Original Research Paper

Identification of Radiomic Features as an Imaging Marker to Differentiate Benign and Malignant Breast Masses Based on Magnetic Resonance Imaging

By
Sachini Udara Wickramasinghe, BSc (Hons) Radiography, Department of Radiography and Radiotherapy, Faculty of Allied Health Sciences, General Sir John Kotelawala Defence University, Rathmalana, Sri Lanka.

Phone Number - + 94714029007
Email – udarawickramasinghe2@gmail.com

Thushara Indika Weerakoon, BSc (Hons) Radiography, Department of Radiography and Radiotherapy, Faculty of Allied Health Sciences, General Sir John Kotelawala Defence University, Rathmalana, Sri Lanka.

Phone Number - +94773690018
Email – mmweerakoonmm@gmail.com

Dr. Pradeep Jayantha Gamage, MBBS, Registrar in Radiology, National Hospital, Sri Lanka

Phone Number - +94715379994
Email – prdp1990@icloud.com

Dr. Muditha Suranga Bandara, PhD, Department of Physics, University of Colombo, Sri Lanka.

Phone Number - +94717051987
Email – muditha@sci.cmb.ac.lk

Dr. Aruna Pallewatte, MBBS, MD, FRCR (Lond) Consultant Radiologist, National Hospital, Sri Lanka.

Phone Number - +94718384832
Email – asp31263@hotmail.com
Abstract

Background - Breast cancer is one of the most common cancers among women globally and early identification is known to increase patient outcomes. Therefore, the main aim of this study is to identify the essential radiomic features as an image marker and compare the diagnostic feasibility of feature parameters derived from radiomics analysis and conventional Magnetic Resonance Imaging (MRI) to differentiate benign and malignant breast masses.

Methods and Material - T1-weighted Dynamic Contrast-Enhanced (DCE) breast MR axial images of 151 (benign (79) and malignant (72)) patients were chosen. Regions of interest were selected using both manual and semi-automatic segmentation from each lesion. 382 radiomic features computed on the selected regions. A random forest model was employed to detect the most important features that differentiate benign and malignant breast masses. The ten most important radiomics features were obtained from manual and semi-automatic segmentation based on the Gini index to train a support vector machine. MATLAB and IBM SPSS Statistics Subscription software used for statistical analysis.

Results - The accuracy (sensitivity) of the models built from the ten most significant features obtained from manual and semi-automatic segmentation were 0.815 (0.84), 0.821 (0.87), respectively. The top 10 features obtained from manual delineation and semi-automatic segmentation showed a significant difference (p<0.05) between benign and malignant breast lesions.

Conclusion - This radiomics analysis based on DCE-BMRI revealed distinct radiomic features to differentiate benign and malignant breast masses. Therefore, the radiomics analysis can be used as a supporting tool in detecting breast MRI lesions.

Keywords: Dynamic Contrast-Enhanced Breast MRI; Manual delineation; Radiomics; Semi-automatic segmentation
Introduction
Breast cancer (BC) can be identified as one of the highly dominant cancers in women globally and its prevention remains as a worldwide challenge [1]. BC has higher mortality rates and is a prominent cause of cancer fatality in women [2]. Breast carcinoma originates from cells of the mammary gland and displays a broad range of morphological characteristics, distinct histochemical contours, and distinctive histopathological sub-categories. More than 95% of breast malignancies are adenocarcinomas. Invasive ductal carcinoma (IDC) is the main category of invasive breast carcinoma. IDC is responsible for 55% of breast carcinoma occurrence after detection [3].

Breast Magnetic Resonance Imaging (BMRI) is an established supplementary tool to assess suspicious breast masses. Dynamic contrast-enhanced MRI (DCE-MRI) has revealed great investigative importance in identifying multifocal, multicentric, or contralateral disorders not identified on mammography or ultrasound (US), recognition of in situ ductal carcinoma, recognition of hidden metastatic axillary swellings, and recognition of malignancy in condensed tissues in the breast [4]. Furthermore, DCE-MR images are the first series of images reviewed by radiologists during image interpretation. The radiologist identifies clinically significant lesion based on the pattern of contrast enhancement. Therefore, DCE-MRI is one of the most clinically significant MRI sequences contrary to other noninvasive approaches [5].

Variability in BMRI techniques, definitions of morphological findings, assessment, and categorization of conventional characteristics of MR breast masses have led to the development of a lexicon on BMRI. This lexicon, the Breast Imaging Reporting and Data System (BI-RADS) MRI of the American College of Radiology [6] contains the vocabulary to define the lesion infrastructure and features. Zhou et al. [7] stated that BI-RADS shows subjectivity, less reproducibility, and demonstrates wide intra- and inter-reader disagreement, particularly
among less-experienced readers. But conventional BI-RADS signifiers are the gold standard schema for conventional BMR lesion categorization when a biopsy or histopathological reports are not available [8]. Therefore, this research will be helpful to find a superior method to improve the accuracy by eliminating subjectivity, documenting the conventional BMRI features, reducing false-positive detection errors, and supporting the findings that have already been diagnosed using the conventional BMRI features based on the BI-RADS lexicon.

Recently, the use of BMRI has emerged an excellent opportunity to establish new approaches to non-invasive benign and malignant tumor diagnosis using radiomics. Radiomics is an explicit method for extracting a wide range of quantifiable characteristics from digital images that are not discernible to human eyes [8]. Radiomic features offer information on patterns of the range of grayscale and relationships between pixels [9]. Therefore, this research will be helpful to evaluate the importance of radiomics features to diagnose lesion pathologies that provide essential details beyond the human eye.

Quantitative radiomic features are usually classified as characteristics based on shape, first-order statistics features, and second-order statistics features. The shape characteristics are centered on the two-dimensional and three-dimensional images reconstructed to determine the geometrical features [10]. The dispersion of specific voxel quantities without spatial correlations is denoted by characteristics of first-order statistics. The characteristics of second-order statistics comprise the textural features obtained by computing the statistical inter-relationships between neighboring voxels or pixels [10]. The highly prominent texture models employed in MRI texture analysis demonstrated to be the gray-level co-occurrence matrix (GLCM) [11] and grey-level run-length matrix (GLRLM) [12]. The GLCM is a square matrix with the aspect of the number of gray intensities in the region of interest (ROI). GLCM is a frequency organization or how frequently a sequence of pixel intensity ideals occurs in an image [13]. The gray-level run is a series of co-linear image points with a similar gray intensity
Radiomics has also been used and assessed in different modalities such as US [14], Computed Tomography (CT) [15], Nuclear Medicine [16], MRI [17], and plain radiography. Lesion segmentation by identification of the ROI in the medical images retrieved is an important phase of the radiomics analysis (18). The entire lesion can be delineated by manual, semi-automated, or automated procedures. According to Loizou et al. [18], the gold standard is now called the physical description of ROIs and is the chosen alternative over automatic processes in many applications [19], [20], [21]. However, such a technique creates intra and interobserver inconsistency and repeatability failures in addition to its inefficient existence since many tumors have doubtful and confusing borders [22]. The expansion and verification of new semi-automatic differentiation algorithms is an accessible area of experiment with fascinating and advanced outcomes. Semi-automatic approaches facilitate specialists to categorize cases that are easy to segment and to deal with difficult lesions. Models based on active contour are extremely valuable functions for semi-automated lesion segmentation [23]. Semi-automatic segmentation of 2D DCE-BMR images of this study was performed using active contour segmentation without edges based on the Chan-Vese technique [24]. Therefore, it is essential to determine the most accurate method for lesion segmentation for feature extraction. This research investigated the accuracy of manual and semi-automatic lesion segmentation methods to differentiate benign and malignant breast masses.

Certain radiomic features that have been extracted from the manual delineation and semi-automatic segmentation may not be precisely appropriate for further evaluation and may not be essential. Therefore, it is essential to determine the most important radiomic features before the construction of the machine learning (ML) models. One of the techniques that can be used to identify the most prominent radiomic features is the random forest (RF) technique. RF utilizes the mean Decrease Gini Index to discover the importance of a specific characteristic to the prediction model. Higher mean decrease Gini features indicate the most influential
predictors for the classifier [25]. RF [26] has superior predictive execution than the principal component analysis [27],[28]. The most prominent radiomic features chosen by the RF from both manual delineation and semi-automatic segmentation were applied to train a linear support vector classifier in the current study. Support vector machine (SVM) is a supervised machine learning method that has exhibited elevated execution in unraveling categorization obstacles in numerous biomedical disciplines [29]. A linear SVM was applied in the current study for the radiomic feature classification.

This research investigated the relationship between the automated features, and conventional MRI features apparent on breast DCE-MR images. It assessed the inter-operator variability between operators and evaluated the usefulness of building machine learning (ML) models for mass phenotype recognition and classification of benign and malignant breast masses. The main aim of our research is to identify the essential radiomic features as an image bio marker so that it is possible compare the diagnostic feasibility of feature parameters derived from radiomics analysis and conventional MRI to differentiate benign and malignant breast masses.

**Methods**

T1-weighted (T1W) DCE axial breast MR images of 151 (benign (79) and malignant (72)) patients assessed according to the 5th edition of the American College of Radiology Breast Imaging Reporting and Data System (ACR BI-RADS) MR lexicon selected from the population. Breast masses of female patients (age>40 years) that were radiologically categorized according to the 5th edition of BIRADS lexicon included for this study. Poor image quality, non-mass-like enhancement, and background parenchymal enhancements excluded from this study. Informed consent of the patients was waived off as the study involved only a retrospective review of existing BMRI with known diagnoses. This study was approved by the institutional ethical clearance committee.
MRI studies were performed using a 3T Ingenia (Philips Medical Systems, Best, The Netherlands), using a dedicated bilateral SENSE 16 channel phase array breast coil with patient, in the prone position. The MRI examinations consisted of a 3D fat suppressed DCE sequence (pMEdyn_eTHRIVE). Breath-holding techniques were not used. DCE axial images with 1.6 mm slice thickness were obtained. Field of View (FOV) was determined according to the size of the patient. The reconstruction matrix was adjusted at the time of FOV adjustment in order to ensure voxel size was constant. For dynamic contrast enhancement, a 0.1 mmol/kg bolus of Magnevist (gadopentetate dimeglumine) was injected according to the bodyweight of the patient (average 5ml) with a flow rate of 2.5-3.0 ml/sec, followed by a 20-mL saline flush. No other MR sequences (functional or morphological) have been included in the analysis.

To avoid inter-reader variability, the MR images of the DCE series were evaluated by two independent radiologists with over 10 years and 5 years of experience in BMR image interpretation using the DICOM_viewer, an in-house software developed on MATLAB [30]. The axial slice, which showed the optimal representation of the tumor area of the DCE image series, was chosen. The radiologists were blinded to the history of the patients.

The selected image from the series was transferred to the MaZda [31] workspace. To avoid inter-reader variability, the same two radiologists independently performed a 2D manual delineation of breast tumor on the DCE-BMR image in MaZda software by defining an irregular ROI to include only the visible tumor and excluding equivocal normal breast tissue and visible necrotic areas possible. The edge voxels were excluded to avoid the partial volume effect (Figure 1(a)).

The semi-automatic segmentation of 2D DCE-BMR images was performed using active contour segmentation without edges based on the Chan-Vese technique to extract tumors from the DCE-BMR images. An in-house software developed on MATLAB was used for the
implementation of semi-automatic segmentation by active contour segmentation without edges based on the Chan-Vese technique. Prior to the segmentation, a rectangular ROI was drawn independently by the same two radiologists on the selected slice, aiming to contain the whole lesion with some normal tissue (Figure 1(b)).

Thereafter, based on this rectangular region, the tumor area with some normal tissue in the slice was cropped for better visualization of the tumor. The active contour algorithm was then applied to segment the tumor using 2000 iterations. The image of the semi-automatically extracted tumor was reshaped to the size of the original cropped image (Figure 2).

Seventy-three (73) shape features explained by Materka (2002) [32] were obtained. Nine (9) first-order statistic features focused on histogram-based properties (mean, variance, skewness, kurtosis, 1st, 10th, 50th, 90th, and 99th percentiles), 11 GLCM features (angular second moment, contrast, correlation, entropy, sum entropy, sum of squares, sum average, sum variance, inverse difference moment, difference entropy, difference variance) obtained in 4 directions (vertical, horizontal, 45°, and 135°) and five interpixel distances (offsets; n = 1 to 5) respectively, and four (4) different GLRLM (short-run emphasis, long-run emphasis, grey level nonuniformity, run-length nonuniformity, and the fraction of image in runs) for four pixel-run directions (vertical, horizontal, at 45° and 135°) [32] were calculated in defined 2D ROIs obtained from both segmentation methods using MaZda software. A total of 382 radiomic features were extracted separately from both segmentation methods.

In order to determine the most important radiomic features before the reconstruction of the ML models, the RF classification technique was applied to obtain the 10 most suitable radiomic features that can distinguish the benign and malignant breast masses centered on the mean decrease Gini index. RF classification was performed using Random-Forest package (function RandomForest ()) available in R software (R Foundation for Statistical Computing, Vienna,
RF classification was employed for radiomic features derived from both manual delineation and semi-automatic segmentation. RF (Number of trees: 300, no. of variables: 10, no of iterations: 100) was applied to find out the most suitable radiomic features to differentiate benign against malignant breast masses. The most prominent 10 radiomic features that were chosen by the RF from both segmentation methods were then applied to train a linear SVM. Selected features detached from the linear SVM and investigated its accuracy with fivefold cross-validation. SV classification was performed using MATLAB.

The most prominent 10 radiomic features obtained from benign and malignant breast masses were tested to evaluate the statistical significance. The independent samples t-test compared the means of the benign and malignant groups to determine statistical evidence that the associated population means are significantly different. The Mann-Whitney U test was used to compare differences between benign and malignant groups when the data was not normal. Since the sample size was greater than 50, the Shapiro-Wilk test was used to determine the normality of the feature parameters. The Intraclass Correlation Coefficient (ICC) was chosen as the reproducibility metric to assess the concordance of analyses done by the two radiologists. The Absolute-agreement and a two-way mixed-effect model used assess the ICC for this study. These statistical analyses were performed using a statistical package for the social sciences (SPSS) 25.0 software with p<0.05 indicating a level of significance.

**Results**

ICC values of Average Measures obtained from 382 radiomic features based on manual delineation and semi-automatic segmentation performed by the two operators were greater than 90%. It showed an excellent agreement between the two radiologists for both segmentation methods. Therefore, the mean values of the 382 radiomic features obtained from the manual and semi-automatic segmentation applied for RF classification.
Based on the highest mean decrease Gini index of the RF classifier, the top 10 features obtained from manual delineation were used to differentiate benign and malignant breast masses. These elements included the prominent features from the shape (GeoW9, GeoW15, GeoW2, GeoW5b, and GeoRc) and first-order statistical features (Mean, Perc.10%, Perc.50%, Perc.90%, and Perc.99). There were no second-order statistics, detected as the most prominent features (Figure 1(c)).

Ten important features obtained from semi-automatic segmentation were assessed based on the highest mean decrease Gini index of the RF classifier to differentiate benign and malignant breast masses. These elements included the prominent features from the shape (GeoW9, GeoW15, GeoW2, GeoW5b, and GeoRc) and first-order statistical features (Mean, Perc.10%, Perc.50%, Perc.90%, and Perc.99). There were no second-order statistics, detected as the most prominent features (Figure 1(d)).

The mean values of all the top 10 radiomic features obtained from manual delineation and semi-automatic segmentation showed a significant difference (p<0.05) between benign and malignant breast lesions.

For the top 10 radiomic features obtained from manual delineation, the SVM (manual model) reached an overall accuracy of 81.5%. The total error rate for the classifier was 0.185. The sensitivity or the true positive rate was 0.84. Specificity or true negative rate was 0.81. The area under (AUC) the ROC curve received as 0.87 (Figure 1(e)).

For the top 10 features obtained from semi-automatic segmentation, the SVM (semi-automatic model) reached overall accuracy of 82.1%. The total error rate for the classifier was 0.179. The sensitivity or the true positive rate was 0.87. Specificity or true negative rate was 0.85. The area under (AUC) the ROC curve received as 0.86 (Figure 1(f)).
To describe the sensitivity and specificity values, the optimal cut-off for each model was determined based on Youden’s index [33]. For manual and semi-automatic models, Youden’s indexes were obtained as 62.5% and 67.1%, respectively.

**Discussion**

In contrast to the canonical computer aided diagnosis in other primary breast imaging modalities, the expansion of automated radiomics analysis for breast DCE-MRI is in its initial phase. Only limited studies have been practiced for semi-automated and/or manual tumor segmentation and feature extraction for tumors detected by breast DCE-MRI to build radiomics ML classification models in a single study. In this study, the linear SVM was capable of categorizing benign and malignant breast masses in both manual and semi-automatic models with an overall accuracy of 0.815 and 0.821, respectively and sensitivity of 0.84 and 0.87, respectively. Therefore, the findings of the present study are significant since this study discusses both the above methods used for radiomics analysis.

These results of the present study can be correlated to the studies conducted by Wedegartner et al. [34] and Huang et al. [35]. Regarding the overall accuracies of the manual and semi-automatic models, the accuracy of the manual model (0.815) was slightly less than the semi-automatic model (0.821). The reasons for this reduction of the accuracy of manual model could be due to the fact that manual delineation is difficult, tiresome and the definition of whole breast tissue containing a vast quantity of separate data collections is less efficient [22]. The reasons for higher accuracy in the semi-automatic model compared to the manual model can be due to semi-automatic approaches facilitate specialists to categorize cases that are easy to segment, able to deal with difficult lesions, and more appropriate for medical practitioners owing to enhanced comprehensibility in the separation procedure [36]. Furthermore, there was a very slight difference between the accuracies of the two segmentation methods. This can be
due to the differences in duration of experience (5 and 10 years) of the two radiologists in BMRI interpretation.

This study assessed the agreement between the two radiologists for both segmentation methods. The inter-operator variability was assessed using ICC estimation for each feature obtained from manual delineation and semi-automatic segmentation. ICC estimation was greater than 90% for the features obtained from manual delineation and semi-automatic segmentation. Therefore, it showed excellent agreement between the two radiologists who had 5 and 10 years of experience in interpreting BMRI. The reason behind this excellent agreement between the two operators in both segmentation methods can be due to their long experience in interpreting BMRI. Huang et al. [35], also mentioned that the performance of manual segmentation depends on the experience of interpreting radiologists.

There were only a very few studies conducted to connect computer-based automated feature extraction of tumors to the conventional MRI features classified according to the BI-RADS lexicon. Therefore, this study investigated the relationship between the automated features and conventional MRI features apparent on breast DCE-MR images. It was assessed using both shape and first-order statistics features. As explained by Materka (2002) [32], all the calculated shape features of the current study were derived based on the surface area, perimeters, various diameters and radix, factors of inscribed circle, circumscribed circle, ellipsis, rectangle, and different ratios of these factors and invariant as elongation, compactness or roundness of the tumor. It is also possible that the benign tumors spread further maintaining their spherical forms than malignant tumors. This outcome of our study is confirmed by a study conducted by Liney et al. [37]. Therefore, all derived shape features were sensitive to the shape and boundary of the tumor. Since all the DCE-BMRI images of this research assessed according to the 5th edition of ACR BI-RADS MR lexicon based on the shape, boundary and phenotype of the masses, each prominent shape feature can be correlated with distinct signifiers characterized
by conventional MRI features based on BI-RADS lexicon to differentiate benign and malignant breast masses.

The internal enhancement characteristics of the BI-RADS lexicon can be contributed to the homogeneous and heterogeneous histogram-based gray-level intensities and the extent of histogram-based gray-level intensities [38] of the extracted quantitative first-order statistic features to differentiate benign and malignant breast masses. The current study found that the malignant tumors had a greater intensity and wider spread in the enhancement histogram that is more heterogeneous related to the benign tumors. A study conducted by Nie et al. [39] agrees with our finding that the malignant tumors had a greater intensity and wider spread in the enhancement histogram that is more heterogeneous compared to the benign tumors. It is one of the explicit features signifying malignant tumors. Consequently, the internal enhancement of the masses of the BI-RADS lexicon can be correlated to the extracted quantitative first-order features to differentiate benign and malignant breast masses in this study. Therefore, these results enhanced the accuracy of features, reduced false-positive detection errors, and supported the findings that have already been diagnosed via the conventional BMRI features based on the BI-RADS lexicon by applying the conventional BMRI features for the radiomics analysis.

The lack of an independent assessment set for model assessment, can be stated as a major limitation and a possible resource of overfitting in the study. The insertion of more subjects and investigators would permit an improved overview of the outcomes and expansion of strong radiomics evaluations. The active contour semi-automatic segmentation without edges based on the Chan-Vese technique was not suitable for intensity in-homogeneity as its inherent characteristic of relatively slowness particularly in handling bulky pictures. A new complex segmentation model or adding several models together can be applied to overcome the intensity in-homogeneity. The findings of the ROC assessment for all the five-folds were not obtained
separately in this study. Therefore, the irregularity in performance metrics between the five folds was not evaluated.

Future studies, assisted by radiomics analyses with MR imaging sequences that do not utilize gadolinium contrast agents will be useful to minimize possibilities of the side effects due to the usage of contrast agents.

**Conclusion**

Even though the location of the tumor and its border can typically be decided, the histopathology of the tumor may be challenging to distinguish only by conventional BMRI features and therefore supplementary tools are essential for precise identification. Furthermore, the classification outcomes of the current study linked very well with the clinical condition of the patients. The findings of the present study suggested that the ML techniques centered on DCE-BMRI radiomic features can be applied to distinguish benign breast masses from malignant breast masses. We believe that the created radiomics models would be a clinically supportive tool in the future, with extensive clinical validations.

**Statements**

- All authors reviewed the final version of the manuscript and agreed to submit it to IMAGING for publication.
- No financial support was received for this study
- The authors have no conflict of interest to disclose
Figures

Segmentation

Feature Classification

Prediction

Accepted Manuscript
**Figure 1.** Diagram of study workflow of Radiomics based classification model constructing phases to discriminate benign and malignant breast masses in DCE-BMR images (a-2D manual delineation of the tumor using MaZda software, b- semi-automatic segmentation by creating a rectangular ROI, c- feature classification based on highest mean decrease Gini index of RF classifier for the features obtained from manual segmentation, d- feature classification based on highest mean decrease Gini index of RF classifier for the features obtained from semi-automatic segmentation, d-Receiver operating characteristic (ROC) curve for the top 10 Radiomics features obtained from manual delineation, e- ROC curve for the top 10 Radiomics features from semi-automatic segmentation).

**Figure 2.** Segmentation of the tumor semi-automatically through active contour algorithm. (A- input image, B- initial mask, C- image showing the number of iterations, D- the binary mask of the cropped image, E- segmented image after reshaping to the size of the original cropped image, F- Output).
References


MATLAB. Natick, Massachusetts, United States: The MathWorks, Inc; 2018.

MaZda. Poland: Institute of Electronics, Technical University of Lodz;

Materka A. MaZda user’s manual. Adres www eletel p lodz


