

# Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

## Neutrosophic segmentation of breast lesions for dedicated breast computed tomography

Juhun Lee  
Robert M. Nishikawa  
Ingrid Reiser  
John M. Boone

# Neutrosophic segmentation of breast lesions for dedicated breast computed tomography

Juhun Lee,<sup>a,\*</sup> Robert M. Nishikawa,<sup>a</sup> Ingrid Reiser,<sup>b</sup> and John M. Boone<sup>c</sup>

<sup>a</sup>University of Pittsburgh, Department of Radiology, Pittsburgh, Pennsylvania, United States

<sup>b</sup>University of Chicago, Department of Radiology, Chicago, Illinois, United States

<sup>c</sup>University of California Davis Medical Center, Department of Radiology, Sacramento, California, United States

**Abstract.** We proposed the neutrosophic approach for segmenting breast lesions in breast computed tomography (bCT) images. The neutrosophic set considers the nature and properties of neutrality (or indeterminacy). We considered the image noise as an indeterminate component while treating the breast lesion and other breast areas as true and false components. We iteratively smoothed and contrast-enhanced the image to reduce the noise level of the true set. We then applied one existing algorithm for bCT images, the RGI segmentation, on the resulting noise-reduced image to segment the breast lesions. We compared the segmentation performance of the proposed method (named as NS-RGI) to that of the regular RGI segmentation. We used 122 breast lesions (44 benign and 78 malignant) of 111 noncontrast enhanced bCT cases. We measured the segmentation performances of the NS-RGI and the RGI using the Dice coefficient. The average Dice values of the NS-RGI and RGI were 0.82 and 0.80, respectively, and their difference was statistically significant ( $p$  value = 0.004). We conducted a subsequent feature analysis on the resulting segmentations. The classifier performance for the NS-RGI (AUC = 0.80) improved over that of the RGI (AUC = 0.69,  $p$  value = 0.006). © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.5.1.014505]

Keywords: neutrosophy; segmentation; quantitative feature analysis; breast CT; CADx.

Paper 17359PR received Dec. 6, 2017; accepted for publication Feb. 12, 2018; published online Mar. 6, 2018.

## 1 Introduction

In image segmentation problems, the main goal is to distinguish the foreground from the background in a given image. However, all natural images (i.e., not simulated or computer-generated) include various types of noise, which are neither foreground nor background in segmentation problems. Such noise degrades the performance of any existing segmentation algorithm.

In dedicated breast computed tomography (bCT), quantum noise is one of the major sources of noise. Quantum noise creates a readily visible salt-and-pepper noise in reconstructed bCT images that can degrade any segmentation algorithms. We can reduce quantum noise by increasing the radiation dose, but this may increase the cancer risk to the patient. Thus, one needs to balance the image quality (or image noise) and radiation dose, to maximize the patient benefits.

Many ways exist to control the noise in bCT images. One can use different reconstruction kernels, e.g., smooth kernels for low noise but low spatial resolution, or sharp kernels for high noise but high spatial resolution. Researchers are developing iterative image reconstruction algorithms for bCT,<sup>1,2</sup> which can suppress the image noise while maintaining the spatial resolution and contrast even in low radiation dose scans. In addition to these noise control methods in the reconstruction domain, we can reduce image noise after reconstruction, directly on bCT images. One may simply smooth the entire image or region of interest of the image to reduce the effect of the noise. However, simply smoothing can remove useful information (e.g., edge of lesion) for segmentation. In this respect, it is beneficial to develop algorithms that suppress image noise while preserving useful information for segmentation.

This study attempted to distinguish and suppress the noise in the image after reconstruction, before applying the segmentation algorithms. Then, we applied a segmentation algorithm to noise-suppressed images to improve segmentation performance. For this study, we adopted the neutrosophy theory to achieve the objective.

Neutrosophy is a branch of philosophy that generalizes dialectics and studies the concept and properties of neutralities.<sup>3</sup> Neutrosophy theory considers entity  $\langle A \rangle$  and its relation to  $\langle \text{Anti-}A \rangle$  and  $\langle \text{Neut-}A \rangle$ , where  $\langle \text{Anti-}A \rangle$  and  $\langle \text{Neut-}A \rangle$  represent the opposite and the neutrality entity of  $\langle A \rangle$ , respectively. Neutrosophy covers various concepts, including neutrosophic logic, neutrosophic probability, neutrosophic set (NS), etc.<sup>3</sup> One can consider neutrosophy logic as the generalized version of fuzzy logic, where it explicitly takes into account the neutrality or indeterminacy of the problem.<sup>4</sup>

We can consider the noise in the image as a neutral or indeterminate element. The classical set and fuzzy set only handle this neutral or indeterminate element partially, as neutrality or indeterminacy is absorbed into either the true or false set (or background or foreground set). Due to the existence of neutrality, one can expect that the neutrosophy set can handle the noise element in the image effectively.

This study, therefore, used the neutrosophy theory, specifically the neutrosophic set, to tackle segmentation problems for dedicated breast CT images. This study adapted and modified the segmentation approach proposed by Guo and Cheng<sup>5</sup> to solve our problem, i.e., segmenting breast lesions in bCT images. Xian et al. used the original method for segmenting breast lesions in two-dimensional (2-D) ultrasound images.<sup>6</sup>

\*Address all correspondence to: Juhun Lee, E-mail: [leej15@upmc.edu](mailto:leej15@upmc.edu)

We extended and modified their method to segment three-dimensional (3-D) bCT images.

## 2 Methods

### 2.1 Dataset

We used a dataset of 122 pathology proven breast lesions (44 malignant and 78 benign) of 111 noncontrast-enhanced bCT cases collected under an approved Institutional Review Board (IRB) protocol (Table 1). The prototype dedicated breast CT system at the University of California at Davis<sup>7</sup> was utilized to obtain all patients' bCT cases, and operated at 80 kVp with variable mAs to provide a similar mean glandular dose of standard two view mammograms. This study used an FeldKamp-Davis-Kress (FDK) reconstruction algorithm<sup>8</sup> to reconstruct each image.

**Table 1** Characteristics of breast CT dataset.

			All
Total number of lesions			122
Subject age (years)		Mean (min, max)	55.6 [35, 80]
Lesion diameter (mm)		Mean (min, max)	13.8 [4, 35]
Breast density (% among lesions considered)		1	12 (10%)
		2	46 (38%)
		3	46 (38%)
		4	18 (14%)
Diagnosis*	Malignant (% among malignant lesions considered)	IDC	55 (71%)
		IMC	11 (14%)
		ILC	7 (9%)
		DCIS	4 (5%)
		Lymphoma	1 (1%)
	Total		44
	Benign (% among benign lesions considered)	FA	18 (41%)
		FC	7 (16%)
		FCC	4 (9%)
		PASH	1 (2%)
CAPPS		2 (4%)	
Other benign lesions such as sclerosing adenosis and cyst		12 (28%)	
Total		78	

\*IDC, invasive ductal carcinoma; IMC, invasive mammary carcinoma; ILC, invasive lobular carcinoma; DCIS, ductal carcinoma *in situ*; FA, fibroadenoma; FC, fibrocystic; FCC, fibrocystic changes; PASH, pseudoangiomatous stromal hyperplasia; CAPPS, columnar alteration with prominent apical snouts and secretions.

### 2.2 Preprocessing: Image Normalization

To reduce false positives, i.e., assigning a nonlesion to foreground, we preprocessed bCT images such that they are within the range of the possible voxel intensity of breast tissues. We assumed a range of Hounsfield unit (HU) values for breast tissue as  $[-500, 300]$  HU.  $-500$  HU and  $300$  HU are the highest HU number for lung<sup>9</sup> and the lowest HU number for cortical bone.<sup>10</sup> In fact, a previous study<sup>11</sup> showed that HU values for breast tissue in bCT images can range from  $-350$  (adipose tissue at low tube voltage) to  $100$  HU (breast cancer at high tube voltage). Another study<sup>12</sup> showed that contrast can enhance malignant breast lesions in bCT images up to  $120$  HU. Thus, we can expect that the range  $[-500, 300]$  HU should include all possible values for breast tissue in bCT images. Any voxels outside of this HU range were replaced with the averaged HU value of other neighboring voxels.

### 2.3 Preprocessing: Neutrosophic Image Enhancement for Breast CT Images

Let  $x(t, i, f)$  be an element of the NS.  $t$ ,  $i$ , and  $f$  refer to the membership (%) of the element  $x$  in the neutrosophic components; true ( $T$ ), indeterminacy ( $I$ ), and false ( $F$ ), respectively. In this study, we treated  $T$ ,  $I$ , and  $F$  as foreground, noise, and background, respectively.

Let  $V(x, y, z)$  be a voxel in the bCT image. The neutrosophic representation of  $V(x, y, z)$  is given as  $V_{NS}(x, y, z) = \{T(x, y, z), I(x, y, z), F(x, y, z)\}$ , where each neutrosophic component is defined as

$$T(x, y, z) = \frac{f_w[p(x, y, z)] - \min\{f_w[p(x, y, z)]\}}{\max\{f_w[p(x, y, z)]\} - \min\{f_w[p(x, y, z)]\}}, \quad (1)$$

$$I(x, y, z) = \frac{g[p(x, y, z)] - \min\{g[p(x, y, z)]\}}{\max\{g[p(x, y, z)]\} - \min\{g[p(x, y, z)]\}}, \quad (2)$$

$$F(x, y, z) = 1 - T(x, y, z), \quad (3)$$

where  $p(x, y, z)$  is the intensity level of  $V(x, y, z)$ , and  $f_w(\cdot)$  represents the mean filtering of the image with a cubic window size of  $w \times w \times w$ , and  $g(k) = |k - f_w(k)|$ . We set  $w$  as three following the choice of a mean filter window size of a previous study on classifying breast lesions in breast CT.<sup>13</sup>

The  $\alpha$ -mean operation,  $\alpha[V_{NS}(x, y, z)] = \{T_\alpha(x, y, z), I_\alpha(x, y, z), F_\alpha(x, y, z)\}$  is defined as

$$T_\alpha(x, y, z) = \begin{cases} T(x, y, z), & I(x, y, z) < \alpha \\ f_w[T(x, y, z)], & I(x, y, z) \geq \alpha \end{cases} \quad (4)$$

$$F_\alpha(x, y, z) = \begin{cases} F(x, y, z), & I(x, y, z) < \alpha \\ f_w[F(x, y, z)], & I(x, y, z) \geq \alpha \end{cases} \quad (5)$$

$$I_\alpha(x, y, z) = \frac{g[T_\alpha(x, y, z)] - \min\{g[T_\alpha(x, y, z)]\}}{\max\{g[T_\alpha(x, y, z)]\} - \min\{g[T_\alpha(x, y, z)]\}}. \quad (6)$$

If the indeterminacy level of a voxel is higher than  $\alpha$ , the  $\alpha$ -mean operation locally smooths the portion around that voxel. We empirically set  $\alpha$  as  $0.9$ .

The  $\beta$ -enhancement operation,  $\beta[V_{NS}(x, y, z)] = \{T_\beta(x, y, z), I_\beta(x, y, z), F_\beta(x, y, z)\}$  is defined as

$$T_\beta(x, y, z) = \begin{cases} T(x, y, z), & I(x, y, z) < \beta \\ h[T(x, y, z)], & I(x, y, z) \geq \beta \end{cases} \quad (7)$$

$$F_\beta(x, y, z) = \begin{cases} F(x, y, z), & I(x, y, z) < \alpha \\ h[F(x, y, z)], & I(x, y, z) \geq \alpha \end{cases} \quad (8)$$

$$I_\beta(x, y, z) = \frac{g[T_\beta(x, y, z)] - \min\{g[T_\beta(x, y, z)]\}}{\max\{g[T_\beta(x, y, z)]\} - \min\{g[T_\beta(x, y, z)]\}}, \quad (9)$$

where  $h(k) = \begin{cases} 2k^2, & k \leq 0.5 \\ 1 - 2(1 - k)^2, & k > 0.5 \end{cases}$ . The  $\beta$ -enhancement operation enhances the contrast of the given volumetric image by reducing the intensity level of a voxel when its corresponding indeterminacy level is higher than  $\beta$ . We empirically set  $\beta$  as 0.5.

The  $\gamma$ -plateau operation,  $\gamma[V_{NS}(x, y, z)] = \{T_\gamma(x, y, z), I_\gamma(x, y, z), F_\gamma(x, y, z)\}$  is defined as

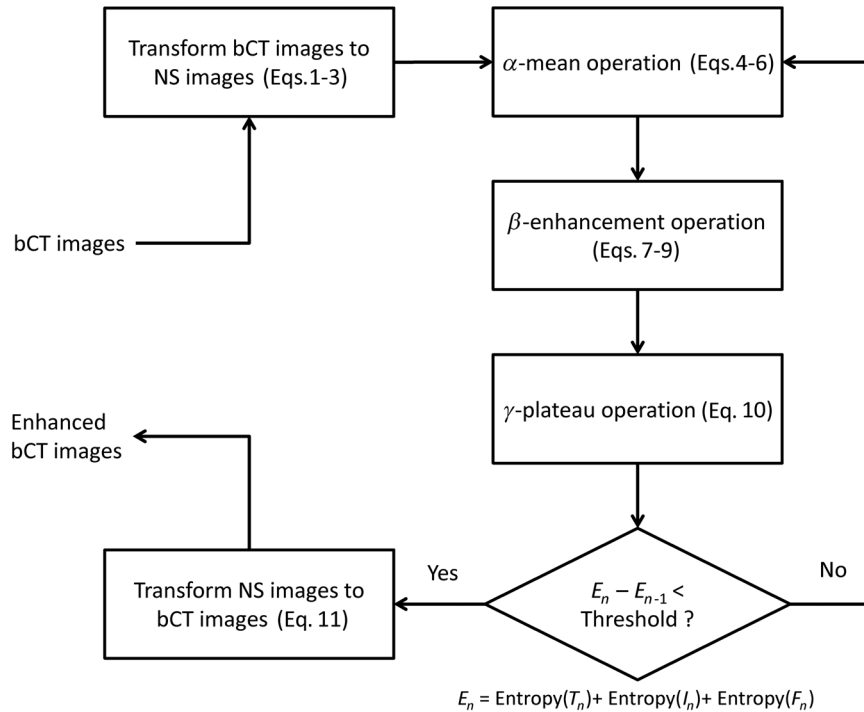
$$I_\gamma(x, y, z) = \frac{u[T(x, y, z)] - \min\{u[T(x, y, z)]\}}{\max\{u[T(x, y, z)]\} - \min\{u[T(x, y, z)]\}}, \quad (10)$$

where  $T_\gamma = T$  and  $F_\gamma = T$ . In Eq. (10),  $u[T(x, y, z)] = |f_w[T(x, y, z)] - \Delta T(x, y, z)|$  and  $\Delta$  is the Laplace operator. The  $\gamma$ -plateau operation is a new addition to the original NS approach proposed by Guo and Cheng that changes the indeterminacy set to an edge-enhanced image. When combined with the  $\alpha$ -mean operation and when the voxel intensity is higher than  $\alpha$ , the  $\gamma$ -plateau operation smooths the volume surrounding the given voxel. As a result, the  $\alpha$ -mean operation with the  $\gamma$ -plateau operation smooths the peaks and valleys in the foreground and, therefore, creates the plateaued (i.e., smoothed) foreground.

Once the true (i.e.,  $T$  or foreground), indeterminacy (i.e.,  $I$  or noise), and false (i.e.,  $F$  or background) components of the given image in the NS domain no longer change, which is measured by the sum of image entropies<sup>14</sup> of true, indeterminacy, and false sets,  $V_{NS}(x, y, z)$  is transformed back to  $V(x, y, z)$  with  $(\lambda, w) = (0.5, 3)$  as

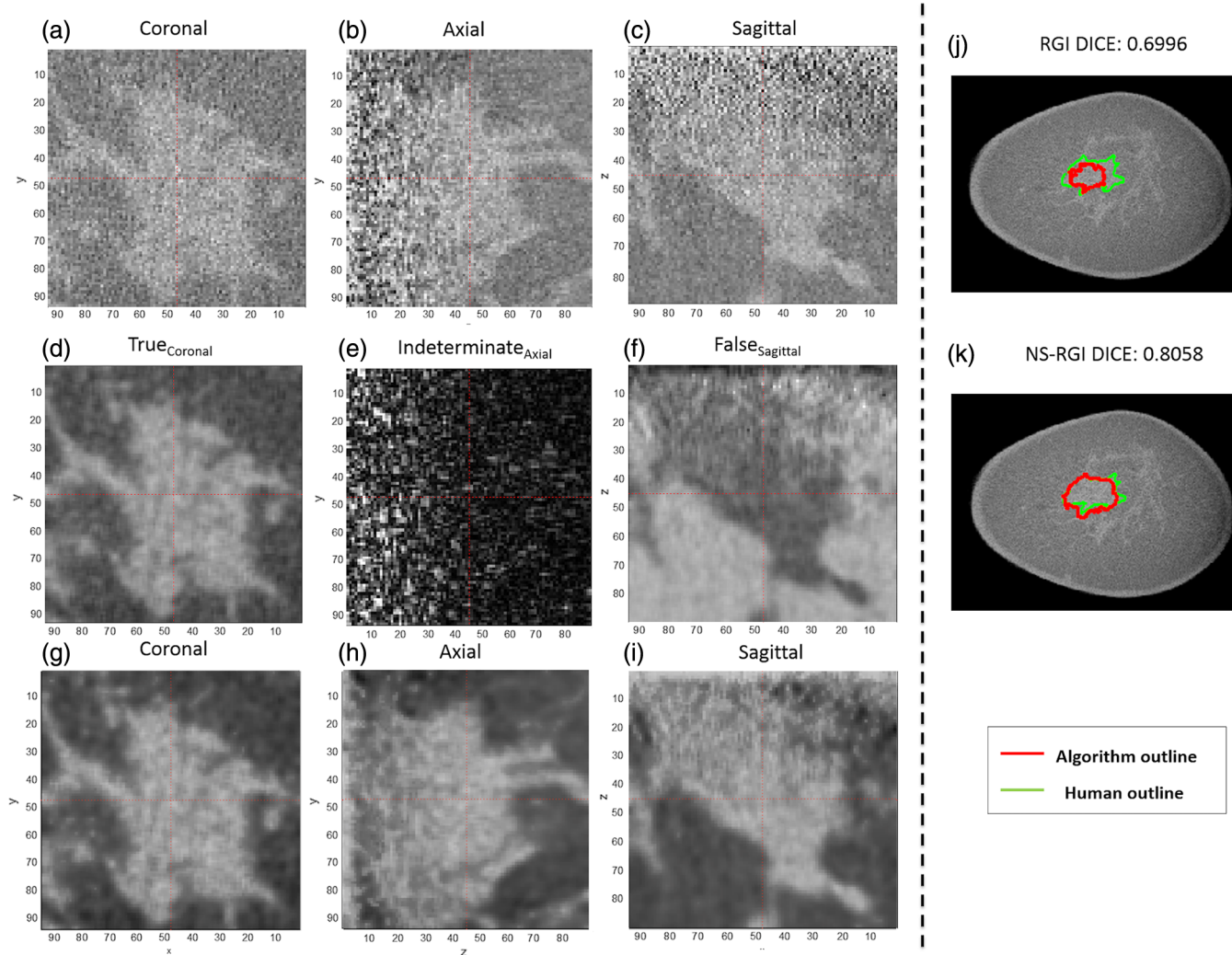
$$V(x, y, z) = \begin{cases} T(x, y, z), & I(x, y, z) \leq \lambda \\ f_w[T(x, y, z)], & I(x, y, z) > \lambda \end{cases} \quad (11)$$

We set the threshold to stop the enhancement as 0.0001. In addition, we set the maximum iteration of applying NS enhancement as 50. Figure 1 shows a flowchart of the procedures of the NS enhancement for bCT images, and Fig. 2 shows the effect of the NS enhancement on an example bCT image. Note, we adapted all parameter values from the original study of Guo and Cheng<sup>5</sup> and further optimized to achieve the best segmentation outcomes.



**Fig. 1** This diagram illustrates the procedures for the proposed neutrosophic image enhancement for bCT images. The algorithm transforms the bCT images into NS domain by assigning each voxel's membership in true (foreground), false (background), and indeterminacy (noise) sets. After that, three operations iteratively smooth and enhance the NS image to increase the contrast between true (breast lesion) and false (other breast tissue) sets by isolating image noise. Once the changes in true, intermediate, and false sets are stabilized, the algorithm transforms the NS images back to create cleaned or enhanced bCT images.





**Fig. 2** This figure illustrates how the proposed method enhances or cleans the given image for segmentation. (a)–(c) Input image in coronal, axial, and sagittal view. (d)–(f) Images in NS domain after one iteration. (g)–(i) Output image in coronal, axial, and sagittal view. (j)–(k) Segmentation results in coronal view for RGI and NS-RGI. It is clear that the NS method was able to clean the noise from the image, while retaining other information (e.g., lesion edge information) in the image, thus resulted in better segmentation.

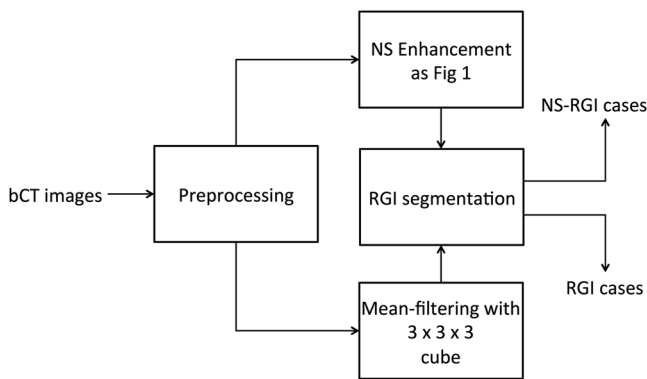
## 2.4 Neutrosophic Breast Lesion Segmentation

Once the image is enhanced or cleaned, we can use any existing segmentation algorithms to segment breast lesions in the bCT images. To show the effectiveness of the NS enhancement, we selected RGI segmentation, which is an existing algorithm,<sup>15</sup> and evaluated the improvement in segmentation performance with or without the NS enhancement.

The RGI segmentation is a semiautomatic algorithm that requires a manually allocated lesion center to search the boundary of that lesion. For each image, a research specialist with over 15 years of experience in mammography provided the lesion center and the lesion boundary. It was shown that the RGI segmentation algorithm can successfully segment breast lesions in bCT images.<sup>15</sup> We refer to the RGI segmentation algorithm applied on NS-enhanced images as the NS-RGI segmentation algorithm and corresponding images used for the NS-RGI segmentation algorithm as NS-RGI images. Similarly, we refer to the RGI segmentation applied on nonenhanced images as the RGI segmentation algorithm and corresponding images used for

the RGI segmentation algorithm as RGI images. Figure 3 shows how we created NS-RGI images and RGI images for the study. Note that we smoothed bCT images for RGI images. For both NS-RGI and RGI images, we selected the volume of interest (VOI) from the lesion center, to reduce false positives and increase the processing speed for the segmentation algorithm. We used a cube with an edge length of 35 mm for the VOI. We used the Dice coefficient<sup>16</sup> between the lesion boundary computed by the algorithm and that of the aforementioned research specialist as our figure of merit. Figures 2(j) and 2(k) show the Dice values for an example malignant breast lesion. Then, we compared the segmentation performance of the RGI segmentation algorithm on NS-RGI images to that of RGI images via bootstrap sampling method.

Once the volumetric segmentation was completed, we created a 3-D surface representation of the segmented result, using an existing algorithm (isosurface in MATLAB®), to compute 3-D surface image features explained in the next section. As the mean resolution of bCT images is around 350  $\mu\text{m}$  (i.e.,

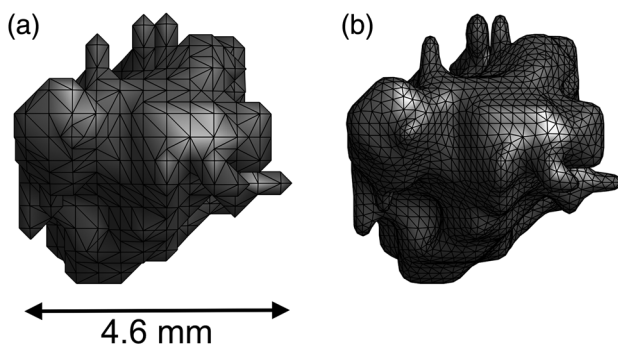


**Fig. 3** This diagram illustrates how we created NS-RGI and RGI cases for this study. For both NS-RGI and RGI cases, bCT images were first preprocessed. For NS-RGI cases, bCT images were cleaned or enhanced via the proposed NS enhancement and then the RGI segmentation algorithm was applied to the resulting enhanced images. For RGI cases, bCT images smoothed with  $3 \times 3 \times 3$  cube to reduce the effect of the noise and then the RGI segmentation algorithm was applied on the smoothed images.

0.35 mm), small lesions with a diameter less than 10 mm only have less than 30 vertices in each dimension for their 3-D surface representation. This resulted in crude 3-D surface representations of small breast lesions [e.g., Fig. 4(a)]. Thus, we interpolated the vertices of the 3-D surface of a small lesion such that it can have at least 30 vertices available in each dimension. For this, we used a mesh-subdivision algorithm<sup>17</sup> to interpolate the vertices of the 3-D surface representation of small lesions. We treated any lesions with maximum diameter smaller than 10 mm as small lesions (following the criteria used in a previous study<sup>18</sup>) and applied the vertex interpolation method explained above. Figure 4 shows the effect of mesh-subdivision on a small benign breast lesion (max diameter is 4.6 mm). One can see the smoother representation of the given lesion after the mesh-subdivision [Fig. 4(b)].

## 2.5 Quantitative Image Feature Analysis

We conducted quantitative feature analysis to determine if the cleaned image and the associated improved segmentation actually resulted in improved classification. We extracted 23 quantitative image features from the segmented lesions, which



**Fig. 4** This figure illustrates how mesh-subdivision improves the surface representation of small breast lesions less than 10 mm. (a) A small benign lesion with a maximum diameter of 4.6 mm. (b) The small lesion after the mesh-subdivision. One can see the improvement in the surface representation of the breast lesion, especially on spiculated margins.

were used in previous studies on breast CT.<sup>13,19,20</sup> The quantitative image features included: (1) four histogram features, which quantify the gray-value distribution in and around a breast lesion, (2) seven shape features, which summarize the overall shape of a breast lesion, (3) five margin features, which quantify the morphology in a breast lesion's margin, (4) four texture features, which are 3-D versions of a gray-level co-occurrence matrix, and (5) surface features, which summarize the variations over a lesion's surface.

Using the leave-one-out cross-validation (LOOCV), we selected the most salient features for classifying breast lesions using a feature selection algorithm (sequentialfs in MATLAB®). Then, we trained a linear discriminant analysis (LDA) classifier using the same training samples in LOOCV. We tested the trained classifier using the held-out sample. We used the area under the receiver operating characteristic curve (AUC) as a figure of merit. We compared the AUC values of the classifier trained on NS-RGI images and those of RGI images. We used the Delong's method<sup>21</sup> to compare the AUC values of two classifiers (i.e., one is trained on NS-RGI cases, and another is trained on RGI cases). In addition, we estimated the 95% confidence interval for AUC values of both NS-RGI and RGI classifiers. We refer to the classifier trained on NS-RGI images and RGI images as the NS-RGI classifier and the RGI classifier, respectively.

## 3 Results

The mean and standard deviation of Dice coefficients for the NS-RGI segmentation algorithm and the RGI segmentation algorithm were [0.80, 0.12] and [0.82, 0.09], respectively. The difference between the Dice values of the NS-RGI and RGI segmentation algorithms was statistically significant ( $p$  value = 0.004) (Table 2).

As breast density can affect the segmentation performance, we compared the segmentation performances of the NS-RGI and RGI segmentation algorithms in terms of breast density levels following BI-RADS.<sup>23</sup> The NS-RGI segmentation algorithm achieved better segmentation performance in terms of the average Dice value than that of the RGI segmentation algorithm for all density levels (Table 2). However, only density level 2 showed a statistically significant improvement.

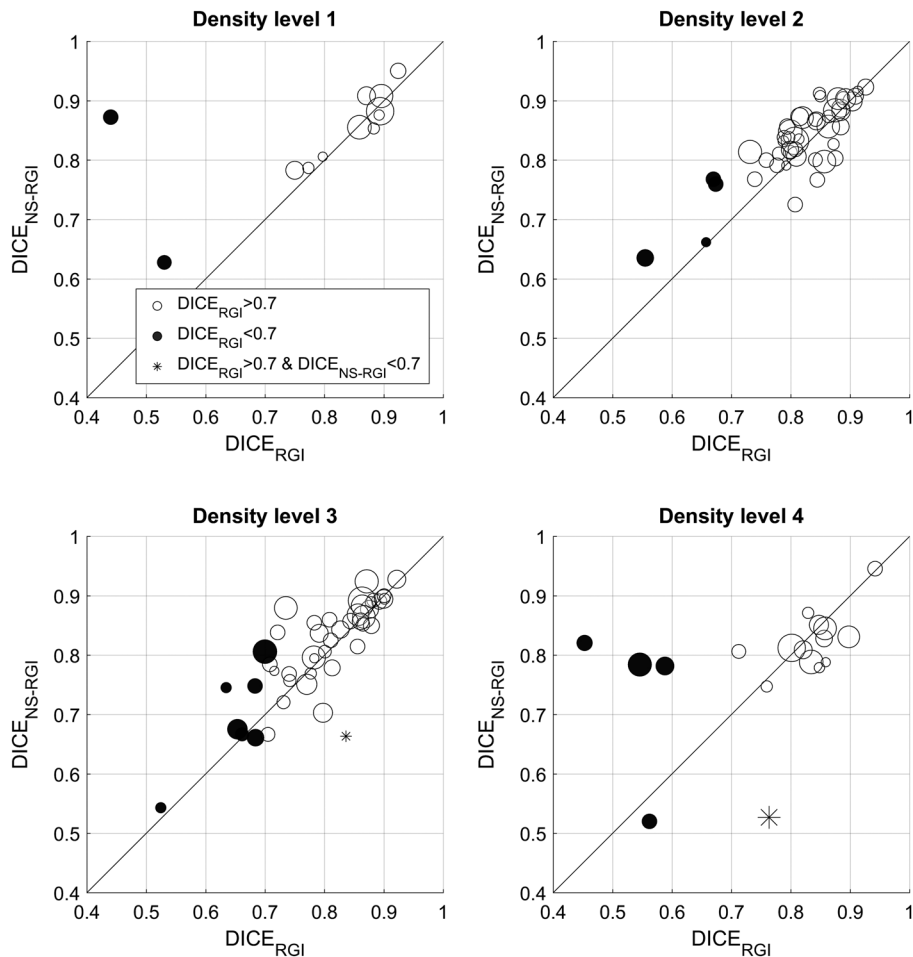
We considered Dice value of 0.7 or higher as a good segmentation outcome.<sup>24</sup> Based on this criterion, the NS-RGI segmentation algorithm showed similar segmentation performances for the breast lesion cases with an RGI segmentation performance of 0.7 or higher. However, the NS-RGI segmentation algorithm showed better segmentation performances for 11 out of 17 cases where the RGI segmentation failed (black circles in Fig. 5, Dice values < 0.7). For the remaining 6 of 17 cases, both the NS-RGI and RGI showed similar performance. From this, we can conclude that the proposed method can clean or enhance the given image such that it allows the RGI algorithm to segment the breast lesions that previously failed without the NS enhancement.

The AUC values for the NS-RGI and RGI classifiers obtained from LOOCV were 0.8 (95% CI: [0.73, 0.88]) and 0.69 (95% CI: [0.6 0.78]), respectively (Table 3). The difference in the AUC values of the NS-RGI and RGI classifiers was 0.11 with 95% CI of [0.032 0.19]. The differences between the two AUC values were statistically significant with a  $p$ -value of 0.006.

Table 2 Segmentation performance comparison for NS-RGI and RGI.

	NS-RGI	RGI	Difference	<i>p</i> -value
	Mean [95%CI]	Mean [95%CI]	Mean [95%CI]	
All	0.82 [0.8, 0.83]	0.80 [0.78, 0.81]	0.019 [0.007, 0.033]	0.004*
Density level 1	0.84 [0.78, 0.88]	0.79 [0.68, 0.85]	0.05 [0.01, 0.2]	0.27
Density level 2	0.83 [0.81, 0.85]	0.82 [0.79, 0.84]	0.016 [0.004, 0.028]	0.01*
Density level 3	0.81 [0.78, 0.83]	0.79 [0.77, 0.82]	0.014 [−0.002, 0.03]	0.057
Density level 4	0.78 [0.72, 0.82]	0.77 [0.69, 0.82]	0.02 [−0.032, 0.093]	0.53

\*Statistically significant with the corrected significance levels by Holm method.<sup>22</sup>



**Fig. 5** This figure shows the scatter plots for Dice values for the NS-RGI and RGI segmentation algorithms in terms of breast density. The diameters of circles in the plot are proportional to the maximum lesion diameter measured by the expert. For all density levels, the average Dice values of the NS-RGI segmentation algorithm were higher than those of the RGI cases, while only density level 2 showed a statistically significant difference. For cases with  $Dice_{RGI}$  higher than 0.7,  $Dice_{NS-RGI}$  showed similar performances. For many cases with  $Dice_{RGI}$  less than 0.7,  $Dice_{NS-RGI}$  showed improved segmentation performance. There were two cases (with asterisk marker) that  $Dice_{NS-RGI}$  showed inferior segmentation performance than  $Dice_{RGI}$ .

Among the 23 features we considered, a total of six features were selected for the NS-RGI classifier and a total of eight features were selected for the RGI classifier. All six features for the NS-RGI classifier were selected 99% to 100% in each LOOCV

loop, and the feature appearance frequencies for the RGI classifier were varied from 39% to 100% (Table 4). These features were irregularity,<sup>19,20</sup> ellipsoid axes min-to-max ratio,<sup>19,20</sup> relative margin distance variation,<sup>20</sup> margin volume,<sup>20</sup> radial

**Table 3** Classification performance of trained LDA classifiers using LOOCV for NS-RGI and RGI cases.

Performance comparison (AUC)		Difference in AUC	
NS-RGI	RGI		
AUC <sub>L</sub> [95% CI]	AUC <sub>R</sub> [95% CI]	AUCL–AUCR [95% CI]	<i>p</i> -value
0.80 [0.73, 0.88]	0.69 [0.6, 0.78]	0.11 [0.032, 0.19]	0.006

**Table 4** List of frequently selected features under LOOCV.

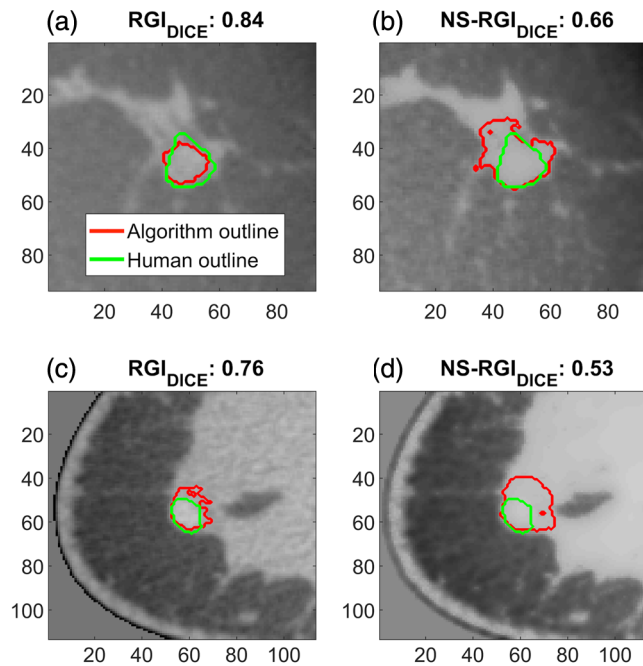
Selected features for NS-RGI	Selection frequency (%)	Selected features for RGI	Selection frequency (%)
Radial gradient index*	100	Ellipsoid axes min to max ratio*	100
Total curvature*	100	Radial gradient index*	100
Ellipsoid axes min to max ratio*	99	Margin volume*	98
Margin volume*	99	Irregularity	97
Gaussian curvature	99	Total curvature*	52
Mean curvature*	98	Mean curvature*	48
		Margin strength 1	44
		Relative margin distance variation	39

\*Features that appeared for both NS-RGI and RGI classifiers.

gradient index,<sup>20,25</sup> margin strength,<sup>20</sup> total curvature,<sup>13</sup> Gaussian curvature,<sup>13,26,27</sup> and mean curvature.<sup>13,26,27</sup> Five features were in common for NS-RGI and RGI cases (Table 4).

## 4 Discussion

Although the segmentation results of NS-RGI showed similar segmentation performances to RGI alone for most cases, there were a couple of cases where NS-RGI showed inferior segmentation performance compared to RGI alone. The cases included one with density level 3 and another with density level 4, where Dice values of the NS-RGI segmentation algorithm were lower than 0.7, and Dice values of the RGI segmentation algorithm were higher than 0.7 (the cases with an asterisk in Fig. 5). For these two cases, we found that the NS enhancement did not stop when it should have, such that the NS algorithm incorrectly included neighboring parenchymal tissue as the foreground and kept applying smoothing operations. The visible lesion boundary was lost as a result, and therefore the NS-RGI algorithm failed to stop at the true lesion boundary (Fig. 6). It is possible that one may observe similar failures for lesions surrounded by complex parenchymal tissues, which is typical for dense breasts (density level 3 or higher). In fact, the above two failure cases were from density level 3 or higher subgroup. As the NS-RGI segmentation algorithm showed at least similar performance compared to the RGI segmentation algorithm for other dense breast cases, we might



**Fig. 6** This figure shows the cases (in coronal view) that NS-RGI showed lower segmentation performances (Dice value lower than 0.7), while RGI showed successful segmentation outcomes (Dice value higher than 0.7). The breast density level of first (a and b) and second (c and d) row lesions are levels 3 and 4, respectively. Left column subimages (a and c) are the bCT cases without NS enhancement and right column subimages (b and d) are those with NS enhancement. As the lesions (highlighted as green outline) are connected with breast parenchymal tissue, the NS enhancement incorrectly includes the neighboring parenchymal tissue as foreground, and therefore it kept applying the enhancement operations to the case. The resulting segmentation outcome (b and d) from NS-RGI showed leaked boundary compared to those of RGI.

conclude that the above two cases are special cases where NS enhancement failed. However, one could reduce these failures by adjusting the maximum number of iterations in the NS enhancement for dense breasts, to terminate the process before it smooths key information for segmentation. It will require a future study to explore the optimal maximum iteration number for dense and fatty breasts.

There are two possible reasons that lesion classification based on NS-RGI segmentation showed better classification performance than the RGI segmentation-based classifier. The first possible reason is the number of strong features used for the classifier. We showed that six and eight features were frequently selected for NS-RGI and RGI classifiers, respectively (Table 4). However, the NS-RGI classifier held only strong features (six features with 98% or higher selection frequency), whereas the RGI classifier included weak features (four weak features with 50% or less selection frequency). The existence of weak features for the RGI classifier is a possible reason for having an inferior classification performance than the NS-RGI classifier. This makes sense as weak features cannot build a classifier that can be generalizable for unseen data. This supports the finding of our previous study,<sup>28</sup> where we showed that a classifier trained with a set of a few strong features can achieve better classification performance than classifiers trained with a set of weak features.

In addition, we note that the selection frequencies of the curvature-related features (e.g., total curvature and mean curvature)



increased from 48% to 50% (for the RGI classifier) to 98% to 100% (for the NS-RGI classifier), respectively. This is another possible reason as to why NS-RGI segmented lesions resulted in better classification performance than RGI segmentation alone. Our previous study<sup>13</sup> showed that curvature features, especially total curvature feature, hold useful information for classifying benign and malignant breast lesions in bCT images. In addition, Kuo et al.<sup>19</sup> showed that increased classification power of morphological features due to improved segmentation of breast lesions can lead to improved performance of the classifier. Note that the curvature features are morphological features, that are specifically related to the surface of the segmented breast lesions. Based on these, we may draw a conclusion that the NS enhancement resulted in changes in segmentation outcomes, and such change improved the morphological representation (especially on the surface) of breast lesions and therefore caused the curvature features to provide more useful information to the NS-RGI-based classifier than the RGI-based classifier. This improvement in curvature features might remove subtle features from consideration during training of the NS-RGI-based classifier, such that the classifier retained only strong features.

There are some limitations to our study. We used the same dataset to select operating variables such as  $\alpha$ ,  $\beta$ , and  $\gamma$  for the NS enhancement, feature selection for the classifier, training, and testing the classifier. Although we used the LOOCV to evaluate the performance of the NS-RGI and RGI classifiers, the fact that we selected operating variables for the enhancement from the same dataset might bias our results. An independent dataset is required to check if the operating variables selected for this study are the global optimum or just a local optimum. In fact, there exists a noise simulator<sup>29,30</sup> that can be used to thoroughly study how the NS enhancement would handle various levels of noise, as well as to search the optimal operating variables for bCT images. It is therefore worthwhile to conduct a future study to find the optimal operating variables for the NS enhancement with an independent dataset and the above noise simulator.

We used one reconstruction algorithm, i.e., FDK reconstruction, which can be an additional limitation of this study. There are other image reconstruction algorithms, specifically, iterative image reconstruction algorithms, available for breast CT.<sup>1</sup> It is possible that state-of-the-art iterative image reconstruction algorithms can successfully reduce the noise in the image such that the proposed NS enhancement is less effective. We or others need to conduct further research to determine how the proposed NS enhancement performs on images reconstructed with different algorithms.

Another limitation is that we tested only one segmentation algorithm, RGI segmentation, as an example to show the effectiveness of the NS enhancement. As there exists other algorithms for bCT images (e.g., extended versions of RGI segmentation using active contour by Kuo et al.<sup>31</sup>), one might get different results using those other algorithms. In addition, deep convolutional neural network (CNN) is becoming state-of-the-art for many image analysis tasks, including segmentation. Although most previous publications on deep CNN for segmentation is on 2-D images,<sup>32,33</sup> 3-D extensions exist.<sup>34</sup> Testing the proposed NS enhancement on other segmentation algorithms, such as breast CT-specific algorithms and more generic CNN algorithms, should be done in the future. Since the NS enhancement works as a preprocessing step to

clean up the image, however, we expect that our proposed enhancement method may improve the performance of other segmentation algorithms.

In conclusion, we introduced an NS enhancement method as a preprocessing step to denoise or enhance bCT images in order to improve the performance of computer segmentation algorithms. We showed that the proposed method could improve the computer segmentation performance, as well as the computer classification performance, when trained on features extracted from segmented breast lesions.

## Disclosures

This study has been supported in part by grants from the National Institutes of Health R21-EB015053 and R01-CA181081. Drs. Lee, Nishikawa, and Reiser have nothing to declare. Dr. Boone has a research contract with Siemens Medical Systems and receives royalties from Lippincott Williams and Wilkins (book).

## References

1. N. Antropova et al., "Efficient iterative image reconstruction algorithm for dedicated breast CT," *Proc. SPIE* **9783**, 97834K (2016).
2. E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.* **53**(17), 4777–4807 (2008).
3. F. Smarandache and A. Unifying, *Field in Logics: Neutrosophic Logic. Neutrosophy, Neutrosophic Set, Neutrosophic Probability*, American Research Press, Rehoboth, New Mexico (2003).
4. M. Zhang, L. Zhang, and H. D. Cheng, "A neutrosophic approach to image segmentation based on watershed method," *Signal Process.* **90**(5), 1510–1517 (2010).
5. Y. Guo and H. D. Cheng, "New neutrosophic approach to image segmentation," *Pattern Recogn.* **42**(5), 587–595 (2009).
6. M. Xian, H. D. Cheng, and Y. Zhang, "A fully automatic breast ultrasound image segmentation approach based on neutro-connectedness," in *22nd Int. Conf. on Pattern Recognition (ICPR)*, pp. 2495–2500 (2014).
7. K. K. Lindfors et al., "Dedicated breast computed tomography: the optimal cross-sectional imaging solution?" *Radiol. Clin. North Am.* **48**(5), 1043–1054 (2010).
8. L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Am. A* **1**(6), 612–619 (1984).
9. H. B. Robert et al., "Lung density changes with growth and inflation," *Chest* **148**(4), 995–1002 (2015).
10. M. N. Naik et al., "Interpretation of computed tomography imaging of the eye and orbit. A systematic approach," *Indian J. Ophthalmol.* **50**(4), 339 (2002).
11. J. M. Boone et al., "Dedicated breast CT: radiation dose and image quality evaluation," *Radiology* **221**(3), 657–667 (2001).
12. N. D. Prionas et al., "Contrast-enhanced dedicated breast CT: initial clinical experience," *Radiology* **256**(3), 714–723 (2010).
13. J. Lee et al., "Local curvature analysis for classifying breast tumors: preliminary analysis in dedicated breast CT," *Med. Phys.* **42**(9), 5479–5489 (2015).
14. R. C. Gonzalez, *Digital Image Processing Using MATLAB*, Dorling Kindersley Pvt. Ltd., New Delhi (2006).
15. I. Reiser et al., "Evaluation of a 3D lesion segmentation algorithm on DBT and breast CT images," *Proc. SPIE* **7624**, 76242N (2010).
16. L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology* **26**(3), 297–302 (1945).
17. C. T. Loop, "Smooth subdivision surfaces based on triangles," Master's Thesis, Department of Mathematics, University of Utah (1987).
18. E. B. Elkin et al., "The effect of changes in tumor size on breast carcinoma survival in the U.S.: 1975–1999," *Cancer* **104**(6), 1149–1157 (2005).
19. H.-C. Kuo et al., "Impact of lesion segmentation metrics on computer-aided diagnosis/detection in breast computed tomography," *J. Med. Imaging* **1**(3), 031012 (2014).

20. I. Reiser et al., "Automated detection of mass lesions in dedicated breast CT: a preliminary study," *Med. Phys.* **39**(2), 866–873 (2012).
21. E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach," *Biometrics* **44**(3), 837–845 (1988).
22. S. Holm, "A simple sequentially rejective multiple test procedure," *Scand. J. Stat.* **6**(2), 65–70 (1979).
23. E. A. Sickles, C. J. D'Orsi, and L. W. Bassett, "ACR BI-RADS® Mammography," in *ACR BI-RADS® Atlas*, C. J. D'Orsi et al., Eds., American College of Radiology, Reston, Virginia (2013).
24. A. P. Zijdenbos et al., "Morphometric analysis of white matter lesions in MR images: method and validation," *IEEE Trans. Med. Imaging* **13**(4), 716–724 (1994).
25. I. Reiser et al., "Computerized detection of mass lesions in digital breast tomosynthesis images using two- and three-dimensional radial gradient index segmentation," *Technol. Cancer Res. Treat.* **3**(5), 437–441 (2004).
26. E. Luders et al., "A curvature-based approach to estimate local gyrification on the cortical surface," *Neuroimage* **29**(4), 1224–1230 (2006).
27. D. M. Cash et al., "Cortical folding analysis on patients with Alzheimer's disease and mild cognitive impairment," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, N. Ayache et al., Eds., pp. 289–296, Springer, Berlin, Heidelberg (2012).
28. J. Lee et al., "Optimal reconstruction and quantitative image features for computer-aided diagnosis tools for breast CT," *Med. Phys.* **44**(5), 1846–1856 (2017).
29. P. Massoumzadeh et al., "Validation of CT dose-reduction simulation," *Med. Phys.* **36**(1), 174–189 (2009).
30. B. R. Whiting et al., "Measurement of bow tie profiles in CT scanners using a real-time dosimeter," *Med. Phys.* **41**(10), 101915 (2014).
31. H.-C. Kuo et al., "Segmentation of breast masses on dedicated breast computed tomography and three-dimensional breast ultrasound images," *J. Med. Imaging* **1**(1), 014501 (2014).
32. J. Lee and R. M. Nishikawa, "Automated mammographic breast density estimation using a fully convolutional network," *Med. Phys.* **2018**, in press.
33. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, N. Navab et al., Eds., pp. 234–241, Springer, Cham (2015).
34. Ö. Çiçek et al., "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016*, S. Ourselin et al., Eds., pp. 424–432, Springer International Publishing (2016).

**Juhun Lee** is a research instructor in the Imaging Research Laboratory in the Department of Radiology at the University of Pittsburgh. He received his PhD in electrical and computer engineering in 2014 from the University of Texas at Austin. His research interests include algorithm developments for computer-aided diagnosis and breast lesion segmentation for bCT and mammography.

**Robert M. Nishikawa** received his PhD in medical biophysics from the University of Toronto in 1990. He is currently a professor and director of the Imaging Research in the Department of Radiology at the University of Pittsburgh. He is a fellow of the AAPM, SBI, and AIMBE. He has over 200 publications in breast imaging concentrating on computer-aided diagnosis, technology assessment, and quantitative imaging.

**Ingrid Reiser** is an assistant professor of radiology at the University of Chicago in Chicago, Illinois, USA. She holds a PhD in physics from Kansas State University. Her research interests include computer-aided detection and diagnosis methods for breast cancer in dedicated breast CT and digital breast tomosynthesis, as well as objective assessment of x-ray tomographic x-ray breast imaging systems.

**John M. Boone** is a professor and vice chair (research) of radiology, and professor of biomedical engineering at the University of California, Davis, California, USA. After receiving his BA degree in biophysics from UC Berkeley, he received his PhD in radiological sciences from UC Irvine. He has research interests in breast imaging, CT, and radiation dosimetry. He is the PI of the breast tomography project, where over 600 women have been imaged on breast CT scanners fabricated in his laboratory.