Classification of Breast Tissue for Mammograms Images Using Intensity Histogram and Statistical Methods

Alyaa Hussain[*1], Alaa Noori Mazher[2], Asraa Razak[1]
[1]Department of physics, College of Science for women, University of Baghdad, Baghdad, Iraq.
[2]Department of computer and information System, University of Technology, Baghdad, Iraq.

Abstract:
In order to classify a image texture it must be segmented into a number of regions. each region have the same properties. a circular mask has been used. after the required region is obtained a intensity histogram and then as set of histogram features has been calculated in order to classify the required images.
The statistical features of a mammogram are extracted using image processing techniques. The extracted statistical features are the mean, standard deviation, smoothness, third moment uniformity and entropy for each of the histogram intensity and first order statistic a compare between the two results in order to obtain the best for classification the Breast tissue.
keyword: statistical feature, histogram equalization, histogram intensity.

تصنيف نسيج الثدي لصور Mammograms باستخدام histogram intensity والطرق الإحصائية

علياء حسين علي*[1], علاء نوري مزهر[2], اسراء رزاق سوادي[1]
[1]قسم الفيزياء, كلية العلوم للبنات, جامعة بغداد, بغداد, العراق
[2]قسم الحاسبات ونظم المعلومات, الجامعة التكنلوجيه, بغداد, العراق

الخلاصه:
لغرض تصنيف أي نسيج نحتاج إلى تقسيم نسيج الصورة الى عدد من المناطق. كل منطقه تحوي على نفس الخصائص. لقد استخدم القناع الدائري لهذا الغرض وبعد ما تم الحصول على المنطقة المطلوبة استخدمنا طريقة histogram intensity و طريقة إحصاء الدرجة الأولى ومنها تم حساب بعض الخصائص الإحصائية لغرض تصنيف المنطقة المطلوبة. الخصائص الإحصائية لصور Mammograms تم الحصول عليها باستخدام معالجة الصور الرقمية وهذه الخصائص هي mean, standard deviation, smoothness, third moment uniformity and entropy حيث تم حسابها لكل من histogram intensity و first order statistic حيث تمت مقارنة النتائج لغرض الحصول على أفضل نتائج لتصنيف نسيج الصورة.
الكلمة ألمفتاحيه: الخصائص الإحصائية, تسوية المنحني, شدة المنحني

Introduction:

Texture is defined as a pattern that is repeated and is represented on the surface of an object. segmentation is a fundamental low-level operation on images, in which homogeneous regions refers to a group of connected pixel in the image that share a common feature [1]. Texture can be observed in many natural images from multispectral satellite images to the microscopic images of cell or tissue samples or tissue samples. Texture is an innate property of virtually all surfaces it contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment. since the textural properties of images appear to carry useful information, for discriminating purpose features have always been calculated for textures[2]. The images segmentation refers to a process of assigning labels to a set of pixels or multiple regions[3]. medical image segmentation problems has been approached with several solution methods range of applicability such as fuzzy C. Breast cancer is the most common type of cancer in among women in the word mammography is regarded as an effective tool for early detection and diagnosis of breast cancer . there are various image texture analysis techniques for the detection of micro calcifications [4].In this work the first order statistics based on and intensity histogram is used for texture analysis.

Selection of Interest Region :

A practical methods that has chosen in this work is the used circular mask used to determine wither the texture pixel is edge or not depending on the neighboring rotating masks inside the running window, in horizontal, vertical, right, diagonal and left diagonal directions.

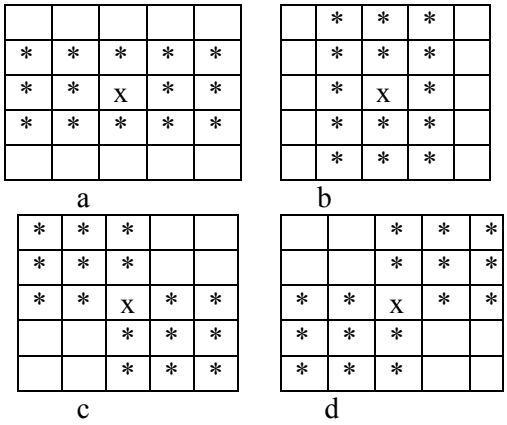


Figure 1-Rotating Window Contain Central Pixel 5x5

Algorithm [5]

1-start by 5×5 window reading the original image

2- A circular masks of 3×3 window that contain the central pixels of 5×5window of the images shown in fig(1)

3- The sum of gray levels of the circular masks are calculated

4-The difference between the two (RD, (LD) ,(V) and (H) directions circular rotating masks rare computed

5- In order to find the edge, maximum of point (4) the central pixel is replaced by 1 if and only if the adjacent differentiating masks maximum, mean, minimum gray level is greater than or equal to central pixel is replaced by 0.

6-Repeating step 1 to 5 by 5×5

Feature Extraction

Statistical texture analysis is based on statistical properties of intensity histogram without considering spatial dependence. The histogram of the image gives summary of the statistical information about the image. The first order statistical information of the image can be obtained using the image histogram of the image. the probability density of occurrence of the intensity ,as determined from the histogram with total number of pixels in the image is given as [6].

P(i)=h(i)/NM

where

i=0,1,2.....G-1

G= gray level tone of an image.

N=number of cell in horizontal domine

M= number of cell vertical domine

A useful features of the image can be obtained from the histogram, including mean is the average value of the intensity of the image. the variance tells the intensity variation around the mean, skewness tells the symmetries of the histogram around the mean, the kurtosis is the flatness of the histogram. the entropy represent the uniformity of the histogram

The following is the list of features obtained using the histogram[7]

mean: $\mu = \sum_{i=1}^{G-1} ip(i)$ 2

variance: $\sigma^2 = \sum_{i=1}^{G-1}(i-\mu)^2 p(i)$ 3

skewness: $\mu_3 = \sigma^{-3}\sum_{i=1}^{G-1}(i-\mu)^3 p(i)$ 4

kurtosis: $\mu_4 = \sigma^{-4}\sum_{i=1}^{G-1}(i-\mu)^4 p(i)$ 5

energy: $E=\sum_{i=1}^{G-1}(p(i))^2$ 6

entropy: $H=-\sum_{i=1}^{G-1} p(i)log_2[p(i)]$ 7

Table 1- Presents The First Order Statistical Features For Normal Images

images	mean	Stander deviation	variance	energy
1	87.2	19.51	380.64	84.99
2	28.3	24.96	623.00	79.41
3	34.6	32.12	1031.69	75.97
4	22.69	38.96	1517.88	74.41
5	45.78	70.4	4956.16	76.11

Table 2- Represent The First Order Statistical Features For Cancel Images

images	mean	Stander deviation	variance	energy
1	77.3	37.24	1386.81	84.54
2	16.4	28.54	814.53	73.59
3	24.5	37.95	1440.20	72.14
4	20.8	44.24	1957.17	73.55
5	33.6	27.87	776.73	75.51

Histogram Equalization

Histogram equalization method which applied for improving the appearance of poor image. It is a function similar to that of a histogram stretching but often provides more visually pleasing result across wide range of images. Histogram equalization is a technique where the histogram of the result image is as flat as possible. the result in a histogram with a mountain grouped closely together to "spreading or fluting histogram masks the dark pixels appear darker and the light pixels appear lighter, however, the pixels that are only slight lighter become much lighter become much lighter, then the darker pixels will appear darker[8].

however, the pixels that are only slight lighter become much lighter become much lighter, then the darker pixels will appear darker[8].

The steps for histogram equalization

1-greating a running sum of histogram values.

2- normalized by dividing by total number of pixels

3- multiply these values by maximum gray level

4- map the original values to the result obtained from step 3 by one to one correspondence. Some statistical features are obtained from histogram equalization see table(3)

the result which has been obtained from the histogram are:

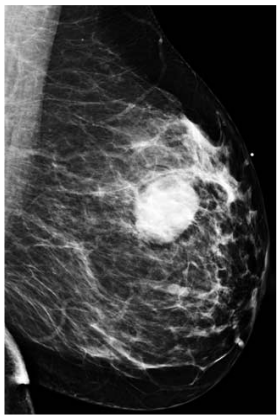
1- third moment which measure the uniformity in the histogram

2- Entropy which is the measure of randomness.

3-mean gives the measure of average intensity

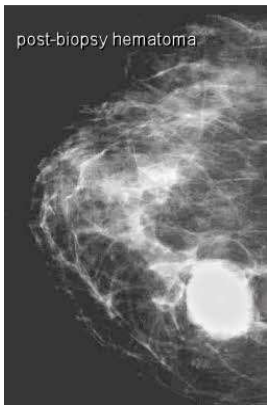
4- standard deviation gives the measure of the average contrast

The image Samples: The images samples has been taken from the internet which is a mammogram images for different cases, we have took five samples these samples have been segmented using the circular mask in order to calculate the first order statistical and the histogram intensity then different features are obtained from the histogram

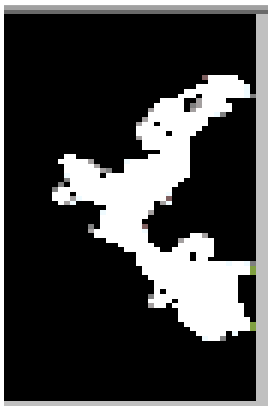


Sample 1

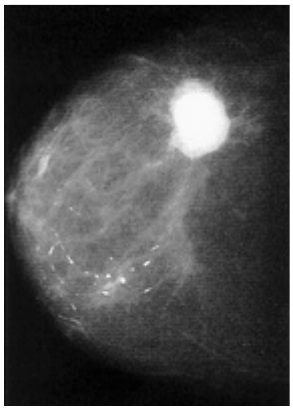
Segmented image



Sample2



Segmented image

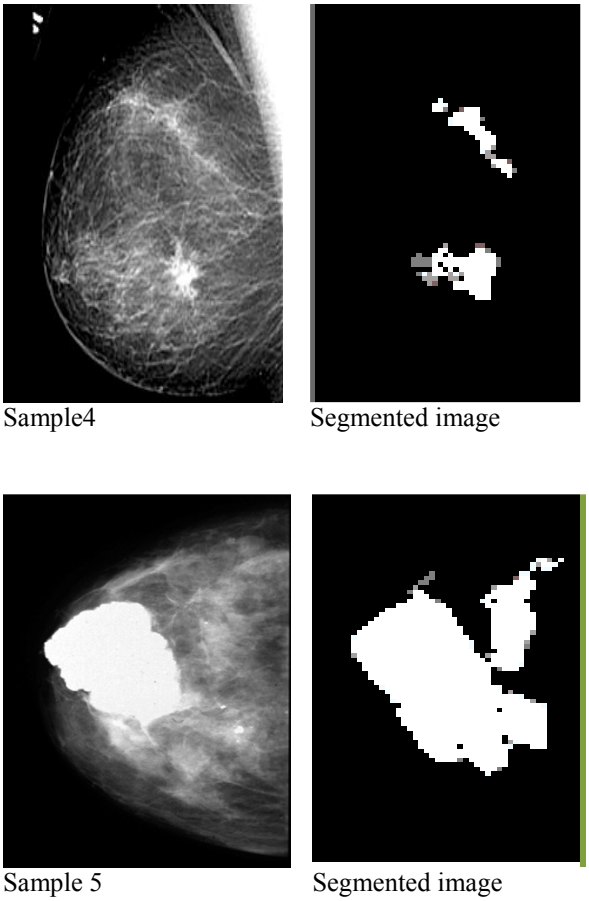


Sample3

Segmented image

Table 3-The Features Obtained From The Histogram Intensity

Images	Average intensity	Average contrast	Smoothness	Third moment	uniformity	Entropy
1	34.67	38.41	0.017	0.520	0.154	3.012
2	47.81	49.71	0.074	5.423	0.201	3.582
3	58.14	58.41	0.069	0.989	0.242	3.897
4	47.42	62.57	0.089	4.02	0.260	3.451
5	67.13	74.61	0.069	9.198	0.407	3.904



Figur 2- The Cancer Images And Its Segments Images

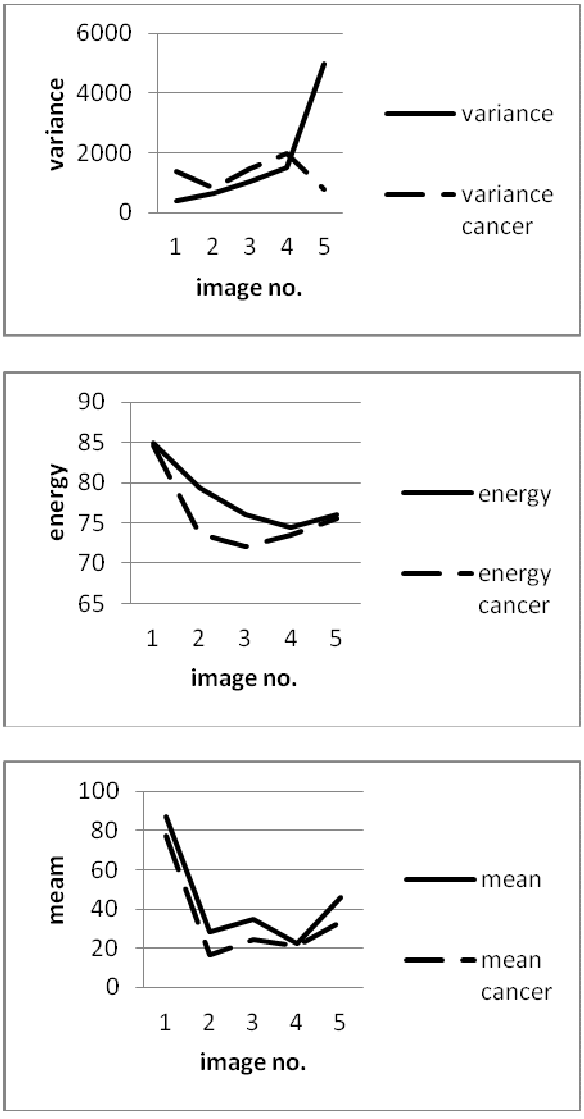


Figure 3- A Graph Between The No. Of Images and First Order Statistical Features For Cancer And Normal Images

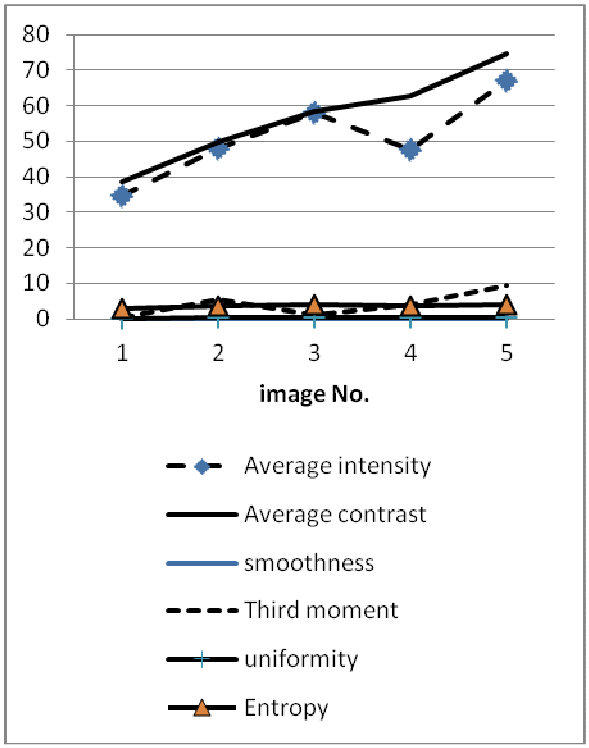


Figure 4- A Graph Between The No. Of Images And Histogram Statistical Features For Cancer Images

Conclusion:

The proposed research is for isolated regions of the breast cancer taken from the mammogram images. This system gives good result for the cancer extraction also the circular mask segmented the images and isolated the defect region depending on the texture properties. The first order extraction features gives information about the cancer region the comparison between the normal breast and the cancer breast with the statistical feature gives indication that the value of variance increase for some samples and decreases for other since the variance tells the intensity variation around the mean, while for the energy the value decrease for the cancer image this can be seen from fig(3) because the energy gives indication about the measure of randomness. we can also note that the mean value for the cancer images decrease because the mean value represent the average intensity see fig(3). so the mean and energy can be used for diagnose the cancer images, the result which obtained from the histogram intensity which is the average, contrast, smoothens and third moment in which everyone gives information about the texture properties which are a useful parameters for tissue classification. the first order statistic feature gives best result for classification .

Reference:

[1] Honeycutt.C.E and Plotnick.R,**2008** Image analysis techniques and Gray-level co-occurrence matrices for calculating bioturbation indices and characterizing biogenic sedimentary structures Computers& Geosciences 34,pp1461-1472.

[2] Bhiwani and Khan **2010**. Texture Based Pattern Classification. International Journal of Computer Applications vol.1, No.1,pp 0975-8887.

[3] Elaiza.N and Ibrahim. Sh. Feb **2011**. MIR Brain Abnormalities Segmentation using K-Nearest Neighbors (K-NN). International Journal on Computer Science and Engineering. vol.3 NO.2.

[4] kulkarni.D.A. and Bhagyashree.S.M. Agust **2010**. Texture Analysis of Mammographic Images. International Journal of Computer Applications. vol. 5, No.6, 0975-8887.

[5] Nagaraja. A. and Lavanya. February **2010** Segmentation Method Based on Circular Edge Magnitude for Classification of Textures. International Journal of Computer and Network Security. vol.2,No.2. pp25-29.

[6] Chevrefits.C and Cheriet.F.July **2009** Texture Analysis for Automatic Segmentation of Intervertebral Disks of Scoliotic Spines from MR Images. IEEE Transactions on information technology in biomedicine, vol.13 No. 4.

[7] Arfan. M., Jaffar and Latif .Classification and Segmentation of Brain Tumor Using Texture Analysis. Recent Advance in Artificial intelligence, Knowledge engineering and data bases.