Disponible en ligne sur
ScienceDirect
www.sciencedirect.com
Elsevier Masson France
EM|consulte
www.em-consulte.com
IRBM 38 (2017) 228–233
IRBM

RITS 2017

Fibroglandular Tissue Quantification in Mammography by Optimized Fuzzy C-Means with Variable Compactness

A.L.M. Pavan [a], A. Vacavant [b], A.P. Trindade [c], D.R. de Pina [c,*]

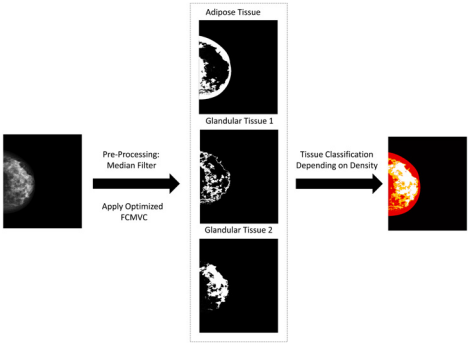
[a] *Department of Physics and Biophysics, Biosciences Institute of Botucatu, São Paulo State University, Distrito de Rubião Junior, Botucatu, São Paulo 18618-000, Brazil*

[b] *Institut Pascal, Université Clermont Auvergne, UMR 6602 UCA/SIGMA/CNRS, F-63171 Aubière, France*

[c] *Department of Tropical Diseases and Diagnostic Imaging, Botucatu Medical School, São Paulo State University, Distrito de Rubião Junior, Botucatu, São Paulo 18618-000, Brazil*

Received 14 March 2017; received in revised form 18 May 2017; accepted 18 May 2017
Available online 7 June 2017

Graphical abstract



Abstract

Background: Mammography is a wordwild image modality used to diagnose breast cancer, even for asymptomatic women. Due to its large availability, mammograms can be used to measure breast density and to predict cancer development.
Methods: We developed a methodology to estimate breast density using post-processed digital mammogram. Our automatic approach utilizes an optimized Fuzzy C-Means with variable compactness algorithm to classify and quantify fibroglandular tissue in mammograms.
Results: Fibroglandular tissue percentage estimation by our method has been compared with BI-RADS assessment from radiologist and achieved 67.8% of correct classification, with Spearman's correlation coefficient of $\rho = 0.618$, for $p < 0.001$. Furthermore, a Bland–Altman statistics showed no significant differences (bias of -0.20 ± 1.52) between both methods, indicating that the assessment widely used in clinical routine is consistent with the results generated by the algorithms. Cohen's kappa coefficient comparing the performance of the algorithm with the visual assessment for the different BI-RADS scores was 0.47 suggesting a moderate agreement.
Conclusion: Then, our methodology showed to be robust and accurate when compared with visual assessment. Furthermore, our methodology is fully automatic and reproducible, avoiding inter and intra observers variation, which has a potential to be implemented in clinical routine.

* Corresponding author.
E-mail address: drpina@fmb.unesp.br (D.R. de Pina).

http://dx.doi.org/10.1016/j.irbm.2017.05.002

1. Introduction

Breast cancer risk in women may be estimated from her age, history of breast biopsy, family history of breast cancer and breast density on an initial mammogram [1]. Studies have reported that women with increased mammographic density have a four- to six-fold increase in their risk of developing breast cancer [2,3].

In clinical routine, radiologists perform subjective visual assessments based on Breast Imaging Reporting and Data Systems (BI-RADS) density classification [2,3]. This is an important task to follow up patients under prevention therapy. However, BI-RADS density categories have limitations [4]. It has been suggested that subjective assessment of mammographic density shows variable intra- and inter-observer agreement [4,5]. Furthermore, BI-RADS categories are too coarse to monitor breast density changes in individual women [4]. For this reason, tasks such as selection of women who may benefit from supplemental screening exams and prediction of breast cancer risk, may be challenging with only subjective density assessment [3,6,7].

Therefore, developing quantitative methods of measuring breast density has become a relevant approach, which is less dependent on the observer [3,8]. Fully automatic methods for breast density (BD) measurements have been proposed to accurately quantify mammographic density [3,4,8,9]. The standard method to estimate BD in scientific research is the Cumulus software, which is an area-based approach [7,10]. However, these methods presents problems with accuracy and reproducibility [11]. In addition, these techniques have some limitations, such as dependence on the parameters of image formation, the need of a small calibration object to be imaged in each mammogram, and compressed breast thickness value [9,12,13]. Furthermore, most of these semi-automated methods necessitate an experience user to define a threshold stage, which is time-consuming [5,14] and implies in measurement variability [7].

Area-based and volumetric methods in literature are often applied in raw images. However, most medical centers achieve only the post-processed images for clinical purposes. Therefore, development of methodology to estimate BD in post-processed mammogram would be beneficial both in terms of direct clinical application and retrospective research-related studies [15,16].

An alternative modality for estimating content of tissues involves Magnetic Resonance Imaging (MRI). MRI has the potential to quantify fibroglandular tissue with a much higher degree of accuracy and precision than mammography [2,17] and, therefore, has been used to estimate BD. But MRI is costly, requires intravenous contrast agent administration, has a risk for contrast agent reactions, and is not well tolerated by all patients [6].

Breast is compressed during mammography exam, producing a projection image, which contains a variable combination of fibroglandular and adipose tissues. In general, pixels are neither pure fibroglandular nor pure adipose tissues [2]. Fuzzy C-Means (FCM) approach allows pixels to belong to multiple clusters with reasonable degree of membership grades [18]. Fuzzy C-Means with Variable Compactness (FCMVC) is an extended version of this algorithm, which improves pixel classification, when compared with popular FCM approach. Therefore, fibroglandular and adipose tissue may be classified with greater accuracy.

The aim of this work is to use an optimized FCMVC [19] algorithm to estimate the percentage of BD using digital mammography. The methodology, presented in Section 2, is applied in post-processed mammograms. Section 3 is dedicated to compare the BD results from the proposed algorithm with BI-RADS system, and shows its accuracy and reproducibility.

2. Materials and methods

2.1. Dataset

The present study was developed with ethical approval from the authors' institutions under protocol number 50547315.8.0000.5411.

Women aged 18 years or older undergoing screening mammography between 2013 and 2015 at Botucatu Medical School were included. For a woman who has been included in the study, she must have had no previous history of breast cancer or breast surgery, and had a BI-RADS assessment of either 1 or 2 (negative or benign finding, respectively). Only mammography of one side, left or right, was used for each patient in craniocaudal view, allowing assessment of different breast compositions. A total of 30 mammograms were evaluated.

The mammography system used was a Senographe 600T (GE Healthcare, Milwaukee, WI) with a CR-85X image digitizer (Agfa-Gevaert Group, Mortsel, BE). An Agfa image plate (18×24 cm^2 or 24×30 cm^2) with a pixel pitch of 50 μm was employed. A Mo/Mo anode/filter combination was used for all X-ray exposures. Processed images were used, since raw data is often not achieved in clinical practice [16].

For each evaluated mammography, an experienced radiologist rated breast density according to the qualitative BI-RADS density category (4th edition): category 1 implies breast tissue that is less than 25% glandular; category 2, breast tissue that is approximately 25–50% glandular; category 3, breast tissue that is approximately 51–75% glandular; and category 4, breast tissue that is more than 75% glandular.

2.2. Developed algorithm

The proposed algorithm has been developed in Matlab software to estimate fibroglandular breast tissue percentage. The algorithm is a fully automatic hybrid method, which uses different image processing techniques. It is based on FCMVC [19,20]. Detailed process is explained in the following subsections.

2.2.1. Preprocessing

In the first stage, the mammogram is segmented into the breast area by thresholding based on the gray-level intensity histogram, as literature employs [15], resulting in a mask with

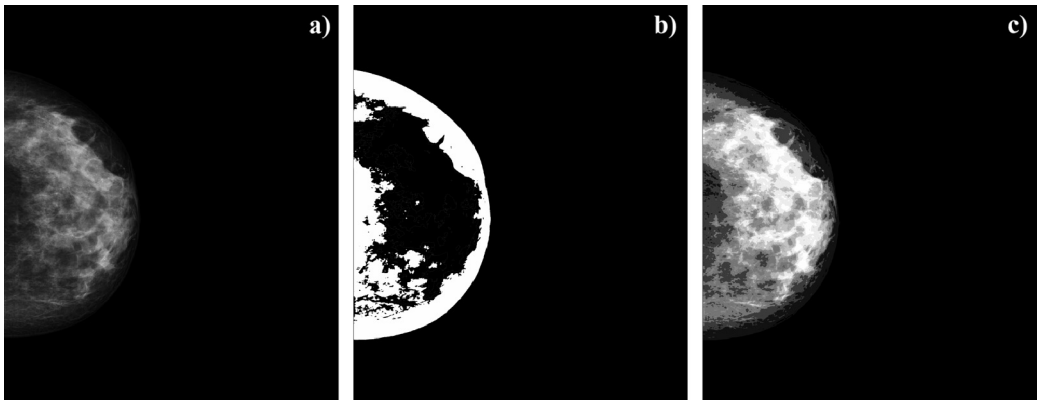


Fig. 1. a) Filtered image; b) Equalized image; and c) Rough fibroglandular mask.

the region of interest. A median filter is then applied in the breast area to reduce image noise.

2.2.2. Optimizing fuzzy C-means with variable compactness

FCMVC approach is based on FCM algorithm, which is an unsupervised clustering algorithm [19,20]. The advantage of FCMVC is that any cluster definition is improved by classifying pixels with a parameter called *compactness* (p_i). FCMVC is formulated as the minimization of the following objective function with respect to the fuzzy membership functions, the centroids and parameter p_i:

$$J_{FCMVC} = \sum_{j=1}^{n} \sum_{i-1}^{c} \hat{\mu}_{ij} (x_j - v_i)^{2p_i}$$

where n is the number of pixels which are classified into c clusters, $\mu_{ij} \in [0, 1]$ is the fuzzy membership of the input pixel x_i and v_i is the ith centroid associated to each cluster. Minimization of J_{FCMVC} gives the membership functions as:

$$\mu_{ij} = \frac{\left(x_j - v_i\right)^{-2p_i}}{\sum_{i=1}^{c} \left(x_j - v_i\right)^{-2p_i}}$$

For the implementation of FCMVC algorithm, it is necessary to introduce values for *compactness* and initial centroids. To do so, we optimize the algorithm by using information from equalized image, as described below.

The parameter p_i is a measure of compactness of a cluster, which should be large for small classes and small for large classes [19,20]. Therefore, for mammographic image, it is necessary to use different compactness values depending on breast type. As an example, adipose breasts has low amount of fibroglandular pixels representing a small group. Thus, in this example, adipose and fibroglandular clusters should have small and large compactness values, respectively. An exhaustive search was performed to estimate $\boldsymbol{p_i}$, such that pixel classification generated could segment fibroglandular and adipose tissue efficiently.

For the purpose of the current study, the algorithm preclassifies mammograms combining the two top and bottom BI-RADS category as "dense" and "fat" groups.

The classification is based on skewness values from equalized histogram image. "Fat" group has greater skewness values than "dense" group. Then, a threshold was stablished to automatically differentiate breast groups. Thus, a different compactness value is applied for each breast group.

In FCMVC, the energy function is minimized if high membership values are assigned to observations close to centroids and low membership values to observations far away from centroids. In each iteration, the centroid values are updated to achieve pixel classification. However, the choice of initial centroid value is important to reduce bias classification, to minimize number of iterations, time and computational cost.

In this work, initial centroid values are estimated for each image. This step is represented in Fig. 1. The segmentation of dense tissue is roughly estimated by a thresholding process in the equalized image (Fig. 1b), resulting in a fibroglandular mask (Fig. 1c). This mask is employed in the filtered image (Fig. 1a) to have the mean pixel intensities values. The mean adipose pixel values are computed with the remaining pixels. These values are then used to estimate the initial centroids in FCMVC.

2.2.3. Applying fuzzy C-means with variable compactness

FCMVC is applied in filtered mammography image to segment tissues in three different clusters, using the centroids and compactness parameters achieved before.

Defuzzification process was made by alpha-cut technique, using 0.95 as thresholding value. According to breast groups, classified in Section 2.2.2, each cluster represents different tissues. For "fat" breast group, clusters 1 and 2 represent pixels from adipose tissue; and cluster 3, fibroglandular tissue. On the other hand, for "dense" breast group, cluster 1 represents adipose tissue, while clusters 2 and 3 correspond to fibroglandular tissue.

As tissues are represented by different clusters in each mammogram groups ("fat" and "dense"), a different weight is given to clusters to calculate the percentage of fibroglandular tissue.

As a result, algorithm estimates the fibroglandular tissue percentage (FTP) in mammograms. FTP was computed by dividing the fibroglandular tissue area by total breast area.

Table 1
Confusion matrix to compare breast groups using results from skewness versus BI-RADS density score.

BI-RADS	Algorithm		
	Fat	**Dense**	
Fat	16	1	17
Dense	4	9	13
	20	10	30

Table 2
Confusion matrix for breast density using our algorithm versus BI-RADS density score.

BI-RADS	Algorithm				
	1	**2**	**3**	**4**	
1	1	3	0	0	4
2	0	12	1	0	13
3	0	1	7	0	8
4	0	3	2	0	5
	1	19	10	0	30

3. Results

From BI-RADS assessment made by the radiologist 4 mammograms were classified as category 1, 13 as category 2, 8 as category 3 and 5 as category 4, totalizing 30 mammograms.

As a result of using skewness to previously separate breasts into "fat" and "dense" groups, we obtained 83.3% of correct classification. A $t-$test, comparing skewness values, shows significant difference ($p < 0.05$) between the "dense" and "fat" breast groups. Table 1 shows the confusion matrix to compare breast groups classified using results from skewness versus BI-RADS density score given by the radiologist. Cohen's kappa coefficient comparing breast groups was 0.65 suggesting good agreement.

The estimated FTP by our method was compared with the assessment made by radiologist using BIRADS system, for each evaluated image, and achieved a 67.8% rate of correct classification.

Agreement between the BI-RADS density category and estimated FTP by the proposed method were determined with the Spearman's correlation coefficient. This results showed a highly significant positive correlation with visual assessment, with Spearman's correlation coefficient of $\rho = 0.618$, for $p < 0.001$. Table 2 shows the confusion matrix for BD using results from our algorithm versus BI-RADS density score given by the radiologist. Furthermore, Cohen's kappa coefficient comparing the performances of the algorithm with the visual assessment for the different BI-RADS scores was 0.47 suggesting a moderate agreement.

In addition, Bland–Altman statistics were used to evaluate agreement between both assessments, to quantify the amount and direction of bias, and to determine the upper and lower limits of agreement (bias ± 1.96 of the difference). Fig. 2 shows the Bland–Altman plot of the score difference between FTP estimated by our method and the radiologist evaluations.

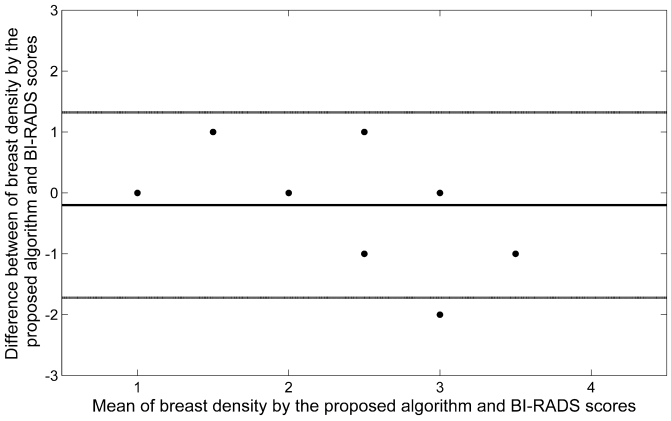


Fig. 2. Bland–Altman plot for brest density assessment. The difference refers to the BI-RADS minus the algorithm assessment. The difference between BI-RADS and algorithm scores was compared with the average score between the radiologist and computational results. Dashed lines represents the interval of 2 standard deviations. The middle line represents bias of -0.20 ± 1.52.

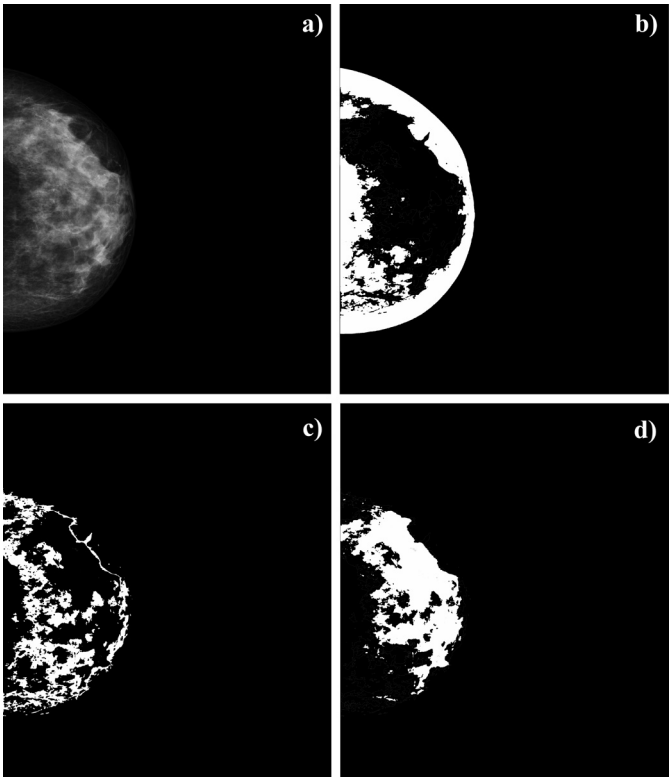


Fig. 3. Mammogram assessed as category 3 by BI-RADS. a) Filtered mammography; b) Cluster 1: adipose tissue; c) Cluster 2: almost completely fibroglandular tissue; and d) Cluster 3: fibroglandular tissue.

An example of FCMVC as a tool to segment tissues is showed in Fig. 3 for a "dense" breast group. Fig. 3a shows a mammogram classified as category 3 by radiologist BI-RADS evaluation. Figs. 3b–d illustrate membership function for the three clusters. In this example, Fig. 3b shows an adipose cluster while Figs. 3c–d shows fibroglandular clusters.

4. Discussion

We developed a methodology to estimate BD based on post-processed digital mammogram. Our automatic approach utilizes an optimized FCMVC algorithm to classify and quantify fibroglandular tissue in mammograms. FTP estimation by our method was compared with BI-RADS assessment from radiologist.

In relation with the results of breast groups pre-classification, skewness values could be used with high accuracy to significantly differentiate breast as "fat" and "dense". As mammographic images has high resolution, statistics may reliably be performed in histogram analysis [15], such as skewness measurement. This classification was important to stablish compactness and initial centroid values. Therefore, compactness values could be correctly assigned for each breast group. Furthermore, initial centroids values resulted in a few iteration in FCMVC, showing their good estimations. The dispended time by radiologists to make BI-RADS assessment may vary in the clinical practice according to radiologist experience, breast composition, and others [21]. The developed method showed be fast and easy to run, which implied in a short computational time ($\sim$ 130 s per exam in a mammogram with approximately 4700×5600 pixels).

Comparing the proposed method with BI-RADS, results suggest that this computational procedure offers a reliable, objective, and precise method that can be used to supplement visual grading, thereby providing a more advanced method for assessing BD in mammograms. Different from BI-RADS assessment, algorithm results presented no score equal to 4, as shows Fig. 2. Literature shows that when subjective visual evaluation is used, radiologists tend to give the maximum value [7]. Quantitative methods to evaluate breast density tend to underestimate values when compared to BI-RADS and MRI assessment [3,4,7], and our results corroborate with this.

From the statistical analyses comparing both methods, there was a positive association between FTP and the BI-RADS density scale. The rate of correct classification and Spearman's correlation coefficient found in this work represents a good association and low dispersion between methods. Results from confusion matrix and Cohen's kappa coefficient comparing the performances of the algorithm with the visual assessment for the different BI-RADS suggests a moderate agreement. The Bland–Altman analysis showed no significant differences (bias of -0.20 ± 1.52) between both methods, indicating that the assessment widely used in clinical routine is consistent with the results generated by the algorithms. These differences were sufficiently small to have the same confidence level for the results for both methods.

Fig. 3 shows an example of original and the three clusters used in our method. Visual examination of the clusters (Figs. 3b–d) revealed good segmentation of tissues.

One of the limitations of our study is that our approach is based in projected area, since it does not take the thickness of dense tissue into account. Dense volume is expected to be a more 'biologically relevant' measure and to be a better predictor of breast cancer risk [3,7,22]. However, no volumetric method for mammography has been used as standard method yet [23]. Furthermore, differences in breast positioning, compression, and technical parameters have also been suggested as factors that could influence the apparent volumetric density of a mammogram [6,14]. Other limitation is that only one radiologist evaluated mammograms accordingly with BI-RADS and, therefore, intra and inter variability could not me assessed. Furthermore, future works should be done to evaluate the performance of our methodology using different mammography systems with different imaging parameters.

Automatic BD measurements would be easier to implement in screening programs due to the tendency to be less time-consuming and labor-intensive than visual assessment with BI-RADS breast density [7]. Objective BD measurements is important to develop breast cancer risk models and may be used in the development of personalized screening protocols [3]. The methodology presented herein, is free from variability inter- and intra-observer. In conclusion, our study shows that it is feasible to obtain automatic measurements of BD from digital mammograms.

This encouraging result invites us to improve our algorithm. We expect that it will have the capability to segment breast tissue objectively and accurately, avoiding subjective assessment by BI-RADS.

Conflict of interest statement

The authors declare that there is no conflict of interest for this study.

Acknowledgements

The authors thank all of the clinical and technical personnel of our institution radiodiagnostic facility. Financial support was provided by CNPq and CAPES under process PDSE n° 88881.132793/2016-01.

References

[1] Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. Ann Intern Med 2011;155:10–20.

[2] Lu LJ, Nishino TK, Johnson RF, Nayeem F, Brunder DG, Ju H, et al. Comparison of breast tissue measurements using magnetic resonance imaging, digital mammography and a mathematical algorithm. Phys Med Biol 2012;57:6903–27.

[3] Gubern-Merida A, Kallenberg M, Platel B, Mann RM, Marti R, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms: a validation study. PLoS ONE 2014;9:e85952.

[4] Wang J, Azziz A, Fan B, Malkov S, Klifa C, Newitt D, et al. Agreement of mammographic measures of volumetric breast density to MRI. PLoS ONE 2013;8:e81653.

[5] Seo JM, Ko ES, Han BK, Ko EY, Shin JH, Hahn SY. Automated volumetric breast density estimation: a comparison with visual assessment. Clin Radiol 2013;68:690–5.

[6] Freer PE. Mammographic breast density: impact on breast cancer risk and implications for screening. Radiographics 2015;35:302–15.

[7] van der Waal D, den Heeten GJ, Pijnappel RM, Schuur KH, Timmers JM, Verbeek AL, et al. Comparing visually assessed BI-RADS breast density

and automated volumetric breast density software: a cross-sectional study in a breast cancer screening setting. PLoS ONE 2015;10:e0136667.

[8] van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. IEEE Trans Med Imaging 2006;25:273–82.

[9] Alonzo-Proulx O, Jong RA, Yaffe MJ. Volumetric breast density characteristics as determined from digital mammograms. Phys Med Biol 2012;57:7443–57.

[10] Kallenberg MG, Lokate M, van Gils CH, Karssemeijer N. Automatic breast density segmentation: an integration of different approaches. Phys Med Biol 2011;56:2715–29.

[11] Malkov S, Mahmoudzadeh AP, Kerlikowske K, Shepherd J. Automated volumetric breast density derived by statistical model approach. In: Breast imaging, vol. 8539. 2014. p. 257–64.

[12] Highnam RP, Brady JM. Mammographic image analysis (computational imaging and vision). Dordrecht: Kluwer Academic Publishers; 1999.

[13] Pawluczyk O, Augustine BJ, Yaffe MJ, Rico D, Yang J, Mawdsley GE, et al. A volumetric method for estimation of breast density on digitized screen-film mammograms. Med Phys 2003;30:352–64.

[14] Alonzo-Proulx O, Mawdsley GE, Patrie JT, Yaffe MJ, Harvey JA. Reliability of automated breast density measurements. Radiology 2015;275:366–76.

[15] Keller BM, Nathan DL, Wang Y, Zheng Y, Gee JC, Conant EF, et al. Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation. Med Phys 2012;39:4903–17.

[16] He W, Hogg P, Juette A, Denton ER, Zwiggelaar R. Breast image preprocessing for mammographic tissue segmentation. Comput Biol Med 2015;67:61–73.

[17] Klifa C, Carballido-Gamio J, Wilmes L, Laprie A, Shepherd J, Gibbs J, et al. Magnetic resonance imaging for secondary assessment of breast density in a high-risk cohort. Magn Reson Imaging 2010;28:8–15.

[18] Kannan SR, Ramathilagam S, Devi R, Sathya A. Robust kernel FCM in segmentation of breast medical images. Expert Syst Appl 2011;38:4382–9.

[19] Vacavant A, Ali A-R, Grand-Brochier M, Albouy-Kissi Aeı, Alfidja A, Chabrot P, et al. Smoothed shock filtered defuzzification with Zernike moments for liver tumor extraction in MR images. In: International conference on image processing theory, tools and applications, IPTA; 2015.

[20] Roy S, Agarwal H, Carass A, Bai Y, Pham DL, Prince JL. Fuzzy C-means with variable compactness. Proc IEEE Int Symp Biomed Imaging 2008;4541030:452.

[21] Ciatto S, Houssami N, Apruzzese A, Bassetti E, Brancato B, Carozzi F, et al. Categorizing breast mammographic density: intra- and interobserver reproducibility of BI-RADS density categories. Breast 2005;14:269–75.

[22] Morrish OW, Tucker L, Black R, Willsher P, Duffy SW, Gilbert FJ. Mammographic breast density: comparison of methods for quantitative evaluation. Radiology 2015;275:356–65.

[23] Tagliafico A, Tagliafico G, Astengo D, Airaldi S, Calabrese M, Houssami N. Comparative estimation of percentage breast tissue density for digital mammography, digital breast tomosynthesis, and magnetic resonance imaging. Breast Cancer Res Treat 2013;138:311–7.