

Automated Non-invasive Diagnosis of Melanoma Skin Cancer using Dermoscopic Images

Huda Khan^{1,*}, Anushka Yadav¹, Reha Santiago¹, and Sangita Chaudhari¹

¹Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, 400706, India

Abstract. Melanoma skin cancer is one of the deadliest cancers today, the rate of which is rising exponentially. If not detected and treated early, it will most likely spread to other parts of the body. To properly detect melanoma, a skin biopsy is required. This is an invasive technique which is why the need for a diagnosis system that can eradicate the skin biopsy method arises. It is observed that the proposed method is successfully detecting and correctly classifying the malignant and non-malignant skin cancer. Finally, a neural network is used to classify benign and malignant images from the extracted features.

Keywords: Melanoma, non-invasive, skin lesion, neural network.

1 Introduction

Skin cancer is a fatal disease that is life-threatening. Recently, skin cancer has become one of the most lethal forms of cancers found in human beings. Out of all the various types of skin cancers, melanoma cell skin cancers are the most common type of skin cancer and is the most unpredictable. Malignant melanoma has become one of the most dangerous and aggressive type of skin cancer and its occurrence amongst humans has been increasing quickly. Melanoma, also known as malignant melanoma, is a variant of cancer that develops from the cells that contains a pigment known as melanocytes. These melanocytes may develop from a mole and bring about a changes that include increase in size, irregular border, and change in color, breakdown of skin or itchiness. Melanoma is usually caused by over-exposure to Ultra-Violet (UV) radiation from the sun and tanning beds. Due to this, it damages the DNA of the skin cells, and the cells can begin to grow out of control. It is one of the most unpredictable skin cancers and therefore detection of melanoma cancer at the early stages could help in curing it quickly and efficiently. Figure 1 shows a dermoscopic image of a skin melanoma.

Some of the symptoms of skin melanoma cancer is the ABCD rule which is the most common method used for detection of melanomas by doctors where the symptoms taken into consideration are:

- **A - Asymmetry:** The first half of a spot/lesion is not equal to the other half.
- **B - Border:** The edges of the lesion are ragged, irregular, or blurred.
- **C - Color:** The color across the lesion is not homogeneous and may include different shades of white, red, brown, blue, gray, or black.

- **D - Diameter:** The lesion is greater than 6 millimeters. (Albeit not all melanomas adhere to this)



Figure 1: Dermoscopic image of a malignant melanoma

Melanoma of the skin is the 19th most commonly occurring cancer in men and women[1]. There were about 300,000 new cases of melanoma skin cancer in patients in 2018. In 2019, there were 104,350 with active melanomas which resulted in 11,650 deaths worldwide. In 2018, the highest rate of melanoma was observed in Australia followed by New Zealand. In 2018, India had about 26% of cancers reported as melanoma of the skin. Skin melanomas are treated according to the stage and location of the melanoma [2]. Most treatment operations require surgery to remove the affected area, radiation therapy, biological therapy, targeted therapy, immunotherapy, and chemotherapy. However, melanoma skin cancer can be treated easily if identified and diagnosed at an early stage. This makes the early diagnosis of melanoma as a critical part, which can be a challenging task for dermatologists since other skin lesions might have similar physical characteristics. Dermoscopic images are a widely common technique used in capturing melanomas accurately. This is done by performing an in-vivo observation of pig-

*e-mail: huda.gm.khan@gmail.com

mented skin lesions. Earlier, for the detection of malignant melanomas, dermoscopic images had great prospective, but their elucidation requires not only time but is also subjective. As a result, the necessity to construct a system that can help dermatologists in making an accurate decision for the early diagnosis of melanoma has become critical. In this paper, we are proposing an automated non-invasive diagnosis of melanoma skin cancer using dermoscopic images. Skin melanoma is represented with the help of simple yet efficient descriptors and further classified into malignant or nonmalignant melanoma cancer using neural network.

2 Literature Survey

Global warming has increased the intensity of solar radiation which has led to rise of skin melanoma cancer in human beings [2]. Melanoma, can be cured if it is detected in the early stages. Traditional approach to melanoma skin cancer detection requires biopsy- an invasive technique which can be a painful, costly and tardy procedure. Therefore, the need for an automated procedure to detect skin melanoma cancer accurately is a necessity in the medical field. Melanomas can exist in varied shapes, colors and sizes which make it difficult to detect it an early stage which is why it is essential to design a system that takes proper features into account for extraction. Diagnosis of skin melanoma cancer using an automated system includes the following steps:

- Image Acquisition: Input of data image sets is collected.
- Image Pre-processing: Removes unwanted noises and distortions in the input images.
- Image Segmentation: Involves extracting the skin lesion from the input image for further analysis.
- Feature Extraction: Selecting and extracting the relevant features from the skin lesion that is required to correctly detect skin melanoma cancer. The most common method of feature extraction is the ABCD rule where:
 - A - Asymmetry: Asymmetry specifies the axis of symmetry. Benign lesions tend to be symmetric while malignant lesions are asymmetric in shape.
 - B - Border: This is a measure of the irregular shape of the border of a lesion. Benign lesions are almost less ragged and have a smooth looking border while malignant lesions are highly irregular in nature.
 - C - Color: A color score of a lesion is a crucial feature. Benign lesions are more homogeneously colored throughout while malignant lesions could contain various undertones of brown, black, red, white or blue colors and are not uniform in color.
 - D - Diameter: A malignant melanoma tends to grow in size and surpasses its diameter of larger than 6 mm.
- Classification: The system classifies the input image as either, malignant (cancerous) or benign (non- cancerous).

Majumder and Ullah[3] proposed a system that considered more features from the fundamental ABCD rule. Image resizing and contrast adjustment were used for image pre-processing. An averaging filter was applied to the RGB input image to remove any noise like air bubbles or hair. Segmentation of the skin lesion is done by Otsu's automatic thresholding after which a gray scale masking is done to remove any small blobs present in the segmented lesion. The features that are extracted are: Asymmetry (across x and y-axis), Border Irregularity, Color Variegation, Lesion Diameter, Difference between Maximum and Minimum Feret's Diameters. On analysis, it is found that higher difference signifies the lesion is malignant and a lower difference signifies benign melanoma. These extracted features are then used to train the Back propagation Neural Network (BNN) for classification of the skin lesion as malignant or benign melanoma.

Eltayef et al.[5] proposed another system that focused on various image pre-processing and segmentation techniques. The image pre-processing in this system aimed at hair and noise detection and removal techniques. In order to eradicate any air bubbles, a thresholding method was used that used the pixel's intensity values. For the hair removal, first, the system uses a bank of 64 directional filters to detect hair in the image after which the image is filtered with the help of each of the aforementioned directional filters. Consequently, a Gaussian filter is implied by calculating each pixel's local maximum. This method of thresholding is used to differentiate between the hairs and the actual background of image. Upon detecting the hair from the image, their binary masks and gray scale image are multiplied. This system used two-step segmentation: Fuzzy c-means (FCM) and Markov Random Field (MRF).

Premaladha and Ravichandran[6] emphasizes the need for more number of features to be extracted for feature extraction. This system also uses a different image enhancement and segmentation technique. An input dataset of 992 dermoscopy images are used. It is then pre-processed of any noise, air bubbles, scars or hairs on the skin lesion. Contrast Limited Adaptive Histogram Equalization technique (CLAHE) is the preprocessing technique used to get a contrasted image which helps in accurately deriving the features. Median filtering is applied to eliminate any noise and smooth out the ragged edges of the skin lesion. Normalized Otsu's Segmentation is used for segmentation. The features extracted are: Mean, Standard deviation, Variance, Entropy, Contrast, Homogeneity, Energy, Correlation, Area, Perimeter, Diameter, Asymmetry index, Circularity index, Fractal dimension, and Compactness index. Deep learning based neural networks (DLNN) and Hybrid AdaBoost algorithms were used for classification of the skin melanomas.

Jain et al.[7] proposed a system that is so potent that it can take the input image from any camera, e.g. mobile camera. Image processing techniques like gamma correction were used that facilitated image resizing and adjustment of contrast and brightness. Image segmentation was done using Otsu's automatic thresholding, binary masking and edge detection. The features extracted in this system included geometric features of the skin lesion, they are:

Table 1: Various methods used in existing systems

| Author | Pre-Processing | Segmentation | Features Extracted | Classification |
|--------------------------------|--|---|---|--|
| Majumder and Ullah[3] | Image resizing, contrast adjustment and averaging filter to RGB input image. | Otsu's automatic thresholding method. | Asymmetry, Border Irregularity, Color Variegation, Lesion Diameter, Difference between Maximum and Minimum Feret's Diameters. | Backpropagation Neural Network (BNN). |
| Eltayef et al.[5] | Directional Gabor filters for hair removal. | Fuzzy C-means (FCM) and Markov Random Field (MRF). | Asymmetry, Border, Color Variegation, and Diameter. | Not specified. |
| Premaladha and Ravichandran[6] | Contrast Adaptive Histogram Equalization (CLAHE) and Median filtering. | Normalized Otsu's Segmentation. | Mean, Standard deviation, Variance, Entropy, Contrast, Homogeneity, Energy, Correlation, Area, Perimeter, Diameter, Asymmetry index, Circularity index, Fractal dimension, Compactness index. | Deep learning based neural networks (DLNN) and Hybrid AdaBoost algorithms. |
| Jain et al.[7] | Image resizing, contrast and brightness adjustment. | Otsu's segmentation, 3-plane binary masking procedure and edge detection. | Area, Perimeter, Major and Minor Axis Lengths, Circularity Index, Irregularity Index. | Done by ABCD rule. |
| Sheha et al.[8] | Image resizing, from RGB to grey level conversion. | Uses Texture Analysis. | Contrast, Correlation, Cluster Prominence, Dissimilarity, Homogeneity, Difference variance, Difference entropy, and Inverse differences. | Multilayer Perceptron (MLP). |
| Isasi et al.[9] | Pattern recognition algorithms used: Globular, Reticular and Blue-vein. | Globular, Reticular and Blue-vein algorithms. | ABCD features along with features extracted from pattern recognition algorithms. | Done by the ABCD rule. |
| She et al.[10] | High-pass filtering and gradients in the horizontal and vertical directions are applied. | Snake-based edge detection technique. | Asymmetry, Border irregularity, Color variegation, Diameter, and one feature set is extracted from the pattern of the skin lesion. | Principal Component Analysis(PCA). |

Area, Perimeter, Major Axis Length, Minor Axis Length, Circularity Index and Irregularity Indices. A pre-defined set of thresholds using the ABCD rule were set for the classification stage. The output would classify the skin melanoma as normal skin mole or melanoma skin cancer.

Sheha et al.[8] encouraged a system that uses texture analysis that eliminates the segmentation step. 102 dermoscopy images with 51 benign and 51 malignant melanomas were used on which image resizing of 512x512 is done. Additionally, features based on grey level co-occurrence matrix is carried out upon converting RGB image to gray-level. The features extracted are: Contrast, Correlation, Cluster Prominence, Dissimilarity, Homogeneity, Difference variance, Difference entropy, Information measure of correlation, Information measure of correlation, Inverse difference homogenous, Inverse difference normalized, and Inverse difference moment normalized. For classification, Multilayer Perceptron (MLP) is used.

Isasi et al.[9] proposed different pattern recognition algorithms based on the nature of the skin melanomas. Three algorithms- globular, reticulated and blue pigmentation- are proposed as these features are recurrent in malignant melanomas. Therefore, the features extracted are those from the fundamental ABCD rule in addition with the features extracted from the pattern recognition algorithms.

She et al.[10] pivots around designing a system that utilizes the ABCD rule to extract features from the skin lesion which helps improve the system's classification accuracy. High-pass filtering and gradients in the horizontal and vertical direction are applied for the image pre-processing after which a snake-based edge detection technique is used to define the lesion boundary. The features extracted are from the ABCD rule and one from the pattern of the skin lesion. Furthermore, Principal Component Analysis (PCA) is used for classification. Table 1 summarizes the various methods used in existing systems.

Table 2: Comparison of existing method and proposed method for feature extraction

| Method | Features Extracted |
|--------------------------------|--|
| Majumder and Ullah[3] | Asymmetry, Border Irregularity, Color Variegation, Lesion Diameter, Difference between Maximum and Minimum Feret’s Diameters. |
| Eltayef et al.[5] | Asymmetry, Border, Color Variegation, and Diameter |
| Premaladha and Ravichandran[6] | Mean, Standard deviation, Variance, Entropy, Contrast, Homogeneity, Energy, Correlation, Area, Perimeter, Diameter, Asymmetry index, Circularity index, Fractal dimension, Compactness index. |
| Jain et al.[7] | Area, Perimeter, Major and Minor Axis Lengths, Circularity Index, Irregularity Index. |
| Sheha et al.[8] | Contrast, Correlation, Cluster Prominence, Dissimilarity, Homogeneity, Difference variance, Difference entropy, and Inverse differences. |
| Isasi et al.[9] | ABCD features along with features extracted from pattern recognition algorithms. |
| She et al.[10] | Asymmetry, Border irregularity, Color variegation, Diameter, Pattern of skin lesion. |
| Proposed System | Asymmetry along x-axis, Asymmetry along y-axis (Concept of connectivity in pixels used to evaluate principal axes for the asymmetry scores), Area to perimeter ratio, Compactness index, Product of Area and Perimeter, Color score evaluated with bitwise AND operation on the HSV image and color mask, Average of lesion diameter, Difference between the major and minor axis’ lengths, and Lesion diameter. |

3 Proposed System

Table 2 gives the comparison of the existing method and proposed method to extracting features. The proposed methodology comprises of the following primary steps: image acquisition, image pre-processing, image segmentation, feature extraction, and classification of the image. An input image is given to the system is a skin lesion that is to be classified as a benign or malignant melanoma. The input images are collected from PH2 database [11] and The International Skin Imaging Collaboration, ISIC. Removal of noises and hairs from the input image is crucial which is why morphological and blackhat filtering, and inpainting algorithm is implied. To segment the skin lesion from the input image, Otsu’s algorithm [12] and Chan-Vese model [13] is used. This model extracts nine features according to the fundamental ABCD rule and further uses ANN classifier for classification of benign or malignant melanoma.

3.1 Image Acquisition

The PH2 database is a publicly available database of dermoscopy images provided by the Pedro Hispano Hospital. From this database, 200 melanocytic images were collected which consists of 160 benign and 40 malignant lesions. Additionally, 219 benign and 325 malignant images were collected from another publicly available database of dermoscopic images from the The International Skin Imaging Collaboration, ISIC. This totals to a dataset of 744 images.

3.2 Image Pre-processing and segmentation

The input image can contain a lot of noise and other artifacts, like air bubbles, hairs and varied textures of the skin. With all these factors, it is essential that the image

is properly preprocessed and correctly segmented to increase the accuracy of the system. Image preprocessing is done by performing various morphological operations on the grayscale image. Upon applying the morphological filter on the grayscale image, a blackhat filtering operation is used that helps enhance the darkened moles on the skin lesion as the region of interest. Finally, an inpainting algorithm is used wherein a hair mask is created using thresholding that highlights the hair contours after which the original image is inpainted depending on the mask. Inpainting algorithm helps restore the background to the hairs detected in the hair mask. Suppose the hair mask, H , specifies the location of hair contours’ pixels in the input image, I as follows:

$$H(x) = \begin{cases} 0, & \text{if pixel}(x,y) \text{ has hair in } I \\ 1, & \text{if pixel}(x,y) \text{ has no hair in } I \end{cases}$$

On locating the hairs, the pixels are then reconstructed by exploiting the information present in the regions with no hair.

Segmentation is an important step in digital image processing. The process of separating the region of interest from the skin lesion is crucial as this segmented image will further be used for feature extraction. Segmentation is performed in two steps using Otsu’s thresholding method [12] and the Chan-Vese model [13]. Furthermore, morphological operations are performed to remove any more noises in the image. This allows accurate segmentation of the image. Otsu’s method is used to separate the foreground and background pixels into different classes using bi-modal histogram. The algorithm then assesses the threshold that is the most optimum such that the two classes partitioned have their inter-class variance as maximum as possible. Algorithm 1 shows the Otsu’s Thresholding Algorithm.

Algorithm 1: Otsu’s Thresholding Algorithm

- 1 Evaluate histogram and probabilities of each intensity level.
- 2 Set the initial values of $\omega_i(0)$ and $\mu_i(0)$.
- 3 Compute class probabilities of the two classes, ω_i , and their class means, μ_i , as:
- 4 $\omega_i(t) = \sum_0^t P(i)$
- 5 $\mu_i(t) = \sum_0^t P(i)x(i)$
- 6 Step through all possible t thresholds from $t = 1$ till maximum thresholding intensity.
- 7 Update $\omega_i(t)$ and $\mu_i(t)$.
- 8 Evaluate $\sigma_b^2(t)$ as:
- 9 $\sigma_b^2(t) = \omega_2(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]$
- 10 Maximum value of $\sigma_b^2(t)$ is the desired threshold.

The Chan-Vese model uses an active contour model. This contour must stop once it reaches the boundary separating the foreground and background. The binary output image after the Otsu’s method, B , is used as an initial contour to the Chan-Vese model. The model minimizes the energy function E as given in the Equation 1:

$$E(c_1, c_2, B) = \mu.Length(B) + v.Area(inside(B)) + \lambda_1 \int_{inside(B)} |u_0(x, y) - c_1|^2 dx dy + \lambda_2 \int_{outside(B)} |u_0(x, y) - c_2|^2 dx dy \quad (1)$$

The values of μ , v , λ_1 , λ_2 and p are $p = 1$, $v = 0$, and $\lambda_1 = \lambda_2 = 1$ according to the Mumford-Shah approach [14]. c_1 and c_2 are the averages of the given image, u_0 . After the two-step segmentation, a median filter of window size 25 and morphological closing operation is applied on the image. The preprocessed and segmented images of two sample images are shown in Figure 2.

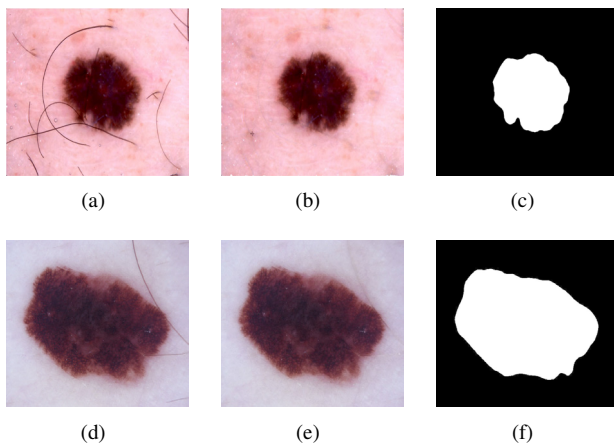


Figure 2: (a), (d) are the original images, (b), (e) are the preprocessed images, and (c), (f) are the segmented images

3.3 Feature Extraction

Once the skin lesion of interest is extracted from the source input image, we have to extract the appropriate and necessary features needed from the image. In this paper, we have heavily emphasized the accuracy of the ABCD rule, therefore, we will continue to use the fundamental rule in addition to more features. Features like Asymmetry and Diameter require principal axes for the evaluation of their scores. In order to evaluate this, the concept of connectivity in pixels in digital space has been used. A pixel p at coordinates (x, y) is connected to every other pixel horizontally, vertically, and diagonally.

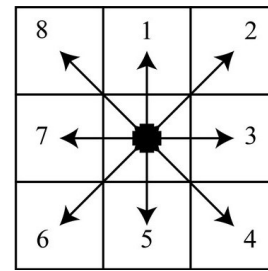


Figure 3: 8-connectivity of a pixel in digital space

As shown in the Figure 3, the pixel at the center p has 8 connected pixels. This means only four axes are possible for this pixel. Upon identifying the centroid of the skin lesion and translating it to the origin of the coordinate system, the length of the four axes passing through the centroid are evaluated and compared. The axis with the greatest length is assigned as the major axis. As principal axes are always perpendicular to each other, the corresponding minor axis of the image. For instance, in the Figure 3, if $7 - 3$ is major axis, then $1 - 5$ is assigned as the minor axis.

3.3.1 Asymmetry

Asymmetry is one of the most important features in order to classify the skin lesion as malignant or benign. It evaluates how similar one half of a region of the lesion is to the other half. A malignant melanoma will tend to be more asymmetric in nature. To calculate the asymmetry, the image obtained upon segmentation (Figure 4(a)) is made to align with the Euclidean coordinate system. The centroid of the image (Figure 4(b)) is translated into the origin of the coordinate system (Figure 4(c)). This image is then rotated with the orientation angle so that the major axis of the blob fits into the x-axis of the image, giving a lesion image, L (Figure 4(d)). This rotated image is then flipped along the x-axis, L_x (Figure 4(e)). The difference between L and L_x will give the non-overlapping region along x-axis, L'_x (Figure 4(f)). Asymmetry along y-axis is carried out in a similar fashion as shown in Figure 4(g), 4(h), and 4(i).

Asymmetry score across x-axis ($A1$) is calculated using Equations (2) and (3). Similarly, asymmetry across y-axis ($A2$) is also calculated using Equations (4) and (5).

$$L'_x = L \oplus L_x \quad (2)$$

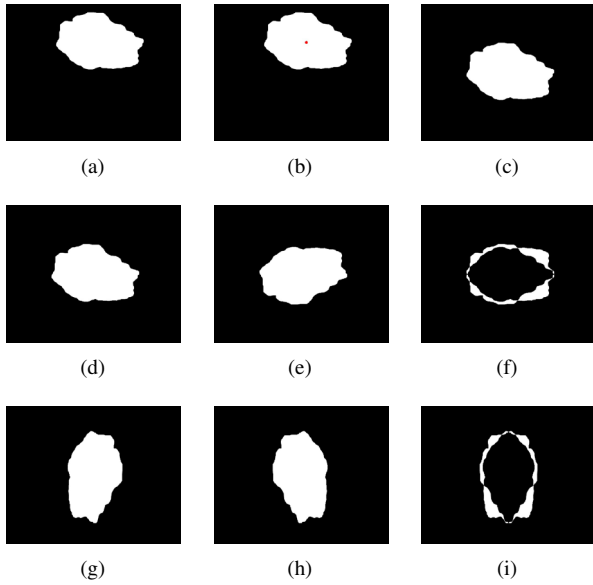


Figure 4: Evaluation of asymmetry across x and y axis

$$A1 = \frac{\text{Area of } L'_x}{\text{Area of } L} \quad (3)$$

$$L'_y = L \oplus L_y \quad (4)$$

$$A2 = \frac{\text{Area of } L'_y}{\text{Area of } L} \quad (5)$$

3.3.2 Border

Border defines the contour of the blob on the skin lesion. A malignant melanoma is highly irregular in nature and can have ragged edges to itself. The border features are $B1$, $B2$, and $B3$ which are Area to perimeter ratio, Compactness index, and Product of Area and Perimeter respectively. $B1$ and $B2$ gives smaller values in case of malignant lesions. $B2$ represents smoothness of the lesion's edges and ranges between 1 and 0. In malignant lesions, the edges are uneven and not smooth thus, giving a $B2$ score of a number closer to 0 whereas a benign lesion will have a $B2$ score of a number closer to 1. Additionally, for malignant melanomas, $B3$ tends to be of greater values. These are calculated by the equations (6), (7), and (8) respectively.

$$B1 = \frac{A}{P} \quad (6)$$

$$B2 = \frac{4\pi A}{P^2} \quad (7)$$

$$B3 = P \times A \quad (8)$$

3.3.3 Color

In determining the color score, six colors are considered important which are: black, dark brown, light brown, blue-gray, red, and white. The segmented image is used as a mask on the original RGB image to get colored segmented image. RGB threshold values for each of the six colors

is initialized and the image is transformed into HSV color space. A mask is created using each corresponding color. Bitwise AND operation is performed on the HSV image in order to determine the color mask and active contours. If the length of active contours is found to be greater than zero then that color is assumed to be present in the lesion. Since there are six colors assumed to be present in the lesion, the color score, C , can range from 1 to 6. Two images having color scores 1 and 3 are shown in the Figure 5(a) and Figure 5(b).

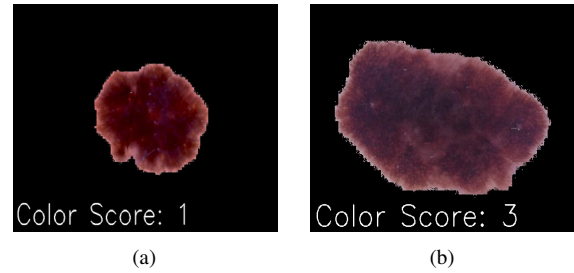


Figure 5: Color scores of two images

3.3.4 Diameter

The diameter of a malignant melanoma may grow in size of over 6mm. The diameter features are $D1$, $D2$, and $D3$ which are Average of lesion diameter, Difference between the major and minor axis' lengths, and Lesion diameter respectively. $D1$ is the average of the diameters Da and Db of the skin lesion. $D2$ is difference of the major axis, D , and minor axis, d . $D3$ is the diameter of the lesion in mm, where M is the magnification factor of the original image and 0.2645 is the factor used to convert the diameter acquired from pixels to mm. These are calculated by the equations (9), (10), and (11) as follows:

$$D1 = \frac{(Da + Db)}{2} \text{ where, } Da = \sqrt{\frac{4A}{\pi}}, Db = \frac{D + d}{2} \quad (9)$$

$$D2 = D - d \quad (10)$$

$$D3 = \frac{Da}{M} \times 0.2645 \quad (11)$$

3.4 Classification

After extracting the features, a classifier is trained that uses the features extracted from the input images. Here, a sequential artificial network with four dense layers has been used. The input layer with nine input dimensions ($A1$, $A2$, $B1$, $B2$, $B3$, C , $D1$, $D2$, and $D3$) and a batch size of 10, and the second and third dense layers having a batch size of 50. The elu activation function has been used in the first two layers and the relu activation function in the third layer. The final output dense layer uses the sigmoid activation function. The classification model has been compiled using the Adam optimizer. The loss of the model has been evaluated using Mean Squared Error function.

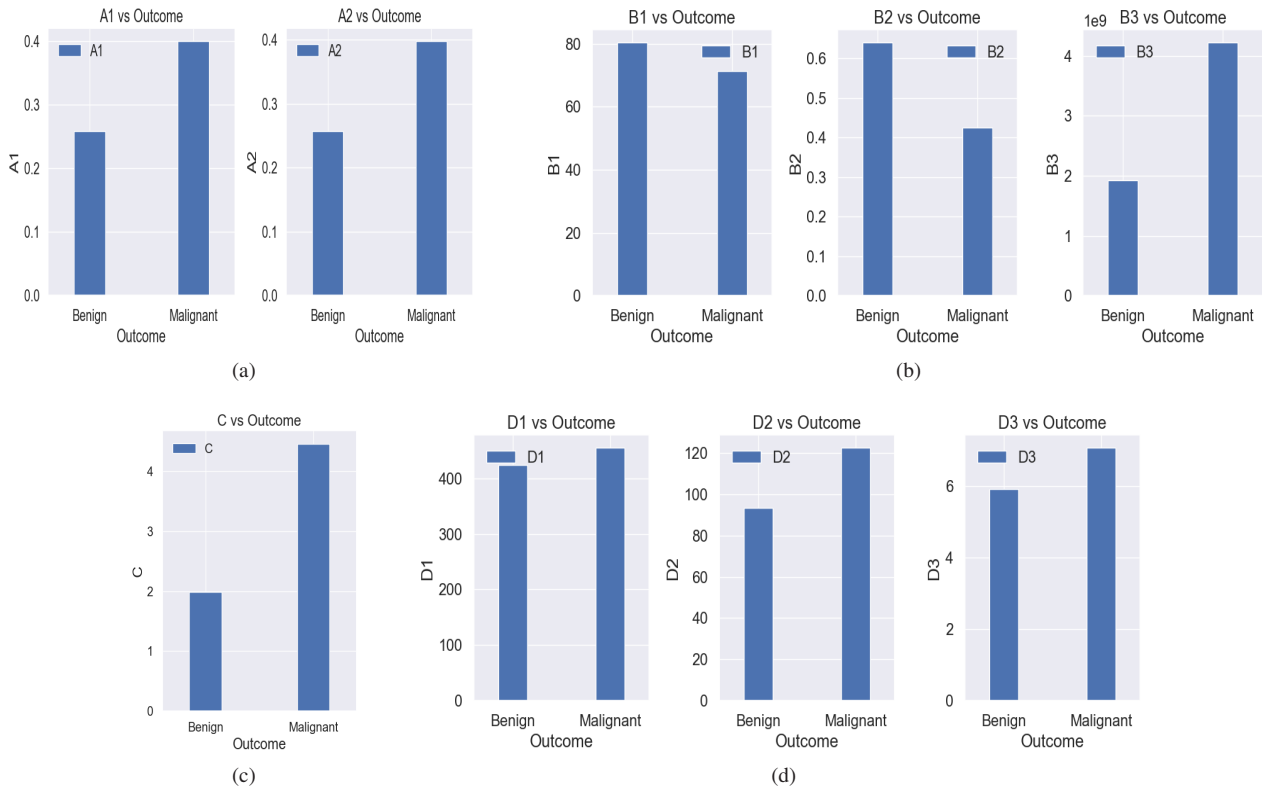


Figure 6: (a): Asymmetry Scores, (b): Border Scores, (c): Color Score, (d): Diameter Scores

4 Results

The diagnostic system proposed in this work has been implemented in Python. For the purpose of testing the system, 25 new images of lesions - 12 benign and 13 malignant - were used. The system could correctly classify the images as 'Benign' or 'Malignant'.

4.1 Feature Analysis

Figure 6(a), 6(b), 6(c), and 6(d) shows a graph representing the asymmetry scores ($A1$ and $A2$), border scores ($B1$, $B2$, and $B3$), color score, C , and diameter scores ($D1$, $D2$, and $D3$) for benign and malignant melanomas. It is observed that in case of malignant melanomas, $A1$ and $A2$ are greater for both axes, due to their asymmetric nature, $B1$ and $B2$ is smaller and $B3$ is greater, $D1$, $D2$, and $D3$ is greater, and finally, C is observed to be greater as the number of colors in them is higher when compared to benign melanomas. It is also noted that malignant melanomas tend to have 2 or more colors present in them.

4.2 System Performance

The performance measures calculated are: Accuracy, Precision, Recall, and F1-score using the Equations 12, 13, 14, and 15 respectively.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (12)$$

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

$$F1 - score = \frac{2TP}{2TP + FP + FN} \quad (15)$$

The system gives an accuracy of 89.93% with test error rate of 0.074, precision of 95%, recall of 82.61%, and F1-score of 88.37%. Figure 7 shows the ROC curve of the ANN classifier.

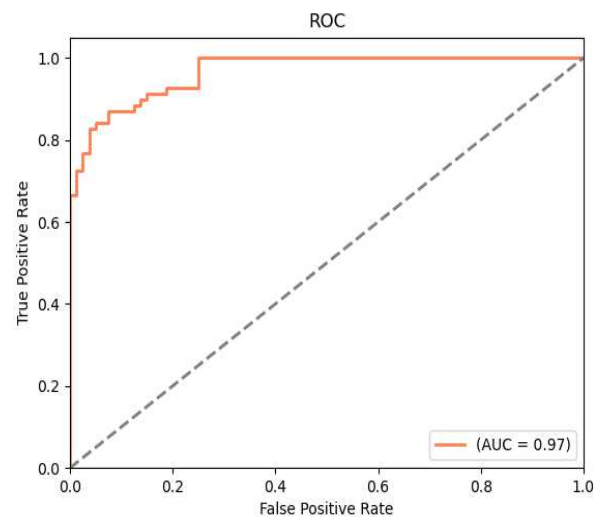


Figure 7: ROC curve of the ANN classifier

4.3 Classification with Different Classifier Models

The features extracted have also been tested with four other classifier models - Decision Tree, Random Forest, Extra Trees, and Gradient Boosting. Figure 8 shows the ROC curves of the different classifiers and Table 3 shows their performance measures.

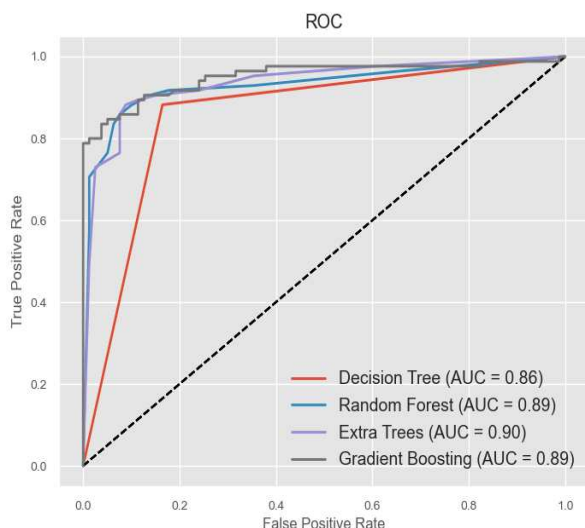


Figure 8: ROC curves of different models

Table 3: Accuracy and AUC of different classifier models

| Model | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|-------------------|--------------|-----------------|-----------------|
| Decision Tree | 85.98 | 87.4 | 86.8 |
| Random Forest | 89.02 | 87.12 | 92.88 |
| Extra Trees | 89.63 | 84.66 | 93.14 |
| Gradient Boosting | 88.41 | 86.03 | 93.14 |

It is observed that all the models have AUC of over 0.85 which is considered good for a classifier model. Also, the accuracy, sensitivity, and specificity scores of the models are also above 85% meaning that they can perform fairly well. This shows that the features extracted are appropriate and can be used to feed various classifier models to assist in classifying the skin lesions as benign or malignant.

5 Conclusion and Further Work

The exponential rise of melanoma skin cancer calls for an alternative approach to detect melanomas at an early stage so that it can be cured. The traditional approach to detecting melanomas comprise of carrying out skin biopsy. This method, although widely used, is painful, costly and time-consuming. There is a dire need of an automated process taking over the traditional method that is just as efficient, thus, making it an important area of research. The computer aided system proposed in this paper aims to do

the same by using various image processing methods and machine learning algorithms that can help accurately detect melanomas. This system provides an automated system with varied methods for pre-processing, segmentation, feature extraction and classification of image. However, there is still some areas that would require further study. One instance is the segmentation of image in dark-skinned patients. Also, a large amount of image datasets would be used for further improving training the classifier neural network in the future works.

References

- [1] World Cancer Research Fund on Skin Cancer Statistics. <https://www.wcrf.org/dietandcancer/cancer-trends/skin-cancer-statistics>. Accessed 07 Mar 2020
- [2] World Health Organization on Climate Change and Human Health. <https://www.who.int/globalchange/climate/summary/en/index7.html>. Accessed 07 Mar 2020
- [3] S. Majumder and M. A. Ullah, *Feature Extraction from Dermoscopy Images for an Effective Diagnosis of Melanoma Skin Cancer*, 2018 10th International Conference on Electrical and Computer Engineering (ICECE), 185-188 (2018). <https://doi.org/10.1109/ICECE.2018.8636712>
- [4] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. S. Marcal, J. Rozeira, 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 5437-5440 (2013). <https://doi.org/10.1109/EMBC.2013.6610779>
- [5] K. Eltayef, Y. Li, X.Liu, *Detection of Melanoma Skin Cancer in Dermoscopy Images*, J. Phys.: Conf. Ser. 787 012034 (2017). <https://doi.org/10.1088/1742-6596/787/1/012034>
- [6] J. Premaladha, K. S. Ravichandran, *Novel Approaches for Diagnosing Melanoma Skin Lesions Through Supervised and Deep Learning Algorithms*, J. Med Syst **40**, 96 (2016). <https://doi.org/10.1007/s10916-016-0460-2>
- [7] S. Jain, V. Jagtap, N. Pise, *Computer aided Melanoma skin cancer detection using Image Processing*, J. Procedia Computer Science **48**, 735-740 (2015). <https://doi.org/10.1016/j.procs.2015.04.209>
- [8] M. A. Sheha, M. S. Mabrouk, A. Sharawy, *Automatic Detection of Melanoma Skin Cancer using Texture Analysis*, J. Comp App (0975 – 8887) **42-No.20**, 22-26 (2012). <https://doi.org/10.5120/5817-8129>
- [9] A. G. Isasi, B. G. Zapirain, A. Mendez *Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms*, J. Comp in Bio and Med **41**, Issue 9, 742-755, (2011). <https://doi.org/10.1016/j.combiomed.2011.06.010>
- [10] Z. She, Y. Liu, A. Damatoa, *Combination of features from skin pattern and ABCD analysis for lesion classification*, J. Skin Research and Tech, **13**, 25-33 (2007). <https://doi.org/10.1111/j.1600-0846.2007.00181.x>

- [11] PH2 Database. <https://www.fc.up.pt/addi/ph2%20database.html>
- [12] N. Otsu, *A Threshold Selection Method from Gray-Level Histograms*, IEEE Transactions on Systems, Man, and Cybernetics, **9**, no. **1**, 62-66 (1979). <https://doi.org/10.1109/TSMC.1979.4310076>
- [13] T. F. Chan and L. A. Vese, *Active contours without edges*, IEEE Transactions on Image Processing, **10**, no. **2**, 266-277, (2001). <https://ieeexplore.ieee.org/document/902291>
- [14] D. Mumford and J. Shah, *Optimal approximation by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math, **42**, 577-685, (1989).