(func(\hat{y}) - func(\hat{z}) \quad \text{Equation 2.4}$$

$$\hat{X} = \frac{X}{X_n}, \hat{Y} = \frac{Y}{Y_n}, \hat{Z} = \frac{Z}{Z_n} \quad \text{Equation 2.5}$$

Where L represent black to white and the value of a and b defines the ranges from green to red and blue to yellow, respectively (Wold JH, 2001). Nisar et al. (2013), present commonly used four colour spaces, namely HSI, CMY, YCbCr, and CIELAB. These colour spaces are employed in both pre-processing and segmentation of skin lesion. This research evaluated these colour spaces to determine the most efficient one for skin lesion pre-processing and segmentation. A new algorithm is proposed that utilized Hue Saturation Value (HSV) based colour transformation (Shetty et al., 2022). Accurate image segmentation requires better pre-processing and for this step author implemented re-sizing, ordering, and data augmentation to avoid overfitting. Segmentation is implemented through HSV-based colour transformation. The proposed method in this study uses the fuzzy c-means clustering method for skin lesion segmentation. The study found that the modified fuzzy c-means clustering method was the most effective for this task, as demonstrated by the results. Sengupta et al. (2022) explore the challenges of identifying melanoma that shares similarities with other skin lesions. Proposed algorithm enhanced skin lesion by self-guided edge smoothing-based colour space filtering. For the segmentation of skin lesion, this method used global thresholding with colour space-based split and merge technique. The algorithm proposed in this article showed better performance even with images having complex texture, low illumination, and poor contrast.

2.3.3 Black-Hat Morphology and Total Variation

Top-hat and black-hat transforms are renowned morphological techniques in digital image processing. These morphological techniques are used to extract tiny components and details from images. Top-hat and black-hat transforms can be differentiated in two ways. The first difference is the input image and its opening by some structuring element, while the second difference is the input image and its closing. These transformations are utilized for a variety of image processing applications, including feature extraction, background equalization, image enhancement, and so on (Sekniashvili, 2022).

To find hair contours, black-hat morphological operation is applied with respect to direction. Morphological operations are applied to shrink or enhance some image regions through opening, closing, erosion, and dilation. Besides hair detection, other computer vision-based problems were solved with black-hat morphological techniques. For example, a proposed algorithm utilized black-hat transform on CT images to detect skull fractures (Yamada et al., 2016). Bone regions are first extracted through CT images of the head. Fracture lines are detected through 3D Laplacian filtering and linear structure are extracted through black-hat transform. Another study presents a denoising algorithm for medical images in which a combination of the Total Variation (TV) minimization and wavelet technique is utilized (Wang and Zhou, 2006). The algorithm demonstrates effective noise reduction in real-world noisy medical images while preserving object sharpness. Additionally, it allows for the implementation of an automatic stopping time criterion, making it a noteworthy solution for denoising medical images. Bhadauria and Dewal (2013) introduces a noise reduction method for CT and Magnetic Resonance (MR) images to address the challenge of interpreting medical images correctly due to noise and artifacts introduced during acquisition. The method fuses denoised images using the total variation

and curvelet methods, along with edge information extracted from the TV method. The proposed method was tested on real brain CT and MR images and showed significant improvement in noise suppression and edge preservation. Medical image fusion combines information from multiple medical images to produce a more comprehensive and accurate representation, which can improve image-guided medical diagnosis and treatment. Zhao and Lu (2017) proposed a novel method that uses a multiscale alternating sequential filter to extract useful information from noisy input images. An adaptive fractional order total variation constraint is utilized to suppress noise while avoiding the staircase effect, resulting in improved fusion of quality and noise reduction as compared to conventional methods. Black-hat transform is utilized in many image processing applications, but limited use of this technique is observed for skin lesion detection, which motivated us to utilize this technique in our research.

2.4 Segmentation of Skin Lesion

Segmentation is the second step in computer-aided systems to diagnose skin cancer. During the computerized process of skin lesion detection, it may result in different diagnoses if the lesion is not segmented accurately. Segmentation is usually the first step in some models where pre-processing is not used initially.

Several statistical methods are presented in the literature for skin lesion segmentation. Otsu's image thresholding method, which is based on maximizing image variance, is a commonly used method for skin lesion segmentation (Otsu , 1979). A cut-off value is chosen in all fixed thresholding approaches, depending on which the image is transformed from grayscale to binary scale**.** Amelard et al. (2012) presented a weighted segmentation method for segmenting pigmented skin lesions detected from macroscopic

images. By collecting regions from different parts of the image and setting a weighted threshold based on the regions of individual features. This proposed method is a variation of Otsu's thresholding approach, they utilized 260 images from the Dermquest collection. Masood and Al-Jumaily (2013) utilized a mixture of the Fuzzy C Means (FCM) algorithm, Image thresholding, and the Level Set (LS) approaches are used to segment the skin lesions. The input image is subjected to a three-class FCM, with each pixel being allocated to one of the classes based on its intensity value. The lowest and maximum of each class are utilized to compute an adaptive thresholding value, which is then used to establish the LS function in evolution. This technique has a 92.6 percent true detection rate, a 4.66 percent false positive rate, and a 7.34 percent false negative rate. This algorithm performed better when compared to the Fuzzy C Means, adaptive thresholding, and region-based active contours techniques. Another study proposes the use of Otsu's thresholding in skin lesion segmentation (Zortea et al., 2017). In this algorithm, segmentation is implemented by combining independent threshold estimates, computed from histograms of different parts of a new intensity image which is designed to separate skin lesions from the background. Here proposed technique exploits the fact that the skin lesion is approximately at the center of the image. The segmentation accuracy achieved through this proposed technique is 94.6%. Another study proposes a saliency method for the segmentation of skin lesions in dermoscopic images (Joseph & Olugbara, 2022). This method efficiently detects regions of skin lesions without preprocessing and investigates its performance in comparison to leading supervised and unsupervised segmentation methods. The proposed method utilizes colour histogram clustering and Otsu thresholding and achieves 92% accuracy. Bi et al. (2022), authors propose a hyper-fusion network (HFN) that refines the segmentation of difficult skin lesions by iteratively fusing user inputs and image features over multiple stages, and the

HFN was tested on three benchmark datasets where it outperformed state-of-the-art methods in terms of by securing segmentation accuracy of 96.41 on ISIC dataset and 96.81 on PH2 dataset. A new method for semi-automatic segmentation of dermoscopic images is presented in (Santos et al., 2022). This proposed method combines the saliency algorithm, grey-level co-occurrence matrix, texture features, and seeded fuzzy C-means clustering to achieve a superior average segmentation accuracy of 96.78% on a large dataset as compared to previous methods. If we compare the accuracy of these segmentation algorithms, Bi et al. (2022) and Santos et al. (2022) presented the best results in terms of performance measures.

A histogram-based image segmentation technique for skin lesion detection is presented (El Abbadi & Miry, 2013). The input image is converted to grayscale, then the image is de-noised using the Weiner filtering method. Following that, histogram-based segmentation is implemented by splitting the image into two groups based on the histogram and picking the best threshold value. In the post-processing stage, morphological techniques are used to remove the black areas inside the segmented image. To compare the outcomes of various segmentation approaches, Hammoude Distance (HM) and True Detection Rate (TDR) were employed. When compared to existing segmentation methods, the suggested approach had an HM of 11.37 percent and a maximum TDR of 96.32 percent. Watershed is another renowned image segmentation method, in which objects are separated by treating pixel values as a local topography (Kornilov & Safonov, 2018). It requires markers (seed points) inside each object, including the background. The markers can be chosen by an operator or by an automatic or heuristic procedure. The algorithm floods the valleys from the markers until they meet. It is useful for extracting touching or overlapping objects. Sagar and Saini (2017) presented an improved watershed transformation approach with non-negative matrix factorization. Here colour channel-based segmentation is utilized for skin

lesion segmentation**.** This algorithm tested over 175 images from different datasets available online and secured an accuracy of 94%. Moussaoui et al. (2023), presented a new technique for detecting skin cancer. This technique used a hybrid approach of fuzzy c-means algorithm, multi-Otsu thresholding algorithm, and marker-controlled watershed algorithm. The effectiveness of this proposed method was evaluated by applying it to various types of skin cancer images and measuring its performance using evaluation metrics. To enhance skin lesion detection in images affected by various factors such as poor contrast, low illumination, and complex textures, the author proposed an improved split and merge process along with a global thresholding segmentation technique (Sengupta et al., 2022). Additionally, the proposed method employs a self-guided edge smoothing-colour space technique, which is compared with other existing thresholding and colour space techniques. The comparative analysis showed better performance achieved by this proposed technique.

In another study, authors utilized a neural network method to segregate melanoma skin lesions (Nik-Zainal et al., 2012). In this proposed algorithm, Texture Distinctiveness Lesion Segmentation (TDLS) was introduced as a segmentation technique that combined a back propagation neural network and texture features of skin lesions to segment the image. The specific metrics are calculated using the image's texture distributions. The image is segmented into distinct areas using a statistical Region Merging (SRM) method based on texture similarities. The lesion in the images is then identified using Otsu thresholding. The output of TDLS is fed into a One-Layer back propagation neural network, which improves the segmented image's accuracy. Other image segmentation techniques, such as L-SRM (M. Celebi et al., 2008), Otsu-R (Cavalcanti et al., 2010), and Otsu-RGB (Cavalcanti & Scharcanski, 2011) are compared to the TDLS algorithm using neural network. The Otsu-R method (Cavalcanti et al., 2010) is similar to the Otsu thresholding method, but it only finds

the threshold for the red channel of the image. Whereas, the Otsu-RGB method (Cavalcanti & Scharcanski, 2011) finds the threshold for all three channels of the image. These methods are also used as benchmarks for comparison with our proposed segmentation algorithm.

In another study, authors propose the use of modified U-Net architecture to segment dermoscopic images accurately (Anand et al., 2022). In this method, the best performance was achieved with an Adam optimizer, a batch size of 8, and 75 epochs, while also testing the model with augmentation techniques to increase the number of images available in the PH2 dataset. A new hybrid FCN-based deep learning architecture named FCN-ResAlexNet was proposed in the study for skin lesion segmentation (Barın & Güraksın, 2022). This proposed method was evaluated on ISIC2017 and ISIC2018 datasets and found to outperform the FCN-AlexNet architecture in accuracy, dice, and Jaccard index performance metrics. Identifying melanoma accurately is vital for clinical diagnosis and treatment, but doctors' accuracy is only 60%. Current segmentation techniques based on fully connected networks and U-Net have limitations, reducing the Jaccard index of segmentation models. To address these limitations and enhance melanoma patient survival, an advanced skin lesion segmentation model based on U-Net++ is proposed (Zhao et al., 2022). This proposed model achieved a high Jaccard index of 84.73% on the ISIC2018 Task I dataset. A deep learning approach for accurate skin lesion segmentation with minimal user input is proposed by Feng et al. (2022). They used ResNet34 as a backbone network under an encoder-decoder framework. This architecture has two parts, Primary Segmentation Network (PSNet) which is used to locate the skin lesion areas, and Auxiliary Boundary Learning Network (ABLNet) which is used to detect the boundaries of skin lesion area based on the low-level features of the encoder. The proposed approach achieves better results than existing methods on established public datasets, including ISIC and PH2 datasets. A U-shaped deep network with

two lightweight attention modules is proposed for skin lesion segmentation (Zhang et al., 2023). It used Adaptive Channel-Context-Aware Pyramid Attention (ACCAPA) module and the global feature fusion (GFF) module. This network can model the characteristics of the lesion areas by dynamically learning the channel information, contextual information, and global structure information. This deep learning-based algorithm achieves better segmentation performance as compared to benchmark methods.

2.4.1 Adaptive Thresholding

In adaptive thresholding method for skin lesion segmentation, the threshold value is modified based on the local image characteristics to differentiate the lesion from the background. This process is insensitive and strong against variations in the lighting and contrast of digital dermoscopic images. An automated algorithm utilizing adaptive thresholding is proposed to segment skin lesions from digital dermoscopic images (Gupta et al., 2017). It used average filtering for noise removal. Later statistical features like mean and standard deviation along with adaptive thresholding are used to segment the foreground of images. This algorithm demonstrated high precision with an average correlation of 90% and an average overlapping score of 83%. To enhance the accuracy of diagnosis and treatment of diseases, particularly skin cancer detection, the precise segmentation of medical images is crucial. Thanh et al. (2019) propose two adaptive thresholding methods to segment skin lesions in dermoscopic images. These methods are based on colour channel normalization models. The effectiveness of the proposed methods is evaluated using dice and Jaccard similarities and tested on the ISIC dataset. Moreover, this paper compares these methods with other similar skin lesion segmentation approaches to prove their efficiency.

In another study, authors proposed the use of adaptive thresholding for removal of unwanted objects and segmentation of skin lesion (Kushimo et al., 2023). The resultant

segmented images were then converted to RGB colour space and superimposed with the original image to obtain better-segmented lesion areas with reduced noise. It is also proposed in this study that improved image quality helped the CNN to discriminate between normal and cancer-affected skin areas. Computer-Aided Diagnosis (CAD) of skin lesion is pivotal in early detection of melanoma. In most of the CAD systems, image retrieval has been developed but current systems are noninteractive and not effectively evaluated. Rahman et al. (2017) propose an interactive CAD system that performs lesion segmentation, feature extraction, classification, and retrieval using adaptive thresholding, region growing, Extreme Learning Machines (ELM), and a similarity fusion approach. The proposed methods show promising results on the ISIC dataset. To support drug development research, fungi images need to be classified based on their genre. Maulana et al. (2022) compared three segmentation methods: Otsu thresholding, adaptive thresholding, and k-means clustering, on images of fungi from three genera, Aspergillus, Trichoderma, and Cladosporium. The evaluation used Dice and Jaccard similarity, and the results showed that adaptive thresholding outperformed the other two methods, achieving the highest Jaccard and Dice scores. The study suggests that adaptive thresholding is the best method for segmenting fungi images in this context.

2.4.2 Delaunay Clustering and Particle Swarm Optimisation (PSO)

Most of the existing clustering methods used in segmentation have some known limitations mainly due to variable size, density value, noise, and shape of the lesions. These techniques use different shape and distance measures that are later compared with the threshold values to form a cluster. These approaches are also flaw-able for non-linear data. There are some methods that address these flaws including Delaunay triangulation, which is a renowned clustering algorithm (Pennisi et al., 2015). Delaunay triangulation converts high

dimensional data of polygons and polyhedral into triangles. Delaunay triangulation divides the overall space by joining the dots on the 2D plane by drawing triangles in a way that the smallest point of the inner angles of connecting triangles is maximized, and every split triangle is almost an equilateral. Consider the set of multiple points $S = \{p_1, \ldots . p_N\}$ with each point having $x_i and\ y_i$. Triangulation of these points can be defined as, $DT(q) = \{d_1, \ldots . d_H\}$. Each T represents the set of points of each triangle. H represents total triangles that are maximum 2N-2 as per point data distribution (Pennisi et al., 2015). This method is incorporated with other techniques as well for the implementation of segmentation. In (Pennisi et al., 2016b), authors proposed a method that consists of four steps. This method implemented segmentation without training. Different techniques are incorporated in this method during the first three stages, and Delaunay triangulation is used to extract the binary mask of the skin lesion at the final stage of the algorithm. It achieved 87.1% specificity and 93.5% sensitivity on 200 images from the PH2 dataset. Experimental results show that this algorithm achieves promising results for clustering and separability problems in applications of computer vision. Sasikaladevi (2022) used Delaunay triangulation with a deep learning-based technique to detect covid from very low radiation X-ray images. With the help of Delaunay-based clustering, the proposed algorithm was able to highlight suspicious regions in the images.

Particle Swarm Optimisation (PSO) is a recognized method for the optimization of segmented region. This method is implemented after the grouping of diseased pixels is done through a clustering method. It does not require any gradient calculation and works based on population (swarm) and candidate (particle). According to the formulation of PSO, the particles are displaced in the allowed range based on their best-known position and best swarm position. Initially, particles within the swarm are ranked based on their fitness values.

The overall population is divided into swarms then with the help of crossover and mutation operators, the velocity of particles is adjusted. The search process iterates until the termination criteria are met (Wang et al., 2018). PSO algorithm is used in multiple applications of computer vision. Mekhmoukh and Mokrani (2015) utilized PSO with outlier rejection for the segmentation of brain MR images. In this algorithm, FCM based clustering is implemented while using PSO to secure optimal centers of the clusters. This reduced the sensitivity in outliers and preserve the spatial information as well. PSO is also used in computer vision problems related segmentation of skin lesion. Tan et al. (2020) used PSO to fine tune the canters of K-means clusters. In this algorithm, features are extracted through CNN based model and PSO is also used for optimised feature selection. This proposed algorithm secured 88.2% accuracy on the ISIC dataset.

2.4.3 Split & Merge

The split & merge method is an image processing approach to segment out the objects, which is skin lesion in this scenario, from dermoscopic images. This process starts by dividing the complete image into multiple small segments that are called quadrants. Afterward, these quadrants are concatenated together to create the segmented image (Chaudhuri, 2010). Integral image is a basic part of the split & merge method. It is used to find the homogeneity in the image. In integral image, every pixel is the sum of pixels above and left to it. Summation of the squared region is done to generate integral image. The overall summation of pixel intensities from (i_0, j_0) to (i_1, j_1), of a squared region is calculated through Equation 2.6. The sums of powered intensities from (i_0, j_0) to (i_1, j_1) of a squared region r is calculated through Equation 2.7 (Veksler, 2003).

$$s = S(i_1, j_1) - S(i_0 - 1, j_1) - S(i_1, j_0 - 1) + S(i_0 - 1, j_0 - 1) \quad \text{Equation 2.6}$$

$$s_r = \sum_{j=j_0}^{j_1} \sum_{i=i_0}^{i_1} I^r(i,j)$$ Equation 2.7

Where i and j represent coordinates of the image. Then arithmetic mean and variance of the squared region is calculated. Later image is divided into quadrants based on homogeneity. Every quadrant is checked through variance and variation threshold which will decide if further splitting is required or not (He et al., 2016). Quadrants are managed through the QuadTree building process. After the splitting of homogenous quadrants is completed the merging process will start. Homogenous regions will be kept on combining and later class labels will be assigned. Merged regions are associated with groups and groups will be connected to classes, which in our case would be foreground and background (Garcia Ugarriza et al., 2009). Ramesh et al. (2021) explores the challenges of identifying melanoma which shares similarities with other skin lesions. Here review is done on image segmentation techniques where split & merge and other methodologies are utilized. Algorithms presented in this article showed better performance even with the images having complex texture, low illumination, and poor contrast. This split & merge method is also used in other computer vision-based problems. Ma et al. (2022) addressed the problem of object detection in high resolution images through the split & merge method. In this proposed algorithm, features of small objects are extracted through the split & merge technique while eliminating the salient information of larger objects. Features of interest are enhanced through offset error rectification as well which suppresses the background information. Proposed algorithm secured improved results on images from multiple datasets.

2.5 Feature Extraction

In feature extraction, distinct class dependent features are extracted. These features are used in the classification process in which skin lesion is assigned to a specific class

depending on its features. In some cases, reduction in feature dimensionality is also implemented. It eliminates duplicate features while making the training process simpler. The process of feature extraction is updated by combining and transforming the existing feature set. Colour, shape, texture, and pixel value are examples of characteristics that may be retrieved as features from medical images for further processing (Amelard et al., 2012).

Benyahia et al. (2022) examine how pre-trained CNN-based architectures can be utilized as feature extractors with machine learning classifiers to categorize skin lesions from two distinct datasets. It concludes that combining DenseNet201 with either Fine K-Nearest Neighbor (KNN) or Cubic SVM provides the highest accuracy for the ISIC and PH2 datasets, while the suggested method proves to be more effective than other techniques. Khan et al. (2019) present a new method for automated skin lesion classification, which uses Deep Convolution Neural Network (DCNN) feature extraction based on transfer learning and kurtosis-controlled principal component for selecting optimal features, resulting in high accuracy on three datasets and outperforming other techniques. Al-masni et al. (2018) propose an algorithm that uses Full Resolution Convolution Networks (FrCN) without any pre-processing or artifacts removal to secure resolution-based features. Analysis was performed on two publicly available datasets, ISIC and PH2. This proposed method exhibits superior performance as compared to other deep learning-based methods like Fully Convolution Networks (FCN), U-net, and SegNet. Another study used statistical features based on the colour evaluation of skin lesion (Sabbaghi Mahmouei et al., 2019). It also incorporated the QuadTree-based method for further processing. It is shown in this research that melanoma lesions contain more colours with a higher frequency of pink, black, and blue-grey as compared to benign lesions. Four classifiers were evaluated from which Support Vector Machine (SVM) provided the best results of 93%. Hosny et al. (2019) presented a

transfer learning-based method was incorporated with deep neural network for feature extraction and classification. Transfer learning is implemented on AlexNet for feature extraction while using softmax as classification layer. Skin lesion was classified between melanoma, nevus, and SK while an accuracy of 93% was achieved on the ISIC dataset. A hybrid of machine learning and statistical method is proposed by Butt et al., (2019). In this proposed algorithm, Visual Geometry Group (VGG) and AlexNet models are used to extract optimized colour features, and later classification is done through a deep CNN-based model. The author claimed segmentation and classification accuracy of 90% and 92% respectively. Anjum et al. (2020) presented a hybrid of deep learning and Newton Raphson, which is an iteration controlled based method. Segmentation is performed through region-based neural networks with feature extraction through deep learning and selection by Newton Raphson-based method. These features were given input to multilayer forward neural networks and secured an accuracy of almost 94%. Al-masni et al. (2020) presented an algorithm that detected the boundary of skin lesion through Full Resolution Convolution Network (FrCN). This method extracted critical class features which were given input to multiple CNN-based classifiers such as DenseNet, Inception, ResNet, and Inception-ResNet-v2. Divya and Ganeshbabu (2020) presented a method that utilized region growing based segmentation on pre-processed images after hair removal. Feature extraction is done through local and colour morphological features. Deep learning-based recurrent neural network model is used for the classification.

2.5.1 Hough Transform Based Feature Extraction

Hough transform is used to detect specific and complex line-based features in the images. It draws the lines and calculates intersections. The intersection points give us the shape features related to the area of interest from the image (Zhao et al., 2022). For skin

lesion detections, shape-based features of segmented lesion are secured through intersection points from the Hough transform. There are many articles in the literature which used such features for image classification in a wide range of applications. El-Wahab et al. (2022) propose a method for accurately segmenting skin images that are obstructed by hair. The proposed method involves several steps, in the first step images are converted into CIELAB colour space. Later Principal Component Analysis (PCA) is used to generate grayscale images. These grayscale images are enhanced using contrast-limited adaptive histogram equalization and an average filter. After this binary image is generated using iterative thresholding method and applying the Hough transform to generate the hair mask. In the last stage, harmonic inpainting is used to remove hair pixels. The system's performance is evaluated on images from ISIC dataset. Javadi and Soltanizadeh (2021) proposes an algorithm in which the colour properties of lesion are identified in RGB colour space. It used Hough transform to determine shape features related to respective skin lesion types as well. Later genetic algorithms are utilized for feature selection. This model tested over 50 digital images and secured 98% accuracy. Besides skin lesion detection, Hough transform is utilized in many other computer vision-based applications. Saleh and Menemencioğlu (2023) introduces an innovative approach for iris segmentation in the presence of eye diseases, which is a complex task in computer vision. The method utilizes an adaptive illumination correction and a modified circular Hough transform algorithm, along with morphological operations and pupil localization.

2.5.2 Colour Histogram Based Feature Extraction

The colour features of a skin lesion play a crucial role in determining the type of skin lesion. In general skin lesion has a very wide range of colour distribution due to which it is very difficult to decide on the specific type by analyzing the colour of lesion. Existing skin

lesion detection methods, either belong to machine or deep learning paradigms, used colour features in the decision making of skin lesion type. In (Pham et al., 2019), authors performed binary classification by utilizing colour, texture, and shape-based features. To aid in further research on skin cancer classification, this study compares six classifiers, seven feature extraction methods, and four data pre-processing steps using the largest skin cancer datasets. This study reveals that the optimal prediction is achieved through the combination of Linear Normalization, HSV for feature extraction, and Balanced Random Forest for classification. Warsi et al., (2019), introduce a novel approach to extract features from dermoscopic images using a multi-direction 3D colour-texture feature (CTF) and a back propagation multilayer neural network classifier. Skin disorders are widespread and difficult to diagnose due to the complexities of skin tone and colour variation caused by hair regions. Jain et al. (2022) propose an Optimal Probability-based Deep Neural Network (OP-DNN) to assist medical professionals in diagnosing skin diseases accurately. Pre-processing and feature extraction are followed by the OP-DNN training process, which classifies clinical images with probability values and utilizes whale optimization for weight optimization. Arora et al. (2022) presented a computer-aided system for the early detection and diagnosis of skin cancer by proposing a feature extraction technique that fuses bag-of-feature and speeded up robust features with a quadratic support vector machine for classification. Alamdar and Keyvanpour (2011) proposed a colour feature extraction method based on QuadHistogram using quadtree decomposition and homogenous blocks with different sizes. The extracted features show acceptable efficiency for image retrieval compared to the global colour histogram.

2.5.3 Shape Based Feature Extraction

Shape based features are very important and widely used in research problems related to classification. Srivastava et al. (2022) presents a texture-based feature extraction algorithm using median-based Local Ternary Pattern and local quantized ternary patterns for the classification of dermatoscopic images, achieving an average accuracy of 96% and outperforming other existing algorithms. Bansal et al. (2022) proposed a handcrafted feature extraction technique to extract features from dermoscopic images. This method used two binary variants of Harris Hawk Optimization (HHO) algorithms for feature selection. The proposed classifier, which used binary HHO for features selection, achieved superior results compared to existing metaheuristic algorithms. Local binary pattern texture features combined with colour features provide higher classification accuracy as compared to global and other local texture feature extraction techniques. Saba (2021) proposed a method utilizing the morphological skeleton and distance transform for shape-based feature extraction in skin lesion classification. This study compares machine and deep learning techniques using handcrafted and automatic features, along with clinical features and parameters. Through comparative analysis, better performance of the proposed algorithm is presented as compared to existing benchmarks.

2.6 Classification of Skin Lesion

The classification is the last and most critical step in skin lesion detection. In this process, features of segmented lesion are utilized to associate with different classes of skin lesion. Image classification is usually implemented through either traditional machine learning algorithms or deep learning algorithms. In the following subsections, we will introduce existing literature that use classifiers from these both paradigms.

2.6.1 Classification of Skin Lesion Through Machine Learning

A survey study of methods for skin lesion classification is presented by Afroz et al. (2022). This study identifies the difficulties in detecting skin lesion from dermoscopic images. It discusses pre-processing and segmentation techniques while presenting comparisons of the latest research in these two steps. It also examines different methods for the classification of skin lesions, such as SVM, KNN, and Logistic Regression based classifiers. It investigates the results applied to well-known skin cancer image datasets. This study also outlines the challenges of skin disease analysis using ISIC dataset. Another study presented the integration of a machine learning-based method, SVM with ResNet50 and VGG-16 models to improve the accuracy (Wang and Zhou, 2006). This method utilizes transfer learning, data augmentation, and various image enhancement techniques, with k-fold cross-validation to verify performance. The ResNet50 model combined with SVM shows the best results, particularly for the ISIC dataset, achieving high accuracy, Area Under the Curve (AUC), sensitivity, precision, F1 score, and computational time. Multi-Layer Perceptron (MLP) based techniques are also used for classification problem. The problem with this technique is the very large data set training required. With a smaller dataset, the performance of MLP based classifier is quite underlying. Lv et al. (2022) proposed a classifier that used multi-dimensional MLP for classification, which helps in training from scratch even for smaller datasets. The attenuation mechanism for MLP is also modified. This proposed algorithm achieved 90.9% accuracy even without data augmentation. Bechelli and Delhommelle (2022) presented a critical comparison between machine and deep learning techniques for skin lesion classification. Machine learning algorithms assessed in this work include Gaussian naïve Bayes, KNN, linear discriminant analysis, logistic regression, and decision tree classifier. Whereas deep learning methods are mainly based on CNN. These

techniques performed binary classification on two datasets, ISIC and Harvard Dataverse. Machine learning methods secure an accuracy of 72% which is improved up to 78% through transfer learning whereas deep learning methods secured an accuracy up to 88%. Bassel et al. (2022) proposed a method that used three folds stacking of existing classifiers which are SVM, KNN, and random forest. Feature extraction is implemented through VGG-16 and ResNet50. This proposed model was trained on 1000 images belonging to two classes of skin lesion. Training and testing were performed on 70% and 30% of the overall dataset. This binary classifier achieved 91% accuracy and it is claimed that these results can be improved through training with the larger dataset. The use of machine learning techniques is not limited to this application only. Kumar and Kumar (2023) presented an algorithm that performed image classification on different types of vegetable leaves. The author used SVM, MLP, decision tree, and linear regression methods to classify 25 classes of vegetable leaf images. A dataset of 7226 images is utilized for training of these machine learning models. MLP base classifier secured the highest accuracy of 90.68%.

2.6.2 Classification of Skin Lesion Through Deep Learning

Deep learning has been used to classify skin lesions in recent years, and it has shown significant improvement over standard classification approaches (Baig et al., 2020). Hekler et al. (2019) presented an algorithm that incorporates neural network and human intelligence. Classification is implemented through a CNN-based model and dermatologists. Later two independently determined diagnoses were combined into a new classifier with the help of a gradient boosting method. Moldovanu et al. (2021) presented a classifier that predict the class of skin cancer by a combination of KNN with 5-fold cross validation and deep learning-based method, Radial Basis Function Neural Network (RBFNN). Improvement in results reported here due to the inclusion of RBFFNN as a final classifier. Dorj et al. (2018)

presented a pre-trained CNN-based model AlexNet which is used for feature extraction. Later classification is implemented through modified SVM. This proposed method is tested on 3753 images collected from internet while classifying images on four types of skin cancer. Here classifier achieved 90.7% accuracy on melanoma class detection. Multiple CNN-based models are proposed in the existing literature. For example, AlexNet (Krizhevsky et al., 2012) in which five convolution layers followed by a pooling layer is used. VGG-16 (Simonyan and Zisserman, 2014) proposed by a research group from Oxford University, this model used a smaller size convolution filter with 16 weighted layers. In VGG-19 (Simonyan & Zisserman, 2017), 19 weighted convolution layers are used in this model. More details of these existing networks are provided in the next section. Mateen et al., (2018), presented an algorithm that used VGG-19 with principal component analysis and singular value decomposition for feature extraction and classification of fundus images. The results of this proposed algorithm are compared with AlexNet-based existing methods. The experiment was performed on the KAGGLE dataset with over 35 thousand images. This proposed classifier achieved 92.2% accuracy. Aldwgeri and Abubacker (2019) presented an algorithm that used multiple CNN-based pre-trained models, including VGG. Later, an ensemble approach is utilized by combining and averaging several CNN-based architectures to classify the skin lesion. Although better accuracies are achieved through CNN-based classifiers, the performance of such models rely heavily on training and overfitting. A review of CNN-based classification methods is presented by Yamashita et al. (2018) where the problems of overfitting and extensive training are discussed as well.

Another study proposes a Shallow Convolutional Neural Network (SCNN) model to classify skin lesions as benign or malignant in dermoscopic images (Montaha et al., 2022). The complexity of the lesions makes it challenging to identify and classify them, leading to

the development of computer-aided solutions. This model is trained using image pre-processing, data augmentation, and a box-blur down-scaling method to limit overfitting. A transfer learning-based deep learning system using convolutional neural networks is proposed where classification is implemented without pre-processing and segmentation using pre-trained ALexNet (Ameri, 2020). Here layer modifications are applied to obtain the final classifier which secures 91% classification accuracy. The objective of this work was to exclude complex tasks of pre-processing and segmentation. Prasad et al. (2022) utilized a custom image dataset with four classes of skin disease to develop a CNN model that was compared to several benchmark algorithms, and a federated learning approach was employed to ensure data privacy, achieving notable performance in medical imaging. Here author presented the importance of deep learning frame works especially in Internet-of-Medical-Things (IoMT). Here the author also raised concerns over data sharing and how proposed models are less efficient due to limited images. It is highlighted here that researchers should focus on such models which can be trained with limited images while providing optimal results. Salma and Eltrass (2022) presented an automated CAD system for skin lesion classification using image processing techniques, pretrained convolutional neural networks, and SVM. This proposed method achieves high classification performance and superior performance as compared to recent techniques, which can aid medical practitioners in classifying different skin lesions.

2.7 Deep Learning Methods

The section discusses frequently used deep learning methods and the effectiveness of these methods in this area of research. These methods are used in multiple applications of image processing due to their ability to learn features directly from data without requiring feature selection. Deep learning-based methods utilize artificial neural networks consisting

of input, hidden, and output layers made up of nodes to process data and produce results. The weights and biases of the hidden layers are adjusted to minimize the loss function and allow the model to converge. Deep learning models work well with large amounts of data and are applied to various tasks such as image processing with convolutional neural networks, natural language processing with recurrent neural networks, robotics, video synthesis, facial recognition, and medical diagnosis (Alzubaidi et al., 2021).

When dealing with images or computer vision tasks, mostly CNN-based networks are used. Numerous CNN-based networks have been created to assist with object detection, segmentation, classification, human face detection, vehicle identification, and so on. CNN is often utilized in the medical field to aid in disease diagnosis. CNNs are one of the most used learning processes for image identification, classification, and object detection, as well as automatic feature extraction. It is based on representative learning stages. Lower-level concepts define Higher-level concepts, and multiple higher-level concepts can be defined using the same lower-lever concepts. It learns various levels of abstraction and representation, because of common weights and adaptability to have a simple structure and fewer training parameters (Liu et al., 2015). CNN has become the go-to model for any image-related issue. They are also successfully applied to recommender systems, natural language processing, and more. The key advantage of CNN over its predecessors is that it automatically finds the relevant features without any human regulation, and CNN is also computationally efficient in learning features.

A basic architecture of CNN is shown in Figure 2.6. CNNs are supervised learning techniques that use data that has been labeled with the proper classifications to train them. CNNs learn the relationship between the input items and the class labels. They are made up

of two parts: hidden layers, where the features are extracted, and fully connected layers, where the actual classification task is implemented after processing, as shown in Figure 2.6. In a typical neural network, each layer consists of a set of neurons, with each layer's neuron linked to the neuron of the layer before it. The design of hidden layers in a CNN varies significantly. The neurons in a layer aren't all linked to the neurons in the previous layer; rather, they're only connected to a handful. Local connections are restricted, and additional pooling layers are added to aggregate local neuron outputs into a single value, resulting in translation-invariant features. Therefore, the training procedure has been simplified, and the model's complexity has decreased. VGG-16 (Simonyan & Zisserman, 2014), VGG-19 (Simonyan & Zisserman, 2014), AlexNet (Krizhevsky et al., 2012), and ResNet (He et al., 2016) are commonly used CNN-based architectures.

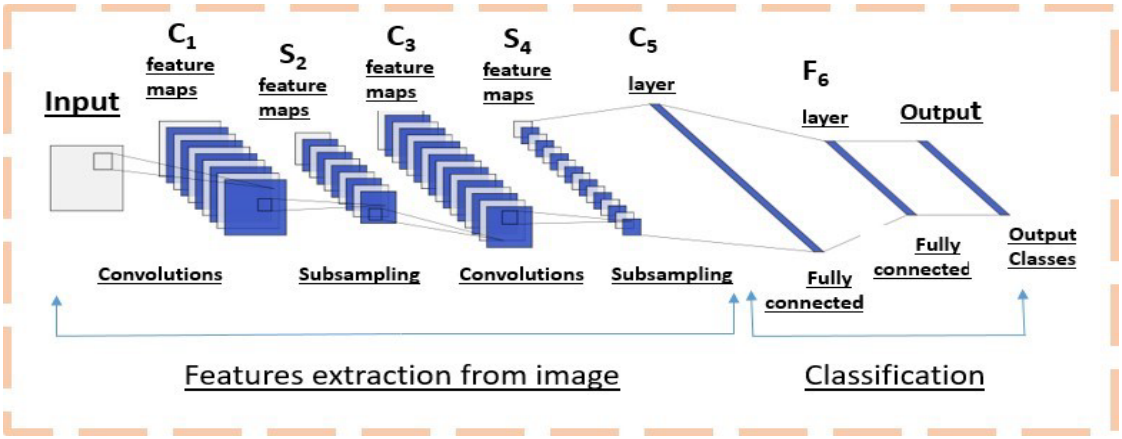


Figure 2.6: General architecture of CNN

2.7.1 AlexNet

AlexNet is a CNN-based architecture with 60 million parameters and 650,000 neurons. AlexNet comprises eight layers, including five convolutional layers and three fully connected layers, culminating in a 1000-way Softmax distribution over 1000 class labels. Every convolutional and fully connected layer received the Relu treatment. The input

dimension of the network is 150,528 (224 x 224 x 3). The first convolutional layer filters the input image with 96 kernels of size 11 x 11 x 3 and a stride of 4 pixels. The pooled output of the first convolutional layer is fed into the second convolutional layer, which filters it with 256 kernels of size 5 x 5 x 48. There are no pooling or normalizing layers between the third, fourth, and fifth convolutional layers. Each of the completely connected layers contains 4096 neurons. The 11 x 11 filter size consumes more RAM and batch normalization is not employed. AlexNet won the ILSVRC-2012 competition to classify 1.2 million images into 1000 different classes, with a top-5 test error rate of 15.3 percent, compared to 26.2 percent for the second-best entry. The basic architecture of AlexNet is shown in Figure 2.7 (Krizhevsky et al., 2012).

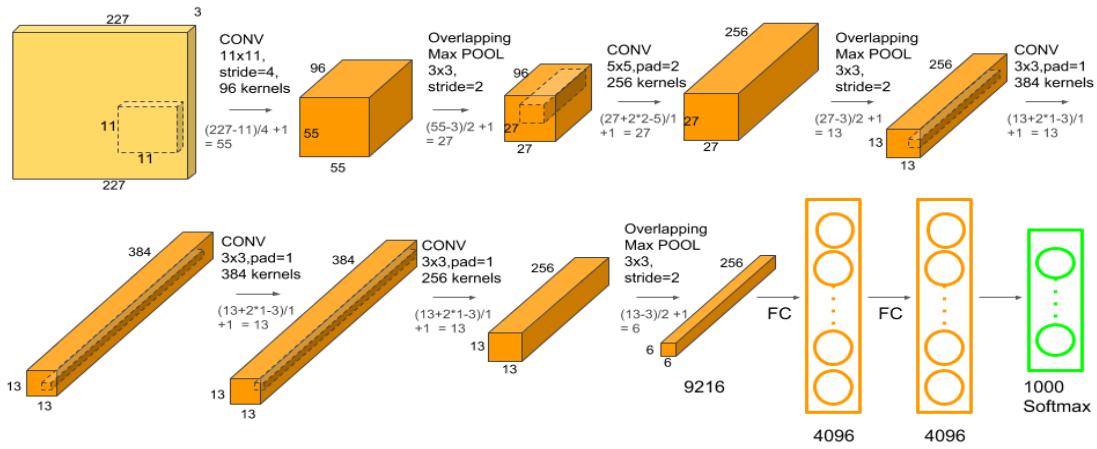


Figure 2.7: Architecture of AlexNet (Krizhevsky et al., 2012)

2.7.2 VGG-16

VGG-16 is a CNN-based architecture that consists of 16 learned layers, including thirteen convolutional layers with a 3 x 3 filter size, five pooling layers that follow some of the convolutional layers, and three fully connected layers with a 1000-way Softmax that

produces a distribution over 1000 class labels. ReLU nonlinearity was used to apply all learned layers. Their key contribution was a careful analysis of increasing-depth networks using an architecture with extremely small (3 x 3) convolution filters convolved with the input at every pixel with a stride size of 1. VGG comes in second position 39 in the ILSVRC-2014 challenge to classify 1.2 million images into 1000 different classes, with a top-five test error rate of 7.3 percent. They didn't use batch normalization to shorten the training time and keep the model stable to learn features by changing the parameter VGG16. The basic architecture of VGG16 is shown in Figure 2.8 (Simonyan & Zisserman, 2014).

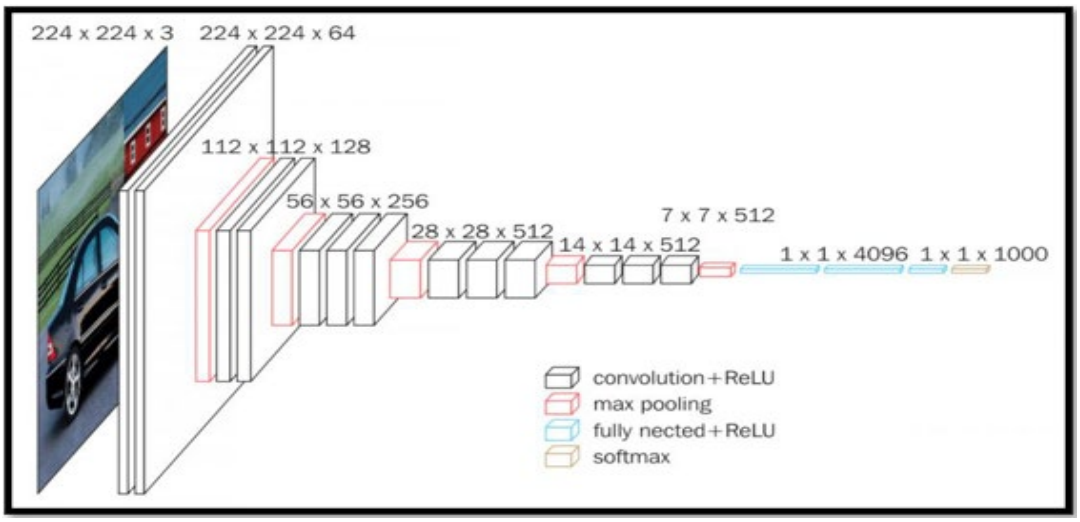


Figure 2.8: Architecture of VGG-16 (Simonyan & Zisserman, 2014)

2.7.3 VGG-19

VGG-19, a deep CNN architecture, was developed for image recognition and classification tasks by the Visual Geometry Group at the University of Oxford. As an extension of the VGG-16 model, VGG-19 includes additional layers to improve performance. This network comprises 19 weight layers, including 16 convolutional layers, 3 fully connected layers, 5 max-pooling layers, and a softmax layer for output. The model utilizes 3x3 convolutional filters and incorporates ReLU (Rectified Linear Unit) activation

functions to introduce non-linearity. The main idea behind the VGG-19 network was to increase network depth by stacking multiple convolutional layers with small filters, allowing the model to capture more intricate features from input images. This deep architecture enables VGG-19 to excel in a variety of image recognition and classification tasks, such as object detection and scene classification. Although VGG-19 achieves high accuracy, its large number of parameters makes it computationally demanding and time-intensive to train. However, VGG-19 continues to be a popular option for transfer learning, where pre-trained weights of the network are adapted for specific tasks or employed as feature extractors for other models (Simonyan & Zisserman, 2014). Figure 2.9 shows the basic architecture of VGG19.

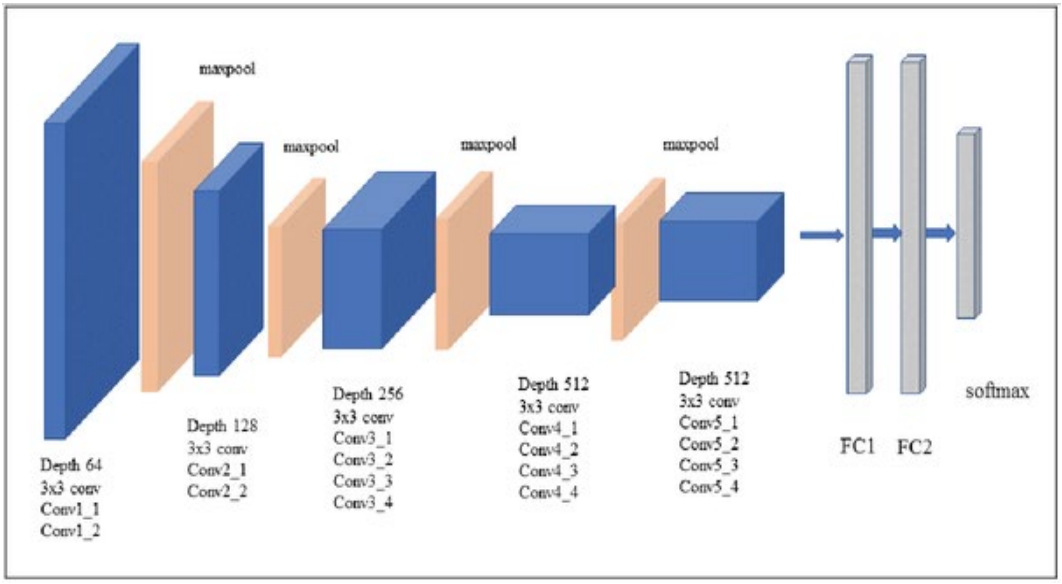


Figure 2.9: Architecture of VGG19 (Simonyan & Zisserman, 2014)

2.7.4 ResNet

ResNet, with 152 layers, is the deepest CNN model among VGG, AlexNet, and GoogleNet (Lin et al., 2019). ResNet took first place in the ILSVRC-2015 classification competition, which required it to classify 1.2 million images into 1000 distinct classes while

maintaining a top-5 test error rate of 3.57 percent. It also took first place in the ImageNet detection, ImageNet localization, COCO detection, and COCO segmentation challenges (Tao et al., 2022). ResNet-152 represents a deep residual neural network with a total of 152 layers, which encompasses both convolutional and fully connected layers. The primary breakthrough of ResNet is the incorporation of residual connections, sometimes referred to as skip or shortcut connections. These connections facilitate a more efficient learning process for the network to map inputs to outputs by enabling the gradient to flow more directly during backpropagation. As a result, the issue of vanishing gradients, commonly encountered in deep neural networks, is mitigated. The architecture is depicted in the following Figure 2.10. More details for parameters of this architecture is provided by He et al. (2016).

Figure 2.10: Architecture of ResNet152 (He et al., 2016)

2.7.5 Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN), a sort of artificial neural network that can recognize and predict data sequences such as text, genomes, calligraphy, spoken word, or numerical time series data. They have loops that allow for a constant flow of data and can

handle sequences of any length. The RNNs attempt to address the 31 requirements of comprehending data in sequences. Recurrent nets differ from feed-forward nets in that they incorporate a feedback loop, in which the result from step n-1 is fed back into the net to influence the outcome of step n, and so on (Gao et al., 2018). The defining characteristic of RNN is that the network contains a cycle that allows information to persist from one step of the sequence to the next. A standard feedforward neural network takes a fixed-sized input and produces a fixed-sized output through the application of hidden layers. However, in RNNs, the hidden layer activations are passed from one time step to the next, allowing the network to maintain an internal state that encodes information about the history of the input sequence. This makes RNNs well-suited for tasks involving sequences, such as language modeling, speech recognition, and machine translation. The following Figure 2.11 shows the architecture in terms of network flow, of RNN. More details for parameters of this architecture is provided by Gao et al. (2018).

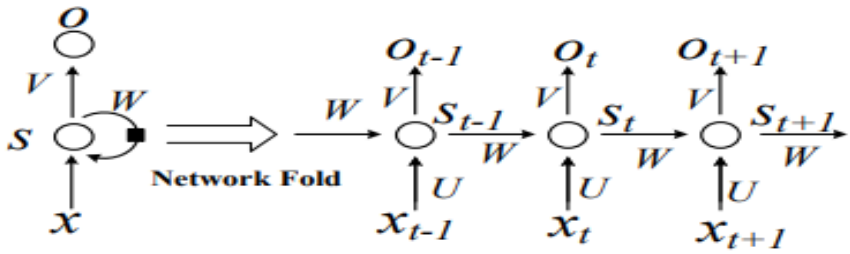


Figure 2.11: General architecture of RNN (Gao et al., 2018)

2.7.6 Long Short-Term Memory Networks

Long Short-Term Memory (LSTM) is a variant of RNN, specifically designed to address the challenge of long-term dependencies faced by basic RNNs. LSTMs leverage the use of memory cells that regulate the entry, output, and forgetting of information by the memory. This is made possible through the incorporation of four gates: the input gate, output

gate, forget gate, and gate. It handles the problem of vanishing gradients in traditional RNNs. LSTMs are commonly used for tasks involving sequences, such as language modeling, speech recognition, and machine translation. The key difference between LSTMs and traditional RNNs is the introduction of memory cells, which can retain information for an extended period. The memory cells are controlled by gates that regulate the flow of information in and out of the cell, allowing the LSTM to decide what information to remember or forget as shown in Figure 2.12. More details for parameters of this architecture is provided by Staudemeyer & Morris, (2019).

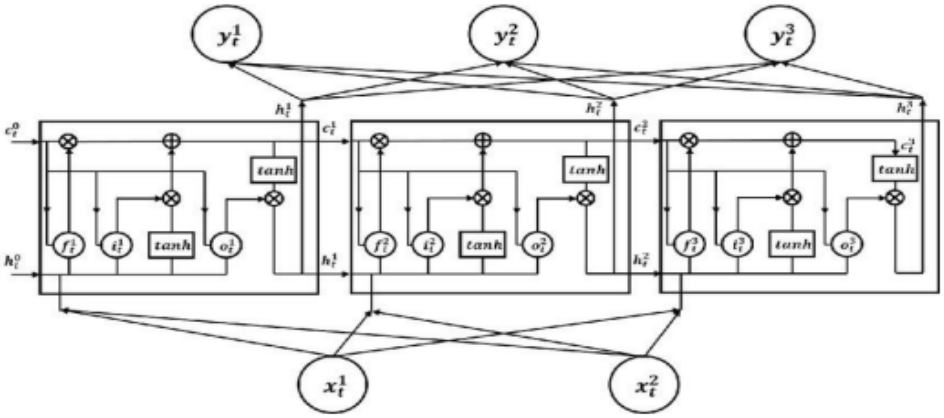


Figure 2.12: General structure of LSTM (Staudemeyer & Morris, 2019)

2.8 Performance Measures

To evaluate the performance of proposed algorithms, Dice Similarity Coefficients (DSC), Jaccard Similarity Coefficients (JSC), Sensitivity, Specificity, and Accuracy are used. These parameters will evaluate the performance of proposed algorithms by identifying the same pixels between ground truth images and segmented images. These parameters use the terms True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). These terms are mainly used to analyze the performance of a segmented image. Ground truth image will be a reference image in which the segmented part is also referred to as foreground or black pixels and the remaining pixels are the background or white pixels.

In this scenario TP is equivalent to several pixels correctly segmented as foreground, TN is the number of pixels segmented correctly as background, FP is several pixels segmented falsely as foreground and FN is equivalent to several pixels falsely segmented as background.

DSC is a validation metric that will measure the percentage of overlapped segmented parts of any two images in the spatial domain. JSC will also measure the similarity between two segmented images by calculating the size of the intersection of a similar segmented part of the image divided by the union of it. DSC and JSC are calculated by the following Equations, as described by T. Li et al., (2018).

$$D(G,I) = \frac{2(G \cap I)}{(G + I)} = \frac{2 \times TP}{((FP + TP) + (TP + FN))} \quad \text{Equation 2.8}$$

$$J(G,I) = \frac{(G \cap I)}{(G \cup I)} = \frac{TP}{FP + TP + FN} \quad \text{Equation 2.9}$$

Where G and I represent the pixels that belong to the segmented foreground in the ground truth image and proposed method image respectively, $\cap$ and $\cup$ are logical AND and OR operators respectively.

Sensitivity shows the performance of an algorithm based on the number of TP pixels identified, which is like the correct estimation of the foreground of the image. Specificity is equivalent to performance measurement based on the correct estimation of TN pixels, which is like the percentage of the correct estimation for the background of the image. Accuracy is the calculation of the overall correct estimation of both TP and TN pixels, which means it will calculate the percentage of overall correct segmentation for both foreground and background of the image. Performance measures for segmentation are sensitivity,

specificity, and accuracy, which are calculated as following equations (Kakadiaris et al., 2018).

$$Sensitivity = TPR = \frac{TP}{TP + FN} \quad \text{Equation 2.10}$$

$$Specificity = TNR = \frac{TN}{TN + FP} \quad \text{Equation 2.11}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{Equation 2.12}$$

To evaluate the classification algorithms, performance parameters such as precision, recall, F1 and FNR are used, which are calculated as following Equations (Gouda et al., 2022).

$$Precison = \frac{TP}{TP + FP} \quad \text{Equation 2.13}$$

$$Recall = \frac{TP}{TP + FN} \quad \text{Equation 2.14}$$

$$F_1 = \frac{2 * (Recall * Precision)}{Recall + Precision} \quad \text{Equation 2.15}$$

$$FNR = \frac{FN}{TP + FN} \quad \text{Equation 2.16}$$

Where TP and TN represent the number of pixels accurately classified as foreground and background respectively. Moreover, FP and FN represent the number of pixels incorrectly considered as foreground and background, respectively. Response time is another parameter that is used to evaluate machine and deep learning-based classification models (Aslam & Curry, 2021). These all are the most commonly used performance measures which

are used in literature to test proposed work (Hasan et al., 2020; Xie et al., 2020; Abayomi-Alli et al., 2021; Gouda et al., 2022;).

2.9 Summary

In this chapter, a review of recent research work in skin lesion pre-processing, segmentation, feature extraction, and classification is presented. This develops the basis for the proposed and implemented work in this thesis. All the techniques related to the methods which are used in the proposed algorithms in this thesis are presented in this chapter. This chapter also covers the deep learning techniques used in pre-processing, segmentation, feature extraction, and classification of skin lesions. The basic functionality of deep learning methods and their different variants are explained as well. More emphasis was provided on deep learning techniques which are related to implemented work in this thesis.

CHAPTER 3

METHODOLOGY

3.1 Introduction

In this chapter, we will discuss the proposed algorithms for all four steps of skin lesion detection. Proposed algorithms are used in the implementation of pre-processing, segmentation, feature extraction, and classification of skin lesions into melanoma, nevus, and SK. In the pre-processing stage, the proposed algorithms employ statistical techniques like ADF (Perona & Malik, 1990), local contrast enhancement, and haze reduction in the CIELAB colour space, which utilize mean and standard deviation to improve the image quality (Wold, 2001). Hair removal is implemented through black-hat morphological processing and total variation inpainting techniques.

The first algorithm for segmentation is based on adaptive thresholding and boundary conditions. Here various objects in the image are considered as part of the lesion based on the distance of these objects from the lesion. The second algorithm combines spatial Delaunay triangulation-based clustering with PSO for segmentation. In the second stage of this algorithm, Hues, Saturation and Value arrangement (HSV) value-based thresholding is integrated as well. The first two segmentation algorithms utilize pre-processed images, whereas the third algorithm directly performs segmentation without pre-processing. The third algorithm employs a deep learning-based model for segmentation. This approach involves two pipelines for feature extraction: split & merge and contextual encoding with attention mechanisms. The features extracted from both architectures are then combined, and the segmented region is predicted using the intersection-over-union mechanism.

The first classification algorithm utilized a CNN model with reduced complexity in terms of the number of layers, parameters, and filters. The second classifier extracts statistical features related to colour, pattern, and shape. Later, these features are fused based on implemented criteria. In the final stage of this classifier, ensemble learning is implemented using Support Vector Machines (SVM) as initial classifiers and Artificial Neural Networks (ANN) as the final classifier to learn the patterns of different skin lesion class features. The third classification algorithm extracts hybrid features using AlexNet and VGG-16 through a transfer learning approach. This hybrid classifier employs three additional fully connected layers while ignoring separate fully connected layers of AlexNet and VGG-16 networks. These proposed algorithms classify images into three classes of skin lesions: melanoma, nevus, and SK. This chapter provides explanation of all these proposed algorithms. Additionally, this chapter also outlines the two datasets used in this research.

Figure 3.1 presents the flow chart of the complete process of skin lesion detection. To summarize and highlight the contributions, Figure 3.1 presents the proposed algorithm for each step, pre-processing, segmentation, feature extraction, and classification of skin lesion. Figure 3.1 also highlights the research questions and research objectives that are achieved through the respective proposed algorithms.

3.2 Datasets

Two publicly available datasets are employed to train and evaluate the performance of the proposed algorithms. Details of these datasets are discussed in the following sections.

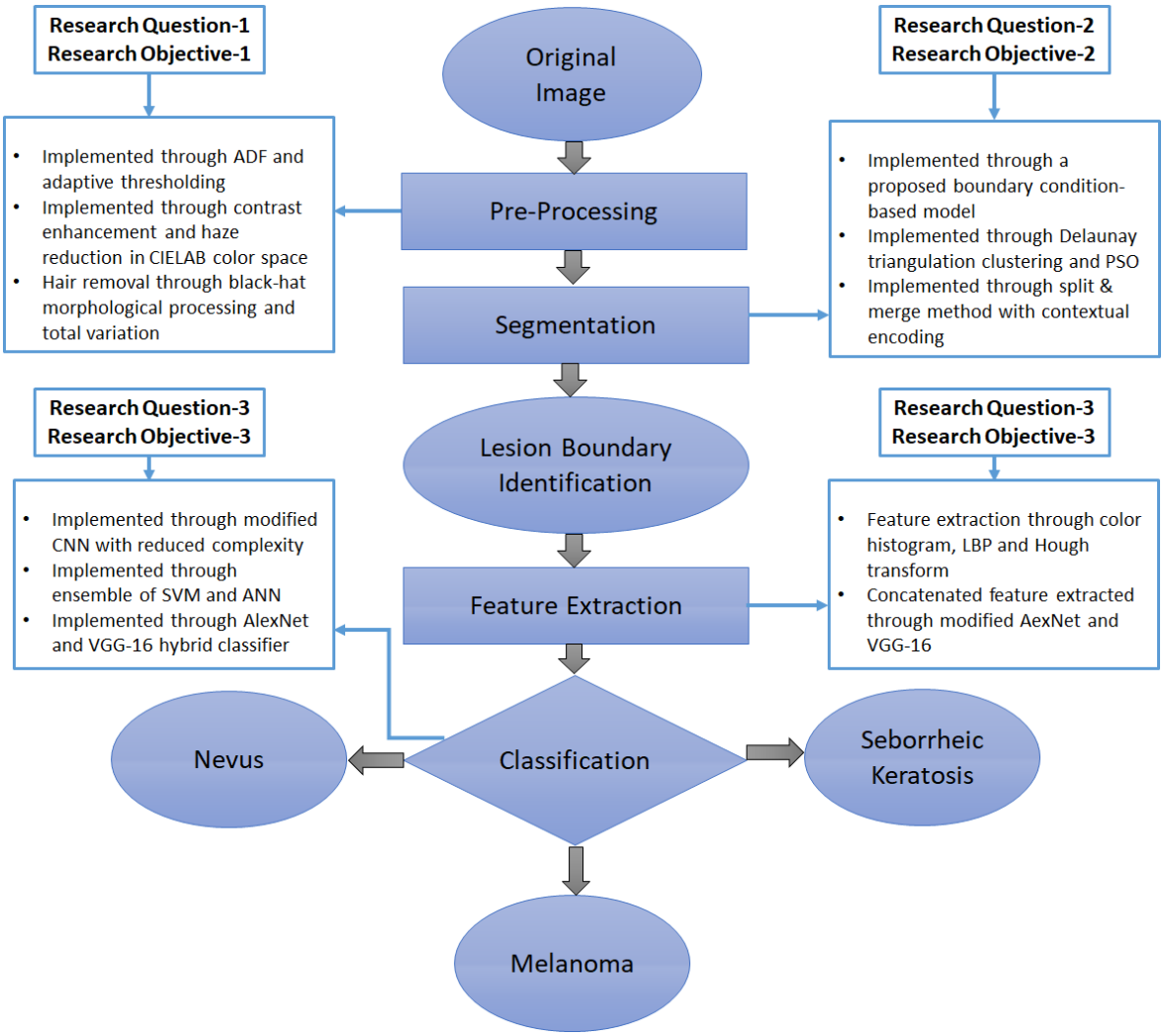


Figure 3.1: Flow chart of complete process for the skin lesion detection with implemented algorithms for each step

3.2.1 International Skin Imaging Collaboration (ISIC) Dataset

International Skin Imaging Collaboration (ISIC) is one of the biggest datasets accessible publicly. This dataset contains 2000 images with classification and segmentation details of skin lesions (Rotemberg et al., 2021). All the images from this dataset are used to validate this research work. There are many proposed algorithms in the literature that are using this dataset to validate their work as well (Khan et al., 2019; Xie et al., 2020; Rout et al., 2021).

ISIC dataset comprises skin lesion images from practically all the speculated body parts for lesions. These dermoscopic images were gathered from various medical clinics. The images' size, aspect ratio, and illumination are unique. In this research work, we have utilized this dataset in experimentations. From the 2000 visual images, 374 images are melanoma, 1372 are nevus and 254 images are classified as SK. Each image has one corresponding original JPEG format image, segmented image, and clinical information in terms of age and sex are available in a CSV record. In addition, another CSV document contains the ground truth of the class for each image; this record has three sections. The first section contains the 7-digit unique serial for each image, while the second and third segments are for the skin lesion class respectively. If the second column of CSV contains '1' at that segment, then the analyzed image is melanoma. If the third column contains '1' at that segment, then the analyzed image is SK and in case of both being 0 then the image is nevus. This dataset has been utilized in examinations of many existing works and a few strategies have been proposed on this dataset, which likewise urges us to utilize this in our research work.

3.2.2 PH2 Dataset

PH2 dataset is provided by Dermatology Service of Hospital Pedro Hispano, Matosinhos, Portugal (Mendonca et al., 2013; Pennisi et al., 2016). It comprises of 200 dermoscopic images, along with clinical explanations of 40 melanoma, 80 atypical nevi, and 80 normal nevi images. Each original image has one corresponding segmented image. All the images from this dataset are used in this research work. The dermoscopic images were deliberately procured utilizing amplification of 20 times under unaltered conditions. They are 8-piece RBG shading images with a goal of 768 × 560 pixels. The arrangement of images accessible in the PH2 database was chosen with certain imperatives, regarding their quality,

goal, and dermoscopic highlights, so that they are sufficiently appropriate to be utilized in a dermoscopic reference database. There are many proposed methods in the literature that are using this dataset for the validation of their work (Abayomi-Alli et al., 2021; Hasan et al., 2020; Hu et al., 2019). The proposed segmentation algorithms have also been tested and evaluated on this dataset.

3.3 Proposed Frameworks for Image Pre-processing

Three algorithms are proposed for pre-processing step. In the first proposed algorithm, pre-processing is implemented through ADF and basic filtering. In the second proposed algorithm, hair and artifact removal is implemented through the proposed six steps. Here colour space transformation, a weighted matrix, black-hat morphological processing, adaptive thresholding, and total variation-based inpainting is utilized to remove the hairs from dermoscopic images. In the third proposed algorithm, local contrast enhancement along with a statistical method of haze reduction is applied on the CIELAB colour space of the images. These proposed algorithms are addressing the research question that how to remove the anomalies such as shadows, hairs, and reflections from dermoscopic images so that skin lesion can be highlighted.

3.3.1 Pre-processing for Smoothing and Enhancement of Skin Lesion

In this proposed algorithm, we used basic ADF along with other filtering processes in pre-processing stage for smoothing and enhancement of skin lesion. The process starts by reading the image I(n, m) and then applying grey scale conversion to it. The image will be filtered by the available ADF (Perona & Malik, 1990) process. A 2-Dimensional Fast Fourier Transform (FFT) is applied over the image, which takes the image from the spatial domain into the frequency domain. The reason to apply FFT is that the frequency domain is

computationally faster as compared to the spatial domain therefore most filtering methods are applied in the frequency domain. In frequency domain image is decomposed into its sinusoidal components which makes it easier to target a specific frequency component.

$I(n, m)$ is the image pixel at position (n, m) and $F(x, y)$ is the frequency component of image at (x, y) , N and M are image dimensions. After FFT the zero-frequency component of the image is shifted to the center of the image spectrum, which brings the average brightness values of the image in the center. In order to focus on the lesion part of the image, la ow pass Gaussian filter is applied over the image with a cutoff frequency of 40, this will further smooth the image by ignoring pixels that are not part of the skin lesion. For this purpose, we set the variance to 40.

After convolution, the Inverse FFT will be applied to obtain the filtered image $I_{n,m}$. To ignore the non-affected part of the image, pixel manipulation is applied to the ground pixels. All the pixels having value greater than 155 will be assigned 255 which will turn the non-lesion part of the image into white. Visual results with details for this framework are provided in Chapter 4.

3.3.2 Pre-processing for Hair and Artifacts Removal

In this algorithm, a new approach is presented to eliminate hair and artifacts from dermoscopic images. The presence of hair and other artifacts in skin lesion images can hinder the complete analysis and accurate identification of the skin lesion category. Human skin is adorned with hairs varying in colour, direction, and texture, making it essential to remove hair from dermoscopic images for accurate classification of skin lesions as melanoma, nevus or SK. This hair and artifact removal technique offers a more straightforward and efficient solution compared to existing methods. The proposed algorithm employs colour space

transformation, fundamental kernel operations, morphological closing, and image inpainting to reconstruct a dermoscopic image containing hair. This proposed algorithm is implemented on the ISIC dataset which includes several images having hair and other artifacts such as dark corners, reflections, and markers etc. Once these obstructions are eliminated, diagnosis becomes easier, both visually and through classification techniques. The proposed algorithm consists of the following six steps.

i. RGB to Gray Scale Conversion
ii. Divide the image into 8 × 8 blocks
iii. Finding hair contours
iv. Intensifying hair contours
v. Detection of dark corners
vi. Image inpainting

Following is a brief explanation of these steps.

i. RGB to Gray Scale Conversion

A RGB image consists of three channels red, green, and blue. It can be visualized in Figure 3.2 (b) that the red channel of the image highlights more information about artifacts. Therefore, to convert an image from 3-channels to 1-channel, the proposed algorithm extracts red channel values from image as shown in Figure 3.2.

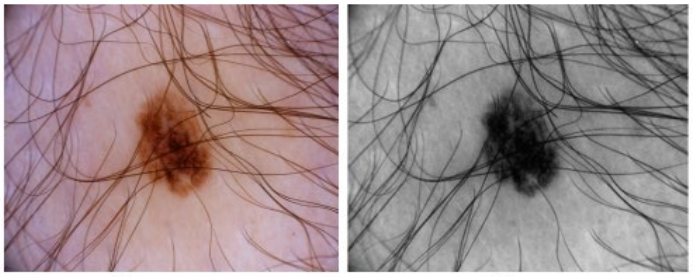


Figure 3.2: (a) RGB image and (b) shows the red channel of the image

ii. Divide Image into 8 × 8 Blocks

Grey scale image I_Grey is divided in to 16 non-overlapping sets of blocks. To generate 16 blocks, image is split up into 8 blocks in each horizontal and vertical direction. It is observed that in skin Lesion images usually the foreground region appears in the center as depicted in Figure 3.3. These 8 × 8 blocks are weighted with an 8 × 8 weight matrix with each value weighs corresponding block. Hence, the skin lesion is intensified towards white with respect to the outer regions to make the detection of hair artifacts accurate. A 3 × 3 weight matrix is described in Equation 3.1.

$$W = \begin{bmatrix} 1.6 & 1.2 & 1.6 \\ 1.2 & 3.6 & 1.2 \\ 1.6 & 1.2 & 1.6 \end{bmatrix} \quad \text{Equation 3.1}$$

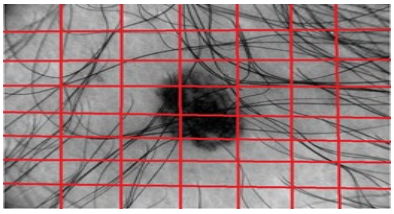


Figure 3.3: Block Division of Skin Lesion Images

iii. Finding Hair Contours

To find hair contours, black-hat morphological operation is applied with respect to direction. Morphological operations are applied to shrink or enhance some image regions through opening, closing, erosion, and dilation (Tan et al., 2020). Black-hat morphology used in this proposed algorithm tends to enhance image components for which the structuring element is larger as well as these components are darker than their surroundings. For hair detection, a structuring element of size 23 × 23 is utilized for morphological operation.

iv. Intensifying Hair Contours

Hair contour images that are obtained from the previous step contain variations of grey scale intensity. To increase the intensity of detected hair regions, binary thresholding is applied which can be described in Equation 3.2, and the resulting image that enhances the hair contours, can be seen in Figure 3.4. Figure 3.4(a) is the output image after applying black-hat transformation and Figure 3.4(b) is the output image after intensifying hairs through Equation 3.2.

$$I(x,y) = \begin{cases} 1 & if\ I(x,y) > threshold \\ 0 & otherwise \end{cases} \quad \text{Equation 3.2}$$

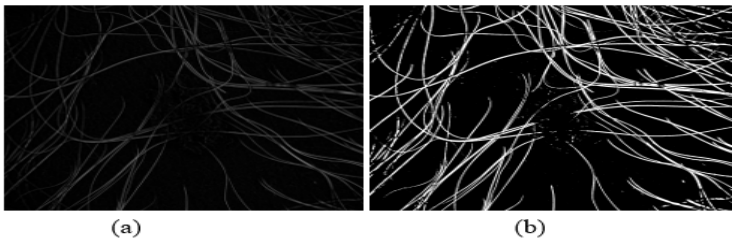


Figure 3.4: (a) Shows the image obtained after applying black-hat transformation to smaller blocks of grey scale images and (b) shows the intensified hair pixels

v. Detection of Dark Corners

When images are captured through dermatoscope, these images sometimes include black round corners. These corners may have same pixel intensities as the skin lesion in images which can affect the classification performance. In order to address this issue, Otsu thresholding (Otsu, 1979) is exploited to identify mask for black corners. Otsu thresholding calculates image histogram and iterations over several thresholds affect the classification performance. This issue can be resolved by deciding the optimal threshold. This optimal threshold reduces the inter-class i.e., foreground and background, variance. Figure 3.5(a) shows an image with dominating dark corners and Figure 3.5(b) shows the detection of round

corners after applying Otsu thresholding (Otsu, 1979). Detected corners are denoted with 1 value that is foreground.

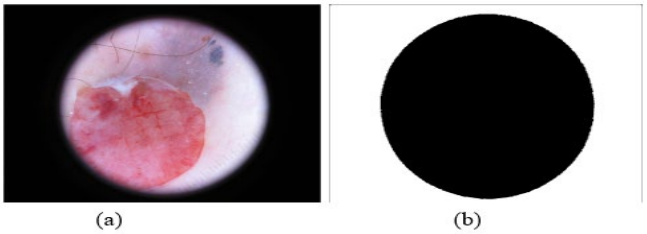


Figure 3.5: (a) Original image with black round corners and (b) shows the detected round corners in the image

vi. Image Inpainting

Image inpainting refers to a computer vision approach that focuses on restoring lost or corrupted segments of images, commonly employing deep learning techniques for this purpose. To recover hair occluded images, total variation inpainting is employed (Pennisi et al., 2016). When total variation inpainting is employed using original image and hair contours, recovered image is obtained that enhances the skin lesion. Figure 3.6 presented the visual result for sample image from ISIC dataset after implementing final step of proposed hair removal algorithm. Through this proposed algorithm we have achieved our research objective in which hairs should be eliminated while highlighting skin lesion, as shown in Figure 3.6.

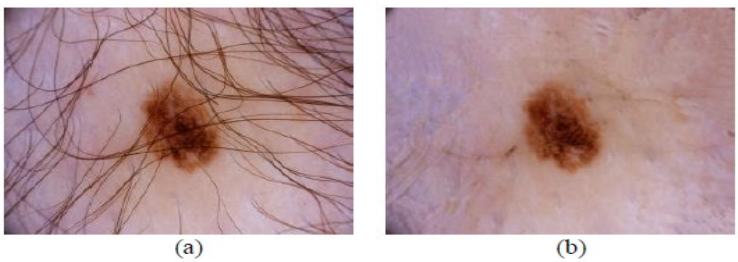


Figure 3.6: (a) presents hairy image and (b) presents image after applying inpainting step

3.3.3 Pre-processing for Enhancing Contrast and Haze Reduction

In this proposed algorithm we have combined multiple techniques for pre-processing the original noisy image. Initially, we applied local contrast enhancement on the image to improve the contrast which makes the lesion more prominent. Following this, a statistical method of haze reduction is applied on CIELAB colour space of the image that uses mean and standard deviation to enhance the overall image. Overall contrast enhancement is carried out through Equation 2.1, where local contrast enhancement is done through mean and nonlinear contrast gains (Chang & Wu, 1998). Incorporating this enhancement method improves the overall contrast of the skin lesion image by making it more prominent.

To increase the visibility of foreground, the RGB image is converted in to CIELAB colour space. CIELAB colour space carries every single colour which can be perceived by human vision (Commission Internationale de l'Éclairage, 1978). It defines the brightness in the form of L from black to white and the value of a and b defines the ranges from green to red and blue to yellow, respectively. From experimentation, it is concluded that channel representation is more suitable to highlight foreground. Direct conversion from RGB to CIELAB is not possible and first, it needs to be transformed from RGB to XYZ tristimulus with the selection of a white point in the image and gamma function. Gamma function is the method utilized to linearize simple RGB image components when transforming them into a linear space. Conversion of images in to CIELAB colour space is carried out through Equation 2.2 till 2.5 (Wold, 2001).

In final stage of pre-processing, haze reduction is applied on CIELAB image to improve the quality of the foreground and make it more separable from the background. For the improved version of the output image, dehazing is applied on the darker channel of the

CIELAB image. The progression of skin lesion enhancement through pre-processing on a sample image from ISIC dataset (Rotemberg et al., 2021) is shown in Figure 3.7. More results are provided in Chapter 4.

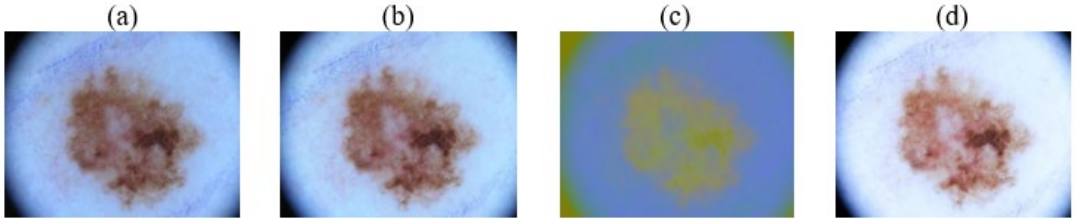


Figure 3.7: Skin lesion enhancement through pre-processing: (a) Original image, (b) Local contrast enhanced image, (c) CIELAB converted image, (d) Haze removed image

With the help of these proposed pre-processing algorithms, we have achieved our research objective 1, in which skin lesion should be highlighted while removing noise.

3.4 Proposed Frameworks for Image Segmentation

Three algorithms are proposed for image segmentation. In the first proposed algorithm, segmentation is implemented using a boundary condition-based model. In this algorithm, final step of segmentation is implemented through two proposed equations. In the second proposed algorithm, segmentation is implemented through Delaunay triangulation clustering in the spatial domain and PSO. In this algorithm, HSV value-based thresholding is integrated at the last stage. In the third proposed algorithm, segmentation is implemented through the split and merge and contextual encoding with attention mechanisms. All the methods mentioned here, are integrated into the proposed algorithms to address the research question that how segmentation can be implemented using statistical, machine, and deep learning-based methods to extract skin lesion from dermoscopic images.

3.4.1 Image Segmentation through Boundary Condition Model

In certain cases, non-uniform illumination can cause irregularities in the background of an image posing a challenge for segmentation. To address this, adaptive thresholding filtering is often employed to distinguish the skin lesion foreground from the background of the image. This filtering method assigns a foreground value to pixels with intensities below a certain threshold, while assigning a background value to the remaining pixels. There are several methods to set the threshold value, in our case best results are achieved by using a locally derived threshold based on the mean (Ahmad & Choi, 1999). To improve the results, a local threshold (Mean-C) is utilized here, where C is a constant. A window $W(x, y)$ is defined and the mean is calculated as shown in Equation 3.3.

$$\mu = \frac{1}{p \times p} \sum_{x=0,y=0}^{x=p,y=p} W(x, y) \qquad \text{Equation 3.3}$$

And local threshold $T = \mu - C$ is utilized. In our case optimal results were achieved with $C = 0.02$. Morphological operations like opening, erosion, and dilations are applied on binary image to preserve the information of the foreground part of the image and to enlarge the boundary pixels of foreground.

In this last part of the algorithm, segmentation is applied based on boundary conditions, which is implemented through two proposed equations which are Equations 3.4 and 3.5. All the foreground parts of the image are referred to as objects. The largest object, which is also referred to as the reference object, is found by calculating the number of black pixel values in all objects and selecting the maximum one, as shown in Equation 3.4.

$$M_O = Max\left(\sum_{i=1}^{p} O_i(n,m) = 0\right)$$ Equation 3.4

Where O_i is the number of objects detected in the image having only black pixels and M_o is the selected reference object which has the highest number of pixels. Distance of all the objects from the reference object is calculated through Equation 3.5.

$$D_{Oi} = \sqrt{(M_O(n) - O_i(n))^2 + (M_O(m) - O_i(m))^2}$$ Equation 3.5

If the distance D_{Oi} of object O_i from the reference object $M_o$is greater than 200 pixels then that object is ignored, and the remaining objects will be considered as the part of the lesion or foreground of the image. Later we calculate the boundaries of the selected objects and the exact segment of the affected part is obtained by mapping it on original image.

The progression of skin lesion segmentation through this proposed algorithm, on a sample image from the ISIC dataset (Rotemberg et al., 2021), is shown in Figure 3.8. Both visual and numerical results with a detailed analysis of multiple samples are presented in Chapter 4. Through this proposed algorithm we have achieved our research objective in which the boundary of skin lesion is detected more accurately while using statistical method.

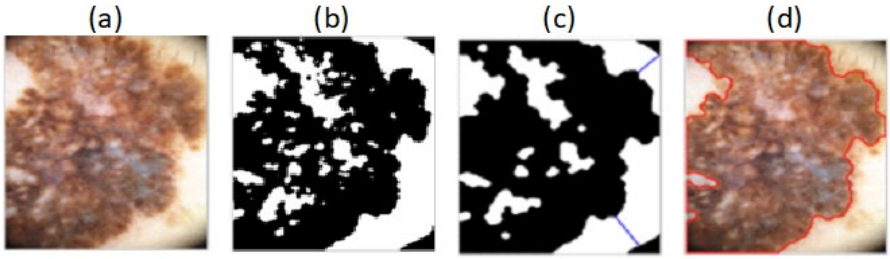


Figure 3.8: (a) Original Image, (b) Image after adaptive thresholding, (c) Image after morphological operations and boundary conditions, (f) Segmented Image

3.4.2 Image Segmentation Through Delaunay Clustering and PSO

The main purpose of the segmentation step is to split the original image into multiple segments, in our case it is known as foreground and background. The foreground is separated from the background based on colour, shape, and other features similarity. In this proposed algorithm we combine Delaunay clustering with PSO. Further, HSV value-based thresholding is incorporated in the final stage of segmentation.

3.4.2.1 Delaunay Triangulation Based Clustering with PSO

Most of the existing clustering methods used in segmentation have some known limitations mainly due to variable size, density value, noise, and shape of the lesions. These techniques use different shape and distance measures that are later compared with the threshold values to form a cluster. These approaches are also flaw-able for non-linear data. We employed an extended version of Delaunay that will perform the clustering in a spatial domain called Delaunay triangulation for spatial clustering. For the problem of skin lesion segmentation, Experimental results show that this technique achieves promising results for clustering and separability problems in applications of computer vision (Wang et al., 2016).

Delaunay triangulation converts high dimensional data of polygons and polyhedral into triangles. Delaunay triangulation divides the overall space by joining the dots on the 2D plane by drawing triangles in a way that the smallest point of the inner angles of connecting triangles is maximized, and every split triangle is almost an equilateral. Consider the set of multiple points $S = \{p_1, \ldots . p_N\}$ with each point having $x_i and\ y_i$. Triangulation of these points can be defined as, $DT(q) = \{d_1, \ldots . d_H\}$. Each T represents the set of points of each triangle. H represents total triangles that are maximum 2N-2 as per point data distribution.

After grouping the pixels of lesion using Delaunay triangulation, we exploited a novel and recognized method for optimization of the segmented region called PSO. It is a mathematical algorithm that makes little assumptions about the algorithm being optimized in our case clusters of Delaunay triangulation. Further, it does not require any gradient calculation and works based on population (swarm) and candidate (particle) (Tan et al., 2020). According to the formulation of PSO, the particles are displaced in the allowed range based on their best-known position and best swarm position. The process of movement is repeated until the best clusters are formed for segmentation. A visual representation of Delaunay triangulation with PSO is shown in Figure 3.9.

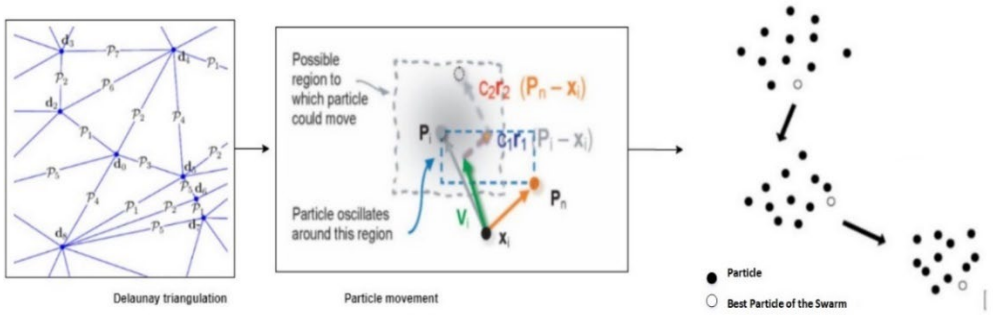


Figure 3.9: Simple representation of Delaunay triangulation with PSO

3.4.2.2 Hues, Saturation and Value (HSV) Based Thresholding

When an RGB image is converted to the HSV colour space, it produces three components Hue (H), which represents the colour in a 360-degree range; Saturation (S), which indicates the concentration of colour between 0 and 100; and Value (V), which represents the brightness of the colour on a scale of 0 to 100 (Gupta et al., 2017). By using specific threshold values (39, 88, 90) and conducting multiple experiments, we were able to identify the lesion which had a distinct dark reddish and brownish appearance. We applied binary thresholding to the extracted region and overlaid it onto the original dermoscopic

image to highlight the segmented lesion. Figure 3.10 presents the step-by-step process and visual outputs of a sample image during the segmentation process. More results are provided in Chapter 4. In Chapter 4, results achieved through this proposed algorithm are compared with renowned benchmarks to show better performance of this algorithm. Through this proposed segmentation algorithm, we have achieved the research objective in which Delaunay-based clustering is incorporated with PSO and HSV to detect the skin lesion boundary more accurately.

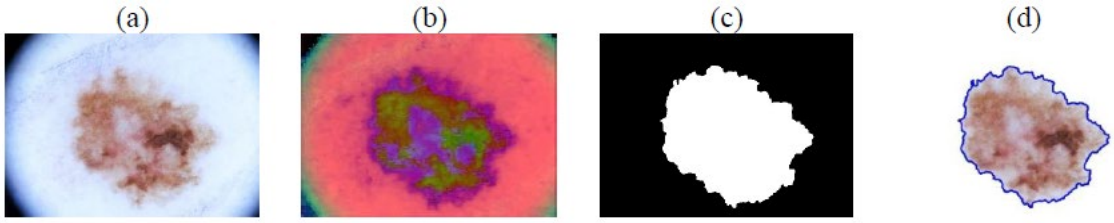


Figure 3.10: Segmentation through proposed algorithm for sample image: (a) Input image, (b). Image after HSV thresholding, (c) Binary Image, (d) Segmented image

3.4.3 Segmentation Through Split & Merge and Contextual Encoding

This proposed algorithm aims to enhance segmentation accuracy by combining traditional machine learning-based method with a deep learning method. Specifically, we integrate Split & Merge, a machine learning method (Chaudhuri, 2010) with the Context Encoding Network (EncNet), a deep learning method (Correa-Tome & Sanchez-Yanez, 2015). Since both techniques have origins in region growing methods, we hypothesize that combining them will produce more favourable results. In the first stage of the algorithm, feature and region extraction are performed using split & merge and context encoding separately. Subsequently, the segmentation results are combined using Intersection over Union (IoU) to produce the final segmented image.

3.4.3.1 Split and Merge Method

Split & merge method based image segmentation is an image processing approach to segment out the objects, which is skin lesion in this scenario, from dermoscopic image (Chaudhuri, 2010) This process starts by dividing the complete image into multiple small segments that are called quadrants. Afterward, these quadrants are concatenated together to create the segmented image. We specifically employed integral method for finding the homogeneity in the image (He et al., 2016). Following, briefly explained the steps are involved in split and merge method.

i. Integral Images

The integral image is also known as the summed area of a region; effectively used for computing sums of pixel intensities in a squared region. Let us consider an input image I with dimensions of width and height as ($w \times h$ having a starting pixel as $(0,0)$ on the top left of the overall image and the $(w-1, h-1$ on the bottom of the right side. The pixel at $(x, y$ position is denoted by $I(x, y)$ for x as column and$y$as the row number (Chaudhuri D, 2010). The integral image extracted from above calculation is employed for computing the pixel intensities summation for a particular squared region. The computation time is constant and independent of the area of the squared image involved. Let's consider multiple points on the image that are $w = (i_0, j_0)$, $x = (i_1, j_0)$, $y = (i_0, j_1)$ and $z = (i_1, j_1)$. Where i and j shows the coordinates of the image as shown in Figure 3.11. The overall summation of pixel values cbe described as in Equation 3.6 and Figure 3.11 (Veksler, 2003).

$$s = S(i_1, j_1) - S(i_0 - 1, j_1) - S(i_1, j_0 - 1) + S(i_0 - 1, j_0 - 1) \quad \text{Equation 3.6}$$

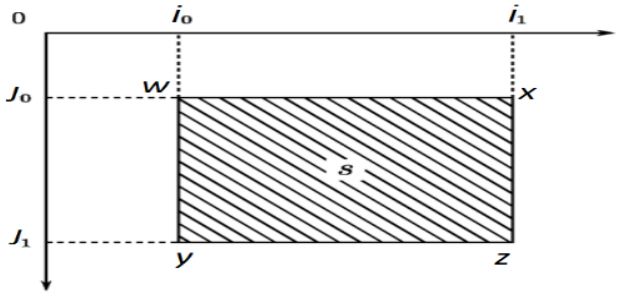


Figure 3.11: Summed of intensities

ii. Statistical Computation

Here we calculate the arithmetic mean and variance of the squared region computed in the previous section. To find the sums of powered intensities from (i_0, j_0) to (i_1, j_1) of a squared region r, Equation 2.7 is utilized. After this, mean and variance of the integral image based on squared region are calculated. The value of mean is represented as μ and the variance as σ^2 as shown in Equation 3.7 and Equation 3.8 respectively. Where N is the number of pixels in squared region r.

$$Mean = \mu = \frac{S_r}{N} \qquad \text{Equation 3.7}$$

$$variance = \sigma^2 = \frac{S_r}{N} - \left(\frac{S_r}{N}\right)^2 \qquad \text{Equation 3.8}$$

iii. Splitting Process

After this region is subdivided into further regions. This step breaks the overall image into homogeneous regions of equal width and height. The complete image is broken into four quad sub-images. If the quad image is not homogeneous in nature, then it is further split into quads. This method is executed in loop until the quads are homogeneous. Considering this situation, a particular threshold is introduced as a tolerance level. The intensity variation allowed can be determined by considering the variation threshold σ^2 for every particular image individually (He et al., 2016).

Before further splitting the image, the input image is processed in a way to enable subdivision easier. This method tries to make the image in a square shape along with the sides that are the power of two. This power method enables that the image is divided at every tree level and output images are in the power of two. If the original dimensions of the input image do not satisfy the power of two condition then the image is padded to meet the requirement $(2^n, 2^n)$. Where the parameter n is the positive whole number. A special value for an empty pixel is computed by checking the remaining pixel values. These padded pixels are different from original pixels to make these extra pixels excluded from the segmentation process.

After upsizing the image using empty pixels, the division process is started by first checking the homogeneity of the complete image. Considering the dimensions of the image as 2^n, the number of intensities would be n. Following this, the overall homogeneity of the image is determined by checking the intensities of the pixels with respect to the threshold parameter σ^2 (He et al., 2016). To determine whether an image is homogenous or not, its intensity variation is compared with the maximum allowed variance, and its variance should be below the threshold. If this condition is not met, then the image is subdivided into four portions with each side of 2^{n-1}. This step is repeated in loop for all four sub-regions until all the regions meet the homogeneity condition or n=0 level is attained, means no more subdivision is possible.

iv. Merging Process

The method of splitting and merging integrals involves dividing an image into several homogeneous regions, which are then combined into larger representations assigned to a single class label (Blaschke et al., 2004). The regions must be connected and meet certain

criteria such as automatic determination of class count, smooth gradients, and distinct regions for merging. In a tree structure, each homogeneous region is a node that can merge with others based on specific conditions such as being unclassified or having similar intensity means. If a region cannot be assigned to an existing class, a new one is created, and if a region matches multiple others, neighbouring regions are reassigned to a single class.

v. Class Selection

Initially, when the process of merging begins, no region has any class assigned to it. In this case, when a region does not have any candidate to attach with then a group structure g is created, and all the neighbouring nodes that satisfy the condition of merging are connected with that group structure as shown in Figure 3.12.

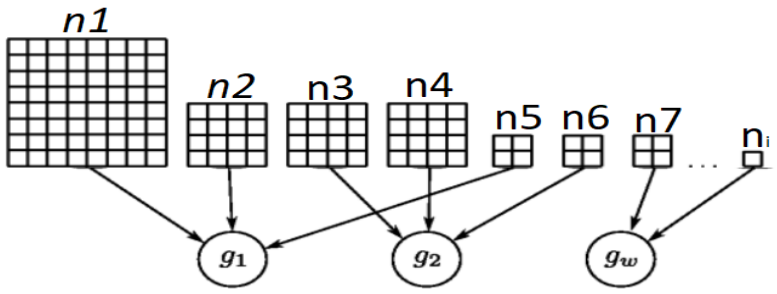


Figure 3.12: Group structure formation from nodes

Whenever a fresh node is created, at the same time a new class structure is defined, and all the respective groups are attached to it. The class maintains the list of all associated groups. The node array is also created that is sorted based on region size, which means initially all classes have the largest possible regions, and small regions are attached later during the process of merging.

The extra number of small regions may cause issues for the corresponding application. If the pixel counts in a region assigned to a class is less than a particular threshold M, then the complete class is removed from the segmented image. These declassified regions are attached to other classes (may be background) using the region growing method. This whole process of split & merge for a sample image is presented in Figure 3.13.

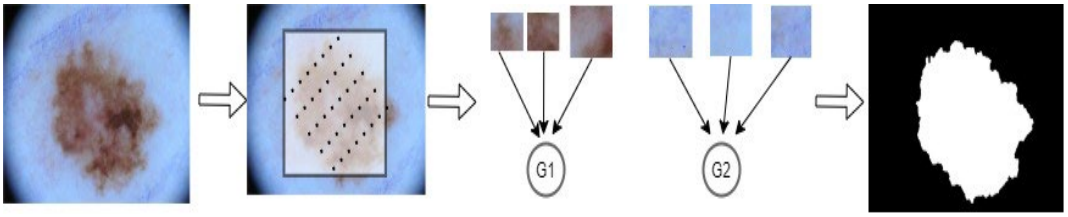


Figure 3.13: Segmentation pipeline based on split and merge method

The second type of network that we employed is context encoding network that utilizes the power of convolution neural network for finding the segmented region. Later, we combined both the results based on mask branch.

3.4.3.2 Contextual Encoding Network for Semantic Segmentation

Segmentation is classified into the two different categories; first one is the semantic segmentation and other one is the instance segmentation. In semantic segmentation pixel is classified based on the particular classes present in the image. For example, if there is a car, bus, tree etc. present in the images. It uniquely identifies and separate each class based on the pixel. Whereas, in instance segmentation, distinct objects of the same class have been separated by giving the different colours. Here we have utilized the semantic segmentation technique to segment out the input image. In recent times, a remarkable progress has been seen in the pixel labelling based on the spatial resolution using the fully convolutional neural network. In this proposed algorithm, we have taken the advantage of the latest deep learning-

based architecture of semantic segmentation. This architecture analyses the impact of the global contextual information on the semantic segmentation. It is based on the contextual encoding module which captures the semantic features of the instances present in the image along with finding the class dependent feature map (Zhang et al., 2018). Details are provided in following subsections.

i. Context Encoding

Contextual information is important in any type of semantic segmentation. To extract the information about the objects, present in the image, a pre-trained CNN has been utilized. This network is trained on the ImageNet (Socher et al., 2009) dataset. This dataset contains the variety of the images related to the different objects and categories. An encoding layer is introduced in this proposed architecture to extract the statistics of the global semantic features. The output which we get from the encoding layer, termed as the encoded semantics. Different scaling factors for different objects present in image are predicted by using the contextual information. This process helps to identify the class dependent feature map.

The defined encoding layer learns the pattern of the contextual information from the dataset and outputs the encoder containing the rich information related to the contextual information. The encoding layer picks up the features map in the shape of the $C \times H \times W$ features which is defined as $F = \{f_1, f_2, f_3 .. f_N\}$. Here C presents the number of channels, whereas H and W represent the height and width, respectively. Furthermore, F contains the total number of features. These features are learnable and trained on the codebook $C = \{c_1, c_2, c_3 .. c_k\}$ using the K number of code words present in it. It also contains the smoothing factor for detected objects as $S = \{s_1, s_2, s_3 .. s_k\}$. This encoding layer outputs the residual by combining with the weights. This residual output obtained after applying the batch

normalization and Relu activation layer (Fred Agarap, 2018). This process helps to reduce the feature dimensionality in order to achieve the better results.

ii. Attention Mechanism on Feature Map

In this mechanism, scaling factors against each class or object's feature map have been predicted. This will give us the idea or feedback to whether to emphasize or not on a particular feature map to identify the class dependability. This process has been achieved by applying the fully connected layer on the top of the encoding layer. After that sigmoid activation function (Sau et al., 2018) has been applied, the algorithm predicts the scaling factors of the feature map. Hosny et al. (2019) presented this strategy of getting feedback against feature map. Here the transfer technique is utilised to tune the feature map scale and generate the statistics.

iii. Loss Function for Semantic Encoding

In most of the semantic segmentation algorithms, the loss is calculated by using the cross-entropy loss function. It has been calculated by measuring the difference of the pixel's value from the input image and ground truth label. The drawback of using only this loss function is elimination of global semantic contextual information, which latterly affects the overall performance and accuracy of the algorithm. So, to overcome this problem, the proposed architecture has additionally utilized the semantic encoding loss function that also calculate the global semantic information. This is considerably noted that additional loss function takes very less computation cost. The additional fully connected layers above the semantic encoding layer calculates the loss value from the prediction of the object presence in the visual scene. Unlike the cross-entropy loss function of per pixel, it considers both large and small objects equally. Practically, by considering the global semantic information

it has been found that we obtain good results on the semantic segmentation of the small objects.

iv. Context Encoding Network (EncNet) Architecture

The Contextual Encoding Network (EncNet) architecture has been developed using the pre-trained ResNet (He et al., 2016). The dilated network strategy has been proposed to take the advantage of the pre-trained network. To improve the performance and regularizing the training process of the context encoding module, a separate layer has been introduced in proposed architecture, to minimize the semantic encoding loss. This layer takes the semantic encoded features as an input and predicts the object in those features. As both context encoding module and semantic encoding loss are very light weight, so we place additional semantic encoding loss layers. This will help to reduce more semantic encoding loss. This loss function is much similar but much cheaper as compared to the auxiliary loss (Zhao et al., 2017). Ground truth mask of semantic encodings loss have been extracted using the ground truth labels mask annotations without providing any additional information. The proposed contextual encoding architecture is very flexible and easily inserted into the already existed fully convolutional neural network architecture (Krizhevsky et al., 2012). This process does not require any modification or necessary adjustment along with the supervision on the original framework. If we look at the computational cost, then it increases only marginally to the original computation cost of the utilized convolution neural network architecture. Figure 3.14 shows the complete architecture used for this deep learning-based segmentation.

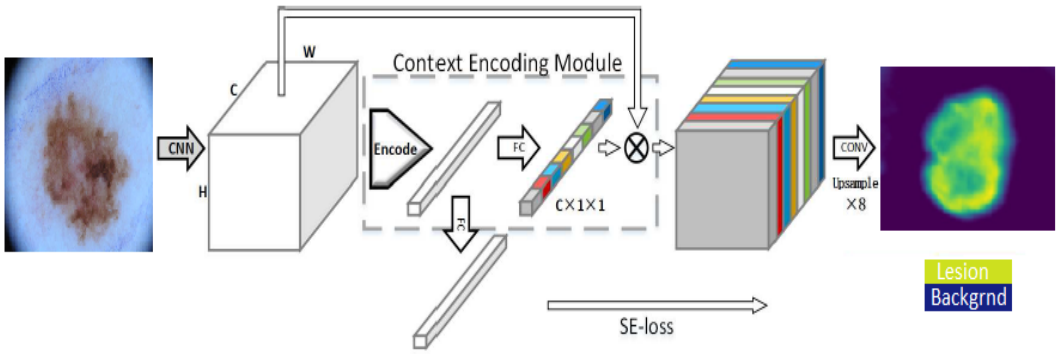


Figure 3.14: Context Encoding Network (EncNet) Architecture

3.4.3.3 Intersection Over Union (IoU)

The final step of the whole segmentation process is to combine the results from split & merge and EncNet. In this proposed algorithm, we have employed a simplest technique to combine the results from both techniques. We exploited Intersection over Union (IoU) for combining tradition results with deep learning results as shown in Figure 3.15. This approach provides the overlap between two areas that is considered as final segmented region, provided their IoU is more than 0.8, meaning very few parts of lesion is left out from both segments.

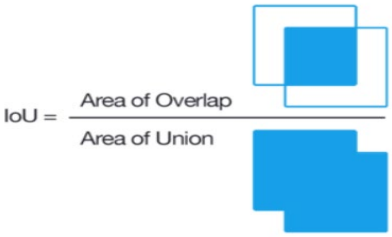


Figure 3.15: Intersection over Union (IoU) for selecting the region

In this proposed segmentation algorithm, we have achieved the research objective in which a traditional machine learning-based method, split & merge is incorporated with deep-learning based method, contextual encoding, with proposed modifications. In proposed

contextual encoding network, an encoding layer is introduced to extract the statistics of the global semantic features, additional semantic encoding loss function is used instead of cross-entropy loss function. Final segmentation results are achieved after implementation of IoU on both streams. These incorporations of existing methods are implemented to detect the skin lesion boundary more accurately as compared to existing benchmarks, which is our research objective as well.

3.5 Proposed Framework for Feature Extraction

Feature extraction is implemented through both, statistical and deep learning-based techniques. In deep learning paradigm feature extraction is part of the classification model therefore the details of feature extraction through AlexNet and VGG-16 based network is provided with respective classification model. Here we will target the research questions that how statistical or deep features would be extracted so that classification can be improved. Proposed statistical features are explained in following section.

3.5.1 Statistical Features

In this proposed framework, we used a fusion of colour, pattern and shape-based features for obtaining the best division of skin lesion images. Techniques employed in proposed algorithm are simple and less resource-hungry yet provide better results. Following features are extracted from the segmented images to perform the classification task.

3.5.1.1 Colour Histograms

Colour histograms are one of the most common and successful technique to extract the colour features from the image. Colour histograms is used to extract the colour features from the images by a different perspective (Nasir et al., 2018). It creates different bins based on the colour frequency distribution within the image. The pixels having similar values are

being collected and then stored. This process has been done with the help of two different types of histograms; the first one is the global histogram and the second one is the local histogram. A global histogram is used to define the globally describable colours or global descriptors of the particular image. A global histogram is particularly used to solve the problems of image rotation, change in translation, or angle view. Whereas in the contradiction local descriptor is used to focus and analyse the individual parts and objects of the image. In the local histogram, the main focus is on the spatial features of the image at the pixel level, and this property is not found and usually missed in the global colour histogram. We have employed the histogram intersection to separate the features of local as well as global histograms. This has been implemented using the following Equation 3.9.

$$HI(x,y) = \sum_{i=1}^{n} MinOrMax\ [(x_i, y_i)] \qquad \text{Equation 3.9}$$

Here $x\ and\ y$ represents the value of the pixels, whereas i is the counter which goes from the first pixel to the end level. As described above, we have global histogram and local histogram. *MinOrMax* function is used collectively for the local and global histogram. By using, colour histograms values which belongs to global histogram have been sorted using the max function of Equation 3.9. Whereas local histogram values sorted using the min function of Equation 3.9 (Ramella & Sanniti Di Baja, 2011). Colour histogram features are very simple, easy to compute, accurate, and smaller in size. These features are particularly useful in segmentation task like skin lesion segmentation because the particular patch of skin lesion in the image has absolute colour scheme. The reason behind the selection of colour histogram features is that, in our proposed solution, we worked on the spatial lesion segmentation. So, due to the nature of our input images is spatial, this performed better in our case. In this proposed work, we focused on the colour histograms using the HSV colour

space. The reason behind the selection of the features of HSV is that it is very resilient against the noise of the image (Pham et al., 2019). We select 90 features by creating the 4 bins of H (360 values) and 25 features for each V and S. This created 140 overall features that are then passed to the fusion layer.

3.5.1.2 Local Binary Patterns (LBP)

Local Binary Pattern (LBP) is a very simple and effective feature extraction technique. It is used to extract the texture information from the image. For skin lesion problem, LBP is used to identify the texture of lesion from segmented area of the image. For this purpose, it labels the pixels by setting the threshold for neighbourhood pixel values, and binary values are used as a result (Mahmoud et al., 2018). The local binary patterns based on the four different parameters which are as follows.

i. In Circular radius, we have defined the radius, which is later used to define the circular area of the local binary pattern. The value of the central pixel in the radius is defined and is set to 1.

ii. In neighbourhood pixels, the neighbour's pixels are used to set the circular boundary around the central pixel. In most of the cases, it has been proved through our experiments that the eight neighbour's points around the circle give the most efficient and accurate results with less computation. If the neighbour pixels would be increased then the computational power also be increased, which makes this worst computationally.

iii. The X grid represents the number of pixels in the cells which are placed in the horizontal direction. The finer LBP features depend upon the how vast is the

dimensionality, and how much we selected the cells. With the experiments it is also proved that value 8 for this parameter is most suitable.

iv. The Y grid represents the pixel present in the cells placed in the vertical direction. Same as the X grid, the Y grid also produces the good local binary features depending upon the dimensionality we select. The finer features depend upon the number of cells selected. Through experimentation it is also set to 8 for good feature selection.

In this phase, the algorithm labelled the pixels of a frame. The kernel with the size of 3×3 has been convolved on it, to find the central pixel point of this filter. Then, the pixel's value of this particular filter is converted into binary values. These binary values are then separated based on some specific threshold into the category's high and low values. In the end, these values are converted into the decimal value for better feature extraction. Each input segmented image has been divided into four 4×4 blocks, and then local edge histograms are computed for each of the blocks as shown in Figure 3.16. From each block, the algorithm selects the 16-dimensional vectors. At the end of we have 64 dimensional vectors or features to represent the edges information in the frame.

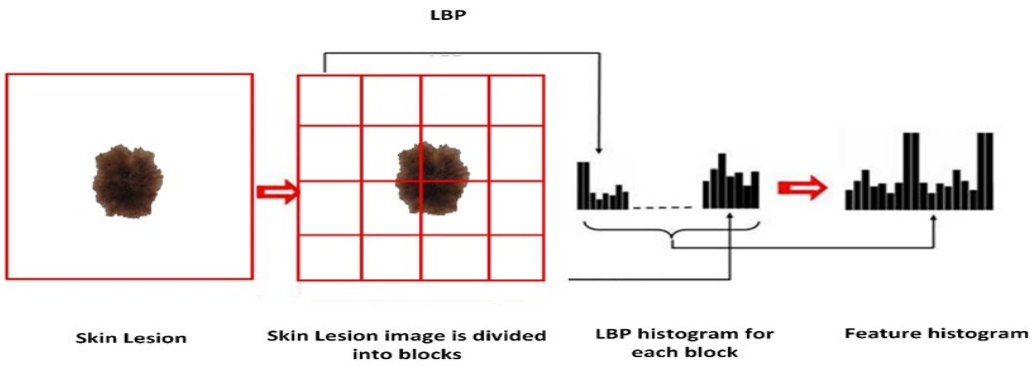


Figure 3.16: Feature extraction through LBP

3.5.1.3 The Hough Transform

The Hough transformation algorithm is used to extract the features of the shapes like circular, oval or rectangle and this is trainable. To identify the shape-based features, the Hough transform is implemented on segmented image (El-Wahab et al., 2022). In Hough transform, we must identify the edge points. We have used the canny edge detector to extract the edge features. A canny edge detector has been utilized for the edge points (Ding & Goshtasby, 2001). It is based on the four different steps which are as follows.

i. Reduction of the Noise
ii. Calculation of the gradient
iii. Non-maximum suppression
iv. Doubling the threshold

Gaussian blurring technique has been utilized to remove the noise from the frame. This process is done with the help of filters of different sizes using the sliding window technique. After that gradient descent technique identifies the edges from the blurred image, which we get from the convolution process. The edge points become prominent, and the rate of change has been measured to identify the edges more accurately on the image. But the problem is that edges of lesion are thick. So, the proposed work utilized the non-maximum suppression to tackle the thick features. The principle behind this algorithm is very simple, the algorithm scans all values of the matrix having the gradient intensity. After that, it searches the values having the maximum edge intensity in the defined direction. Then a matrix is initialized, having the same dimensions as the gradient intensity matrix. After that it checks the direction of the edge pixels. In the end it processed only those pixels which are in the same direction as the edge pixel has, whereas the rest of the pixels are discarded. So,

we secured thin edges in the end that represent the shape of the skin lesion present in the image.

Later, the Hough transform is applied on edge features, which were identified earlier and then draw a line on it. After drawing the line, we find out the intersection (El-Wahab et al., 2022). The intersection points give us the shape feature which is the skin lesion in our case, and we identify shape-based features of the segmented lesion. We have set the threshold value for the selection of the shape features based on the intersection value. We have selected the 10 points (x, y) on this criterion, making 10 features, and then passed it to the fusion layer.

3.5.1.4 Early Fusion

All of the statistical features presented in above Section 3.5.1, were combined to create a hybrid vector that represents the overall heterogenic feature space to include colour, texture and shape properties of skin lesion. From the colour histogram 140 features are combined with the 64 features of LBP Histogram and 10 features from the Hough transform, making the total features count to 214. All these features collectively represent the heterogenic properties of the skin lesion. Feature extraction of a 1D array of 214 sizes is used by the classification part which finally decides the actual type of skin lesion.

In this proposed algorithm for statistical feature extraction, we have extracted colour-based features by utilising HSV colour space, in which proposed values of H, S and V are used for best colour-based features. Texture based features are extracted through LBP, in which proposed values of multiple parameters, such as, X-grid, Y-grid, kernel, window and filter size etc. are used after experimentation to secure best texture-based features. A matrix with edge features extracted through canny edge detector then apply Hough transform to

secure shape-based features. In the end feature fusion is implemented to secure feature vector of size 214. Details of deep feature is provided with proposed hybrid classifier based on ALexNet and VGG-16.

3.6 Proposed Framework for Image Classification

As mentioned earlier that classification is implemented through three proposed algorithms. In first proposed algorithm it is implemented through CNN in which number of parameters, filters and layers are modified to reduce the complexity. In the second proposed algorithm, classification is implemented through SVM as initial classifiers and ANN as second classifier via boost ensemble learning. This algorithm is termed as ensemble classifier because it generates multiple predictions and later combine these predictions. In the third proposed algorithm, hybrid features are extracted using AlexNet and VGG-16 through transfer learning approach. This classifier is termed as hybrid because features from both architects are concatenated with additional fully connected layer to perform the classification. Here the aim is to answer the research question of how multi-class classification can be implemented through simplified machine and deep learning-based techniques after features are extracted.

3.6.1 CNN Based Classifier

In this proposed algorithm, a modified CNN based classifier is implemented. Details of basic CNN based architects such as VGG-16, VGG-19 and AlexNet is already provided in Chapter 2. Total number of layers in VGG-16 and VGG-19 are sixteen and nineteen respectively. AlexNet is identical to LeNet but deeper in terms of convolution layers and number of filters in each layer (Krizhevsky et al., 2012a). Performing experiments reveals that to classify skin lesion require simplified network to achieve better accuracy. It is

identified that AlexNet produces better results for categorization of skin lesion because of the fact that fine detailed features are present in the dermoscopic images (Yamashita et al., 2018). This fact motivated us to introduce more simplification in the model by introducing appropriate changes to secure optimal results.

To classify skin lesion, we propose a less complex network than AlexNet. VGG-16 and VGG-19 to improve the classification accuracy. As accuracy increases with decreased convolution operations, this proposed network incorporates three convolution layers followed by two fully connected layers. Proposed network is simplified to retain maximum information as input. Layers of proposed network are defined as following.

i. Input image of variable dimension is resized to 224 × 224 and is passed to first convolution layer having 48 filters of 5 × 5 which are convolved over input image to compute features of size 220 × 220 × 48. Convolution layer is followed by a pooling layer of kernel size three and stride two to generate output features of size 110 × 110 × 48.
ii. Second convolution layer takes output of previous layer and apply 64 filters of size 3×3. After convolution, maxpooling is applied to reduce in feature space to 54 × 54 × 64.
iii. Last convolution layer applies 128 filters of size 3× 3 on input features followed by pooling layer to generate feature space of 26 × 26 × 128.
iv. Output features from convolution layers are passed through two fully connected layers of 2048 and 1024 units, respectively.
v. Softmax activation function is employed at the end of the network for classification of input image.

This proposed network is modified by keeping in mind AlexNet (Krizhevsky et al., 2012a), VGG-16 and VGG-19 (Simonyan & Zisserman, 2014). Table 3.1 shows parametric comparison of proposed network with AlexNet, VGG-16 and VGG-19, with respect to number of layers, parameters and filters. Form numerical analysis presented in this table, number of convolution layers, filters, and parameters decrease gradually from VGG- 19 to the proposed network. Number of fully connected layers in all the networks are same. The proposed network has reduced the number of filters, parameters, and layers significantly as compared to existing networks, ALexNet, VGG-16 and VGG-19 as presented in Table 3.1. Proposed network reduced the network parameters from millions to thousands. It has minimum weighted layers, but similar pooling and fully connected layers as compared to AlexNet, VGG-16, and VGG-19.

Table 3.1: Parametric comparison of AlexNet, VGG-16, VGG-19, and proposed classifier

Network	Number of Parameters	Number of Convolutional Layers	Number of Pooling Layers	Number of Filters
AlexNet	4.7 million	5	3	1376
VGG-16	138 million	13	5	4224
VGG-19	143 million	16	5	5504
Proposed Network	60k	3	3	240

Figure 3.17 presents the layer configuration of the proposed network. The proposed networks implemented Softmax activation function (Hosny et al., 2019), two fully connected layers, and three pooling layers, as shown in Figure 3.17. The distribution of layers parameters and filters for the proposed network, as shown in Figure 3.17, is employed after multiple experiments so that optimal results are achieved with reduced complexity. Hence from the numerical analysis presented in Table 3.1, it is evident that the proposed classifier

reduced the overall network complexity significantly by reducing the number of parameters in thousands as compared to millions in existing networks. The same reduction can be seen with respect to number of filters and convolution layers. This simplified proposed classifier helped in achieving the research objective that classification will be implemented through a less complex neural network-based classifier.

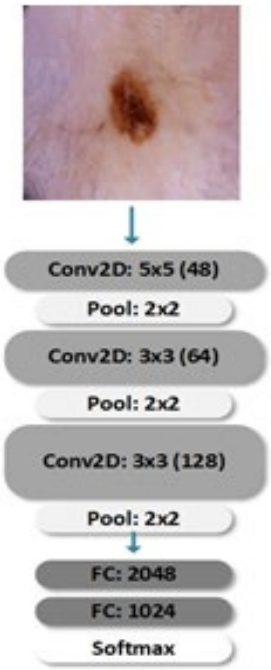


Figure 3.17: Layer configuration of the proposed classification network

3.6.2 Ensemble Classifier with SVM and ANN

A boost ensemble learning algorithm using SVM as the initial classifiers and ANN as a final classifier is employed to learn the patterns of different skin lesion class features. This is termed as an ensemble classifier because it generates multiple predictions and later combine these predictions. In this proposed boost ensemble learning, initially classifiers are used as weak learners and later utilized to improve classification accuracy. More details are provided in the following sub-sections.

3.6.2.1 Support Vector Machine (SVM)

SVM is the most widely used classifier for multi-class classification in pattern recognition problems. SVM draws a hyperplane to separate the data points. The hyperplane is a multi-dimensional space that enables the classification of an N-dimensional feature set. There can be multiple hyperplanes that potentially separate the points, but an optimal plane is that maximizes the margin in a way that the distance between support vectors of opposite classes should be maximum (Murugan et al., 2019). This proposed classifier used statistical features, which we extracted earlier as explained in section 3.5.1, as input. Sometimes features are not well classified in the input dimensions and they need to be converted into other feature dimensions by using some kernel trick. It's a kind of function that takes the input feature set in low-dimension space and converts it into higher dimensions and makes the separable features. There are different options in SVM for selecting kernel like polynomial, radial basis, linear, and hyperbolic (Tarjoman et al., 2013). We employed a Gaussian Radial basis for transforming the feature set from low dimensions to high dimensions. SVM with the radial base kernel is employed for training on half of the ISIC dataset images. The following Equation 3.10 presents the radial basis kernel.

$$K(a_1, a_2) = exp\,(\gamma\,(a_1 - a_2)^2) \qquad \text{Equation 3.10}$$

Where K is the kernel and $a_1\ and\ a_2$ are feature vectors, and γ is gamma the hyperparameter that is used for smoothing of the decision boundary (H. Wang & Hu, 2005). We have employed SVM as our first classifier which is initially trained on half of the ISIC dataset with 214 features extracted earlier and fused, explained in Section 3.5.1. All the instances are passed from the SVM model with the gamma parameter set to 0.1, this value is used after different attempts. Half of the training examples are passed to train two different SVM models. After generalization, the model predicted correct labels on most of these

images, and the remaining images are separated to be trained on the second classifier type. Further explanation is given in the section explaining ensemble learning.

3.6.2.2 Artificial Neural Network (ANN)

The artificial neural network makes computers to learn in the same way as a human does based on historical examples (Gardner & Dorling, 1998). A single layer neural network is known as a perceptron and the visual representation of a neuron is represented in Figure 3.18.

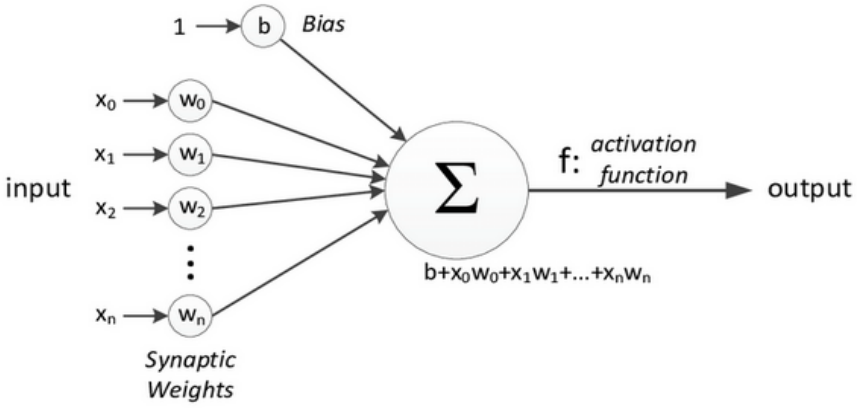


Figure 3.18: Representation of a simple neuron in a neural network

In Figure 3.18, x vector shows the input to the neuron. Each explicit input is further multiplied with the weight matrix as represented by $w_1, w_2, w_3, \ldots, w_n$. After multiplication with the weight matrix, the resultant product is summed over all the products with the sum of bias factor b. After the process of summation and activation function is applied to the resultant signal (Hekler et al., 2019). The main purpose of the activation function is to decide whether a neuron should be fired or not. The basic motive behind using the activation function is to introduce non-linearity in the signal. There are different functions available that can be used as an activation function to introduce the non-linearity. In this algorithm,

we have used Rectified Linear Unit (RELU) as the activation function (Fred Agarap, 2018). This activation function is defined by Equation 3.11.

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \qquad \text{Equation 3.11}$$

The gradient of the loss function is computed during backpropagation to update the weights of the model. The gradient is computed with respect to the weights of the model using the chain rule. The gradient is then used to update the weights of the model using an optimization algorithm such as stochastic gradient descent (SGD) or Adam optimizer, we utilized the Adam optimizer function for optimizing the weights of hidden layers (Reddi et al., 2019).

We employed an ANN based four-layered Multi-Layer Perceptron (MLP) (Gardner & Dorling, 1998) method as a classifier after SVM. Each layer has a separate number of hidden units for training. A set of 214 features are transformed using a set of linear layers to 128 features using the first layer of 512, followed by 256 and 128 neurons units in the second and third layers, respectively.

3.6.2.3 Ensemble Learning

This proposed process of boost ensemble learning is used to generate multiple predictions and later combine these predictions. This is the type of ensemble learning technique in which initially models are learned as weak learners with their optimization process. At the same time, the third classifier is built at every step to improve the overall accuracy. The basic flow of the boosting process is as follows,

i. Extract a random sample set $s_1$with no replacement from a bigger training set and train the first weak learner that is SVM in our case.

ii. Again, extract a second randomized sample set s_2 from the original set. Using this small set to train the second weak learner (SVM).

iii. Finally, extract another subset to train the third weak learner (MLP).

iv. In the end, combine all the trained weak learners based on weighted majority voting.

We have used SVM as the first and second learner and ANN (MLP) as the third learner. Initially, both SVM classifiers train on a subset followed by an ANN network trained on the data where both SVMs contradict because of data randomness. For the loss function, we have employed cross-categorical cross-entropy as the loss function (Aribisala & Claridge, 2005). In this proposed classifier we have incorporated SVM and ANN through proposed ensemble learning for multi-class classification, which is our research objective as well.

3.6.3 AlexNet and VGG-16 Based Hybrid Classifier

In this proposed framework, features of two pre-trained architectures which are VGG-16 (Simonyan & Zisserman, 2014) and Alexnet (Krizhevsky et al., 2012a) are extracted and later convolution of these features is implemented. This classifier is termed as hybrid because features from both architects are concatenated with an additional fully connected layer to perform the classification. Both architectures are pre trained on the ImageNet dataset. This dataset contains 15 million images of 1000 classes (Socher et al., 2009). Our proposed algorithm will classify the images into three skin lesion classes: melanoma, nevus, and SK. In deep learning-based convolution neural network architectures, one class required at least 10 thousand instances to train the model from scratch. Since this defined dataset contains the limited number of instances per class, so to achieve good results transfer learning mechanism has been utilized.

In both described architectures, the input images are pre-processed into the 224x224 size. Our first pre-trained architecture is AlexNet. This architecture contains a total 8 layers, which contain 5 convolution layers and 3 fully connected layers. Convolutions and max Pooling have been performed on the filter size of 11x11, 5x5, and 3x3. In addition to that, each convolution layer also contains the Relu activation and dropout function. Optimization of the weights is done with the stochastic gradient descent (Woodworth et al., 2020). This architecture only picks the convolution features from the last convolution layer and the rest of the fully connected layers of this network have been discarded.

On the other hand, VGG-16 consists of total of 16 layers. In VGG-16, there are 6 different blocks. The first five blocks contain the convolution layers whereas the sixth block has the three fully connected layers. The proposed classification architecture only utilized the convolution blocks for feature extraction. Each convolution layer utilized the small size filters of 3x3 and 2x2 in the whole architecture. Each convolution layer also consists of the max pooling layer and Relu activation function (Fred Agarap, 2018). This architecture has also been pre-trained on the ImageNet dataset (Socher et al., 2009). Likewise, to AlexNet, from this architecture, we picked the convolution features from the last convolution block and discard the remaining layers.

To classify our given dataset ISIC, the algorithm contains the linear classification model using the concatenated convolution features of both architectures. After that, there are three fully connected layers. Which contains the 512, 256, and 128 nodes respectively. Each fully connected layer has adjacent Relu activation function along with the dropout layer. This will prevent the model to go towards overfitting. The cross entropy loss function has been utilized to find out the difference between the actual and the predicted value. Whereas

stochastic gradient descent has been utilized for the weights optimization with the momentum weight of 0.9. The learning rate has been changed dynamically. Initially, we have set 10^{-1}, and it has been increased with -1 factor depending upon the loss value. At last, we have three nodes and the softmax classification function (Hosny et al., 2019) gives us the probability against all three classes present in ISIC dataset. A complete framework diagram for classification is presented in Figure 3.19 which depicts the hybrid nature of this proposed classifier, as feature concatenation is implemented from both networks, AlexNet and VGG-16.

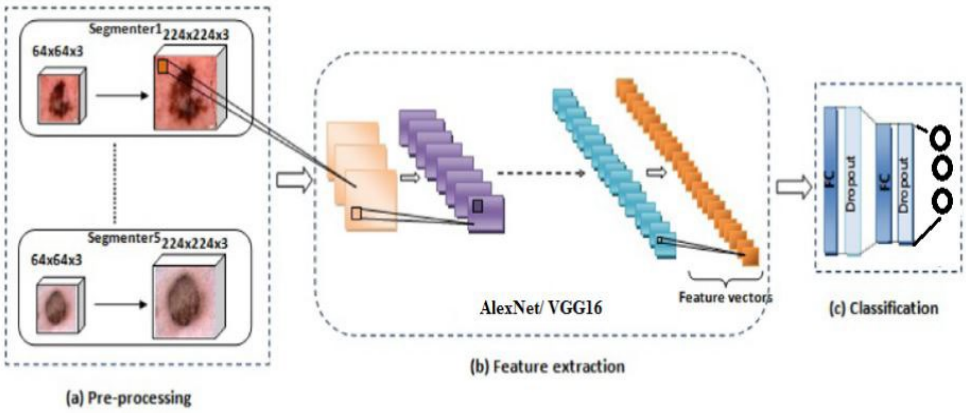


Figure 3.19: Proposed hybrid classifier architecture

Here proposed classification algorithm is simplified by ignoring fully connected layers of AlexNet and VGG-16 and including only three fully connected layers at the end of this hybrid classifier, which reduced the complexity in terms of layers to perform multi-class classification, which is our research objective as well.

3.7 Proposed Frameworks of Complete System for Skin Lesion Detection

In this research work, two end to end systems are proposed and published for skin lesion detection (Khan et al., 2021, 2022). These systems used proposed frameworks for pre-processing, segmentation, feature extraction, and classification which are already explained

in this chapter. The first complete system (Khan et al., 2021) is shown in Figure 3.20. In this system, local contrast enhancement and haze reduction in the CIELAB colour space of the images is implemented in the pre-processing stage. Segmentation is implemented by combining Delaunay triangulation clustering and PSO. Further, HSV value-based thresholding is incorporated in the second stage of segmentation. Colour, pattern, and shape based statistical features are extracted and fused as per implemented criteria. A boost ensemble learning algorithm is implemented for classification using SVM as initial and ANN as a final classifier to learn the patterns of different skin lesion and classify images as melanoma, nevus, and SK.

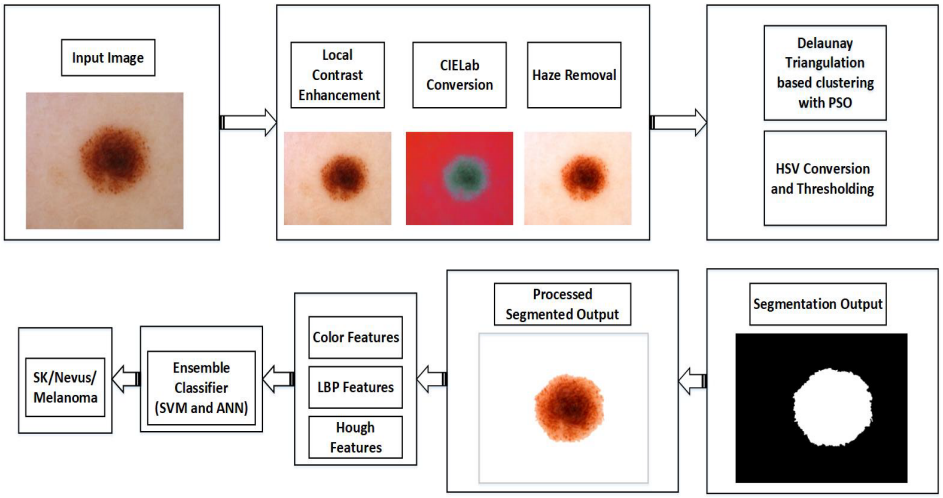


Figure 3.20: Complete proposed System-1 for skin lesion detection

The second complete proposed system (Khan et al., 2022) is presented in Figure 3.21. Here segmentation is performed without any pre-processing. In this algorithm, we have two pipelines for feature extraction. We extract the features through split & merge and contextual encoding approaches. Later we combine the features of both architectures and predict the segmented region through intersection over union mechanism. For the classification step, hybrid features are extracted using Alex-Net and VGG-16 through a transfer learning

approach and this simplified hybrid classifier is used to classify the dermoscopic image into melanoma, nevus, and SK.

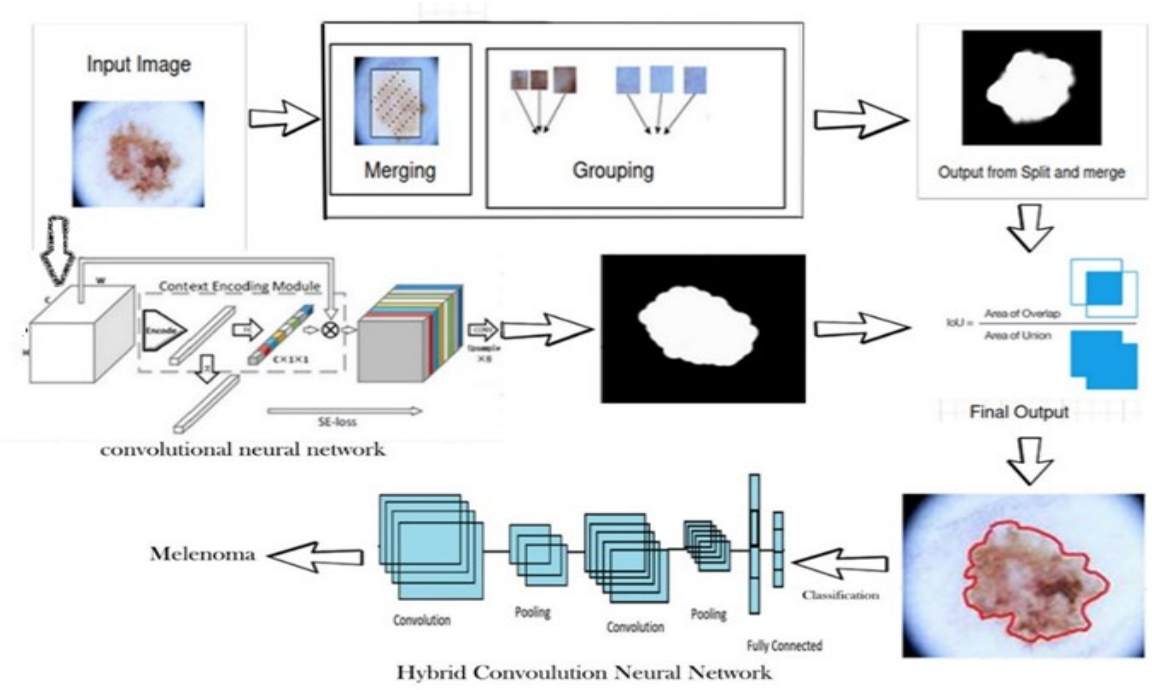


Figure 3.21: Complete proposed system-2 for skin lesion detection

3.8 Summary

This chapter explained all the proposed algorithms for pre-processing, segmentation, feature extraction, and classification. All the techniques and methods which are incorporated in these algorithms with proposed modifications and simplifications, for respective step is also explained in detail. In the pre-processing stage, proposed algorithms used ADF, local contrast enhancement, and haze reduction in CIELAB colour space to enhance the overall image quality. Hair and artifact removal is implemented through the proposed six steps in which a weighted matrix is utilized with black-hat morphology and subsequent techniques.

In the first segmentation algorithm, we proposed a boundary condition-based method that used two newly proposed equations. In the second segmentation algorithm, spatial

clustering was introduced with Delaunay triangulation and PSO to enhance the performance of the overall algorithm. In the third segmentation algorithm, split & merge, and contextual encoding network are utilized. In this proposed contextual encoding network, an encoding layer is introduced to extract the statistics of the global semantic features. In the first classification algorithm, a modified CNN-based network is used. The complexity of this proposed network is reduced with respect to the number of layers, parameters, and filters. For the second classifier, statistical features based on colour, pattern, and shape are extracted with proposed values. Later a proposed boost ensemble learning algorithm is implemented using SVM and ANN. In the third algorithm for classification, hybrid features are extracted using AlexNet and VGG-16 through a transfer learning approach. The complexity of this proposed classifier is reduced by ignoring fully connected separate layers of AlexNet and VGG-16. Visual and numerical results are presented in the next chapter to show better performance of the proposed algorithms as compared to benchmarks.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

In this chapter, we explained and discussed all the numerical and visual results generated after implementing the proposed algorithms for pre-processing, segmentation, and classification of skin lesion. We have used two datasets, ISIC (Berseth, 2017) and PH2 (Mendonca et al., 2013). Details of these both datasets are already provided in Section 3.2. To evaluate the proposed work, we have used diverse performance measures, for which details are already provided in Section 2.8. The assembled results are compared with the renowned benchmark techniques available in the existing literature.

4.2 Results and Discussion of Proposed Frameworks for Pre-processing

This section discusses the results of pre-processing using three proposed algorithms. The first algorithm used ADF and adaptive thresholding. In the second algorithm, hair removal is implemented through black hat morphology with total variation. In the third algorithm, haze reduction is implemented in CIELAB colour space. All the proposed algorithms are implemented to achieve the first research objective that we need to highlight the skin lesion after hair and other artifact removal.

4.2.1 Results of Pre-processing for Smoothing and Enhancement of Skin Lesion

This proposed algorithm for pre-processing is implemented in Matlab and tested on the ISIC dataset (ISIC, 2016). From this dataset, we have used 300 images with different skin cancer histories, phototypes, sizes, subtypes, and localization. Visual results for five selected samples are presented in Figure 4.1. Sample-1 image is from benign, SK skin cancer

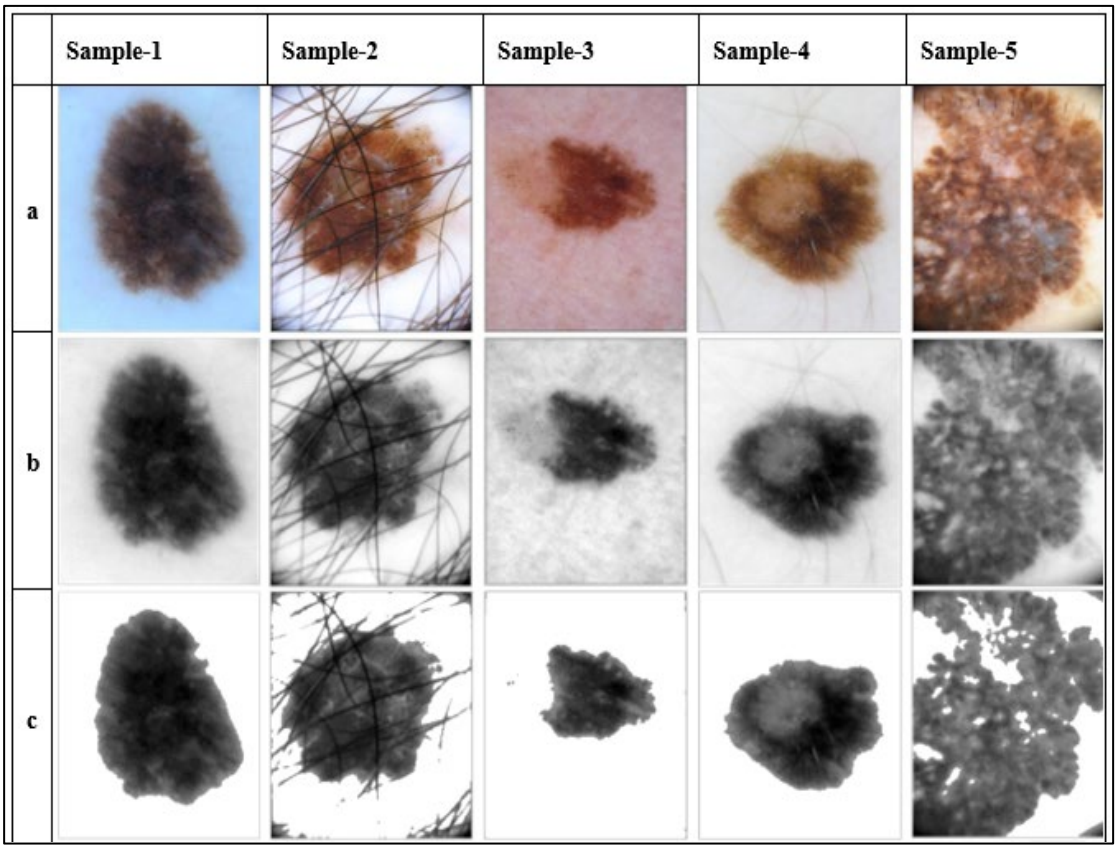


Figure 4.1: Five samples of pre-processed images by the proposed algorithm, (a) Original Image, (b) Gaussian filtered image, (c) Image after ground pixel manipulation

type, which is localized at the abdomen, it has uniform lesion for the affected part. In sample-2 melanoma was diagnosed and located at the pubis. This image has bad illumination due to the poor quality of the image as it is dark at the corners, also skin lesion in this sample is covered by body hair. Sample-3 was diagnosed as nevus but localized at the breast and it has a non-uniform lesion as the right side of the foreground of the image is more affected as compared to the left side. In sample-4 nevus was diagnosed and localized at the upper limb, it has a uniform lesion only at the boundary of the foreground of the image. Sample-5 is diagnosed as melanoma and located at the upper extremities of the body, it has rough skin lesion with degraded image resolution and illumination.

In all the samples, the foreground is considered as skin lesion or affected part, and the rest of the image parts are considered as background. In sample 1 and 4, the lesion was uniform mostly without any hair or dark corners in the image therefore proposed algorithm was able to highlight almost every single part of the foreground in the image as shown in Figure 4.1 (c). In sample 2, there were a lot of hairs, but still foreground was highlighted and most of the hairs were ignored as shown in Figure 4.1 (c). In sample 5 although there were dark corners near lesion, the proposed algorithm was able to highlight the foreground in such a way that dark corners can be differentiated easily from lesion. In sample 3, the lesion was non-uniform due to which the left side of it was ignored after pre-processing. This shows that pre-processing through this proposed algorithm is less effective for images having high non-uniformity.

4.2.2 Results of Pre-processing for Hair and Artifacts Removal

This algorithm incorporated black hat morphology and total variation filtering to remove the hairs from dermoscopic images. Details of these techniques and how we are using these methods in our proposed algorithm is already provided in Section 3.3.2. This algorithm is implemented in Python and has been tested on 2000 images on the ISIC dataset (ISIC, 2016). Here is a brief explanation of this proposed algorithm, morphological operations are applied to shrink or enhance some image regions through opening, closing, erosion, and dilation. Black hat morphology used in this algorithm tends to enhance image components for which the structuring element is larger, as well as these components, are darker than their surroundings. For hair detection, a structuring element of size 23 × 23 is utilized for morphological operation. Hair contour images that are obtained from the previous step contain variations of grey-scale intensity. To increase the intensity of detected hair regions, binary thresholding is applied, and the resulting image enhances the hair

contours. In the end, total variation filtering is applied for image inpainting. Details of these steps is already provided in Section 3.3.2. Figure 4.2 presents three sample images over which this proposed algorithm is applied to remove the hairs. In all the output images hairs are removed efficiently without distorting useful information about the lesion, which is very critical at the later stage of the classification. The image in Figure 4.2 (a) has a higher number of hairs, still proposed algorithm removed every single hair efficiently. The image in Figure 4.2 (e) has a marked line which is also removed through the proposed algorithm.

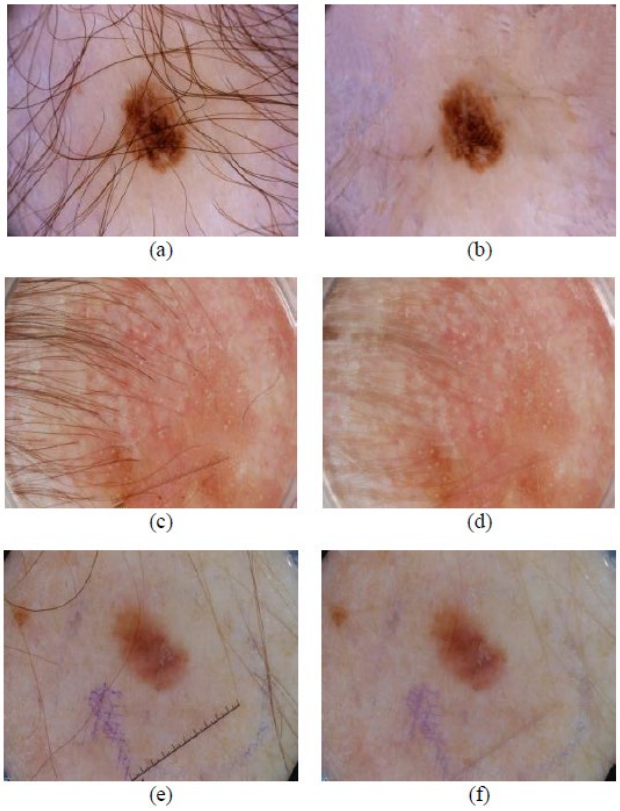


Figure 4.2: (a), (c) and (e) present the original images with hair (b), (d) and (f) present the recovered images without hair and other artifacts

The efficiency of this proposed algorithm for hair removal can be observed through numerical analysis as well, provided in Table 4.1. Classification is implemented through a CNN-based classifier with the reduced number of layers and compared with traditional AlexNet, VGG-16, and VGG-19-based classifiers. Details of the proposed CNN-based

Table 4.1: Accuracy (%) based comparison of proposed algorithms against the ISIC dataset images with and without hair

Methods	Accuracy With Hair (%)	Accuracy Without Hair (%)
Proposed	83	96
AlexNet	78	84
VGG-16	77	81
VGG-19	67	79

classifier with reduced complexity is provided in Section 3.6.1. The numerical analysis presented in Table 4.1 shows how classification results are improved after hair removal is implemented through our proposed algorithm. For classification dataset has been divided in to 80% training set, 10% validation set, and 10% test set. Single-pipelined architecture is trained. For optimizing the trained model, the stochastic gradient descent function is used. Learning rate is set to 0.0001 and momentum to 0.3. For training of the proposed architecture, GTX 1080 Ti having 11 GB memory is utilized. Training of this model took approximately 4 days for experiments with all three chunks of datasets. For calculating network loss mean square error is employed. Table 4.1 contains an accuracy-based comparison of all the classifiers when implemented on images with the hair and on images when pre-processed with the proposed hair removal algorithm. In Table 4.1, the CNN-based classifier is simplified after proposed changes, whereas AlexNet and VGG-16 are pre-trained networks. Details and results related to these classifiers are provided in more detail in Section 4.4.1. From the results provided in Table 4.1, it can be stated that the classification accuracy of all the classifiers improved after implementing our proposed hair removal algorithm.

4.2.3 Results of Pre-processing for Enhancing Contrast and Haze Reduction

This proposed algorithm combined multiple techniques for pre-processing the original noisy image for which details are already provided in Section 3.3.3. Here is a brief overview of all the steps. Initially, we applied local contrast enhancement on the image to improve the contrast which makes the diseased region more prominent. Following this, a statistical method of haze reduction is applied to the CIELAB colour space of the image that uses mean and standard deviation to enhance the overall image. Direct conversion from RGB to CIELAB is not possible and first, it needs to be transformed from RGB to XYZ tristimulus with the selection of a white point in the image and gamma function. The gamma function is a method utilized to linearize simple RGB image components when transforming them into a linear space. This algorithm is implemented in Python and tested over 2000 images of the ISIC dataset. The results of the sample images are shown in Figure 4.3. The contrast of all the objects in the images is enhanced after applying local contrast enhancement, as shown in Figure 4.3 (b). Figure 4.3 (c) is the blue channel representation of images in CIELAB colour space, the reason for selecting this channel is already provided in Chapter 3 for this proposed algorithm. Usually, the colour of skin lesion, especially melanomas, are very dark, and channel b of the CIELAB colour space clearly highlighted the lesion part of the images. Figure 4.3 (d) represents the images after haze reduction, and it is evident from this visual analysis that now images are more suitable for segmentation as compared to the original form of the images.

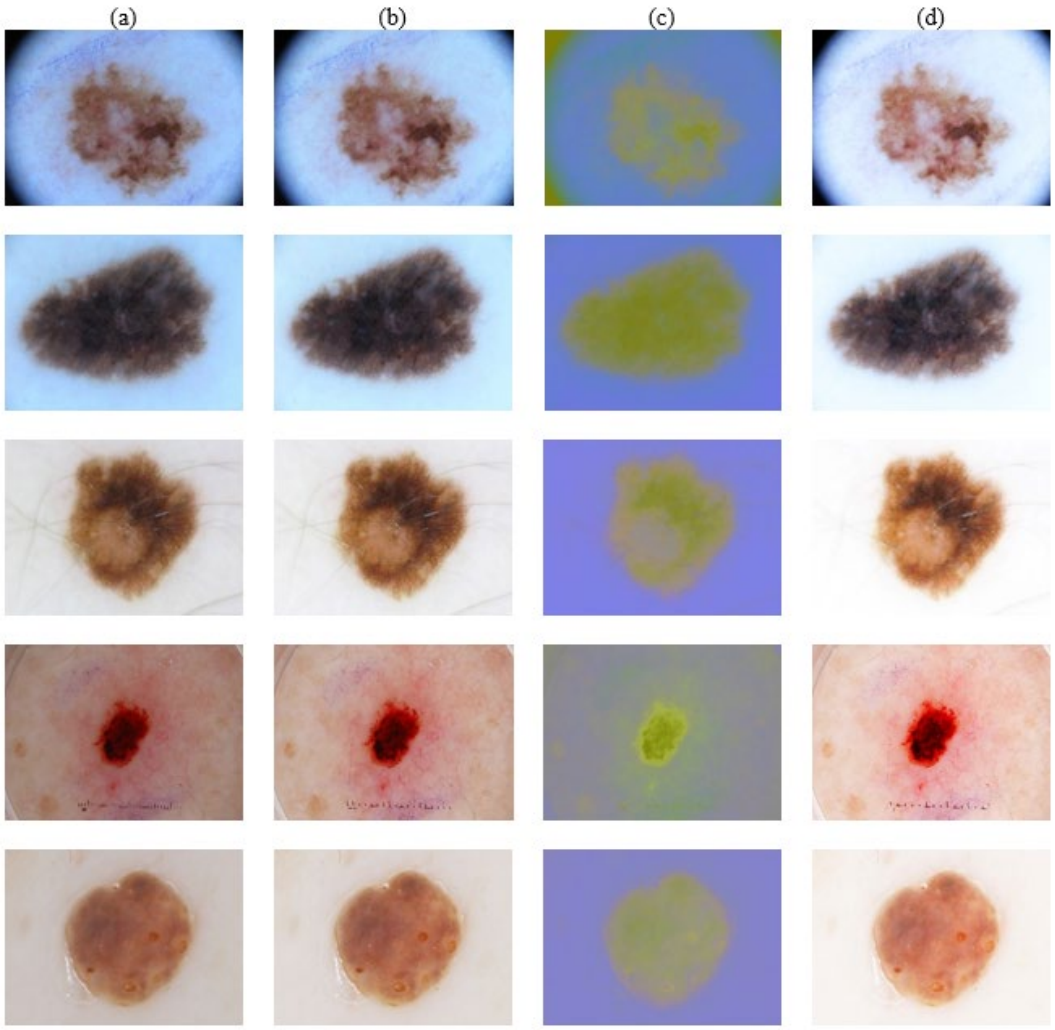


Figure 4.3: Skin lesions enhancement by proposed pre-processing: (a) Original image, (b) Local contrast-enhanced image, (c) CIELAB converted image, (d) Haze removed Image

4.3 Results and Discussion of Proposed Frameworks for Segmentation

This section discusses the results of image segmentation by proposed algorithms. In the first algorithm, a new method based on boundary condition is used. The second algorithm used Delaunay clustering with PSO. The third algorithm used spilt & merge and contextual encoding network for image segmentation. In these proposed segmentation algorithms, we have targeted the second research question in which we had to implement and incorporate statistical, machine, and deep learning-based methods to extract skin lesion from dermoscopic images, which is implemented here to achieve the second research objectives.

4.3.1 Results of Image Segmentation Through Boundary Condition Model

This proposed method is implemented in Matlab and tested over the ISIC dataset. Numerical results for five selected samples are presented in Table 4.2 and visual results for sample images are presented in Figure 4.4. Performance measures used here are DSC, JSC, Sensitivity (TPR), Specificity (TNR), and Accuracy. Details of these are already provided in Section 2.8; Equation 2.8 to Equation 2.12 (Balafar, 2014; Kakadiaris et al., 2018). The proposed algorithm segmented all the sample images like the ground truth images, except for sample 3, in which only the right side of the lesion is segmented, shown in Figure 4.4 (f). In this sample image, the left part of the lesion was ignored due to pixel manipulation. For same sample-3, lower numerical results were observed in Table 4.2 where JSC and sensitivity are 75.6% and 75.5% respectively and overall accuracy is also the lowest which is 94.3% as compared to other sample images.

Sample-1 and 4 secured the highest accuracy as the skin lesion was uniform and image quality was better therefore segmented part is almost like a ground truth image and achieved the highest DSC, JSC, and accuracy, as shown in Table 4.2. In sample 2, the image has underbody hair and due to bad illumination, there were black spots in the corners, but this proposed technique even segmented this image with an accuracy of 95.2%.

Table 4.2: Statistical results for sampled images using the proposed algorithm

Samples	DSC	JSC	Sensitivity	Specificity	Accuracy
Sample-1	98.3%	96.6%	96.9%	99.8%	98.9%
Sample-2	92.0%	85.3%	95.0%	95.3%	95.2%
Sample-3	86.1%	75.6%	75.5%	98.7%	94.3%
Sample-4	98.5%	97.1%	99.8%	99.0%	99.2%
Sample-5	85.2%	74.1%	92.8%	94.8%	94.9%

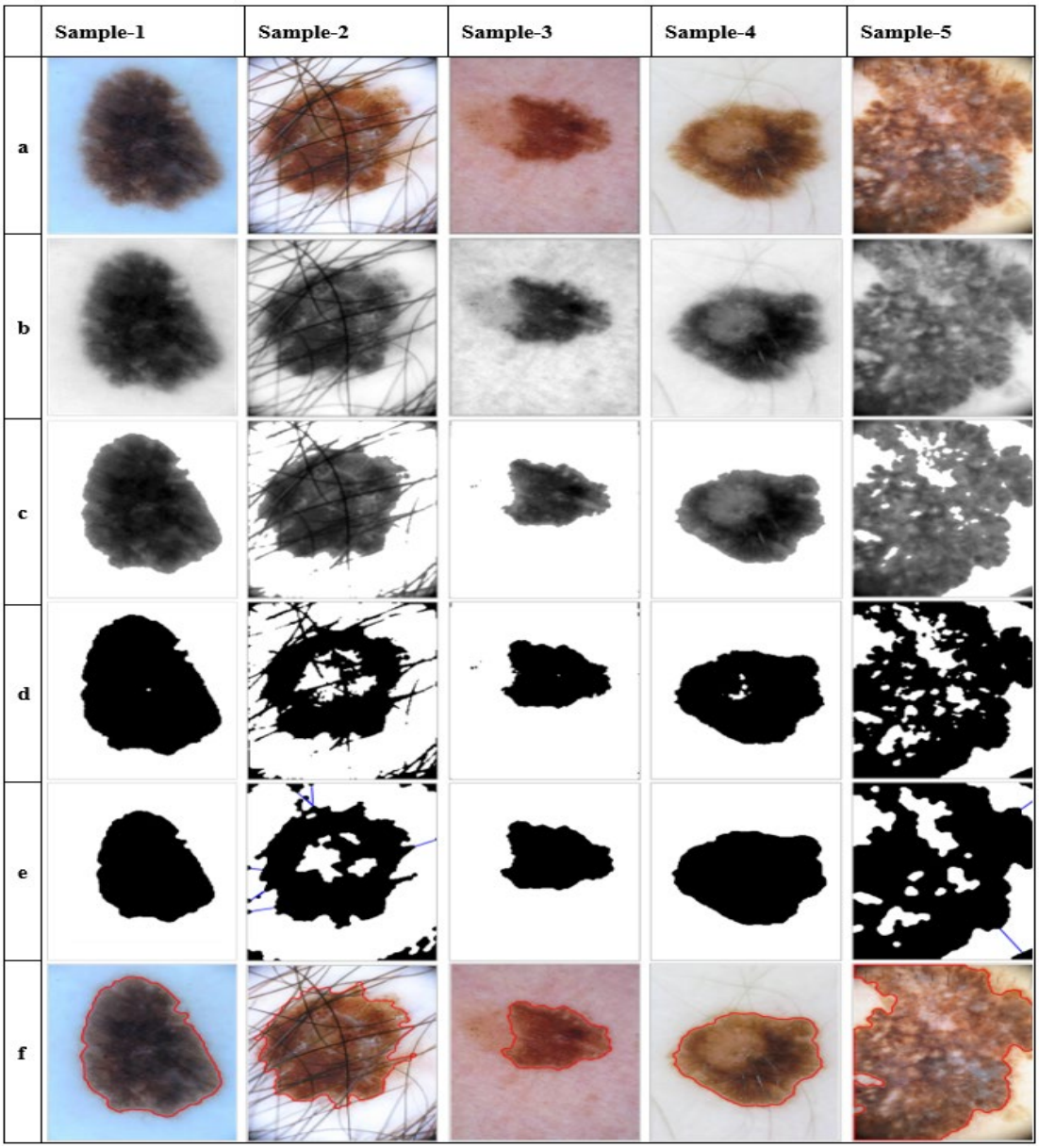


Figure 4.4: Five samples of segmented images using the proposed algorithm, (a) Original Image, (b) Gaussian filtered image, (c) Image after ground pixel manipulation, (d) Image after adaptive thresholding, (e) Image after morphological operations and boundary conditions, (f) Segmented Image

This proposed algorithm ignored the hair and other objects in an image through boundary condition implementation, as shown in Figure 4.4 (e), because the distance of these objects from the skin lesion was more than 200 pixels, distance measurement can be seen by blue lines in Figure 4.4 (e) for sample 2. For sample 5, non-skin lesion parts at the corner of this image were ignored by the same boundary condition, as shown in Figure 4.4 (e). The

inner part of skin lesion for sample-2 and 5 got ignored due to morphological functions and adaptive thresholding but it did not affect the overall segmentation results due to the proposed boundary condition algorithm. Overall, it can be stated that the proposed algorithm performed better segmentation for all conditions except when the skin lesion is non-uniform at the boundaries, due to which the less affected part of the skin lesion is ignored.

The averaged statistical results for 300 images are presented in Table 4.3. In this table, the performance of the proposed method is also compared with averaged results of other techniques in the literature; L-SRM (M. E. Celebi et al., 2008), Otsu-R (Cavalcanti et al., 2010), Otsu-RGB (Cavalcanti & Scharcanski, 2011) and TDLS (Cavalcanti et al., 2014). Comparison is done based on sensitivity, specificity, and accuracy only as results for DSC and JSC were not available for benchmark techniques. The proposed algorithm achieved the highest specificity and accuracy. As for sensitivity, the proposed algorithm is in third place as shown in Table 4.3. The main reason for lesser sensitivity is due to the images having higher non-uniformity in skin lesions. This proposed algorithm is implemented in Matlab software and a screenshot of the sample segmented image is shown in Figure 4.5. All the results presented in this section are aligned with the second research objective in which the boundary of skin lesion is detected successfully through a statistical model while exhibiting

Table 4.3: Averaged statistical results for all 300 images for the proposed method and benchmark techniques

Methods	Sensitivity	Specificity	Accuracy	DSC	JSC
Proposed Method	88.4%	99.4%	96.6%	91.56%	84.97%
L_SRM	89.4%	92.7%	92.3%	--	--
Otsu-R	87.3%	85.4%	84.9%	--	--
Otsu-RGB	93.6%	80.3%	80.2%	--	--
TDLS	88.8%	98.0%	96.0%	--	--

better performance as compared to related existing methods.

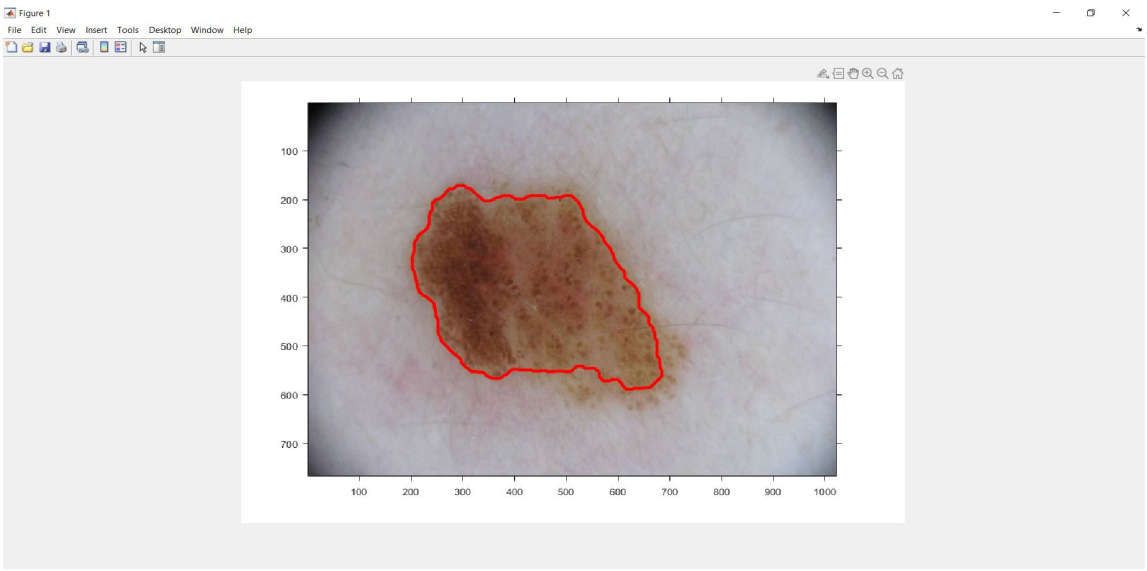


Figure 4.5: Matlab screenshot of a segmented image

4.3.2 Results of Image Segmentation Through Delaunay Clustering and PSO

This algorithm is implemented over two datasets, 2000 images of ISIC and 200 images of PH2 datasets. The step-by-step visual output of sample images during the segmentation process is shown in Figure 4.6. Images after applying HSV thresholding are presented in Figure 4.6 (b). As mentioned earlier that H (hue) is the actual colour pigment in 360 rotations, S (saturation) shows the concentration of colour from 0 to 100, and V (value) tells how bright the colour is from the range 0 to 100. From the output images in Figure 4.6 (d), it can be easily considered that most lesions have specific brunt dark reddish and brownish representations. Using this specific threshold (39, 88, 90) after multiple experiments, we extracted the skin lesion region and applied binary thresholding as shown in Figure 4.6 (c) which is translated over the original dermoscopic image to outline the segmented lesion as shown in Figure 4.6 (d). For some images, several pixels are segmented falsely as foreground, but overall proposed segmentation algorithm performed well.

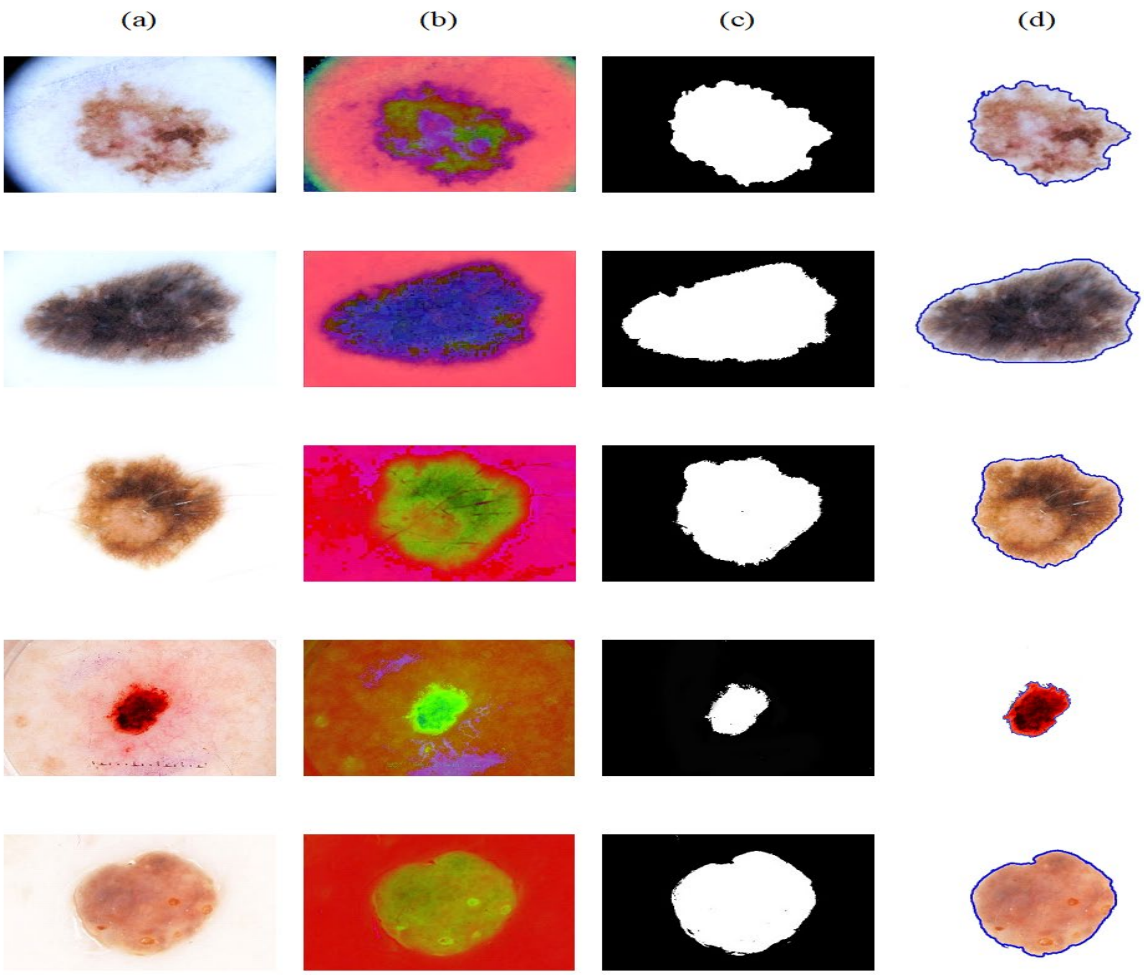


Figure 4.6: The segmentation results by proposed algorithm: (a) Input image for the segmentation process, (b). Image after HSV thresholding, (c) Binary Image, (d) Segmented image

Table 4.4 and 4.5 present the confusion matrix of the ISIC and PH2 datasets respectively. A confusion matrix is formulated by averaging the ratio of each TP, FN and FP, TN with the number of actual positive and negative pixels respectively. Performance measures that are used for evaluation are presented with details in Section 2.8; Equations 2.10, 2.11, and 2.12 for TPR, TNR, and accuracy respectively (Hasan et al., 2020; Xie et al., 2020; Abayomi-Alli et al., 2021; Gouda et al., 2022).

Table 4.4: Confusion matrix for segmentation on ISIC dataset using the proposed algorithm

ISIC	Foreground	Background
Foreground	97.6%	2.4%
Background	4.1%	95.9%

Table 4.5: Confusion matrix for segmentation on PH2 dataset using the proposed algorithm

PH2	Foreground	Background
Foreground	95.3%	4.7%
Background	11.2%	89.8%

Table 4.4 shows that the proposed segmentation algorithm accurately classifies foreground pixels with true positive accuracy of 97.6% and background pixels with true negative accuracy of 95.9%. Such results are demanding because no information regarding skin lesions will be dropped. Similarly, Table 4.5 presents the segmentation results on the PH2 dataset which depicts that 95.3% of foreground pixels are filtered out correctly while 4.7% of pixels from the foreground are misclassified as background, hence some loss of information. Likewise, in results on ISIC, more background pixels are regarded as foreground pixels.

A comparison of the proposed segmentation technique with some renowned benchmarks from the literature concerning the accuracy, TNR, and TPR are presented in Table 4.6 and 4.7 for ISIC and PH2 datasets respectively. From Table 4.6 and 4.7 we can interpret that the proposed algorithm has achieved segmentation accuracy of 96.8% and 92.1% for the ISIC and PH2 datasets respectively, better than statistical methods and deep learning-based models.

Table 4.6: Comparison of the proposed algorithm with benchmarks for the ISIC dataset

Paper	Accuracy (%)	TNR (%)	TPR (%)
(Jaisakthi et al., 2018)	91.9	98.5	80.5
(Afza et al., 2019)	96.0	-	-
(Sultana et al., 2018)	94.6	-	-
(Khan et al., 2019)	94.4	90.0	92.0
(Xie et al., 2020)	93.8	96.4	87.0
Proposed	96.8	97.6	95.9

Table 4.7: Comparison of the proposed algorithm with benchmarks for the PH2 dataset

Paper	Accuracy (%)	TNR (%)	TPR (%)
(Fan et al., 2017)	93.6	-	87.0
(Pennisi et al., 2016a)	89.6	97.2	80.2
Proposed	92.1	95.0	89.5

In terms of sensitivity (TPR) and specificity (TNR), the proposed techniques achieved better results than traditional Delaunay triangulation (Pennisi et al., 2016) and deep learning-based method (Khan et al., 2019) for segmentation. The sensitivity value indicates the ability of an algorithm to accurately distinguish skin lesion pixels from other skin region pixels. Whereas the specificity tells about the ability to accurately discard the skin region pixels which are not related to skin lesions, which can later cause misclassification during skin lesion type recognition. Hence, accurate segmentation is necessary for correct classification. Other than comparable accuracy with existing techniques, the proposed algorithm does not require any computational complexity as in deep learning-based algorithms. We can interpret from Table 4.7 that the performance of the proposed algorithm on PH2 dataset is compelling in that is trained using ISIC dataset. If we employ the PH2 dataset to train the same algorithm, we can achieve better performance on the PH2 dataset

as well. To validate the results, here is an example of accuracy calculation is given for the ISIC dataset. As mentioned earlier, TP is equivalent to several pixels correctly segmented as foreground, TN is the number of pixels segmented correctly as background, FP is several pixels segmented falsely as foreground and FN is equivalent to several pixels falsely segmented as background. From the confusion matrix presented in Table 4.4, TP=97.6%, TN=95.9%, FP=2.4% and FN=4.1%. Accuracy is calculated by following Equation 4.1.

$$Accuracy = \frac{97.6+95.9}{97.6+95.9+2.4+4.1} = 96.8\% \qquad \text{Equation 4.1}$$

This accuracy result is same as the accuracy of the proposed algorithm provided in Table 4.6.

Visual segmentation results on ISIC and PH2 datasets are presented in Figure 4.7 and Figure 4.8 respectively. In these figures, input images are presented in the first column. The second and third columns present the predicted output by the proposed algorithm and ground truth image respectively. The last column shows the difference in the ground truth and predicted output by drawing red lines as the predicted output and green lines as the ground truth image. When analyzed closely it is observed that the false positives are higher than the false negatives, same was observed through the numerical results. Hence, the impact on false classification will be reduced. All the results presented in this section are aligned with the second research objective in which the boundary of the skin lesion is detected successfully by incorporating the existing statistical model, Delaunay and PSO, while exhibiting better performance as compared to related existing techniques.

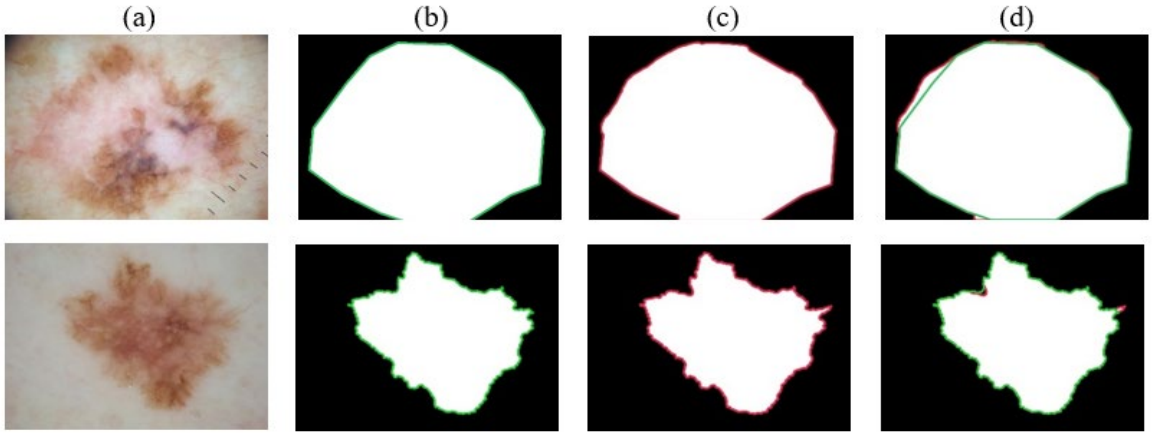


Figure 4.7: Segmentation results through proposed algorithm for ISIC dataset: (a) original image, (b) ground truth output, (c) predicted output, (d) difference in the ground truth and predicted output

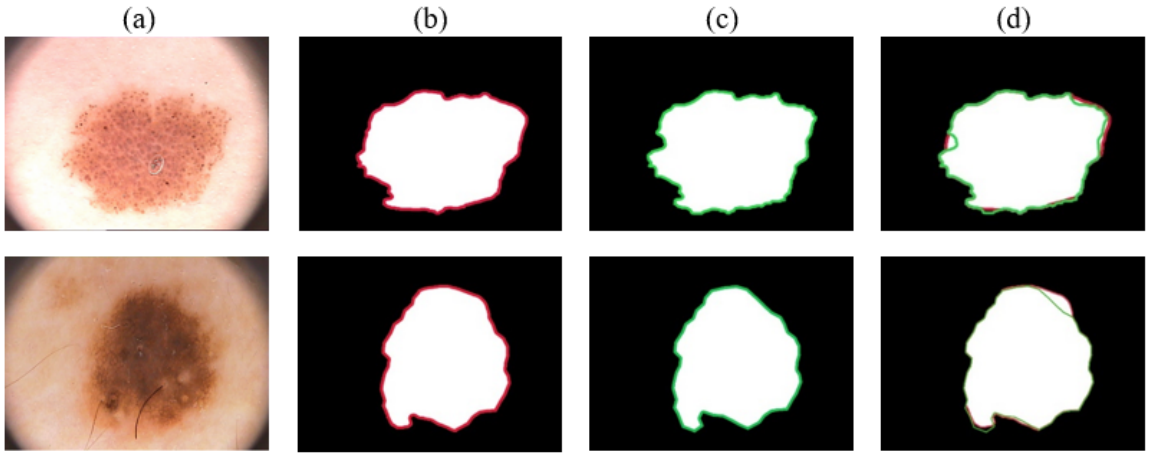
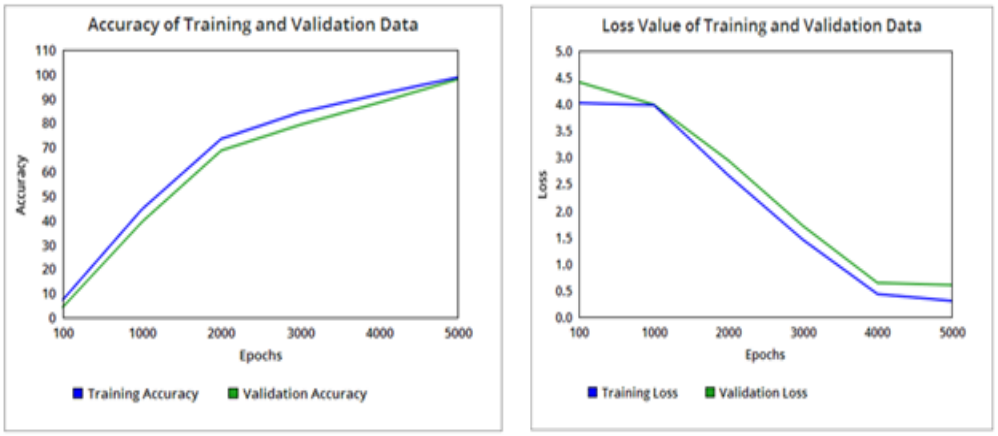


Figure 4.8: Segmentation results through proposed algorithm for PH2 dataset: (a) original image, (b) predicted output (c) ground truth output, (d) difference in the ground truth and predicted output

4.3.3 Results of Image Segmentation by Split & Merge and Contextual Encoding

This proposed algorithm has been implemented using the deep learning library tensor-flow in the Python language. Whereas the dataset splits into 80 percent training portion and 20 percent images on the validation side. Training and validation graphs of both accuracy and loss for segmentation are shown in Figure 4.9. This proposed architecture is trained till the convergence of the loss. Figure 4.9 (a) presents the training and validation accuracy of the contextual encodings. The accuracy is started with 10.2% at the 100^{th} epochs,

after that, it ramped exponentially and ends with 98.25% training accuracy and 97.21% validation accuracy. Similarly, Figure 4.9 (b) depicts the training and validation loss, both values decrease gradually which depicts the robustness of used hyperparameters in our proposed architecture.



(a) Training and Validation Accuracy on Contextual Encoding

(b) Training and Validation loss on Contextual Encoding

Figure 4.9: Training and validation for accuracy and loss on contextual encoding in the proposed algorithm

This algorithm is implemented over two datasets, 2000 images of ISIC and 200 images of PH2 datasets. The step-by-step visual output of sample images during the segmentation process is shown in Figure 4.10. The first column in Figure 4.10 presents the input images, second and third one shows the output result after the split & merge method and contextual encodings respectively. Whereas the last column depicts the output image after taking the Intersection over Union (IoU) of the split & merge method with the contextual encoding. It is clear from the visual results presented in Figure 4.10 (d) that the skin lesion of the image is segmented in almost all the images with reasonable accuracy. The issue of over-segmentation is observed while segmentation with the split & merge method as shown in the 4th image in Figure 4.10 (b), where the non-affected skin part is also included in the skin lesion. But due to the proposed IoU with contextual encoding, it did not affect the

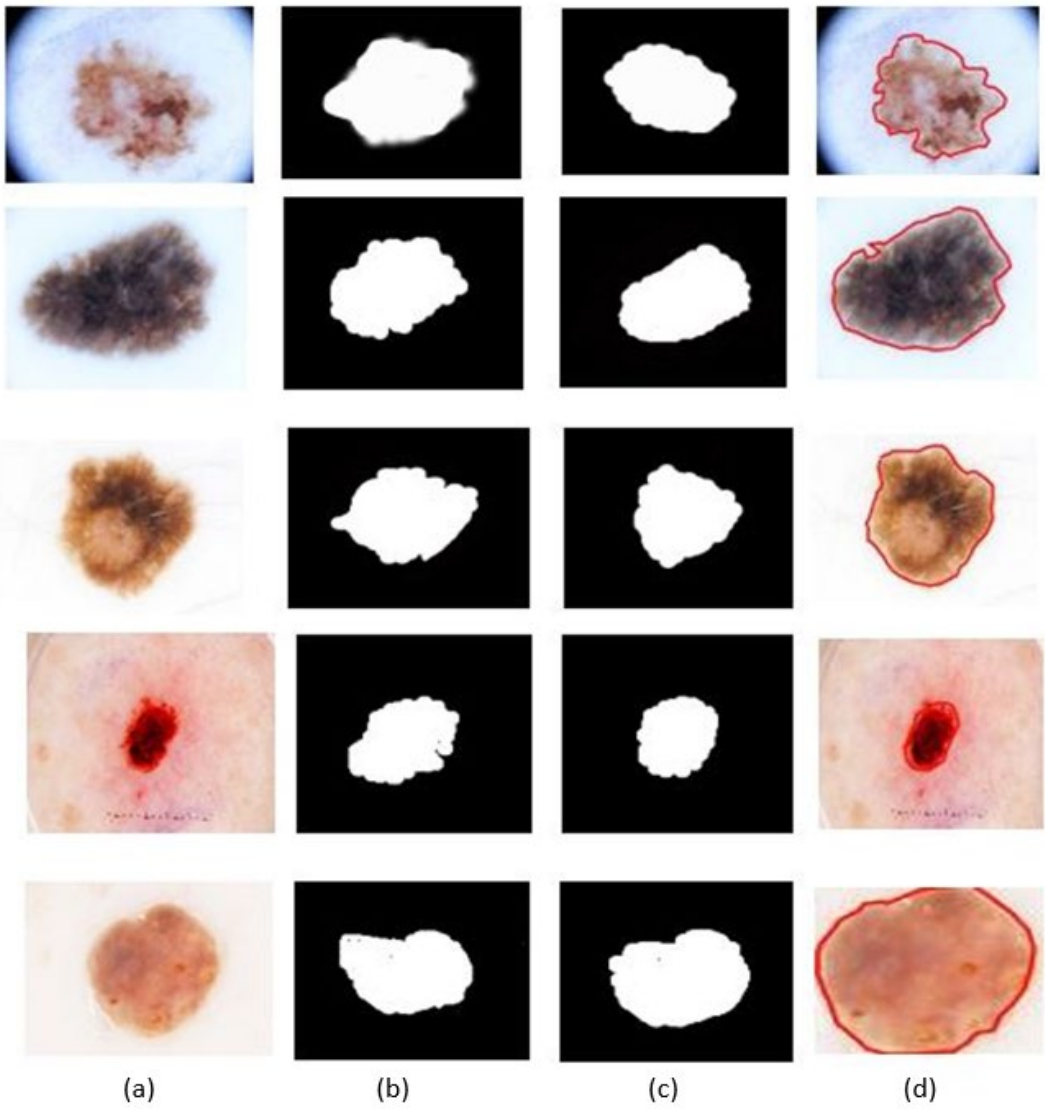


Figure 4.10: Segmentation results by proposed algorithm (a) input image, (b). image after split & merge, (c) image after contextual encoding, (d) segmented image after IoU

final segmented image, as shown in the 4th image in Figure 4.10 (d). While implementing segmentation through contextual encoding, some portion of lesion is left out, as shown in the 2nd image in Figure 4.10 (c). But it did not affect the final segmented image because of IoU with split & merge segmentation.

Moreover, Figure 4.11 and Figure 4.12 show the results of our proposed segmentation algorithm with the ground truth labels of the ISIC and Ph2 datasets respectively, for three selected images. The blue rounded line depicts the ground-truth label

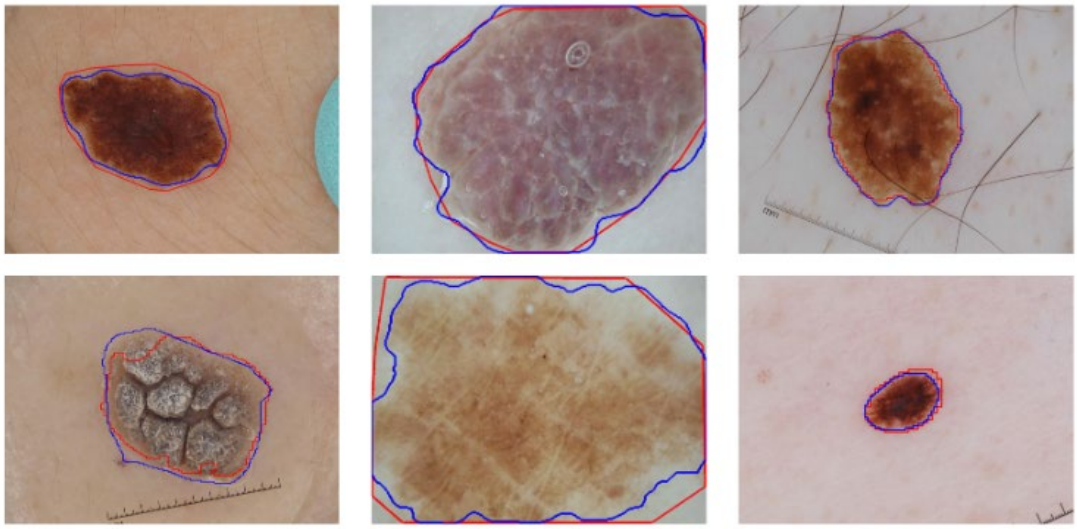


Figure 4.11: Predicted segmentation mask (red line) Vs Ground truth masks (blue line) on ISIC dataset

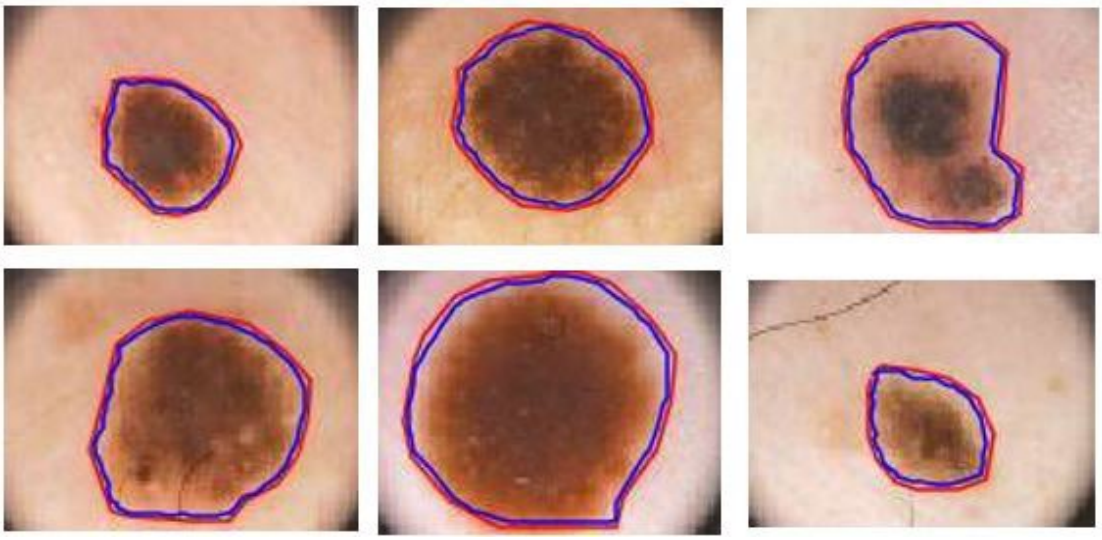


Figure 4.12: Predicted segmentation mask (red line) Vs Ground ruth masks (blue line) on PH2 dataset

from respective datasets and the red line shows the region which our proposed algorithm predicts. From Figures 4.11 and 4.12, we can state that false negatives are reduced now, which means pixels segmented falsely as background whereas they belong to foreground, are reduced now. Predicted and ground truth masks are reasonably similar. In the skin lesion segmentation problem, the more important issue is that portion of the lesion should not be left out, and here we have achieved that with the proposed algorithm. Hence from the visual

analysis presented in Figures 4.10, 4.11, and 4.12, it can be stated that this proposed segmentation algorithm almost clearly identifies the skin lesion boundary with minimum error.

The confusion matrix on segmentation for both datasets ISIC and PH2 datasets has been shown in Table 4.8 and Table 4.9 respectively. A confusion matrix is formulated by averaging the ratio of each TP, FN and FP, TN with the number of actual positive and negative pixels respectively. Details of these terms are already provided in Section 2.8. Our proposed segmentation algorithm accurately classifies the foreground pixels with true positive accuracy of 98.2% and background pixels with true negative accuracy of 97.4% as shown in Table 4.8. The results given in Table 4.8 show that more background pixels are classified as foreground pixels as compared to foreground pixels classified as background. These results are quite demanding as no information related to skin lesion will be dropped. Similar results were observed in visual analysis from Figure 4.11 and Figure 4.12 where false negatives are reduced as no information regarding skin lesion is lost. Likewise, segmentation results on the PH2 dataset are presented in Table 4.9 which shows that 97.1%

Table 4.8: Confusion Matrix on ISIC dataset using the proposed algorithm

ISIC	**Foreground**	**Background**
Foreground	98.2%	1.8%
Background	2.6%	97.4%

Table 4.9: Confusion Matrix on PH2 dataset using the proposed algorithm

PH2	**Foreground**	**Background**
Foreground	97.1%	2.9%
Background	3.7%	96.3%

of foreground pixels are classified correctly whereas 2.9% of foreground pixels are misclassified as background pixels, resulting in a very minor loss of useful information.

Moreover, the proposed algorithm is compared with existing segmentation methods and results are presented in Table 4.10 and Table 4.11 for ISIC and PH2 datasets respectively. Performance measures that are used for evaluation are presented with details in Section 2.8; Equations 2.8 to 2.12 for DSC, JSC, TPR, TNR, and accuracy respectively.

Table 4.10: Proposed segmentation algorithm comparison with benchmarks on the ISIC dataset

Paper	Accuracy (%)	TNR (%)	TPR (%)	DSC	JSC
(Jaisakthi et al., 2018)	91.9	98.5	80.5	81.2	79.3
(Afza et al., 2019)	96.0	91.3	83.4	84.2	81.1
(Sultana et al., 2018)	94.6	80.3	87.4	85.4	83.2
(M. A. Khan, Javed, et al., 2019)	94.4	90.0	92.0	83.2	81.9
(Xie et al., 2020)	93.8	96.4	87.0	85.1	82.4
(Hasan et al., 2020)	77.5	78.2	76.8	82.4	76.3
(Berseth, 2017)	92.9	98.1	81.5	84.2	76.6
(Yuan, 2017)	93.6	97.8	82.0	85.4	77.9
(Bi et al., 2017)	93.7	99.0	80.2	84.8	76.9
(Jahanifar et al., 2017)	93.6	98.5	82.1	85.0	76.8
Proposed	97.8	98.2	97.4	87.1	79.2

The proposed algorithm achieved the highest accuracy as compared to all benchmark techniques i.e., 97.8% and 96.7% for ISIC and PH2 datasets respectively. Moreover, it has been observed from Table 4.10 and Table 4.11, that on five different evaluation measures our proposed algorithm outperformed benchmark algorithms. The benchmark algorithms used in both tables are based on statistical, traditional machine learning, and the latest neural network-based techniques with reasonable accuracies. In Table 4.10 for ISIC, (Jaisakthi et

Table 4.11: Proposed segmentation algorithm comparison with benchmarks on the PH2 dataset

Paper	Accuracy (%)	TNR (%)	TPR (%)	DSC	JSC
(Fan et al., 2017)	93.6	91.9	87.0	82.4	76.3
(Pennisi et al., 2016)	89.6	97.2	80.2	_	_
(Hasan et al., 2020)	87.0	89.0	85.0	82.4	76.3
(Berseth, 2017)	90.0	94.3	80.5	85.4	74.3
(Yuan, 2017)	92.1	95.2	81.3	85.2	76.2
(Bi et al., 2017)	93.2	96.3	81.3	81.5	77.4
(Jahanifar et al., 2017)	94.2	97.2	83.2	83.3	77.3
(Hu et al., 2019)	91.9	91.3	92.5	_	_
(Abayomi-Alli et al., 2021)	92.2	95.1	80.8	_	_
Proposed	96.7	97.1	96.3	88.1	81.2

al., 2018) a statistical method base segmentation is presented. The accuracy of our proposed algorithm is much better as compared to these methods. TNR for these methods is better as compared to the proposed algorithm and the same was observed from visual analysis. These results also depict that our proposed incorporation of existing techniques is better as compared to existing combinational techniques. Most of the benchmark techniques presented in Table 4.10 and Table 4.11 belong to the deep learning paradigm. The proposed algorithm with reasonable computational cost achieved better accuracy as compared to all these existing methods. Also, we can interpret from Table 4.11 that performance on the PH2 dataset is compelling in that is trained using ISIC dataset. If we employ the PH2 dataset to train the same algorithm, we can achieve better results as well on the PH2 dataset.

To validate the result, here is an example of accuracy calculation is given for the ISIC dataset. As mentioned earlier, TP is equivalent to several pixels correctly segmented as foreground, TN is the number of pixels segmented correctly as background, FP is several

pixels segmented falsely as foreground and FN is equivalent to several pixels falsely segmented as background. From the confusion matrix presented in Table 4.8, TP=98.2, TN=97.4, FP=1.8, and FN=2.6. Accuracy is calculated by following Equation 4.2.

$$Accuracy = \frac{98.2+97.4}{98.2+97.4+1.8+2.6} = 97.8\% \qquad \text{Equation 4.2}$$

This accuracy result is similar to the accuracy of the proposed algorithm provided in Table 4.10. All the results presented in this section are aligned with the second research objective in which the boundary of the skin lesion is successfully detected by the proposed incorporation of split and merge with contextual encoding network and better results are achieved as compared to existing methods.

4.4 Results and Discussion of Proposed Frameworks for Classification

This section discusses the results of algorithms proposed for classification. In the first algorithm, classification is implemented through CNN in which number of layers are modified to reduce the complexity. In the second algorithm, classification is implemented through SVM as the initial classifier and ANN as the second classifier via proposed boost ensemble learning. In the third algorithm, hybrid features are extracted using Alex-Net and VGG-16 through a transfer learning approach. This classifier is termed as a hybrid because features from both architects are concatenated with additional fully connected layers to perform the classification. Classification results are tested only on the ISIC dataset because our main objective was to classify skin lesions into melanoma, nevus, and SK only. Whereas classes in PH2 dataset are completely different. In the proposed classification algorithms, we have targeted the third research objective in which we had to extract statistical and deep features, which will be used in the simplified deep learning-based algorithms to perform

multi-class classification. In this section, all the results and discussion are presented for these classifiers.

4.4.1 Results of CNN-based Classifier

This proposed algorithm is CNN based with a reduced number of convolution layers and filters. Three convolution layers have 48 filters of 5 × 5, 64 filters of 3 × 3, and 128 filters of 3 × 3 sizes respectively. There is a pooling layer after every convolution layer followed by two fully connected layers and Softmax activation at the end. More details regarding this proposed network is already provided in Section 3.6.1. As mentioned earlier this proposed network is evaluated on a publicly available ISIC dataset where we have divided selected images into three chunks: comprising 80% training images, 10% validation images, and 10% test images. Single-pipelined architecture is trained. For optimizing the trained model, stochastic gradient descent function is used. Learning rate is set to 0.0001 and momentum to 0.3. For training of the proposed architecture, GTX 1080 Ti having 11 GB memory is utilized. Training of the model took approximately 4 days for experiments with all three chunks of datasets. For calculating network loss mean square error is employed. Accuracy-based comparisons of the proposed network with available benchmark networks is presented in Table 4.12 where the accuracy of the proposed network exceeds then all the commonly used benchmarks. Accuracy is calculated as given in Equation 2.12 from Section

Table 4.12: Accuracy based comparison of the proposed algorithm with benchmarks

Method	Accuracy (%)
(Hekler et al., 2019)	89
(Dorj et al., 2018)	94
(Almansour & Netw, 2016)	90
Proposed	96

2.8. This high accuracy is achieved by using supervised convolution providing feature templates to the CNN network. One of the reasons to achieve higher accuracy is because we implemented the proposed hair and artifact removal algorithm, presented in Section 3.3.2, before classification. This numerical analysis proved that better classification results are achieved by using the proposed hair removal method along with less complex CNN. Figure 4.13 presents the training accuracy achieved by removing hair and artifact while preserving information about the lesion. Training accuracy crosses 96% while validation accuracy reaches up to 89% after completing sixty epochs. For the proposed network loss initiated from 2.4 drops up to 0.32 and becomes stagnant in succeeding epochs.

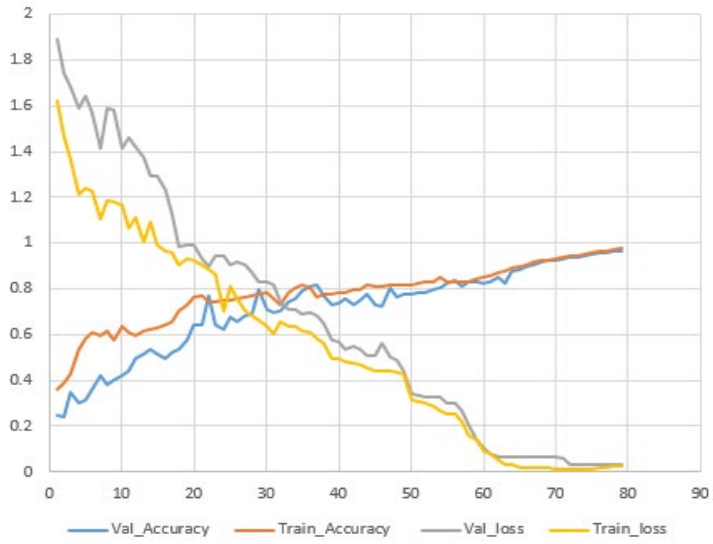


Figure 4.13: Training graph of proposed classification model

Table 4.13 shows the normalized confusion matrix for skin lesion classes. Details of confusion matrix calculation and associated terms are already provided in Section 2.8. We have achieved nearly 96% of validation accuracy on 3 classes of skin lesion datasets which are melanoma, nevus, and SK. Table 4.13, also proved that classification accuracy substantially improved after hair removal implemented through the proposed black-hat morphological algorithm.

Table 4.13: Confusion matrix of accuracies (%) of three classes with hair artifacts and without hair artifacts using the proposed algorithm

Dataset with Details of Hair and Artifacts				**Dataset without Details of Hair and Artifacts**			
	Melanoma	Nevus	SK		Melanoma	Nevus	SK
Melanoma	80	16	4	Melanoma	96	3	1
Nevus	14	81	5	Nevus	5	94	1
SK	9	5	86	SK	2	1	97

Table 4.13 shows that the trained system mixes some of the closely related classes like melanoma and nevus because of similar features. Such similar features of distinct classes affect the overall accuracy of the classifier. Still, the proposed algorithm achieved very reasonable classification accuracy.

It is necessary to highlight the fact that this accuracy is achieved through a simplified CNN-based classifier. The proposed classifier has used only 240 filters, 60K parameters, and 6 layers which is substantially less as compared to existing networks including AlexNet (Krizhevsky et al., 2012), VGG-16, and VGG-19 (Simonyan & Zisserman, 2014). This was the third research objective as well that multi-class classification should be implemented through a simplified deep learning-based classifier.

4.4.2 Results of Ensemble Classifier with SVM and ANN

This proposed algorithm for classification used a set of 214 feature vectors. From the colour histogram, 140 features are secured, 64 features of the LBP Histogram, and 10 features from the Hough method are used. The fusion of these features is performed to train the ensemble classifier. More details regarding feature extraction is already provided in Section 3.5.1. This algorithm is used to generate multiple predictions and later combine these predictions. This is the type of ensemble learning technique in which initially first and

second models are learned as weak learners during the optimization process. At the same time, the third classifier is built at every step to improve the overall accuracy. The classification results of the proposed algorithm are measured in the following three ways.

- Single SVM Classifier
- Single MLP Classifier
- Ensemble SVM and MLP

For all these three ways, a set of features remains the same. To train these classifiers, 10-fold cross-validation is employed. This proposed classifier is implemented in Python and tested on the ISIC dataset in which 2000 images are divided into ten equal number of sets; one set is kept for testing and the classifier is trained for the rest of nine sets. The process is adapted 10 times and results are computed as an average of all 10 rounds. A confusion matrix for the classification performance is shown in Table 4.14. Details of confusion matrix calculation and performance measures used here for evaluation, accuracy, precision, recall, F1 score, and FNR are presented in Equations 2.12 to 2.16 respectively, in Section 2.8.

Table 4.14: Confusion matrix for classification using proposed ensemble classifier

ISIC	Melanoma	Nevus	SK
Melanoma	97.6%	0.8%	1.6%
Nevus	1.5%	97.3%	1.2%
SK	0.7%	0.5%	98.8%

From the results presented in Table 4.14, we can state that similarity between melanoma and nevus is higher, the same was observed from images available in the ISIC dataset where the physical appearance of melanoma and nevus is comparable. Table 4.15 presents the results of the ISIC dataset when experiments are made on three classification

Table 4.15: Comparison of classification between single and proposed ensemble classifier

Classifier	Accuracy (%)	FNR (%)	Precision	Recall	F1-Score
SVM	92.2	7.8	92.0	92.4	92.1
MLP	93.5	6.5	93.2	93.8	93.4
Ensemble SVM and MLP	97.9	2.1	97.2	98.0	97.0

algorithms with the same feature set. Results in Table 4.15 show the superiority of the proposed ensemble classifier over single classifiers like SNM and MLP. The proposed classifiers achieve the highest accuracy of 97.9% on the ISIC dataset. As precision is the highest of the proposed algorithm which means that the correct classification of melanoma skin lesion type is highest as compared to the single classifiers. Better performance is achieved in other performance measures as well provided in Table 4.15.

Moreover, to prove the performance of the proposed classification algorithm along with robust features, a comparison is made with the previous techniques from the literature. Results have been presented in Table 4.16. These performance measures are already presented in Section 2.8. The classification results of the proposed ensemble classifier of SVM and MLP, are better than the recognized benchmark techniques as shown in Table 4.16.

Table 4.16: Comparison of proposed ensemble classifier with benchmark techniques

Paper	Accuracy (%)	TNR (%)	TPR(%)	Response Time(s)
(Khan et al., 2019)	96.5	97.0	96.5	16.46
(Afza et al., 2019)	93.7	95.0	93.3	38.61
(Yap et al., 2018)	72.1	-	-	-
Proposed Ensemble Classifier	97.9	98.0	97.2	20.44

The comparison has been made with two deep learning-based techniques (Yap et al., 2018; Khan et al., 2019) and one statistical technique (Afza et al., 2019) on the ISIC dataset. The proposed algorithm achieves 97.9% of classification accuracy which is better than the deep learning-based techniques presented in Table 4.16. Additionally, the response time of the proposed algorithm is slightly higher than that (Khan et al., 2019) but is less than the statistical approach (Afza et al., 2019). It is logical to have a slightly higher response time because we are implementing an ensemble of two machine learning methods. This trade-off is acceptable if we secure better accuracy as compared to renowned benchmarks. The proposed classifier also exhibits much better performance concerning TNR and TPR as compared to both deep learning and statistical-based methods. In Figure 4.14, true labels and output predicted labels through the proposed classifier on the ISIC dataset for sample images, are presented. Hence all the results presented in this section proved that the proposed ensemble of existing machine learning-based methods provided better classification results as compared to existing statistical and deep learning-based methods, which is our third research objective as well.

Figure 4.14: Classification results on ISIC dataset: (a) True label, (b) Predicted label

4.4.3 Results of AlexNet and VGG-16 Based Hybrid Classifier

This proposed architecture has been implemented using the deep learning library tensor-flow in the Python language and tested on the ISIC dataset. Whereas the dataset splits

into 80 percent training images and 20 percent images for validation. Training and validation graphs of both accuracy and loss for the proposed classification are shown in Figure 4.15. In both cases, convergence has been done moderately. We have achieved 98.40% training accuracy and 97.01% validation accuracy on our proposed classification network.

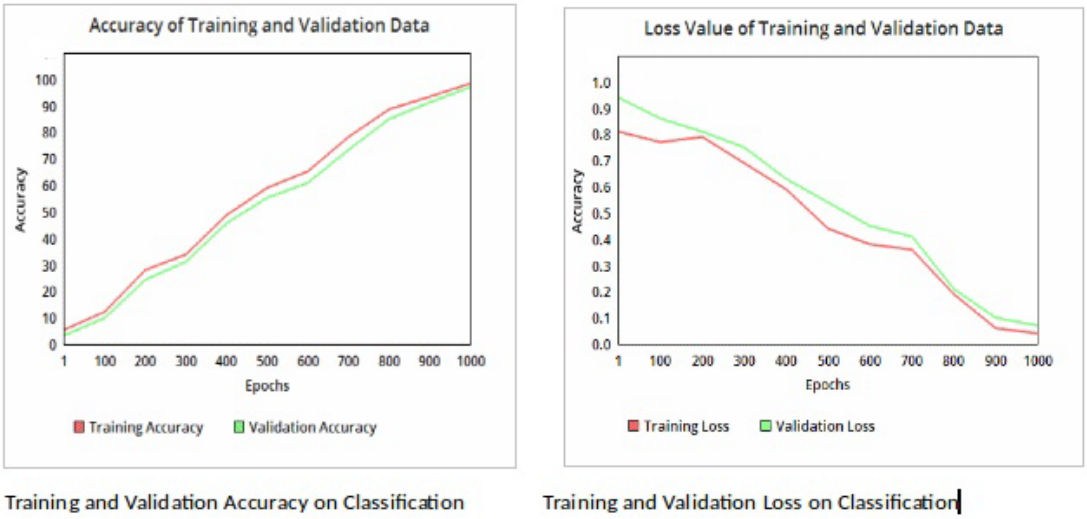


Figure 4.15: Training & validation of accuracy and loss for proposed classifier

In this proposed classifier, features of two pre-trained architectures which are VGG-16 and AlexNet are extracted, and later convolution of these features is implemented. Both architectures are pre-trained on the ImageNet dataset. To classify our given dataset, the algorithm contains the linear classification model using the concatenated convolution features of both architectures using the transfer learning approach. More details regarding this proposed hybrid classifier is already presented in Section 3.6.3. A confusion matrix for the classification performance is shown in Table 4.17. Details of confusion matrix calculation and performance measures used here for evaluation, accuracy, precision, recall, F1 score, and FNR is presented in Equations 2.12 to 2.16 respectively, in Section 2.8.

Table 4.17: Confusion matrix for classification using proposed hybrid classifier

ISIC	Melanoma	Nevus	SK
Melanoma	98.2%	1.8%	0%
Nevus	2.6%	97.4%	0%
SK	0%	1.0%	99.0%

This proposed classifier able to differentiate between melanoma and SK skin lesion class every time it is tested which became possible after the feature concatenation of these two classes through the proposed classifier, as presented in Table 4.17. Accuracies presented in Table 4.17 showed high similarities among melanoma and nevus skin lesion classes and the same was observed from images available in the ISIC dataset where the physical appearance of melanoma and nevus is comparable but overall reasonable accuracies are achieved through the proposed classifier.

Table 4.18 illustrates the results of different experiments done on single classifiers, AlexNet, VGG-16, and the proposed hybrid classifier. It has been observed from Table 4.18 that our proposed hybrid architecture achieved 98.2% accuracy which is considerably better as compared to existing single architectures and the main reason for this improvement was feature concatenation in the proposed hybrid model. As precision is significantly improved in this Table 4.18 which means correct detection of melanoma class is improved. Also, FNR is reduced in Table 4.18 which means number of times, melanoma is classified as another skin lesion type, is reduced as well. Better performance is achieved in other performance measures as well provided in Table 4.18. These all are very encouraging results and show the superiority of the proposed hybrid network.

Table 4.18: Comparison of classification between single and proposed hybrid classifier

Classifier	Accuracy (%)	FNR (%)	Precision	Recall	F1-Score
AlexNet	93.2	6.8	93.0	93.4	93.1
VGG-16	94.5	5.5	94.2	94.8	94.4
Hybrid of AlexNet and VGG-16	98.2	1.8	98.0	98.5	98.2

Moreover, to prove the performance of the proposed classification algorithm, a comparison is made with existing techniques from the literature. These existing techniques mainly belong to machine and deep learning paradigms and results have been presented in Table 4.19. Performance measures used in Table 4.19 are already presented in Section 2.8. Yap et al. (2018) present a simplified ResNet-50 model where fully connected layers are removed at the end of the network, but the accuracy of our proposed hybrid classifier is better than this technique as presented in Table 4.19. An SVM-based radial function is used by Khan et al. (2019) and this technique secured better response time as compared to our proposed classifier but at the cost of reduced accuracy. A recurrent neural network-based model is utilized for classification by Alom et al. (2019) but our proposed classifier exhibited significantly better accuracy.

Table 4.19: Comparison of proposed classification algorithm with benchmark techniques

Paper	Accuracy (%)	TNR (%)	TPR (%)	Response Time(s)	Dataset
(M. A. Khan, Javed, et al., 2019)	96.5	97.0	96.5	16.46	ISIC
(Afza et al., 2019)	93.7	95.0	93.3	38.61	ISIC
(Yap et al., 2018)	72.1	-	-	-	ISIC
(Alom et al., 2019)	87.0	-	-	-	ISIC
Proposed Classifier	98.2	98.3	97.5	22.44	ISIC

The proposed classifier also exhibits much better performance concerning TNR and TPR as compared to benchmark methods presented in Table 4.19. It is worth mentioning here that the proposed architecture is simplified by ignoring fully connected layers of AlexNet and VGG-16, which reduced the overall complexity of this proposed hybrid classifier. Still encouraging results are achieved as compared to existing methods, even after simplification, as presented in Table 4.19. In Figure 4.16, true labels and output predicted labels through the proposed classifier on the ISIC dataset for sample images, are presented. Hence all the results presented in this section proved that the proposed hybrid classifier performed reasonably well compared to existing benchmarks, which was our third research objective as well.

Figure 4.16: Classification results on ISIC dataset: (a) true label, (b) predicted label

4.5 Summary

This chapter presented the results of all the proposed algorithms for skin lesion pre-processing, segmentation, and classification. There is no separate section usually for the results of feature extraction because features are used for classification for which details are already provided in the respective section of classification results. In the pre-processing stage, results for proposed algorithms using ADF, local contrast enhancement, and haze reduction in CIELAB colour space and black-hat morphology with total variation for hair removal are presented. Visual results provided in these sections confirmed that better performance in pre-processing is achieved with the proposed algorithms. Visual and

numerical results for all proposed segmentation algorithms are provided in this chapter. Segmentation performed through the proposed boundary condition-based method on the ISIC dataset achieved 96.6% accuracy. The second proposed algorithm which is based on Delaunay triangulation clustering and PSO, achieved segmentation accuracy of 96.8% and 92.1% for the ISIC and PH2 datasets respectively. In the third proposed algorithm, segmentation is implemented through split and merge and contextual encoding methods. This algorithm achieved segmentation accuracy of 97.8% and 94.2% for the ISIC and PH2 datasets respectively. Numerical results for all three classification algorithms are provided, which were evaluated on the ISIC dataset. In the first proposed algorithm, classification is implemented through modified CNN. This algorithm achieved 96% accuracy. The second proposed algorithm for classification is based on boost ensemble learning, implemented through SVM and ANN. It used statistical features based on colour, pattern, and shape. This algorithm achieved 97.9% of classification accuracy. In the third proposed algorithm for classification, hybrid features are extracted using Alex-Net and VGG-16. Here classification is implemented through a transfer learning approach. This algorithm achieved 98.2% of classification accuracy. Detailed performance measures are used to compare the results of proposed algorithms with renowned benchmarks as well.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Introduction

Skin cancer which is one of the deadliest cancer types triggered by the uncontrolled growth of cells that produce pigments in human skin. In general, skin cancer can be either malignant or benign. Melanoma belongs to the malignant type of skin cancer whereas nevus and SK belong to the benign skin cancer type. American Cancer Society has provided the statistics that almost 197,700 new cases of skin cancer diagnosed in 2022, out of which 97,920 were belonged to benign skin lesion and 99,780 were malignant melanoma. Skin cancer is expected to be the fifth most diagnosed cancer for both males and females in 2022 (Siegel et al., 2022). Melanoma skin cancer has been on the rise for the past few years. Melanoma can be cured easily with surgical procedures if it is diagnosed at earlier stages. But in later stages, this cancer can penetrate the skin and spread through other organs which may result in death of the patient. Fatality rates caused by melanoma skin lesions are the highest among skin cancer patients.

Dermoscopy is a non-invasive process through which dermatologists diagnose the actual skin lesion type and cure accordingly. Dermatologists need to screen every patient thoroughly for possible infection. But due to the similarity among skin lesion types and the presence of artifacts in dermoscopic images, accurate diagnosis is very challenging even for experienced dermatologists. Therefore, a computer-based diagnosis is required to overcome these challenges. Image processing techniques can help dermatologists by automatically filtering melanoma-affected patients.

A complete process of skin lesion detection consists of four steps, pre-processing, segmentation, feature extraction, and classification. In pre-processing stage different anomalies and artifacts, including hairs needs to be removed to highlight the skin lesion from dermoscopic images. In the segmentation step, foreground which is the skin lesion in the image, is separated from the background to detect the boundary of the skin lesion. Once a segmented skin lesion is acquired then colour, shape, and size-based features are extracted. In the last step, these features are fed to a classifier that predicts skin lesion type as melanoma, nevus, or SK. We have incorporated statistical, machine, and deep learning methods to propose algorithms for these four stages of skin lesion detection.

In this chapter, we present thesis contributions made in each of the four steps of skin lesion detection: pre-processing, segmentation, feature extraction, and classification. Limitations of the proposed work and future directions in this area of research are also presented in this chapter.

5.2 Contributions

Our main objective was to implement pre-processing (including hair removal), segmentation, feature extraction, and classification of skin lesion from dermoscopic images by incorporating statistical, machine, and deep learning-based techniques. We have implemented the combination, ensemble, and hybrid of existing methods in the proposed algorithms. Also, we simplified existing CNN-based networks like AlexNet and VGG-16, after parameter manipulation and layer reduction to perform multi-class classification of dermoscopic images into melanoma, nevus, and SK. Here is the summary of the contributions.

5.2.1 Pre-processing Including Hair Removal

Pre-processing is applied to remove low contrast, variable background, illumination variation like shadows, and mainly hair. Following are the contributions in the pre-processing stage addressing research question 1 and research objective 1.

i. In the first proposed algorithm, pre-processing is implemented through ADF then basic transformation and Gaussian filtering are applied to enhance the skin lesion for subsequent steps.

ii. In the second proposed algorithm, hair and artifact removal is implemented through the proposed six steps in which colour space transformation and a weighted matrix are utilized initially. Later it incorporates black-hat morphological processing, adaptive thresholding, and total variation inpainting techniques to remove the hairs from dermoscopic images.

iii. In the third proposed algorithm, local contrast enhancement is applied to the images to improve the contrast, making the skin lesion more prominent. Following this, a statistical method of haze reduction is applied to the CIELAB colour space of images, which uses mean and standard deviation to enhance the foreground of dermoscopic images.

5.2.2 Segmentation

In proposed algorithms, segmentation is implemented through the integration of machine and deep learning-based method. Also, it has been implemented through the proposed boundary condition-based model. Following are the contributions in the segmentation stage addressing research question 2 and research objective 2.

i. In the first proposed algorithm, adaptive thresholding is applied to segment the skin lesion of the image by binarizing it. The morphological operations are applied for further enhancement and segmentation is implemented in the final stage through the proposed boundary condition-based model. This model is implemented through two proposed Equations (3.4 and 3.5). This algorithm is tested over the ISIC dataset with 96% of segmentation accuracy.

ii. In the second algorithm, segmentation is implemented by combining Delaunay triangulation-based clustering with PSO. Here spatial clustering is introduced with Delaunay triangulation. Further, HSV value-based thresholding is incorporated in the second stage of segmentation. This enhanced implementation of Delaunay triangulation achieved better results on ISIC and PH2 datasets by achieving accuracy of 96.8% and 92.1% respectively.

iii. In the third proposed algorithm for segmentation, we extract the features through a machine learning approach split & merge, and a deep learning-based context encoding network along with attention mechanism. In the first stage of the algorithm, feature and region extraction is performed using split & merge and context encoding network separately. Subsequently, the segmentation results are combined using IoU to produce the final segmented image. Since both techniques have origins in the region growing methods and it is assumed that combining these will produce more favourable results as compared to existing methods. Segmentation accuracy of 97.8% and 96.7% is achieved for ISIC and PH2 datasets respectively.

5.2.3 Feature Extraction

In this step, statistical and deep features were extracted. These features were later fed to the proposed classifiers. Deep features were also utilized in the third proposed

segmentation algorithm. Following are the contributions in the feature extraction step addressing research objective 3.

i. In the first proposed algorithm, the fusion of statistical features based on colour, pattern, and shape is implemented through the proposed values, for the best division of skin lesion images. Methods employed in this proposed algorithm are simple and less resource-hungry yet achieved optimal results.

ii. In the second proposed algorithm, deep features are extracted using two pre-trained architectures, AlexNet and VGG-16 through a transfer learning approach. This proposed approach only picks the convolution features from the last convolution layer of ALexNet and VGG-16 while discarding the remaining layers. This reduces the computational cost significantly while simplifying the overall architecture.

5.2.4 Classification

In the research, the classification step was implemented through an ensemble of machine learning-based methods. It has been implemented through modified neural networks with a reduced number of parameters and layers. Following are the contributions in the classification stage addressing research question 3 and research objective 3.

i. The first proposed classification algorithm used a CNN-based architecture with reduced complexity in terms of network parameters, filters, and number of layers. The proposed architecture utilizes only 60K parameters, 240 filters, 3 convolution layers, and 3 pooling layers which are significantly less as compared to the existing architectures such as AlexNet, VGG-16, and VGG-19. This proposed algorithm implements multi-class classification achieving 96% of accuracy on the ISIC dataset.

ii. In the second proposed algorithm, fused colour, pattern, and shape-based features, which were extracted earlier, are given as input. SVM and ANN are incorporated here through proposed ensemble learning where SVM is used as an initial classifier and ANN as a final classifier to learn the patterns of different skin lesion class features. This proposed algorithm implements multi-class classification and achieved 97.9% of accuracy on the ISIC dataset.

iii. In the third proposed algorithm, deep features which were extracted using AlexNet and VGG-16 through the transfer learning approach, are used for classification. Here proposed architect is simplified by ignoring separate fully connected layers of AlexNet and VGG-16 while including only 3 fully connected layers at the end. This classifier is termed as a hybrid because features from both architectures are concatenated with additional fully connected layers to perform the classification. This proposed algorithm implements multi-class classification and achieved 98.2% of accuracy on the ISIC dataset.

5.3 Limitations in Proposed Work

Although the proposed algorithms performed well as compared to existing methods still there are some areas in which improvement can be done. In the proposed algorithm for hair removal, we implemented the inpainting process through total variation but for some images, inpainting produced blocky artifacts and this step made the whole algorithm computationally expensive. The split and merge method used in the third proposed segmentation algorithm caused over-segmentation due to the image being divided into a higher number of regions during the splitting process. Although, this issue is resolved by adjusting the threshold variation parameter but still it can be improved through more research. In the third proposed hybrid classifier, response time was higher as compared to

existing methods because it utilized two feature streams from AlexNet and VGG-16. In the proposed algorithms where deep learning was used, although overfitting was not a great concern, because of the proposed reduction in the number of layers, parameters, and filters still this problem requires more research in designing architectures.

5.4 Future Works

There is no doubt that more research requires in this area of biomedical imaging as skin cancer is spreading rapidly all over the world because UV rays are becoming more dangerous due to the weakening of the ozone layer. In the proposed algorithms, it is observed that better results were achieved by combining methods from different paradigms, such as statistical, machine, and deep learning. Therefore, more combinations, fusions, or ensembles from these paradigms can be tried for skin lesion detection. In the future, larger datasets are expected to be available publicly. Training of algorithms on these datasets will help in achieving more reliable and accurate results. As spatial features of skin lesion such as shape, colour, size, etc. are very critical for accurate decision-making therefore in the future 3D modeling of skin lesion can be implemented which of course would be more helpful in decision-making. An automatic decision support system can be integrated with the existing skin lesion detection system. This system can recommend the required surgery decision to doctors by assessing the size, shape, colour, lesion type, fatality rate, and severity of the lesion. Also, more improvement can be done in the proposed algorithms by incorporating different statistical, machine, or deep learning methods. For example, image inpainting in the proposed hair removal algorithm can be integrated with other diffusion or patch-based methods.

5.5 Summary

In this chapter, we described detailed contributions of this work, limitations of the proposed algorithms, and future directions in this area of research. This work primarily contributed to four stages of skin lesion detection, which are pre-processing, segmentation, feature extraction, and classification. Algorithms for each step are proposed while implying modifications to existing statistical, machine, and deep learning methods. These modifications are applied, mostly to reduce the complexity of the overall framework. Researchers in the field of biomedical image processing will find the work presented in this thesis of great interest.

REFRENCES

Abayomi-Alli, O. O., Damaševičius, R., Misra, S., Maskeliūnas, R., & Abayomi-Alli, A. (2021). Malignant skin melanoma detection using image augmentation by over sampling in nonlinear lower-dimensional embedding manifold. *Turkish Journal of Electrical Engineering and Computer Sciences*, *29*(8), 2600–2614.

Abbas, Q., Garcia, I. F., Emre Celebi, M., & Ahmad, W. (2013). A feature preserving hair removal algorithm for dermoscopy images. *Skin Research and Technology*, *19*(1), 27–36.

Adobe Stock. (2022). *UVB and UVA explained. Filtering of rays Sun Protection. Penetration into the human skin. Stock Vector | Adobe Stock*. Adobe Systems Software Ireland Limited: https://stock.adobe.com/sa/images/uvb-and-uva-explained-filtering-of-rays-sun-protection-penetration-into-the-human-skin/196524870

Afroz, A., Zia, R., Garcia, A. O., Khan, M. U., Jilani, U., & Ahmed, K. M. (2022). Skin lesion classification using machine learning approach: A survey. *2022 Global Conference on Wireless and Optical Technologies, GCWOT 2022*. 1-8. IEEE.

Afza, F., Khan, M. A., Sharif, M., & Rehman, A. (2019). Microscopic skin laceration segmentation and classification: A framework of statistical normal distribution and optimal feature selection. *Microscopy Research and Technique*, *82*(9), 1471–1488.

Ahmad, M. B., & Choi, T. S. (1999). Local threshold and boolean function based edge detection. *IEEE Transactions on Consumer Electronics*, *45*(3), 674–679.

Alamdar, F., & Keyvanpour, M. R. (2011). A new Colour feature extraction method based on quadhistogram. *Procedia Environmental Sciences*, *10*, 777–783.

Aldwgeri, A., & Abubacker, N. F. (2019). Ensemble of deep convolutional neural network for skin lesion classification in dermoscopy Images. In *International Visual Informatics Conference: Vol. 11870 LNCS* (pp. 214–226). Springer.

Almansour, E., & Netw, M. J. (2016). Classification of dermoscopic skin cancer images using color and hybrid texture features. *IJCSNS International Journal of Computer Science and Network Security*, *16*(4), 135–139.

Al-masni, M. A., Al-antari, M. A., Choi, M. T., Han, S. M., & Kim, T. S. (2018). Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Computer Methods and Programs in Biomedicine*, *162*(7), 221–231.

Al-masni, M. A., Kim, D. H., & Kim, T. S. (2020). Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. *Computer Methods and Programs in Biomedicine*, *190*(1), 105-351.

Alom, M. Z., Aspiras, T., Taha, T. M., & Asari, V. K. (2019). Skin cancer segmentation and classification with NABLA-N and inception recurrent residual convolutional networks. *ArXiv Preprint ArXiv:1904.11126*. http://arxiv.org/abs/1904.11126

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1), 1–74.

Amelard, R., Wong, A., & Clausi, D. A. (2012). Extracting morphological high-level intuitive features (HLIF) for enhancing skin lesion classification. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 4458–4461. IEEE.

Ameri, A. (2020). A deep learning approach to skin cancer detection in dermoscopy images. *Journal of Biomedical Physics & Engineering*, *10*(6), 801-834.

Anand, V., Gupta, S., & Koundal, D. (2022). Skin disease diagnosis: challenges and opportunities. *In Proceedings of Second Doctoral Symposium on Computational Intelligence: DoSCI 2022* (pp. 449-459). Springer Singapore.

Anand, V., Gupta, S., Koundal, D., Nayak, S. R., Barsocchi, P., & Bhoi, A. K. (2022). Modified U-NET architecture for segmentation of skin lesion. *Sensors 2022, 22*(3)*,* 867-895.

Anjum, M. A., Amin, J., Sharif, M., Khan, H. U., Malik, M. S. A., & Kadry, S. (2020). Deep semantic segmentation and multi-class skin lesion classification based on convolutional neural network. *IEEE Access*, *8*(1), 129668–129678.

Aribisala, B. S., & Claridge, E. (2005). A border irregularity measure using a modified conditional entropy method as a malignant melanoma predictor. *In International Conference Image Analysis and Recognition* (pp. 914-921). Berlin, Heidelberg: Springer Berlin Heidelberg.

Arora, G., Dubey, A. K., Jaffery, Z. A., & Rocha, A. (2022). Bag of feature and support vector machine based early diagnosis of skin cancer. *Neural Computing and Applications*, *34*(11), 8385–8392.

Ashwini, A., & Murugan, S. (2020). Automatic skin tumour segmentation using prioritized patch based region–a novel comparative technique. *IETE Journal of Research, 69*(1), 137-148.

Aslam, A., & Curry, E. (2021). Investigating response time and accuracy in online classifier learning for multimedia publish-subscribe systems. *Multimedia Tools and Applications*, *80*(9), 13021–13057.

Baig, R., Bibi, M., Hamid, A., Kausar, S., & Khalid, S. (2020). Deep learning approaches towards skin lesion segmentation and classification from dermoscopic images - a review. *Current Medical Imaging Formerly Current Medical Imaging Reviews*, *16*(5), 513–533.

Balafar, M. A. (2014). Gaussian mixture model based segmentation methods for brain MRI images. *Artificial Intelligence Review*, *41*(1), 429–439.

Bansal, P., Vanjani, A., Mehta, A., Kavitha, J. C., & Kumar, S. (2022). Improving the classification accuracy of melanoma detection by performing feature selection using binary Harris hawks optimization algorithm. *Soft Computing*, *26*(17), 8163–8181.

Barın, S., & Güraksın, G. E. (2022). An automatic skin lesion segmentation system with hybrid FCN-ResAlexNet. *Engineering Science and Technology, an International Journal*, *34*(5), 101-174.

Bassel, A., Abdulkareem, A. B., Alyasseri, Z. A. A., Sani, N. S., Mohammed, H. J., Bassel, A., Abdulkareem, A. B., Abdi, Z., Alyasseri, A., Sani, N. S., & Mohammed, H. J. (2022). Automatic malignant and benign skin cancer classification using a hybrid deep learning approach. *Diagnostics 2022, 12*(10), 24-72.

Bazulin, E. G. (2021). Application of adaptive anisotropic diffusion filter to improve the reflector image quality when performing ultrasonic non destructive testing. *Russian Journal of Nondestructive Testing*, *57*(5), 343–351.

Bechelli, S., & Delhommelle, J. (2022). Machine learning and deep learning algorithms for skin cancer classification from dermoscopic images. *Bioengineering 2022, 9*(3), 97-110.

Benyahia, S., Meftah, B., & Lézoray, O. (2022). Multi-features extraction based on deep learning for skin lesion classification. *Tissue and Cell*, *74*(2), 101-117.

Berseth, M. (2017). ISIC 2017 - Skin lesion analysis towards melanoma detection. *ArXiv:1703.00523*. http://arxiv.org/abs/1703.00523

Bevan, P. J., & Atapour-Abarghouei, A. (2022). Detecting melanoma fairly: skin tone detection and debiasing for skin lesion classification. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *13542 LNCS*, 1–11. https://doi.org/10.1007/978-3-031-16852-9_1/Cover.

Bhadauria, H. S., & Dewal, M. L. (2013). Medical image denoising using adaptive fusion of curvelet transform and total variation. *Computers & Electrical Engineering*, *39*(5), 1451–1460.

Bi, L., Feng, D. D., Fulham, M., & Kim, J. (2021). Multi label classification of multi-modality skin lesion via hyper-connected convolutional neural network. *Pattern Recognition*, *107*(3), 107-502.

Bi, L., Fulham, M., & Kim, J. (2022). Hyper-fusion network for semi-automatic segmentation of skin lesions. *Medical Image Analysis*, *76*(9), 102-334.

Bi, L., Kim, J., Ahn, E., & Feng, D. (2017). Automatic skin lesion analysis using large-scale dermoscopy images and deep residual networks. In *arXiv preprint arXiv:1703.04197*. arXiv. https://challenge.kitware.com/#phase/584b08eecad3a51cc66c8e1f

Blaschke, T., Burnett, C., & Pekkarinen, A. (2004). *Image Segmentation Methods for Object-based Analysis and Classification*. 211–236. Dordrecht: Springer Netherlands.

Butt, M. O., Al-Asad, J. F., Khan, A. H., & Awang Iskandar, D. N. F. (2019). Ultrasound image denoising using orthogonal decomposition in frequency domain. *2019 IEEE 9th International Conference on System Engineering and Technology, ICSET 2019 - Proceeding*, 349–353.

Carrera, C., Segura, S., Aguilera, P., Scalvenzi, M., Longo, C., Barreiro, A., Broganelli, P., Cavicchini, S., Llambrich, A., Zaballos, P., Thomas, L., Malvehy, J., Puig, S., & Zalaudek, I. (2017). Dermoscopic clues for diagnosing melanomas that resemble seborrheic keratosis. *JAMA Dermatology*, *153*(6), 544–551.

Cavalcanti, P. G., & Scharcanski, J. (2011). Automated prescreening of pigmented skin lesions using standard cameras. *Computerized Medical Imaging and Graphics*, *35*(6), 481–491.

Cavalcanti, P. G., Scharcanski, J., & Lopes, C. B. O. (2010). Shading attenuation in human skin color images. In *Advances in Visual Computing: 6th International Symposium, ISVC 2010, Las Vegas, NV, USA, November 29-December 1, 2010. Proceedings, Part I 6.* (pp. 190-198). Springer Berlin Heidelberg.

Cavalcanti, P. G., Scharcanski, J., Martínez, C. E., & Di Persia, L. E. (2014). Segmentation of pigmented skin lesions using Non-negative Matrix Factorization. *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, 72–75.

Celebi, M. E., Kingravi, H. A., Iyatomi, H., Aslandogan, Y. A., Stoecker, W. V., Moss, R. H., Malters, J. M., Grichnik, J. M., Marghoob, A. A., Rabinovitz, H. S., & Menzies, S.

W. (2008). Border detection in dermoscopy images using statistical region merging. *Skin Research and Technology*, *14*(3), 347–353.

Celebi, M., Iyatomi, H., & Stoecker, W. (2008). Automatic detection of blue-white veil and related structures in dermoscopy images. *Computerized Medical Imaging and Graphics*, *32*(8), 670–677.

Chang, D. C., & Wu, W. R. (1998). Image contrast enhancement based on a histogram transformation of local standard deviation. *IEEE Transactions on Medical Imaging*, *17*(4), 518–531.

Chaudhuri D, A. A. (2010). Split-and-merge procedure for image segmentation using bimodality detection approach. *Defence Science Journal*, *60*(3), 290–301.

Chi, H. N. H. (2010). Segmentation of light and dark hair in dermoscopic images: a hybrid approach using a universal kernel designing radiology workstations for softcopy reading view project eye hand coordination of surgeons view project segmentation of light and dark hair. *Medical Imaging 2010: Image Processing.* 1436-1443. SPIE.

Choudhary, P., Singhai, J., & Yadav, J. S. (2021). Curvelet and fast marching method-based technique for efficient artifact detection and removal in dermoscopic images. *International Journal of Imaging Systems and Technology*, *31*(4), 2334–2345.

Commission Internationale de l'Éclairage. (1978). Recommendations on uniform color spaces, color-difference equations, psychometric color terms. *CIE Publ. 15 Suppl. 2Commission Internationale de l'Éclairage Paris, France*, 135–178.

Correa-Tome, F. E., & Sanchez-Yanez, R. E. (2015). Integral split-and-merge methodology for real-time image segmentation. *Journal of Electronic Imaging*, *24*(1), 007–013.

Dai, D., Dong, C., Xu, S., Yan, Q., Li, Z., Zhang, C., & Luo, N. (2022). Ms RED: A novel multi-scale residual encoding and decoding network for skin lesion segmentation. *Medical Image Analysis*, *75*, 102–293. https://doi.org/10.1016/J.MEDIA.2021.102293

Das, J. B. A., Mishra, D., Das, A., Mohanty, M. N., & Sarangi, A. (2022). Skin cancer detection using machine learning techniques with ABCD features. *2nd Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology, ODICON 2022*. 1-6. IEEE.

Dermatology, I. L.-B. J., & Leigh. (2014). Progress in skin cancer: the UK experience. *British Journal of Dermatology*, *171*(3), 443–445.

Diepgen, T. L., & Mahler, V. (2002). The epidemiology of skin cancer. *British Journal of Dermatology*, *146*(61), 1–6.

Ding, L., & Goshtasby, A. (2001). On the Canny edge detector. *Pattern Recognition*, *34*(3), 721–725.

Divya, D., & Ganeshbabu, T. R. (2020). Fitness adaptive deer hunting-based region growing and recurrent neural network for melanoma skin cancer detection. *International Journal of Imaging Systems and Technology*, *30*(3), 731–752.

Dorj, U. O., Lee, K. K., Choi, J. Y., & Lee, M. (2018). The skin cancer classification using deep convolutional neural network. *Multimedia Tools and Applications*, *77*(8), 9909–9924.

Eedy, D. (2020). Dermatology: A specialty in crisis. *Clinical Medicine, Journal of the Royal College of Physicians of London*, *15*(6), 509–510.

El Abbadi, N. K., & Miry, A. H. (2013). Automatic segmentation of skin lesions using histogram thresholding. *Journal of Computer Science*, *10*(4), 632–639.

El-Wahab, B. S., S. A., Wahba, M. A., Mansour, D. E. A., Hodeib, A. A. E., Khedr, R. A. E. G., & Hassan, G. F. (2022). Cascaded hough transform based hair mask generation and harmonic inpainting for automated hair removal from dermoscopy images. *Diagnostics 2022, 12*(12), 30-40.

Fan, H., Xie, F., Li, Y., Jiang, Z., & Liu, J. (2017). Automatic segmentation of dermoscopy images using saliency combined with Otsu threshold. *Computers in Biology and Medicine*, *85*(7), 75–85.

Feng, R., Zhuo, L., Li, X., Yin, H., & Wang, Z. (2022). BLA-Net:Boundary learning assisted network for skin lesion segmentation. *Computer Methods and Programs in Biomedicine*, *226*(7), 107-190.

Filali, Y., Ennouni, A., Sabri, M. A., & Aarab, A. (2018). A study of lesion skin segmentation, features selection and classification approaches. *International Conference on Intelligent Systems and Computer Vision, ISCV 2018*, *2018 May*, 1–7.

Fred Agarap, A. M. (2018). *Deep Learning using Rectified Linear Units (ReLU)*. https://arxiv.org/abs/1803.08375v2

Gao, M., Shi, G., & Li, S. (2018). Online prediction of ship behavior with automatic identification system sensor data using bidirectional long short-term memory recurrent neural network. *Sensors (Switzerland)*, *18*(12), 11-42.

Garcia Ugarriza, L., Saber, E., Vantaram, S. R., Amuso, V., Shaw, M., & Bhaskar, R. (2009). Automatic image segmentation by dynamic region growth and multiresolution merging. *IEEE Transactions on Image Processing*, *18*(10), 2275–2288.

Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron) a review of applications in the atmospheric sciences. *Atmospheric Environment*, *32*(14), 2627–2636.

Goldstein, A., & Tucker, M. (2013). Dysplastic nevi and melanoma. *Cancer epidemiology, biomarkers & prevention, 22*(4), 528-532.

Gouda, W., Sama, N. U., Al-Waakid, G., Humayun, M., & Jhanjhi, N. Z. (2022). Detection of skin cancer based on skin lesion images using deep learning. *Healthcare*, *10*(7), 1007-1183.

Gulzar, Y., & Khan, S. A. (2022). Skin lesion segmentation based on vision transformers and convolutional neural networks a comparative study. *Applied Sciences 2022, 12*(12), 59-90.

Gupta, A., Issac, A., Dutta, M. K., & Hsu, H. H. (2017). Adaptive thresholding for skin lesion segmentation using statistical parameters. *Proceedings - 31st IEEE International Conference on Advanced Information Networking and Applications Workshops, WAINA 2017*, 616–620.

H Thanh, D. N., Thi Thanh, L., Sergey, D., Surya Prasath, V. B., & Quang San, N. (2019). Adaptive thresholding segmentation method for skin lesion with normalized color channels of NTSC and YCbCr. *14th International Conference on Pattern Recognition and Information Processing*, 21–23.

Haenssle, H. A., Winkler, J. K., Fink, C., Toberer, F., Enk, A., Stolz, W., Deinlein, T., Zukervar, P. (2023). Skin lesions of face and scalp – Classification by a market-approved convolutional neural network in comparison with 64 dermatologists. *European Journal of Cancer*, *144*(3), 192–199.

Hasan, M. K., Ahamad, M. A., Yap, C. H., & Yang, G. (2023). A survey, review, and future trends of skin lesion segmentation and classification. *Computers in Biology and Medicine*, *155*(7), 106-624.

Hasan, M. K., Dahal, L., Samarakoon, P. N., Tushar, F. I., & Martí, R. (2020). DSNet: Automatic dermoscopic skin lesion segmentation. *Computers in Biology and Medicine*, *120*(2), 103-738.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. IEEE.

Hearn, J. (2008). *Competitive Medical Image Segmentation With The Fast Marching Method*. Case Western Reserve University.

Hekler, A., Utikal, J., & Enk, A. (2019). Superior skin cancer classification by the combination of human and artificial intelligence. *European Journal of Cancer*, *120*(9), 114–121.

Hosny, K. M., Kassem, M. A., & Foaud, M. M. (2019). Classification of skin lesions using transfer learning and augmentation with Alex-net. *PLOS ONE*, *14*(5), 217-293.

Hu, K., Niu, X., Liu, S., Zhang, Y., Cao, C., Xiao, F., Yang, W., & Gao, X. (2019). Classification of melanoma based on feature similarity measurement for codebook learning in the bag-of-features model. *Biomedical Signal Processing and Control*, *51*(2), 200–209.

ISIC Archive. (2016). *ISIC Archive*. The Shore Family Fund. https://www.isic-archive.com/#!/topWithHeader/wideContentTop/main.

Jahanifar, M., Tajeddin, N. Z., Gooya, A., & Asl, B. M. (2017). Segmentation of lesions in dermoscopy images using saliency map and contour propagation. In *arXiv*.

Jain, A., Rao, A. C. S., Jain, P. K., & Abraham, A. (2022). Multi-type skin diseases classification using OP-DNN based feature extraction approach. *Multimedia Tools and Applications*, *81*(5), 6451–6476.

Jaisakthi, S. M., Mirunalini, P., & Aravindan, C. (2018). Automated skin lesion segmentation of dermoscopic images using GrabCut and kmeans algorithms. *IET Computer Vision*, *12*(8), 1088–1095.

Javadi, N., & Soltanizadeh, H. (2021). Automated detection, 3D position of facial skin lesions using genetic algorithm and Kinect camera. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 10*(1), 48–54.

Jenkins, R. W., & Fisher, D. E. (2021). Treatment of advanced melanoma in 2020 and beyond. *Journal of Investigative Dermatology*, *141*(1), 23–31.

Jeyakumar, J. P., Jude, A., Priya, A. G., & Hemanth, J. (2022). A Survey on computer aided intelligent methods to identify and classify skin cancer. *Informatics 2022, 9*(4), 40-99.

Johr, R. H. (2002). Dermoscopy: Alternative melanocytic algorithms - The ABCD rule of dermatoscopy, menzies scoring method, and 7-point checklist. *Clinics in Dermatology*, *20*(3), 240–247.

Joseph, S., & Olugbara, O. O. (2022). Preprocessing effects on performance of skin lesion saliency segmentation. *Diagnostics 2022, 12*(2), 344-384.

Kakadiaris, I. A., Vrigkas, M., Yen, A. A., Kuznetsova, T., Budoff, M., & Naghavi, M. (2018). Machine learning outperforms ACC/AHA CVD risk calculator in MESA. *Journal of the American Heart Association*, *7*(22), 118-197.

Kassianos, A. P., Emery, J. D., Murchie, P., & Walter, F. M. (2015). Smartphone applications for melanoma detection by community, patient and generalist clinician users: A review. *British Journal of Dermatology*, *172*(6), 1507–1518.

Khan, A. H., Awang Iskandar, D. N. F., Al-Asad, J. F., El-Nakla, S., & Alhuwaidi, S. A. (2021). Statistical feature learning through enhanced delaunay clustering and ensemble classifiers for skin lesion segmentation and classification. *Journal of Theoretical and Applied Information Technology*, *15*(5), 790-841.

Khan, A. H., Awang Iskandar, D. N. F., Al-Asad, J. F., Mewada, H., & Sherazi, M. A. (2022). Ensemble learning of deep learning and traditional machine learning approaches for skin lesion segmentation and classification. *Concurrency and Computation: Practice and Experience*, *34*(13), 690-724.

Khan, M. A., Javed, M. Y., Sharif, M., Saba, T., & Rehman, A. (2019). Multi-model deep neural network based features extraction and optimal selection approach for skin lesion classification. *2019 International Conference on Computer and Information Sciences, ICCIS 2019*. 1-7. IEEE.

Khan, M. A., Sharif, M. I., Raza, M., Anjum, A., Saba, T., & Shad, S. A. (2019). Skin lesion segmentation and classification: A unified framework of deep neural network features fusion and selection. *Expert Systems*, *39*(7), 124-197.

Kornilov, A. S., & Safonov, I. V. (2018). An overview of watershed algorithm implementations in open source libraries. *Journal of Imaging 2018, 4*(10), 123-176.

Korotkov, K., & Garcia, R. (2012). Computerized analysis of pigmented skin lesions: A review. *Artificial Intelligence in Medicine*, *56*(2), 69–90.

Kretzler, M. (2013). *Automated curved hair detection and removal in skin images to support automated melanoma detection*. Case Western Reserve University.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*(9), 1097-1105.

Kumar, C., & Kumar, V. (2023). Vegetable Plant Leaf Image Classification Using Machine Learning Models. *In Proceedings of Third International Conference on Advances in Computer Engineering and Communication Systems*, Singapore: Springer Nature Singapore, 31-45.

Kushimo, O. O., Salau, A. O., Adeleke, O. J., & Olaoye, D. S. (2023). Deep learning model to improve melanoma detection in people of color. *Arab Journal of Basic and Applied Sciences*, *30*(1), 92–102.

Lee, T., Ng, V., Gallagher, R., Coldman, A., & McLean, D. (1997). DullRazor: a software approach to hair removal from images. *Computers in Biology and Medicine*, *27*(6), 533–543.

Li, L. F., Wang, X., Hu, W. J., Xiong, N. N., Du, Y. X., & Li, B. S. (2020). Deep learning in skin disease image recognition: a review. *IEEE Access*, *8*(2), 208264–208280.

Li, T., Horton, R. M., Bader, D. A., Liu, F., Sun, Q., & Kinney, P. L. (2018). Long-term projections of temperature-related mortality risks for ischemic stroke, hemorrhagic stroke, and acute ischemic heart disease under changing climate in Beijing, China. *Environment International*, *112(9)*, 1–9.

Li, W., Joseph Raj, A. N., Tjahjadi, T., & Zhuang, Z. (2021). Digital hair removal by deep learning for skin lesion segmentation. *Pattern Recognition*, *117*(6), 107–994.

Lin, C. J., Lin, C. H., Wang, S. H., & Wu, C. H. (2019). Multiple convolutional neural networks fusion using improved fuzzy integral for facial emotion recognition. *Applied Sciences 2019, 9*(13), 25-93.

Liu, T., Fang, S., Zhao, Y., Wang, P., & Zhang, J. (2015). Implementation of training convolutional neural networks. *arvXiv preprint arXiv:1506.01195*.

Lv, T., Bai, C., & Wang, C. (2022). MDMLP: Image classification from scratch on small datasets with MLP. *arvXiv preprint arXiv:2205.14477*.

Ma, W., Li, N., Zhu, H., Jiao, L., Tang, X., Guo, Y., & Hou, B. (2022). Feature split merge enhancement network for remote sensing object detection. *IEEE Transactions on Geoscience and Remote Sensing*, *60*(1), 1-17.

Mahmoud, H., Abdel-Nasser, M., & Omer, O. A. (2018). Computer aided diagnosis system for skin lesions detection using texture analysis methods. *Proceedings of 2018 International Conference on Innovative Trends in Computer Engineering, ITCE 2018*, *2018-March*, 140–144. IEEE.

Malik, S., Akram, T., Ashraf, I., Rafiullah, M., Ullah, M., & Tanveer, J. (2022). A hybrid preprocessor DE-ABC for efficient skin lesion segmentation with improved contrast. *Diagnostics*, *12*(11), 1211-2625.

Marchetti, M. A., Liopyris, K., Dusza, S. W., Codella, N. C. F., Gutman, D. A., & Malvehy, J. (2021). Computer algorithms show potential for improving dermatologists' accuracy to diagnose cutaneous melanoma: Results of the International Skin Imaging Collaboration 2017. *Journal of the American Academy of Dermatology*, *82*(3), 622–627.

Marghoob, N. G., Liopyris, K., & Jaimes, N. (2019). Dermoscopy: A review of the structures that facilitate melanoma detection. *Journal of the American Osteopathic Association, 119*(6), 380-390.

Masood, A., & Al-Jumaily, A. A. (2013). Fuzzy c mean thresholding based level set for automated segmentation of skin lesions. *Journal of Signal and Information Processing*, *04*(03), 66–71.

Massone, C., Di Stefani, A., & Soyer, H. P. (2005). Dermoscopy for skin cancer detection. *Current Opinion in Oncology*, *17*(2), 147–153.

Mateen, M., Wen, J., Nasrullah, Song, S., & Huang, Z. (2018). Fundus image classification using VGG-19 architecture with PCA and SVD. *Symmetry*, *11*(1), 1109-1157.

Maulana, B., Wibowo, M., Putra, G., Pinem, J. G., Chasanah, U., & Nugroho, A. S. (2022). Image segmentation for aspergillus, cladosporium, and trichoderma fungus. *ACM International Conference Proceeding Series*, 111–116.

Mekhmoukh, A., & Mokrani, K. (2015). Improved Fuzzy C-Means based Particle Swarm Optimization (PSO) initialization and outlier rejection with level set methods for MR brain image segmentation. *Computer Methods and Programs in Biomedicine*, *122*(2), 266–281.

Mendonca, T., Ferreira, P. M., Marques, J. S., Marcal, A. R. S., & Rozeira, J. (2013). PH2 - A dermoscopic image database for research and benchmarking. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 5437–5440.

Moldovanu, S., Michis, F. A. D., Biswas, K. C., Culea-Florescu, A., & Moraru, L. (2021). Skin lesion classification based on surface fractal dimensions and statistical color

cluster features using an ensemble of machine learning techniques. *Cancers 2021, 13*(21), 52-67.

Montaha, S., Azam, S., Rakibul Haque Rafid, A. K. M., Islam, S., Ghosh, P., & Jonkman, M. (2022). A shallow deep learning approach to classify skin cancer using down-scaling method to minimize time and space complexity. *PLOS ONE*, *17*(8), 269-826.

Moussaoui, H., El Akkad, N., & Benslimane, M. (2023). A hybrid skin lesions segmentation approach based on image processing methods. *Statistics, Optimization & Information Computing*, *11*(1), 95–105.

Murugan, A., Nair, S. A. H., & Kumar, K. P. S. (2019). Detection of skin cancer using SVM, random forest and KNN classifiers. *Journal of Medical Systems*, *43*(8), 1–9.

Murugan, A., Nair, S. A. H., Preethi, A. A. P., & Kumar, K. P. S. (2022). Diagnosis of skin cancer using machine learning techniques. *Microprocessors and Microsystems*, *81*(2), 103-727.

Mustafa, W. A., & Yazid, H. (2017). Image enhancement technique on contrast variation: a comprehensive review. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *9*(3), 199–204.

Mutlag, W. K., Ali, S. K., Aydam, Z. M., & Taher, B. H. (2020). Feature extraction methods: a review. *Journal of Physics: Conference Series*, *1591*(1), 12-28.

Nasir, M., Attique Khan, M., Sharif, M., Lali, I. U., Saba, T., & Iqbal, T. (2018). An improved strategy for skin lesion detection and classification using uniform segmentation and feature selection based approach. *Microscopy Research and Technique*, *81*(6), 528–543.

Nik-Zainal, S., Alexandrov, L. B., Wedge, D. C., Van Loo, P., Greenman, C. D., & Stratton, M. R. (2012). Mutational processes molding the genomes of 21 breast cancers. *Cell*, *149*(5), 979–993.

Nisar, H., Ch'ng, Y. K., Chew, T. Y., Yap, V. V., Yeap, K. H., & Tang, J. J. (2013). A color space study for skin lesion segmentation. *ICCAS 2013 - 2013 IEEE International Conference on Circuits and Systems: "Advanced Circuits and Systems for Sustainability,"* 172–176. IEEE.

Otsu N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66.

Papageorgiou, V., Apalla, Z., Sotiriou, E., Papageorgiou, C., Lazaridou, E., Vakirlis, S., Ioannides, D., & Lallas, A. (2018). The limitations of dermoscopy: false-positive and false-negative tumours. *Journal of the European Academy of Dermatology and Venereology*, *32*(6), 879–888.

Pennisi, A., Bloisi, D. D., Nardi, D., Giampetruzzi, A. R., Mondino, C., & Facchiano, A. (2015). Melanoma detection using delaunay triangulation. *2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI)*, *2016-January*, 791–798. IEEE.

Pennisi, A., Bloisi, D. D., Nardi, D., Giampetruzzi, A. R., Mondino, C., & Facchiano, A. (2016). Skin lesion image segmentation using delaunay triangulation for melanoma detection. *Computerized Medical Imaging and Graphics*, *52*(4), 89–103.

Pereira, P. M. M., Fonseca-Pinto, R., Paiva, R. P., Assuncao, P. A. A., Tavora, L. M. N., Thomaz, L. A., & Faria, S. M. M. (2020). Dermoscopic skin lesion image segmentation

based on local binary pattern clustering: Comparative study. *Biomedical Signal Processing and Control*, *59*(7), 101-924.

Perona, P., & Malik, J. (1990). Scale space and edge detection using anisotropic diffusion. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, *12*(7), 56-205.

Perrinaud, A., Gaide, O., French, L. E., Saurat, J. H., Marghoob, A. A., & Braun, R. P. (2007). Can automated dermoscopy image analysis instruments provide added benefit for the dermatologist? A study comparing the results of three systems. *British Journal of Dermatology*, *157*(5), 926–933.

Pham, T. C., Tran, G. S., Nghiem, T. P., Doucet, A., Luong, C. M., & Hoang, V. D. (2019). A comparative study for classification of skin cancer. *Proceedings of 2019 International Conference on System Science and Engineering, ICSSE 2019*, 267–272.

Prasad, V. K., Bhattacharya, P., Maru, D., Tanwar, S., Verma, A., Singh, A., Tiwari, A. K., Sharma, R., Alkhayyat, A., Țurcanu, F. E., & Raboaca, M. S. (2022). Federated Learning for the Internet-of-Medical-Things: A Survey. *Mathematics 2023, 11*(1), 151-183.

Proksch, E., Brandner, J. M., & Jensen, J. M. (2008). The skin: an indispensable barrier. *Experimental Dermatology*, *17*(12), 1063–1072.

R Gordon. (2020). Skin cancer: an overview of epidemiology and risk factors in Seminars in oncology nursing. *Journal of Oncology Research and Treatment*, *8*(9), 208–252.

Rahman, M., Alpaslan, N., & Bhattacharya, P. (2017). Developing a retrieval based diagnostic aid for automated melanoma recognition of dermoscopic images. *Proceedings - Applied Imagery Pattern Recognition Workshop*. 1-7. IEEE.

Rahman, Md. M., Nasir, M. K., A-Alam, N., Khan, S. I., Band, S., Dehzangi, I., Beheshti, A., & Rokny, H. A. (2022). Hybrid feature fusion and machine learning approaches for melanoma skin cancer detection. *Preprints.org 2022,* 2022010258.

Ramella, G., & Sanniti Di Baja, G. (2011). Color histogram-based image segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6854 LNCS*(PART 1), 76–83.

Ramesh, K. K. D., Kiran Kumar, G., Swapna, K., Datta, D., & Suman Rajest, S. (2021). A Review of Medical Image Segmentation Algorithms. *EAI Endorsed Transactions on Pervasive Health and Technology*, *7*(27), 29–47.

Reddi, S. J., Kale, S., & Kumar, S. (2019). On the Convergence of Adam and Beyond. *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*. 209-245.

Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., & Soyer, H. P. (2021). A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Scientific Data 2021*, *8*(1), 1–8.

Rout, R., Parida, P., Alotaibi, Y., Alghamdi, S., & Khalaf, O. I. (2021). Skin lesion extraction using multiscale morphological local variance reconstruction based watershed transform and fast fuzzy c means clustering. *Symmetry 2021, 13*(11), 20-85.

Saba, T. (2021). Computer vision for microscopic skin cancer diagnosis using handcrafted and non-handcrafted features. *Microscopy Research and Technique*, *84*(6), 1272–1283.

Sabbaghi Mahmouei, S., Aldeen, M., Stoecker, W. V., & Garnavi, R. (2019). Biologically inspired quadtree color detection in dermoscopy images of melanoma. *IEEE Journal of Biomedical and Health Informatics*, *23*(2), 570–577.

Sagar, C., & Saini, L. M. (2017). Color channel based segmentation of skin lesion from clinical images for the detection of melanoma. *1st IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems, ICPEICES 2016*. 109-122. IEEE.

Saleh, A. H., & Menemencioğlu, O. (2023). A dynamic circular hough transform based iris segmentation. *Lecture Notes on Data Engineering and Communications Technologies*, *147*(6), 9–20.

Salma, W., & Eltrass, A. S. (2022). Automated deep learning approach for classification of malignant melanoma and benign skin lesions. *Multimedia Tools and Applications*, *81*(22), 32643–32660.

Santos, E. S. dos, de M S Veras, R., R T Aires, K., M B F Portela, H., Braz Junior, G., Santos, J. D., & Tavares, J. M. R. S. (2022). Semi-automatic segmentation of skin lesions based on super pixels and hybrid texture information. *Medical Image Analysis*, *77*(6), 102-363.

Sasikaladevi, N. (2022). Delaunay triangulation based intelligent system for the diagnosis of covid from the low radiation CXR images. *Journal of Ambient Intelligence and Humanized Computing 2022*, *109*(4), 1–10.

Sau, K., Maiti, A., & Ghosh, A. (2018). Preprocessing of skin cancer using anisotropic diffusion and sigmoid function. *Advances in Intelligent Systems and Computing*, *706*(11), 51–61.

Schmid-Saugeon, P., Guillod, J., & Thiran, J. (2003). Towards a computer-aided diagnosis system for pigmented skin lesions. *Computerized Medical Imaging and Graphics*, *27*(1), 65–78.

Sekniashvili, D. (2022). A morphology based methods of correcting printed and handwritten georgian text images. *Gospodarka i Innowacje.*, *20*(4), 1–6.

Sengupta, S., Mittal, N., & Modi, M. (2022). Colour space based thresholding for segmentation of skin lesion images. *International Journal of Biomedical Engineering and Technology*, *39*(4), 347-663.

Shetty, B., Fernandes, R., Rodrigues, A. P., Chengoden, R., Bhattacharya, S., & Lakshmanna, K. (2022). Skin lesion classification of dermoscopic images using machine learning and convolutional neural network. *Scientific Reports*, *12*(1), 81-34.

Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2022). Cancer statistics, 2022. *CA: A Cancer Journal for Clinicians*, *72*(1), 7–33.

Siegel, R. L., Miller, K. D., Wagle, N. S., & Jemal, A. (2023). Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, *73*(1), 17–48.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. 155-164.

Simonyan, K., & Zisserman, A. (2017). Very deep convolutional networks for large-scale image recognition. *ArXiv Preprint ArXiv:1409.1556*.

Socher, R., Li, L.-J., Li, K., Deng, J., Dong, W., & Fei-Fei, L. (2009). ImageNet: A large scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. IEEE.

Srivastava, V., Kumar, D., & Roy, S. (2022). A median based quadrilateral local quantized ternary pattern technique for the classification of dermatoscopic images of skin cancer. *Computers and Electrical Engineering*, *102*(9), 108-259.

Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks. *arXiv preprint arXiv:1909.09586*. https://arxiv.org/abs/1909.09586v1

Sultana, N. N., Mandal, B., & Puhan, N. B. (2018). Deep residual network with regularised fisher framework for detection of melanoma. *IET Computer Vision*, *12*(8), 1096–1104.

Sunsuhi, G. S., & Albin Jose, S. (2022). An adaptive eroded deep convolutional neural network for brain image segmentation and classification using Inception ResnetV2. *Biomedical Signal Processing and Control*, *78*(2), 103-863.

Tan, T. Y., Zhang, L., & Lim, C. P. (2020). Adaptive melanoma diagnosis using evolving clustering, ensemble and deep neural networks. *Knowledge-Based Systems*, *187*(12), 104-807.

Tao, Z., Yuncan, L., Huiling, L., Xinyu, Y., Xiaoyu, C., Tao, Z., Yuncan, L., Huiling, L., Xinyu, Y., & Xiaoyu, C. (2022). ResNet and its application to medical image processing: research progress and challenges. *Journal of Electronics & Information Technology*, *44*(1), 149–167.

Tarjoman, M., Fatemizadeh, E., & Badie, K. (2013). An implementation of a CBIR system based on SVM learning scheme. *Journal of Medical Engineering & Technology*, *37*(1), 43–47.

Thomsen, K., Christensen, A. L., Iversen, L., Lomholt, H. B., & Winther, O. (2022). Deep learning for diagnostic binary classification of multiple lesion skin diseases. *Frontiers in Medicine*, *7*(5), 604-631.

Toossi, M. T. B., Pourreza, H. R., Zare, H., Sigari, M.-H., Layegh, P., & Azimi, A. (2013). An effective hair removal algorithm for dermoscopy images. *Skin Research and Technology*, *19*(3), 230–235.

Veksler, O. (2003). Fast variable window for stereo correspondence using integral images. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *1*. 121-140. IEEE.

Wang, D., Pang, N., Wang, Y., & Zhao, H. (2021). Unlabeled skin lesion classification by self-supervised topology clustering network. *Biomedical Signal Processing and Control*, *66*(5), 102–428.

Wang, D., Tan, D., & Liu, L. (2018). Particle swarm optimization algorithm: an overview. *Soft Computing*, *22*(2), 387–408.

Wang, H., & Hu, D. (2005). Comparison of SVM and LS-SVM for regression. *Proceedings of 2005 International Conference on Neural Networks and Brain Proceedings, ICNNB'05*, *1*, 279–283.

Wang, X., Liu, Y., Chen, Y., & Liu, Y. (2016). An adaptive density based time series clustering algorithm: A case study on rainfall patterns. *ISPRS International Journal of Geo-Information 2016*, *5*(11), 245-271.

Wang, Y., & Zhou, H. (2006). Total variation wavelet-based medical image denoising. *International Journal of Biomedical Imaging*, *206*(4), 302-345.

Warsi, F., Khanam, R., Kamya, S., & Suárez-Araujo, C. P. (2019). An efficient 3D color-texture feature and neural network technique for melanoma detection. *Informatics in Medicine Unlocked*, *17*(3), 1001-1176.

Witkowski, A., Pellacani, G., Gonzalez, S., & Longo, C. (2017). Dermoscopy: Basic knowledge of an innovative imaging tool. *Agache's Measuring the Skin.* 211–228. Springer International Publishing.

Wold JH, V. A. (2001). The derivation of XYZ tristimulus spaces: A comparison of two alternative methods. *Color Research & Application, Wiley Online Library*, *26*(3), 222–224.

Wollina, U. (2018). Seborrheic Keratoses – The most common benign skin tumor of humans. Clinical presentation and an update on pathogenesis and treatment options. *Open Access Macedonian Journal of Medical Sciences*, *6*(11), 2270–2275.

Woodworth, B., Kshitij Patel, K., Stich, S. U., Dai, Z., Bullins, B., Brendan McMahan, H., Shamir, O., & Srebro, N. (2020). *Is Local SGD Better than Minibatch SGD?,* 10334–10343. PMLR.

Xie, F., Yang, J., Liu, J., & Jiang, Z. (2020). Skin lesion segmentation using high-resolution convolutional neural network. *Computer Methods and Programs in Biomedicine*, *17*(2), 105–241.

Yamada, A., Teramoto, A., Otsuka, T., Kudo, K., Anno, H., & Fujita, H. (2016). Preliminary study on the automated skull fracture detection in CT images using black-hat transform. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 6437–6440.

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. In *Insights into Imaging, 1*(4), 611–629.

Yap, J., Yolland, W., & Tschandl, P. (2018). Multimodal skin lesion classification using deep learning. *Experimental Dermatology*, *27*(11), 1261–1267.

Yousef, H., Alhajj, M., & Sharma, S. (2022). Anatomy, Skin (Integument), Epidermis. *In: StatPearls. StatPearls Publishing, Treasure Island (FL).* PMID: 29262154.

Yuan, Y. (2017). Automatic skin lesion segmentation with fully convolutional-deconvolutional networks. *International Skin Imaging Collaboration (ISIC) 2017 Challenge at the International Symposium on Biomedical Imaging (ISBI)*. arXiv preprint arXiv:1703.05165.

Zaqout, I. (2019). Diagnosis of skin lesions based on dermoscopic images using image processing techniques. In *Pattern Recognition - Selected Methods and Applications*. IntechOpen. 1320.

Zhang, H., Dana, K., Shi, J., Zhang, Z., Wang, X., Tyagi, A., & Agrawal, A. (2018). Context encoding for semantic segmentation. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7151–7160.

Zhang, W., Lu, F., Zhao, W., Hu, Y., Su, H., & Yuan, M. (2023). ACCPG-Net: A skin lesion segmentation network with adaptive channel context aware pyramid attention and global feature fusion. *Computers in Biology and Medicine*, *154*(9), 106-580.

Zhao, C., Shuai, R., Ma, L., Liu, W., & Wu, M. (2022). Segmentation of skin lesions image based on U-Net + +. *Multimedia Tools and Applications*, *81*(6), 8691–8717.

Zhao, K., Han, Q., Zhang, C. Bin, Xu, J., & Cheng, M. M. (2022). Deep hough transform for semantic line detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(9), 4793–4806.

Zhao, W., & Lu, H. (2017). Medical image fusion and denoising with alternating sequential filter and adaptive fractional order total variation. *IEEE Transactions on Instrumentation and Measurement*, *66*(9), 2283–2294.

Zhou, H., Chen, M., Gass, R., Rehg, J. M., Ferris, L., Ho, J., & Drogowski, L. (2008). Feature-preserving artifact removal from dermoscopy images. In J. M. Reinhardt & J. P. W. Pluim (Eds.), *SPIE Medical Imaging, 69(*14*),* 209-241.

Zortea, M., Flores, E., & Scharcanski, J. (2017). A simple weighted thresholding method for the segmentation of pigmented skin lesions in macroscopic images. *Pattern Recognition*, *64*(7), 92–104.

APPENDICES

Journal Publications

1. **Khan, A. H.,** Iskandar, D. N. F., Al-Asad, J. F., & El-Nakla, S. (2021). Classification of skin lesion with hair and artifacts removal using black-hat morphology and total variation. *International Journal of Computing and Digital Systems*, *10*(1), 597-604.
2. **Khan, A. H.**, Iskandar, D. A., Al-Asad, J. F., & El-Nakla SAMIRAnd Alhuwaidi, S. A. (2021). Statistical feature learning through enhanced delaunay clustering and ensemble classifiers for skin lesion segmentation and classification. *Journal of Theoretical and Applied Information Technology*, *99*(5), 1122-1138.
3. **Khan, A. H.,** Iskandar, D. A., Al-Asad, J. F., Hiren Mewada and Muhammad Abid Sherazi, Ensemble learning of deep learning and traditional machine learning approaches for skin lesion segmentation and classification, *Concurrency and Computation Practice and Experience*, *34*(13), 6907-6953.

Conference Publications

1. **Khan, A. H.,** Latif, G., Iskandar, D. A., Alghazo, J., & Butt, M. (2018). Segmentation of melanoma skin lesions using anisotropic diffusion and adaptive thresholding. In CBEES, *Proceedings of the 2018 8th International Conference on Biomedical Engineering and Technology*, Indonesia, 23-25 April 2018.