

International Conference on Information and Communication Technologies (ICICT 2014)

Fuzzy clustering algorithms for cDNA microarray image spots segmentation.

Biju V.G.^{a,*}, Mythili. P^b

^aAssociate Professor, C.E. Munnar and Research Scholar, Division of Electronics, School of Engineering, CUSAT, Cochin-22, India

^bAssociate Professor, Division of Electronics, School of Engineering, CUSAT, Cochin-22, India

Abstract

cDNA microarray image provides useful information about thousands of gene expressions simultaneously. Hence microarray image segmentation is an important task. In this paper, existing fuzzy clustering image segmentation methods in the literature have been tested for its suitability to perform segmentation of noisy cDNA microarray images. The algorithms considered for this purpose include fuzzy clustering based methods like, Fuzzy c-means (FCM), Possibilistic c means (PCM), Possibilistic fuzzy c means (PFCM) and Fuzzy local information c means (FLICM). The results of segmentation shows that FLICM is better in segmenting microarray spots compared to the other under the presence of noise.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)

Keywords: gene expression; clustering; spatial information; local information; image segmentation

1. Introduction

The initiation of microarray imaging technology has helped the scientist to effectively view the expression levels of thousands of gene at a time, which led to the expansion of studies in the life science.^{1,2,3} The spots on a microarray are segmented from the background to compute the gene expression. The three basic operations to compute the spot intensities are gridding, segmentation and intensity extraction.

* Corresponding author. Tel.: +919447234843.

E-mail address: bvgpillai@gmail.com

These operations are used to find the accurate location of the spot, separate spot foreground (FG) from background

(BG) and the calculation of the mean red and green intensity ratio for gene expression.

For microarray image analysis several software packages and algorithms were developed during the last decade. ScanAlyze Software uses a fixed circle segmentation algorithm for segmenting microarray spots, analysed each spot with a circle of fixed radius.⁴ An adaptive circle segmentation technique was employed in the GenePix software⁵, where the radius of each spot was not considered constant but adapts to each spot separately. Dapple software estimated the radius of the spot using the laplacian based edge detection.⁶ An adaptive shape segmentation technique was used in the Spot software.⁷ A histogram-based segmentation method was used in the ImaGene software⁸, Later Watershed⁹ and the Seeded region growing algorithms were employed¹⁰. The disadvantage of the above mentioned software packages and algorithms were either the spots were considered to be circular in shape or a priori knowledge of the precise position of the spot's center was a prerequisite.¹¹ Further segmentation algorithms based on the statistical Mann–Whitney test were also used¹², which assess the statistical significant difference between the FG and BG. The K-means, Fuzzy c mean (FCM) and Genetic algorithm based fuzzy c mean algorithms (GAFCM) were the clustering algorithms used for microarray spot segmentation.^{13,14,15}

The existing microarray spot segmentation algorithms show poor performance under noise. In order to avoid these limitations, in this paper, existing fuzzy clustering image segmentation algorithm have been tested for its suitability to segment cDNA microarray images in the presence of noise. The fuzzy clustering algorithms such as FCM¹⁴, Possibilistic c mean (PCM)¹⁶, Possibilistic fuzzy c means (PFCM)¹⁶, and Fuzzy local information c mean (FLICM)¹⁷ are used for microarray spot segmentation. Gridding of microarray image is done based on an original genetic algorithm approach for automatic gridding.¹⁸ For evaluation and testing of the algorithm both simulated and real microarray images are used. The performance of the algorithms are tested by evaluating the segmentation matching factor (SMF), Probability of error (p_e) and Normal mean square error (NMSE).

2. Fuzzy clustering

The aim of microarray image processing is to separate the FG and BG of each spot and then extract the intensity information from it. This is done through 3 steps 1. Gridding 2.Segmentation and 3.Intensity extraction. To address each spots in the image, gridding is done. Segmentation will separate the FG and BG of each spot. The intensity extraction will find the intensity of red and green information of the FG. The logarithmic value (base 2) of the ratio of intensities is the gene expression. These results are useful for accurate microarray analysis which involves data normalization, filtering and data mining. Fuzzy clustering is one of the most significant techniques that is used for segmentation and is applied in microarray images. The idea of clustering application is to divide the pixels of the image into several clusters (usually two clusters) and characterize these clusters as signal or background. The fuzzy clustering algorithms such as FCM, PCM, PFCM and FLICM are coded using matlab¹⁹ and used for spot segmentation of cDNA microarray images.

2.1 Possibilistic fuzzy c means (PFCM)

The PFCM¹⁶ is a hybrid version of FCM and PCM.^{14,16} It enjoys the benefits of both models. It solves the noise sensitivity defect of FCM and overcomes the coincident clustering problem of PCM. The FCM clustering algorithm was first introduced by Dunn and later extended by Bezdek. Let x_i $i = 1$ to N be the pixels of a single microarray spot, where N is the total number of pixels present in the spot image. These pixels have to be clustered in two classes BG and FG. Let c_j $j = 1, 2$ be the prototype cluster centers of the FG and BG pixels respectively. A membership function u_{ij} represents the membership value of each pixel to be in different clusters. Based on the maximum value of the membership function each pixel is grouped. The cluster centers are updated iteratively based on the grouped pixel. FCM is an iterative clustering algorithm that produces an optimal c partitions by minimizing the weighted within group sum of squared error objective function F^t

$$F^t = \sum_{i=1}^N \sum_{j=1}^c u_{ij}^m d_{ij} \quad m \in [1, \infty] \quad (1)$$

where d_{ij} is the Euclidean distance from a pixel to a cluster center and is given by

$$d_{ij} = \|x_i - c_j\|^2 \quad (2)$$

Hence the algorithm aims at iteratively improving the membership degree function until there is no change in the cluster centers. The sum of the membership values of a pixel belonging to all clusters should satisfy the following equation.

$$\sum_{j=1}^c u_{ij} = 1 \quad \forall i = 1, 2, \dots, N \quad (3)$$

The aim of this method is to minimize the absolute value of the difference between the two consecutive objective functions F^t and F^{t+1} given by the equation 4.

$$\|F^{t+1} - F^t\| \leq \varepsilon \quad (4)$$

where m is the fuzziness parameter and ε is error which has to be minimized. Iteratively in each step, the updated membership u_{ij} and the cluster centers c_j are given by the following equations.

$$u_{ij} = \frac{1}{\sum_{k=1}^c (d_{ij}/d_{ik})^{2/(m-1)}} \quad (5)$$

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m x_i}{\sum_{i=1}^N u_{ij}^m} \quad (6)$$

In FCM the membership of each pixel x_i is inversely related to the relative distance between x_i and c_j . In microarray images $j = 2$, i.e. FG and BG, so when $\|x_i - c_1\| = \|x_i - c_2\|$ these pixel x_i has to be given equal membership in each clusters i.e. 0.5 irrespective of the low value or high value of these pixels and these creates noise points or outliers in FCM. To avoid this, Krishnapuram and Kellar proposed a new clustering model named PCM.²⁰

In PCM, which is an improved version of FCM, the constraint in equation 2 is relaxed to facilitate a possibilistic interpretation of the membership function. In other words, each pixels of the j^{th} column can be any number between zero and one, so long as at least one of them is positive. The value of u_{ij} has to be interpreted as the typicality of x_i relative to cluster j (rather than its membership in the cluster). Each row of t_{ij} was interpreted as a possibility over x . The objective function in PCM is given by equation 7.

$$F^t = \sum_{i=1}^N \sum_{j=1}^c t_{ij}^m d_{ij} + \sum_{i=1}^c \gamma_i \sum_{j=1}^N (1 - t_{ij})^m m \in [1, \infty] \quad (7)$$

where t_{ij} is called the typicality of the pixel, each row of t_{ij} is interpreted as the possibility distribution over x . t_{ij} is calculated by equation 8.

$$t_{ij} = \frac{1}{1 + (d_{ij}/\gamma_i)^{1/(m-1)}} \quad (8)$$

where γ is a constant and is given by equation 9.

$$\gamma_i = K \frac{\sum_{j=1}^N u_{ij}^m d_{ij}}{\sum_{j=1}^N u_{ij}^m} \quad K > 0 \quad (9)$$

where K is a constant greater than zero and u_{ij} is obtained using equation 5. The cluster centers c_j is given by equation 10.

$$c_j = \frac{\sum_{i=1}^N t_{ij}^m x_i}{\sum_{i=1}^N t_{ij}^m} \quad (10)$$

PFCM is a hybrid of FCM and PCM and enjoys the benefits of both models. PFCM solves the noise sensitivity defect of FCM and overcomes the coincident clustering problem of PCM. It is an iterative algorithm that tries to minimize the objective function given by equation 11. The cluster centers c_j is given by equation 12.

$$F^t = \sum_{i=1}^N \sum_{j=1}^c (a u_{ij}^m + b t_{ij}^\eta) d_{ij} + \sum_{i=1}^c \gamma_i \sum_{j=1}^N (1 - t_{ij})^\eta \quad (11)$$

$$c_j = \frac{\sum_{i=1}^N (a u_{ij}^m + b t_{ij}^\eta) x_i}{\sum_{i=1}^N (a u_{ij}^m + b t_{ij}^\eta)} \quad (12)$$

where u_{ij} is the membership function given by equation 5, t_{ij} is called the typicality of the pixel obtained by equation 8 and 9, m & γ are the fuzziness parameters, a & b are constants with value equal to or greater than one.

In PFCM¹⁶, the advantages and disadvantages of fuzzy clustering algorithms such as FCM and PCM methods were mathematically analyzed and a hybrid method PFCM was presented. The paper had compared various aspects of FCM, PCM and PFCM with four numerical data sets $X_{10}, X_{12}, X_{400}, X_{550}$ and one image dataset IRIS. For all data sets the maximum number of iterations was taken as 100. The number of clusters was taken as 3 for IRIS and 2 for all other data sets.

2.2 Fuzzy local information c means (FLICM)

FLICM is a modified version of FCM which makes use of fuzzy local similarity measure, aiming to guarantee noise insensitiveness and image details preservation. In order to enhance the insensitiveness to noise, a new factor G_{ij} is included in FCM objective function.¹⁷ This factor incorporates local gray level and local spatial information in a fuzzy way so as to obtain robustness and noise insensitiveness, and also control the influence of the neighbourhood pixels depending on their distance from the central pixel.

$$G_{ij} = \sum_{i=N_k} \frac{1}{d_{ki}+1} (1 - u_{ij})^m \|x_i - c_j\|^2 \quad (13)$$

where the k^{th} pixel is the center of the local window, j is the reference cluster and the i^{th} pixel belongs to the set of neighbours falling into a window around the k^{th} pixel N_k . d_{kj} is the spatial Euclidean distance between pixels k & j , u_{ij} is the degree of membership of the i^{th} pixel in the j^{th} cluster, m is the weighting exponent on each fuzzy membership, and c_j is the prototype of the centre of cluster j . Local grey level and spatial information are included in the objective function and is given by equation 14.

$$F^t = \sum_{i=1}^N \sum_{j=1}^c [u_{ij}^m d_{ij} + G_{ij}] \quad (14)$$

where the membership function and u_{ij} center c_j are obtained by equation 15 and 16.

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij} + G_{ij}}{d_{ik} + G_{ik}} \right)^{1/(m-1)}} \quad (15)$$

$$c_j = \frac{\sum_{i=1}^N (u_{ij}^m) x_i}{\sum_{i=1}^N (u_{ij}^m)} \quad (16)$$

The objective function F^t has to be minimized to find the optimum cluster centers. It can be obtained through an iterative process. The FLICM algorithm is given as follows.

- Step 1. Initialize the cluster centers c_j , fuzzification parameter(m) and the stopping condition(ε).
- Step 2. Find u_{ij} from equation 5 of FCM.
- Step 3. Set the loop count $iter=0$
- Step 4. Compute fuzzy partition membership degree matrix u_{ij} using equation 15.
- Step 5. Calculate c_j using equation 16.
- Step 6. Compute objective function F^t using equation 14.
- Step 7. If the max $\{||F^{t+1} - F^t||\} \leq \varepsilon$ then stop otherwise $iter = iter+1$ and go to step 4.

In FLICM¹⁷, the efficiency and the robustness of the method was compared with six fuzzy algorithms FCM S1, FCM S2, EnFCM, FGFCM S1, FGFCM S2, FGFCM, and two well-known non-fuzzy algorithms, k-means and SLINK algorithm. The performance of the algorithm was compared by presenting numerical results and examples on various synthetic and real images, with different types of noise. The synthetic test image used was a 128×128 pixel image with two gray level values taken as 20 and 120 and corrupted by different levels of Gaussian, Uniform and Salt & Pepper noise respectively. The number of clusters was taken as 2. The real image includes coin, wheel and flower images. The de noising performances of the above nine algorithms were compared with respect to the optimal segmentation accuracy (SA) and fuzzy similarity measure (r). SA is defined as the sum of the correctly classified pixels divided by the sum of the total number of pixels. r is a measure indicating the degree of equality between A_i and C_i , A_i represents the set of pixels belonging to the i^{th} class found by the algorithm, while C_i represents the set of pixels belonging to the i^{th} class in the reference segmented image.

3. Database used

3.1 Synthetic database

A set of 40 microarray images, each with 225 spot, are simulated by the authors as mentioned in the literature^{21,22} for numerically evaluating and comparing the various segmentation methods. In order to generate spots with realistic characteristics, the following procedure is adopted. Real cDNA sub array images are used as templates. Its binary version is produced by employing a suitable threshold. The address of each spot are identified by gridding. The intensities of each FG region is drawn from a uniform distribution whose mean value is taken as original spot mean value. The remaining BG pixel intensities are drawn from a uniform distribution whose mean intensity is determined from the original image. Note, all the BG intensities is drawn from a single distribution while FG intensities of each target region are drawn from an uniform distribution whose mean is estimated separately from the original respective spot region. The spots in microarray images sometimes exhibited doughnut-like shapes. During the simulation the doughnut holes which are identified as BG during thresholding, have the same intensity distributions as the BG.

3.2 Real microarray images

Real microarray images with category cell line & sub category drug treatment are downloaded from the UNC microarray data base.²³ Each image consists of 34 blocks or sub arrays with each block containing 625 spots. From the downloaded images we have arbitrary selected 25 microarray blocks, i.e. total 15625 spots are used for segmentation. Although the ground truth is not known, it is clear from the segmentation result that the proposed method is more efficient in segmenting the real microarray spots.

4. Measures used for evaluation

To compare the performance of the fuzzy clustering algorithm applied on a synthetic database the following parameters are used. The synthetic database is evaluated with three parameter such as probability of error p_e , segmentation matching factor (SMF), and Normal mean square error (NMSE). The synthetic database images are corrupted by Additive white gaussian noise (AWGN)^{21,22} with the Signal-to-noise ratio (SNR) ranging from 1 to 10 dB. The segmentation ability of the FLICM algorithm is compared with FCM, PCM and PFCM by

finding the Segmentation matching factor (SMF), Probability of error (p_e), and Normal mean square error (NMSE) for every binary spots produced by these clustering algorithms.

The pixel-level accuracy of the segmentation is examined with the statistical parameter probability of error p_e , which measures the missegmented pixels, and is defined as^{21,24}

$$p_e = P(F)P\left(\frac{B}{F}\right) + P(B)P\left(\frac{F}{B}\right) \quad (17)$$

where $P\left(\frac{B}{F}\right)$ is the probability of error in classifying foreground pixel as background pixels, $P\left(\frac{F}{B}\right)$ is the probability of error in classifying background pixels as foreground pixels, $P(F)$ and $P(B)$ are a priori probabilities of foreground and background pixels in the image. The minimum value of zero occurs for p_e when all of the pixels of the spots are segmented correctly. A maximum value of one for p_e indicates a situation where all of the pixels of the background are segmented as foreground and vice versa. The Segmentation matching factor (SMF)^{21,22} for every binary spot, produced by the clustering algorithm is given by

$$SMF = \frac{(A_{seg} \cap A_{act})}{(A_{seg} \cup A_{act})} * 100 \quad (18)$$

where A_{seg} is the area of the spot, as determined by the proposed algorithm and A_{act} is the actual spot area. A perfect match in the case of SMF is indicated by a 100% score, any score higher than 50% indicates reasonable segmentation where as a score less than 50% indicate poor segmentation.^{21,22}

For a simulated image with known ground truth, another metric called Normalized mean square error (NMSE)²⁵ is used to measure the performance of the proposed approach which is given by

$$NMSE = \frac{\sqrt{\frac{1}{MN} \sum_i^M \sum_j^N (x_{ij} - \bar{x}_{ij})^2}}{\frac{1}{MN} \sum_i^M \sum_j^N (x_{ij})} \quad (19)$$

where M and N are the dimensions of the image. x_{ij} and \bar{x}_{ij} are the original and clustered image pixels respectively. NMSE is calculated for varying noise levels in the input image. A minimum value of zero is desirable for better segmentation.

5. Results and Discussion

After validating the fuzzy clustering algorithms and ensuring its correctness, it is applied on a synthetic database. The synthetic microarray dataset includes a set of 40 images, each with 225 spots which is simulated as mentioned in the literature.^{21,22} Table 1 summarizes the average performance results obtained for segmenting spots of 40 simulated microarray images (10000 spots). The Fig. 1.a shows a simulated microarray image and Fig. 1. b shows the segmentation result obtained using the FLICM algorithm. The segmentation algorithm is applied on each image after applying the AWGN noise. The SNR value of noise is varied from 1 to 10 dB. The performance measurement parameter such as Segmentation matching factor (SMF), Probability of error (p_e), and Normalized mean square error ($NMSE$) achieved for all simulated spots corresponding to different SNR levels are presented in Table 1. Regarding the SMF , the FLICM algorithm resulted in higher spot area identification accuracy than FCM, PCM and PFCM.

The ultimate goal of the segmentation process in microarray image processing is to obtain intensity measurement. Accurate segmentation of spot has a great impact on the intensity calculation. Measurements based on the pixel intensity, rather than the segmentation area such as Probability of error (p_e) and Normalized mean square error ($NMSE$), support the superiority of the FLICM against FCM, PCM and PFCM. When evaluating the results in intensity extraction perspective a lower value of p_e and $NMSE$ are expected for the better performance of the algorithm.¹² Hence the FLICM is better compared to other algorithms

Real microarray images downloaded from the UNC microarray data base²⁴ is used for segmenting spots FG from BG. All downloaded images are having 34 sub arrays with 625 spots in each sub array. The FLICM algorithm is applied on 25 such sub arrays. The Fig. 2. a shows a real image sub array obtained from the UNC microarray data base and Fig. 2 .b shows segmentation result using FLICM algorithm for real cDNA microarray image.

Table 1. The comparison of FCM, PCM, PFCM, FLICM algorithm based on segmentation matching factor (SMF), Probability of error (p_e) and Normalized mean square error ($NMSE$) for simulated microarray images with different levels of additive white Gaussian noise SNR(dB)

SNR(dB)	SMF				p_e				$NMSE$			
	FCM	PCM	PFCM	FLICM	FCM	PCM	PFCM	FLICM	FCM	PCM	PFCM	FLICM
1	65.689	65.329	66.282	81.864	0.140	0.145	0.126	0.049	0.363	0.396	0.337	0.181
2	71.321	70.821	72.513	86.230	0.118	0.120	0.096	0.030	0.303	0.323	0.275	0.138
3	77.217	76.622	78.322	90.359	0.084	0.091	0.071	0.027	0.235	0.256	0.217	0.096
4	83.197	82.557	84.153	92.641	0.041	0.060	0.038	0.020	0.178	0.196	0.158	0.074
5	88.594	87.794	89.329	94.677	0.032	0.034	0.028	0.015	0.128	0.143	0.107	0.053
6	92.304	90.786	93.799	95.854	0.018	0.020	0.017	0.009	0.069	0.073	0.062	0.041
7	95.142	94.326	96.363	96.780	0.009	0.011	0.008	0.007	0.040	0.044	0.036	0.034
8	96.226	96.164	97.032	97.332	0.006	0.007	0.006	0.006	0.031	0.032	0.030	0.027
9	97.342	97.243	98.169	98.469	0.005	0.005	0.004	0.004	0.019	0.020	0.018	0.015
10	97.975	97.967	98.552	98.652	0.004	0.004	0.003	0.003	0.015	0.016	0.014	0.013

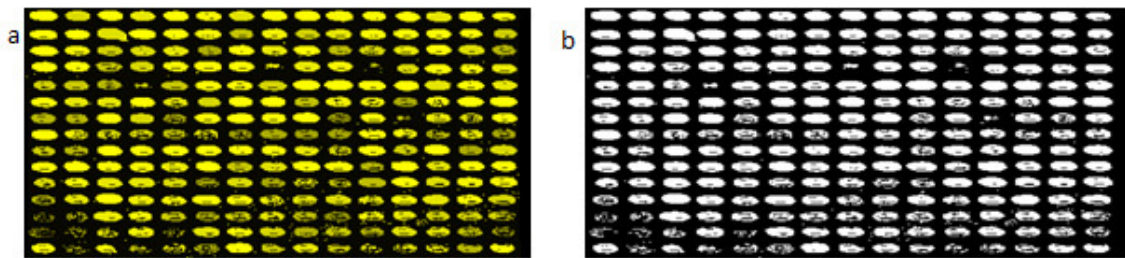


Fig. 1(a) A microarray simulated image with 225 spots; (b) Segmentation result obtained for the FLICM algorithm

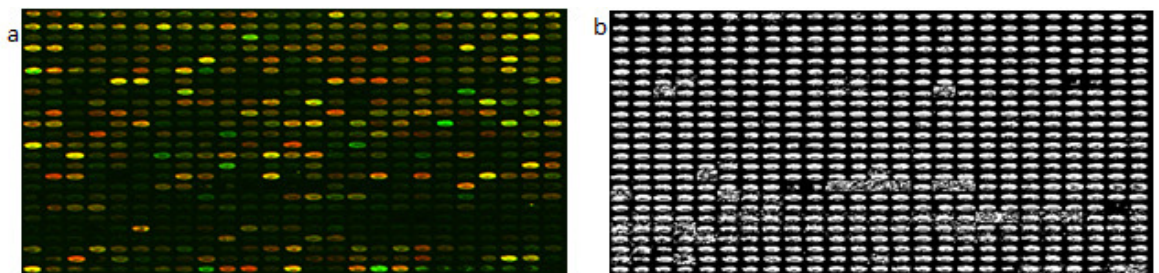


Fig. 2(a). Real cDNA microarray image; (b) Segmentation result using FLICM algorithm for real cDNA microarray image.

6. Conclusion

In this paper, existing fuzzy clustering image segmentation methods available in the literature have been tested for its suitability to perform for better microarray spot segmentation under noise. The algorithms are tested on both simulated and actual cDNA microarray images. The number of spots in the synthetic images used for the evaluation purpose and measures obtained support the superiority of the FLICM method over other existing fuzzy clustering methods for microarray image processing.

References

1. Yang YH, Buckley MJ, Dubois S, Speed TP. Comparison of methods for image analysis on cDNA microarray data. *J.Comput. Graphical Statist*, 2002; 11: p.108–136.
2. Lehmussola A, Ruusuvaari P, YliHarja A. Evaluating the performance of microarray segmentation algorithms. *Bioinformatics*, 2006; 22: p.29102917.
3. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 270, 1995; p.467- 470.
4. Eisen M B, ScanAlyze, 1999, <http://rana.lbl.gov/EisenSoftware.htm>.
5. GenPix 4000, A User's Guide, Axon Instruments, Inc, Foster City, CA.6.Buhler J, Ideker T, Haynor, D. *Dapple: improved techniques for finding spots on DNA microarrays*, Technical Report. UWTR 2000; 08, 05, UV CSE, Seattle, Washington, USA.
6. Buhler J, Ideker T, Haynor, D. *Dapple: improved techniques for finding spots on DNA microarrays*, Technical Report. UWTR 2000; 08, 05, UV CSE, Seattle, Washington, USA
7. Buckley M.J. The spot user's guide CSIRO Mathematical and Information Science,2000, <http://www.cmis.csiro.au/IAP/Spot/spotmanual.html>.
8. ImaGene, ImaGene 6.1 User Manual, http://www.biodiscovery.com/index/pappsweb_files/action.
9. Beucher S, Meyer F, Themorphological approach to segmentation The watershed transformation, *Opt. Eng.*, 1993; 34, p. 433–481.
10. Adams R, Bischof L, Seeded region growing, *IEEE Trans. Pattern Anal. Mach. Intel.*, 1994; 16, (6), p. 641–647.
11. Bozinov D, Rahenfuhrer J, Unsupervised technique for robust target separation and analysis of DNA microarray spots through adaptive pixel clustering, *J. Bioinform*, 2002; 18, p. 747-756.
12. Chen Y, Dougherty ER, Bittne ML. Ratio based decisions about the quantitative analysis of cDNA microarray images, *J. Biomed.Opt*, 1997; 2, p.264–374.
13. Wu S, Yan H. Microarray Image Processing Based on Clustering and Morphological Analysis in *Proc. Of First Asia Pasific Bioinformatics Conference, Adelaide, Australia, 2003*; p. 111-118.
14. Volkan U, Dhsan, OB. Microarray image segmentation using clustering methods, *Mathematical and Computational Applications* 15, (2) p.2 40-247.
15. Biju VG, Mythili P. A Genetic Algorithm based Fuzzy C Mean Clustering Model for Segmenting Microarray Images, *International Journal of Computer Applications*, 2012; 52, (11), p.42-48.
16. Nikil RP, Kuhu P, James MK, James CB. A Possibilistic Fuzzy CMeans Clustering algorithm, *IEEE Trans on Fuzzy systems*, 2005;13,(4), p. 517-530.
17. Krinidis S, Vassilios C.A Robust Fuzzy Local information C Means Clustering algorithm, *IEEE Transaction Image Processing* 2010;19,5.
18. Zacharia E, Maroulis D. An original Genetic approach to fully automatic gridding of microarray images, *IEEE transaction on medical imaging*, 2008; 2796, p.805-813.
19. The Math Works, Inc., Software, MATLABR (2010a), Natick, MA.
20. Krishnapuram R, Kellar JA. Possibilistic approach to clustering, *IEEE Trans, Fuzzy Systems*, 1993; 4,(3), p.98-110.
21. Athanasiadis EI, Cavouras DA, Spyridonos PP, Glotsos DT, Kalatzis IK, Nikiforidis GC. DNA microarray image processing based on the Fuzzy Gaussian mixture model, *IEEE Transaction on Information Technology in Biomedicine*, 2009;13, issue 4.
22. Athanasiadis EI, Cavouras DA, Spyridonos PP, Glotsos DT, Kalatzis IK, Nikiforidis GC. A Wavelet based markov random field segmentation mode in segmenting microarray experiments, *Computer methods and programs in biomedicine*, 2011;104, p. 307- 315.
23. UNC Microarray database. <https://genome.unc.edu>
24. Zacharia, E., Maroulis, D. 3 D Spot Modeling for Automatic Segmentation of cDNA Microarray Images, *IEEE, Transactions on Nano Bioscience*, 2010, 9, (3), p 181-192
25. Ping WYU, Maheshwar G. A comparison of fuzzy clustering approaches for quantification of Microarray Gene expression *Journal for signal processing systems*, 2008;50,(3), p. 305-320.