INDIAN STATISTICAL INSTITUTE

MacPso: Machine assisted analysis of Psoriatic skin

by

Anabik Pal

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

Under the supervision of
Prof. Utpal Garain
Computer Vision and Pattern Recognition Unit

October 2019

Declaration of Authorship

I, Anabik Pal, declare that this thesis titled, ‘MacPso: Machine assisted analysis of Psoriatic skin’ and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Anabik Pal

15/10/2019

"For a successful technology, reality must take precedence over public relations, for Nature cannot be fooled."

Richard Feynman

Abstract

Psoriasis is a chronic, immune-mediated, relapsing, inflammatory skin disease and usually associated with itch. The prevalence of psoriasis varies 1%-12% among different populations worldwide. This disease develops when the immune system mistakes a normal skin cell for a pathogen and sends out faulty signals that cause overproduction of new skin cells. This disease can be diagnosed by visual and haptic inspection. The visual changes of the outer skin surface due to this disease include the development of elevated red scaly dry patches with well-demarcated borders on the skin surface. However, the shape, size, color, and distribution of these patches vary. In dermatology, these patches are termed as psoriatic plaque.

No drug is available yet to cure psoriasis completely but the severity can be controlled by suitable drug doses. However, the drug response varies among different patients. So, the development of reliable severity assessment procedure is required to decide the type and dosage of treatment required as well as to measure disease progress and drug efficacy. So, dermatologists use Psoriasis Area Severity Index (PASI) for estimating severity. This procedure considers four parameters: (i) percentage of affected skin area, (ii) redness (erythema) level of the affected area, (iii) silveriness (scaling) of the affected area and (iv) the elevation (induration) formed on the skin for this disease. The PASI scoring is an eye estimation procedure and done by the dermatologists in a subjective manner.

Another challenge in psoriasis management is to diagnose psoriasis. It is often difficult to differentiate psoriasis from other erythemato-squamous diseases like Seborrheic dermatitis, Leprosy, Lichen planus, Tinea corporis, Pityriasis, Eczema, etc. as they carry similar clinical features. Hence, histopathological examination is considered for confirmation. But the quantitative analysis of biopsy images by a human is time-consuming and laborious. Moreover, issues like over-staining, tissue overlapping in the biopsy slide make the task challenging, even for highly experienced pathologists.

This thesis deals with four different research problems to cope up with dermatologists' limitations. The very first task is to localize plaque regions in skin images. For this task, a new clustering-based segmentation algorithm is developed and used. The second task is to develop machine learning algorithm for automatic severity scoring of psoriatic plaque. For this task a multi-task learning based

Deep Convolutional Neural Network (DCNN) is developed. Furthermore, a new loss function is designed to train a DCNN for ordinal classification. These two approaches are used for automatic severity scoring of psoriatic plaques. The third task is the development of convolutional neural network-based segmentation approach for segmenting biopsy images into skin layers. The final task is to detect Munro's Microabscess in skin biopsy images. Munro's Microabscess is considered as diagnostic hallmark of psoriasis. Separation of stratum corneum layer, extraction of patches from the layer followed by classification of patches with respect to presence or absence of neutrophils form the basis of the overall approach which is effected through an integration of a capsule network-based segmentation and classification network. The development of annotated datasets for conducting all experiments is another contribution of this research.

Acknowledgements

I am eternally grateful to my supervisor, Professor Utpal Garain. This thesis would not have been possible without his guidance, expertise and endless enthusiasm and optimism. It has been a pleasure working with him for the past four years.

I would also like to thank my collaborators, Dr. Swapan Senapati, Dr. Raghunath Chatterjee, and Dr. Aditi Chandra. They spent countless hours in generating image datasets and explaining the biology to me.

I would also like to thank Dr. Anandarup Roy and Mr. Akshay Chaturvedi who taught me how to understand complex machine learning algorithms. Without them, it would have been difficult for me to explore the advances in machine learning algorithms for my research. Thanks to Mr. Sounak Ray for his effort in implementing the GUI (Graphical User Interface).

Thanks to faculty members (especially, Prof. Bidyut Baran Chaudhuri, Prof. Swapan Senapati, Dr. Mandar Mitra, Dr. Ujjwal Bhattarcharya, Dr. Sarbani Palit) and office staffs (especially, Sankar Da, Bhabani Di, Manik Da, and Sunil Da) of CVPR unit for their kind co-operation.

Thanks to all research scholars and project-linked persons of CVPR Unit, past, and present, for making it such a fun place to work. Special thanks go to Abhishek, Apurba Da, Arjun, Onkar, Rupayan Da, Soumen, Dwaipayan, Chandan, Partha Da, Kushal, Sayan, Ayan Da, Ankan Da, Oendrila Di, Chanchal, Sudip, Suchana, Jayati, Sumit, Tanmoy for being a good friend, great sounding board for ideas and problems. I wish them all the best and hope we stay in touch. Thanks to some of the hostel mates, including, but not limited to Ankush, Sourav, Ankita, Bidesh, Sarbendu, Biltu, Gopal, Sanchayan, Manjari, Raju Da, Tapas Da, etc.

Last not the least, I would like to thank my parents and sister for their years of mental support and encouragement.

Contents

List of Figures

List of Tables

Abbreviations

FCM	**F**uzzy **C**- **M**eans
GMM	**G**aussian **M**ixture **M**odel
SWGMM	**S**emi- **W**rapped **G**aussian **M**ixture **M**odel
JCLMM	**J**oint **C**ircular **L**inear **M**ixture **M**odel
MAP	**M**aximum **A** **P**osteriori
EM	**E**xpectation **M**aximization
SLIC	**S**imple **L**inear **I**terative **C**lustering
PRI	**P**robabilistic **R**and **I**ndex
GCE	**G**lobal **C**onsistency **E**rror
VoI	**V**ariation **o**f **I**nformation
LBP	**L**ocal **B**inary **P**attern
GLCM	**G**ray **L**abel **C**o-occurence **M**atrix
HOG	**H**istogram of **O**riented **G**radient
KNN	**K**- **N**earest **N**eighbour
SVM	**S**upport **V**ector **M**achine
RBF	**R**adial **B**asis **F**unction
RF	**R**andom **F**orest
CNN	**C**onvolutional **N**eural **N**etwork
DCNN	**D**eep **C**onvolutional **N**eural **N**etwork
ReLU	**Re**ctified **L**inear **U**nit
STL	**S**ingle **T**ask **L**earning
MTL	**M**ulti **T**ask **L**earning
EMD	**E**arth **M**over's **D**istance
FCN	**F**ully **C**onvolutional **N**eural Network

MRF	**M**arkov **R**andom **F**ield
RCPC	**R**atio of **C**orrect **P**ixel **C**lassification
JC	**J**accard's **C**oefficient
CAD	**C**omputer **A**ided **D**iagnosis
ROC	**R**eceiver **O**perating **C**urve
PASI	**P**soriasis **A**rea **S**everity **I**ndex
WSI	**W**hole **S**lide **I**mage
SC	**S**tratum **C**orneum
SG	**S**tratum **G**ranulosum
SS	**S**tratum **S**pinosum
SB	**S**tratum **B**asale
Ne	**Ne**utrophils
Nu	**Nu**cleus
MM	**M**unro's **M**icroabcess

Dedicated to my Parents

Chapter 1

Introduction

Psoriasis is a life long skin disease characterized by reddish or silvery elevated patches on the human body. According to Chandran and Raychaudhuri, the prevalence of psoriasis varies from 1%-12% among different populations worldwide [17]. There are several research challenges related to psoriasis. However, the important computer vision challenges related to psoriasis are: analysis of skin images for standardizing the severity assessment procedure and extraction of histopathological characteristic from skin biopsy images. Hence, this thesis focuses on the machine-assisted analysis of psoriasis images. Before explaining the research conducted for this thesis, this chapter will provide a brief introduction about psoriasis, motivation for development of computer vision algorithm for analyzing the disease and the available related research works.

The organization of this chapter is as follows. The clinical background of psoriasis skin diseased is given in Section 1.1. Section 1.2 discusses the clinician's approach for the severity assessment of psoriasis patients. Section 1.3 discusses the computer vision task related to psoriasis biopsy analysis. The research motivation and the previous research works are given in Section 1.4 and Section 1.5. Section 1.6 summarizes the contributions of this thesis and finally, in Section 1.7 the organization of the thesis is written.

1.1 Psoriasis

Psoriasis is a chronic, immune-mediated, relapsing, inflammatory skin disease and usually associated with itch. This disease develops when the immune system mistakes a normal skin cell for a pathogen and sends out faulty signals that cause overproduction of new skin cells. Due to this disease, both visual changes of the outer skin, as well as histopathological disorder of human skin occurs.

The visual changes of outer skin surface due to psoriasis include the development of elevated red scaly dry patches with well-demarcated borders on the skin surface. The shape, size, color, and distribution of these patches vary. In dermatology, these patches are termed as psoriatic plaque. The silvery-white scale of psoriatic plaque reveals bleeding points when removed [96]. For the visual representation, two images collected from the psoriasis patients are shown in Fig. 1.1 and two cropped image samples obtained from the marked regions of Fig. 1.1 are shown in Fig. 1.2.

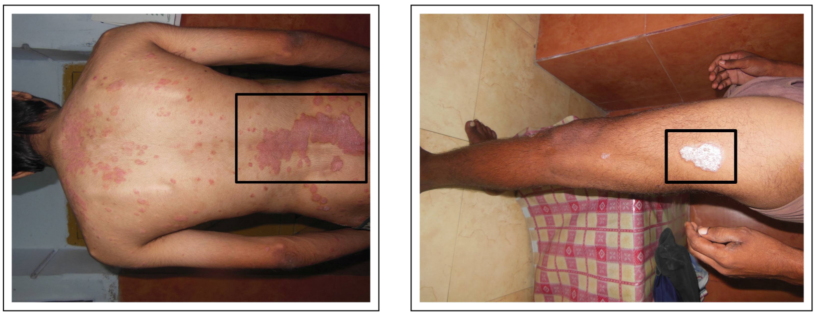


Figure 1.1: Images from two psoriasis affected patients.

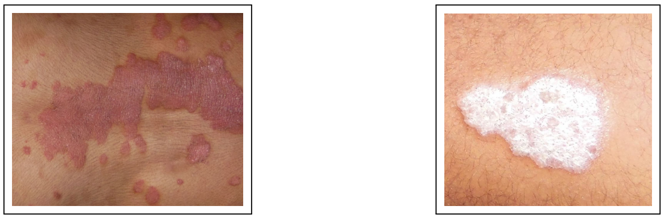
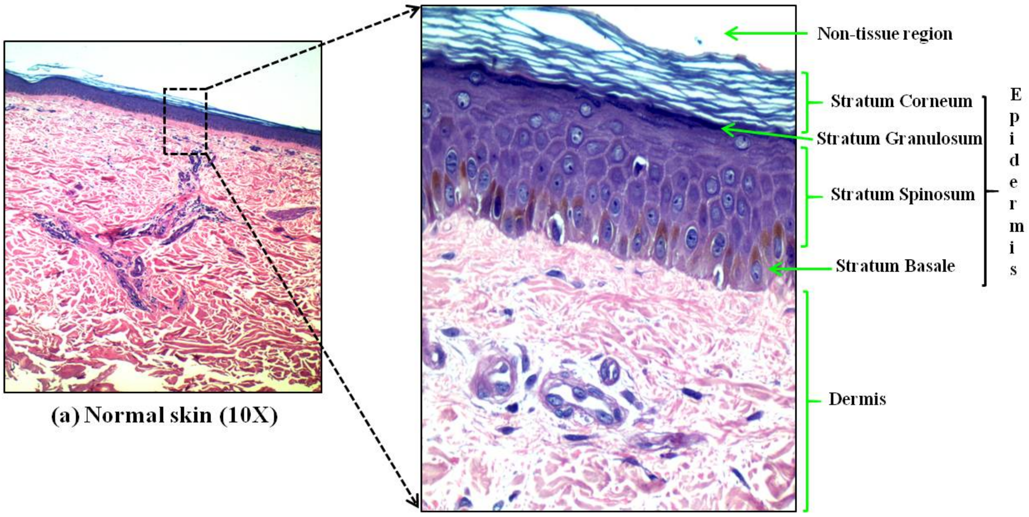


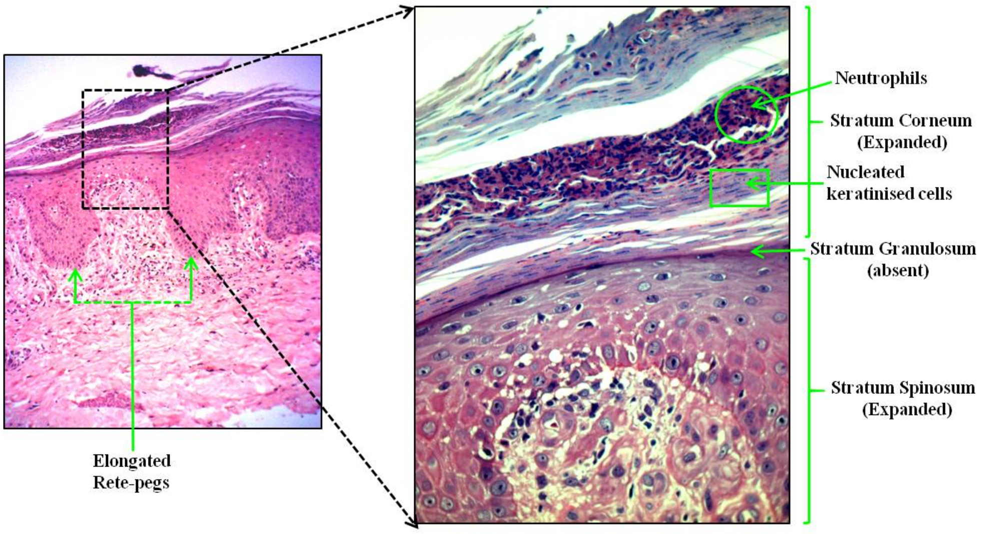
Figure 1.2: Cropped skin image samples from the marked regions of Fig. 1.1: reddish due to erythema and silvery due to scaling.

Psoriasis can be diagnosed by the clinical features by looking at the outer skin surface. But instead of psoriasis, there are other erythemato-squamous diseases like Seborrheic dermatitis, Leprosy, Lichen planus, Tinea corporis, Pityriasis, and Eczema, etc. which have similar clinical features [42, 43, 61]. Therefore, the diagnosis of psoriasis as a different and distinct disease had been a controversial



(a) Normal skin (10X)

(b) Normal skin (40X)



(c) Psoriatic skin (10X)

(d) Psoriatic skin (40X)

Figure 1.3: Comparison of normal and psoriasis affected skin biopsy images kept under the same field of view: Epidermis is thin uniform layer in normal skin but, in a diseased skin, it is thickened, becomes irregular and generates elongation of rete-pegs; Stratum Corneum (SC) is intact in the normal skin but expanded in the diseased skin with occasional neutrophils (Ne) and abnormal retention of Nucleus (Nu); Stratum Granulosum (SG) layer is present in normal skin but absent in the diseased skin; Stratum Spinosum (SS) layer is elongated in the diseased skin; Stratum Basale (SB) does not appear in the field of view (40X) for the diseased skin.

subject since ancient times. It is worth to mention that the disease was confused with leprosy on several occasions which was finally solved when Camille Melchoir Gibert (1792-1866) described psoriasis as a separate disease and its difference from other papulosquamous diseases [93] by using a set of histopathological characteristics. Hence, in doubtful cases histopathological examination is carried out for confirmation.

1.2 Psoriasis severity assessment

The Psoriasis Area Severity Index (PASI), first described in [35], is a popularly used severity estimation technique for quantifying the disease severity for psoriasis management and clinical research. PASI considers two major aspects of the disease: percentage of body surface area affected by this disease and the severity of the plaques formed on the skin surface. The four body parts: (i) head (h), (ii) trunk (t), (iii) upper extremity (u) and (iv) lower extremity (l) are considered. These body parts nearly correspond to 10%, 20%, 30% and 40% of the total body surface area. The extent of area involved in each part (A_h, A_t, A_u and A_l) is given a score between $0-6$ (as shown in Table 1.1). The severity of the plaques is measured based on three parameters: degree of redness or erythema (E_h, E_t, E_u and E_l), thickness or the induration (I_h, I_t, I_u and I_l) and scaling or desquamation (S_h, S_t, S_u and S_l) and given a value between $0-4$ according to level of severity (as shown in Table 1.2). Finally, the PASI score is calculated using the following formula : $0.1(E_h+I_h+S_h)A_h+0.2(E_u+I_u+S_u)A_u+0.3(E_t+I_t+S_t)A_t+0.4(E_l+I_l+S_l)A_l$. The PASI score varies in gradation of 0.1 units between 0.0 to 72.0. Patients with PASI score > 10 are considered as suffering from the severe form of the disease.

Apart from the area of involvement, all of these severity factors can take one of the five (5) severity grades in between $0-4$. Hence, every image has three severity scores on the basis of erythema, scaling, and induration. To get a better understanding of the problem, psoriasis image samples for different severity grading and severity factors are shown in Table 1.3.

Table 1.1: Scoring of skin surface area involvement

Area involved (%)	Score
0	0
< 10	1
10-29	2
30-49	3
50-69	4
70-89	5
90-100	6

Table 1.2: Scoring of each severity parameter

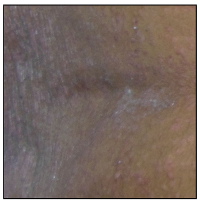
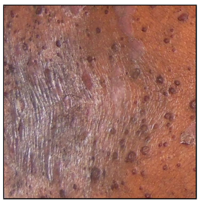
Severity	Score
Absent	0
Mild	1
Moderate	2
Severe	3
Very severe	4

1.3 Psoriasis histopathology

Histopathology refers to the microscopic examination of the tissue specimen for studying the manifestations of a disease. For histopathological analysis, tissue samples are collected under the supervision of experts. Formalin-fixed tissues are dehydrated and embedded in paraffin blocks. Thin sections are used for slide preparation and then stained with chemicals to prepare the histopathological slides. Afterward, the prepared histopathological slides are kept under a microscope for examination.

The histopathological changes of psoriasis skin biopsy include thickening of the epidermal layer or epidermis (outer layer of human skin), causing partial invasion of epidermis inside dermal layer or dermis (inner layer of human skin), retention of nuclei in the stratum corneum (parakeratosis), elongation of rete-pegs, suprapapillary thinning, dilation of blood vessels (angiogenesis), infiltration of large number of immune cells into the dermis and often into the epidermal compartments. Extensions of the epidermal layer, called rete-pegs, often fuse and trap parts of dermal layer forming several sub-compartments. The infiltration of nucleated keratinocytes along with accumulation of neutrophils in Stratum Corneum (outer sub-layer of epidermis) is termed as Munro's Microabscess (MM). This infiltration occurs either in confluent (throughout the Stratum Corneum layer) or in

Table 1.3: Sample images from our database. The second column contains five (5) images with five (5) different severity labels with respect to Erythema. Erythema score is increasing when the skin becomes redder. The third column contains five (5) images with five (5) different severity labels with respect to Scaling. Scaling score is increasing when the skin becomes more silvery. The fourth column contains five (5) images with five (5) different severity labels with respect to Induration. Induration score is increasing when the skin becomes more elevated.

Severity Score	Erythema	Scaling	Induration
Absent (0)			
Mild (1)			
Moderate (2)			
Severe (3)			
Very Severe (4)			

focal (not confluent) manner. In clinical pathology, Munro's Microabscess is considered as the diagnostic hallmark of psoriasis [79]. Figure 1.3 illustrates a pictorial representation of normal and psoriasis affected human skin biopsy images.

1.4 Research Challenge and Motivation

Psoriasis is a recurrent disease. No drug is available yet to cure psoriasis completely but some drugs are there to control the severity. Still, these drugs show variable degrees of response and adverse effects in different patients. The severity assessment is important to decide the type and dosage of treatment required as well as to measure disease progress and drug efficacy. But the current practice of severity assessment depends on the experts' experience, efficiency and suffers from both inter and intra-observer variations. So, clinicians often need to standardize the severity assessment procedure. This motivates computer vision researchers to conduct research for developing imaging algorithms for this purpose.

On the other hand, histopathological analysis of biopsy slides is a very tedious task and it is error-prone due to staining variation, inter-observer variability among pathologists, and several other factors. Most importantly, the histopathological study is subjected to high degree of variability because of lack of any quantitative approach. In that respect, an automated system can deal with the pathologists' limitations in a better way. Moreover, the availability of such a system will relieve the workload of pathologists by sieving out obvious healthy cases allowing specialists to spend more time on more sophisticated cases.

1.5 Literature review

The existing works related to psoriasis skin image processing are described in this section. The section is divided into two subsections. The first subsection discusses the existing research for camera held psoriasis skin image analysis and the next subsection discusses the existing research for psoriasis skin biopsy image analysis.

1.5.1 Camera held psoriasis skin image analysis

Literature reports several research efforts for the automatic analysis of psoriasis images. The very beginning task for an automatic camera held psoriasis image processing is the segmentation of the affected regions in the skin images. Several attempts are already reported for segmentation of affected regions [7, 11, 12, 58, 108, 109] and segmentation of scaling regions [75]. The available psoriatic plaque segmentation methods can be broadly categorized into four groups: (a) use of color and texture feature for disease region identification, (b) contour evaluation approaches, (c) thresholding based approaches, and (d) watershed algorithm. Taur used the fuzzy texture spectrum and the two-dimensional fuzzy color histogram in the hue-saturation space as feature descriptor and perform classification of skin region using neuro-fuzzy classifier [108], and multi-resolution orthogonal subspace classifier (MSSC) [109]. Caliman et al. used extended mathematical morphology based feature descriptor and then applied k-means clustering to segment the plaque regions [12]. Bogo et al. employed Geometric Active Contours for segmentation of plaque regions [7]. In this work, the initial contour is decided based on chromatic information and the contour is evolved with the image forces. Jailani et al. used thresholding based approach for lesion border detection [58]. In this work, firstly, the original images were pre-processed by several filtering techniques, morphological operations and then the diseased regions were identified by thresholding. Căliman and Ivanovici used watershed segmentation approach to segment affected region [11]. In this work, color lacunarity based feature was used to locate the initial markers. Lu et al. attempted segmentation of scaling regions in the skin images [75]. In this work, the pixel-wise classification of image pixels is done by using color and texture feature. For every image, individually trained classifiers are prepared using support vector machine. Finally, a Markov random field (MRF) is used to smooth a pixel-wise classification outcome. All these approaches are tested on their in-house data set having a few images and achieves promising performance. But all of these approaches have their own algorithmic limitations. For region classification based approaches, the segmentation performance is highly dependent on the used feature set. On the other hand, the performances of the thresholding based and contour evolution based segmentation approaches are hindered by the presence of noise in the images.

Determination of the lesion area of psoriasis affected skin from full-body images are reported in [56, 57]. For these research, images from 8 patients were collected by following the PASI standard (images of face, anterior and posterior side from trunk and both left and right upper limbs and lower limbs). Ihtatho et al., firstly, segment the body region and then segment lesional areas in the input images [56]. In this work, the segmentation task is done by color analysis and thresholding methods. The experimental result shows that in most cases, human equivalent performance is achieved by this system. Later, Ihtatho et al. used the same data set and investigated the possibility of using centroids obtained from three different groups based on their skin color (namely fair, brown, and dark) to avoid calculating centroids for each patient [57]. In this method, CIE L*a*b* color space is used to represent the skin color and for each group, the centroids of four body regions (head, trunk, arm, and leg) are calculated from normal skin and lesion areas of all patients in the group. The limitation of this approach is that the performance of the system is highly dependent on the segmentation result.

The computer-aided lesion area determination required intelligent imaging technique. From that motivation, an enhanced imaging technique is proposed in [45]. The determination of body surface area (BSA) and the lesion area is improved by the proposed imaging method. In this work, the severity assessment procedure (PASI) is termed as α-PASI method. The fuzzy c-means clustering algorithm was applied and the membership functions of lesions area for α-PASI area scoring was determined. The performance of the proposed α-PASI area scoring had been tested in an in-house image data set collected from 46 patients. The proposed system produced BSA accuracies obtained at four body regions were 97.80% (lower limb), 92.41% (trunk), 87.72% (upper limb), and 83.82% (head). The experimental result shows that the Kappa coefficient of the proposed α-PASI system was greater than or equal to 0.72 for all body regions (Head - 0.76, Upper limb - 0.81, Trunk - 0.85, Lower limb - 0.72) and the overall kappa coefficient for the α-PASI area is 0.80.

In literature, few approaches are available where automatic erythema scoring is reported. Fadzil and Ihtatho [30] and Hani et al. [46] used Chromameter for Skin color analysis. Fadzil and Ihtatho modeled the color of psoriasis lesion by hue (H), saturation (S), and value (V) [30]. In this research, thirty-three (33) patients with plaque psoriasis from Malay, Chinese and Indian ethnic origin were involved to prepare a database of 3 skin types (fair, brown and dark). Chromameter CR-400 (Konica Minolta, Osaka) was used to assess the skin color of patients. Their experiment revealed that among all three color components of HSV (Hue, Saturation, Value) color-space hue and saturation are the two most correlated parameters with PASI erythema scores. Hani et al., firstly, obtained the skin color (CIE L*a*b*) data of representative lesions and their surrounding normal skin from the different body areas (head, trunk, lower and upper extremities) using a chromameter (Konica Minolta CR-400). Then they grouped the patients into dark, brown, light brown or fair skin tones based on the L* values of their normal skin. Fuzzy C-Means (FCM) algorithms are applied consecutively to classify normal skin tone and to score PASI erythema. The experiment was performed on an in-house psoriasis skin image data set. The training data set contains 2,323 normal skin and 1,462 diseased skin and the test data set contains 430 diseased skin on which erythema scoring was tested. Raina et al. proposed color image analysis approach for erythema scoring [95]. In this work, color calibration based image analysis method is chosen. Several color-based features are extracted from the representative area associated with erythema and a linear discriminant analysis based classifier is built. The proposed system is tested on an in-house data set of 80 clinical images taken from 20 patients. A 4x6 color card (CameraTrax, Las Vegas, Nevada) with 24 colors was used to calibrate the coloration. Gupta et al. proposed a multi-spectral image analysis approach for automatic erythema grading [41]. In this work, the Skellam distribution statistics is used as feature vectors and this feature extraction method extends the tissue-photon interaction model to make it skin tone independent. A random forest based classification technique has opted for classification task. The proposed method was evaluated on an in-house data set of multi-spectral images collected from 30 patients. The limitation of using Chromameter is that it can take the color intensity from the skin but it cannot store the skin texture which may be useful for better performance. Hence, the performance is highly dependent on skin color intensity acquisition. On the other hand, the image analysis based approaches are highly dependent on the used feature descriptors.

Automatic evaluation of the severity of scaling and erythema in psoriasis lesions is attempted in [26]. The proposed system has three steps: segmentation of the lesion, segmentation of scales region and finally, scoring the disease. Three input features are used to develop the model: the area of the scaling, the ratio between the area of scaling and the area of the lesion, and the ratio between the area of scaling and the area of redness. The computational pipeline shows that the performance of this system depends on three factors- (i) accurate lesion segmentation,

(ii) accurate scaling region segmentation and (iii) robustness of the feature used for scaling and erythema grading.

Classification of psoriasis images into one of the five severity grades: healthy, mild, moderate, severe and very severe is attempted in [106]. In this system, firstly, for every image the region of interests (ROI) are outlined in an interactive way. Then several color and texture features are extracted from the ROIs and important discriminating features are chosen by feature selection algorithms. Finally, a classifier is built using a classification algorithm. Principal Component Analysis (PCA) and Fisher Discriminant Analysis (FDA) algorithms are used for feature reduction. Several classification schemes are developed and the performance is compared. Support Vector Machine (SVM) and Decision Tree (DT) based classifiers are chosen for classifier building. The performance comparison among the constructed classifiers is reported on the basis of an in-house data set of 848 psoriasis images with five severity grades: healthy, mild, moderate, severe and very severe, consisting of 383, 47, 245, 145, and 28 images respectively. The limitation of this approach is that the proposed approach requires expert's intervention to mark the plaque regions from which the feature is extracted.

Automatic identification of the presence of psoriatic lesion in dermatology images is attempted in [104, 105]. Shrivastava et al. used several grayscale space, color space and aggressiveness of psoriatic disease such as redness and chaoticness as feature descriptor [104]. In this work, Support Vector Machine (SVM) classifier is used for classification. The region of interest from the images are manually marked before feature extraction. The performance of the proposed system using several experimental protocols is reported on the basis of an in-house data set of 540 image patches (270 healthy and 270 diseased) captured from 30 psoriasis patients of Indian ethnic origin. The feature extraction strategy successfully chose 15 optimal features and the proposed system provided the classification accuracy of 99.53%, 99.66% and 99.81% for 5-fold, 10-fold, and Jack Knife protocols respectively. Later, Shrivastava et al. used four different types of features namely, Higher-Order Spectra (HOS) features, texture features, and color features for classification in [105]. A Principal Component Analysis (PCA) based feature selection was opted for dominant feature and SVM is used for classification. The performance of the proposed system using several experimental protocols was reported on the basis of the data set used in [104]. The best model produced from this research achieves 100% classification accuracy, 100% sensitivity, and specificity. Although the proposed approach achieves perfect classification accuracy, it is highly dependent on expert's intervention for marking the plaque regions.

Some attempts have been done for severity change assessment of psoriasis plaque as reported in [25, 40, 76]. Gomez et al., firstly, registered the images and then transformed the original images into sets of variables that exhibit decreasing degree of similarity, based on correlation measures [40]. The proposed system was tested on an in-house data set containing images from three (3) different psoriasis patients collected in five (5) consecutive weeks. Delgado et al. used a two-stage approach for registering psoriasis images [25]. In this method, at the first stage, the diseased area is segmented and in the second stage segmentation information is used to align the image based on the first two statistical moments of the area. In order to prove the efficacy of the proposed system a data set consisting of four psoriasis images of same lesion was collected during several weeks and registered with respect to another image of the same lesion. Lu et al. proposed an automatic computer aided image analysis system for quantitatively assessing the changes of erythema and scaling severity of psoriatic lesions in long-term treatment [76]. In this method, firstly, scaling and erythema regions are segmented separately. The segmentation outcome is used for calculation of erythema area, scaling area and the whole lesion area which can be further used to assess the change of affected plaque regions. In order to validate the strength of the proposed system, an in-house image data set (14 for scaling change assessment, 17 for Erythema change assessment) was used. Obviously, the accurate change assessment depends on three main factors: (a) the imaging standard used to capture the images (important for alignment), (b) the efficiency of the used segmentation algorithm and (c) the feature representing the diseased regions.

1.5.2 Skin biopsy analysis

There exists a little research on the automatic analysis of psoriasis skin biopsy. Most of the works attempt to segment epidermis in a Whole Slide Histopathology Images (WSI) [44, 73, 74]. Lu and Mandal, firstly, applied a linear transformation on the RGB color vectors of every pixel for providing better discriminant information between the epidermis and dermis areas. Then the segmentation of the WSIs was done by using a global threshold method and shape analysis. The experiment was conducted on an in-house data set of 16 WSIs captured under 40X magnification on a Carl Zeiss MIRAX MIDI Scanning system. The proposed approach achieved about 92% sensitivity rate, 93% precision, and 97% specificity rate. Later, Lu and Mandal proposed a slightly advanced approach where an initial segmentation was done with the help of global thresholding and shape analysis and then the segmentation output was refined by template matching with the use of adaptive template intensity value. The performance of the proposed system was tested using an in-house data set of 62 skin WSIs and they got sensitivity of 97.99%, and precision of 96.00%. Haggerty et al. transformed the image pixels into the L*a*b* color space and then applied color normalization technique to enhance the algorithm's robustness to variations in the lighting and staining of the input images. Finally, thresholding based approach is chosen for segmentation and the segmentation is fine-tuned using a combination of morphological processing and object classification rules. This algorithm was tested on an in-house data set of 40 hematoxylin & eosin (H&E) stained skin biopsy images. This method achieved mean specificity of 97.7%, mean sensitivity of 89.4% and a mean accuracy of 96.5% on their data set. However, when a simple user interaction for selecting the tissue region is used, the specificity increases to 98.0%, the sensitivity to 91.0% and the accuracy to 96.8%. In summary, it can be commented that few age-old thresholding based approaches combined with suitable pre and post-processing are used for psoriasis biopsy image segmentation. These algorithms produce good performance on the test data set. But thresholding based approaches are not suitable for the biopsy images having imaging artifacts, staining variation, tissue variation, etc..

1.6 Thesis Outline and contribution

The goal of this thesis is to develop computer vision algorithms for analyzing psoriasis images. Following four distinct problems are addressed in this thesis:

- The importance of psoriasis image segmentation and the available approaches are described in the first paragraph of Section 1.5.1. Although several attempts are reported, still more attention is needed to achieve better results. Hence, the first problem chosen for this thesis is the segmentation of psoriatic plaques from skin images. This module can be used as the backbone for estimating the percentage of psoriasis affected area of human skin which is necessary for PASI area scoring.

- The available approaches for severity scoring from digital image of psoriasis are described in second to the last paragraph of Section 1.5.1. From this discussion, it is clear that although erythema scoring is attempted by some researchers, efforts for scoring of scaling and induration are rare. With the best of our knowledge, no work is available for induration scoring. Hence, the next problem chosen for this thesis is the severity assessment of psoriatic plaque with respect to erythema, scaling, and induration.

- Psoriasis causes structural abnormalities in the skin anatomy (see Section 1.3) which are reflected in biopsy sample. The automatic segmentation of skin layers in biopsy images is necessary for further analysis of histopathological conditions. So, this thesis in one of its chapters addresses the segmentation of skin layers in biopsy images.

- Munro's Microabscess is an important histopathological characteristic of psoriatic skin. According to clinical practice Munro's Microabscess is considered as the hallmark sign of psoriasis (see Section 1.3). So, detection of Munro's Microabscess in skin biopsy image is attempted in this thesis.

The key contributions of this thesis are:

- A new mixture model-based clustering algorithm is developed and a fully automated computational framework is designed for psoriatic plaque localization in skin images. The performance of the developed system is compared with several baselines segmentation approaches.

- Deep Convolutional Neural Networks (DCNN) are trained to produce automatic severity grades from skin images. Development of DCNN based Multi-task Learning framework and training of DCNN models with ordinal loss function are key contributions of this research.

- Two different approaches namely, (i) superpixel classification and (ii) Fully Convolutional Neural (FCN) network-based semantic segmentation are explored for automatic segmentation of a skin biopsy image into dermis, epidermis and non-tissue regions. The superpixel classification approach classifies the superpixel with a trained Deep Convolutional Neural Network (DCNN).

On the other hand, the trained U-Net produces pixel-wise semantic map for the pixels of the biopsy image. Development of DCNN based superpixel classification algorithms and development of Fully Convolutional Neural Network (FCN) for skin biopsy segmentation can be considered as significant contributions of this research.

- The pathological challenges of detection of Munro's Microabscess in skin biopsy have been addressed by developing an automated computational framework incorporating the latest advances in deep learning. The capsule network has been designed in such a way that it reduces the number of parameters drastically without sacrificing the performance. Moreover, the trained capsule network can detect/locate multiple copies of neutrophils in a weakly supervised approach. Furthermore, mega-pixel images are used instead of usual Giga-pixel ones to reduce the computational burden and thereby supporting a low-cost imaging system.
- A significant initiative is taken to develop a data set suitable for research on psoriasis image analysis. So, the development of the image data set is another contribution of this research. The skin images and biopsy slides of psoriasis are collected from a private clinic in Kolkata, India. There are camera held skin images from 80 patients and biopsy slides from 120 patients.

1.7 Organization of the thesis

In the remaining part of the thesis, there are five chapters. A brief summary of the chapter organization is given below:

- Chapter 2 provides a brief introduction about the finite mixture model for multivariate data vector. Next, it develops a finite mixture model of bi-variate circular-linear data vectors which is further referred as JCLMM (**J**oint **C**ircular **L**inear **M**ixture **M**odel). Finally, a JCLMM based computational framework is designed for psoriatic plaque localization in skin images. The applicability of JCLMM for the segmentation task is experimented and compared with the state of the art segmentation approaches.
- Chapter 3 discusses the problem of automatic severity scoring of psoriasis plaques. In this chapter, the effectiveness of Deep Convolutional Neural

Network is studied and compared with the classifiers build with texture feature.

- Chapter 4 focusses on skin layer segmentation in biopsy images. In this chapter, several deep neural architectures are experimented for segmenting psoriasis skin biopsy images into dermis, epidermis and non-tissue regions.
- Chapter 5 presents an automatic approach for detection of Munro's Microabcess in Stratum Corneum (SC) of human skin biopsy in order to realize a machine-assisted diagnosis of psoriasis. The proposed framework is built with a CapsNet based segmentation and classification network. The advantages of using CapsNet over traditional Deep Convolutional Neural Network (DCNN) are shown and an in-depth analysis of the capsule network for tissue specimen classification is performed.
- Finally, Chapter 6 concludes the findings of this thesis and highlights some scopes of future research in the related domain.

Chapter 2

Psoriatic plaque localization in skin images

Localization of psoriatic plaques in skin images is considered as an important step for doing further analysis like PASI area scoring and assessing the degree of erythema, scaling, and induration of the plaques. As psoriatic plaques create discoloration of human skin, the segmentation problem is treated as a color clustering problem. The target is to cluster image pixels with respect to color similarity. In literature, several color spaces are proposed for representing color of image pixels. Among them, some color spaces are linear (like RGB, CYMK, L*u*v*, L*a*b*) and some color spaces contain an angular component (like CIE LCH) subsequently, make the color vector circular-linear. Due to the scarcity of the clustering algorithm for clustering circular-linear data, only linear color spaces are used for color clustering.

This chapter introduces a new finite mixture model that incorporates a joint distribution of a circular and a linear variable by means of *circular copula* and offers a flexible architecture that can deal with heterogeneous margins for different mixture components. The model is termed as JCLMM (Joint Circular Linear Mixture Model). Later, the clustering of psoriatic image pixels in circular-linear color space is explored with the JCLMM. Finally, for plaque localization, a minimally supervised approach is employed.

The chapter is organized as follows. A brief introduction of the necessary mathematical background to understand the proposed mixture model is given in Section 2.1. Section 2.2 presents description of the proposed mixture model and methods for its parameter estimation. Section 2.3 discusses how the proposed mixture model is utilized for the present task. The experimental protocol, result, and discussion are given in Section 2.4 and 2.5. Finally, Section 2.6 summarizes the chapter.

2.1 Background

In pattern recognition, finite mixture models [82] are popularly used for clustering and classification. Perhaps the most common approach for mixture modeling is the Gaussian mixture model (GMM) [33, 65, 91]. This model uses Gaussian (univariate or multi-variate) distribution to describe each of its components. However, when the data are non-Gaussian in nature (i.e., follow a non-Gaussian distribution), the Gaussian mixture model (GMM) may produce poor performance. Noticing this fact, a number of non-Gaussian mixture models were also developed. In this context, some notable examples are the mixture of student's t-distributions [90], the generalized Dirichlet mixture model [9, 10] and copula-based approach [98] etc..

Most of the well-known mixture models, including the above examples, can handle only linear data. Contrary to that, in real-life applications, we sometimes come across certain multivariate data that possess circular characteristics along with the linear counterparts. For example, consider model-based color image segmentation. For this task, linear color spaces (e.g., RGB, CIE L*a*b*) are widely used in a number of articles [9, 33, 62]. On the other hand, only a few recent articles (e.g. [99, 100]) consider circular color spaces, such as CIE Lch, HSV, and design suitable mixture distributions. Among such studies, Roy et al. [99] first considered hue and chroma as independent and later proposed a semi-wrapped Gaussian mixture model (SWGMM) [100] which offers a correlated model for one circular and multiple linear variables. The SWGMM provides a mixture model for correlated circular-linear data. Each mixture component of SWGMM follows a multivariate semi-wrapped Gaussian distribution [5] which forms a joint distribution of one circular and multiple linear variables. The prime restriction of SWGMM is that it allows only predefined margins. But there are some scenarios when the mixture components themselves are different in terms of the underlying distribution. In such cases the inter-component heterogeneity cannot be captured by SWGMM and hence it may not always be an appropriate choice.

However, for the construction of a heterogeneous bi-variate distribution, statisticians often use the "copula" theory [88] to de-couple the margins from the joint distribution. In other words, any two different (or same) margin families can be coupled using a copula to form a joint distribution. The dependence structure is defined by different types of bi-variate copulas [88]. In this sense, copulas facilitate designing a heterogeneous joint distribution and which captures inter-component heterogeneity in a mixture distribution.

The proposed mixture model uses *circular copula* for construction of a heterogeneous bi-variate distribution. The necessary mathematical background to understand the proposed mixture model is given in the following subsections.

2.1.1 Introduction to Directional Statistics

A circular random variable Θ has domain $[0, 2\pi)$ and can be represented as a point on a unit circle [78]. The mean of circular random variables (*circular mean*) is the direction of the resultant vector by treating every variable as a unit vector. Suppose, $\theta_1, \ldots, \theta_n$ are circular observations. Their resultant vector is given by:

$$\boldsymbol{R} = (C, S) \text{ where } C = \sum_{i=1}^{n} \cos\theta_i \text{ and } S = \sum_{i=1}^{n} \sin\theta_i. \tag{2.1}$$

The direction of the resultant vector, i.e., the *circular mean* is computed by $\overline{\Theta}$ and defined as:

$$\overline{\Theta} = \arctan\left(\frac{S}{C}\right). \tag{2.2}$$

Besides the circular mean, a reasonable measure of *circular distance* between two points A and B with angles θ_1 and θ_2 is given by,

$$d(A, B) = 1 - \cos(\theta_1 - \theta_2). \tag{2.3}$$

The random variables having both circular and linear component are referred to as circular-linear variables. The definition of circular mean and distance are combined with linear mean and distance for circular-linear cases. Suppose $\boldsymbol{p_i}$ denotes the vector $(\theta, x_1, \ldots, x_d)^T$, where, $\boldsymbol{p_i} \in [0, 2\pi) \times \Re^d$. The mean ($\boldsymbol{\mu}^{cl}$) and distance for the vectors $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$, are computed using the following equations:

$$\boldsymbol{\mu}^{cl} = \left(\overline{\Theta}, \overline{\boldsymbol{X}}\right)^T \tag{2.4}$$

$$d(\boldsymbol{p}_i, \boldsymbol{p}_j) = 1 - \cos(\theta_i - \theta_j) + \|\boldsymbol{x}_i - \boldsymbol{x}_j\|. \tag{2.5}$$

Here the standard deviation of each $\boldsymbol{x}_k$ is assumed to be unity.

2.1.2 Distributions involving circular variables

One of the widely used circular distribution is the *von-Mises* distribution [78]. This distribution has many properties that are analogous to Gaussian distribution on the real line. The density of von-Mises distribution is given as follows.

$$f(\theta|\mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos(\theta - \mu)) \tag{2.6}$$

Here, $I_0(\kappa)$ is the modified Bessel function of the first kind of order zero with concentration parameter $\kappa \in \Re^+ \cup \{0\}$ and mean $\mu \in [0, 2\pi)$.

An alternative way to construct a circular distribution is by wrapping a linear variable around a circle. A linear variable X can be transformed to a circular variable Θ by reducing it by modulo 2π as: $\Theta = X(\text{mod } 2\pi)$. This operation corresponds to wrapping the real line around a circle of unit radius, accumulating probabilities over all the overlapping points $x = \theta, \theta \pm 2\pi, \theta \pm 4\pi, \ldots$. Consequently, for an uni-variate Gaussian distribution $\mathcal{N}_{\mu,\sigma}(x)$, the wrapped uni-variate Gaussian distribution [78] is defined as follows.

$$\mathcal{N}^w_{\mu^c,\sigma^c}(\theta) = \sum_{w \in \mathbb{Z}} \mathcal{N}_{\mu,\sigma}(\theta + 2w\pi), \; 0 \leq \theta < 2\pi. \tag{2.7}$$

$\mathcal{N}^w_{\mu^c,\sigma^c}(.)$ is uni-modal and symmetric about the circular mean μ^c. Here, σ^c is the circular standard deviation. In a similar way, a multivariate Gaussian distribution can be wrapped around a hyper-sphere to obtain its wrapped version. The mixture of multivariate wrapped Gaussian distributions was proposed by Agiomyrgiannakis and Stylianou in the context of modeling speech data [2].

In literature, a limited number of attempts are available that deal with circular-linear joint distribution. One of the former attempts to model such data was done by Mardia and Sutton [77]. They proposed a circular-linear distribution with a von-Mises margin for Θ. The margin for the linear variable, on the other hand, is complicated. A more flexible model was proposed by Johnson and Wehrly [59] in 1978 which provides a suitable framework for constructing bi-variate distributions of circular variables. This framework was further used by Shieh et al. [103] and Fernández-Durán [32] to construct circular-circular and circular-linear bi-variate distributions respectively. This framework is exploited to construct the JCLMM. Apart from this, the semi-wrapped Gaussian distribution, proposed by Bahlmann [5], is a recent circular-linear distribution. This distribution is the basis of SWGMM proposed by Roy et al. [100] in 2014. Although SWGMM permits correlated circular and linear distributions, its margins are pre-specified. But the developed copula-based joint distribution is free from such restrictions.

2.1.3 Copula based joint distribution

Copulas are used to describe the dependence between random variables. The following theorem due to Sklar [88] provides an excellent basis for the concept of the copula.

Theorem 1 (Sklar). Let F be a joint distribution function with marginal distributions F_1 and F_2. Then there exists a copula C such that for all $x, y \in [-\infty, \infty]$,

$$F(x, y) = C(F_1(x), F_2(y)). \tag{2.8}$$

If F_1 and F_2 are continuous, then C is unique; otherwise, C is uniquely determined on $Range(F_1) \times Range(F_2)$. Conversely, if C is a copula and F_1 and F_2 are distribution functions, then the function $F(.,.)$ defined by Eq. 2.8 is a joint distribution function with margins F_1 and F_2.

Since $F(.,.)$ is the distribution function, the joint density function $f(.,.)$ can be obtained by differentiating $F(.,.)$ with respect to X and Y. It is given by:

$$f(x, y) = c(F_1(x), F_2(y))f_1(x)f_2(y), \tag{2.9}$$

where, $f_1(.)$ and $f_2(.)$ are the density functions for X and Y respectively. The function $c(.,.)$ is the density of the copula $C(.,.)$.

2.1.4 Johnson and Wehrly's joint circular-linear distribution

Let Θ be a circular random variable with $\theta \in [0, 2\pi)$ and X be a linear random variable with $x \in (-\infty, \infty)$. Let us assume that Θ and X have distribution functions $F(\theta)$ and $F(x)$ respectively and the corresponding densities are $f(\theta)$ and $f(x)$. Then, if $g(.)$ is a circular distribution, according to Johnson and Wehrly [59], the joint distribution of Θ and X takes the following form.

$$f_J(\theta, x) = 2\pi g[2\pi F(\theta) + 2\pi F(x)] f(\theta) f(x). \tag{2.10}$$

The intrinsic property of Eq. 2.10 is the de-coupling of margins from the joint distribution. One can easily construct a bi-variate joint distribution for any choice of circular and linear marginal distributions. The density of the corresponding copula of Eq. 2.10 can be found by transforming the variables $(\Theta, X) \rightarrow (U, V)$, where $U = F(\theta)$ and $V = F(x)$. According to the probability integral transform, U and V follow the uniform distribution, as required to form a copula [88].

Since the Jacobian determinant $\left|\frac{\delta(u,v)}{\delta(\theta,x)}\right| = f(\theta)f(x)$, the density $c(u, v)$ of the copula is given by:

$$c(u, v) = 2\pi g[2\pi u + 2\pi v]. \tag{2.11}$$

In this article, we assume von-Mises distribution (see Section 2.1.2) with parameters $\tilde{\mu}$ and $\tilde{\kappa}$, as the joining density $g(.)$ of the two marginal densities. Hence, Eq. 2.11 takes the following form.

$$\psi(u, v|\tilde{\mu}, \tilde{\kappa}) = \frac{1}{I_0(\tilde{\kappa})} \exp\left[\tilde{\kappa} \cos(2\pi u + 2\pi v - \tilde{\mu})\right] \tag{2.12}$$

The copula $\psi(., .)$ is further referred to as *circular copula*. Using the circular-linear distribution $f_J(.)$, we construct a mixture distribution in the next section.

2.2 JCLMM: Joint Circular Linear Mixture Model

Let Θ and X be a circular and a linear random variable respectively. The mixture of K circular-linear distributions is defined by the following pdf.

$$f(\theta, x|\Xi) = \sum_{h=1}^{K} P_h f_J(\theta, x|h, \Xi_h), \tag{2.13}$$

where P_h ($0 < P_h < 1$ and $\sum\limits_{h=1}^{K} P_h = 1$) is the h^{th} component mixing proportion (prior probability), $f_J(\theta, x|h, \Xi_h)$ is the joint distribution representing the h^{th} component of the mixture and Ξ_h is the set of parameters of the h^{th} component. $\Xi = (\Xi_1, \ldots, \Xi_K, P_1, \ldots, P_K)$ is the set of parameters to be estimated.

2.2.1 Maximum likelihood parameter estimation

Let $\boldsymbol{\aleph} = \boldsymbol{p}_1, \ldots, \boldsymbol{p}_N$ be a finite set of N samples drawn independently from JCLMM. Here $\boldsymbol{p}_i = (\theta_i, x_i)$ is a circular-linear data. Then the parameters Ξ can be estimated using Expectation Maximization (EM) algorithm. Let us start with the distribution of hidden variables expressed as follows.

$$q(h|\boldsymbol{p}_i) = \frac{P_h f_J(\boldsymbol{p}_i|h, \Xi_h)}{\sum\limits_{l=1}^{K} P(l) f_J(\boldsymbol{p}_i|l, \Xi_l)}. \tag{2.14}$$

Computing $q(h|\boldsymbol{p}_i)$ constitutes the E-step of the algorithm. With the inclusion of the missing variables, the log likelihood function of the mixture distribution becomes:

$$\Phi(\boldsymbol{\aleph}, \Xi, \lambda) = \sum_{i=1}^{N} \sum_{h=1}^{K} \left[\ln\left(P_h f_J(\boldsymbol{p}_i|h, \Xi_h)\right)\right] q(h|\boldsymbol{p}_i) + \lambda \left(1 - \sum_{h=1}^{K} P_h\right), \tag{2.15}$$

where λ is the Lagrange multiplier. During the M-step of EM procedure, we maximize $\Phi(\boldsymbol{\aleph}, \Xi, \lambda)$ with respect to its parameters Ξ.

In the context of parameter estimation, García-Portugués et al. [37] designed a two-step procedure to estimate $f_J(.)$ parameters. In the first step, the margin parameters are obtained and in the next step the parameters of $g(.)$ are estimated. The same strategy is followed for estimating mixture parameters and each marginal likelihood are maximized separately. Afterwards, the two margins (u and v) are put into the copula $\psi(u,v)$ (Eq. 2.12) and estimate the parameters $\tilde{\mu}$ and $\tilde{\kappa}$. Now, after separating the marginal distributions, the mixture log likelihood (Eq. 2.15) can be written as follows.

$$\begin{aligned}\Phi(\boldsymbol{\aleph}, \boldsymbol{\Xi}, \lambda) &= \sum_{i=1}^{N}\sum_{h=1}^{K} q(h|\boldsymbol{p}_i)\left[\ln\left(\psi(F(\theta_i), F(x_i)|\tilde{\mu}_h, \tilde{\kappa}_h)\right)\right] \\ &+ \sum_{i=1}^{N}\sum_{h=1}^{K} q(h|\boldsymbol{p}_i)\ln f(\theta_i) + \sum_{i=1}^{N}\sum_{h=1}^{K} q(h|\boldsymbol{p}_i)\ln f(x_i) \\ &+ \sum_{i=1}^{N}\sum_{h=1}^{K} q(h|\boldsymbol{p}_i)\ln P_h + \lambda\left(1 - \sum_{h=1}^{K} P_h\right). \qquad (2.16)\end{aligned}$$

Here, the first term of Eq. 2.16 indicates the mixture log likelihood involving only the copula. The next two terms describe mixture log likelihood with only the margins. Then a priori probabilities can be found with $P_h = \frac{1}{N}\sum_{i=1}^{N} q(h|\boldsymbol{p}_i)$. Concerning the marginal and copula parameters, the estimation procedure is described in the following subsections.

After all the mixture parameters are estimated by the EM algorithm, the different components indicate different clusters in the data. Let these clusters be denoted by $C_1, C_2, \ldots, C_K$. The cluster assignment of a data point $\boldsymbol{p}_i$ is performed using the following maximum a posteriori (MAP) criterion.

$$\boldsymbol{p}_i \in C_j \text{ if } q(j|\boldsymbol{p}_i) > q(h|\boldsymbol{p}_i)\ \forall h \neq j. \qquad (2.17)$$

2.2.1.1 Parameter estimation for the margins

The construction of JCLMM allows any circular distribution for the first marginal and any linear distribution for the second marginal. However, this section describes the parameter estimation of JCLMM when the circular random variable Θ follows a von-Mises distribution (Eq. 2.6) and linear random variable X allows inter-component heterogeneity and can follow any of the two distributions - Gaussian or Beta distribution only.

To estimate μ_h and κ_h of the h^{th} mixture component of von-Mises distribution, the second term of Eq. 2.16 is maximized with respect to μ_h and κ_h and the expression for μ_h is as follows:

$$\tan \mu_h = \frac{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i) \sin \theta_i}{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i) \cos \theta_i}. \tag{2.18}$$

The expression for κ_h is, however, given in terms of the ratio of the modified Bessel functions:

$$A_2(\kappa_h) = \frac{I_1(\kappa_h)}{I_0(\kappa_h)} = \frac{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i) \cos(\theta_i - \mu_h)}{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i)}. \tag{2.19}$$

Since $A_2(\kappa_h)$ involves the ratio of Bessel functions, it is not possible to get an analytical solution. $A_2(\kappa_h)$ is a non-decreasing function, thus one may obtain κ_h by applying Newton-Raphson. Instead, to save computation time, we use the approximation suggested by Banerjee et al. [6]. Accordingly, $\kappa_h = (2\overline{r} - \overline{r}^3) \diagup (1 - \overline{r}^2)$, where $\overline{r} = A_2(\kappa_h)$.

The marginal distribution for the linear random variable X can be obtained by maximizing the third term in Eq. 2.16. As X allows inter component heterogeneity and can follow any of the two distributions, namely, Gaussian and Beta distribution, an appropriate distribution is to be selected during the iteration of EM procedure. The detail description of this selection method is explained in Sect. 2.2.1.4. Besides selecting the distribution, their respective parameters are to be estimated. For the Gaussian distribution, the parameters, ν_h and σ_h are estimated by the following expressions:

$$\nu_h = \frac{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i) x_i}{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i)} \tag{2.20}$$

$$\sigma_h^2 = \frac{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i)(x_i - \nu_h)^2}{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i)}. \tag{2.21}$$

For the other alternative, namely, the Beta distribution, the method of moments is used to maximize the third term of Eq. 2.16. Given the estimated moments ν_h

and σ_h, the parameters α_h and β_h are estimated by the following expressions.

$$\alpha_h = \nu_h \left(\frac{\nu_h(1-\nu_h)}{\sigma_h} - 1 \right), \tag{2.22}$$

$$\beta_h = (1-\nu_h) \left(\frac{\nu_h(1-\nu_h)}{\sigma_h} - 1 \right) \tag{2.23}$$

2.2.1.2 Parameter estimation for the circular copula

In order to estimate the circular copula parameters are plugged in the margins to the first term of Eq. 2.16. So all the margins are known for the copula and estimation of the parameters $\tilde{\mu}$ and $\tilde{\kappa}$ of $\psi(u, v)$ is needed. After differentiating the first term of Eq. 2.16 with respect to $\tilde{\mu}_h$ and $\tilde{\kappa}_h$ and setting them to zero, we obtain the following estimates.

$$\tan \tilde{\mu}_h = \frac{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i) \sin\left[2\pi F(\theta_i) + 2\pi F(x_i)\right]}{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i) \cos\left[2\pi F(\theta_i) + 2\pi F(x_i)\right]} \tag{2.24}$$

$$A_2(\tilde{\kappa}_h) = \frac{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i) \cos\left[2\pi F(\theta_i) + 2\pi F(x_i) - \tilde{\mu}_h\right]}{\sum_{i=1}^{N} q(h|\boldsymbol{p}_i)} \tag{2.25}$$

As before, we use approximation to invert $A_2(\tilde{\kappa}_h)$.

2.2.1.3 EM Initialization

The initialization of the clustering procedure can be done by random cluster assignment. However, for psoriatic plaque segmentation problem, firstly, K-Means algorithm is applied with (Θ, X) values to obtain an initial clustering of the data. The K-Means procedure uses the circular-linear version of mean and distance (i.e. Eq. 2.4 and 2.5). The number of mixture components is the same as the number of clusters. The initial estimation equations are similar to the equations 2.18- 2.23 presented in Sect. 2.2.1.1. However, since the K-Means perform a hard clustering of the data, therefore the expression for $q(h|\boldsymbol{p}_i)$ is defined as:

$$q(h|\boldsymbol{p}_i) = \begin{cases} 1 & \text{if } \boldsymbol{p}_i \in C_h \\ 0 & \text{otherwise} \end{cases} \tag{2.26}$$

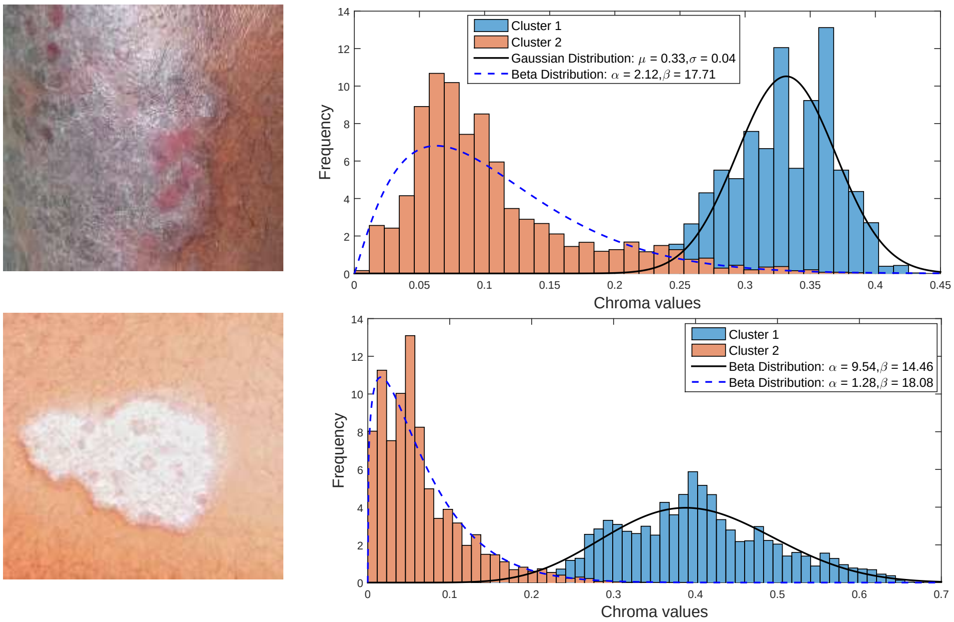
With this definition of $q(h|\boldsymbol{p}_i)$, the margin parameters are estimated using the method described in Sect. 2.2.1.1. The parameters for the von-Mises copula are estimated by Eq. 2.24 and Eq. 2.25. Finally, the prior probabilities are estimated using $P_h = \frac{N_h}{N}$, where N_h is the number of points in the cluster C_h.

2.2.1.4 Online selection of the distributions

In this thesis, for clustering skin pixels two alternative distributions (Gaussian and Beta) for chroma vector is assumed. Subsequently, each component of the mixture model can be formed using any of these two alternatives. This flexibility allows inter-component heterogeneity for the mixture model. However, considering each combination, M^K different mixture models can be formed assuming M margin candidates (here $M = 2$). The obvious issue arises is among all these mixture models which is the best fit. The simplest but laborious way is to examine all of them. Instead of this brute force approach, the following strategy is opted in this thesis. After E-step, the marginal distribution which have maximum (between Gaussian and Beta) log-likelihood (the second term of Eq. 2.16) is chosen. For K clusters, the M-step requires $K(M + 1)$ margin estimations. This selection procedure, however, does not guarantee to produce the best-fitted mixture model. Yet, compared to the brute force approach (which requires $2K \times M^K$ estimations), this method requires much less number of estimations. Also, this strategy is easily applicable for a large set of margin alternatives.

2.3 Psoriatic plaque localization using JCLMM

In this thesis, the plaque localization task is carried out in two steps consisting of skin image segmentation followed by a minimally supervised plaque region detection algorithm. The segmentation task is performed by the proposed clustering algorithm (JCLMM) and the plaque region is detected by a probabilistic approach. The detail description of them is described in the following subsections.



(a) (b)

Figure 2.1: Examples of heterogeneous nature of chroma distribution. Beta and Gaussian distributions can adequately approximate the heterogeneous structure.

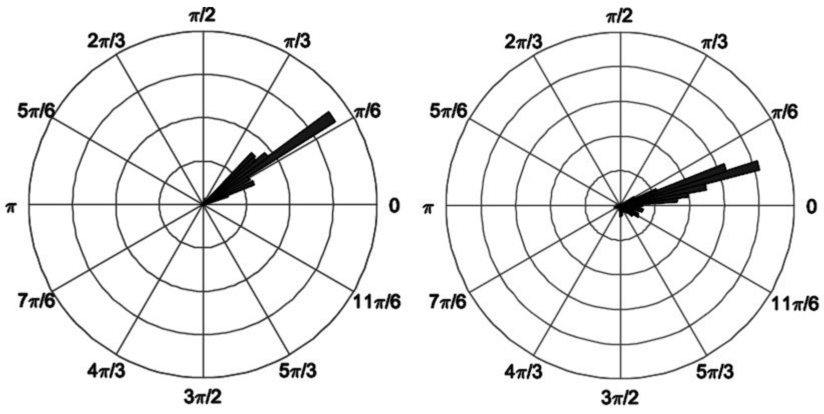
2.3.1 Plaque segmentation

Hue carries an important color information so a color space containing hue (like CIE Lch) as one of the components is chosen for clustering the image pixels. However, due to its directional property, a circular distribution is required to characterize the hue component. Previously introduced Semi-Wrapped Gaussian Mixture Model (SWGMM [100]) can deal with the circular-linear data vectors. But the limitation of SWGMM is that this algorithm assumes that circular marginal follows Wrapped-Gaussian distribution whereas other linear marginals follow Gaussian distribution. But this assumption is not true for all scenarios of the present image segmentation task. For example, consider the skin images shown in Fig. 2.1 (a). The histogram for corresponding chroma values is shown in Fig. 2.1 (b). It is somewhat clear that the chroma values can be approximated by the mixture of a Gaussian and a Beta distribution and some scenarios by the mixture of only Beta distributions. Therefore, the mixture components themselves may be different in terms of underlying distribution. This inter-component heterogeneity can not be captured by SWGMM and hence the proposed JCLMM is an intelligent choice.

The hue and chroma components of an image pixel carry crucial information so image pixels represented by the hue, chroma pair are considered to perform clustering. The image pixels are considered to be independent random samples following a K component JCLMM. The parameters of JCLMM are estimated using EM methodology as described in Sect. 2.2.1. The E-step and M-step are iterated until the corresponding log-likelihood stabilizes. After clustering, all pixels are assigned to a specific cluster using Eq. 2.17.

2.3.2 Plaque localization

After segmentation, when respective image pixels are mapped to spectral clusters, the image contains several spatially disconnected regions where each region is represented by a connected component. Each individual component corresponds either to plaque or to normal skin. Hence, the remaining task is to identify all components correspond to the plaque. The region properties of every connected zone are exploited for tagging the diseased regions. Average color vectors (hue and chroma) are chosen as region properties as they are computed easily after clustering the image pixels. A binary classifier is suitable for the identification task. But there is no well-known binary classifier that classifies circular-linear data. Comparison of probability densities with respect to the diseased and non-diseased pixel color (hue and chroma) distributions cannot be done because of the unknown margin of the second component (chroma). So, the problem is treated as an anomaly detection problem.



(a) (b)

Figure 2.2: The distribution of hue for (a) healthy skin pixels and (b) psoriasis skin pixels.

Fig. 2.2 shows that the hue distributions of the healthy and the diseased skin pixels differ significantly. So, for simplicity, a statistical hypothesis test is developed on

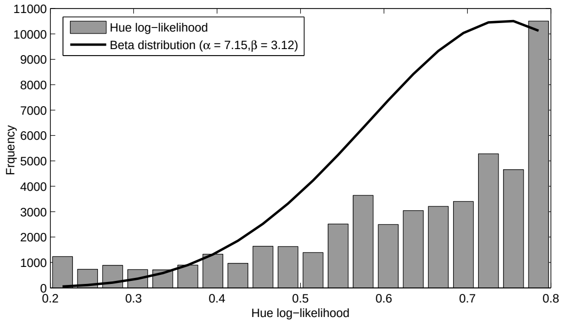


Figure 2.3: The distribution of log-likelihood statistic for ground truth hue values and the corresponding Beta distribution of normal skin pixels.

the basis of average hue values to detect plaque regions. This is a minimally supervised way to perform the classification task.

As the range of the hue of the normal healthy skin pixels is narrow with respect to psoriasis pixels so a confidence region is constructed on the hue distribution of normal skin. To do so, a repository of normal skin pixels is developed and their hue values are extracted. Since hue is circular in nature, the von-Mises distribution is used to characterize it. From each connected component or region of a segmented image, its mean hue value is obtained. Afterward, a statistical test is performed under the null hypothesis that if the mean hue lies *inside* the confidence region of normal skin. If this becomes *false*, the mean hue does not indicate normal skin and hence the corresponding component is identified as a diseased region.

The problem with this approach is that no analytical expression for the confidence interval is available for von-Mises distribution. Generally, a bootstrap method is employed to get the confidence interval. Instead, an alternative methodology is followed. Since, each normal skin pixel is assumed to follow the same von-Mises distribution, log-likelihood statistics is computed for these pixels. The distribution of these log-likelihood statistics is shown in Fig. 2.3. In the next step, a parametric model is obtained to approximate this distribution. It is observed that this distribution is left tailed. Without loss of generality, the statistic values are normalized in $(0, 1)$ and a Beta distribution is fitted to model them (Fig. 2.3). For the concerned data set, the parameters α and β of the Beta distribution are obtained as 7.1560 and 3.1153 respectively. Evidently, $\alpha > \beta$ which implies a left tailed distribution. Given this distribution, if the log-likelihood statistic of the mean hue of a component lies at the tails, the null hypothesis is rejected. In such a situation, the corresponding component is tagged as the diseased region (a plaque).

Given a hue value, the proposed hypothesis assesses whether or not this hue comes from a skin portion. Therefore the aim is to successfully reject all hue values that do not correspond to skin. These rejected hue values are assumed to correspond to diseased regions. The proposed strategy is advantageous in many ways. First, normal skin images are easily available as compared to diseased skin images. Moreover, this strategy may help in building a personalized normal skin pixel database, which may later be used for detecting other kinds of skin disease. Again, skin color is usually different in different geographical regions. From this point of view, a normal skin database can also be customized.

2.4 Experimental Protocol

2.4.1 Data set

The skin images used in this experiment are captured using digital cameras from the plaque psoriasis patients of the eastern part of India. The images are captured in an uncontrolled environment by layman photographers with different viewing angles, distances, varying backgrounds, and lighting conditions. The present data set contains images collected from eighty (80) patients of different age groups (0-20: 6, 21-40: 20, 41-60: 40, Older than 60: 14). We consider both male (60) and female (20) patients having different skin color tone. These images are manually cropped to remove background by a photo editing software. Afterward, dullrazor [66] algorithm is applied for removing skin hair. Every cropped image contains diseased skin zone along with normal skin zone. The images vary in their sizes (154×151 to 1800×1700). For quantitative evaluation, manual annotation (each pixel is labeled with either *normal skin* or *plaque*) is performed and verified by the dermatologists. The time required for annotating an image depends on several factors like size of the image, area of the affected regions, distribution of the affected regions in the image, expertise of the human annotator. In the present scenario, the annotator having moderate expertise in handling the annotating tool takes about 30 minutes to prepare the ground-truth for an image of size 1800×1700. The data set contains ground truth information for seventy-five (75) images[1]. Among them fifteen (15), images are used for training pixel database generation and test data set is prepared taking images from the rest of the sixty (60) images.

---

[1] A part of the data set is available at `http://www.isical.ac.in/~utpal/resources.php`

2.4.2 Baselines

In order to make a comparative assessment of JCLMM with the competing approaches, a number of well-known approaches are considered as baseline results. The clustering techniques, namely, Fuzzy C-Means (FCM) [21, 92], Gaussian Mixture Model (GMM) and Semi-Wrapped Gaussian Mixture Model (SWGMM) [100]. FCM and GMM are performed in linear (CIE L*u*v*) color space whereas SWGMM is performed in circular-linear (CIE Lch) color space.

Apart from clustering approaches, three widely used image segmentation algorithms are considered. These are JSEG [27–29], Mean-shift [20] and level-set approach of Chan Vese (LS-CV) [14]. Both JSEG and Mean-shift approaches use clustering followed by region merging. The level set is a contour evolution-based approach. An initial contour (rough object boundary) is provided as input and this contour evolves with the image forces. Although for the present problem this method needs user interaction, no post-classification approach is needed to localize the disease regions.

Besides the unsupervised approaches, the performance of JCLMM is also compared with supervised pixel classification-based approaches. In this regard, the K- Nearest Neighbour (KNN), Support Vector Machine (SVM) and Random Forest (RF) classifiers are considered. For all of these approaches, the same training images are used to develop pixel repositories of healthy skin pixels and diseased skin pixels. All the above classifiers are trained on spectral information represented in CIE L*u*v* color space to perform semantic labeling of the image pixels.

2.4.3 Implementation Details

The in-house implementations of the clustering algorithms (FCM, GMM, and SWGMM) is used for this research. The number of clusters in a particular image was chosen empirically and remain same for an image for all the clustering methods. Moreover, to minimize initialization effects, all clustering algorithms were replicated 15 times and the best results are reported.

For comparing with JSEG color image segmentation approaches, the implementation of Vision Research Lab, UC Santa Barbara [2] is used. On the other hand, the

---

[2] JSEG code available at: `http://vision.ece.ucsb.edu/segmentation/jseg/software`

Edge Detection and Image SegmentatiON (EDISON) software[3], is used for Mean-shift. The parameters of JSEG, Mean-shift algorithms are obtained by multiple trials.

For KNN, RF and level set procedures, Matlab R2016a implementations are used. The best result for the KNN classifier is found at $K = 1$. The number of decision trees in the RF is empirically set to 200. For SVM-based classification, LIBSVM [18] package is used with Radial Basis Function (RBF) kernel using default parameters.

2.4.4 Evaluation Metrics

The proposed system performs the psoriatic plaque localization task in a two-step process. Firstly, color partitioning and then localizing the diseased regions. So, both the color partitioning (called evaluation of segmentation) and localization are evaluated separately using a different set of evaluation metrics. Finally, the performance of the proposed method is compared to other baselines with respect to the same metrics. Some approaches which are efficient to localize the diseased regions in a single process are exempted from the first type of evaluation (evaluation of segmentation) and performance is compared only on the basis of the set of metrics used for evaluating localization performance.

2.4.4.1 Evaluation of Segmentation

In order to make a quantitative comparison of different algorithms with respect to ground-truthed segmentation, three metrics, namely, the Probabilistic Rand Index (PRI) [111], the Global Consistency Error (GCE) [81], and the Variation of Information (VoI) [83] are used. The PRI is a measure of the similarity between two different ways for cluster assignment. It is the probability that an arbitrary pair of data points has consistent labels in both ways of cluster assignment. A higher value for PRI implies better performance. The GCE measures the extent to which one segmentation can be viewed as a refinement of the other. A lower value for GCE implies better performance. The VoI measures the sum of information

---

[3] Mean-Shift code available at: http://coewww.rutgers.edu/riul/research/code/EDISON/

loss and information gain between the two different ways of cluster assignment. Lower values of VoI indicate greater similarity.

2.4.4.2 Evaluation of plaque localization

Five metrics (Four metrics adopted from [75] and Precision) are used for quantitative performance evaluation of plaque localization. All these metrics lie in $[0, 1]$ and higher values indicate better performance. Let the detected diseased region be Q_D, non-diseased region be Q_{ND} and the corresponding ground truths are G_D and G_{ND}, respectively. With these definitions, the evaluation metrics are explained below.

Sensitivity or Recall: In the present problem, the target is to correctly identify the diseased pixels in the skin images. So, Recall is chosen as an accuracy metric which is represented by the percentage of the diseased pixels that are correctly identified among all the diseased pixels and defined as:

$$\text{Recall } = \frac{|Q_D \cap G_D|}{|G_D|} \tag{2.27}$$

Precision: In the present problem, the algorithm which is biased towards classifying an image pixel as diseased pixel is capable of providing good Recall but in that case false positives will be increased. So, Precision is considered. Precision measures the probability that a detected diseased region is truly a diseased region and defined as:

$$\text{Precision } = \frac{|Q_D \cap G_D|}{|Q_D|} \tag{2.28}$$

Dice's coefficient: In the present problem, the algorithm which can classify the diseased pixels correctly with minimal false positives is intended. So, the Dice coefficient is chosen as an accuracy measure. It is a function of the above two accuracy metrics. Dice's coefficient is used to measure the similarity between actual and detected diseased pixel sets and defined as:

$$\text{Dice } = 2 \times \frac{|Q_D \cap G_D|}{|Q_D| + |G_D|} \tag{2.29}$$

Specificity: Specificity is an accuracy metric which measures the percentage of the non-diseased pixels that are correctly identified among all the non-diseased pixels and defined as:

$$\text{Specificity} \; = \frac{|Q_{ND} \cap G_{ND}|}{|G_{ND}|} \tag{2.30}$$

Spatial Support (SS): The definition of Spatial Support (SS) for given c partitions of the data (SS) is defined as:

$$\text{SS} \; = \frac{1}{c} \times \sum_{i=1}^{c} \frac{|Q_i \cap G_i|}{|Q_i \cup G_i|} \tag{2.31}$$

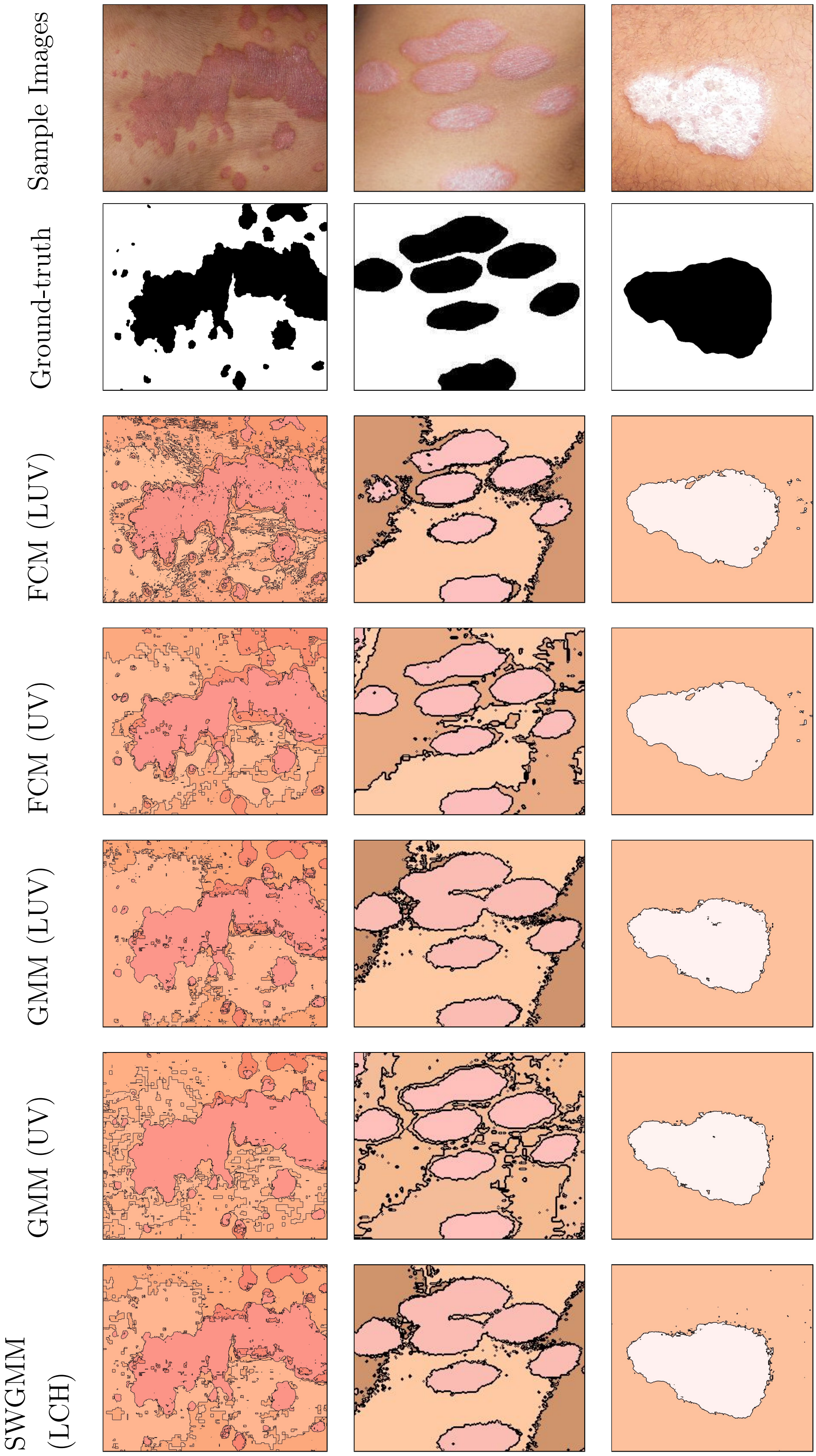
Here, the result for i^{th} partition is Q_i and the corresponding ground truth is G_i. For our experiments $c = 2$ (normal skin and plaque).

2.5 Results and Discussion

The experimental results are organized into two halves in order to emphasize segmentation and plaque localization, respectively. Consequently, sect. 2.5.1 includes segmentation results and Sect. 2.5.2 presents localization outcome. Results are analyzed qualitatively as well as quantitatively using the commonly used evaluation metrics.

2.5.1 Segmentation results

The segmentation outcomes of all competing clustering algorithms along with the original images and the corresponding ground-truths for three images are shown in Fig. 2.4. In Fig. 2.4, It is clear that the clustering algorithms partition the input images with respect to color similarity.



Continued on next page

Figure 2.4– continued from previous page

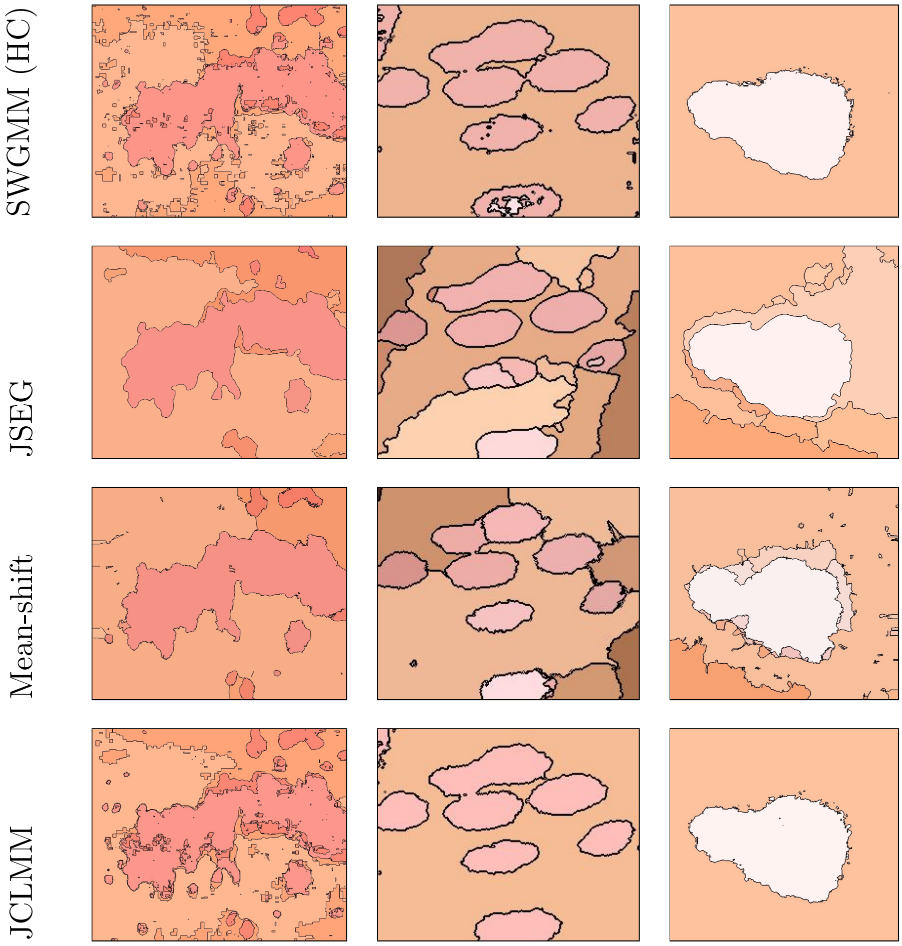


Figure 2.4: Segmentation outcomes produced by different algorithms, for three example images.

In order to make a quantitative comparison among the above algorithms, the mean of the evaluation scores (i.e., PRI, GCE, and VoI) for all the test images are considered and. The average scores are presented in Table 2.1. According to Table 2.1, JCLMM performs either better or equivalent than all other comparing algorithms with respect to all evaluation metrics. The result shows that the mixture model-based clustering approaches (GMM, SWGMM, JCLMM) perform much better than FCM clustering. This is due to the fact that the mixture-model based approaches are less sensitive to the noise than the fuzzy clustering. In comparison between GMM and SWGMM, it can be stated that using same number of color components [4] both methods are performing with the same efficiency. This

---

[4] For both cases: GMM (LUV) Vs SWGMM (LCH) & GMM (UV) Vs SWGMM (HC)

Table 2.1: Average segmentation performance. The symbols =, > and < respectively indicate whether JCLMM is equivalent, superior or inferior. These statistics are obtained based on a Wilcoxcon signed rank test at 95% significance level.

Method	PRI	GCE	VoI
FCM(LUV)	$0.6677^{>}$	$0.1474^{>}$	$1.3526^{>}$
FCM(UV)	$0.6958^{>}$	$0.1287^{>}$	$1.2239^{>}$
GMM(LUV)	$0.7258^{=}$	$0.1003^{=}$	$1.1271^{=}$
GMM(UV)	$0.7544^{=}$	$0.0900^{=}$	$1.0589^{=}$
SWGMM(LCH)	$0.7309^{>}$	$0.1096^{=}$	$1.1219^{=}$
SWGMM(HC)	$0.7535^{>}$	$0.1004^{=}$	$1.0609^{=}$
JSEG	$0.7467^{=}$	$0.1034^{=}$	$1.3820^{>}$
Mean-shift	$0.7476^{=}$	$0.0955^{=}$	$1.4896^{>}$
JCLMM	0.7604	0.0962	1.0433

shows that clustering image pixels using circular-linear color space (CIE Lch) is a viable alternative way.

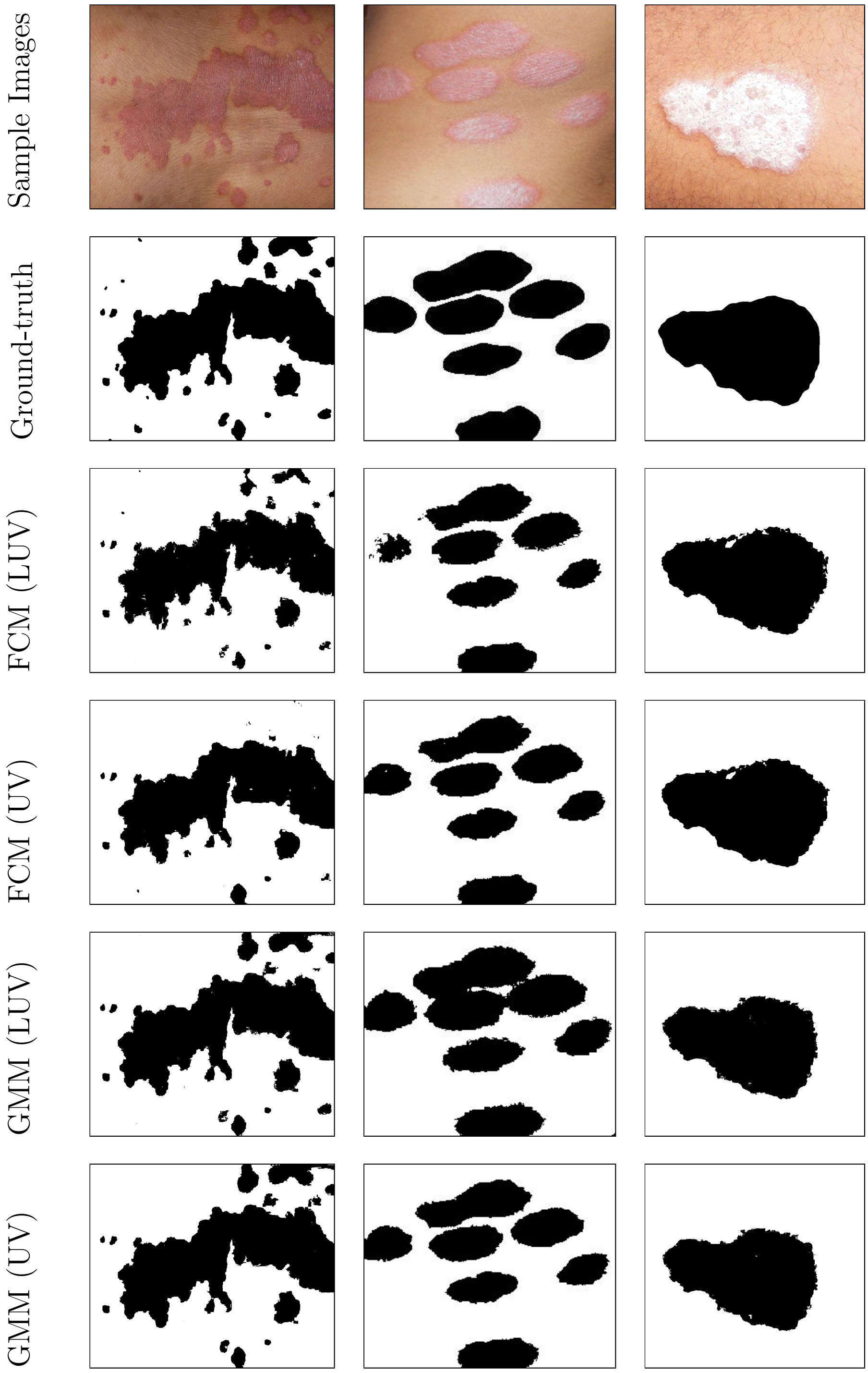
The experiment shows that in case of FCM, GMM, and SWGMM, the inclusion of the lightness component does not improve the performance at all, rather performance is degraded sometimes. This justifies the application of JCLMM for the chosen problem. According to the present data set among seventy-five (75) images, forty (40) images contain at least one cluster where the chroma of the image pixels follows the Beta distribution. This fact explains the importance of treating the heterogeneity of the chroma distribution separately for the present problem.

In comparison with the other segmentation algorithms, i.e., JSEG, Mean-shift it is observed that JCLMM performs better with respect to VoI and equivalent with respect to the other two metrics.

The clustering approaches studied in this paper, do not take into account the spatial information of the image pixels. Each pixel is assumed to be sampled independently from a particular model. In particular, for the mixture models, the EM algorithm employs a pixel color based clustering, producing several small regions inside an image. On the other hand, JSEG and Mean-shift consider the spatial closeness of the pixels and hence yield a smaller number of regions. However, the used plaque localization methodology considers each spatially connected components i.e., each region to examine whether it represents a plaque. Thus, spurious regions do not significantly affect the performance of localization procedure.

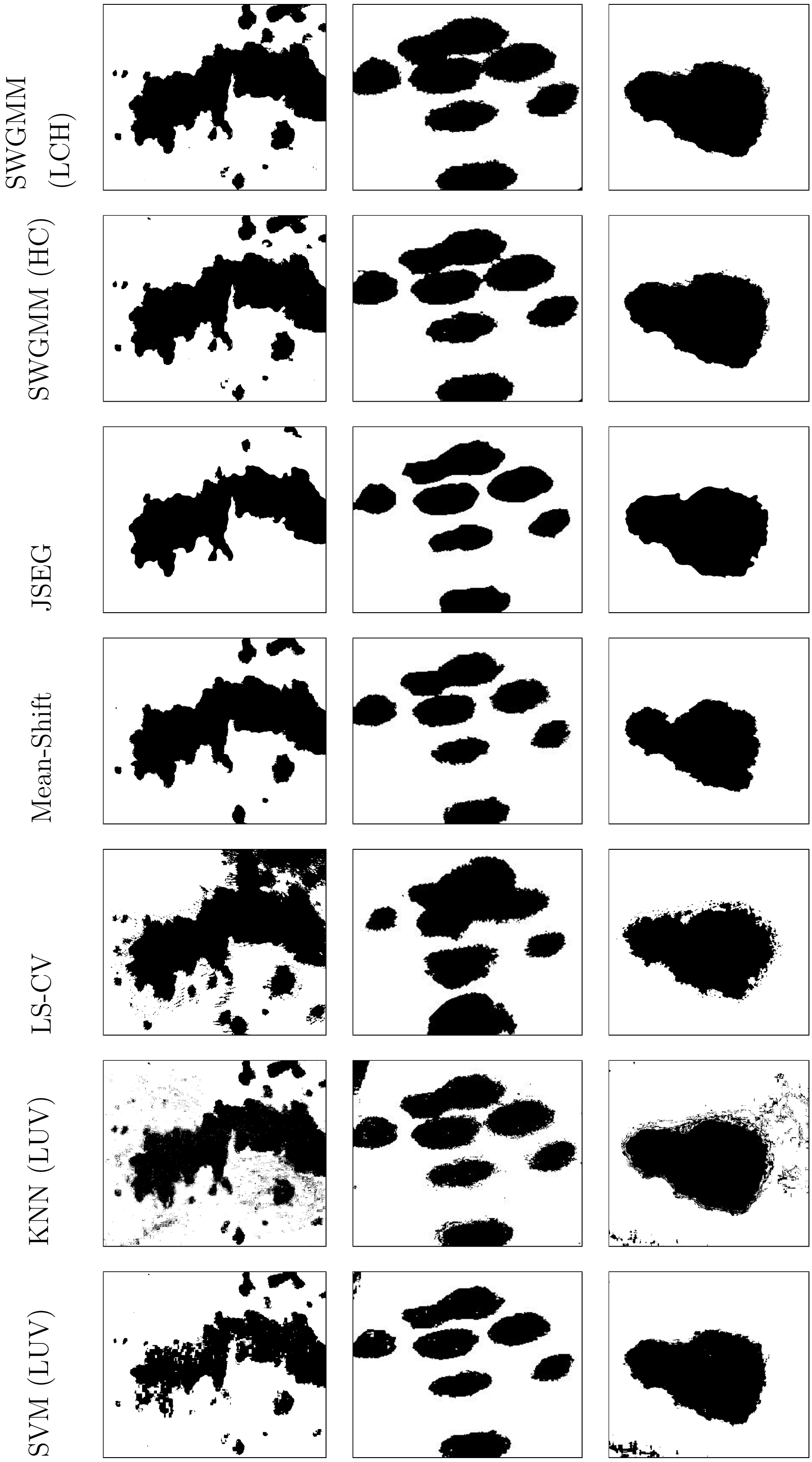
2.5.2 Plaque localization results

In order to evaluate the plaque localization performance, at first, we make a visual inspection of the example images given in the first row of Fig. 2.4. For those images, the detected plaque regions are presented in Fig. 2.5.



Continued on next page

Figure 2.5– continued from previous page



Continued on next page

Figure 2.5– continued from previous page

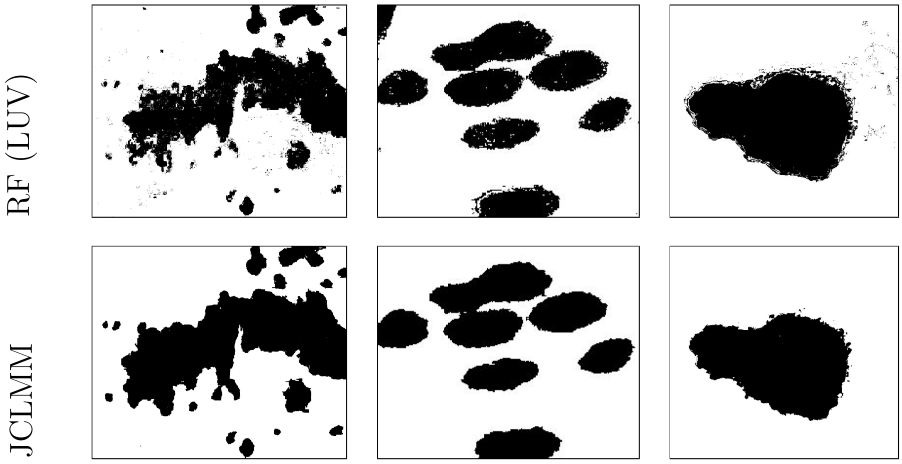


Figure 2.5: Plaque localization outcomes produced by different algorithms, for some example images.

By visual comparison with the ground truth, it can be said that the proposed localization procedure successfully retrieves the spatial components corresponding to the psoriatic plaque. However, since it is difficult to compare different algorithms visually, the quantitative evaluation (see Sect. 2.4.4.2) is performed. Table 2.2 presents the average evaluation results. Note that bi-channel color clustering makes use of UV components for FCM and GMM and HC components for SWGMM, whereas tri-channel color clustering makes use of LUV components for FCM and GMM and LCH components for SWGMM.

Table 2.2 shows that the performance of JCLMM is better than both bi-channel or tri-channel color component based FCM algorithm, however, the performance of FCM is equivalent with respect to JCLMM in terms of specificity. JCLMM achieves equivalent performance with respect to both bi-channel and tri-channel GMM. JCLMM is performing better than SWGMM (LCH) if Dice coefficient, specificity, and spatial support are considered. However, JCLMM is performing better with respect to SWGMM (HC) if the Dice coefficient, specificity, and spatial support are considered. But no clustering algorithm could outperform JCLMM for the present problem.

In comparison with the popularly used color image segmentation algorithms, JCLMM beats JSEG with respect to all metrics except average recall and specificity. However, according to the statistical significance test, the performance of JCLMM is

Table 2.2: Average plaque localization performance. The symbols =, > and < respectively indicate whether JCLMM is equivalent, superior or inferior. These statistics are obtained based on a Wilcox-con signed rank test at 95% significance level.

Approaches	Method	Recall	Precision	Dice	Specificity	SS
Clustering based approaches	FCM(LUV)	0.8113 $^{>}$	0.8580 $^{=}$	0.8057$^{>}$	0.8285 $^{>}$	0.7283 $^{>}$
	FCM(UV)	0.8365 $^{>}$	0.8685 $^{=}$	0.8168 $^{>}$	0.8220 $^{>}$	0.7412 $^{>}$
	GMM(LUV)	0.8789 $^{=}$	0.8704 $^{=}$	0.8381 $^{=}$	0.8871 $^{=}$	0.7930 $^{=}$
	GMM(UV)	0.8693 $^{=}$	0.8788 $^{=}$	0.8406 $^{=}$	0.8824 $^{=}$	0.7946 $^{=}$
	SWGMM(LCH)	0.8977 $^{=}$	0.8631 $^{=}$	0.8620 $^{>}$	0.8603 $^{>}$	0.7889 $^{>}$
	SWGMM(HC)	0.8984 $^{=}$	0.8619 $^{=}$	0.8593 $^{=}$	0.8702 $^{>}$	0.7944 $^{>}$
Segmentation approaches	JSEG	0.9145 $^{>}$	0.8687 $^{=}$	0.8710 $^{=}$	0.8907 $^{=}$	0.8124 $^{=}$
	Mean-shift	0.7829 $^{>}$	0.9106 $^{<}$	0.8354 $^{=}$	0.9080 $^{=}$	0.8125 $^{=}$
	LS-CV	0.8109 $^{>}$	0.8830 $^{=}$	0.8052 $^{>}$	0.8244 $^{>}$	0.7157 $^{>}$
Supervised approaches	KNN	0.8282 $^{=}$	0.8860 $^{=}$	0.8505 $^{=}$	0.8249 $^{>}$	0.7280 $^{>}$
	SVM	0.9016 $^{=}$	0.9010 $^{=}$	0.8763 $^{=}$	0.7818 $^{>}$	0.7490 $^{>}$
	RF	0.8965 $^{=}$	0.9015 $^{=}$	0.8536 $^{=}$	0.7919 $^{>}$	0.7415 $^{>}$
Proposed approach	JCLMM	0.9041	0.8817	0.8797	0.8855	0.8180
	JCLMM-MF	0.8998	0.9384	0.9042	0.9028	0.8289

better when the recall is considered otherwise there is no statistically significant difference between these two algorithms.

In comparison with Mean Shift, a density-based color clustering approach, the performance is equivalent with respect to Dice Coefficient, Specificity and Spatial Support whereas Recall is improved but Precision is worst. Poor Recall but good Precision justifies the fact that for present data set, in a skin image, the Mean Shift is detecting only a subset of the diseased pixels with few false positives.

In comparison with the basic Chan Vese level set (LS-CV) approach, the proposed method performs better for all metrics except Precision (performs with the same efficiency). The heterogeneous nature of the plaque regions makes it difficult for the LS-CV in guiding the contour in the appropriate direction.

Next, the supervised approaches are considered. These approaches directly assign a pixel to plaque or normal skin based on training samples. From Table 2.2, it is observed that the performance of SVM, RF, and KNN are equivalent with JCLMM with respect to recall, precision and dice coefficient. However, these algorithms achieve poor performance with respect to Specificity and Spatial Support. This justifies the fact that these supervised approaches are capable of recognizing a smaller subset of the diseased pixels.

According to the discussion in Section 2.4.4.2, Spatial Support is a metric which evaluates how much an algorithm is efficient in detecting diseased pixels with less false positives. The proposed approach is performing better than all of the other algorithms considered in this paper. But the performance of JCLMM is equivalent to GMM(LUV), GMM (UV), JSEG, Mean Shift. However, one may be able to get better performance by Mean-shift or JSEG approach using better parameters but finding a right set of parameters is a daunting task whereas the proposed approach does not depend on such parameter tuning. So, it can be concluded that JCLMM, not only performs with equal efficiency with respect to several popularly used approaches but also outperforms several baseline approaches for the chosen problem.

In Table 2.2, it is observed that several smaller (dotted) normal regions are there inside a diseased region. Therefore, a median filter is applied, which reassigns those regions (image pixels) as diseased pixels. The corresponding localization results are presented in the last row (JCLMM-MF) of Table 2.2. As compared to the naive version of JCLMM, although the opted post-processing technique reduces the recall marginally but improves the overall localization performance.

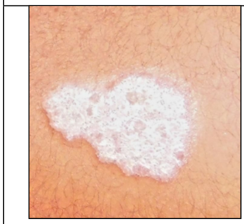
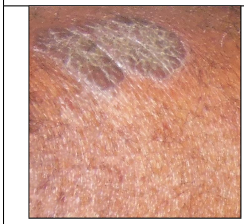
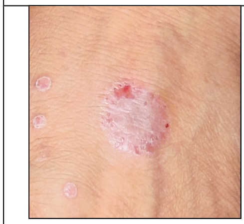
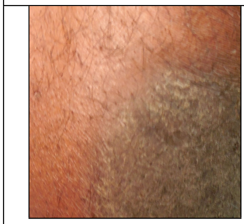
2.5.3 JCLMM: Engineering challenges

This chapter develops a theoretically efficient algorithm for clustering circular-linear data having heterogeneous marginal distribution. However, from the engineering perspective, several challenges are there to execute the algorithms. The number of parameters used in our model and the complexity in terms of mathematical modeling are higher than traditional linear multi-variate data clustering algorithms. Hence, in some cases algorithms are not converging and in many cases execution is suffering from data overflow or underflow problem. To get an idea about the performance of the present implementation, Five (5) images and the mean and standard deviation of their performance for 15 trials with respect to GCE, PRI and VoI are shown in Figure 2.3. So, we can conclude that there is a huge scope to focus on such engineering challenges. For the time being, the clustering algorithms are executed multiple trials (15) to discard such anomalous outcomes and the best results are compared with the chosen baselines.

2.6 Summary

In this chapter, a mixture model-based clustering algorithm (JCLMM) for circular-linear data is presented. Further, this model is used for developing a fully probabilistic framework for the localization of psoriatic plaque in skin images. In the proposed system, the JCLMM is used to perform an image segmentation in a circular-linear color space, consisting of hue and chroma. On the basis of these experiments, it can be concluded that a circular-linear color space may provide a viable alternative to the commonly used linear color spaces. Moreover, JCLMM can consider heterogeneous marginal distributions, which occurs frequently for chroma channel for the present data set. The present study shows that the proposed mixture model can play a crucial role to deal with circular-linear color spaces.

Table 2.3: Few sample images and mean and standard deviation (std) of PRI, GCE, and VoI scores for 15 trials.

Images	PRI (mean± std)	GCE (mean± std)	VoI (mean± std)
	0.7762±0.0889	0.0582±0.0185	0.9231±0.1135
	0.7545±0.0919	0.0406±0.0028	0.9275±0.1066
	0.6410±0.0818	0.1166±0.0044	1.4634±0.1364
	0.6410±0.0739	0.0368±0.0082	1.0197±0.0936
	0.8161±0.0755	0.1078±0.0199	0.9524±0.1047

Chapter 3

Severity Grading of Psoriatic Plaques

Psoriasis is a life-long disease. No drug is available yet to cure psoriasis completely but some drugs are there to control the severity. Still, these drugs show variable degrees of response and have adverse effects on different patients. In this regard, severity grading of psoriatic plaque is an important task for clinical decision making and clinical outcome evaluation. But the current practice of severity grading depends on the experts' experience, efficiency and suffers from both inter and intra-observer variation. Hence, the development of robust system for standardizing the severity assessment procedure is highly intended.

This chapter discusses the problem of automatic machine analysis based severity scoring of psoriasis disease. Three different disease parameters namely, erythema, scaling, and induration are considered for such severity grading. Given an image containing a psoriatic plaque the task is to predict severity scores for all the three parameters. As a pattern recognition problem, this is indeed a challenging task due to the unavailability of good feature descriptors, the diversity of the human skin color, presence of hair, wrinkles on the human skin, etc.

The severity scoring for each disease parameter can be modeled as a classification task using an effective feature descriptor. However, due to the scarcity of such robust feature descriptors, the development of Deep Convolutional Neural Network (DCNN) based approach is attempted for achieving the task. These disease severity parameters (i.e. erythema, scaling, and induration) are not completely uncorrelated to each other. Hence, apart from viewing this task as three different single-task learning (STL) problems (i.e. three different classification problems), a new Multi-Task learning (MTL) is also presented where the three classification tasks are treated as interdependent and thereby the neural net is trained accordingly. The trained network grade an image with respect to all the three parameters simultaneously.

In the present severity assessment task, there exists an ordinal relationship among the severity grades. Suppose, the actual and predicted severity score of a misclassified image is K and K_1 respectively. Then, the classifier would have the least possible absolute difference $|K - K_1|$. Motivated from this, several DCNN based ordinal classifiers are experimented further for the present severity assessment tasks.

In literature, several approaches are available for designing neural network-based ordinal classifier. The most common strategies are conversion to ordinal regression [34], decomposition to multiple binary classification problems, and loss adjustment or loss transform strategies. The ordinal regression approaches are highly sensitivity towards the presence of noise in test images. On the other hand, the success of binary decomposition method depends on the performance of every binary classifier and in many scenarios, the marginal classifiers suffer from data imbalance problems. Hence, the loss adjustment strategies are very much attractive research field for ordinal classifier design. In this regard, 2-Wasserstein distance-based CNN trained for NIR-VIS face [50], 1-Wasserstein distance or Earth Mover's Distance based CNN trained for age estimation [52] are significant contributions available in the literature. For the present problem, we train DCNN model with Earth Mover's Distance based loss function due to its simplicity, applicability, and effectiveness. This is the first time where an ordinal classifier is designed for severity assessment task for dermatological images.

The organization of this chapter is as follows. Section 3.1 presents a brief introduction of the necessary background knowledge to understand the proposed methodologies. Dataset description is given in section 3.2. The detail descriptions of the two different approaches considered for this thesis are given in a separate section. Section 3.3 discusses the DCNN based Multi-Task Learning approach used in this thesis, on the other hand Section 3.4 discusses the DCNN based ordinal classification approach used in this thesis. Finally, Section 3.5 concludes the paper along with a few remarks related to the future extension of the present research.

3.1 Background

In this section, the necessary machine learning methods for understanding the research presented in this chapter are discussed.

3.1.1 Multi-Task Learning

In machine learning, Multi-Task Learning (MTL) has emerged as a promising approach to improve generalization performance by learning the correlated tasks in parallel using a shared representation; i.e. the features which are learned for a particular task can help other tasks to learn better [13].

In the recent past, some research works have been reported where deep CNN based MTL has been explored [116][67][114]. In these approaches, the network models learn different targets for related multiple tasks but share some common layers. Furthermore, these models are also flexible to deal with different error functions for each task. The key advantage of the MTL model is the reduction of testing time along with less memory requirement due to presence of shared layers as well as it exploits the interdependence and thereby improves the overall accuracy.

3.1.2 Image classification by DCNN

Nowadays, Convolutional Neural Network (CNN) [49, 63, 107] is widely used for image classification tasks as it relieves the researchers from designing hand-engineered feature descriptors and automatically develops powerful mathematical models directly from the training images. The efficiency of CNN attracts medical image analysis researchers for developing automated tools for clinical assistance [71, 80]. Recent literature shows that CNN based approaches are efficient to deal with medical image analysis problems like disease recognition [22, 38, 39, 60, 84, 117], severity gradation [36, 54], etc.

A CNN model is made up of multiple processing units and each processing unit consists of trainable weights and biases. In the training phase, the network parameters are updated by comparing the distribution of predicted class labels with the actual class labels of the training images. A brief description of the traditional categorical cross-entropy (CCE) loss and the mean square error (MSE) loss functions are given below.

Suppose, for a C-class ($C > 2$) single-label image classification problem, the ground truth of a k^{th}-class image is given by a binary vector G of length C such that $G_i = 1$ whenever $i = k$ and 0 otherwise. The output of the CNN is a probability distribution P of length C such that its i^{th} entry (P_i) represents the predicted probability of the i^{th} class. Now the definition of CCE loss and the MSE loss are given in Equation 3.1 and 3.2.

$$\mathcal{L}_{CCE} = -\sum_{i=1}^{C} G_i \, ln(P_i) \qquad (3.1)$$

$$\mathcal{L}_{MSE} = \sum_{i=1}^{C} (P_i - G_i)^2 \qquad (3.2)$$

It can be seen from Equation 3.1 and 3.2, CCE and MSE loss ignores the ordinal relationship since CCE only considers the probability of the correct class and MSE is invariant to permutation of probabilities of incorrect classes.

3.1.3 Ordinal Classification

Machine learning algorithms for classification problems commonly assume that the class values are unordered. However, in many practical applications, the class values do exhibit a natural order. For example, when learning how to grade a skin image with respect to severity class. The goal of such classifier is to produce the least possible absolute difference between the actual and the predicted class. From this motivation, in this thesis, following three different approaches are experimented for making the DCNN as ordinal classifier.

Conversion to regression problem: The very simple approach to ordinal classification is the conversion of the class value into a numeric quantity and applies a regression learner to the transformed data, translating the output back into a discrete class value in a post-processing step [34]. In this thesis, the severity scores are projected into C equal partitions in $[0, 1]$ and the CNN is trained in such a way that the i^{th} class ($i = 1, 2, ...C$) image outputs a value in $[\frac{i-1}{C}, \frac{i}{C}]$.

Conversion to binary decomposition problem: In this approach, a C class classification problem is decomposed into $C-1$ binary classification problems where the i^{th} classifier predicts whether an image has the classification label more than i or not. Then these trained classifiers are used to predict class labels of the test images. The key disadvantage of this method is that the classification scheme produces the correct classification scheme if all of the constructed binary classifiers produce correct classification labels. Moreover, the decomposed classifiers which are trained for detecting the marginal class labels will suffer from data imbalance problem.

Training DCNN with Earth Mover's Distance (EMD) based loss function: In this thesis, for the severity assessment task, the Earth Mover's Distance (EMD) or 1-Wasserstein distance based loss function is proposed to use for training the deep models [52]. The target of this method is to transform the softmax output obtained from the classification layer of the DCNNs into a uni-modal distribution so that it can capture the ordinal relationship among the class labels.

Mathematically, let X_i^{CDF} denote the i^{th} element of the cumulative distribution of X then the loss function is as follows:

$$\begin{aligned}
\mathcal{L}_{EMD} &= \sum_{i=1}^{C}(P_i^{CDF} - G_i^{CDF})^2 \\
&= \sum_{i=1}^{C}(\sum_{j=1}^{i} P_j - \sum_{j=1}^{i} G_j)^2 \\
&= \underbrace{\sum_{i=1}^{k-1}\left(\sum_{j=1}^{i} P_j\right)^2}_{\mathcal{A}} + \underbrace{\sum_{i=k}^{C}\left(\sum_{j=1}^{i} P_j - 1\right)^2}_{\mathcal{B}} \qquad (3.3)
\end{aligned}$$

where k is the correct class. According to Equation 3.3, when $i < k$, increasing the value of P_i increases the value of $\mathcal{A}$ whereas when $i \geq k$, increasing the value of P_i decreases the value of $\mathcal{B}$. Since, in $\mathcal{A}$, P_i occurs $(k-i)$ times hence, the value of $\mathcal{L}_{EMD}$ increases as $|i-k|$ increases. Similarly, in $\mathcal{B}$, for $i \geq k$, P_i occurs $(C-i)$ times hence, the value of $\mathcal{L}_{EMD}$ increases as $|i-k|$ increases. Thus the proposed loss function trains the network in such a way that the class label farthest from the actual class gets less probability.

3.2 Dataset

The skin images used in this experiment are collected from psoriasis patients using commonly used digital cameras. These images are captured in an uncontrolled environment by layman photographers with different view angles, distance, and varying backgrounds. The data set is built with images collected from eighty (80) patients of different age groups (0-20: 6, 21-40: 20, 41-60: 40, Older than 60: 14). We consider both male (60) and female (20) patients having different skin color tone. These images are captured from every body region of a patient where the psoriasis lesion is found. So, several images are collected from most of the patients.

These images are manually cropped into several non-overlapping images containing only the skin portion. The cropped images vary in their sizes. For evaluation purposes, manual annotation (severity grading with respect to erythema, scaling, and induration) is done by a domain expert. The data set contains ground truth information for seven hundred and seven (707) images.

3.3 Severity grading of psoriatic plaque using DCNN based Multi-Task Learning

Deep models are data-hungry. So, in many real-life applications where data volume is limited pre-trained models are used for fine-tuning. Pre-trained models are nothing but a DCNN model trained with the Imagenet dataset. In the present application number of training images is limited. So, instead of designing a custom deep architecture, we use AlexNet as backbone network for the Multi-Task Learning (MTL) based computational framework used in this research. The pre-trained weights are used for network initialization. In this section, the detail description of the Multi-Task Learning (MTL) based computational framework designed for automatic severity grading of psoriatic plaque is given.

3.3.1 Network Architecture

The block diagram of the proposed Multi-Task learning network architecture is shown in Fig 3.1. The diagram shows that, firstly, the training images are given as

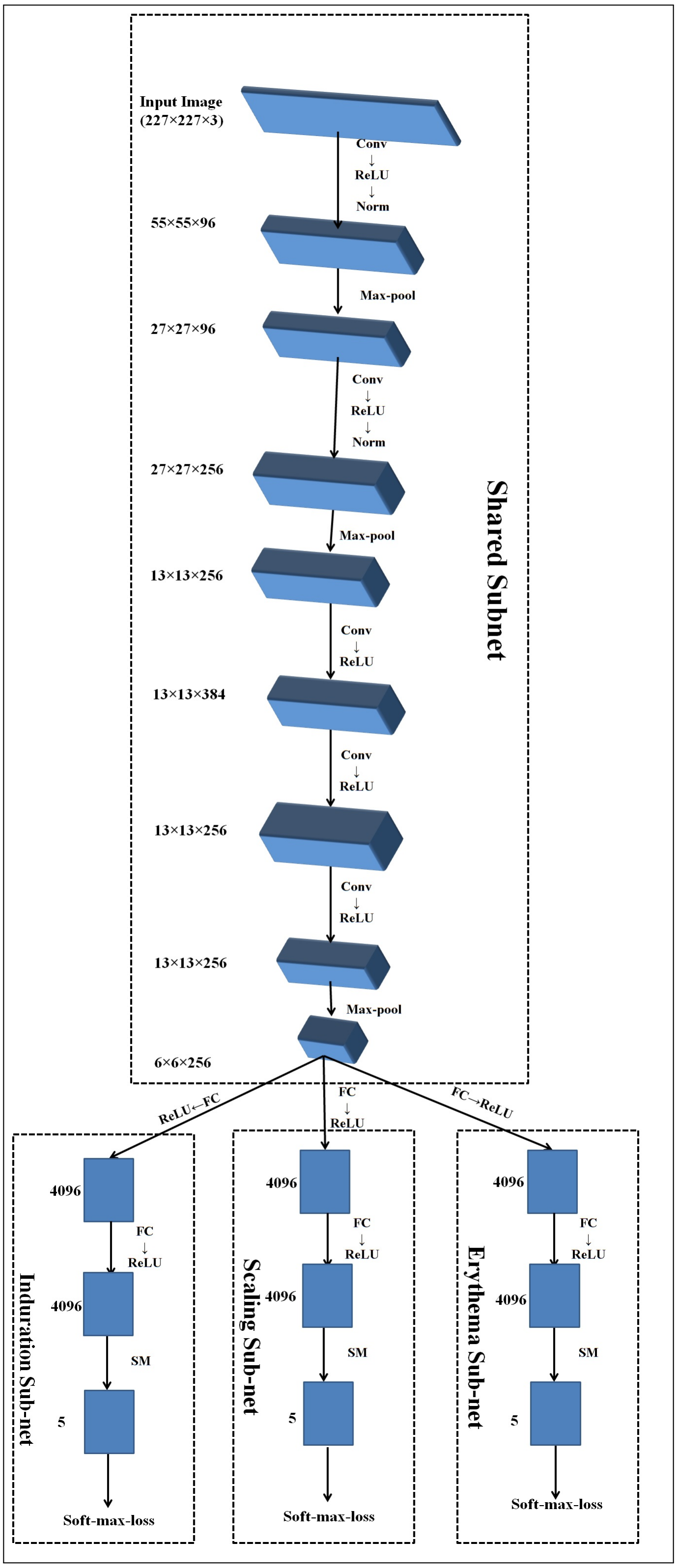


Figure 3.1: DCNN structure used for Multi-Task Learning

input to the shared sub-net. Then the produced feature map is given as input to the three different sub-nets namely erythema sub-net, scaling sub-net, induration sub-net. Finally, erythema, scaling and induration scores are obtained from the last layer of respective sub-nets.

The shared sub-net of our proposed MTL architecture contains five convolutional layers, five ReLU[1] activation layers, three pooling layers, two normalization layers. The first convolutional layer has 96 kernels of size $11 \times 11 \times 3$ with a stride of 4 pixels. The feature map from the first convolution layer is passed through ReLU activation layer, normalization layer, and max-pooling (3×3 with a stride of 2 pixels) layer. Then the feature map from the max-pooling layer is given as input to the second convolution layer. The second convolution layer has 256 kernels of size $5 \times 5 \times 48$ with stride of 1 pixel. Then the feature map from the second convolution layer is passed through ReLU activation layer, normalization layer, and max-pooling (3×3 with stride of 2 pixels) layer. The feature map from the pooling layer is given as input to the third convolutional layer. The third convolutional layer has 384 kernels of size $3 \times 3 \times 256$ with stride of 1 pixel. The output from ReLU layer after third convolution layer is given as input to fourth convolution layer which has 384 kernels of size $3 \times 3 \times 192$ with stride of 1 pixel. The output from ReLU layer after fourth convolution layer is given as input to fifth convolution layer which has 256 kernels of size $3 \times 3 \times 192$ with stride of 1 pixel. Then the feature map from fifth convolution layer is passed through ReLU activation layer and max-pooling (3×3 with stride of 2 pixels) layer.

The erythema, scaling and induration sub-nets of our MTL architecture have same network structure. They consist of two fully-connected layers and a soft-max classification layer. The two fully connected layers have 4096 nodes and the soft-max layer has 5 nodes corresponding to five severity scores. After fully-connected layers ReLU activation layers are kept. Note that the present DCNN architecture gains motivation from the AlexNet [63].

3.3.2 Cost Function

The proposed network is flexible to allow different loss functions for each task (or sub-net). However, for our problem domain we consider same cost function for

---

[1] ReLU(x)=max(0,x)

all tasks. The cross-entropy loss is maximized in this thesis. The mathematical expression for the loss is given below:

$$\epsilon = -\sum_{i=0}^{4} t_i \log y_i \tag{3.4}$$

where, $t_i = 1$ if the sample has ground-truth severity i, otherwise $t_i = 0$; y_i is the computed soft-max probability of the data for i^{th} class.

The following aggregated loss is minimized for the proposed Multi-Task Learning (MTL) architecture:

$$\epsilon_{tot} = (\lambda_1 \times \epsilon_E) + (\lambda_2 \times \epsilon_S) + (\lambda_3 \times \epsilon_I) \tag{3.5}$$

where ϵ_E is the loss from the erythema sub-net, ϵ_S is the loss from the scaling sub-net; ϵ_I is the loss from the induration sub-net; λ_i, $i = 1, 2, 3$ are linear weights, currently, all are set to 1.0 for the present problem. Obviously, if only one of the λ_is is non-zero and others are set to 0 then this network will be exactly same as Single task Learning (STL) network.

3.3.3 Data pre-processing and Training

The proposed network takes a 3-channel image of size 227×227. So, the psoriasis images are normalized to 227×227 before giving it as input to the network. Images are represented in RGB color space and stochastic gradient descent is used for learning network parameters.

3.3.4 Network Initialization, Parameter Settings

The used network architecture as explained in Section 3.3.1. The weight vectors of both the shared sub-net and the three non-shared sub-nets are to be initialized. As the volume of the data set is very smaller with respect to the number of parameters in the network, the entire network is initialized by the weight vectors obtained from the trained AlexNet using Imagenet challenge data set [63]. Then the feature map from the last layer of the shared sub-net is computed. Using this feature map the parameters of the three non-shared sub-nets are tuned separately. Next, the network is trained in a Multi-Task Learning (MTL) framework.

Every model is trained using a batch size of 256 images, momentum of 0.9, weight decay of 0.0005 and a learning rate of 0.001. The images in every batch are randomly shuffled before training. The learning process is stopped when the network parameters are saturated (when the training error and objective function are not changing significantly for more than 10 epochs). In case of the present MTL, it is found that 200 epochs are enough for saturating the model parameters.

3.3.5 Baselines

The performance of the proposed system is evaluated and compared with three relevant baseline approaches. A brief description of them is given below.

Experiment1: The performance of the proposed system is compared with **Local Binary Pattern (LBP)** [89] texture feature-based approach. This texture descriptor is a widely used hand-crafted feature descriptor. The classifier models are developed using Support Vector Machine (SVM) and K-Nearest Neighbour (KNN) classifiers to classify the test images using the LBP feature.

Experiment2: The feature maps from **Fully-connected Layer6 (FC6)** and **Fully-connected Layer7 (FC7)** of the pre-trained model of AlexNet (trained on Imagenet challenge data set) are also extracted and used as image feature descriptors. These extracted features are used separately to build the classifier models using Support Vector Machine (SVM) and K-Nearest Neighbour (KNN) classifiers to classify a test image.

Experiment3: The MTL framework is used as **Single task Learning (STL)** framework by taking only one sub-net (erythema sub-net, scaling sub-net or induration sub-net) at a time for dealing with a particular type of scoring. Thus three STL frameworks are developed to get three labels of an image separately. The initialization of the weight vectors is done in the same way as it was done for the MTL framework. The network weights are tuned using the training images in our data set. The learning process is stopped when the network parameters are saturated (when the training error and objective function are not changing significantly for more than 10 epochs). In case of STL, it is found that 100 epochs are enough for saturating the model parameters.

3.3.6 Evaluation strategy

To compare the performance of the proposed model with the other approaches, 7-fold cross-validation is done and for each fold, the percentage of the correctly classified images is computed. It is important to mention that for each approach same data samples are kept in a specific fold and every fold is trained with same hyper-parameters.

Table 3.1: Experimental Results (mean ± standard deviation): DCNN based Multi-Task Learning

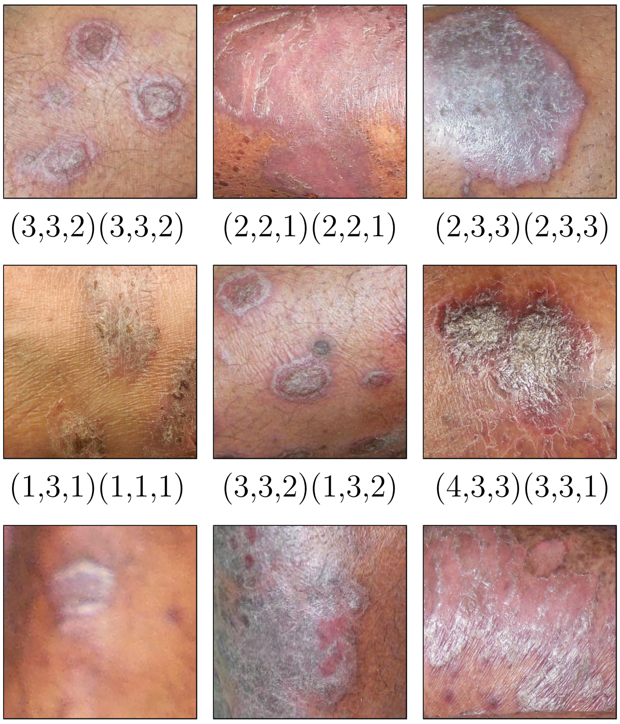
Method	Erythema Acc(%)	Scaling Acc(%)	Induration Acc(%)
LBP+KNN	24.2857 ±6.9621	28.1429 ±2.2113	34.5714 ±4.2132
LBP+SVM	39.6040 ±2.2135	38.8967 ±2.3467	45.6860 ±3.2129
FC6+KNN	25.8571 ±3.7642	25.0000 ±3.1673	35.8571 ±2.1816
FC6+SVM	53.6068 ±8.9521	51.2023 ±4.1265	57.2843 ±5.2137
FC7+KNN	26.8571 ±9.0105	27.7143 ±9.0012	32.0000 ±7.1719
FC7+SVM	57.5672 ±9.2135	49.7878 ±4.5756	58.9816 ±3.3457
STL	59.6888 ±10.5360	**58.9816** ±4.3955	60.6789 ±3.8224
Proposed MTL	**60.6789** ±4.5985	54.8797 ±4.1107	**61.1032** ±2.8826

3.3.7 Result and Discussion

Results based on the mean and standard deviation of percentage of the correctly classified images over 7-folds are reported in Table 3.1. Table 3.1 shows that for the present problem, an SVM classifier provides much better accuracy than a KNN classifier. While evaluating the effectiveness of the feature sets, it shows that the KNN classifier is not able to shed any light whereas the SVM classifier shows that Local Binary Pattern (LBP) feature descriptor performs inferior to the off-the-self DCNN feature descriptors (using FC6 and FC7 Fully Connected Layers of AlexNet tuned by the Imagenet data set). This fact justifies the use of DCNN based approaches where feature learning is done automatically.

While comparing FC6 and FC7 feature sets, the SVM classifier shows that the FC7 feature set performs better for erythema and induration whereas scaling accuracy is better if the FC6 feature set is used. Finally, it is cleared from Table 3.1 that the accuracies are further improved when the pre-trained AlexNet is further trained using our own training data.

Generally, a deep model needs a large volume of data to build a classifier but Table 3.1 shows that the use of pre-trained model deals with the data scarcity problem and outperforms traditional feature-based classifier in a huge margin. This justifies the use of pre-trained deep model for the present research problem. It is worth to mention that all 7-classifiers (for every fold) are trained with same hyper-parameters and the mean and standard deviation of classification accuracy are reported in Table 3.1.The low standard deviation of classification accuracy for all 7 folds justifies that the trained models are generalized.



(3,3,2)(3,3,2) (2,2,1)(2,2,1) (2,3,3)(2,3,3)

(1,3,1)(1,1,1) (3,3,2)(1,3,2) (4,3,3)(3,3,1)

(2,3,2)(2,2,0) (3,4,2)(1,3,1) (3,3,2)(2,2,1)

Figure 3.2: Sample Images and their actual scores and corresponding predicted scores. The first vector under every image represents actual scores (in Erythema, Scaling and Induration order) and the second vector represents the corresponding predicted scores by the proposed MTL framework.

3.3.7.1 STL vs MTL

The proposed MTL model takes 0.05 sec for grading a single test image whereas equivalent STL models take 0.09 sec for the same task. So it can be concluded that the proposed MTL framework is not only time (testing phase) and space-efficient than three different STL frameworks, it also performs better in scoring erythema and induration.

For a given image, an ideal system should make correct severity score prediction for all the three severity parameters. So, a comparative study of STL and MTL frameworks is conducted with respect to this criterion. The experimental result shows that MTL achieved an average accuracy of 25.60% which is 1.13% more than the score given by the STL framework. This shows that though the MTL or STL shows moderate efficiency (about 60%) in predicting severity when the parameters are considered separately but predicting severity for all the three parameters together is indeed a difficult task and further research is needed to improve the performance. Some sample images and their actual and machine predicted severity scores are given in Fig 3.2. The figure shows images where predictions are perfect for all the three disease parameters, images where one or two parameters and some images where the severity of no parameters is predicted correctly.

3.3.7.2 Results with (±1) Tolerance

As severity rating of psoriatic plaque suffers from inter and intra-observer variability. So, prediction of severity scores with a maximum deviation of ±1 (for any of erythema, scaling and induration scoring) from actual scores may be considered. This motivates us to compute the results allowing ±1 tolerance in severity prediction. The results show the accuracies are substantially increased if such tolerance is allowed. Average accuracies of 93.64% for erythema scoring, 93.78% for scaling scoring and 93.78% for induration scoring are achieved by the MTL framework. While comparing with STL results with tolerance, it shows that MTL outperforms STL in scoring erythema by 2.97% and induration by 0.42%, whereas STL outperforms MTL in scoring scaling by 0.99%.

Given an image, when accuracies in correctly predicting all the three disease parameters are considered together, it is found that MTL achieves an accuracy of

85.29% with ±1 tolerance and this score is 2.12% more than that of STL framework with same tolerance. Note that the ±1 tolerance has been given to each of the three parameters.

3.3.7.3 Combining STL and MTL Results

According to the experiment, the erythema and induration score prediction networks from the MTL framework and scaling score prediction network from the STL network provide the best results. Based on this observation, if one more experiment is conducted where MTL is used for scoring erythema and induration and STL is used for scoring only scaling and then the results are combined. The best accuracies for scoring the three disease parameters separately are shown in boldface in Table 3.1. A slight improvement is possible for correctly scoring all the three parameters together. Average accuracy of 27.58% is achieved by following such a combination and when the ±1 tolerance is allowed this accuracy is increased to 86.28%.

3.4 Severity grading of psoriatic plaque using DCNN based ordinal classifier

In this section, the detail description of the ordinal classification based computational framework designed for automatic severity grading of psoriatic plaque is given.

3.4.1 Network:

As the data volume is small, the training of a Deep Convolutional Neural Network (DCNN) from scratch does not produce satisfactory performance. Fine-tuning of pre-trained network is opted for the present classification task. Two pre-trained networks ResNet-50 [49] and Mobile Net [53] trained on the Image-net dataset are considered for fine-tuning. ResNet-50 is chosen due to its impressive performance on the Image-net classification. The Mobile Net is chosen as it contains comparatively fewer parameters but produces a good performance on Image-net classification.

3.4.2 Data pre-processing and Training

Both ResNet-50 and Mobile Net takes a 3-channel image of size 224 × 224. So, the psoriasis images are normalized to 224 × 224 before giving it as input to the network. Images are represented in RGB color space.

In this thesis, the performance of the developed system is reported on the basis of 7-fold cross-validation. The model is trained with the stochastic gradient descent optimizer using a batch size of 4 images, momentum of 0.9, weight decay of 10^{-6} and with the learning rate of 0.001. For every fold, the network is trained 10 times and the trained model which ends with minimal loss is chosen for prediction of test images. Horizontal and vertical flipping augmentation are used for improving the generalization ability of the classifiers.

3.4.3 Baselines

In this thesis, the performance of the CNN trained with proposed ordinal loss minimization is compared with four baselines. The first two CNNs are trained with traditional categorical cross-entropy ($\mathbf{CNN_{CCE}}$) and mean-square error ($\mathbf{CNN_{MSE}}$) loss minimization. In the third approach ($\mathbf{CNN_{Regr}}$), the severity scores are projected into C equal partitions in $[0, 1]$ and the CNN is trained in such a way that the i^{th} class ($i = 1, 2, ...C$) image outputs a value in $[\frac{i-1}{C}, \frac{i}{C}]$. The last approach is the decomposition ($\mathbf{CNN_{Decomp}}$) of the C class classification problem into $C - 1$ binary classification problems where the i^{th} classifier predicts whether an image has classification label more than i or not. Then these trained classifiers are used to predict class labels of the test images. It is worth mentioning that the binary CNNs are trained with binary cross-entropy loss minimization. Among all considered baselines, only the last two classifiers can capture the ordinal relationship among the labels.

3.4.4 Performance evaluation metrics

The performance of the trained CNN is measured with three different evaluation metrics- (i) Mean Accuracy (MA), (ii) Mean Absolute Error (MAE) and (iii) Kendall's τ_b. The value of MA lies in [0,1] and a higher value represents better performance. A lower value of MAE represents better performance. On the other hand, Kendall's τ_b measures the association or rank correlation between two measured quantities. The τ_b value lies in $[-1, +1]$, where, $+1$ is the maximum agreement between the prediction and the ground truth class labeling, 0 represents no correlation between them and -1 represents maximum disagreement. MAE and Kendall's τ_b are used since MA ignores the ordinal relationship between predicted and actual class for a misclassified image.

Suppose, there are N test images having a discrete class label in $[1, C]$. Let Y_i^p, Y_i^g represent the predicted and the ground-truth class label of the i^{th} test image respectively. Then the mathematical expressions of these metrics are shown in Equation 3.6, 3.7, 3.8.

$$MA = \frac{1}{N}\sum_{i=1}^{N}\delta(Y_i^g, Y_i^p); \qquad \delta(x,y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Y_i^g - Y_i^p| \tag{3.7}$$

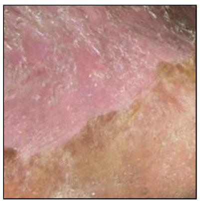
$$\tau_b = \frac{\sum_{i,j=1}^{N}\hat{C}_{i,j}C_{i,j}}{\sqrt{\sum_{i,j=1}^{N}\hat{C}_{i,j}^2 \sum_{i,j=1}^{N}C_{i,j}^2}}, \text{where}$$

$$\hat{C}_{ij} = \begin{cases} 1, & \text{if } Y_i^p > Y_j^p \\ -1, & \text{if } Y_i^p < Y_j^p \\ 0, & \text{otherwise.} \end{cases} \qquad C_{ij} = \begin{cases} 1, & \text{if } Y_i^g > Y_j^g \\ -1, & \text{if } Y_i^g < Y_j^g \\ 0, & \text{otherwise.} \end{cases} \tag{3.8}$$

3.4.5 Results and Discussion

The average performance (metrics described in Section 3.4.4) of Mobile Net and ResNet-50 for erythema, scaling and induration scoring using considered approaches are listed in Table 3.2. According to Table 3.2, the performance of the chosen networks trained with proposed loss function outperforms the same network trained with CCE or MSE loss minimization. However, the networks trained with CCE and MSE loss minimization produce comparable performance. We receive poor performance when the CNN is trained for regression (Regr) output. This justifies the fact that the sensitivity of this method towards the presence of noise in test images affects the performance badly. So, this approach is unsuitable for the present task. On the other hand, the binary decomposition approach outperforms the CNN models trained with CCE and MSE loss minimization. However, in most cases, this approach is beaten by the proposed method. Obviously, the success of the binary decomposition approach depends on the robustness of all decomposed classifiers and a weak classifier may affect the whole classification scheme adversely. According to Table 3.2, among all considered approaches, the best performance is achieved when ResNet-50 is fine-tuned with EMD loss minimization. Some images in our dataset along with their actual and predicted severity scores with respect to erythema, scaling and induration predicted by the best models are given in Table 3.3.

Table 3.3: Psoriasis Images and their ground-truthed (GT) and Predicted (Pred) severity scores achieved from the best classifier (ResNet-50 trained with the proposed loss function). The scores are given in (Erythema, Scaling, Induration). The errors are highlighted in yellow.

 GT:(1,2,1) Pred:(1,2,1)	 GT:(1,2,2) Pred:(1,2,2)	 GT:(1,1,1) Pred:(1,1,1)	 GT:(3,1,2) Pred:(3,1,2)
 GT:(2,1,2) Pred:(2,1,2)	 GT:(2,2,1) Pred:(2,2,1)	 GT:(4,1,1) Pred:(4,1,1)	 GT:(2,3,2) Pred:(2,3,2)
 GT:(1,2,1)	 GT:(1,2,1)	 GT:(2,3,1)	 GT:(1,2,2)

Continued on next page

Figure – continued from previous page

Pred:(1,2,1) Pred:(1,2,1) Pred:(2,3,1) Pred:(1,2,2)




GT:(3,2,2) GT:(1,2,1) GT:(0,0,0) GT:(1,1,1)

Pred:(3,2,2) Pred:(1,2,1) Pred:(0,0,0) Pred:(1,1,1)




GT:(2,3,2) GT:(1,1,2) GT:(3,4,2) GT:(2,3,1)

Pred:(1,3,2) Pred:(2,1,2) Pred:(4,4,2) Pred:(3,3,1)




GT:(1,1,1) GT:(3,2,2) GT:(3,1,2) GT:(2,4,2)

Pred:(1,2,1) Pred:(3,3,2) Pred:(3,0,2) Pred:(2,3,2)




GT:(3,3,1) GT:(2,3,3) GT:(2,3,3) GT:(1,1,1)

Pred:(3,3,2) Pred:(2,3,2) Pred:(2,3,2) Pred:(1,1,2)

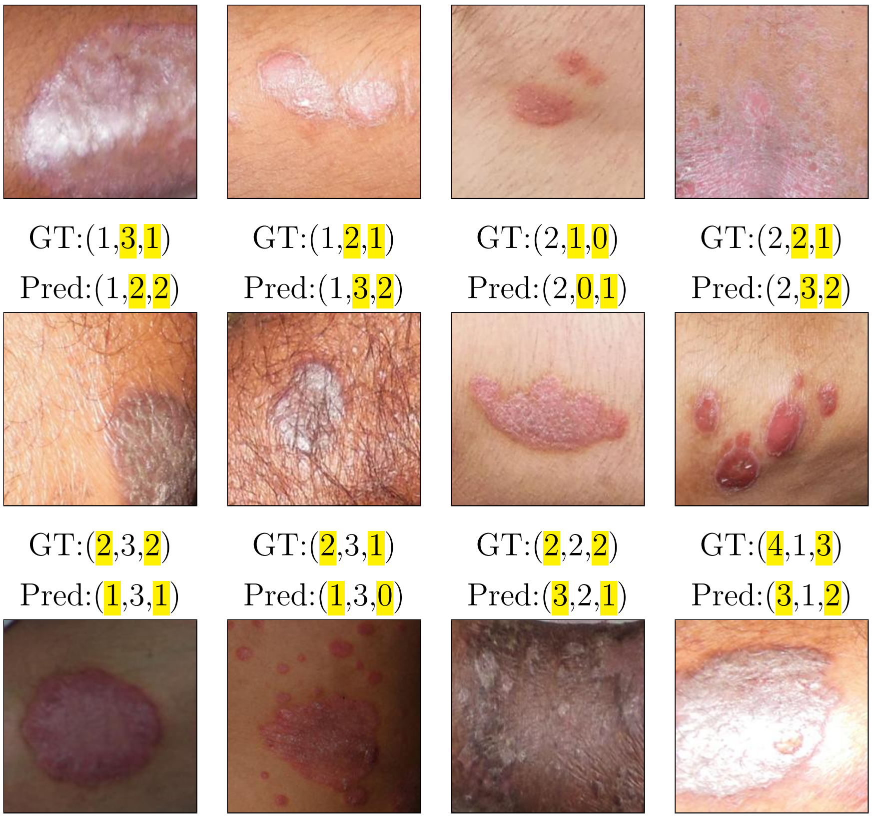



GT:(2,2,1) GT:(2,2,2) GT:(3,2,2) GT:(2,1,2)

Pred:(1,0,1) Pred:(3,3,2) Pred:(2,3,2) Pred:(3,2,2)

Continued on next page

Figure – continued from previous page



GT:(1,3,1) Pred:(1,2,2)
GT:(1,2,1) Pred:(1,3,2)
GT:(2,1,0) Pred:(2,0,1)
GT:(2,2,1) Pred:(2,3,2)

GT:(2,3,2) Pred:(1,3,1)
GT:(2,3,1) Pred:(1,3,0)
GT:(2,2,2) Pred:(3,2,1)
GT:(4,1,3) Pred:(3,1,2)

GT:(4,3,3) Pred:(3,0,2)
GT:(3,1,1) Pred:(4,2,2)
GT:(0,2,1) Pred:(4,4,2)
GT:(2,3,2) Pred:(1,2,1)

3.4.6 Comparison with MTL based approach

In the MTL based approach discussed in the previous section (Section 3.3), the best models for erythema and induration were obtained from the AlexNet based MTL network and for scaling it was from the AlexNet based STL network. The performance was evaluated with average correct classification accuracy, without and with ±1 tolerance[2], combined average classification accuracy[3] without and with ±1 tolerance[2]. In Table 3.3, the first row contains the best result obtained from DCNN based Multi-Task Learning (Section 3.3) and the second row contains the result produced by the best model obtained from DCNN based ordinal classifier (ResNet-50 trained with proposed loss). According to Table 3.3, ResNet-50 trained with proposed loss becomes the current state-of-the-art for the present dataset.

---

[2] maximum deviation of predicted scores for an image lies in [-1,1]

[3] erythema, scaling and induration all scores are correct

Table 3.2: Experimental Result (mean ± standard deviation): DCNN based Ordinal Classification

	Method	Erythema			Scaling			Induration		
		MA (%)	MAE	τ_b	MA (%)	MAE	τ_b	MA (%)	MAE	τ_b
MobileNet	CNN_{MSE}	57.14 ±6.99	0.521 ±0.080	0.763 ±0.047	57.28 ±2.88	0.523 ±0.048	0.775 ±0.022	59.97 ±5.42	0.443 ±0.063	0.757 ±0.034
	CNN_{CCE}	57.57 ±6.35	0.501 ±0.058	0.779 ±0.033	57.14 ±4.53	0.502 ±0.058	0.790 ±0.037	58.84 ±4.78	0.457 ±0.064	0.748 ±0.036
	CNN_{Regr}	47.52 ±4.75	0.605 ±0.072	0.737 ±0.031	45.54 ±3.13	0.634 ±0.039	0.734 ±0.027	55.45 ±3.43	0.504 ±0.038	0.721 ±0.027
	CNN_{Decomp}	56.01 ±5.27	0.508 ±0.060	0.777 ±0.026	58.84 ±3.83	**0.471** ±0.027	**0.804** ±0.022	**60.54** ±5.22	**0.431** ±0.071	**0.766** ±0.048
	Proposed	**59.69** ±4.49	**0.488** ±0.063	**0.781** ±0.034	**59.26** ±1.01	0.478 ±0.047	0.800 ±0.025	60.11 ±4.49	0.446 ±0.058	0.758 ±0.037
ResNet-50	CNN_{MSE}	58.27 ±4.52	0.478 ±0.036	0.791 ±0.021	58.56 ±2.65	0.487 ±0.038	0.790 ±0.016	58.84 ±4.42	0.444 ±0.044	0.764 ±0.036
	CNN_{CCE}	59.26 ±5.03	0.467 ±0.048	0.796 ±0.021	59.83 ±3.70	0.457 ±0.053	0.809 ±0.026	61.39 ±7.19	0.410 ±0.075	0.772 ±0.035
	CNN_{Regr}	45.97 ±4.85	0.632 ±0.067	0.728 ±0.056	43.71 ±2.77	0.655 ±0.053	0.729 ±0.032	48.23 ±5.50	0.581 ±0.078	0.686 ±0.057
	CNN_{Decomp}	58.56 ±4.96	0.474 ±0.056	0.792 ±0.025	59.83 ±7.11	0.465 ±0.080	0.803 ±0.041	62.52 ±2.52	0.409 ±0.043	0.775 ±0.026
	Proposed	**61.10** ±3.46	**0.440** ±0.050	**0.812** ±0.028	**62.66** ±0.01	**0.430** ±0.060	**0.820** ±0.034	**63.51** ±3.81	**0.390** ±0.038	**0.782** ±0.032

Table 3.3: Comparison with the DCNN based MTL approach. WoT refers **w**ith**o**ut **t**olerance and WT refers **w**ith **t**olerance.

Method	Individual Mean Accuracy						Overall Mean Accuracy	
	Erythema		Scaling		Induration			
	WoT	$WT(\pm1)$	WoT	$WT(\pm1)$	WoT	$WT(\pm1)$	WoT	$WT(\pm1)$
MTL	60.68	93.64	58.98	94.77	61.10	93.78	27.58	86.28
Ordinal Classifier	**61.10**	**95.47**	**62.66**	**95.19**	**63.51**	**97.45**	**28.43**	**89.11**

3.5 Summary

Deep neural network-based machine learning approaches are explored for automatic severity grading of psoriatic plaque in skin images. Development of DCNN based Multi-Task Learning approach and design of a novel loss function for making CNN suitable for ordinal classification are two key contributions of this chapter. The work presented in this chapter can be considered as a pioneering attempt where machine-assisted grading of the camera held psoriasis images is done on the basis of erythema, scaling, and induration. This study would eventually lead towards the development of some machine-assisted tools which are crucial for further advancement in the field of psoriasis image analysis. Image-based PASI score estimation system and patient's disease region monitoring system can be developed using the present work as one of the modules. Furthermore, no data set was previously available for severity analysis of psoriasis images. Therefore, the annotated data set (will be made available for research/non-profit purposes) prepared under the present work will undoubtedly facilitate the further research on machine-assisted analysis of psoriasis disease.

Chapter 4

Psoriasis skin biopsy image segmentation

Psoriasis can be diagnosed by the clinical features by looking at the outer skin surface. But instead of psoriasis, there are other erythemato-squamous diseases like Seborrheic dermatitis, Leprosy, Lichen planus, Tinea corporis, Pityriasis, and Eczema, etc. which have similar clinical features [42, 43, 61]. Therefore, the diagnosis of psoriasis as a different and distinct disease had been a controversial subject since ancient times. Hence, in doubtful cases histopathological examination is carried out for confirmation. For this purpose, doctors need to check for several characteristic features of the disease in histopathological images. This includes (i) epidermal changes like thickening, which is non-uniform extension of rete-pegs into dermis at some points, while at others there is thinning of the epidermis at the tip of dermal papillae, (ii) presence of immune cells within the epidermal layers, (iii) presence of nucleated cells within stratum corneum. While dermal changes include increased number and length of blood vessels and accumulation of huge number of immune cells around the blood vessels. As a computer vision task, before identification of these features, the most crucial issue is to segment the epidermal and dermal layers accurately.

Nowadays, Convolutional Neural Networks (CNNs) are widely used in several medical image segmentation tasks [68–70, 86, 87, 97, 101, 113, 115] etc. The most common way of CNN based medical image segmentation is the classification of every image pixel [68, 70, 86]. Generally, a local window around every pixel is considered for feature extraction and semantic labeling of the center pixel is done with a trained classifier based on the computed local feature. However, these methods are tardy as computational complexity depends on number of pixels as all pixels are labeled individually. Hence, this approach is not suitable for segmenting high dimensional images as real-time segmentation will not be possible. In order to cope up with this limitation two different approaches are devised. Firstly, instead of computing pixel-wise semantic maps superpixel wise semantic maps are computed [8, 31, 94, 110]. Superpixel divides an image into perceptual regions and for computing semantic map a local window around the centroid of the superpixels is considered for feature extraction. In contrast to pixel classification method, this approach reduces the computational time as all pixels in a superpixel are assigned same class label. Different strategies are devised for sampling training patches, selection of number of superpixels, etc. Another solution to seep up the computation is to devise a learning algorithm that can be trained end-to-end and will classify all image pixels together [3, 97, 101, 113, 115].

This chapter discusses the development of automatic approaches for robust segmentation of psoriasis skin biopsy images. Two different deep neural networks are trained for segmenting psoriasis skin biopsy images into dermis, epidermis and non-tissue regions. The deep model of the first kind is trained for classifying the superpixels generated by Simple Linear Iterative Clustering. The second deep model is a U-shaped Fully Convolutional Neural Network (FCN) which is trained in an end-to-end learning fashion where input is the original color image and the output is the segmentation class map for the skin layers. The performances of these approaches are compared using real psoriasis skin biopsy images.

The organization of this chapter is as follows. Section 4.1 discusses about the segmentation approaches. Section 4.2 presents experimental protocols. Experimental results and analysis are presented in Section 4.3. Finally, section 4.4 summarizes the chapter.

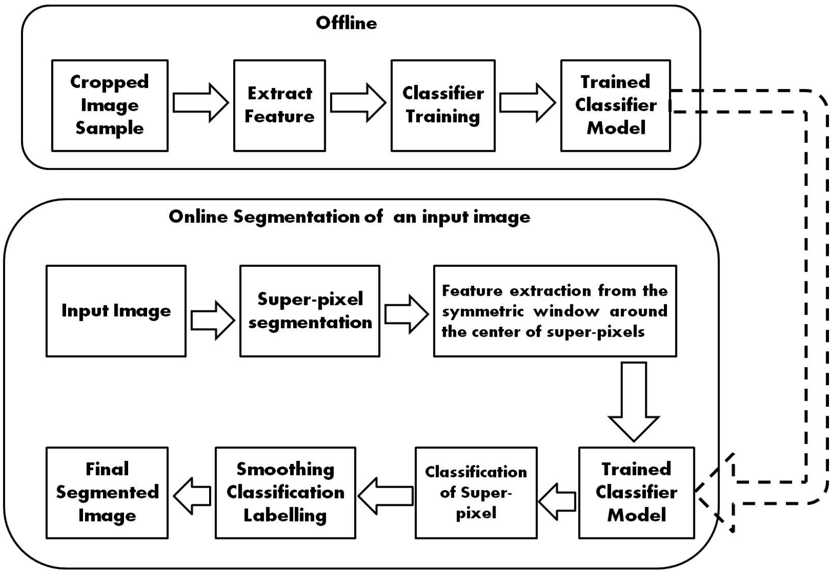


Figure 4.1: Block diagram of the superpixel classification based segmentation system

4.1 Methodology

4.1.1 DCNN based superpixel classification for segmentation

Fig 4.1 presents the block diagram of the proposed superpixel classification based system. The development of the classification model for classifying dermis, epidermis and non-tissue regions is termed as off-line process and superpixel classification is termed as on-line process. During off-line process, firstly, random square patches of size 63 × 63 are cropped from the original color (RGB) images. These image patches are cropped in such a way that every pixel in a patch belongs to same class (either from dermis, epidermis or neither from any skin layer). Then these patches are labeled accordingly and used for training the DCNN. During the on-line process, three tasks are done- (i) superpixel generation and (ii) classification of the superpixels into respective classes (called semantic labeling) (iii) smoothing the superpixel class labels. A more detailed description of the system is given in the following subsections.

4.1.1.1 superpixel generation

The success of the segmentation approach depends on the robustness of superpixel's boundary adherence with respect to objects in an image. So, the number of superpixels is assigned in such a way that all pixels in a superpixel belong to the same class label. In general, more superpixels guarantee more accurate boundary adherence but with the increase of number of superpixels the segmentation process becomes tardy. Therefore, there is a trade-off among the number of superpixels, speed, and segmentation accuracy.

4.1.1.2 Classifier model development

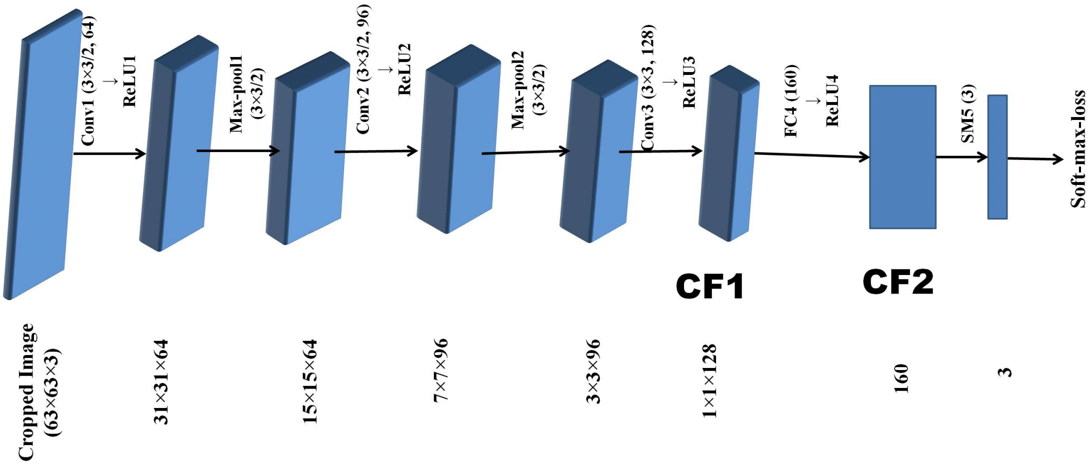


Figure 4.2: Patch classification CNN Architecture. Convolution layers' parameters written as (height×width /stride, number of filters). Max-pool layers' parameters written as (height×width /stride). The numbers inside the first bracket after a fully connected layer or a softmax layer represents the number of output nodes. After training of the CNN, the feature maps from CF1 and CF2 are used to build classifier models with the help of traditional classification algorithms.

The detail description of the proposed CNN architecture is shown in Fig 4.2. In this model, the Rectified Linear Unit (ReLU) is used as activation function due to its faster execution compared to other activation functions. The classification layer has three nodes so that it can deal with a three-class classification problem. The soft-max (SM) loss is used due to its popularity for dealing with classification loss.

The key idea behind superpixel based segmentation is that the output from the last layer of the trained Convolutional Neural Network (CNN) provides the softmax class probability of the considered regions (or superpixels). The feature maps obtained from the intermediate layers (CF1 and CF2 as shown in Figure 4.2) of the trained CNN (Chopped Feature) are also considered and used for constructing classifier models. K-Nearest Neighbour (KNN), Support Vector Machine (SVM) and Random Forest (RF) are considered for developing classifier models.

4.1.1.3 Smoothing

The segmentation result after superpixel classification often requires refinement as underlying spatial constraints among neighboring superpixels may not be dealt properly by the classification method. In this research, Markov Random Field (MRF) energy minimization is applied for smoothing the segmentation outcome after superpixel classification. This is a commonly used segmentation refinement strategy.

The goal of this method is to label the image superpixels (P) in such a way that the labeling ($\mathcal{L}$), minimizes the energy E($\mathcal{L}$) where

$$\begin{aligned} E(\mathcal{L}) &= E_{data}(\mathcal{L}) + E_{smooth}(\mathcal{L}) \\ &= \sum_{p \in P} D_p(f_p) + \sum_{p,q \in \mathcal{N}} V_{p,q}(f_p, f_q) \end{aligned} \tag{4.1}$$

where f_p and f_q are the feature representations for p^{th} and q^{th} superpixels respectively. $D_p(f_p)$ depends on class assignment of p^{th} superpixel. $V_{p,q}(f_p, f_q)$ depends on the relative class assignment of p^{th} and q^{th} superpixels. $V_{p,q}(f_p, f_q)$ is responsible to impose spatial smoothness among the neighbouring superpixels.

4.1.2 FCN based semantic segmentation

The success of superpixel classification based approaches is highly dependent on the boundary adherence ability of the chosen superpixel generation algorithm. Moreover, as the superpixels are classified independently, there is a chance of losing the spatial smoothness. Considering these shortcomings, a viable alternative is to segment an image considering all of its pixels together. This section explores this alternative approach where a whole image is fed as input to a Deep Convolutional Neural Network (DCNN) that gives a segmentation class map as output.

In literature, there are several approaches for developing segmentation DCNN model. Long et al. [72] convert classification net into a fully convolutional net that produces coarse output maps at three different resolutions which are then connected together to generate pixel-wise object labeling. But in this method, it is difficult to reconstruct complex structures of object boundaries accurately. So, designing of a CNN architecture with an original color image as input and corresponding semantic segmentation labeling as output is considered as a smarter choice [97]. In these CNN architectures, there will be a contracting path to capture the context and symmetric expanding path for precise localization of the objects. In this approach, high-resolution features from the contracting path are combined with the up-sampled output for localization [3, 55, 97]. Hence, the net is able to learn complex and finer object structures. In these Fully Convolutional Neural Network also the average per-pixel classification loss is minimized using the gradient-based optimization algorithm.

The detailed description of the proposed U-shaped Fully Convolutional Neural Network (FCN) architecture is shown in Fig 4.3. In this model, the Rectified Linear Unit (ReLU) is used as activation function due to its faster execution compared to other activation functions. The average cross-entropy error from the image pixels is considered as the classification loss. The network firstly contracts the size of the image 16 times using four 2 × 2 max-pooling operations and then expands to get the original dimension using four 2 × 2 up-sampling layers. The copy path is used to concatenate the feature response obtained from the previous layer (as shown in Fig 4.3) to the present layer. A learned FCN network is used to generate a segmentation class-map for an unseen image.

4.2 Experimental Protocol

4.2.1 Data acquisition and preparation

For histopathological analysis, after clinical confirmation of psoriasis affected tissues are collected in 10% formalin under the supervision of an expert dermatologist. Formalin-fixed tissues are dehydrated and embedded in paraffin blocks. Thin

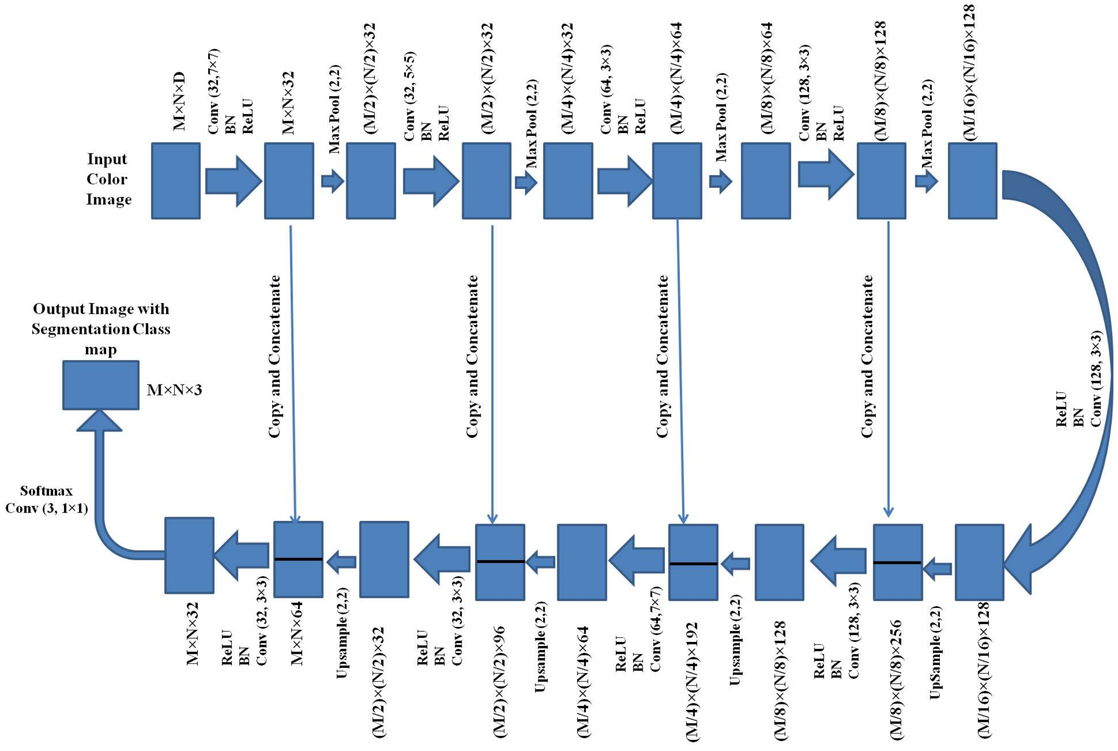


Figure 4.3: The proposed U-shaped FCN architecture. $M \times N$ represents image dimension and D represents the number of color channels of the input images. Convolution layers' parameters written as (number of filters, height×width). Max-pool and Up-sampling layers' parameters are written as (height×width). BN denotes batch normalization layer.

sections (5 μM) are used for slide preparation and then stained with hematoxylin and eosin to prepare the histopathological slides. Afterward, prepared histopathological slides are kept under a microscope (ZEISS Axiostar Light Microscope) with 10X magnification and the images are collected from the microscope using a digital camera (TUCSEN Camera USB 2.0 series). The size of the captured images is 1936×2584 pixels. Then the segmentation annotation (semantic labeling of all the pixels as dermis, epidermis, and non-tissue) is prepared by an expert pathologist. Written informed consent was obtained from all patients before recruiting them for the study. This study is conducted after obtaining ethical approval from the 'Review Committee for Protection of Research Risks to Humans' of Indian Statistical Institute, Kolkata.

The present data set contains annotations for ninety (90) images. This data set is divided into two disjoint sets- Training set and Test set each containing 45 images.

For the superpixel segmentation based approach, the classification model for semantic labeling of the superpixels is built using the images from the Training set. A database of 63×63 sized patches is developed for training the classifier. These patches are cropped in such a way that all the pixels containing in a patch belong to the same class (either dermis, epidermis or neither from any skin layer) in the original image. It is worth mentioning that the training patches are cropped in such a way that it's centroid coincide with a superpixel's centroid. Our database consists of $9,060$ non-skin patches, $11,900$ dermis layer patches and $11,600$ epidermis layer patches. Note that the same training images are used for training the Fully Convolutional Neural Network (FCN).

4.2.2 Hand-crafted Feature

In order to make a comparative assessment of our Deep Convolutional Neural Network-based classifiers with other popularly used classification algorithms (using KNN, SVM, and RF), the following commonly used hand-crafted color and texture (C&T) features are considered.

Color Feature

Though several color spaces are there in the literature, RGB color space is considered as all the other color spaces are obtained by some linear or non-linear manipulation on RGB vectors. In addition, CIE Lab and HSV color spaces are also considered as they provide distinguishing visual clues among the different tissue patterns available in the biopsy sample. For each color component for all of the three color spaces three statistics namely, mean, standard deviation and entropy are computed to build the color feature.

Texture feature:

Three widely used texture descriptors namely, Gray Label Co-occurrence Matrix (GLCM) [47], Local Binary Pattern (LBP) [89] and Histogram of Oriented Gradient (HOG) [23] are considered. Note that rotation invariant texture descriptors are extracted for the present problem as the tissue samples may be aligned in any direction.

4.2.3 Implementation

The CNN based classifier used in this work is built with the **MatConvNet** [112], an open-source implementation of Convolutional Neural Network. The weight vectors of the CNN model are initialized randomly. The model is trained using a batch size of 256 images, momentum of 0.9, weight decay of 0.005, and the learning rate of 0.001. The images in every batch are randomly shuffled before training. According to our experiments, 100 epochs are enough for saturation of the network parameters. **Matlab R2016a** in-built implementations of SLIC, RF, KNN are used and for SVM-based classification, LIBSVM [18] package is used with Radial Basis Function (RBF) kernel. The MRF energy minimization is performed by the implementation of the graph-cut algorithm by Bagon [4]. In order to develop hand-crafted feature-based classifiers, firstly, the color and texture feature described in Sect 4.2.2 are extracted and then concatenated. The best parameters for the classifier algorithms (by KNN, SVM, and RF) are chosen through multiple trials. In order to make a comparative study among the competing classifiers same training patches are used for training purposes. For hand-crafted features, since the range of all feature dimensions are not equal so, feature-wise min-max scaling is applied.

After superpixel classification, similar smoothing strategy is opted for all approaches. In our experiment, if a superpixel (p) is classified into i^{th} class label, then the energy of assigning p into i^{th} class set to zero i.e. $D_p(f_p) = 0$ otherwise, $D_p(f_p) = \epsilon$ (ϵ is an empirically decided positive value) and the Euclidean distance is used to measure the smoothness energy ($V_{p,q}(f_p, f_q)$).

For the FCN based approach, Keras [19] is used. The model is trained using a batch size of 1 image, momentum of 0.9 and with the learning rate of 0.001. The images in every batch are randomly shuffled before training. According to our experiments, 100 epochs are enough for saturation of the network parameters.

An Intel® Core™ i7-4770 processor-based workstation with GTX 1080 GPU is used for performing all the experiments. For the implementation of the U-shaped FCN architecture, because of the resource constraint, we down-sampled the images (which are originally 1936×2584) to 960×1280 and the corresponding ground-truthed images are changed accordingly.

4.2.4 Evaluation Metrics

For measuring the overall image segmentation performance, the Ratio of Correct Pixel Classification (RCPC) accuracy measure is used. RCPC represents the ratio of the correctly classified pixels to the number of pixels in an image. Mathematically RCPC is defined as follows:

$$RCPC = \frac{Number\ of\ Pixel\ Correctly\ classified}{Number\ of\ pixel\ in\ the\ image} \tag{4.2}$$

Jaccard's Coefficient (JC), a popularly used metric is also used for evaluating the segmentation performance. JC provides an object wise (in this work non-tissue, dermis and epidermis region) performance measure of a segmentation algorithm. Mathematically, JC is defined as follows:

$$E(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{4.3}$$

where A is the ground-truth set of pixels of an object, B is the set of machine detected pixels of the same object and $|*|$ represents the set cardinality.

4.3 Results and Discussion

In our experiment, two different Deep Convolutional Neural Networks are used for segmenting the psoriasis skin biopsy images in two different ways. The analysis and comparison of the experimental outcomes are given in the following subsections.

4.3.1 DCNN based superpixel classification

4.3.1.1 Model development for classification

The CNN architecture shown in Fig. 4.2 is trained using the database of cropped patches. Moreover, the trained CNN model is used as a feature extraction engine and extract feature from intermediate (CF1, CF2) layers and build classifiers with KNN, SVM, and RF. Finally, these CNN provided features are fed to the traditional classifiers and classification accuracies are compared with traditional hand-crafted feature-based classifiers. 10-fold cross-validation is performed and the best classification results (among 10 folds) are listed in Table 4.1. According to Table 4.1, the CNN given feature-based classifiers are more efficient than the hand-crafted color and texture (C&T) feature-based classifiers.

Table 4.1: Performance of cropped image patch classification

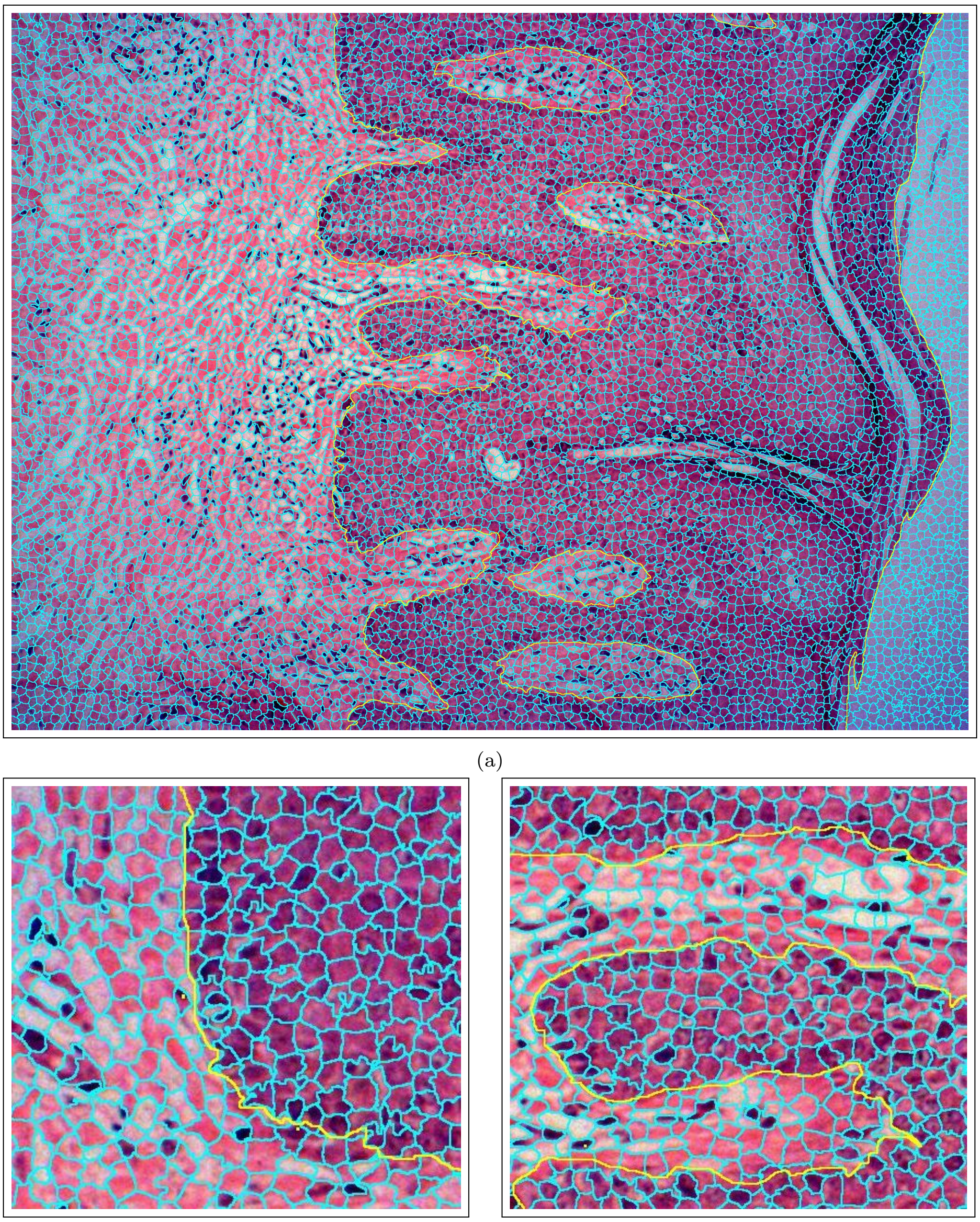
Feature	Classifier Accuracy (%)		
	KNN	SVM	RF
C&T	88.58	92.17	92.14
CF1 Feature	92.15	94.66	92.85
CF2 Feature	92.59	94.87	93.15
CNN	95.17		

4.3.1.2 superpixel generation

Due to the limited WSIs, we are not interested to select the number of superpixels (N) in any supervised settings. Rather, our goal is to set the value of N in such a way that bigger sized superpixel is obtained but all the pixels of a superpixel will be in the same class (same skin layer). In other words, the tissue region boundaries maximally overlap with superpixel boundaries. This will ensure that the correct classification of the selected patch around a superpixel will provide good segmentation outcomes. Hence, we compute Global Consistency Error (GCE) between the ground-truth segmentation of skin layers and superpixel segmentation. GCE gives the extent to which superpixelation can be viewed as a refinement (or over-segmentation) of the ground-truth skin layers. A lower value for GCE implies a better performance and if GCE=0 when all superpixels contain pixels from the same class.

4.3.1.3 superpixel classification

After generation of superpixels, the next task is to classify the superpixels using the trained classifiers. For visual representation, the segmentation outcomes of the competing classification algorithms for a test image are shown in Fig. 4.5. Fig. 4.5 shows that all superpixel segmentation based approaches are able to detect skin layers. Comparison among the chosen algorithms in qualitative manner (as shown in Fig. 4.5) is quite difficult so a quantitative evaluation is devised next. The evaluation metrics described in Section 4.2.4 are computed for every algorithm and for all images in our data set. In order to make a comparative analysis,

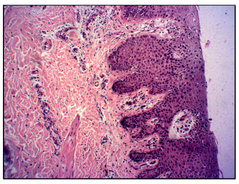


(a)

(b) (c)

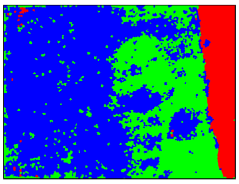
Figure 4.4: superpixel generation: Ground-truthed region (skin layer) boundaries are represented by yellow lines and SLIC region boundaries are represented by cyan lines. (a) An image, corresponding ground-truth and region boundaries produced by SLIC algorithm ; (b) and (c) two magnified portions from Figure 4.4(a).

segmentation performance is reported in terms of mean and standard deviation for both the training (Table 4.2) and the test images (Table 4.3).

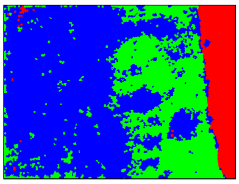


Sample Images

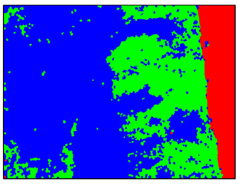
Ground-Truth



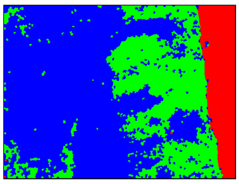
KNN (C&T)



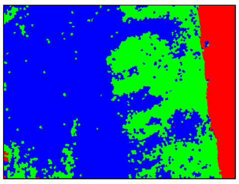
Smoothed KNN (C&T)



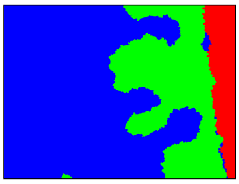
KNN (CF1)



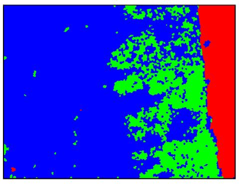
Smoothed KNN (CF1)



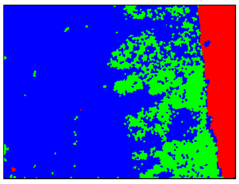
KNN (CF2)



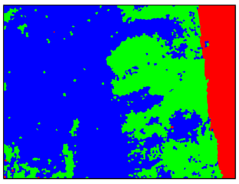
Smoothed KNN (CF2)



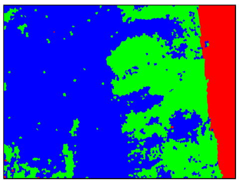
SVM (C&T)



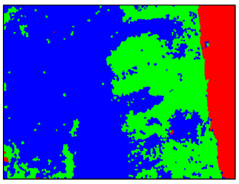
Smoothed SVM (C&T)



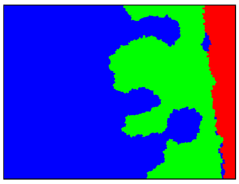
SVM (CF1)



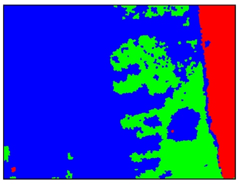
Smoothed SVM (CF1)



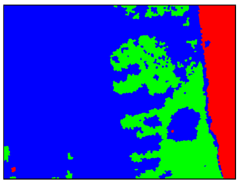
SVM (CF2)



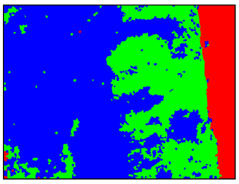
Smoothed SVM (CF2)



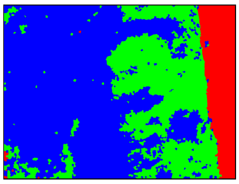
RF (C&T)



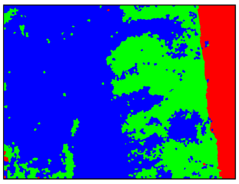
Smoothed RF (C&T)



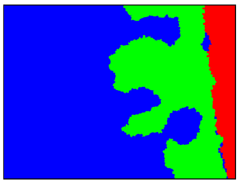
RF (CF1)



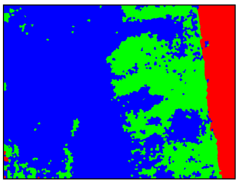
Smoothed RF (CF1)



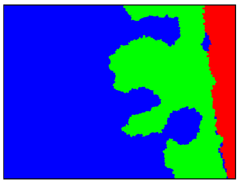
RF (CF2)



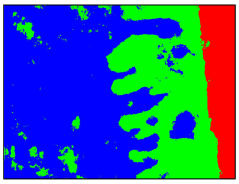
Smoothed RF (CF2)



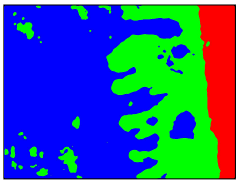
CNN



Smoothed CNN



FCN



Smoothed FCN

Figure 4.5: An image from the test data set, corresponding segmentation ground-truth and the segmentation result with respect to the competing approaches are shown. In the ground-truth image as well as in the segmented images the red portion represents the non-tissue region, the blue portion represents the dermis region and the green portion represents the epidermis region. The text under the segmented images represents the respective segmentation approach. C&T represents hand-crafted color and texture features, CF1 and CF2 correspond to chopped CNN features and FCN represents Fully Convolutional Neural Network.

Note that the proposed superpixel classification approaches produce promising segmentation (psoriasis biopsy) results both for the training (Table 4.2) and test images (Table 4.3). This fact justifies the generalization ability of the developed classifiers in segmenting psoriasis biopsy images.

For superpixel based classification methods, the best results achieved with respect to RCPC and JC (for three different regions) are star-marked (*) in Table 4.2 and Table 4.3. From these tables, it can be concluded that the DCNN model-based classifiers (either using it as feature extractor or using it as feature extractor cum classifier) are more efficient than the traditional hand-engineered color and texture (C&T) feature-based classifiers built on KNN, SVM, and RF. Hence, our experiment attests the importance of using DCNN based superpixel classification for the present segmentation task.

4.3.2 Fully Convolutional Neural Network-based approach

In order to explore the efficiency of the Fully Convolutional Neural Network (FCN), a U-shaped Deep Convolutional Neural Network is trained (as described in Fig. 4.3). The segmentation outcome of a test image (first column of the first row of Fig. 4.5) is shown in the third column of the sixth row of Fig. 4.5. The training accuracy of FCN is shown in the last row of Table 4.2 and its accuracy on the test data set is given in the last row of Table 4.3. Table 4.2 and Table 4.3 show that for both the training and test data sets noticeable performance gains in terms of average accuracies are achieved by using the FCN based approach. This justifies the importance of developing FCN model for the present segmentation task. However, for some cases, FCN is unable to capture the structure (skin layers) and produces very inferior segmentation that in turn results in slightly higher standard deviation.

Table 4.2: Segmentation Performance (mean ± standard deviation) on the training data set

Method	RCPC	Non-Tissue (JC)	Dermis (JC)	Epidermis (JC)
C&T+KNN	0.8281 ± 0.0573	0.8271 ± 0.1495	0.6762 ± 0.1274	0.6878 ± 0.0955
C&T+SVM	0.8479 ± 0.0697	0.8541 ± 0.1382	0.7477 ± 0.1273	0.6806 ± 0.1189
C&T+RF	0.8403 ± 0.0765	0.8370 ± 0.1461	0.6753 ± 0.1264	0.7133 ± 0.1434
CF1+KNN	0.8590 ± 0.0640	0.8696* ± 0.1334	0.7407 ± 0.1282	0.7222 ± 0.0983
CF1+SVM	0.8625 ± 0.0623	0.8624 ± 0.1406	0.7420 ± 0.1274	0.7343 ± 0.0959
CF1+RF	0.8655* ± 0.0622	0.8631 ± 0.1412	0.7500 ± 0.1286	0.7380* ± 0.0961
CF2+KNN	0.8583 ± 0.0622	0.8614 ± 0.1410	0.7387 ± 0.1278	0.7233 ± 0.0960
CF2+SVM	0.8638 ± 0.0637	0.8613 ± 0.1432	0.7463 ± 0.1286	0.7356 ± 0.0991
CF2+RF	0.8653 ± 0.0629	0.8602 ± 0.1412	0.7513 ± 0.1274	0.7363 ± 0.0975
CNN	0.8606 ± 0.0701	0.8618 ± 0.1401	0.7532* ± 0.1281	0.7204 ± 0.1094
FCN	**0.9126** ± 0.0824	**0.8843** ± 0.1725	**0.8397** ± 0.1479	**0.8295** ± 0.1230

Table 4.3: Segmentation Performance (mean ± standard deviation) on the test data set

Method	RCPC		Non-Tissue (JC)		Dermis (JC)		Epidermis (JC)	
	Raw Output	Smoothed	Raw Output	Smoothed	Raw Output	Smoothed	Raw Output	Smoothed
C&T+KNN	0.7757 ± 0.0861	0.797 ± 0.0882	0.8019 ± 0.1698	0.8362* ± 0.1713	0.5928 ± 0.1536	0.6819 ± 0.1597	0.6165 ± 0.1295	0.5629 ± 0.1319
C&T+SVM	0.7963 ± 0.0927	0.799 ± 0.0921	0.8338 ± 0.1348	0.8269 ± 0.1335	0.6812 ± 0.1276	0.6773 ± 0.1270	0.5621 ± 0.2107	0.5946 ± 0.2102
C&T+RF	0.7981 ± 0.1164	0.799 ± 0.1168	0.8244 ± 0.1643	0.8269 ± 0.1645	0.6762 ± 0.1638	0.6773 ± 0.1648	0.5932 ± 0.2281	0.5946 ± 0.2291
CF1+KNN	0.8386 ± 0.0638	0.8596 ± 0.0542	0.8648* ± 0.1090	0.7988 ± 0.1949	0.7077 ± 0.1205	0.7590 ± 0.1142	0.6853 ± 0.1060	0.7076 ± 0.1088
CF1+SVM	0.8437 ± 0.0651	0.8549 ± 0.0574	0.8534 ± 0.1230	0.8053 ± 0.1403	0.7117 ± 0.1225	0.7389 ± 0.1199	0.7047 ± 0.1055	0.7183 ± 0.1025
CF1+RF	0.8461* ± 0.0677	0.8656* ± 0.0676	0.8507 ± 0.1193	0.7957 ± 0.1854	0.7193 ± 0.1262	0.7640 ± 0.1381	0.7059 ± 0.1088	0.7285* ± 0.1091
CF2+KNN	0.8395 ± 0.0674	0.8594 ± 0.0599	0.8449 ± 0.1226	0.7957 ± 0.1441	0.7094 ± 0.1252	0.7528 ± 0.1234	0.6936 ± 0.1095	0.722 ± 0.1028
CF2+SVM	0.8461* ± 0.0660	0.8617 ± 0.0568	0.8484 ± 0.1189	0.8069 ± 0.1852	0.7183 ± 0.1238	0.7541 ± 0.1243	0.7078* ± 0.1096	0.7285* ± 0.0989
CF2+RF	0.8458 ± 0.0661	0.8618 ± 0.0576	0.8389 ± 0.1275	0.8080 ± 0.1374	0.7220 ± 0.1227	0.7572 ± 0.1175	0.7050 ± 0.1092	0.7254 ± 0.1023
CNN	0.8412 ± 0.0623	0.8622 ± 0.0605	0.8473 ± 0.1249	0.8058 ± 0.1394	0.7229* ± 0.1122	0.7706* ± 0.113	0.6814 ± 0.1136	0.7082 ± 0.1177
FCN	**0.8801** ± 0.0780	**0.8845** ± 0.0781	**0.8900** ± 0.1092	**0.8962** ± 0.1083	**0.7717** ± 0.1476	**0.7792** ± 0.1482	**0.7710** ± 0.1267	**0.7785** ± 0.1274

4.3.3 Segmentation Refinement

After superpixel classification, segmented outcome contains spurious holes, i.e, inside an epidermis region, some smaller regions are segmented as dermis or background regions, as shown in Figure 4.5. In this scenario, segmentation refinement plays a crucial role in reducing these spurious holes by enforcing the spatial smoothness of the superpixels' class labels. Figure 4.5 shows that the chosen Markov Random Field (MRF) energy minimization based refinement approach is successfully able to reduce these spurious holes. This is also reflected in Table 4.3 and a notable performance gain is achieved across all the superpixel based segmentation algorithms (note the RCPC values in Table 4.3). For JC values, smoothing works well for dermis and epidermis segmentation though it sometimes degrades non-tissue segmentation accuracy.

In case of FCN based approach, pixel relabelling by median filtering (due to resource constraints the pixel relabelling with MRF is not explored) also improves the segmentation performance. Hence, the segmentation refinement module should be considered as an important part of the segmentation system.

4.4 Summary

The proposed deep convolutional neural network-based approaches successfully address the problem of epidermal and dermal segmentation, required as the first stage of image analysis of psoriasis histopathology. This research can be considered as a pioneering attempt for developing a successful automatic machine-assisted tool for psoriasis biopsy image analysis by marching several image processing and computer vision techniques. From the machine learning perspective, the experimental results show that the deep neural architectures could provide viable segmentation algorithms when compared with the traditional methods. On the other hand, from the medical image analysis viewpoint, the success of the present research provides new avenues for further research for developing automated identification of other characteristic disease features that can distinguish psoriasis from different skin conditions. Moreover, no annotated image data set of psoriasis skin biopsy was available before and therefore, the data set (which will be made available for research/non-profit purposes) prepared for the present work would undoubtedly facilitate the future research endeavors.

The present research points some more technical issues that require future attention. Experience from the superpixel based segmentation approaches reveals that the errors in segmentation are mostly arising due to two reasons: (i) occasional non-overlapping of superpixel boundaries and tissue boundaries and (ii) classification errors. In order to deal with the first problem, we need to have a more efficient approach for superpixel generation, which generates superpixels in such a way that the region boundaries are overlapped with superpixels' boundaries. However, the use of a pixel classification based approach cannot be an alternative because such a method could be very time-consuming as compared to the superpixel based method (in our case 120h versus 15min). For dealing with the second problem, i.e. to reduce classification errors, further effort is required to make the classifiers more robust. On the other hand, the present FCN has been trained with limited number of training images and therefore, use of data augmentation techniques may improve the performance further.

Chapter 5

Munro's Microabscess Detection

Munro's Microabscess is a histopathological characteristic of psoriatic skin characterized by the presence of neutrophils along with nucleated keratinocytes in the Stratum Corneum (SC) of the epidermis. In clinical pathology, Munro's Microabscess is considered as the hallmark sign of psoriasis. The pathologists perform microscopic inspection of skin biopsy for detecting it. The manual inspection of skin biopsy is a tedious job and depends on expert's experience. Moreover, staining variation in histopathology slides often creates confusion to differentiate neutrophils from skin keratinocytes.

This chapter develops an automated system for solving the unmet clinical needs of neutrophils' detection in the SC layer in the presence of skin keratinocytes. Specifically, this chapter focuses on designing a capsule network (CapsNet) based framework for the detection of multiple copies of the neutrophils available in the SC layer in the high-dimensional histopathology images using a weakly supervised approach. Segmentation of SC layer and analysis of sampled SC regions (SC patches) with classification network are two key sequential steps of the proposed system. CapsNet based segmentation and classification networks are used in this framework. We choose CapsNet owing to their robustness towards better hierarchical object representation.

The organization of this chapter is as follows. In Section 5.1, a brief introduction about the clinical and technical background is given. Section 5.2 discusses the proposed framework. The experimental protocol and analysis of experimental results are given in Section 5.3 and Section 5.4 respectively. Finally, Section 5.5 concludes the chapter.

5.1 Background

In this section, firstly, a brief introduction about Munro's Microabscess is provided and then the necessity for automated system for detecting it is discussed. This section also presents an introduction about the deep neural network models needed to understand the work presented in this chapter.

5.1.1 Munro's Microabscess

Psoriasis develops when the immune system mistakes a normal skin cell for a pathogen and sends out faulty signals to yield the overproduction of new skin cells. This disease causes several notable disorders in skin anatomy which can be inspected by skin biopsy analysis (see Section 1.3). The retention of nucleated keratinocytes in the Stratum Corneum and infiltration of large number of immune cells into the dermis and often into epidermal compartments are important histopathological characteristics for psoriasis. The infiltration of nucleated keratinocytes and immune cells in Stratum Corneum occurs either in confluent (throughout the Stratum Corneum layer) or in focal (not confluent or localize) manner. The infiltration of nucleated keratinocytes in Stratum Corneum is termed as parakeratosis and the accumulation of neutrophils in Stratum Corneum along with parakeratosis is termed as Munro's Microabscess (MM). In clinical pathology, Munro's Microabscess is considered as the diagnostic hallmark of psoriasis [15, 16, 24, 79].

Munro's Microabscess is detected by inspecting the Haematoxylin & Eosin (H&E) stained skin biopsy slides keeping under a microscope. The manual inspection procedure is error-prone due to the observer's experience and mental fatigue, staining variation in the prepared biopsy slides, etc. So, the Munro's Microabscess detection in psoriasis suffers from inter-observer variability among pathologists. The key source of such disagreement among pathologists lies in the correct recognition of neutrophils in the Stratum Corneum layer. In Stratum Corneum layer, the nuclei of the keratinocytes appears as a light stained and oval-shaped object and the neutrophils appear as circular shaped and dark stained objects (see Fig 5.1). But due to over-staining sometimes nuclei in a keratinocyte is misclassified as neutrophil and vice-versa. In that scenario, an automated system can

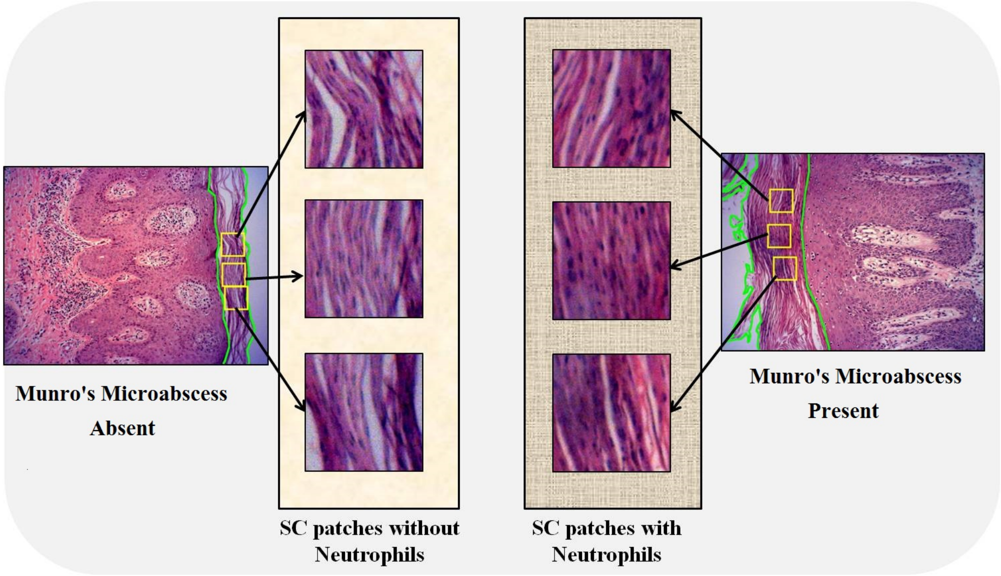


Figure 5.1: Two Whole Slide biopsy Images (WSIs) from our data set. The Stratum Corneum (SC) layer is outlined with a green border. The right side image contains the sign of Munro's Microabscess. The magnified SC patches from both images are shown between these two WSIs.

be designed to complement the expert's limitations. It is observed that parakeratosis (i.e. presence of keratinocytes in the Stratum Corneum layer) occurs in many erythemato-squamous diseases (i.e. Pityriasis lichenoides chronia, Pityriasis rubra pilaris, Lichenification, etc.) but the presence of neutrophils in the Stratum Corneum layer in addition to the parakeratosis occurs only for psoriasis [79]. Hence, this chapter attempts to develop computational method for automatic detection of neutrophils in Stratum Corneum layer in the presence of keratinocytes.

5.1.2 Capsule Network

Recently, Convolutional Neural Network (CNN) [49, 63, 107] become the state-of-the-art for several medical image classification [22, 22, 36, 38, 38, 39, 39, 54, 60, 60, 84, 84, 117, 117], and segmentation [68–70, 86, 87, 97, 101, 113, 115] tasks. However, the max-pooling layer in a CNN model loses important spatial hierarchies between object parts in the lower layer and object in the higher layer as sub-sampling by "max-pooling" can only produce translational invariance feature but does not provide translational equivariance feature. In order to mitigate the problems associated with CNN, Capsule network or CapsNet is proposed [51, 102]. The capsule networks replace the "max-pooling" layers with convolutional strides and dynamic routing. The information at the neuron level is stored as vectors

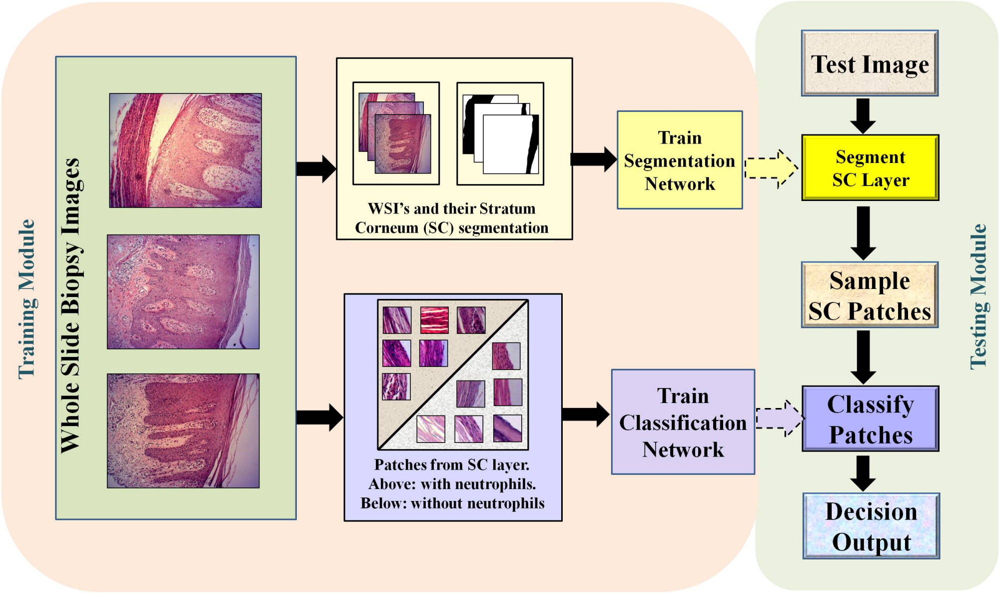


Figure 5.2: Proposed system architecture.

(called capsule), rather than scalars. Hence, in a capsule network, a capsule can preserve both the spatial orientation and magnitude of an object or object part. The probability of presence of an object or object part in the current capsule is represented by the magnitude of the vector. The sets of neurons in a capsule network are “routed” to the capsules in higher layer via a dynamic routing algorithm. The presence of an object in higher layer is decided by the agreement amongst the capsule vectors in the lower layer. Thus in capsule networks, a meaningful part-to-whole relationship is preserved. Hence, recently researchers started to use capsule networks in medical image classification [85] and segmentation [64] purposes.

5.2 Proposed Methodology

The target of the proposed system is to detect the presence of neutrophils in the Stratum Corneum layer of the human skin. But in a biopsy image, the Stratum Corneum layer lies in a small portion of the Whole Slide Biopsy Image (approx. 2-15%). Therefore, identification of the Stratum Corneum layer followed by neutrophil detection is the logical approach. Hence, the proposed system firstly segments the Stratum Corneum layer and then looks for the presence of neutrophil in a selected region (referred as patch) inside Stratum Corneum.

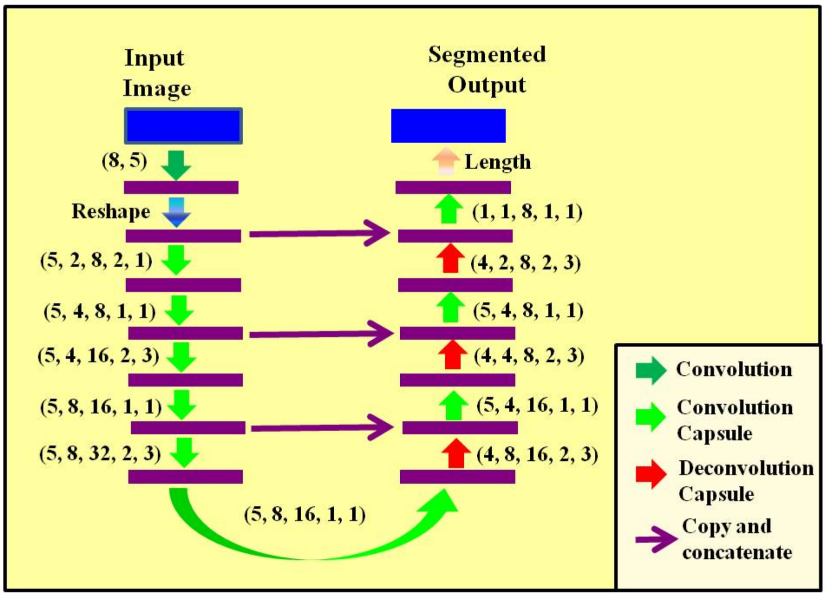


Figure 5.3: UCaps architecture. Parameters of the convolutional layer is kept in (Number of filters, Filter size) order, convolutional capsule layer is kept in (Filter size, number of capsule, number of atom, stride, num of routing) order and deconvolutional capsule layer is kept in (Filter size, number of capsule, number of atom, scaling, num of routing) order.

The proposed framework is shown in Figure 5.2. In the training module, a segmentation network is trained for Stratum Corneum segmentation as well as a classification network is trained for classifying Stratum Corneum patches. The test phase consists of three steps. Firstly, the Stratum Corneum layer is segmented using the trained segmentation network, then patches (sub-images) from the segmented regions are sampled and finally, the sampled patches are classified on the basis of the presence of neutrophils using the trained classification network. The conclusion about the Munro's Microabscess is made on the basis of aggregated classification results of the sampled patches.

In the proposed framework, both segmentation and classification network are built with capsule networks. The reason for choosing the capsule network is its robustness towards the preservation of the object's part-whole relationships. The detail description of the computational units of the proposed framework is given in the following subsections.

5.2.1 Segmentation of Stratum Corneum layer

The pictorial representation of the proposed U-shaped capsule network trained for Stratum Corneum segmentation is shown in Figure 5.3. According to Figure 5.3, the segmentation network has a contracting path and a symmetric expanding path. The contracting path captures the context and the symmetric expanding path precisely localizes the desired skin layer. Unlike traditional UNet [97], where max-pooling is used for contraction and up-sampling is used for expansion, we use convolutional capsules in contracting path and deconvolutional capsules in the expanding path. In a convolutional capsule, only the capsules present within a defined spatial region are routed together and the transformation matrices within a capsule type are shared. This is done since the “dense” routing used in [102] is infeasible due to large spatial dimension of the whole slide biopsy image. The deconvolutional capsules are built with transposed convolutions with locally-constrained routing. The high-resolution features from the contracting path are combined with the up-sampled output (in the expansion path) with a skip connection to make the network suitable to learn the desired complex and finer skin layer. This network is further referred to as UCaps.

During training, the proposed network minimizes the Dice loss. The mathematical expression for dice loss function L_{dice} given by

$$L_{dice} = \frac{2 \times \sum_{i,j} G_{i,j} S_{i,j}}{\sum_{i,j} G_{i,j} + \sum_{i,j} S_{i,j}} \tag{5.1}$$

where $S_{i,j}$ and $G_{i,j}$ denote the segmentation output and the ground-truth respectively at (i, j).

5.2.2 Selection of Stratum Corneum patches

The analysis of the whole Stratum Corneum (SC) layer for neutrophil detection is done by selecting several patches and then classifying them by the trained network. The patch selection approach should satisfy the following two important criteria- (a) all pixels in an SC patch should belong to a perceptually similar region and (b) the union of the selected patches should cover the entire SC layer. In the proposed framework, the generation of perceptual regions in the whole slide biopsy images is achieved by the application of Simple Linear Iterative Clustering

(SLIC) [1] superpixel algorithm. After superpixel decomposition, the symmetric region around the centroids of the superpixels which lie in the SC layer are selected for classification.

5.2.3 Classification of Stratum Corneum patches

The architecture of the capsule network trained for classification of Stratum Corneum (SC) patches is shown in Figure 5.4. The capsule makes a strong representational assumption that at each location in the image, there is at most one instance of the type of entity that a capsule represents. So, the network is designed in such a way that the receptive fields of the capsules in the secondary capsule does not contain multiple neutrophils. In the proposed network, the secondary capsule does routing by agreement between the capsules in the same spatial region of the primary capsule. This routing strategy reduces the model parameters drastically. The probability of neutrophil in a particular portion of the image patch is denoted by the magnitude of the output vector of the secondary capsule. There can be multiple neutrophils in a particular image patch. Hence, the $top - K$ average pooling layer is kept after secondary capsule layer to get the average of the top-K probabilities. The output from the $top - K$ average pooling layer is the probability of neutrophil in the input SC patch. This network is further referred to as CapsDeMM (**Caps**ule network for **De**tection of **M**unro's **M**icroabcess).

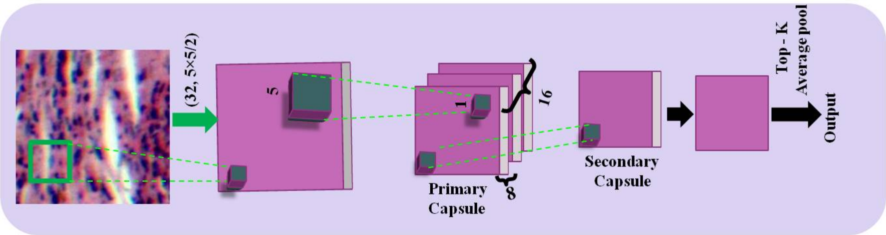


Figure 5.4: The proposed four layered capsule network for neutrophil detection: (i) Convolution layer (number of filters 32, filter size 5, stride 2), (ii) Primary capsule layer (8 convolutional units of kernel size 5 × 5, stride 2, number of channels 16), (iii) Secondary capsule layer and (iv) $top - K$ Average Pooling layer.

Given an input SC patch P, let P_{ne} denote the probability of **ne**utrophil in the patch. Then, during training, CapsDeMM minimizes the binary cross entropy (BCE) loss function L_{BCE} given by

$$L_{BCE} = -y_P * log(P_{ne}) - (1 - y_P) * log(1 - P_{ne}) \tag{5.2}$$

where y_P equal to 1 if neutrophil accumulation is observed in the patch and y_P equal to 0 otherwise.

Advantages of CapsDeMM over traditional CNNs

For the present problem of Stratum Corneum (SC) patch classification, the use of CapsDeMM provides two key advantages over traditional Convolutional Neural Networks (CNN). First of all, CapsDeMM generates the probability map of neutrophil's existence for every spatial location. So, unlike other CNNs (like ResNet), it is able to explain the classification output. Subsequently, the generation of neutrophils' probability map can be utilized for the segmentation of neutrophils in the SC patches. The segmentation (coarser segmentation) of neutrophils in SC patches without using any segmentation ground-truth is an additional advantage of CapsDeMM. Furthermore, the CapsDeMM is free from the restriction that the training and test images must be of equal size.

5.3 Experimental Protocol

5.3.1 Data set

For the data set used in this experiment, the help of an expert dermatologist is sought. On clinical confirmation of psoriasis, affected tissues are collected in 10% formalin under the supervision of the dermatologist. Next, the formalin-fixed tissues are dehydrated and then embedded in paraffin blocks. The histopathological slides are prepared from tissue sections of 5μM, which are stained with hematoxylin and eosin. Images of these slides are taken using a digital camera connected to a microscope with 10X magnification. While taking the images, our goal was to fit the entire tissue sample in the field of view of the camera. Note that 10X is the maximum possible magnification that achieves this. The images are of 1936×2584 pixels. The present research is carried out following the Declaration of Helsinki principles. Written informed consent was taken from all the recruited patients.

The ‘Review Committee for Protection of Research Risks to Humans’ of Indian Statistical Institute, India gave necessary approval towards the data collection.

In total, 120 patients participated in this study and gave samples. All serial sections of skin tissue that are present in a slide are imaged separately. In some images, the Stratum Corneum layer is lost during tissue processing. Such images are discarded. Two experts are involved for labeling the remaining images. The data set is built with those images for which both experts' labels match. The data set consists of 273 images among which 88 images have Munro's Microabscess and the remaining 185 images do not have Munro's Microabscess. An expert is employed for ground-truth annotation of SC segmentation. For constructing the patch classification data set, several squared patches (224×224 pixel-sized) are cropped from stratum corneum regions. The existence of neutrophil in these patches is labeled by two experts. Cases, where there is a disagreement among experts, are discarded. Altogether, there are 886 patches with neutrophils and 1700 patches without any neutrophil. Thereafter, SC patches with and without neutrophils are termed as positive and negative patches respectively.

5.3.2 Experimental Setting

Three-fold cross-validation is used for evaluating the performance of the proposed system. In every fold, equal (91) number of biopsy images are kept and the images for a fold is selected by random sampling. Every fold contains nearly equal (29, 29, 30) number of biopsy images where Munro's Microabscess is present. The cropped SC patches are grouped fold-wise (862 patches/fold) to build the fold-wise SC patch classifiers. In order to set the training hyper-parameters, 10% training images (original WSI) are used in UCaps and 20% training SC patches are used in CapsDeMM.

5.3.3 Evaluation strategy

In this paper, before evaluation of the WSI diagnosis performance, both capsule network-based SC segmentation module and SC patch classification module are evaluated individually. The quantitative evaluation of the SC segmentation performance is done by using four metrics- (i) Ratio of Correct Pixel Classification (RCPC), (ii) Dice Coefficient (DC), (iii) True Positive Rate (TPR) and (iv) False

Positive Rate (FPR). On the other hand, the quantitative evaluation of SC patch classification is done by (i) Recall, (ii) Precision, (iii) F1 Score and (iv) ACC (average classification accuracy).

Ideally, the detection of a single SC patch having neutrophils indicates that the biopsy image contains Munro's Microabscess. But it is difficult to develop a classifier that will classify all SC patches correctly. The potential error in Munro's Microabscess detection in WSIs due to the limitation of the SC patch classifier is dealt with deciding an optimal threshold T. The decision criteria for having Munro's Microabscess in WSIs will be the detection of more than T number of positive patches. Two different strategies are analyzed for selecting the value of T: Strategy I- The proposed system should produce best WSI classification result: T is selected in such a way that it will maximize the WSI classification accuracy; Strategy II- The proposed system will reduce the workload of pathologists by discarding the negative cases (WSIs without having Munro's Microabscess): T is selected in such way that True Negative Rate (TNR) is maximized. Average classification accuracy (ACC) is used to evaluate the performance when Strategy I is used. On the other hand, True Negative Rate (TNR) and Precision are used to evaluate the performance when Strategy II is used. T is decided by analyzing the training WSIs and the value of T varies across the folds for a particular thresholding strategy.

5.4 Results and Discussion

5.4.1 Stratum Corneum Segmentation

Table 5.1: Comparison of Stratum Corneum segmentation performance.

Metric	UNet	UNet (PP)	UCaps	UCaps (PP)
RCPC	0.9629 ± 0.0080	0.9676 ± 0.0057	0.9690 ± 0.0077	0.9698 ± 0.0077
DC	0.8493 ± 0.0247	0.8614 ± 0.0140	0.8778 ± 0.0185	0.8801 ± 0.0203
TPR	0.8716 ± 0.0212	0.8622 ± 0.0236	0.8927 ± 0.0356	0.8888 ± 0.0380
FPR	0.0217 ± 0.0101	0.0149 ± 0.0068	0.0166 ± 0.0010	0.0149 ± 0.0009
Parameters	723457		354984	

The UCaps architecture shown in Figure 5.3 is used for Stratum Corneum layer segmentation. The segmentation performance of the proposed network is compared with the UNet shown in Fig. 5.5. In many cases, both segmentation networks

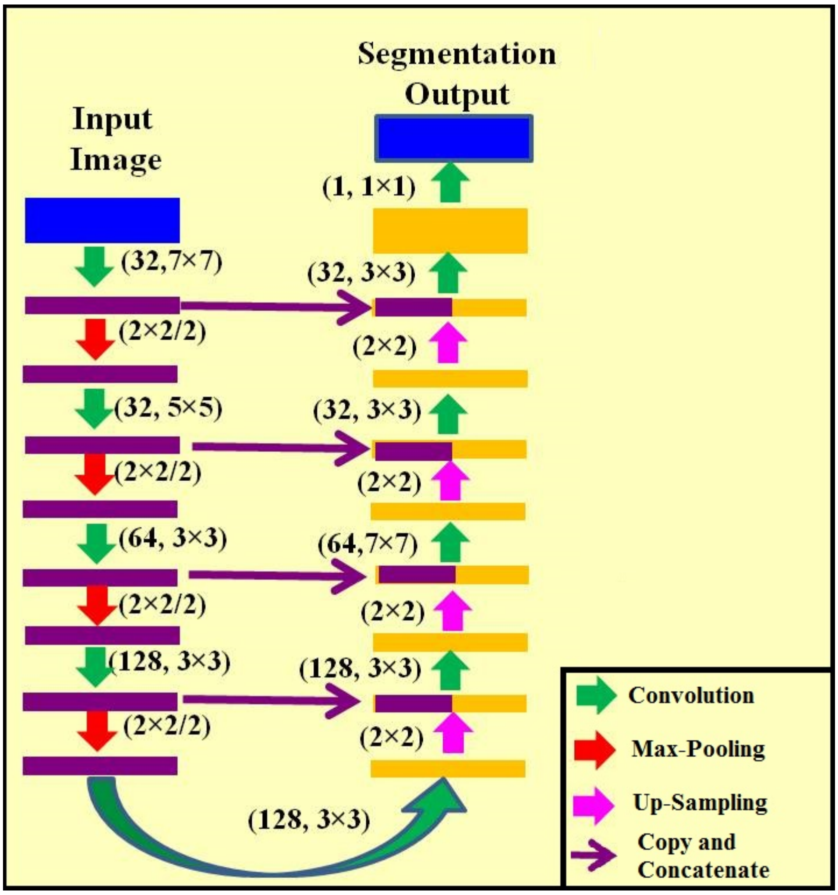


Figure 5.5: UNet architecture. Parameters of the convolutional layer is kept in (Number of filters,Filter size) order, max-pooling layer is kept in (window size/stride) order and up-sampling layer is kept in (Up-sample Ratio) order.

generate several spurious holes. These unwanted holes are removed by the hole-filling algorithm. The overlay segmentation boundary after post-processing (PP) by hole-filling is shown in Figure 5.6. It is difficult to compare the performance of the chosen models using visual inspection. Hence, quantitative evaluation of the SC segmentation outcomes is done by evaluation metrics mentioned in section 5.3.3. The average performance over all three folds for both UNet and UCaps are listed in Table 5.1. According to Table 5.1, both networks produce comparable segmentation performance but the performance of the proposed CapsNet based segmentation network is better than the competing network (UNet). Moreover, UCaps uses less than half of the number of parameters of UNet for producing the SC segmentation result. Thus the proposed segmentation algorithm not only improves the segmentation performance but also reduces the memory requirement for storing the model parameters.

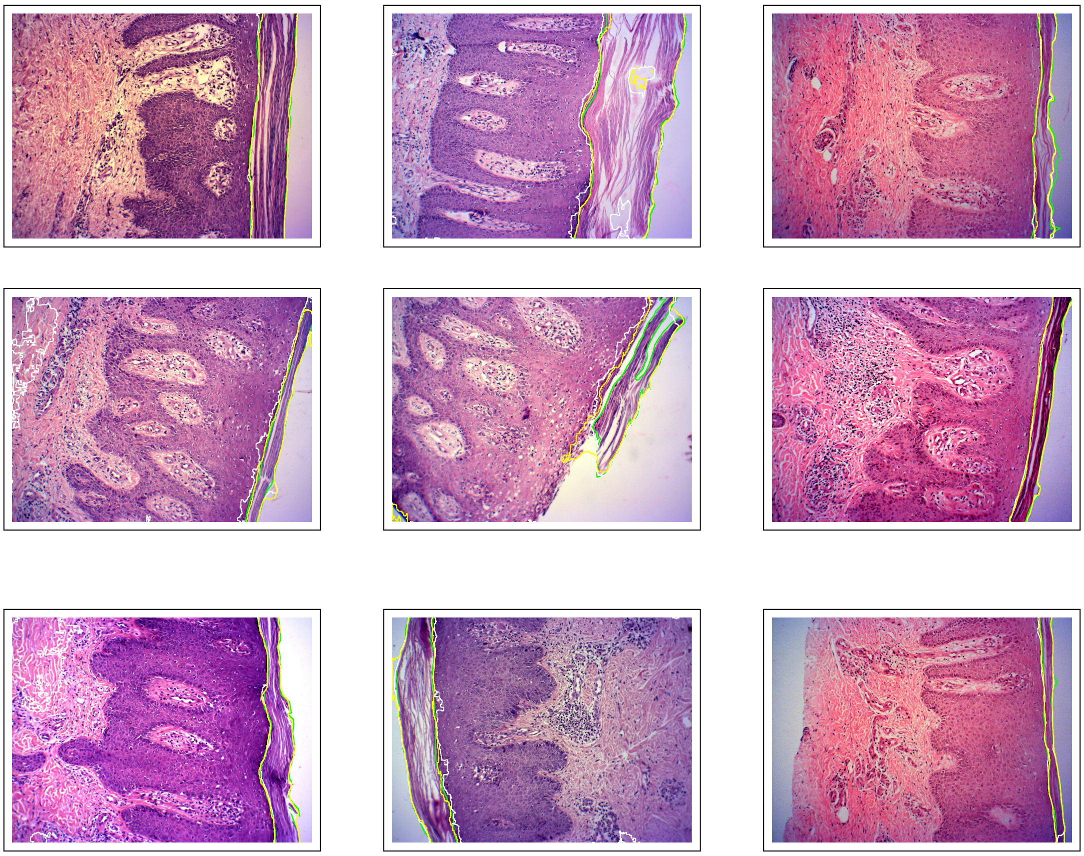


Figure 5.6: The Stratum Corneum segmentation output. Ground-truth segmentation border is outlined with Green, the U-Net segmentation border is outlined with White and the proposed approach's segmentation border is outlined with Yellow.

5.4.2 Stratum Corneum patch classification

The capsule network shown in Figure 5.4 is used for Stratum Corneum patch classification. In $top-K$ average pooling, lower values of K might misclassify an image patch as positive due to some portions of the patches getting high probabilities whereas higher values of K might overcompensate this effect leading to positive samples being classified as negative. To get the optimum value of K, the ROC curves for different values of K were compared. The comparison of ROC curves for first fold for five different K values (1, 3, 5, 7 and 9) is shown in Figure 5.7. In the proposed system, $K = 5$ is chosen for pooling since it provides the best AUC score (average of three folds). The diagnosis for the presence or absence of neutrophils in a patch is made by comparing the cut off value obtained from the ROC analysis. In case the network output for a patch is less than the cut off value, the predicted output is negative, otherwise, the predicted output is positive. This SC patch classification network is further referred as **CapsDeMM-5**.

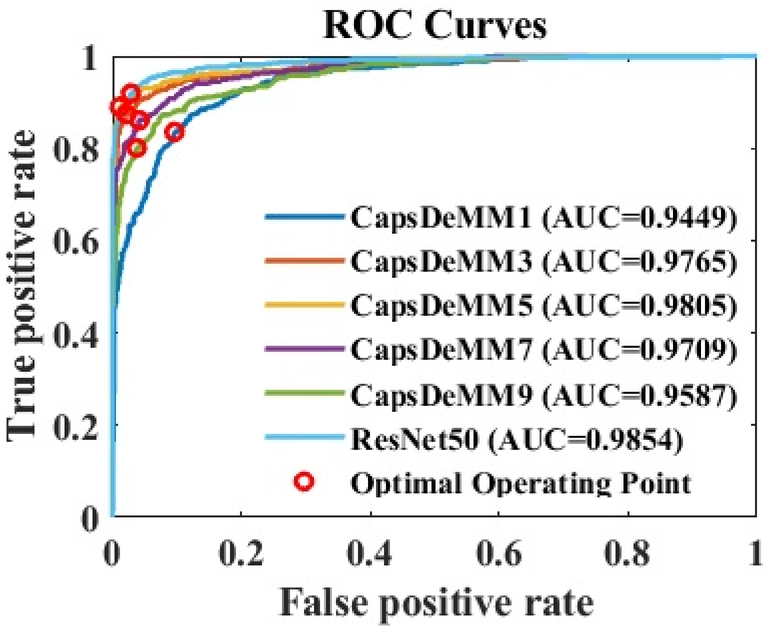


Figure 5.7: Comparison of ROC curves.

The diagnosis for the presence or absence of neutrophils in a patch is made by comparing the optimal cut off obtained from the Receiver Operating Curve (ROC). Only those patches where output from the classification network is greater than or equal to the cut off are predicted as positive patches. The evaluation metrics mentioned in section 5.3.3 are used for quantitative evaluation of the classification performance. The mean and standard deviation of the evaluation metrics for all three folds are listed in Table 5.2. Finally, the performance of the proposed capsule network is compared with a trained CNN i.e. ResNet-50, trained on the same data set. According to Table 5.2, **CapsDeMM-5** achieves comparable accuracy to ResNet-50 despite having orders of magnitude less parameters (23.5 million vs 0.1 million).

Table 5.2: Stratum Corneum patch classification. ResNet-50 vs CapsDeMM-5. M = Million, s = second, s/img= second/image.

Evaluation Metric	ResNet-50 (mean± std)	CapsDeMM-5 (mean± std)
Recall	0.87 ± 0.078	0.84 ± 0.064
Precision	0.95 ± 0.015	0.94 ± 0.027
F1 Score	0.91 ± 0.036	0.88 ± 0.027
ACC	0.94 ± 0.020	0.92 ± 0.013
Parameters	23.5 M	0.1 M
Load Time	2.15 s	0.26 s
Prediction Time	0.021 s/img	0.087 s/img

Table 5.3: Performance comparison of WSI classification using 224 × 224 sized test patches. Strategy I: To maximize classification accuracy. Strategy II: To maximize true negative rate.

Number of Superpixel	Classifier	UNet			UCaps		
		Strategy I	Strategy II		Strategy I	Strategy II	
		ACC (%)	TNR	Precision	ACC (%)	TNR	Precision
300	ResNet-50	86.13	1.0000	0.4649	86.85	0.9782	0.4736
	CapsDeMM-5	88.71	0.9606	0.4373	91.98	0.9885	0.4391
500	ResNet-50	85.76	0.9837	0.4479	86.49	0.9774	0.4277
	CapsDeMM-5	88.70	0.9656	0.4291	89.06	0.9861	0.4649
700	ResNet-50	87.59	0.9651	0.5068	87.95	0.9775	0.5103
	CapsDeMM-5	89.06	0.9333	0.4000	89.79	0.9865	0.4120

5.4.3 Neutrophil localization

According to the architecture of CapsDeMM, the output obtained from the second last layer (after secondary capsule layer) is nothing but the segmentation maps of the neutrophils present in the input patches. The pictorial representation of the Stratum Corneum patches along with their coarse segmentation maps (after up-scaling) obtained from CapsDeMM-5 is shown in Figure 5.8. Figure 5.8 shows that in most cases, CapsDeMM-5 successfully localizes (spots) the neutrophils present in the SC patches. Subsequently, the proposed research provides a noble direction for producing neutrophil's segmentation map in SC patches without using any segmentation ground-truth.

5.4.4 Whole Slide Biopsy Image (WSI) Diagnosis

The correct diagnosis of the Whole Slide Biopsy Images (WSI) is dependent on the Stratum Corneum segmentation outcome and the robustness of the classifier used for classifying the SC patches. In this chapter, both the number of superpixels and the size of the sampled test patches (SC patches) are varied. The performance assessment of the proposed system is done by the evaluation strategies explained in section 5.3.3. The WSI diagnosis performance is analyzed for two different situations separately: (i) when the test patch size is equal to the training patch size and (ii) when the test patch size is not equal to the training patch size. The detailed analysis of them is given in the following subsections.

5.4.4.1 Equal sized train and test patches

224 × 224 sized images are used to train the SC patch classifier. So, after SC segmentation, the test patches (SC patches) of size 224×224 are sampled from the

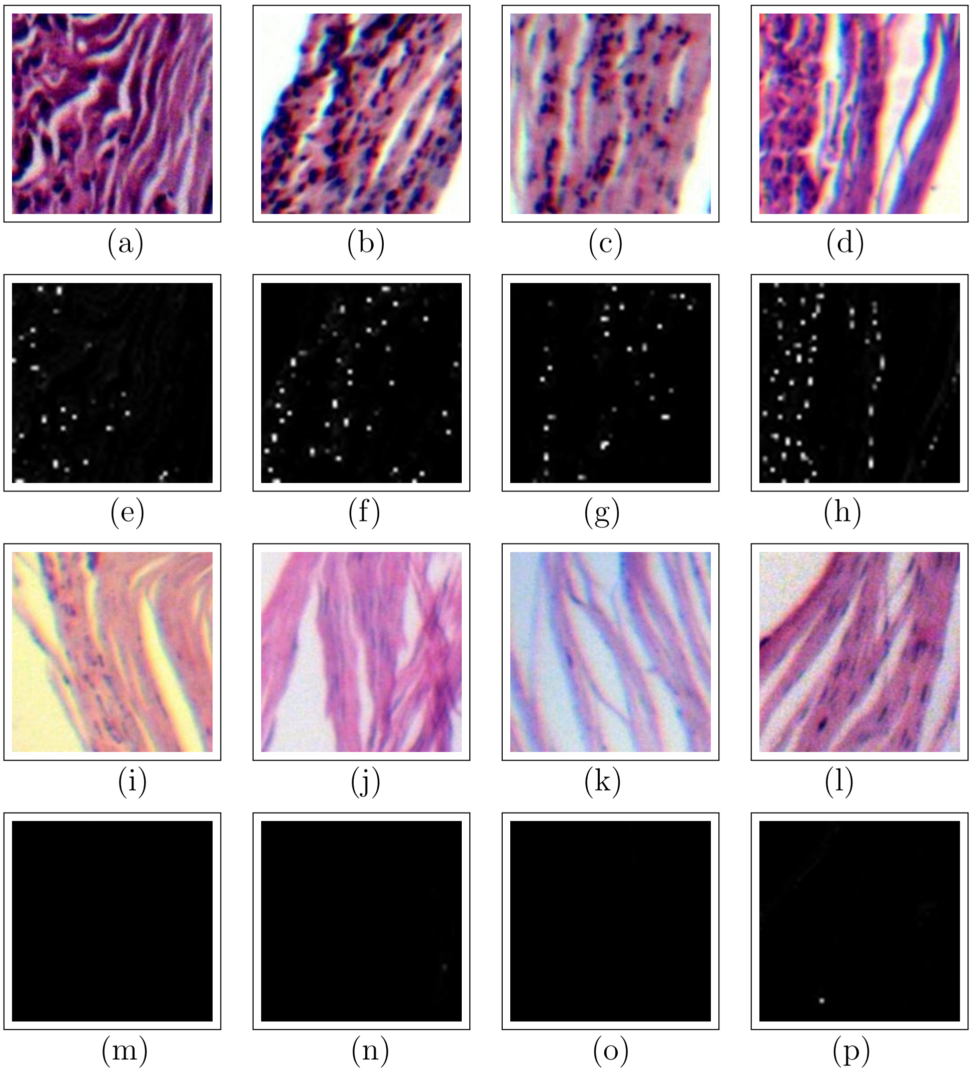


Figure 5.8: Neutrophil localization. Images (a)-(d) are Stratum Corneum patches where neutrophils are present. Images (e)-(h) are the neutrophil segmentation map of the images(a)-(d) using CapsDeMM-5. Images (i)-(l) are Stratum Corneum patches where neutrophils are absent. Images (m)-(p) are the neutrophil segmentation map of the images(i)-(l) using CapsDeMM-5.

segmented region and classified to produce the diagnosis result. For comparison purposes, we use two different segmentation networks (UNet and UCaps) and two different SC patch classification networks (ResNet-50 and CapsDeMM-5). The average WSI diagnosis performance (on the basis of evaluation strategies described in section 5.3.3) across three folds are reported in Table 5.3.

When the performance comparison is done on the basis of Strategy I, (see Table 5.3), the UCaps based system consistently gives better performance than UNet based system when classification network and the number of superpixels are fixed. Furthermore, CapsDeMM-5 based system consistently gives better performance

Table 5.4: CapsDeMM-5 based performance comparison of WSI classification using rectangular patches

Number of Superpixel	Test Patch Size = (224,112)			Test Patch Size = (112,224)		
	Strategy I	Strategy II		Strategy I	Strategy II	
	ACC (%)	TNR	Precision	ACC (%)	TNR	Precision
300	88.71	0.9887	0.4373	87.60	0.9843	0.5258
500	88.70	0.9868	0.4291	87.24	0.9773	0.4992
700	89.06	0.9859	0.4000	88.70	0.9817	0.4983

than ResNet-50 based system when the segmentation network and the number of superpixels are fixed.

When the performance comparison is done on the basis of Strategy II, (see Table 5.3), it is observed that for a fixed number of superpixels the CapsDeMM-5 based system gives comparable performance to ResNet-50 based system when UNet is used for SC segmentation. On the other hand, when UCaps is used for SC segmentation, in most cases the CapsDeMM-5 based system gives better performance than ResNet-50 based system. The promising result of WSI diagnosis using the capsule network justifies the use of it for constructing both segmentation and classification network in the proposed framework.

5.4.4.2 Unequal sized train and test patches

In this section, we shall analyze the WSI diagnosis performance of our system when sampled SC patches (test) are of size 224×112 or 112×224 but SC patch classifier is trained with patches of size 224×224. In this analysis, we consider UCaps for SC segmentation as it gives better result than UNet, and use CapsDeMM-5 for SC patch classification as it is flexible to predict classification output from the chosen patch size. We vary the number of superpixels and compute the average WSI diagnosis performance (evaluation strategies described in section 5.3.3) across three folds. The average results for all considered combinations are reported in Table 5.4. Table 5.4 shows that our system gives promising performance for different test patch sizes.

5.5 Summary

This chapter presents a machine-assisted approach for the detection of Munro's Microabscess in psoriatic skin biopsy images. This research may be considered as a pioneering attempt towards the development of an automated system for detection of Munro's Microabscess in psoriatic skin biopsy images. The proposed capsule network-based framework not only reduces the model parameters but also improves the diagnosis performance on the basis of classification accuracy. The promising performance of the Munro's Microabscess detection system presented in the paper shows their applicability for diminishing the workload of the pathologists by a huge margin. The use of Mega-pixel images reduces the overall computational burden as well as attests the use of a low-cost system consisting of a light microscope (without digital scanner) and a digital camera. Moreover, unlike traditional CNNs (e.g. ResNet-50), CapsDeMM can explain the neutrophil prediction result through localization.

Chapter 6

Conclusion

This thesis deals with the development of computer vision algorithms for machine-assisted analysis of psoriasis. The research presented here is significant both for the computer vision researchers and the dermatologists. The developed algorithms, use of existing algorithms for the chosen tasks attract the computer vision researchers. On the other hand, this research is important to the dermatologists as it addresses the clinical goals for studying psoriasis. In many cases, psoriasis makes a patient mentally disabling as it destroys the beauty of the human skin. It is a life-long disease as no drug is there which can cure it permanently. The response of the drugs which are applied to control the severity of this disease varies from patient to patient. So, for the selection of proper drug and drug dosage, severity assessment is important. However, the present way of severity assessment is suffering from both inter- and intra-observer variability. So, a novel attempt is made to develop an algorithmic way to standardize the severity assessment procedure. On the other hand, this thesis develops computer vision algorithms for analyzing the skin biopsy for histopathological inspection for detecting important characteristics of psoriasis. The following subsections summarize the technical contribution of the thesis and a set of issues for future research in the related field.

6.1 Achievements

The key goals of computer vision researchers for psoriasis image analysis are two-fold: (i) development of an automatic system to track the severity of the psoriasis patient and (ii) automatic histopathological report generation from skin biopsy images. The available technical innovation related to psoriasis image analysis is discussed in Section 1.5 of Chapter 1. The literature shows that although several researchers show their research interest in this field during the last twenty (20) years, we are far apart from developing a usable image-based severity tracking system and automatic histopathological report generation. The key reasons are the unavailability of skin image data set, lack of skin imaging standard as well as unavailability of suitable computer vision algorithms. In this respect, this thesis is an important addition to the state-of-the-art research in psoriasis image analysis field. The major outcomes of the research conducted under this thesis are as follows:

- This thesis develops four different annotated image data sets for attempting four distinct computer vision challenges related to psoriasis. These data sets are publicly available for academic research. Hence, it can be claimed that this thesis takes a significant step to develop psoriasis skin image data set for conducting research in this field.

- Literature shows that several attempts have been made for the development of an automatic psoriasis image analysis system with the help of traditional computer vision techniques. However, most of the attempts restrict themselves to narrow research goals. This thesis plays a leading role where the latest advancement of image analysis techniques are used for solving diverse computer vision challenges associated with psoriasis.

- Dealing with multivariate circular-linear data is useful in many practical applications. For such problems, due to the scarcity of algorithmic approaches, in many cases, researchers ignore the correlation between the components of the data vectors. In this respect, this thesis develops a robust algorithm for mixture modeling of circular-linear data which considers the underlying correlation between the components of the data vectors. Another advantage of this algorithm is that it allows heterogeneous mixture components. Although the proposed mixture modeling technique is used for the purpose of psoriatic plaque segmentation, it can be used for any color image segmentation problem or clustering of bi-variate circular-linear data.

- The present thesis for the first time considers to predict erythema, scaling and induration score simultaneously using the Multi-task learning paradigm.

Moreover, the present study reveals the advantage of treating the severity assessment task as an ordinal classification problem.

- A pioneering attempt is presented for skin layer segmentation in biopsy images using the latest advances in machine learning algorithms. The design of a deep convolutional neural network for superpixel classification and the design of the UNet for semantic segmentation are two important contributions of this research.
- Manual labeling of multiple smaller objects in high-dimensional images is very expensive, time-consuming, and error-prone; and hence, it may not be available. Hence, the proposed capsule network-based framework developed for detection/localization of multiple copies of neutrophils in the high-dimensional skin histopathology images in a weakly supervised manner is an important outcome from this thesis.
- Detection of Munro's Microabscess in biopsy images is attempted for the first time. The proposed framework uses a low-cost imaging system and the latest advancement of the deep neural network for analyzing Mega-pixel images in an efficient way. The initial research for Munro's Microabscess detection reported in Medical Image Computing and Computer-Assisted Intervention (MICCAI, 2018) has already created a significant impact on the related research field.

6.2 Scope of future works

The current thesis opens up several opportunities for conducting further research for the development of useful CAD tools for clinical use. Such scopes are summarized as below:

- In this thesis, for conducting psoriatic plaque segmentation and severity assessment tasks, a limited effort is given in maintaining the sample population with respect to age-group, plaque region, skin color tone, etc. and experimental results are reported on the basis of a small size data set. A detailed analysis of sample characteristics for customized model production is the immediate future scope of this work. On the other hand, the study of the impact of variation of cameras, lighting conditions, camera operators will provide an idea about how well the method generalizes to be addressed next.
- From the pattern recognition point of view, JCLMM based clustering approach can be extended in several directions. The construction of JCLMM opens a scope for inserting other circular distribution as well. Other variants of circular distributions may also be studied. JCLMM provides a satisfactory model in two-dimension, however, the present version cannot be directly extended to multi-dimensions. Future research can explore this issue.
- From the application point of view, the proposed psoriatic plaque segmentation approach can be extended in the following ways. Firstly, the JCLMM algorithm can be used for other color image segmentation purposes. Secondly, although the proposed localization approach is applied to psoriasis images, it can be applied to other human skin diseases lesion localization purposes.
- Designing an evaluation scheme to assess the impact of color calibration or specular reflections in the skin images are considered as the future scope of the JCLMM based skin image segmentation task.
- The segmentation algorithm and the subsequent disease localization method presented for psoriatic plaques can be used for computing PASI score and severity analysis. It can be further used for disease monitoring too so that a dermatologist can assess whether a given treatment is responding well.

Further, the development of color tone and geographic region-specific skin color distribution model may be explored for achieving better localization results.

- The present study develops classification models for automatic severity scoring of psoriasis using cropped skin patches. The developed model may have biased towards specific gender (male), color-tone, age-group or body parts where it is located. Development of custom models or evaluation of performance when a deep model is trained with uniformly sampled data can be explored next.

- Increment of data volume for evaluating severity assessment task, analysis explain-ability of the models and the role of diverse data augmentation models for robust classifier design will be considered as important future work of this research.

- For ordinal classifier design, the present thesis uses Earth Mover Distance-based loss function. However, recently introduced 2-Wasserstein distance-based CNN [50] can also be explored next for the ordinal classifier.

- The present study scores the severity of psoriasis image with respect to erythema, scaling, and induration. The classifier can be applied for severity assessment for other skin diseases where erythema, scaling, or induration are present. The reduction of model complexity and deployment of the developed models in mobile devices for clinical use are certainly issues to be addressed in the future.

- In order to cope up with the challenges (in terms of producing real-time output) behind the semantic segmentation of skin biopsy, the present thesis proposed to use two different approaches namely, superpixel classification based approach and Fully Convolutional Network (FCN) based approach. However, training an FCN in the superpixel setting will speed up the execution further. Moreover, superpixel based FCN will reduce the memory requirement. But present scenario, deep models can not be trained with superpixels. Hence, training of end-to-end deep model in a superpixel setting will be considered as an interesting challenge to the vision community.

- Present thesis uses commonly used Markov Random Field-based approach for smoothing the segmentation outcome after superpixel classification. However, recently introduced Mask-RCNN [48] for superpixel smoothing to be explored next.

- This thesis initiates the psoriatic skin biopsy analysis task with layer segmentation and Munro's Microabscess detection but remaining histopathological feature extraction can be attempted next. For instance, finding the maximum and minimum length of the epidermis, the number of taped dermis region inside the epidermis, etc. Analysis of dermis can be conducted for computing the number and length of blood vessels, detecting the accumulation of immune cells around the blood vessels, etc.

- This thesis design a capsule network (CapsNet) based framework for the detection of multiple copies of the neutrophils available in the SC layer in the high-dimensional histopathology images using a weakly supervised approach. The proposed method is used for neutrophil localization in skin stratum corneum. However, the proposed CapsNet can be utilized for different imaging applications where the task is to find multiple copies of an object.

- Analysis of histopathological condition for psoriasis is important for clinical decision making. This thesis starts with the segmentation of the skin layer and the detection of Munro's Microabscess. The biopsy image data set can be further utilized for addressing other histopathological analysis problems like the presence and nature of rete-pegs, detection of parakeratosis (presence of nucleated cells within stratum corneum) classification of parakeratosis (confluent/focal), detection of Kogoj Microabscesses, etc.

- The present work assumes that psoriasis images are fed to the system. However, from the practical viewpoint, skin disease classification task could be a nice addition to the existing framework. Given an affected skin, it would be of great clinical help if the kind of skin disease (e.g., eczema, leprosy, lichen planus, psoriasis, etc.) can be categorized in the beginning.

Publication related to this thesis

Journal

- Anandarup Roy, **Anabik Pal**, Utpal Garain, “JCLMM: A finite mixture model for clustering of circular-linear data and its application to psoriatic plaque segmentation”, Pattern Recognition, Volume 66, June 2017, Pages 160-173, ISSN 0031-3203.
- **Anabik Pal**, Utpal Garain, Aditi Chandra, Raghunath Chatterjee, Swapan Senapati, “Psoriasis skin biopsy image segmentation using Deep Convolutional Neural Network”, Computer Methods and Programs in Biomedicine, Volume 159, 2018, Pages 59-69, ISSN 0169-2607.

Conference

- **Anabik Pal**, Anandarup Roy, Kushal Sen, Raghunath Chatterjee, Utpal Garain, and Swapan Senapati, “Mixture Model based Color Clustering for Psoriatic Plaque Segmentation”, in 3rd IAPR Asian Conf on Pattern Recognition (ACPR), Kuala Lumpur, November, 2015, pp. 376-380.
- **Anabik Pal**, Utpal Garain, Raghunath Chatterjee and Swapan Senapati, “Psoriatic Plaque Segmentation in Skin Images”, in 5th National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Patna, December, 2015, pp. 1-4.
- **Anabik Pal**, Akshay Chaturvedi, Utpal Garain, Aditi Chandra and Raghunath Chatterjee, “Severity grading of psoriatic plaques using deep CNN based multi-task learning”, 23rd International Conference on Pattern Recognition (ICPR), Cancun, 2016, pp. 1478-1483.

- **Anabik Pal**, Akshay Chaturvedi, Utpal Garain, Aditi Chandra, Raghunath Chatterjee and Swapan Senapati "CapsDeMM: Capsule network for Detection of Munro's Microabscess in skin biopsy images", 21st International Conference On Medical Image Computing & Computer Assisted Intervention (MICCAI), Granada, Spain, 2018, pp.389-397.

Workshop

- **Anabik Pal**, Akshay Chaturvedi, Utpal Garain, Aditi Chandra, Raghunath Chatterjee and Swapan Senapati: "Severity assessment of psoriatic plaques using deep CNN based Ordinal classification", International workshop on Skin Lesion Analysis Towards Melanoma Detection, Held in Conjunction with MICCAI, Granada, Spain, 2018, pp. 252-259.

Bibliography

[1] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282.

[2] Agiomyrgiannakis, Y. and Stylianou, Y. (2009). Wrapped gaussian mixture models for modeling and high-rate quantization of phase data of speech. *IEEE Transactions on Audio, Speech, and Language Processing*, 17(4):775–786.

[3] Badrinarayanan, V., Kendall, A., and Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12):2481–2495.

[4] Bagon, S. (2006). Matlab wrapper for graph cut.

[5] Bahlmann, C. (2006). Directional features in online handwriting recognition. *Pattern Recognition*, 39(1):115–125.

[6] Banerjee, A., Dhillon, I. S., Ghosh, J., and Sra, S. (2005). Clustering on the unit hypersphere using von mises-fisher distributions. *Journal of Machine Learning Research*, 6:1345–1382.

[7] Bogo, F., Samory, M., Fortina, A. B., Piaserico, S., and Peserico, E. (2012). Psoriasis segmentation through chromatic regions and geometric active contours. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 5388–5391. IEEE.

[8] Boschetto, D., Mirzaei, H., Leong, R. W., and Grisan, E. (2016). Superpixel-based automatic segmentation of villi in confocal endomicroscopy. In *Biomedical and Health Informatics (BHI), 2016 IEEE-EMBS International Conference on*, pages 168–171. IEEE.

[9] Bougulia, N., Ziou, D., and Vaillancourt, J. (2004). Unsupervised learning of a finite mixture model based on the dirichlet distribution and its application. *IEEE Transactions on Image Processing*, 13(13):1533–1543.

[10] Boutemedjet, S., Bouguila, N., and Ziou, D. (2009). A hybrid feature extraction selection approach for high-dimensional non-gaussian data clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(8):1429–1443.

[11] Căliman, A. and Ivanovici, M. (2012). Psoriasis image analysis using color lacunarity. In *Optimization of Electrical and Electronic Equipment (OPTIM), 2012 13th International Conference on*, pages 1401–1406. IEEE.

[12] Caliman, A., Ivanovici, M., and Coliban, R.-M. (2013). Pseudo-granulometry and morphological covariance for color psoriasis image segmentation. In *E-Health and Bioengineering Conference (EHB)*, pages 1–4. IEEE.

[13] Caruana, R. (1997). Multitask learning. *Machine Learning*, 28(1):41–75.

[14] Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277.

[15] Chandra, A., Ray, A., Senapati, S., and Chatterjee, R. (2015). Genetic and epigenetic basis of psoriasis pathogenesis. *Molecular Immunology*, 64(2):313–323.

[16] Chandra, A., Senapati, S., Roy, S., Chatterjee, G., and Chatterjee, R. (2018). Epigenome-wide dna methylation regulates cardinal pathological features of psoriasis. *Clinical epigenetics*, 10(1):108.

[17] Chandran, V. and Raychaudhuri, S. P. (2010). Geoepidemiology and environmental factors of psoriasis and psoriatic arthritis. *Journal of Autoimmunity*, 34(3):J314–J321.

[18] Chang, C.-C. and Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[19] Chollet, F. et al. (2015). Keras. `https://github.com/fchollet/keras`.

[20] Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619.

[21] Cui, W., Wang, Y., Fan, Y., Feng, Y., and Lei, T. (2013). Localized fcm clustering with spatial information for medical image segmentation and bias field estimation. *Journal of Biomedical Imaging*, 2013:13.

[22] Dai, L., Fang, R., Li, H., Hou, X., Sheng, B., Wu, Q., and Jia, W. (2018). Clinical report guided retinal microaneurysm detection with multi-sieving deep learning. *IEEE Transactions on Medical Imaging*, 37(5):1149–1161.

[23] Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 886–893. IEEE.

[24] Das, A., Chandra, A., Lahiri, A., Datta, S., Senapati, S., and Chatterjee, R. (2016). Genetics of psoriasis. *eLS*.

[25] Delgado, D., Ersbøll, B., and Carstensen, J. (2004a). Sharp: A smart hierarchical algorithm to register psoriasis. In *Proc. of Int. Workshop on Systems, Signals and Image Processing*, pages 43–46.

[26] Delgado, D., Ersbϕll, B., and Carstensen, J. M. (2004b). An image based system to automatically and objectively score the degree of redness and scaling in psoriasis lesions. In *Proceedings fra den 13. Danske Konference i*, page 130.

[27] Deng, Y. and Manjunath, B. (2001). Unsupervised segmentation of color-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):800–810.

[28] Deng, Y., Manjunath, B. S., and Shin, H. (1999). Color image segmentation. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, volume 2, pages 446–451. IEEE.

[29] Emre Celebi, M., Alp Aslandogan, Y., Stoecker, W. V., Iyatomi, H., Oka, H., and Chen, X. (2007). Unsupervised border detection in dermoscopy images. *Skin Research and Technology*, 13(4):454–462.

[30] Fadzil, M. A. and Ihtatho, D. (2008). Modeling psoriasis lesion colour for pasi erythema scoring. In *Information Technology, 2008. ITSim 2008. International Symposium on*, volume 2, pages 1–6. IEEE.

[31] Farag, A., Lu, L., Roth, H. R., Liu, J., Turkbey, E., and Summers, R. M. (2017). A bottom-up approach for pancreas segmentation using cascaded superpixels and (deep) image patch labeling. *IEEE Transactions on Image Processing*, 26(1):386–399.

[32] Fernández-Durán, J. J. (2007). Models for circular-linear and circular-circular data constructed from circular distributions based on nonnegative trigonometric sums. *Biometrics*, 63:579–585.

[33] Figueiredo, M. A. T. and Jain, A. K. (2002). Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):381–396.

[34] Frank, E. and Hall, M. (2001). A simple approach to ordinal classification. In *Proceedings of the 12th European Conference on Machine Learning*, EMCL '01, pages 145–156, London, UK, UK. Springer-Verlag.

[35] Fredriksson, T. and Pettersson, U. (1978). Severe psoriasis–oral therapy with a new retinoid. *Dermatology*, 157(4):238–244.

[36] Gao, X., Lin, S., and Wong, T. Y. (2015). Automatic feature learning to grade nuclear cataracts based on deep learning. *IEEE Transactions on Biomedical Engineering*, 62(11):2693–2701.

[37] García-Portugués, E., Crujeiras, R. M., and González-Manteiga, W. (2013). Exploring wind direction and so_2 concentration by circular?linear density estimation. *Stochastic Environmental Research and Risk Assessment*, 27(5):1055–1067.

[38] Gerard, S. E., Patton, T. J., Christensen, G. E., Bayouth, J. E., and Reinhardt, J. M. (2018). Fissurenet: A deep learning approach for pulmonary fissure detection in ct images. *IEEE Transactions on Medical Imaging*, pages 1–1.

[39] Gessert, N., Lutz, M., Heyder, M., Latus, S., Leistner, D. M., Abdelwahed, Y. S., and Schlaefer, A. (2018). Automatic plaque detection in ivoct pullbacks using convolutional neural networks. *IEEE Transactions on Medical Imaging*, pages 1–1.

[40] Gomez, D. D., Maletti, G., Nielsen, A. A., and Ersboll, B. (2004). Multiset multitemporal canonical analysis of psoriasis images. In *Biomedical Imaging:*

Nano to Macro, 2004. IEEE International Symposium on, pages 1151–1154. IEEE.

[41] Gupta, M. D., Srinivasa, S., Madhukara, J., and Antony, M. (2015). Random forest based erythema grading for psoriasis. In *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, pages 819–823. IEEE.

[42] Güvenir, H. and Emeksiz, N. (2000). An expert system for the differential diagnosis of erythemato-squamous diseases. *Expert Systems with Applications*, 18(1):43–49.

[43] Güvenir, H. A., Demiröz, G., and Ilter, N. (1998). Learning differential diagnosis of erythemato-squamous diseases using voting feature intervals. *Artificial Intelligence in Medicine*, 13(3):147–165.

[44] Haggerty, J. M., Wang, X. N., Dickinson, A., O'Malley, C. J., and Martin, E. B. (2014). Segmentation of epidermal tissue with histopathological damage in images of haematoxylin and eosin stained human skin. *BMC Medical Imaging*, 14(1):7.

[45] Hani, A. F. M., Prakasa, E., Nugroho, H., Affandi, A. M., and Hussein, S. H. (2012a). Body surface area measurement and soft clustering for pasi area assessment. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 4398–4401. IEEE.

[46] Hani, A. F. M., Prakasa, E., Nugroho, H., and Asirvadam, V. S. (2012b). Implementation of fuzzy c-means clustering for psoriasis assessment on lesion erythema. In *Industrial Electronics and Applications (ISIEA), 2012 IEEE Symposium on*, pages 331–335. IEEE.

[47] Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621.

[48] He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask r-cnn. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2980–2988.

[49] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition.*

[50] He, R., Wu, X., Sun, Z., and Tan, T. (2019). Wasserstein cnn: Learning invariant features for nir-vis face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(7):1761–1773.

[51] Hinton, G. E., Sabour, S., and Frosst, N. (2018). Matrix capsules with EM routing. In *International Conference on Learning Representations*.

[52] Hou, L., Yu, C.-P., and Samaras, D. (2016). Squared earth mover's distance-based loss for training deep neural networks. *arXiv preprint arXiv:1611.05916*.

[53] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). Mobilenets: efficient convolutional neural networks for mobile vision applications. *CoRR*, abs/1704.04861.

[54] Hu, J., Chen, Y., Zhong, J., Ju, R., and Yi, Z. (2018). Automated analysis for retinopathy of prematurity by deep neural networks. *IEEE Transactions on Medical Imaging*, pages 1–1.

[55] Huang, L., Xia, W., Zhang, B., Qiu, B., and Gao, X. (2017). Msfcn-multiple supervised fully convolutional networks for the osteosarcoma segmentation of ct images. *Computer Methods and Programs in Biomedicine*, 143:67–74.

[56] Ihtatho, D., Fadzil, M. A., Affandi, A. M., and Hussein, S. (2007a). Area assessment of psoriasis lesion for pasi scoring. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 3446–3449. IEEE.

[57] Ihtatho, D., Fadzil, M. A., Affandi, A. M., and Hussein, S. (2007b). Automatic pasi area scoring. In *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on*, pages 819–822. IEEE.

[58] Jailani, R., Hashim, H., Taib, M. N., and Sulaiman, S. (2004). Border segmentation on digitized psoriasis skin lesion images. In *TENCON 2004. 2004 IEEE Region 10 Conference*, volume 100, pages 596–599. IEEE.

[59] Johnson, R. A. and Wehrly, T. E. (1978). Some angular-linear distributions and related regression models. *Journal of the American Statistical Association*, 73(363):602–606.

[60] Kawahara, J., Daneshvar, S., Argenziano, G., and Hamarneh, G. (2018). 7-point checklist and skin lesion classification using multi-task multi-modal neural nets. *IEEE Journal of Biomedical and Health Informatics*, pages 1–1.

[61] Kim, B. Y., Choi, J. W., Kim, B. R., and Youn, S. W. (2015). Histopathological findings are associated with the clinical types of psoriasis but not with the corresponding lesional psoriasis severity index. *Journal of Autoimmunity*, 27(1):26—-31.

[62] Kim, S. C. and Kang, T. J. (2007). Texture classification and segmentation using wavelet packet frame and gaussian mixture model. *Pattern Recognition*, 40(4):1207–1221.

[63] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 1097–1105.

[64] LaLonde, R. and Bagci, U. (2018). Capsules for object segmentation. *arXiv preprint arXiv:1804.04241*.

[65] Law, M. H., Figueiredo, M. A., and Jain, A. K. (2004). Simultaneous feature selection and clustering using mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1154–1166.

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

[67] Li, S., Liu, Z.-Q., and Chan, A. B. (2014). Heterogeneous multi-task learning for human pose estimation with deep convolutional neural network. *International Journal of Computer Vision*, 113(1):19–36.

[68] Li, W., Manivannan, S., Akbar, S., Zhang, J., Trucco, E., and McKenna, S. J. (2016). Gland segmentation in colon histology images using hand-crafted features and convolutional neural networks. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, pages 1405–1408. IEEE.

[69] Li, X., Chen, H., Qi, X., Dou, Q., Fu, C., and Heng, P. (2018). H-denseunet: Hybrid densely connected unet for liver and tumor segmentation from ct volumes. *IEEE Transactions on Medical Imaging*, pages 1–1.

[70] Liskowski, P. and Krawiec, K. (2016). Segmenting retinal blood vessels with deep neural networks. *IEEE Transactions on Medical Imaging*, 35(11):2369–2380.

[71] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A., Van Ginneken, B., and Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88.

[72] Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440.

[73] Lu, C. and Mandal, M. (2012). Automated segmentation and analysis of the epidermis area in skin histopathological images. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 5355–5359. IEEE.

[74] Lu, C. and Mandal, M. (2014). Efficient epidermis segmentation for whole slide skin histopathological images. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 5546–5549. IEEE.

[75] Lu, J., Kazmierczak, E., Manton, J. H., and Sinclair, R. (2013a). Automatic segmentation of scaling in 2-d psoriasis skin images. *IEEE Transactions on Medical Imaging*, 32(4):719–730.

[76] Lu, J., Kazmierczak, E., Manton, J. H., and Sinclair, R. (2013b). A quantitative technique for assessing the change in severity over time in psoriatic lesions using computer aided image analysis. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 2380–2383. IEEE.

[77] Mardia, K. and Sutton, T. (1978). A model for cylindrical variables with applications. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 229–233.

[78] Mardia, K. V. and Jupp, P. E. (2009). *Directional statistics*, volume 494. John Wiley & Sons.

[79] Marks, R. M., Knight, A., and Laidler, P. (2012). *Atlas of Skin Pathology*, volume 11. Springer Science & Business Media.

[80] Marsh, J. N., Matlock, M. K., Kudose, S., Liu, T., Stappenbeck, T. S., Gaut, J. P., and Swamidass, S. J. (2018). Deep learning global glomerulosclerosis

in transplant kidney frozen sections. *IEEE Transactions on Medical Imaging*, pages 1–1.

[81] Martin, D., Fowlkes, C., Tal, D., and Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 416–423. IEEE.

[82] McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*. John Wiley & Sons.

[83] Meilă, M. (2005). Comparing clusterings: an axiomatic view. In *Proceedings of the 22nd international conference on Machine learning*, pages 577–584. ACM.

[84] Mesejo, P., Pizarro, D., Abergel, A., Rouquette, O., Beorchia, S., Poincloux, L., and Bartoli, A. (2016). Computer-aided classification of gastrointestinal lesions in regular colonoscopy. *IEEE Transactions on Medical Imaging*, 35(9):2051–2063.

[85] Mobiny, A. and Van Nguyen, H. (2018). Fast capsnet for lung cancer screening. In *MICCAI Proceedings, Granada, Spain, Sept. 16-20.*

[86] Nasr-Esfahani, E., Samavi, S., Karimi, N., Soroushmehr, S. M. R., Ward, K., Jafari, M. H., Felfeliyan, B., Nallamothu, B., and Najarian, K. (2016). Vessel extraction in x-ray angiograms using deep learning. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 643–646.

[87] Naylor, P., Laé, M., Reyal, F., and Walter, T. (2018). Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Transactions on Medical Imaging*, pages 1–1.

[88] Nelsen, R. B. (2007). *An introduction to copulas*. Springer Science & Business Media.

[89] Ojala, T., Pietikainen, M., and Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987.

[90] Peel, D. and Mclachlan, G. J. (2000). Robust mixture modelling using the t–distribution. *Statistics and Computing*, 10:339–348.

[91] Permuter, H., Francos, J., and Jermyn, I. (2006). A study of gaussian mixture models of color and texture features for image classification and segmentation. *Pattern Recognition*, 39(4):695–706.

[92] Pham, D. L., Xu, C., and Prince, J. L. (2000). Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2(1):315–337.

[93] Pusey, W. A. (1933). *The history of dermatology.* Springfield, III. ; Baltimore, Md. : Charles C. Thomas.

[94] Qin, W., Wu, J., Han, F., Yuan, Y., Zhao, W., Ibragimov, B., Gu, J., and Xing, L. (2018). Superpixel-based and boundary-sensitive convolutional neural network for automated liver segmentation. *Physics in Medicine & Biology*, 63(9):095017.

[95] Raina, A., Hennessy, R., Rains, M., Allred, J., Hirshburg, J. M., Diven, D., and Markey, M. K. (2016). Objective measurement of erythema in psoriasis using digital color photography with color calibration. volume 22, pages 375–380. Wiley Online Library.

[96] Roenigk, H. H. (1998). *Psoriasis Basic and Clinical Dermatology.* CRC Press; 3 edition.

[97] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241, Cham. Springer International Publishing.

[98] Roy, A. and Parui, S. K. (2014). Pair-copula based mixture models and their application in clustering. *Pattern Recognition*, 47(4):1689–1697.

[99] Roy, A., Parui, S. K., and Roy, U. (2011). Color image segmentation using a semi-wrapped gaussian mixture model. In *Procedings of International Conference on Pattern Recognition and Machine Intelligence*, pages 148–153. Springer-Verlag.

[100] Roy, A., Parui, S. K., and Roy, U. (2016). Swgmm: a semi-wrapped gaussian mixture model for clustering of circular–linear data. *Pattern Analysis and Applications*, 19(3):631–645.

[101] Roy, A. G., Navab, N., and Wachinger, C. (2018). Recalibrating fully convolutional networks with spatial and channel 'squeeze amp; excitation' blocks. *IEEE Transactions on Medical Imaging*, pages 1–1.

[102] Sabour, S., Frosst, N., and Hinton, G. E. (2017). Dynamic routing between capsules. In *Advances in Neural Information Processing Systems*, pages 3859–3869.

[103] Shieh, G. S., Zheng, S. R., and Shimizu, K. (2006). A bivariate generalized von mises distribution with applications to circular genomes. Technical report, Institution of Statistical Science, Academia Sinica, Taiwan.

[104] Shrivastava, V. K., Londhe, N. D., Sonawane, R. S., and Suri, J. S. (2015). Reliable and accurate psoriasis disease classification in dermatology images using comprehensive feature space in machine learning paradigm. *Expert Systems with Applications*, 42(15-16):6184–6195.

[105] Shrivastava, V. K., Londhe, N. D., Sonawane, R. S., and Suri, J. S. (2016a). Computer-aided diagnosis of psoriasis skin images with hos, texture and color features. *Computer Methods and Programs in Biomedicine*, 126(C):98–109.

[106] Shrivastava, V. K., Londhe, N. D., Sonawane, R. S., and Suri, J. S. (2016b). A novel approach to multiclass psoriasis disease risk stratification: Machine learning paradigm. *Biomedical Signal Processing and Control*, 28:27–40.

[107] Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[108] Taur, J.-S. (2003). Neuro-fuzzy approach to the segmentation of psoriasis images. *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, 35(1):19–27.

[109] Taur, J. S., Lee, G.-H., Tao, C.-W., Chen, C.-C., and Yang, C.-W. (2006). Segmentation of psoriasis vulgaris images using multiresolution-based orthogonal subspace techniques. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 36(2):390–402.

[110] Tian, Z., Liu, L., Zhang, Z., and Fei, B. (2016). Superpixel-based segmentation for 3d prostate mr images. *IEEE Transactions on Medical Imaging*, 35(3):791–801.

[111] Unnikrishnan, R., Pantofaru, C., and Hebert, M. (2007). Toward objective evaluation of image segmentation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):929–944.

[112] Vedaldi, A. and Lenc, K. (2015). Matconvnet – convolutional neural networks for matlab.

[113] Xiong, Z., Fedorov, V. V., Fu, X., Cheng, E., Macleod, R., and Zhao, J. (2018). Fully automatic left atrium segmentation from late gadolinium enhanced magnetic resonance imaging using a dual fully convolutional neural network. *IEEE Transactions on Medical Imaging*, pages 1–1.

[114] Yan, Z., Zhan, Y., Peng, Z., Liao, S., Shinagawa, Y., Zhang, S., Metaxas, D. N., and Zhou, X. S. (2016). Multi-instance deep learning: Discover discriminative local anatomies for bodypart recognition. *IEEE Transactions on Medical Imaging*, 35(5):1332–1343.

[115] Yang, X., Yu, L., Li, S., Wen, H., Luo, D., Bian, C., Qin, J., Ni, D., and Heng, P. (2018). Towards automated semantic segmentation in prenatal volumetric ultrasound. *IEEE Transactions on Medical Imaging*, pages 1–1.

[116] Zhang, C. and Zhang, Z. (2014). Improving multiview face detection with multi-task deep convolutional neural networks. In *Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on*, pages 1036–1041. IEEE.

[117] Zhou, Y., Xu, J., Liu, Q., Li, C., Liu, Z., Wang, M., Zheng, H., and Wang, S. (2018). A radiomics approach with cnn for shear-wave elastography breast tumor classification. *IEEE Transactions on Biomedical Engineering*, 65(9):1935–1942.