

# **Skin Lesion Segmentation and Classification Using Multi-Level Feature Fusion and Selection Framework**



*By*

**Farzana Kousar**

**Reg No. CIIT/FA19-REE-018/WAH**

**MS Thesis**

**In**

**Electrical Engineering**

**COMSATS University Islamabad,  
Wah Cantt- Pakistan**

**Spring, 2021**



**COMSATS University Islamabad**

# **Skin Lesion Segmentation and Classification Using Multi-Level Feature Fusion and Selection Framework**

A Thesis Presented to

**COMSATS University Islamabad, wah cantt**

In partial fulfillment

of the requirement for the degree of

## **MS Electrical Engineering**

By

**Farzana Kousar**

**CIIT/FA19-REE-018/WAH**

**Spring, 2021**

# Skin Lesion Segmentation and Classification Using Multi-Level Feature Fusion and Selection Framework

A Post Graduate Thesis submitted to the Department of Electrical Engineering as partial fulfillment of the requirement for the award of Degree of MS in Electrical Engineering.

Name	Registration Number
Farzana Kousar	CUI/FA19-REE-018/WAH

## Supervisor

Dr. Tallha Akram  
Associate Professor,  
Department of Electrical and Computer Engineering  
COMSATS University Islamabad (CUI), Wah

## Final Approval

---

This thesis titled

# **Skin lesion Segmentation and Classification using Multi-level Feature Fusion and Selection Framework**

**By**

*Farzana Kousar*

*Registration No CUI/FA19-REE-018/WAH*

**Has been approved For the COMSATS University Islamabad, Wah Cantt**

External Examiner: \_\_\_\_\_

Dr xxxxx xxxxxxxxxx  
Affiliation

Supervisor: \_\_\_\_\_

Dr. Tallha Akram  
Department of Electrical and Computer Engineering, CUI Wah

HoD: \_\_\_\_\_

Dr. Muhammad Iqbal  
Department of Electrical and Computer Engineering, CUI Wah

## Declaration

I Farzana Kousar reg no. CUI/FA19-REE-018/WAH hereby declare that I have produced the work presented in this thesis, during the scheduled period of study. I also declare that I have not taken any material from any source except referred to wherever due. If a violation of HEC rules on research has occurred in this thesis, I shall be liable to punishable action under the plagiarism rules of HEC

Date: 06 July, 2021  
(Date of thesis submission)

Signature of the student:

---

**(Farzana Kousar)**  
**(CUI/FA19-REE-018/WAH)**

## Certificate

It is certified that Author Name CUI/FA19-REE-018/WAH has carried out all the work related to this thesis under my supervision at the department of Electrical Engineering COMSATS University Islamabad, WahCantt, and the work fulfills the requirement for award of MS Degree. Certified that the originality of the contents of the research work is verified through plagiarism software Turnitin and found similarities within acceptable range.

Date: 28 June 2021  
(Date of thesis submission)

Supervisor:

-----

Dr. Tallha Akram  
Associate Professor,  
Department of Electrical and Computer Engineering,  
CUI Wah

**Head of Department:**

\_\_\_\_\_  
Dr. Muhammad Iqbal  
Department of Electrical and Computer Engineering, CUI Wah

## DEDICATION

**To my**

**Great Parents,**

**My supervisor,**

**My teachers.**

## **ACKNOWLEDGEMENTS**

I would like to thank Dr. Tallha Akram my supervisor at CUI Wah Campus for their constant support and guidance for this research work. They shared their knowledge and provided many useful references and friendly encouragements.

Finally, Allah almighty has always been very kind.

Farzana Kousar

CUI/FA19-REE-018/WAH



## ABSTRACT

### **Skin lesion Segmentation and Classification using Multi-level Feature Fusion and selection framework**

The most deviant and ornery form of skin lesion is cancer, the topmost dangerous and life-threatening skin lesion. In the proposed research work, we are dealing with classification of skin lesion using machine learning techniques, and specially using deep convolution neural networks. In this work we present Multi-level Feature Fusion and selection framework to Segment and classify benign & malignant type of skin lesion. The proposed framework incorporates three different modules starting from segmentation, feature extraction and finally classification. For this we shall be utilizing clinical images from set of benchmark datasets including PH2, ISCI16, ISIC17 both for training and testing. The main objective of proposed plan is, initially enhance the image contrast and later extracting Region of Interest (ROI) using Saliency estimation. Deep Network i.e., Deeplabv3+ is used to segment lesion images independently which then combined to generate fused image version of two segmentation techniques. In the second phase, features will be extracted after applying Transfer Learning using multiple pre-trained models such as Inception-Resnet-v2, & Inception V3. Feature vectors are optimized using biological inspired Algorithm i.e., Grey Wolf Optimization (GWO) Technique. These features, after fusion fed into different classifiers i.e., Support Vector Machines (SVM), K-Nearest Neighbor (K-NN). The maximum achieved segmentation accuracy is on  $PH^2$  dataset with average Jaccard of 88.54 %, whereas the minimum segmentation accuracy is on ISCI 2017 database with 81.24 % Jaccard. SVM performs exceptionally as compared to other classifiers by achieving the classification accuracy of 96.1%, 92.5%, and 90.4% on PH2, ISBI 2016 and ISIC 2017.

# TABLE OF CONTENTS

<b>1</b>	<b>Introduction.....</b>	<b>1</b>
1.1	Skin Structure and Skin lesion .....	2
1.1.1	Types of Skin cancer .....	3
1.2	Automated Diagnostic System and Machine Learning .....	4
1.3	Motivation.....	7
1.4	Significance.....	7
<b>2</b>	<b>Literature Review .....</b>	<b>8</b>
	Literature review.....	9
2.1	Segmentation .....	9
2.1.1	Conventional Methods.....	9
2.1.2	Deep Learning Methods .....	10
2.2	Feature Extraction & selection .....	12
2.3	Transfer learning: .....	14
2.4	Classification:.....	15
2.4.1	Conventional methods:.....	15
2.4.2	Deep learning methods:.....	17
<b>3</b>	<b>Proposed Segmentation and classification of skin lesion.....</b>	<b>20</b>
3.1	Image dataset collection: .....	21
3.2	Segmentation .....	21
3.2.1	Saliency detection-based Segmentation technique .....	22
3.2.2	CNN based Segmentation: Deeplabv3Plus.....	30
3.2.3	Image Fusion Framework. ....	34
3.3	Transfer Learning, Feature selection and feature vector optimization. ....	36
3.3.1	Grey Wolf Optimization.....	37
3.4	Classification.....	41
<b>4</b>	<b>Simulation and Results .....</b>	<b>43</b>
4.1	Performance evaluation metrics .....	45
4.1.1	Dice: .....	45
4.1.2	Jaccard:.....	46
4.1.3	Sensitivity (Sen),.....	46
4.1.4	Specificity (Spe).....	46
4.1.5	Precision (Pre),.....	46

4.1.6	False negative rate (FNR),.....	46
4.1.7	False positive rate (FPR), .....	46
4.1.8	Accuracy (Acc).....	46
4.2	Performance Evaluation: Segmentation .....	46
4.3	Performance Evaluation: Classification .....	49
4.4	Conclusion:.....	52
<b>5</b>	<b>References</b> .....	<b>53</b>

## LIST OF FIGURES

---

Figure 1-1 Human Skin cross section view .....	2
Figure 1-2 Melanocytes and abnormal growth of cell .....	3
Figure 1-3 Melanoma; a: Malignant, b: nevus (source HAM dataset).....	4
Figure 1-4 Types of Skin Cancer .....	6
Figure 2-1 Conventional methods for segmentation .....	10
Figure 2-5 Deep learning methods for segmentation. ....	17
Figure 2-6: Conventional and Deep learning classification approaches .....	19
Figure 3-1 Proposed Framework.....	21
Figure 3-2. Skin lesion segmentation. a) Dermoscopy image in input. b) Binary mask in output. c) The border of the mask overlaid on the enhanced image. ....	21
Figure 3-3 Saliency estimation using information weight content .....	23
Figure 3-4 Semantic segmentation (source) .....	30
Figure 3-5 Deeplabv3 architecture[62].....	32
Figure 3-6 Pyramid pooling .....	33
Figure 3-7 atrous convolution .....	33
Figure 3-8 DCNN's structure (Source: MathWorks).....	36
Figure 3-9: Detailed flow diagram for Lesion classification .....	41
Figure 3-10: Entropy-Controlled binary GWO Feature Selection .....	42
Figure 4-1 Segmentation results .....	45
Figure 4-2 Segmentation performance: Dice index for selected samples .....	48
Figure 4-3 Segmentation Performance: Jaccard Index for sample data .....	48
Figure 4-4: Average Jaccard Indices for Selected Dataset .....	49
Figure 4-5: Performance parameters for PH2 Dataset .....	50
Figure 4-6: Performance Parameters for ISIC-2016 .....	51
Figure 4-7: Performance Parameters for ISIC-2017 .....	51
Figure 4-8 Classification performance parameters at three datasets.....	52

## LIST OF TABLES

---

Table 4-2 Dataset Used.....	<b>Error! Bookmark not defined.</b>
Table 4-3: Dice & Jaccard indices .....	47
Table 4-4: Classification Performance Parameters .....	49

## **LIST OF ABBREVIATIONS**

ABC	Artificial Bee Colony
ANN	Artificial Neural Network
CNN	Convolutional Neuro Network
DCNN	Deep convolution Neural Network
ECNCA	Entropy-Controlled Neighborhood component analysis
FCN	Fully Convolutional Network
FrCN	Full resolution Convolutional Network
GA	Genetic Algorithm
GHT	Gradient Histogram Thresholding
GWO	Grey Wolf Optimization
IMFO	Improved moth flame optimization
INR	iterative Newton-Raphson
KNN	K-Nearest Neighborhood
LBP	Local Binary Patterns
LOG	Laplacian of a gaussian
MLP	Multilayer Perceptron
MMCA	multiset maximum correlation analysis
MMSE	Minimal Mean square error
NFCA	Neighborhood feature component analysis
PP	Product of probabilities
RCNN	Regional based convolutional neural network
ROC	receiver operating characteristic
ROI	Region of interest
SA	Simulated annealing
SMP	Sum of the maximal Probabilities
SMV	Simple Majority Voting
SP	Sum of Probabilities
IOU	Intersection Over union

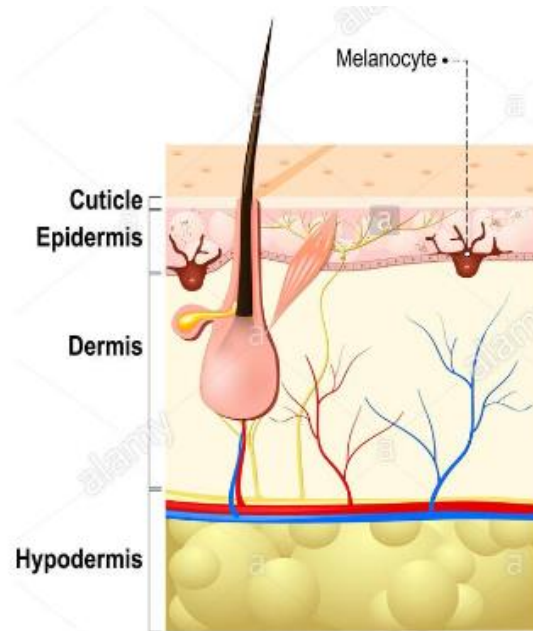
SVM	Support Vector machine
TL	Transfer learning
YOLO	You Only Look Once

# **Chapter 1**



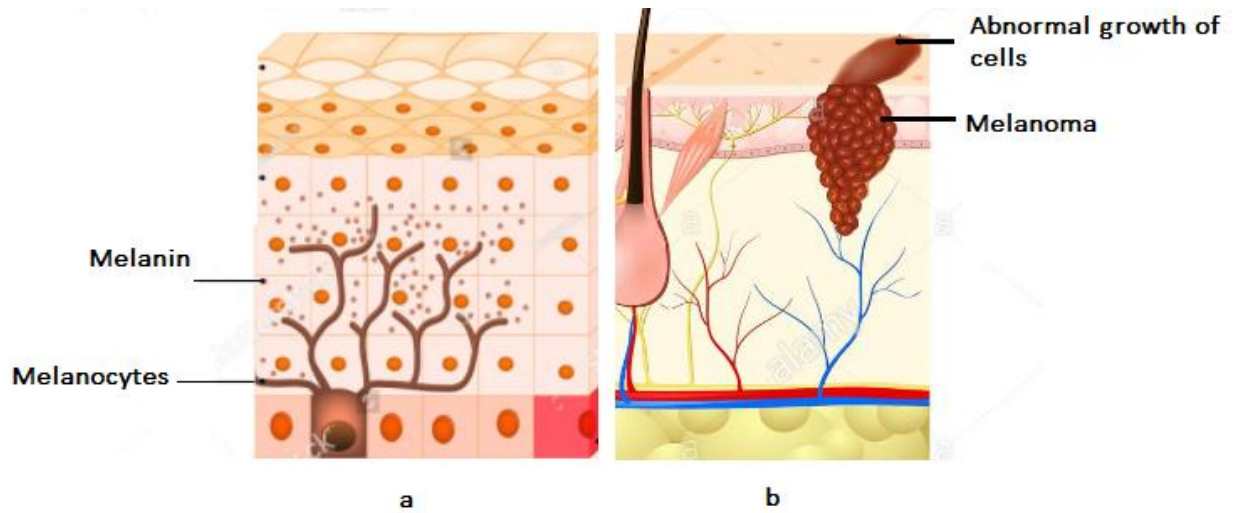
## 1.1 Skin Structure and Skin lesion

Skin is a natural barrier that protects human body from the external environment like sunlight, temperature, virus etc. Basically, skin is stack of three layers, top layer is epidermis, central is dermis, and last layer, which is adjacent to muscle is hypodermis. There are some special cells which are rooted in in epidermis called melanocytes as shown in Figure 1-1. These melanocytes are the key reason to generate the pigment melanin in the skin, which is accountable for skin color.



**Figure 1-1 Human Skin cross section view**

Production of melanin by melanocytes is excited by exposure of skin to strong radiations i.e. Ultraviolet radiations[1]. Skin color changes if there comes certain variation in the normal amount of melanin. Disturbance in the level of melanin results pigment disorders, usually known as skin lesions as shown in Figure 1-2. The two-fundamental driver of the skin disease are the sun's destructive ultraviolet (UV) beams and the utilization of UV tanning machines. These strong radiation result in uncontrolled growth of abnormal cells in the epidermis, where theses mutations are triggered by unrepaired DNA damage [2]. These abnormal skin cells are multiplied and tend to form malignant tumors. Skin lesions can be categorized as primary and secondary skin lesion. Further classification depends upon the stage of diversity to which skin tissues perverted.



**Figure 1-2 Melanocytes and abnormal growth of cell**

Skin lesion may be as harmless as scrap and may be as swear as cancer. Consequences of skin lesion vary from no worry mark to sever such as disability, death and depression which may lead to suicide[3]. Skin lesion is becoming global disease burden as it causes to increase disability rate worldwide. The most perverted form of skin lesion is skin cancer which may lead to death.

### **1.1.1 Types of Skin cancer**

Skin cancer is the dangerous most and life threatening skin lesion.[4]. It is also the world most common human cancer. According to American Academy of Dermatology Association 2020 [5], there are four kinds of skin cancer: Basal Cell Carcinoma (BCC), Squamous Cell Carcinoma (SCC), Merkle Cell Carcinoma (MCC) and Melanoma.

Melanoma is the lethal skin cancer which can be further classified into benign and Malignant Melanoma (MM) as in Figure 1-3. Although Melanoma is rare as compared to non-Melanoma, but it is deadly aggressive, Only 22- 30% of Melanoma are found in existing moles where 70 to 80% arises on a normal skin [2]. Different types of skin cancer are shown in Figure 1-4. Non-melanoma skin cancer refers to all kinds of skin cancer not melanoma.

Benign melanoma used to develop slowly in epidermis layer of skin and do not spread as such. It is very common type of skin cancer, where MM is uncommon but life threatening. Malignant melanoma penetrates deep into the skin tissues and tends to effect nearby skin



**Figure 1-3 Melanoma; a: Malignant, b: nevus (source HAM dataset)**

(WHO 2018)[6]. MM is susceptible to spreads rapidly through the skin tissues. According to world Cancer Research fund[7] melanoma is 19th most common cancer worldwide and it is fifth common type of cancer in UK which is expected to increase, whereas 123000 cases are diagnosed worldwide every year [1].

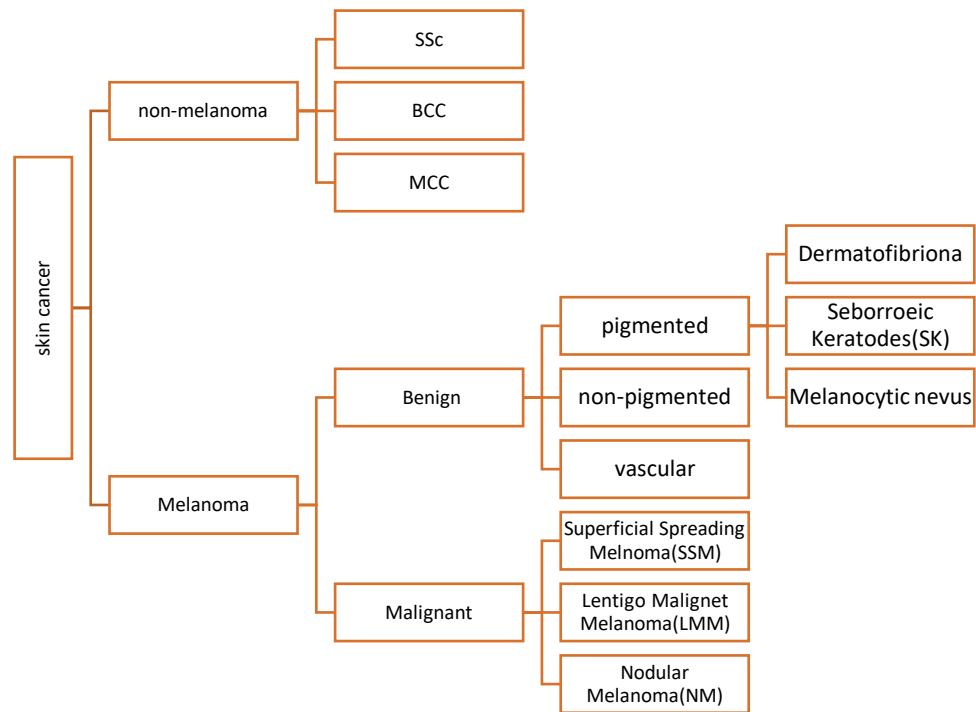
Skin cancer is treatable if diagnoses made at early age. Early diagnosis reduces death rate significantly, even, chances of survival increase more than 90 % [8]. According to American cancer society (Cancer Facts & Figures 2020), there were about 16.9 million American citizens by Jan, 2019, who survived from cancer just because of early diagnosis and respective cure [9]. According to world Cancer Research fund, melanoma is 19th most common cancer worldwide and it is 5th common type of cancer in UK which is expected to increase. According to Union of international cancer control (UICC), It has risen to 19.3 million cases and 10 million cancer deaths in 2020.

## ***1.2 Automated Diagnostic System and Machine Learning***

Conventional clinical approaches namely Biopsy & Dermoscopy methods are used for diagnosis of skin lesion. Biopsy is an Invasive method of diagnosis in which a part of the affected skin is removed as sample to examine the type and extent of disease. Dermoscopy

is non-invasive method of diagnosis in which affected skin is illuminated and magnified many 100 times to make it clear to be viewed by naked eye as well as taking & storing the images of the skin for later use. Dermoscopy, have been in practice for lesion diagnosis by clinicians since many years. The results are much better then visual inspection. It is used to examine skin by means of surface microscopy. It is carried out with help of high-class magnifying lens and a powerful lighting system. It examine the highlighted skin pattern in terms of asymmetry (A), border (B), color (C), and differential structure (D) of the affected skin (ABCD of dermoscopy)[10]. Any proper device attachment i.e., digital camera or smart phone, with surface microscopy can produce still skin photographs to examine the skin later. The ABCD rule when applied to clinical/dermatological images, evidenced to improve diagnostic accuracy.

Despite dermoscopy provide clear image still, this manual inspection is tedious and time taking as well as need intense practice and expertise[11] [12]. There lies a wide room for improvements for accurate diagnosis and classification using computer aided design techniques (CAD). Also, there is a need of reliable automated non-invasive diagnostic system for early diagnosis of skin lesion in less time with the accuracy at a level of competence comparable to domain experts. Also, there is a need of reliable automated non-invasive diagnostic system for early diagnosis of skin lesion in less time with the accuracy at a level of competence comparable to domain experts. The amalgam of computer vision and machine learning have proven to develop amazing, automated Computer Aided diagnostic system.



**Figure 1-4 Types of Skin Cancer**

Machine learning (ML) techniques are aiding to solve many challenging problems of medical industry including automated diagnostic systems [13]. With the headway of ML, researchers have tremendously worked out various automated diagnostic systems to overcome the bottleneck left by conventional clinical methods. These automated systems help early and accurate detection of skin lesion with a level of competence comparable to domain experts who are using conventional methods of classifying skin diseases. With the development of massive computing capacity and big data, ML have evolved into deep learning and result in revolutionized image classification. Deep learning methods i.e. Convolutional Neural Networks (CNN) are capable to model high-level abstractions in data by using many processing layers. This really made DL good enough to solve challenging image classification. To manage DL models for different applications it needs training a DL model. A big data and huge time are required to train the model, here Transfer Learning (TL) comes into the play. Transfer learning is an approach used in a process of training a model, where a

model already trained for a task is utilized as the starting point for the training of new model to be applied on some other related tasks

### **1.3 Motivation**

Skin cancer is becoming global disease burden as it causes disability worldwide with high mortality rate. Early detection and screening of skin cancer is important before it reaches lethal stage e. g., Malignant Melanoma. Early diagnosis increases the survival rate more than 90%. Conventional clinical approaches used for diagnosis of skin lesion namely Biopsy (Invasive) & Dermoscopy (ABCDE rule) methods are time consuming, more prone to human error, needs intense practice, experience, and expertise. Conventional ML system need third party intervention and do only for what they are designed for; nothing more, nothing less. Deep Learning achieves great performance power and flexibility with big data. With combination of Transfer learning, it further reduces computational complexity. This work proposes an automatic classification of Skin cancer which is helpful in assisting dermatologists in their diagnosis.

### **1.4 Significance**

When Cancer is diagnosed and treated at earliest stage, the chances of survival increase more than 90 %. The proposed model will be helpful to dermatologist to determine skin lesion in less time with the accuracy at a level of competence comparable to domain experts who are using conventional methods of classifying skin diseases. These algorithms may be used to develop Android based application to help user to monitor and diagnose these types of skin lesion. The develop model may also be helpful in digital health monitoring.

## **CHAPTER 2**

## *Literature review*

The automated diagnostic systems involve four stages: Pre-processing, Segmentation, feature extraction and classification [14][15]. Where, Segmentation and classification are two important critical procedures for the accurate development of automated diagnostic model.

This chapter is about related work for classifying skin cancer. This review will cover following

1. Segmentation (Conventional ML and DL methods)
2. Feature vector extraction and selection.
3. Transfer Learning
4. Classification (Conventional ML and DL methods)

## *2.1 Segmentation*

### **2.1.1 Conventional Methods**

Segmentation extracts the attributes related to clinical images. It is the key step in automated diagnosis as it discriminates Skin lesion region from the normal and healthy skin and pave the path of diagnostic process to achieve accurate classification results. The journey of segmentation started from hand crafted features with traditional image processing and machine learning methods like Histogram threshold[16], edge detection[17], Clustering <sup>1</sup>, and supervised methods such as Support vector machine (SVM) [18]. These methods still posed the challenges for classification because of tough contrast and boundary clearance of lesion. Different conventional [19] [20]and deep learning [20][21]method for segmentation are summarized in **Error! Reference source not found.**Figure 2-2

---

<sup>1</sup> Emre Celebi M, Wen Q, Hwang S, Iyatomi H, Schaefer G (2013) Lesion border detection (**library**)



Pedro M.M. Pereiraa et al in [22] proposed a hybrid model from local binary patterns (LBP) clustering and Gradient Histogram Thresholding (GHT) to create segmentation mask. They incorporated boarder line features in segmentation of images in addition to traditional hand-crafted features i.e. shape, color and texture to improve the classification of melanoma and nevus. Their model showed improved Diagnostic performance and reliability to classify melanoma at the accuracy above 90% while managing commonly occurring clinical imbalanced datasets.

Figure 2-3 shows conventional as well as deep learning methods, which have been reported in literature for the classification [23][24][25][26][27].

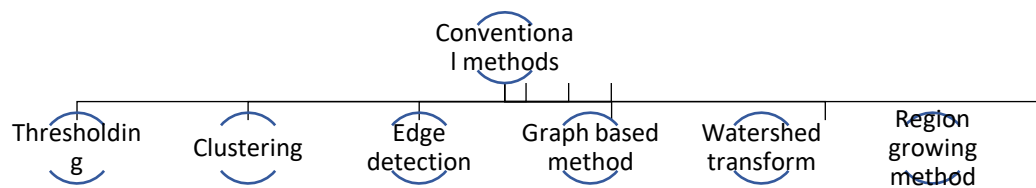


Figure 2-1 Conventional methods for segmentation

### 2.1.2 Deep Learning Methods

In research[28], a unique combination of a deep learning model and iterative Newton-Raphson (IcNR) based feature selection approach is suggested for skin cancer localization and detection. The extract ROI using Fast region-based convolutional neural network (RCNN) followed by the IcNR technique for selection of highly detailed features. they use contrast stretching strategy founded on the Artificial Bee Colony method (ABC) in preprocessing and segmentation process. They applied transfer learning on DenseNet201 pre-trained model, after which the selected models are tested with the IcNR strategy. Multilayered feed-forward neural networks employed upon the selected most discriminant features for classification. Their model delivers accuracies of 94.5% and 93.4% on ISBI2016 and ISBI2017 datasets.

In [29] authors proposed a blend of deep features extracted from three CNNs at different abstraction level. they utilize a model composed of an SVM and a convolutional network for classification. Improved classification performance is gained by the combination of various prediction probability classification vectors, achieved from diversity of training models. They used ISIC 2017 dataset for performance evaluation.

In [15] The authors proposed a novel and pipeline for skin lesion segmentation in dermoscopic images by combining a deep convolutional neural network i.e. You Only Look Once (YOLO) with the Grab Cut algorithm. Their method performs lesion segmentation using a clinical image. Their proposed method was evaluated on two publicly well-known datasets, that is the PH2 and the ISBI 2017. The proposed pipeline method has achieved a 90% sensitivity and outperformed other deep learning-based methods

In recent years Deep learning methods like Neural Network are emerged as a promising option in contrast to different traditional ML methods and, have been widely used for the medical image segmentation and classification to aid automated diagnostic system [15].

Although conventional ML techniques are extensively used for segmentation but due to complex visual nature of lesion images i.e. low contrast between healthy and affected region etc., still there exist multiple challenges to accurately extract true region of interest[30][31]. Here, deep learning techniques came into the ground, perform intelligently, and keep scoring good success in automating the diagnostic system.

Lei Bi [32] use fully convolutional network (FCN) for lesion segmentation. To enhance lesion boundary, they take average of multiple scaled images and achieve Accuracy of 95.10%.

Similarly, Mohammed A. Al-masni [12] use full resolution convolutional network (FrCN) to segment the lesion followed by multiple deep classifiers to automate the early diagnosis. Lesion boundaries were extracted using deep FrCN

segmentation. Traits of input pixels were preserved to avoid losing spatial resolution. This was done by eliminating sub-sampling layers in FrCN and exploiting the convolutional layers. The boundary features hauled from segmentation mask are followed by CNN classifiers i.e., Inception-v3, ResNet-50, Inception-ResNet-v2, and DenseNet-201 to classify melanoma. ResNet-50 reported to the best of all four classifiers used.

A hybrid of conventional and deep ML models was also reported in literature to further improve classification of melanoma. Similar work is done by Nazia Hameed et.al. and reported 96.45% accuracy [13]. They used hybrid technique of clustering and thresholding for segmentation followed by multi-layered feed-forward artificial neural network (ANN) to classify three classes of skin lesions i.e. benign, malignant, and eczema. The best results were achieved with four hidden layers of ANN. They used deep learning convolutional neural network (CNN) for classification, AlexNet network. The traditional machine learning model achieved 61.64% & 63.79% accuracy for single level & multilevel classification. While these efficiencies improved to 96.03% & 96.47% for the hybrid deep learning CNN model.

To make use of different deep network capabilities, many researchers follow ensemble of different techniques. Manu Goyal et.al proposed such ensemble model to do segmentation [33]. Initially they pre-process images by applying shades of grey algorithm to stabilize the color contrast. In second step, two deep ML technique for segmentation namely DeeplabV3+ and MASk R-CNN were employed. Before ensemble of segmentation of these two, a post processing for DeeplabV3+ segmentation was done to remove artefacts i.e., broken regions. Then, in third step, ensemble of two segmentations is done to produce final segmentation mask. Their model prove accuracy of 94.1% for the ISIC dataset.

## **2.2 Feature Extraction & selection**

Features which are extracted after segmentation process plays a vital role in classifying lesion type. Quality of features extracted, directly manipulates the

classification accuracy. Recently feature fusion practices are introduced in literature. A primary goal of fusion approaches is to increase the amount of information from numerous sources. In pursuance of feature fusion, feature vector length gets to increase as well. However, the amount of computing time is affected by this adjustment and problem of overfitting may also rise. Feature selection provides a means for finding the essential features while excluding those that are superfluous from the feature vector <sup>2</sup>. Thus, feature selection will cut down on the number of dimensions, improve prediction, and good data interpretation for classification[34]. Many researchers attempt to treat the feature selection as a problem of optimization, leading towards the finest feature space. Several classical <sup>3</sup> [35]and evolutionary algorithm [36] [37] [38] have been reported in literature for feature optimization.

Genetic Algorithm (GA) was employed in [3] to determine the appropriate feature vector. GA proved to be good enough in reducing feature vector length and improved accuracy.

Author [39] optimize their feature selection using improved moth flame optimization (IMFO) algorithm followed by multiset maximum correlation analysis (MMCA) to finally generate fused feature vector. They analyze segmentation performance on ISBI 2016-2017 and ISIC 2018 and achieve accuracy of 95%, 95.79%, 92.69%, and 98.70%, respectively. While performance evaluation of classification is made with HAM10000 dataset with 90.67% accuracy.

In research [40] Author used Alex-net and VGG-16 model to extract feature vector which followed by serial fusion and cut on dimension with PCA. They

---

<sup>2</sup> B. Chizi, L. Rokach, O. Maimon, "A Survey of Feature Selection Tech-niques", Encyclopedia of Data Warehousing and Mining, Second Edi-tion, IGI Global, pp. 1888-1895, 2009.(dd library)

<sup>3</sup> R. O. Duda, P. E. Hart, D. G. Stork, "Pattern Classi\_cation" (2<sup>nd</sup> Edition), Wiley-Interscience, 2000.(add library)

analyze classification performance on combined dataset of ISBI 2016-2017, and PH2 through classification learner.

In [41], Author report segmentation method with maximum accuracy of 95.32% for the set of ISBI 2016 images. Dull Razour is used to remove noise i.e., hairs and features were extracted based on extended texture feature analysis. Neighborhood feature component analysis (NFCA) used for feature reduction. Evaluation results are processed on Multi-SVM classifier with PH2, ISBI 2016 and ISIC data sets. Classification accuracy achieved with ISBI 2016 is 99.2 & PH2 is 97.5 %.

### ***2.3 Transfer learning:***

Although deep learning methods need a lot of data for training but handling that much data for training is another challenge. Also, large computing resources and memory is an extra head for the system to deal with. Transfer learning (TL) is now becoming popular with deep learning methods when it comes to deal with intense data. TL starts with base model already trained for a task as the starting point for the training of new model. Base data is used to train network and extracting the feature map, which is to feed the second model to train for another task. Similar challenges have addressed by Talha Akram et.al. [42] and Mahbod et.al [43]. They also incorporate feature fusion techniques to reduce computational complexities and overfitting. Talha et.al proposed framework is exploiting feature fusion techniques along with feature dimension reduction using Entropy-Controlled neighborhood component analysis (ECNCA). Their proposed model not only achieve overall accuracy but also reduces computational time and memory requirement. At first, they extract features by applying transfer learning approach on various pre-trained CNN models i.e. inception-V3, Inception-Resnet-V2. Later a compact feature map was generated using ECNCA by truncating irreverent features while keeping original feature space well-preserved. They validated their model with accuracy of 98.8% at PH2, 99.2% at ISIC MSK, 97.1% at ISIC UDA and 95.9% at ISBI-2017 dataset.

In [43], Authors use three CNN models, AlexNet, VGG16, ResNet-18 . They produce feature vector from fully connected layers at different abstraction level. They fit SVM classifiers over that feature vector and then fused the output of each SVM classifier. they achieve area under receiver operating characteristic (ROC) of 83.83% and 97.55%for melanoma and benign skin cancer.

## **2.4 Classification:**

Classification is decision making step which, decides the object class of the under considered skin lesion samples. This section will go through different conventional and DL techniques and methods used for classification step.

### **2.4.1 Conventional methods:**

Once feature extraction is done next is classification. Custom machine learning methods as Decision Tree & nearest neighbor methods[44] , support vector machines (SVMs) [45], logistic regression, and Artificial Neural Networks (ANNs), have been reported for classification with modest success [23]. Figure 5 shows conventional as well as deep learning methods, which have been reported in literature for the classification[24][25][26][27].

In [34] Nasir et.al. submitted a method for classifying skin lesion, including melanoma. They applied DullRazor hair removal procedures in the pre-processing. To increase the contrast of the lesions, they employ lesion texture and colour information. In lesion segmentation, a hybrid methodology that uses both additive law of probability and exponential distribution has been implemented. the property of colour, texture, and HOG are successively extracted and fused using a serial technique. In the end, they adopt a Boltzmann-entropy-based technique to pick the fused features. After completing the features selection, the features are care passed to Support Vector Machine. The strategy presented is tested on the PH2 public data set and report 97.7% sensitivity, a 96.7% specificity, an accuracy of 97.5%, and an F-score of 97.5%.

There is an increasing tendency to use deep convolutional neural networks (CNNs) for medical image processing, such as skin lesion classification. With

CNNs, which are optimized using a large dataset of real pictures, used as feature extractors for skin lesion pictures, the bottlenecks of conventional techniques may be overcome, and furthermore, for specified tasks, tiny datasets may be used without any effect on the outcome.

[46] , Researcher proposed classification model using CNN in combination with one-versus all. they input skin lesion images without any preprocessing. To classify the skin diseases, researchers use Convolutional Neural Network model, and its combination with one- versus- all. They use HAM10000 dataset. They gained and achieved accuracy of 77 % with CNN only while it got improved to 92.90% with the combination of one-versus-all methodology.

Ghasem et al. presented a meta learning based hybrid structure of heterogeneous classifiers to categorize skin lesions [11]. They presented two level classification where, at each level combination of three conventional classifier is employed. Four well known classifiers i.e., K-Nearest Neighborhood (KNN), Support Vector Machine (SVM), and Multilayer Perceptron (MLP) Neural Network was used. In first stage, to differentiate between melanoma and non-melanoma, SVM, KNN & MLP are combined. In second stage SVM, MLP & ENN were combined to further sort dysplastic and benign type of skin lesion. Training data set was partitioned into different subsets for two stage classifiers to avoid overfitting. Output of first stage becomes input feature space for second stage classifiers, whose training is done with different training dataset. The model attains 96% accuracy in classifying melanoma, and 93.33% accuracy in further categorizing dysplastic and benign melanoma.

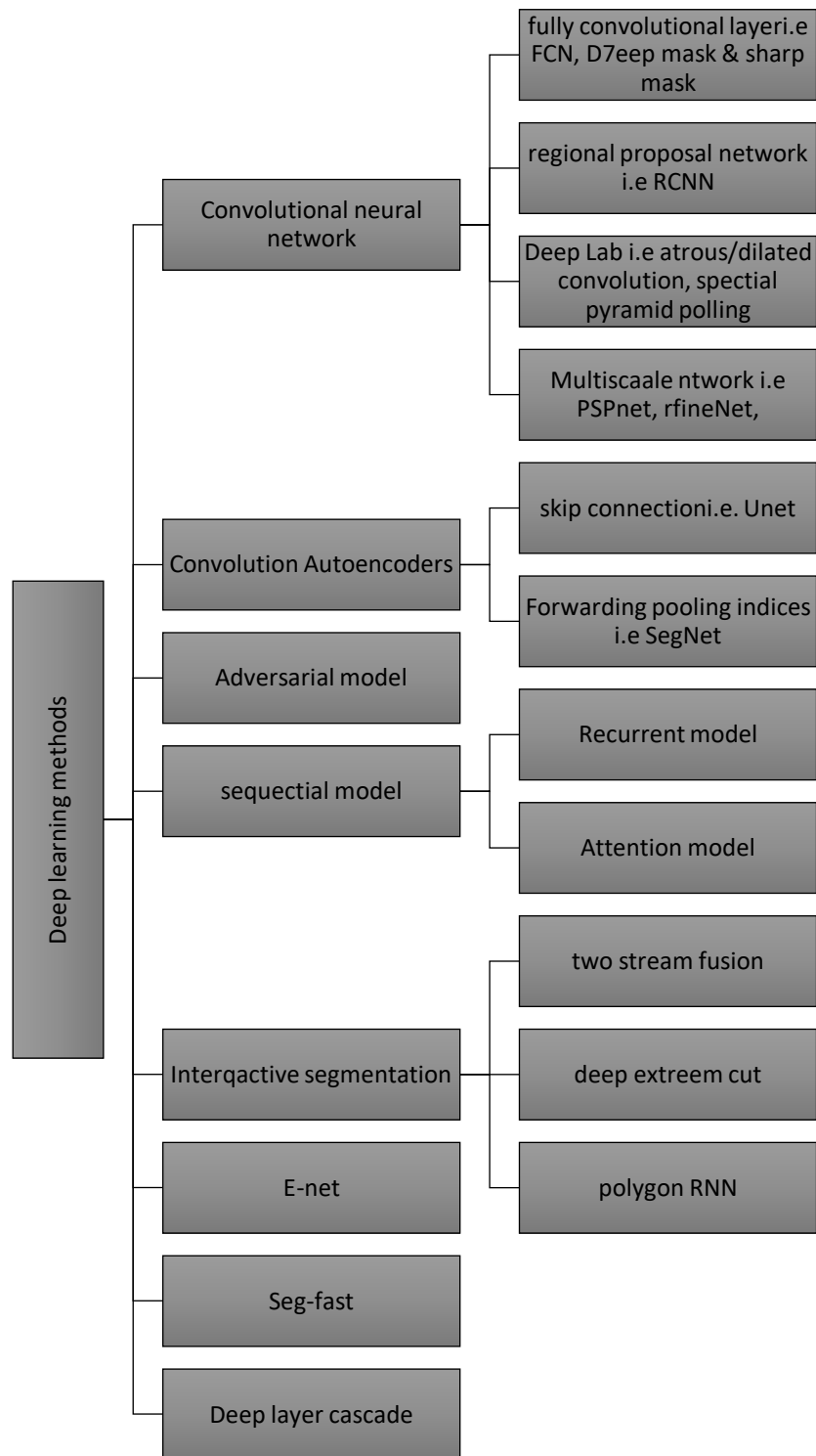


Figure 2-2 Deep learning methods for segmentation.

#### 2.4.2 Deep learning methods:

D. A. Gavrilov and et.al [47] develop a deep convolution neural network (DCNN) model based on Inception.v3 network approach for lesion classification. Their



trained model achieves 91% accuracy when compared to diagnostic result of qualified dermatologist.

Although deep learning methods have outperformed in automated lesion classification, However, unavailability of clinical data on vast and variety is somewhat causing to create bottleneck the pathway. To open this bottleneck, Balazs Harangi[1] proposed a collaborative use of deep neural networks to get advantage of individual exactitudes of deep networks to enhance the classification capability of skin lesions. He created amalgamate of four well known CNNs: GoogLeNet, AlexNet, ResNet, VGGNet to categorizing three different classes of lesion namely: melanoma, nevus, and seborrheic keratosis. He used pre-trained CNNs using ImageNet with maximum of 2000 images. Confidence value of each trained model was used to create fusion of these CNNs based on different ensemble approaches: Sum of the probabilities (SP), Product of the probabilities (PP), Simple majority voting (SMV) & Sum of the maximal probabilities (SMP). Simple Majority Voting (SMV) prove best of others. Weights of CNNs were optimized using Simulated Annealing (SA). These ensemble models have proven their performance better than individual CNN used for classification.

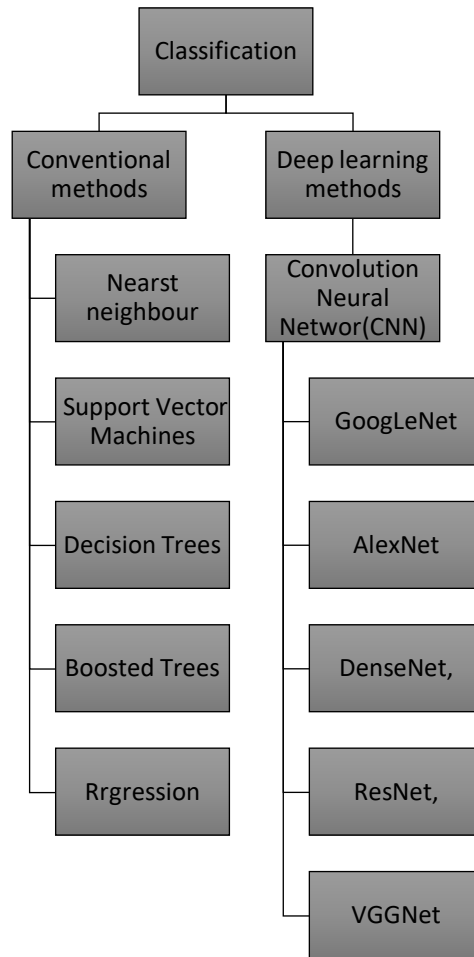
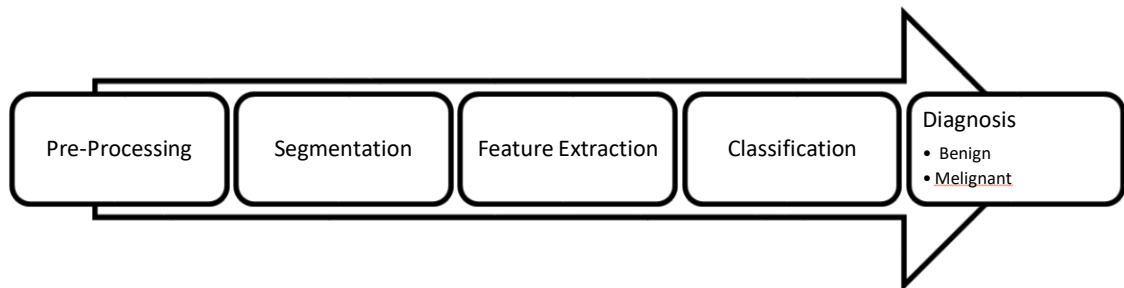


Figure 2-3: Conventional and Deep learning classification approaches

## **CHAPTER 3**

This chapter explained the proposed method used for skin cancer classification. It follows Multi-level Feature Fusion and selection framework for Skin lesion Segmentation and Classification. Following is the framework.



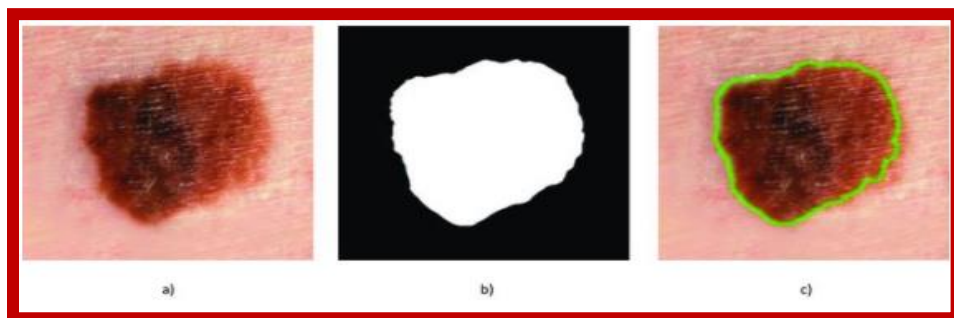
**Figure 3-1 Proposed Framework**

### 3.1 Image dataset collection:

For the proposed research I used clinical images from set of benchmark datasets including PH2, ISBI16, ISIC17-19, and HAM10000 both for training and testing.

### 3.2 Segmentation

The partitioning of an image into multiple segments is referred to as 'image segmentation for the Extraction of the attributes and features related to clinical image. With this in place, analyzing the image will be far easier. s. Segmentation of the lesions is done using deep CNN network i.e., Deeplabv3plus and Saliency estimation method to generate the mask to extract region of Interest (ROI). Segmented Mask then mapped to Enhanced RGB images to generate final ROI.



**Figure 3-2. Skin lesion segmentation. a) Dermoscopy image in input. b) Binary mask in output. c) The border of the mask overlaid on the enhanced image.**

The details of Proposed Segmentation Algorithm are given below.

### **3.2.1 Saliency detection-based Segmentation technique**

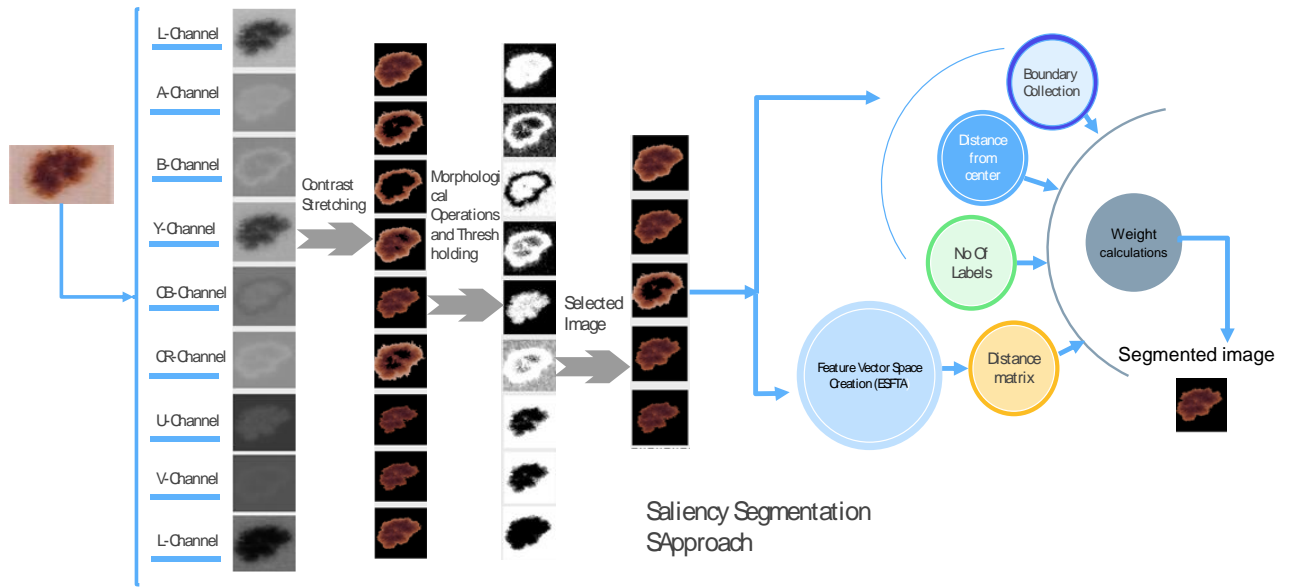
#### **3.2.1.1 Salient Regions:**

Those image Regions, that stand out visually due to their contrast with other neighboring regions are called salient. Recently medical imaging has included research on human visual system and applied in combination of neural networks to identify and extract Region of interest[48]. The detection of salient zones of an image proves to be healthy way to deal with the segmentation chore of classification problems[49]. Saliency algorithms develop saliency maps, which are aimed to depict the pixels' intensity values as an indication of locales with similar “salient” features, while comparable to other locales such as change of contrast. Many researchers have work out at segmentation through saliency maps[50][51]. Saliency estimation is based on various salient region properties[52] as mentioned below

1. Salient zone is distinct in colors and levels of contrast when compared to its neighboring regions.
2. When compared to the background, salient states have various textural features and location.
3. Extraneous features appear often in the salient region as compared to background.
4. They lie usually at the center of image.
5. They are homogeneous and non-recurrent.

#### **3.2.1.2 Saliency detection-based Segmentation technique:**

This section describes the proposed skeleton for Segmentation based on extracting the attributes of salient precincts as shown in Figure 3-3. Since in specific color spaces, significant regions are always visible at least in one color or more than one color channels of these spaces. Following this perspective, we choose three color modes ( $L^*ab$ ,  $L^*uv$ ,  $Y^*CbCr$ ), where each mode further comprises of three-color channels.



**Figure 3-3 Saliency estimation using information weight content**

The idea behind the proposed method mainly revolves around in finding the most relevant color channel with the most important salient regions, which we did by using weighting criteria built on global and spatial coordinates. First, the gray-level information histogram is used to maximize the contrast of all channels, followed by morphological procedures to highlight foreground stuffs and to drop background features. In second stage, a feature vector is constructed by calculating the Euclidean distance matrix using the suggested ESFTA algorithm, which determines the images' similarity with regard to determined features. Image features comparison, distance from the center, border connection, and number of connected labels are the key factors which are found to allot weights. Finally, we increase saliency by reducing background details and increasing foreground features.

Saliency based Segmentation is four-fold step. A detail of each step is given below.

#### **3.2.1.2.1 Image Enhancement:**

Image enhancement follows the following two step process

##### **A: Defining color Space:**

As discussed previously, significant regions are always visible in one or more color channels of the color spaces. For detecting salient region of an image, we must select the color

channels to be used throughout different image processing procedures to find the region of interest. Color space selection is critical step in saliency segmentation. Many researchers work out to find best color space for their image segmentation methods[53]. In this method we choose three color spaces (L'ab, L'uv, Y'CbCr). Color channels of each color space are treated as autonomous channel (i.e, L', a, b, L', u v, Y, Cb, Cr) for contrast stretching.

**B: Contrast stretching:**

In literature various Contrast stretching techniques have been deployed[54]. Skin lesion images have the challenges of identification of low contrast between affected area and healthy skin. Using gradient of an image we generate grey level information histogram Inspired by [55] . The output that was needed is enhanced details of the foreground (i.e., lesion region), and reducing the background detail from all the color channels. This fact is prevailed by high gradient areas which is object while gradient is low with the background.

**3.2.1.2.2 Enhancement and grey level histogram**

Enhancement and grey level histogram follow the below mentioned procedural steps.

1. Applying Gaussian filter: Gaussian filter is applied to smooth the image. Sobel Operator is applied to approximate Image gradient at each pixel. Sobel filter with the kernel size of 3x3, accounts for calculating the gradient of image intensity of each pixel resulting in image emphasizing edges. Gradient D at each pixel with spatial location (x,y) is defined as

$$M(x,y) = \sqrt{[D_r(x,y)]^2 + [D_c(x,y)]^2} \quad (3.1)$$

Where,  $D_r(x,y)$  represents row gradient and  $D_c(x,y)$  shows column gradient.

2. Next, to develop grey-level histogram, we set four threshold limits of the gradient intensities of pixels of image to define gradient intervals. Now the image is divided into equal size regions (i.e., N= 4, 8, 12,) based on ascending order of pixel gradient intensities lying in already set gradient intervals. All

pixels laying in a particular gradient interval, are assigned weight accordingly as follows

$$W_D(x, y) = \begin{cases} w_{r1} & M(x, y) \leq T_1 \\ w_{r2} & T_1 < M(x, y) \leq T_2 \\ w_{r3} & T_1 < M(x, y) \leq T_N \\ w_{r(N+1)} & otherwise \end{cases} \quad (3.2)$$

Where  $T_1$  through  $T_N$  are threshold values for defining gradient interval. Based on these Gradient Intervals,  $w_{ri}$  ( $i= 1 \dots N$ ) is statistical coefficient assigned to each pixel.

#### a. Weighing criteria

These weights are assigned based upon following two factors naming

##### I. Defining Region size

##### II. Finding Maximum information region

Following the fact describes in [55], region of interest is usually 20-80 % of the image area. So at least lowest limit should be followed while choosing number of regions, to make histogram insensitive to No. of regions. Where, higher No. of region may cause complex computations. By trading off these facts, we divide image into eight groups in the proposed method.

As described earlier, Region with high intensity values is contributing to ROI. So, region of maximum information is extracted using Laplacian of a Gaussian (LOG). Regions with more edge information and hence more pixel count points to maximum information region. We already following this fact in dividing region: regions are grouped in ascending order of gradient magnitude. This shows the region N will have maximum information.

Utilizing the above two-point information, weights are assigned according to maximum No of edge points E (i.e., Maximum information region N) as follows

$$W_{ri} = \frac{E_i}{E_N} \quad i = 1, 2, \dots, N \quad (3.3)$$



Where  $E_i$  represents Number of edges of  $i^{\text{th}}$  region and  $E_N$  is number of edge points of maximum information region i.e., N. This weighting criteria tends to enhance the grey level histogram generation process.

3. Now, for each of eight regions evaluate cumulative sum of all weighted grey values as follows

$$G_l(W) = \sum_{i=1}^N w_{ri} n_{ri}(w) \quad (3.4)$$

Where,  $n_{ri}$  is number of those pixels in a region  $ri$  having grey level value define with statistical weight  $w_{ri}$ .  $G_l(W)$  Shows cumulative sum of grey levels of all regions.

### **b. Thresholding**

To extract region of interest, Otsu's thresholding is carried out. Before thresholding is done morphological operations i.e., dilation & erosion, opening and closing[56] are utilized to prominent the foreground from background. Let  $M(x,y)$  is grey scale image and  $S(m,n)$  is structural element then dilation and erosion are defined as follows

#### **Dilation:**

Dilation tends to increase grey level information of resultant image.

$$(M \oplus S) = \max_{m,n} (M(x - m, y - n) + S(m, n)) \quad (3.5)$$

$$= \max_{m,n} (M(x - m, y - n)) \quad \text{if } S(m, n) = 0 \forall (m, n) \quad (3.6)$$

#### **Erosion:**

Erosion works exactly opposite to Dilation while produce same results

$$(M \ominus S) = \min_{m,n} (M(x + m, y + n) + S(m, n)) \quad (3.7)$$

$$= \min_{m,n} (M(x + m, y + n)) \quad \text{if } S(m, n) = 0 \forall (m, n) \quad (3.8)$$

#### **Opening:**

Opening and closing are called opening & closing by Morphological reconstruction.

Using above mention dilation and erosion opening is define as

$$M \odot S = ((M \ominus S) \oplus S) \quad (3.9)$$

**Closing:**

$$M.S = ((M \oplus S) \ominus S) \quad (3.10)$$

### 3.2.1.3 Feature extraction

Next is proposed extended segmentation method based on fractal texture analysis (ESFTA). To extract salient features, we pay attention to textural characteristics so to diminish background and enhance foreground. To extract the features using ESFTA, first we divide the images into stack of binary images. To divide the image into binary we utilized proximal plane clustering[57].

#### a. Proximal plane clustering (PPC):

For the image  $M$ , which need to be divided into  $k$  clusters around the clustering hyperplane. Programming problem of Proximal plane clustering becomes

$$\min_{w_i, b_i \neq 0} \|M_i w_i + b_i e\|_2^2 - c \|B_i w_i + b_i e\|_2^2 \quad \text{s.t } \|w_i\|_2^2 = 1 \quad i=1, 2, \dots, k \quad (3.11)$$

Where  $M_i$  is the  $i$ th cluster of image data points and  $B_i$  are other data points,  $c$  is weighting term,  $\|\cdot\|_2$  is the  $L_2$  norm,  $e$  represents vector of ones. After  $k$  cluster hyperplane is attained the allocation of data point  $x$  belongs to  $i$ th cluster is given as below

$$t_r = \arg \min_i \|w_i^T + b_i\|, \quad \text{where } i = 1, 2, \dots, k \quad (3.12)$$

For the given image data points  $M$ ,  $k$ -plane clustering initializes randomly ( $w_i, b_i$ ) &  $i = 1, 2, \dots, k$  with  $\|w_i\| = 0$

After clustering is done next step is to set pair of threshold values from each cluster i.e.,  $T = [t_1, t_2]$

$t_1$  and  $t_2$  are lower and upper limits to define threshold interval for each cluster.

Stack of binary image will be as follows

$$M_b(x, y) = \begin{cases} 1 & \text{if } t_1 < M(x, y) \leq t_2 \\ 0 & \text{otherwise} \end{cases} \quad (3.13)$$

#### **ESFTA extraction algorithm:**

Binary image stacked from clustering will be used to initiate feature extraction. Grey level of binary image along with pixel count and boundary fractal components are used to extract features. The boundaries of a binary image's regions are represented as border image as below

$$\Delta(x, y) = \begin{cases} 1, & \text{if } (x', y') \in N_8[(x, y)]; \\ M_b(x', y') = 0 \quad \wedge \quad M(x, y) = 1; \\ 0, & \text{otherwise} \end{cases} \quad (3.14)$$

$N_8[(x, y)]$  shows 8-connected pixels.  $\Delta(x, y)=1$ , if pixel at spatial position  $(x, y)$  of  $M_b(x, y)$  has Value =1 or at least one zero at its neighboring pixel location. Linear time box-counting algorithm is used to find fractal dimensions [58]. The Hausdorff's fractal dimension  $D$  is calculated as below

$$D_0 = \lim_{\epsilon \rightarrow 0} \frac{\log N(\epsilon)}{\log \epsilon^{-1}} \quad (3.15)$$

Where,  $N(\epsilon) = \text{numbe of image regions (atleast one pixel)}$   $\epsilon = \text{length of unit region}$ .

Following steps are followed for box-counting algorithm.

1. Image division was created with a square grid of size  $\epsilon \times \epsilon$ .
2.  $N(\epsilon)$ , counting the number of blocks where at least one pixel is present
3. Grid size is varied to plot  $\log N(\epsilon)$  vs  $\log \epsilon^{-1}$
4. The best fit is achieved by approximating the curve using a linear minimal mean square error estimator (MMSE).

#### **Weight Calculation:**

In Fig. 2 shows proposed framework. We picked three color spaces with nine color channels, and each with a different weighting. Feature vector computation excluded four color channels due to maximum and lowest detail levels. We use five remained color channels to determine final segmented image where, we use the distance matrix,

boundary connections, distance from the center, and the number of linked components to get weights.

### **Boundary connection:**

To find boundary connection, let image border that is unit pixel wide having the mask as  $M_{b,mask}(x_{[i]}, y_{[j]}) = 1$ . And  $M_{b,mask}(x_{[i']}, y_{[j']}) = 0$  for the pixel excluded from image border. For Binary image  $S_b(x, y)$ , there exist boundary connection between image pixel and mas pixel if  $(M_{b,mask}(x_{[i]}, y_{[j]}) \cap S_b(x_i, y_i)) = 1$ . The weight is then calculated as

$$w_{bc} = \frac{1}{\sum_{c=1}^n n_c} \quad (3.16)$$

$n_c$  is number of common pixels to binary image and border.

### **Connected components:**

According to the assumption stating that there must be a minimum number of linked components in a color channel with noticeable salient areas. From this premise, we may derive the theory of linked components labeling, which asserts that when two points p and q are deemed to be adjacent, eight-connected paths must exist between them (i.e.,  $q \in N_8(p)$ ). Thus weight has been determined to be the following

$$w_{conn} = \frac{1}{\sum_{i=1}^k C_i} \quad C_i = \text{connected component labels} \quad (3.17)$$

### **Distance from center:**

Additionally, when it is well known that objects of interest are generally located near the center, we use this assumption and apply weights by calculating the connection between a Gaussian mask with the center  $G(\delta_x, \delta_y)$  where,  $\delta_x = (No. of rows)/8$ , and  $\delta_y = \frac{No. of coloumn}{8}$ , with the maximum sized blob.

$$W_{cp} = G\left(\delta_x, \delta_y \cap \arg \max_{l \in [1, k]} \sum_{i=1}^k c_{max}\right) \quad (3.18)$$

Where,  $c_{max}$  represents labels with maximum size connection.

Finally, we use the ESFTA technique to generate the Euclidean distance matrix using the feature vector. Those images with the most repetition is found using the mode operation. For example, if the third image has the shortest distance with the fourth

image out of the five, image four will be assigned weight 2. The closer an image is to the minimum distance, the more weight it will receive. This condition leads to choose Maximum repeated image or maximum weighted image. The total weight of all four conditions is computed as follows:

$$w_{final} = w_{bc} + w_{conn} + w_{cp} + w_{rept}^{\max} \quad (3.19)$$

The  $w_{final}$  will choose the color channel ( $\max w_{final}$ ) to proceed with segmentation.

### Salient information improvement:

We present a new method to enhance salient information extraction which tends to improve and strengthen the results of saliency identification. We estimated the geometric center or centroid of the item using the segmentation result of cumulative weighting criterion and applied Gaussian as below

$$M_{Enh} = 1 - \exp\left(\frac{G(\sigma_{Ex}, \sigma_{Ey})}{\max(M_{seg}(x, y))}\right) \quad (3.20)$$

Where  $M_{seg}$  is the normalized segmented grey image and,  $G(\sigma_{Ex}, \sigma_{Ey})$  is a Gaussian with the object's mean as the center and a variance of one.

### 3.2.2 CNN based Segmentation: Deeplabv3Plus

In the current era, Deep learning's incredible breakthroughs, mainly CNN, have provided an immense boost to the development of visual object recognition. In this section we explore CNN architecture i.e., Deeplabv3Plus based segmentation method for lesion detection and classification.

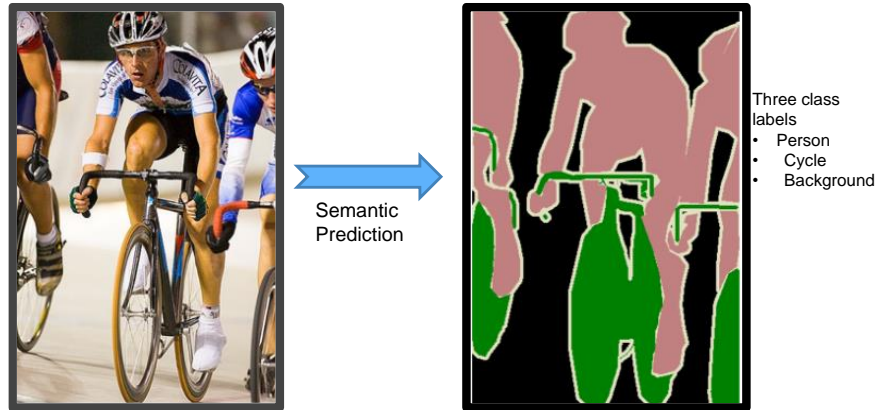


Figure 3-4 Semantic segmentation (source<sup>4</sup>)

<sup>4</sup> robot.ox.ac.uk

The implementation of the semantics segmentation technique we are working with is known as Deeplabv3. In this approach, every pixel in the image belongs to one of several classes, which include persons, trees, buildings, windows, etc. Color assignment is applied to all pixels that correspond to a certain class. Each class is thus represented by different color. This type of Classification of every pixel in an image is called semantic segmentation.

Recently, like in other domains of life ,CNN has been extensively applied in medical image analysis research [59]. Interest in deep learning segmentation techniques for segmentation, object detection, and image classification has recently increased [60]. CNNs (also known as deep artificial neural networks) are an increasingly common form of deep artificial neural network. CNNs (as well as other state-of-the-art neural networks) perform better in numerous machine learning applications, which is why CNNs outperform other state-of-the-art neural networks[59][61].

Many other CNN architectures have also been utilized for image segmentation i.e., R-CNN, Mas RCNN, U-net, YOLO etc. Deep Lab is a cutting-edge semantic segmentation model that was released in 2016 by Google, who released it as open-source software. Since the model has been through several upgrades, such as DeepLab V2, DeepLab V3, and the current DeepLab V3, many new features have been included. Now that we've analysed the DeepLab V3 architecture, we will learn how to apply it to our dataset.

### **3.2.2.1 Deeplabv3+ Architecture:**

DeepLab is a two-step process.

#### **3.2.2.1.1 Encoding Phase:**

This phase is all on discovering the key features of the image. For a more complete explanation, this is accomplished via a pre-trained Convolutional Neural Network (CNN). Convolutional layers pass feature information from one layer to the next. This is an effective strategy for image classification and is especially useful for finding objects in images.

#### **3.2.2.1.2 Decoding Phase:**

In this process, the information that was retrieved in the encoding step is employed to rebuild the output.

Main architecture is shown as follows

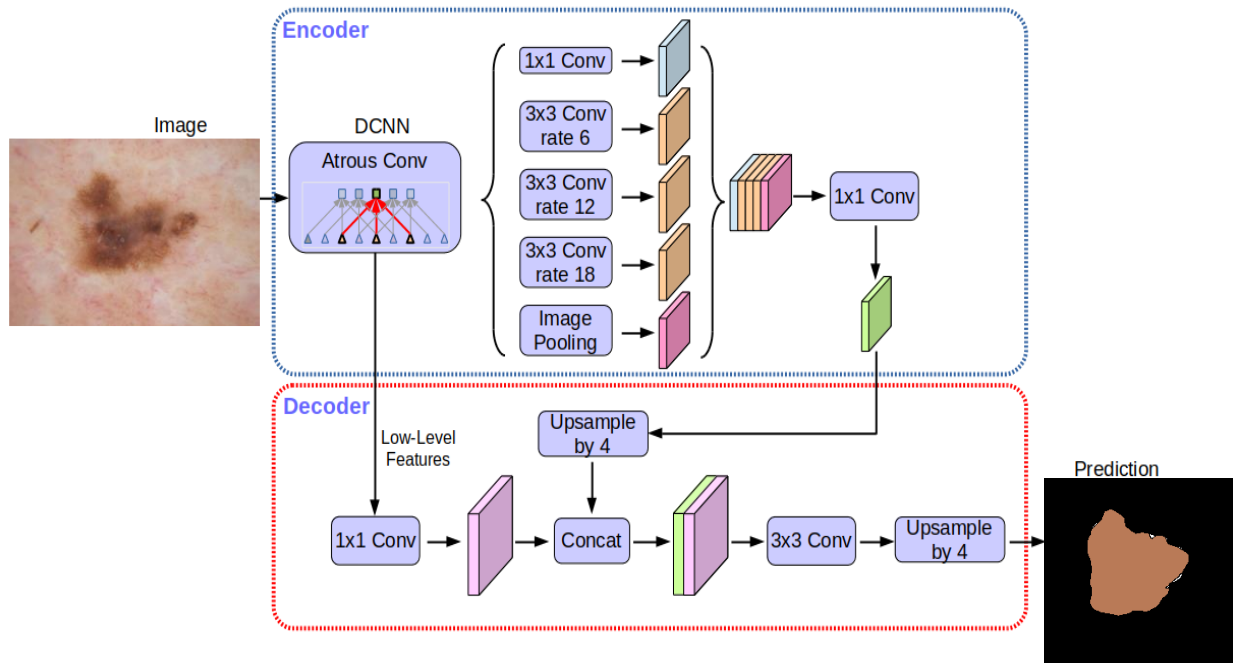
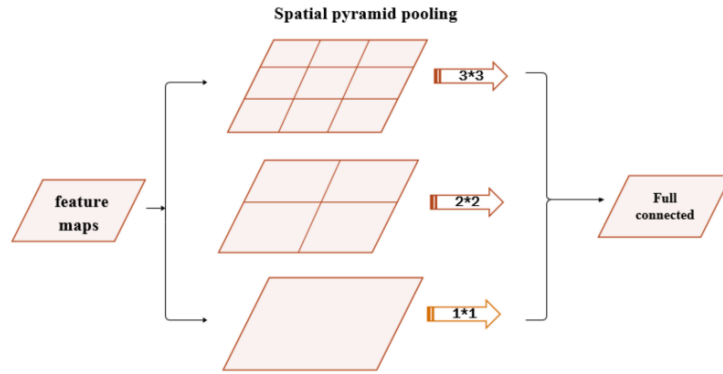


Figure 3-5 Deeplabv3 architecture[62]

The DeepLab architecture is based on combining two popular neural network architectures:

- **Spatial Pyramid Pooling**
- **Encoder-decoder networks**

Using spatial pyramid pooling networks will help address customization of data used for problem. Multiple versions of the input are used for training, and that means multi-scale information is considered i.e., pyramid pooling. Spatial pyramid pooling networks generally use parallel versions of the same underlying network to train on inputs at different scales and combine the features at a later step.

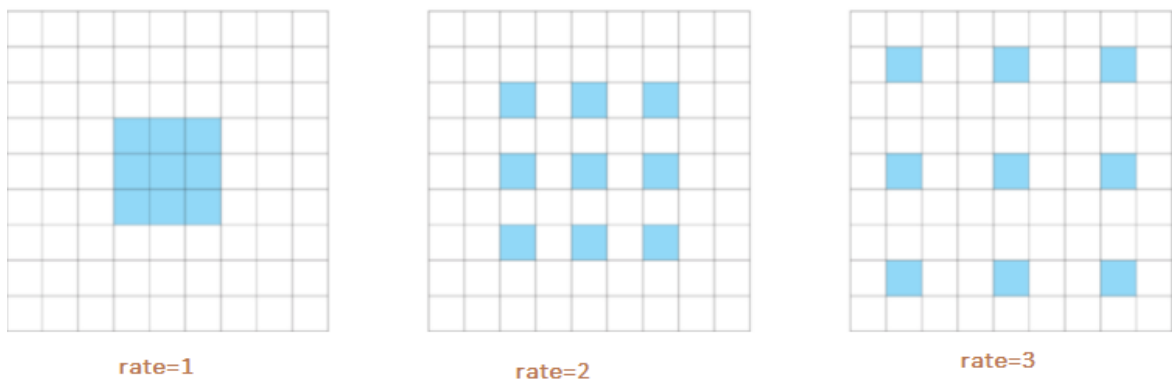


**Figure 3-6 Pyramid pooling**

Several instances of the same design are used in partial pyramid pooling, increasing the computational complexity and memory needs of training increases training costs. Google has solved the problem by introducing the concept of atrous convolutions. This is a generalized form of the convolution operation. Atrous convolutions need a parameter named : rate, which is used to explicitly control the effective field of view of the convolution. The generalized form of atrous convolutions is given as:

$$y[i] = \sum_k x[i + r \cdot k]w[k]$$

For each location  $i$  on the output  $y$  and a filter  $w$ , atrous convolution is applied over the input feature map  $x$  where the atrous rate  $r$  corresponds to the stride with which we sample the input signal.



**Figure 3-7 atrous convolution**



the normal convolution is a special case of atrous convolutions with  $r = 1$ . DeepLab uses atrous convolution with rates 6, 12 and 18. Therefore, since atrous convolutions have a wider effective field of vision than other methods while utilizing the same amount of computing resources, they are advantageous.

DeepLab V3 uses ImageNet's pretrained Resnet-50 with atrous convolutions as its main feature extractor. In the modified ResNet model, the last ResNet block uses atrous convolutions with different dilation rates. It uses Atrous Spatial Pyramid Pooling and upsampling for the decoder module on top of the modified ResNet block.

### 3.2.3 Image Fusion Framework.

A contrast-reinforced image fusion framework is employed that consumes resultant contrast enhanced RGB images generated from the saliency approach and Deeplabv3plus network. The contrast stretched grey images from both saliency techniques, & Deeplabv3plus are used as inputs in the proposed contrast-reinforced image fusion architecture. The final binary image is generated by the fusion framework using direct correlation of both images. The structuring filter traverses the image based on an initial set of parameters to determine the mean and variance of each block, and then assigns a weight value to each block based on the selected parameters. The blocks with the highest weight are then fused to produce the final fused binary image. As a result, we present multi-modal picture fusion methods in this paper; its strength is that it preserves structural information while being simple to implement. It is explained in the following section.

A Gradient of the image is generated to locate edges that occur in different image sections. To spot texture and boundary details, we use the gradient magnitude. Assume we have two contrast stretched images  $M_1(x, y)$  &  $M_2(x, y)$ , the magnitude of gradient in coordinates  $(x, y)$  is determined using Eq. (3-23) for both images as below

$$\|\nabla M(x, y)\| = \sqrt{\left(\frac{\partial M}{\partial x}\right)^2 + \left(\frac{\partial M}{\partial y}\right)^2} \quad (3-23)$$

Where  $\left(\frac{\partial M}{\partial x}\right)$  calculated as

$$\frac{\partial M}{\partial x} = \frac{\partial M(x,y)}{\partial x} = M(x+1, y) - M(x, y) \quad (3-24)$$

As in later the component operation's postprocessing, the saliency map is normalized. For the reason Gradient  $\nabla M(x, y)$  is needed to normalize. Normalization is done using the linear transformation function as demonstrated in Eq. 3-25

$$M'(x, y) = \frac{\nabla M(x, y) - \min(\nabla M(x, y))}{\max(\nabla M(x, y)) - \min(\nabla M(x, y))} \quad (3-25)$$

The saliency operations may result in gaps and holes, morphological Operation are used next. The morphological closing operation is used to fill those gaps and holes in homogeneous portions of images. Closing by Morphological reconstruction based on Erosion and Dilation[56] is depicts in following Eq. (3-26)

$$M_m = (M' \oplus S) \odot S \quad (3-26)$$

Where S is Structural element and  $M_m$  is resultant image after morphological closing. The gradient magnitude is large if the pixel value is decisive in reflecting the scene, and low if the pixel value is irrelevant. The weighted map is created by comparing the saliency map using the following formula

$$\bar{\omega} = [\beta'_1, \beta'_2] \quad (3-27)$$

If the corresponding element value of  $\beta'_1$  is larger than  $\beta'_2$ , returns 1 for element of  $\bar{\omega}$ ; else, returns 0. Since the linear filter's (box filter's) input is with window size  $\rho$ :

$$\bar{w} = \text{boxfilter} \left[ \frac{\bar{\omega} u}{\rho} \right] \quad (3-28)$$

Now, mean ( $\mu$ ) and variance ( $\sigma^2$ ) are calculated as shown in below equations

$$\mu = \frac{\text{boxfilter}(\bar{w}, u)}{\rho} \quad (3-29)$$

$$\sigma^2 = \frac{\text{boxfilter}(\bar{w} * \bar{w}, u)}{\rho} - \mu * \mu \quad (3-30)$$

Using mean and variance weight maps are calculated as follows

$$\hat{w} = \mu + \frac{\sigma^2}{\sigma^2 + \gamma} * (\bar{w} - \mu) \quad (3-31)$$

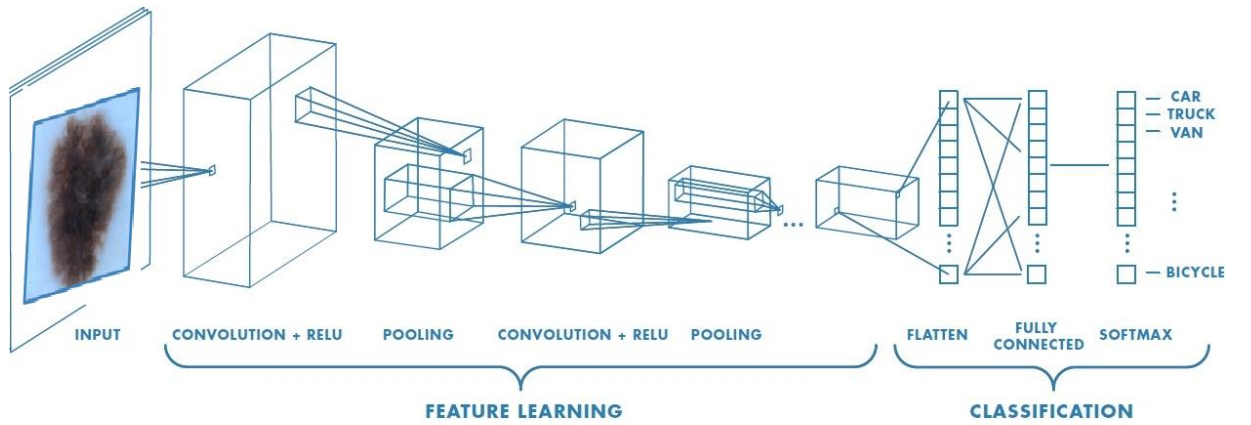
Finally, applied weight are applied to obtain fuse image as follows

$$M_f(x, y) = \hat{w}M_{m1}(x, y) + (1 - \hat{w})M_{m2}(x, y) \quad (3-32)$$

### 3.3 Transfer Learning, Feature selection and feature vector optimization.

Quality of features extracted directly affects the classification accuracy. Features will be extracted using transfer learning. Transfer learning is an approach used in a process of training a model, where a model already trained for a task is utilized as the starting point for the training of new model to be applied on some other related tasks.

Usually, classification DCNNs have four main operations. Convolutions, activation function, pooling, *and* fully connected layers. Passing an image through a series of these operations outputs a feature vector containing the probabilities for each class label. Note that in this setup, we categorize an image either nevus or malignant lesion type. That is, we assign a single label to an entire image.



**Figure 3-8 DCNN's structure (Source: MathWorks)**

Inception-v3 and Inception-Resnet V2 is used to extract feature vector. Feature Vector from each pre-trained CNN model is then serially fused to generate final feature vector. Which is then followed by optimization algorithm to optimize feature vector length before feeding into classifier. We used binary Grey wolf

optimization followed by Entropy-controlled feature selection. Following is the description of Grey Wolf Optimization Technique.

### 3.3.1 Grey Wolf Optimization

Grey Wolf Optimization(GWO) is a modern novel metaheuristic optimization technique initially presented by Mirjalili in 2014.[63]. GWO has the advantages of being easy to use, stable and converges quickly, and being useful for approximating the global optimal solution.

GWO resembles the predatory behavior of grey wolves in haunting. In Grey Wolf pack, the wolves are separated into four different crowds, namely Alpha ( $\alpha$ ), Beta( $\beta$ ), Delta( $\delta$ ), and Omega( $\omega$ ). Alpha wolf is the main leader of the pack, which makes decisions and leads the haunting Process. Beta is the second in command who aids the alpha wolf in making decisions and various other duties. The third-in-command are Delta's wolf. Deltas are responsible for many duties lie care taking, watching the territory, and securing the pack. The rest belong to omega group whipped by other dominating groups in their social hierarchy.

Haunting process is a three-step process, Searching, Encircling, and attacking. Alpha leads the haunting, with the help of beta and delta while, Omega follows all dominating groups. The mathematical model of haunting is described by following equations (1)-(4).

$$\vec{W}(t+1) = \vec{X}_p(t) + \vec{A} \cdot \vec{D} \quad (3.33)$$

Where  $\vec{W}$  shows Grey Wolf position.  $X_p$  Denotes position of prey,  $\vec{A}$  &  $\vec{D}$  are co-efficient vectors defined in eq (3) & (4).  $\vec{D}$  is defined by equation (2), and t shows current iteration

$$\vec{D} = |\vec{C} \cdot \vec{X}_p(t) - \vec{W}(t)| \quad (3.34)$$

$$\vec{A} = 2a \cdot \vec{r}_A - a \quad (3.35)$$

$$\vec{C} = 2\vec{r}_C \quad (3.36)$$

Where ‘a’ is encircling coefficient and get reductions linearly from 2 to 0.  $\vec{r}_A$ ,  $\vec{r}_C$  are random vectors  $\in [0, 1]$ .

Alpha, beta, and delta wolves are thought to be most cunning and have a better understanding of the surroundings of apparent prey positions in GWO. As a result, the omega wolves are being directed to the best possible location by the leaders. The new updated position of the wolf is mathematically shown in Equation (3.37)

$$\vec{W}(t + 1) = \frac{\vec{W}_1 + \vec{W}_2 + \vec{W}_3}{3} \quad (3.37)$$

Where  $\vec{W}_1$   $\vec{W}_2$   $\vec{W}_3$  are delimited by equation 6,7,8 in terms of first obtained best three solution at  $t$  iteration in the Swarm i.e.  $\vec{W}_\alpha$   $\vec{W}_\beta$   $\vec{W}_\delta$

$$\vec{W}_1 = |\vec{W}_\alpha - \vec{A}_1 \cdot \vec{D}_\alpha| \quad (3.38)$$

$$\vec{W}_2 = |\vec{W}_\beta - \vec{A}_2 \cdot \vec{D}_\beta| \quad (3.39)$$

$$\vec{W}_3 = |\vec{W}_\delta - \vec{A}_3 \cdot \vec{D}_\delta| \quad (3.40)$$

$\vec{D}_\alpha$   $\vec{D}_\beta$   $\vec{D}_\delta$  are also defined in terms of best three solution equation (3.28) (3.29) & (3.30)

$$\vec{D}_\alpha = |\vec{C}_1 \cdot \vec{W}_\alpha - \vec{W}| \quad (3.41)$$

$$\vec{D}_\beta = |\vec{C}_2 \cdot \vec{W}_\beta - \vec{W}| \quad (3.42)$$

$$\vec{D}_\delta = |\vec{C}_3 \cdot \vec{W}_\delta - \vec{W}| \quad (3.43)$$

And finally, the coefficient a is updated as in equation (3.31)

$$a = 2 - 2 \frac{t}{T} \quad (3.44)$$

T is showing maximum number of iterations allowed.

### Binary GWO:

GWO is commonly used to solve problems that require constant optimization. It is necessary to model a binary version of GWO for binary optimization such as feature selection. Emary et al. [37] proposed two binary grey wolf optimization approaches, bGWO1 and bGWO2, respectively.

### bGWO1:

For the bGWO1, Crossover operator is used to update and reframe the wolf position as follows

$$W(t + 1) = Crossover(w_1, w_2, w_3), \quad (3.45)$$

binary vectors  $w_1, w_2$ , and  $w_3$  showing the effect of wolf's movement with respect to  $\alpha, \beta$ , &  $\delta$  grey wolves, respectively.  $Crossover(w_1, w_2, w_3)$  is the fitting crossover between solutions. Equations (14), (15) and (16), are used in bGWO1 to define  $w_1, w_2$ , and  $w_3$ .

$$w_1^d = \begin{cases} 1 & \text{if } (w_\alpha^d + bstep_\alpha^d) \geq 1 \\ 0 & \text{otherwise} \end{cases} \quad (14) \quad (3.46)$$

Where  $w_\alpha^d$  denotes the alpha's location through dimension d, and  $bstep_\alpha^d$  denotes a binary step through dimension d that can be determined by equation (3.36)

$$bstep_\alpha^d = \begin{cases} 1 & \text{if } cstep_\alpha^d \geq r1 \\ 0 & \text{otherwise} \end{cases} \quad (3.47)$$

Where  $r1$  is a random number  $\in [0, 1]$ , and  $cstep_\alpha^d$  is the step size which is continuous valued through dimension d, which may be determined via sigmoid function in equation 16.

$$cstep_\alpha^d = \frac{1}{1 + e^{-10(A_1^d D_\alpha^d - 0.5)}} \quad (3.48)$$

Where  $A_1^d, D_\alpha^d$  are calculated in the dimension d using already defined equations (3) and (9). Similarly binary vectors  $w_2$ , and  $w_3$  are defined by equation (17) through equation (22) in dimension d as follows

$$w_2^d = \begin{cases} 1 & \text{if } (w_\beta^d + bstep_\beta^d) \geq 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.49)$$

$$bstep_\beta^d = \begin{cases} 1 & \text{if } cstep_\beta^d \geq r2 \\ 0 & \text{otherwise} \end{cases} \quad (3.50)$$

$$cstep_\beta^d = \frac{1}{1 + e^{-10(A_1^d D_\beta^d - 0.5)}} \quad (3.51)$$

$$w_3^d = \begin{cases} 1 & \text{if } (w_\delta^d + bstep_\delta^d) \geq 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.52)$$

$$bstep_\delta^d = \begin{cases} 1 & \text{if } cstep_\delta^d \geq r3 \\ 0 & \text{otherwise} \end{cases} \quad (3.53)$$

$$cstep_\delta^d = \frac{1}{1 + e^{-10(A_1^d D_\delta^d - 0.5)}} \quad (3.54)$$

Where  $A_1^d D_\beta^d$  and  $A_1^d D_\delta^d$  are calculated in the dimension d using already defined equations (3) and (9).

The wolf's new position is updated because of the crossing procedure, once the crossover candidates i.e., binary vectors  $w_1, w_2$ , and  $w_3$  are obtained. Where  $d$  is the search space dimension.

$$W^d(t+1) = \begin{cases} w_1^d & \text{if } r4 < \frac{1}{3} \\ w_2^d & \frac{1}{3} \leq r5 < \frac{2}{3} \\ w_3^d & \text{otherwise} \end{cases} \quad (3.55)$$

Where  $r2$  through  $r5$  is a random number  $\in [0, 1]$ .

### **bGWO2:**

In the BGWO2 technique, the wolf's location is updated by converting its position into a binary vector, i.e. only the updated grey wolf position vector is compelled to be binary, as indicated in Equation (24).

$$w_d^{t+1} = \begin{cases} 1 & \text{if } \text{sigmoid}\left(\frac{w_1 + w_2 + w_3}{3}\right) \geq rand \\ 0 & \text{otherwise} \end{cases} \quad (3.56)$$

Where random number  $rand$  is  $\in [0, 1]$  and,  $d$  is the dimension of search space, and sigmoid is defined through equation (25).

$$\text{sigmoid}(w) = \frac{1}{1 + e^{-10(w-0.5)}} \quad (3.57)$$

### **Entropy-Controlled Feature selection:**

Feature selection is one of the most important steps in classification. Most relevant features are needed to select to reduce the complexity and computation overhead. Entropy based feature selection is proposed in combination with GWO. Let fused feature vector after transfer learning is  $FL$ . Initially  $FL$  may have irrelevant and redundant features. To select most distinguished features, concept of Entropy is utilized. Let  $FR$  is represented as follows

$$FL = \{(f_1, l_1), \dots (f_k, l_k), \dots \dots (f_N, l_N)\} \quad (3.58)$$

where,  $F \in \{f_i\}_{i=1}^N \in \mathbb{R}$  is a feature vector with the labels as  $L = \{l_i\}_{i=1}^N$  are labels. As classification problem is binary problem so  $l \in [0,1]$ . Feature selection is two step procedures.

1. Entropy calculation
2. Features with maximum entropy are selected.

The Entropy for feature space  $\varphi$  is shown as follows

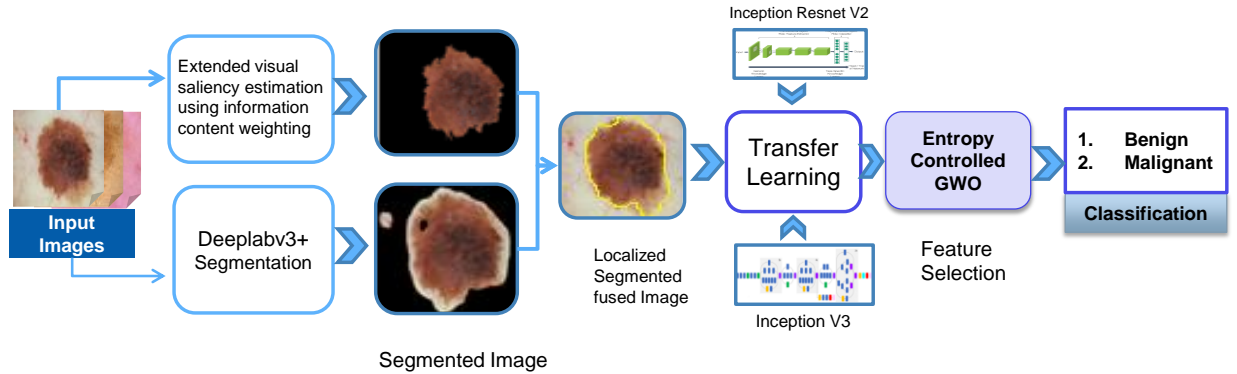
$$\tilde{E}(F) = - \sum_{i=1}^N P(f_i) \log P(f_i) \quad (3-59)$$

Where,  $P(f_i)$  is probability of  $i$ th feature to be selected.

Following is flow chart of proposed Entropy controlled grey wolf Optimization.

### 3.4 Classification.

Three Classifiers are reported based on their performance evaluation, Cubic SVM, Linear SVM, Cosine KNN.



**Figure 3-9: Detailed flow diagram for Lesion classification**



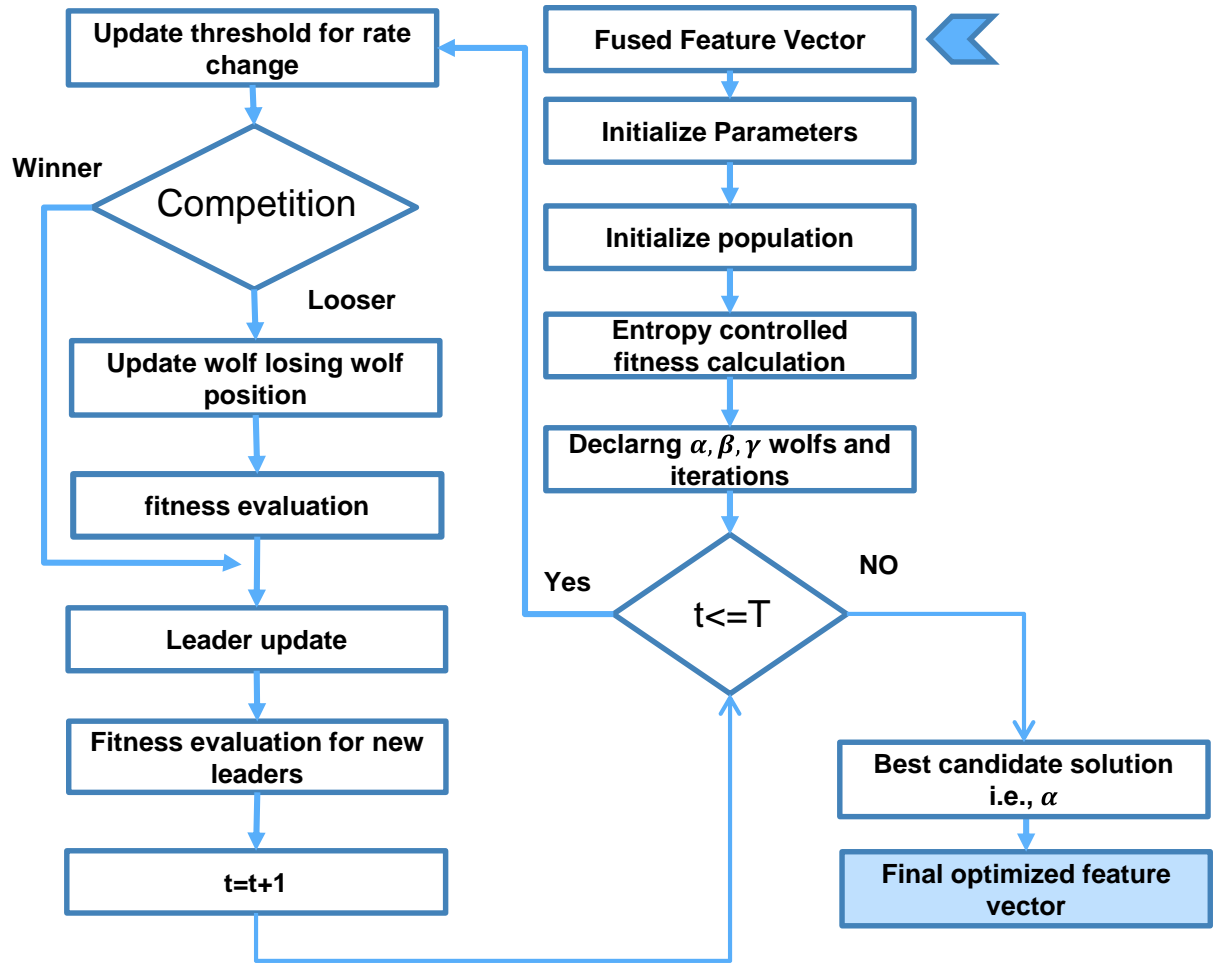


Figure 3-10: Entropy-Controlled binary GWO Feature Selection

## **CHAPTER 4**

In this chapter, we describe the experimental setup used, performance metrics and present our results and findings. The discussion includes Comparative analysis with other related work.

### **Simulation and Results:**

The proposed framework incorporates four different modules starting from Pre-processing, then segmentation, feature extraction and finally classification. For this we shall be utilizing clinical images from set of benchmark datasets including *PH<sup>2</sup>*, ISBI 2016, and ISIC 2017, both for training and testing. The main objective of proposed plan is: 1) segmentation; 2) Classification. At first, we apply Salient estimation based segmentation to extract ROI. Then extracting (ROI) using Deep Networks Deeplabv3+ method which then combined to produce final fused images for feature extraction. In the third phase, features will be extracted after applying Transfer Learning using multiple pre trained models such as Inception-v3, Inception-Resnet V2. Feature vectors were optimized using biological inspired Algorithm i.e., Grey Wolf Optimization (GWO) Techniques. These features, after fusion fed into different classifiers i.e., Support Vector Machines (SVM), K-Nearest Neighbor (K-NN) and Convolutional Neural Network (CNN).

To validate the suggested strategy, three publicly available benchmark datasets are simulated to determine the results. Simulations are performed in MATLAB 2020b and 2021a. Two performance metrics including DICE index, and Jaccard Index are employed to calculate the segmentation accuracy. Sensitivity, specificity, precision, FNR, AUC and FPR are used to test classification results, about 80% of the data is utilized for training in all three datasets, whereas 20% of the data is used for testing. The technique of ten-fold cross validation is applied to ensure the categorization findings are correct.


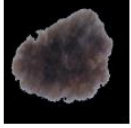
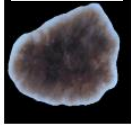

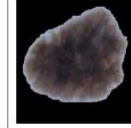
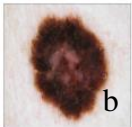
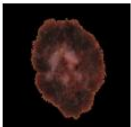
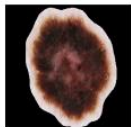
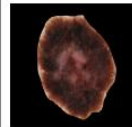
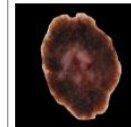






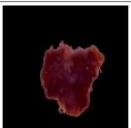
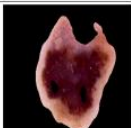


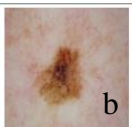




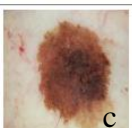
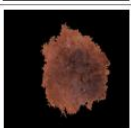
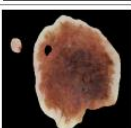






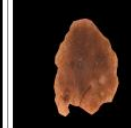
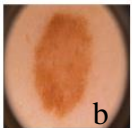




Dataset	Original image	Weighted Saliency	Deepla bv3+	Mask Fusion	Ground Truth
ISBI2016	 a				
	 b				
	 c				
ISIC2017	 a				
	 b				
	 c				
PH <sup>2</sup>	 a				
	 b				

Figure 4-1 Segmentation results

#### 4.1 Performance evaluation metrics

The Performance parameters are described below.

##### 4.1.1 Dice:

Recalling from Chapter 3, using expression of final fused image Dice is defined as below

$$Dice \left( M_f(x, y), M_f^{GT}(x, y) \right) = \frac{2|M_f(x, y) \cdot M_f^{GT}(x, y)|}{|M_f(x, y) + M_f^{GT}(x, y)|} \quad (4.1)$$

Where,  $M_f(x, y)$  is final fused image and  $M_f^{GT}(x, y)$  is the ground truth image.

#### 4.1.2 Jaccard:

$$Jaccard \left( M_f(x, y), M_f^{GT}(x, y) \right) = \frac{|M_f(x, y) \cdot M_f^{GT}(x, y)|}{|M_f(x, y)| + |M_f^{GT}(x, y)| - |M_f(x, y) \cdot M_f^{GT}(x, y)|} \quad (4.2)$$

#### 4.1.3 Sensitivity (Sen),

$$sen = \frac{TP}{TP+TF} \quad (4.3)$$

#### 4.1.4 Specificity (Spe)

$$Spe = \frac{TN}{TN+FP} \quad (4.4)$$

#### 4.1.5 Precision (Pre),

$$Pre = \frac{TP}{TP+FP} \quad (4.5)$$

#### 4.1.6 False negative rate (FNR),

$$FNR = \frac{FN}{P} \quad (4.6)$$

#### 4.1.7 False positive rate (FPR),

$$FPR = \frac{FP}{N} \quad (4.7)$$

#### 4.1.8 Accuracy (Acc)

$$Acc = \frac{TP+TN}{TP+TN+TF+FP} \quad (4.8)$$

Where TP, TN, FP, and FN are the true positive, true negative, false positive, and false negative which are obtained from the confusion matrix.

Table 4-1 Dataset Used

Datasets	Total images	Training Images	Testing Images
PH <sup>2</sup>	200	160	40
ISBI 2016	900	720	180
ISIC 2017	2750	2200	550

## 4.2 Performance Evaluation: Segmentation

To evaluate performance proposed segmentation method, two performance parameters are used namely Dice and Jaccard are given by Eq 4-1 and Eq. 4-2. The achieved Jaccard and Dice indices for selected images shown in **Error! Reference source not**

**found.** are summarize in Table 4-2. The obtained Dice and Jaccard indices are between 65.12% and 93.11%. On overall, through the whole database, the average Jaccard index is calculated to 83.32%, 81.24%, and 88.54% for ISBI 2016, ISIC 2017, respectively. The experimental analysis is telling us that the best segmentation accuracy has been recorded on the  $PH^2$  dataset, while the worst is on the ISCI 2017 dataset. These results clearly witness the challenge of complex data in accurately classifying skin lesion classification. Jaccard and Dice are summarized in Figure 4-4 & .....

**Table 4-2: Dice & Jaccard indices**

<b>Dataset</b>	<b>Image Sample</b>	<b>Dice</b>	<b>Jaccard</b>	<b>Average Jaccard index</b>
<b><math>PH^2</math></b>	a	84.45%	85.07%	88.54%
	b	81.62%	80.23%	
ISBI 2016	a	87.22%	86.21%	83.23%
	b	92.4%	93.11%	
	c	67.86%	65.12%	
ISBI 2017	a.	82.01%	81.76%	81.24%
	b.	80.22%	80.14%	
	c.	87.21%	86.33%	

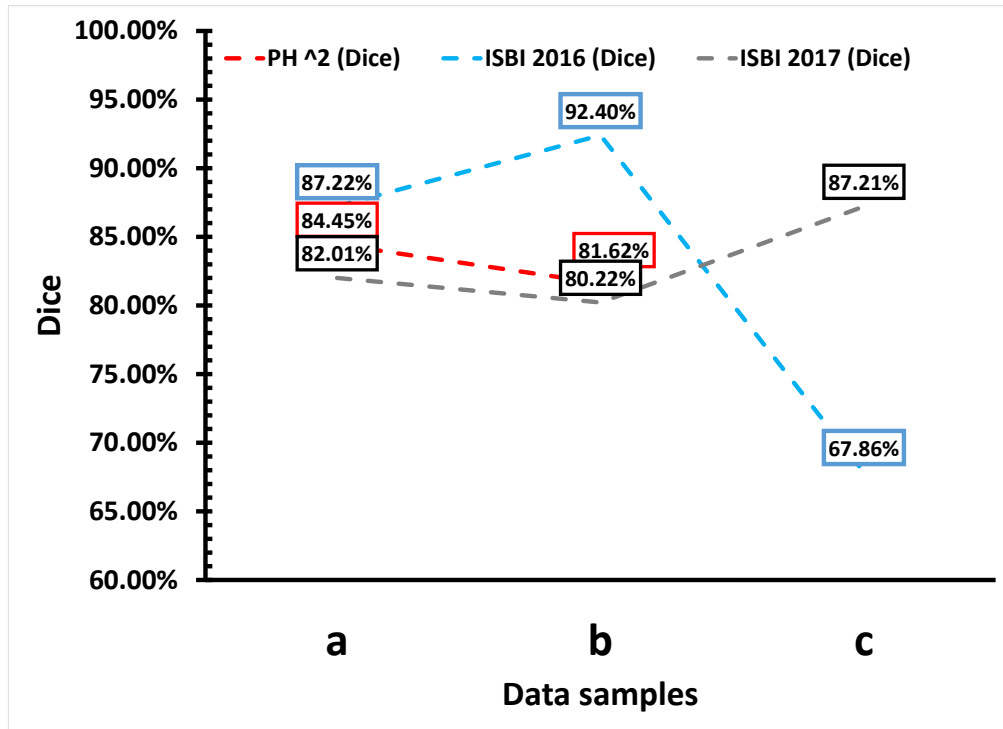


Figure 4-2 Segmentation performance: Dice index for selected samples

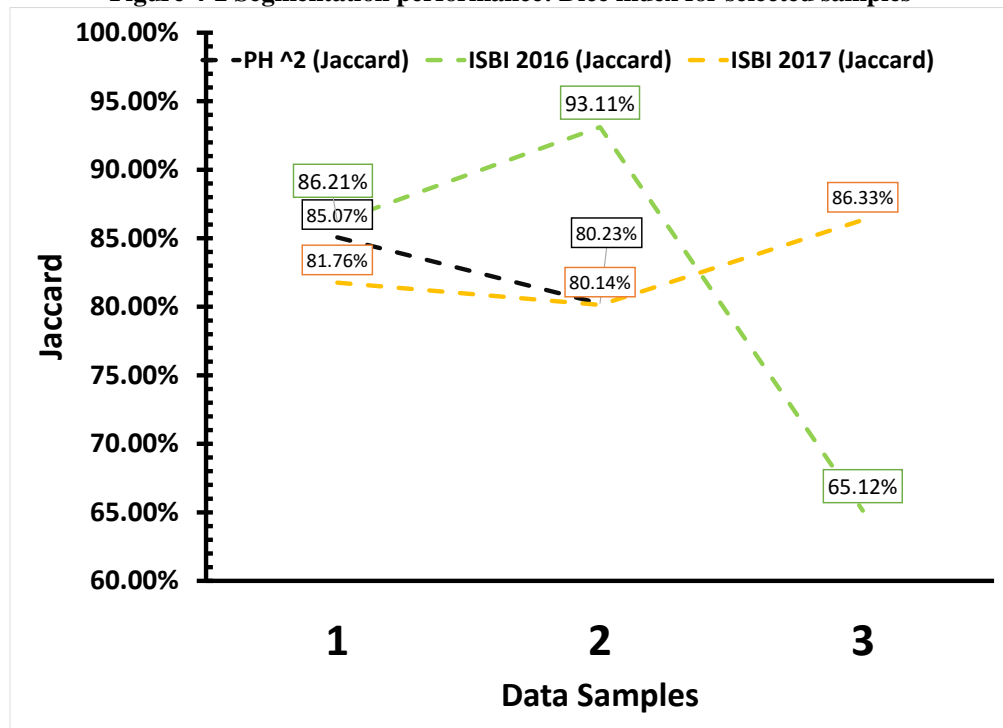


Figure 4-3 Segmentation Performance: Jaccard Index for sample data

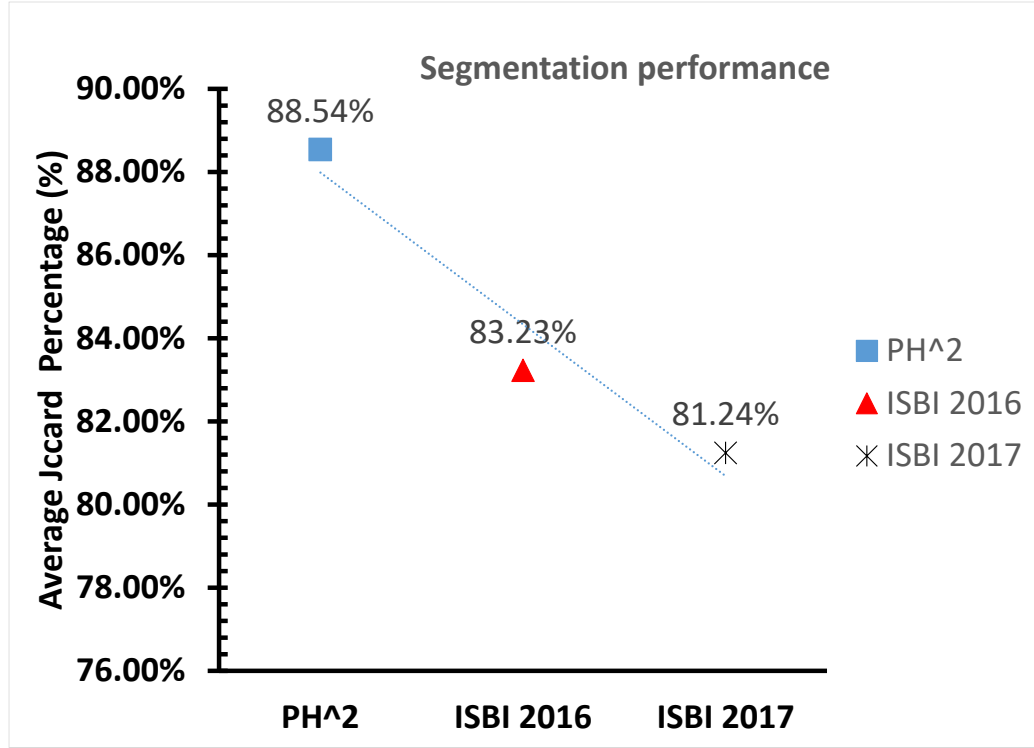


Figure 4-4: Average Jaccard Indices for Selected Dataset

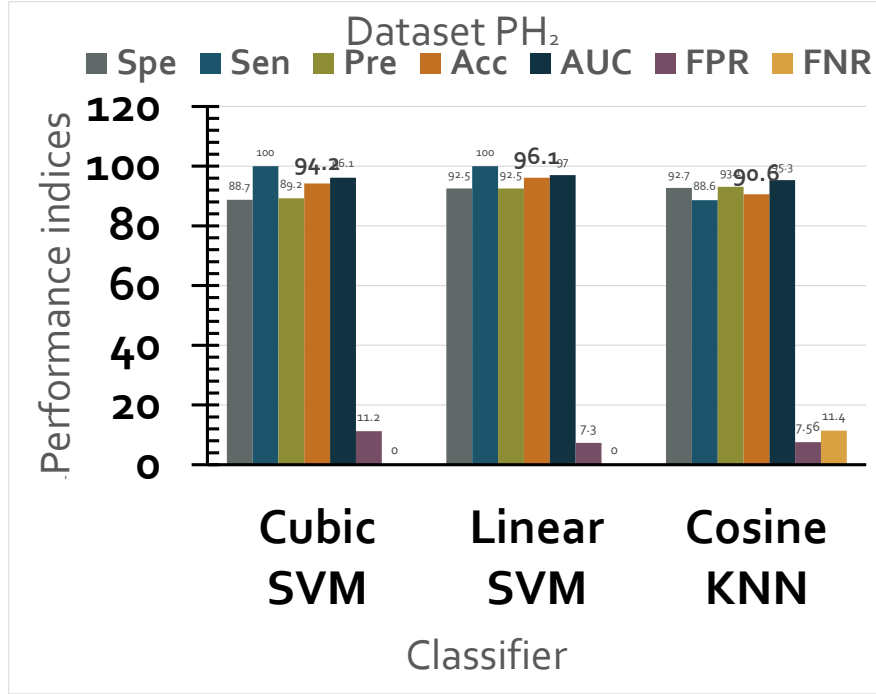
#### 4.3 Performance Evaluation: Classification

Transfer learning is utilized to extract features from two pre-trained models Inception-Resnet V2 & Inception-V3, and then feature fusion is done serially. Then feature vector is optimized using Entropy-Controlled GWO. The final feature vector passed through classifiers to classify the lesion localization. Multiple classifiers are used. Based on their better results three classifiers are selected i.e., Cubic SVM, Linear SVM, Cosine KNN. Table 4-3 demonstrates the classification results on the three Dataset the  $PH^2$ , ISIC 2016, ISIC 2017.

Table 4-3: Classification Performance Parameters

Datasets	Classifier	Sen	Spe	Pre	FNR	FPR	Acc	AUC
PH <sup>2</sup>	Cubic SVM	100	88.7	89.2	0	11.2	94.2	96.1
	Linear SVM	100	92.5	92.5	0	7.3	<b>96.1</b>	97
	Cosine KNN	88.6	92.7	93.1	11.4	7.56	90.6	95.3
ISBI 2016	Cubic SVM	93.9	84.1	88.57	6.12	16.1	90.1	94.7
	Linear SVM	96.6	87.5	90.63	3.8	12.5	<b>92.5</b>	95.2
	Cosine KNN	92.3	65.23	87.27	7.9	33.1	84.7	91.8
ISIC 2017	Cubic SVM	95.8	25.9	89.4	2.3	74.04	88.1	91.1
	Linear SVM	98.8	56.5	90	11.1	43.5	<b>90.4</b>	92.6
	Cosine KNN	95.7	55.9	84.9	2.9	44.1	85.96	90.2





**Figure 4-5: Performance parameters for PH2 Dataset**

To validate the proposed classification framework seven performance parameters i.e., Sensitivity, Accuracy, Specificity, Precisions, AUC (Area under the ROC Curve) False positive rate, False negative rate, are used.

It can observe from the Table 4-3 & Figure 4-5, Figure 4-6, Figure 4-7, that the linear SVM performs exceptionally compared to other classifiers by achieving the classification accuracy of 96.1%, 92.5%, and 90.4% on PH2, ISBI 2016 and ISIC 2017, respectively. Similarly, AUC greater than 95% clearly shows the average accuracy bounds on all these selected datasets.

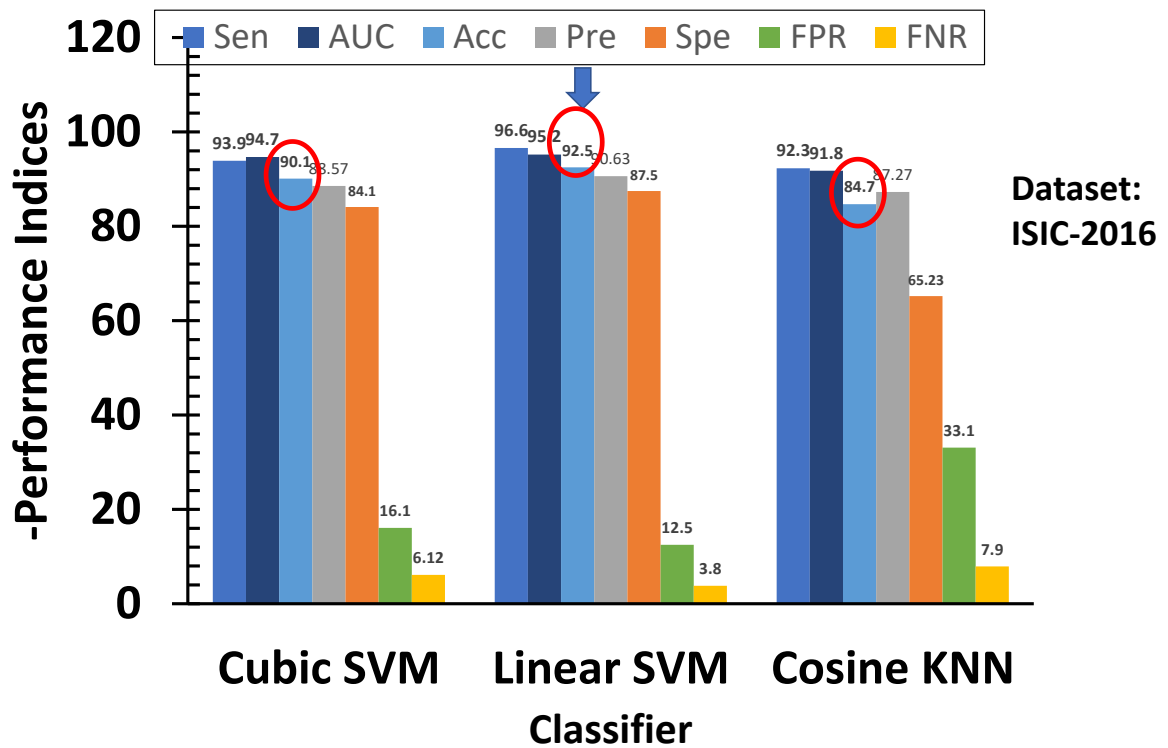


Figure 4-6: Performance Parameters for ISIC-2016

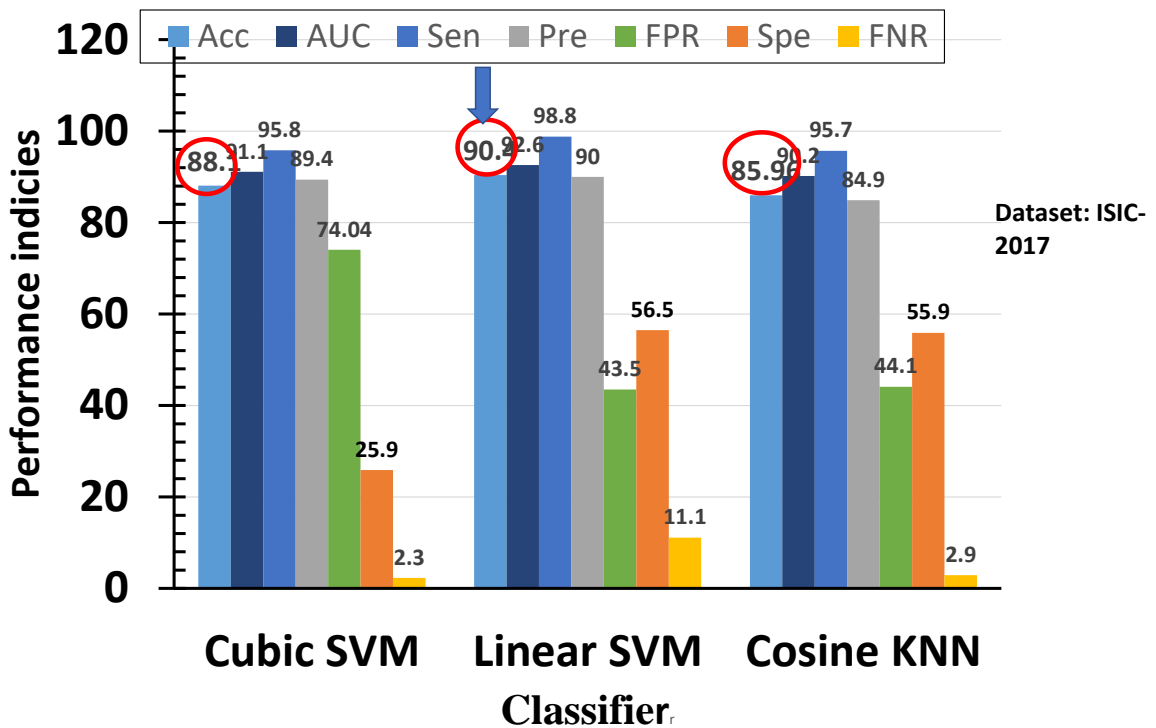


Figure 4-7: Performance Parameters for ISIC-2017

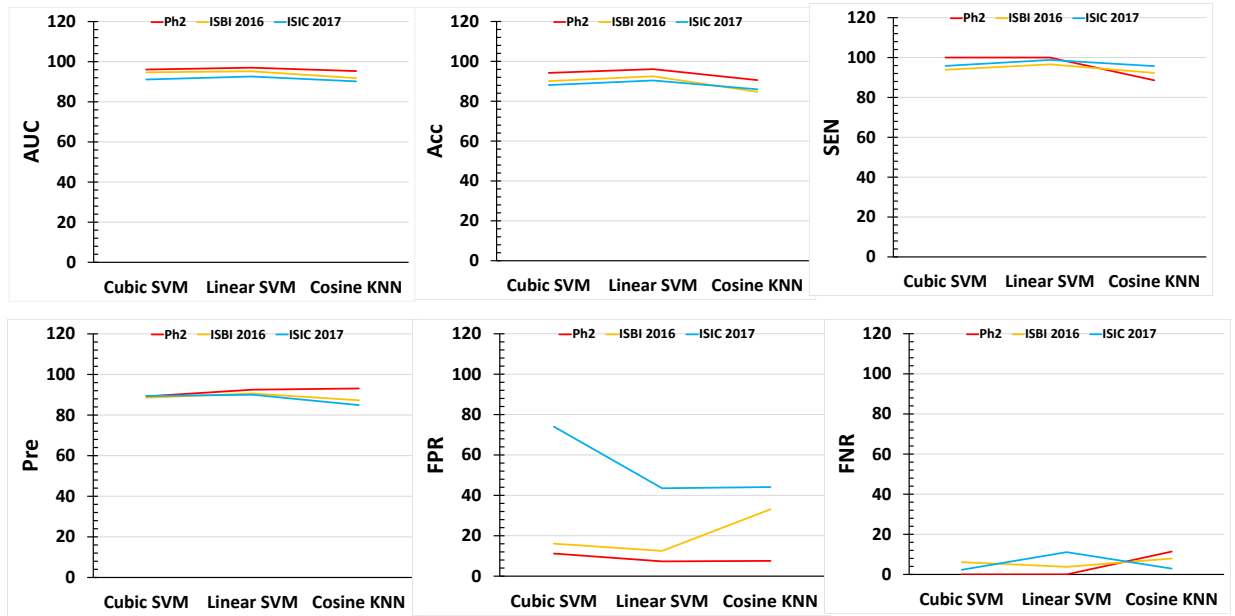


Figure 4-8 Classification performance parameters at three datasets

#### 4.4 Conclusion:

The proposed model perform relatively well with a good( 88.54 % ) segmentation Accuracy and 96.2% (PH2 dataset) Classification Accuracy. Although Segmentation Accuracy is satisfied but still there is room of improvement for segmentation method to cope with the data complexity. Combining these results with other existing architectures i.e., U-Net architecture, Mas-RCNN & VGG19-UNet may provide a boost in performance.

## **REFERENCES**

**S**

## References

- [1] B. Harangi, “Skin lesion classification with ensembles of deep convolutional neural networks,” *J. Biomed. Inform.*, vol. 86, no. June, pp. 25–32, 2018, doi: 10.1016/j.jbi.2018.08.006.
- [2] “Skin Cancer Information - The Skin Cancer Foundation.” [Online]. Available: <https://www.skincancer.org/skin-cancer-information/>. [Accessed: 20-Feb-2020].
- [3] A. Picardi, I. Lega, and E. Tarolla, “Suicide risk in skin disorders,” *Clinics in Dermatology*, vol. 31, no. 1. Elsevier, pp. 47–56, 01-Jan-2013, doi: 10.1016/j.clindermatol.2011.11.006.
- [4] “5 Most Common Skin Disorders,” *Fox News*, Mar. 2015.
- [5] “American Academy of Dermatology.” [Online]. Available: <https://www.aad.org/>. [Accessed: 20-Feb-2020].
- [6] A. EGDAHL, “WHO: World Health Organization,” *Ill. Med. J.*, vol. 105, no. 5, pp. 280–282, 1954, doi: 10.5260/chara.12.4.54.
- [7] J. Ferlay *et al.*, “Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods,” *International Journal of Cancer*, vol. 144, no. 8. Wiley-Liss Inc., pp. 1941–1953, 15-Apr-2019, doi: 10.1002/ijc.31937.
- [8] K. A. Freedberg, A. C. Geller, D. R. Miller, R. A. Lew, and H. K. Koh, “Screening for malignant melanoma: A cost-effectiveness analysis,” *J. Am. Acad. Dermatol.*, vol. 41, no. 5 I, pp. 738–745, Nov. 1999, doi: 10.1016/S0190-9622(99)70010-1.
- [9] C. Facts, “Cancer Facts & Figures 2020,” 2020.
- [10] A. B. Cagnetta, T. Vogt, M. Landthaler, O. Braun-Falco, and G. Plewig, “The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions,” *J. Am. Acad. Dermatol.*, vol. 30, no. 4, pp. 551–559, 1994, doi: 10.1016/S0190-9622(94)70061-3.
- [11] G. S. Ghalejoogh, H. M. Kordy, and F. Ebrahimi, “A hierarchical structure based on Stacking approach for skin lesion classification,” *Expert Syst. Appl.*, vol. 145, p. 113127, 2020, doi: 10.1016/j.eswa.2019.113127.
- [12] M. A. Al-masni, D. H. Kim, and T. S. Kim, “Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification,” *Comput. Methods Programs Biomed.*, vol. 190, p. 105351, 2020, doi:

- 10.1016/j.cmpb.2020.105351.
- [13] N. Hameed, A. M. Shabut, M. K. Ghosh, and M. A. Hossain, “Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques,” *Expert Syst. Appl.*, vol. 141, p. 112961, 2020, doi: 10.1016/j.eswa.2019.112961.
  - [14] M. A. Khan *et al.*, “An implementation of normal distribution based segmentation and entropy controlled features selection for skin lesion detection and classification,” *BMC Cancer*, vol. 18, no. 1, pp. 1–20, 2018, doi: 10.1186/s12885-018-4465-8.
  - [15] H. M. Ünver and E. Ayan, “Skin lesion segmentation in dermoscopic images with combination of yolo and grabcut algorithm,” *Diagnostics*, vol. 9, no. 3, 2019, doi: 10.3390/diagnostics9030072.
  - [16] S. Member, “Short papers,” vol. 17, no. 4, pp. 27–27, 2009, doi: 10.1109/vissof.2009.5336427.
  - [17] K. Møllersen, H. M. Kirchesch, T. G. Schopf, and F. Godtliebsen, “Unsupervised segmentation for digital dermoscopic images,” *Ski. Res. Technol.*, vol. 16, no. 4, pp. 401–407, 2010, doi: 10.1111/j.1600-0846.2010.00455.x.
  - [18] 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.
  - [19] S. Ghosh, N. Das, I. Das, and U. Maulik, “Understanding deep learning techniques for image segmentation,” *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–58, 2019, doi: 10.1145/3329784.
  - [20] M. Emre Celebi, Q. Wen, S. Hwang, H. Iyatomi, and G. Schaefer, “Lesion Border Detection in Dermoscopy Images Using Ensembles of Thresholding Methods,” *Ski. Res. Technol.*, vol. 19, no. 1, pp. 252–258, 2013, doi: 10.1111/j.1600-0846.2012.00636.x.
  - [21] Q. Abbas, M. E. Celebi, I. Fondón García, and M. Rashid, “Lesion border detection in dermoscopy images using dynamic programming,” *Ski. Res. Technol.*, vol. 17, no. 1, pp. 91–100, 2011, doi: 10.1111/j.1600-0846.2010.00472.x.
  - [22] P. M. M. Pereira *et al.*, “Skin lesion classification enhancement using border-line features – The melanoma vs nevus problem,” *Biomed. Signal Process. Control*, vol. 57, p. 101765, 2020, doi: 10.1016/j.bspc.2019.101765.
  - [23] B. D. Ripley, “Neural Networks and Related Methods for Classification,” *J. R. Stat. Soc.*

- Ser. B*, vol. 56, no. 3, pp. 409–437, 1994, doi: 10.1111/j.2517-6161.1994.tb01990.x.
- [24] T. F. Gonzalez, “Handbook of approximation algorithms and metaheuristics,” *Handb. Approx. Algorithms Metaheuristics*, pp. 1–1432, 2007, doi: 10.1201/9781420010749.
  - [25] A. Garcia-Perez, F. Gheriss, D. Bedford, A. Garcia-Perez, F. Gheriss, and D. Bedford, “Measurement, Reliability, and Validity,” *Des. Track. Knowl. Manag. Metrics*, pp. 163–182, 2019, doi: 10.1108/978-1-78973-723-320191012.
  - [26] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
  - [27] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 2261–2269, 2017, doi: 10.1109/CVPR.2017.243.
  - [28] M. A. Khan, M. Sharif, T. Akram, S. A. C. Bukhari, and R. S. Nayak, “Developed Newton-Raphson based deep features selection framework for skin lesion recognition,” *Pattern Recognit. Lett.*, vol. 129, pp. 293–303, 2020, doi: 10.1016/j.patrec.2019.11.034.
  - [29] A. Mahbod, G. Schaefer, I. Ellinger, R. Ecker, A. Pitiot, and C. Wang, “Fusing fine-tuned deep features for skin lesion classification,” *Comput. Med. Imaging Graph.*, vol. 71, pp. 19–29, 2019, doi: 10.1016/j.compmedimag.2018.10.007.
  - [30] A. Pennisi, D. D. Bloisi, D. Nardi, A. R. Giampetruzzi, C. Mondino, and A. Facchiano, “Skin lesion image segmentation using Delaunay Triangulation for melanoma detection,” *Comput. Med. Imaging Graph.*, vol. 52, pp. 89–103, 2016, doi: 10.1016/j.compmedimag.2016.05.002.
  - [31] 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.
  - [32] L. Bi, J. Kim, E. Ahn, D. Feng, and M. Fulham, “Semi-automatic skin lesion segmentation via fully convolutional networks,” *Proc. - Int. Symp. Biomed. Imaging*, pp. 561–564, 2017, doi: 10.1109/ISBI.2017.7950583.
  - [33] M. Goyal, A. Oakley, P. Bansal, D. Dancey, and M. H. Yap, “Skin lesion segmentation in dermoscopic images with ensemble deep learning methods,” *arXiv*, pp. 4171–4181, 2019.
  - [34] M. Nasir, M. Attique Khan, M. Sharif, I. U. Lali, T. Saba, and T. Iqbal, “An improved

- strategy for skin lesion detection and classification using uniform segmentation and feature selection based approach,” *Microsc. Res. Tech.*, vol. 81, no. 6, pp. 528–543, Jun. 2018, doi: 10.1002/jemt.23009.
- [35] R. Y. M. Nakamura, L. A. M. Pereira, K. A. Costa, D. Rodrigues, J. P. Papa, and X. S. Yang, “BBA: A binary bat algorithm for feature selection,” *Brazilian Symp. Comput. Graph. Image Process.*, pp. 291–297, 2012, doi: 10.1109/SIBGRAPI.2012.47.
  - [36] G. Venugopal, M. Navaneethakrishna, and S. Ramakrishnan, “Extraction and analysis of multiple time window features associated with muscle fatigue conditions using sEMG signals,” *Expert Syst. Appl.*, vol. 41, no. 6, pp. 2652–2659, 2014, doi: 10.1016/j.eswa.2013.11.009.
  - [37] E. Emary, H. M. Zawbaa, and A. E. Hassanien, “Binary grey wolf optimization approaches for feature selection,” *Neurocomputing*, vol. 172, pp. 371–381, 2016, doi: 10.1016/j.neucom.2015.06.083.
  - [38] J. Too, A. R. Abdullah, N. M. Saad, N. M. Ali, and W. Tee, “A new competitive binary grey wolf optimizer to solve the feature selection problem in EMG signals classification,” *Computers*, vol. 7, no. 4, 2018, doi: 10.3390/computers7040058.
  - [39] M. A. Khan, M. Sharif, T. Akram, R. Damaševičius, and R. Maskeliūnas, “Skin Lesion Segmentation and Multiclass Classification Using Deep Learning Features and Improved Moth Flame Optimization,” *Diagnostics*, vol. 11, no. 5, p. 811, 2021, doi: 10.3390/diagnostics11050811.
  - [40] J. Amin *et al.*, “Integrated design of deep features fusion for localization and classification of skin cancer,” *Pattern Recognit. Lett.*, vol. 131, pp. 63–70, 2020, doi: 10.1016/j.patrec.2019.11.042.
  - [41] T. Akram, M. A. Khan, M. Sharif, and M. Yasmin, “Skin lesion segmentation and recognition using multichannel saliency estimation and M-SVM on selected serially fused features,” *J. Ambient Intell. Humaniz. Comput.*, vol. 0, no. 0, p. 0, 2018, doi: 10.1007/s12652-018-1051-5.
  - [42] T. Akram *et al.*, “A multilevel features selection framework for skin lesion classification,” *Human-centric Comput. Inf. Sci.*, vol. 10, no. 1, 2020, doi: 10.1186/s13673-020-00216-y.
  - [43] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, and I. Ellinge, “Skin Lesion Classification Using Hybrid Deep Neural Networks,” *ICASSP, IEEE Int. Conf. Acoust. Speech Signal*



- Process. - Proc.*, vol. 2019-May, pp. 1229–1233, 2019, doi: 10.1109/ICASSP.2019.8683352.
- [44] 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.
  - [45] H. Iyatomi *et al.*, “Computer-based classification of dermoscopy images of melanocytic lesions on acral volar skin,” *J. Invest. Dermatol.*, vol. 128, no. 8, pp. 2049–2054, 2008, doi: 10.1038/jid.2008.28.
  - [46] K. Polat and K. Onur Koc, “Detection of Skin Diseases from Dermoscopy Image Using the combination of Convolutional Neural Network and One-versus-All,” *J. Artif. Intell. Syst.*, vol. 2, no. 1, pp. 80–97, 2020, doi: 10.33969/ais.2020.21006.
  - [47] D. A. Gavrilov, N. N. Shchelkunov, and A. V. Melerzanov, “Deep learning based skin lesions diagnosis,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 42, no. 2/W12, pp. 81–85, 2019, doi: 10.5194/isprs-archives-XLII-2-W12-81-2019.
  - [48] A. Kumar, L. Bi, J. Kim, and D. D. Feng, *Machine learning in medical imaging*. Elsevier Inc., 2020.
  - [49] H. Liu, *Rail transit inspection unmanned aerial vehicle (UAV) systems*. 2020.
  - [50] A. Borji, M. M. Cheng, Q. Hou, H. Jiang, and J. Li, *Salient object detection: A survey*, vol. 5, no. 2. 2019.
  - [51] E. Ahn *et al.*, “Saliency-Based Lesion Segmentation Via Background Detection in Dermoscopic Images,” *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 6, pp. 1685–1693, 2017, doi: 10.1109/JBHI.2017.2653179.
  - [52] Q. Duan, T. Akram, P. Duan, and X. Wang, “Visual saliency detection using information contents weighting,” *Optik (Stuttg.)*, vol. 127, no. 19, pp. 7418–7430, 2016, doi: 10.1016/j.ijleo.2016.05.027.
  - [53] L. Busin, N. Vandenbroucke, and L. Macaire, *Color Spaces and Image Segmentation*, vol. 151, no. 07. Elsevier Masson SAS, 2009.
  - [54] A. Pradesh and A. Pradesh, “A New Method of Image Enhancement in Spatial Domain Using Histogram Equalization , Smoothing and Fuzzy Technique,” vol. 8491, no. c, pp. 77–79, 2011.
  - [55] M. Zeng, Y. Li, Q. Meng, T. Yang, and J. Liu, “Improving histogram-based image contrast enhancement using gray-level information histogram with application to X-ray

- images,” *Optik (Stuttg.)*, vol. 123, no. 6, pp. 511–520, 2012, doi: 10.1016/j.ijleo.2011.05.017.
- [56] H. M. Wechsler, “Digital image processing, 2nd ed.,” *Proceedings of the IEEE*, vol. 69, no. 9, pp. 1174–1175, 2008, doi: 10.1109/proc.1981.12153.
- [57] Y. H. Shao, L. Bai, Z. Wang, X. Y. Hua, and N. Y. Deng, “Proximal plane clustering via eigenvalues,” *Procedia Comput. Sci.*, vol. 17, pp. 41–47, 2013, doi: 10.1016/j.procs.2013.05.007.
- [58] R. H. Abraham, “Book review Minutes from an infinite paradise ’,” *Science (80-. )*, vol. 18, pp. 1041–1043, 1994.
- [59] K. S. & A. Zisserman+, “VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION Karen,” *Am. J. Heal. Pharm.*, vol. 75, no. 6, pp. 398–406, 2018.
- [60] M. I. Razzak, S. Naz, and A. Zaib, “Deep Learning for Medical Image Processing: Overview, Challenges and the Future BT - Classification in BioApps: Automation of Decision Making,” *Springer*, pp. 323–350, 2018.
- [61] T. Brosch, Y. Yoo, D. K. B. Li, A. Traboulsee, and R. Tam, “Modeling the Variability in Brain Morphology and Lesion Distribution in Multiple Sclerosis by Deep Learning,” in *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2014*, 2014, pp. 462–469.
- [62] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation.” .
- [63] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey Wolf Optimizer,” *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.