AN EXPERT SYSTEM FOR SKIN CANCER CLASSIFICATION AND DETECTION

Thesis Submitted for the Award of the Degree of

DOCTOR OF PHILOSOPHY

In

Computer Science and Engineering

By

Uzma Saghir

Registration Number: 11720107

Supervised By

Dr. Shailendra Kumar Singh

Assistant Professor

Department of Computer Science and Engineering

Lovely Professional University, Punjab India

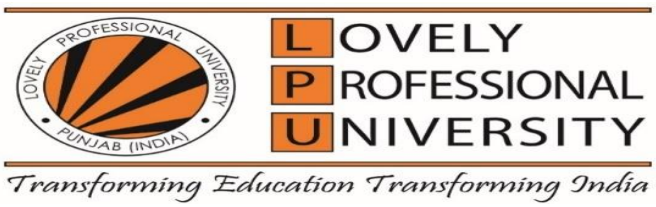
Co-Supervised by

Dr. Moin Hasan

Assistant Professor

Department of Computer Science and Engineering

Jain Deemed-to-be University, Bengaluru India



LOVELY PROFESSIONAL UNIVERSITY, PUNJAB

2023

DECLARATION

I, hereby declare that the presented work in the thesis entitled "**An Expert System for Skin Cancer Classification and Detection**" in fulfilment of the degree of **Doctor of Philosophy (PhD in Computer Science and Engineering)** is the outcome of research work carried out by me under the supervision of Dr. Shailendra Kumar Singh, working as an Assistant Professor in Department of Computer Science and Engineering at Lovely Professional University, Punjab, India and under the Co-supervision of Dr. Moin Hasan working as an Assistant Professor in Department of Computer Science and Engineering at Jain Deemed-to-be University, Bengaluru in keeping with the general practice of reporting scientific observations, due acknowledgements have been made whenever work described here has been based on findings of another investigator. This work has not been submitted in part or full to any other University or Institute for the award of any degree.

Uzma Saghir

Signature of Scholar

Name of the Scholar: Uzma Saghir

Registration No: 11720107

Department/School: Computer Science and Engineering

Lovely Professional University,

Punjab, India

CERTIFICATE

This is to certify that the work reported in the Ph. D. thesis entitled "An Expert System for Skin Cancer Classification and Detection" submitted in fulfilment of the requirement for the reward of the degree of **Doctor of Philosophy (Ph.D.)** in Computer Science and Engineering, is a research work carried out by Uzma Saghir (Registration No 11720107), is bona fide record of his/her original work carried out under my supervision and that no part of thesis has been submitted for any other degree, diploma or equivalent course.

Signature of Supervisor

Name of supervisor: Dr. Shailendra Kumar Singh
Assistant Professor
Department of Computer Science and Engineering
Lovely Professional University, Punjab

Signature of Co-Supervisor

Name of Co-Supervisor: Dr. Moin Hasan
Assistant Professor
Department of Computer Science and Engineering
Jain Deemed-to-be University, Bengaluru

ABSTRACT

Skin cancer affects millions of people worldwide, and extensive research is dedicated to detecting its root causes. The skin, being the body's largest and most sensitive organ, can develop cancer due to inappropriate treatment or environmental factors. The skin plays a crucial role in supporting the body's structure and appearance, and various diseases can have adverse effects on it. During the last few years, the incidence of skin cancers has increased worldwide. According to the World Health Organization (WHO), skin cancer is becoming more common than all other types of cancer. This accounts for one in every three cancer diagnoses, with 132,000 newly diagnosed cases identified each year. Additionally, it is reported that melanoma incidence rises by 4% to 6% annually in regions with people who have light complexions.

Being among the deadliest malignancies, melanoma is becoming more prevalent worldwide. Dermatologists inspect skin lesions using a diagnostic method known as dermoscopy. This involves using a magnification tool and a source of light to inspect structures beneath the skin that would be otherwise imperceptible. Dermoscopy, despite its efficiency, is a talent that takes years to master. Furthermore, the diagnosis is frequently subjective and difficult to duplicate. As a result, there is a need to develop automated ways to aid dermatologists in providing a more reliable diagnosis.

Computer-based detection (CAD) techniques show potential, improving melanoma diagnosis rates by between five and thirty per cent over the naked eye. Because visual perception is prone to errors, a second opinion is required for increased precision and reliability. Furthermore, these systems relieve clinicians of tasks and responsibilities, allowing for a more efficient approach to diagnosis. Computer-aided diagnosis has grown in popularity in the medical industry. Skin cancer is a common and potentially fatal ailment that a CAD system could potentially detect. Given that it manifests visibly on the skin, the diagnosis can rely solely on photographs of skin lesions.

The primary objective of this thesis is to introduce an intelligent approach to skin cancer classification that utilizes image processing techniques and artificial intelligence tools, including machine learning. The important aim of this study is to investigate and achieve the highest possible accuracy and effectiveness in medical classification applications. This research aims to detect skin cancer using a novel system for skin cancer classification with a customized classification phase.

The proposed methodology is divided into several phases. In the initial phase, data collection is performed, which includes benchmark ISIC datasets. After data collection, pre-processing is applied to remove noise from the collected dataset. The preprocessing phase involves a series of techniques aimed at cleaning, normalizing, and enhancing dermoscopic images. This phase is critical for the accurate classification of skin cancer. During pre-processing, the image is resized, hairy parts are

removed, and discrete cosine transformation (DCT) and color space conversion are to enhance and restore the image.

The segmentation stage is critical and challenging in image processing, requiring both speed and precision. The effectiveness of succeeding procedures, particularly feature extraction and classification, is heavily dependent on its proficiency. In this context, our research takes a fresh look at the segmentation step, combining background subtraction with a midpoint analysis to define the area of interest (ROI). The proposed strategy not only speeds up segmentation but also improves the performance of several classifiers.

Each dermoscopy image can produce hundreds of features that serve as image descriptors. However, not all of these characteristics are important for lesion classification. Unnecessary characteristics can complicate the classifier, raise computational needs, and perhaps impair classification accuracy. It is critical to choose the most relevant elements to portray the distinctive characteristics of skin cancer images. As a result, the goal is to extract the fewest features that effectively discriminate across the images. In this work a novel method Differential Analyzer Algorithm (DAA) for feature extraction and selection is proposed. The adaptation of the DAA approach for feature extraction and selection involves iteratively moving from one feature set to another while assessing the impact on the model's performance.

Finally, a customized classifier is employed for skin cancer classification, and the results are validated based on classification accuracy, specificity, and sensitivity. The research attains a high classification accuracy of 96.35% with the customized classifier, 96.21% with SVM, 94.00% with KNN, 93.50% with a decision tree, and 91.00% with the Random Forest approach, thereby confirming the validity of the proposed approach.

ACKNOWLEDGMENT

I am grateful to Almighty Allah for giving me the life and wisdom to pursue my Ph.D. in Computer Science and Engineering at Lovely Professional University, India. I want to thank my supervisors, Dr. Shailendra Kumar Singh and Dr. Moin Hasan, for their guidance and support throughout my Ph.D. journey. I would like to extend my gratitude to Dr. Rekha, Dr. V Devendran and Mrs. Tanu Mahajan for their support and guidance.

I am deeply thankful to my mother Akhtoob Begum and father Mohd Saghir for their encouragement, prayers, and support during my studies. To my sisters Dr. Fozia Saghir, Tranum Sultana, Taiba Sultana and my brother Abrar Ahmed I am grateful for their patience and prayers.

I want to acknowledge my friends Satvinder Kour, Mohd Asif, Shagufta Gorsi, Refit Amir and Adwithiya DeBarama and other colleagues for their immense encouragement and contributions to my success, though their names are too many to mention individually. I will always remember and appreciate their help, Thank you all.

TABLE OF CONTENT

DECLARATION i
CERTIFICATE ii
ABSTRACT iii
ACKNOWLEDGMENT v
TABLE OF CONTENT vi
LIST OF FIGURES ix
LIST OF TABLES x
LIST OF ABBREVIATION xi
CHAPTER 1 1
INTRODUCTION 1
1.1 Skin Cancer 1
1.1.1 Different Types and Stages of Skin Cancer 2
1.1.2 Clinical Presentation 4
1.2 Existing Diagnostic Methods 5
1.2.1. Dermoscopy 5
1.2.2. Biopsy 6
1.2.3. Total Body Photography 6
1.2.4. Reflectance Confocal Microscopy 7
1.2.5. Ultrasound Imaging 7
1.2.6. Teledermatology 8
1.2.7. Machine Learning and Artificial Intelligence 8
1.3 Machine Learning for Skin Cancer Detection 10
1.3.1 Supervised Learning 11
1.3.2 Unsupervised Learning 12
1.3.3 Ensemble Methods 12
1.3.4 Deep Learning for Skin Cancer Detection 13
1.3.5 Computer-Aided Diagnosis Systems for Skin Cancer Detection 13
1.3.6 Hybrid Approaches 13
1.4 Problem Statement 15
1.5 Objective of the Study 17
1.6 Thesis Organization 17
1.7 Summary 18
CHAPTER 2 19
LITERATURE REVIEW 19

2.1 Background of Cancer Detection Systems 19
2.1.1 Skin Cancer Image Preprocessing 27
2.1.2 Skin Cancer Image Segmentation 32
2.1.3 Skin Cancer Image Feature Extraction 37
2.1.4 Skin Cancer Image Classification 46
2.2 Summary 52
CHAPTER 3 53
METHODOLOGY AND SYSTEM FRAMEWORK 53
3.1 Proposed System 53
3.2 Publicly Available ISIC Dataset 54
3.2.1 Applications and Contributions 54
3.2.2 Diverse Challenges and Opportunities 55
3.3 Pre-Processing of Dermoscopic Images 55
3.3.1 Image Resizing 56
3.3.2 Hair Removal Using Bottom Hat Filter 56
3.3.3 Image Enhancement Applying DCT and Color Space Conversion 57
3.4 Segmentation of Dermoscopic Images 60
3.4.1 Proposed Segmentation Method 61
3.4.2 Advantages and Challenges 64
3.5 Feature Extraction and Selection for Dermoscopic Images 64
3.5.1 Proposed Method for Feature Extraction and Selection 66
3.5.2 Steps Involved in DAA Algorithm 66
3.5.3 Benefits and Considerations of DAA Algorithm 71
3.5.4 Normalization 71
3.6 Summary 72
CHAPTER 4 73
CLASSIFICATION 73
4.1 Models for Skin Cancer Classification and Detection 73
4.1.1 Support Vector Machine 73
4.1.2 K- Nearest Neighbor 74
4.1.3 Random Forest 74
4.1.4 Decision Tree Classifier 74
4.1.5 Artificial Neural Networks 75
4.1.6 Convolutional Neural Networks 75
4.1.7 Customized Classifier 76
4.2 Proposed Classification Model 77

4.3 Summary 82
CHAPTER 5 83
RESULTS AND DISCUSSION 83
5.1 Evaluation of the Proposed System 83
5.2 Dataset and Tools 83
5.3 Metrics for Performance Evaluation 83
5.4 Results of Preprocessing Phase 84
5.5 Results of Segmentation Phase 85
5.6 Results of Feature Extraction and Selection Phase 88
5.6.1 Results of Normalization 90
5.6.2 Performance Evaluation of Different Classifiers on Different Feature Sets 90
5.7 Result Analysis of Different Classifiers 91
5.8 Results Comparisons with the State of Art Methods 94
5.9 Results and Performance Evaluation on the ISIC Dataset 96
5.10 Results and Comparison with Deep Learning Methods 97
5.11 Summary 100
CHAPTER 6 101
CONCLUSION AND FUTURE SCOPE 101
6.1 Contribution of Proposed System to Society 101
6.2 Conclusion 101
6.3 Future Scope 103
REFERENCE 104
APPENDIX A 135
APPENDIX B 136

LIST OF FIGURES

Figure 1.1 Different Skin's Layers and Cells.....2
Figure 1.2 Stages of Skin Cancer.....3
Figure 1.3 Types of Melanomas (Cancerous).....5
Figure 2.1 Phases of Skin Cancer Detection System 27
Figure 2.2 Preprocessing Methods for Dermoscopic Skin Cancer Images 28
Figure 2.3 Segmentation Methods for Skin Cancer Images 32
Figure 2.4 Feature Extraction and Selection Methods for Skin Cancer Images 38
Figure 3.1 Block Diagram for Proposed Frameworks of Skin Cancer Classification and Detection.....54
Figure 3.2 ISIC Dataset Images (Cancerous and Non- Cancerous).....55
Figure 3.3 Preprocessing Phase of Proposed System.....57
Figure 3.4 Image Enhancement during Preprocessing Phase.....58
Figure 3.5 Working of Background Subtraction Model.....61
Figure 3.6 Different Types of Features Extracted.....65
Figure 3.7 Mechanism of DAA Algorithm.....70
Figure 4.1 Proposed Correlation and Linearity Model 77
Figure 5.1 Results of Preprocessing Phase.....85
Figure 5.2 Results of Segmentation Phase.....86
Figure 5.3 Shows Feature Extraction and Selection Using DAA.....89
Figure 5.4 Colour Feature Matrix After Normalization.....90
Figure 5.5 Performance Analysis of Different Classifiers using Grayscale Normalization.....92
Figure 5.6 Performance Analysis of Different Classifiers using Min-Max Normalization.....93
Figure 5.7 Performance Analysis of Different Classifiers using Depth based Normalization..93

LIST OF TABLES

Table 1.1 Existing Technologies for Skin Cancer Detection 9
Table 1.2 Classification Methods for Skin Cancer Classification and Detection 14
Table 2.1 Different Techniques for Skin Cancer Segmentation 36
Table 2.2 Pattern Analysis Method for Feature Extraction 39
Table 2.3 ABCD Rule Of Dermoscopy 40
Table 2.4 Seven-Point Checklist Method for Feature Extraction 40
Table 2.5 Types of Features used by Menzie's Method for Feature Extraction 41
Table 2. 6 CASH Algorithm for Feature Extraction 41
Table 2.7 Comparative Analysis of Different Feature Extraction Method 44
Table 2.8 Comparative Analysis of Different Correlation Method 50
Table 3.1 The Range of Values Obtained Through DAA Approach 69
Table 3.2 Constraint Settings for DAA 70
Table 5.1 Comparison Analysis of Proposed Segmentation Method with Different State-of-Art Methods Using SVM Classifier 87
Table 5.2 Comparison Analysis of Proposed Segmentation Methods with Different Segmentation Methods 88
Table 5.3 Performance Evaluation of Different Classifiers on Different Feature Sets 91
Table 5.4 Level of Correlation with Different Features of Image 91
Table 5.5 Metrics for Performance Evaluation 92
Table 5.6 Comparison of Various Techniques for Skin Cancer Classification and Detection on Different Dataset 95
Table 5.7 Comparative Analysis of the Proposed SCCC System with State of Art Methods ...97
Table 5.8 Comparative Analysis of the Proposed SCCC System with Deep Learning Methods 98
Table 5.9: A Comparative Analysis of Deep Learning Methods for Skin Cancer Classification and Detection 98

LIST OF ABBREVIATION

ALL	Acute Lymphoblastic Leukemia
AI	Artificial Intelligence
ANN	Artificial Neural Networks
ABCD	Asymmetric, Border, Color, Diameter
CASH	Colour, Architectural, Symmetry, Homogeneity
CAD	Computer-Aided Diagnostic
CCA	Correlation Component Analysis
CF	Correction Factor
CFG	Correlation-Based Feature Group
CFS	Correlation-Based Feature Selection
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNMD	Consensus Meeting of Dermoscopy
CNN	Convolutional Neural Networks
DAA	Differential Analyzer Algorithm
DBDS	Duke Breast Cancer Dataset
DCNN	deep convolutional neural network
DCT	Discrete Cosine Transformation
DNN	Deep Neural Network
DT	Decision Tree
FCEDN	Full Convolutional Encoder-Decoder Network
FCM	Fuzzy c-mean
FD	Feature Dimension
FrCN	Full-Resolution Convolutional Networks

GA	Genetic Algorithm
GLCM	Grey-Level Co-Occurrence Matrix
GLRLM	Grey Level Run Length Matrix
GVF	Gradient Vector Flow
HRF	High-Resolution Fundus
HSV	Hue, Saturation, Value
HOG	Histograms of Oriented Gradients
IEDCA	Intra-Class and Extra-Class Discriminative Correlation Analysis
ISIC	International Skin Imaging Collaboration
KNN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
LM	Linearity model
LSTM	Long-Short-Term-Memory
ML	Machine Learning
MMN	Min-Max Normalization
NOS	Normalised Otsu's Segmentation
PCA	Principal Component Analysis
PNN	Probabilistic neural network PNN
RBF	Radial Basis Function
sRCM	Reflectance Confocal Microscopy
RF	Random Forests
ROC	Receiver Operating Characteristic
SC	Spectral Clustering
SGWT	Spectral Graph Wavelet Transform
SVM	Support Vector Machine

SVM-RFE	Support Vector Machines Recursive Feature Elimination
TBP	Total Body Photography
TDS	Total Dermoscopy Score
TCSPC	Time-Correlated Single-Photon Counting
t-SNE	t-Distributed Stochastic Neighbor Embedding
WHO	World Health Organization
UV	Ultraviolet
UVR	Ultraviolet Radiation

CHAPTER 1

INTRODUCTION

Skin cancer is a widespread disease pervaded by uncontrollable cellular proliferation. Its ubiquity emphasizes the critical importance of increased ultraviolet (UV) radiation exposure, which is mostly emitted by solar radiation or artificial sources like as tanning beds. This section aims to provide a fundamental understanding of skin cancer, including an in-depth examination of the skin's complex structure, crucial technical terminology for comprehending the complexities of skin cancer, and the diagnostic techniques utilized by medical specialists. Furthermore, we will explore function of artificial intelligence (AI) and computer-assisted diagnosis (CAD) technologies in skin cancer detection, explaining numerous approaches to improve accuracy and efficiency.

1.1 Skin Cancer

Skin cancer is becoming an increasingly prominent global health issue, experiencing a rise in incidence in recent years. Afflicting millions of individuals worldwide annually, it stands as one of the most widespread forms of cancer. The condition emerges when skin cells undergo abnormal expansion and uncontrolled proliferation [1]. The human body's skin protects it from heat, injury, infection, and UV radiation damage. It can also produce vitamin D, hold water and fat.

During the last few years, the incidence of skin cancers has increased worldwide. According to the World Health Organization (WHO), skin cancer is becoming more common than all other types of cancer. This accounts for one in every three cancer diagnoses, with 132,000 newly diagnosed cases of identified each year. Additionally, it is reported that melanoma incidence rises by 4% to 6% annually in regions with people who have light complexions [2]. While affecting barely two per cent of all diagnoses, the malignant type of melanoma contributes to eighty per cent of the disease's mortality. Based on data from the American Cancer Society, the estimated total number of new cases of cancer in 2022 will probably exceed 1,900,000, with 609,360 deaths [3]. UV radiation emitted by the sun is the major cause of skin cancer, and the rise in popularity can be ascribed to a variety of causes, such as more sun exposure, changing lifestyles, and ozone layer depletion [4].

As a result of these findings, numerous campaigns and initiatives focusing on skin cancer prevention and early diagnosis have been started. The risks of tanning booths, the necessity of limiting ultraviolet radiation (UVR) exposure during the peak of the day, the significance of seeking

out shade, the requirement for sun-protective clothing, and the significance of using sunscreen should all be discussed with patients by primary care physicians and dermatologists [5]. Even with the implementation of these public health measures, skin cancer may be on the rise in several countries.

The human skin is an intricate structure comprising multiple layers, with the epidermis and dermis serving as its primary components. The epidermis, the outermost layer, consists of squamous cells, which are flat in shape. Just beneath the squamous cells, basal cells are found. Melanocytes are cells found in the epidermis's lowest levels, between the basal cells. Melanocytes in the skin create pigment (colour). The dermis is the skin's second major layer, located underneath the epidermis. It contains a variety of cell types, including lymph tubes, blood vessels, and glands. Some glands help to dry the epidermis, while others help regulate temperature and soothe the body. Figure 1.1 illustrate the skin's layers and cells [6].

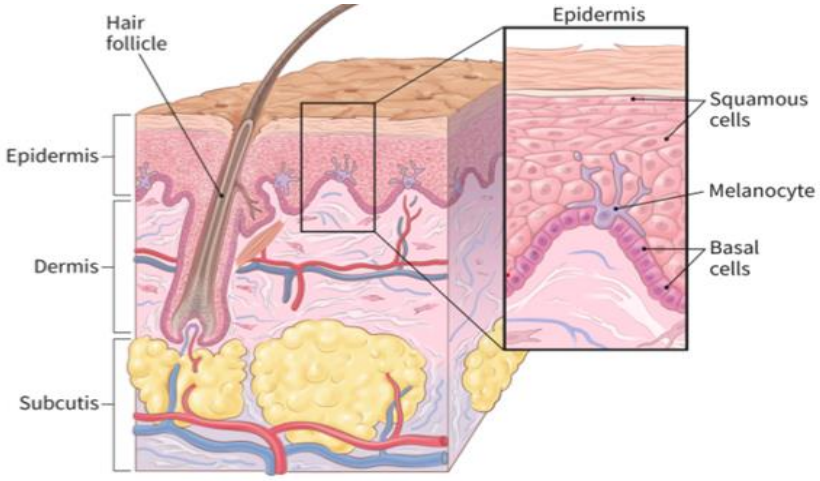


Figure 1.1 Different Skin's Layers and Cells

1.1.1 Different Types and Stages of Skin Cancer

Trillions of living cells make up the human body. These cells spread, split, and decease in an orderly fashion in normal bodies. Adults divide their cells to replace those that are worn out, wounded, or dead. Cancer arises when the proliferation of abnormal cells in a specific part of the body becomes uncontrollable [7]. Cancer cell proliferation generates more cancer cells capable of infecting different tissues. Skin cancer is caused by UV radiation that is harmful to the skin and abnormal development of skin cells known as melanocytes. Melanocytes produce melanin pigments in melanosomes, which are then transported to the neighbouring epidermis via the dendritic process [8]. Melanocytes develop from the neural crest and travel to various places across the skin during the growth stage. Skin abnormalities or tumours are typically classified as benign or malignant.

- **Benign Tumours:** These are non-cancerous growths such as cysts or moles. They usually stay localized, do not infect surrounding tissues, and have no chance of smattering to further areas of the body. Benign tumours are usually non-threatening and can be removed safely and easily.
- **Malignant Tumours:** In contrast, malignant tumours are cancerous and pose a major health danger. They develop uncontrollably and have the ability to infiltrate neighbouring tissues. Their capacity to metastasise, or spread to distant organs and areas of the body, makes them particularly hazardous. Malignant skin tumours can be fatal if left untreated, necessitating rapid medical care and complete treatment. Melanoma is a common type of malignant skin cancer.

Melanoma is categorised into 5 stages from Stage 0 to Stage 4 as shown in Figure 1.2 [9].

- **Stage 0:** This is the first stage of melanoma and specifies that the cancerous cells are present but only in the epidermis layer of the skin and have not grown to any other part of the body.
- **Stage 1:** This stage indicates that cancer has reached outside the epidermis but not to any other part of the body, and it lies between 0.8 mm and 2.0 mm in thickness.
- **Stage 2:** In this stage, the thickness of the melanoma is bigger than 2.0 mm but less than 4.0 mm, just like in stage 1. This also does not spread to any other part.
- **Stage 3:** The cancer cells have started to grow in the dermic layer and have taken place into contiguity with the nerves at this point. The thickness of melanoma at this stage is greater than 4.0 mm.
- **Stage 4:** This stage indicates that the melanoma has grown in other organs of the body. At this stage, it is arduous to treat melanoma [10].

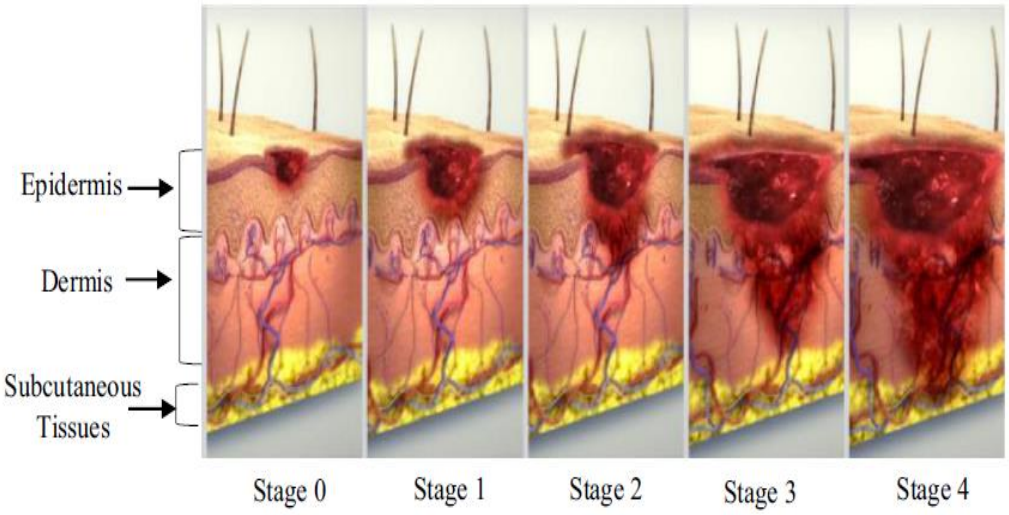


Figure 1.2 Stages of Skin Cancer

The background of skin cancer detection encompasses various factors, including its prevalence, risk factors, clinical presentation, and existing diagnostic methods. Understanding the background is essential to recognize the significance of developing advanced techniques for early detection and reliable diagnosis.

1.1.2 Clinical Presentation

Skin cancer presents various clinical features that dermatologists assess during the examination. These features include changes in the size, shape, color, and texture of moles or lesions as shown in Figure 1.3. Suspicious signs may comprise anomalies, uneven boundaries, variations in color or diameter, and evolving characteristics over time. However, visual inspection alone can be subjective and prone to errors, leading to the need for advanced diagnostic techniques [11].

Malignant melanomas are categorized as:

- **Nodular Melanoma**: This type of tumor accounts for approximately fifteen per cent of all malignancies and grows rapidly after diagnosis. Nodular melanomas are nodules or blisters that are dusky bronze to dark in color or resemble blood vessels; however, five per cent of them are amelanotic. These melanomas are commonly found on different parts of the body, but especially on men's hips [12].
- **Superficial Spreading:** The most common kind of melanoma, which accounts for more than 70% of all melanomas, is superficial spreading melanoma. This melanoma grows horizontally on the skin for a long time (years) before becoming invasive. Superficial spreading melanomas are brown lesions with irregular asymmetric borders that are flat or slightly elevated with the staining that is black, blue, or pink. These melanomas can appear on anybody's surface, but they are most common on men's heads, necks, and trunks and on females' lower extremities [13].
- **Lentigo Maligna and lentigo Maligna melanoma:** These cancers commonly appear on the face of elderly people. It resembles a huge and uneven mole that grows slowly [14].
- **Acral Lentiginous Melanoma:** It typically develops on hairless skin, such as the palms and soles. It nearly never receives a timely diagnosis, making it the form of malignant melanoma with the worst prognosis [15].
- **Amelanotic Melanoma:** It is an extremely rare kind of skin cancer. These types of cutaneous melanomas are challenging to identify and are frequently misdiagnosed as normal lesions on the skin [16].

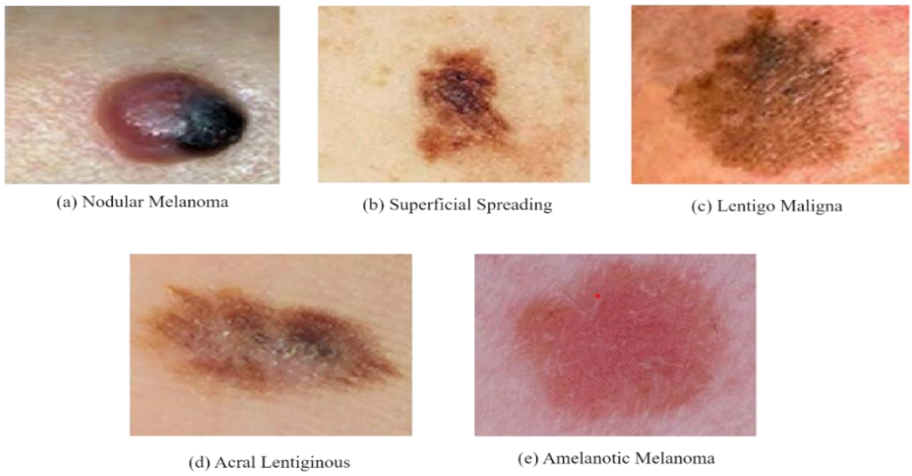


Figure 1.3 Types of Melanomas (Cancerous)

1.2 Existing Diagnostic Methods

Traditionally, skin cancer diagnosis relies on visual inspection and clinical expertise [17]. Dermatologists examine suspicious lesions, often using dermoscopy, a technique that magnifies the skin surface for better visualization. Biopsy, where a tissue sample is obtained for pathological examination, is performed if malignancy is suspected. However, these methods have limitations, including subjectivity, inter-observer variability, and dependence on expertise, leading to the exploration of automated and computer-aided diagnostic systems [18].

The significance of skin cancer detection lies in its potential to save lives, improve treatment outcomes, and enhance accessibility to early diagnosis. Early detection allows for less invasive treatment options, better chances of successful outcomes, and reduced healthcare costs. Moreover, implementing automated skin cancer detection methods can help overcome geographical and socioeconomic barriers to accessing dermatological expertise [19].

The development of advanced techniques for skin cancer detection, such as the utilization of digital differential analyzers coupled with different classifiers, can provide objective and quantitative analysis to support clinical decision-making. These techniques aim to improve accuracy, sensitivity, and specificity in identifying malignant lesions, enabling timely intervention and improved patient care.

1.2.1. Dermoscopy

A non-invasive diagnostic method used to examine skin lesions is known as dermoscopy. Epiluminescence microscopy and dermatoscopy are other names for it [20]. It entails the use of a

dermatoscopy, which is a handheld equipment with magnification and illumination that allows healthcare practitioners to see the skin at a considerably better resolution than the naked eye. This handheld device provides a higher level of detail, allowing healthcare professionals to visualize subsurface skin structures and identify characteristic features associated with skin cancer [21]. Dermatologists and other skin disease specialists regularly use this approach to aid in the identification of various skin problems such as moles, melanomas, benign growths, and other skin tumors [22].

1.2.2. Biopsy

A biopsy is a common and important process used to establish a definitive diagnosis of skin cancer. When an alarming skin lesion is identified, a biopsy is performed to obtain tissue samples from the afflicted area for laboratory examination. The biopsy helps to detect if the lesion is malignant or benign, and if it is malignant, it identifies the specific kind of skin cancer present [23]. There are several methods for collecting tissue samples for the biopsy, considering the proportions and features of the lesion. The following are the most common techniques for skin cancer biopsies [24]:

- **Punch Biopsy:** A small, cylindrical sample is extracted from the skin lesion using a round tool.
- **Shave Biopsy:** A tiny film of the lesion is removed with a knife or other equivalent equipment.
- **Excisional Biopsy:** Using a scalpel, the entire lesion or a significant piece of it is removed.

It is critical to emphasise that a biopsy is still the optimal procedure for detecting skin cancer. While computer-aided diagnostics and noninvasive imaging tools can help discover and classify skin lesions, a sample is required to confirm the diagnosis. Skin cancer identification and diagnosis are critical for starting the right treatment and getting better results.

1.2.3. Total Body Photography

Total Body Photography (TBP) is a thorough imaging technology used to monitor and identify skin cancer in people who are at high risk of getting it. It entails taking high-resolution photographs of a person's whole skin surface, including the front, back, and sides, as well as particular close-up images of individual skin lesions or moles [25]. TBP is commonly performed by specialized clinics or dermatologists and is quite beneficial to patients. While TBP is a useful tool for high-risk individuals, it is not a replacement for routine skin checks or biopsies when necessary. TBP is part of a broader skin cancer surveillance program that includes regular skin exams, patient education,

and, if necessary, biopsies for final diagnosis [26]. TBP can help in this endeavour by improving the surveillance procedure for high-risk patients.

1.2.4. Reflectance Confocal Microscopy

Reflectance Confocal Microscopy (RCM), a non-invasive technique for imaging, is used to look at the skin's live tissue at the cellular level. RCM provides doctors with high-resolution, instantaneous images of the structure of cells of the skin, enabling them to observe skin lesions and determine whether or not they are benign [27]. A portable probe with a laser light source is placed on the skin surface during an RCM examination. The laser light penetrates the skin and is reflected to the probe's detector. The detected signals are subsequently translated into photographs of the skin's cellular architecture in real-time [28].

RCM captures individual skin cells and their architectural patterns in great aspect. This enables dermatologists to identify biological traits linked to particular skin conditions, such as skin cancer. Dermatologists can detect changes in the cellular architecture of a lesion by comparing RCM pictures recorded at different points in time, allowing for early identification of probable skin cancer. Dermatologists employ RCM to differentiate between malignant and non-malignant skin lesions by leveraging cellular features. This distinction is critical for effective treatment planning, ensuring targeted interventions, and avoiding needless therapies.

1.2.5. Ultrasound Imaging

Ultrasound imaging or sonography, is a sort of imaging procedure which employs sound waves with a high frequency to generate real-time images of interior organs and tissues. Although ultrasound is more frequently used to image internal body parts and cutaneous tissues, it may also be employed to diagnose skin cancer in some circumstances [29]. A skin tumour's thickness and depth can be evaluated using ultrasound. Given that tumour thickness is a crucial element in predicting prognosis and therapy options for melanomas, this is very helpful for staging these cancers.

Ultrasound can be used to examine surrounding lymph nodes for symptoms of metastasis (cancer spread) in some skin malignancies, including melanoma. In some circumstances, ultrasound can be applied to track a patient's reaction to treatment for skin cancer. It might aid in determining how tumour features and size have changed over time [30]. It's important to remember that, in contrast to other imaging techniques like dermoscopy, ultrasonic imaging is not the primary tool for diagnosing the majority of skin malignancies.

1.2.6. Teledermatology

Teledermatology is a medical practice that uses telecommunication technology to perform remote diagnosis, evaluation, and treatment of dermatological problems such as skin cancer. It enables people to receive dermatological care without having to travel to a clinic or hospital. Patients can use video conferencing or telecommunication techniques to consult with dermatologists or skin cancer specialists remotely. This is especially useful for people who live in distant or underdeveloped locations and may not have easy access to specialized treatment [31]. Dermatologists can evaluate the photos and provide early detection and monitoring of problematic lesions, which is critical for prompt skin cancer diagnosis and treatment. It is crucial to emphasize, however, that teledermatology has limitations. Some skin disorders, particularly those that are complex or fast-changing, may still necessitate in-person inspection and treatment.

1.2.7. Machine Learning and Artificial Intelligence

Machine Learning (ML) and Artificial Intelligence (AI) are rapidly being used in skin cancer detection to enhance accuracy, efficiency, and early diagnosis. The image data for skin cancer is acquired by employing a dermatoscopy or other imaging equipment. These photos are processed by AI and ML algorithms to extract relevant properties such as colour, texture, structure, and shape. These extracted features are fed into classification models as inputs [32]. To identify patterns associated with diverse forms of skin lesions, these techniques are refined on a huge dataset of labelled dermoscopic images.

The ML model can classify unseen dermoscopic images into a range of classes, including benign lesions, malignant melanomas, and specific types of skin cancer. The model, which assigns a probability to each class, aids dermatologists' decision-making. Dermatologists may employ AI and ML models to help them make decisions [33]. In complex cases or when a dermatologist has limited experience with a specific type of skin lesion, they may provide a second opinion or additional information to aid in the diagnosis.

Diagnosing skin cancer using technology has significantly advanced in recent years, enabling more accurate and efficient detection. Several technologies are employed to aid the diagnosis process, as discussed in Table 1.1.

Table 1.1 Existing Technologies for Skin Cancer Detection

Approaches	**Purposes**	Limitations
Dermoscopy [20] [21]	Magnifies the skin surface for better visualization.	• Operator Dependent • Expensive • Error-prone
Biopsy [23][24]	A tissue sample is obtained for pathological examination, and is performed if malignancy is suspected.	• Sampling error • Expensive • Time Taking
Total Body Photography (TBP) [25],[26]	A comprehensive visual record of the patient's skin.	• Slow process • Expensive • Multiple-time sample collection and storage • Dependence on Imaging Quality
Reflectance Confocal Microscopy (RCM) [27], [28]	A non-invasive technique for imaging is used to look at the skin's live tissue at the cellular level.	• Error-prone • Expensive
Ultrasound imaging or Sonography [29], [30]	Imaging procedure which employs sound waves with a high frequency to generate	• Subjective

	real-time images of interior organs and tissues.	• Image Artifacts • Limited Soft Tissue Differentiation
Teledermatology [31]	Use of telecommunications technology to provide dermatological care and consultation remotely.	• Relies on visual assessments • Technological Barriers • Limited Diagnostic Tools • Miscommunication
Computer-Aided Diagnosis Systems [34], [35], [36], [37], [38], [39], [40]	These systems analyze dermoscopic images and provide a quantitative assessment or a probability score for malignancy.	• The quality of the image dataset • User dependency

1.3 Machine Learning for Skin Cancer Detection

Skin cancer detection has greatly benefited from advancements in machine learning techniques. ML-based mechanisms have shown promise in automating the diagnosis process, improving accuracy, and aiding healthcare professionals in making informed decisions [41]. The combination of these technologies improves the precision, effectiveness, and affability of skin cancer diagnosis. ML models can quickly and accurately analyse large volumes of dermatological pictures. They can detect minor changes in moles or skin lesions that may be suggestive of skin cancer early on, resulting in more effective therapy and better outcomes [42].

Skin lesions can be evaluated objectively using ML algorithms. The models retain consistency in their assessments, unlike human judgement, which can be impacted by factors such as exhaustion or bias. These models can aid in lowering the rate of false positive outcomes, which can be

problematic in skin cancer detection. These models can learn from new data indefinitely, improving their diagnostic accuracy over time [43]. This recurrent learning process guarantees that they are up to date on the most recent advancements in skin cancer detection. Remote consultations can be facilitated by ML, allowing patients in underserved areas to receive professional advice without having to travel large distances. The use of machine learning techniques in the diagnosis of skin cancer has the potential to greatly increase public awareness of the significance of routinely checking one's skin and getting medical help when needed [44].

However, while technology can aid healthcare practitioners in the diagnosing process, a final diagnosis often necessitates a histological study of a biopsy material. Therefore, technology serves as a valuable tool in aiding the diagnostic process but should not replace clinical expertise and judgment. Early detection and accurate diagnosis are essential for efficient treatment and improved results for patients. The background of skin cancer detection encompasses various factors, including its prevalence, risk factors, clinical presentation, and existing diagnostic methods. The development of advanced techniques for skin cancer detection holds great significance in enhancing early detection, reducing mortality rates, and improving accessibility to healthcare services. Here, we discuss various ML-based mechanisms for skin cancer detection in this section.

1.3.1 Supervised Learning

Supervised learning algorithms require labelled training data to learn patterns and classify new instances accurately. In skin cancer detection, these algorithms are trained on dermoscopic images with corresponding labels (benign or malignant) [45]. The supervised learning algorithms mentioned below are commonly used for skin cancer detection:

- **Support Vector Machines (SVM):** SVMs are used in skin cancer detection to find lesions by figuring out which hyperplane is best for differentiating between different classifications, including benign and malignant. These approaches improve class separation by transforming data into higher-dimensional spaces via a kernel function. SVMs use extracted features to provide accurate classification and have demonstrated efficacy in differentiating among lesions [46].
- **Artificial Neural Networks (ANNs):** Artificial Neural Networks (ANNs), specifically advanced deep learning architectures like Convolutional Neural Networks (CNNs), have sparked a transformative breakthrough in the realm of skin cancer detection. CNNs exhibit the capacity to autonomously acquire hierarchical features from dermoscopic images, effectively capturing nuanced patterns essential for precise classification. Transfer learning, a method that entails refining

pre-existing CNN models on skin cancer datasets, has demonstrated remarkable efficacy in achieving high levels of performance within this field [47].

- **Random Forests (RF):** Random Forests construct an ensemble of decision trees and use majority voting to classify instances. They can handle high-dimensional feature spaces and are robust against overfitting. Random Forests have been used effectively for skin cancer detection, leveraging feature importance to identify discriminative features [48].

1.3.2 Unsupervised Learning

Unsupervised learning represents a facet of machine learning that involves algorithms exploring structures, patterns, or relationships within data without relying on explicit guidance from labelled samples. This facet holds significant importance in both data analysis and machine learning, as it enables the extraction of valuable insights from unstructured or unlabeled data [49]. In the context of skin cancer detection, prevalent methods within unsupervised learning are frequently utilized for this purpose.

- **Clustering Algorithms:** These techniques play a crucial role in grouping similar instances based on their features, facilitating the identification of distinct classes or patterns within skin lesion data. Employing algorithms such as K-means clustering and hierarchical clustering has proven effective in analyzing dermoscopic images, potentially unveiling concealed structures or subtypes of skin cancer [50].
- **Dimensionality Reduction:** The goal of dimensionality reduction methods is to simplify high-dimensional data while retaining essential information. Widely adopted approaches such as t-distributed Stochastic Neighbor Embedding (t-SNE) and Principal Component Analysis (PCA) are utilized to visualize and explore dermoscopic images. These techniques can aid in feature selection and anomaly detection, contributing to a more comprehensive understanding of the data [51].

1.3.3 Ensemble Methods

Ensemble approaches improve machine learning prediction accuracy by merging many models and exploiting collective knowledge to aggregate forecasts. Two commonly used ensemble approaches in the detection of skin cancer include [52]

- **Bagging:** Bagging (Bootstrap Aggregating) is integrating numerous machine learning models trained on different data samples used for training. Each model makes an independent prediction, which is determined by average or majority voting. This method is efficient in reducing model variance and improving overall accuracy [53].

- **Boosting**: Boosting trains, a sequence of weak ML models, where each subsequent model focuses on instances that previous models struggled with, thereby improving overall performance. AdaBoost and Gradient Boosting are popular boosting algorithms used for skin cancer detection, achieving superior classification results [54].

1.3.4 Deep Learning for Skin Cancer Detection

In the field of skin cancer diagnosis, deep learning architectures—in particular, Convolutional Neural Networks (CNNs)—have demonstrated exceptional performance. These CNNs eliminate the need for manually created features by automatically extracting hierarchical features from dermoscopic images. Their success in this field is largely attributed to their ability to instinctively recognize complex patterns, textures, and spatial relationships. The limitations posed by the scarcity of labelled data for training have led to a broad acceptance of transfer learning, a technique that involves fine-tuning pre-trained CNN models using datasets related to skin cancer [55].

In the domain of skin cancer, machine learning has significant advantages over deep learning. Machine learning algorithms can be used in healthcare settings while maintaining data privacy. They can give precise and quick skin cancer diagnosis, which is critical for effective therapy. Furthermore, machine learning algorithms are capable of optimising interaction rounds, resulting in improved rates of accuracy with smaller interactions [56]. Deep learning models, on the other hand, such as CNNs, are created specifically for the analysis of images and can classify skin lesions with more accuracy.

1.3.5 Computer-Aided Diagnosis Systems for Skin Cancer Detection

Computer-aided diagnosis (CAD) systems integrate ML algorithms into clinical workflows to aid healthcare professionals in diagnosing skin cancer. These systems analyze dermoscopic images and provide a quantitative assessment or a probability score for malignancy. As a second opinion, CAD systems can help dermatologists diagnose patients more accurately and reduce subjectivity.

1.3.6 Hybrid Approaches

Hybrid approaches combine different ML techniques to leverage their respective strengths. For example, a hybrid approach might employ dimensionality reduction techniques like PCA or t-SNE to reduce the feature space, followed by a supervised learning algorithm like SVM or ANN for classification. To enhance overall robustness and accuracy in skin cancer detection, a hybrid

approach integrates multiple machine-learning algorithms. This strategy often involves combining feature extraction techniques, dimensionality reduction methods, and classification algorithms [57].

In a hybrid approach, the results obtained from the dimensionality reduction stage are utilized as input for the classification algorithm. This transformation of the feature space is designed to enhance its suitability for classification, leveraging the discriminative information extracted from dermoscopic images. The hybrid approach effectively amalgamates the advantages of dimensionality reduction methods, feature extraction techniques, and classification methods to boost the correctness and dependability of melanoma detection. By seamlessly integrating these components, the hybrid approach can significantly enhance the overall performance of the skin cancer detection system, contributing to early diagnosis and timely intervention [58].

In conclusion, various ML-based mechanisms have been employed for skin cancer detection. Supervised learning algorithms like SVMs, ANNs (particularly CNNs), and Random Forests have achieved remarkable performance. Unsupervised learning, ensemble methods, deep learning, CAD systems, and hybrid approaches also contribute to accurate and efficient skin cancer detection. Supervised, Unsupervised and Hybrid classification methods are three fundamental approaches in machine learning as presented in Table 1.2.

Table 1.2 Classification Methods for Skin Cancer Classification and Detection

Fundamental Approaches	**Classification Methods**	**Limitations**
Supervised Classification [59],[60],[61], [62], [63], [64], [65], [66]	• Decision Trees • Support Vector Machines (SVM) • K-Nearest Neighbors (k-NN) • Logistic Regression • Neural Networks	• Prone to overfitting, especially with insufficient training data. • Requires large amounts of computational resources for training, especially for deep networks.
Unsupervised Classification [67], [68], [69], [70], [71], [72], [73], [74], [75]	• K-means clustering • Hierarchical clustering • Density-based clustering (DBSCAN) • Gaussian Mixture Models (GMM) • Self-organizing maps	• Sensitive to the choice of initial parameters. • Sensitive to the choice of initial centroids. • May struggle with datasets of varying densities or with high-dimensional data

	(SOM)	
Hybrid Classification [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [88], [89]	• Semi-supervised Learning • Active Learning • Transfer Learning • Ensemble Methods • Self-supervised Learning	• Performance heavily depends on the quality and representativeness of the labelled and unlabeled data. • Fine-tuning pre-trained models may require significant computational resources. • Interpretability may be compromised due to the combination of multiple models. • Designing effective pretext tasks may be challenging and domain-specific.

Each classification method has its unique strengths and weaknesses, and the choice of method depends on various factors such as the nature of the data, task requirements, computational resources, and interpretability considerations. Hybrid methods aim to overcome the limitations of individual approaches by combining multiple techniques to achieve superior performance in classification tasks.

1.4 Problem Statement

Skin cancer is becoming more prevalent than all other types of cancer, making it a foremost public health concern across the globe. Current approaches for detecting and classifying skin lesions rely mainly on dermatologists' visual examination, which is arduous and time-consuming. Furthermore, clinicians' visual screening cannot ensure entire detection and may result in potential harm, such as unneeded treatments or undetected lesions that could lead to fatal results. Additionally, the traditional approach to detecting skin cancer lacks accurate criteria for identifying lesions, resulting in differences in diagnosis. A system to classify and detect skin lesions should therefore be developed that can enhance dermatologists' ability to diagnose them and improve their accuracy in detecting and categorizing carcinomas with a view to reducing the potential adverse effects and risks associated with manual visual screening.

This research focuses on addressing the challenge of developing a precise and efficient expert system for the classification and detection of skin cancer, utilizing image processing and machine learning techniques. Despite notable advancements, current methodologies still lack a satisfactory means to differentiate between healthy and cancerous skin lesions.

One of the main challenges in the development of skin cancer CAD systems is image segmentation, which plays a critical role. The primary goal of image segmentation is to identify and isolate the region of interest (ROI) within the image. The use of image segmentation highlights the significant features of images while hiding the unnecessary ones. Accurate segmentation is essential to get more precise and obvious features for melanoma classification. It is crucial for accurate melanoma categorization and detection to eliminate normal tissue and extract more illustrative information from the disease using segmentation. Overall, the similarity and dissimilarity of numerous image properties throughout the segmentation technique determine the region of interest in a segmented image. There are several challenges to consider when implementing segmentation, including selecting effective segmentation algorithms, monitoring algorithm performance and its effects on image analysis, and recognizing specific portions in an object.

An additional significant challenge lies in extracting pertinent features from dermoscopic images. Given the diverse shapes, colors, and textures of skin lesions, it becomes imperative to pinpoint discriminative features capable of effectively distinguishing between benign and malignant lesions. Current feature extraction methods frequently depend on manually crafted features, which may not completely grasp the intricate and nuanced patterns indicative of malignancy. Consequently, there is a demand to delve into advanced feature extraction techniques that can autonomously learn and portray relevant features from dermoscopic images.

A pivotal challenge involves the meticulous selection and optimization of machine-learning algorithms for skin cancer diagnosis. Previous studies have explored classifiers like ANNs, random forests, and SVMs. However, the quest for the most effective classifier for accurate skin cancer classification remains ongoing. These classifiers' success is closely linked to crucial variables such as the quality and amount of the training dataset, the prudent selection of appropriate features, and the cautious selection of hyperparameters. In the realm of skin cancer diagnostics, optimising these factors to achieve excellent accuracy poses a significant difficulty.

Furthermore, skin cancer detection techniques must be made more efficient. Real-time diagnosis is crucial for timely intervention and treatment. However, some existing approaches may suffer from high computational costs and slow processing speeds, hindering their practical implementation. It is essential to explore techniques that can improve the efficiency of machine learning algorithms without sacrificing diagnostic accuracy. Additionally, the availability of annotated and diverse dermoscopic datasets poses a challenge in the creation and assessment of adaptable machine-learning techniques. Critical for training and evaluating skin cancer detection systems is access to extensive, representative datasets with precise annotations.

1.5 Objective of the Study

There has been a lot of interest in using computer technologies to detect skin cancer over the last several decades. The aim of such systems is to give an additional conclusion on the diagnosis with higher accuracy, fewer faults and dependability than a human expert would ordinarily achieve. Several studies have been conducted to improve the automated identification of melanoma. The potential advantages of such research are enormous and incomprehensible. Furthermore, the difficulties are numerous, and new contributions in the field are highly valued. However, it is commonly acknowledged that more reliable and precise detection systems demand more precision. Hence, a novel system for skin cancer detection using machine learning is proposed. The objectives of this research are defined below based on research gaps discussed in chapter 2.

1. Development of the segmentation method to find ROI and extract the actual shape of the lesion.
2. Development of a feature extraction method which extracts the structural and textural data.
3. To design a reliable expert system for the classification and detection of skin cancer.
4. Performance analysis based on various measures.

1.6 Thesis Organization

This section includes supplementary information that supports the main content of the thesis. The outline of the thesis is presented in this section.

The thesis unfolds in a structured manner, commencing with Chapter 1 where the research topic, skin cancer detection, is introduced, emphasizing its significance. This chapter also defines the problem, outlines objectives, and specifies the investigation's scope and limitations. Chapter 2 delves into a comprehensive literature review, exploring various aspects of existing research on skin cancer detection, critically analysing strengths, limitations, and gaps in the literature. Moving to Chapter 3, the proposed methodology and system are elucidated, detailing data collection, dermoscopic image acquisition, segmentation using background subtraction, and feature extraction with the Differential Analyser algorithm. Chapter 4 navigates classification methods, including customized classifiers for skin cancer detection and an exploration of other considered classifiers. Chapter 5 presents experimental results, describing the setup, dataset, and evaluation metrics, with a comparative analysis of classifier performance. Chapter 6 synthesizes findings and draws conclusions, summarizing main discoveries, highlighting contributions, and suggesting future research directions for the field of skin cancer detection.

1.7 Summary

The objective of this chapter is to offer insights into human skin, skin cancer, and a review of past research discoveries and diagnostic tools. It underscores the importance of a meticulous diagnostic assessment for achieving accurate results and preventing misinterpretation. The prevalence of skin cancer is on the rise globally, making early detection crucial for increasing patients' chances of survival. Consequently, prioritizing early diagnosis emerges as a promising strategy for reducing the mortality rate associated with skin cancer. The subsequent chapter 2 delves into various state-of-the-art methods for skin cancer classification and detection. Researchers stand to benefit from a comprehensive understanding of human skin, diagnostic techniques, and previous research findings in this field.

CHAPTER 2

LITERATURE REVIEW

Understanding the advances, obstacles, and advancements in automatic skin cancer diagnosis begins with a review of the existing literature in this field. It allows researchers to build on existing information and contribute to the continued development of precise and dependable diagnostic tools. Continuous information collection and the adoption of innovative approaches are required for more effective skin cancer diagnosis and, eventually, better patient outcomes. This section examines the pertinent literature, emphasizing the important research, findings, and methodology used. Furthermore, it emphasizes the significance of collecting and using the necessary knowledge in order to develop a reliable skin cancer detection system.

2.1 Background of Cancer Detection Systems

Several experimental studies have been launched in order to improve skin cancer diagnosis and develop automated detection technologies. Precision in skin cancer classification and detection is critical in the field of dermatology. The ability to quickly and accurately identify the kind and stage of skin cancer is critical for developing effective treatment programs and ensuring excellent patient outcomes. Several previous research have been performed to address the difficulty of dermatological classification, with a specific focus on skin cancer diagnosis.

He et al., 2016 [90] introduced a groundbreaking architecture, ResNet has significantly impacted diverse image recognition tasks, notably in skin lesion classification. ResNet addresses the problem of deep neural network training by incorporating residual blocks. These blocks use skip connections to allow data from previous layers to bypass particular layers and go directly to subsequent ones. This novel design improves the network's capacity to acquire and retain important characteristics during training, boosting its performance in picture classification, including the difficult task of skin lesion analysis. This design mitigates the vanishing gradient problem and enables training very deep networks with improved accuracy. In skin lesion classification, ResNet's ability to handle deep architectures is particularly valuable. Skin lesion images are complex, with fine-grained textures and intricate patterns. ResNet's skip connections facilitate the extraction of informative features, capturing both low-level and high-level image details, essential for accurate lesion classification.

Szegedy et al., 2016 [91] presented a new perspective on the Inception architecture and its applications in computer vision tasks, which may have implications for skin lesion classification models. One of the key contributions is the introduction of "Inception-v3," a refined version of the Inception network with significant adjustments. The authors incorporate factorized convolutions, which break down standard convolutions into smaller, separate convolutions, thereby reducing computation and improving efficiency without compromising accuracy. They also introduce "bottleneck" layers that decrease the computational burden by reducing the number of input channels while retaining important information.

Xie et al., 2017 [92] the researchers proposed an ensemble model that integrates multiple neural networks, collaboratively contributing to the decision-making process in classification. This innovative approach aims to harness the collective power of diverse neural networks to enhance the accuracy and reliability of melanoma classification from dermoscopy pictures. The study represents a significant step forward in leveraging ensemble techniques for improved skin lesion classification and sets the stage for further advancements in the field.

Zhang et al., 2017 [93] introduced an efficient technique for precisely calculation the Hausdorff distance in 3D point sets, with potential applications in assessing spatial disparities in skin lesion images. In the realm of medical imaging and dermatology, the precise separation of skin lesions from 3D images is critical for analysis and treatment preparation. Evaluating the quality of segmentation is paramount to ensuring the accuracy and reliability of computer-aided diagnostic systems. The Hausdorff distance, as employed in this context, serves as a metric for quantifying the maximum spatial variation between two-point sets. In the context of skin lesion analysis, this would correspond to the variances among the segmented lesion and the actual data obtained from medical experts.

Liang et al., 2017 [94] presented a review of deep learning methods that may have transferable applications in skin lesion classification. Deep learning techniques have demonstrated remarkable success in processing and understanding complex data, such as natural language text. The paper investigated several deep learning architectures, encompassing Recurrent Neural Networks (RNNs), Transformer models, and CNNs, underscoring their efficacy in extracting significant features from textual data. Within the domain of dermatology and healthcare, the accurate classification of skin lesions stands as a pivotal task, where precise diagnosis plays a critical part in the primary detection and treatment of skin problems.

Jana et al., 2017 [95] the paper delves into the detection of skin cancer cells by means of image processing techniques, which are pertinent to the classification of skin lesions. Preprocessing techniques were applied to improve image quality and reduce noise, ensuring that subsequent

analyses were based on reliable data. Segmentation played a pivotal role in identifying individual cancer cells by isolating them from surrounding tissue. Sophisticated segmentation algorithms were likely explored to accurately delineate cancerous regions, aiding in lesion boundary detection and subsequent classification. Feature extraction emerged as a crucial step where relevant characteristics of cancer cells, such as texture, shape, and color, were quantified. The extracted features served as inputs for classification algorithms capable of distinguishing between benign and malignant cells or classifying different types of skin cancer.

Farooq et al., 2019 [96] the study focused on sophisticated deep-learning approaches for the prodromal stages of skin cancer, to contribute to the early detection of skin abnormalities. Detecting skin cancer in its early stages is critical because it allows for timely diagnosis and treatment, increasing the likelihood of a successful outcome. The article digs into a variety of advanced deep-learning approaches used in skin cancer classification. These strategies included cutting-edge transfer learning, CNNs, attention mechanisms, recurrent neural networks (RNNs), and others.

Brinker et al., 2019 [97] study compared deep neural networks' performance to skin specialist in malignance classification, demonstrating the potential of AI in this field. The study addressed the challenge by evaluating the efficacy of DNNs in automated melanoma classification. The paper detailed the architecture of the DNN model used, which involve CNNs or other deep learning architectures. These models are capable of automatically learning intricate patterns and features from melanoma images, contributing to their superior performance in classification tasks.

Zhang et al., 2019 [98] author explored attention residual learning, particularly in the context of skin lesion classification, emphasizing the potential of attention mechanisms to enhance model performance. Attention mechanisms empowered the network to selectively emphasize crucial regions within the input image. This selective focus enabled the model to concentrate on discriminative features while effectively ignoring irrelevant or noisy information. Incorporating attention mechanisms into the residual learning architecture further enhances the network's ability to learn and represent complex features effectively. The residual connections facilitate the flow of information across layers, enabling the model to learn residual features from the input and refine them through attention mechanisms.

Shorten et al., 2019 [99] author delved into diverse image data augmentation techniques employed in deep learning, particularly crucial for expanding limited skin lesion datasets. While CNNs have established notable achievement in image classification tasks, their efficacy is heavily dependent on a substantial amount of labelled data for training. In medical domains like dermatology, acquiring such labelled data can be both scarce and costly. Therefore, the exploration

of effective image data augmentation methods becomes essential to enhance the demonstration of deep learning models when faced with inadequate training data.

Wu et al., 2020 [100] author described a method for detecting skin lesions utilizing highly linked convolutional networks with attention residual learning. For enhanced classification accuracy, the approach uses attention processes to focus on relevant features. The suggested method combines two highly effective deep learning techniques: densely linked convolutional networks (DenseNet) and attention residual learning. DenseNet is known for its ability to alleviate vanishing gradient problems and enhance feature reuse, making it well-suited for image classification tasks.

Basly et al., 2020 [101] described approach for human activity recognition, which combines CNN and SVM, may apply to skin lesion categorization problems. CNNs have proven to be particularly effective in picture identification tasks in the realm of skin lesion categorization, utilizing their capability to autonomously acquire hierarchical features from datasets. The combination of CNNs and SVM, both of which are well-known for their performance in classification tasks, has the potential to provide a robust and versatile methodology for reliably classifying skin lesions based on image data. They excel at capturing visual patterns and features in skin lesion images, making them a popular choice for such tasks. However, CNNs can sometimes suffer from overfitting, especially when dealing with inadequate training dataset. However, SVMs are effective in handling small datasets and are known for their strong generalization capabilities.

Chaturvedi et al., 2020 [102] the study investigated the practice of advanced convolutional neural networks for multi-class cancer classification, perhaps providing insights for lesion classification. In the field of skin cancer classification, accurate identification of distinct forms of skin lesions is critical for timely diagnosis and treatment. Traditional approaches usually require manual feature design, which is a time-consuming procedure that may fail to capture complicated patterns. Deep CNNs, on the other hand, have shown substantial success in image recognition tasks, making them well-suited for the automated classification of lesions because of their capacity to learn detailed patterns and features straight from data. The author explored the potential of CNNs in a multi-class setting, where the model is trained to distinguish between various types of skin cancers. This approach could be beneficial in real-world scenarios where lesions may have overlapping features, and a comprehensive classification system is required. The use of deep CNNs enables the model to learn hierarchical representations of skin lesion images, automatically extracting relevant features at different levels. This capacity allows the CNN to capture intricate patterns, textures, and spatial information that may be indicative of specific skin cancer types.

Labani et al., 2021 [103] the author presented comprehensive epidemiological data on the occurrence of these cancers in Indian and global populations, providing a basis for comparative

analysis. The study may delve into factors that contribute to the incidence of skin cancers, such as exposure to ultraviolet (UV) radiation, genetic predisposition, and lifestyle behaviours. Understanding these risk factors could help identify high-risk groups and guide preventive measures to reduce the prevalence of skin lesions and skin cancer cases.

Wu et al., 2022 [104] the paper dealt with salient object detection, the use of bi-stream networks with a small training dataset could be valuable in skin lesion classification. Skin lesion classification require robust and accurate feature extraction methods to distinguish between different types of lesions. Bi-stream networks are designed to process and combine information from multiple streams, leading to enhanced feature representations. By leveraging this network architecture, researchers can potentially extract more discriminative features from skin lesion images, even with a small training dataset. The advantage of using bi-stream networks is their ability to capture different aspects of skin lesions. For instance, one stream could focus on fine-grained texture details, while another could emphasize global shape information. By fusing the information from both streams, the model can better comprehend the diverse characteristics of skin lesions, potentially leading to improved classification performance.

Wu et al., 2022 [105] focused on salient object detection, its use of recursive multi-model complementary deep fusion and parallel sub-networks could inspire advancements in skin lesion classification methods. The key contributions of the paper lie in its novel approach to salient object detection, which involves a recursive multi-model complementary deep fusion technique. Multiple sub-networks are built in parallel in this method, each optimised for obtaining various features from the given input data. These sub-networks then complement each other's outputs through recursive fusion, leading to enhanced feature representations and improved detection accuracy. The concept of using parallel sub-networks for feature extraction and then combining them to exploit their complementary strengths can be translatable to skin lesion classification tasks. The high complexity and diversity of skin lesion images often require a robust feature representation to accurately differentiate between various types of lesions.

Sundari et al., 2023 [106] introduced an automatic classification system designed for skin lesions, employing deep-learning neural networks, with a specific focus on DCNNs. The authors enhanced the performance of CNNs by fine-tuning layers through various methods, including VGG-16, InceptionResNet, InceptionV3, and DenseNet. Notably, the proposed system attained an accuracy level comparable to that of dermatologists in detecting malignancy, underscoring the efficacy of deep learning algorithms within this domain. The significance of timely recognition and treatment of skin diseases was emphasized in the paper, acknowledging their profound impact on patient's quality of life and the potentially life-threatening nature of certain conditions. The

integration of DCNNs in cancer classification and detection was identified as a promising approach that surpassed the capabilities of traditional image processing methods.

In the concurrent year, Abdelhakim et al., 2023 [107] proposed a model that combined deep learning and reinforcement learning algorithms to facilitate the early detection and classification of skin cancer using dermoscopic images. In the application of this model, various pre-processing techniques were employed, alongside the utilization of the watershed algorithm for segmentation, and a DCNN for classification purposes. Remarkably, the proposed model achieved a notable accuracy rate of 80% in effectively classifying skin cancer into seven distinct types. The study underscored the successful amalgamation of reinforcement learning with deep learning, particularly in the context of skin cancer classification tasks. To train and evaluate the proposed model, dermoscopy images sourced from the ISIC database were employed, encompassing seven types of skin cancers. The research aimed to address the investigation of the CNN-DQN model's performance concerning the classification and detection of skin lesions. The methodology adopted was rooted in deep reinforcement learning, presenting a novel and innovative approach for the classification and detection of skin cancer.

Tembhurne et al., 2023 [108] proposed an ensembled method for lesion detection and achieved an accuracy of 93% on the ISIC dataset. In the proposed framework, a novel segmentation method “background subtraction with midpoint analysis” for finding ROI and DAA for feature extraction has been proposed. For the classification of skin cancer images, different classifiers have been implemented. It is observed that SVM perform better among other classifiers, then the state-of-the-art methods and achieves an accuracy of 95.30%.

Avanija et al., 2023 [109] this study delved into the critical issue of skin cancer detection, emphasizing its life-threatening nature and the imperative need for prompt and accurate identification. Skin cancer arose from unchecked cell proliferation due to DNA damage, leading to mutations and the formation of malignant tumors. Traditionally, machine learning had been employed for cancer identification, but manual feature creation could be laborious. Deep Learning offered an automated solution through convolutional-based deep neural networks, leveraging the ISIC Skin Cancer Dataset for detection. The study underscored the significance of precise identification, as inaccuracies could have severe consequences. Standalone machine learning models might have lacked reliability and accuracy, and their generalizability to real-world scenarios remained a concern. To address these challenges, the study emphasized comprehensive data curation, enhancement, regularization, and model design. Moreover, the research proposed the use of ensemble learning to improve prediction precision, acknowledging the importance of

amalgamating decisions from individual learners. Through these efforts, the study identified crucial factors for skin cancer classification, offering a dependable solution for this critical healthcare issue.

Sharafudeen and Chandra 2023 [110] introduced of a novel method offered a promising avenue in the realm of skin lesion classification. By harnessing features derived from both patient data and EfficientNets, the researchers put forth a comprehensive approach aimed at enhancing classification accuracy. Their method, as reported, yielded notable results, achieving an accuracy rate of 94.13% on the ISIC 2018 dataset and 91.93% on the ISIC 2019 dataset. This study contributed to the growing body of literature on skin lesion classification by proposing a methodology that integrated diverse sources of information. By leveraging features extracted from patient data alongside those from EfficientNets, the researchers demonstrated a holistic approach to classification that capitalized on both clinical and computational insights. The reported accuracy rates underscored the efficacy of the proposed method in accurately discerning between benign and malignant skin lesions across different datasets. Such findings not only validated the utility of the approach but also suggested its potential applicability in clinical settings for aiding dermatological diagnosis.

Nadiger et al., 2023 [111] the study demonstrated the effectiveness of neural networks in improving skin cancer identification and diagnosis. One study achieved an overall accuracy rate of 90.73% by utilizing neural networks and proposed an automatic melanoma diagnosis system. Another study developed a Convolutional Neural Network (CNN) with a 90% accuracy rate in classifying skin rashes as benign or malignant. Additionally, a methodological approach presented in one study utilized neural networks to classify skin lesions with a similar overall accuracy rate of 90.73%. The aim across these studies was to enhance classification results and overall accuracy, with the potential to increase the number of output classes as more data became available. Furthermore, a proposed automatic melanoma diagnosis and inspection system achieved a correct categorization rate of 86% using photos of skin rashes taken with a consumer-level digital camera. These findings underscored the promising potential of neural networks in automated skin cancer diagnosis and identification systems.

Avanija et al., 2023 [112] delved into the critical issue of skin cancer detection, emphasizing its life-threatening nature and the imperative need for prompt and accurate identification. Skin cancer arose from unchecked cell proliferation due to DNA damage, leading to mutations and the formation of malignant tumors. Traditionally, machine learning had been employed for cancer identification, but manual feature creation could be laborious. Deep Learning offered an automated solution through convolutional-based deep neural networks, leveraging the ISIC Skin Cancer Dataset for detection. The study underscored the significance of precise identification, as inaccuracies could have severe consequences. Standalone machine learning models might have

lacked reliability and accuracy, and their generalizability to real-world scenarios remained a concern. To address these challenges, the study emphasized comprehensive data curation, enhancement, regularization, and model design. Moreover, the research proposed the use of ensemble learning to improve prediction precision, acknowledging the importance of amalgamating decisions from individual learners. Through these efforts, the study identified crucial factors for skin cancer classification, offering a dependable solution for this critical healthcare issue.

Chaugule et al., 2023 [113] in the realm of skin cancer detection and categorization, recent research has aimed to develop an automated system utilizing machine learning algorithms. This endeavor relies on a dataset comprising dermatoscopic images sourced from diverse origins, covering various forms of skin lesions, including malignant melanoma, basal cell carcinoma, and squamous cell carcinoma. Throughout these studies, several key phases have been identified. Initially, researchers focus on preprocessing methods such as noise reduction, normalization, and feature extraction to enhance image quality. Subsequently, a comprehensive array of characteristics such as color, texture, and portrait features are extracted from the processed images. These attributes serve as crucial inputs for a range of machine learning models employed in the research, including Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), and Random Forests.

Yaqoob et al., 2023 [114] examining the utilization of federated learning in the detection and categorization of skin cancer. Shedding light on the efficacy of conventional machine learning methods in diagnosing skin cancer and the utilization of federated learning techniques in healthcare systems that prioritize privacy. Identifying relevant search terms in line with the research inquiries and creating search queries using logical operators. Providing a comprehensive overview of federated learning and its integration into the healthcare sector. Assessing solutions devised using traditional machine learning for predicting skin cancer and federated learning algorithms for detecting and categorizing skin cancer.

Jadhav et al., 2023 [115] the non-invasive medical computer vision techniques were implemented for automated analysis of images in melanoma detection. The process involved preprocessing, segmentation, feature extraction, and classification through the utilization of a convolutional neural network (CNN), which resulted in an impressive classification accuracy of 92.1%.

Melanoma has garnered a significant role in the research arena due to an upsurge in the count of new incidents worldwide each year. Particularly in developing software solutions to recognize and evaluate dermoscopy images of patients' skin lesions. The range of methodologies is wide, and each study covers a distinct combination of techniques such as preprocessing, segmentation, low-level

feature extractions and various machine-learning approaches. The basic steps involved in classifying skin cancer images is presented in Figure 2.1.

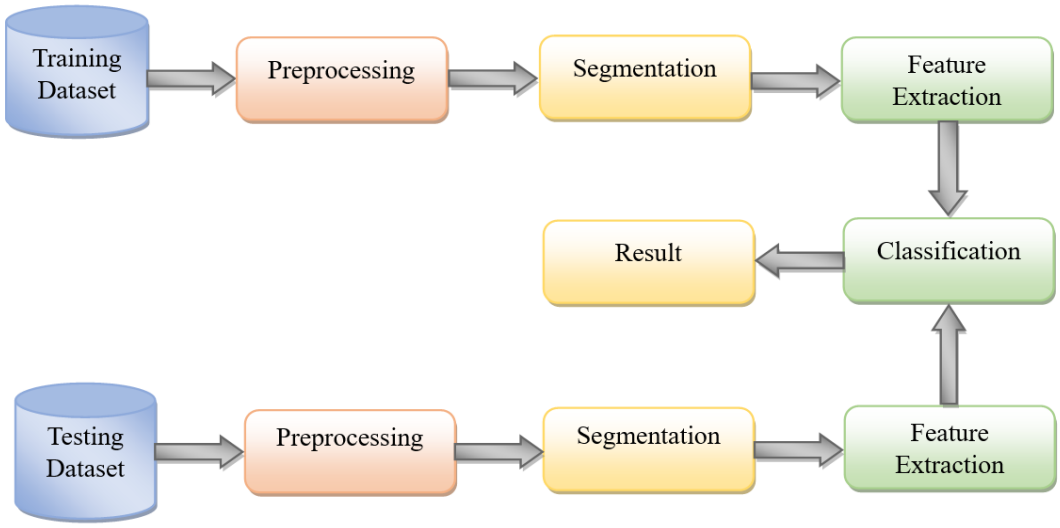


Figure 2.1 Phases of Skin Cancer Detection System

The authors conducted an impressive analysis to offer a baseline for the classification of skin cancer. They accomplished this by fine-tuning a previously trained deep convolutional neural network. Machine learning models have shown considerable potential in dermatology, particularly in skin cancer diagnosis. Though these algorithms produce hopeful results, the great variety of real-world skin problems makes using automated differential diagnosis in real-world scenarios difficult.

2.1.1 Skin Cancer Image Preprocessing

The preprocessing of dermoscopic images is a crucial step that plays a pivotal role in advancing automated systems for diagnosing skin conditions. This initial stage encompasses a sequence of operations intended to enhance the quality and relevance of the input data before it undergoes further analysis [116]. Effective preprocessing significantly contributes to the accuracy and reliability of automated diagnostic systems in dermatology. By improving the quality of dermoscopic images by preprocessing, automated diagnostic systems can extract more useful information, resulting in more accurate and reliable skin problem detection [117]. This emphasises the critical significance of preprocessing in creating the groundwork for the success of automated skin health assessment systems. These steps typically include artefact removal and image enhancement [118], [119], [120], [121] as shown in Figure 2.2.

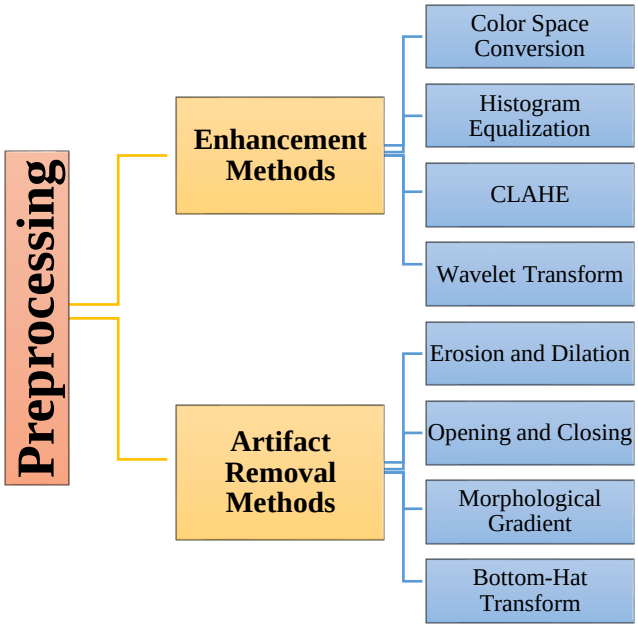


Figure 2.2 Preprocessing Methods for Dermoscopic Skin Cancer Images

Artefact removal is an essential step in preprocessing as it helps to eliminate any unwanted elements or noise present in the image. This can include removing artefacts such as hair, dust, scratches, or shadows that may affect the accuracy of the diagnosis. Image enhancement techniques are used to upgrade the brightness and visual of the skin lesions, making it easier for dermatologists to identify and classify them accurately. A review of various preprocessing methods proposed in the past for dermoscopic images is presented in this section.

Lee et al., 1997 [122] proposed as a new technique Dull Razor hair removal that employs a bilinear interpolation method to eliminate hair pixels. This technique comprises three basic steps: (1) morphological closing operation to locate the location of dark hairs; (2) bilinear interpolation to replace hair pixels; and (3) an adaptive median filter to smooth the final output.

Celebi et al., 2009 [19] has demonstrated a method for enhancing dermoscopy images. This approach determines the optimal weights for converting color images into corresponding grey-level images by maximizing Otsu's histogram bimodality measure.

Zhou et al., 2011 [123] introduced an innovative approach that leverages mean shift-based gradient vector flow (GVF) to guide both inner and outer dynamics in the correct direction. This technique initiates by computing force vectors in the image domain as the initial step for energy force. The mean shift of pixels within a defined region constrains the deformation of the area enclosed by the evolving boundary. In order to enhance image segmentation, the researchers integrated a mean-field term into the conventional GVF objective function.

Barata et al., 2012 [124] has developed a new method for extracting melanin regions from dermoscopy images that requires a series of steps. The dermoscopy image is first subjected to a preprocessing technique. A set of directed filters and linked component analysis (CCA) are then used to identify the pigmented network "lines". Finally, attributes from the discovered region are retrieved and applied to train an AdaBoost algorithm to identify each lesion in terms of the melanin region's presence.

Abbas et al., 2013 [125] have demonstrated a hair removal technique that identifies and removes hair in the CIE L*u*v* colour space using morphological operations and thresholding. On these fine systems that are based on their illuminance, a hair mask was constructed using a defined thresholding method. Finally, a morphological operation replaced each masked pixel with an average of its non-masked neighbours.

Huang et al., 2013 [126] evaluated the DullRazor software's efficacy for hair removal using clinical photos and discovered that it was ineffective for getting rid of little or darkened hairs. The author offered a different strategy to deal with this problem. In order to remove hair, this method combines conventional matched filters with region-growing algorithms and uses linear discriminant analysis based on a colour similarity criterion. The suggested method additionally makes use of multiscale matching filters to improve the ability to identify dark and fine hairs in clinical images. The two-stage thresholding method with hysteresis was also used by the author for edge detection and hair segmentation. To distinguish between classes and maximise the ratio of inter-class measures to intra-class measurements, the author also used a matrix of inter-class scatter measures using linear discriminant analysis (LDA).

Aswin et al., 2014 [127] a novel approach for melanoma detection was introduced, integrating genetic algorithm (GA) and ANN techniques. Hair removal from images was executed using the Dull-Rozar technique, while segmentation of region was identified through the Otsu thresholding method. Additionally, the Gray-Level Co-occurrence Matrix (GLCM) technique was applied to extract unique features from the segmented results. Following this, combining ANN and GA, a hybrid classification algorithm was implemented, to classify lesion images into malignant and benign classes.

Lee and Chen, 2014 [128] a technique has been presented to precisely identify the optimal threshold value for defining malignant boundaries in skin scans. The proposed method incorporates the utilization of the fuzzy c-means approach (FCM) in tandem with a type-2 fuzzy set algorithm. In the preprocessing phase, the researchers also implemented the 3D color constancy method to alleviate the impact of shadows and variations in skin tone present in the images. This integrated strategy significantly improves the precision of boundary delineation for malignant skin lesions.

Youssef et al., 2017 [129] presents a multiscale strategy for analysing and identifying skin lesions using dermoscopic pictures, automated segmentation, and classification tools. A pre-processing phase is included in the segmentation procedure to partition the textured image into discrete texture and geometric components using colour texture image decomposition. The geometry component is used for segmentation to obtain ROI, whereas the texture component archives the lesion's textural features. A SVM is employed to classify the features. The proposed method's efficacy and performance are assessed by comparing it to contemporary and robust dermoscopic techniques published in the literature.

Khalid et al., 2017 [130] describes a novel method for identifying melanoma in dermoscopy images. The suggested method employs a preprocessing phase where noise and hair in images are detected using directional filters, as well as an image colouring method to fill in unidentified areas in the images. To outline the border of the lesion in the images, the method employs Markov random field and fuzzy C-Means methods. The research reveals that the strategy performs better other methods for identifying melanoma skin cancer. The approach performs well in automatic image segmentation when tested on dermoscopic image dataset.

Waghulde et al., 2019 [131] suggests using image processing techniques to identify melanoma in digital photos. To remove noise from the image, preprocessing techniques such as median filters are used. For additional analysis, the image is turned into an HSI colour image. The image is segmented using texture segmentation algorithms and active shape segmentation algorithms. For feature extraction, the GLCM Feature Extraction algorithm is used. The image is classified as normal or melanoma using the probabilistic neural network (PNN) classifier.

Netala et al., 2020 [132] presented various approaches and methods for each phase of pre-processing including image enhancement, restoration, and removal of hair. The paper emphasizes the importance of automated skin cancer detection for early diagnosis and saving lives. The study emphasises the potential advantages of pre-processing approaches that can improve the reliability and effectiveness of the systems for skin cancer diagnosis.

Hajiarbabi et al., 2023 [133] proposed a methodology for the preprocessing of images involving the elimination of noise and illumination effects, succeeded by the utilization of transfer learning from the ImageNet dataset to train a convolutional neural network. The subsequent step includes the fine-tuning of the network to specialize in the identification of melanoma as opposed to other benign cancers, with a specific emphasis on the central region of the image where the suspicious area is situated.

Jeslin et al., 2023 [134] delved into the critical role of image preprocessing in refining attributes extracted from images, particularly in the realm of dermoscopic imaging. It highlighted the use of

curvelet features as the foundation of a comprehensive preprocessing algorithm aimed at enhancing accuracy and precision. Various techniques were discussed, including CLAHE for contrast enhancement, the Frangi vessel filter for artefact identification and elimination, and the Fast Marching Method (FMM) for inpainting missing regions. Denoising methods such as median filtering, frequency domain filtering, wavelet transform, and curvelet were also explored, each balancing noise reduction with the preservation of critical image features like edges and textures.

Mukadam et al., 2023 [135] underscored the critical role of preprocessing methods, with a particular focus on the Enhanced Super-Resolution Generative Adversarial Network (ESRGAN). Their study highlighted how employing ESRGAN can significantly enhance image quality, thereby facilitating more effective feature extraction processes. Furthermore, the author showcased a comprehensive approach by integrating various preprocessing techniques. This methodological strategy aimed to optimize the dataset for subsequent analysis and modelling tasks. By leveraging a combination of preprocessing methods, including ESRGAN, the researchers demonstrated a commitment to enhancing data quality and improving the performance of machine learning models in image-related tasks.

Gururaj et al., 2023 [136] addressed the crucial task of image preprocessing within the context of their study. This preparatory phase encompassed several key steps aimed at optimizing the raw image dataset for subsequent analysis. Initially, the researchers conducted sampling of the raw image data, an essential procedure to manage and streamline the dataset for further processing. Subsequently, author applied an approach termed the "dull razor" method, which involved employing specific filtering techniques to effectively eliminate noise artifacts present in the images.

Aydin et al., 2023 [137] the preparatory stage involves a series of processing steps aimed at optimizing the input data for subsequent analysis. Among these steps are normalization, which standardizes the data distribution, and the removal of unnecessary data, which streamlines the dataset for improved model efficiency. By standardizing data through normalization, researchers ensure that features are on a comparable scale, thereby mitigating issues related to varying magnitudes and facilitating more effective model training.

In preprocessing, various challenges are observed in the literature. (i) The original datasets contain images with high resolution, which impose a high cost of calculation. As a result, images must be scaled because direct scaling can cause structural deformation in lesions. (ii) In cases where the lesion and skin tone are very similar, the images may have low contrast. It's perplexing to identify the dissimilarity between an ordinary and a malicious mole in these instances. (iii) Another issue is the existence of noise in the images, such as hair and noise, which might make it difficult to analyse the imaged skin lesions properly.

2.1.2 Skin Cancer Image Segmentation

The process of dividing an image into multiple regions or components is referred to as segmentation. Skin cancer image segmentation is a critical step in the detection and diagnosis of this prevalent illness. The primary goal of this segmentation is to precisely delineate and indicate the boundaries of skin lesions from surrounding healthy skin or background [138]. This segmentation procedure is significant in dermatology and medical image analysis for a range of applications, including skin disease identification, treatment strategy development, and tracking the progression of skin cancer over time. Due to the complex structures of lesions, variations in shape, size, borders, and colour, as well as the presence of artefacts such as air bubbles, hair, and ink markings, skin cancer segmentation is difficult. Previous research has sought to develop computer-based systems for skin cancer detection by segmenting lesion regions from surrounding skin and extracting valuable features [139]. These models, however, were based on traditional machine learning approaches, and their accuracy was restricted. For the separation of ROI in medical images, various segmentation approaches have been suggested and are discussed in this section. The methods including thresholding-based, edge-based, region-based, artificial intelligence-based, and deformable model based [118], [140], [141], [142], [143] as shown in Figure 2.3.

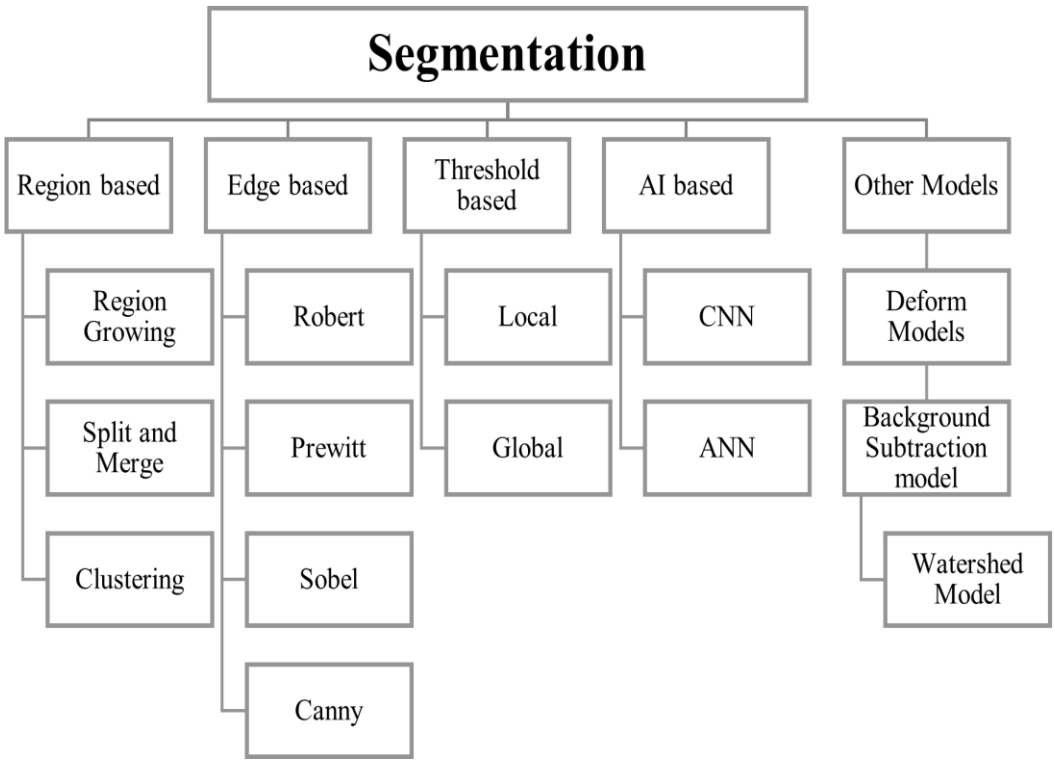


Figure 2.3 Segmentation Methods for Skin Cancer Images

Premaladha et al., 2016 [144] has developed a novel segmentation method known as Normalized Otsu's Segmentation (NOS). NOS addresses the challenge of constant luminance and effectively separates damaged skin tissues from normal tissues. Additionally, they implemented the contrast-limited Adaptive Histogram Equalization (CLAHE) approach and a median filter to enhance image contrast. Furthermore, author developed classification techniques, including Deep Learning-based Neural Networks and a Hybrid Ad Boost- SVM algorithm. These classification methods are provided with fifteen features that have been developed and extracted from the ROI within the images.

Dhane et al., 2016 [145] has proposed a novel technique for scar detection by using digital photographs acquired with a portable optical camera. The method for clustering that is suggested uses a spectral approach and is based on an affinity matrix. The spectral clustering (SC) process entails creating a Laplacian similarity matrix based on the Ng-Jorden-Weiss method.

Zortea et al., 2017 [146] proposed a simple weighted Otsu thresholding method that builds upon the Otsu threshold. The initial threshold takes into account all of the pixels in the image, including cross-diagonal pixels, which are weighted based on an independent estimate derived from the skin pixels in the peripheral region. This method proves to be both simple and effective, particularly when dealing with lesions that are obscured by coarse hairs.

Xie et al., 2017 [147] introduced a skin lesion classification system that divided lesions into noncancerous and cancerous groups. The proposed paradigm is divided into three stages. Initially, a self-generating NN was employed to recover diseases from images. The data on boundary, appearance and colour details were retrieved in the second phase. The dimensionality of the features was reduced using PCA, permitting the optimal amount of characteristics to be determined. In the last step, lesions were classified using a NN ensemble approach.

Satheesa et al., 2017 [148] has presented a non-invasive computerized dermoscopy system for diagnosing skin lesions that take into account the predicted depth of lesions. It is proposed to use 2D dermoscopic pictures to rebuild a 3D skin lesion. Adaptive snake technique is applied to obtain ROI from the data. The calculated depth map is fitted to the existing 2D surface to produce the 3D reconstruction. The depth and 3D form features are extracted based on the generated 3D tensor structure. Feature selection is used to inspect the impact of quality and combinations of features on decision-making.

Choudhary and Biday, 2017 [149] have suggested an ANN-based skin cancer diagnosis model which employs a maximum entropy thresholding approach to segment images. To retrieve distinctive properties of melanomas, a GLCM was used. The input images were categorised into either malignant or benign stages of melanoma by employing a feed-forward ANN.

Esteva et al., 2017 [150] presented a CAD system that is implemented on an ISIC dataset using image processing and machine learning techniques to identify melanoma. Author have used the Histogram equalization technique to enhance image contrast and a dilation technique to eliminate hairs or skin defects. Xie et al., 2017 [147] proposed a self-generating neural net to separate skin lesions and coloured edges; and texture-based features from the lesion are extracted for feature extraction. The proposed work divides skin cancer into two categories such as: cancerous and non-cancerous.

Maity et al., 2018 [151] employed the optimum technique for chromatic aberration, to determine RGB pictures of skin ulcers in a comparative analysis of colour constancy approaches and found that the weighted grey edge method is best. Al-Masni et al., 2018 [152] a unique segmentation approach based on full-resolution convolutional networks (FrCN) was presented in this paper. The proposed FrCN method is designed to capture full-resolution features from every pixel in the input data directly, without the requirement for intermediate steps

Mobeen et al., 2018 [153] lay the framework using segmentation and a convolution neural network on dermoscopy pictures to detect skin lesions with a malignant predisposition. As a dataset, images from ISIC-2016 were used.

Mahmoud et al., 2018 [154] presented a new skin cancer CAD system based on texture analysis methodologies in their research. Hair removal, filtering, extraction of features, and classification are the four processes of the proposed CAD system. The CAD system was used to distinguish between non-malignant skin lesions (common nevi or dysplastic nevi) and malignant skin lesions. Extraction of histograms of oriented gradients (HOG) features after artefact removal offers the best classification results, according to the findings of their experiments.

Pathan et al., 2018 [155] presented a comprehensive review of state-of-the-art approaches used in computer-aided diagnostic systems. The author initiated the review by delving into the domain-specific aspects of melanoma and then proceeded to explore the primary techniques utilized in each phase of the diagnostic process.

Artificial Bee Colony (ABC) was developed by Aljanabi et al., 2019 [156] improved melanoma detection can be achieved through the use of skin lesion segmentation as a means of identifying lesions in dermoscopy images. Chouhan et al., 2019 [157] conducted an extensive survey on skin lesion segmentation using computational intelligence techniques. The survey covered a range of methods, including fuzzy C-means, CNN, and GA. These techniques are commonly applied for image segmentation in various fields, including medical imaging, scientific analysis, engineering, and the humanities. Monika et al., 2020 [158] presented a Colour-based k-means clustering for segmentation since colour is a significant characteristic in determining the kind of malignancy.

Kaymak et al., 2020 [159] FCN-8, FCN-16, FCN-32, and FCN-Alex Net are four FCN architectures generated by using the CNN model for the segmentation of skin cancer images. The network was first put through its paces on the ISIC 2017 dataset. Plot and Jacquard coefficients are employed to evaluate the degree of agreement among the partitioned result and the authentic ground truth. Even though the multiple FCN frameworks performed proportionately well, FCN-8 was discovered to be more effective in terms of segmentation.

Mohakud and Dash 2021 [160] for dermoscopic image segmentation, the author proposes a hyperparameter-optimized Full Convolutional Encoder-Decoder Network (FCEDN). The new Exponential Neighbourhood Gray Wolf Optimization (EN-GWO) algorithm was used to optimize the network hyperparameters. The neighbourhood search strategy in EN-GWO is determined by combining the individual wolf hunting strategy with the global search strategy, and it reflects the appropriate balance of exploration and exploitation. Murugan et al., 2021 [161] the median filter is applied in pre-processing phase, and for the separation of the region of interest, the mean shift method is used.

Gururaj et al., 2023 [136] delved into the process of image segmentation, a fundamental aspect of image analysis and computer vision. Their study highlighted the utilization of encoder and decoder technologies in this segmentation process, emphasizing the adoption of advanced computational techniques to address complex image-processing tasks. The essence of image segmentation, as elucidated by author involved the delineation of distinct regions within an image. This encompassed the separation of foreground elements from the background, as well as the grouping of pixels based on shared characteristics such as color or shape similarity. By employing encoder and decoder technologies, the researchers aimed to enhance the precision and efficiency of this segmentation process.

Aydin et al., 2023 [137] proposed an innovative method that combined histogram-based descriptors from various color spaces, offering a novel and highly effective strategy for categorizing different types of cancer. While color imagery often relied on key point-based features, the incorporation of global features in this context had received limited attention in prior research. The findings highlighted a notable improvement in accuracy rates when employing color images compared to grayscale counterparts, suggesting the potential advantages of leveraging color information for enhanced classification outcomes.

The summary of different existing techniques for skin cancer segmentation is presented in Table 2.1. The table highlights the outcomes of various methods in terms of pros and cons from the past few years.

Table 2.1 Different Techniques for Skin Cancer Segmentation

Author	Dataset	Method	Contribution and Performance Measures	Remarks
Giotis et al., (2015) [162]	Digital archives	• K-means • ABCD rule • CLAM	Proposed a system for melanoma detection with an accuracy of 80%	• Noise sensitive • Depends on the size of the data
Satheesha et al., (2017) [148]	ISIC PH2 ATLAS	• Adaptive snake • 3D Tensor	A non-invasive computerized dermoscopy technique for diagnosing skin lesions based on their estimated depth was described.	• For reconstruction, the actual depth of the lesion cannot be computed
Yuan et al., (2017) [163]	ISIC PH2	• Deep CNN	Proposed an automatic segmentation approach for skin cancer images with a 95.50% accuracy.	• A deeper network needs more training samples to avoid overfitting
Singh et al., (2018) [164]	Images from different sources	• Computerized tools for melanoma detection	For computerised devices, dermoscopic images are the most suitable solution.	• Segmentation feature extraction and accuracy are major issues in computerized tools
Al Masni et al., (2018) [152]	ISIC PH2	• Fully resolution convolution network	A method for automatic segmentation of skin cancer images was proposed, with an accuracy of 84.97%.	• Highly sensitive • The large data set cannot be handled • Over and under-segmentation are not taken care of
Seeja et al., (2019) [165]	ISIC	• DCNN • FCN	Proposed a segmentation method for skin lesion segmentation and achieved an accuracy of 85.19%	• Notation of dataset
Zaffar et al., (2020)	ISIC PH2	• CNN	Suggests Automated method for lesion	• The dataset was not handled correctly

[166]			boundary segmentation by combining U-Net and ResNet and achieving an accuracy of 85.40%	• Data overfitting
Vani et al., (2021) [167]		• Deep Learning • Active Contour	A deep learning-based strategy for melanoma diagnosis was proposed, and an accuracy of 90.00% was attained.	• Noise is not moderated
Ashraf et al., (2022) [168]	ISIC	• Deep Learning	Proposed deep learning-based automated skin segmentation method and achieved a Jaccard index of 90.20%.	• Low-resolution images were not identified • Skin lesions lose their ground truth annotations.
Tahir et al., (2023) [169]	ISIC HAM10000 DermIS	• Deep learning	A deep learning-based network for skin cancer was proposed, with an accuracy of 94.17%.	• Proposed system is suited for only people with fair skin tone.

In segmentation, images exhibiting higher inconsistencies at the lesion centre compared to the boundary inconsistencies yields poor segmentation results. Furthermore, in some situations, a higher threshold value set for segmentation causes a regression area at the image centre, resulting in the false detection of lesions. Many features that aren't required for classification complicate the classifier and increase computational time, lowering classification accuracy. The best aspects of the skin cancer photographs should depict the region's characteristics. As a result, the finest method to identify lesion features is required. Therefore, a novel system is required to develop which can handle these all challenges and achieve better performance than the state-of-the-art methods.

2.1.3 Skin Cancer Image Feature Extraction

Feature extraction is a critical step in image processing. Each dermoscopy image can produce hundreds of features that serve as image descriptors. However, not all of these characteristics are

important for lesion classification. Unnecessary characteristics can complicate the classifier, raise computational needs, and perhaps impair classification accuracy.

Extraction of feature is an important step in obtaining a discriminatory illustration of the skin lesions. Achieving the right feature is a challenging procedure and has been extensively researched in this area, which may lead to the identification of a wide range of features describing the skin lesion system [170]. The major features that are extracted are texture feature, shape features and colour features. Texture features are basically a kind of spatial variation function in the pixel value of the image. These values are the pixel intensities or greyscale values in the image which are useful in several applications. The texture is also a type of visual content in the image. Shape features are a type of structural information of an image like asymmetry and border irregularities of a lesion. Colour features represent the variation in red, green and blue components of an image.

These features can be classified into 4 categories: i) Handcrafted that include shape, colour, symmetry and texture of the lesion, these features are global image descriptor and are very common one, ii) dictionary-based features, in which to acquire local image descriptors of lesion approaches like bags of features or sparse coding are used, iii) deep learning features those who acquire decent image descriptors automatically by using neural networks or convolution neural networks, and iv) clinical features aimed at providing medical explanations for features used by a computer-assisted diagnostic system [11]. In this section, a brief review of various feature extraction method for skin cancer detection are addressed as shown in Figure 2.4.

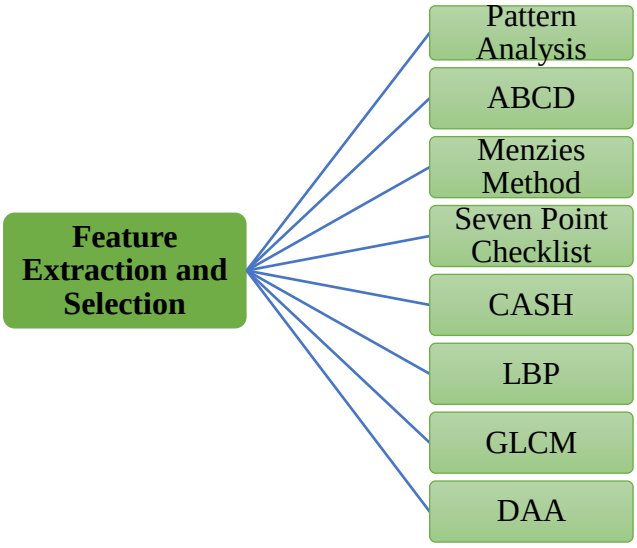


Figure 2.4 Feature Extraction and Selection Methods for Skin Cancer Images

A. Pattern Analysis

In 1987 Pehamberger et. al presented the pattern analysis method and then it is updated in 2000 in the Consensus Meeting of Dermoscopy (CNMD) [171]. For extraction of features in skin lesion diagnosis pattern analysis method have been used. The presence of uneven shapes indicates the global patterns in lesion images. It comprises the reticular formation, bulbous, cobbled, regular, stardust, analogous, multimodule, lacunar and nonspecific shapes. Local patterns of dermoscopic arrangements include pigment structures, spots or globs, strips, blue-white cloak, blemishes or pigment, vitiligo, recession and vascular bundle situated in a particular section of the lesion [172] as shown in Table 2.2.

Table 2.2 Pattern Analysis Method for Feature Extraction

Global pattern	Local pattern
Reticular formation	• Pigmented structure (existing or non-existent/usual or abnormal)
Bulbous	• Spots/globs (existing or non-existent /even or uneven)
Cobbled	• Strips (existing or non-existent/ even or uneven)
Regular	• Blue Whitish cloak (existent or non-existent)
Stardust	• Blemishes or pigment (existing or non-existing/ even or uneven)
Analogous	• Vitiligo (existing or non-existing)
Multimodule	• Recession Structure (existing or non-existing)
Nonspecific (missing)	• Vascular Structure (existing or non-existing)

B. Asymmetric, Border, Color, Diameter Rule

ABCD method of dermoscopy was the next method afterwards pattern analysis [173]. This method is based on some criteria which include asymmetric (A) that is calculated by dividing image horizontally and vertically, border (B) an uneven, blurry and shabby boundary indicate melanoma, color (C) like blue-grey, black, red, white, dark brown, light brown present in lesion indicate melanoma and diameter (D) if the diameter of a lesion id greater than 6nm then it indicates melanoma's presence as shown in Table 2.3. All these criteria are assigned with some score and a total dermoscopy score (TDS) is computed the formulae for calculating TDS given in equation 2.1.

$$TDS = A * 1.3 + B * 0.1 + C * 0.5 + D * 0.5 \quad (2.1)$$

Table 2.3 ABCD Rule Of Dermoscopy

Features	Definition	Score	Weight factor
Asymmetry (A)	• Vertically or horizontally (even or uneven)	0-2	1.3
Boundary (B)	• An irregular, blurred and ragged border	0-8	0.1
Colour (C)	• Presence of 6 key colors (blue-grey, red, dark brown, white, black, light brown)	1-6	0.5
Diameter (D)	• More Than 6nm	1-5	0.5

C. Seven-Point Checklist

This method has been used to attain high accuracy of dermoscopic images in computational diagnosis. This method is based on seven criteria that include three major (unusual pigmented pattern, Blue-whitish cloak and unusual vascular structure) and four minor criteria (Asymmetrical strips, Asymmetrical colouration, Asymmetrical spots/globs and Recession pattern as shown in Table 2.4. In this method, a total score is calculated and if the value of the final score is 3 or greater than that means melanoma is present otherwise not. To calculate a total score each minor criteria is given one point and each major criteria is given two points [174].

Table 2.4 Seven-Point Checklist Method for Feature Extraction

Minor criteria	Score	Major criteria	Score
Asymmetrical Strips	1	Unusual pigmented Pattern	2
Asymmetrical colouring	1	Blue whitish cloak	2
Asymmetrical spots/globs	1	Unusual Vascular structure	2
Recession pattern	1		

D. Menzies's Method

This method is based on nine positive (Blue-white cloak, Various brown spots, Pseudopodium pattern, radiated coursing, white discolouration, bordering black spots/globs, multicolours, Various blue/grey spots, Prolonged structure) and two negative features (Point and axial symmetry of pigmentation, Presence of a single color) present in structure as shown in Table 2.5. A lesion is said

to be melanoma when there is at least one or more positive features are present and none of the negative features is present [175].

Table 2.5 Types of Features used by Menzie's Method for Feature Extraction

Negative features	Positive features
Point and axial symmetry of lesion	• Blue Whitish Cloak (existing or non-existing)
The presence of a single-color	• Various brown spots • Pseudopodium • Pattern radiated • Coursing white • Discolouration • Bordering black spots/globs Multicolours • Various blue/grey spots

E. Color, Architectural, Symmetry, Homogeneity Algorithm

Henning et al., [176] give an updated type of pattern analysis known as the CASH algorithm. This algorithm includes architectural features which were not colours by any other criteria previously. CASH algorithm adds Colour (C), Architectural (A), Symmetry(S) and Homogeneity (H) features for lesion and each feature is given a score, then a total dermoscopy score (TDS) is calculated on the basis of these features if the value of TDS is greater than 7 then the lesion is melanoma as shown in Table 2.6.

Table 2. 6 CASH Algorithm for Feature Extraction

Features	Definition	Score
Colour (C)	• Presence of 6 key colors (blue-grey, red, dark brown, white, black, light brown)	1 for each
Architectural (A)	• Undefined • Visible variation	1 2
Symmetry (S)	• Monoaxial symmetry • Biaxial asymmetry	1 2
Homogeneity (H)	• Network; spots/globs Stripe/pseudopodium Pattern Blue-white cloak	1 for each

	• Recession systems blots • Heterogeneous venule	

Colour present in the CASH algorithm is bright brown, blue, dark black and white, dark brown and red with a score of 1 point each. The colour red, white, and, mainly, blue are probably very crucial in differentiating non-cancerous lesion from the cancerous lesion. Architecture refers to the structural and colour of a lesion a modest with a score of 1 point or noticeable change with a score of 2 points in lesion have a high probability of being a melanoma. Structural features refer to the presence of shape of the lesion which is Monoaxial symmetry with a score of 1 point and biaxial asymmetry with a score of 2 points indicate the presence of melanoma.

Homogeneousness/heterogeneousness is primarily established on previously indication of the multiple patterns, which are characterized as three or more dermoscopic arrangements exist in the lesion. These patterns are network; spots/globs; strips/pseudopodium; blue-white cloak; recession arrangements; blots; and heterogeneous venule each with a score of 1 point [177].

Choudhary and Biday, 2014 [149] suggested an ANN-based skin cancer diagnosis model which employs a maximum entropy thresholding approach to segment images. To retrieve distinctive properties of melanomas, a grey-level co-occurrence matrix (GLCM) was used. The input images were categorised into either malignant or benign stages of melanoma by employing a feed-forward ANN.

Aswin et al., 2014 [127] introduced is a novel approach for melanoma detection that employs a blend of GA and ANN techniques. To eliminate hair from the images, the Dull-Rozar technique was implemented, and ROI were isolated using the Otsu thresholding technique. Moreover, the GLCM method was utilized to derive unique features from the segmentation results. Following this, a hybrid classification algorithm that integrates ANN and GA was applied to classify lesion images into malignant and benign categories.

Premaladha et al., 2016 [144] devised classification methodologies involving Deep Learning-based Neural Networks and Hybrid Ad Boost-SVM algorithms. These approaches were provided with a set of fifteen features created and extracted from the ROIs within the images.

Xie et al., 2017 [147] introduced a skin lesion classification system that categorizes lesions into noncancerous and cancerous groups. The proposed paradigm consists of three stages. Firstly, a self-generating Neural Network (NN) was employed to identify diseases from images. In the second phase, data on boundary, appearance, and color details were extracted. The dimensionality of the features was then reduced using PCA, enabling the determination of an optimal set of characteristics. In the final step, lesions were classified using a neural network ensemble approach.

Satheesa et al., 2017 [148] describe a non-invasive computerized dermoscopy system designed for diagnosing skin lesions, taking into consideration the predicted depth of the lesions. The approach involves utilizing 2D dermoscopic images to reconstruct a 3D representation of the skin lesion. To achieve this, an adaptive snake technique is applied to extract the ROI from the data. Subsequently, a calculated depth map is fitted to the existing 2D surface, resulting in the creation of a 3D reconstruction of the skin lesion. The depth and 3D form features are then extracted based on the generated 3D tensor structure.

Al-Masni et al., 2018 [152] a distinctive segmentation approach was introduced, leveraging full-resolution convolutional networks (FrCN). The suggested FrCN technique captures full-resolution features directly from each pixel in the input data, without relying on intermediate steps. This eliminates the necessity for pre- or post-processing procedures, such as artifact removal, illumination adjustments, or additional strengthening of the segmented skin lesion boundaries. The use of FrCN allows for a streamlined and efficient segmentation process, emphasizing a direct and comprehensive extraction of features at the full resolution of the input data.

Gulati et al., 2019 [32] proposed a system for melanoma and non-melanoma classification is. Preprocessing measures are first carried out to clarify and improve dermoscopic images by eliminating unwanted artefacts. Then, using active contour-based segmentation, the primary area of interest is constructed. Moreover, colour, structure, and texture are retrieved and fed into the classification model (SVM) for efficient and precise malignant and non-malignant tumour categorization.

Mahbod et al., 2020 [178] a framework is introduced for extracting deep features from several well-established and pre-trained deep CNNs. The deep features are derived from pre-trained models such as AlexNet, ResNet-18, and VGG16. These features are subsequently fed into a multi-class SVM classification model. The outcomes of the individual classifiers are combined to make a final classification decision. The proposed model is validated on the ISIC 2017 database, demonstrating its effectiveness in skin lesion classification.

Monika et al., 2020 [158] in the segmentation phase, a color-based k-means clustering technique is applied, leveraging the significance of color as a crucial characteristic in discerning the type of malignancy. The features extracted include ABCD, and the Grey Level Co-occurrence Matrix (GLCM). These features encompass both statistical and textural aspects, providing a comprehensive set of characteristics for further analysis and classification of skin lesions.

Murugan et al., 2021 [161], the proposed approach employs an image processing method for the classification and detection of skin melanoma. In the preprocessing phase, a median filter is applied to the ROI, and the mean shift method is utilized to separate the ROI effectively. Feature extraction

from the ROI involves Moment Invariant features such as Gray Level Run Length Matrix and Gray Level Co-occurrence Matrix. These extracted features are then utilized for the classification of lesions into malignant and non-malignant categories. The combined use of preprocessing and feature extraction techniques contributes to an effective and accurate skin melanoma detection system.

There are both advantages and disadvantages to feature extraction and selection methods in the analysis of skin cancer images. Though these methods are important to improve the accuracy of machine learning models in detecting skin cancer, they do come with a number of limitations. While dealing with huge datasets or high-resolution images, feature extraction can be highly computational. Handcrafted features that are predetermined and constructed based on prior information are used in many classic feature extraction methods. These traits may not be appropriate for every case of skin cancer. Overfitting can occur when feature selection procedures result in the model performing well on training data but poorly on unknown data.

Table 2.7 Comparative Analysis of Different Feature Extraction Method

Author	**Method**	**Findings**	**Remarks**
Giotis et al., (2015) [162]	• ABCD • CLAM	• Accuracy= 80% • Specificity= 81% • Sensitivity = 80%	• Noise sensitive • It depends on the size of the data
Premaladha et al., (2016) [144]	• Normalised Otsu's • GLMC	• Accuracy = 93%	• Data overfitting • Decision parameters not defined accurately • The dataset was not handled correctly
Kasmi and Mokrani (2016) [179]	• ABCD	• Accuracy= 94.0% • Sensitivity = 91.25% • Specificity = 95.83%	• Small-size melanoma is not detected
Oliveira et al., (2016) [180]	• Canny's edge detector	• Low error probability • It dodges the discovery of double-edge	• Borders of the lesion are not completely spotted • Large sensitivity to noise

	• Thresholding based	• Lesion boundaries correctly detected	• Leads Irregular lesion edges • Sensitive to artefacts
	• Region-based	• Low contrast boundaries detected	• Computationally expensive • Sensitive to noise
	• AI based	• Good accuracy and sensitivity	• Complexity of implementation • Unnecessary steps involved • Time-consuming
	• Active counter	• Identified low-contrast boundaries • Overcome image noise	• The success of segmentation is determined by the original curve's appropriateness.
Pathan et al., (2018) [155]	• Pattern Analysis	• Uses local and global patterns of Dermoscopic	• The criteria used for feature extraction are not sufficient
	• ABCD	• TDS is calculated by using the weight factor	• Lesions less than 4.75 TDS are not detected
	• Menzies Method	• Uses positive and negative features	• Less accurate
	• Seven Point Checklist	• Based on three major and four minor criteria • Melanoma is more likely if you have a total score of three or more points.	• The low specificity and accuracy
	• CASH	• Include a feature architecture	• Low specificity
Satheesha et al., (2018) [148]	• Adaptive snake • Tensor flow	• Sensitivity = 98% • Specificity = 99%	• For reconstruction, the actual depth of the

			lesion cannot be computed
Singh and Gupta (2018) [164]	• Literature survey of computerized tools for melanoma detection	• Dermoscopic images are the most viable option for computerized tools	• Segmentation feature extraction and accuracy are major issues in computerized tools
Al Masni et al., (2018) [152]	• Fully resolution convolution network for segmentation	• Accuracy = 84.97%	• Highly sensitive • The large data set cannot be handled • Over and under-segmentation is not taken care of
Zaqout et al., (2019) [181]	• Thresholding • ABCD	• Accuracy=94.5%, • Sensitivity=82.5% • Specificity=97.5%,	• Sensitive to artefacts • More datasets are needed for validation

2.1.4 Skin Cancer Image Classification

Over the past ten years, there has been a wealth of research focused on the detection and differentiation of both malignant and benign skin cancers. Numerous datasets have been made accessible to the academic community, enabling researchers to employ various methods such as partitioning, merging, grouping, and classification algorithms for the purpose of detecting and managing skin cancer. Every method comes with its particular limitations and has contributed to significant advancements in the medical field, aiding healthcare professionals in their decision-making processes. In the field of image classification, correlation-based methods have grabbed significant recognition due to their capability to capture the underlying relationships between features and classes. Common similarity measures such as mutual information, cross-correlation, entropy, and pattern intensity have been employed in these methods [182].

One of the correlation-based methods used for image classification is mutual information. Mutual information has been extensively employed in machine learning for various tasks, like feature selection and independent component analysis [129]. Mutual information is also frequently used in medical image analysis to align images to the same coordinate system [183]. Another commonly used correlation-based method in image classification is cross-correlation. Cross-correlation is a technique that measures the similarity between two signals by sliding one signal

over the other and calculating the correlation at each position [184]. Unlike mutual information, which focuses on the statistical relationship between two variables, cross-correlation directly compares the intensity of the image pixels [185]. A review of state-of-the-art approaches based on correlation is presented in this section.

Dimou et al., 2006 [186] the author evaluates the use of ensemble classifiers and fusion techniques in diagnostic cancer models to improve accuracy, confidence, and feature space coverage. In two medical situations, the author assesses the performance of an ensemble of eight classifiers based on 15 alternative fusion algorithms. The study uses 11 commonly accepted metrics to measure the correlation of the base classifiers and provides insights for selecting an improved hyper-classifier. The author also discusses the concept of multi-classifier systems and the benefits of combining the outputs of multiple classifiers to achieve higher accuracy.

Truong et al., 2014 [187] the author performed an extensive analysis of double Debye model parameters for the purpose of categorizing non-melanoma skin cancer. Utilizing Pearson correlation, the study aimed to unveil the sensitivity of these parameters and assess the impact of variations in their combinations on the tumor percentage in skin samples. Receiver Operating Characteristic (ROC) plots were employed to pinpoint and assess the parameters with the highest sensitivity, gauging their effectiveness in discriminating between skin cancer and normal skin.

Kavitha et al., 2016 [188] proposed a model to classify breast cancer genes using microarray data, called Correlation-based Support Vector Machine Recursive Multiple Feature Elimination (CSVM-RMFE). With the SVM-RFE approach, several irrelevant genes are removed in a single iteration to lessen the computing burden of dimension reduction. Prior to using SVM-RFE, it also focuses on identifying linked genes and isolating a new gene from them in order to increase the classifier's accuracy.

Chutia et al., 2017 [189] author introduced a novel ensemble classification framework that combined Random Forest (RF) with a correlation-based feature selection (CFS) technique for the precise classification of heterogeneous land surfaces displaying homogeneous land cover classes in satellite images. The CFS method was applied to evaluate the significance of features, enhancing the performance of the RF classifier by selecting the most relevant feature set. The research not only compared the effectiveness of the RF classifier against other supervised classifiers but also demonstrated its superiority.

Rajesh et al., 2018 [190] the author proposed a framework for the identification and categorization of various skin disorders. PCA and Linear Discriminant Analysis (LDA) were employed to extract features from the original data. The retrieved features were then condensed to a subset with sufficient discriminatory power for classification, achieved through the application of

the Fisher ratio approach, a feature selection technique. To categorize the diverse skin illnesses within the dataset, ensemble-based classifiers, including Bayesian, self-organized map, and SVM, were utilized.

Sinayobye et al., 2019 [191] the author employed a hybrid classification model that incorporated a correlation-based filter feature selection technique with machine learning classifiers to identify relevant features and assess their performance. A smart meter dataset was used to train the model, predict outcomes, and evaluate the classification performance of various classifiers. In the feature selection process, features were grouped based on their relative correlation coefficient with respect to the class attribute, and the top K features were chosen to generate a reduced dataset. According to the results, the Random Forest classifier outperformed the other classifiers used in the experiment.

Murugan et al., 2019 [192] author applied a watershed segmentation approach to segment the region of interest. Shape, ABCD rule, and GLCM were among the features collected from the segmented regions. For classification, three classifiers were used: Random Forest, KNN, and SVM. The SVM classifier outperformed the other classifiers in terms of skin lesion categorization.

Rajasekhar et al., 2020 [193] presented a method for classifying skin lesions as melanoma from dermoscopic images. This method makes use of Shape, Colour, and Texture data, as well as the brand-new Spectral Graph Wavelet Transform (SGWT) for texture feature extraction. The SGWT is better than traditional wavelet transformations because it can handle images with irregular shapes by employing weighted graphs that are produced as meshes. The suggested method classified characteristics derived from the dermoscopic pictures using Nave Bayes, SVM and K-Nearest Neighbor (KNN) classifiers.

Murugan et al., 2021 [161] introduced is a system for identifying malignant skin cancer utilizing a support vector machine. The skin images underwent pre-processing with a median filter and subsequent segmentation through Mean Shift segmentation. Feature extraction from the segmented images employed three strategies: GLCM, Grey Level Run Length Matrix (GLRLM), and Moment Invariants. Classification of the extracted features utilized RF, SVM, and PNN classifiers, along with a combination of SVM+RF classifiers. In direct comparison with alternative classifiers, the combined RF+SVM classifier demonstrated superior performance.

Abbas et al., 2022 [194] employed the Wisconsin Breast Cancer Dataset (WBCD)[1] and the Duke Breast Cancer Dataset (DBDS)[2] in this research to predict breast cancer tumors. Linear Discriminant Analysis (LDA) feature selection was integrated with a variety of machine learning classifiers, including Neural Networks, SVM, RF, and Decision Tree (DT).

[1] https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data/discussion/62297

[2] https://sites.duke.edu/mazurowski/resources/breast-cancer-mri-dataset/

Hajiarbabi et al., 2023 [133] author proposed a methodology that encompasses the creation and execution of a multi-scale architecture for the detection of skin cancer, which integrates the results from three distinct Convolutional Neural Networks (CNN) and channels them into a Fully Connected Network (FCN). Furthermore, the application of image processing methodologies is adopted to ameliorate both the quality and quantity of training images by incorporating techniques such as cropping and scaling to significantly augment the dataset size.

Abdelhafeez et al., 2023 [195] proposed a novel approach for identifying skin lesions in computer-aided diagnostic (CAD) systems, addressing the complexities arising from variations in melanoma size and texture. The method combined deep learning-based layer fusion with neutrosophic set techniques. Off-the-shelf networks, including GoogleNet and DarkNet, were evaluated using transfer learning on the ISIC 2019 skin lesion datasets. Initially, individual networks achieved accuracies of 77.41% and 82.42%, respectively. Through a feature fusion methodology, the accuracies were boosted to 79.2% and 84.5%. In the second stage, error-correcting output codes (ECOC) were employed to construct well-trained support vector machine (SVM) classifiers from fused feature maps. Neutrosophic techniques were then applied to resolve ambiguities in classification scores, resulting in an improved accuracy of 85.74%. This outperformed recent proposals and offered trained models and single-valued neutrosophic sets (SVNSs) for public use in relevant research fields.

A comparative analysis of correlation-based techniques in different research areas is presented in Table 2.8. Singhal and Singh, 2015 [196] proposed an approach for automatically detecting Acute Lymphoblastic leukemia (ALL) using shape information taken from pictures of lymphocyte cells. The Correlation based feature group (CFG) technique is used to create a set of sixteen features that may accurately predict whether a lymphocyte cell is normal or blasted, with a 92.30% accuracy. Maleki et al., 2015 [197] the author employs a method known as CFG, in which features are organized into groups based on their correlation values. Following that, unique feature vectors are created for each of these groupings. The KNN and SVM algorithms are then used to perform classification jobs and attain 93.55% accuracy. Jiang et al., 2017 [198] proposed a correlation-based strategy for optimizing the bag-of-visual-words (BOVW) model for image classification by reducing dictionary size while keeping characteristics with strong category relevance. The method entails creating a visual dictionary containing features that have a high correlation to categories and then training it with an SVM classifier.

Zhang et al., 2017 [199] proposed a method for an apple-diseased leaf image database containing three types of apple leaf diseases: powdery mildew, mosaic, and rust. The proposed method was based on image processing techniques and pattern recognition methods, combined with GA and

CFS, and achieved an accuracy of 94.28%. Arora et al., 2017 [200] proposed a hybrid classification technique that employs correlation-based feature selection and regression classification to categorize segmented chromosomes into five categories: straight, overlapping, bending, touching, or noise. On 1592 segmented chromosomes, the approach obtains an overall accuracy of 94.78%. Nahid et al., 2020 [201] implemented a Deep Neural Network (DNN) algorithm guided by structural and statistical information proposed for classifying breast cancer images exhibiting high precision, accuracy, and F-measure values. The author focuses on the Break His dataset and offers a combination of Long-Short-Term-Memory (LSTM) models and CNN for breast cancer image classification, together with SVM and softmax decision-making layers.

Liu et al., 2020 [202] presented. a normal vector correlation-based image denoising method for lung cancer. The lung CT image is split into many multiscale sub-images and pre-processed. The first-level sub-image is subjected to a modified super pixel segmentation algorithm to create a collection of super pixels. In order to separate the lungs, a random forest classifier is used to categorize the super pixels of each sub-image based on the attributes that were retrieved from them. Nasir et al., 2021 [203] proposed a real-time supervised learning strategy for document categorization that employs a deep convolutional neural network (DCNN) and feature selection using the Pearson correlation coefficient. The technique incorporates phases such as feature extraction using pre-trained NN models, data augmentation, feature selection and feature fusion. On the Tobacco3482 dataset, the proposed technique achieves a classification accuracy of 93.1% using a cubic SVM classifier. Khan et al., 2021 [204] employed pattern recognition and machine learning approaches for identifying and recognizing fruit conditions. It addresses obstacles such as convex edges, colour inconsistency, variation, accessibility, scale, and source. On the Plant Village dataset, the suggested method obtains an average accuracy of 93.74%.

Table 2.8 Comparative Analysis of Different Correlation Method

Author	**Different Applications**	**Feature selection method**	**Classifier**	**Performance**
Singhal and Singh (2015) [94]	Acute Lymphoblastic Leukemia	Correlation-based Feature Selection	SVM	92.30%
Maleki et al., (2015) [95]	Protein-protein interactions (PPIs) and breast cancer	Correlation-based Feature grouping	SVM KNN	92.39%

Jiang et al., (2017) [96]	Documents	Correlation-based model bag-of-visual-words	SVM	87.50%
Zhang et al., (2017) [97]	Leaf disease	Feature selection based on correlation and GA	SVM	94.28%
Arora et al., (2017) [98]	Genetic defects	Correlation-based Feature Selection	CVR classifier	94.78 %
Nahid et al., (2018) [99]	Breast cancer	K-Means Mean-Shift	Deep Neural Network	91.00%
Liu et al., (2020) [100]	Lung Disease	Gaussian Mixture Model	RF	92.30%
Nasir et al., (2020) [101]	Document classification	Pearson correlation coefficient	Deep convolutional neural network	93.1%
Khan et al., (2021) [102]	Grape rot leaves, apple rust, grape powdery mildew, apple scab	Entropy-rank correlation-based feature selection	M-SVM	93.74%,

After the study of various research articles [205], [206] [207] and review papers [208] [124], [205],[209], [210] some research gaps have been identified and summarized below:

(i) The original datasets include high-resolution images, which require significant computational resources.

(ii) Images may have little contrast when there is a lot of similarity between the lesion and skin tone.

(iii) Another issue is the noise within the image, e.g., Hair may interfere with the accurate analysis of skin lesions that have been imaged.

(iv) Poor segmentation results are caused by images with more central lesion irregularities than border irregularities.

(v) Additionally, in some situations, a higher segmentation threshold can result in a region of regression in the image's center, which can result in the false detection of lesions.

(vi) While dealing with huge datasets or high-resolution images, feature extraction can be highly computational.

(vii) Handcrafted features that are predetermined and constructed based on prior information are used in many classic feature extraction methods.

(viii) Feature selection approaches can sometimes result in overfitting, in which the model performs well on training data but badly on unknown data.

2.2 Summary

This chapter focuses on a review of the literature on skin cancer detection devices. It provides insights and information to aid in the design of components for automated skin cancer detection systems, with the ultimate goal of enhancing and improving skin cancer detection accuracy. The strategies and methodologies chosen from the studied literature are highlighted in each section of the chapter. The proposed methodologies will be used in the proposed skin cancer detection system in Chapter 3, contributing to the system's development and improvement.

CHAPTER 3

METHODOLOGY AND SYSTEM FRAMEWORK

The automated diagnosis of skin cancer poses a challenging and crucial problem within the realm of medical image processing. The accurate differentiation between benign and malignant skin melanomas holds considerable importance for patient outcomes and the reduction of diagnostic errors. Researchers have devised a range of methods and algorithms to tackle this challenge, with significant contributions from advancements in machine learning, computer vision, and deep learning, all pivotal in enhancing diagnostic accuracy.

The study proposes a unique contribution in the form of an innovative medical expert system designed to categorize skin cancer lesions autonomously and with high accuracy. The major goal is to produce classification results that are not only accurate but also contain fewer errors than those produced by human specialists.

3.1 Proposed System

The proposed workflow initiates with the data acquisition phase, where information is gathered from the publicly available ISIC Dataset. Subsequently, a pre-processing mechanism is implemented to eliminate any potential noise within the dataset, and image enhancement techniques are applied to improve image contrast. The segmentation of the ROI is accomplished by utilizing background subtraction along with midpoint analysis. Feature extraction and selection are carried out using a differential analyzer algorithm (DAA) to identify crucial features for skin prediction. Finally, the system concludes with the classification stage, employing a correlation and linearity model. Figure 3.1 illustrates the flow of the proposed system.

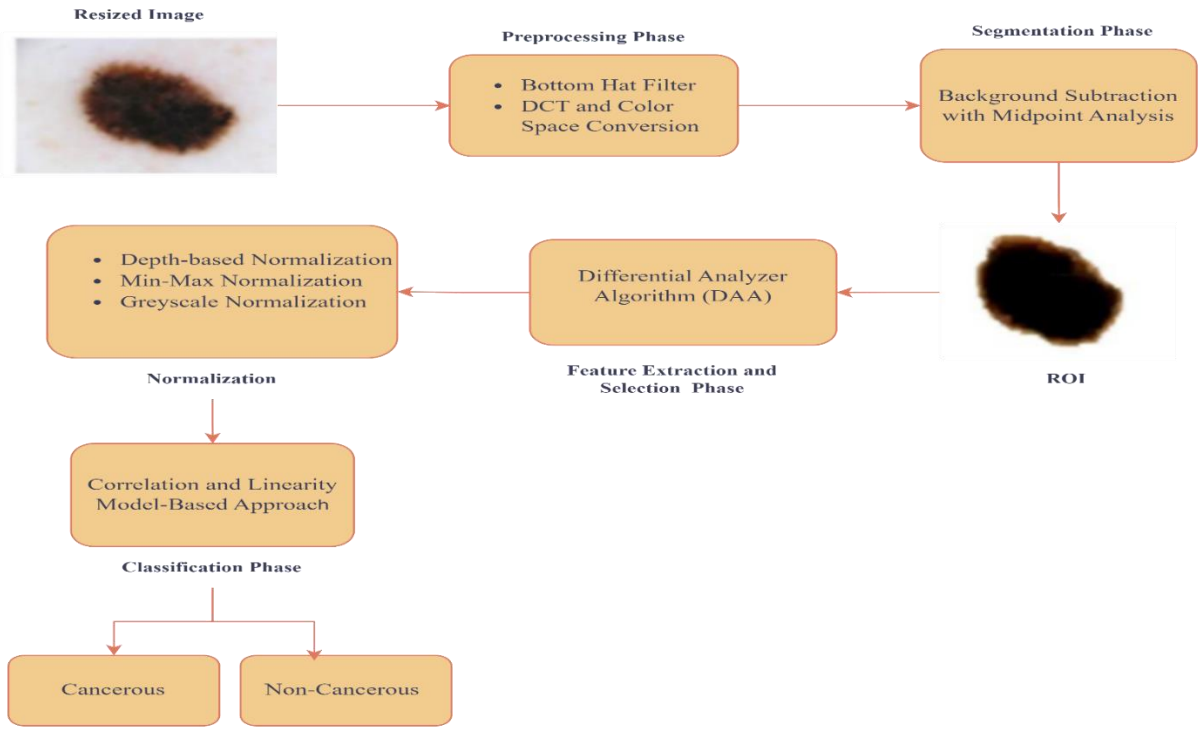


Figure 3.1 Block Diagram for Proposed Frameworks of Skin Cancer Classification and Detection

3.2 Publicly Available ISIC Dataset

The International Skin Imaging Collaboration (ISIC) dataset stands as a critical resource in the realm of dermatology, specifically for research concerning skin cancer [211]. Skin cancer, a significant global health concern, necessitates comprehensive datasets to advance diagnostic and prognostic methodologies. The ISIC dataset, renowned for its diverse and extensive collection of dermoscopic images, serves as a cornerstone in addressing this challenge and propelling advancements in skin cancer research. For the study, as shown in Figure 3.2, a dataset of SIM ISIC Datasets [212] is used which is a publicly available dataset designed to detect skin cancer.

3.2.1 Applications and Contributions

Researchers worldwide leverage the ISIC skin cancer dataset to develop and fine-tune algorithms capable of distinguishing benign from malignant lesions [213]. These algorithms aid clinicians in accurately diagnosing skin cancers, facilitating timely interventions and improving patient outcomes. Moreover, the dataset serves as a standard for assessment of different algorithms, fostering healthy competition and continuous refinement of diagnostic tools [214].

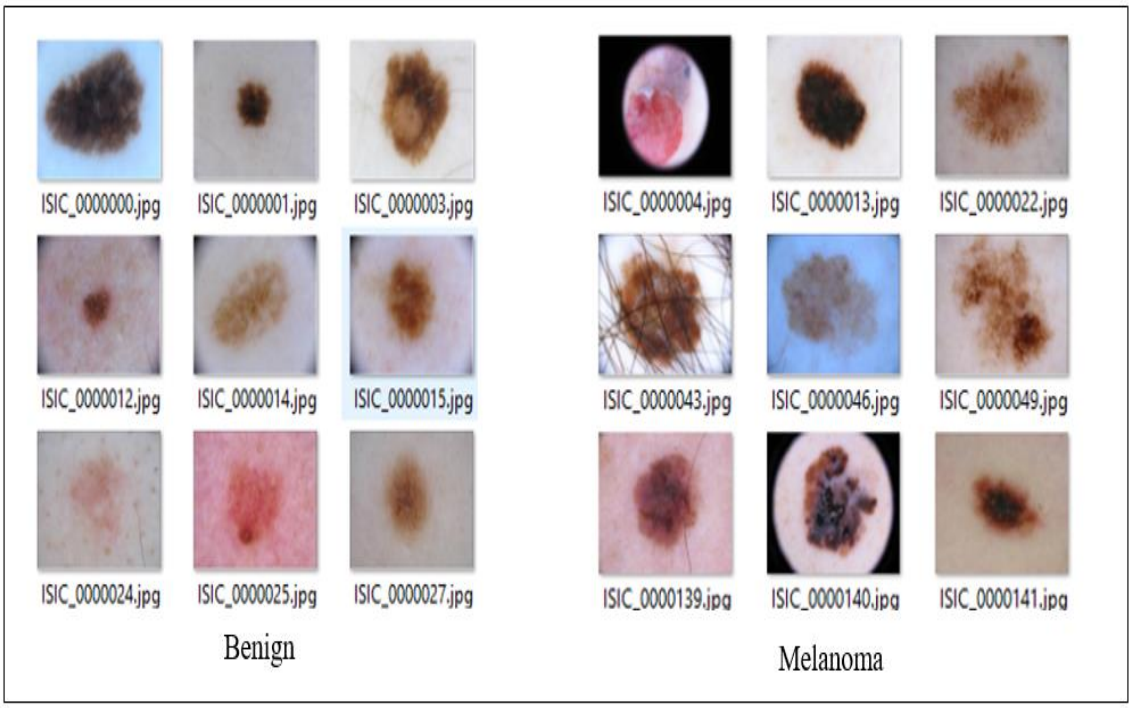


Figure 3.2 ISIC Dataset Images (Cancerous and Non- Cancerous)

3.2.2 Diverse Challenges and Opportunities

However, working with the ISIC dataset comes with its set of challenges. Variability in image quality, diversity of skin types, and the need for rigorous validation are factors that necessitate careful consideration. Despite these challenges, the dataset offers researchers a unique opportunity to harness the potential of machine learning and artificial intelligence in healthcare. By researching novel approaches, researchers can contribute to more reliable, effective, and easily accessible skin cancer diagnoses.

3.3 Pre-Processing of Dermoscopic Images

In the realm of dermoscopic image analysis, preprocessing serves as a crucial stage that significantly influences the accuracy and reliability of subsequent processing steps. The preprocessing phase involves a series of techniques aimed at cleaning, normalizing, and enhancing dermoscopic images. This phase is critical for the accurate classification of skin cancer. Several image processing techniques are applied to the input images during the preprocessing phase of skin cancer analysis before segmentation or other diagnostic activities are conducted.

The preprocessing is required to improve image quality and subsequent analysis, but it has some drawbacks. The following are some of the most common problems associated with the preprocessing step of skin cancer analysis: Data loss may occur during some preprocessing

processes [215]. When image resolution is lowered, details and small details may be blurred or destroyed, limiting the reliability of subsequent examination or diagnosis [216]. It is crucial to strike a balance between reducing noise and retaining key information during preprocessing [217]. To handle these issues during pre-processing, the image is resized, hairy parts are removed, and discrete cosine transformation (DCT) [218] and color space conversion are to enhance and restore the image during this pre-processing stage.

3.3.1 Image Resizing

Dermoscopic images often vary in resolutions and scales, necessitating the use of image resizing to standardize dimensions for compatibility with analysis algorithms. The careful execution of resizing is crucial to prevent distortion or the loss of vital details. In certain scenarios, dermoscopic images may have resolutions exceeding 1000 × 700, incurring a high computational cost. To address this, the original images are resized to [296, 296] using a MATLAB function, striking a balance between computational efficiency and preserving essential information in the images. This resizing process ensures that the images maintain compatibility with subsequent analysis algorithms while minimizing the computational resources required.

3.3.2 Hair Removal Using Bottom Hat Filter

Pre-processing is carried out using a dual strategy. First, the bottom hat filtering mechanism was applied to remove the hair from the image by reading the image and converting the RGB image into grayscale as shown in Figure 3.3. Bottom-hat filtering is an image processing technique used to enhance the visibility of small, darker structures or features in an image that are smaller than the structuring element used. It can be useful in the context of skin cancer image analysis for highlighting subtle details, such as hairs or darker areas on the skin.

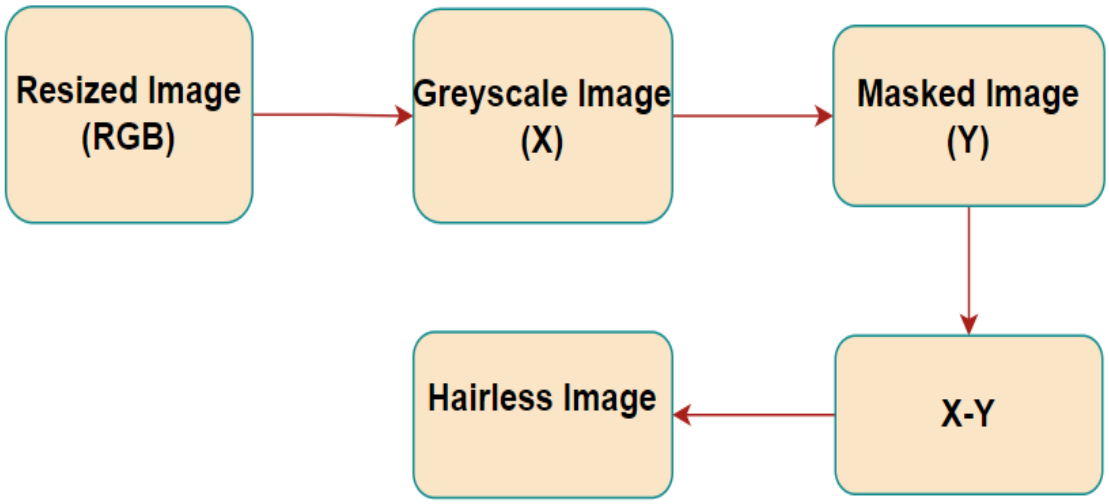


Figure 3.3 Preprocessing Phase of Proposed System

During the proposed methodology, we don't need the extensive amount of color information found in RGB images, therefor RGB images are converted into greyscale images. The outcome of a conversion is kept in [u, v]. The cross-shaped structured element's results will be used to determine how well the image measures up to the region boundaries. The maximum size of an object is specified by the T_size variable. If the maximum size (T_Size) is exceeded, these values will be replaced with the neighbourhood intensity levels denoted by Object-1(u,v). The working of the mechanism is explained in Algorithm-3.1.

Algorithm 3.1: Bottom_Hat(Image)

- [u,v]=imread(rgb2gray(Image))
- $Objects_i$=Boundaries([u,v])
- If($Object_i$>Threshold)then
  I=i+1
  Else if($Object_i$>T_Size)
  $Object_i(u,v)=Object_{i-1}(u,v)$
  End
- Repeat the above steps until all hair is removed.

3.3.3 Image Enhancement Applying DCT and Color Space Conversion

Standardizing the brightness and contrast levels of dermoscopic images is essential to ensure consistent visual characteristics. Variations in lighting conditions during image capture can lead to disparities in image quality. Preprocessing methods such as gamma correction, histogram

stretching, and dynamic range compression help normalize these attributes, ensuring that images are comparable and conducive to accurate recognition.

Image enhancement highlights the useful information present within the image and reduces the redundant information within the image as well. The proposed model used individual colour-checking and stretching mechanisms to introduce uniformity at each point within the image. To overcome the non-uniformity illuminations color space conversion strategy with DCT is applied. One of the key discussions in the realm of image enhancement centres around the utilization of the DCT method [219]. DCT is highly regarded for its speed and effectiveness in processing images. Its fundamental purpose is to convert signals from their spatial domain into a frequency domain. The overall operation of enhancement with DCT and contrast enhancement is given as under Figure 3.4.

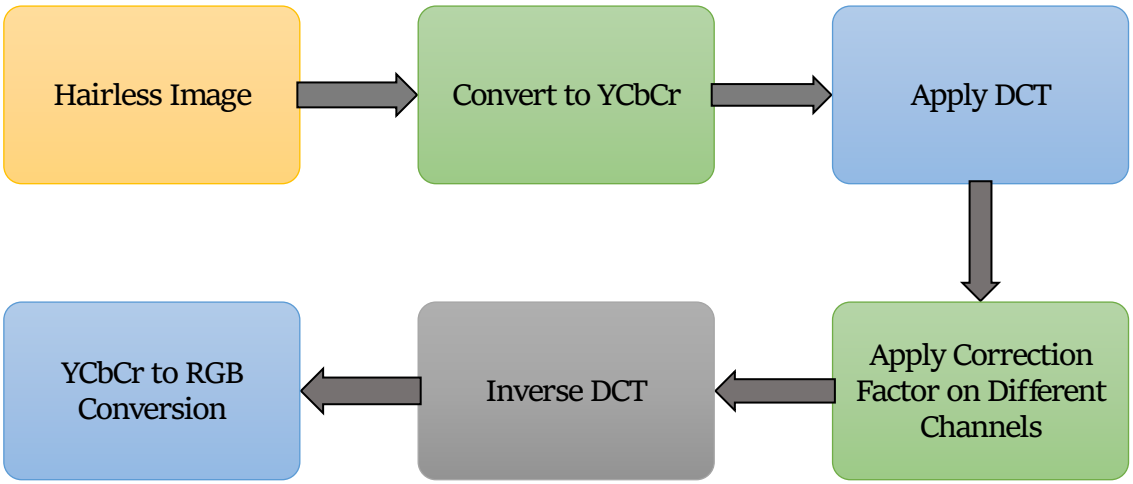


Figure 3.4 Image Enhancement during Preprocessing Phase

Chen et al., [220] introduced a modest modification to the Discrete Cosine Transform (DCT) algorithm, which led to substantial performance improvements, particularly when dealing with significant variations in illumination. Their insight was that by discarding low-frequency DCT coefficients within the logarithmic domain, the method demonstrated a remarkable ability to mitigate the adverse effects of illumination variations. One of the notable advantages of this approach is that it does not require modelling or the use of bootstrap sets. Moreover, it is known for its speed and ease of implementation, making it particularly suitable for recognition systems.

The Discrete Cosine Transform can be employed for image enhancement in cancer images, similar to how it is used in various image processing tasks. DCT-based image enhancement can help improve the quality and visibility of relevant features within cancer images. This section explains the methodology for image enhancement using DCT and color space conversion.

After replacing the hairy part from the image with neighborhood pixels, the image is required to be enhanced. Convert the hair removal RGB cancer image into HSV (Hue, Saturation, Value) that separates the information needed to be enhanced. Apply the 2D DCT transformation to the chosen colour channel (e.g., the Value channel in HSV) of the cancer image. Apply a 2D DCT transformation to the LL (low-low) frequency bands of the image. The LL frequency bands represent the low-frequency components and are typically found in the top-left corner of the DCT coefficient matrix. Most of the information within the signal is concentrated towards lower frequency components and hence LL frequency bands will be transformed through 2D DCT. The correction coefficient for this purpose is required to be evaluated. This is given through equation 3.1.

$$CF = \max\left(\frac{\sum LL\underset{i}{\rightarrow}}{\sum LL_i}\right) \tag{3.1}$$

Here "CF" is the correction factor. $LL\underset{i}{\rightarrow}$ is the similarity-valued matrix of the adaptive histogram equivalent image and LL_i is the singular valued matrix of the input image.

After obtaining the DCT coefficients of the LL bands, Correction is done with these coefficients to control the contrast enhancement. The correction factor determines the amount of enhancement. It should be less than 1 to avoid over-amplification. The correction coefficient can be calculated based on the correction factor you want to apply. The coefficient represents the reciprocal of the correction factor and is used to restore the image. The correction factor (CF) will be multiplied by the colour channel. Apply the inverse 2D DCT transformation to the modified DCT coefficients. This step restores the enhanced image back to the spatial domain. Clip and scale pixel values as necessary to ensure they fall within a suitable intensity range (e.g., [0, 255] for 8-bit images) to create the final enhanced image. Apply contrast stretching to further refine the quality of the enhanced image.

- Step 1: Convert RGB image to HSV
- Step 2: Apply DCT
- Step 3: Apply correction factor
- Step 4: H=CF*H
  - S=CF*S
  - V=CF*V
- Step 5: Apply inverse DCT
- Step 6: Convert HSV to RGB

It is observed that the scale factor of "CF" gives better clarity and hence we use CF as the scale factor in contrast enhancement.

3.4 Segmentation of Dermoscopic Images

Segmentation in image processing involves isolating the critical elements of an image while removing unnecessary portions. In the domain of skin cancer images, segmentation is applied to identify the boundary between a lesion and the surrounding healthy skin [221]. Its primary goal is to distinguish sets of related pixels within a ROI so that structural transitions between these sets can be more easily detected. Segmentation accuracy is critical to attaining lower error rates in the later assessment of skin lesion shape, border, and size parameters [222]. One of the most crucial and complex operations in image processing is segmentation. It has to be simultaneously fast and precise. The effectiveness of subsequent procedures, such as feature extraction and classification, is heavily reliant on the efficiency of the segmentation phase.

As previously stated, segmenting dermoscopic pictures can be difficult due to factors such as poor contrast between the lesion and normal skin tone, colour fluctuations within the pigmented region, and the presence of artefacts. Despite these difficulties, dermatologists find segmentation extremely useful since it provides critical information on asymmetry, border irregularity, colour, and diameter—all of which are important determinants in melanoma diagnosis. As a result, using a proper segmentation approach to outline the complete lesion area from photographs considerably improves the diagnosis procedure. In medical image segmentation, various segmentation approaches have been used, including active contours-based methods, thresholding-based, artificial intelligence-based, region-based and edge-based. [208], [223]-[224].

- **Threshold-based techniques:** These approaches depend on the selection of one or more histogram threshold values to distinguish objects from the background [225].
- **Edge-based:** These approaches locate the boundaries separating the regions using edge operators [226], [227].
- **Region-based**: By merging, splitting, or both, pixels are organized into homogeneous regions [228].
- **AI-based:** These approaches depict images as random fields, with parameters determined through various optimization techniques [229].
- **Active contours**: These techniques identify object contours by using curve evolution [179], [230].

Though segmentation approaches are important in the research and detection of skin cancer, they do have certain limitations. Skin cancer segmentation procedures involve manual input or monitoring. This procedure can be tedious and costly because it frequently includes physically outlining the lesion's limits [231]. Skin lesions can have a wide range of appearances, including

shape, size, texture, and colour. Some segmentation methods may struggle to represent the complex and diverse features of many types of skin lesions [232]. Segmentation methods may suffer from over- or under-segmentation. Both scenarios can result in incorrect and inadequate segmentation data, which can affect future analysis and diagnosis [233]. To consider these kinds of issues in this study a novel method has been proposed for the segmentation of medical images. The proposed midpoint segmentation method using a background subtraction approach, enhances the performance of different classifiers. It helps segment the image into regions of interest, allowing for a more focused analysis of specific objects. By extracting only ROI, the background subtraction reduces the computational load associated with processing the entire image or video stream.

3.4.1 Proposed Segmentation Method

Background subtraction is a popular method for identifying video frames in a sequence taken by stationary cameras. The basic idea is to detect objects by determining the disparity between the current frame and a reference frame, which is sometimes referred to as the 'Background Image' or 'Background Model.' This is generally accomplished by detecting the ROI in a picture, with ROI detection serving as the primary goal of this approach. In our proposed methodology we have implemented a background subtraction algorithm with midpoint analysis [234] . Figure 3.5 depicts a schematic overview of the proposed strategy.

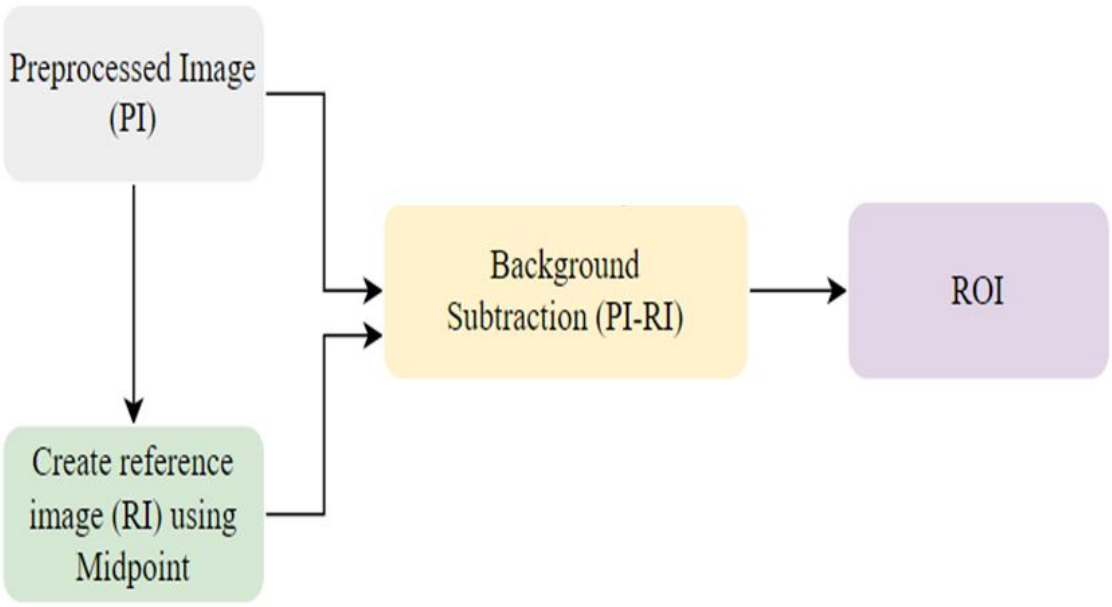


Figure 3.5 Working of Background Subtraction Model

In our proposed methodology, we have created a reference image by applying midpoint analysis. This reference image will be used to extract the region of interest from the input image. Region extraction is very significant since it will be used to examine irregularities in separate regions. Typically, the region extraction step is done with an inside-outside test that categorises the boundaries of the image. These concerns have been addressed in [56]. In our proposed work, the process is reserved. The boundary separation of the image in region-based separation mechanisms is a problem. Therefore, extraction of the cancerous region may not be appropriate. The advantages of reorientation include the shortest time to identify regions of interest in subsequent segmentation processes.

The process, which starts from the inside and moves outward, cannot be completed until it encounters the boundary, but using the proposed method, the control will move from the boundary of this region and move inward to the centre, creating a threshold point at which changes in intensity levels occur. If no abnormalities are detected when scanning 50% of the area, the entire area will be marked as non-critical. The overall workings of the proposed method are given below in Algorithm 3.2.

In the beginning, the proposed method uses the levels of color intensity (C_i) to locate the edge (E). Color is extremely important in identifying the region of the image (R). The equation 3.2 is used to identify the color.

$$E_{x_e,y_e} = \begin{cases} C_i! = C_i + 1\ R\_E \\ 0\ othrwise \end{cases} \tag{3.2}$$

The region edges (RE), which are identified by a color comparison with the adjacent pixel, are stored in the $E_{x,y}$ variable, where ‘E’ stands for the region's edges. As a result, multiple edges corresponding to different regions will be detected, and the edge variable requires the subscripted variable "e."

To determine the region's dimension, the diameter of the region is estimated. The equation 3.3 is used to determine the region's centre (CN_i), which is obtained by dividing the region's dimension by "2" to obtain the radius. The centre of the region will be determined by the region's edges. The normal flow of the system will now come from the centre.

$$CN_i = \begin{cases} (max(E_{x_i,y_i}) - min(E_{x_i,y_i}))\ /2\ > 0 \\ \leq 0\ RE\ extraction\ not\ possible \end{cases} \tag{3.3}$$

Algorithm 3.2

Image_Region_Cut_BS(image, Threshold) // image is the training dataset

* Threshold: Values corresponding to the boundaries of the image

* Color is the value of the colour for the inner region filling

- [x,y]=imread(image)

// Module to identify the boundary of the image

- For i=1:size(image)

    If (intensity(x,y)==threshold)

    B (i) = intensity(x,y);

    Cord(x,y)=[x,y]

    End of if

  End of for

// Region Cut and Midpoint analysis

- [X,Y]=B(i)
- For i=1:size(B)

    For j=i: size (B)

    $[Mid_x, Mid_y]=(X_i/2,Y_i/2)$

    For k=1:X_i

    For k1:Y_i

    putpixel(k,k1,color)

    end of for

    end of For

    end of for

  end of for

The region extraction phase is crucial in the proposed segmentation process. The process begins with the identification of labelled segments. Regions corresponding to critical points have edge values.

A set of intensity values (Int) will be stored in buffers (Bfi) corresponding to these edges. Intensity values corresponding to these regions will be used to remove the background and extract the region of interest. The centre intensity values greater than 0 are stored in the buffer for further processing, and values less than or equal to 0 are discarded. The primary equation used for the region extraction is given in equation 3.4.

$$Bf_i = \begin{cases} Int(CN > 0), Valid\ centre\ values\ will\ be\ stored\ within\ buffer \\ Discard\ Int(CN \leq 0)\ invalid\ intensities\ will\ be\ rejected \end{cases} \quad (3.4)$$

The background colour with the skin image has a maximum of 255 intensity levels. The region obtained from the region separation and the intensities from the region separation are used in the background subtraction process. This is expressed in equation 3.5.

$$Without_{Bg_i} = \lfloor 255 - Int(Bf_i) \tag{3.5}$$

This intensity will be subtracted from the rest of the image to extract a background-free image. After background subtraction, the region of interest is obtained. The most positive matching segments will be kept, and the rest will be rejected. To acquire relationships between components, bwconncomp is used. Regionprops is applied to get the coordinates of the region with the highest correlation.

3.4.2 Advantages and Challenges

Background subtraction is advantageous for scenarios where the foreground and background have a clear contrast, and it can work well in real-time applications like surveillance and object tracking. However, its accuracy may decrease when lighting conditions change or when there are gradual changes in the background.

Midpoint and Region cut, on the other hand, provides more refined results by considering multiple attributes for segmentation. It can handle more complex scenarios where objects have varying colors, textures, and shapes. However, region cut methods can be computationally intensive and may require parameter tuning.

In the realm of image segmentation, background subtraction and midpoint analysis are essential techniques that play a pivotal role in extracting relevant information from images. Background subtraction helps identify the foreground objects by highlighting them against the background, while midpoint further refines the segmentation process by grouping pixels with similar characteristics. Both methods contribute to various applications across industries, enabling more accurate analysis and understanding of visual data.

3.5 Feature Extraction and Selection for Dermoscopic Images

Feature extraction is employed specifically on the extracted region, enhancing processing speed by focusing on relevant portions rather than the entire image. The chosen method for feature extraction is the bag of features approach, which proves advantageous in this context [235]. This algorithm considers both local and global solutions. Local solutions are ideal within a certain part of the search space, whereas global solutions are optimal globally. If the local solution outperforms the

global solution in a given case, the global solution will be replaced by the superior local solution. Color, shape, and texture features were discovered to be the most relevant for the detection of chronic melanoma images [236]. These features are mentioned in Figure 3.6.

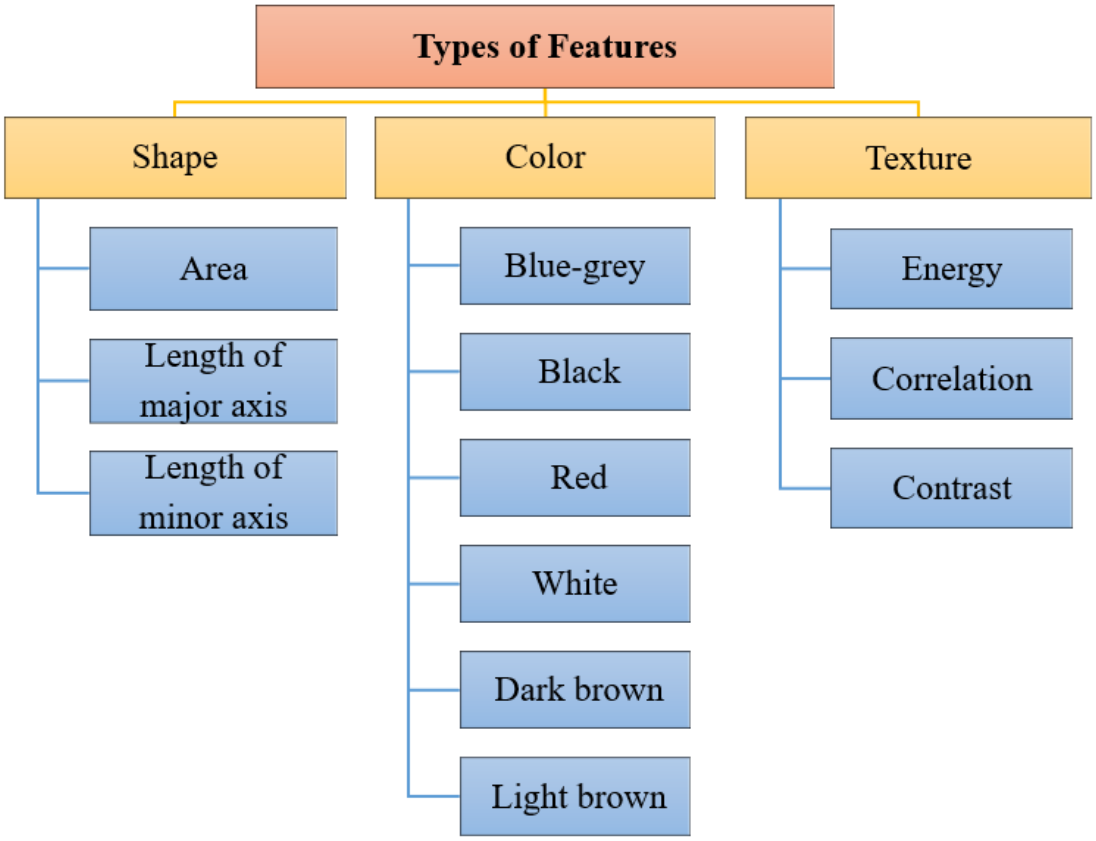


Figure 3.6 Different Types of Features Extracted

These features provide valuable information that can aid in differentiating between benign and malignant skin lesions. Contribution of each feature to the detection process is given in this section:

Shape: The shape of a skin lesion can provide important clues about its nature. Malignant lesions often have irregular, asymmetrical shapes, with uneven borders [237]. In contrast, benign lesions tend to have more regular, symmetrical shapes [238]. Analyzing the shape of a lesion can help dermatologists identify potential signs of malignancy.

Color: Color variation is a key characteristic of many skin lesions, and analyzing color patterns can aid in diagnosis [239]. Malignant lesions may exhibit uneven pigmentation, with areas of dark brown, black, red, or even blue. Benign lesions, on the other hand, often have more uniform coloring, such as light brown, pink, or tan [240]. Evaluating color distribution and changes over time can provide important diagnostic information.

Texture: The texture of a skin lesion refers to its surface characteristics, such as smoothness, roughness, or scaliness [241]. Malignant lesions may exhibit irregular texture patterns, including ulceration, crusting, or scaling [242]. Benign lesions, in contrast, tend to have smoother, more uniform textures [243]. Assessing texture features can help distinguish between benign and malignant lesions and guide further diagnostic evaluation.

In computer-aided skin cancer detection systems, shape, color, and texture features are often quantified and analyzed using image processing techniques and machine learning algorithms [244]. By extracting and analyzing these features from digital images of skin lesions, these systems can assist healthcare providers in identifying suspicious lesions and determining the likelihood of malignancy.

While shape, color, and texture features are important components of skin cancer detection, they are typically considered alongside other clinical features, such as size, symmetry, and evolution of the lesion, to form a comprehensive assessment [245]. Integrating multiple features allows for more accurate and reliable diagnosis of skin cancer lesions, ultimately improving patient outcomes.

3.5.1 Proposed Method for Feature Extraction and Selection

The Differential Analyzer Algorithm (DAA) algorithm is primarily used for drawing lines on a computer screen. It calculates the coordinates of points along a straight line between two given points (usually the endpoints of the line) and then plots those points to create the line. The algorithm is based on the idea of using incremental changes along the x-axis to determine corresponding changes along the y-axis.

3.5.2 Steps Involved in DAA Algorithm

In this iterative and differential feature selection strategy, each step represents a deliberate modification of the feature set, assessing the impact on the model's performance. The assumption of linearity in each feature aligns with the idea of fitting feature values over linear dimensions, highlighting the importance of how each feature contributes to the overall model.

- Calculate the slope of the line.
- Determine the number of steps required.
- Calculate the incremental changes in x and y.
- Initialize a starting point.
- Perform a loop for i from 1 to steps:
- Plot the point

- Increment x by dx and calculate the new value of y.

The conventional DAA algorithm is commonly employed in graphics, but a comparable incremental approach holds potential for adaptation in the realm of machine learning and data analysis. In this context, Feature Extraction and Selection play a crucial role, emphasizing the selection of a subset of pertinent features from the initial feature set. The objective is to enhance model performance and alleviate computational complexity. The adaptation of the DAA approach for feature selection involves iteratively moving from one feature set to another while assessing the impact on the model's performance. The incremental adjustments mimic the incremental changes along the x-axis in the traditional DAA algorithm. Each feature is supposed to possess a certain linearity. A differential approach used will try to fit the feature values over linear dimensions. Features not fitting the dimension will be rejected. The steps entailed in the feature selection process are outlined in Algorithm 3.3.

Algorithm 3.3

```
Feature Selection (Image)
I=imread(image)
        For i=1:rows(image)
                For(j=1:cols(image)
                        Fit extracted value on straight line
                        y=mx+b
                                End of if
                        End of for
                End of for
```

To choose the relevant features, the differential analyzer algorithm (DAA) is used [246]. DAA is a meta-heuristic technique that works in the same way as genetic algorithms [247]. In this approach, each unit within the population is designated as either male or female, and each child is regarded as a potential solution to the given problem. The male population is represented by "X," while the female population is denoted as "Y." The determination of the required number of iterations for convergence is based on the computation of Max and Min values derived from the extracted parameters, as depicted in equations (3.6) and (3.7).

$$dx = x_{max} - x_{min} = \begin{cases} \max(X_i), where\ X_i > 0 \\ \min(X_i), where\ X_i > 0 \end{cases} \quad (3.6)$$

$$dy = y_{max} - y_{min} = \begin{cases} \max(Y_i), where\ Y_i > 0 \\ \min(Y_i), where\ Y_i > 0 \end{cases} \quad (3.7)$$

Where "I" is the size of the population, dx denotes the highest statistical variance between the maximum and minimum population of males, and dy represents the highest statistical variance between the maximum and minimum population of females. The total number of iterations allowed is calculated using equation (3.8).

$$s = \begin{cases} dx, if\ dx > dy \\ dy, otherwise \end{cases} \quad (3.8)$$

The 'S' variable indicates the steps, the total number of repetitions after which the algorithm automatically terminates. To derive offspring indicative of potential solutions, the crossover factor is computed from both the male and female populations. The equations (3.9) and (3.10) are used to compute the crossover factors represented by X_C and Y_C.

$$X_c = \frac{dx}{steps} \quad (3.9)$$

$$Y_c = \frac{dy}{steps} \quad (3.10)$$

The population's selection for offspring is determined by X_C and Y_C. The equations (3.11) and (3.12) are used to produce the next generation of offspring.

$$X_n = X_n + X_c \quad (3.11)$$

$$Y_n = Y_n + Y_c \quad (3.12)$$

The subsequent male and female individuals chosen to produce offspring are denoted as Xn and Yn. Their fitness is evaluated using a linear fitness equation, as illustrated by equation (3.13), where "m" represents the slope, and "c" represents the intercept.

$$f(x, y) = m * (X_n, Y_n) + c \quad (3.13)$$

The error function serves as the objective function to determine whether to retain the current offspring. The error rate, calculated using equation (3.14), must be less than the specified tolerance for the offspring to be accepted.

$$error_{x,y} = \{\sum f_{measured}(X_n, Y_n) - f_{model}(X_n, Y_n) \quad (3.14)$$

If the generated offspring satisfies the optimization function, it is retained. Equation (3.15) defines the error rate, which must be less than the specified tolerance for the offspring to be preserved:

$$\varepsilon = \begin{cases} error_{x,y}, if\,(dx < pt\ and\ dy < pt \\ NA, Otherwise \end{cases} \quad (3.15)$$

Here, 'ε' represents the specified tolerance, which must be less than 0.001. If this condition is not met, optimality is not achieved, and the process is reiterated.

The feature set extracted from the image equivalent to the specific region determines the optimal values that can be plotted linearly. For example, the range of values extracted from the region image has local maxima represented as dx = 10 and dy = 20. The minimum set of values obtained is given as (10,20) and the maximum value is given as (20,40). We calculate steps = dy = 20. The value of x_{incr} will be 0.5, and the value of y_{incr} will be 1. The range of values obtained through the DAA approach is listed in Table 3.1.

Table 3.1 The Range of Values Obtained Through DAA Approach

X	Y	(X, Y)
10	20	(10,20)
10.5	21	(10,21)
11	22	(11,22)
.	.	.
.	.	.
.	.	.
20	40	(20,40)

These values of X and Y are compared against the extracted values. Out of the range of values, values closely matched with the line coordinates are retained, and the rest of the values not satisfying the line coordinates are rejected. Optimal feature extraction will be achieved after the end of this algorithm.

This algorithm undertakes the extraction of shape, size, and texture features from region images. Subsequent to the extraction process, each feature undergoes various normalization mechanisms to mitigate differences among feature values. The optimization function is employed for the evaluation process, with a predefined tolerance set. The obtained results are then compared to this specified tolerance to identify the most favourable offspring. Each of the fittest offspring is associated with a specific feature. The crossover factor is modified to the weight factor acquired for optimizing the Differential Analyzer Algorithm (DAA). Notably, in comparison to a standard genetic algorithm, this approach exhibits a remarkable convergence rate. The number of iterations is variable, ranging from 10, 20, and 30, up to 100, thereby enhancing the DAA's performance towards achieving the best possible solution. The constraint settings for the DAA are outlined in Table 3.2.

Table 3.2 Constraint Settings for DAA

Constraints	Values
Feature reduction ratio	55
Repetitions	100
Replication	5
Appropriate Tolerance	0.001

Every optimised feature in this process is assigned a gender during optimisation. Which male and female traits will be used to produce offspring depends heavily on the crossover factor. The best feature is represented by the progeny, which is kept in the final feature matrix. Figure 3.7 shows how the Differential Analyzer Algorithm (DAA) works.

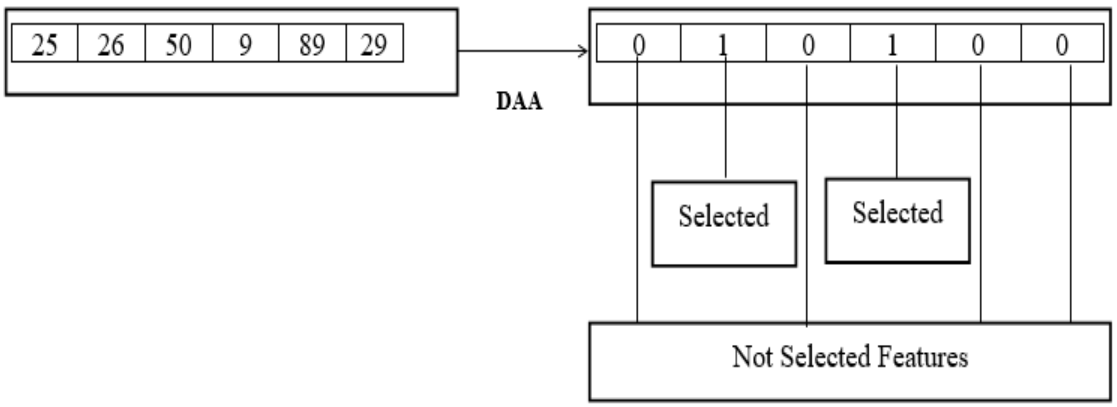


Figure 3.7 Mechanism of DAA Algorithm

In this context, the pivotal role of feature extraction and selection cannot be overstated, as it underscores the significance of choosing a subset of pertinent features from the initial feature set. The overarching goal is to augment model performance while concurrently mitigating computational complexity. Feature selection, as a core component of this process, entails the meticulous curation of a subset of relevant features from the original set. This strategic approach is designed to not only enhance the efficacy of the model but also to alleviate the computational burden associated with an exhaustive feature set. The adaptation of the DAA approach for feature selection involves iteratively moving from one feature set to another while assessing the impact on the model's performance. The incremental adjustments mimic the incremental changes along the x-axis in the traditional DAA algorithm.

3.5.3 Benefits and Considerations of DAA Algorithm

- **Efficiency**: Similar to the DAA algorithm's efficiency in approximating lines, this approach can help efficiently navigate through different feature combinations without exhaustive search.
- **Flexibility**: This approach allows for gradual changes in the feature set, which can be beneficial when transitioning between various subsets of features.
- **Complexity**: Feature selection is a complex problem, and this approach simplifies the process by breaking it down into incremental steps. However, the algorithm's design and stopping criteria need careful consideration.

In conclusion, while the Differential Analyzer Algorithm is traditionally used for drawing lines, its incremental nature can be creatively adapted to solve other problems like feature extraction and selection. This showcases the versatility of algorithms and their potential for inspiring innovative solutions in various domains of computer science and data analysis.

3.5.4 Normalization

This phase is crucial for ensuring that the features used in a machine learning model are on a similar scale, preventing certain features from disproportionately influencing the model due to differences in their magnitude [248]. The overarching goal of these normalization strategies is to close the gap between the highest and lowest extracted feature values. By ensuring a consistent and standardized scale across features, the normalization phase contributes to the model's ability to generalize well and improve classification accuracy. This is particularly important in machine learning, where the effectiveness of models can be influenced by the relative scales of input features. The three normalization mechanisms mentioned—depth-based, min-max normalization (MMN), and grey-level scaling—are applied to generate a feature matrix linearly.

- Depth-based Normalization: To compare two different feature sets, this step estimates the depth of sequencing.
- Min-Max Normalization (MMN): This technique linearly adjusts the raw data to predetermined lower and upper bounds. Normally, the data is rescaled between 0 and 1 or -1 and 1[248].
- Greyscale normalization: This likely involves scaling pixel values to a common range, making it easier for the machine learning algorithm to work with the data. It can be performed either by normalizing each column or row to a range of values from 0 to 1.

3.6 Summary

This chapter provides full details of proposed segmentation and feature extraction methods used in automated skin cancer detection systems. The proposed method for segmentation is applied especially for obtaining ROI. The proposed feature extraction and selection algorithm is applied for various features to classify benign from malignant lesions. The feature sets extracted in this chapter part will be given to the different classifiers to trace their role in differentiating benign from melanoma. This will help the appropriate features attain their weight according to the performance related to the classification of melanoma as discussed in next chapter 4.

CHAPTER 4

CLASSIFICATION

The act of classifying skin lesion images into several classes or categories according to the type of skin cancer or the probability of malignancy is known as classification in skin cancer imaging. The classification of a lesion as malignant or non-cancerous is a crucial and final stage in computer-aided diagnostics. Pattern recognition, visual analysis, and artificial intelligence all rely heavily on classification. Training or learning and testing are the two processes in supervised classification. During the training stage, the classifier is taught or trained using a dataset with features as input. When a machine or algorithm is trained, it is put through testing with reserved features, which implies that training is done using distinct images or feature vectors. Classifier performance is thus assessed in this manner. Before explaining the proposed lesion classification method, a brief description of numerous classification algorithms is presented.

4.1 Models for Skin Cancer Classification and Detection

The use of machine learning and artificial intelligence in healthcare is critical for skin cancer categorization and diagnosis. The classifying models used for this purpose are chosen based on the unique qualities of the data and the application's specific requirements. This section investigates and discusses some commonly used models for skin cancer categorization and detection.

4.1.1 Support Vector Machine

A Support Vector Machine (SVM) comes under the group of supervised learning approach and is accompanied by specific learning algorithms. Serving as a non-probabilistic binary linear classifier, it finds application in data analysis for both classification and regression tasks [249]. Created using an SVM training algorithm, this model is crafted to classify new cases into one of two predetermined categories. As a member of the supervised learning models, it employs associated learning algorithms. Fundamentally, the SVM-generated model, trained via a dedicated algorithm, assigns new instances to one of the two predefined categories, drawing from a set of training examples [250]. SVMs are strong algorithms in the field of supervised machine learning, notably for grouping datasets into meaningful groupings. These models rely on decision plane and decision

boundary principles. The basic purpose of SVMs is to find the best hyperplane that maximises the distance between the closest points of different data clusters. Instances located near this ideal hyperplane are referred to as 'Support Vectors.'

4.1.2 K- Nearest Neighbor

The K-nearest neighbour (KNN) algorithm is a slow learning method and one of the foundational techniques in machine learning. Its adaptability allows it to be used in both classification and regression problems. Notably, KNN is sensitive to the peculiarities of the input data. This method has various advantages, including high accuracy in specific cases, resistance to outlier influence, and the absence of underlying assumptions about the nature of the data.

This extensively used classification strategy includes classifying unknown occurrences based on their similarity or distance to records in the training set [251]. The approach computes the similarity or distance of each test record to each record in the training set. These records are then ordered depending on the proximity function (similarity or distance) computed [252]. The top-k records are then chosen as the test record's k-nearest neighbours. When a record's k-nearest neighbours (KNN) include instances from several classes, the decision is made by a majority vote [253].

4.1.3 Random Forest

As the name of this algorithm illustrates, a forest is a collection of trees. Random Forest is a methodology that builds an ensemble of decision trees that provide the output class for the input variable collectively. Random Forest is used in our work to address the overfitting difficulties that are often associated with training sets in decision tree models. To generate many tree learners, this technique leverages bootstrap aggregation. To solve overfitting problems, feature bagging is accomplished by randomly picking a subset of features.

This method falls within the domain of ensemble classifiers [254]. The process involves creating multiple random subsample spaces from the training set, each acting as a sub-training set. For each of these subsets, a decision tree is constructed, resulting in a collection of decision trees termed random forests. Each decision tree assigns a class label to a test sample, and the ultimate decision is determined through a majority vote among the ensemble of decision trees [192].

4.1.4 Decision Tree Classifier

Decision trees (DT) have a structure analogous to flowcharts, with each node denoting an attribute test and each branch denoting the test's result. The leaf node denotes the final class label, which is

the result of assessing all attribute computations. Applications for this predictive modelling technique can be found in data mining, statistics, and machine learning. It predicts outcome values based on many input elements. Decision trees are frequently represented visually to improve understanding and interpretation [161].

A training dataset is used to generate DT, with intermediate nodes dividing the dataset into subgroups and leaf nodes signifying class labels. The training set is divided in this manner until the decision tree covers every record in the training set [255].

4.1.5 Artificial Neural Networks

Artificial Neural Networks (ANNs), specifically advanced deep learning architectures like Convolutional Neural Networks (CNNs), have sparked a transformative breakthrough in the realm of skin cancer detection [256], [257]. CNNs exhibit the capacity to autonomously acquire hierarchical features from dermoscopic images, effectively capturing nuanced patterns essential for precise classification [258]. Transfer learning, a method that entails refining pre-existing CNN models on skin cancer datasets, has demonstrated remarkable efficacy in achieving high levels of performance within this field [133].

4.1.6 Convolutional Neural Networks

Convolutional Neural Networks (CNNs)—have demonstrated exceptional performance [259]. These CNNs eliminate the need for manually created features by automatically extracting hierarchical features from dermoscopic images [260]. Their success in this field is largely attributed to their ability to instinctively recognize complex patterns, textures, and spatial relationships [261]. The limitations posed by the scarcity of labelled data for training have led to a broad acceptance of transfer learning, a technique that involves fine-tuning pre-trained CNN models using datasets related to skin cancer [262].

In the domain of skin cancer, machine learning has significant advantages over deep learning [263]. They can give precise and quick skin cancer diagnosis, which is critical for effective therapy. Furthermore, machine learning algorithms are capable of optimising interaction rounds, resulting in improved rates of accuracy with smaller interactions [240]. Deep learning models, on the other hand, such as CNNs, are created specifically for the analysis of images and can classify skin lesions with more accuracy [264]

4.1.7 Customized Classifier

The customized classifier is tailored to the specific dataset and task requirements. It can be a combination of various algorithms, data preprocessing techniques, and feature selection methods. Its adaptability allows it to exploit the dataset's unique characteristics and yield optimized results. This classifier's strength lies in its flexibility to be finely tuned, but it requires domain expertise to create and optimize.

There are numerous other classification algorithms and techniques, each with its strengths, weaknesses, and suitability for different types of data and tasks [265]. The ones listed are among the most commonly used and well-known in the field of machine learning [266]. However, there are several reasons why other methods may not have been included in the list:

Scope and Focus: The list provided focuses on a representative selection of supervised, unsupervised, and hybrid classification methods [267]. Including every possible classification algorithm would make the list exhaustive and less focused [268].

Relevance and Popularity: The methods listed are widely used and studied in both academia and industry. They represent some of the most established and effective approaches for classification tasks across various domains [269].

Space Limitations: Providing an exhaustive list of classification methods and their descriptions would require a significant amount of space and may overwhelm the reader [270]. The goal is to provide a concise overview of the main approaches while still covering a broad range of techniques.

Evolution of Techniques: The field of machine learning is continuously evolving, with new algorithms and techniques being developed and published regularly. While the listed methods are well-established, newer approaches may not have been included due to their relatively recent emergence or limited adoption at the time of listing [271].

Application Specificity: Some classification methods are tailored to specific domains or types of data and may not be as widely applicable or general-purpose as the ones listed. Including highly specialized methods may not be relevant for a general overview [272].

In practice, the choice of classification method depends on factors such as the nature of the data, the complexity of the problem, the availability of labelled data, computational resources, and the specific goals of the analysis [273]. Researchers and practitioners often experiment with various algorithms to find the most suitable one for their particular task and dataset.

4.2 Proposed Classification Model

The proposed model skin cancer customized classifier (SCCC) represents the binary classification problem. In detection, there were only different possibilities: affected or not affected. The classification is executed through a customized-based approach, integrating tailored features, correlation analysis, and optimization to enhance the performance of classification. This approach incorporates global and local phases, fitness evaluation, and probability-based selection, adding resilience to the classifier and rendering it suitable for real-world applications. Notably, the Linearity Model Classifier exhibited the lowest degree of misclassification, indicating the highest classification accuracy. This observation aligns with the performance of the trained model. Moreover, the customized trained model demonstrated proficiency in testing scenarios with SVM, random forest, decision tree, and naïve Bayes approaches. The proposed method leverages a correlation and linearity model-based approach, as illustrated in Figure 4.1.

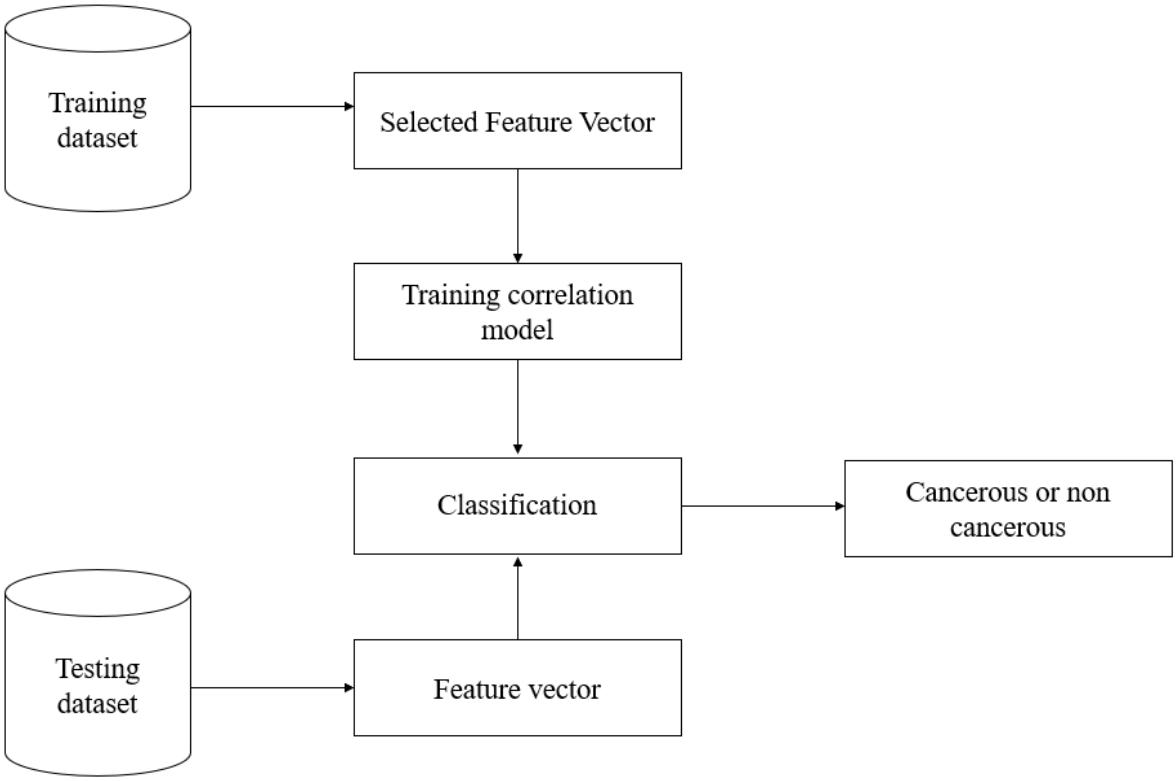


Figure 4.1 Proposed Correlation and Linearity Model

The classifier performs operations on an optimization basis. The initial phase is based on correlation. The features having the highest correlations with the target or output variables will be retained and the rest of the feature vectors will be iterated. The obtained solution is termed a local solution. For the next iteration, the local solution is converted to a global solution by comparing it against the objective function. There exist two solutions corresponding to feature selection. The first

solution is termed a local solution and the other solution is termed a global solution. The values of local and global solution are represented with 'g'. The switching probabilities is used to select the next feature vector for correlation analysis. The process continues until features having highest correlation with the target value is obtained. After obtaining highest correlation feature vector, classification result is printed.

The novelty in this approach is introduced by first of all modelling the assignments of local descriptors. These are termed contributor functions. After this multiple assignment strategy is used for reconstructing local features using neighbouring visual words from the vocabulary. Reconstruction weights are calculated from quadratic programming. These weights will be used to form contributor functions. Features possessing the highest weight will be extracted and retained. The process generates the local solution during the local leader phase. This phase is in charge of creating local solutions. It is often a search or optimizing method that investigates a confined segment of the solution space to identify an optimal solution inside that segment. The current search space is changed in this phase based on the results of the local leader phase. Its goal is to create a reference position that will be used by a differential technique to determine the location of a feature within an image.

The reference location (RF_l) within the sample space is updated by taking into account the contemporary energy location, the leadership position, and a randomly picked item from the group. The perturbation rate (per_{rate}) limits the dimension of the solution space [274]. If the perturbation rate (Per_{rate}) is greater than or equal to a random number x(0, 1), the decision at position (a, b) is determined by equation (4.1). The decision is updated if the random number x(0, 1) satisfies the requirements; otherwise, it is set to 1.

$$RF_{(l)} = \begin{cases} RF_{l(a1,b)} + x(0,1) * \left(N_{i(a1)} - Local_{sol(a1,b)}\right) + x(-1,1) * \left(N_{i(a1)} - Global_{sol(a1,b)}\right) where\ x(0,1) \geq Per_{rate} \\ RF_{l(a1,b)}\ where\ x(0,1) < Per_{rate} \end{cases} \quad (4.1)$$

In this scenario, x(0,1) represents a random sample drawn from the sample space. The sample space takes into account the perturbation rate. The combination of the splitting and shifting method, along with perturbation, improves the accurateness of feature extraction. The target function is influenced by factors such as throughput, energy efficiency, and the degree of imbalance. If the newly proposed solution demonstrates superior fitness values compared to the previous generation, it supersedes the old one. The acceptability of the current solution for region selection is determined by its fitness value, as defined in equation (4.2).

$$Fitness = \begin{cases} 1 \; if \; N_{g(a1)} > 0 \\ 0 \; if \; N_{g(a1)} \leq 0 \end{cases} \quad (4.2)$$

The following fitness equation implies that the region selected must contain certain properties related to the target vector, or else the present solution would be rejected. Equation (4.3) represents a target function (T_f) related to the total solution.

$$T_f = \sum \max (E_c) Max(T) Min(Im_{deg}) \quad (4.3)$$

In this context, 'Ec' represents the residual features that need to be maximized, 'Im_{deg}' signifies the degree of imbalance that should be minimized and 'T' denotes throughput (number of features).

The global leadership phase holds significant importance and is tasked with updating the search space, akin to the local leadership phase but with a distinct approach. The selection or update of the differential (image) is determined by the probability value (Py) calculated in (4.4).

$$Py_{Nselect} = 0.9 * \frac{fitness_{n(i)}}{\text{maxp} - \text{fitness}_{\text{cuurent}}} + 0.1 \quad (4.4)$$

Max_p stands for the maximum predicted fitness value, while, $fitness_{current}$ represents the current fitness function value. $Fitness_{n(i)}$ indicates the fitness value of the current region. To assess the likelihood of region selection, the total probability is set to "1," which is subsequently divided into two parts: 0.9 and 0.1.

At this point, a new attempt is performed utilizing the fitness equation. Following that, the fitness value of the existing region is compared to the freshly computed region. The region with the most potential is picked for the construction of the feature vector.

The proposed classifier is based on the Linearity model (LM) and correlation mechanism. The linearity model (LM) for the classification of skin cancer aims to find an optimal function that fits the data points while minimizing the error within a certain tolerance. The basic formulation of LM involves solving an optimization problem. The key components and steps involved in this algorithm are presented in this section.

- **Input data**

Let's assume we have N training samples with features X = {x1, x2, ..., xn} and corresponding continuous target values y = {y1, y2, ..., yn}.

- **Model**

The proposed classifier is based on a correlation mechanism and the Linearity Model (LM). The LM is a classification approach applied for skin cancer classification. Its goal is to create an optimal function that fits the data points while minimizing errors within a particular tolerance. The LM

model seeks to find a function f(x) that predicts the target values y based on the input features. Given a new test sample x*, the LM predicts the corresponding target value y* using the learned model.

- **Training**

The algorithm uses the N training samples to learn the optimal function f(x). This process involves adjusting the model parameters to minimize the errors between the predicted values and the true target values within the defined tolerance. Once the optimal function f(x) has been learned during the training phase, it can be applied to new, unseen data for skin cancer classification. The model makes predictions about whether a given sample represents skin cancer or not based on its features.

- **Optimization problem**

The goal of the LM optimisation problem-solving is to determine the ideal values for the model's weights (w) while regulating the model's complexity and minimising the training error. This is accomplished by utilising the following equation (4.5).

$$minimize\ \left(\frac{1}{2}\right) * \left|\left|w\right|\right|^2 + C * \Sigma\xi_i + \Sigma\xi_i^* \quad (4.5)$$

This expression denotes the objective function's regularisation component. The objective is to reduce the square of the norm of the weight vector 'w.' By penalising big weight values, the regularisation term prevents the model from becoming overly complex. For mathematical simplicity, the factor of 1/2 is frequently given. Σ represent sum of the slack variables ξ_i. The variable ξ_i accounts for some mistakes in the model's predictions. 'C' is a regularisation parameter that governs the trade-off between minimising training errors and minimising model complexity.

Constraints are included in the optimisation problem to ensure that the model's predictions stay within a specific margin (ε) of the true values. The difference between the true target value 'Y_i' and the anticipated value '$f(X_i)$' must be less than or equal to ε plus the slack variable ξ_i, according to this requirement. In other words, it assures that the prediction is not too far off from the genuine value given in equation (4.6).

$$Yi - f(Xi) <= \varepsilon + \xi i \quad (4.6)$$

The constraint defined in equation (4.7) is identical to the preceding one, except it covers the situation where the anticipated value is less than the true value. It also assures that the prediction is within ε of the actual value.

$$f(Xi) - Yi <= \varepsilon + \xi i * \quad (4.7)$$

The constraints in defined equation (4.8) compel the slack variables to be non-negative. This implies they can take values greater than or equal to zero, showing how far a data point deviates from the margin (ε).

$$\xi i, \xi i * >= 0 \quad (4.8)$$

- **Kernel trick**

LM often employs the kernel trick to handle non-linear relationships between features and target values. The input features X_i and X_j are transformed into a higher-dimensional feature space using a kernel function, typically a radial basis function (RBF) kernel, which enables linear separation in that space.

- **Solution**

The LM optimization issue tries to determine the best model weights while balancing model complexity and training mistakes. It includes slack variables to provide some flexibility in dealing with data points that are not completely predicted by the model. The regularisation parameter governs the relevance of minimizing training mistakes about model complexity. The constraints ensure that the model's predictions remain within a certain margin (ε) of the true values, while also preventing the slack variables from going negative. To discover the best answer, this optimization problem is often tackled using quadratic programming techniques.

- **Classification Threshold**

After training the Linearity Model (LM) and using it to make predictions, a classification threshold is applied to the predicted continuous values. The categorization threshold is used to divide the anticipated values into risk levels or categories.

The overall process described involves training a Linearity Model with normalization, dealing with slack variables for classification flexibility, ensuring that predictions remain within a margin of true values, and applying a suitable classification threshold to make practical risk assessments for skin cancer. Several aspects, including data quality, model selection, and the selection of suitable hyperparameters and thresholds would determine the effectiveness of this strategy.

4.3 Summary

This chapter introduces a methodology for the automatic diagnosis of skin cancer, a complex challenge in the field of medical image processing. Its primary objective is to support medical professionals in determining whether a skin melanoma is benign or malignant. Thus, the quest for more effective detection methods to reduce error rates is a critical focus for researchers. The significant contribution of this work lies in the proposal and evaluation of novel medical expert systems capable of autonomously classifying skin cancer tumors as either benign or dangerous, with a level of precision that potentially surpasses that of human experts. The proposed algorithm in this thesis employs an innovative approach for the classification of skin melanomas. It accomplishes this by improving the quality of skin cancer images, isolating tumors from the surrounding skin, extracting and selecting the most relevant tumor features, and ultimately classifying them as melanoma or non-melanoma. Various methods are applied at different stages to create an efficient system. This research aims to make contributions to various facets of the system. The algorithms are designed to expedite detection while minimizing errors compared to traditional methods. The intention is for these proposed algorithms to have a positive impact on public health systems and assist medical experts in the early-stage screening of tumors.

CHAPTER 5

RESULTS AND DISCUSSION

A novel framework has been created and deployed for the detection of Melanoma and Non-Melanoma, specifically designed for the automated identification and analysis of skin lesions or infected patches in dermoscopic images. Comprising pre-processing stages, segmentation algorithms, feature extraction methods, and classification, the framework aims to recognize and characterize these areas for subsequent analysis or diagnosis.

5.1 Evaluation of the Proposed System

The proposed framework initiates pre-processing, addressing noise through hair removal using a bottom hat filter. Contrast enhancement is performed through Discrete Cosine Transform (DCT) and color space conversion, facilitating the detection and isolation of infected regions within the provided dataset. Segmentation employs the background subtraction method, tracking lesions using region properties and a midpoint analysis approach for extracting the Region of Interest (ROI). Various features are extracted from the ROI utilizing the DAA algorithm. In the classification of skin lesions, the system utilizes an ensemble classifier based on a correlation and linearity model, achieving an impressive accuracy of 96.35%.

5.2 Dataset and Tools

The dataset employed in this study is ISIC, a publicly available dataset specifically curated for skin cancer detection. Test results are also provided alongside the dataset. This dataset comprises images depicting both affected and non-affected skin conditions, with specified coordinates for the affected skin outlined in separate files. The simulation is carried out using MathWorks Matlab 2019a. In designing our proposed system, we utilized the visualization tools inherent to Matlab, enabling the creation of an interactive interface that is user-friendly.

5.3 Metrics for Performance Evaluation

The suggested method is evaluated using three parameters: accuracy, sensitivity and specificity. They are discussed as follows.

Accuracy (Ac) refers to the overall performance of a system. Accuracy ensures that the results produced by the system are true and accurate to avoid misdiagnosis. It is calculated using equation (5.1).

$$Ac = \frac{TP+TN}{TP+TN+FP+FN} \tag{5.1}$$

Sensitivity (Sn) denotes the capability of a system to recognize people who have the disease. It shows how good the system is in identifying the person with an actual disease, if the numerical data is high then the possibility of identifying the person with the disease also increases. It is calculated using equation (5.2).

$$Sn = \frac{TP}{TP+FN} \tag{5.2}$$

Specificity (Sp) denotes the capability of a system to recognize people who do not have the disease. It shows how good the system is in identifying an individual who does not have a disease, if the numerical data is high then the possibility of identifying healthy persons also increases. It is calculated using equation (5.3).

$$Sp = \frac{TN}{TN+FP} \tag{5.3}$$

TP means true positive rate (detection of abnormality in a person with disease), TN means true negative rate (in a healthy individual, no abnormalities are detected.), FP means false positive (identification of an anomaly in a healthy individual) and FN means false negativity (no abnormality detected in a person with disease).

5.4 Results of Preprocessing Phase

In this part of our proposed process, we transform the original RGB colour image into a grayscale image. This transformation is an important phase in our process since it helps us identify lesion margins and other important aspects in dermatology. Grayscale images have various advantages in this scenario, especially because they are simpler than full-color ones. Grayscale photos represent differences in grayscale shades, necessitating less information per pixel and making analysis more efficient.

The proposed framework first performs pre-processing in which noise in the form of hair removal is accomplished with the help of a bottom hat method. In this situation, the bottom-hat operation is especially helpful since it highlights darker things of importance against lighter backgrounds. The selection of the shape and size parameters for the structuring element is crucial in

order to correctly distinguish hair structures from the remaining grayscale image. To remove undesirable hair artefacts without compromising the integrity of the underlying image careful selection of the structuring element is essential. DCT and Color space conversion are used for performing image enhancement, this enables the detection and separation of infected regions from the sample (i.e., given dataset).

Figure 5.1 (A) depict the original input image from ISIC dataset, (B) shows the grey scale image of original image, (C) shows the result of bottom hat operation (D) depict the output image without any hair in it and (E) shows the results of image enhancement phase of preprocessing.

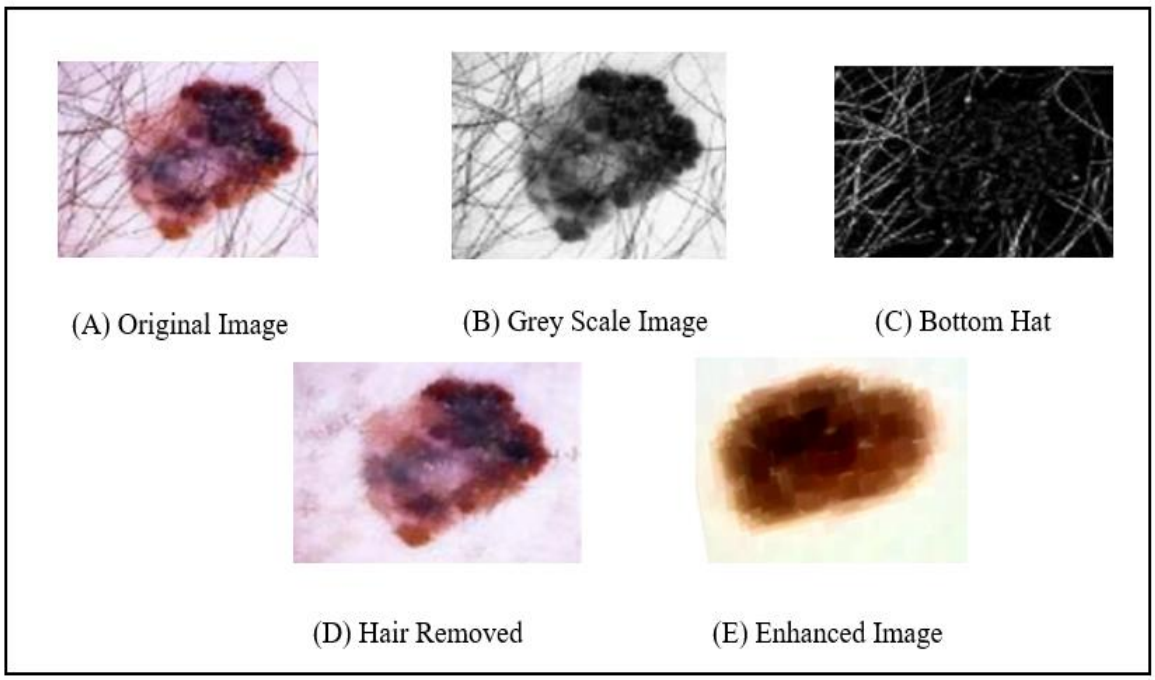


Figure 5.1 Results of Preprocessing Phase

5.5 Results of Segmentation Phase

For the segmentation, the output image of preprocessing phase is given as input for further segmentation operations. The segmentation is performed by applying the background subtraction with midpoint analysis approach to track only lesions from a given image and region props is used to extract the ROI. The result in terms of region separation and extraction is more accurate through improved inside-outside tests and region extraction mechanisms. Using these proposed approach regions overlapping issue is resolved. The Skin images and normal and abnormal separation mechanisms enclose the abnormal shape with boundaries. The grid structure is formed using boundary value analysis, and the outcome accurately separates lesions from normal skin.

Figure 5.2 (A) depicts the input image i.e., enhanced image, (B) shows the result of background subtraction with midpoint analysis method (C) represent the ROI obtained applying region prop mechanism and (D) final output image with ROI.

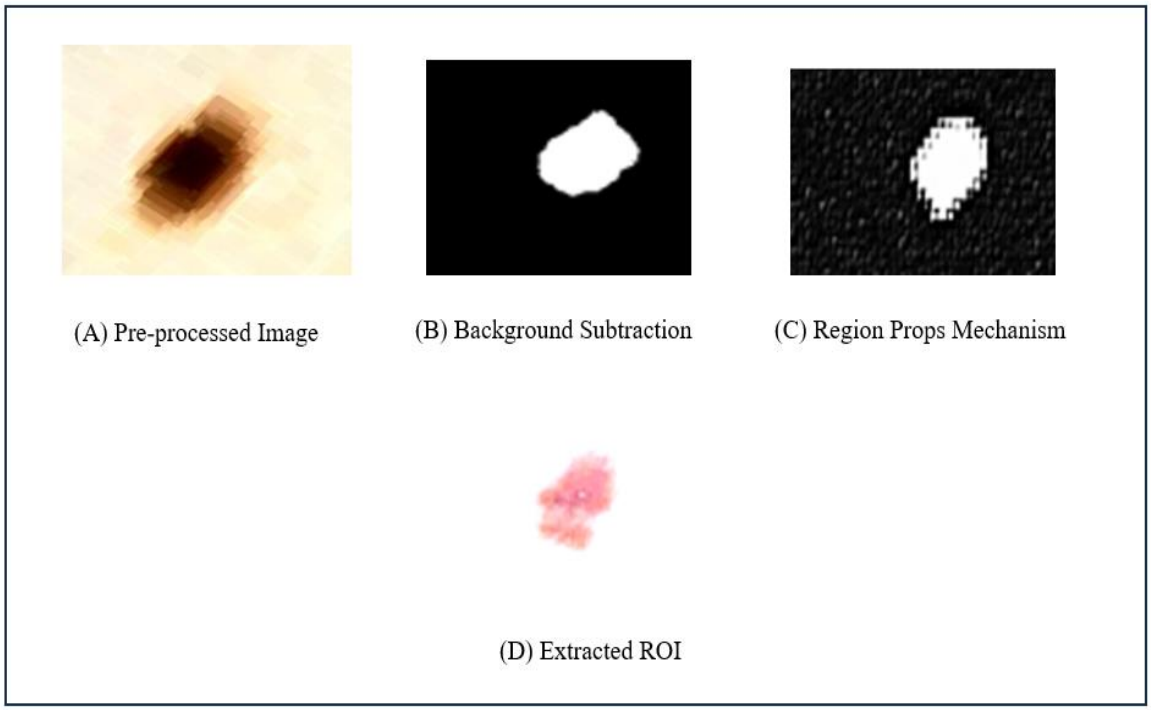


Figure 5.2 Results of Segmentation Phase

The Jacobi index [73] is used to evaluate segmentation performance. The Jacobi index compares the true and estimated results. Jacobi index is given in equation (5.4).

$$JI = \frac{TP}{TP+FN+FP} \tag{5.4}$$

The JI index ranges from 0 to 1. True positive values were set to 1 if both the estimation and the fact were 1. In the incident of an inappropriate prediction, the false-positive and negative values are determined to 1. The proposed method has TP=238, FN=10, TN=8, and a segmentation accuracy of 92%. In some cases, the accuracy is even greater than 98%A comparative study of SVM classifier on the ISIC dataset with different segmentation and feature extraction methods is presented in Table 5.1. Carrera and Dominguez [275] applied the optimal threshold for segmentation and SVM for classification on the ISIC dataset and achieved an accuracy of 75.00%. Thaajwer and Ishanka [276] achieved an accuracy of 85.00% on the ISIC dataset by applying Otsu segmentation and the Watershed method for image segmentation and classifying by implementing an SVM classifier. Kassem, Hosny, and Fouad [277] extracted features from images using Google Net and achieved an accuracy of 81.00% by applying multi-SVM to the ISIC dataset. Pushpalatha et al., [278] achieved

an accuracy of 70.00% by applying deep learning for segmentation on the ISIC dataset and SVM for the classification of cancer images. Keerthana and Venugopal [279] used the ISIC dataset and extracted features using CNN, achieving an accuracy of 87.43% by implementing an SVM classifier. Tembhurne et al., [108] proposed an ensembled method for lesion detection and achieved an accuracy of 93.00% on the ISIC dataset. In the proposed framework, a novel segmentation method "background subtraction with midpoint analysis" for finding ROI and DAA for feature extraction has been proposed. For the classification of skin cancer images, different classifiers have been implemented. It is observed that SVM perform better among other classifiers, then the state-of-the-art methods and achieves an accuracy of 96.21%.

Table 5.1 Comparison Analysis of Proposed Segmentation Method with Different State-of-Art Methods Using SVM Classifier

Author	Dataset	Methods	Classifier	Accuracy
Carrera and Dominguez (2018) [275]	ISIC	•Optimal threshold	SVM	75.00
Thaajwer and Ishanka (2019) [276]	ISIC	•Otsu segmentation and Watershed method	SVM	85.00
Kassem et al., (2020) [277]	ISIC	•Google net	Multiclass SVM	81.00
Pushpalatha et al., (2021) [278]	ISIC	•Deep Learning	SVM	70.00
Keerthana and Venugopal (2022) [279]	ISIC	•CNN	SVM	87.43
Tembhurne et al., (2023) [108]	ISIC	•Deep learning and machine learning	SVM	93.00
Proposed method	**ISIC**	**Background subtraction with midpoint analysis**	**SVM**	**96.21**

A comparison of the proposed methods with the different classifiers using different segmentation methods is presented in Table 5.2. Murgan et al., [192] applied the watershed algorithm for segmentation of skin cancer images and extracted GLMC-based features. For classification, SVM, RF, and KNN were implemented and attained the maximum accuracy of 85.72%. Bassel et al., [280] proposed an approach for classifying melanoma and benign skin cancers based on the layering of classifiers and method achieved a notable accuracy of 90.00%. Wang et al., [281] proposed a methodology for cancer detection using the ISIC dataset, leveraging a deep learning-based approach. The model demonstrated a commendable accuracy of 92.00%, showcasing the

effectiveness of the proposed technique. Das et al., [282] implemented the ABCD rule for feature extraction and various machine learning methods in the context of skin cancer detection on the ISIC dataset. The approach exhibited a high accuracy of 93.51%, highlighting the efficacy of combining feature extraction methodologies with machine learning techniques for improved diagnostic outcomes.

Table 5.2 Comparison Analysis of Proposed Segmentation Methods with Different Segmentation Methods

Author	**Methods**	**Performance measure (Accuracy %)**			
		SVM	**KNN**	**RF**	**DT**
Murugan et al., (2021) [192]	•Watershed segmentation •ABCD rule •GLMC	85.72	69.54	74.32	NA
Bassel et al., (2022) [280]	Xception, VGG16 Resnet50,	86.70	81.00	80.30	75.70
Wang et al., (2022) [281]	•Deep learning	81.00	76.00	NA	NA
Das et al., (2022) [282]	•ABCD Rule	93.51	91.00	90.67	NA
Proposed Method	• Background Subtraction with Midpoint Analysis	**96.21**	**92.20**	**90.0**	**89.60**

The proposed segmentation approach (background subtraction with midpoint analysis) is implemented using SVM, KNN, RF, and DT classifiers. The outcomes demonstration that the proposed technique performed better with all classifiers in terms of accuracy.

5.6 Results of Feature Extraction and Selection Phase

Developing an efficient classification system for identifying and categorizing skin lesions requires careful consideration of both feature extraction and selection in the context of skin cancer image analysis. In order to help classify benign and malignant skin lesions, these procedures include extracting pertinent information from the image.

To discover the best features in proposed system, a DAA algorithm is used to extract and select the feature set. This methodology tries to automatically discover the most important traits, which can lead to more accurate and efficient skin lesion diagnosis and classification. The performance of the proposed algorithm will be checked based on the selected features. It works by iteratively developing a population of candidate solutions (feature sets) over numerous generations to discover the best set of features that meet specific performance criteria.

The DAA develops offspring during the optimisation process by changing or combining existing feature sets. These offspring are assessed using a fitness function that assesses how well they execute the task. Offspring who do not produce the strongest results may be removed in each iteration, assisting in the refinement of the feature set. Each fittest offspring is branded with a unique trait. The actual point plotted as shown in Figure 5.3 represents the multiple features in a single image and the DAA plotted points are the selected features for the proposed methodology. Color (Blue-grey, Black, Red, White, Dark brown, Light brown), Shape (Area, Length of major axis, Length of minor axis) and Texture (Energy, Correlation, Contrast).

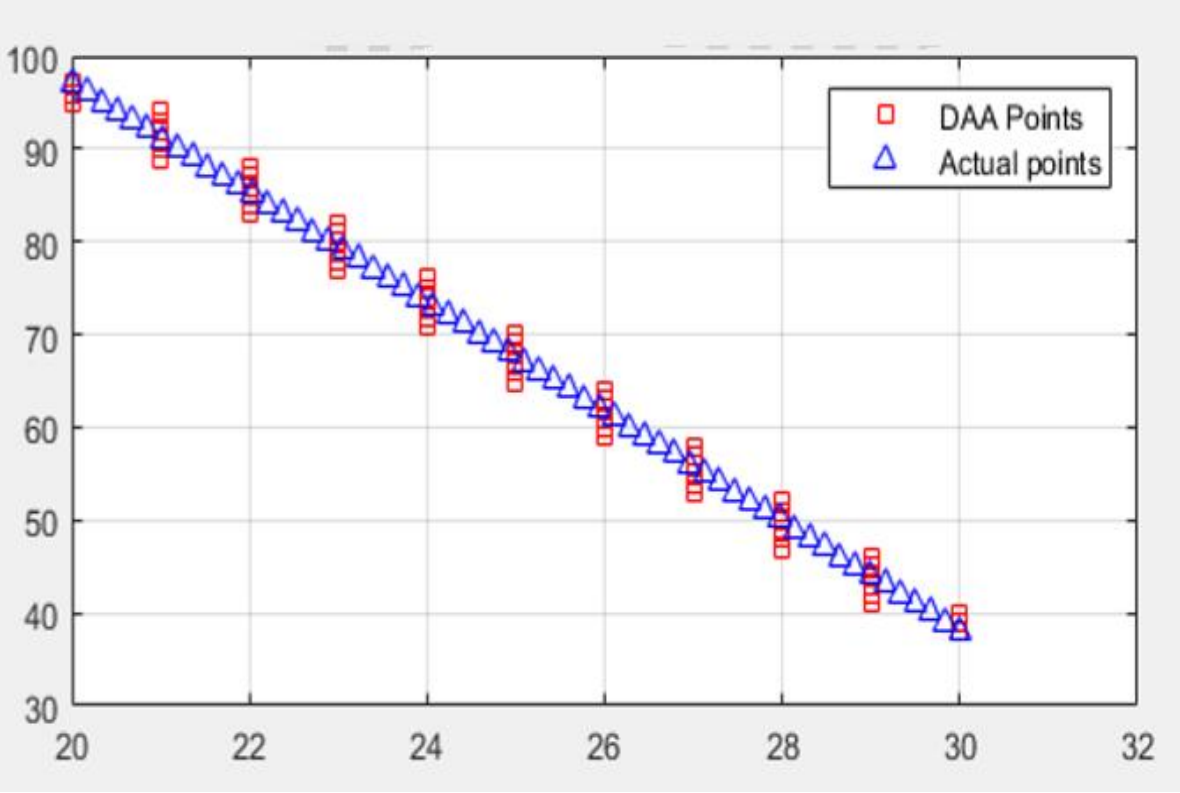


Figure 5.3 Shows Feature Extraction and Selection Using DAA

5.6.1 Results of Normalization

This phase seeks to increase the accuracy of classification by merging all of the feature sets (colour, texture, and shape) into a feature vector. The process involves scaling the values of each feature within the matrix to fit within a predefined range. By doing so, the model ensures that all features are on a level playing field and that no single feature dominates the classification process due to differences in scale. During the normalization phase, three primary normalization mechanisms are employed: depth-based normalization, min-max normalization, and gray scale normalization. Following the specified operation, the three normalization matrices are shown in Figure 5.4.

10815	394.556	138.776	11415	385.65
9663	376.56	135.656	10160	389.65
8024	325.633	106.562	8215	320.565
7656	319.52	109.562	7896	315.265
8225	320.25	102.365	8230	315.248
8583	332.074	102.458	8596	327.458

a) Without Normalization

0.0815	0.039456	0.013878	0.1415	0.038565
0.9663	0.037656	0.013566	0.016	0.038965
0.8024	0.032563	0.010656	0.8215	0.032057
0.7656	0.031952	0.010956	0.7896	0.031527
0.8225	0.032025	0.010237	0.823	0.031525
0.8583	0.033207	0.010246	0.8596	0.032746

b) Gray Scaling

0.1815	0.139456	0.113878	0.2415	0.138565
0.0663	0.137656	0.113566	0.116	0.138965
0.9024	0.132563	0.110656	0.9215	0.132057
0.8656	0.131952	0.110956	0.8896	0.131527
0.9225	0.132025	0.110237	0.923	0.131525
0.9583	0.133207	0.110246	0.9596	0.132746

c) Min-Max

0.2015	0.159456	0.133878	0.2615	0.158565
0.0863	0.157656	0.133566	0.136	0.158965
0.9224	0.152563	0.130656	0.9415	0.152057
0.8856	0.151952	0.130956	0.9096	0.151527
0.9425	0.152025	0.130237	0.943	0.151525
0.9783	0.153207	0.130246	0.9796	0.152746

d) Depth Based

Figure 5.4 Colour Feature Matrix After Normalization

5.6.2 Performance Evaluation of Different Classifiers on Different Feature Sets

The presented results in Table 5.3 highlight the performance of different classifiers with distinct sets of features, revealing varying accuracies for each combination of feature set and classifier. This comprehensive analysis contributes to a thorough understanding of the effectiveness of different feature sets in the context of skin cancer detection with diverse classifiers. The performance evaluation of various classifiers using different feature sets is categorized into two cases, namely Case 1 and Case 2:

Table 5.3 Performance Evaluation of Different Classifiers on Different Feature Sets

Case	Type of features	Classifier Performance (Accuracy %)			
		SVM	**KNN**	**RF**	**SCCC**
1	Shape: (Area, Length of major axis, Length of minor axis) Color: (Blue-grey, Black, Red, White)	95.30	92.20	90.0	95.87
2	Shape: (Area, Length of major axis, Length of minor axis) Color: (Blue-grey, Black, Red, White, Dark brown, Light brown) Texture: (Energy, Contrast, Correlation)	96.21	93.75	92.03	96.35

5.7 Result Analysis of Different Classifiers

This section presents the performance evaluation corresponding to different classifiers within the methodology. The proposed approaches for segmentation, feature extraction, and classification contribute to the enhanced accuracy of the overall system. To validate the integrity of the image, correlation analysis is employed by comparing the results with the original image. The outcomes of the correlation analysis are detailed in Table 5.4.

Table 5.4 Level of Correlation with Different Features of Image

Images	Values of Correlation Analysis	Level of correlation	Result
Image A	0.90	0-1	High correlated
Image B	0.87	0-1	High correlated
Image C	0.24	0-1	Positively correlated
Image D	0.77	0-1	High correlated
Image E	0.2	0-1	Less correlated

This section provides a performance analysis of classifier accuracy, sensitivity and specificity using features selected by the DAA approach with various normalization techniques in Table 5.5.

Table 5.5 Metrics for Performance Evaluation

Classifier	Specificity (%)	Sensitivity (%)	Accuracy (%)
KNN	76.6	86.7	93.75
RF	96.4	92.1	92.03
SVM	95.3	92	96.21
SCCC	97.6	95.2	96.35

The proposed approach incorporates three different normalizations, recording the best, average, and worst possible classification accuracy. Simulation results with diverse normalization mechanisms are documented in Figure 5.5, Figure 5.6, and Figure 5.7. The proposed approach used a customized classifier-based approach in achieving the classification of test images. The best possible approach in terms of classification accuracy will be selected by implementing and analyzing results of five different classifier (SVM, KNN, RF, DT and SCCC)). As multiple classifiers predicted similar results, classification accuracy using a customized classifier-based approach was high.

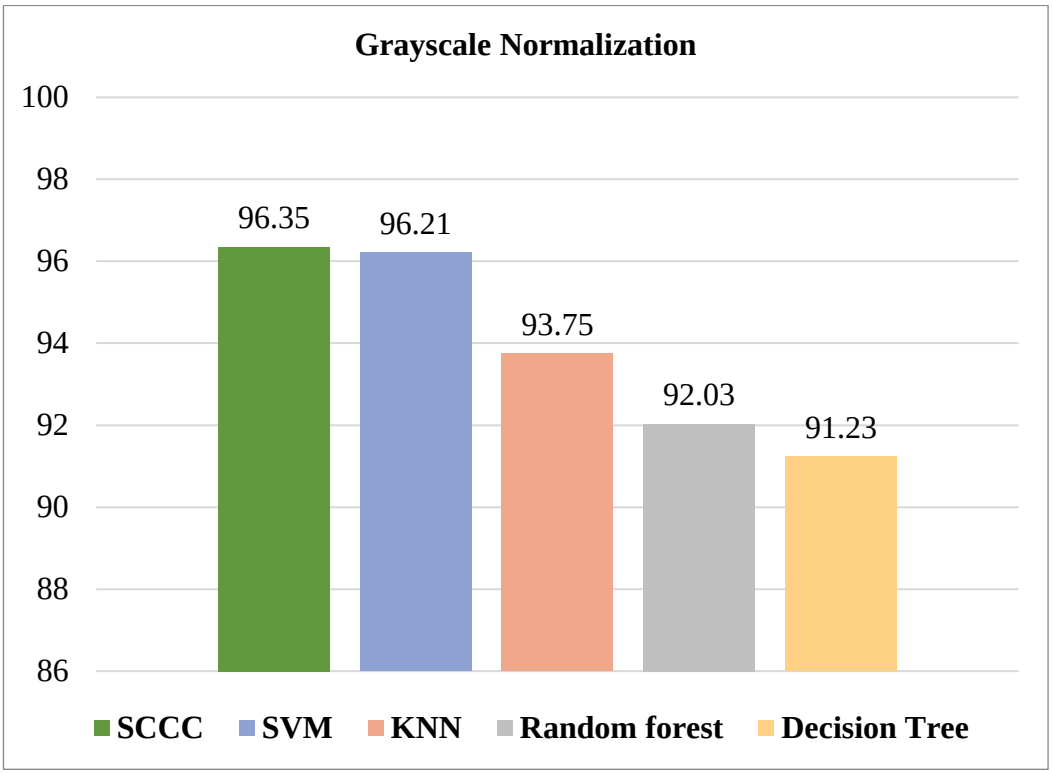


Figure 5.5 Performance Analysis of Different Classifiers using Grayscale Normalization

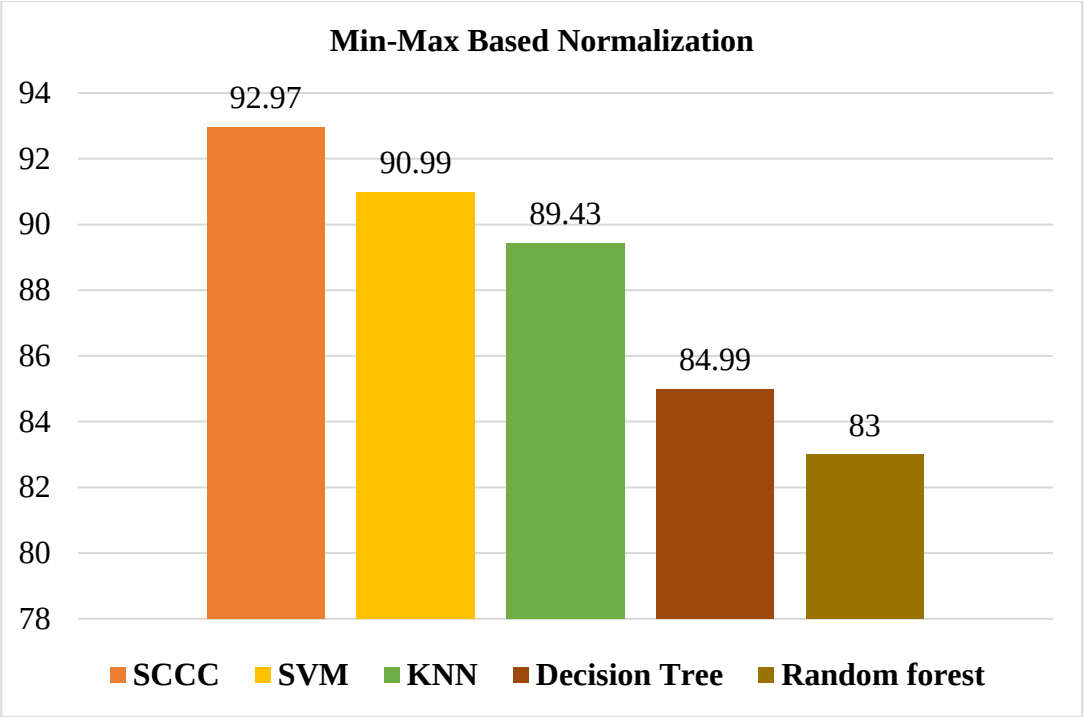


Figure 5.6 Performance Analysis of Different Classifiers using Min-Max Normalization

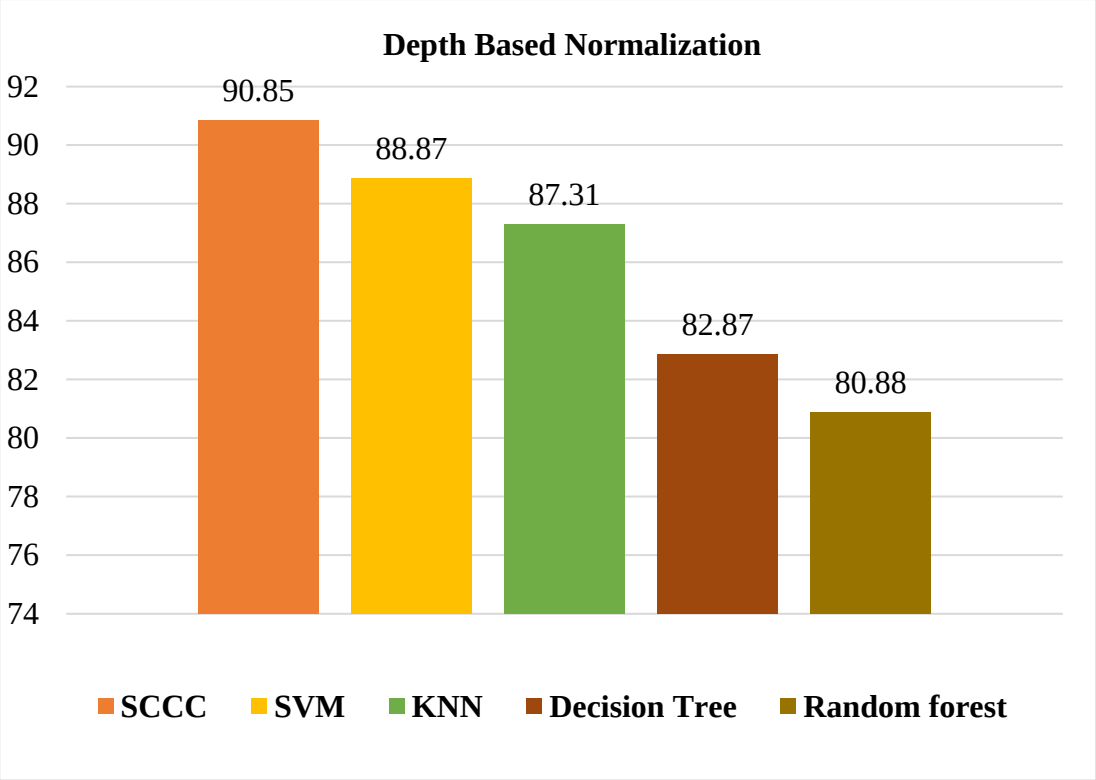


Figure 5.7 Performance Analysis of Different Classifiers using Depth based Normalization

The customized classifier (SCCC) exhibits superior performance compared to all other classifiers, as illustrated in Figure 5.5, Figure 5.6, and Figure 5.7. Specifically, the performance of the customized classifier under greyscale normalization surpasses that of min-max and depth-based normalization. Consequently, the classification outcome of SCCC (accuracy 96.35%) in the context of DAA will be employed in the subsequent section for comparison with state-of-the-art techniques. It is important to note a limitation of our work, which involves utilizing only one dataset (ISIC) with two cases (melanoma and benign) for classification and detection. Results may vary based on the dataset employed.

5.8 Results Comparisons with the State of Art Methods

A comparative study of various techniques for skin cancer classification and detection on the different dataset are presented in Table 5.6. Satheesa et al., [148] has used three different datasets ISIC, PH2 and ATLAS and applied adaptive snake technique for segmentation, 3D reconstruction of dermoscopic images for feature extraction and SVM, AdaBoost and BOF for classification. Codella et al., [283] has implemented FCN for segmentation and CNN, U-NET and DRN for classification on the ISIC dataset and achieved an accuracy of 76%. Zortea et al., [146] has used microscopic images and applied SWOT for segmentation. Smaoui and Derbel, [284] has applied Multi-threshold Otsu for segmentation, ABCD rule for feature extraction and TVD rule for classification on the PH2 dataset and achieved an accuracy of 90%. Carrera and Dominguez, [275] has achieved an accuracy of 75% by applying the Optimal threshold for segmentation and SVM, and DT for classification on the ISIC dataset. Al-Masni et al., [152] has applied FCRN on the ISIC and PH2 data set for segmentation of cancer images with 95% accuracy. Murugan et al., [192] has achieved an accuracy of 85% by applying the Watershed approach for segmentation, ABCD rule for feature extraction and SVM, KNN, and RF for classification on the ISIC dataset. Thaajwer and Ishanka, [227] has used the ISIC dataset and achieved an accuracy of 85% by implementing Otsu segmentation and Watershed method for segmentation and NN, SVM, and CNN for classification. Khan et al., [285] has proposed an Improved K-means clustering for segmentation and SVM for classification on the DERMIS dataset and achieve an accuracy of 96%. Kassem et al., [277] has applied Google net for feature extraction and Multiclass SVM for classification on the ISIC dataset with an accuracy of 81%. Albert, [286] has applied CNN and SVM on the Med Node dataset and achieved an accuracy of 91%. Hohn et al., [287] achieve an accuracy of 92% by applying deep learning CNN. Jiang et al., [288] has applied DRA-Net on the Histopathological Image set and

achieved an accuracy of 86%. Murgan et al., [161] has used the mean shift method for segmentation and SVM, probabilistic NN, and random forests for classification and achieved an accuracy of 89%.

Table 5.6 Comparison of Various Techniques for Skin Cancer Classification and Detection on Different Dataset

Authors	**Datasets**	**Techniques**	**Accuracy %**
Satheesa et al., (2017) [148]	ISIC PH2 ATLAS	Adaptive snake for segmentation 3D reconstruction of dermoscopic images for feature extraction SVM, AdaBoost and BOF for classification	NA
Codella et al., (2017) [283]	ISIC	Fully convolution network (FCN) for segmentation CNN, U-Net and deep residual network for classification	76.00
Zortea et al., (2017) [146]	Microscopic images	Shading attenuation approach for preprocessing A simple weighted Otsu thresholding (SWOT) method for segmentation Morphological operation for post-processing	NA
Smaoui, and Derbel, (2018) [284]	PH2	Median filtering for preprocessing Multi-threshold Otsu for segmentation ABCD rule for feature extraction TVD rule for classification	90.00
Carrera and Dominguez (2018) [275]	ISIC	Dilation approach and histogram equalization for preprocessing The optimal threshold for segmentation SVM, DT for classification	75.00
Al-Masni et al., (2018) [152]	ISIC PH2	Full-resolution convolutional networks for segmentation	95.62
Murugan et al., (2019) [192]	ISIC	A median filter for preprocessing Watershed approach for segmentation ABCD rule for feature extraction SVM, KNN, and RF for classification	85.72
Thaajwer and Ishanka (2019) [227]	ISIC	Hough transformation for hair removal Otsu segmentation and watershed method for segmentation	85.00

		NN, SVM, CNN	
Khan et al., (2019) [285]	DERMIS	Gaussian filter for preprocessing Improved K-means clustering for segmentation SVM for classification	96.00
Kassem et al., (2020) [277]	ISIC	Google net for feature extraction Multiclass SVM for classification	81.00
Albert (2020) [286]	Med node	CNN SVM	91.00
Hohn et al., (2020) [287]		Deep learning CNN	92.60
Jiang et al., (2021) [288]	Histopathological Image set	Deep learning framework DRA-Net	86.60
Murgan et al., (2021) [161]	ISIC	A median filter for pre-processing Mean shift method for segmentation SVM, probabilistic NN, and random forests for classification	89.00
Proposed System	ISIC	Bottom hat filter and DCT for preprocessing Background subtraction with midpoint and region props for segmentation DAA for feature extraction SCCC	**96.35**

5.9 Results and Performance Evaluation on the ISIC Dataset

In Table 5.7, the performance of the proposed system is compared with state-of-the-art methods using the publicly available ISIC dataset. Previous studies, such as those conducted by authors in[275]**,**[192], [227], [277] and [161] utilized the ISIC dataset and employed various techniques for preprocessing, segmentation, feature extraction, and classification, achieving the highest accuracy of 89%. Notably, the proposed system surpasses this performance, demonstrating superior accuracy of 96.35%.

Table 5.7 Comparative Analysis of the Proposed SCCC System with State of Art Methods

Authors	Datasets	Techniques	Accuracy %
Carrera and Dominguez (2018) [275]	ISIC	Dilation approach and histogram equalization for preprocessing The optimal threshold for segmentation SVM, DT for classification	75.00
Murugan, Nair, and Kumar (2019) [192]	ISIC	A median filter for preprocessing Watershed approach for segmentation ABCD rule for feature extraction SVM, KNN, and RF for classification	85.72
Thaajwer and Ishanka (2019) [227]	ISIC	Hough transformation for hair removal Otsu segmentation and Watershed method for segmentation NN, SVM, CNN	85.00
Kassem, Hosny, and Fouad (2020) [277]	ISIC	Google net for feature extraction Multiclass SVM for classification	81.00
Murgan et al. (2021) [161]	ISIC	A median filter for pre-processing Mean shift method for segmentation SVM, probabilistic NN, and random forests for classification	89.00
Proposed System	ISIC	Bottom hat filter and DCT for preprocessing Background subtraction with midpoint and region props for segmentation DAA for feature extraction SCCC	**96.35**

5.10 Results and Comparison with Deep Learning Methods

To evaluate the performance of the customized classifier on the ISIC dataset in the proposed system comparison with the different existing classifiers on the same dataset is shown in Table 5.8 and with different dataset is shown in Table 5.9. Authors in [289][289], [290], [118], and [178] use the same dataset and different classifiers for melanoma classification and achieved the highest accuracy of

96.00%. The proposed system is reliable, resilient, and efficient which achieves an accuracy of 96.35% on the same dataset.

Table 5.8 Comparative Analysis of the Proposed SCCC System with Deep Learning Methods

Author	Dataset	Classifier	Accuracy (%)
Perez et al., 2018 [116]	ISIC	CNN	87.40
Serte and Demirel 2019 [117]	ISIC	Gabor wavelet-based CNN	96.00
Zhang et al., 2019 [118]	ISIC	Deep CNN	87.50
Mahbod et al., 2020 [178]	ISIC	Multi-CNN	86.20
Kamrul et al., 2022 [207]	ISIC	Hybrid-CNN	96.00
Proposed System	**ISIC**	**SCCC**	**96.35**

Table 5.9: A Comparative Analysis of Deep Learning Methods for Skin Cancer Classification and Detection

Author	Technique	Classifiers	Accuracy (%)
Xie F et al., 2017 [92]	Deep Learning	SGNN	91.11
Zhang et al., 2019 [292]	Deep Learning	ARL-CNN	91.70
Hassan et al., 2019 [293]	Deep Learning	CNN	93.70
Wu et al., 2020 [100]	Deep Learning	DenseNet	83.70
Javaid et al., 2021 [294]	Deep Learning	OTSU, PCA, SVM, RF	93.89
Gouda et al., 2022 [295]	Deep Learning	ESRGAN, CNN, Transfer Learning	83.20
Reis et al., 2022 [296]	Deep Learning	InSiNet, CNN	94.59
Nadiger et al., 2023 [297]	Deep Learning	CNN	90.00
Sundari et al., 2023 [298]	Deep Learning	InceptionV3, Inception ResNet, DenseNet, and VGG-16,	95.00
Pangsibidang et al.,	Deep Learning	CNN, K-means clustering,	78.00

2023 [299]		Genetic algorithm, Naive Bayes and SVM	
Viknesh et al., 2023 [300]	Deep Learning	AlexNet, LeNet, VGG-16 models and SVM	91.00
Tembhurne et al., 2023 [301]	Deep Learning	Machine learning, Deep learning	93.00
Keethna et al., 2023 [302]	Deep Learning	CNN, SVM	88.02
Proposed method	**Machine Learning**	**SCCC**	**96.35**

Almost every existing work lacked optimization in feature selection. The selection ratio in the proposed work was 50%. The reduction can be furthered in the feature selection mechanism to determine variation. As seen from the result section, exponential variation was observed. There are the following reasons for the improvement in the results:

In the proposed method, some important features are extracted and employed. Each element influences a different aspect of skin lesion because each type of skin lesion has a distinct colour. Differentiation can be recognized using colour features, Shape features and textural features.

A differential analyser algorithm is used since it has the least calculation complexity. The extracted features were scaled to fit onto a line. Due to optimised space exploration, DAA ensures a better selection of features and avoids local minima. This model selects optimised features by verifying these features against optimised functions.

For qualitative evaluation, the experimental results of applying the suggested technique to dermoscopy images are compared with actual data. A comparison analysis is performed against several additional approaches in addition to qualitative assessment. Our strategy produced the best results based on the evaluation metric. As a result, we believe that our effective and dependable technique has the potential to assist dermatologists in making more accurate and timely diagnoses.

5.11 Summary

In this chapter, we delve into the outcomes of the proposed system. The ISIC dataset has proven to be a valuable source of images, yielding excellent results. To prepare the images for preprocessing, certain adjustments were necessary to meet the specific requirements. We provide an in-depth analysis of the results achieved in segmentation, feature extraction, and the application of the customized classifier.Our assessment covers texture, shape, and color features, which were evaluated using training sets in the context of traditional machine learning models. The performance of this classifier is noteworthy, with all models consistently demonstrating accuracy ranging between 94% and 96.35%. It becomes evident that the segmentation and feature extraction methods proposed play a crucial role in the model's construction and overall succe

CHAPTER 6

CONCLUSION AND FUTURE SCOPE

A novel approach for detecting skin cancer was introduced in this study. Before executing the proposed segmentation method, some image processing techniques were used to improve recognition accuracy. To acquire colour, shape, and texture features, the system employs a feature extraction and selection method, which is then used in conjunction with a customized classifier for lesion categorization.

6.1 Contribution of Proposed System to Society

A skin cancer expert system contributes significantly to society by functioning as a beneficial instrument in the fields of healthcare and dermatology. It plays a critical role in the early identification and accurate diagnosis of skin cancer, resulting in timely intervention and improved patient outcomes. These expert systems improve healthcare practitioners' efficiency in analysing enormous volumes of dermatological data by utilising new technologies such as machine learning and deep learning. As a result, faster and more precise diagnoses are possible, lowering the load on healthcare systems and potentially saving lives. Furthermore, the availability of such expert systems can extend healthcare services to remote or underdeveloped locations, allowing individuals to receive timely and trustworthy skin cancer examinations.

6.2 Conclusion

In conclusion, the proposed model presents a significant advancement in dermatological image analysis, particularly in the context of melanoma detection. Through a comprehensive approach that involves preprocessing, segmentation, feature extraction, and classification, the system achieves commendable results, showcasing its potential for practical application in the medical field.

One of the notable strengths of the proposed model lies in its effective segmentation phase, which employs background subtraction with a midpoint and region prop mechanism. This process is crucial for isolating and highlighting relevant features within the images, a pivotal step in dermatological analysis. The Jacobi index is utilized as a performance metric, comparing true and estimated results. With a segmentation accuracy of 95.30%, the proposed model demonstrates its

proficiency in accurately identifying and segmenting skin lesions. This accuracy is further underscored by the normalization phase, which aims to merge color, texture, and location features into a cohesive feature vector. The three normalization techniques grey scaling, min-max, and depth-based—contribute to refining the feature matrices.

In terms of classification, the model employs different classifiers, with customized classifiers consistently outperforming others. The classification accuracy of the proposed model, reaching 96.35%, is a testament to its efficacy in distinguishing between malignant and benign skin lesions. The comparative analysis with other classifiers and normalization techniques reaffirms the superiority of the proposed approach.

Looking beyond the experimental results, the proposed model is positioned favourably when compared with existing state-of-the-art methods. This underscores its competitiveness and potential for integration into clinical settings. The model's high accuracy, coupled with its ability to adapt to various normalization techniques, positions it as a robust tool for dermatologists and healthcare professionals.

In terms of future scope, there are several avenues for refinement and expansion. The feature reduction mechanism could be further optimized to enhance performance and reduce computational complexity. Exploring and validating the model with diverse datasets encompassing a broader range of dermatological conditions would contribute to its generalizability. Additionally, real-world clinical validation and collaboration with dermatologists could provide valuable insights into the model's practical utility and potential integration into diagnostic workflows.

Continued research into emerging technologies and methodologies in dermatology, such as advancements in imaging techniques or incorporation of artificial intelligence for real-time analysis, could inform future iterations of the model. Collaboration with the medical community and adherence to evolving standards and protocols would be essential for the model's seamless integration into the healthcare landscape.

In essence, the proposed model represents a promising step forward in leveraging computational techniques for dermatological diagnosis. Its accuracy, adaptability, and potential for further enhancement position it as a valuable tool in the ongoing efforts to improve early detection and diagnosis of skin lesions, particularly melanoma.

6.3 Future Scope

The proposed dermatological image analysis model, while showcasing promising results, opens up a realm of future possibilities and areas for further exploration. The dynamic landscape of medical imaging and artificial intelligence presents several avenues for improvement and expansion, positioning the model as a foundation for ongoing research and innovation.

One prominent area for future development lies in diversity in datasets. Future research should involve the integration of larger and more diverse datasets, encompassing a broader spectrum of dermatological conditions beyond melanoma. This expansion would contribute to a more comprehensive understanding of the model's capabilities and limitations across various skin lesions.

Real-world clinical validation is paramount for the translation of computational models into practical healthcare applications. Collaborations with dermatologists and medical practitioners can provide valuable insights, ensuring that the model aligns with clinical workflows and meets the stringent standards of the medical community. Integration into existing diagnostic processes and electronic health record systems is an essential step toward seamless adoption.

Artificial intelligence (AI) and machine learning (ML) are continuously evolving, and incorporating the latest advancements into the proposed model can enhance its diagnostic capabilities. Ongoing research in AI-driven image analysis, including deep learning architectures and transfer learning, presents opportunities for increasing the model's sophistication and accuracy.

In conclusion, the future scope of the proposed dermatological image analysis model is rich with possibilities. Fine-tuning feature reduction mechanisms, diversifying datasets, engaging in clinical collaborations, exploring advanced imaging technologies, embracing AI advancements, and extending the model's coverage to various skin conditions collectively represent a roadmap for continued research and innovation. The ongoing synergy between computational approaches and clinical expertise holds the key to advancing dermatological diagnostics and improving patient outcomes.

REFERENCE

[1]N. H. Khan *et al.*, "Skin cancer biology and barriers to treatment: Recent applications of polymeric micro/nanostructures," *Journal of Advanced Research*, vol. 36. Elsevier B.V., pp. 223–247, Feb. 01, 2022. doi: 10.1016/j.jare.2021.06.014.

[2]K. Urban, S. Mehrmal, P. Uppal, R. L. Giesey, and G. R. Delost, "The global burden of skin cancer: A longitudinal analysis from the Global Burden of Disease Study, 1990–2017," *JAAD Int*, vol. 2, pp. 98–108, Mar. 2021, doi: 10.1016/j.jdin.2020.10.013.

[3]"2022-cancer-facts-and-figures".

[4]K. Saginala, A. Barsouk, J. S. Aluru, P. Rawla, and A. Barsouk, "Epidemiology of Melanoma," *Medical sciences (Basel, Switzerland)*, vol. 9, no. 4. NLM (Medline), Oct. 20, 2021. doi: 10.3390/medsci9040063.

[5]D. C. Grossman *et al.*, "Behavioral counseling to prevent skin cancer: US preventive services task force recommendation statement," *JAMA - Journal of the American Medical Association*, vol. 319, no. 11, pp. 1134–1142, Mar. 2018, doi: 10.1001/jama.2018.1623.

[6]"About Melanoma Skin Cancer What Is Melanoma Skin Cancer?"

[7]A. Jiang *et al.*, "Skin cancer discovery during total body skin examinations," *Int J Womens Dermatol*, vol. 7, no. 4, pp. 411–414, Sep. 2021, doi: 10.1016/j.ijwd.2021.05.005.

[8]M. Krishna Monika, N. Arun Vignesh, C. Usha Kumari, M. N. V. S. S. Kumar, and E. Laxmi Lydia, "Skin cancer detection and classification using machine learning," in *Materials Today: Proceedings*, Elsevier Ltd, 2020, pp. 4266–4270. doi: 10.1016/j.matpr.2020.07.366.

[9]U. Saghir and M. Hasan, "Skin cancer detection and classification based on differential analyzer algorithm," *Multimed Tools Appl*, 2023, doi: 10.1007/s11042-023-14409-x.

[10] E. Z. Keung and J. E. Gershenwald, "The eighth edition American Joint Committee on Cancer (AJCC) melanoma staging system: implications for melanoma treatment and care," *Expert Review of Anticancer Therapy*, vol. 18, no. 8. Taylor and Francis Ltd, pp. 775–784, Aug. 03, 2018. doi: 10.1080/14737140.2018.1489246.

[11] U. Saghir and V. Devendran, "A Brief Review of Feature Extraction Methods for Melanoma Detection," in *2021 7th International Conference on Advanced Computing and Communication*

Systems, ICACCS 2021, Institute of Electrical and Electronics Engineers Inc., Mar. 2021, pp. 1304–1307. doi: 10.1109/ICACCS51430.2021.9441787.

[12] L. E. Hernandez, F. S. Frech, N. Mohsin, I. Dreyfuss, and K. Nouri, "Nodular Melanoma: A Review of Pathogenesis, Presentation, Diagnosis, and Treatment," 2021.

[13] F. M. Trindade, M. L. P. de Freitas, and F. V. Bittencourt, "Dermoscopic evaluation of superficial spreading melanoma," *An Bras Dermatol*, vol. 96, no. 2, pp. 139–147, Mar. 2021, doi: 10.1016/j.abd.2020.06.012.

[14] H. Iznardo, C. Garcia-Melendo, and O. Yélamos, " Lentigo Maligna: Clinical Presentation and Appropriate Management ," *Clin Cosmet Investig Dermatol*, vol. Volume 13, pp. 837–855, Nov. 2020, doi: 10.2147/CCID.S224738.

[15] X. Hao *et al.*, "Acral lentiginous melanoma of foot and ankle: A clinicopathological study of 7 cases," *Anticancer Res*, vol. 39, no. 11, pp. 6175–6181, 2019, doi: 10.21873/anticanres.13825.

[16] "Amelanotic Melanoma," in *Color Atlas of Melanocytic Lesions of the Skin*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 204–212. doi: 10.1007/978-3-540-35106-1_30.

[17] J. Dinnes *et al.*, "Visual inspection for diagnosing cutaneous melanoma in adults," *Cochrane Database of Systematic Reviews*, vol. 2018, no. 12, Dec. 2018, doi: 10.1002/14651858.CD013194.

[18] M. E. Vestergaard, P. Macaskill, P. E. Holt, and S. W. Menzies, "Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: A meta-analysis of studies performed in a clinical setting," *British Journal of Dermatology*, vol. 159, no. 3, pp. 669–676, 2008, doi: 10.1111/j.1365-2133.2008.08713.x.

[19] M. E. Celebi, H. Iyatomi, G. Schaefer, and W. V. Stoecker, "Lesion border detection in dermoscopy images," *Computerized Medical Imaging and Graphics*, vol. 33, no. 2, pp. 148–153, 2009, doi: 10.1016/j.compmedimag.2008.11.002.

[20] J. Kato, K. Horimoto, S. Sato, T. Minowa, and H. Uhara, "Dermoscopy of Melanoma and Non-melanoma Skin Cancers," *Front Med (Lausanne)*, vol. 6, no. August, pp. 1–7, 2019, doi: 10.3389/fmed.2019.00180.

[21] B. Nirmal, "Dermatoscopy: Physics and principles," *Indian Journal of Dermatopathology and Diagnostic Dermatology*, vol. 4, no. 2, p. 27, 2017, doi: 10.4103/ijdpdd.ijdpdd_13_17.

[22] I. Zalaudek *et al.*, “Dermoscopy in general dermatology,” *Dermatology*, vol. 212, no. 1. pp. 7–18, Nov. 2005. doi: 10.1159/000089015.

[23] D. M. Elston, E. J. Stratman, and S. J. Miller, “Skin biopsy: Biopsy issues in specific diseases,” *Journal of the American Academy of Dermatology*, vol. 74, no. 1. Mosby Inc., pp. 1–16, Jan. 01, 2016. doi: 10.1016/j.jaad.2015.06.033.

[24] D. Manocha, N. Bansal, and R. S. Farah, “3 Types and Selection Criteria for Various Skin Biopsy Procedures.” [Online]. Available: www.intechopen.com

[25] L. T. Dengel *et al.*, “Total body photography for skin cancer screening.”

[26] “THE ROLE OF TOTAL BODY PHOTOGRAPHY IN EARLY MELANOMA DETECTION A literature review Student: Ann Praveena Rameshkumar,” 2019.

[27] R. Hofmann-Wellenhof *et al.*, “Reflectance Confocal Microscopy-State-of-Art and Research Overview,” *Seminars in Cutaneous Medicine and Surgery*, vol. 28, no. 3. pp. 172–179, Sep. 2009. doi: 10.1016/j.sder.2009.06.004.

[28] N. Shahriari, J. M. Grant-Kels, H. Rabinovitz, M. Oliviero, and A. Scope, “Reflectance confocal microscopy,” *J Am Acad Dermatol*, vol. 84, no. 1, pp. 1–14, Jan. 2021, doi: 10.1016/j.jaad.2020.05.153.

[29] B. H. Oh, K. H. Kim, and K. Y. Chung, “Skin Imaging Using Ultrasound Imaging, Optical Coherence Tomography, Confocal Microscopy, and Two-Photon Microscopy in Cutaneous Oncology,” *Front Med (Lausanne)*, vol. 6, Nov. 2019, doi: 10.3389/fmed.2019.00274.

[30] J. Dinnes *et al.*, “High-frequency ultrasound for diagnosing skin cancer in adults,” *Cochrane Database of Systematic Reviews*, vol. 2018, no. 12, Dec. 2018, doi: 10.1002/14651858.CD013188.

[31] P. Pala, B. S. Bergler-Czop, and J. Gwiżdż, “Teledermatology: idea, benefits and risks of modern age – a systematic review based on melanoma,” *Advances in Dermatology and Allergology*, vol. 37, no. 2, pp. 159–167, 2020, doi: 10.5114/ada.2020.94834.

[32] S. Gulati and R. K. Bhogal, *Detection of malignant melanoma using deep learning*, vol. 1045. Springer Singapore, 2019. doi: 10.1007/978-981-13-9939-8_28.

[33] D. Kumar *et al.*, “Automatic Detection of White Blood Cancer from Bone Marrow Microscopic Images Using Convolutional Neural Networks,” *IEEE Access*, vol. 8, no. Mm, pp. 142521–142531, 2020, doi: 10.1109/ACCESS.2020.3012292.

[34] S. Khattar and R. Kaur, “Computer assisted diagnosis of skin cancer: A survey and future recommendations,” *Computers and Electrical Engineering*, vol. 104, p. 108431, Dec. 2022, doi: 10.1016/j.compeleceng.2022.108431.

[35] Z. Xu, F. R. Sheykhahmad, N. Ghadimi, and N. Razmjooy, “Computer-aided diagnosis of skin cancer based on soft computing techniques.,” *Open Med (Wars)*, vol. 15, no. 1, pp. 860–871, 2020, doi: 10.1515/med-2020-0131.

[36] Z. Xu, F. R. Sheykhahmad, N. Ghadimi, and N. Razmjooy, “Computer-aided diagnosis of skin cancer based on soft computing techniques,” *Open Medicine*, vol. 15, no. 1, pp. 860–871, Sep. 2020, doi: 10.1515/med-2020-0131.

[37] S. Jain, V. jagtap, and N. Pise, “Computer Aided Melanoma Skin Cancer Detection Using Image Processing,” *Procedia Comput Sci*, vol. 48, pp. 735–740, 2015, doi: 10.1016/j.procs.2015.04.209.

[38] S. Bakheet and A. Al-Hamadi, “Computer-Aided Diagnosis of Malignant Melanoma Using Gabor-Based Entropic Features and Multilevel Neural Networks,” *Diagnostics*, vol. 10, no. 10, p. 822, Oct. 2020, doi: 10.3390/diagnostics10100822.

[39] N. Hameed, A. Shabut, and M. A. Hossain, “A Computer-aided diagnosis system for classifying prominent skin lesions using machine learning,” in *2018 10th Computer Science and Electronic Engineering (CEEC)*, IEEE, Sep. 2018, pp. 186–191. doi: 10.1109/CEEC.2018.8674183.

[40] E. V. Carrera and D. Ron-Domínguez, “A Computer Aided Diagnosis System for Skin Cancer Detection,” 2019, pp. 553–563. doi: 10.1007/978-3-030-05532-5_42.

[41] V. Venugopal, N. I. Raj, M. K. Nath, and N. Stephen, “A deep neural network using modified EfficientNet for skin cancer detection in dermoscopic images,” *Decision Analytics Journal*, vol. 8, p. 100278, Sep. 2023, doi: 10.1016/j.dajour.2023.100278.

[42] H. ORHAN and E. YAVŞAN, “Artificial intelligence-assisted detection model for melanoma diagnosis using deep learning techniques,” *Mathematical Modelling and Numerical Simulation with Applications*, vol. 3, no. 2, pp. 159–169, Jun. 2023, doi: 10.53391/mmnsa.1311943.

[43] A. Boudhir Abdelhakim, B. Ahmed Mohamed, and D. Yousra, “A New Approach using Deep Learning and Reinforcement Learning in HealthCare,” *International journal of electrical*

and computer engineering systems, vol. 14, no. 5, pp. 557–564, Jun. 2023, doi: 10.32985/ijeces.14.5.7.

[44] M. Naqvi, S. Q. Gilani, T. Syed, O. Marques, and H.-C. Kim, "Skin Cancer Detection Using Deep Learning—A Review," *Diagnostics*, vol. 13, no. 11, p. 1911, May 2023, doi: 10.3390/diagnostics13111911.

[45] "Skin Cancer Detection Based on Deep Learning," *J Biomed Phys Eng*, vol. 12, no. 6, Dec. 2022, doi: 10.31661/jbpe.v0i0.2207-1517.

[46] J. Requa *et al.*, "High-fidelity detection, subtyping, and localization of five skin neoplasms using supervised and semi-supervised learning," *J Pathol Inform*, vol. 14, p. 100159, 2023, doi: 10.1016/j.jpi.2022.100159.

[47] V. Venugopal, N. I. Raj, M. K. Nath, and N. Stephen, "A deep neural network using modified EfficientNet for skin cancer detection in dermoscopic images," *Decision Analytics Journal*, vol. 8, p. 100278, Sep. 2023, doi: 10.1016/j.dajour.2023.100278.

[48] A. Almutairi and R. U. Khan, "Image-Based Classical Features and Machine Learning Analysis of Skin Cancer Instances," *Applied Sciences*, vol. 13, no. 13, p. 7712, Jun. 2023, doi: 10.3390/app13137712.

[49] M. Zafar, M. I. Sharif, M. I. Sharif, S. Kadry, S. A. C. Bukhari, and H. T. Rauf, "Skin Lesion Analysis and Cancer Detection Based on Machine/Deep Learning Techniques: A Comprehensive Survey," *Life*, vol. 13, no. 1, p. 146, Jan. 2023, doi: 10.3390/life13010146.

[50] M. Nawaz *et al.*, "Skin cancer detection from dermoscopic images using deep learning and fuzzy *k* -means clustering," *Microsc Res Tech*, vol. 85, no. 1, pp. 339–351, Jan. 2022, doi: 10.1002/jemt.23908.

[51] P. Ghosh *et al.*, "SkinNet-16: A deep learning approach to identify benign and malignant skin lesions," *Front Oncol*, vol. 12, Aug. 2022, doi: 10.3389/fonc.2022.931141.

[52] T. G. Dietterich, "Ensemble Methods in Machine Learning," 2000, pp. 1–15. doi: 10.1007/3-540-45014-9_1.

[53] A. Mohammed and R. Kora, "A comprehensive review on ensemble deep learning: Opportunities and challenges," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 2, pp. 757–774, Feb. 2023, doi: 10.1016/j.jksuci.2023.01.014.

[54] G. Renith and A. Senthilselvi, “An efficient skin cancer detection and classification using <scp>Improved Adaboost Aphid–Ant Mutualism</scp> model,” *Int J Imaging Syst Technol*, vol. 33, no. 6, pp. 1957–1972, Nov. 2023, doi: 10.1002/ima.22932.

[55] V. A. O. Nancy, P. Prabhavathy, M. S. Arya, and B. S. Ahamed, “Comparative study and analysis on skin cancer detection using machine learning and deep learning algorithms,” *Multimed Tools Appl*, vol. 82, no. 29, pp. 45913–45957, Dec. 2023, doi: 10.1007/s11042-023-16422-6.

[56] H. Basly, W. Ouarda, F. E. Sayadi, B. Ouni, and A. M. Alimi, “CNN-SVM Learning Approach Based Human Activity Recognition,” 2020, pp. 271–281. doi: 10.1007/978-3-030-51935-3_29.

[57] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *J Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[58] D. Keerthana, V. Venugopal, M. K. Nath, and M. Mishra, “Hybrid convolutional neural networks with SVM classifier for classification of skin cancer,” *Biomedical Engineering Advances*, vol. 5, p. 100069, Jun. 2023, doi: 10.1016/j.bea.2022.100069.

[59] A. Imran, A. Nasir, M. Bilal, G. Sun, A. Alzahrani, and A. Almuhaimeed, “Skin Cancer Detection Using Combined Decision of Deep Learners,” *IEEE Access*, vol. 10, pp. 118198–118212, 2022, doi: 10.1109/ACCESS.2022.3220329.

[60] N. V. Kumar, P. V. Kumar, K. Pramodh, and Y. Karuna, “Classification of Skin diseases using Image processing and SVM,” in *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*, IEEE, Mar. 2019, pp. 1–5. doi: 10.1109/ViTECoN.2019.8899449.

[61] “Applications of Support Vector Machine (SVM) Learning in Cancer Genomics,” *Cancer Genomics Proteomics*, vol. 15, no. 1, Jan. 2018, doi: 10.21873/cgp.20063.

[62] P. Thanh Noi and M. Kappas, “Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery,” *Sensors*, vol. 18, no. 2, p. 18, Dec. 2017, doi: 10.3390/s18010018.

[63] E. Y. Boateng, J. Otoo, and D. A. Abaye, “Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review,” *Journal of Data Analysis and Information Processing*, vol. 08, no. 04, pp. 341–357, 2020, doi: 10.4236/jdaip.2020.84020.

[64] Y. Zhang *et al.*, “Comparison of machine learning methods for stationary wavelet entropy-based multiple sclerosis detection: decision tree, *k* -nearest neighbors, and support vector machine,” *Simulation*, vol. 92, no. 9, pp. 861–871, Sep. 2016, doi: 10.1177/0037549716666962.

[65] M. Bansal, A. Goyal, and A. Choudhary, “A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning,” *Decision Analytics Journal*, vol. 3, p. 100071, Jun. 2022, doi: 10.1016/j.dajour.2022.100071.

[66] T. A. Assegie, “Support Vector Machine And K-Nearest Neighbor Based Liver Disease Classification Model,” *Indonesian Journal of electronics, electromedical engineering, and medical informatics*, vol. 3, no. 1, pp. 9–14, Feb. 2021, doi: 10.35882/ijeeemi.v3i1.2.

[67] F. Iglesias, T. Zseby, and A. Zimek, “Clustering refinement,” *Int J Data Sci Anal*, vol. 12, no. 4, pp. 333–353, Oct. 2021, doi: 10.1007/s41060-021-00275-z.

[68] M. Z. Rodriguez *et al.*, “Clustering algorithms: A comparative approach.,” *PLoS One*, vol. 14, no. 1, p. e0210236, 2019, doi: 10.1371/journal.pone.0210236.

[69] J. Isohanni, “Recognising small colour changes with unsupervised learning, comparison of methods,” *Advances in Computational Intelligence*, vol. 4, no. 2, p. 6, Jun. 2024, doi: 10.1007/s43674-024-00073-7.

[70] R. Wehrens, “Data Mapping: Linear Methods versus Nonlinear Techniques,” in *Comprehensive Chemometrics*, Elsevier, 2009, pp. 619–633. doi: 10.1016/B978-044452701-1.00065-X.

[71] I. H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions,” *SN Comput Sci*, vol. 2, no. 3, p. 160, May 2021, doi: 10.1007/s42979-021-00592-x.

[72] F. Ge and X. Liu, “Density Peaks Clustering Algorithm Based on a Divergence Distance and Tissue—Like P System,” *Applied Sciences*, vol. 13, no. 4, p. 2293, Feb. 2023, doi: 10.3390/app13042293.

[73] F. Ge and X. Liu, “Density Peaks Clustering Algorithm Based on a Divergence Distance and Tissue—Like P System,” *Applied Sciences*, vol. 13, no. 4, p. 2293, Feb. 2023, doi: 10.3390/app13042293.

[74] M. Chaudhry, I. Shafi, M. Mahnoor, D. L. R. Vargas, E. B. Thompson, and I. Ashraf, “A Systematic Literature Review on Identifying Patterns Using Unsupervised Clustering

Algorithms: A Data Mining Perspective," *Symmetry (Basel)*, vol. 15, no. 9, p. 1679, Aug. 2023, doi: 10.3390/sym15091679.

[75] M.-S. Yang, C.-Y. Lai, and C.-Y. Lin, "A robust EM clustering algorithm for Gaussian mixture models," *Pattern Recognit*, vol. 45, no. 11, pp. 3950–3961, Nov. 2012, doi: 10.1016/j.patcog.2012.04.031.

[76] Y. Wu, B. Chen, A. Zeng, D. Pan, R. Wang, and S. Zhao, "Skin Cancer Classification With Deep Learning: A Systematic Review," *Front Oncol*, vol. 12, Jul. 2022, doi: 10.3389/fonc.2022.893972.

[77] M. Darvishnezhad, "A multi-view ensemble model based on semi-supervised feature learning for small sample classification of PolSAR images," *Int J Remote Sens*, vol. 45, no. 3, pp. 981–1031, Feb. 2024, doi: 10.1080/01431161.2024.2305627.

[78] K. S. Shakya, A. Alavi, J. Porteous, P. K, A. Laddi, and M. Jaiswal, "A Critical Analysis of Deep Semi-Supervised Learning Approaches for Enhanced Medical Image Classification," *Information*, vol. 15, no. 5, p. 246, Apr. 2024, doi: 10.3390/info15050246.

[79] A. Kumar and R. K. Aggarwal, "An exploration of semi-supervised and language-adversarial transfer learning using hybrid acoustic model for hindi speech recognition," *J Reliab Intell Environ*, vol. 8, no. 2, pp. 117–132, Jun. 2022, doi: 10.1007/s40860-021-00140-7.

[80] H. Bi, M. Perello-Nieto, R. Santos-Rodriguez, P. Flach, and I. Craddock, "An active semi-supervised deep learning model for human activity recognition," *J Ambient Intell Humaniz Comput*, vol. 14, no. 10, pp. 13049–13065, Oct. 2023, doi: 10.1007/s12652-022-03768-2.

[81] S. Yang, "A review of research and development of semi-supervised learning strategies for medical image processing," *EAI Endorsed Transactions on e-Learning*, vol. 9, Jan. 2024, doi: 10.4108/eetel.4822.

[82] S. V. Mahadevkar *et al.*, "A Review on Machine Learning Styles in Computer Vision—Techniques and Future Directions," *IEEE Access*, vol. 10, pp. 107293–107329, 2022, doi: 10.1109/ACCESS.2022.3209825.

[83] Z. Solatidehkordi and I. Zualkernan, "Survey on Recent Trends in Medical Image Classification Using Semi-Supervised Learning," *Applied Sciences*, vol. 12, no. 23, p. 12094, Nov. 2022, doi: 10.3390/app122312094.

[84] L. Qu, S. Liu, X. Liu, M. Wang, and Z. Song, “Towards label-efficient automatic diagnosis and analysis: a comprehensive survey of advanced deep learning-based weakly-supervised, semi-supervised and self-supervised techniques in histopathological image analysis,” *Phys Med Biol*, vol. 67, no. 20, p. 20TR01, Oct. 2022, doi: 10.1088/1361-6560/ac910a.

[85] Y. Li, Y. Lv, S. Wang, J. Liang, J. Li, and X. Li, “Cooperative Hybrid Semi-Supervised Learning for Text Sentiment Classification,” *Symmetry (Basel)*, vol. 11, no. 2, p. 133, Jan. 2019, doi: 10.3390/sym11020133.

[86] J. E. van Engelen and H. H. Hoos, “A survey on semi-supervised learning,” *Mach Learn*, vol. 109, no. 2, pp. 373–440, Feb. 2020, doi: 10.1007/s10994-019-05855-6.

[87] A. Safonova, G. Ghazaryan, S. Stiller, M. Main-Knorn, C. Nendel, and M. Ryo, “Ten deep learning techniques to address small data problems with remote sensing,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 125, p. 103569, Dec. 2023, doi: 10.1016/j.jag.2023.103569.

[88] J. M. Duarte and L. Berton, “A review of semi-supervised learning for text classification,” *Artif Intell Rev*, vol. 56, no. 9, pp. 9401–9469, Sep. 2023, doi: 10.1007/s10462-023-10393-8.

[89] Z. Zhao, L. Alzubaidi, J. Zhang, Y. Duan, and Y. Gu, “A comparison review of transfer learning and self-supervised learning: Definitions, applications, advantages and limitations,” *Expert Syst Appl*, vol. 242, p. 122807, May 2024, doi: 10.1016/j.eswa.2023.122807.

[90] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-December, pp. 770–778, Dec. 2016, doi: 10.1109/CVPR.2016.90.

[91] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-December, pp. 2818–2826, Dec. 2016, doi: 10.1109/CVPR.2016.308.

[92] F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng, and A. Bovik, “Melanoma Classification on Dermoscopy Images Using a Neural Network Ensemble Model,” *IEEE Trans Med Imaging*, vol. 36, no. 3, pp. 849–858, Mar. 2017, doi: 10.1109/TMI.2016.2633551.

[93] D. Zhang, F. He, S. Han, L. Zou, Y. Wu, and Y. Chen, “An efficient approach to directly compute the exact Hausdorff distance for 3D point sets,” *Integr Comput-Aided Eng*, vol. 24, no. 3, pp. 261–277, 2017, doi: 10.3233/ica-170544.

[94] H. Liang, X. Sun, Y. Sun, and Y. Gao, “Text feature extraction based on deep learning: a review,” *EURASIP J Wirel Commun Netw*, vol. 2017, no. 1, pp. 1–12, Dec. 2017, doi: 10.1186/s13638-017-0993-1.

[95] E. Jana, R. Subban, and S. Saraswathi, “Research on Skin Cancer Cell Detection Using Image Processing,” *2017 IEEE International Conference on Computational Intelligence and Computing Research, ICCIC 2017*, Nov. 2018, doi: 10.1109/ICCIC.2017.8524554.

[96] M. A. Farooq, A. Khatoon, V. Varkarakis, and P. Corcoran, “Advanced deep learning methodologies for skin cancer classification in prodromal stages,” *CEUR Workshop Proc*, vol. 2563, pp. 40–51, 2019.

[97] T. J. Brinker *et al.*, “Deep neural networks are superior to dermatologists in melanoma image classification,” *Eur J Cancer*, vol. 119, pp. 11–17, Sep. 2019, doi: 10.1016/j.ejca.2019.05.023.

[98] J. Zhang, Y. Xie, Y. Xia, and C. Shen, “Attention residual learning for skin lesion classification,” *IEEE Trans Med Imaging*, vol. 38, no. 9, pp. 2092–2103, Sep. 2019, doi: 10.1109/tmi.2019.2893944.

[99] C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *J Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[100] J. Wu, W. Hu, Y. Wen, W. Tu, and X. Liu, “Skin Lesion Classification Using Densely Connected Convolutional Networks with Attention Residual Learning,” *Sensors*, vol. 20, no. 24, p. 7080, Dec. 2020, doi: 10.3390/s20247080.

[101] H. Basly, W. Ouarda, F. E. Sayadi, B. Ouni, and A. M. Alimi, “CNN-svm learning approach based human activity recognition,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12119 LNCS, pp. 271–281, 2020, doi: 10.1007/978-3-030-51935-3_29.

[102] S. S. Chaturvedi, J. V. Tembhurne, and T. Diwan, “A multi-class skin Cancer classification using deep convolutional neural networks,” *Multimed Tools Appl*, vol. 79, no. 39, pp. 28477–28498, Oct. 2020, doi: 10.1007/s11042-020-09388-2.

[103] S. Labani, S. Asthana, K. Rathore, and K. Sardana, “Incidence of melanoma and nonmelanoma skin cancers in Indian and the global regions,” *J Cancer Res Ther*, vol. 17, no. 4, pp. 906–911, Jul. 2021, doi: 10.4103/JCRT.JCRT_785_19.

[104] Z. Wu, S. Li, C. Chen, A. Hao, and H. Qin, “Deeper Look at Image Salient Object Detection: Bi-Stream Network with a Small Training Dataset,” *IEEE Trans Multimedia*, vol. 24, pp. 73–86, 2022, doi: 10.1109/TMM.2020.3046871.

[105] Z. Wu, S. Li, C. Chen, A. Hao, and H. Qin, “Recursive multi-model complementary deep fusion for robust salient object detection via parallel sub-networks,” *Pattern Recogn*, vol. 121, p. 108212, Jan. 2022, doi: 10.1016/j.patcog.2021.108212.

[106] S. S. Sundari, Dr. S. AK, and Dr. M. Islabudeen, “Skin Lesions Detection using Deep Learning Techniques,” *Int J Res Appl Sci Eng Technol*, vol. 11, no. 5, pp. 2546–2548, May 2023, doi: 10.22214/ijraset.2023.52129.

[107] A. Boudhir Abdelhakim, B. Ahmed Mohamed, and D. Yousra, “A New Approach using Deep Learning and Reinforcement Learning in HealthCare,” *International journal of electrical and computer engineering systems*, vol. 14, no. 5, pp. 557–564, Jun. 2023, doi: 10.32985/ijeces.14.5.7.

[108] J. V. Tembhurne, N. Hebbar, H. Y. Patil, and T. Diwan, “Skin cancer detection using ensemble of machine learning and deep learning techniques,” *Multimed Tools Appl*, Jul. 2023, doi: 10.1007/s11042-023-14697-3.

[109] J. Avanija, C. Chandra Mohan Reddy, C. Sri Chandan Reddy, D. Harshavardhan Reddy, T. Narasimhulu, and N. V. Hardhik, “Skin Cancer Detection using Ensemble Learning,” in *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, IEEE, Jun. 2023, pp. 184–189. doi: 10.1109/ICSCSS57650.2023.10169747.

[110] M. Sharafudeen and V. C. S. S., “Detecting skin lesions fusing handcrafted features in image network ensembles,” *Multimed Tools Appl*, vol. 82, no. 2, pp. 3155–3175, Jan. 2023, doi: 10.1007/s11042-022-13046-0.

[111] P. Nadiger, “Skin Cancer Detection and Classification Using Deep Learning,” *Int J Res Appl Sci Eng Technol*, vol. 11, no. 5, pp. 7182–7185, May 2023, doi: 10.22214/ijraset.2023.53390.

[112] J. Avanija, C. Chandra Mohan Reddy, C. Sri Chandan Reddy, D. Harshavardhan Reddy, T. Narasimhulu, and N. V. Hardhik, “Skin Cancer Detection using Ensemble Learning,” in *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, IEEE, Jun. 2023, pp. 184–189. doi: 10.1109/ICSCSS57650.2023.10169747.

[113] Prof. Balaji Chaugule, Kiran Bomble, Shruti Jundare, Nilesh Maske, and Vikas Gagare, "Skin Melanoma Cancer Detection and Classification using Machine Learning," *Int J Sci Res Sci Technol*, pp. 519–524, May 2023, doi: 10.32628/IJSRST523103110.

[114] A. Yaqoob, R. Musheer Aziz, and N. K. verma, "Applications and Techniques of Machine Learning in Cancer Classification: A Systematic Review," *Human-Centric Intelligent Systems*, vol. 3, no. 4, pp. 588–615, Sep. 2023, doi: 10.1007/s44230-023-00041-3.

[115] Prof. R. Jadhav, R. Kale, S. Kadam, T. Gund, and S. Jadhav, "Skin Cancer Detection using Image Processing," *Int J Res Appl Sci Eng Technol*, vol. 11, no. 5, pp. 6781–6785, May 2023, doi: 10.22214/ijraset.2023.53282.

[116] N. Kavitha and M. Vayelapelli, "A Study on Pre-processing Techniques for Automated Skin Cancer Detection," 2020, pp. 145–153. doi: 10.1007/978-981-15-2407-3_19.

[117] R. M. Reshma M and S. H. Ataulla, "Preprocessing of Skin Cancer Images using Non-Local Means (Nlm) Filter," *Biomedical and Pharmacology Journal*, vol. 15, no. 4, pp. 2203–2208, Dec. 2022, doi: 10.13005/bpj/2557.

[118] S. Khattar, R. Kaur, and G. Gupta, "A Review on Preprocessing, Segmentation and Classification Techniques for Detection of Skin Cancer," in *2023 2nd Edition of IEEE Delhi Section Flagship Conference (DELCON)*, IEEE, Feb. 2023, pp. 1–6. doi: 10.1109/DELCON57910.2023.10127546.

[119] B. Sreedhar, M. Swamy B.E, and M. Sunil Kumar, "A Comparative Study of Melanoma Skin Cancer Detection in Traditional and Current Image Processing Techniques," in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, IEEE, Oct. 2020, pp. 654–658. doi: 10.1109/I-SMAC49090.2020.9243501.

[120] M. Vidya and M. V. Karki, "Skin Cancer Detection using Machine Learning Techniques," in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, IEEE, Jul. 2020, pp. 1–5. doi: 10.1109/CONECCT50063.2020.9198489.

[121] A. N. Hoshyar, A. Al-Jumaily, and A. N. Hoshyar, "The Beneficial Techniques in Preprocessing Step of Skin Cancer Detection System Comparing," *Procedia Comput Sci*, vol. 42, pp. 25–31, 2014, doi: 10.1016/j.procs.2014.11.029.

[122] T. Lee, V. Ng, R. Gallagher, A. Coldman, and D. McLean, “Dullrazor®: A software approach to hair removal from images,” *Comput Biol Med*, vol. 27, no. 6, pp. 533–543, 1997, doi: 10.1016/S0010-4825(97)00020-6.

[123] H. Zhou, G. Schaefer, M. E. Celebi, F. Lin, and T. Liu, “Gradient vector flow with mean shift for skin lesion segmentation,” *Computerized Medical Imaging and Graphics*, vol. 35, no. 2, pp. 121–127, 2011, doi: 10.1016/j.compmedimag.2010.08.002.

[124] C. Barata, J. S. Marques, and J. Rozeira, “A system for the detection of pigment network in dermoscopy images using directional filters,” *IEEE Trans Biomed Eng*, vol. 59, no. 10, pp. 2744–2754, 2012, doi: 10.1109/TBME.2012.2209423.

[125] Q. Abbas, I. F. Garcia, M. Emre Celebi, W. Ahmad, and Q. Mushtaq, “A perceptually oriented method for contrast enhancement and segmentation of dermoscopy images,” *Skin Research and Technology*, vol. 19, no. 1, pp. 1–8, 2013, doi: 10.1111/j.1600-0846.2012.00670.x.

[126] A. Huang, S.-Y. Kwan, W.-Y. Chang, M.-Y. Liu, M.-H. Chi, and G.-S. Chen, *A Robust Hair Segmentation and Removal Approach for Clinical Images of Skin Lesions*. 2013. doi: 10.0/Linux-x86_64.

[127] R. B. Aswin, J. A. Jaleel, and S. Salim, “Hybrid genetic algorithm - Artificial neural network classifier for skin cancer detection,” *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies, ICCICCT 2014*, pp. 1304–1309, 2014, doi: 10.1109/ICCICCT.2014.6993162.

[128] H. Lee and Y. P. P. Chen, “Skin cancer extraction with optimum fuzzy thresholding technique,” *Applied Intelligence*, vol. 40, no. 3, pp. 415–426, 2014, doi: 10.1007/s10489-013-0474-0.

[129] M. Youssef, K. Thierry, and K. Zoubir, “Cross Correlation Method for Images Alignment: Application to 4 Buckets Calculation in Thermoreflectance,” in *2020 36th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM)*, IEEE, Mar. 2020, pp. 137–142. doi: 10.23919/SEMI-THERM50369.2020.9142843.

[130] K. Eltayef, Y. Li, B. I. Dodo, and X. Liu, “Skin Cancer Detection in Dermoscopy Images Using Sub-Region Features,” 2017, pp. 75–86. doi: 10.1007/978-3-319-68765-0_7.

[131] M. Waghulde, S. Kulkarni, and G. Phadke, “Detection of Skin Cancer Lesions from Digital Images with Image Processing Techniques,” in *2019 IEEE Pune Section International Conference (PuneCon)*, IEEE, Dec. 2019, pp. 1–6. doi: 10.1109/PuneCon46936.2019.9105886.

[132] N. Kavitha and M. Vayelapelli, “A Study on Pre-processing Techniques for Automated Skin Cancer Detection,” 2020, pp. 145–153. doi: 10.1007/978-981-15-2407-3_19.

[133] M. Hajiarbabi, “Skin cancer detection using multi-scale deep learning and transfer learning,” *J Med Artif Intell*, vol. 6, pp. 23–23, Nov. 2023, doi: 10.21037/jmai-23-67.

[134] M. C. P, J. P. Jo, and Vinitha. V, “Computer-Aided Detection of Skin Cancer Detection from Lesion Images Via Deep Learning Techniques: A Review,” *Int J Res Appl Sci Eng Technol*, vol. 11, no. 4, pp. 4455–4461, Apr. 2023, doi: 10.22214/ijraset.2023.51271.

[135] S. B. Mukadam and H. Y. Patil, “Skin Cancer Classification Framework Using Enhanced Super Resolution Generative Adversarial Network and Custom Convolutional Neural Network,” *Applied Sciences*, vol. 13, no. 2, p. 1210, Jan. 2023, doi: 10.3390/app13021210.

[136] H. L. Gururaj, N. Manju, A. Nagarjun, V. N. M. Aradhya, and F. Flammini, “DeepSkin: A Deep Learning Approach for Skin Cancer Classification,” *IEEE Access*, vol. 11, pp. 50205–50214, 2023, doi: 10.1109/ACCESS.2023.3274848.

[137] Y. Aydin, “A Comparative Analysis of Skin Cancer Detection Applications Using Histogram-Based Local Descriptors,” *Diagnostics*, vol. 13, no. 19, p. 3142, Oct. 2023, doi: 10.3390/diagnostics13193142.

[138] B. Shetty, R. Fernandes, A. P. Rodrigues, R. Chengoden, S. Bhattacharya, and K. Lakshmanna, “Skin lesion classification of dermoscopic images using machine learning and convolutional neural network,” *Sci Rep*, vol. 12, no. 1, p. 18134, Oct. 2022, doi: 10.1038/s41598-022-22644-9.

[139] H. Liang, X. Sun, Y. Sun, and Y. Gao, “Text feature extraction based on deep learning: a review,” *EURASIP J Wirel Commun Netw*, vol. 2017, no. 1, p. 211, Dec. 2017, doi: 10.1186/s13638-017-0993-1.

[140] S. Nofallah *et al.*, “Segmenting Skin Biopsy Images with Coarse and Sparse Annotations using U-Net,” *J Digit Imaging*, vol. 35, no. 5, pp. 1238–1249, Oct. 2022, doi: 10.1007/s10278-022-00641-8.

[141] D. A. Okuboyejo and O. O. Olugbara, “A Review of Prevalent Methods for Automatic Skin Lesion Diagnosis,” *Open Dermatol J*, vol. 12, no. 1, pp. 14–53, Apr. 2018, doi: 10.2174/1874372201812010014.

[142] A. Victor, B. S. Gandhi, M. R. Ghalib, and A. J. M, "A Review on Skin Cancer Detection and Classification using Infrared images," *International Journal of Engineering Trends and Technology*, vol. 70, no. 4, pp. 403–417, Apr. 2022, doi: 10.14445/22315381/IJETT-V70I4P235.

[143] G. Nasreen, K. Haneef, M. Tamoor, and A. Irshad, "Review: a comparative study of state-of-the-art skin image segmentation techniques with CNN," *Multimed Tools Appl*, vol. 82, no. 7, pp. 10921–10942, Mar. 2023, doi: 10.1007/s11042-022-13756-5.

[144] J. Premaladha and K. S. Ravichandran, "Novel Approaches for Diagnosing Melanoma Skin Lesions Through Supervised and Deep Learning Algorithms," *J Med Syst*, vol. 40, no. 4, pp. 1–12, 2016, doi: 10.1007/s10916-016-0460-2.

[145] D. M. Dhane, V. Krishna, A. Achar, and C. Bar, "Spectral Clustering for Unsupervised Segmentation of Lower Extremity Wound Beds Using Optical Images," *J Med Syst*, 2016, doi: 10.1007/s10916-016-0554-x.

[146] M. Zortea, E. Flores, and J. Scharcanski, "A simple weighted thresholding method for the segmentation of pigmented skin lesions in macroscopic images," *Pattern Recognit*, vol. 64, pp. 92–104, 2017, doi: 10.1016/j.patcog.2016.10.031.

[147] F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng, and A. Bovik, "Melanoma classification on dermoscopy images using a neural network ensemble model," *IEEE Trans Med Imaging*, vol. 36, no. 3, pp. 849–858, 2017, doi: 10.1109/TMI.2016.2633551.

[148] T. Y. Satheesha, D. Satyanarayana, M. N. G. Prasad, and K. D. Dhruve, "Melanoma Is Skin Deep: A 3D Reconstruction Technique for Computerized Dermoscopic Skin Lesion Classification," *IEEE J Transl Eng Health Med*, vol. 5, no. c, 2017, doi: 10.1109/JTEHM.2017.2648797.

[149] S. Choudhari and S. Biday, "Artificial Neural Network for SkinCancer Detection," *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, vol. 3, no. 5, pp. 147–153, 2014, [Online]. Available: www.ijettcs.org

[150] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017, doi: 10.1038/nature21056.

[151] S. Bhattacharyya *et al.*, *Advanced Computational and Communication Paradigms*, vol. 475. Springer Singapore, 2018. doi: 10.1007/978-981-10-8240-5.

[152] M. A. Al-masni, M. A. Al-antari, M. T. Choi, S. M. Han, and T. S. Kim, “Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks,” *Comput Methods Programs Biomed*, vol. 162, pp. 221–231, 2018, doi: 10.1016/j.cmpb.2018.05.027.

[153] M. U. Rehman, S. H. Khan, S. M. Danish Rizvi, Z. Abbas, and A. Zafar, “Classification of Skin Lesion by Interference of Segmentation and Convolotion Neural Network,” *2018 2nd International Conference on Engineering Innovation, ICEI 2018*, pp. 81–85, 2018, doi: 10.1109/ICEI18.2018.8448814.

[154] S. Dubey and K. Shah, “a Nalysis of V Arious F Laws D Etection Using,” *2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE)*, vol. 3, no. 2, pp. 765–774, 2012.

[155] S. Pathan, K. G. Prabhu, and P. C. Siddalingaswamy, “Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—A review,” *Biomed Signal Process Control*, vol. 39, pp. 237–262, 2018, doi: 10.1016/j.bspc.2017.07.010.

[156] M. Aljanabi, Y. E. Özok, J. Rahebi, and A. S. Abdullah, “Skin lesion segmentation method for dermoscopy images using artificial bee colony algorithm,” *Symmetry (Basel)*, vol. 10, no. 8, 2018, doi: 10.3390/sym10080347.

[157] S. S. Chouhan, A. Kaul, and U. P. Singh, *Soft computing approaches for image segmentation: a survey*, vol. 77, no. 21. Multimedia Tools and Applications, 2018. doi: 10.1007/s11042-018-6005-6.

[158] M. Krishna Monika, N. Arun Vignesh, C. Usha Kumari, M. N. V. S. S. Kumar, and E. Laxmi Lydia, “Skin cancer detection and classification using machine learning,” *Mater Today Proc*, vol. 33, no. xxxx, pp. 4266–4270, 2020, doi: 10.1016/j.matpr.2020.07.366.

[159] Ç. Kaymak and A. Uçar, *A brief survey and an application of semantic image segmentation for autonomous driving*, vol. 136. 2019. doi: 10.1007/978-3-030-11479-4_9.

[160] R. Mohakud and R. Dash, “Skin cancer image segmentation utilizing a novel EN-GWO based hyper-parameter optimized FCEDN,” *Journal of King Saud University - Computer and Information Sciences*, no. xxxx, 2022, doi: 10.1016/j.jksuci.2021.12.018.

[161] A. Murugan, S. A. H. Nair, A. A. P. Preethi, and K. P. S. Kumar, “Diagnosis of skin cancer using machine learning techniques,” *Microprocess Microsyst*, vol. 81, no. October 2020, p. 103727, 2021, doi: 10.1016/j.micpro.2020.103727.

[162] I. Giotis, N. Molders, S. Land, M. Biehl, M. F. Jonkman, and N. Petkov, “MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images,” *Expert Syst Appl*, vol. 42, no. 19, pp. 6578–6585, 2015, doi: 10.1016/j.eswa.2015.04.034.

[163] Y. Yuan, M. Chao, and Y. C. Lo, “Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks with Jaccard Distance,” *IEEE Trans Med Imaging*, vol. 36, no. 9, pp. 1876–1886, 2017, doi: 10.1109/TMI.2017.2695227.

[164] N. Singh and S. K. Gupta, “Recent advancement in the early detection of melanoma using computerized tools: An image analysis perspective,” *Skin Research and Technology*, vol. 25, no. 2, pp. 129–141, 2019, doi: 10.1111/srt.12622.

[165] R. D. Seeja and A. Suresh, “Deep learning based skin lesion segmentation and classification of melanoma using support vector machine (SVM),” *Asian Pacific Journal of Cancer Prevention*, vol. 20, no. 5, pp. 1555–1561, 2019, doi: 10.31557/APJCP.2019.20.5.1555.

[166] K. Zafar *et al.*, “Skin lesion segmentation from dermoscopic images using convolutional neural network,” *Sensors (Switzerland)*, vol. 20, no. 6, pp. 1–14, 2020, doi: 10.3390/s20061601.

[167] R. Vani, J. C. Kavitha, and D. Subitha, “Novel approach for melanoma detection through iterative deep vector network,” *J Ambient Intell Humaniz Comput*, no. 2018, 2021, doi: 10.1007/s12652-021-03242-5.

[168] H. Ashraf, A. Waris, M. F. Ghafoor, S. O. Gilani, and I. K. Niazi, “Melanoma segmentation using deep learning with test-time augmentations and conditional random fields,” *Sci Rep*, vol. 12, no. 1, pp. 1–16, 2022, doi: 10.1038/s41598-022-07885-y.

[169] M. Tahir, A. Naeem, H. Malik, J. Tanveer, R. A. Naqvi, and S. W. Lee, “DSCC_Net: Multi-Classification Deep Learning Models for Diagnosing of Skin Cancer Using Dermoscopic Images,” *Cancers (Basel)*, vol. 15, no. 7, Apr. 2023, doi: 10.3390/cancers15072179.

[170] S. Chatterjee, D. Dey, and S. Munshi, “Integration of morphological preprocessing and fractal based feature extraction with recursive feature elimination for skin lesion types classification,” *Comput Methods Programs Biomed*, vol. 178, pp. 201–218, Sep. 2019, doi: 10.1016/j.cmpb.2019.06.018.

[171] A. S. K. W. H. Pehamberger, “In vivo epiluminescence microscopy of pigmented skin lesions. I. Pattern analysis of pigmented skin lesions, J. Am.Acad. Dermatol. 17 (4) (2016) 571–583.”.

[172] M. E. C. J. S. Marques. Catarina Barata, "' A Survey of Feature Extraction in Dermoscopy Image Analysis of Skin Cancer', IEEE Journal of Biomedical and Health Informatics, 2019.".

[173] W. S. T. M. A. B. C. T. V. M. L. et al. , F. Nachbar, "The ABCD rule of dermatoscopy: high prospective value in the diagnosis ofdoubtful melanocytic skin lesions, J. Am. Acad. Dermatol. 30 (4) (1994)551–559.".

[174] F. M. Walter *et al.*, "Using the 7-point checklist as a diagnostic aid for pigmented skin lesions in general practice: a diagnostic validation study," *British Journal of General Practice*, vol. 63, no. 610, pp. e345–e353, May 2013, doi: 10.3399/bjgp13X667213.

[175] S. W. Menzies, K. A. Crotty, and W. H. McCarthy, "The morphologic criteria of the pseudopod in surface microscopy.," *Arch Dermatol*, vol. 131, no. 4, pp. 436–40, Apr. 1995.

[176] J. S. . (2007) Henning, "The CASH (colour, architecture, symmetry, and homogeneity) algorithm for dermoscopy".

[177] J. S. Henning *et al.*, "CASH Algorithm for Dermoscopy Revisited," *Arch Dermatol*, vol. 144, no. 4, Apr. 2008, doi: 10.1001/archderm.144.4.554.

[178] A. Mahbod, G. Schaefer, C. Wang, G. Dorffner, R. Ecker, and I. Ellinger, "Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification," *Comput Methods Programs Biomed*, vol. 193, p. 105475, 2020, doi: 10.1016/j.cmpb.2020.105475.

[179] R. Kasmi, K. Mokrani, R. K. Rader, J. G. Cole, and W. V. Stoecker, "Biologically inspired skin lesion segmentation using a geodesic active contour technique," *Skin Research and Technology*, vol. 22, no. 2, pp. 208–222, 2016, doi: 10.1111/srt.12252.

[180] R. B. Oliveira, M. E. Filho, Z. Ma, J. P. Papa, A. S. Pereira, and J. M. R. S. Tavares, "Computational methods for the image segmentation of pigmented skin lesions: A review," *Comput Methods Programs Biomed*, vol. 131, pp. 127–141, 2016, doi: 10.1016/j.cmpb.2016.03.032.

[181] I. Zaqout, "Diagnosis of Skin Lesions Based on Dermoscopic Images Using Image Processing Techniques," *Pattern Recognition - Selected Methods and Applications*, pp. 1–18, 2019, doi: 10.5772/intechopen.88065.

[182] H. Liu, J. Zhang, and C. Mou, "Image Denoising Based on Correlation Adaptive Sparse Modeling," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and*

Signal Processing (ICASSP), IEEE, Jun. 2021, pp. 2060–2064. doi: 10.1109/ICASSP39728.2021.9414213.

[183] D. Sengupta, P. Gupta, and A. Biswas, "A survey on mutual information based medical image registration algorithms," *Neurocomputing*, vol. 486, pp. 174–188, May 2022, doi: 10.1016/j.neucom.2021.11.023.

[184] S. Yamaki, S. Seki, N. Sugita, and M. Yoshizawa, "Performance Evaluation of Cross Correlation Functions Based on Correlation Filters," in *2021 20th International Symposium on Communications and Information Technologies (ISCIT)*, IEEE, Oct. 2021, pp. 145–149. doi: 10.1109/ISCIT52804.2021.9590596.

[185] N. S. Vadivu, G. Gupta, Q. N. Naveed, T. Rasheed, S. K. Singh, and D. Dhabliya, "Correlation-Based Mutual Information Model for Analysis of Lung Cancer CT Image," *Biomed Res Int*, vol. 2022, pp. 1–12, Aug. 2022, doi: 10.1155/2022/6451770.

[186] I. N. Dimou, G. C. Manikis, and M. E. Zervakis, "Classifier Fusion Approaches for Diagnostic Cancer Models," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Aug. 2006, pp. 5334–5337. doi: 10.1109/IEMBS.2006.260778.

[187] B. C. Q. Truong, H. D. Tuan, A. J. Fitzgerald, V. P. Wallace, and H. T. Nguyen, "High correlation of double Debye model parameters in skin cancer detection," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Aug. 2014, pp. 718–721. doi: 10.1109/EMBC.2014.6943691.

[188] K. R. Kavitha, G. S. Rajendran, and J. Varsha, "A correlation based SVM-recursive multiple feature elimination classifier for breast cancer disease using microarray," in *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, Sep. 2016, pp. 2677–2683. doi: 10.1109/ICACCI.2016.7732464.

[189] D. Chutia, D. K. Bhattacharyya, J. Sarma, and P. N. L. Raju, "An effective ensemble classification framework using random forests and a correlation based feature selection technique," *Transactions in GIS*, vol. 21, no. 6, pp. 1165–1178, Dec. 2017, doi: 10.1111/tgis.12268.

[190] M. Rajesh Babu and K. Thenmozhi, "Classification of skin disease using ensemble-based classifier," *Int J Biomed Eng Technol*, vol. 28, no. 4, p. 377, 2018, doi: 10.1504/IJBET.2018.10017204.

[191] J. O. Sinayobye, K. S. Kaawaase, F. N. Kiwanuka, and R. Musabe, “Hybrid Model of Correlation Based Filter Feature Selection and Machine Learning Classifiers Applied on Smart Meter Data Set,” in *2019 IEEE/ACM Symposium on Software Engineering in Africa (SEiA)*, IEEE, May 2019, pp. 1–10. doi: 10.1109/SEiA.2019.00009.

[192] A. Murugan, S. A. H. Nair, and K. P. S. Kumar, “Detection of Skin Cancer Using SVM, Random Forest and kNN Classifiers,” *J Med Syst*, vol. 43, no. 8, 2019, doi: 10.1007/s10916-019-1400-8.

[193] K. S. Rajasekhar and T. R. Babu, “Analysis and classification of dermoscopic images using spectral graph wavelet transform,” *Periodica polytechnica Electrical engineering and computer science*, vol. 64, no. 3, pp. 313–323, Jul. 2020, doi: 10.3311/PPee.15048.

[194] M. Abbas and H. Ghous, “Early Detection of Breast Cancer Tumors using Linear Discriminant Analysis Feature Selection with Different Machine Learning Classification Methods,” *Computer Science & Engineering: An International Journal*, vol. 12, no. 1, pp. 171–186, Feb. 2022, doi: 10.5121/cseij.2022.12117.

[195] A. Abdelhafeez, H. K. Mohamed, A. Maher, and N. A. Khalil, “A novel approach toward skin cancer classification through fused deep features and neutrosophic environment,” *Front Public Health*, vol. 11, Apr. 2023, doi: 10.3389/fpubh.2023.1123581.

[196] V. Singhal and P. Singh, “Correlation based Feature Selection for Diagnosis of Acute Lymphoblastic Leukemia,” in *Proceedings of the Third International Symposium on Women in Computing and Informatics*, New York, NY, USA: ACM, Aug. 2015, pp. 5–9. doi: 10.1145/2791405.2791423.

[197] M. Maleki and L. Rueda, “Classification via correlation-based feature grouping,” in *2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, IEEE, Aug. 2015, pp. 1–6. doi: 10.1109/CIBCB.2015.7300293.

[198] J. Jiang, D. Wu, and Z. Jiang, “A correlation-based bag of visual words for image classification,” in *2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC)*, IEEE, Oct. 2017, pp. 891–894. doi: 10.1109/ITOEC.2017.8122482.

[199] Z. S. Y. J. S. Y. and C. J. Z. Chuanlei, “‘‘Apple leaf disease identification using genetic algorithm and correlation based feature selection method,’” *Int. J. Agricult. Biol. Eng., vol. 10, no. 2, pp. 74–83, 2017*.

[200] T. Arora and R. Dhir, “Correlation-based feature selection and classification via regression of segmented chromosomes using geometric features,” *Med Biol Eng Comput*, vol. 55, no. 5, pp. 733–745, May 2017, doi: 10.1007/s11517-016-1553-2.

[201] A.-A. Nahid, M. A. Mehrabi, and Y. Kong, “Histopathological Breast Cancer Image Classification by Deep Neural Network Techniques Guided by Local Clustering,” *Biomed Res Int*, vol. 2018, pp. 1–20, 2018, doi: 10.1155/2018/2362108.

[202] C. Liu, R. Zhao, and M. Pang, “A fully automatic segmentation algorithm for CT lung images based on random forest,” *Med Phys*, vol. 47, no. 2, pp. 518–529, Feb. 2020, doi: 10.1002/mp.13939.

[203] I. M. Nasir *et al.*, “Pearson Correlation-Based Feature Selection for Document Classification Using Balanced Training,” *Sensors*, vol. 20, no. 23, p. 6793, Nov. 2020, doi: 10.3390/s20236793.

[204] M. A. Khan, T. Akram, M. Sharif, M. Alhaisoni, T. Saba, and N. Nawaz, “A probabilistic segmentation and entropy-rank correlation-based feature selection approach for the recognition of fruit diseases,” *EURASIP J Image Video Process*, vol. 2021, no. 1, p. 14, Dec. 2021, doi: 10.1186/s13640-021-00558-2.

[205] M. Ruela, C. Barata, J. S. Marques, and J. Rozeira, “A system for the detection of melanomas in dermoscopy images using shape and symmetry features,” *Comput Methods Biomech Biomed Eng Imaging Vis*, vol. 5, no. 2, pp. 127–137, 2017, doi: 10.1080/21681163.2015.1029080.

[206] C. Barata, M. Ruela, M. Francisco, T. Mendonca, and J. S. Marques, “Two systems for the detection of melanomas in dermoscopy images using texture and color features,” *IEEE Syst J*, vol. 8, no. 3, pp. 965–979, 2014, doi: 10.1109/JSYST.2013.2271540.

[207] M. K. Hasan, M. T. E. Elahi, M. A. Alam, M. T. Jawad, and R. Martí, “DermoExpert: Skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation,” *Inform Med Unlocked*, vol. 28, p. 100819, 2022, doi: 10.1016/j.imu.2021.100819.

[208] N. M. Zaitoun and M. J. Aqel, “Survey on Image Segmentation Techniques,” *Procedia Comput Sci*, vol. 65, no. Iccmit, pp. 797–806, 2015, doi: 10.1016/j.procs.2015.09.027.

[209] P. Dubai, S. Bhatt, C. Joglekar, and S. Patii, “Skin cancer detection and classification,” *Proceedings of the 2017 6th International Conference on Electrical Engineering and*

Informatics: Sustainable Society Through Digital Innovation, ICEEI 2017, vol. 2017-Novem, pp. 1–6, 2018, doi: 10.1109/ICEEI.2017.8312419.

[210] H. Bhatt, V. Shah, K. Shah, R. Shah, and M. Shah, "State-of-the-art machine learning techniques for melanoma skin cancer detection and classification: a comprehensive review," *Intelligent Medicine*, 2022, doi: 10.1016/j.imed.2022.08.004.

[211] D. Wen *et al.*, "Characteristics of publicly available skin cancer image datasets: a systematic review," *Lancet Digit Health*, vol. 4, no. 1, pp. e64–e74, Jan. 2022, doi: 10.1016/S2589-7500(21)00252-1.

[212] "isic." [Online]. Available: https://www.isic-archive.com/

[213] B. Cassidy, C. Kendrick, A. Brodzicki, J. Jaworek-Korjakowska, and M. H. Yap, "Analysis of the ISIC image datasets: Usage, benchmarks and recommendations," *Med Image Anal*, vol. 75, p. 102305, Jan. 2022, doi: 10.1016/j.media.2021.102305.

[214] T. Saba, M. A. Khan, A. Rehman, and S. L. Marie-Sainte, "Region Extraction and Classification of Skin Cancer: A Heterogeneous framework of Deep CNN Features Fusion and Reduction," *J Med Syst*, vol. 43, no. 9, p. 289, Sep. 2019, doi: 10.1007/s10916-019-1413-3.

[215] M. K. Hasan, M. A. Ahamad, C. H. Yap, and G. Yang, "A survey, review, and future trends of skin lesion segmentation and classification," *Computers in Biology and Medicine*, vol. 155. Elsevier Ltd, Mar. 01, 2023. doi: 10.1016/j.compbiomed.2023.106624.

[216] S. L. Lee and C. C. Tseng, "Image enhancement using DCT-based matrix homomorphic filtering method," *2016 IEEE Asia Pacific Conference on Circuits and Systems, APCCAS 2016*, pp. 1–4, 2017, doi: 10.1109/APCCAS.2016.7803880.

[217] T. F. Sanam and H. Imtiaz, "A DCT-based noisy speech enhancement method using teager energy operator," *Proceedings of the 2013 5th International Conference on Knowledge and Smart Technology, KST 2013*, pp. 16–20, 2013, doi: 10.1109/KST.2013.6512780.

[218] R. Rajagopal, "Ch01-P373624.tex Discrete Cosine and Sine Transforms 1.1 Introduction," 2006.

[219] V. Britanak, P. C. Yip, and K. R. Rao, "Fast DCT/DST Algorithms," in *Discrete Cosine and Sine Transforms*, Elsevier, 2007, pp. 73–140. doi: 10.1016/B978-012373624-6/50006-0.

[220] R. Chen, S. Krueger-Ziolek, A. Lovas, B. Benyó, S. J. Rupitsch, and K. Moeller, “Structural priors represented by discrete cosine transform improve EIT functional imaging,” *PLoS One*, vol. 18, no. 5, p. e0285619, May 2023, doi: 10.1371/journal.pone.0285619.

[221] L. Xu *et al.*, “Segmentation of skin cancer images,” *Image Vis Comput*, vol. 17, no. 1, pp. 65–74, 1999, doi: 10.1016/s0262-8856(98)00091-2.

[222] N. S. Zghal and N. Derbel, “Melanoma Skin Cancer Detection based on Image Processing,” *Current Medical Imaging Formerly Current Medical Imaging Reviews*, vol. 16, no. 1, pp. 50–58, 2018, doi: 10.2174/1573405614666180911120546.

[223] M. K. Hasan, M. A. Ahamad, C. H. Yap, and G. Yang, “A survey, review, and future trends of skin lesion segmentation and classification,” *Comput Biol Med*, vol. 155, no. January, 2023, doi: 10.1016/j.compbiomed.2023.106624.

[224] M. A. Siddique and S. K. Singh, “A Survey of Computer Vision based Liver Cancer Detection,” *Int J Bioinform Res Appl*, vol. 18, no. 6, p. 1, 2022, doi: 10.1504/IJBRA.2022.10053584.

[225] D. Divya and T. R. Ganesh Babu, “A Survey on Image Segmentation Techniques,” *Lecture Notes on Data Engineering and Communications Technologies*, vol. 35, pp. 1107–1114, 2020, doi: 10.1007/978-3-030-32150-5_112.

[226] A. T. Beuren, R. Valentim, C. Palavro, R. Janasieivicz, R. A. Folloni, and J. Facon, “Skin Melanoma Segmentation by Morphological Approach,” pp. 972–978, 2012.

[227] M. A. Ahmed Thaajwer and U. A. Piumi Ishanka, “Melanoma skin cancer detection using image processing and machine learning techniques,” *ICAC 2020 - 2nd International Conference on Advancements in Computing, Proceedings*, pp. 363–368, 2020, doi: 10.1109/ICAC51239.2020.9357309.

[228] R. Javid, M. S. M. Rahim, T. Saba, and M. Rashid, “Region-based active contour JSEG fusion technique for skin lesion segmentation from dermoscopic images,” *Biomedical Research*, vol. 30, no. 6, pp. 1–10, 2019.

[229] S. Garg and B. Jindal, “Skin lesion segmentation using k-mean and optimized fire fly algorithm,” *Multimed Tools Appl*, vol. 80, no. 5, pp. 7397–7410, 2021, doi: 10.1007/s11042-020-10064-8.

[230] F. Riaz, S. Naeem, R. Nawaz, and M. Coimbra, “Active Contours Based Segmentation and Lesion Periphery Analysis for Characterization of Skin Lesions in Dermoscopy Images,” *IEEE J Biomed Health Inform*, vol. 23, no. 2, pp. 489–500, 2019, doi: 10.1109/JBHI.2018.2832455.

[231] A. Bassel, A. B. Abdulkareem, Z. A. A. Alyasseri, N. S. Sani, and H. J. Mohammed, “Automatic Malignant and Benign Skin Cancer Classification Using a Hybrid Deep Learning Approach,” *Diagnostics*, vol. 12, no. 10, Oct. 2022, doi: 10.3390/diagnostics12102472.

[232] M. Dildar *et al.*, “Skin cancer detection: A review using deep learning techniques,” *Int J Environ Res Public Health*, vol. 18, no. 10, 2021, doi: 10.3390/ijerph18105479.

[233] R. Javid, M. S. M. Rahim, T. Saba, and M. Rashid, “Region-based active contour JSEG fusion technique for skin lesion segmentation from dermoscopic images,” *Biomedical Research*, vol. 30, no. 6, pp. 1–10, 2019.

[234] U. Saghir and S. K. Singh, “Segmentation of Skin Cancer Images Applying Background Subtraction with Midpoint Analysis,” 2024, doi: 10.1201/9781003405580-93.

[235] H. Sharma, G. Hazrati, and J. C. Bansal, “Spider monkey optimization algorithm,” *Studies in Computational Intelligence*, vol. 779, pp. 43–59, 2019, doi: 10.1007/978-3-319-91341-4_4/COVER.

[236] U. Saghir and V. Devendran, “A Brief Review of Feature Extraction Methods for Melanoma Detection,” *2021 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021*, pp. 1304–1307, 2021, doi: 10.1109/ICACCS51430.2021.9441787.

[237] S. Nazari and R. Garcia, “Automatic Skin Cancer Detection Using Clinical Images: A Comprehensive Review,” *Life*, vol. 13, no. 11, p. 2123, Oct. 2023, doi: 10.3390/life13112123.

[238] S. Arslan, E. Ozyurek, and C. Gunduz-Demir, “A color and shape based algorithm for segmentation of white blood cells in peripheral blood and bone marrow images,” *Cytometry Part A*, vol. 85, no. 6, pp. 480–490, 2014, doi: 10.1002/cyto.a.22457.

[239] F. K. Nezhadian and S. Rashidi, “Melanoma skin cancer detection using color and new texture features,” in *2017 Artificial Intelligence and Signal Processing Conference (AISP)*, IEEE, Oct. 2017, pp. 1–5. doi: 10.1109/AISP.2017.8324108.

[240] J.-A. Almaraz-Damian, V. Ponomaryov, S. Sadovnychiy, and H. Castillejos-Fernandez, “Melanoma and Nevus Skin Lesion Classification Using Handcraft and Deep Learning Feature

Fusion via Mutual Information Measures," *Entropy*, vol. 22, no. 4, p. 484, Apr. 2020, doi: 10.3390/e22040484.

[241] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit*, vol. 29, no. 1, pp. 51–59, Jan. 1996, doi: 10.1016/0031-3203(95)00067-4.

[242] S. Oukil, R. Kasmi, K. Mokrani, and B. García-Zapirain, "Automatic segmentation and melanoma detection based on color and texture features in dermoscopic images," *Skin Research and Technology*, vol. 28, no. 2, pp. 203–211, Mar. 2022, doi: 10.1111/srt.13111.

[243] F. Warsi, R. Khanam, S. Kamya, and C. P. Suárez-Araujo, "An efficient 3D color-texture feature and neural network technique for melanoma detection," *Inform Med Unlocked*, vol. 17, p. 100176, 2019, doi: 10.1016/j.imu.2019.100176.

[244] C. Barata, J. S. Marques, and T. Mendonça, "Bag-of-features classification model for the diagnose of melanoma in dermoscopy images using color and texture descriptors," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7950 LNCS, pp. 547–555, 2013, doi: 10.1007/978-3-642-39094-4_62.

[245] N. Quantology, "A Robust technique for Pre-Processing Liver CT images using Double Stage Gaussian Filtering with Texture and Contrast Enhancement Sentiment Analysis of Social Issues View project Mohammad Anwarul Siddique/A Robust technique for Pre-Processing Liver CT images using Double Stage Guassian Filtering with Texture and Contrast Enhancement A Robust technique for Pre-Processing Liver CT images using Double Stage Guassian Filtering with Texture and Contrast Enhancement," *Neuro Quantology*, vol. 20, no. 9, pp. 1963–1969, 2022, doi: 10.14704/nq.2022.20.9.NQ44227.

[246] Muhammad. U. Khan, Md. R. Beg, and M. Z. Khan, "Improved Line Drawing Algorithm: An Approach and Proposal," no. November, pp. 322–327, 2012, doi: 10.3850/978-981-07-1403-1_713.

[247] M. Rostami, K. Berahmand, and S. Forouzandeh, "A novel community detection based genetic algorithm for feature selection," *J Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-020-00398-3.

[248] D. Singh and B. Singh, “Investigating the impact of data normalization on classification performance,” *Appl Soft Comput*, vol. 97, no. xxxx, p. 105524, 2020, doi: 10.1016/j.asoc.2019.105524.

[249] Y. Zhang, “Support vector machine classification algorithm and its application,” *Communications in Computer and Information Science*, vol. 308 CCIS, no. PART 2, pp. 179–186, 2012, doi: 10.1007/978-3-642-34041-3_27.

[250] A. Masood, A. Al-jumaily, and K. Anam, “System for Classification of Skin Cancer Using SA-SVM,” pp. 101–109, 2014.

[251] N. B. Linsangan and J. J. Adtoon, “Skin cancer detection and classification for moles using K-nearest neighbor algorithm,” *ACM International Conference Proceeding Series*, pp. 47–51, 2018, doi: 10.1145/3309129.3309141.

[252] L. Ballerini, R. B. Fisher, B. Aldridge, and J. Rees, “A color and texture based hierarchical K-NN approach to the classification of non-melanoma skin lesions,” *Lecture Notes in Computational Vision and Biomechanics*, vol. 6, pp. 63–86, 2013, doi: 10.1007/978-94-007-5389-1_4.

[253] K. Taunk, S. De, S. Verma, and A. Swetapadma, “A brief review of nearest neighbor algorithm for learning and classification,” *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, no. May 2019, pp. 1255–1260, 2019, doi: 10.1109/ICCS45141.2019.9065747.

[254] P. D. Caie, N. Dimitriou, and O. Arandjelović, “Precision medicine in digital pathology via image analysis and machine learning,” *Artificial Intelligence and Deep Learning in Pathology*, pp. 149–173, 2021, doi: 10.1016/b978-0-323-67538-3.00008-7.

[255] R. Panigrahi and S. Borah, *Classification and Analysis of Facebook Metrics Dataset Using Supervised Classifiers*. Elsevier Inc., 2019. doi: 10.1016/b978-0-12-815458-8.00001-3.

[256] C. A. Ronao and S.-B. Cho, “Deep Convolutional Neural Networks for Human Activity Recognition with Smartphone Sensors,” 2015, pp. 46–53. doi: 10.1007/978-3-319-26561-2_6.

[257] M. Kumar, M. Alshehri, R. AlGhamdi, P. Sharma, and V. Deep, “A DE-ANN Inspired Skin Cancer Detection Approach Using Fuzzy C-Means Clustering,” *Mobile Networks and Applications*, vol. 25, no. 4, pp. 1319–1329, Aug. 2020, doi: 10.1007/s11036-020-01550-2.

[258] L. H. S. Vogado, R. M. S. Veras, F. H. D. Araujo, R. R. V. Silva, and K. R. T. Aires, "Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification," *Eng Appl Artif Intell*, vol. 72, no. October 2017, pp. 415–422, 2018, doi: 10.1016/j.engappai.2018.04.024.

[259] M. M. Ahsan, S. A. Luna, and Z. Siddique, "Machine-Learning-Based Disease Diagnosis: A Comprehensive Review," *Healthcare*, vol. 10, no. 3, p. 541, Mar. 2022, doi: 10.3390/healthcare10030541.

[260] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[261] V. Nyemeesha, M. Kavitha, and B. Mohammed Ismail, "Detection and Classification of Skin Cancer Using Unmanned Transfer Learning Based Probabilistic Multi-Layer Dense Networks," *Int J Comput Intell Appl*, vol. 21, no. 04, Dec. 2022, doi: 10.1142/S1469026822500274.

[262] H. Bhatt, V. Shah, K. Shah, R. Shah, and M. Shah, "State-of-the-art machine learning techniques for melanoma skin cancer detection and classification: a comprehensive review," *Intelligent Medicine*, vol. 3, no. 3, pp. 180–190, Aug. 2023, doi: 10.1016/j.imed.2022.08.004.

[263] V. A. O. Nancy, P. Prabhavathy, M. S. Arya, and B. S. Ahamed, "Comparative study and analysis on skin cancer detection using machine learning and deep learning algorithms," *Multimed Tools Appl*, vol. 82, no. 29, pp. 45913–45957, Dec. 2023, doi: 10.1007/s11042-023-16422-6.

[264] A. Masood and A. Ali Al-Jumaily, "Computer Aided Diagnostic Support System for Skin Cancer: A Review of Techniques and Algorithms," *Int J Biomed Imaging*, vol. 2013, pp. 1–22, 2013, doi: 10.1155/2013/323268.

[265] M. Dildar *et al.*, "Skin cancer detection: A review using deep learning techniques," *International Journal of Environmental Research and Public Health*, vol. 18, no. 10. MDPI AG, May 02, 2021. doi: 10.3390/ijerph18105479.

[266] S. A. Taghanaki, K. Abhishek, J. P. Cohen, J. Cohen-Adad, and G. Hamarneh, "Deep Semantic Segmentation of Natural and Medical Images: A Review," Oct. 2019, [Online]. Available: http://arxiv.org/abs/1910.07655

[267] M. K. Monika, N. Arun Vignesh, Ch. Usha Kumari, M. N. V. S. S. Kumar, and E. L. Lydia, "Skin cancer detection and classification using machine learning," *Mater Today Proc*, vol. 33, pp. 4266–4270, 2020, doi: 10.1016/j.matpr.2020.07.366.

[268] J. V. Tembhurne, N. Hebbar, H. Y. Patil, and T. Diwan, "Skin cancer detection using ensemble of machine learning and deep learning techniques," *Multimed Tools Appl*, vol. 82, no. 18, pp. 27501–27524, Jul. 2023, doi: 10.1007/s11042-023-14697-3.

[269] T. T. Nguyen, D. M. Nguyen Huu, T.-H. Le, Q.-H. Phan, and T.-T.-H. Pham, "Skin cancer detection using effective optical parameters and the classification and regression tree algorithm: A novel framework," *Ministry of Science and Technology, Vietnam*, vol. 65, no. 1, pp. 63–69, Mar. 2023, doi: 10.31276/VJSTE.65(1).63-69.

[270] D. Hemalatha, K. N. Latha, and Pandala. M. Latha, "Skin Cancer Detection Using Deep Learning Technique," in *2023 2nd International Conference for Innovation in Technology (INOCON)*, IEEE, Mar. 2023, pp. 1–5. doi: 10.1109/INOCON57975.2023.10101344.

[271] M. Shorfuzzaman, "An explainable stacked ensemble of deep learning models for improved melanoma skin cancer detection," *Multimed Syst*, vol. 28, no. 4, pp. 1309–1323, Aug. 2022, doi: 10.1007/s00530-021-00787-5.

[272] H. Ganster, P. Pinz, R. Rohrer, E. Wildling, M. Binder, and H. Kittler, "Automated melanoma recognition," *IEEE Trans Med Imaging*, vol. 20, no. 3, pp. 233–239, Mar. 2001, doi: 10.1109/42.918473.

[273] S. S. Chaturvedi, J. V. Tembhurne, and T. Diwan, "A multi-class skin Cancer classification using deep convolutional neural networks," *Multimed Tools Appl*, vol. 79, no. 39–40, pp. 28477–28498, Oct. 2020, doi: 10.1007/s11042-020-09388-2.

[274] H. H. Kumar, T. Sabherwal, N. Bongale, and M. K. Nair, "Spider Monkey Optimization Algorithm in Data Science: A Quantifiable Objective Study," pp. 115–152, 2020, doi: 10.1007/978-981-15-3689-2_6.

[275] E. V. C. B and D. Ron-dom, *Technology Trends*, vol. 895. in Communications in Computer and Information Science, vol. 895. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-05532-5.

[276] M. A. Ahmed Thaajwer and U. A. Piumi Ishanka, "Melanoma skin cancer detection using image processing and machine learning techniques," *ICAC 2020 - 2nd International Conference*

on Advancements in Computing, Proceedings, pp. 363–368, 2020, doi: 10.1109/ICAC51239.2020.9357309.

[277] M. A. Kassem, K. M. Hosny, and M. M. Fouad, "Skin Lesions Classification into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning," *IEEE Access*, vol. 8, no. June, pp. 114822–114832, 2020, doi: 10.1109/ACCESS.2020.3003890.

[278] A. Pushpalatha, P. Dharani, R. Dharini, and J. Gowsalya, "Retraction: Skin Cancer Classification Detection using CNN and SVM," *J Phys Conf Ser*, vol. 1916, no. 1, 2021, doi: 10.1088/1742-6596/1916/1/012148.

[279] D. Keerthana, V. Venugopal, M. K. Nath, and M. Mishra, "Hybrid convolutional neural networks with SVM classifier for classification of skin cancer," *Biomedical Engineering Advances*, vol. 5, no. December 2022, p. 100069, 2023, doi: 10.1016/j.bea.2022.100069.

[280] A. Bassel, A. B. Abdulkareem, Z. A. A. Alyasseri, N. S. Sani, and H. J. Mohammed, "Automatic Malignant and Benign Skin Cancer Classification Using a Hybrid Deep Learning Approach," *Diagnostics*, vol. 12, no. 10, 2022, doi: 10.3390/diagnostics12102472.

[281] X. Wang, "Deep Learning-based and Machine Learning-based Application in Skin Cancer Image Classification," *J Phys Conf Ser*, vol. 2405, no. 1, 2022, doi: 10.1088/1742-6596/2405/1/012024.

[282] J. Das, D. Mishra, A. Das, M. Mohanty, and A. Sarangi, *Skin cancer detection using machine learning techniques with ABCD features*. 2022. doi: 10.1109/ODICON54453.2022.10009956.

[283] N. C. F. Codella *et al.*, "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM J Res Dev*, vol. 61, no. 4–5, pp. 1–15, 2017.

[284] N. Smaoui and N. Derbel, "Simple but efficient approach for image based skin cancer diagnosis," *2018 15th International Multi-Conference on Systems, Signals and Devices, SSD 2018*, pp. 274–280, 2018, doi: 10.1109/SSD.2018.8570526.

[285] M. Q. Khan *et al.*, "Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer," *IEEE Access*, vol. 7, pp. 90132–90144, 2019, doi: 10.1109/ACCESS.2019.2926837.

[286] B. A. Albert, “Deep Learning from Limited Training Data: Novel Segmentation and Ensemble Algorithms Applied to Automatic Melanoma Diagnosis,” *IEEE Access*, vol. 8, pp. 31254–31269, 2020, doi: 10.1109/ACCESS.2020.2973188.

[287] J. Höhn *et al.*, “Combining CNN-based histologic whole slide image analysis and patient data to improve skin cancer classification,” *Eur J Cancer*, vol. 149, pp. 94–101, 2021, doi: 10.1016/j.ejca.2021.02.032.

[288] A. Jiang *et al.*, “Skin cancer discovery during total body skin examinations,” *Int J Womens Dermatol*, vol. 7, no. 4, pp. 411–414, 2021, doi: 10.1016/j.ijwd.2021.05.005.

[289] F. Perez, C. Vasconcelos, S. Avila, and E. Valle, “Data augmentation for skin lesion analysis,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11041 LNCS, pp. 303–311, 2018, doi: 10.1007/978-3-030-01201-4_33.

[290] S. Serte and H. Demirel, “Gabor wavelet-based deep learning for skin lesion classification,” *Comput Biol Med*, vol. 113, no. August, p. 103423, 2019, doi: 10.1016/j.compbiomed.2019.103423.

[291] J. Zhang, Y. Xie, Y. Xia, and C. Shen, “Attention Residual Learning for Skin Lesion Classification,” *IEEE Trans Med Imaging*, vol. 38, no. 9, pp. 2092–2103, 2019, doi: 10.1109/TMI.2019.2893944.

[292] J. Zhang, Y. Xie, Y. Xia, and C. Shen, “Attention Residual Learning for Skin Lesion Classification,” *IEEE Trans Med Imaging*, vol. 38, no. 9, pp. 2092–2103, Sep. 2019, doi: 10.1109/TMI.2019.2893944.

[293] M. Hasan, S. Das Barman, S. Islam, and A. W. Reza, “Skin cancer detection using convolutional neural network,” in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Apr. 2019, pp. 254–258. doi: 10.1145/3330482.3330525.

[294] A. Javaid, M. Sadiq, and F. Akram, “Skin Cancer Classification Using Image Processing and Machine Learning,” in *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST)*, IEEE, Jan. 2021, pp. 439–444. doi: 10.1109/IBCAST51254.2021.9393198.

[295] W. Gouda, N. U. Sama, G. Al-Waakid, M. Humayun, and N. Z. Jhanjhi, “Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning,” *Healthcare*, vol. 10, no. 7, p. 1183, Jun. 2022, doi: 10.3390/healthcare10071183.

[296] H. C. Reis, V. Turk, K. Khoshelham, and S. Kaya, “InSiNet: a deep convolutional approach to skin cancer detection and segmentation,” *Med Biol Eng Comput*, vol. 60, no. 3, pp. 643–662, Mar. 2022, doi: 10.1007/s11517-021-02473-0.

[297] P. Nadiger, “Skin Cancer Detection and Classification Using Deep Learning,” *Int J Res Appl Sci Eng Technol*, vol. 11, no. 5, pp. 7182–7185, May 2023, doi: 10.22214/ijraset.2023.53390.

[298] S. S. Sundari, Dr. S. AK, and Dr. M. Islabudeen, “Skin Lesions Detection using Deep Learning Techniques,” *Int J Res Appl Sci Eng Technol*, vol. 11, no. 5, pp. 2546–2548, May 2023, doi: 10.22214/ijraset.2023.52129.

[299] R. C. A. Pangsibidang and S. Tuba, “Optimization Detection of Skin Cancer using Deep Learning,” *Fidelity : Jurnal Teknik Elektro*, vol. 5, no. 1, pp. 39–42, Jan. 2023, doi: 10.52005/fidelity.v5i1.139.

[300] C. K. Viknesh, P. N. Kumar, R. Seetharaman, and D. Anitha, “Detection and Classification of Melanoma Skin Cancer Using Image Processing Technique,” *Diagnostics*, vol. 13, no. 21, p. 3313, Oct. 2023, doi: 10.3390/diagnostics13213313.

[301] J. V. Tembhurne, N. Hebbar, H. Y. Patil, and T. Diwan, “Skin cancer detection using ensemble of machine learning and deep learning techniques,” *Multimed Tools Appl*, vol. 82, no. 18, pp. 27501–27524, Jul. 2023, doi: 10.1007/s11042-023-14697-3.

[302] D. Keerthana, V. Venugopal, M. K. Nath, and M. Mishra, “Hybrid convolutional neural networks with SVM classifier for classification of skin cancer,” *Biomedical Engineering Advances*, vol. 5, p. 100069, Jun. 2023, doi: 10.1016/j.bea.2022.100069.

APPENDIX A

Algorithm 4.1: Skin Cancer Customized Classifier (SCCC)

Customized classifier (n,S_prob,$Prob_{nodeselection}$,N_e)

"n" is the number of images

"SF" selected feature

S_prob indicates the switching probabilities

$Prob_f$ indicates the probability of feature selection

N_e is the image features on training data

Defining linearity function f(n).

Define the number of images (n), switching probabilities {0,1} and the maximum number of generations.

- While ($g<iteration_{max}$) or(termination_criteria)
  - Select a feature from N_e
  - SF=max(corr(Ng,Target))
  - Verify the global solution(G_g)=f(SF^{g+1}) for validating the fitness
- While the fitness of the n_g is not improved or ($g<iteration_{max}$) or(termination_criteria)
  - Perform learning based on the linearity function
  - Update local and global counter
- If($G_g>G_{g+1}$) then
  - G_{g+1}<-G_g //The old solution is replaced with the new solution

  End if
- If(r(0,1)<S_prob) then
  - Initialize the dataset again
  - Perform correlation again to handle a better global solution
  - Increase the value of the local and global counter by 1.

End of if

- If($G_g<G_{g+1}$) then
  - Select N_e(g) for selection
  - Assign rank to the feasible solution Feature_Vec(t)=N_e(g)
  - t=t+1

end of while

- return value

APPENDIX B

PUBLICATIONS

1. U. Saghir and M. Hasan, “Skin cancer detection and classification based on differential analyzer algorithm,” *Multimed Tools Appl*, vol. 82, no. 26, pp. 41129–41157, Nov. 2023, doi: 10.1007/s11042-023-14409-x. **(SCIE Indexed)**
2. U. Saghir, S. K. Singh, and M. Hasan, “Skin Cancer Image Segmentation Based on Midpoint Analysis Approach,” *Journal of Imaging Informatics in Medicine*, Apr. 2024, doi: 10.1007/s10278-024-01106-w. **(SCIE Indexed)**
3. U. Saghir and S. K. Singh, “Segmentation of Skin Cancer Images Applying Background Subtraction with Midpoint Analysis,” in *Computer Science Engineering and Emerging Technologies*, London: CRC Press, 2024, pp. 574–579. doi: 10.1201/9781003405580-93. **(SCOPUS Indexed)**
4. U. Saghir and V. Devendran, “A Brief Review of Feature Extraction Methods for Melanoma Detection,” in *2021 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021*, Institute of Electrical and Electronics Engineers Inc., Mar. 2021, pp. 1304–1307. doi: 10.1109/ICACCS51430.2021.9441787. **(SCOPUS Indexed)**
5. Saghir, U., Singh, S.K. & Hasan, M (2024). “Classification of Skin Cancer-Based on Dermoscopic Images Using Customized and Machine Learning Classifier”. **Communicated**

CONFERENCES ATTENDED

1. 7th International Conference on Advanced Computing and Communication Systems (ICACCS-2021).
2. Booth100 6th International Joint Conference On Computing Sciences (ICCS-2022).