BIOSCIENCES BIOTECHNOLOGY RESEARCH ASIA, April 2015. Vol. **12**(1), 653-660

Multi-class Abnormal Breast Tissue Segmentation Using Texture Features and Analyzing the Growth Factor Using Power Law

B. Monica Jenefer[1]* and V. Cyrilraj[2]

[1]Computer Science & Engineering, Sathyabama University, Chennai, India.
[2]Department of Computer Science & Engineering,
Dr. M.G. R. Educational and Research Institute, Chennai, India.

DOI: http://dx.doi.org/10.13005/bbra/1709

(Received: 20 December 2014; accepted: 23 January 2015)

This paper motivated to design and develop an automatic model for multi-class breast tissue segmentation and find the growth of the cancer in breast mammogram images. Various breast tissues are categorized by a novel texture features such as PTPSA-[Piecewise Triangular Prism Surface Area], intensity difference and regular-intensity in mammogram images. Using CRF-[Classical Random Forest] method segmentation and classification of the features can be obtained in mammogram images. The input image feature values are compared to the ground-truth values for confirming the true positive rate of the proposed approach. Efficacy of abnormal breast tissue segmentation is evaluated using publicly available MIAS training dataset.In this paper, we investigate the consequences of an option, yet just as conceivable, suspicion of tumor growth, to be specific power growth law, acknowledged in direct expand in tumor breadth. We exhibit a simple model for tumor growth, whose global flow demonstrates power law growth of the tumor, much under boundless supplement supply. For corroboration, it is carried out and examined one-, two- and three-dimensional tumor growth tests both in vitro, in MCF-7 cells (breast malignancy cell line) and in vivo, in mouse xenografts.After successful tissue segmentation, the growth of the tissue is analyzed using Power Law. Performance evaluation of the proposed approach can be obtained by comparing the simulation output with the ground truth data. The accuracy of the proposed approach reaches up to 97% for MIAS database in term of tumor detection. Also, simulating radiotherapy under power law, Gompertz and exponential tumor growth, it is indicated that the power law model predicts profoundly diverse conclusions for the usually used treatment. This shows the significance of utilizing the proper tumor growth model when computing ideal measurement fractionation plan for radiotherapy.

Key words: Breast Cancer, Mammogram Images, Texture Features, Image Classification, MIAS dataset, Power Law Growth.

Nowadays breast cancer is one of the leading reasons for death among women and it is the second main reason after lung cancer[1-4]. Due to breast cancer the death rate is getting increased all over world surveyed by the World Health Organization. There is no efficient way to prevent the breast cancer[5, 6]. Mammography is a most important tool helping early noticing the breast cancer[7]. Since earlier treatment create most successful records in the medical field, analyzing, and finding earlier stage of breast cancer via mammogram is too good and necessary today. Various imaging techniques are available for investigating breast, consists of MRI, ultrasound, PET, CT, OPT, SY and X-ray based images. Within this Mammography [X-ray image] is the most important and common method for radiologists to

* To whom all correspondence should be addressed.
E-mail: bmonicajenefer@gmail.com

detect and analyze breast cancer[8,9]. Film mammography and Digital mammography are the two types of mammography presently used. Digital mammography is accepted happily by physicians because of the better quality of the image provides better performance in terms of True positive detection and analyzation.

Numerous numbers of approaches are already available to detect and classify the abnormality of mammogram images. In[15] a two-stage algorithm is applied for enhancing and detecting masses due to the pixel weight information. But the weight of the pixels may vary and not normalized in the image. In[16] region clustering method is proposed for image segmentation. Massive lesion classification using features was introduced in[17]. Apriori Algorithm combined with PCA based feature selection technique is used for detecting breast cancer[19]. One of a novel clustering method is proposed for segmenting the breast cancer[20], where this novel approach divides the image into pixels, then features of this pixel are mapped with the feature space. According to the mapping weight, the cancer portion is identified and detected.

Tarassenko *et al*[10], proposed an image segmentation technique based on region clustering. The mammogram is partitioned into clusters on the basis of data density. In each region the probability density is calculated using Parzen estimator, and the result of the image segmentation procedure is an image containing all possible regions of interest. The regions of interest are then presented to the human expert for further analysis. Bottigli *et al*[11] presented a comparison of some classification system for massive lesion classification. An algorithm based on morphological lesion differences was used to extract the features. The two classes (pathological or healthy ROIs) were differentiated by utilizing the features. A supervised neural network was employed to check the discriminating performances of the algorithm against other classifiers and the ROC curve was used to present the results [12,13]. A rough K-means clustering method is used for cancer segmentation in mammogram images[18].

Tumor growth has been intensively studied mathematically, as its dynamics crucially determine the success of oncotherapy. In radiotherapy and chemotherapy, where treatment is often given in fractions and the tumor regrows in the time intervals between cycles, it is important to take a tumor growth pattern into account when planning the scheduling of treatment. However, definitive understanding of macroscopic tumor growth dynamics is still lacking.

It is widely accepted that tumor growth may obey exponential, Gompertz or logistic growth laws (e.g.,[22, 23]) and clinical treatments are usually performed according to standard schedules, which were set based on these assumptions. However, these assumptions have never been biologically proven. A recently published paper doubts the validity of Gompertz growth, which motivates the current policy of adjuvant chemotherapy, and raises serious doubts about its underlying experimental evidence[24]. Moreover, new experimental data imply constant growth of tumor diameter, i.e., power law tumor growth[25], as was also inferred from mathematical analysis of mammography screening trial data[26]. Although several mathematical models for chemotherapy schedule planning considered different growth laws (e.g.,[27]), only few previous studies have suggested a power law tumor growth rate[28, 29, 26, 30]. In radiotherapy, one can analyze the dynamic effects of different radiation schedules using the relatively good clinical documentation and the relevant mathematical models. Available mathematical models of cancer radiotherapy have two parts: radiation effect module, for assessing tumor cell survival after a single dose of irradiation, and tumor re-growth model, which describes what happens to the tumor between successive radiation doses.

Power law model - linear growth of tumor diameter

In an earlier work[30], tumor growth was explored using a *CA* model, in the beginning presented in[30] for telling rising tissues. Numerical simulations of this model were conducted to attain the macroscopic dynamics of tumor growth. For the two-dimensional (2D) CA, it is observed that the time course of total number of cells in the unsaturated stage of growth (before space/ nutrients limitations are imposed) could be well approximated by a parabola, i.e., it is proportional to the square of time, suggesting power law tumor growth[30]. We also simulated a one-dimensional (1D) CA, which showed linear growth of the total

number of cells. Thus, we observed that both in 1D and 2D cases, tumor diameter grows linearly in time.

The fundamental idea of this paper is based on breast cancer segmentation works. In the previous work, SVM classifier modified AdaBoost algorithm was applied for cancer detection and segmentation. In contrast, this paper utilizes random forests[14] for multi-class abnormality of breast cancer classification and obtains it through a sequence of processes. Since MIAS dataset is already used benchmarking, it is restricted that the preprocessing steps to bias and inhomogeneity correction only. The entire functionality diagram of this paper is depicted in Fig.1.

MATERIALS AND METHODS

Preprocessing

To reduce the intensity bias of the mammogram images, intensity normalization is applied. The image processing gets failed due to intensity inhomogeneity. In this paper level-set based bias correction on a mammogram is applied for correcting the intensity inhomogeneity. And the histogram equalization method is applied to normalize the image.

PTPSA - Feature Extraction

From the input image spatial as well as non-local features are extracted to identify the cancer. To obtain accurate results a doctor suggested ground-truth images [one is tumor and the other is non-tumor image] are taken for comparison. Initially extract the pixel intensities from input image as well as from the ground-truth images and find out the differences between the pixel. The intensities of the breast mammogram areextracted as I1, I2,…, In. These difference features [d1 = I1 – GTI1] captures the amount of intensity variation at each point among the mammogram images. Also texture features as fractals, pixel intensities are extracted for analyzing the surface of the mammogram image.

Classification

Finally, both features of the input imageare compared with the ground truth image using Mahalanobis distance and compute the score. The less score used to classify and identify the input image is cancer affected image or normal image. If the image is decided as cancer affected image, then the segmentation method is applied to segment the cancer portion.

Segmentation

The threshold segmentation method is used to segment the cancerous portion of the input image. Since, conventional methods do segmentation on top of threshold value; the intensity inhomogeneity is corrected with the help of the bias correction method. Also the threshold value makes the difference between the foreground and background of the image. By choosing an appropriate and accurate threshold value, the estimation of the image for segmentation is easy and it helps to locate the cancer place, shape of the object correctly.

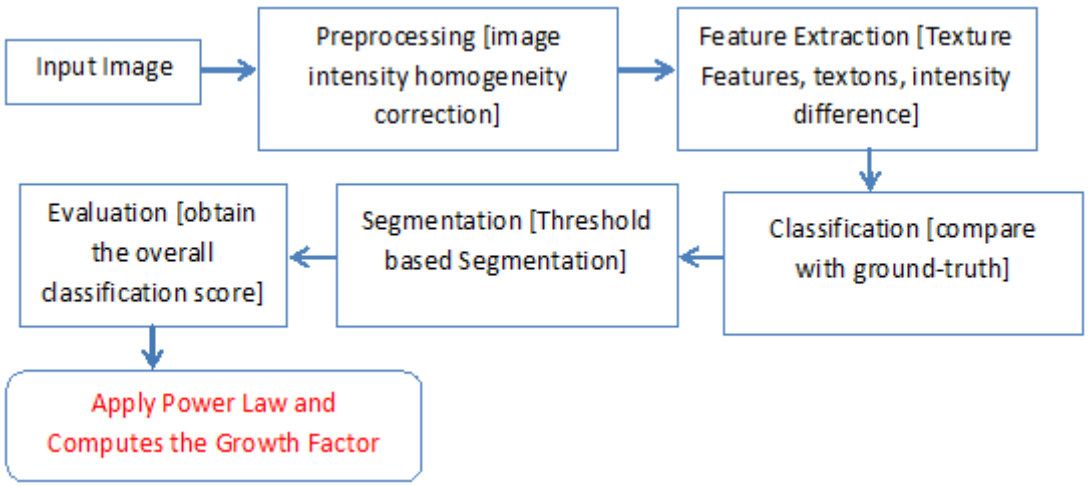


Fig. 1. Overall Functionality of the Proposed Approach

Abnormal Breast Cancer Tissue Classification

RF[14] method is used for classification of abnormal tissues in breast image. The RF is one of an ensemble method which generates different classifiers and aggregates their output to take a right decision. RF follows Tree-based-classification and includes an extra layer of randomness to bagging. Each branch of the tree is splited using the best from a randomly selected subset of predictors. This process helps RF to give better performance comparing with other classifiers like SVM, Multi-SVM and neural networks.

Also CFs is ensembles of binary classification tree. In each node$n$, the classification tree randomly takes a subset of training samples X_n and predicts a class $p_t^n(\omega|x)$, where p_t is the probability of the sample x in class ω. According to the features CF continuously divides the training samples at each node, then the division again X_L and X_R to the left/right nodes.The division is obtained in a random dimension on the feature set. Tree length continues until a defined depth, D_T. During the testing process, data points to be classified are dumped through each tree **t**, with the learned division function. The last child node probability is straightly used as the tree probability i.e.

$$p_t(\omega|x) = p_t^l(\omega|x)$$

Where p_t^l indicates the probability of at last-child node, l of sample x in class . The entire probability is calculated by the following equation,

$$p(\omega|x) = \frac{1}{T}\sum_{t=1}^{T} p_t(\omega|x)$$

Where $p(\omega|x)$, is the average probability of sample in class and is the total number of trees. At last the class with largest probability is estimated and it is the original class, i.e.

$$\hat{\omega} = \arg max_{\omega}\, p\,(\omega|x)$$

The cross-verification and cross validation is applied on entire training data. From the simulation, it is obtained that the 2D segmented tissue, which are predicted pixels and labeled from RF. These 2D abnormal tissues are segmented and shown in the following Figure.

Fractal Growth

The tumor growth model is developed using Power-Law[18], First, we define the three different growth models that we consider. The exponential model is given by the equation

$$\mathbf{N = rN}$$

Where N is the tumor size and r is the constant growth rate. The solution of this equation yields

$$\mathbf{N(t) = N(0)\exp(rt)}$$

where N(0) is the initial tumor size at t = 0. The Gompertz equation for tumor growth is

$$\mathbf{\dot{N} = -gN\ln(N/K)}$$

where$g$ is the growth rate parameter, and K is the tumor carrying capacity, i.e., the maximal size a tumor would reach without treatment. The solution for this equation is

$$\mathbf{N(t) = K\exp\left[\exp(-gt)\cdot\ln\left(\frac{N(0)}{K}\right)\right]}$$

The power law model, resulting from our previously described assumptions, takes the following 3D form:

$$\mathbf{\dot{N} = 3AN^{\frac{2}{3}}}$$

where$A$ is proportional to the linear growth parameter for the radius of the tumor. The solution for the latter equation is

$$\mathbf{N(t) = \left(At + N(0)^{\frac{1}{3}}\right)^3}$$

Table 1 shows simulation results for SF values after a conventional fractioned regimen, that is irradiation of 2Gy every 24hrs, for 30 days[12]. Tumor parameters α/β and $t2$ were varied, and SF values at the end of treatment were calculated for each of the three growth models. The α/β values were taken between 1 and 20, representing the known range of different tumors in humans[9]. The values of t2 ranged from 10 to 120 days, as mentioned above. Generally, under all growth laws, the fractionation of the dosage becomes less

efficacious with increasing tumor proliferation rate (smaller t2 values) and decreasing repair capacity. For exponential and Gompertz tumor growth models, our results are comparable to those of McAneney *et al*. [8]. The small differences are due to a slightly different definition of SF, calculated.

RESULTS AND DISCUSSION

In this paper the MATLAB software is used to implement the proposed approach algorithm, due to MATLAB is an efficient and appropriate language for education and research. The first step of the proposed approach is reading

Nomenclature

Symbols	Description
PL	Power law
CA	Cellular automata
D	diameter
t	time
N	Tumor size
A, â	Tumor parameters
G	Growth rate
K	Tumor carrying capacity
p	Growth density

Step 1: *Read the image.*

Step 2: *Convert image to gray [if necessary].*

Step 3: *Resize the image .*

Step 4: *intensity homogeneity correction applied*

Step 5:Feature extraction on the surface image

Step 6: *Apply random thresholding on the image.*

Step 7: *extract pixel intensities for input image as well as ground – truth image.*

Step 8: *compare both features and find similarity measurement*

using mahalanobis distance

Step 9: *apply cancer detection if the score is less for cancer image*

Step 10: *Detected Cancer portion is compared with the Growth Model*

3.3 ***ALGORITHM FOR Classification***

Step 1: *using CF method classify the two classes as normal or cancer affected*

Proposed Algorithm

the input image and the ground truth image; convert them into a two-dimensional matrix f(x, y), where x, y is the spatial coordinates of the image. Column-1 in table-1 shows the input image, and column-2 shows the ground truth image.

In this paper the total number of images taken is 100 where 50 images are normal and 50 images are cancer affected. According to the proposed approach each image is executed in MATLAB and find out the image class is shown in table-2.

In this paper the total number of images taken is 100 where 50 images are normal and 50 images are cancer affected. According to the proposed approach each image is executed in MATLAB and find out the image class is shown in table-2.

From table-2 it is clear that the proposed approach detects and classifies the normal images is 50 out of 50 input images and cancer affected images is 47 images out of 50 images from the database. From the table the sensitivity, specificity and accuracy are computed for evaluating the proposed approach.The performance can be evaluated by calculating the metrics True Positive Rate and False Positive Rate. The TPR, FPR can be calculated using:

$$TPR = \frac{Number\ of\ classification\ correctly\ obtained}{Total\ number\ of\ images\ to\ be\ classified}$$

$$TNR = \frac{Number\ of\ Normal\ images\ identified}{Total\ Number\ of\ normal\ images}$$

$$FPR = \frac{number\ of\ classification\ wrongly\ obtained}{Total\ Number\ of\ images\ to\ be\ classified}$$

$$FNR = \frac{number\ of\ abnormal\ images\ incorrectly\ identified\ as\ Normal}{Total\ Number\ of\ abnormal\ images}$$

From Table-2, the TPR, FPR can be calculated

$$TPR = \frac{47}{50} + \frac{50}{50} = 97\%$$

$$TNR = \frac{50}{50} = 100\%$$

$$FPR = \frac{3}{100} = 3\%$$

$$FNR = \frac{3}{50} = 6\%$$

Detecting and differentiating normal as well as cancer affected, and the obtained TPR, FPR values. The efficacy of the proposed method can be analyzed using sensitivity and specificity metrics. And it can be obtained using the following formula:

$$Sensitivity = \frac{TPR}{TPR + FNR} = \frac{97}{97 + 6} = 0.94\%$$

Table 1. Input image, Pre-Processed Image, Cancer Detected Image, Image Classification

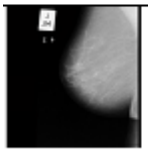
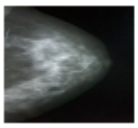
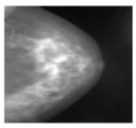
Input Image	Preprocessed Image	Cancer Detection	Type
			Normal
			Abnormal [Cancer Affected]

Table 2. Performance of Proposed Approach

Image Type	Data base Image	Proposed Image
Normal Image	50	50
Cancer Affected Image	50	47
Total	100	97

Table 3. Accuracy Comparison

Methods	Accuracy [in %]
CART [Ref-21]	94.63
CHAID [Ref-21]	93.66
QUEST [Ref-21]	91.22
Proposed Approach	95.63

Table 4. Growth of the Detected Tumor

Detected Tumor	<10 fractals	10 to 20 Fractals	20 to 35 fractals	35 to 50 fractals	50 and above fractals
47	3	7	17	15	5

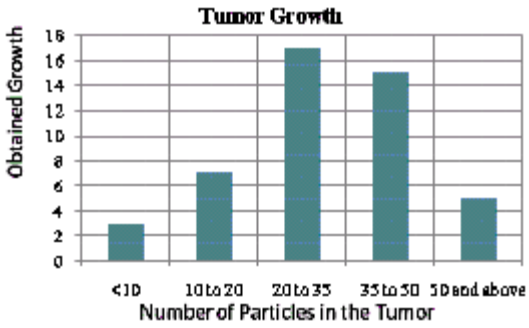


Fig. 2. Tumor Growth Analysis

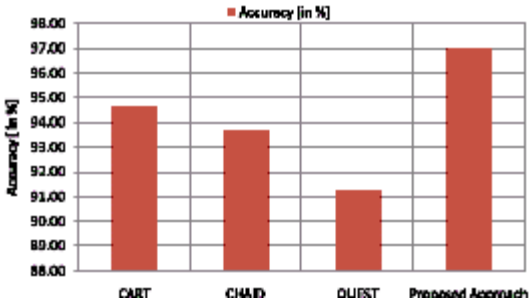


Fig. 3. Accuracy Comparison

$$Specificity = \frac{TNR}{TNR + FPR} = \frac{100}{100 + 3} = 97.08\%$$

Since the sensitivity and the specificity are comparatively better than the existing approaches it is decided that the proposed approach is efficient.

$$Accuracy = \frac{number\ of\ true\ postives + number\ of\ true\ negatives}{number\ of\ true\ positives + false\ positives + false\ negatives + true\ negatives}$$

$$Accuracy = \frac{97 + 100}{97 + 3 + 6 + 100} = \frac{197}{206} = 95.63\%$$

The obtained result in terms of accuracy of the proposed approach is compared with the existing methods CART, CHAID, QUEST [21] is given in Table-3 and depicted in Fig.3. From the Table-3, and Fig.3 it is clear that the proposed approach provides better accuracy in detecting cancer form mammogram images than the existing approaches.

The detected cancerousportion is compared with the power law growth model generated using MATLAB software is shown in table-4. From the comparison, it can be understood the growth of the tumor [severity] to take necessary action. The growth can be obtained by computing the number of fractals available in the tumor. There are 3 tumors are having less than 10 fractals, 7 tumors are having more than 10 and lesser than 20 fractals, 17 tumors are having more than 20 and lesser than 35 fractals, 15 tumors are having more than 35 and lesser than 50 fractal and finally, there are 5 tumors are having more than 50 fractals and it shows the severity of the tumor detected from the breast images.

CONCLUSION

There were 100 images [50 normal and 50 abnormal] were used for investigating the proposed approach. The bestperformances are achieved in classifying the pixels are 97% for the sensitivity and 97.03% for the specificity. This approach applied for classifying and detection finally comparing with the ground truth image to reduce the false rate of detection and classification. The results obtained from MATLAB code are satisfied and it further enhanced with classifying benign, malignant, non-tumor, tumor-starting, sever kind of images more accurately.Also, this paper can provide a significant way of finding the growth of the tumor detected from the mammogram images. This growth model helps the people to take necessary fast actions for preventing from major surgery and death.

REFERENCES

1. R.Mata, E.Nava,F. Sendra, "Micro calcifications detection using multi resolution methods, pattern Recognition,"2000,proceeings,15th International Conference, 2000; **4**; 344-347.
2. X. P. Zhang, "Multistage tumor detection and segmentation in mammograms," in *Proc. IEEE Int. Symp. Biomed.Imag.*, 2002; 213–216.
3. G.BharathaSreeja, Dr. P. Rathika, Dr. D. Devaraj" Detection of Tumors in Digital Mammograms Using Wavelet Based Adaptive Windowing Method''*international journal modern education and computer science*, 2012: **3**; 57-65
4. S.SahebBasha, Dr.K.Satya Prasad'' automatic detection of breast cancer mass in Mammograms using morphological operators And fuzzy c –

means clustering'' *journal of theoretical and applied information technology*

5. D. E. Stewart, A. M. Cheung, S. Duff, F. Wong, M. McQuestion, T. Cheng, L. Purdy, and T. Bunston, "Attributions of cause and recurrence in long-term breast cancer survivors," *Psycho-Oncology,* 2001; **10**: 179–183.
6. H. D. Cheng and H. Xu, "A novel fuzzy logic approach to mammogram contrast enhancement," *Inf. Sci.,* 2002; **148**; 167–184.
7. H. D. Cheng and Muyi Cu, "Mass Lesion Detection with a Fuzzy Neural Network", *J. Pattern Recognition,* 2004; **37**; 1189-1200.
8. R. Mousa, Q. Munib, and A. Moussa, "Breast cancer diagnosis system based on wavelet analysis and fuzzy-neural," *Expert Syst. Appl.,* 2005; **28**: 713–723.
9. X. Gao, Y. Wang, X. Li, and D. Tao, "On combining morphological compo-nent analysis and concentric morphology model for mammographic mass detection,"*IEEE Trans. Inf. Technol. Biomed.,* 2010; **14**(2): 266–273.
10. Tarassenko, L., Hayton, P., Cerneaz, N., Brady, M.: "Novelty detection for the identification of masses in mammograms". In: *Fourth International Conference on Artificial Neural Networks,* Cambridge, 1995 : 442–447.
11. Bottigli, U., Cascio, D., Fauci, F., Golosio, B., Magro, R., Masala, G.L., Oliva, P., Raso, G., Stumbo, S.: "Massive Lesions Classification using Features based on Morphological Lesion Differences". *Proceedings of World Academy of Science, Engineering and Technology,* 2006; **12**.
12. Zhang, W., *et al*.: "An improved shift-invariant artificial neural network for computerized detection of clustered micro calcifications in digital mammograms". *Medical Physics,* 1998; **23**: 595–601.
13. Lai, S.M., Li, X., Bischof, W.F.:"On techniques for detecting circumscribed masses in mammograms". *IEEE Transactions on Medical Imaging,* 1989; **8**: 377–386.
14. A. Cutler and L. Breiman, "Random forests-classification description," *Technical report,* University of California, Berkeley, 2004.
15. R.Mata, E.Nava,F. Sendra, "Micro calcifications detection using multi resolution methods, pattern Recognition,"2000,proceeings,*15th International Conference,* 2000; **4**: 344-347.
16. X. P. Zhang, "Multiscale tumor detection and segmentation in mammograms," in *Proc. IEEE Int. Symp. Biomed.Imag,* 2002: 213–216.
17. G.BharathaSreeja, Dr. P. Rathika, Dr. D. Devaraj'' Detection of Tumours in Digital Mammograms Using Wavelet Based Adaptive Windowing Method''international journal modern education and computer science 2012; **3**: 57-65.
18. R. Subash Chandra Boss, K. Thangavel, D. Arul Pon Daniel, "*Mammogram Image Segmentation Using Rough Clustering"*, Ijret-2013.
19. OnurInan, Mustafa SerterUzer and NihatY_lmaz, "A New Hybrid Feature Selection Method Based On Association Rules and PCAFor Detection Of Breast Cancer", ICIC International-2013.
20. Marius Iacomi, DonatoCascio, Francesco Fauci and Giuseppe Raso, "Mammographic images segmentation based on chaotic map clustering algorithm", BMC Medical Imaging-2014.
21. PushpalataPujari, JyotiBala Gupta, "Improving Classification Accuracy by Using Feature Selection and Ensemble Model", *IJSCE* – 2012.
22. Kozusko, F., Bajzer, Z.: Combining Gompertzian growth and cell population dynamics. *Math Biosci,* 2003; **185**: 153-67.
23. McAneney, H., O'Rourke, S.F.C.: Investigation of various growth mechanisms of solid tumour growth within the linear-quadratic model for radiotherapy. *Phys. Med. Biol,* 2007; **52**: 1039-54.
24. Retsky, M.: Metronomic Chemotherapy was Originally Designed and first used in 1994.
25. Bru, A., Albertos, S., Luis Subiza, J., Garca-Asenjo ,J.L., Bru, I. The universal dynamics of tumor growth. *Biophys J,* 2003; **85**: 2948-61.
26. Hart, D., Shochat, E., Agur, Z.: The growth law of primary breast cancer as inferred from mammography screening trials data, *Br. J. Cancer,* 1998; **78**; 382-387.
27. Panetta, J.C., Adam, J.: A mathematical model of cycle-specific chemotherapy Math. *Comput.Modelling,* 1995; **22**: 67-82.
28. Drasdo, D., Hoehme, S.: A single-cell based model to tumor growth in-vitro: Mono-layers and spheroids. *PhysBiol,* 2005; **2**: 133-47.
29. Gatenby, R.A., Frieden, B.R.: Information dynamics in carcinogenesis and tumor growth. *Mutat Res,* 2004; **568**; 259-73.
30. Vainstein, V., Kirnasovsky, O.U., Kogan, Y., Agur, Z.: "Strategies for cancer stem cell elimination: Insights from mathematical modeling". *J. Theor. Biol,* 2012; **298**: 32-41.