Spring 2013

Fuzzy logic color detection: Blue areas in dermoscopy melanoma images

Mounika Lingala

Department:

FUZZY LOGIC COLOR DETECTION:

BLUE AREAS IN DERMOSCOPY

MELANOMA IMAGES

by

MOUNIKA LINGALA

A THESIS

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE IN COMPUTER ENGINEERING

2013

Approved by

Ronald J. Stanley, Advisor
William V. Stoecker
Randy H. Moss

ABSTRACT

Malignant melanoma is one of the deadliest forms of skin cancer and is one of the most rapidly increasing cancers in the world. Image analysis techniques for the early detection of melanoma are dependent upon the detection of multiple color shades in melanoma images. Use of magnified visible-light imaging by dermoscopy allows detailed investigation of these color shades. This research explores the fuzzy logic methods for the detection and characterization of blue color shades to discriminate melanoma from benign skin lesions. Standard thresholds are placed on the RGB dermoscopy images firstly and then the restrictiveness is explored by implementing fuzzy logic. Multiple shades of blue (lavender blue, light blue and dark blue) are segmented based on applying alpha cuts, for different combinations of color fuzzy sets, whereby, a pixel is included in the segmented blue area mask if the pixel satisfies the alpha-cut membership constraints for all of the color fuzzy sets. An accuracy of 82.7% was obtained at alpha cuts of 0.30, 0.40 and 0.65 from the features derived on an 866-image dataset using logistic regression model of Statistical Analysis Software and an accuracy of 81.4% was obtained at an alpha cut of 0.25 using the SVM classifier.

ACKNOWLEDGMENTS

I am greatly indebted to my advisor, Dr. R. Joe Stanley for providing me with an opportunity to work under him. I would like to express my gratitude for him for all his guidance, encouragement and continued patience throughout this research. His constant support, both professional and personal has been incredible. I am also thankful to him for supporting me financially.

I would like to thank Dr. Stoecker for sharing his knowledge. I wish to express my gratefulness to him for reviewing my work regularly and providing helpful comments for improving the work. I also wish to thank my committee member Dr. Randy H. Moss for his academic guidance and support.

I would take this opportunity to thank my parents for having faith in me and for their continual support in all times. I would also like to thank my sister, Warlu and Shourya for sharing every moment throughout my stay in United States.

Finally, I would like to thank all my friends especially Uday, Sadhana and Harshini for sharing my good and not so good times.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

Malignant melanoma is one of the deadliest forms of skin cancer that is curable if detected and treated in its early stages. It is very difficult to differentiate malignant melanoma and a benign skin lesion as they can appear similar. Malignant melanoma exacts the highest total mortality among all skin cancers and is one of the most rapidly increasing cancers in the world, with an estimated incidence of 76,250 and an estimated total of 9,180 deaths in the United States in 2012 alone [1]. The identification of specific patterns of colors and structures can better distinguish malignant and benign skin lesions. Dermoscopy is a non-invasive imaging technique that uses optical magnification to visualize features of melanocytic lesions unseen by the naked eye, such as blue pigmentation [2]. Figure 1.1 presents an example of a dermoscopy skin lesion image with blue pigmentation.

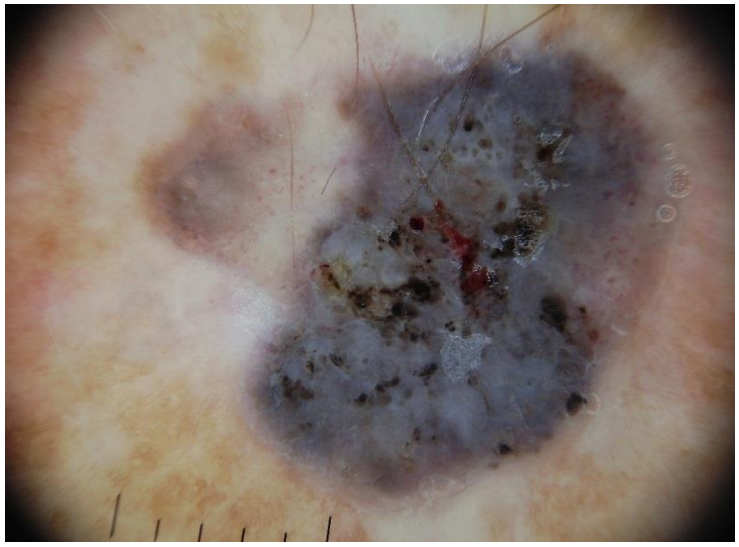


Figure 1.1. Dermoscopy skin lesion example

Blue coloration is present within several dermoscopic features: Homogeneous blue pigmentation; blue-white veil, and blue-black color. Homogeneous blue pigmentation is a first-step criterion in the 2-step procedure for examining pigmented skin lesions [3]. Blue-white veils are defined as irregular structureless areas of confluent blue pigmentation with an overlying white "ground-glass" appearance [4]. These structures allow semi-automatic discrimination of melanoma from benign with a sensitivity of 78.2% and a specificity of 89.97% [5], and are one of the positive criteria within the 7-point checklist for suspicious melanocytic skin lesions [2]. A blue-black color rule has been suggested as an indicator of pigmented nodular melanoma, yielding a sensitivity of 78.2% [6]. In this research, 195 melanoma images were examined and it was found that 22 had some blue coloration. This contrasts with earlier studies that recently showed a few lesions presented in a clinical setting (Dr. Stoecker; Rolla, MO) had fully developed blue veils or blue-black areas. In this research, chosen for analysis were three shades of blue: light blue, dark blue, and lavender blue. The arrows in Figure 1.2 point to the areas with different shades of blue.

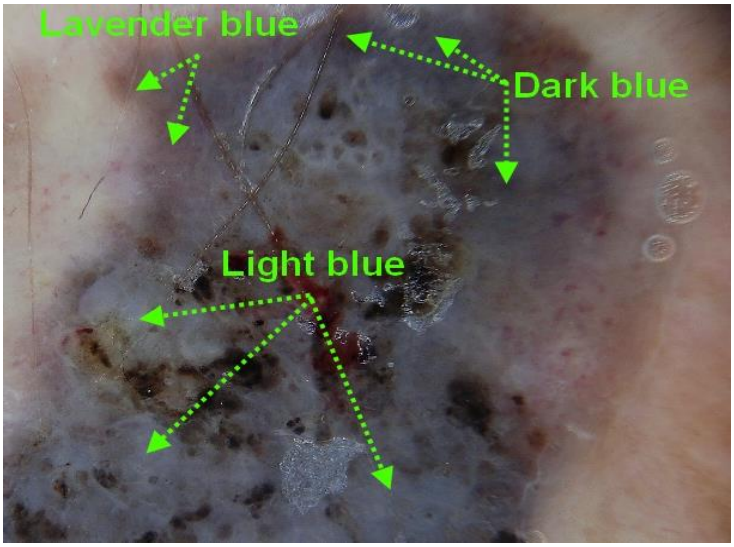


Figure 1.2. Dermoscopy skin lesion with three shades of blue

Melanoma classification has been studied for different shades of blue using different color features in RGB color space [7-10]. Chromaticity is found to be a very important feature for classification of blue areas. To equalize the changes occurring due to different skin types and the lighting in the photography, a relative intensity feature is used in which the average background skin color is subtracted from each pixel of the skin lesion. Other features like 'value' from HSV color space and the ratio of blue: green features are used to restrict areas other than blue areas.

The basic approach for blue area segmentation is to place standard thresholds on the RGB dermoscopy images to segment blue areas using color features and then to explore the importance and restrictiveness of these color parameters by constructing fuzzy sets on these parameters. The image processing algorithm for segmenting the blue areas for different shades of blue is presented. Color features, contour features, and texture features are computed from the blue area mask. Malignant melanoma versus benign lesion discrimination is performed to evaluate discrimination capability for the different shades of blue. The dermoscopy image set examined for lesion discrimination in this study consists of skin lesions with and without blue areas. A support vector machine classifier and logistic regression model for statistical analysis are used for melanoma discrimination analysis.

The other sections of this thesis include 1) Methodology including image set description and pre-processing, feature characterization and feature extraction, 2) Results and 3) Conclusion and future scope.

2. METHODOLOGY

Figure 2.1 illustrates an overview of the image analysis process for determining parameters for blue shade segmentation using a training set of 22 melanoma dermoscopy skin lesion images. It also illustrates an overview of the image analysis and lesion discrimination process for a test set of 866 dermoscopy images (693 benign images and 173 melanoma images). The overall blue area detection process is performed in two stages. The first stage consists of the manual blue area markings, determination of absolute threshold parameters and fuzzy set implementation. This is examined on a subset of 22 melanoma images. The next stage consists of mask generation, feature extraction and classification. These are performed on a set of 866 images which includes malignant and benign images. The following sections present the different steps of the skin lesion image analysis and classification process.

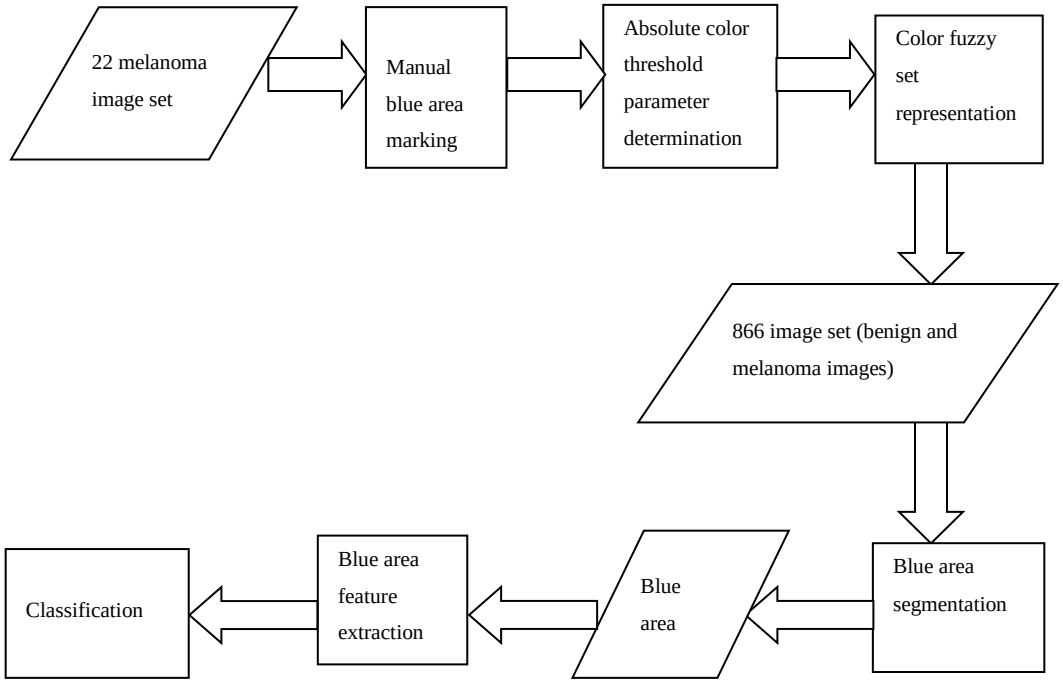


Figure 2.1. Blue area analysis approach

IMAGE SET DESCRIPTION AND PRE-PROCESSING

2.1.1. Description of Experimental Data Set A dataset of 888 contact, non-polarized dermoscopic images (195 melanoma and 693 benign), were acquired in the course of NIH CA153927-02A2 study. A subset of 22 melanoma images with significant blue areas was selected from the 195 melanoma images dataset for initial study. This set with 22 images was considered a training set to find the absolute thresholds. The data set of 866 images with 693 benign images and 173 melanoma images (195 melanoma images – 22 training set images) was used for both feature extraction and classification.

Contact non-polarized dermoscopy images were taken with a 3Gen DermLite Fluid Dermatoscope (3Gen Inc., Dana Point, CA) and with a Sony DSC-W70 7.2 megapixel digital camera with dermoscopic adapter. This research was approved according to the Helsinki Declaration by Phelps County Regional Medical Center Institutional Review Board, Rolla, Missouri USA. Informed consent was obtained from all participants.

2.1.2. Manual Blue Area Marking The initial target was to determine all possible blue areas, irrespective of the shades. Thus under Dr. William V. Stoecker's (Rolla, MO) supervision, three uniform-sized ellipses with blue areas were marked manually in each lesion of the 22 melanoma image data set. These areas were marked with a tool called winshow. With this tool, points could be marked around the area of interest. This tool automatically generates masks with the area of interest being white and all the other parts of the images being black. These three, small, uniform, elliptical areas (representing the widest range of blue areas) were used for both pre-processing and threshold determination. It was then determined that the blue areas on skin lesions can be classified into three shades: light blue, dark blue and lavender blue (see Figure 2.2). An attempt was made again to mark three ellipses with each elliptical area covering a different shade of blue.

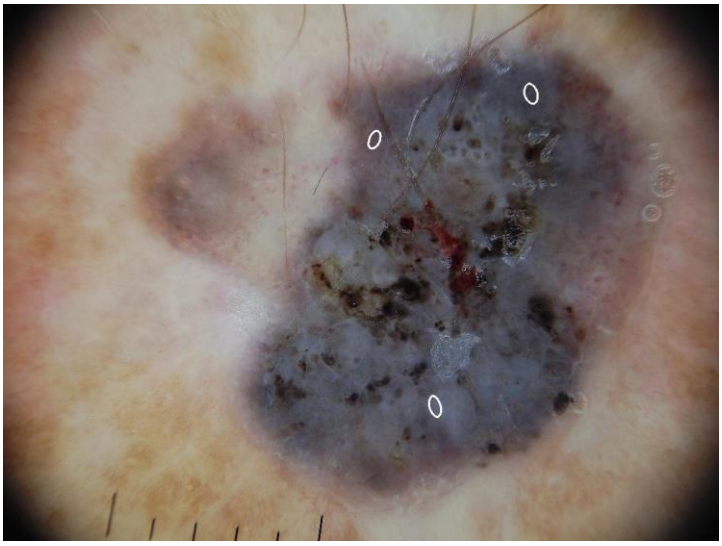


Figure 2.2. Manually marked blue areas

FEATURE CHARACTERIZATION

2.1.3. Absolute Thresholds Manually marked elliptical areas were examined to identify different shades of blue in the lesion. Some color features, such as average R, G, B values, maximum and minimum R, G, B values, relative intensity, and chromaticity, were examined on the manually marked elliptical areas. The range of variation for chromaticity was found to be far less when compared to any other feature. Chromaticity was thus selected as a primary criterion to classify the blue areas.

The relative intensity approach was used to equalize the color changes due to different skin types and processing techniques. For the further classification of blue areas to different shades, brightness and blue: green ratio was used.

2.2.1.1. Chromaticity The chromaticity for any color in an RGB model can be given by the following equation:

$$ChrX = X/(R+G+B);\ X \in \{R, G, B\} \quad (1)$$

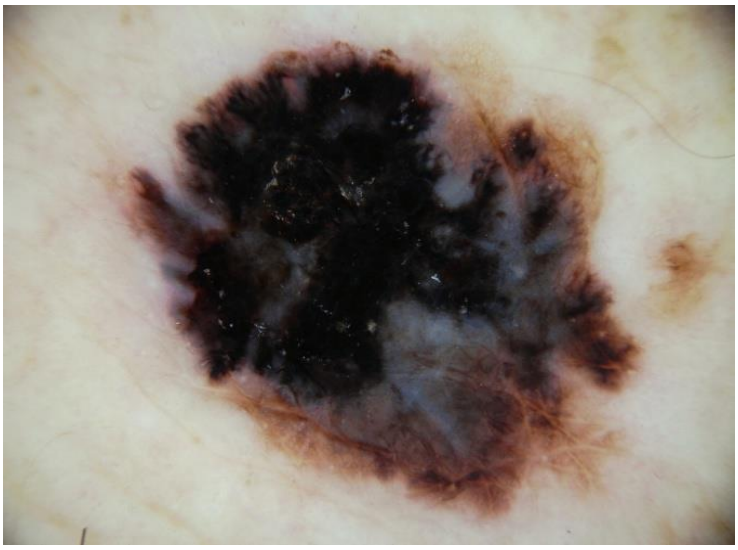
The three chromaticity components for a pixel are chrR, chrB, and chrG. Upper and lower thresholds were initially determined from a training set of 22 melanoma images for the chromaticity of red and blue colors. A single threshold was determined from a training set of 22 melanoma images for chromaticity of green color. Let the thresholds be represented as $chrR_{opt,}$ $ChrG_{opt}$, and $ChrB_{opt}$. The conditions the pixels must satisfy to be considered as blue area will be the following:

$$ChrR_{opt1} < ChrR < ChrR_{opt2} \quad (2)$$

$$ChrG < ChrG_{opt2} \quad (3)$$

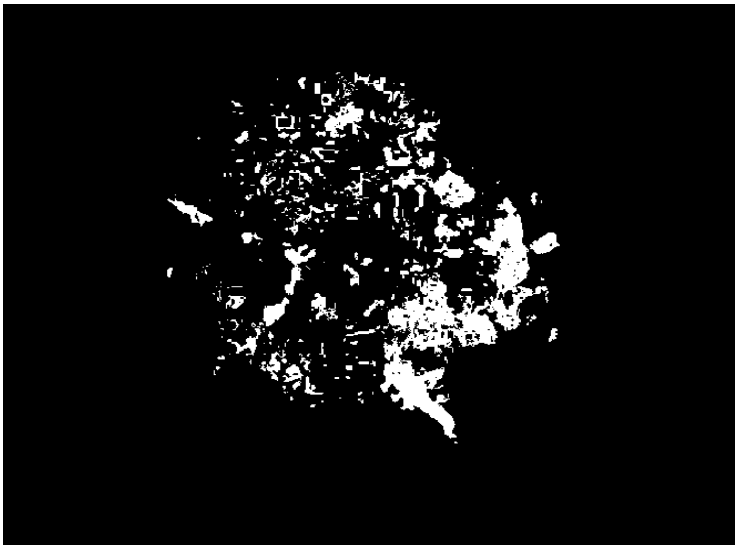
$$ChrB_{opt1} < ChrB < ChrB_{opt2} \quad (4)$$

Opt1 and opt2 refer to the thresholds set empirically from the 22 melanoma image training set. Figure 2.3 (a) is a training set melanoma image example. Figure 2.3 (b) is the corresponding blue area mask obtained by applying the thresholds (described above). The mask illustrates blue areas determined as white pixels and all the other pixels being black when chromaticity thresholds were placed on the skin lesion image.



(a) Melanoma image with blue areas

Figure 2.3. Blue area segmentation based on chromaticity thresholds



(b) Mask with segmented blue areas

Figure 2.3. Blue area segmentation based on chromaticity thresholds (cont.)

2.2.1.2. Relative Intensity The relative intensity was calculated from the difference in actual pixel value inside the skin lesion and the average value of some surrounding skin pixels. The surrounding skin region was an area equal to 20% of the lesion area just outside the border plus a 10-pixel wide safety margin. The value of the relative intensity lies in between -255 and 255. This neutralizes any problem caused by the variations in the lighting and processing techniques.

2.2.1.3. Brightness For the red, green, and blue chromaticity threshold rules examined, lavender blue and combined regions of light blue and dark blue could be separated, but could not further separate light and dark blue. Thus, HSV (hue, saturation and value) plane is determined from RGB image by a matlab command 'rgb2hsv'. An optimal brightness threshold V_{opt}, with optimized weights based on the HSV plane, was used for all of the light and dark blue shaded areas

$$V_{opt} = 0.5\ V_m + 0.224 \qquad (5)$$

where V_m corresponds to the median brightness from combined light and dark blue areas (determined from each skin lesion). Value 0.224 corresponds to 50% of the average median found from the manually marked elliptical areas in the training set.

2.1.4. Color Fuzzy Set Specification To explore the importance and restrictiveness of these color parameters for blue area determination, fuzzy set representations were examined to characterize each color parameter [11]. A fuzzy set was specified to provide high membership values to pixels within the skin lesion image that perceptually contributed to identifying moderate blue areas. This type of specification gives degrees of membership to pixels perceived to be representative of either lighter or darker blue areas. From a dermatologic perspective, providing some degree of membership to lesion pixels that are perceived as either lighter or darker blue permits color variation between lesion images as well as variation in types of blue areas detected.

Trapezoidal membership functions were used for both the red and blue chromaticity thresholds that had both upper and lower color thresholds. The membership value of 'one' could be given to all the values between absolute thresholds and other degree of memberships to the ramp on both sides of absolute thresholds. Z-shaped and S-shaped membership functions were given to green chromaticity and relative intensity, respectively. These two parameters had a single absolute threshold value. Thus, a membership value of one could be given to all the values specified by the single threshold and a degree of membership could be given to the ramp.

The trapezoidal membership function (see figure 2.4.) for red chromaticity $\mu_{red\_chrom}(x)$, which depends on four parameters, is given as

$$\mu_{red\_chrom}(x) = \begin{cases} 0, x \le a \\ \dfrac{x-a}{b-a}, a \le x \le b \\ 1, b \le x \le c \\ \dfrac{d-x}{d-c}, c \le x \le d \\ 0, d \le x \end{cases} \quad (6)$$

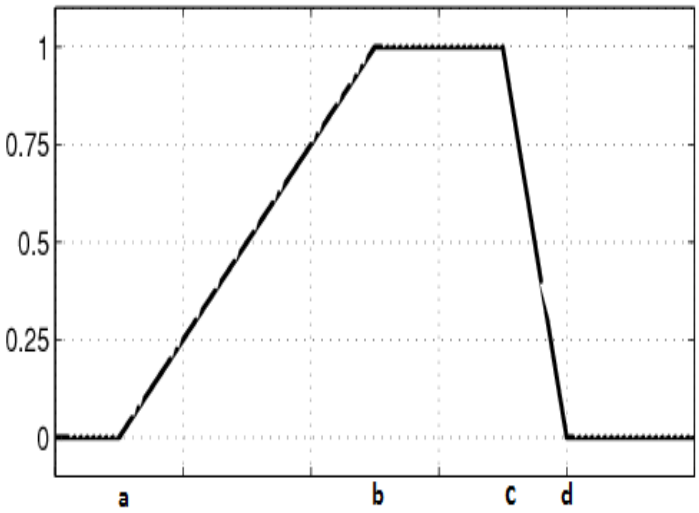


Figure 2.4. Trapezoidal Membership function

Let RX be the set of red chromaticity values for three different shades, with X ∈ {light blue, dark blue and lavender blue}. Fuzzy sets were chosen based on the observations on the training set. Different thresholds were placed on all the training set images and the results were examined. The empirical red chromaticity sets: R_lightblue ranges from 0.21 to 0.4, R_darkblue also ranges between 0.21 and 0.4 and R_lavenderblue ranges from 0.33 to 0.5. The fuzzy set is the pair (RX, μ_{red_chrom}), where $\mu_{red\_chrom} : RX \rightarrow [0,1]$. All of the values in the RX sets have a certain degree of membership.

Four parameters (a, b, c, and d) of the trapezoidal function listed in Table 2.1 determine the degree of membership. The parameter ‘a’ is empirically determined minimum chromaticity and ‘d’ is empirically determined maximum chromaticity of the all the pixels in manually marked areas. The range between ‘b’ and ‘c’ is empirically determined range in which more than 60% of the pixels from blue areas in training images are present. All of the values between the parameters ‘b’ and ‘c’ are fully included in the fuzzy set and given a value of $\mu_{red\_chrom}(x) = 1$. Any value either less than ‘a’ or greater than‘d’ were not included in the fuzzy set and given a value of

$\mu_{red\_chrom}(x) = 0$. Values between 'a' and 'b' have a membership value that depends on the alpha cuts. Similarly, the membership values between 'c' and 'd' depend on the alpha cuts.

Fuzzy sets were also constructed based on the trapezoidal membership functions for blue chromaticity thresholds of lavender blue and blue chromaticity thresholds of combined light and dark blue. This membership function for blue chromaticity $\mu_{blue\_chrom}(x)$ is given as

$$\mu_{blue\_chrom}(x) = \begin{Bmatrix} 0, x \le a \\ \dfrac{x-a}{b-a}, a \le x \le b \\ 1, b \le x \le c \\ \dfrac{d-x}{d-c}, c \le x \le d \\ 0, d \le x \end{Bmatrix} \quad (7)$$

The fuzzy sets for blue chromaticity can be given with the pair (BX, μ_{blue_chrom}), where $\mu_{blue\_chrom} : BX \rightarrow [0,1]$ and BX are the sets for blue chromaticity values of different shades with X ∈ {light blue, dark blue and lavender blue}. The parameter 'a' is empirically determined minimum chromaticity and 'd' is empirically determined maximum chromaticity of the all the pixels in manually marked areas. The range between 'b' and 'c' is empirically determined range in which more than 60% of the pixels from blue areas in training images are present. For light blue, the empirical blue chromaticity set was between 0.24 and 0.49. The same range, 0.24 to 0.49 was for dark blue and the fuzzy set range for lavender blue was between 0.24 and 0.42. Four parameters of the trapezoidal membership functions for blue chromaticity of different shades are listed in Table 2.1.

The empirical fuzzy set for green chromaticity is given as the fuzzy set pair (GX, μ_{green_chrom}), where $\mu_{green\_chrom} : GX \rightarrow [0,1]$ and GX is the green chromaticity set ranging from 0 to 0.38 for lavender blue. A Z-shaped membership function (see figure 2.5.) denoted as $\mu_{green\_chrom}(x)$ is defined by two parameters and is given as

$$\mu_{green\_chrom}(x) = \begin{cases} 1, x \le a \\ 1 - 2\left(\frac{x-a}{b-a}\right)^2, a \le x \le \frac{a+b}{2} \\ 2\left(\frac{x-b}{b-a}\right)^2, \frac{a+b}{2} \le x \le b \\ 0, x \ge b \end{cases} \quad (8)$$

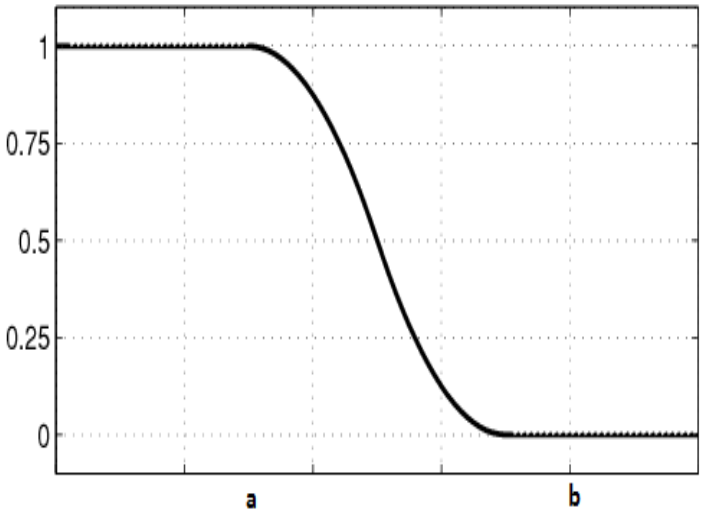


Figure 2.5. Z-shaped Membership function

where ‘a’ locates the left extreme of the sloped portion of the curve, and ‘b’ locates the right extreme of the curve. Parameter ‘a’ is empirically determined value below which more than 60% of the pixels in manually marked areas from training set are present. Parameter ‘b’ is the maximum green chromaticity value of all the pixels from blue areas in training set.

The final fuzzy sets for relative intensity are defined using S-shaped membership functions. Let RelativeX denote the set with all possible relative intensity values ranging from -255 to 255. Then, the fuzzy set is the pair (RelativeX, $\mu_{\text{rel\_intensity}}$) where

$\mu_{\text{rel\_intensity}} : \text{Relative} X \rightarrow [0,1]$. The S-shaped membership function (see figure 2.6.) $\mu_{\text{rel\_intensity}}(x)$ is given as

$$\mu_{\text{rel\_intensity}}(x) = \begin{cases} 0, x \le a \\ 2\left(\frac{x-a}{b-a}\right)^2, a \le x \le \frac{a+b}{2} \\ 1-2\left(\frac{x-b}{b-a}\right)^2, \frac{a+b}{2} \le x \le b \\ 1, x \ge b \end{cases} \quad (9)$$

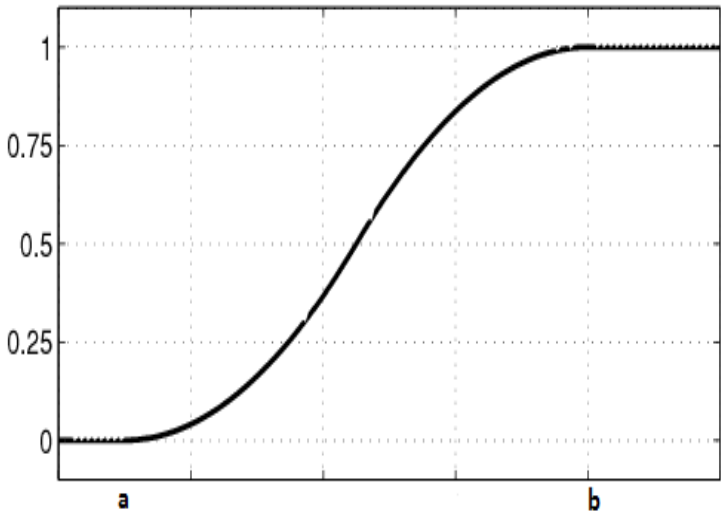


Figure 2.6. S-shaped Membership function

where '*a'* and '*b'* define the left and right extremes of the sloped portion of the curve respectively. Parameter 'b' is the empirically determined relative intensity value above which more than 60% of all the pixels in blue areas from training set are present. Parameter 'a' is the minimum relative intensity value of all the pixels in manually marked blue areas. For the relative intensity sets of light blue, dark blue, and lavender blue, all the relative intensity values above 'b' are fully included in the fuzzy set and the relative intensity values below 'a' are not included in the fuzzy set. The relative intensity values

between ‘a’ and ‘b’ have membership values depending on the alpha cut. Table 2.1 lists the parameters (a, b, c, and d) used for different fuzzy sets to segment the blue areas determined from the training set of melanoma images.

Table 2.1 Corner points of membership functions for color parameters

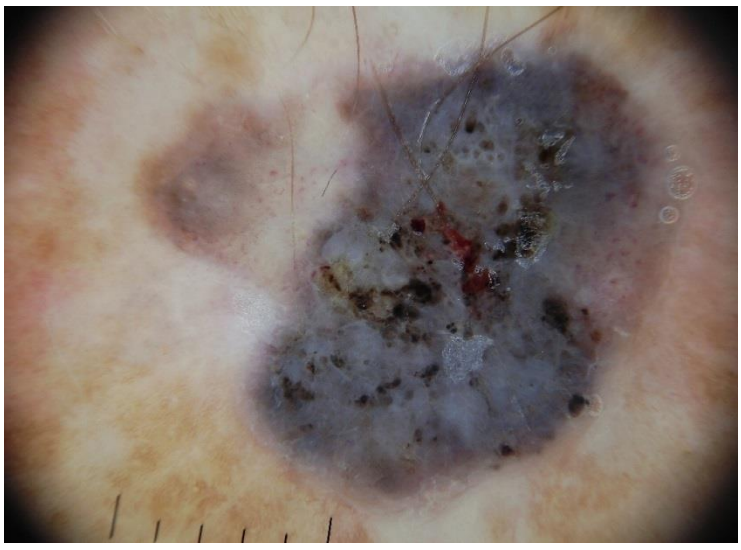
	Membership Function	a	b	c	d
Red chromaticity (light blue)	$\mu_{red\_chrom}(x)$	**0.21**	**0.27**	**0.34**	**0.4**
Red chromaticity (dark blue)	$\mu_{red\_chrom}(x)$	**0.21**	**0.27**	**0.34**	**0.4**
Red chromaticity (lavender blue)	$\mu_{red\_chrom}(x)$	**0.33**	**0.34**	**0.38**	**0.5**
Blue chromaticity (light blue)	$\mu_{blue\_chrom}(x)$	**0.24**	**0.3**	**0.43**	**0.49**
Blue chromaticity (dark blue)	$\mu_{blue\_chrom}(x)$	**0.24**	**0.3**	**0.43**	**0.49**
Blue chromaticity (lavender blue)	$\mu_{blue\_chrom}(x)$	**0.24**	**0.32**	**0.38**	**0.42**
Green chromaticity (lavender blue)	$\mu_{green\_chrom}(x)$	**0.32**	**0.38**	**-**	**-**
Relative intensity (light blue)	$\mu_{\mathrm{rel\_intensity}}(x)$	**0**	**10**	**-**	**-**
Relative intensity (dark blue)	$\mu_{\mathrm{rel\_intensity}}(x)$	**0**	**10**	**-**	**-**
Relative intensity (lavender blue)	$\mu_{\mathrm{rel\_intensity}}(x)$	**15**	**30**	**-**	**-**

2.1.5. Blue Shade Segmentation Blue areas were segmented according to applied alpha cuts for different combinations of color fuzzy sets, whereby, a pixel is included in the segmented blue area mask if the pixel satisfies the alpha cut membership constraints for all of the color fuzzy sets. An alpha cut is a set of values whose membership value for the fuzzy set is equal to or greater than the specified membership cut off value [12].

For a specific alpha cut, the membership values of the red chromaticity set, the blue chromaticity set and the relative intensity set of dark blue determines the masked

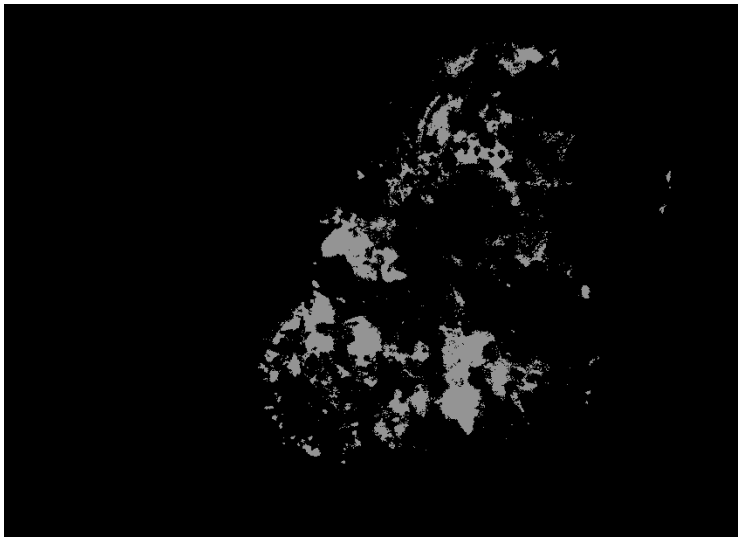
regions based on the alpha cut for dark blue. These masks are logically ANDed along with the mask determined from brightness threshold to get the final dark blue area mask. For the same alpha cut, the membership values of the red chromaticity set, the blue chromaticity set and the relative intensity set of light blue determines the masked regions and these masks are logically ANDed along with brightness to get the final light blue mask. Similarly, the membership values of the red chromaticity set, the blue chromaticity set, the green chromaticity set and relative intensity set of lavender blue determines the masked regions for the same alpha cut that are logically ANDed to find the final lavender blue mask.

Alpha cuts with values between 0.25 and 0.75 were observed in increments of 0.05 on the 866-image set for the classification. Fig 2.7 illustrates the three shades of blue determined for an alpha cut of one.

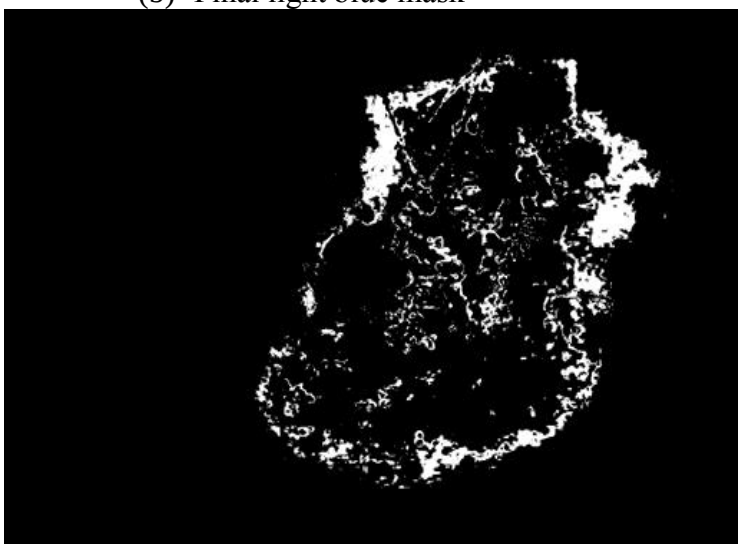


(a) Dermoscopy melanoma image

Figure 2.7. Final blue area masks for three shades of blue

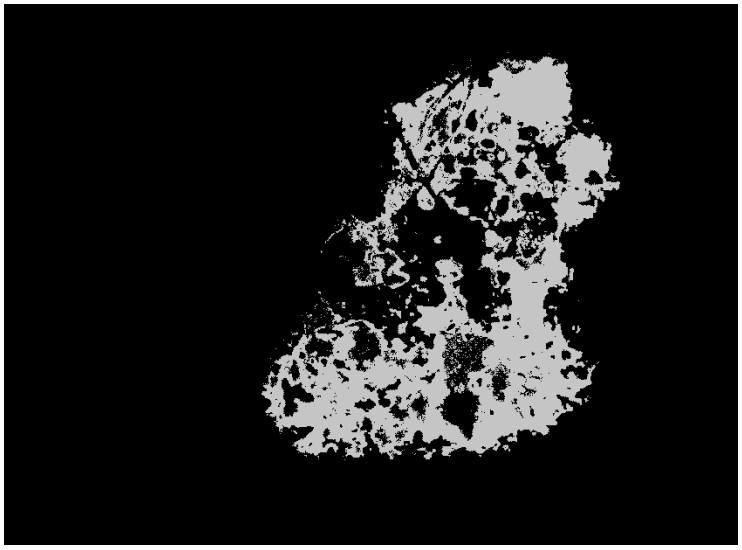


(b) Final light blue mask



(c) Final lavender blue mask

Figure 2.7. Final blue area masks for three shades of blue (cont.)



(d) Final dark blue mask

Figure 2.7. Final blue area masks for three shades of blue (cont.)

2.2. FEATURE EXTRACTION

Different alpha-cut combinations were investigated to generate segmented blue area regions for both feature and lesion discrimination analysis. Various contour features, color features, and texture features (listed in the following section) were extracted from the final dark blue masks, final light blue masks and final lavender blue masks at different alpha cuts for the analysis.

2.2.1. Contour Features The following contour features were computed using the blue area mask and the original color lesion image:

- Eccentricity of the largest blob (E): Returns the eccentricity value by calculating the Euclidean distance (D) between the centroid of the largest blob and the centroid of the lesion. It then divides it by the square root of the lesion area (A):

$$E = \frac{D}{\sqrt{A}} \tag{10}$$

- Relative size of all blobs (R): Measures the relative size by dividing the sum of all blob areas (B_i) by the area of the lesion (A):

$$R = \frac{\sum_{i=1}^{n} B_i}{A} \quad (11)$$

where n is the number of blobs in the lesion.

- Relative size of largest blob (S): Ratio of the largest blob area (B_{max}) within both the lesion and the area of the lesion (A):

$$S = \frac{B_{max}}{A} \quad (12)$$

- Absolute size of largest blob: Gives the area of the largest blob from the blue area within the lesion.
- Number of blobs in the lesion

2.2.2. Color Features The following color features were computed using the blue area mask and the original color lesion image:

- Blue areas by quintiles: Measures the blue areas present within each lesion quintile created by the Euclidean distance transform [13].
- Average RGB values: The average Red, green and blue values are determined for all the blue areas.
- Maximum and minimum values: The maximum red, green and blue values are determined in the masked blue areas. Similarly, minimum red, green and blue values are determined in the masked blue areas.
- Standard Deviation: Measures the variation of red, green and blue values in blue areas
- Chromaticity: Measures the red, green and blue chromaticity values for all the blue areas. ChrX = X/(R+G+B); X ∈ {R, G, B}

2.2.3. Texture Features The following texture features were computed using both the blue area mask and the original color lesion image for the three planes of RGB image and also for grayscale image. Let (r_i, g_i, b_i) be the RGB values and $gray_i$ be the grayscale value for histogram bin 'i', where separate histograms are determined for the R, G, B planes and grayscale in the masked color image region:

- Mean (μ_R, μ_G, μ_B, μ_{Gray}): Measured for grayscale and RGB planes and calculated as

$$\mu_R = \sum_{l=0}^{L-1} r_i \, p(r_i) \quad (13)$$

where $p(r_i)$ is the probability of the occurrence of r_i. L is the number of possible red levels. μ_G , μ_B and μ_{Gray} are similarly defined.

- Variance ($\sigma^2_{R,}$ $\sigma^2_{G,}$ $\sigma^2_{B,}$ σ^2_{Gray}): Measures the average contrast of blue areas in all of the three planes of RGB (red plane, green plane, and blue plane), as well as in grayscale.

$$\sigma^2{}_R = \sum_{l=0}^{L-1} (r_i - \mu)^2 \, p(r_i) \tag{14}$$

σ^2_G , σ^2_B and σ^2_{Gray} are similarly defined.

- Smoothness (S): Measures the smoothness of the blue areas for RGB planes and grayscale; its value is near zero for very smooth areas and near one for unsmooth areas:

$$S = 1 - \frac{1}{(1+\sigma^2)} \tag{15}$$

- Skewness (κ): Provides the skewness of blue area histograms for RGB planes and grayscale. It returns a value of zero for symmetric histograms about the mean, a positive value for right-skewed histograms; and a negative value for left-skewed histograms. The following equation returns skewness for red plane:

$$\kappa = \sum_{l=0}^{L-1} (r_i - \mu)^3 \, p(r_i) \tag{16}$$

Skewness for green plane, blue plane and grayscale are similarly defined.

- Uniformity (U): Returns a maximum value when either all gray levels or all red, green, and blue values are equal in all of the blue areas:

$$U = \sum_{l=0}^{L-1} p^2(r_i) \tag{17}$$

- Entropy (*e)*: Measures the randomness or uncertainty of values over the blue areas for RGB planes and grayscale. The following equation measures entropy for red plane. Entropy for other planes is defined similarly:

$$e = -\sum_{l=0}^{L-1} p(r_i) \log_2 p(r_i) \tag{18}$$

2.3. CLASSIFICATION

Different combinations of color parameters are used to determine final dark blue area masks, final light blue area masks and final lavender blue area masks. For every final blue mask, contour, color, and texture features are computed. These features are provided as inputs to classification algorithms for melanoma discrimination. Melanoma discrimination was performed on a test set of 866 skin lesions (see Section 2.1.1). Platt's sequential minimal optimization algorithm was implemented to train a support vector machine (SVM) classifier [14]. For the 866 image set, a ten-fold cross-validation approach was used to generate training and testing set using 195 melanomas and 693 benign lesions. This approach randomizes the data initially and then generates the fold. Every run in the ten-fold cross-validation received a new, but defined seed value. The classification results were obtained for different alpha cuts for all the three shades of blue.

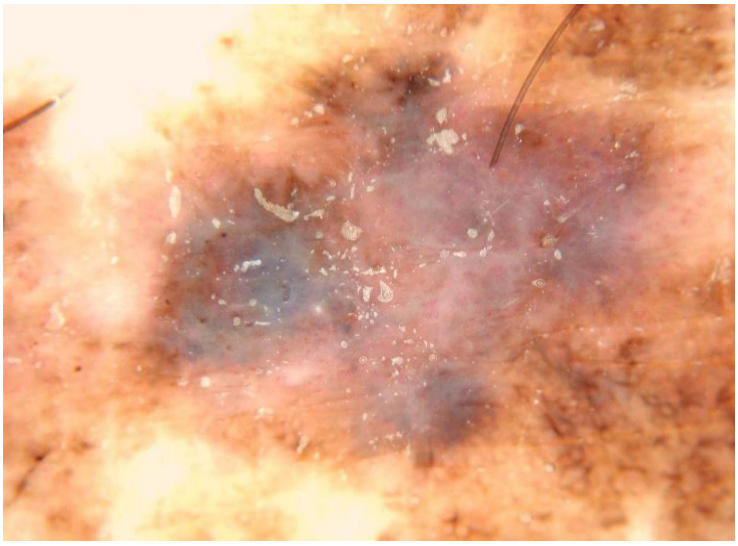
3. RESULTS

Experiments were performed for different alpha cuts on the 866 image data set and diagnostic accuracy was measured. Table 3.1 lists the diagnostic results obtained for different alpha cuts using a support vector machine (SVM) classifier. Alpha cuts ranging from 0.25 to 0.75 were observed in increments of 0.05 and the alpha cut of 1 in the table represents the absolute threshold. Classification results were obtained by using the SVM classifier on the features extracted from final dark blue masks, features extracted from final light blue masks and features extracted from final lavender blue masks at different alpha cuts.

Table 3.1. SVM Results based on different alpha cuts for three shades of blue

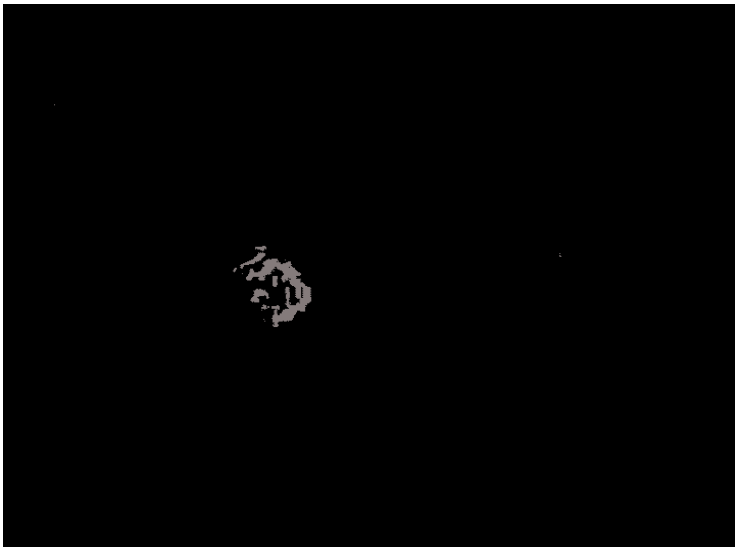
Alpha cuts applied to color features	Light blue	Dark Blue	Lavender Blue
0.25	80.14	80.02	80.02
0.30	80.02	80.02	80.02
0.35	80.02	80.02	80.02
0.4	80.02	79.90	80.02
0.45	79.90	79.90	80.02
0.5	80.02	79.79	80.02
0.55	79.90	79.79	80.02
0.6	80.02	79.79	80.02
0.65	80.14	79.90	80.02
0.7	80.14	79.90	80.02
0.75	80.02	80.02	79.90
1	79.90	80.14	79.90

The classification results for different alpha cuts seem to be very close and the peak accuracy is at different alpha cuts for the three shades. Different combinations and different color parameters have been used for determining these three shades which resulted in the variation of alpha cut for peak accuracy. As the alpha cuts were decremented from 0.75 to 0.25, the blue areas determined were expanded around the areas detected by the higher alpha cut, but in very few cases new areas were detected in the images where nothing was found at the previous high alpha cut which might be the reason for close accuracy results. Figure 3.1 shows an example of applying different alpha cuts for determining lavender blue areas. It can be observed from the figures that, as the alpha cut value is decremented, the restrictiveness on the thresholds is relaxed and more area is determined.

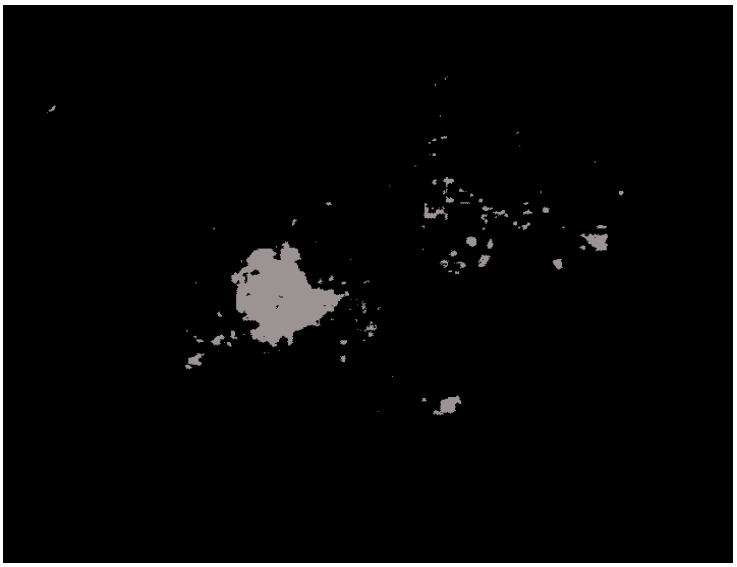


(a) Dermoscopy image with blue area

Figure 3.1. Final lavender blue mask with different alpha cut values

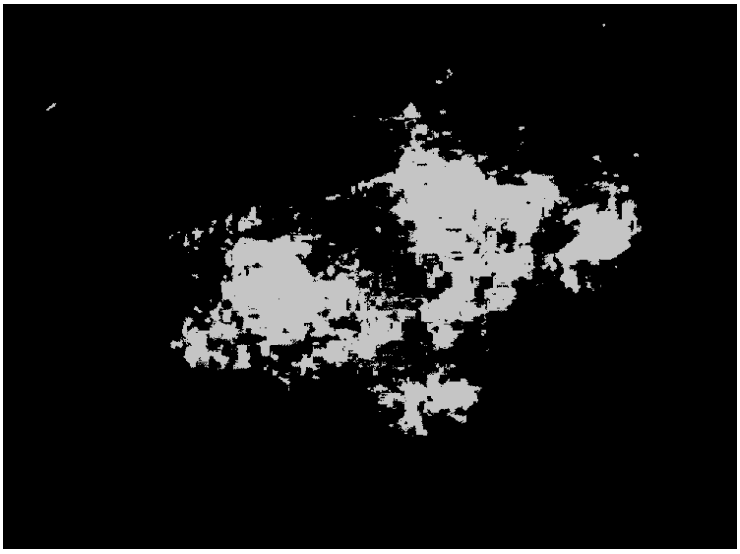


(b) Final lavender blue mask at alpha cut value 1

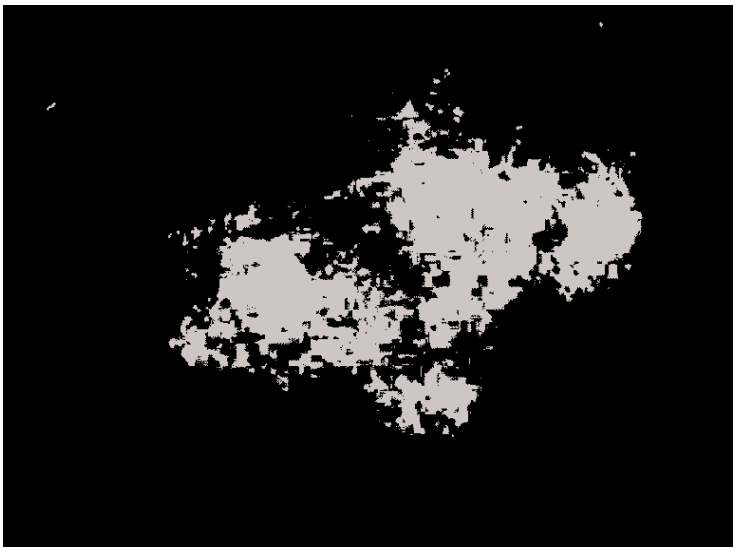


(c) Final lavender blue mask at alpha cut value 0.75

Figure 3.1. Final lavender blue mask with different alpha cut values (cont.)



(d) Final lavender blue mask at alpha cut value 0.5



(e) Final lavender blue mask at alpha cut value 0.25

Figure 3.1. Final lavender blue mask with different alpha cut values (cont.)

The diagnostic accuracy was measured for the 866-image data set. The accuracy for individual shades ranged from 79.79%-80.14%: the peak accuracy of 80.14% for dark blue occurred at an alpha cut of 1; the peak accuracy of 80.02% occurred for lavender blue at many different alpha cuts and the accuracy for light blue is maximized (80.14%) at alpha cut of at 0.25, 0.65, and 0.70.

The features obtained from final dark blue masks, final light blue masks and final lavender blue masks were all combined together to represent all the blue areas in the lesion irrespective of different shades. Diagnostic accuracy was measured at different alpha-cuts for the combined features. The accuracy of the combined features gave a higher accuracy than the individual shades at many alpha cuts. 81.41% accuracy was obtained for an alpha cut of 0.25 using the SVM classifier for all features combined. The logistic regression model of Statistical Analysis Software (SAS Corp., Cary, North Carolina) was also used for analysis which resulted in the diagnostic accuracy of 82.7% at alpha cuts of 0.30, 0.40 and 0.65 for the combined features. The results of SVM and SAS analysis gave similar results and the accuracy based on SAS results were within 3% of the accuracy for SVM at all the alpha cuts. The accuracy results for the combined features at different alpha cuts and logistic regression estimate for area under curve by 'c' statistic are shown in Table 3.2.

Table 3.2. Logistic Regression and SVM results based on different alpha cuts for combined features

Alpha cuts applied to color features	**SVM maximum correct**	**Logistic regression maximum correct**	**Logistic regression estimate for area under curve by 'c' statistic**
0.25	81.4%	82.2%	0.769
0.30	81.2%	82.7%	0.786
0.35	80.7%	82.3%	0.753
0.40	80.6%	82.7%	0.765
0.45	80.6%	82.3%	0.771
0.50	80.4%	82.3%	0.791
0.55	80.0%	82.4%	0.761
0.60	79.9%	81.4%	0.715
0.65	79.8%	82.7%	0.788
0.70	79.7%	81.9%	0.801

Table 3.2. Logistic Regression and SVM results based on different alpha cuts for combined features (cont.)

0.75	79.3%	81.9%	0.773
1.00	80.3%	81.8%	0.68

The receiver operation characteristic (ROC) curves which plot the sensitivity versus one minus specificity were also examined using SAS. These curves are obtained at different alpha cuts from the combined features that represent all the blue areas in the lesion. The ROC curves which depict the results of different alpha cuts ranging from 0.30 to 0.70 with 0.1 increments are shown in Figure 3.2.

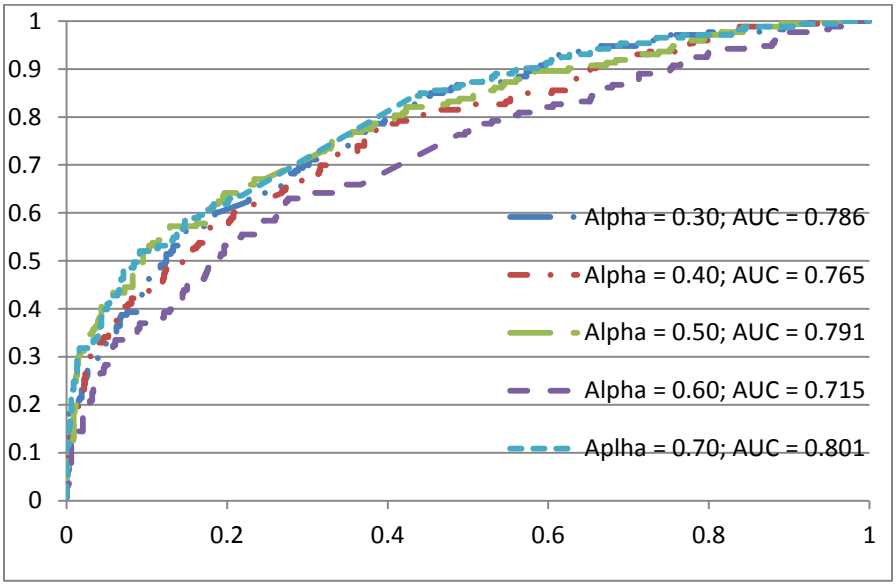


Figure 3.2. Receiver Operating Characteristic Curves at different alpha cuts

4. CONCLUSION AND FUTURE SCOPE

This study analyzes the implementation of fuzzy sets on blue area segmentation for melanoma discrimination from benign lesions. A diagnostic accuracy of 82.7% was obtained at alpha cuts of 0.30, 0.40 and 0.65 from the features derived on 866 image dataset and an accuracy of 81.4% was obtained at 0.25 using the SVM classifier. The diagnostic accuracy was improved for the combined shades with more features over the accuracy of individual shades. Fuzzy logic allowed optimization over alpha cuts, with most of the alpha cuts showing higher AUC than the crisp case.

For future work, additional features can be included to improve the accuracy and the corner points of the membership functions can be varied to further optimize melanoma discrimination.

There are clinical correlates for the three most important features found by logistic regression. Blue itself corresponds to the Tyndall effect, the optical phenomenon of preferential reflection of short wavelengths from melanin present deep within the lesion. Lavender blue is newly described here. We hypothesize that lavender blue represents erythema coinciding with blue, giving a red tinge to the blue area. The erythema represents physiologic expansion of blood vessels, shown by subsurface transillumination to be a marker for skin cancer [15]. The second and third most important features in logistic regression, light blue in the 1^{st} and 2^{nd} quintiles (periphery) of the lesion, correspond to areas of regression in the immunologically active lesion periphery.

Weaknesses in the current study include the relatively small set of lesions analyzed, which limit the power of conclusions drawn here. Future work includes the inclusion of more lesions in the analysis and the use of automatic borders in lesion analysis.

BIBLIOGRAPHY

[1] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin*. 2012; 62(1):10-29.

[2] Argenziano G, Fabbrocini G, Carli P, De Giorgi V, Sammarco E, Delfino M. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis. *Arch Dermatol*. 1998; 134(12): 1563-70.

[3] Argenziano G, Soyer HP, Chimenti S, Talamini R, Corona R, Sera F, et al. Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet. *J Am Acad Dermatol*. 2003; 48(5):679-93.

[4] Argenziano G, Soyer HP, De Giorgi V, et al. *Interactive atlas of dermoscopy*. Milan, Italy: EDRA Medical Publishing & New Media, 2002.

[5] Celebi ME, Iyatomi H, Stoecker WV, Moss RH, Rabinovitz HS, Argenziano G, et al. Automatic detection of blue-white veil and related structures in dermoscopy images. *Comput Med Imaging Graph*. 2008; 32(8): 670-7.

[6] Argenziano G, Longo C, Cameron A, Cavicchini S, Gourhant JY, Lallas A, et al. Blue-black rule: a simple dermoscopic clue to recognize pigmented nodular melanoma. *Br J Dermatol*. 2011; 165(6): 1251-5.

[7] Cheng Y, Swamisai R, Umbaugh SE, Moss RH, Stoecker WV, Teegala S, et al. Skin lesion classification using relative color features. *Skin Res Technol*. 2008; 14(1):53-64.

[8] Stanley RJ, Stoecker WV, Moss RH. A relative color approach to color discrimination for malignant melanoma detection in dermoscopy images. 2007; 13(1):62-72.

[9] Stanley RJ, Stoecker WV, Moss RH, Rabinovitz HS, Cognetta AB Jr, Argenziano G. A basis function feature-based approach for skin lesion discrimination in dermatology dermoscopy images. *Skin Res Technol*. 2008; 14(4): 425-35.

[10] Stoecker WV, Wronkiewiecz M, Chowdhury R, Stanley RJ, Xu J, Bangert A, et al. Detection of granularity in dermoscopy images of malignant melanoma using color and texture features. *Comput Med Imaging Graph*. 2011; 35(2):144-7.

[11] Stanley RJ, Moss RH, Van Stoecker W, Aggarwal C. A fuzzy-based histogram analysis technique for skin lesion discrimination in dermatology clinical images. *Comput Med Imaging Graph*. 2003; 27(5):387-96.

[12] Khan A, Gupta K, Stanley RJ, Stoecker WV, Moss RH, Argenziano G, et al. Fuzzy logic techniques for blotch feature evaluation in dermoscopy images. *Comput Med Imaging Graph*. 2009; 33(1):50-7.

[13] Dalal A, Moss RH, Stanley RJ, Stoecker WV, Gupta K, Calcara DA, et al. Concentric decile segmentation of white and hypopigmented areas in dermoscopy images of skin lesions allows discrimination of malignant melanoma. *Comput Med Imaging Graph*. 2011;35(2): 148-54.

[14] Platt JC. Sequential minimal optimization: A fast algorithm for training support vector machines. Technical Report MSR-TR-98-14, Microsoft Research, 1998.

[15] Terushkin V, Dusza SW, Mullani NA, Duvic M, Zouridakis G, Weinstock M, et al. Transillumination as a means to differentiate melanocytic lesions based on their vascularity. *Arch Dermatol*. 2009; 145(9): 1060-2.

VITA

Mounika Lingala was born on May 29, 1990 in Andhra Pradesh, India. She received her Bachelor of Technology degree in Electronics and Communications Engineering in 2007 from Jawaharlal Nehru Technological University, India. She received her Master's degree in Computer Engineering from Missouri University of Science and Technology in August, 2013.