International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 2 (2019) pp. 397-403

Follicle Detection in Digital Ultrasound Images using Adaptive K-means Clustering Algorithm

M.Jayanthi Rao[@], Dr.R.Kiran Kumar[#]

[@]*Research Scholar, Department of CS, Krishna University, Machilipatnam, India.*

[#]*Department of Computer Science, Krishna University, Machilipatnam, India.*

Abstract

Ultrasound Imaging is one of the technique used to study inside human body with images generated using high frequency sounds waves. The applications of ultrasound images include examination of human body parts such as Kidney, Liver, Heart and Ovaries. This paper mainly concentrates on ultrasound images of ovaries. The detection of follicles in ultrasound images of ovaries is concerned with the follicle monitoring during the diagnostic process of infertility treatment of patients. Monitoring of follicle is important in human reproduction. This paper presents a method for follicle detection in ultrasound images using Adaptive K-means clustering algorithm. The main requirements for any clustering algorithm is the number of clusters K. Estimating the value of K is difficult task for given data. This paper presents adaptive K-means clustering algorithm which generates accurate segmentation results with simple operation and avoids the interactive input K (number of clusters) value for segmentation of ultrasound image. The qualitative and quantitative results shows that adaptive K-means clustering algorithm is more efficient than normal K-means clustering algorithm in segmenting the ultrasound image. After segmentation, using the region properties of the image, the follicles in the ovary image are identified. The proposed algorithm is tested on sample ultrasound images of ovaries for identification of follicles and with the region properties, the ovaries are classified into three categories, normal ovary, cystic ovary and polycystic ovary with its properties. The experiment results are compared qualitatively with inferences drawn by medical expert manually and this data can be used to classify the ovary images.

Keywords: Ovarian Classification, Image Processing, Histogram Equalization, Bi-dimensional Empirical Mode Decomposition.

1. INTRODUCTION

PCOS (Polycystic Ovarian Syndrome) is one of the common disorder seen in females of reproductive age. The disorder is characterized by the formation of many follicular cysts in the ovary. The main cause of this disorder in females is due to menstrual problems, hirsutism, endocrine abnormalities, acne, obesity etc [1]. The detection of ovarian follicle is done using ultrasound images of ovaries. Object recognition in an ultrasound image is a challenging task which includes the detection of follicles in ovary, growth of the foetus, monitoring of proper development of the foetus and presence of tumor [2]. Now a days the diagnosis performed by doctors is to manually count the number of follicular cysts in the ovary, which is used to judge whether PCOS exists or not. This manual counting may lead to problems of variability, reproducibility and low efficiency. Automating this mechanism will resolve these problems. In literature, less work done in automating PSOC diagnosis.

In case of normal ovary, under influence of right levels of hormones FSH and LH (i.e. Follicle Stimulating Hormone and Luteinizing Hormone), only one follicle grows in size to about 20 mm in diameter, matures and becomes ready for ovulation. In PCOS affected ovary, due to reduced levels of FSH and LH and high levels of prolactine, follicles fail to grow and attain maturity. Thus, in ultrasound image of PCOS affected ovary, large number of small follicles (typically 12 or more and about 2-9 mm in diameter) can be seen distributed along the periphery of the ovary, classically described as 'necklace formation [3]. Moreover, the ovarian volume in such patients is typically increased over 10 cm^3. Figure 1 shows ultrasound image of normal ovary showing only one follicle ready for ovulation. Whereas, figure 2 shows numerous small follicles present along the periphery of the ovary [4]. In this paper, an algorithm for identification of follicles in ovarian images is presented using Adaptive K-means clustering algorithm. The experiment results are compared with inferences drawn by medical expert manually and identifies the ovary images more accurately than manual segmentation.

The paper is organized as follows: Section II presents contrast Enhancement using Bi-dimensional Empirical Mode Decomposition (BEMD), Section III presents K-means clustering algorithm, Section IV presents Adaptive K-means Clustering Algorithm, Section V presents identification of follicles in Ovary Ultrasound Images using BEMD and Adaptive K-means, Section V presents Experimental results, and finally Section VI reports conclusion.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 2 (2019) pp. 397-403

Figure 1: Normal Ovary

Figure 2: PCOS affected Ovary

2. CONTRAST ENHANCEMENT USING BI-DIMENSIONAL EMPIRICAL MODE DECOMPOSITION

Image Contrast Enhancement: If the contrast of the ovary image is low, the quality of the edges of follicles extracted from the ovary image will be poor. The quality of the ovary image can be improved by applying this BEMD based contrast enhancement algorithm to the original image prior to the computation of identification of follicles. The BEMD [5-7] process for generation of IMFs of the ovary ultrasound image are shown in figure 3 and the BEMD based contrast enhancement is given below.

i. First divide the Ovary image into Intrinsic Mode Functions (IMF) by using Bi-dimensional Empirical Mode Decomposition.

$$BEMD(F) = [imf_1, imf_2, ..., imf_n] .\backslash$$

ii. The enhanced image is obtained by the summation of IMFs, with each IMF multiplied by weight w_i. The summation of the weights should be equal to one and the weights are determined using Genetic Algorithm.

$$EMI = \sum_{i=1}^{n} w_i * imf_i \tag{1}$$

Where w_i denotes the weight of ith imf and EMI represents the enhanced image.

iii. The Genetic Algorithm is applied as follows.

a) Randomly generate chromosomes representing initial population with values indicating the weights of IMFs. The length of chromosome is equal to the number of IMFs generated from BEMD, and each element represents the corresponding IMF weight. The values of weights differ between 0 and 1 and the summation of weights is equal to 1.

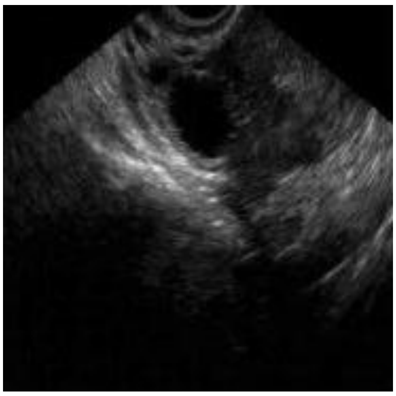
b) Calculate the fitness value of each chromosome using information entropy of image as objective function. The information entropy is defined as

$$\text{Entropy} = -\text{sum}(p.*\log2(p)) \tag{2}$$

c) Based on Entropy values, the chromosomes are selected using roulette wheel technique.

d) To generate new chromosomes from the current population, two operators such as crossover and mutation are used with probabilities 0.7 and 0.07.

e) Maximum number of iterations is used as stopping criteria. The values in the chromosome are used as optimum weights in constructing the enhanced image.

Ultrasound Ovary Image	Ultrasound Ovary Image (Contrast Enhanced)
	

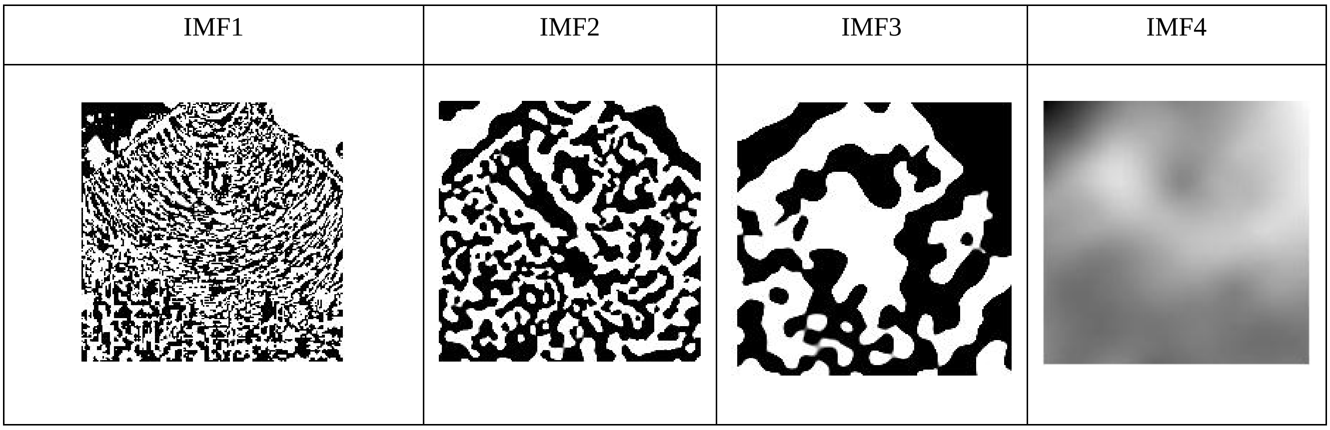


Figure 3: IMF of ultrasound image

3. K-MEANS CLUSTERING ALGORITHM

The K-means clustering algorithm is a partitioning method which assign pixels to a user-defined mutually exclusive number of clusters (k) in such a way that maximizes the separation of those clusters while minimizing intra-cluster distances relative to the cluster's mean or centroid and returns the index value of cluster, for which the pixel is assigned. The k-means clustering is always better than hierarchical clustering for large amount of data. The algorithm mainly dependent on three parameters 1) numbers of clusters k 2) Selection of initial values of centroids 3) the distance metric (optimization function) used in assignment of pixels to different clusters. This paper presents the implementation of k-means on ultrasound image with different parameters available in k-means function in MATLAB and shows how best the clusters are obtained using silhouette index.

The k-means clustering is given as follows:

1. Initialization of k centroid values. This initialization in MATALB by using name –'Start' and values 'plus', 'cluster', 'sample', 'uniform' and 'numeric' in the k-means function. The default value is 'plus', which selects k pixel values from image randomly as centroid values.
2. Assignment of pixels to the corresponding cluster based on the minimum value of distance measure from corresponding pixel to centroid.

   The k-means function in MATLAB has following distance measures with name –'Distance' and values – 'sqeuclidean', 'cityblock', 'cosine', 'correlation' and 'hamming'.
3. Up-dation of centroid values by taking mean of pixels belonging to cluster generating new centroids.
4. This process is repeated until there is no change in old and new centroid values or by specifying maximum number of iterations in kmeans function using name – 'MaxIter' and value – number of iterations (numeric value).

The k-means function in MATLAB returns cluster indices, centroid locations, distances from pixel to centroid and sum of distances from pixel to centroid within cluster. To know the information how given image is segmented into clusters, draw silhouette plot using the cluster indices returned from k-means function. The silhouette plot displays a measure of how close each pixel in one cluster is to points in the neighboring clusters. The silhouette ranges from -1 to +1. +1 indicates pixels are assigned perfectly to one cluster and -1 indicates wrong assignment of pixels to cluster. This measure is calculated using silhouette function in MATLAB with three parameters, input image, cluster indices returned from k-means function and distance measure.

4. ADAPTIVE K-MEANS CLUSTERING ALGORITHM

The basic idea of any clustering algorithm is to cluster the objects closest to them by clustering the K points in the space. Iteratively, the values of centroid of clusters are updated one by one until the best clustering results are obtained. Determining the correct K value is the key to the success of the any clustering algorithm. In this paper we have implemented Adaptive k-means clustering algorithm for estimation of K-value. The K-means algorithm takes Euclidean distance as the similarity measure, which is to find the optimal classification of an initial cluster center vector, so that the evaluation index is minimum. The error square sum criterion function is used as a clustering criterion function. Although the algorithm of K-means is efficient, value of K should be given in advance, and the selection of K value is very difficult to estimate. In the proposed method, we start with the selection of K = 2, that is, image segmentation starts from two clusters, and then the image is segmented. Finally, we determine the number of segmentation results based on the maximum connected domain algorithm [8, 9]. If the image number of the final segmentation result matches the K value, the K value is selected correctly. If the K value does not match, the K value at the beginning will be increased until the above two values match.

The adaptive k-means clustering algorithm is presented below

For K=2 to 10

{

Randomly consider K initial clusters $\{C_1, C_2, \ldots\ldots, C_k\}$ from the m*n image pixels $\{I_1, I_2, I_3, \ldots\ldots, I_{m*n}\}$.

1. Assign each pixel to the cluster C_j {j=1,2,…..K} if it satisfies the following condition

$$D(I_i, C_j) < D(I_i, C_q), q = 1, 2, ..., K \quad j \neq q \tag{3}$$

Where D(. , .) denotes the dissimilarity measure.

2. Find new cluster centroid as follows

$$C_i^{\wedge} = \frac{1}{n_i} \sum_{I_j \in C_i} I_j, i = 1, 2, ...K \tag{4}$$

Where n_i is the number of pixels belonging to cluster C_i.

3. If

$$C_i^{\wedge} = C_i, i = 1, 2, ..K \tag{5}$$

Then stop.
Else continue from step 2.
Compare the maximum connected domain results
If equal to K print segmented result and break;
else continue with incremented value of K;
}

5. FOLLICLE IDENTIFICATION IN OVARY ULTRASOUND IMAGE

The main steps of the identification of follicles in ovary image are summed up as follows:

1. The ultrasound image is processed by BEMD based contrast enhancement algorithm, then an image with sharper contrast is obtained.
2. Apply Adaptive K-means algorithm to segment the contrast enhanced ovary image
3. Convert the segmented image into binary image and edges are extracted using canny operator.
4. Perform Mathematical Morphology operations on the resulting image to remove isolated pixels and to replace thick lines with thin lines.
5. Use imfill function to fill the holes. These holes denote follicles in ovary image.
6. Using the region properties, we can extract the features of follicles such as major axis length, minor axis length, area, centroids etc. With this features, ovaries are classified into normal ovary, cystic ovary and polycystic ovary.

6. EXPERIMENTAL RESULTS

In this section, the proposed method is used to detect the follicles in ovary ultrasound images. Ovary images are obtained from the publicly available websites www.radiologyinfo.com [11,12], www.ovaryresearch.com [10] . The image1 and image2 is taken from dataset D1 with size 256*256.. First the image is segmented using K-means clustering algorithm in MATLAB. The silhouette plots for the k-means function with different distance measures are shown in figure 4 with different parameters in k-means function. The efficient way to compare the performance of k-means clustering with different functions is done by calculating the mean of silhouette values. The larger mean value means better segmentation. The mean of silhouette values for the above functions executed on ovary image shown in figure 1 is shown in table 1. The qualitative analysis of the proposed method is shown in figure 5 compared with manual segmentation. Table 2 shows the quantitative evaluation of clustering algorithm on Ovary Image 1 using MSE [13].

7. CONCLUSION

The information about the status of female reproductive system is important for problems related to fertility and family planning. The ultrasound imaging is an effective tool in infertility treatment. Monitoring the follicles in ultrasound images in terms of number, size, shape and position is important in human reproductive system. In this paper, a methodology of identification of follicles in ultrasound images is presented. The proposed method uses Bi-dimensional empirical mode decomposition procedure, to extract first intrinsic mode functions and each IMF is multiplied by weight w_i which are estimated using GA. The summation of weighted IMFs gives the contrast enhanced image. After enhancement, the image is segmented using Adaptive k-means algorithm. By using region properties, the follicle regions in the ultrasound image is identified. This proposed method is compared qualitatively with manual segmentation. In future this follicle information is used to classify the ovary images into three classes, normal ovary, cystic ovary and polycystic ovary using the region properties of identified follicles.

Table 1: Silhouette values with different distance measures

Distance measure	**Silhouette value (mean)**
Sqeuclidean	0.8534
City Block	0.7847
Cosine	0.4224
Hamming	0.0735

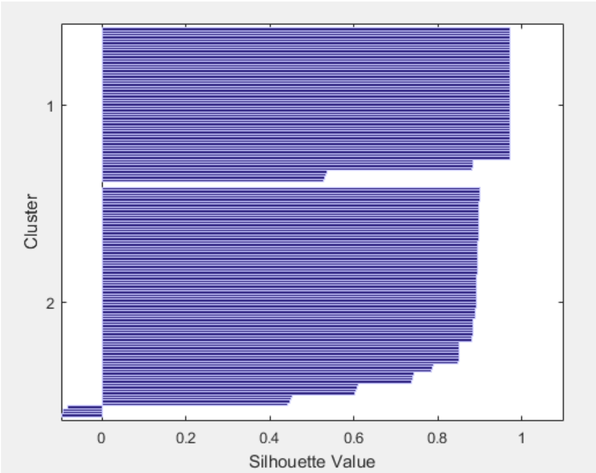
Table 2: MSE value

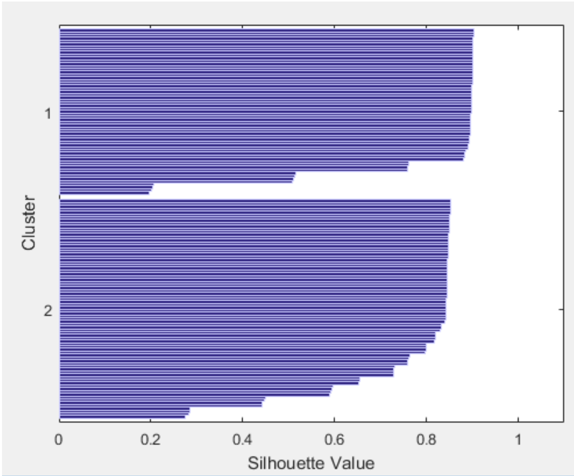
Method / Ovary Image 1	Normal Clustering	Adaptive Clustering
K-means	95.8	94.5

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 2 (2019) pp. 397-403


```
i1=imread('hg1.jpg');
figure,imshow(i1);
i2=rgb2gray(i1);
figure,imshow(i2);
[id11,C] =
kmeans(double(i2),2,'Distance','sqeuclidean');
[silh2, h] = silhouette(double(i2),
id11,'sqeuclidean');
h = gca; h.Children.EdgeColor = [0.8 0.8 1];
xlabel 'Silhouette Value'
ylabel 'Cluster'
```

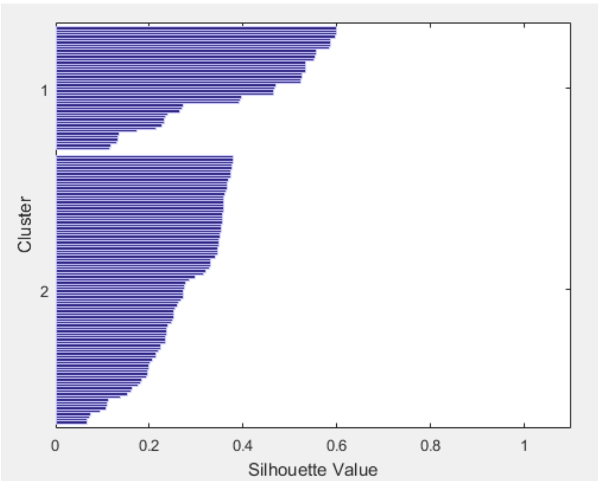
```
[id12,C] =
kmeans(double(i2),2,'Distance','cityblock');
[silh2, h] = silhouette(double(i2), id12,'cityblock');
h = gca; h.Children.EdgeColor = [0.8 0.8 1];
xlabel 'Silhouette Value'
ylabel 'Cluster'
```






```
[id13,C] = kmeans(double(i2),2,'Start', 'sample',
'Distance','cosine');
[silh2, h] = silhouette(double(i2), id13,'cosine');
h = gca; h.Children.EdgeColor = [0.8 0.8 1];
xlabel 'Silhouette Value'
ylabel 'Cluster'
```

```
[id14,C] =
kmeans(im2bw(i2),2,'Distance','hamming',
'MaxIter', 80);
[silh2, h] = silhouette(double(i2),
id14,'hamming');
h = gca; h.Children.EdgeColor = [0.8 0.8 1];
xlabel 'Silhouette Value'
ylabel 'Cluster'
cluster(i)=mean(silh2);
```



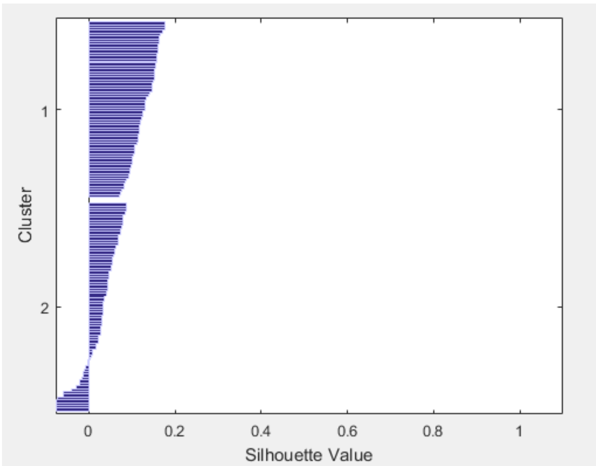
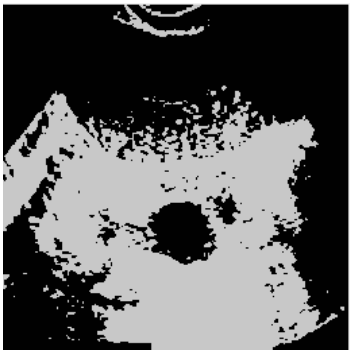
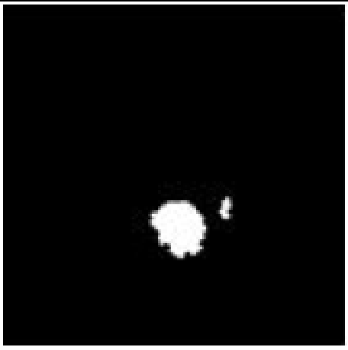
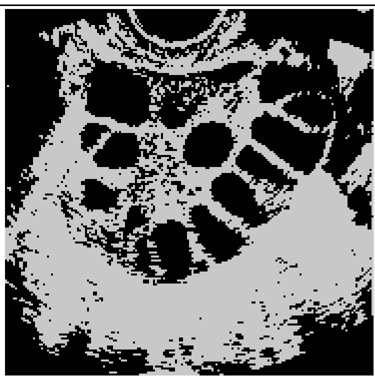
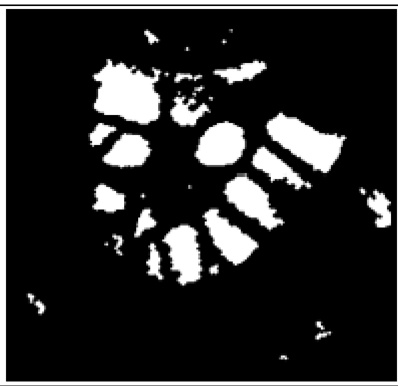


Figure 2: Silhouette Index

Ovary Image 1	Contrast Enhanced using BEMD
	
Segmented using FCM	Follicle identification
	
Ovary Image 2	Contrast Enhanced using BEMD
	
Segmented using Adaptive K-means	Follicle identification
	
Manual Segmentation Ovary Image 1	Manual Segmentation Ovary Image 2

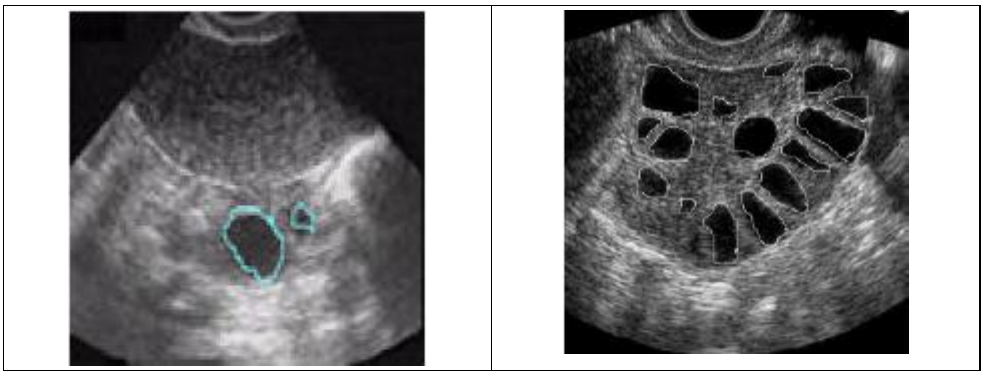


Figure 5: Follicle detection in ovary ultrasound image

REFERENCES

[1] Yinhui Deng, Yuanyuan Wang, Ping Chen, "Automated Detection of Polycystic Ovary Syndrome from Ultrasound Images", 30th Annual International IEEE EMBS Conference, IEEE 2008, PP. 4772-4776.

[2] Kiruthika.V, M.M.Ramya, "Automated Segmentation of Ovarian Follicle using K-menas Clustering", 2014, Fifth International Conference on Signal and Image Processing", pp.137-142.

[3] Palak Mehrotra, Chandana Chakraborthy, Biswanath Ghoshdastidar, "Automated Ovarian Follicle recognition for Polycystic ovary syndrome", Proceedings of International conference on Image Information processing, ICIIP 2011.

[4] Sharvari S.Deshpande, Asmita Wakankar," Automated dectection of Polycystic Ovarian Syndrome using Follicle Recognition", Proceedings of International Conference on Advanced Communication Control and Computing Technologies, IEEE 2014, pp. 1341-1347.

[5] B.Saichandana, J.Harikiran, Dr.K.Srinivas, Dr.R.KiranKumar, "Application of BEMD and Hierarchical Image Fusion in Hyperspectral Image Classification", International Journal of Computer Science and Information Security, 2016; 14(5): 437-445.

[6] J.Harikiran, et.al**.** "Multiple Feature Fuzzy C-means Clustering Algorithm for Segmentation of Microarray image", IAES International Journal of Electrical and Computer Engineering", 2015; 5(5): 1045-1053.

[7] J.Harikiran, et.al. "Fuzzy C-means with Bi-dimensional empirical Mode decomposition for segmentation of Microarray Image", International Journal of Computer Science issues, 2012; 9(5): 273-279.

[8] W Zuo, Research on connected region extraction algorithms [J]. Comp. Appl. Softw. 23(1), 97–98 (2006).

[9] J.Harikiran, Dr.P.V.Lakshmi, Dr.R.Kiran Kumar, "Multiple Feature Fuzzy C-means Clustering Algorithm for Segmentation of Microarray image", IAES International Journal of Electrical and Computer Engineering", Vol. 5, No. 5, pp. 1045-1053, October 2015.

[10] B.Saichandana, Dr.K.Srinivas, Dr.R.Kiran Kumar , "Clustering Algorithm combined with Hill climbing for Classification of Remote Sensing Image", International Journal of Electrical and Computer Engineering,vol 4, No 6, pp.923-930, December 2014 .

[11] P.S.Hiremath and Jyothi R.Tegnoor, "Follicle Detection and Ovarian Classification in Digital Ultrasound Images of Ovaries", INTECH http://dx.doi.org/10.5772/56518.

[12] Sandy RIHINA et.al. "Automated Algorithm for Ovarian Cysts Detection in Ultrasonogram", Proceedings of International Conference on Advances in Biomedical Engineering IEEE 2013.

[13] J.Harikiran, Dr.P.V.Lakshmi, Dr.R.Kiran Kumar, "Fast Clustering Algorithms for Segmentation of Microarray Images", International Journal of Scientific & Engineering Research, Volume 5, Issue 10, pp 569-574, 2014.