Shape analysis for classification of breast nodules on digital ultrasound images

Hanung Adi Nugroho[1], Hesti Khuzaimah Nurul Yusufiyah[2], Teguh Bharata Adji[3], Widhia K.Z Oktoeberza[4]
[1,2,3,4]Department of Electrical Engineering and Information Technology, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia.
[2]Department of Physic, Faculty of Science and Technology, UIN Walisongo Semarang, Indonesia.
[4]Department of Informatics, Faculty of Engineering, Universitas Bengkulu, Indonesia.

Article Info

Article history:

Received Sep 14, 2018
Revised Nov 15, 2018
Accepted Nov 29, 2018

Keywords:

Breast nodules
Invariant moments
Shape analysis
Ultrasound image

ABSTRACT

One of the imaging modalities for early detection of breast cancer malignancy is ultrasonography (USG). The malignancy can be analysed from the characteristic of nodule shape. This study aims to develop a method for classifying the shape of breast nodule into two classes, namely regular and irregular classes. The input image is pre-processed by using the combination of adaptive median filter and speckle reduction bilateral filtering (SRBF) to reduce speckle noises and to eliminate the image label. Afterwards, the filtered image is segmented based on active contour followed by feature extraction process. Nine extracted features, i.e. roundness, slimness and seven features of invariant moments, are used to classify nodule shape using multi-layer perceptron (MLP). The performance of the proposed method is evaluated using 105 breast nodule images which comprise of 57 regular and 48 irregular nodule images. The results of classification process achieve the level of accuracy, sensitivity and specificity at 96.20%, 97.90% and 94.70%, respectively. These results indicate that the proposed method successfully classifies the breast nodule images based on shape analysis.

Corresponding Author:

Hanung Adi Nugroho,
Department of Electrical Engineering and Information Technology,
Faculty of Engineering, Universitas Gadjah Mada,
Jl. Grafika 2 Kampus UGM, Yogyakarta 55281, Indonesia.
Email: adinugroho@ugm.ac.id

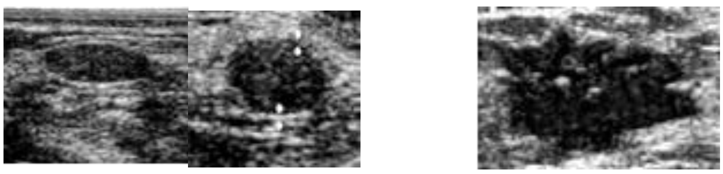
1. INTRODUCTION

Nowadays, the ultrasound examination results have a low accuracy rate of diagnosis due to the different interpretations of sonogram readings among radiologists. This issue is caused by the presence of speckle noise in the sonogram image [1, 2]. Therefore, a decision support system that can minimise the differences among radiologist interpretations is deemed necessary to objectively distinguish between benign and malignant nodules. A computer aided diagnosis (CAD) has been developed to assist radiologists in making a diagnosis in which the results of CAD are able to provide an objective information to radiologists [1]. The general principle of CAD consists of the following stages: pre-processing, segmentation, feature extraction and feature selection and the classification.

Various research works related to enhance breast ultrasound image have been developed. Wu et al. used speckle reduction bilateral filtering (SRBF) to overcome speckle noises while preserving the information of image. However, the SRBF method cannot eliminate the image label [3]. Otherwise, image label can be removed by using adaptive median filter as shown in [4, 5]. For segmentation, many research works have proven that the active contour had well performance to be applied in ultrasound images [6-10]. For classification, analysis of texture features such as histogram statistic, grey level co-occurrence matrices

(GLCM) and grey level run length matrices (GLRLM) were used by Nugroho et al. for classifying thyroid nodules into solid and cystic classes [11]. Some features including roundness, convexity, solidity and aspect ratio, play important role for recognising nodule shape. In several studies, Zernike and invariant moments were also used as features for nodule shape analysis [5, 12-15]. However, not all features have significant correlation to yield accurate classification. Some classifiers, such as support vector machine (SVM) [16, 17] artificial neural network (ANN) [1, 12], k-nearest neighbour (KNN), random forest and Naïve Bayes [12] have been widely used for classifying ultrasound images.

The objective of this study is to develop a method for classifying the nodule shapes of breast ultrasound images into two classes, i.e. regular and irregular classes as illustrated in Figure 1. Characteristics of breast nodule determines the malignancy level of breast cancer. The regular nodule indicates the benign cancer while irregular nodule indicates the malignant one. The combination of adaptive median filter and SRBF is proposed to preserve the information quality of images. Moreover, nine shape features consisting of roundness, slimness and seven features of invariant moments, are extracted to classify the nodule shape of breast ultrasound image.



(a) (b)

Figure 1. The characteristic of breast nodule (a) Regular (b) Irregular

2. RESEARCH METHOD

The digital scanning of breast ultrasound images taken from the Radiology Department of Sardjito and Hardjolukito Hospitals, Yogyakarta, Indonesia, were used in this study. The data were acquired using USG Logic c5 Premium, Voluson. The dataset consisted of 105 RGB images in bitmap format, with 57 regular and 48 irregular nodule images. The 57 regular nodule images comprised of 11 oval and 46 round shapes. A number of 67 images were used as data training and 38 images as data testing.

The radiologists were involved in this study as expert for validating the dataset. The nodule shape appears dark collectively than others. This study consists of four main stages, i.e. pre-processing, segmentation, features extraction and classification as shown in Figure 2.

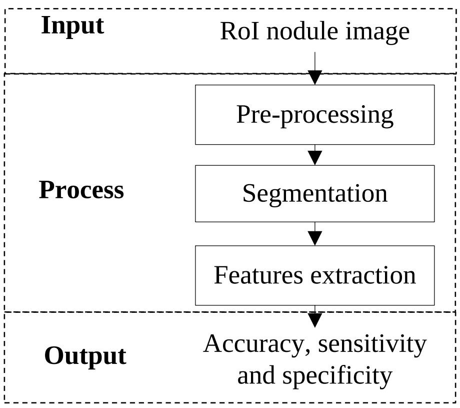


Figure 2. Block diagram of the proposed scheme

2.1. Pre-processing

In the breast ultrasound images, nodule shape appears as a concentrated dark hole area. Firstly, the radiologist marked the specific nodule area of the original image to obtain region of interest (RoI) and focused on the area analysis as depicted in Figure 3. Then, RoI RGB image was converted to greyscale followed by filtering process using the combination of adaptive median filter and SRBF to overcome label and speckle noises.

Adaptive median filter works based on adjacent pixels and is capable to preserve the detailed information of the focused object while conducting noise reduction [14, 15]. Similar to Gaussian filter, SRBF is started by measuring weight of Gaussian in spatial domain and range of intensity. The SRBF is mathematically expressed by (1). Here, f is the input image and h is the output image. The spatial adjacent of coordinate p pixel is described by Ω (p), while ξ is the variable combination representing of Ω coordinate pixels. The spatial weight of Euclidean distance between c and ξ is functioned by c, while s is the weight that operates intensity domain (weight intensity).

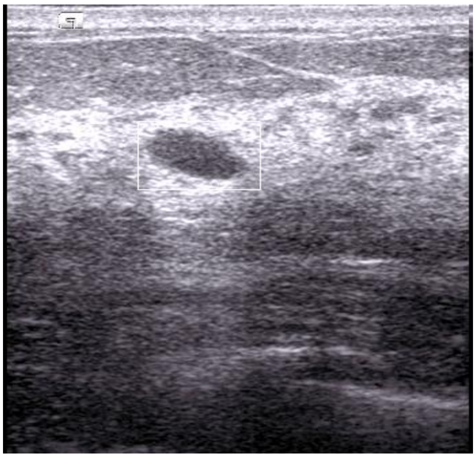


Figure 3. The RoI marked by radiologists

$$h(p) = \Gamma^{-1}(p) \int_{\Omega(p)} f(\xi)\, c(\xi, p) s\big(f(\xi), f(p)\big) d\xi \tag{1}$$

2.2. Segmentation

Segmentation based on active contour is conducted for segmenting nodule area and separating from its background. The concept of segmentation process is employed by grouping similar pixels or sub-region into the larger region. It is started by determining the iteration number used and setting the starting point which is known also as seed point. Then the starting point spreads gradually in sub region (region growing) based on the number of iteration used. In this work, the maximum number of iteration is set to fifty iterations. Finally, segmented nodule known as foreground area is obtained; otherwise it can be categories as background area.

In almost all cases, nodule areas obtained from radiologists mostly appear in the centre RoI image. Therefore, the starting point is determined from the centre of the RoI image. However, false positive area still occurs in the segmented image of active contour. Thus, opening morphological operation is applied to overcome this problem.

2.3. Feature extraction

Feature extraction and feature selection are conducted to obtain important features related to shape analysis. These features are used for classification process. The extracted features are Zernike moments, invariant moments, roundness and slimness. Zernike moments are the basis of Zernike polynomials from x2 + y2 ≤ 1 circle [18] as formulated in (2). Notation r is the radius of the (y, x) to the centre of mass, θ is the angle between r and the x-axis, and Rpq is radial orthogonal polynomials.

$$V_{pq}(x, y) = U_{pq}(r \cos\theta, r \sin\theta) = R_{pq}(r).exp(jq\theta) \tag{2}$$

Seven features of invariant moment [19] are denoted in (3) – (9).

$$\emptyset 1 = (\eta 20 + \eta 02) \tag{3}$$

Roundness known as compactness feature is a simple and popular measure of the efficiency of a shape contour. Slimness or aspect ratio is obtained from the ratio between width and length of the shape. Roundness and slimness is formulated by the (4) and (5).

$$roundness = \frac{4\pi \times Area}{perimeter^2} \quad (4)$$

$$slimness = \frac{width}{length} \quad (5)$$

2.4. Classification

A major task after feature extraction is to classify the object into one of the several categories. Multi-layer perceptron (MLP) is an example of an artificial neural network that is used extensively for the solution of a number of different problems. It is a layered network comprising input nodes, hidden nodes and output nodes as shown in Figure 4.

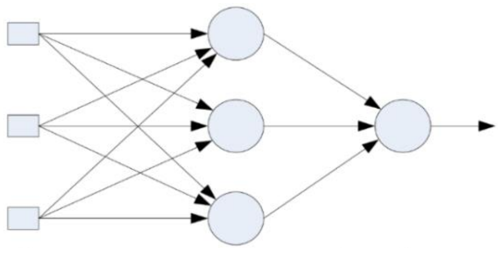
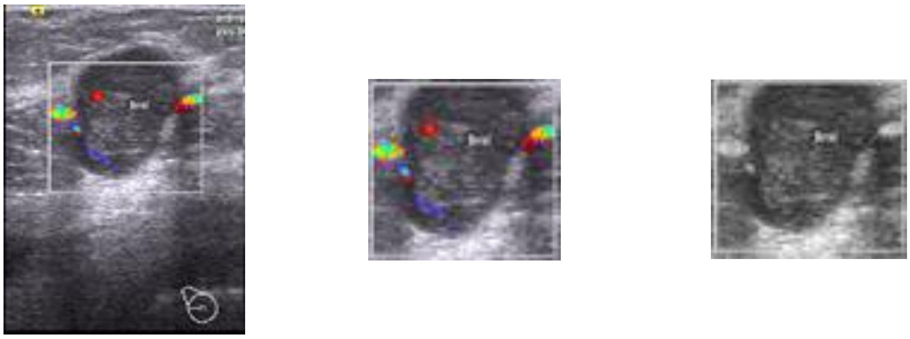


Figure 4. Multi-layer perceptron

3. RESULTS AND ANALYSIS

To facilitate the processing of breast ultrasound image, original RGB format was converted into greyscale as described in Figure 5. Filter process was conducted to overcome some common problems in ultrasound images, such as presence speckle noises and labels. Selection of filtering technique was also based on measurement speckle index (SI) as shown in Figure 6.



(a) (b) (c)

Figure 5. The conversion result (a) original image (b) RoI RGB image (b) greyscale of RoI image

As shown in Figure 6, the combination of adaptive median filter and SRBF obtains the lowest speckle index (SI). It indicates that combination of adaptive median filter and SRBF is more appropriate to overcome speckle noise than the other filtering techniques. In addition, the combination of adaptive median and SRBF method is capable to eliminate the image label as shown in Figure 7.

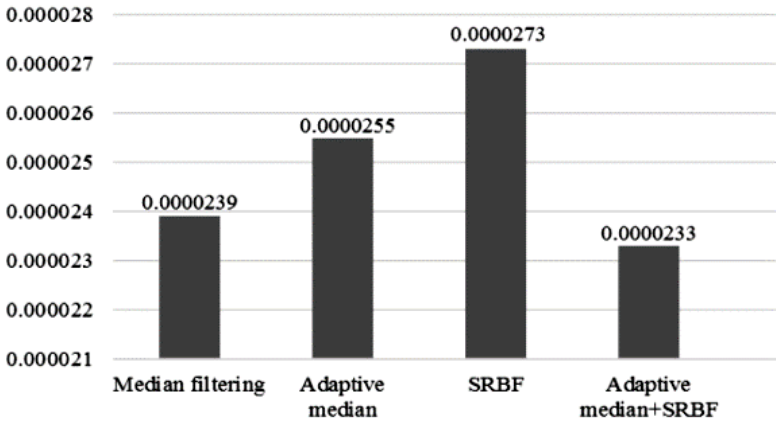


Figure 6. The comparison speckle index of several filtering methods

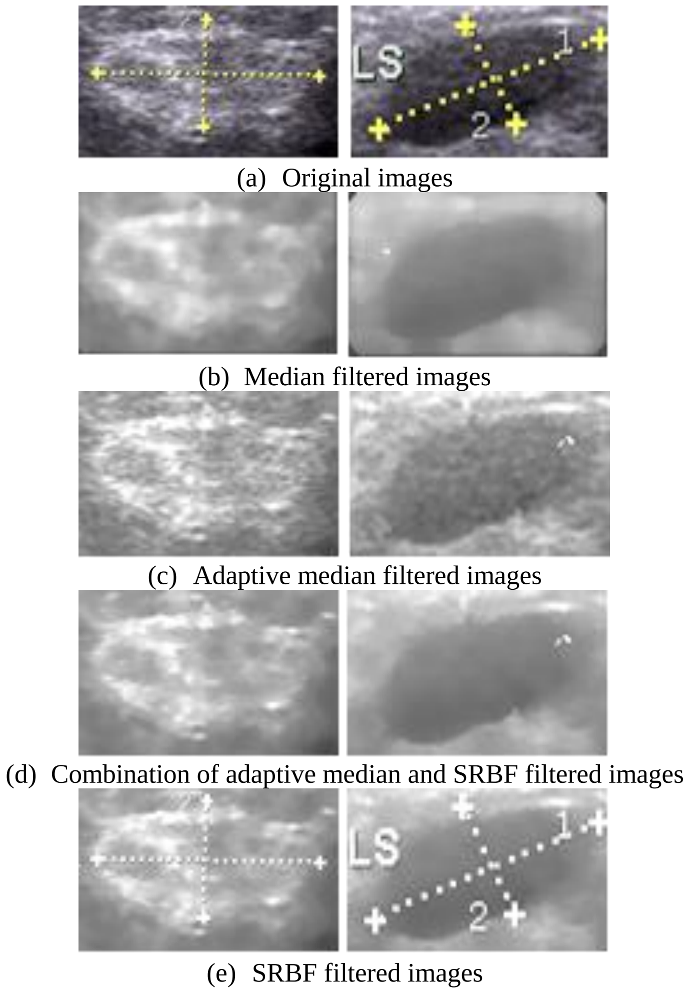


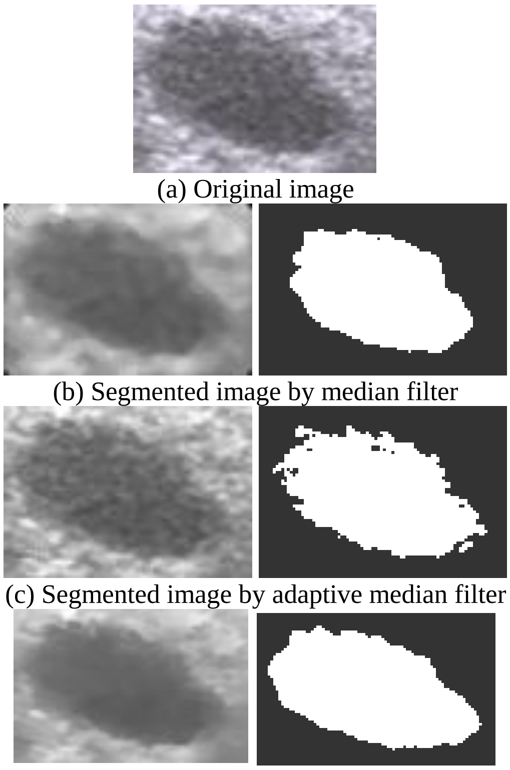
Figure 7. Comparison of original images and filtered images using various filtering methods

Filtered images subsequently underwent active contour-based segmentation. The visual comparison of segmentation result can be seen in Figure 8. The segmented images of adaptive median and SRBF filtered have the closest result to the actual nodule images assessed by radiologists. Segmented nodule was then processed to take its features. Feature extraction process obtained nine features including of roundness, slimness and seven features of invariant moment. Further, extracted features are classified based on MLP classifier. The performance of training and testing features was evaluated by measuring some statistical parameters such as accuracy, sensitivity and specificity using (6) - (7).

$$\text{accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \times 100\% \quad (6)$$

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP+FN}} \times 100\% \quad (7)$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN+FP}} \times 100\% \quad (8)$$



(a) Original image

(b) Segmented image by median filter

(c) Segmented image by adaptive median filter

(d) Segmented image by adaptive median + SRBF

Figure 8. The comparison of image segmentation results

There are six types of classification based on extracted features. They are Zernike moment, invariant moment, roundness and slimness, and the combination of these three kinds of features. The number of features describes the number of input layers in the MLP classifier. Comparison of classification results of these features is shown in Table 1.

Table 1. The Classification Results of the Features of Zernike Moments, Invariant Moments, Roundness and Slimness Parameters

Extracted features	Number of features	Accuracy (%)	Sensitivity (%)	Specificity (%)
Zernike moment	28	85.70	79.20	91.20
Invariant moment	7	90.50	87.50	93.00
Roundness and slimness	2	95.20	97.90	93.00
Zernike and invariant moments	35	85.70	81.30	89.50
Invariant moment, roundness, slimness	9	96.20	97.90	94.70
Invariant moment, Zernike, roundness, slimness	37	95.20	93.80	96.50

As depicted in Table 1, the combination of nine extracted features consisting of roundness, slimness and seven features of invariant moment obtains the best classification results with accuracy, sensitivity and specificity 96.20%, 97.90% and 94.70%. Moreover, Zernike moments do not have significant effect of classification either in self or combination with other features. The combination of Zernike moments with invariant moments, roundness and slimness even reduces the results of classification compared to that of without Zernike moments. Moreover, a comparison of our results to that of other published methods is presented in Table 2. As shown in Table 2, the proposed method is comparable to other existing methods by maintaining high accuracy, sensitivity and specificity rates.

Table 2. Comparison of Results to Other Methods

	Accuracy (%)	Sensitivity (%)	Specificity (%)
Huang et al. [16]	82.2	94.10	N/A
Tahmasbi et al. [13]	N/A	97.60	97.50
Rouhi et al. [12]	96.47	96.87	95.94
Proposed approach	96.20	97.90	94.70

4. CONCLUSION

This paper proposes a scheme to classify nodule shape of breast ultrasound images. The scheme consists of four main stages, i.e. pre-processing, segmentation, feature extraction and classification. At the image enhancement stage, the combination method of adaptive median filter and SRBF is able to eliminate the labels of the ultrasound breast nodule images. The classification of nodule shapes was divided into two classes, i.e. regular and irregular classes. The best performance of classification is obtained by using nine extracted features which comprise of seven features of invariant moments, roundness and slimness. The results of classification process achieve the level of accuracy, sensitivity and specificity at 96.20%, 97.90% and 94.70%, respectively. These results indicate that the proposed has a potential to be implemented in a computerised aided diagnosis system.

In future research, the other malignancy parameter such as posterior acoustic parameter can be used to analyse the breast nodules for further improvement of the radiologist diagnosis. In addition, to improve the performance of this research, the other selection methods can be developed with the other features for a better classification result and accelerate the computation process.

ACKNOWLEDGEMENTS

This research work is funded by Directorate General of Higher Education, Ministry of Research, Technology and Higher Education, Republic of Indonesia. The authors would like to thank radiologists from the Department of Radiology, RSUP Sardjito and Hardjolukito Hospital for cooperating and sharing experiences. We would also like to thank the Intelligent Systems research group members in our Department for sharing all of meaningful knowledge.

REFERENCES

[1] C.-M. Chen, Y.-H. Chou, K.-C. Han, G.-S. Hung, C.-M. Tiu, H.-J. Chiou, *et al.*, "Breast lesions on sonograms: computer-aided diagnosis with nearly setting-independent features and artificial neural networks 1," *Radiology,* vol. 226, pp. 504-514, 2003.
[2] Y. L. Huang, D. R. Chen, Y. R. Jiang, S. J. Kuo, H. K. Wu, and W. Moon, "Computer aided diagnosis using morphological features for classifying breast lesions on ultrasound," *Ultrasound in Obstetrics & Gynecology,* vol. 32, pp. 565-572, 2008.
[3] S. Wu, Q. Zhu, and Y. Xie, "Evaluation of various speckle reduction filters on medical ultrasound images," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2013, pp. 1148-1151.
[4] D. A. Khusna, H. A. Nugroho, and I. Soesanti, "Performance analysis of edge and detailed preserved speckle noise reduction filters for breast ultrasound images," in *International Conference on Information Technology, Computer, and Electrical Engineering*, 2015, pp. 76-80.
[5] H. K. N. Yusufiyah, H. A. Nugroho, T. B. Adji, and A. Nugroho, "Feature extraction for classifying lesion's shape of breast ultrasound images," in *2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 2015, pp. 102-106.
[6] A. K. Jumaat, W. E. Z. W. A. Rahman, A. Ibrahim, and R. Mahmud, "Segmentation of masses from breast ultrasound images using parametric active contour algorithm," *Procedia-Social and Behavioral Sciences,* vol. 8, pp. 640-647, 2010.
[7] A. Nugroho, H. A. Nugroho, and L. Choridah, "Active contour bilateral filter for breast lesions segmentation on ultrasound images," in *International Conference on Science in Information Technology*, 2015, pp. 36-40.
[8] M. Savelonas, D. Maroulis, D. Iakovidis, S. Karkanis, and N. Dimitropoulos, "A variable background active contour model for automatic detection of thyroid nodules in ultrasound images," in *IEEE International Conference on Image Processing*, 2005, pp. I-17-20.

[9] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on image processing,* vol. 10, pp. 266-277, 2001.

[10] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," *IEEE transactions on image processing,* vol. 17, pp. 2029-2039, 2008.

[11] A. Nugroho, H. A. Nugroho, N. A. Setiawan, and L. Choridah, "Internal content classification of ultrasound thyroid nodules based on textural features," *Communications in Science and Technology,* vol. 1, 2016.

[12] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, "Benign and malignant breast tumors classification based on region growing and CNN segmentation," *Expert Systems with Applications,* vol. 42, pp. 990-1002, 2015.

[13] A. Tahmasbi, F. Saki, and S. B. Shokouhi, "Classification of benign and malignant masses based on Zernike moments," *Computers in Biology and Medicine,* vol. 41, pp. 726-735, 2011.

[14] M. Yang, K. Kpalma, and J. Ronsin, "A survey of shape feature extraction techniques," *Pattern recognition,* pp. 43-90, 2008.

[15] H. A. Nugroho, H. K. N. Yusufiyah, T. B. Adji, and A. Nugroho, "Zernike moment feature extraction for classifying lesion's shape of breast ultrasound images," in *International Conference on Information Technology and Electrical Engineering,* 2015, pp. 458-463.

[16] Y.-L. Huang, "Computer-aided diagnosis using neural networks and support vector machines for breast ultrasonography," *Journal of Medical Ultrasound,* vol. 17, pp. 17-24, 2009.

[17] W.-J. Wu and W. K. Moon, "Ultrasound breast tumor image computer-aided diagnosis with texture and morphological features," *Academic radiology,* vol. 15, pp. 873-880, 2008.

[18] M. R. Teague, "Image analysis via the general theory of moments," *JOSA,* vol. 70, pp. 920-930, 1980.

[19] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE transactions on information theory,* vol. 8, pp. 179-187, 1962.