Melanoma Classification on Dermoscopy Skin Images using Bag Tree Ensemble Classifier

Nay Chi Lynn, Nu War
University of Information Technology; University of Computer Studies, Mandalay
naychilynn@uit.edu.mm, nuwar@ucsm.edu.mm

Abstract

Melanoma classification on dermoscopy skin images is a demanding task because of the low contrast of the lesion images, the intra-structural variation of melanomas, the high degree of visual similarity between melanoma and non-melanoma lesions, and the existence of hair and ruler marker artifacts. In this study, the malignant melanoma skin cancer classification system is proposed with the aid of correctly classify melanoma skin cancer. The system involves three main steps: segmentation, feature extraction and classification. Ahead of the segmentation step, the preprocessing skin lesion images is processed for getting rid of the covered hair artifacts. In the segmentation step, the input preprocessed lesion image is segmented by using the proposed texture filter-based segmentation method. Then, the features according to the underlying ABCD (Asymmetry, Border, Color, Differential Texture) dermatology rules using shape, edge, color and texture features are extracted from the segmented region. Finally, the extracted features are classified to identify whether the skin image is malignant melanoma or non-melanoma with the use of bag tree ensemble classifier. The system performance is evaluated with the use of the benchmarking datasets: PH2 dataset, ISBI2016 dataset and ISIC2017 dataset. According to the experimental results, the proposed design allows for both reliable classification of real world dermoscopy images and feasible operation time with today's standard PC computing platforms. To address the class imbalance in the dataset and to yield the improved classification performance, the experiments are also analyzed not only on original imbalanced dataset but also on balancing datasets: undersampled and oversampled datasets. The system works well and provides both high sensitivity and specificity according to the experimental results on the oversampled dataset with bag tree ensemble classifier to leading to statistically better performance compared to original imbalanced dataset.

Key Words- Melanoma, Dermoscopy Skin Image, Segmentation, Feature Extraction, Ensemble Classification

1. Introduction

Skin dermoscopy is a non-invasive diagnostic technique for the inspection of pigmented skin lesions, which makes the structures more visible. This diagnosis allows the recognition of morphological structures that are invisible to the naked eye and opens a new dimension in the analysis of the clinical-morphological features of pigmented skin lesions. Recently, several systems have been described in the literature for computer-aided analysis of skin endoscopy-derived digital images, which are considered an active field of research.

This study describes an image processing approach for computer-aided melanoma detection on dermoscopy images. The proposed methods provide lesion border detection and extract relevant technical features of dermoscopic structures for melanoma detection.

The work developed to meet the aim belongs to the following four objectives as follows:

- To develop a computerized skin cancer detection system for diagnoses of melanoma
- To extract the efficient features for classification of skin lesion image
- To analyze the problem of imbalanced medical dataset classification and the influence of classifiers on melanoma detection
- To help patients to prevent the melanoma in its early stages

The benchmarking datasets: PH2, ISBI2016 and ISIC2017 which are publicly available collection of quality controlled dermoscopy images of skin lesions.

2. Preprocessing Image

This enhancement technique is an image processing technique that maximizes the average contrast of local images and produces the output image with a high-quality display. This is an adjustment method that lets image intensity values to display in a new enhanced range. Low contrast images including histograms and all the gathered values, are centered on the intensity area.

An initial preprocessing step with the aim of enhancing the skin images by removing unwanted hairs which

partially shade the main region of interest. Ruler marks are dealt with similarity and they mimic hair like structure. Hair and ruler maker detection and removal is the one which replace hair pixels by neighboring pixels. The dark hair locations need to identify using generalized grayscale morphological closing operation. The shape of the hair pixels as thin and long structure is verified and replaced by using bilinear interpolation and the Dull Razor approach is used in this research.

3. Segmentation

Lesion segmentation is a very important step in the analysis of dermoscopic skin images as it allows the identification of various features specific to the lesion. Indeed, segmentation of the lesion means separating the lesion (region of interest) from the normal surrounding skin region.

3.1. Intensity Adjustment

Image adjustment is used for intensity transformations of gray scale images by increasing intensity while leaving the contrast unchanged in the whole image. This also means that increasing intensity can be viewed as brightening the image. The shape and edges of image are enhanced to increase the image clarity, obtain better performance and get high contrast. In addition, intensity adjustment can sharpen the image border and improve the accuracy for segmentation. Intensity adjustment is used for intensity transformations of image by increasing image's intensity values to a new range while leaving the contrast unchanged in the whole image.

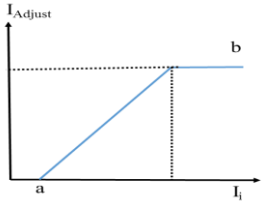

Figure 1: Image Intensity Adjustment

The input RGB is separated as: Img[]]] = R[] G[] B[]
The intensity values from each channel are transformed to a new range, and the new intensity channels are then merged to get the transformed RGB image.
∑(R_Adjust[],G_Adjust[],B_Adjust[])=Img_Adjust[]]]
I = Intensity original image
S = T(I) = Transformed image, T is Transform function

$$I_{Adjust} = \begin{cases} 0, & i \le a \\ 255, & i \ge b \\ 255 * (I_i - a)/(b - a), & otherwise \end{cases}$$

a=min(I_i); minimum of original image intensity values
b=max(I_i); maximum of original image intensity values
where, I_i= Original image on each separated channel (e.g. R, G, B)
I_{Adjust}= Transformed/Adjusted Image
I_{Adjust}= T(I), where T is Transform function

After the intensity values of each separated color channel are enhanced, each color channel is then combined again to get intensity adjusted multi-channel image. This operation is generally used to make objects in an image more distinguishable, thus making it much easier to see the details.

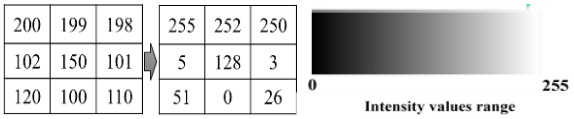

Figure 2: Image Intensity Adjustment Range

3.2. Integral Image Calculation

Integral image is a useful image representation that can be calculated by computing the sum of local images quickly to improve the texture rendering rate. The sum area table is calculated to obtain an integral image. The summed area table can be created with a single pass over a given image. In this research, the intensity adjusted image is converted to grayscale image and therefore the summed area table is computed over the intensity adjusted grayscale image.

$I_{Int}(j,k) = I_{Gray}(j,k) + I_{Gray}(x-1,y) + I_{Gray}(x,y-1) - I_{Gray}(x-1,y-1)$

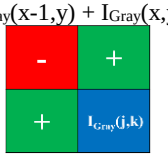

Figure 3. Example of Integral Image Calculation

That is, the original pixel value i(x,y) is resulted from the image, and then it needs to add the values directly above this pixel, and directly left to this pixel from the Summed Area Table at s(x-1, y) and s(x, y-1). Finally, the value is subtracted directly top-left of i(x,y) from the summed area table – that is, s(x-1, y-1). On the left there is the given image, with its corresponding pixel values. On the right there will be the images with the corresponding summed area table.

3.3. Box Smooth Filtering

Given an integral image, the integral filter filters the image. The filter size does not affect the speed of the filtering operation. After computing the sum area table, the sum of intensities is evaluated over any rectangular area with four array references exactly. Then, average integral filter is applied to image with the threshold filter size. The box smooth filter performs average smoothing on an image. This filter is effective at attenuating noise because averaging removes small variations. The effect is identical to that of averaging a set of data to help reduce the effect of outliers.

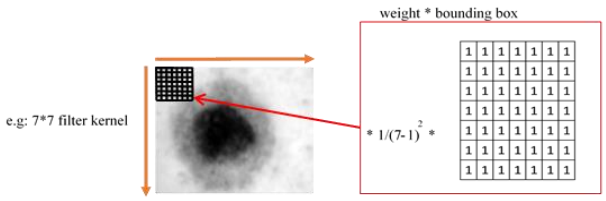

Figure 4: Box Smooth Filter on Integral Image

3.4. Entropy Filtering

Entropy is the "statistical measure of disorder, randomness or uncertainty" and in image processing, the entropy of an image is the average information generated from its pixels. These statistics can characterize the texture of an image because they provide information about the local variability of the intensity values of pixels in an image.

The local entropy of specified moving window region is computed and stored it in each output pixel. The output pixel is every pixel from the top leftmost to bottom rightmost pixel in the image. The specified neighboring pixel region is calculated as a 9 × 9 neighborhood pixel window size in this work. The entropy is calculated using the following equation:

$E(9,9) = -\sum_{i=1}^{k} (p_i \, log_2 p_i)$,

where, k = the total number of gray levels occurred in specified window ($1 \leq k \leq 81$)

p_i = the occurrence probability of gradient magnitude of pixel:

$p_i = n_o/n_i$,

n_o = the occurrence count of the intensity level i

n_i = the total number of intensity levels (e.g. $n_i = 81$)

For pixels on the borders of image, symmetric padding is used. In symmetric padding, the values of padding pixels are a mirror reflection of the border pixels in I.

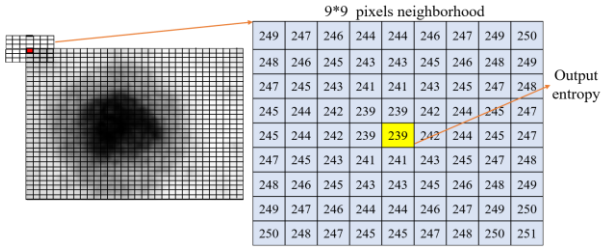

Figure 5: Entropy Filtering on smoothed filtered image using 9*9 neighborhood

3.5. Adaptive Thresholding

Find a scalar threshold value 'th' representing the whole entropy value of the filtered image. Create a binary image from entropy filtered image by replacing all values above a determined threshold by thresholding with 1s and setting all other values to 0s.

$$F(x,y) = \begin{cases} 1 & if\ H(x,y) \geq th \\ 0 & otherwise \end{cases}$$

F(x,y) is output thresholding image

H(x,y) is input entropy filtered image

$th = -\sum (i=1)^{k} (p_i \, log_2 p_i)$

where, k = the total number of gray levels occurred in the whole input image

($1 \leq k \leq$ [number of rows* number of columns])

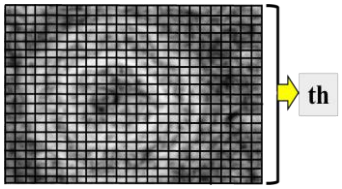

Figure 6: Box Smooth Filter on Integral Image

3.6. Morphological Operation

Morphology, in general, deals with forms or structures by fixing the structuring element A which defines region of interest or neighborhood around a pixel.

Dilation by A: $\partial A(X) = X \oplus A = \{z | (A)z \cap X \neq \emptyset\}$.

Erosion by A: $\epsilon A(X) = X \ominus A = \{z | (A)z \subseteq X\}$.

Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries.

The opening is an erosion followed by a dilation: $(X \ominus A) \oplus A$. The closing is a dilation followed by an erosion: $(X \oplus A) \ominus A$. Morphological closing is used in this work to fill gaps in an image with a circular disk structuring element of size 6.

3.7. Dark Corner Removal

Dark corner vignettes and color charts have different sizes and numbers per image. However, they used to appear in the corner regions. In this work, the lesions are firstly checked if their centroids are likely to be in likelihood rectangular region by calculating 'shortest' Euclidean distance D between centroids of binary regions with the center of image.

If the first point is (a, b) and the second point is (c, d), then the distance D between first and second point

$$D = \sqrt{((a-c)^2 - (b-d)^2)}$$

4. Feature Extraction

Features are extracted using the so-called ABCD (Asymmetry, Border structure, variegated Color, and the Differential Structures of the skin lesion) diagnosis rule which is known as early signs of melanoma and is the basis for a diagnosis by a dermatologist.

Asymmetry: One half of the tumor does not match the other.

Border Irregularity: The edges are ragged, notched,

blurred.

Color Variation: The color (pigmentation) is not uniform. Shades of tan, brown, and black are present. Dashes of red, white, and blue add to a mottled appearance.

Diameter of Lesion: Melanomas are usually greater than 1/4 inch (6mm) when diagnosed and growing.

Table 1: Feature Set and its Descriptions

	Feature Set	
	Description	**Dimension**
Asymmetry Shape	-Asymmetry Index -Lengthening Index	12
Border Edge	-Variance of the Border Rim Distance on Adaptive Extended Border Rim -Mean and Variance of Sobel Gradient -Mean and Variance of Laplacian of Gaussian Gradient	19
Color Variation	-Maximum, minimum and mean and intensity ratio on RGB, HSV, YCbCr and LabColor Spaces -Color Cooccurrence Counter on RGB Color Space	59
Differential Texture	-Maximum, Mean and Variance of Textural Variability -Contrast, Energy, Entropy and Homogeneity on Multidirectional Eight-Distances Color Cooccurrence Matrices	19

To check for the degree of symmetry, there are two values of asymmetry feature i.e. Asymmetry Index (AI) and Lengthening Index (LI). The major axis L1 of the lesion is aligned with its longest diameter, passing through its center; the minor axis L2 is orthogonal to L1 and passes through the shape center. Convex Hull: Line enclosing a set of points in a plane with no concavities.

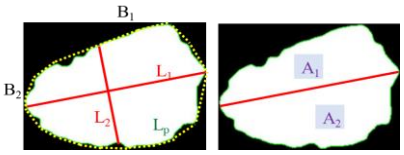

Figure 7: Segmented Lesion and its Parameters

- Lesion Binary Area (L_A)
- Major (L_1) and Minor Axis (L_2) (red solid),
- Convex Hull Area (L_C) (yellow dashed),
- Lesion Perimeter: L_P (green solid),
- Areas around axis (A_1, A_2) and
- Bounding box width (B_1), height (B_2)

Texture analysis quantify by terms such as rough, smooth, silky, or bumpy as a function of the spatial variation in pixel intensities. To characterize the lesion border irregularity, the segmented border rim is firstly dilated by octagon shaped structuring element and one-third portion of the segmented image is used. All the border features are quantified by the average and variance magnitudes of the gradient and GLCM feature at each pixel from the extended border rim. Sobel edge detectors are designed to respond maximally to edges running vertically and horizontally relative to the pixel grid, one kernel for each of the two perpendicular orientations.

Figure 8: Extended Lesion Rim

5. Classification

This step is to classify a lesion either benign (normal) or malignant (cancerous) using bagging decision tree ensemble classifier. The main principle is that a group of weak learners are predicted multiple times to form a strong learner to achieve better prediction performance.

- Model: Bag Ensemble
- Learner Type: Decision Tree
- Number of Bags: 30
- Maximum no of splits: size of training samples – 1
- N: total number of training samples
- n_1, n_2, … n_{30}: number of samples in each bag (less than N)
- T_1, T_2, … T_n: Randomly selected training subset samples

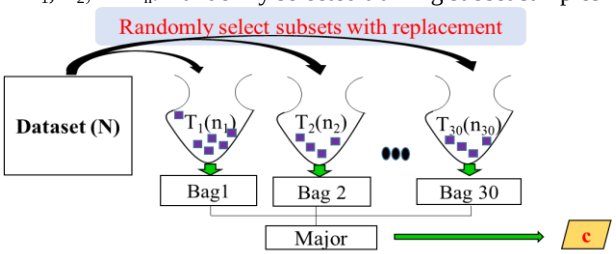

Figure 9: Sample Bag Ensemble Model

5.1. Imbalanced Dataset Classification

Real world datasets treat imbalanced data which contains inequivalent observations in one class. For example, normal observation samples are generally more than cancer patient observation samples in medical datasets. Data sampling is a statistical analysis technique used to analyze a representative subset of data to identify patterns and trends in the large or unbalanced dataset to produce accurate findings.

The common solutions are: *Undersampling* the majority class: to randomly select n samples from the majority class (where n is in approximately equal size with minor class), and use only those as training data, combined with all samples from the minority class. *Oversampling* the minority class: to re-sample the minority class, i.e. to augment data which is exactly same as what it has already.

Image augmentation involves creating transformed versions of images in the training dataset that belong to

the same class as the original image. Transforms include arrange of operations from the field of image manipulation, such as shifts, flips, zooms, rotations, and much more. In this work, image transformations: flipping (horizontal, vertical, both) rotating (45, 135, 225, 315) degrees are used to augment more samples of minority class. Then, classification is performed to original, undersampled and oversampled data.

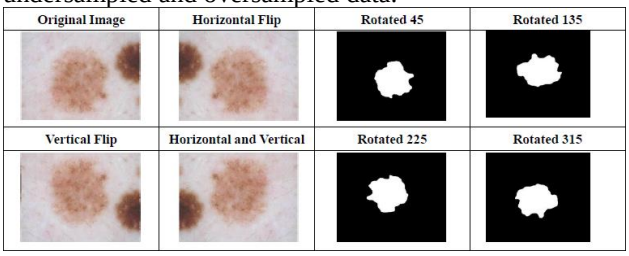

Figure 10: Transforming Images of Minority Class for Image Augmentation

In undersampling, samples of majority class are chosen by means of high Jaccard index score of segmentation to get the equal samples of classes. In oversampling, samples of minority class are also chosen by means of high Jaccard index score of segmentation first, and the image augmentation is performed to those samples.

6. Evaluation Measures and Performances

Evaluation criteria for segmentation are measured by means of the five common segmentation metrics while evaluation criteria for classification are measured by means of the three common classification metrics.

6.1. Segmentation Measures

Evaluation criteria for segmentation are Accuracy, Sensitivity, Specificity, Dice Coefficient, and Jaccard Index.

Pixel-level accuracy (AC): It computes the number of correct predictions with respect to the total number of samples.

Dice Coefficient /F1-Score (DC): This metric is the harmonic mean of precision and sensitivity and computed by comparing the pixel-wise agreement between the ground truth and its corresponding predicted segmentation.

Jaccard Index (JI): It is known as the Jaccard similarity coefficient to compare predictions with the ground truth to see which samples are shared and which are distinct.

Pixel-level sensitivity/Recall (SE): It is also known as recall, computed as the fraction of true positives rate that are correctly identified.

Pixel-level specificity (SP): It is computed as the fraction of true negatives rate that are correctly identified.

where TP, TN, FP, FN, refer to true positive (Pigmented lesion segmented as pigmented lesion), true negative (Normal skin segmented as normal skin), false positive (Normal skin segmented as pigmented lesion), and false negatives (Pigmented lesion segmented as normal skin), at the pixel level respectively.

6.2. Segmentation Performances

The following table shows the performance results of the segmentation part.

Table : Segmentation Experimental Results

	Datasets		
	PH2	**ISBI2016**	**ISIC2017**
AC (%)	90.7257	91.9049	90.7257
DC (%)	85.8833	81.8054	75.0683
JI (%)	75.2592	69.2125	60.0874
SE (%)	86.0499	73.6116	83.3702
SP (%)	93.0069	97.9126	91.7648
second	0.0202	0.0214	0.0218

6.3. Classification Measures

Classification evaluation parameters are:

Accuracy= (TP+TN)/(TP+TN+FP+FN)

Sensitivity(Recall)= TP/(TP+FN)

Specificity= TN/(TN+FP)

where TP, TN, FP, FN, refer to true positive (Melanoma classified as melanoma), true negative (Non-melanoma classified as non-melanoma), false positive (Non-melanoma classified as melanoma), and false negatives (Melanoma classified as non-melanoma) respectively.

6.4. Classification Performance

Classification performance results are as follows:

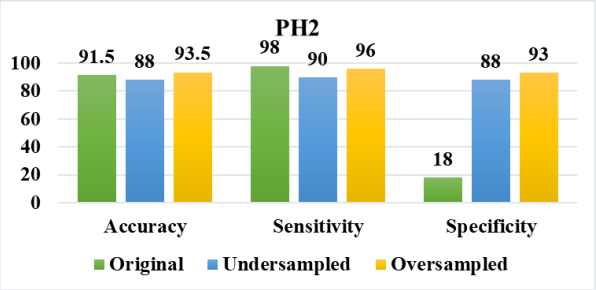

Figure 11: Classification Results of PH2 dataset

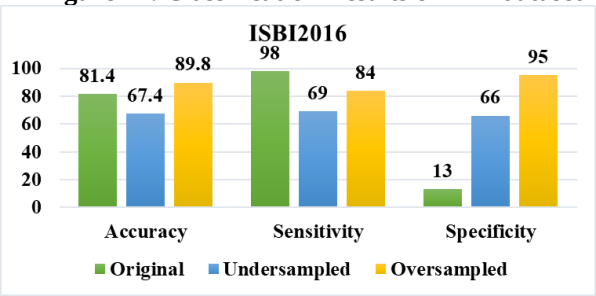

Figure 12: Classification Results of ISBI2016 dataset

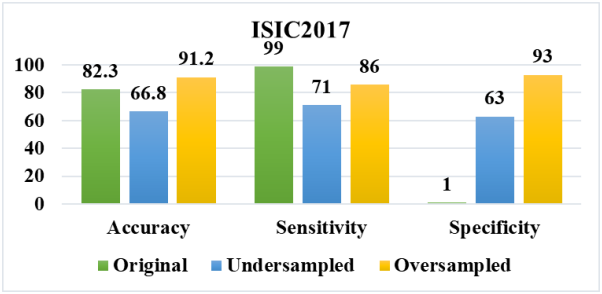

Figure 13: Classification Results of ISIC2017 dataset

The classification results show that the skin lesion images can be classified with promising accuracy, sensitivity and specificity with the use of ensemble classifier by image oversampling.

7. Conclusion and Discussion

Early detection of melanoma skin cancer is advantageous to patients, by which patients can identify the skin cancer without going to hospital or without the help of a doctor. Melanoma detection in dermoscopy images is a challenging task due to the existence of artifacts, the low contrast of skin lesions, the high degree of color variability in melanoma skin lesion images and the huge intra-structural variation. In this study, a model for the detection of malignant melanoma skin cancer is proposed and experiments are conducted on the PH2, ISBI2016 and ISIC2017 benchmarking datasets.

The segmentation algorithm addresses the following challenges to some extent: the low contrast challenge using intensity adjustment step in segmentation, the color variability challenge by the extraction of color features from various color spaces, the existence of hair and ruler marks artifacts in the preprocessing step

The proposed approach is not difficult to implement and requires no additional changes in the skin parameters because of unsupervised nature. It also provides good classification results with 93.5% accuracy, 96% sensitivity and 93% specificity for PH2 dataset, 89.8% accuracy, 84% sensitivity and 95% specificity for ISBI2016 and 91.2% accuracy, 86% sensitivity and 93% specificity for ISIC2017 dataset respectively.

Overall, the experimental results clearly demonstrate that a bag tree ensemble classifier to the balanced oversampled datasets provides effective by addressing the class imbalance dataset nature. It can be concluded that the proposed system can be effectively used by patients and medical experts to diagnose.

12. References

[1] C. M, “Border detection in dermoscopy images using statistical region merging”, 2008 The Authors Journal compilation, Skin Research and Technology.

[2] B. Bozorgtabar1, M. Abedini1, and R. Garnavi, “Sparse Coding Based Skin Lesion Segmentation using Dynamic Rule-based Refinement”, Conference Paper in Lecture Notes in Computer Science October 2016.

[3] E Celebi, M., Wen, Q., Hwang, S., Iyatomi, H. Schaefer, G., 2013. “Lesion border detection in dermoscopy images using ensembles of thresholding methods”. Skin Research and Technology, 19(1), Pp. e252-e258.

[4] J.C. Kavitha, Suruliandi A and Nagarajan D, “Melanoma Detection in Dermoscopic Images using Global and Local Feature Extraction”, International Journal of Multimedia and Ubiquitous Engineering, Vol.12, No.5 (2017), pp.19-28.

[5] O. Abuzaghleh, B. D. Barkana and M. Faezipour, “Automated Skin Lesion Analysis Based on Color and Shape Geometry Feature Set for Melanoma Early Detection and Prevention”, 978-1-4577-1343-9/12/$26.00 ©2014 IEEE.

[6] Khalid Eltayef, Yongmin Li and Xiaohui Liu, “Knowledge Transfer for Melanoma Screening with Deep Learning”, 2016 International Conference on Communication, Image and Signal Processing (CCISP 2016).

[7] Lequan Yu, Hao Chen, Qi Dou, Jing Qin, and Pheng-Ann Heng, “Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks”, IEEE Transactions on Medical Imaging, DOI 10.1109, April 2016.

[8] Masood, A. and Ali Al-Jumaily, A., 2013. “Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms”. International journal of biomedical imaging, 2013.

[9] Mendonça, T., Ferreira, P.M., Marques, J.S., Marcal, A.R. and Rozeira, J., 2013, July. “PH 2- A dermoscopic image database for research and benchmarking”. In 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC) (pp. 5437-5440). IEEE.

[10] N.C. Lynn, Z.M. Kyu, “Segmentation and Classification of Skin Cancer Melanoma from Skin Lesion Images”, The 18th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT’17) in Taipei Taiwan, on December 18-20, 2017.

[11] N.C. Lynn, N. War, “Segmentation of Skin Lesion towards Melanoma Skin Cancer Classification”, International Journal of Computer Science and Network, Volume 8, Issue 3, June 2019, ISSN (Online): 2277-5420.

[12] Schmid-Saugeona, P., Guillodb, J. and Thirana, J.P., 2003. “Towards a computer-aided diagnosis system for pigmented skin lesions”. Computerized Medical Imaging and Graphics, 27(1), pp.65-78.