*Smart Computing and Systems Engineering, 2020*
*Department of Industrial Management, Faculty of Science, University of Kelaniya, Sri Lanka*

**Paper No: SE-12** **Systems Engineering**

# Novel computational approaches for border irregularity prediction to detect melanoma in skin lesions

D. V. D. S. Abeysinghe
*Department of Mathematics*
*Eastern University, Sri Lanka*
dharaabeysinghe@gmail.com

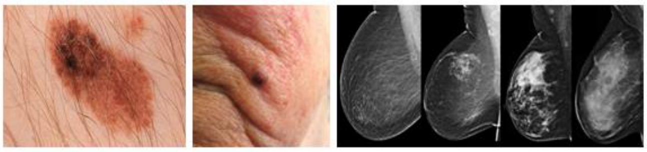
S. Sotheeswaran*
*Department of Mathematics*
*Eastern University, Sri Lanka*
sotheeswarans@esn.ac.lk

***Abstract*: Medical image detection has been a rapidly growing field of study during the last few years. There are different challenges associated with it. Many works have been done in order to provide solutions for key challenges. This study of work is focused on melanoma detection by using Asymmetry, *B*order irregularity, Colour textures, and Diameter (ABCD) feature along with proposing two new approaches for border irregularity detection. The proposed two new approaches are distance difference method and gradient method, which follows the main concept as traversing along the continuous borderline of the lesion. Further, this study varies from the existing studies, since it has been taken counts of distances from the centroid to the borderline without considering the distance from the image border to the borderline of the lesion. It was able to achieve a classification rate of 79% and 78.5% using distance difference method and gradient method, respectively whereas the classification without the border irregularity feature achieved 78% of accuracy performing on PH$^2$ dataset. Further, this study can be stated as most appropriate to classify non-melanoma rather than melanoma. It is contributed by generating simple computer science-based approaches rather than complex mathematical methods to detect border irregularity and makes the medical image detection easy.**

***Keywords:*** *ABCD features, Border irregularity, Distance difference method, Gradient method, Medical image detection, Melanoma and non-melanoma classification.*

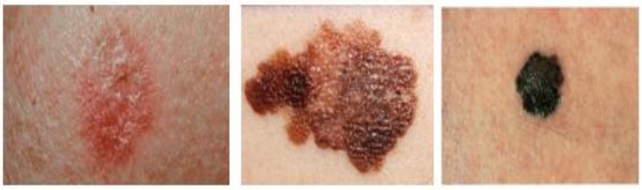
## I. INTRODUCTION

Medical image detection or the task of identifying medical-based objects in a medical image is an important analysis to perform since medical images contain different characteristics rather than normal images. Many key challenges with regards to medical image detection exist such as preprocessing, segmentation, detection and extraction of features, classification into classes and finally to make crucial decisions about the prediction as shown in Fig. 1. These challenges can be highly impacting the dermoscopic images used in skin cancer detection when determining the region of interest and also in the extraction of features to predict skin cancers. Cancers are one of the most common diseases in the world that can develop almost everywhere in the body and over a hundred and more types of cancers skin cancers are a type that can be prevented and treated if found early. Among skin cancers, melanoma also known as malignant melanoma begins in cells called melanocytes that are in the upper layer of the skin which are responsible to produce a pigment called melanin to give the color to the skin. Some dermoscopic images of melanoma and non-melanoma are shown in Fig. 2.



(a) (b) (c)

Fig 1. Challenges to medical image detection: (a) Skin cancer image with hair [1] (b) Dermoscopic image with wrinkles [2] and (c) Scans of Breasts with different densities [3].



(a) (b) (c)

Fig 2. Melanoma and non-melanoma skin cancers : (a) Non-melanoma [4] (b) melanoma [5] and (c) melanoma [6] skin cancers.

In order to cure melanoma, it is an essential factor to enhance the prediction of skin cancers. Hence, to enhance the diagnosis of melanoma dermatologists tend to aid dermoscopy techniques where removal of unnecessary features of skin and get an improved visual effect of skin to produce more related details [7]. Melanoma looks tricky to identify at once due to the complexities in shapes, colors, size, etc. However, constant skin examinations may lead to earlier and correct detection. When focusing on the detection of melanoma, several criteria are available to identify a lesion as cancerous. Ugly duckling sign is one such property to locate a spot that is different from all other spots in the body. Further in dermoscopy, pigment networks, streaks, blue-whitish veil, regression structures, dots and globules, vascular structures, pink shades and starburst pattern etc. are used in the prediction of melanoma. ABCDE features or **A**symmetry, **B**order irregularity, **C**olour textures, **D**iameter and **E**volution of a mole are the most prominent set of features among those features. There can exist moles having one or two above properties or all of the above in it.

Hence this study has focused on improving the detection of melanoma, a quite common type of skin cancer, by providing two new approaches for border irregularity detection which is a feature among the set of geometric features, ABCD to predict melanoma.

The rest of this paper is organized as follows. In Section 2, we summarise different techniques that are closely related to melanoma detection. Section 3 provides the background of skin cancer detection. In Section 4, the proposed techniques are described in detail. Section 5 describes the experimental setup. Section 6 discusses the results of the experiment. Finally, Section 7 concludes the paper with a discussion of the findings towards future extensions.

## II. Literature review

When considering skin cancer detection there are ample studies done related to it in the last decade. This Literature review is focused on the melanoma detection and the border irregularity skin cancer detection of medical images.

In [8], authors have proposed a computer-aided method for melanoma detection using Image Processing tools with ABCD parameters. Most importantly the proposed method pre-processes the images to enhance image quality. Then threshold values were generated by an automatic thresholding process and a binary mask has been created using the threshold values to locate the biggest blob to locate the lesion. Hence along with the location and with edge detection, the lesion has been segmented. Then standard geometry features such as area, perimeter, greatest diameter, circularity index and irregularity index were extracted and classification has been done to produce effective results. Overall this system has been concluded to be a user-friendly and robust system for images under any condition and diagnosis can be expected to be in a higher degree of accuracy.

In [9], authors have proposed a method to measure border irregularity of a skin lesion by locating indentations and protrusions along the lesion border and by creating a new area-based index called irregularity index for each indentation and protrusion. Initially, the scale-space has been extended from binary to a three-valued image and all the images have been preprocessed to remove hair using DullRazor and borders have been detected automatically. Then Gaussian smoothing has been performed and at every smoothing scale, the zero-crossings of the first derivative have been recorded. Finally, the sum of all the irregularity index values has been calculated for the overall prediction. Through this study, it has been able to pinpoint the problem areas of the lesions and also this method can be further used for other medical applications even.

In [10], authors have proposed an approach to detect and classify border irregularity of a skin lesion. This proposed method is consisting of four major steps as image enhancement, lesion segmentation, border irregularity detection and classification. Mainly, in the detection of border irregularity, the proposed study has implemented a four-step algorithm. In the algorithm, the first step computes the bounding box of the lesion. Then the boundary pixels upon the lines joining the center of mass with the vertices have been found. Thereafter, the distance between the border and the image edge has been calculated by soothing ragged edges, if there exists any using the Gaussian filter. Finally, to detect the border irregularity the derivative has been calculated to find the local maximums of the function by detecting the sign changes across the zero point. Then authors have performed a border irregularity classification by the overall score obtained from the scores of each segment of eight segments. This study has used 300 images taken from the Interactive Atlas of Dermoscopy image repository and has achieved a detection accuracy of 79% and a classification accuracy of 90%.

In [11], the author has suggested a method based on lesion rotation and borderline division. Initially, black frame removal, hair detection and hair inpainting have been done to enhance the images. Secondly, the lesion has been segmented using a seeded region-growing algorithm. With the region growing techniques, the proposed method has been able to detect borders correctly even where edges are noisy. Then for the border irregularity assessment, the skin lesion has been rotated keeping the major axis of the lesion parallel to the horizontal line. Since happening a spatial transformation, one set of pixels are remapped to another. According to the proposed method, if the images are continuous then the remapping will not use interpolation, whereas if the images are discrete remapping requires interpolation. Also, for the process of interpolation, a bilinear method has been chosen after comparing it with the nearest neighbor method. After rotating the skin lesion borderline has been detected precisely and then the borderline function has been calculated and boundary pixels on the lines joining the center of mass with the vertices were found. With the use of those four points, the boundary has been divided into four unequal parts. Then the distance between the border and the image edge for each boundary part has been calculated. Further, the gap between the functions which got the results from the distance to the edges has been subtracted to get the border irregularity function. When detecting the border irregularities, a turning point or a point where the sign of the derivative changes has been taken into consideration. Finally, with those points, the border irregularity has been calculated. This proposed method has been tested on 350 dermoscopic images, 280 begin and 70 malignant cases that have been collected by two university hospitals. Further, this database contains 140 cases with borer irregularity less than 3.0 while 210 cases have border irregularity greater than 4.0. Overall the system has achieved 91% accuracy and 89% precision.

Finally, can be stated that many studies have done preprocessing of images to remove unnecessary objects such as hair, black frame, ink marks etc. initially to achieve an accurate segmentation. And the use of geometric features such as ABCD features have been widely used and immensely supportive in melanoma detection. And border irregularity is a crucial factor to be determined and identification of borderline correctly will affect the results. Further, most of the border irregularity techniques have accounted for the distance between the borderline and the image edge to identify the irregularity and not have taken the distance between the centroid of the lesion and the borderline into the account.

Through this study, attempts to fill the gap between the requirements and the existing solutions. Typically, for melanoma detection, ABCD can be stated as a set of pioneer geometric features, and it always generates the requirement to have novel techniques to extract the features. Hence through this study, it focuses to enhance geometric features for melanoma detection (ABCD) by proposing novel methods for border irregularity. Further, this study varies from the existing border irregularity detection techniques since it has been taken counts of distances from the centroid to the borderline without considering the distance from the image border to the borderline of the lesion. Also, the study uses continuous borderline without using the discrete parts of the borderline in irregularity prediction.

## III. BACKGROUND

### A. Skin cancer detection

In the detection of melanoma, some traditional medical methods like histopathological examinations and biopsies are used. Further, in the detection of non-invasive skin cancers, different other methods such as photography, ultrasound, confocal microscopy, Raman spectroscopy, fluorescence spectroscopy, terahertz spectroscopy, optical coherence tomography, the multispectral imaging technique, thermography, electrical bio-impedance and tape stripping are used [12]. Moreover, with the advancement of technology detection using dermoscopic images with Computer-Aided Diagnosis (CAD) systems has been a promising method. In CAD, features are extracted and images are being analyzed to diagnose skin cancers with the aid of numerous computers science approaches such as image processing, machine learning, neural networks and data mining etc. Though CAD systems are highly sensitive for the detection, it has been found that the false-positive rate in CAD predictions is high and the traditional diagnostic methods cannot be still replaced [13]. However, CAD can be a perfect aid for a medical doctor to make his traditional predictions accurate since dermatologists' inspection of dermoscopic images is a complex task that requires ample time and expert skills. Some types of traditional biopsy methods are shown in Fig. 3.

### B. Corner points detection

#### 1) Harris corner detection

1. It is a combined algorithm to detect both border and corner points which have been developed based on Moravec's corner detector by revising it [14]. Harris corner detection algorithm relies on an autocorrelation matrix and then eigenvalues of the matrix values have been found to locate strong intensity variations among the neighborhood [15]. Initially, the images are smoothed using Gaussian filters and the gradient of the image has to calculate. Then the autocorrelation matrix must be generated. And using the autocorrelation matrix, eigenvalues for them have been calculated computing the corner strengths by Harris measure in (1). Then using non-maximum suppression the highest intensity points are being selected. Finally, among the selected points further, the output corner/edge points are chosen and subpixel accuracy has been calculated. (See Fig. 4).

$$R_H = \lambda_1\lambda_2 - \kappa.(\lambda_1 + \lambda_2)^2 \qquad \textbf{(1)}$$

Where,
$\lambda_1$, $\lambda_2$: two real non-negative eigenvalues, κ: value between 0.04 or 0.06 and $R_H$: corner response [15].

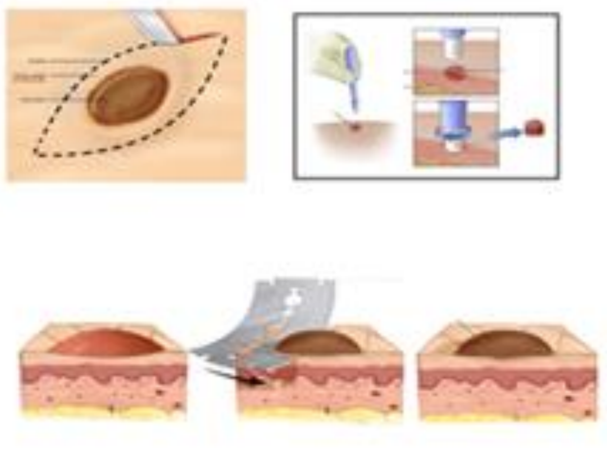


Fig 3. Types of biopsies: Excision biopsy, Punch biopsy and Shave biopsy [16].

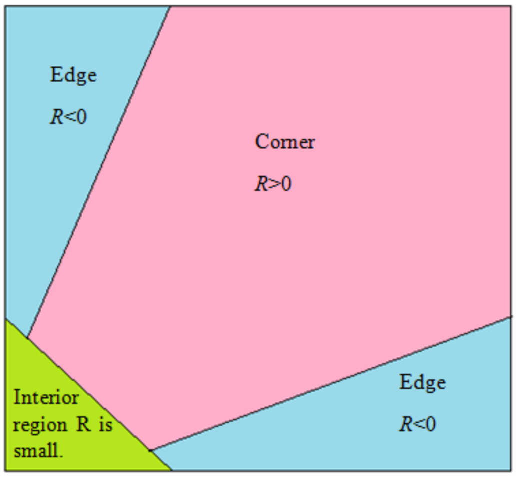



Fig. 4. Harris measure

### C. Lesion classification

#### 1) Naive Bayes classifier

1. Naive Bayes classifier is one of the most prominent classifiers that can be used for medical diagnosis which consists of a group of machine learning algorithms. It is a classifier that uses probability concept and classifying based on Bayes' theorem. The classifier is considered a naive approach as it considers all the data to be independent with each other. This classifier will get to know all the features of the relevant classes and will consider one by one feature for the classification. The Bayes' probability theorem for classification is given below in (2).

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \qquad \textbf{(2)}$$

## IV. METHODOLOGY

### *A. Binarisation*

The obtained dataset has 3 classes of RGB color images of skin cancers along with binarised, well-segmented lesions of each of the images. Then the relevant images were separated as the RGB and binary images into different groups.

### *B. Border irregularity determination*

Two different kinds of methods of the border irregularity were proposed and implemented.

#### *1) Distance difference method*

The centroid of the binary image and the corner points of the lesion were determined. Hence the distances between all corner points and the centroid have been calculated. Using the above-calculated distance, all height differences of the consecutive neighbor points have been calculated while traversing along with the detected corner points. Then considering a set of threshold values ranging from 15 to 30 (16 values) as T1, the extreme height differences above the thresholds were determined for each value. Then the percentage of the outward or inward points concerning the total number of corner points was determined. Another threshold value T2 of values 20, 25 and 30 were assigned to filter the highest percentages to determine the border irregularity. Images having the percentage greater than the second threshold value (T2), was considered to be irregular borders whiles others were considered as regular borders (See Fig. 5). Both threshold values were determined after observing the obtained results for the distance differences and then the percentages of distance differences.

#### *2) Gradient method*

1. In this approach, the gradients of the consecutive corner points were considered. Similar to the above method corner points were determined and the gradients of the neighbour corner points were determined. Using the gradient values rise and fall of the borderline was determined. Hence if the number of rises and falls in a particular image is above a set of threshold values (T), ranging from 45 to 55 (11 values), that was selected after observing the obtained results for gradients and then counting the number of sign changes to determine a local maximum and local minimum. Hence the threshold value T was considered as a benchmark predicting border irregularity. If an image has a ratio between no. of turning points to no. of corner points higher than T, that image was predicted as an image with an irregular border (See Fig. 6).

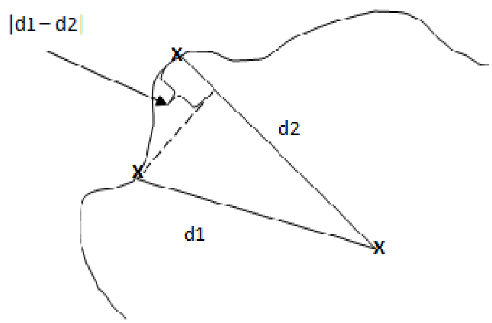



Fig 5: Distance difference method.

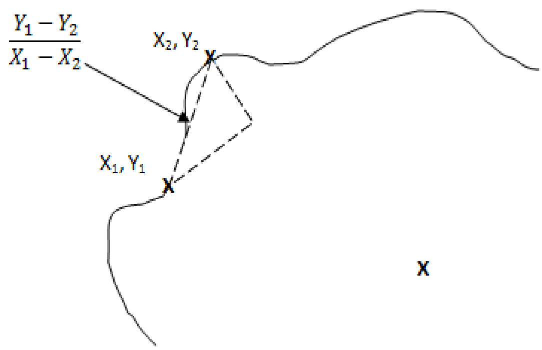



Fig 6: Gradient method.

### *C. ABCD features*

The acquired dataset was having asymmetry, colour features along with it. The diameter of the lesions was also calculated in pixel values using 'region props' function in MATLAB and converted to millimeters. Then using the above-calculated border irregularity predictions, 48 predictions from the distance difference method by two threshold values (T1×T2=16×3) and 11 predictions from gradient method by threshold value different set of ABCD feature values were generated. For each set of ABCD features the classifications were conducted to determine the most appropriate threshold values for each approach particularly for the used dataset. Further, a control set with only using **A**symmetry, **C**olour features and **D**iameter (ACD), was also generated other than the ABCD sets to compare whether the obtained results were able to enhance the prediction or not.

### *D. Classifications*

Naive Bayes classifier in WEKA tool was used to classify the dataset. The classification was done using ten-fold cross-validations making the whole dataset as training and testing to train and test the classifier. Initially with ACD features classification was performed and it was done as the control parameter. Then all features for the above generated 48+11 datasets from both methods were used to classify. Finally, the obtained classification results were compared with each other with respect to the control parameter. When comparing the results, the confusion matrices, number of correctly classified and incorrectly classified instances, True Positive rate (TP), False Positive rate (FP), precision, Root Mean Square Error (RMSE) and also the contribution of border irregularity feature for the atypical, common and melanoma class classification were considered.

## V. EXPERIMENTAL SETUP

### A. *Dataset*

In this study, $PH^2$ dataset which includes 200 dermoscopic images in three classes as atypical nevus, common nevus and melanoma were used. The dataset contains 80 atypical nevus images, 80 common nevus and 40 melanoma images. All the images are in the same size and magnification (See Table I).

### B. *WEKA tool*

Weka tool version 3.9 was used in the classification. It was operated in core i7 4500U CPU @ 1.80GHz 2.40GHz processor with 8GB RAM, 64-bit operating system.

### C. *MATLAB*

In this study MATLAB 2015a, a 32-bit version was operated in a core i7 6700U CPU @ 3.40GHz 3.41GHz processor with 4GB RAM, 64-bit operating system.

TABLE I. CLASSES OF $PH^2$ DATASET.

| | Atypical nevus | Common nevus | Melanoma |
|---|---|---|---|
| **No of images** | 80 | 80 | 40 |

## VI. RESULTS AND DISCUSSION

Through this study, it was able to obtain 78% of correctly classified instances using the ACD features. Further, using the distance difference method 79% of correctly classified instances and by the gradient method it was able to achieve the correct prediction of 78.5% (See Table II.) And when discussing the classifications, it was observed that the proposed methods work well with the non-melanoma classes, atypical nevus and common nevus identification. (See Table III-V). That can be stated because the classifications of atypical and common nevus are correct than the melanoma classification.

Further, according to the results, sample 44 (T1=29, T2=20) shows the best performance in the distance difference method. Moreover, the most accurate class prediction ratios of distance difference method 55/80, 69/80 and 34/40 for the three classes atypical, common and melanoma respectively can only be seen together in the confusion matrix of sample 44. Hence confusion matrix of sample 44 again proves that it has the best performance in the distance difference method. Similarly, sample 10 (T=54) shows the best performance in the gradient method. The most accurate class prediction ratios 52/80, 72/80 and 33/40 for the three classes atypical, common and melanoma respectively can only be seen together in the confusion matrix of sample 10, it again proves sample 10 has the best performance in the gradient method.

TABLE II. CLASSIFICATION RESULTS.

| Method used | Accuracy | TP | Precision | F-measure | RMSE |
|---|---|---|---|---|---|
| Only with ACD features | 78% | 0.780 | 0.790 | 0.778 | 0.334 |
| ABCD features with distance difference method | 79% | 0.790 | 0.796 | 0.789 | 0.336 |
| ABCD features with gradient method | 78.5% | 0.785 | 0.795 | 0.783 | 0.335 |

TABLE III. BEST RESULTS OF CONFUSSION MATRIX FOR DISTANCE DIFFERENCE METHOD

| Class | Atypical | Common | Melanoma |
|---|---|---|---|
| **Atypical** | 55 | 23 | 2 |
| **Common** | 11 | 69 | 0 |
| **Melanoma** | 5 | 1 | 34 |

TABLE IV. BEST RESULTS OF CONFUSSION MATRIX FOR GRADIENT METHOD

| Class | Atypical | Common | Melanoma |
|---|---|---|---|
| **Atypical** | 52 | 26 | 2 |
| **Common** | 8 | 72 | 0 |
| **Melanoma** | 6 | 1 | 33 |

TABLE V. CONFUSION MATRIX WITH ONLY ACD FEATURES

| Class | Atypical | Common | Melanoma |
|---|---|---|---|
| **Atypical** | 51 | 27 | 2 |
| **Common** | 9 | 71 | 0 |
| **Melanoma** | 5 | 1 | 34 |

Overall, when illustrating the border irregularity predictions of skin lesions through the proposed methods, few sample results can be given as follows. The given sample images and respective edge distance difference patterns for consecutive borderline points are shown below in figures. (see Fig. 7-10). Considered threshold value T1=29 is the best T1 threshold value obtained by the distance difference method. Moreover, the following images also have been predicted as the same type as irregular and regular by both gradient and distance difference method.

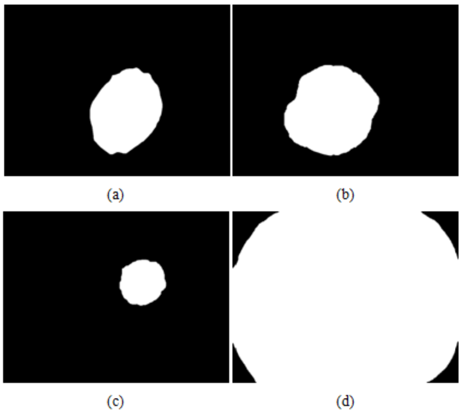



Fig 7: Sample images with regular borders.

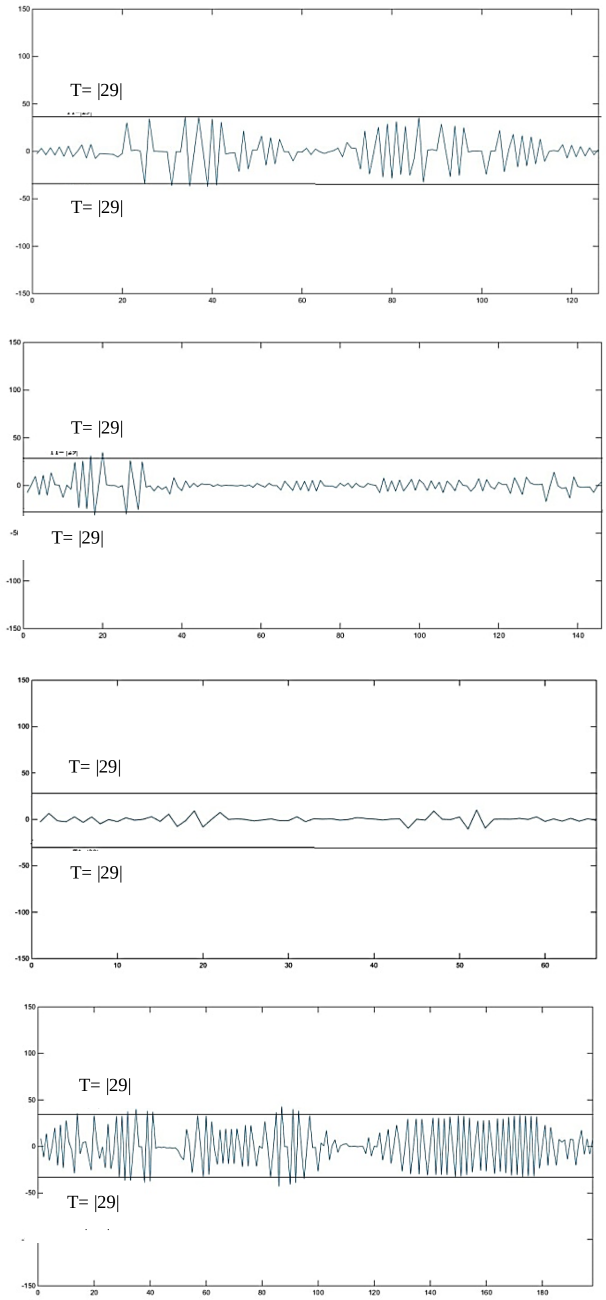



Fig 8: Distance difference of consecutive points for regular border lesions for irregular border skin lesions (a), (b), (c) and (d) in Fig. 7 respectively.

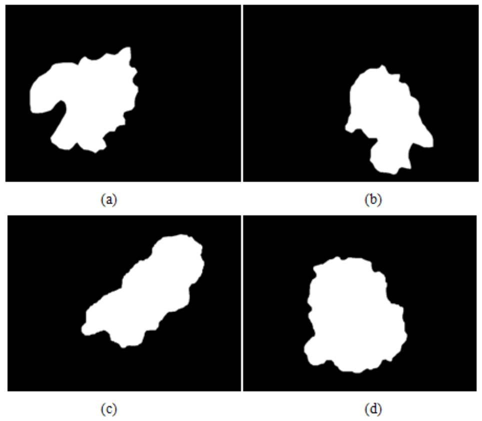



Fig 9: Sample images with irregular borders.

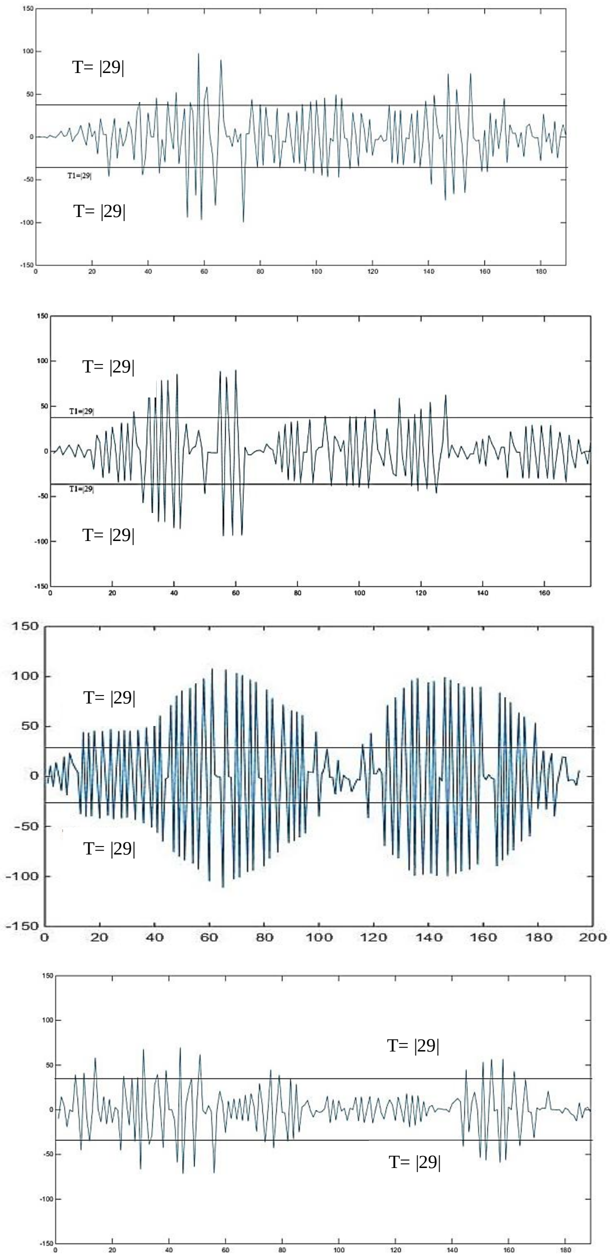



Fig 10: Distance difference of consecutive points for irregular border lesions for irregular border skin lesions (a), (b), (c) and (d) in Fig. 9 respectively.

The results can be varied with the dataset and also with the quality of data. Since this dataset has 200 images and most of the images are fallen to those two categories rather than melanoma it can be a reason for the above observation. Since the data have been segmented well and also the region of interest have been identified clearly, it has improved the precision of the results. However, the use of binary images can be in effect for the obtained results as the patch-based or color intensity features cannot be obtained rather than the number of colors was taken into account. Moreover, this study has obtained commendable results with only using the geometrical features.

## VII. Conclusion

Through this study, it was able to propose two simple approaches to determine the border irregularity of a skin lesion. Further, the obtained results were used in melanoma classification to validate the obtained results. Overall from the study, it can be concluded that the classifications perform well in non-melanoma rather than melanoma classifications. Also, without the extracted border irregularity feature, the classification gives an accuracy of 78% whereas with the results it gives a slight improvement in classification as 79% and 78.5% for distance difference method and gradient method respectively. Further from the suggested methods, for the particular dataset distance difference methods shows remarkable performances in border irregularity detection and can be stated as the best method out of the two.

The future works of this study are commenced in two ways which are bag-of-features and deep learning approaches. The main intention of that work is to compare the success of using geometrical features and bag-of-features in melanoma detection. Further hoping to repeat the proposed approaches in detecting border irregularity for a large dataset and hence compare the performance with this study to validate the proposed methods.

## References

[1] T. Allison, "6 Steps to Check for Melanoma." Healthcentral.Com, HealthCentral, 4 Oct. 2012, sOnline]. Available at: www.healthcentral.com/slideshow/6-steps-to-check-for-melanoma. [Accessed 6 October 2019].

[2] "Another Pigmented BCC." Dermoscopy, [Online]. Available at: dermoscopy.wordpress.com/2014/09/15/another-pigmented -bcc/, 15 September 2014. [Accessed October 2019].

[3] "Breast Density: Are You Informed?", Imaging Technology News, [Online]. Available at: www.itnonline.com/article/ breast-density-are-you-informed, 30 April 2014, [Accessed 6 October 2019].

[4] Editorial Team, "Skin Cancer (Nonmelanoma).", Academic Association of Medicine, [Online]. Available at: aaom.org/health/skin-cancer-nonmelanoma/, 7 December 2017, [Accessed 6 October 2019].

[5] Wikipedia Contributors. "Melanoma." *Wikipedia*, Wikimedia Foundation, [Online]. Available at: en.wikipedia.org /wiki/Melanoma, 12 October 2019, [Accessed 6 October 2019].

[6] The Skin Cancer Foundation. Skin Cancer Information - The Skin Cancer Foundation. [Online]. Available at: https://skincancer.org/skin-cancer-information 2019. [Accessed 6 October 2019].

[7] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks", IEEE Transactions on Medical Imaging, vol. 36, no: 4, pp. 994-1004, April 2017.

[8] J. Shivangi, et al. "Computer-Aided Melanoma Skin Cancer Detection Using Image Processing.", International Conference on Intelligent Computing, Communication and Convergence (ICCC 2015), Procedia Computer Science, vol. 48, pp. 736–741, 2015, India.

[9] T. K. Lee and M. S. Atkins, "A new approach to measure border irregularity for melanocytic lesions", Medical Imaging 2000: Image Processing, 2000.

[10] J. Jaworek-Korjakowska and R. Tadeusiewicz, "Determination of border irregularity in dermoscopic color images of pigmented skin lesions," 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2014.

[11] J. Jaworek-Korjakowska, "Novel method for border irregularity assessment in dermoscopic color images," Computational and Mathematical Methods in Medicine, vol. 2015, pp. 1–11, 2015.

[12] V. Narayanamurthy et.al. ,"Skin cancer detection using non-invasive techniques", Review Article, RSC Advances, vol. **8**, no.49, pp. 28095-28130, 2018.

[13] L. F. di Ruffano et.al. "Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults", Cochrane Database of Systematic Reviews, Apr. 2018.

[14] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," *Proceedings of the Alvey Vision Conference 1988*, 1988.

[15] J. Sánchez, N. Monzón, and A. Salgado, "An Analysis and Implementation of the Harris Corner Detector," Image Processing On Line, vol. 8, pp. 305–328, Mar. 2018.

[16] B. U. W. Lei and T. W. Prow, "A Review of Microsampling Techniques and Their Social Impact.", Biomedical Microdevices, vol. 21, no. 4, 15 August 2019