TIPTEKNO'19
TIP TEKNOLOJİLERİ KONGRESİ
MEDICAL TECHNOLOGIES CONGRESS
3-5 Ekim 2019
Kuşadası / Aydın

Lesion Detection from the Ultrasound Images Using K-Means Algorithm

Fatma Günseli YAŞAR ÇIKLAÇANDIR
Computer Engineering Department
İzmir Kâtip Çelebi University
İzmir, Turkey
fatmagunseli.yasar@ikc.edu.tr

Akın ERTAYLAN, Uğur BİNZAT, Alp KUT
Dokuz Eylul University
İzmir, Turkey
akin.ertaylan@deu.edu.tr, ugur.binzat@deu.edu.tr,
alp@cs.deu.edu.tr

***Abstract*—It is well known that, early diagnosis is also very important for cancer patients. One of the imaging techniques used in the diagnosis of breast cancer is ultrasonography. In this study, a system that helps the doctor to detect the lesion in the breast has been suggested. We used K-Means clustering algorithm to detect the lesion in the images. The effects of three different filters (Median, Laplace, Sobel) have been examined in the study. Also, different partitioning effects are considered. According to the accuracy rates we obtained, it was concluded that the accuracy increased as the number of partitions increased. In addition, the Median filter was the best filter compared to other filters.**

Keywords — cancer; lesion detection; image processing.

I. Introduction

In a few words, a group of diseases involving abnormal cell growth with potential to invade or spread to other parts of the body is called cancer. With high potential to cause damage, cancer is the second leading cause of death in the world. According to the data from World Health Organization, 9.6 million people die from cancer in 2018. It is well known that, early diagnosis is also very important for cancer patients before cancer cause an irremediable damage. Thanks to the computers, millions of lives are saved by machine learning algorithms such as object detection, object segmentation, object tracking, and pattern recognition.

Lesion inspection is very helpful in cancer diagnosis. One of the imaging techniques used in the diagnosis of breast cancer is ultrasonography. In this study, a system that helps the doctor to detect the lesion has been aimed. Although high accuracy rates of lesion inspection, we still need human experience in order to give the final decision. This topic is studied well by many authors in the literature and some of them are inspected briefly below.

Breast cancer is one of the health problems that are becoming more common in the World. MRI, mammography, and ultrasonography techniques are used in imaging the breast lesions. Automated systems that make our lives easier have also been used in breast cancer diagnosis [1-7]. The study in [1] uses ultrasound images for lesion detection while the studies in [2-5] uses mammography images, and the studies in [6-7] uses MRI images.

Xu and Nishimura developed a lesion segmentation system using a spatial model of FCM clustering [1]. The information of intensity and texture were included to the clustering algorithm. According to their experimental results, the spatial FCM gives better results than FCM algorithm.

Lesions and their radius information were detected from mammography images using the Vision Development Module in the study of Cura et al. [2]. They used 10 images from MIAS database for testing their system. Thresholding, morphological filtering, histogram equalization and shape detection functions were used in the module. Total error value in finding the radius of lesions was found as 15% for the system.

Singh et al. proposed a study for detection and segmentation of breast cancer mass in mammograms [4]. Different edge detection techniques (Prewitt, Sobel and Robert) were applied in the study. Then, K-Means and FCM were used for segmentation of the mass. As a result, this system was recommended as an aid system to the diagnosis of the cancer.

Kamil and Salih applied the algorithms of K-Means and Fuzzy C-Means (FCM) for segmentation of tumors [7]. They used the images from mini-MIAS database to compare the performances of the algorithms. Compared to the K-Means, the FCM algorithm has a higher success in finding the tumor.

There are many clustering algorithms used to divide data into clusters. In this study we performed K-means clustering (like studies in [3-7]) on ultrasonography images. We conducted experiments on the images without filter and the images with three different filters (Median, Laplace, Sobel) to examine the effects of the filters. The system were tested for 2×2, 4×4, 8×8, 1×4, 1×8, 4×1, 8×1 partitions.

II. K-Means Clustering

There are many clustering algorithms used to divide data into clusters. The K-Means is one of the most frequently used clustering algorithms proposed by J. B. MacQueen [6]. The number of clusters, called k, must be known before applying the algorithm.

In this algorithm, each element of the database used is assigned to the nearest cluster center. Then, coordinates of each cluster center are updated by getting averages of the elements assigned to that cluster. Element assignments and cluster centers

978-1-7281-2420-9/19/$31.00 ©2019 IEEE

are updated as long as there is a change in the coordinates of the cluster centers.

Initializing the coordinates of cluster centers is a problem in the algorithm. Initial values of the coordinates change the result clusters. There are several methods for initializing the centers [7].

III. PROPOSED SYSTEM

The system was developed using Python. Input image was taken from Internet. Steps followed in this study are as in Fig. 1.

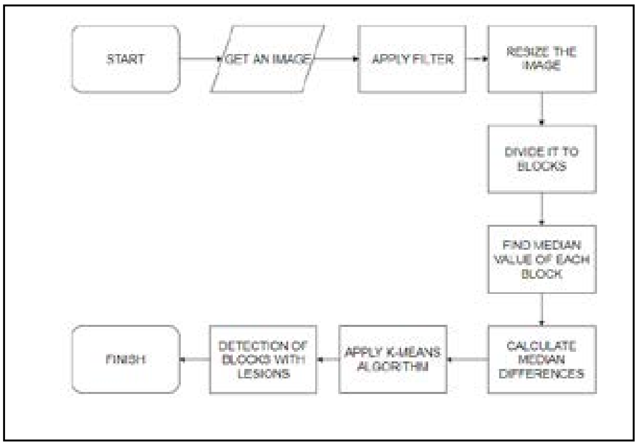


Figure 1. Flowchart of the system

Images taken as inputs were applied a filter and resized to 256×256 pixels. Then, they were divided into partitions as in Fig. 2, Fig.3 and Fig. 4. Then, median value (M) was calculated for each block (Bi) (Eq. 1). In the equation, m is the number of rows and n is the number of columns. Average of all medians was calculated (Eq. 2) and difference between the average and the median of each partition was found as in Eq. 3. These differences were clustered using K-Means algorithm.

B(1,1)
B(2,1)
B(3,1)
B(4,1)

Figure 2. Horizontal division (for 4 partitions)

B(1,1)	B(1,2)	B(1,3)	B(1,4)

Figure 3. Vertical division (for 4 partitions)

B(1,1)	B(1,2)	B(1,3)	B(1,4)
B(2,1)	B(2,2)	B(2,3)	B(2,4)
B(3,1)	B(3,2)	B(3,3)	B(3,4)
B(4,1)	B(4,2)	B(4,3)	B(4,4)

Figure 4. Division to blocks (for 4×4 blocking size)

$$M(i,j) = median\big(B(i,j)\big) \quad i = 1, \dots, n \quad j = 1, \dots, m \qquad (1)$$

$$A = \frac{1}{n*m} \sum_{i=1}^{n} \sum_{j=1}^{m} M(i,j) \qquad (2)$$

$$D(i,j) = |M(i,j) - A| \qquad (3)$$

IV. EXPERIMENTAL RESULTS

We performed experiments on the images without filter and the images with three different filters (Median, Laplace, Sobel) to examine the effects of the filters. The system were tested for 2×2, 4×4, 8×8, 1×4, 1×8, 4×1, 8×1 partitions. The accuracy values of the obtained results after clustering process were calculated using the values of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

Figures 5-7 show the partitions and the results of K-Means clustering. In the results, 1 indicates the region without lesion, and 0 indicates the region with lesion.

If the lesioned blocks were found correctly, the number of blocks with lesion gives the TP value. The block with lesion correctly defined as lesioned block, the number of these blocks gives TP. If healthy block was defined as lesioned block, the number of these blocks gives FP. If healthy block was defined as healthy, the number of these blocks gives TN. If the lesioned block is defined as healthy, the number of these blocks gives FN.

Accuracy is the ability of the test of the presence or absence of the lesion (Eq. 5).

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN) \qquad (4)$$

Table 1 shows the accuracy rates obtained from the experiments. According to these results, Median filter is the best filter compared to the others. In addition to that, 8×8 partitioning gives the best clustering accuracy compared with the other six partitioning type.

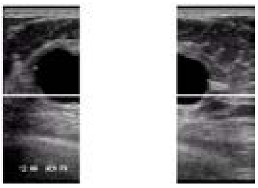
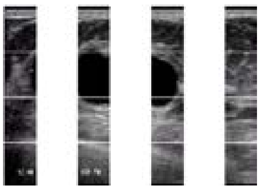
After partitioning into blocks	Obtained Clustering Results
	1 0 1 1
	0 0 0 0 0 1 1 0 0 1 1 1 1 0 0 0
	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 1 1 1 1 0 0 0 0 1 1 1 0 0 0 0 0 1 1 1 1 0 1 1 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0

Figure 5. Partitioning of images without filter (2×2, 4×4, 8×8)

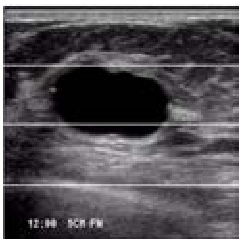
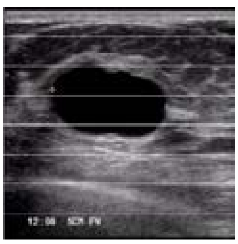
Horizontal partitioning	Obtained Clustering Results
	0 1 1 0
	0 0 1 1 0 1 0 0

Figure 6. Horizontal partitioning of images without filter (4×1, 8×1)

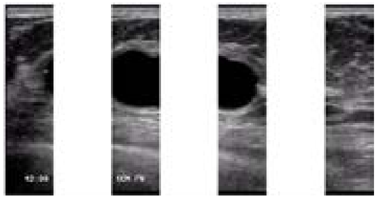
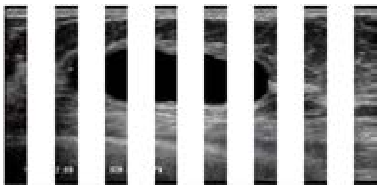
Vertical partitioning	Obtained Clustering Results
	0 1 1 1
	0 1 1 1 1 1 1 1

Figure 7. Vertical partitioning of images without filter (1×4, 1×8)

TABLO I. ACCURACY VALUES OF THE SYSTEM

Input / Filter	Without Filter	Laplace	Median	Sobel	Average acc. rate
2×2 partitioning	0.75	0.75	0.75	0.50	**0.6875**
4×4 partitioning	0.8125	0.8125	0.8125	0.6875	**0.7813**
8×8 partitioning	0.7969	0.7969	0.8333	0.80	**0.8068**
1×4 partitioning	1.00	0.5	1.00	0.5	**0.75**
1×8 partitioning	0.75	0.75	0.75	0.75	**0.75**
4×1 partitioning	0.5	0.75	0.50	1.00	**0.6875**
8×1 partitioning	0.75	0.625	0.75	0.75	**0.7188**
Average acc. rate	**0.7656**	**0.7121**	**0.7708**	**0.7125**	**0.7403**

V. CONCLUSION

It is well known that, early diagnosis is also very important for cancer patients. In this study, K-means clustering algorithm is conducted on lesion images. We performed experiments on the images without filter and the images with three different filters (Median, Laplace, Sobel) to examine the effects of the filters. The system were tested for 2×2, 4×4, 8×8, 1×4, 1×8, 4×1, 8×1 partitions.

According to average accuracy rates, Median filter gave the best results. Also, without filter, close average accuracy rates to the median is obtained. However, Laplace and Sobel filters have the lowest average accuracy rate. 8×8 partitioning gives the best clustering accuracy compared with the other six partitioning type. In addition, 4x4 partitioning has the second best average accuracy rate which is also close to 8×8 partitioning. In conclusions, we suggest to use Median filter with 8x8 partitioning to get the best results.

REFERENCES

[1] Xu, Y. and Nishimura, T., "Segmentation of Breast Lesions in Ultrasound Images Using Spatial Fuzzy Clustering and Structure Tensors", *World Acad. Sci. Eng. Technol*., 53, 2009.

[2] Karabiber Cura, Ö., Şahin, S., Akan, A., "Lesion Detection from the Mammography Image Using the Vision Development Module", *Medical Technologies National Congress*, 2017.

[3] Singh, N., Mohapatra, A.G. and Kanungo, G., "Breast Cancer Mass Detection in Mammograms using K-means and Fuzzy C-means Clustering", *Int. J. of Comput. Appl*., 22(2):15-21, 2011.

[4] Kamil, M. Y. and Salih, A. M., "Mammography Images Segmentation via Fuzzy C-mean and K-mean", *Int. J. of Intelligent Engineering and Systems*, 12(1):22-29, 2019.

[5] Patel, B.C., Sinha, G.R., "An Adaptive K-means Clustering Algorithm for Breast Image Segmentation", *International Journal of Computer Applications*, 10(4):35-38, 2010.

[6] Mandwe, A., Anjum, A., "Detection of Brain Tumor Using K-Means Clustering", *International Journal of Science and Research*, 5(6):420-423, 2016.

[7] Riahi, A., "MRI Image Segmentation by K-Means Clustering Method and Detection of Lesions", *International Journal of Science and Research*, 4(6):2484-2492, 2015.

[8] MacQueen, J.B., "Some Methods For Classification and Analysis of Multivariate Observations", *Proc. Symp. Math. Statist. and Probability*, 281– 297, 1967.

[9] Li, C.S., "Cluster Center Initialization Method for K-means Algorithm Over Data Sets with Two Clusters", *International Conference on Advances in Engineering*, 24:324-328, 2011.