Journal of AL-Qadisiyah for computer science and mathematics
Vol.7 No.1 Year 2015

Hind.R

Page 160-172

Boundaries Object Detection for Skin Cancer Image using Gray-Level Co-Occurrence Matrix (GLCM) and features points

Professor Hind Rustum Mohammed
Computer Department
Faculty of Computer Science and Mathematics
Kufa University, Iraq
hindrustum.shaaban@uokufa.edu.iq

Recived :26\3\2014 **Revised : 18\12\2014** **Accepted :11\1\2015**

Abstract

In the present paper, boundaries object detection(BOD) for skin cancer image using connected components is proposed. We propose connected components algorithm which that capable of Segment with Extraction of connected boundaries for skin cancer image segmentation . The algorithm is proposed to create a color label image using the local features e points in skin cancer as objects image . A new boundaries object detection(BOD) technique is introduced based on the gray-level co-occurrence matrix (GLCM). GLCM represents the distributions of the intensities and the information about relative positions of neighboring pixels of an image.

Database consists of 120 images (60 image for satisfactory skin cancer ,60 image for unsatisfactory skin),types for image .jpg , .png and .bmp image formats. database prepared in our conditions ,images obtained from in Al-Seder Hospital(30 image for satisfactory skin cancer ,30 image for unsatisfactory skin), other images obtained from internet(30 image for satisfactory skin cancer ,30 image for unsatisfactory skin).

Training stage consists of 80 images(20 image for satisfactory skin cancer ,20 image for unsatisfactory skin cancer) from in Al-Seder Hospital and(20 image for satisfactory skin cancer ,20 image for unsatisfactory skin cancer) from internet. Testing stage consists of 40 images(10 image for satisfactory skin cancer ,10 image for unsatisfactory skin cancer) from in Al-Seder Hospital and(10 image for satisfactory skin cancer ,10 image for unsatisfactory skin cancer) from internet.

The performance of object detection with Connected components which are surround influence . The proposed scheme can serve as a easy preprocessing for high level tasks such shape based recognition and image retrieval. The experimental results confirm the effectiveness of the proposed algorithm.

Keywords
Skin cancer image, image segmentation, Object detection, Extraction of connected boundaries ,Connected components .

Journal of AL-Qadisiyah for computer science and mathematics
Vol.7 No.1 Year 2015

Hind.R

الخلاصة

تم الكشف عن حدود كل كائن في صور سرطان الجلد باستخدام مكونات متصلة. تبين الطريقة المقترحة خوارزمية المكونات المتصلة التي تكون قادرة علي استخراج حدود متصلة لصور سرطان الجلد من خلال تقطيع الصورة. وتقترح الخوارزمية تكوين الاجزاء المتقطعة من خلال اللون باستخدام النقاط وميزات تفصيلات المحلية في صور سرطان الجلد والكائنات المكونه الصورة. تستند تقنية كشف الحدود للكائن (BOD) إلى مصفوفة توارد على مستوى الرمادي (GLCM). يمثل GLCM توزيعات شدة والمعلومات حول المواقع النسبية للبكسل المجاورة للصورة.
تتكون قاعدة البيانات من ١٢٠ صور (٦٠ صورة لسرطان الجلد تبين حالات مرضية، ٦٠ صورة للجلد لغير المرضي)، وكانت أنواع تنسيق الصور jpg png.. وBMP.
تم الحصول على قاعدة البيانات من مستشفى الصدر (٣٠ صورة لسرطان الجلد للحالات مرضية، ٣٠ صورة للجلد غير المرضي)، والصور الأخرى التي تم الحصول عليها من الإنترنت (٣٠ صورة لسرطان الجلد مرضية، ٣٠ صورة للجلد حالات غير مرضية) .
وتتكون مرحلة التدريب من ٨٠ صور (٢٠ صورة لسرطان الجلد حالات مرضية، ٢٠ صورة للجلد حالات غير مرضية) من مستشفى الصدر (٢٠ صورة لسرطان الجلد حالات مرضية، ٢٠ صورة الجلد للحالات غير مرضية) من الإنترنت. وتتكون مرحلة الاختبار من ٤٠ صور (١٠ صورة لسرطان الجلد مرضية، ١٠ صورة لسرطان الجلد غير مرضية) من مستشفى الصدر و (١٠ صورة لسرطان الجلد حالات مرضية، ١٠ صورة لسرطان للجلد حالات غير مرضية) من الإنترنت.
يمكن أن تكون الخطة المقترحة بمثابة معالجة ابتدائية بسيطة لمهام عالية المستوى مثل الشكل لتمييز واسترجاع الصور. تؤكد النتائج التجريبية فعالية الخوارزمية المقترحة.

كلمات
صورة سرطان الجلد ، تقطيع الصورة، كشف الوجوه ، استخراج حدود متصلة، مكونات متصلة.

1- Introduction

Skin has the advantage of being non-sensitive to directions, so we can separate skin regions other parts of the color images and segment face regions accurately through post-processing. The application of color can provide valuable candidate region when detecting stationary targets[1].

Skin color and textures are important cues that people use consciously or unconsciously to infer variety of culture-related aspects about each other. Skin color and texture can be an indication of race, health, age, wealth, beauty, etc. [2]. Skin detection is one of the basic subjects in image processing. In many cases such as human detection and tracking, visual identification and face detection, a skin detection stage is needed. The concept of "skin detection" in an image is the classification of the existence pixels in that image into two skin and Non-skin classes. In this direction, several methods have been presented until now. In most of the proposed methods, researchers have tried to define and extract a feature vector for each pixel of image and in the end, classify the feature vectors[3].

Most of the currently applied anti-cancer agents do not greatly differentiate between cancerous and normal cells. In addition cancer is often diagnosed and treated too late, when the cancer cells have already invaded and metastasized into other parts of the body. At the time of clinical presentation, a great percentage of patients with breast, lung, colon, prostate, and ovarian cancer have hidden and over metastatic colonies[4].

Therefore image processing become our choice for an early detection of the skin cancer, as it is non-expensive technique. The identification of the edges of an object in an image scene is an important aspect of the human visual system because it provides information on the basic topology of the object from which an interpretative match can be achieved. In other words, the segmentation of an image into a complex of edges is a useful prerequisite for object identification. However, although many low-level processing methods can be applied for this purpose, the problem is to decide which object boundary each pixel in an image falls within and which high level constraints are necessary[5].

2.Related work

Several methods have been proposed for boundaries object detection segmenting images; this part related studies carried out in this research topic. In [10], suggested unconstrained object segmentation system using Saliency map, in [11] suggested method used Gabor filter and Grab cut, the goal is to generate an initial rectangle automatically for Grab cut. In order to create initial rectangle, author uses Gabor filter and Saliency map and then they uses 4 features(amount of area, variance, amount of class with boundary area , amount of class with saliency map) to categorize foreground and background.

In [12], suggested method used an object segmentation method based on saliency map, Mean shift and level set method. First, a histogram based contrast method is used to generate the saliency map of the input image. Second, the input image is segmented into clusters using Mean shift. Used the saliency map, segmented clusters are classified into background and foreground clusters. After that, an initial contour for level set method is determined by applying morphological erosion on foreground clusters.

In [13], suggested method flower segmentation technic by a level set based on GMM initialization. First, a GMM [11] is applied to the input image to roughly specify flower area. Then a level set searches for the exact flower area based on the initial contour provided by the GMM.

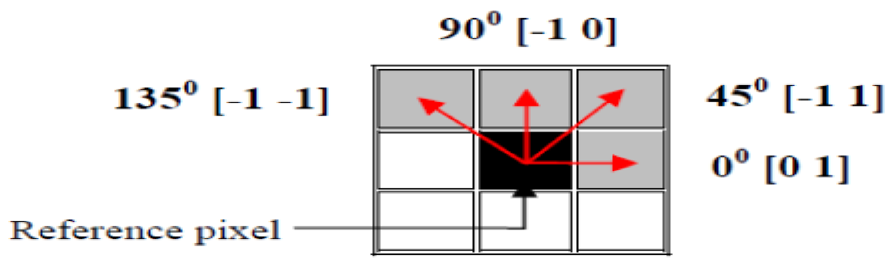
In [14] object detection for RGB-D images used contrast, focusses on instances, but the typical output is a bounding box.

In[15] suggested the problems models are initialization and poor convergence to deep concavities of the boundary and particularly equilibrium issues in the vector field. In the proposed method first the ISAF field is generated which is composed of swirling and attracting component (DETF & DEPF respectively). Then with the help of dynamical system limits cycles are set / located by Newton Raphson algorithm in the Poincare section. And integral equations are used to connect these boundaries .

3. Gray-Level Co-Occurrence Matrix (GLCM)

Grey Level Co-occurrence Matrices (GLCM) are one of the earliest techniques used for image texture analysis. The level up to which the subdivision is carried out depends on the problem being solved. A human skin color model is used to decide either a pixel is skin color or non-skin-color[6]. Texture is an important characteristics used in identifying regions of interest in an image. Grey Level Co-occurrence Matrices (GLCM) is one of the earliest methods for texture feature extraction proposed by Haralick et.al. back in 1973. Since then it has been widely used in many texture analysis applications and remained to be an important feature extraction method inthe domain of texture analysis[7].

The details of the process to generate four symmetrical co-occurrence matrices considering a 4×4 image represented with four gray-tone values from 0 to 3. For the purpose we considered one neighboring pixel (d=1) along four possible directions as {[0 1] for 0^0; [-1 1] for 45^0; [-1 0] for 900 and [-1 -1] for 135^0]}[8].



Figure(1) Co-occurrence matrix directions for extracting texture features

The quantization level is an equally important consideration for determining the co-occurrence texture features. Also, neighboring co-occurrence matrix elements are highly correlated as they are measures of similar image qualities .Statistics applied to co-occurrence probabilities are [9]:

1- Contrast: It is the difference between the highest and the lowest values of a contiguous set of pixels. It measures the amount of local variations present in the image. A low contrast image presents GLCM concentration term around the principal diagonal and features low spatial frequencies.

Contrast=$\sum_i \sum_j (i-j)2\ gij$ ………..…..……………………..(1)

2- Correlation :Measures the joint probability occurrence of the specified pixel pairs**.**

The correlation feature is a measure of gray tone linear dependencies in the image.

$$\text{Correlation}=\frac{\sum \sum (ij)gij-\partial(ij)gij-\mu x\mu y}{\sigma x\sigma y}2 \quad \dots\dots\dots\dots\dots\dots.(2)$$

where μ_x, μ_y, σ_x and σ_y are the means and standard deviations of g_x and g_y

3-Energy: Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment It measures the textural uniformity that is pixel pair repetitions. It detects disorders in textures. Energy reaches a maximum value equal to one.

$$\text{Energy}=\sum_i \sum_j gij2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Where gij = (i,j) th entry in GLCM , $g_x(i)$ = i th entry in marginal probability matrix obtained by summing rows of $gij = \sum_{j=1}^{Ng} g(i,j)$

4- Homogeneity :It measures image homogeneity as it assumes larger values for smaller gray tone differences in pair elements. It is more sensitive to the presence of near diagonal elements in the GLCM. It has maximum value when all elements in the image are same. GLCM contrast and homogeneity are strongly, but inversely, correlated in terms of equivalent distribution in the pixel pairs population. It means homogeneity decreases if contrast increases while energy is kept constant.

$$\text{Homogeneity}=\sum_i \sum_j \frac{1}{1+(i-j)2} \text{ gij} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

The use of these four properties(contrast ,correlation, energy and homogeneity) because it helped the researcher to identify and detect the boundaries of the object in images.

In the next step we find connected components in binary Skin image .The basic steps in finding the connected components are:

1- Search for the next unlabeled pixel, p.

2- Use a flood-fill algorithm to label all the pixels in the connected component containing p.

2- Repeat steps 1 and 2 until all the pixels are labeled

The four structure field for components are :

1-Connectivity: Connectivity of the connected components (objects)

2-ImageSize: Size of image

3-NumObjects: Number of connected components (objects) in image

4-PixelIdxList: 1-by-NumObjects cell array where the k^{th} element in the cell array is a vector containing the linear indices of the pixels in the k^{th} object.

Where

- Contrast for image Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image. Range = [0 (size(GLCM,1)-1)^2] Contrast is 0 for a constant image where GLCM =graycomatrix in matlab .
- Correlation image Returns a measure of how correlated a pixel is to its neighbor over the whole image. Range = [-1 1], Correlation is 1 or -1 for a perfectly positively or negatively correlated image.
- Energy image returns the sum of squared elements in the GLCM. Range = [0 1] Energy is 1 for a constant image.
- Homogeneity Range = [0 1] Homogeneity is 1 for a diagonal GLCM

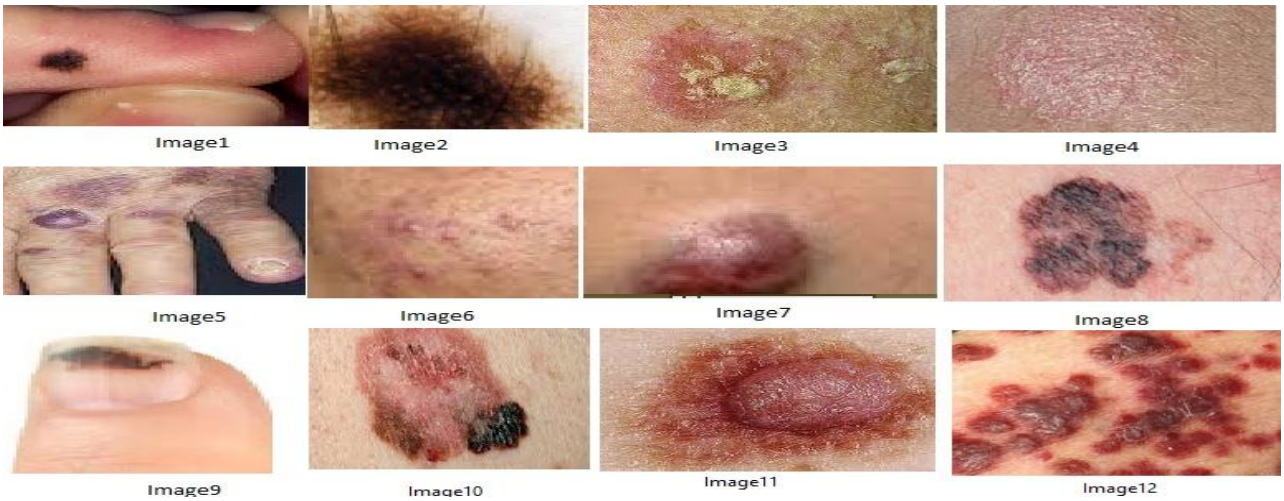
4.System Design and Experimental Results

Skin cancer is a disease in which cancer (malignant) cells are found in the outer layers of the skin. The skin protects the body against heat, light, infection, and injury. The skin has two main layers and several kinds of cells. The top layer of skin is called the epidermis. It contains three kinds of cells: flat, scaly cells on the surface called squamous cells; round cells called basal cells; and cells called melanocytes, which give the skin its color[6]. Three processes are done in a sliding window where size is already defined, which is 3x3. For that reason, the original input image matrix must be added with one pixel width of pixel on each side, so that the output of the pixels at the edge of the original image can be calculated.

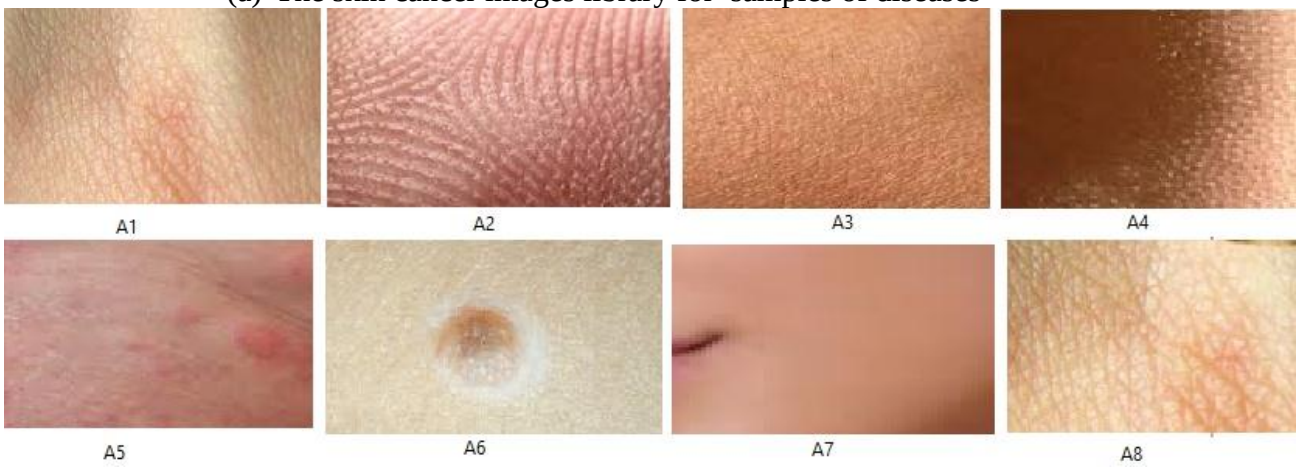
The steps for Algorithm work are:
1- Start
2- Input skin cancer images
3- Statistics applied to co-occurrence probabilities
4-Find features e points for skin cancer images contains(point number, e coordinates(x,y), point direction).
5- Comparison between the images by find relationship between extracted points.
6- Draw lines connected components
7- Extraction of connected boundaries to segment the skin cancer image.
8- Find Boundaries object detection for skin cancer image
9-End

An algorithm is capable of Segment with Extraction of connected boundaries for skin cancer image segmentation has been presented. We have used two types skin mages databases(consists of 120 images):

(1) database prepared in our conditions ,images obtained from in Al-Seder Hospital(30 image for satisfactory skin cancer ,30 image for unsatisfactory skin).
(2) Skin database [16,17] and some other images obtained from internet(30 image for satisfactory skin cancer ,30 image for unsatisfactory skin).



(a) The skin cancer images library for samples of diseases



(b)

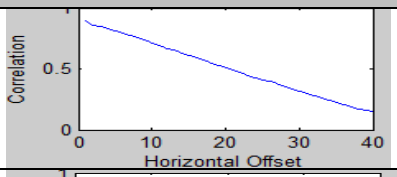
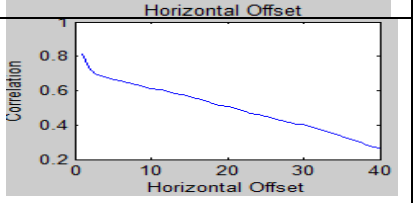
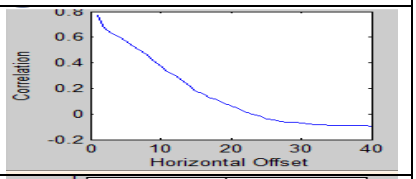
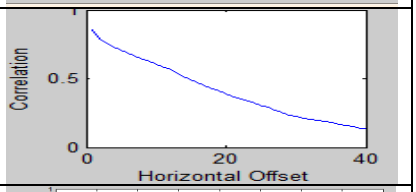
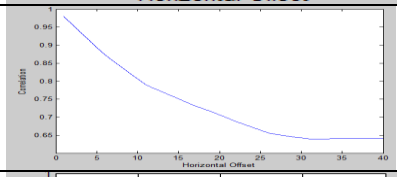
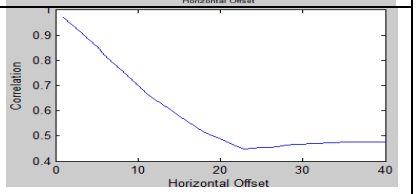
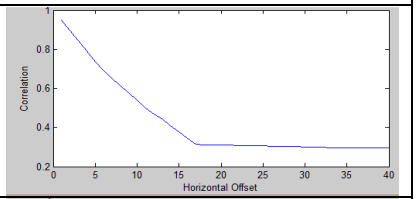
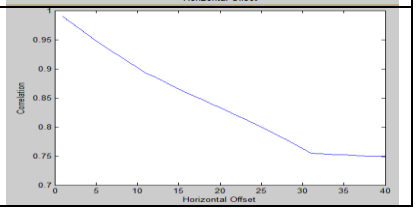
Figure(2): (a)The skin cancer images library for samples of diseases ,(b)The skin images library samples

Table(1) showed Statistics applied to co-occurrence probabilities for sample images and Table(2) Plot correlation as a function of offset for sample images.

Table(1) : Statistics applied to co-occurrence probabilities for sample images

index	Image	Contrast	Correlation	Energy	Homogeneity
1	Image 10	0.3298	0.9270	0.1646	0.9141
2	Image 3	0.1558	0.8275	0.3045	0.9223
3	Image 8	0.0528	0.8105	0.7824	0.9736
4	Image 11	0.0889	0.8611	0.4754	0.9556

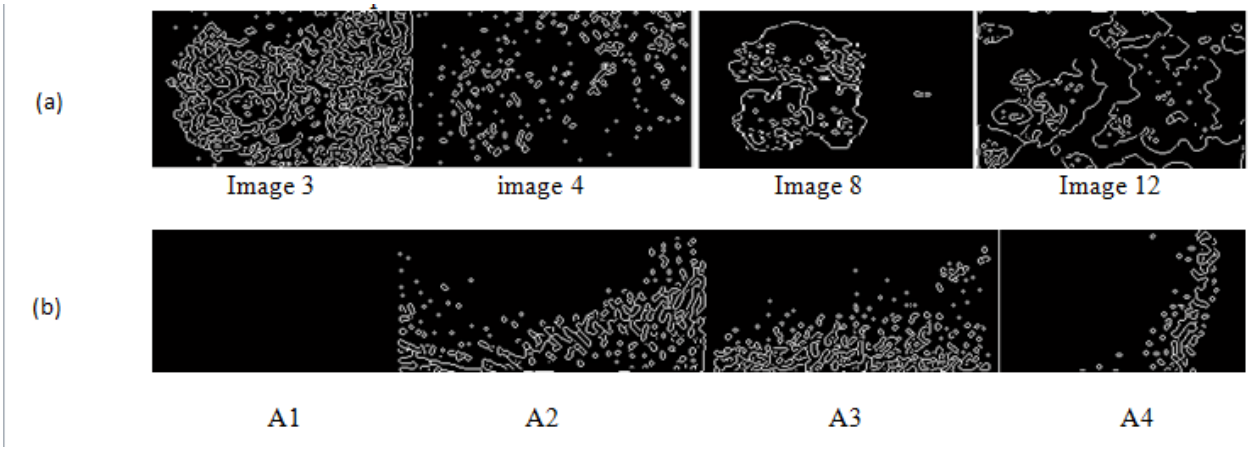
Table(2) Plot correlation as a function of offset

index	Image	
1	Image 10	
2	Image 3	
3	Image 8	
4	Image 11	
5	A1	
6	A2	
7	A3	
8	A4	

For extraction of connected components let **M** represent for all extraction point and **Y** represent a connected component contained in a set **A** and assume that a point p of **Y** is known. Then the following iterative expression yields all the elements of Y

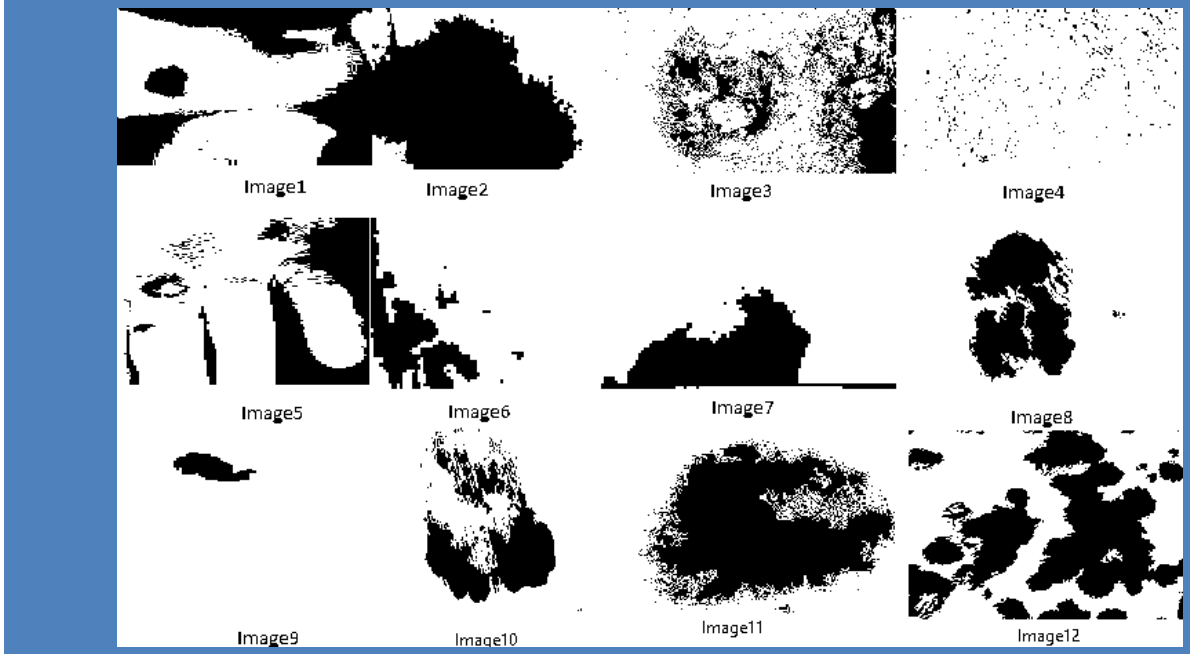
$$Xk = (x_{k-1} \oplus N) \cap x_{y^2} M_k \quad \text{.....(5)}$$

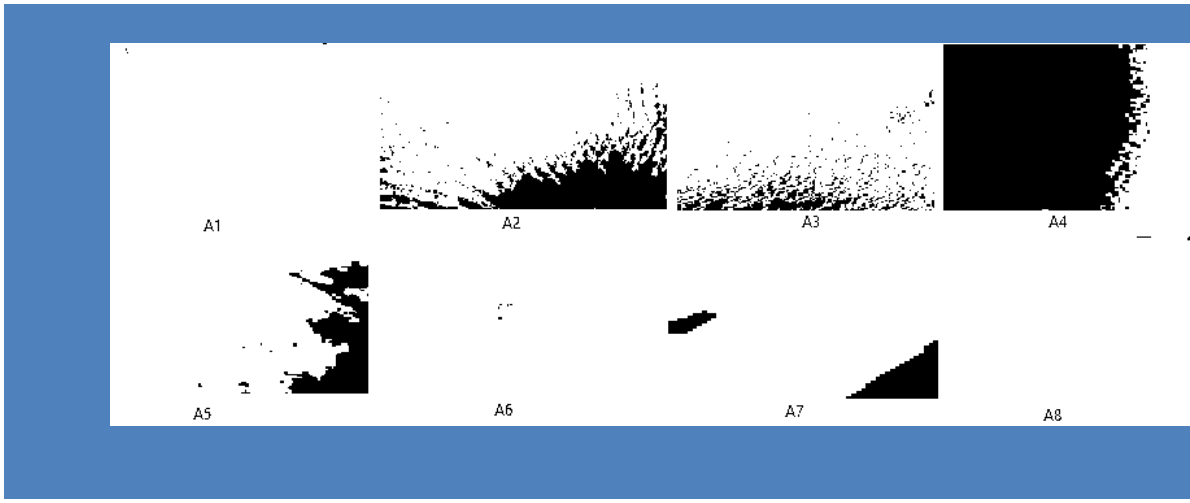
where k= 1, 2, 3,……,n and X0 = p, and N is a suitable structuring element. If Xk = Xk−1, the algorithm has converged and we let Y = Xk. This algorithm is applicable to any finite number of sets of connected components contained in A, assuming that a point is known in each connected component.



Figure(3): (a) Edge for skin cancer images library for samples of diseases
(b) Edge for skin images library samples

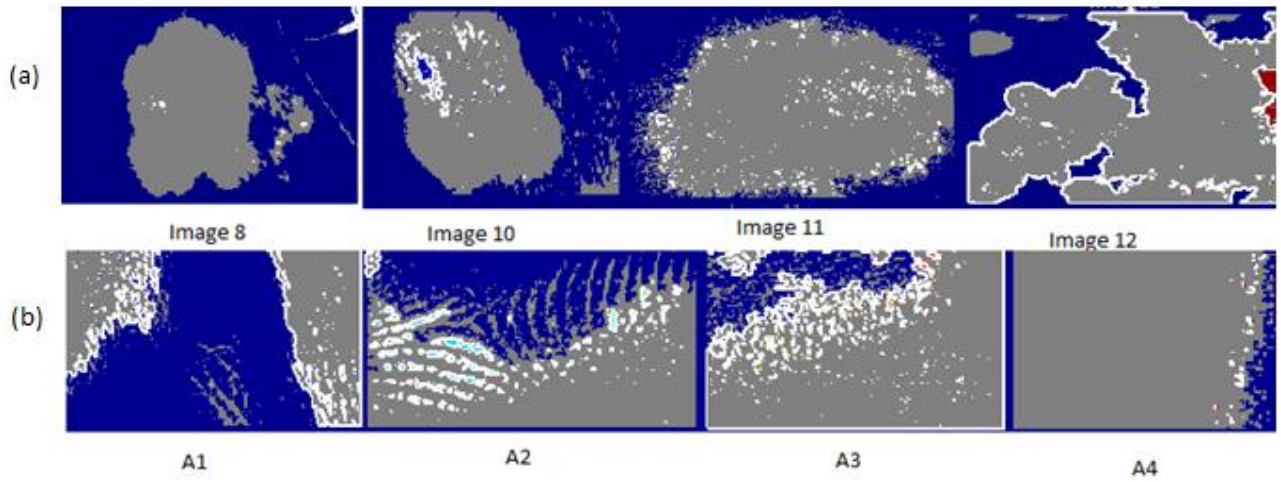
To Remove small objects from image based two-dimensional connectivity 4-connected neighborhood and 8-connected neighborhood for each matrix will be 256x256 for an 8 bit image. Quantizing the image will result in smaller matrices.





Figure(4):Objects for all region in
(a) skin cancer images library samples ,(b) skin images library samples

The main strategy is to keep object regions from merging and merge background regions as many as possible. In this stage we merge background regions with their adjacent region in skin images.



Figure(5): Boundaries Object Detection for (a)skin cancer images samples and (b)skin images samples

5. Conclusion

We propose new method using the proposed Gray-Level Co-Occurrence Matrix (GLCM) and features e points are capable of Segment with Extraction of connected boundaries for Skin cancer image segmentation . We have proposed an automatic scalable object boundary detection algorithm based on edge detection, and region growing techniques. Using smaller number of gray levels (bins) shrinks the size of GLCM which reduces the computational cost of the algorithm and at the same time preserves the high boundaries object detection rates. The algorithm higher accuracy can be achieved by either increasing number of samples per class(points features) in the training phase or considering the of GLCM features.

References

[1] BaozhuWang, Xiuying Chang, Cuixiang Liu ,"Skin Detection and Segmentation of Human Face in Color Images", International Journal of Intelligent Engineering and Systems, Vol.4, No.1, 2011.

[2] Bernhard Fink, K.G., Matts, P.J.," Visible skin color distribution plays a role in the perception of age", attractiveness, and health in female faces. Evolution and Human Behavior 27(6), 2006.

[3] Shervan Fekri-Ershad, Mohammad Saberi ,and Farshad Tajeripour,"Using colors based image retrieval technique", The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.3 ,2012.

[4] Menon U, Jacobs IJ. ," Recent developments in ovarian cancer Screening", Curr. Opin. Obstet. Gynecol. 12,2000.
[5] Mahmoud Elgamal," Automatic skin cancer images classification", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 3, 2013.

[6] T. Sawangsri, V. Patanavijit, and S. Jitapunkul ,"Face Segmentation Using Novel Skin-Color Map And Morphological Technique", World Academy of Science,2003.

[7] R. Haralick, K. Shanmugam, and I. Dinstein "Textural Features for Image Classification",
IEEE Trans. on Systems, Man and Cybernetics, SMC–3(6):pp 610–621, , 1973.

[8] Biswajit Pathak, Debajyoti Barooah, "Texture analysis based on the gray-level co-occurrence matrix considering possible orientations", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 2, Issue 9, 2013.

[9] Dhanashree gadkari, " Image Quality analysis using GLCM",Master of Science in Modeling and Simulation in the College of Arts and Sciences, University of Central Florida Orlando, Florida,2000.

[10] I.-S. Na, K.-H. Oh and S.-H. Kim, "Unconstrained Object Segmentation Using GrabCut Based on Automatic Generation of Initial Boundary", International Journal of Contents, vol. 9, no. 1, 2013.

[11] S.-H. Lee and J. Keun Seo, "Level set-based bimodal segmentation with stationary global minimum", Image Processing, IEEE Transactions, Vol. 15, 2005.

[12] H. Le, S.-H. Kim and I.-S. Na, "Object Segmentation Based on Adaptive Contour Initialization for Level Set Method", Proc. IEEE Int. Symp. Signal Processing and Information Technology 2012 (CD-Pub), Hochiminh City, Vietnam, 2012.

[13] S. C. Park, S.-H. Kim, G. S. Lee, C. W. Lee, I. S. Na, A. V. Ha Le and D. Yen, "Flower Area Segmentation by a Level Set based on GMM Initialization", Proc. KSII 3rd Int. Conf. Internet (ICONI), Sepang, Malaysia, 2011.

[14] Soo Kim, B., Xu, S., Savarese, S.: Accurate localization of 3d objects from RGB-D data using segmentation hypotheses. In: CVPR) ,2013.(

[15] Ruchita V. Indulkar, Sanjay D. Jondhale, Boundary Extraction Using Poincare Map methods International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, Special Issue 4, 2014.

http://guides.library.stonybrook.edu/medicineclinical/images[16]
(Stony Brook University libraries)

[17] http://mobihealthnews.com/17863/university-of-michigan-health-system-offers-skin-cancer-app/
(University of Michigan Health System)