Detection of Bone Metastases Using FCM and Edge Detection Algorithm

A. Aslantaş, E. Dandil, and M. Çakiroğlu

***Abstract*—Cancer diseases are one of the worst diseases threatening the human lives for many years. There is not an effective method of preventing these cancer diseases. However, early cancer detection and diagnosis play an important role in reducing the death rates. Bone scintigraphy scan is one of the most common diagnostic procedures in nuclear medicine. In general, physician experts evaluate the images manually. In these cases, the variation of diseases, which are measured for different times, cannot be determined properly. For that reason, we proposed a method to diagnose the diseases by using Fuzzy C-Means clustering and image processing techniques. In experimental studies, we evaluated the bone scintigraphy of twelve patients who are dealing with breast and prostate cancers. We determined the locations and areas of bone involvements. Also, the deviations of diseases along the times are automatically showed in the software.**

***Index Terms*—Bone scintigraphy, cancer, computer aided diagnosis, edge detection, metastasis.**

I. Introduction

Cancer, today's most dangerous disease, is currently incurable. According to statistics published by international health organization, this disease will be an important matter in human history due to the causing many deaths within the next 20 years. 13% of people who died in 2004 died because of cancer. This figure is remarkably increasing. It has been estimated that 12 million people will be death due to cancer by 2030 [1]. Therefore, early detection and diagnosis of cancer is extremely important for the treatment of this disease. We need to take advantage of technological developments at the highest level for early detection and diagnosis.

Rapidly evolving technologies as ultrasound, functional and metabolic imaging has led to seemingly limitless horizon about diagnosis of diseases in the imaging field [2]. However, the interpretation of images to be used for diagnostic purposes is difficult and is a complex process. Technological developments in this area from two-dimensional images to three-dimensional dynamic images or from anatomical images to molecular images have led to an increase in demand. Therefore, radiologists are faced with excessive work load. However, they still should continue to be very careful during the process of diagnosing and shouldn't allow wrong referrals caused by incorrect negative possibilities during the course of the disease [3].

Manuscript received April 5, 2014; revised June 5, 2014.

A. Aslantaş is with Mehmet Akif Ersoy University, Burdur, 15100 Turkey (e-mail: aslantas@mehmetakif.edu.tr).

E. Dandil is with Bilecik Şeyh Edebali University, Bilecik, Turkey (e-mail: emre.dandil@bilecik.edu.tr).

M. Çakiroğlu is with the Mechatronics Engineering Department, Sakarya University, Sakarya, Turkey (e-mail: muratc@sakarya.edu.tr).

Surveys suggest that radiologists overwork more than 39 percent between 1999 and 2003. Considering the complexity of the process and time, a large part of efficiency of radiologists is spent on diagnosis. The more radiologists examine diagnostic images, the more the risk of errors and concerns increase [2].

The vast majority of comments on diagnosis depend on the person handling the images. Differences of interpretations depend on the person and/or random examinations. An expert can extract more abnormal findings than the other, because these differences are depending on the person who is examining the images as mentioned above. Some radiologists examining the images overestimate the importance of diagnosis, while some think they are inconsiderable. One of the differences in dependence on the person is the words they use. The correct interpretation of the images is still at higher rate. One of the most effective methods to resolve the bugs is to develop a computer system which will warn the radiologist about suspicious diagnosis.

There are a limited number of computer-aided diagnosis (CAD) systems which are capable of scanning the bone system of the whole body. There is a CAD system, reported by Erdi *et al*. [4] that uses the region growing method in order to make an estimate on the survival time of a patient who is facing with prostate cancer disease. This is a semi-automatic system in which the physicians choose a particular bone lesion to put a seed into it, and then the computer system zooms in to the lesion area. In a recent study conducted by Yin and Chiu *et al*. [5], a characteristic-point-based fuzzy inferences system (CPFIS) is used in order to locate the bone lesions in human body. In this study, the bone structure of human body is segmented into six parts manually as follows: chest, vertebra, hand, head, pelvis, and legs. Each part is applied to a different CPFIS. Brightness and radioactivity asymmetry are selected as two info inputs for the characteristic-point-based fuzzy inferences system. These inputs are considered as "expertise and training" by the physicians. The sensitivity of the system is 91.5% (227 of 248) where the mean number of false positives (FPs) is 37.3 lesions per image. The nuclear medicine physicians spent a lot of time in order to discard the false positive lesions where in the high false positive results are occurred. The first completely automatic method which can scan the bones and diagnose the whole body came from Sajn *et al*. [6]. He proposed a robust knowledge-based system in order to detect the reference points in the main regions of the skeletal. The SVM algorithms and ArTeX were used for the diagnosis and these reference points in the main regions of the skeleton are used for the segmentation. The sensitivity ratio was 79.6%, while the specificity was 85.4% in Sajn's method. The nuclear medicine physicians spent a lot of time to find the

DOI: 10.7763/IJIEE.2014.V4.477

undetected lesions which is because of the low sensitivity. Some artificial neural networks are used by Sadik *et al*. [7] in order to detect the bone lesions after adopting some segmentation algorithms. The sensitivity was 90%, while the specificity was 74% in this method. Therefore, this method shows some similarities with the method developed by of Yin and Chiu in terms of low specificity, which have the same insufficiencies. As far as we know, there isn't any study in the literature which covers the cancer's decreased or increased areas of human bodies through using the images of body scans which are taken at a regular basis from the cancer patients. Horikoshi *et al*. [8] proposed the CAD software with a Japanese database, where in the Japanese patients are used as sample of the study for the analysis of the body scan. When we compare his proposal with the corresponding CAD software trained with a European database, we can see that the software proposed by Horikoshi shows significantly higher performance. The reason of this difference in results might be due to the physical differences between European and Japanese patients.

In accordance with the requirements of the system mentioned above, we realized that the Fuzzy C-Means are based on CAD software to facilitate identification and monitor the cancer disease from bone scintigraphy images

II. Bone Scintigraphy

Bone scintigraphy constitutes approximately one-third of all applications of nuclear medicine and effective in determining treatment strategies by providing access to important physiological information such as bone blood flow, bone metabolism. As well as high level of sensitivity, the advantage of bone scintigraphy is to be free of the additional risk of radiation. Bone scintigraphy is applied to the whole-body imaging technique such as investigation of all metastatic diseases, skeletal of malignant metastases in bone tumours, bone and soft tissue infections. Gamma cameras used to view them. The camera systems are gradually developing and these systems give opportunity to scan the bone structure of whole body. Radioactive material is injected into the body through the bloodstream and distributed throughout the body. The substance occurs in bone involvement. There are two types of imaging: one of them is standard late static imaging and the other is three-phase bone scintigraphy. Whole-body is displayed 2-6 hours after the application of substance. Bone scintigraphy applications are faster than any other method in detecting the presence of metastases. It is easy to apply and it is a high sensitive imaging method. Most of the metastatic diseases form a significant pathology in the form of increased or decreased involvement activity in the bone scintigraphy [9].

III. Finding Metastatic Regions by Using Fuzzy C-Means Clustering Method

In this section, cancer diagnostics and monitoring methods based on Fuzzy C-means are discussed.

A. Fuzzy C-Means

Fuzzy C-means (FCM) algorithm is the most well-known and widely used method of fuzzy division clustering techniques. Fuzzy C-means algorithm was put forward by Dunn in 1973 and developed by Bezdek in 1981which works based on the objective function. FCM method permits the objects belonging to two or more clusters. According to the principle of fuzzy logic, all data belongs to each of the clusters with membership value of ranging from [0, 1]. The sum of all classes of membership values in the data must be "1". The object, which is close to the center of a cluster, that the possibility of belonging to this cluster will be larger than the possibility of belonging to other clusters. Clustering process is completed by approaching to the objective function with minimum progress value [10].

The algorithm operates on the principle of minimization of the shifted objective function, which is the generalization of the least squares method. FCM partitions of X_i, i=1,…,N into c fuzzy groups.

$$Jm = \sum_{i=1}^{N}\sum_{j=1}^{c} U_{ij}^{m} \left\| X_i - C_j \right\|^2, 1 \le m \le \infty. \quad (1)$$

In (1), U denotes the membership function matrix of samples; J is the objective function, is the cluster center of fuzzy group of j; N denotes the number of input samples. Algorithm starts by assigning random U membership matrix. In the second step, the center vectors are calculated. Centers are calculated by (2).

$$C_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} . X_i}{\sum_{i=1}^{N} u_{ij}^{m}} \quad (2)$$

According to the calculated cluster centers, the U matrix is re-calculated by using the (3). Old U-matrix is compared to the new U-matrix and process continues until the difference is smaller than ε error rate.

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - c_k \right\|} \right)^{2/(m-1)}} \quad (3)$$

As a result of clustering, U membership matrix containing fuzzy values reflects the result. If desired, these values can be converted to 0 and 1s by clarification [10].

B. Method

Algorithm of the processes performed on the images is shown in Fig. 1. As the first step in the algorithm, we read image as dicom. In the second step, we apply fuzzy-c command by using matlab. Third step is finding maximum and minimum pixel in the cluster which represent metastatic region. Last step is calculating all pixels which indicate metastatic region.

C. Read Image

Original patient data is DICOM images. We used the data without converting it to the jpg format. DICOM (Digital Imaging and Communications in Medicine) is medical images obtained from different devices that are used all over the world. DICOM image format supports 65536 (16-bit)

grey levels of the image. In this respect, converting the images in DICOM format to the JPEG or BMP format will result in the loss of data.

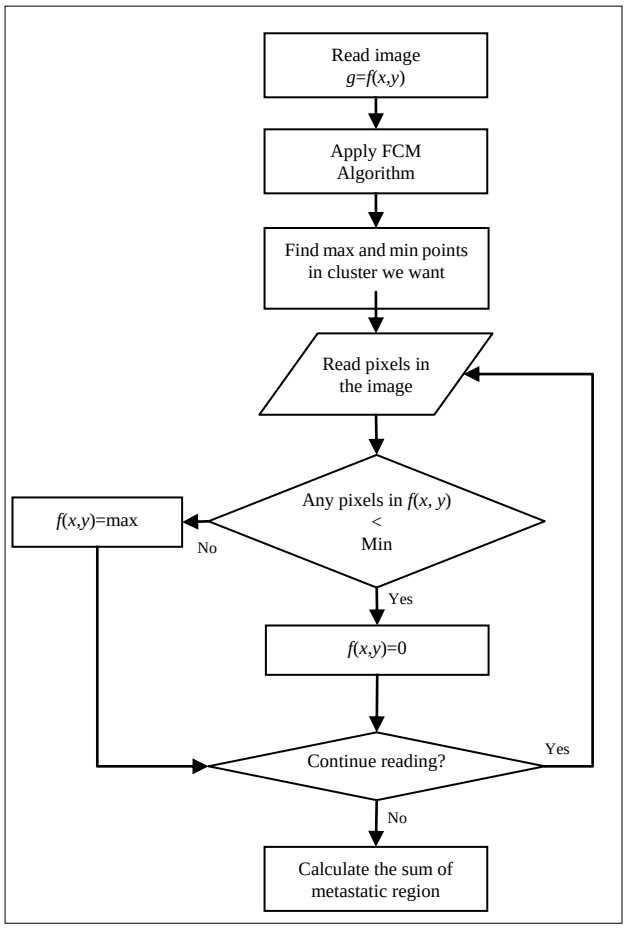


Fig. 1. Flowchart of algorithm.

Fig. 2 is an example of bone scintigraphy. Note that the image is composed of three main regions including the dark, light-coloured and the grey-tone pixels. Fuzzy C-means clustering algorithm is used to obtain these fields.

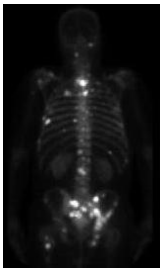


Fig. 2. Bone scintigraphy image.

D. Apply FCM

After reading the image, the method of Fuzzy c-means has been applied. Image is separated based on the number of determined clusters. Resulting clusters and connected pixels are transferred to a variable. This variable has shown the value of membership of each pixel in 4 clusters. In Table I, columns show some of the pixels in the whole body images. Rows indicate membership values. Maximum membership value shows which cluster the pixel belongs to. It can be seen in Table I, the pixel in the first column belongs to fourth Cluster and so on.

TABLE I: MEMBERSHIP VALUES FOR EACH PIXEL IN A CLUSTER

Cluster Number	Membership Values for a Pixel	Membership Values for a Pixel	Membership Values for a Pixel
Cluster-I	0.0029	0.7729	0.0029
Cluster-II	3.3360e-05	2.4560e-05	3.3280e-05
Cluster-III	2.8161e-04	2.8161e-04	0.8884
Cluster-IV	0.9968	3.6578e-04	1.2567e-05

We still don't know which cluster shows the metastatic regions. Average values of each cluster have been estimated to find metastatic regions. Metastatic region is the highest average cluster. Metastatic regions are the brightest pixels in the image. It is clearly seen that the highest average value of the subscript which we need to do processing is row 3 in Table II.

TABLE II: AVERAGE VALUE OF EACH CLUSTER

Average For Each Cluster	Value
First Cluster	58.6036
Second Cluster	18.2362
Third Cluster	169.1914
Fourth Cluster	0.9912

E. Find Maximum and Minimum Pixel in the Cluster Which We Want

Maximum and minimum values were determined in the largest average cluster. Smaller than the minimum values are equal to 0. Other regions are equal to the maximum value. Clusters obtained from the use of the algorithm appear in Fig. 3.

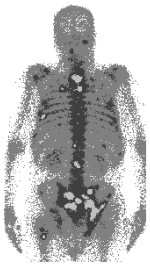


Fig. 3. Illustration of body segments by using FCM algorithm.

F. Calculate the Sum of Metastatic Region

Total pixel count of metastatic regions are calculated. We have the follow-up images of patients. Total metastatic regions are compared in follow-up images. Progression or regression of disease is determined. High rate of detection of metastasis has been clearly seen in Fig. 4 when compared with the original picture.

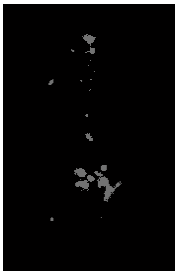


Fig. 4. Determination of metastatic regions.

IV. Finding Metastatic Regions by Using Edge Detection Algorithm

The most important action is obtaining a suitable threshold value in detection of diseased regions of bone scintigraphy. Chances of selecting a good threshold are enhanced considerably if the histogram peaks are tall, narrow, symmetric, and separated by deep valleys.

We are examining the image histogram as an example given in Fig. 5. We can see clearly that histogram is dominated by dark pixels. With the histogram of the images such as Fig. 5, achieving a healthy threshold value is difficult with the classical threshold methods.

One approach for improving the shape of histogram is to consider only those pixels that lie on or near the edges between background and the objects. If only the pixels on or near edges between objects and background were used, the resulting histogram would have peaks approximately with the same height. Using the pixels that satisfy some simple measures based on gradient and Laplacian operators has a tendency to deepen the valley between histogram peaks. [11]

Detecting changes in intensity for purpose of finding edges can be accomplished by using first and second order derivatives.

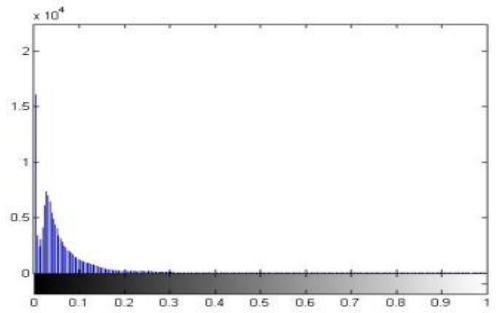
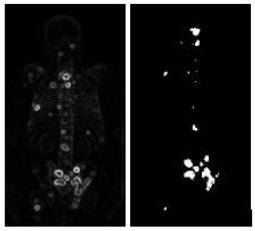


Fig. 5. Histogram of original image.

These derivatives can be implemented by filtering an image with the sobel mask**.** The preceding discussion is summarized in the following algorithm, where $f(x, y)$ is the input image:

1) Smooth image. Because edge detection can be made more selective by smoothing the image prior to computing gradient.
2) Compute an edge image as either the magnitude of the gradient or absolute value of the laplacian of $f(x, y)$
3) Specify a threshold value, T.
4) Threshold the image from step 1 using the threshold from step-2 to produce a binary image, $g_T(x, y)$. This image is used as a mask image in the following step to select pixels from $f(x, y)$ corresponding to "strong" edge pixels.
5) Compute a histogram using only the pixels in $f(x, y)$ that correspond to the locations of the 1-valued pixels in $g_T(x, y)$.
6) Use the histogram from step 4 to segment $f(x, y)$, for example: Otsu's method is being used globally.

If T is set to any value less than minimum value of the edge image, then $g_T(x, y)$ will consist of all 1s, implying that all pixels of (x, y) will be used to compute the image histogram. In this case, the preceding algorithm becomes global thresholding in which the histogram of original image is used. It is customary to specify the value of T corresponding to a percentile, which typically is set high so that few pixels in the gradient/Laplacian image will be used in the computation [11]. Fig. 6 shows the gradient image which is the starting point of this method and the result image. Metastatic regions can be seen clearly in the picture.



(a) (b)

Fig. 6. (a) Gradient image (b) Metastatic regions.

V Experimental Results

In practice, we worked on images of the 12 patients for progress and regressions of metastatic region. These images are composed of whole body, the rib cage and pelvis. The resolution of the images is 1024x256 (width and height). Images were taken from Medical Faculty of Suleyman Demirel University. We received some help from an expert radiologist for identification of metastases in the images. We printed out all the images. Radiologist outlined all hot spots for each image to be used as a gold standard. Metastases outlined by radiologist were segmented by using ROI editor. Both segmentation methods applied to all images respectively.

Fuzzy-C means gave better results. Number of clusters was tested as 3, 4, and 5 for fuzzy-c means. We have seen that the value of 4 gave the best results for metastatic regions. For 50 and 100 epochs, the results were same. The clusters produced incorrect results less than 50 epochs. Total numbers of pixels of metastasis regions in images for each patient are shown in Table III. There are two columns. The first is the prior evaluation of bone scans. The last denotes evaluation of follow up bone scans. By this way, the regression or progression of disease on patients is obtained. The last column denotes the variation as percentage. While "+" operator denotes the increment of diseases, "-" denotes the decrement of diseases. Because of edge detection algorithm results are not good, the progression and regression of disease were calculated for only Fuzzy-C means method.

TABLE III: Results of Total Metastatic Region for Fuzzy-C Means Method

Patient Name	Metastasis Area (Prior)	Metastasis Area (Follow-up Scan)	Variation Rate
Patient-1	73185	265965	+263
Patient-2	39270	13770	-64
Patient-3	56355	6885	-87
Patient-4	510	7905	+1450
Patient-5	1020	35445	+3350
Patient-6	7650	55845	+630
Patient-7	20400	46410	+127
Patient-8	9690	510	-94
Patient-9	4335	255	-94
Patient-10	26010	9945	-61
Patient-11	50235	33150	-34
Patient-12	7905	35445	+348

Both segmentation algorithms compared with each other. The differences between the segmentation algorithms were calculated and shown with percentages. Segmented areas were compared by areas but not by pixels. Their values were estimated on a number of overlapped connected areas. Results can be seen in Table IV.

TABLE IV: DIFFERENCES BETWEEN SEGMENTATION ALGORITHMS AND EXPERT

Images Name	Expert- Edge Detection	Expert- FCM
Patient-1	45	90
Patient-2	50	89
Patient-3	69	52
Patient-4	44	97
Patient-5	45	82
Patient-6	90	96
Patient-7	50	37
Patient-8	69	89
Patient-9	66	86
Patient-10	54	79
Patient-11	60	82
Patient-12	42	89

VI. CONCLUSIONS

When the results of threshold value obtained by using edge detection algorithms evaluated with radiologists, it is observed to reach the wrong conclusions about the detection of metastatic regions. Edge detection validation results didn't show good correspondence with regions marked by a radiologist.

The reason the algorithm results in incorrect data, the lack of standardization. There is no standardization in images because of the amount of a substance given to patients and low resolution cameras. This greatly affects the accuracy of the obtained results. The process of segmentation is an important step for finding metastasis correctly. This study is first step for a CAD system. Good segmentation is important for CAD system. But furthermore**,** it has a different significance. Marking all the hot spots on the image will help physicists to save time to interpret the image.

The results obtained by using the Fuzzy-C clustering algorithm are better. The disease detection and monitoring of regional metastases in 6 of 12 patients were identified correctly. In this research, the disease follow-up of the following periods is based on a total area of the diseased regions. In future, this study can give a higher sensitivity to detect areas of metastasis and improve the method of examining parts of regional metastasis.

REFERENCES

[1] J. Tang, R. M. Rangayyan, J. Xu, I. Naqa, and Y. Yang, "Computer aided detection and diagnosis of breast cancer with mammography: Recent advances," *IEEE Trans. on Information Technology in Biomedicine*, vol. 13, pp. 236-251, 2009.

[2] M. Sadık, "Computer-assisted diagnosis for the interpretation of bone scintigraphy: A new approach to improve diagnostic accuracy," Phd Thesis, Institute of Medicine at Sahlgrenska Academy University of Gothenburg, Sweden.

[3] M. Sadık, M. Suurkula, P. Höglund, A. Jarund, and L. Edenbrant, "Improved classifications of planar whole-body bone scans using a computer assisted diagnosis system: A multi-center, multiple-reader, multiple-case study," *The Journal of Nuclear Medicine*, vol. 50, pp. 368-375, 2009.

[4] Y. E. Erdi, J. L. Humm, M. Imbriaco, H. Yeung, and S. M. Larson, "Quantitative bone metastases analysis based on image segmentation," *J. Nucl. Med.*, vol. 38, no. 99, pp. 1401–1406, 1997.

[5] T. K. Yin and N. T. Chiu, "A computer-aided diagnosis for locating abnormalities in bone scintigraphy by a fuzzy system with a three-step minimization approach," *IEEE Trans. Med. Imag.*, vol. 23, pp. 639–654, 2004.

[6] L. Sajn, M. Kukar, I. Kononenko, and M. Milcinski, "Computerized segmentation of whole-body bone scintigrams and its use in automated diagnostics," *Comput. Methods Programs Biomed.* vol. 80, pp. 47–55, 2005.

[7] M. Sadik, D. Jakobsson, F. Olofsson, M. Ohlsson, M. Suurkula, and L. Edenbrandt, "A new computer-based decision-support system for the interpretation of bone scans," *Nucl. Med. Commun.*, vol. 27, no. 5, pp. 417–423, 2006.

[8] H. Horikoshi, A. Kikuchi, M. Onoguchi, K. Sjöstrand, and L. Edenbrandt, "Computer-aided diagnosis system for bone scintigrams," *Ann Nucl Med*, vol. 26, pp. 622–626, 2012.

[9] Ö. Ömür, C. Kapulu, and Ö. Uğur., "Kas iskelet sistemi hastalıklarında nükleer tıp yöntemleri," *Türk Ortopedi ve Travmatoloji Birliği Derneği Dergisi*, vol. 6, Sayı:1-2, 2007.

[10] M. Işık and A. Y. Çamurcu, "K-means, K-medoids ve bulanık C-means algoritmalarının uygulamalı olarak performanslarının tespiti," *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, vol. 11, pp. 31-45, 2007.

[11] R. C. Gonzalez and R. E. Woods, "Using Edges to improve global thresholding" in *Digital Image Processing*, 3rd ed., M. Horton, Ed., New Jersey: Prentice Hall, 2008, pp. 749-751.

A. Aslantaş was born in Ankara, Turkey, on July 16, 1979. He graduated from Suleyman Demirel University, computer system education in 2001, and he had his M.S. degree in same department from Suleyman Demirel University in 2006. He is interested in image and signal processing. Mr. Aslantaş is currently a lecturer in the Department of Computer Programming at Mehmet Akif Ersoy University.

E. Dandil was born in Osmaniye, Turkey, on August 14, 1983. He graduated from Suleyman Demirel University, Computer System Education in 2007, and he had his M.S. degree in same department from Suleyman Demirel University in 2010. He is interested in artificial intelligence, image and signal processing, pattern recognition, machine learning, data mining and optimization. Mr. Dandıl is currently a lecturer in the department of Computer Programming at Bilecik Şeyh Edebali University since 2010.

M. Cakiroglu received his Ph.D., M.S and BS degrees in computer science from the Sakarya University, Sakarya, Turkey. Dr. Çakıroğlu worked in Computer Science Department at Sakarya University as assistant professor from 2008 to 2011. He conducted various research and development about Reconfigurable Systems, Embedded System Design, and Wireless Senor Networks. Since 2012, He has been working in Mechatronics Department at Sakarya University as Associated Professor and He teaches CS courses. He served as referee for several journals and he published a number of papers and books. His current research interests include wireless sensor networks, image processing, biomedical applications and embedded system design.