IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 1, 2020, accepted July 8, 2020, date of publication July 13, 2020, date of current version July 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3008927

Impact of Image Contrast Enhancement on Stability of Radiomics Feature Quantification on a 2D Mammogram Radiograph

SITI FAIRUZ MAT RADZI[1], MUHAMMAD KHALIS ABDUL KARIM[1], M IQBAL SARIPAN[2], MOHD AMIRUDDIN ABD RAHMAN[1], NURUL HUDA OSMAN[1], ENTESAR ZAWAM DALAH[3], AND NORAMALIZA MOHD NOOR[4]

[1]Department of Physics, Faculty of Science, Universiti Putra Malaysia, Selangor 43400, Malaysia
[2]Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia, Selangor 43400, Malaysia
[3]Clinical Support Service and Nursing Sector, Dubai Health Authority, Dubai, United Arab Emirates
[4]Department of Radiology, Faculty of Medicine, Universiti Putra Malaysia, Selangor 43400, Malaysia

Corresponding author: Muhammad Khalis Abdul Karim (mkhalis@upm.edu.my)

This work was supported by the Ministry of Higher Education Grant under Grant FRGS/5540112.

ABSTRACT The present work aimed to evaluate the reproducibility of radiomics features derived from manual delineation and semiautomatic segmentation after enhancement using the Contrast Limited Adaptive Histogram Equalization (CLAHE) and Adaptive Histogram Equalization (AHE) techniques on a benign tumor of two-dimensional (2D) mammography images. Thirty mammogram images with known benign tumors were obtained from The Cancer Imaging Archive (TCIA) datasets and were randomly selected as subjects. The samples were enhanced for semiautomatic segmentation sets using the Active Contour Model in MATLAB 2019a before analysis by two independent observers. Meanwhile, the images without any enhancement were segmented manually. The samples were divided into three categories: (1) CLAHE images, (2) AHE images, and (3) manual segmented images. Radiomics features were extracted using algorithms provided by MATLAB 2019a software and were assessed with a reliable intra-class correlation coefficient (ICC) score. Radiomics features for the CLAHE group ($ICC = 0.890 \pm 0.554$, $p < 0.05$) had the highest reproducibility compared to the features extracted from the AHE group ($ICC = 0.850 \pm 0.933$, $p < 0.05$) and manual delineation ($ICC = 0.673 \pm 0.807$, $p > 0.05$). Features in all three categories were more robust for the CLAHE compared to the AHE and manual groups. This study shows the existence in variation for the radiomics features extracted from tumor region that are segmented using various image enhancement techniques. Semiautomatic segmentation with image enhancement using CLAHE algorithm gave the best result and was a better alternative than manual delineation as the first two techniques yielded reproducible descriptors. This method should be applicable for predicting outcomes in patient with breast cancer.

INDEX TERMS Breast cancer, radiomics, contrast limited adaptive histogram equalization (CLAHE), adaptive histogram equalization (AHE), semiautomatic segmentation.

I. INTRODUCTION

Breast cancer has been acknowledged as the most prevalent and common cause of death among Malaysian woman over the age of 40 [1]. Several studies emphasize the need and urgency for early detection in reducing breast cancer morbidity and mortality [2]–[4]. Medical imaging techniques, such as mammography, play an important role in non-invasively assessing breast tissues for detection, diagnostic, staging, and management purposes [2]. In an attempt to improve the mortality rate among the population, a mammography screening program was proven to be the most cost-effective program for providing useful details about the presence of abnormal breast tissues [2].

Studies have shown the potential of radiomics feature extraction in providing consistent and unbiased descriptions

The associate editor coordinating the review of this manuscript and approving it for publication was Abdel-Hamid Soliman.

of tumor structure parallel with the increment in medical data [5]–[7]. Radiomics analysis applies advanced computational approaches to convert image data from the selected region into high dimensional feature data, assuming that these data provide information that will be useful for prognosis and could be used as a potential predictive biomarker. The most critical problem in radiomics is reproducibility. Reproducibility describes similar performances of radiomics measurements using different techniques or observers or even from different diagnostic centers [8], [9]. In order to obtain accurate results, extracted features should be optimized to estimate patient survival analysis and boost treatment selection and monitoring for each patient [10], [11].

Tumor segmentation plays a vital role in quantitative image extraction. Although physicians commonly use manual segmentation, this method is a time-consuming process that has more substantial inter-observer variability. However, to ensure higher accuracy in semiautomatic segmentation, pre-processing image enhancement is vital. A computer-aided diagnosis (CAD) system for analyzing mammogram images starts with image pre-processing for contrast enhancement while still conserving image brightness. This technique does not cause any loss or degradation of the image details. On the contrary, contrast resolution is still a primary feature that helps radiologists during the diagnosis. This feature is particularly important when diagnosing a dense classified breast tissue in which malignancy can go undetected [12], [13].

Contrast enhancement using a conventional algorithm is more complex and challenging when compared to the Histogram Equalization (HE) [14], [15]. HE allows the band of contrast in the high histogram region to stretch-out and shrink the contrast of the low histogram region. However, this process becomes less effective when the contrast characteristics vary across images. Adaptive Histogram Equalization (AHE) addresses this issue by generating mapping for each pixel from the histogram in the surrounding pixels [15]. However, performing AHE in a relatively small intensity range can increase image noise in that region, leading to the appearance of artifacts in those regions [16]. Contrast Limited AHE (CLAHE) can overcome the drawbacks of AHE and HE since results were achieved in the cases where noise become too prominent by enhancing contrast.

Notably, the artifacts were produced when using the Adaptive Neighborhood method as contrast enhancement for digital mammograms that were susceptible to higher noise [17]. For contrast improvement, it is recommended to use the Adaptive Neighbourhood Contrast Enhancement (ANCE), which results in improving breast tumor detection [18]. First, derivative and local statistics were used by Kim *et al.* to enhance mammograms [19]. However, the nature of mammogram images is not suitable for this method. Thus, the CLAHE algorithm was introduced and yields excellent contrast enhancement of digital mammogram images [12], [18]. CLAHE has been applied extensively to enhance computer vision and pattern recognition applications. It was proven in previous studies that CLAHE could be successfully applied in medical fields, such as mammography image enhancement. Mammogram pre-processing might have a significant impact on the entire process of radiomics analysis.

Generally, CLAHE helps prevent contrast enhancement of AHE and prevents over-enhancement of noise [20]. Application of CLAHE to ultrasound images improves accuracy, sensitivity, and specificity on liver ultrasounds up to 92.95%, 90.80%, and 97.44, respectively [21]. Contrast enhancement for both CLAHE and RMSHE techniques were found to offer better enhancement of masses and microcalcifications present on low contrast mammograms [22]. Notably, use of CLAHE in pre-processing images offers better visualization for image segmentation, feature extraction, and classification [23].

Accurate segmentation may reduce variability, thus allows radiomics features conveniently become a prognostic or predictive biomarker. It has been hypothesized that imaging feature extracted using semi-automatically segmented tumors implementing CLAHE have lower variability and robust when compared to AHE technique. Numerous studies have been conducted to investigate the reproducibility and reliability of the radiomics features in order to reduce the error while training a predictive model [24]–[26]. The stability of radiomics features usually addresses the tumor delineated from three-dimensional (3D) imaging modalities, such as positron emission tomography (PET), X-ray computed tomography (CT), and magnetic resonance imaging (MRI), and no studies have addressed the stability of radiomics features in 2D mammography. In this study, we implemented the same techniques used to investigate the reproducibility and reliability in 3D imaging modality to 2D mammogram.

The application of ICC as a statistical analytical method for evaluating radiomics features as imaging biomarkers has been proven to work very well when selecting reproducible features in all 3D imaging modalities. This finding clarified the possibility of evaluating radiomics features on a 2D mammogram. Moreover, 2D tumor segmentation can be performed within a few minutes using the method of semi-automatic segmentation and might be easily applicable in clinical settings [6], [27]. Hence, in this study, we aimed to evaluate the reproducibility of radiomics features derived from the segmented tumor and to determine the robustness of certain features to improve automated diagnosis for breast cancer detection.

II. MATERIALS AND METHODS

A. STUDY SUBJECTS

Thirty 2D mammogram images with confirmed benign tumors were collected from a pool of mammogram images from The Cancer Imaging Archive (TCIA) at http://www.cancerimagingarchive.net [28]. The features extracted for this research were from the Curated Breast Imaging Subset of DDSM Digital Database for Screening Mammography

(CBIS-DDSM) [29]. Only mediolateral oblique (MLO) imaging sets were included in this study. Experienced radiologists manually annotated regions of interest (ROIs) for every suspicious mass in each imaging dataset.

B. IMAGE PREPROCESSING

All images were pre-processed with different image enhancement. CLAHE has two important hyperparameters: (1) clip limit (CL) and (2) number of tiles (NT). The CL hyperparameters were determined while the NT was automatically adjusted according to the input image. CL is a numeric value that controls noise amplification. Once the histogram of each sub-area was calculated, they were redistributed in such a way that the histogram height did not exceed a desired ''clip limit''. The CL was set at 0.9 within the range from 0 to 1 and found to be the most suitable level of contrast compared to CL set below 0.9. AHE was also implemented and compared with the CLAHE results to see which method allows easy distinction of the image components through an appropriate upsurge in its contrast. Figure 1 shows the flowchart of reproducibility analysis in which the images were pre-enhanced with CLAHE and AHE.

C. SEMIAUTOMATIC TUMOR SEGMENTATION

The Active Contour Model (ACM) technique, a semiautomatic segmentation, was used for image data-enhancement using AHE and CLAHE. The ACM technique is an iterative region-growing image segmentation algorithm that uses energy forces to separate the pixels of interest for the analysis purpose. The active contour can be described as an active model for the segmentation pre-process, while contours are boundaries that separate the ROIs from the background in an image. Contour also was described as the cumulative pixels that undergo various interpolation processes in terms of linear, splines, or polynomials that define the curve in the image [30]. Several different models of active contours can be used as a segmentation technique in image processing. Herein, the snake model was used as an ACM.

Initially, the image enhancement, CLAHE and AHE algorithm were implemented to enhance the image data. The suspected lesion was marked manually before activating the ACM function using the region-growing algorithm as illustrate in Figure 2. ROIs were segmented into both anterior and posterior regions and converted into binary images. The tumor area in enhanced images were delineated twice by two independent observers for evaluating intra-observer reproducibility. The number of iterations for semi-automatic segmentation were standardized and set to 100 iterations.

D. QUANTITATIVE IMAGING FEATURE EXTRACTION

Thirty-seven radiomics features were extracted from the data enclosed in the area of the segmented tumor and categorized into three types of features: (1) intensity histogram, (2) textural-based features, and (3) shape-based features. Six first-order statistical features specify the distribution values of the specific area without considering the spatial relationship. Twenty-two textural features specify the spatial arrangement of area evaluated by using Gray-Level Co-Occurrence Matrix (GLCM). Nine shape-based features indicate the geometrical of the tumor area. Overall, the extracted imaging features compromised six features describing intensity of the tumor, nine shape features, and 22 textural features. Mathematical equations for all radiomics features can be found in the Supplementary file, and several fundamental formulas, such as energy, contrast, entropy, correlation, and homogeneity are presented in equation below:

$$Energy = \sum_{i,j=0}^{N-1} (P_{ij})^2 \tag{1}$$

where P_{ij} = element of the normalized symmetrical GLCM and N = number of gray levels in the image as a specified number of levels in quantization. Energy feature measures the texture uniformity in the pixels and indicates image homogeneity. The greater the energy delivery, higher intensity value pairs can be seen in the images. Each of the intensity pairs borders each other at high frequencies.

$$Contrast = \sum_{i,j=0}^{N-1} P_{ij}(i-j)^2 \tag{2}$$

in which contrast feature represents the spatial frequency of the image. The larger the value of the contrast, the higher the disparity of intensity values in neighboring pixels. Entropy measures the disorder or complexity of an image in neighborhood intensity values:

$$Entropy = \sum_{i,j=0}^{N-1} \ln(P_{ij})P_{ij} \tag{3}$$

The correlation can explain the linear dependency of gray-levels of neighboring pixels and can be shown as the equation below:

$$Correlation = \sum_{i,j=0}^{N-1} P_{ij}\left(\frac{(i-\mu)(j-\mu)}{\sigma^2}\right) \tag{4}$$

where μ = the mean GLCM and σ^2 = variance of the intensities of all reference pixels in the relationships that contributed to the GLCM. The value of correlation is between 0 and 1, in which 0 indicate uncorrelated while 1 indicates perfectly correlated. These values show the linear dependency of gray level values to their respective pixels. In addition, the variation in image intensity can be defined by the homogeneity feature and is shown in the equation:

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i+j)^2} \tag{5}$$

E. STATISTICAL ANALYSIS

The reproducibility of each extracted feature was assessed using ICC. ICC is known as a statistical measure with reliability values ranging between 0 and 1. Values closer to 1 have

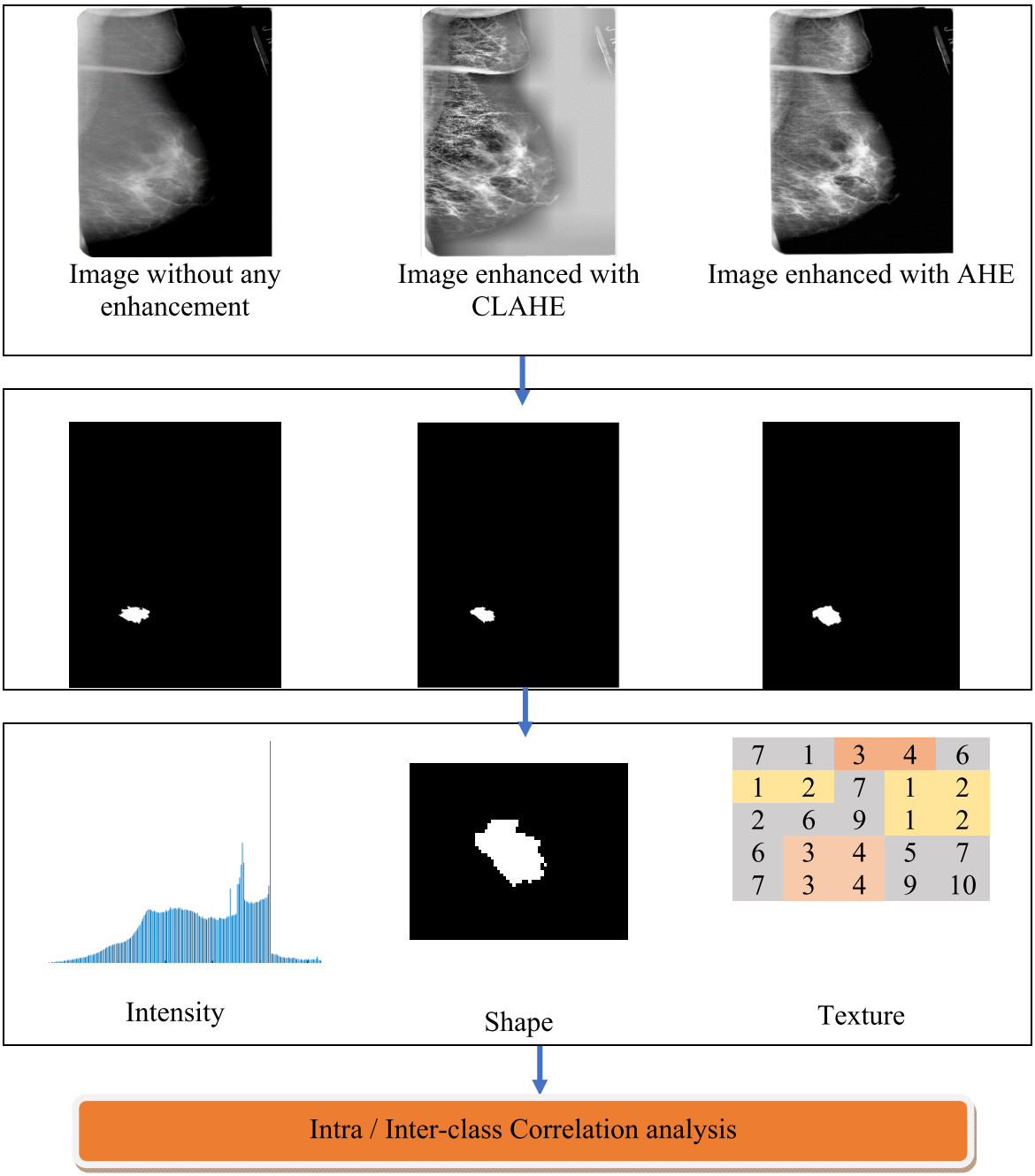

FIGURE 1. Flowchart of reproducibility analysis of radiomics features in the 2D mammograms. (a) Three datasets were selected in our study with different image enhancement. (b) Three datasets were segmented using semi-automatic segmentation and manual delineation. (c) Intensity, shape and textural transformed features were extracted from every dataset. (d) The reproducibility of the radiomics features was measured by two indicators.

the most robust reliability. Subsequently, the inter-observer reliability was estimated using a two-way mixed effect model of analysis of variance (ANOVA) and is given as equation (6):

$$ICC(A,1) = \frac{(MS_R - MS_E)}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)} \quad (6)$$

where MS_R = mean square for rows (observations, fixed factor), MS_E = mean square error, MS_C = mean square for columns (observers, random factor), and k and n represent the number of observers and subjects, respectively. The variance estimates among intra-observer segmentation ICC (C,1) for the reproducibility were obtained from one-way ANOVA using formula (7):

$$ICC(C,1) = \frac{MS_R - MS_W}{MS_R + (k-1)MS_W} \quad (7)$$

where MS_W = mean square for residual sources of variance. The Z-score normalization was implemented for comparing

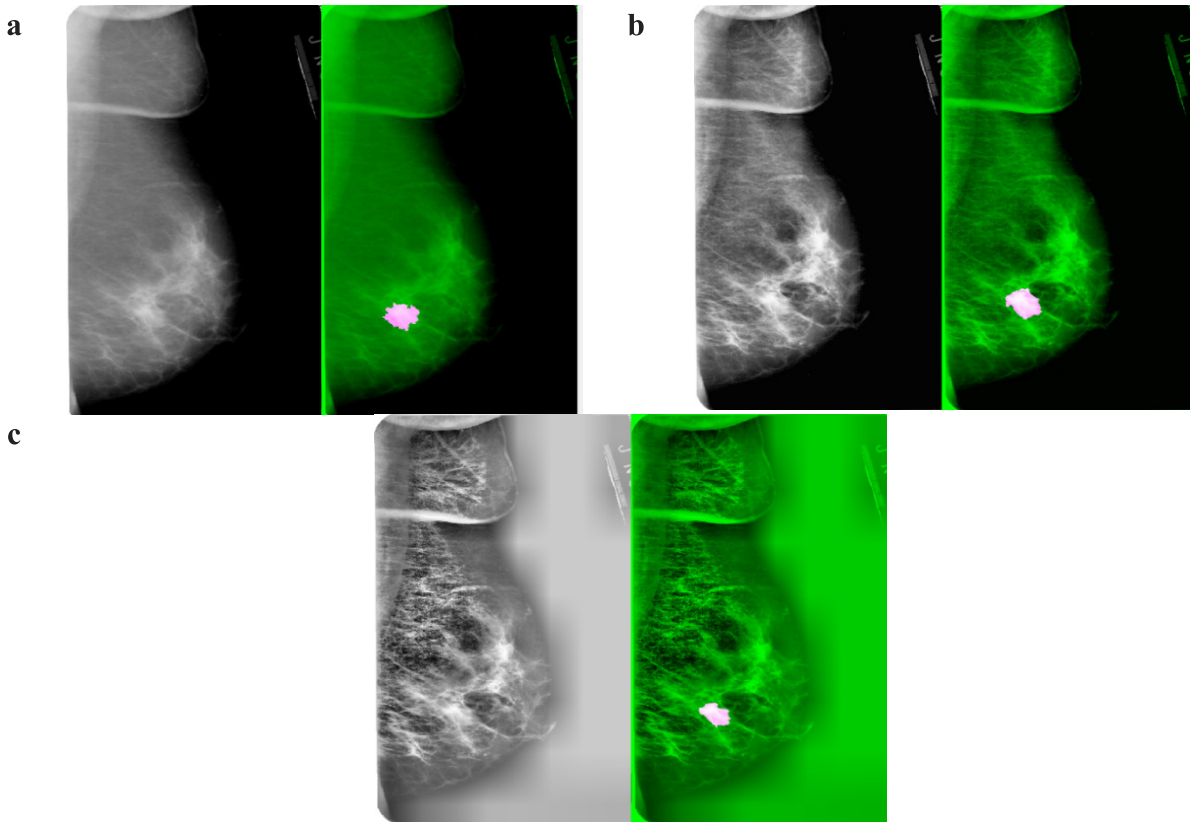

FIGURE 2. Three dataset a) without any image enhancement b) with AHE image enhancement c) with CLAHE image enhancement and segmented tumor part for each dataset.

the feature range between manual and two semiautomatic segmentation for standardization and is defined in the equation:

$$Z = \frac{x - \mu}{\sigma} \tag{8}$$

in which μ and σ are the mean value and standard deviation of radiomic features, respectively. The differences in ICC for segmentation methods were statistically estimated using the Wilcoxon–Rank test with a p-value set at 0.05. All data were expressed as the mean ± SD. All data were statistically analyzed using Statistical Package for Social Sciences (SPSS, also known as IBM SPSS statistics) version 25 (SPSS Chicago, IL, USA)

III. RESULTS

The total of 37 quantitative image features was evaluated to assess the robustness of segmented tumors by using three different image enhancement methods. Figure 3 shows comparison of ICC between manual delineation and two semiautomatic segmentation in term of several parameters: (1) intensity histogram based-features (Figure 3A), (2) shape-based features (Figure 3B), and (3) textural features (Figure 3C).

Table 1 summarizes the number of features in four reproducibility groups across three segmentation techniques.

TABLE 1. Number of features in four reproducibility groups across three segmentation techniques.

Reproducibility groups	Manual	CLAHE	AHE
Poor (ICC<0.4)	10	0	1
Fair (0.4 ≤ ICC ≤ 0.6)	3	0	0
Good (0.6 ≤ ICC ≤ 0.75)	1	0	3
Excellent (0.75 ≤ ICC ≤ 1)	23	37	33

ICC, intra-class coefficient

In general, the ICC of CLAHE image data segmentation (ICC = 0.890 ± 0.554, $p < 0.05$) was higher when compared to features extracted from AHE image data segmentation (ICC = 0.850 ± 0.933, $p < 0.05$) and manual delineation segmentation (ICC = 0.673 ± 0.807, $p > 0.05$). In total, 36 = radiomic features showed higher ICC for image data-enhancement using CLAHE and 35 showed higher ICC values for image data-enhanced using AHE segmentation sets compared to the manual method. Twenty-two features with higher ICC values were found from CLAHE datasets when compared between with AHE.

Hence, image datasets with CLAHE contrast enhancement were found to be reproducible. Nine of 9 (100%) shape features (ICC = 0.931 ± 0.597, $p < 0.05$), 22 of 22 (100%) of textural features (ICC = 0.746 ± 0.564, $p < 0.05$), and all

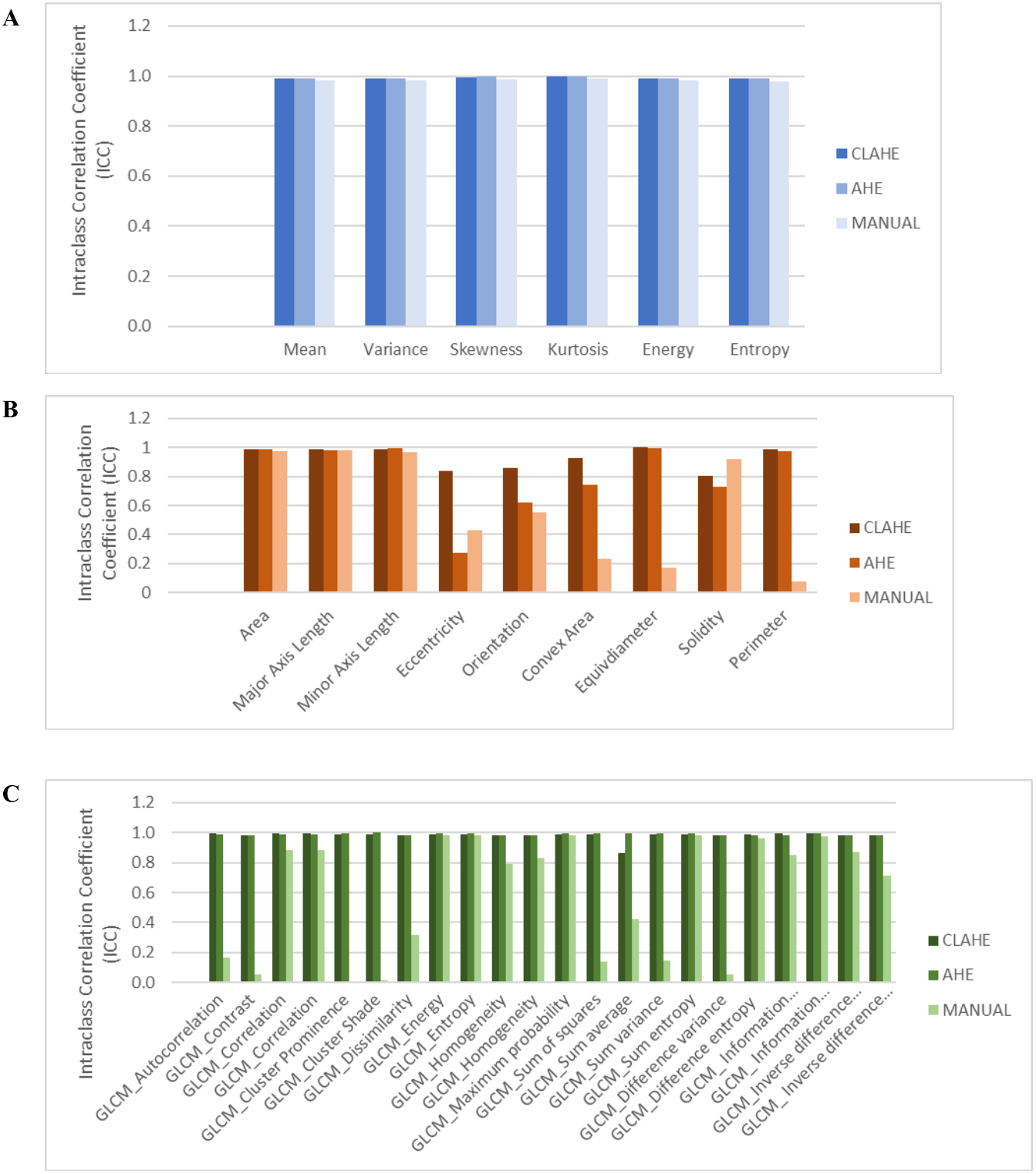

FIGURE 3. Feature comparison of intra-class correlation coefficient (ICC) between manual delineation and two semiautomatic segmentation with two techniques of image enhancements. (A) Intensity histogram-based features; (B) shape-based features; (C) textural features.

six (100%) intensity features (ICC $= 0.993 \pm 0.502$, $p < 0.05$) showed the highest ICC values ($0.75 \leq \text{ICC} \leq 1.0$) compared to other image datasets. In addition, the dataset with AHE contrast enhancement also showed high reproducibility (100%) in which the ICC values for all textures were ICC $= 0.75 \pm 0.5085$ ($p < 0.05$) and intensity features were ICC $= 0.992 \pm 0.502$ ($p < 0.05$). However, only one out of nine shape features (11.1%) features have poor reproducibility.

TABLE 2. Intra-class Correlation Coefficient (ICC) value of radiomic features from different segmentation technique.

Features	CLAHE Segmentation Techniques	AHE Segmentation Techniques	Manual Segmentation
GLCM_Autocorrelation	0.997	0.986	0.166
GLCM_Contrast	0.985	0.983	0.052
GLCM_Correlation	0.993	0.991	0.882
GLCM_Correlation	0.993	0.991	0.882
GLCM_Cluster Prominence	0.991	0.992	0.002
GLCM_Cluster Shade	0.991	1.000	0.011
GLCM_Dissimilarity	0.983	0.983	0.315
GLCM_Energy	0.991	0.992	0.981
GLCM_Entropy	0.991	0.992	0.980
GLCM_Homogeneity	0.985	0.983	0.794
GLCM_Homogeneity	0.985	0.983	0.831
GLCM_Maximum probability	0.991	0.992	0.981
GLCM_Sum of squares	0.991	0.992	0.140
GLCM_Sum average	0.866	0.992	0.419
GLCM_Sum variance	0.991	0.994	0.145
GLCM_Sum entropy	0.991	0.992	0.980
GLCM_Difference variance	0.985	0.983	0.052
GLCM_Difference entropy	0.986	0.983	0.962
GLCM_ Information measure of correlation1	0.993	0.985	0.853
GLCM_ Information measure of correlation2	0.993	0.995	0.978
GLCM_Inverse difference normalized	0.985	0.983	0.869
GLCM_ Inverse difference moment normalized	0.985	0.983	0.715
Mean	0.991	0.989	0.981
Variance	0.991	0.989	0.981
Skewness	0.995	0.998	0.988
Kurtosis	0.997	0.998	0.992
Energy	0.991	0.989	0.981
Entropy	0.990	0.990	0.980
Area	0.986	0.988	0.976
Major Axis Length	0.988	0.979	0.979
Minor Axis Length	0.988	0.992	0.969
Eccentricity	0.837	0.275	0.427
Orientation	0.859	0.620	0.552
Convex Area	0.928	0.744	0.232
Equivdiameter	1.00	0.992	0.172
Solidity	0.806	0.728	0.920
Perimeter	0.991	0.978	0.078

TABLE 3. Inter-observer reproducibility of radiomic features (ICC).

Features	CLAHE RUN_1	CLAHE RUN_2	AHE RUN_1	AHE RUN_2	MANUAL
GLCM_Autocorrelation	0.996	0.988	0.996	0.999	0.189
GLCM_Contrast	0.961	1.000	1.000	1.000	0.031
GLCM_Correlation	0.993	1.000	1.000	1.000	0.576
GLCM_Correlation	0.993	1.000	1.000	1.000	0.576
GLCM_Cluster Prominence	0.991	0.996	0.999	1.000	0.001
GLCM_Cluster Shade	0.991	0.996	1.000	1.000	0.005
GLCM_Dissimilarity	0.980	0.967	1.000	1.000	0.124
GLCM_Energy	0.991	1.000	1.000	1.000	0.971
GLCM_Entropy	0.991	0.996	1.000	1.000	0.970
GLCM_Homogeneity	0.980	1.000	1.000	1.000	0.573
GLCM_Homogeneity	0.980	1.000	1.000	1.000	0.508
GLCM_Maximum probability	0.991	1.000	1.000	1.000	0.971
GLCM_Sum of squares	0.991	1.000	0.992	0.998	0.021
GLCM_Sum average	0.993	1.000	1.000	1.000	0.267
GLCM_Sum variance	0.990	1.000	1.000	1.000	0.064
GLCM_Sum entropy	0.991	0.996	1.000	1.000	0.970
GLCM_Difference variance	0.980	0.967	1.000	1.000	0.016
GLCM_Difference entropy	0.982	0.981	1.000	0.642	0.921
GLCM_ Information measure of correlation1	0.993	1.000	1.000	1.000	0.456
GLCM_ Information measure of correlation2	0.992	0.994	0.999	0.991	0.961
GLCM_Inverse difference normalized	0.98	1.000	1.000	1.000	0.994
GLCM_ Inverse difference moment normalized	0.98	1.000	1.000	1.000	0.994
Mean	0.990	0.996	0.990	0.992	0.988
Variance	0.990	0.996	0.996	0.997	0.989
Skewness	0.995	0.943	0.999	0.980	0.933
Kurtosis	0.997	0.871	0.999	0.970	0.995
Energy	0.990	0.996	0.995	0.996	0.989
Entropy	0.990	0.995	1.000	0.990	0.989
Area	0.985	0.992	0.989	0.999	0.986
Major Axis Length	0.985	0.991	0.981	0.997	0.989
Minor Axis Length	0.988	0.980	0.992	0.998	0.984
Eccentricity	0.835	0.679	0.810	0.963	0.589
Orientation	0.871	0.982	0.629	0.553	0.832
Convex Area	0.931	0.995	0.740	0.990	0.862
Equivdiameter	0.989	0.988	0.993	0.999	0.500
Solidity	0.996	0.790	0.877	0.774	0.127
Perimeter	0.995	0.990	0.979	0.937	0.659

In contrast to manual delineation, all three types of features had the worst reproducibility compared to other datasets. For textural features (0.450 ± 0.980, $p > 0.05$), 12 of 22 (54.5%) features were classified as excellent, one of 22 (4.54%) as fair, and eight of 22 (36.4%) as poor. Intensity features (ICC $= 0.983 \pm 0.502$, $p > 0.05$); however, they

have the highest reproducibility (100%) features compared to textural and shape features. For shape features (ICC = 0.589 ± 0.940, $p < 0.05$), only four of nine (44.4%) have excellent reproducibility, two of nine (22.2%) have fair reproducibility, and three of nine (33.3%) have poor reproducibility. Tables 2 and 3 tabulate the reproducibility analysis of intra-class CC and inter-observer from one subject.

Reproducibility for intensity histogram features were significantly higher when observed in images enhanced by CLAHE segmentation sets (ICC = 0.993 ± 0.502, $p < 0.05$) and image enhanced by AHE segmentation sets (ICC = 0.992 ± 0.502, $p < 0.05$) compared to the manual technique (ICC = 0.983 ± 0.502, $p > 0.05$). A similar trend was also observed for reproducibility in GLCM in which CLAHE and AHE had higher reproducibility. Hence, the reproducibility of radiomic features for image enhanced by CLAHE, AHE, and manual segmentation was 100%, 89% and 62%, respectively. Based on the observations, images enhanced by CLAHE had excellent reproducibility for all the features compared to other segmentations set.

The robustness of the technique was assessed by analyzing the ICC features extracted from inter- and intra-observers. In Figure 4A, we observed ICC of features extracted for the inter-observer segmentation group from image enhanced by CLAHE is the highest (mean ICC = 0.977 ± 0.579). In Figure 4B, higher ICC values were obtained in an image enhanced by CLAHE intra-observer segmentations (mean ICC = 0.979 ± 0.620). Figure 5 presents the normalized

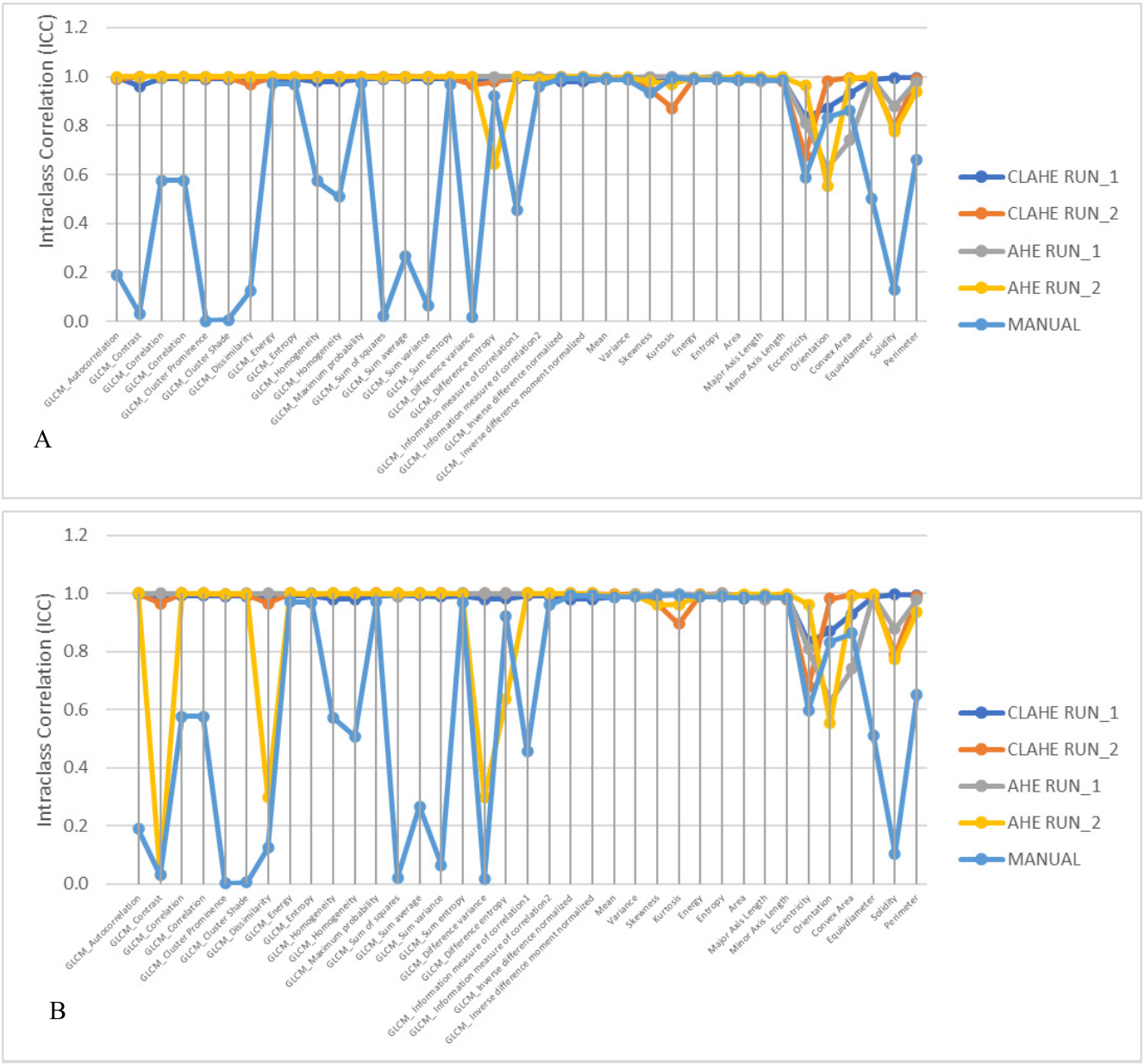

FIGURE 4. Line graph comparing (A) inter- and (B) intra-observer reproducibility of radiomic features. Run1 and run2 are different segmentation sets defined by different observers.

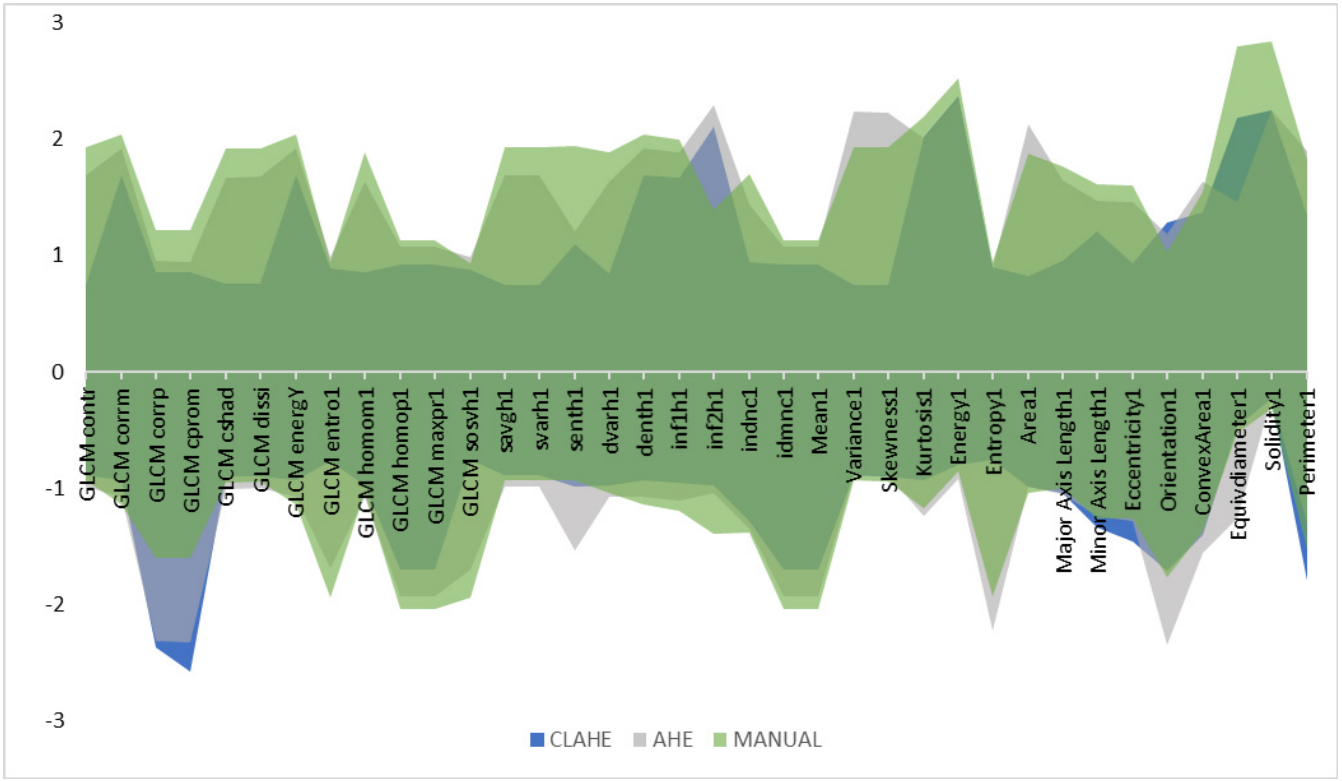

FIGURE 5. **Comparison of normalized feature range between manual delineation and two semiautomatic segmentations with two techniques of image enhancements.**

range of all features for 12 segmentations set (four manual, four images enhanced by CLAHE, and four images enhanced by AHE). It can be observed that the range of feature for image enhanced by CLAHE segmentations set was smaller than other segmentation sets.

IV. DISCUSSION

In this study, we assessed the robustness reproducibility of radiomics features of breast cancer through 2D mammograms. The results showed that most of these features achieved high reproducibility scores when contrast enhancement and semi-automatic segmentation were applied to the image dataset. To date, no studies have comprehensively evaluated the robustness of reproducibility of radiomics features using a 2D mammogram imaging modality. Medical imaging is frequently used and plays an important role in cancer staging, treatment planning, and treatment response monitoring in clinical oncology. With recent developments in CAD, it is notable that radiomics is essential for data mining, and predictive analysis is essential for evaluating the characteristics of tumor phenotype to widen the scope of imaging in clinical settings and potential for treatment planning and monitoring. The radiomic-extracted features from post-treatment compared to the pre-treatment CT images can be an early indicator for progression to local recurrence within six months after stereotactic ablative radiotherapy in early-stage lung cancer [34].

Pre-processing imaging, such as image enhancement, is extremely important to accurately diagnose tumor area by providing better contrast for evaluating the details in the image under investigation. Tumor segmentation is *de facto* in any radiomics study. With the aid of image enhancement, tumor segmentation can be accurately performed. In this study, CLAHE and AHE enhancement were used to segment breast tumor. Since manual delineation consumes much time and is prone to higher inter-observer variability, semiautomatic segmentation with image enhancement can provide more stable segmentation in a shorter time [6]. Reproducibility of tumor segmentation using semiautomatic segmentation is higher than manual delineation with image enhancement using AHE technique with respect to providing a good result [10]. However, despite its advantages, this technique also enhanced the background noise present in the image. Therefore, the CLAHE technique is implemented to overcome this drawback and reduce over-amplification of noise that is present in the AHE technique. The contrast in the image can be adjusted accordingly; hence, a clear enhanced image can be obtained without noise by using the CLAHE technique [15].

In this study, 37 radiomics imaging features were included and classified into three main features (6 tumor intensity histogram-based features, 22 textural features and 9 shape-based features). Reproducibility and robustness of these features were analyzed using manual delineation (without image enhancement), semiautomatic segmentation implemented

with CLAHE and AHE technique. The result show that the CLAHE technique segmentations had significantly higher ICC values than AHE technique segmentation and manual delineations. AHE techniques segmentations had slightly lower values when compared to CLAHE techniques but significantly higher and better robustness compared to manual delineations. The semiautomatic segmentation algorithm was performed by algorithm initialization, and the tumor region was digitized thus allowing for accurate segmentation without observer intervention [6].

Three particular quantitative imaging features were robust and had higher reproducibility than semiautomatic tumor segmentation with image enhancement CLAHE and AHE techniques ($p < 0.05$) compared to manual delineations. The CLAHE technique segmentation set had higher reproducibility for most GLCM-based features and shape-based features compared to AHE techniques. However, there was no significant difference observed in tumor intensity histogram-based features between CLAHE techniques and AHE techniques. These findings indicate that radiomics features derived from a semiautomatic algorithm with CLAHE image enhancement techniques produce higher reproducibility for a benign tumor.

The inter- and intra-observer reproducibility were assessed to evaluate the implementation from three different segmentation techniques. The ICC values for both inter- and intra-observer in CLAHE was higher indicates more reproducible features from the techniques. AHE techniques and manual delineation segmentations were inconsistent and fluctuating for both inter- and intra-observers. Both observed features were significantly lower than those seen in the CLAHE techniques.

Moreover, it is important to determine whether the features extracted from semiautomatic segmentations with image enhancement captured the same tumor image properties as seen with manual delineation. Therefore, we compared the normalized range for all features between these three images dataset (Fig. 5). We normalized every feature value with respect to all 12 (four CLAHE technique segmentations, +4 AHE technique segmentations, and +4 manual segmentations) segmentations using Z-score normalization. We observed that most features extracted from CLAHE and AHE technique segmentations were spread over significantly smaller ranges across observer range compared to those of the manual delineations.

Our findings concluded that CLAHE provides higher reproducibility of radiomic features extracted from semiautomatic segmentation. The only limitation of the study is not applying image descriptors with prediction models due to the small patient cohort even though clinical data are highly available and accessible.

V. CONCLUSION

In conclusion, this study should be applicable for predicting outcomes in patient with breast cancer. This work can be improved by collecting data from a multicenter with large prospective patient cohorts.

REFERENCES

[1] *Malaysia National Cancer Registry Report 2012–2016*, Nat. Cancer Registry Dept., Ministry Health Malaysia, Putrajaya, Malaysia, 2019.
[2] J. M. Seely and T. Alhassan, "Screening for breast cancer in 2018—What should we be doing today?" *Current Oncol.*, vol. 25, p. 115, Jun. 2018.
[3] L. Margolies, M. Salvatore, C. Eber, A. Jacobi, I.-J. Lee, M. Liang, W. Tang, D. Xu, S. Zhao, M. Kale, J. Wisnivesky, C. I. Henschke, and D. Yankelevitz, "The general radiologist's role in breast cancer risk assessment: Breast density measurement on chest CT," *Clin. Imag.*, vol. 39, no. 6, pp. 979–982, Jan. 2015.
[4] M. E. Suleiman, P. C. Brennan, and M. F. McEntee, "Diagnostic reference levels in digital mammography: A systematic review," *Radiat. Protection Dosimetry*, vol. 167, no. 4, pp. 608–619, Dec. 2014.
[5] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R. G. P. M. van Stiphout, P. Granton, C. M. L. Zegers, R. Gillies, R. Boellard, A. Dekker, and H. J. W. L. Aerts, "Radiomics: Extracting more information from medical images using advanced feature analysis," *Eur. J. Cancer*, vol. 48, no. 4, pp. 441–446, Mar. 2012.
[6] Q. Qiu, J. Duan, G. Gong, Y. Lu, D. Li, J. Lu, and Y. Yin, "Reproducibility of radiomic features with GrowCut and GraphCut semiautomatic tumor segmentation in hepatocellular carcinoma," *Transl. Cancer Res.*, vol. 6, no. 5, pp. 940–948, Oct. 2017.
[7] R. T. H. M. Larue, G. Defraene, D. De Ruysscher, P. Lambin, and W. van Elmpt, "Quantitative radiomics studies for tissue characterization: A review of technology and methodological procedures," *Brit. J. Radiol.*, vol. 90, no. 1070, Feb. 2017.
[8] J. P. B. O'Connor *et al.*, "Imaging biomarker roadmap for cancer studies," *Nature Rev. Clin. Oncol.*, vol. 14, no. 3, pp. 169–186, Feb. 2017.
[9] V. Kumar, Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, K. Forster, H. J. W. L. Aerts, A. Dekker, D. Fenstermacher, D. B. Goldgof, L. O. Hall, P. Lambin, Y. Balagurunathan, R. A. Gatenby, and R. J. Gillies, "Radiomics: The process and the challenges," *Magn. Reson. Imag.*, vol. 30, no. 9, pp. 1234–1248, Nov. 2012.
[10] C. Parmar, E. R. Velazquez, R. Leijenaar, M. Jermoumi, S. Carvalho, R. H. Mak, S. Mitra, B. U. Shankar, R. Kikinis, B. Haibe-Kains, P. Lambin, and H. J. W. L. Aerts, "Robust radiomics feature quantification using semiautomatic volumetric segmentation," *PLoS ONE*, vol. 9, no. 7, 2014.
[11] A. M. Garcia, A. S. Castrejón, M. Amo-Salas, J. F. L. Fidalgo, M. M. M. Sanchez, R. A. Cabellos, R. E. Aunion, and V. M. Madero "Glycolytic activity in breast cancer using (18)F-FDG PET/CT as prognostic predictor: A molecular phenotype approach," *Rev. Esp. Med. Nucl. Imagen Mol.*, vol. 35, no. 3, pp. 152–158, 2016.
[12] S. Jenifer, S. Parasuraman, and A. Kadirvelu, "Contrast enhancement and brightness preserving of digital mammograms using fuzzy clipped contrast-limited adaptive histogram equalization algorithm," *Appl. Soft Comput.*, vol. 42, pp. 167–177, May 2016.
[13] T.-L. Ji, M. K. Sundareshan, and H. Roehrig, "Adaptive image contrast enhancement based on human visual properties," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 573–586, Dec. 1994.
[14] K. Santhi and R. S. D. W. Banu, "Contrast enhancement using brightness preserving histogram plateau limit technique," *Int. J. Eng. Technol.*, vol. 6, no. 3, pp. 1447–1453, 2014.
[15] A. Yechur, "Contrast enhancement of MRI brain images using histogram equalization Techniques," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 7, no. 8, pp. 81–89, Aug. 2019.
[16] P. S. Patil and P. P. Pawade, "Biomedical image brightness preservation and segmentation technique using CLAHE and Wiener filtering," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 5, no. 4, pp. 808–812, 2016.
[17] R. Gordon and R. M. Rangayyan, "Feature enhancement of film mammograms using fixed and adaptive neighborhoods: Correction," *Appl. Opt.*, vol. 23, no. 13, p. 2055, Jul. 1984.
[18] R. M. Rangayyan, L. Shen, Y. Shen, J. E. L. Desautels, H. Bryant, T. J. Terry, N. Horeczko, and M. S. Rose, "Improvement of sensitivity of breast cancer diagnosis with adaptive neighborhood contrast enhancement of mammograms," *IEEE Trans. Inf. Technol. Biomed.*, vol. 1, no. 3, pp. 161–170, Sep. 1997.
[19] J. Hwang, H.-J. Kim, P. Lemaillet, H. Wabnitz, D. Grosenick, L. Yang, T. Gladytz, D. McClatchy, III, D. Allen, K. Briggman, and B. Pogue, "Polydimethylsiloxane tissue-mimicking phantoms for quantitative optical medical imaging standards," *Proc. SPIE*, vol. 10056, Mar. 2017, Art. no. 1005603.

[20] R. Shanker and M. Bhattacharya, "An automated computer-aided diagnosis system for classification of MR images using texture features and gbest-guided gravitational search algorithm," *Biocybernetics Biomed. Eng.*, vol. 40, no. 2, pp. 815–835, Apr. 2020.
[21] U. R. Acharya, J. E. W. Koh, Y. Hagiwara, J. H. Tan, A. Gertych, A. Vijayananthan, N. A. Yaakup, B. J. J. Abdullah, M. K. B. M. Fabell, and C. H. Yeong, "Automated diagnosis of focal liver lesions using bidirectional empirical mode decomposition features," *Comput. Biol. Med.*, vol. 94, pp. 11–18, Mar. 2018.
[22] K. Akila, L. S. Jayashree, and A. Vasuki, "Mammographic image enhancement using indirect contrast enhancement techniques—A comparative study," *Procedia Comput. Sci.*, vol. 47, pp. 255–261, Jan. 2015.
[23] M. M. Jadoon, Q. Zhang, I. U. Haq, S. Butt, and A. Jadoon, "Three-class mammogram classification based on descriptive CNN features," *BioMed Res. Int.*, vol. 2017, pp. 1–11, Jan. 2017.
[24] L. Yu, G. Tao, L. Zhu, G. Wang, Z. Li, J. Ye, and Q. Chen, "Prediction of pathologic stage in non-small cell lung cancer using machine learning algorithm based on CT image feature analysis," *BMC Cancer*, vol. 19, no. 1, pp. 1–12, Dec. 2019.
[25] M. Wu, H. Tan, F. Gao, J. Hai, P. Ning, J. Chen, S. Zhu, M. Wang, S. Dou, and D. Shi, "Predicting the grade of hepatocellular carcinoma based on non-contrast-enhanced MRI radiomics signature," *Eur. Radiol.*, vol. 29, no. 6, pp. 2802–2811, Jun. 2019.
[26] P. Lambin, R. T. H. Leijenaar, T. M. Deist, J. Peerlings, E. E. C. de Jong, J. van Timmeren, S. Sanduleanu, R. T. H. M. Larue, A. J. G. Even, A. Jochems, Y. van Wijk, H. Woodruff, J. van Soest, T. Lustberg, E. Roelofs, W. van Elmpt, A. Dekker, F. M. Mottaghy, J. E. Wildberger, and S. Walsh, "Radiomics: The bridge between medical imaging and personalized medicine," *Nature Rev. Clin. Oncol.*, vol. 14, no. 12, pp. 749–762, Dec. 2017.
[27] N. Al-Najdawi, M. Biltawi, and S. Tedmori, "Mammogram image visual enhancement, mass segmentation and classification," *Appl. Soft Comput.*, vol. 35, pp. 175–185, Oct. 2015.
[28] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox, and F. Prior, "The cancer imaging archive (TCIA): Maintaining and operating a public information repository," *J. Digit. Imag.*, vol. 26, no. 6, pp. 1045–1057, Dec. 2013.
[29] R. S. Lee, F. Gimenez, A. Hoogi, K. K. Miyake, M. Gorovoy, and D. L. Rubin, "A curated mammography data set for use in computer-aided detection and diagnosis research," *Sci. Data*, vol. 4, no. 1, Dec. 2017, Art. no. 170177.
[30] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.
[31] J. Zhou, H. Tan, Y. Bai, J. Li, Q. Lu, R. Chen, M. Zhang, Q. Feng, and M. Wang, "Evaluating the HER-2 status of breast cancer using mammography radiomics features," *Eur. J. Radiol.*, vol. 121, Dec. 2019, Art. no. 108718.
[32] E. Segal, C. B. Sirlin, C. Ooi, A. S. Adler, J. Gollub, X. Chen, B. K. Chan, G. R. Matcuk, C. T. Barry, H. Y. Chang, and M. D. Kuo, "Decoding global gene expression programs in liver cancer by noninvasive imaging," *Nature Biotechnol.*, vol. 25, no. 6, pp. 675–680, Jun. 2007.
[33] S. S. F. Yip and H. J. W. L. Aerts, "Applications and limitations of radiomics," *Phys. Med. Biol.*, vol. 61, no. 13, pp. R150–R166, Jul. 2016.
[34] S. A. Mattonen, D. A. Palma, C. Johnson, A. V. Louie, M. Landis, G. Rodrigues, I. Chan, R. Etemad-Rezai, T. P. C. Yeung, S. Senan, and A. D. Ward, "Detection of local cancer recurrence after stereotactic ablative radiation therapy for lung cancer: Physician performance versus radiomic assessment," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 94, no. 5, pp. 1121–1128, Apr. 2016.

SITI FAIRUZ MAT RADZI received the B.Sc. degree (Hons.) in applied physics from Universiti Tun Hussein Onn Malaysia (UTHM), in 2019. She is currently pursuing the master's degree with Universiti Putra Malaysia (UPM). Her research interests include medical image processing and machine learning.

MUHAMMAD KHALIS ABDUL KARIM received the B.Sc. degree (Hons.) in diagnostic imaging and radiotherapy from Universiti Kebangsaan Malaysia (UKM), in 2008, and the M.Sc. and Ph.D. degrees in physics from Universiti Teknologi Malaysia, in 2013 and December 2016, respectively. He has expertise in medical physics particularly radiation dosimetry in medical imaging. He has been appointed as an Advisory Member in several technical committee. He is also actively involved giving lecture in several workshops on radiation protection and nuclear security. He has authored/coauthored 35 international journals articles (with impact factor ranging from 0.2 to 5.1), ten international proceedings, and three national proceedings. His research interests include medical imaging, radiation dosimeters, and image processing.

M IQBAL SARIPAN, photograph and biography not available at the time of publication.

MOHD AMIRUDDIN ABD RAHMAN received the B.S. degree in electrical engineering from Purdue University, USA, in 2006, the M.S. degree in sensors and instrumentation from Universiti Putra Malaysia (UPM), in 2011, and the Ph.D. degree in electronic and electrical engineering from The University of Sheffield and UPM, in 2016. He has been as Post-graduate Researcher with Alcatel-Lucent Bell Labs, Ireland, as a working on co-localization and tracking algorithm for WLAN networks. He has also served as an Invited Researcher with Bell Labs Belgium to apply the localization system for Bell-Labs Future-X day. He is currently a Senior Lecturer with the Department of Physics, UPM. He leads the research in indoor localization within radio frequency and microwave research group. His research interests include signal processing, pattern recognition, and machine learning algorithms.

NURUL HUDA OSMAN, photograph and biography not available at the time of publication.

ENTESAR ZAWAM DALAH, photograph and biography not available at the time of publication.

NORAMALIZA MOHD NOOR, photograph and biography not available at the time of publication.