Medical breast ultrasound image segmentation by machine learning

Yuan Xu[a], Yuxin Wang[a], Jie Yuan[a,*], Qian Cheng[b], Xueding Wang[b,c], Paul L. Carson[c]

[a] Department of Electronic Science and Engineering, Nanjing University, Nanjing 210093, China
[b] Department of Physics, Tongji University, Shanghai 200000, China
[c] Department of Radiology, University of Michigan, Ann Arbor, MI 48109, USA

ARTICLE INFO

ABSTRACT

Breast cancer is the most commonly diagnosed cancer, which alone accounts for 30% all new cancer diagnoses for women, posing a threat to women's health. Segmentation of breast ultrasound images into functional tissues can aid tumor localization, breast density measurement, and assessment of treatment response, which is important to the clinical diagnosis of breast cancer. However, manually segmenting the ultrasound images, which is skill and experience dependent, would lead to a subjective diagnosis; in addition, it is time-consuming for radiologists to review hundreds of clinical images. Therefore, automatic segmentation of breast ultrasound images into functional tissues has received attention in recent years, amidst the more numerous studies of detection and segmentation of masses. In this paper, we propose to use convolutional neural networks (CNNs) for segmenting breast ultrasound images into four major tissues: skin, fibroglandular tissue, mass, and fatty tissue, on three-dimensional (3D) breast ultrasound images. Quantitative metrics for evaluation of segmentation results including Accuracy, Precision, Recall, and $F1_{measure}$, all reached over 80%, which indicates that the method proposed has the capacity to distinguish functional tissues in breast ultrasound images. Another metric called the Jaccard similarity index (JSI) yields an 85.1% value, outperforming our previous study using the watershed algorithm with 74.54% JSI value. Thus, our proposed method might have the potential to provide the segmentations necessary to assist the clinical diagnosis of breast cancer and improve imaging in other modes in medical ultrasound.

1. Introduction

The purpose of this work was to segment the breast ultrasound images into all major functional tissues automatically. Breast cancer is one of the most common cancers accounting for 30% of cancer diagnoses in women, with increasing incidence in recent years [1]. It has become the leading cause of cancer mortality among women younger than 45 years old [2]. Ultrasound imaging is a widely employed imaging method to characterize breast masses [3] due to its avoidance of ionizing radiation, real-time visualization, relatively inexpensive price and non-invasive diagnosis [4]. Ultrasound imaging is also a tool for breast density assessment [5]. Breast density is often taken as a predictor of breast cancer risk assessment and prevention [6]. The percentage of breast density is calculated by dividing the area of the fibroglandular tissue by the total area of the breast. The odds ratio of developing breast cancer for women with most dense breasts is 2 to 6-fold greater than women with normal breast density [7].

Segmentation of breast ultrasound images could help radiologists discriminate different functional tissues, which provides valuable references on image interpretation, tumor localization, and clinical diagnosis of breast cancer. For breast density assessment, quantitative segmentation should be more consistent and precise than the qualitative approach where the analysis is limited by the reader's experience and training [5]. Moreover, different tissue properties have effects on the propagation of ultrasound waves. Segmentation of the bulk of the breast tissues could also facilitate correction of aberrations in those tissues and provide *a priori* information to improve ultrasound imaging in modes other than pulse echo, for example, transmission tomography and limited angle transmission tomography. With the source pulse modeled as a sinusoidal waveform of 0.5 MHz center frequency, Rungroj et al [8] enhanced limited angle tomography to produce speed of sound images by utilizing the image segmentations, and Hooi et al [9] applied a regularized least squares algorithm with data acquired at a center frequency of 3.75 MHz to construct the attenuation image by allowing inclusions of *a priori* information on major structures in the breast, such as the size and shape.

However, manual semantic analysis of breast ultrasound images is dependent on the theoretical basis and clinical experience of
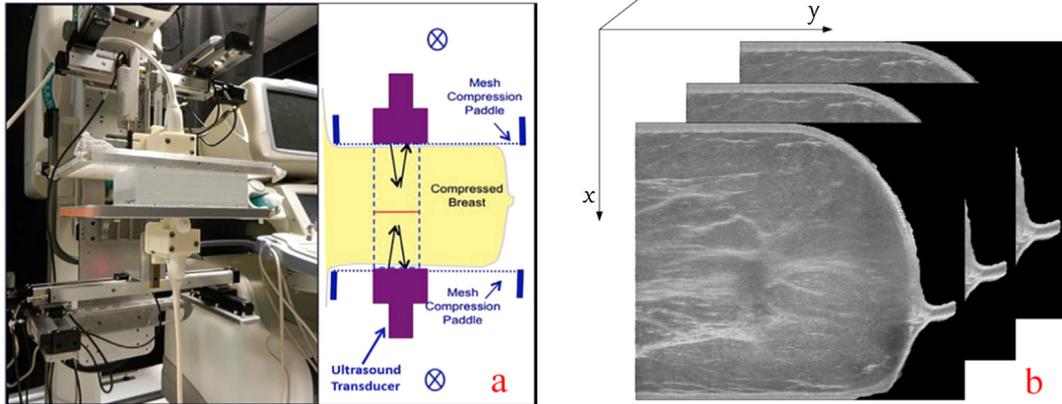
Fig. 1. (a) The BLUCI system; (b) 2D breast ultrasound image sequence.

radiologists, which results in subjective interpretation and inter-observer variability. Such manual analysis is also labor intensive and time-consuming for large scale clinical ultrasound images. To eliminate the operator dependency and reduce the burden on radiologists, the computer-aided detection (CAD) system, one of the major research subjects in medical imaging [10], could be utilized for automatic breast image segmentation [11]. Full breast segmentation is in contradistinction to the much more frequently studied detection and segmentation of breast masses. Among the studies focusing on automated segmentation of breast images in pulse echo ultrasound imaging, various segmentation algorithms have been proposed. These include threshold segmentation [12–14], the active contour model (ACM) [15–18], clustering algorithms [11,19–21], and some other methods [22–25]. Huang et al [16] proposed a method which used a watershed algorithm to obtain the initial active contour model first, and then minimized the energy function to accurately define the tumor boundary. Kekre et al [19] proposed another method for breast image segmentation, which used Vector Quantization (VQ) based clustering technique with Linde Buzo Gray (LBG), Kekre's Proportionate Error (KPE), and Kekre's Error Vector Rotation (KEVR) codebook generation algorithm followed by sequential cluster merging. Gu et al [24] used the Sobel operators on the morphological reconstructed images for edge information extraction to obtain gradient magnitude images and then used the watershed algorithm to demarcate all major tissues in 3D breast ultrasound volumes. However, ultrasound imaging has various inherent artifacts and noises [26,27], such as speckle noise [28], attenuation (absorption and scattering), blurry boundaries [29], and intensity inhomogeneity [27]. All of these artifacts increase the difficulty of the semantic analysis of breast ultrasound images. Furthermore, extraction of effective features for image interpretation is a challenging task, as well as the selection of parameters (e.g. thresholds). Thus, it is difficult to obtain promising and fully automatic segmentation results using existing traditional methods, where thresholds, features, and region of interest (ROI) are often hand-crafted and experience-determined.

Deep learning has emerged as a powerful alternative for supervised image segmentation in recent years [30]. Convolutional Neural Networks (CNNs) in the deep learning field have the ability to capture nonlinear mappings between inputs and outputs and learn discriminative features for the segmentation task without manual intervention. These features generated automatically through CNNs often outperform hand-crafted and pre-defined feature sets. In particular, CNN algorithms have been used to handle a variety of biomedical imaging problems, including medical image segmentation [30–32]. These studies have gained encouraging results, which indicates the potential of CNNs on medical breast image segmentation.

Conventional 2-dimensional (2D) segmentation may be affected by the limitations of 2D ultrasound imaging, such as suboptimal projection angle. 3D breast images can provide stereoscopic information and more image features, which is conducive to accurately distinguish different tissues. However, to classify a target pixel, 3D CNNs need to process a cubic input; this requires an increased number of network parameters, storage memories, and computation iterations compared with 2D CNNs. In our study, to incorporate 3D contextual information and reduce the computational burden of processing 3D ultrasound images, we propose a CNN-based method for breast ultrasound image segmentation on 3 orthogonal image planes. Specifically, the CNNs take image blocks centered at a pixel as inputs and produced the tissue class of the center pixel as the output. Taking the images segmented by clinicians as the ground truth, we trained the CNN models to distinguish skin, fibroglandular tissues, and masses. The segmentation process with the trained CNNs is completely automated without any human intervention. Visualization results showed that the proposed CNNs can demarcate all the major functional tissues consistent with the ground truth. When evaluated by the quantitative assessment metrics, we can yield satisfactory results of more than 80%, which indicate the value of our method in assisting breast cancer interpretation and diagnosis.

The paper is organized as follows: in Section 2, the segmentation material and the method proposed are illustrated; in Section 3, the experimental results are presented; then, discussion and conclusion are given in Section 4.

2. Material and methods

2.1. Data acquisition

The 3D breast ultrasound images in this work were acquired by a dual-sided automated breast ultrasound imaging device at the Department of Radiology, University of Michigan, USA — the Breast Light and Ultrasound Combined Imaging (BLUCI) System. The BLUCI system allows for the acquisition of 3D ultrasound and photoacoustic volumetric imaging with the method analogous to classic mammography. The structure of the BLUCI system is shown in Fig. 1(a). A breast is fixed between two mesh compression paddles, two GE M12L transducers (GE Health Systems, Milwaukee, WI) with 12 MHz center frequency positioned above and below the paddles sweep across the breast with the scan parameters given by a control computer. Previously, solid paddles were used and the breast motion due to lubrication of the skin and paddle surface by the ultrasound coupling gel was counteracted by using an adhesive hair spray for acoustic coupling. With the mesh paddles, the 8 mm high rim compresses the retro-mammary fat trapping the rest of the breast that distends the mesh a few millimeters. There is a 2 mm displacement between the surface of the breast and the transducer when the transducer passes over the compressing membrane. As this is happening on each successive image

plane, there is negligible differential motion between one image and the next. When the image volume is resliced in either the axial/elevational or lateral/elevational plane (i.e., with one axis in the direction of the scan), the images show none of the visible motion that you would expect if there were discrete tissue shifts between slices, except in rare cases where the patient moves the breast substantially during the scan.

The system generally acquires 250 slices over 120 s with a resolution of 0.1153 mm × 0.1153 mm × 0.8 mm. The slices are stitched together in sequence to form a complete 3D breast image, as shown in Fig. 1(b). Our data set acquired by the BLUCI system contains 21 cases including asymptomatic volunteers and patients with masses prior to biopsy. The proposed segmentation scheme utilizes the acquired 21 cases and is not concerned with the frequency of image acquisition. All the 21 human imaging exams involved in this study were performed with the subjects' consent under an Institutional Review Board (IRB) approved protocol.

2.2. Data preparation

To generate the ground truth for training, manual segmentation and labeling were carefully operated by experienced radiologists. Each case used in our study typically has 250 slices and we selected slices which comprised all integral tissues for handcraft annotation. The labels are annotated on the raw images including skin, glands, and masses, each of which was marked with a different grayscale value in a new layer corresponding to the raw image as shown in Fig. 2. With the manual annotation, we generated more than 20,000 images blocks centered at target pixels for each slice. These image blocks were divided into the training set and testing set in our study. Data training was carried out on the training set and the testing set was used to evaluate the segmentation performance. Pixels in these image blocks are each labeled according to their true classes as described above for network training.

2.3. Network design

2.3.1. CNN-I for Pixel-centric patch classification

In our study, we designed an 8-layer CNN to perform pixel labeling in three orthogonal ultrasound image planes including convolution layers 1–3, pooling layers 1–3, a fully connected (FC) layer, and a softmax layer. The CNN architecture and detailed configuration parameters are shown in Fig. 3 and Table 1.

The inputs of the CNN are 128 × 128 normalized image blocks. The size of 128 × 128 was determined based on experimentation. In preliminary experiments, we conducted the segmentation task on 2D slices with 5 different network input sizes. Results are shown in Fig. 4 and Table 2. The evaluation metrics used in Table 2 (e.g., Accuracy) are defined in Section 3.2. As we can see in these results, CNN with small

input size made more misjudgments on the skin and glandular tissues which resulted in a poor visualization. When using a larger input size, the segmentation result was improved and steady. Therefore, we chose the input image size of 128 × 128 in our studies.

After the inputs had been forward propagated throughout the network, the 4-way softmax layer utilized Softmax function to finally generate a probability distribution array with 4 output elements. Softmax function transforms the original input to a K-length distribution vector of values in the range (0, 1) that add up to 1. Given a sample vector x and a weighting vector w, the predicted probability for the j'th class can be calculated as the Eq. (1):

$$P(y = j|x) = \frac{e^{x^{T}w_j}}{\sum_{k=1}^{K} e^{x^{T}w_k}}. \tag{1}$$

The four elements of the output array corresponded to the four tissues in the breast ultrasound images: fibroglandular tissue, mass, skin, and subcutaneous fat respectively. Values of the 4 outputs contributed a distribution indicating the probabilities over the four class labels. The index of the maximum element determined the category of the central pixel in the image block. For example, [0.80, 0.10, 0.10, 0.00] denotes the pixel belongs to glandular tissue while the array of [0.15, 0.70, 0.05, 0.10] indicates a mass pixel. The size of each kernel and the number of layers are experimentally determined in the pre-experiments. To capture nonlinear mappings between inputs and outputs, all the convolution layers and the fully-connected layer are activated by Rectified Linear Units (ReLU) function.

2.3.2. CNN-II for comprehensive evaluation

To ascertain the final classification of the target pixel from the aforementioned three CNN output arrays, a second small classification CNN was applied, where the input 12 neurons (generated by 3 CNN-Is) connected to the 4 output neurons through a convolution layer, a fully connected layer, and a softmax layer. Detailed configuration parameters are shown in Fig. 5. With this small network, we could finally obtain the predicted class of the target pixel. As the CNN-Is, ReLU activation function is applied after the convolution layer.

2.3.3. Cnns training

CNN is a model composed of several layers of nonlinearities, where each layer contains a large number of trainable parameters. As mentioned in Section 2.2, we trained the networks using training set consisting of image blocks of size 128 × 128 extracted from three orthogonal image planes and the corresponding manually labeled ground truth of the center pixel. The training of CNNs is similar to the traditional multilayer feed-forward neural network, where the weight parameters and bias parameters are initialized randomly. In the training process, image blocks are propagated forward through the
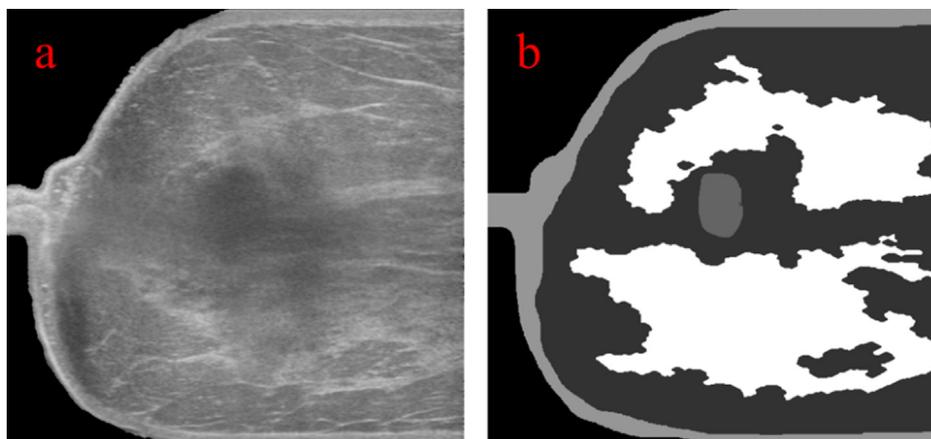

Fig. 2. (a) The raw breast ultrasound image; (b) The annotated image where different tissues are marked with different grayscale values.
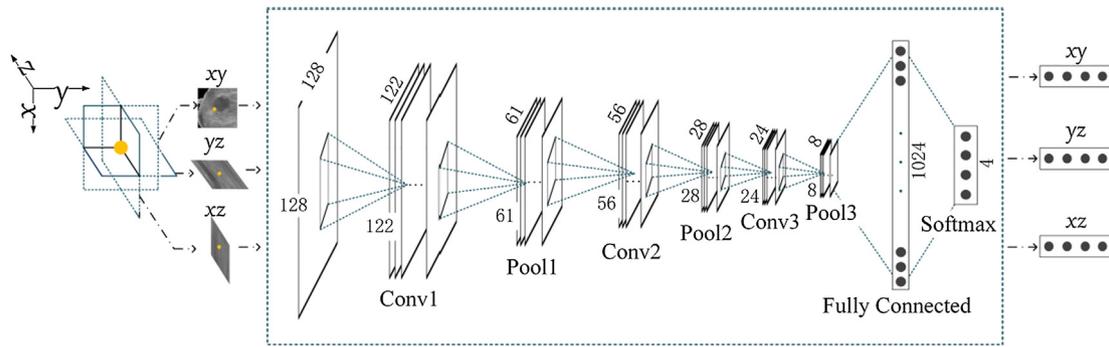
Fig. 3. CNN-I architecture.

Table 1
Configuration parameters of CNN-I.

Layers	Input size	Kernel size	Stride length	Output size
Conv 1	$128 \times 128 \times 1$	$7 \times 7 \times 36$	1	$122 \times 122 \times 36$
Pool 1	$122 \times 122 \times 36$	$2 \times 2 \times 36$	2	$61 \times 61 \times 36$
Conv 2	$61 \times 61 \times 38$	$5 \times 5 \times 36$	1	$56 \times 56 \times 36$
Pool 2	$56 \times 56 \times 36$	$2 \times 2 \times 64$	2	$28 \times 28 \times 64$
Conv 3	$28 \times 28 \times 64$	$3 \times 3 \times 64$	1	$24 \times 24 \times 64$
Pool 3	$24 \times 24 \times 64$	$3 \times 3 \times 64$	3	$8 \times 8 \times 64$
FC	4096×1	–	–	1024×1
Softmax	1024×1	–	–	4×1

Table 2
Quantitative assessment of fibroglandular segmentation with different input sizes.

Input size	Accuracy	Precision	Recall	F1-measure	MHD[1]
48×48	0.8451	0.7225	0.7899	0.7547	159.154
64×64	0.8208	0.6535	0.8643	0.7443	166.433
96×96	0.8956	0.7837	0.9031	0.8392	143.450
108×108	0.8984	0.7943	0.8949	0.8416	59.666
128×128	0.9013	0.8038	0.8898	0.8446	59.034

[1] MHD: Modified Hausdorff Distance.

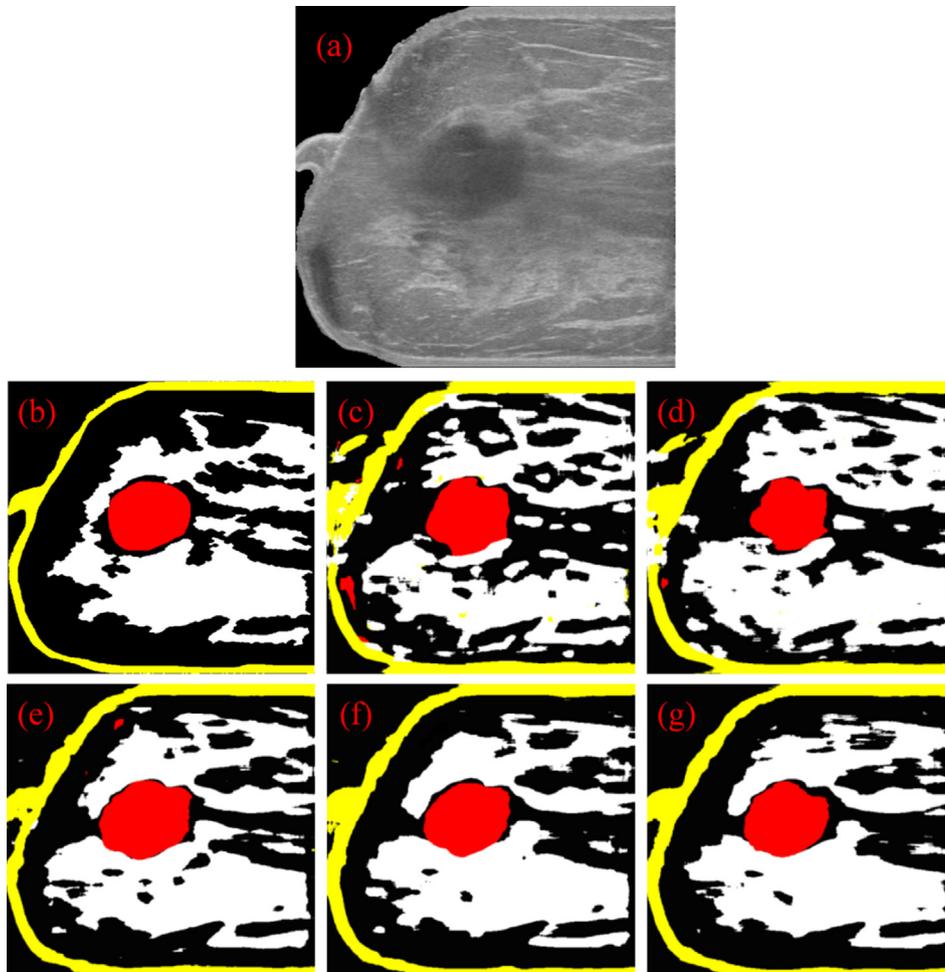**Fig. 4.** (a) Is the raw image; (b) is the manual segmentation result; (c–g) are automatic segmentation results with the input size of 48×48, 64×64, 96×96, 108×108, 128×128 respectively.
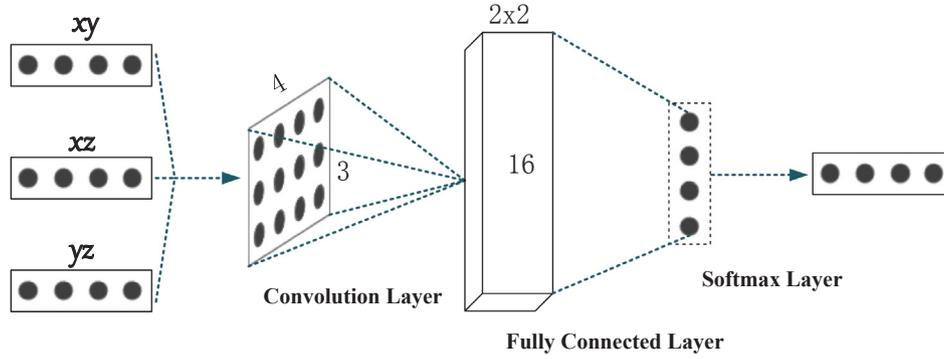
Fig. 5. CNN-II architecture.

network to obtain the outputs. Then the outputs are compared with the corresponding ground truth to calculate the errors using the loss function. At the stage of error back propagation, parameters of the network are gradually updated with a gradient descent algorithm. In this work, both CNN-I and CNN-II took the cross entropy as the loss function, which can be calculated by the following equation:

$$H_{y'}(y) = -\sum_i y'_i \log(y_i),$$ (2)

where y is the predicted probability distribution, and y' is the actual distribution. Cross entropy of every 100 samples was used for error back propagation to avoid the randomness of using one sample. The weight and bias parameters were gradually updated with the gradient descent algorithm named adaptive moment estimation (Adam) optimization algorithm. Dropout methods were applied in training to prevent over-fit. As the input size discussed in Section 2.3.1, the selections of the training hyper-parameters, such as learning rate, were determined experimentally.

3. Experiments and analysis

3.1. Experiment setup

In our experiments, breast ultrasound image segmentation was realized by a pixel classification CNN, where the input was an image block of size 128×128, and the output is the class of the block's central pixel including fibroglandular tissue, tumor, skin, and subcutaneous fat. Since we have 2D image slices stitched together to form 3D breast images, blocks were extracted from three orthogonal planes for each pixel. By minimizing the loss function with the Adam optimizer, the CNN parameters were trained using corresponding manually segmented ground truth. When the CNNs are trained to convergence, it can be tested to segment the breast ultrasound images. As shown in Fig. 6, three image blocks around the pixel are calculated throughout the CNN-Is separately, where the parameters were trained with data from corresponding planes. Then the generated distributions are fed as the input of CNN-II for comprehensive analysis to obtain the final 4-way output. The index of the maximum value in the output indicates the predicted

classification of the target pixel. The segmentation results are visualized by coloring pixels from different classes.

3.2. Quality assessment

3.2.1. Region-based assessment method

In order to measure the image segmentation performance quantitatively, four region-based assessment metrics: Accuracy, Precision, Recall, and $F1_{measure}$ are used in our study. They can be calculated based on the region overlap between manually and automatically segmented results. Taking the tumor segmentation as an example and as explained in Table 3, TP, FP, FN and TN are the numbers of pixels corresponding to the four categories listed in the table below.

According to Table 3, Accuracy, Precision, Recall, and $F1_{measure}$ are defined as:

$$Accuracy = \frac{the\,number\,of\,correct\,predicted\,pixels}{the\,total\,number\,of\,pixels\,in\,the\,image} = \frac{TP + TN}{TP + FP + FN + TN},$$ (3)

$$Precision = \frac{the\,number\,of\,correct\,predicted\,tumor\,pixels}{the\,total\,number\,of\,predicted\,tumor\,pixels} = \frac{TP}{TP + FP},$$ (4)

$$Recall = \frac{the\,number\,of\,correct\,predicted\,tumor\,pixels}{the\,total\,number\,of\,actual\,tumor\,pixels} = \frac{TP}{TP + FN},$$ (5)

$$F1_{measure} = 2 \times \frac{Pricision \times Recall}{Pricision + Recall}.$$ (6)

These evaluation metrics comprehensively evaluate the segmentation performance from different aspects. All of them are values between 0 and 1. In addition, we use Jaccard similarity index (JSI) mentioned in the work by Gu et al [24] to measure the general accuracy of the mass segmentation:

$$JSI = \frac{Segmentation_{Manual} \cap Segementation_{Automated}}{Segmentation_{Manual} \cup Segementation_{Automated}} = \frac{TP}{TP + FP + FN}.$$ (7)

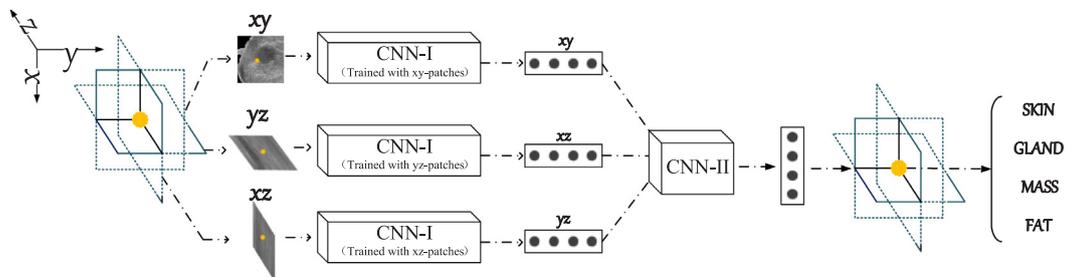**Fig. 6.** Segmenting process.

Table 3
Definition of the abbreviations.

Category	Actual tumor	Actual non-tumor
Predicted tumor	True Positive (TP)	False Positive (FP)
Predicted no tumor	False Negative (FN)	True Negative (TN)

3.2.2. Distance-based assessment method

The evaluation metrics above are based on the region overlap between the automated segmentation and the ground truth. We also adopted another distance-based evaluation metric called the Modified Hausdorff Distance (MHD) used in [33] as a supplementary measure for assessment. Given two sets of pixels (A and B) for a certain tissue, the MHD is defined as:

$$H(S, G) = max\{\max_{a \in A}\{\min_{b \in B}\{d(a, b)\}\}, \max_{b \in B}\{\min_{a \in A}\{d(a, b)\}\}\},\qquad(8)$$

where $d(a, b)$ represents the Euclidean distance between pixels a and b. A smaller value of MHD generally indicates a higher proximity of the two-point sets, thus implying better segmentation performance [30].

3.3. Experimental results

In the experiments, breast ultrasound images were divided into training set and test set. We trained the networks using the training set and the corresponding ground truth according to the experimental methods introduced in Section 3.1. Test sets that had not been trained before were automatically segmented by the trained CNNs to evaluate the segmentation performance. We used over 5 million ultrasound image pixels to evaluate the performance of our proposed method.

Automated segmentation results are compared to the ground truth segmentations by radiologists for evaluation. In order to observe the visual similarity of the shape, the contours of the fibroglandular tissue in the manual and automatic segmentations are extracted and marked with different colors. The results are shown in Fig. 7, where the red lines in Fig. 7(a) and (b) represent the contours from the manually segmented images, and the yellow lines represent the contours from the automatically segmented results. As we can see, in terms of glandular contours, the results of automatic segmentation show good consistency with the results of manual segmentation. In Fig. 7, pixels belonging to fat are drawn as a background colored in black for the fatty tissues wrap around the fibroglandular tissue and fill the entire breast. From the
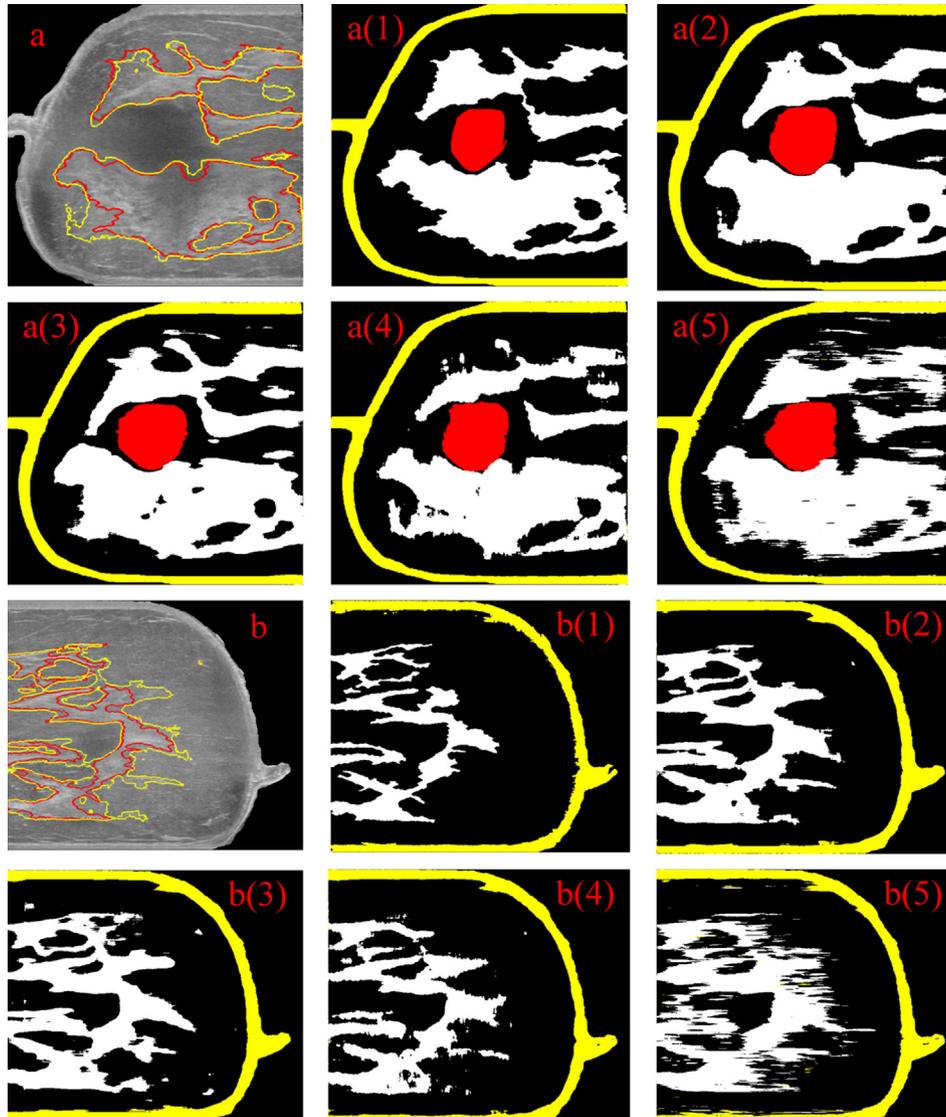

Fig. 7. Automated segmentation results compared with manual segmentation results; (a) and (b) are raw images with gland contours marked; (a1) and (b1) are manual segmentations; (a2) and (b2) are the final outputs of the classification network; (a3) and (b3) are segmentation results with CNN-I trained with XY-patches; (a4) and (b4) are segmentation results with CNN-I trained with XZ-patches; (a5) and (b5) are segmentation results with CNN-I trained with YZ-patches.
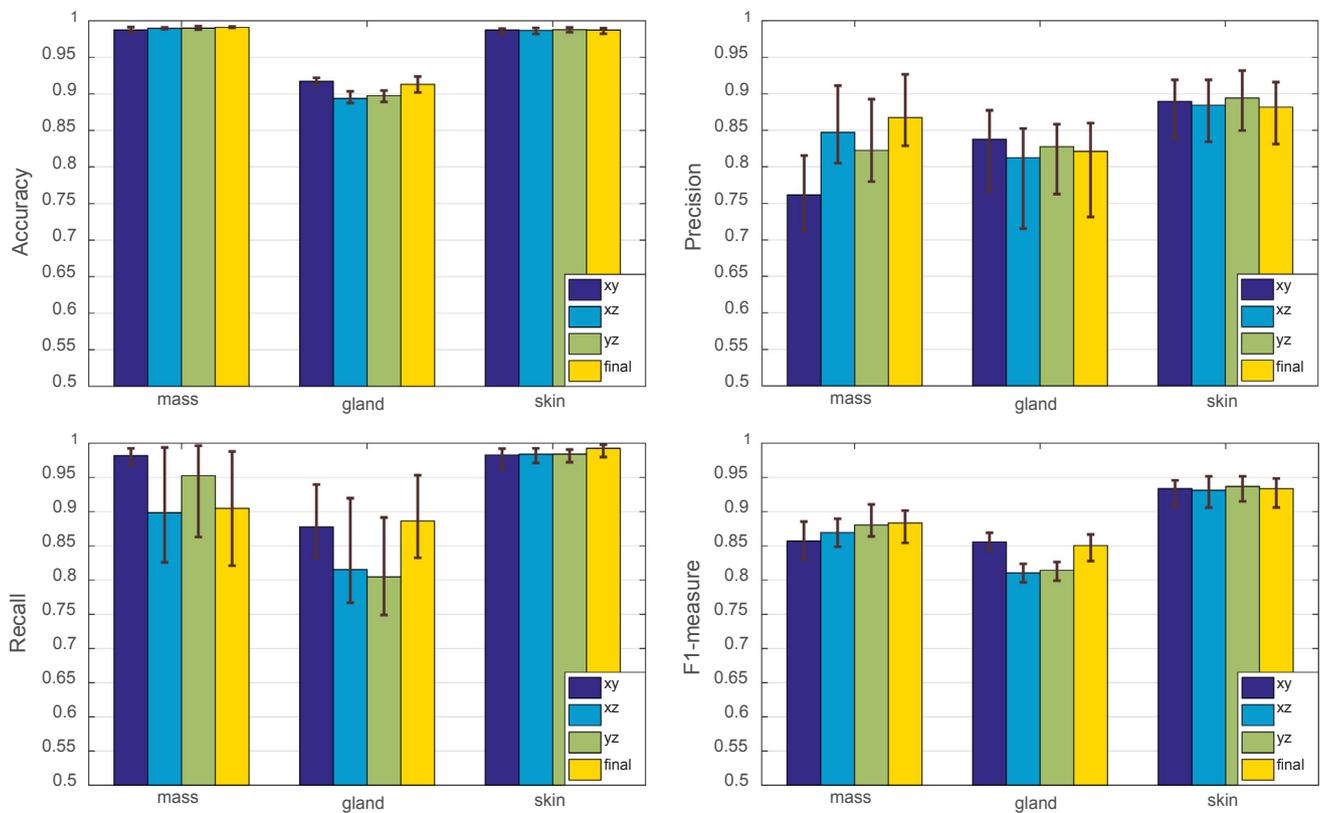
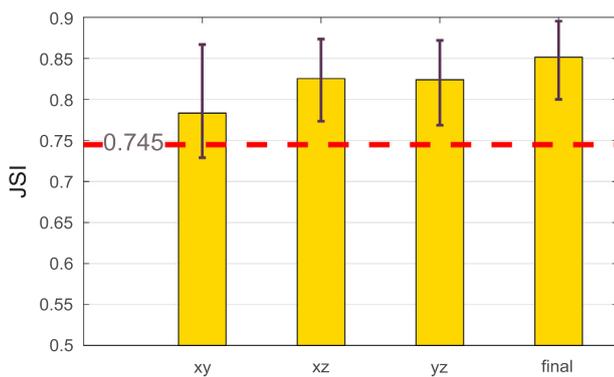**Fig. 8.** Quantitative assessment of automated segmentation results.

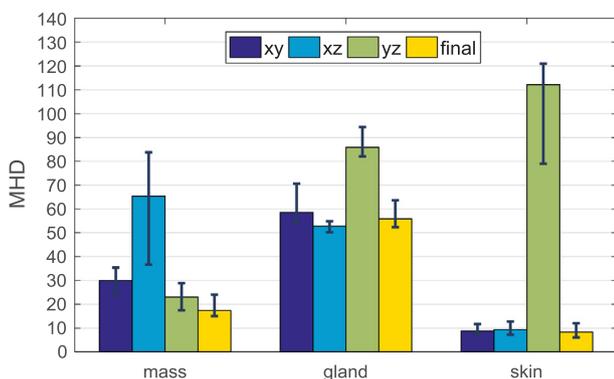**Fig. 9.** Jaccard similarity index of mass segmentation.

Fig. 10. MHD of the automated segmentation results.

results we have obtained, the comprehensive results of CNN-II are in good agreement visually with the ground truth. The intermediate

outputs of the three CNN-Is are also similar to the hand-crafted ground truth. However, as shown in Fig. 7, the final outputs of CNN-II are better visualized than the intermediate outputs by correcting some errors in the results segmented from a single plane, which illustrates the advantages of the 3D analysis in tissue segmentation.

Furthermore, the results of our automated segmentation method were quantitatively evaluated by using the evaluation metrics including Accuracy, Precision, Recall and $F1_{measure}$, which are described in Section 3.2. The results are reported in Fig. 8 using histograms with error bars. We calculated the four metrics using the intermediate and final results. The columns in Fig. 8 show the average performance of the automatically segmented slices and the error bars indicate the stability.

In terms of accuracy, the calculated values have almost reached about 90% with small numerical fluctuations, which verify the validity of the network for segmentation in general. For the other indexes, it is shown that the calculation results of the CNN-II all reached over 80%, indicating that the proposed method is feasible and effective. Comparing the outputs of CNN-II with those of the three trained CNN-Is, although the intermediate results may perform well on some evaluation metrics, the final results are more balanced in all aspects for the synthesis of CNN-II. Since the glandular tissues are more complex compared to other tissues such as skin, which are characterized by complicated shape, irregular boundary, and non-obvious pixel intensity, the performance of gland segmentation is not as good as in the mass and skin segmentation, as shown in Fig. 8. When classifying pixels on the boundary of glandular tissues, the CNNs are prone to misjudgment due to the similarity of image blocks of adjacent pixels. In addition, glandular tissues are distributed throughout the breast images and contain the largest number of pixels except for fatty tissues. Therefore, the number of erroneous judgments is greater, even with the same fraction of errors, leading to the worst performance among all evaluation metrics.

Results of Jaccard similarity index (JSI) for mass segmentation are shown in Fig. 9. In the study by Gu et al [24], the JSI results were

calculated on eight images, taken approximately every 10 slices in each image stack. In our method, the JSI results were calculated on two images that had not been trained before, each of which took 10 slices. Compared to the method of Gu, the average JSI values of the three CNN-I results and the final CNN-II results are all better than the JSI value in that study, which was 74.5%. The final segmentation reached an average similarity value of 85.1%, which is increased by about 10% and is capable to provide objective references for radiologists on breast image interpretation. Moreover, the skin and fibroglandular tissue were taken as the same tissue in the segmentation method proposed in Gu's work. The two tissues, however, can be distinguished separately with our method for they are annotated as two classes for CNNs training.

In addition to quantitatively demonstrating the four region-based metrics of the proposed CNN method, we also use another distance-based measure, the MHD, described in Section 3.2 to evaluate the performance in different tissues of the proposed CNN models (illustrated by Fig. 10).

For the MHD, smaller values mean better segmentation performance. It can be observed that the final results are better and more stable than the XY-, XZ-, YZ-results. Similar to the region-based results in Fig. 8, the performance of fibroglandular tissue segmentation is inferior to that of the mass and skin segmentation due to the complexity of the glandular tissues as discussed earlier.

Performing breast ultrasound image segmentation into different type of tissues in this study leads to a pixel classification problem. Because a 553×579 image slice contains over 320,000 pixels, the segmentation process took about 40 min in our previous attempts. We found that the serial computation of each pixel resulted in slow segmentation, therefore to speed it up, the pixels for classification were divided into 300 batches and the parallel computing approach was explored. Using the parallelization schemes, the segmentation procedure took less than 5 min, which is similar to the performance in the study by Gu. We expect that for further improvement of the segmentation efficiency, up-sampling and deconvolution layers could be added to the CNNs to train end-to-end classifiers for pixelwise prediction [34]. Furthermore, as efficient acceleration algorithm and more clinical data become available, 3D CNNs can be applied for direct 3D segmentation. The spatial relationship of pixels can be utilized more sufficiently by taking the 3D cubes rather than 2D image blocks from 3 orthogonal planes as the network inputs.

4. Discussion and conclusion

In this study, we presented an automatic segmentation method for breast ultrasound images based on machine learning algorithms. The automatic segmentation results demonstrate good consistency with manually segmented ground truth visually, which indicates that the proposed method has the same potential as physicians to distinguish different tissues from the breast ultrasound images. Compared with the all-tissue segmentation method proposed in the study by Gu, experimental results showed that the proposed method performed with better segmentation visualization and quantitative evaluation. Quantitative evaluation achieved better results than our previous study with an increase of 10% in the JSI metric, indicating the capacity of the method to provide segmentation references for radiologists. Furthermore, the proposed method performs fully automatic segmentation with trained CNNs while Gu's method used empirical thresholds in region classification. In Gu's study, the skin and fibroglandular tissue were taken as the same tissue and cannot be separated from each other. Our proposed method can determine the specific type of each pixel and distinguish the skin and fibroglandular tissue as two classes separately, using CNNs as classifiers. Thus, the proposed automated segmentation method might have the potential to provide an objective reference for radiologists on breast image segmentation, so as to help breast cancer diagnosis and breast density assessments.

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2018, Ca-a Cancer J. Clin. 68 (1) (2018) 7–30.

[2] W.Q. Chen, et al., Cancer statistics in China, 2015, Ca-a Cancer J. Clin. 66 (2) (2016) 115–132.

[3] B. Sahiner, et al., Malignant and benign breast masses on 3D US volumetric images: effect of computer-aided diagnosis on radiologist accuracy, Radiology 242 (3) (2007) 716–724.

[4] W.A. Berg, et al., Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer, Jama-J. Am. Med. Assoc. 299 (18) (2008) 2151–2163.

[5] W.H. Kim, et al., Ultrasonographic assessment of breast density, Breast Cancer Res. Treat. 138 (3) (2013) 851–859.

[6] M.T. Mandelson, et al., Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers, J. Natl. Cancer Inst. 92 (13) (2000) 1081–1087.

[7] J.A. Harvey, V.E. Bovbjerg, Quantitative assessment of mammographic breast density: relationship with breast cancer risk, Radiology 230 (1) (2004) 29–41.

[8] R. Jintamethasawat et al., Limited angle breast ultrasound tomography with a priori information and artifact removal, in: Medical Imaging 2017: Ultrasonic Imaging and Tomography, 2017, 10139.

[9] F.M. Hooi, O. Kripfgans, P.L. Carson, Acoustic attenuation imaging of tissue bulk properties with a priori information, J. Acoust. Soc. Am. 140 (3) (2016) 2113–2122.

[10] K. Doi, Computer-aided diagnosis in medical imaging: historical review, current status and future potential, Comput. Med. Imaging Graph. 31 (4–5) (2007) 198–211.

[11] C. Lo, et al., Computer-aided multiview tumor detection for automated whole breast ultrasound, Ultrason. Imaging 36 (1) (2014) 3–17.

[12] M. Xian, Y.T. Zhang, H.D. Cheng, Fully automatic segmentation of breast ultrasound images based on breast characteristics in space and frequency domains, Pattern Recogn. 48 (2) (2015) 485–497.

[13] M.H. Yap, E.A. Edirisinghe, H.E. Bez, Fully automatic lesion boundary detection in ultrasound breast images, in: Medical Imaging 2007: Image Processing, Pts 1–3, 2007, 6512.

[14] P.S. Rodrigues, G.A. Giraldi, Improving the non-extensive medical image segmentation based on Tsallis entropy, Pattern Anal. Appl. 14 (4) (2011) 369–379.

[15] A. Rodtook, S.S. Makhanov, Multi-feature gradient vector flow snakes for adaptive segmentation of the ultrasound images of breast cancer, J. Vis. Commun. Image Represent. 24 (8) (2013) 1414–1430.

[16] Y.L. Huang, D.R. Chen, Automatic contouring for breast tumors in 2-D sonography, 2005 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2005, pp. 3225–3228.

[17] L. Moraru, S. Moldovanu, A. Biswas, Optimization of breast lesion segmentation in texture feature space approach, Med. Eng. Phys. 36 (1) (2014) 129–135.

[18] R. Rodrigues, et al., A two-step segmentation method for breast ultrasound masses based on multi-resolution analysis, Ultrasound Med. Biol. 41 (6) (2015) 1737–1748.

[19] H.B. Kekre, P. Shrinath, Tumor delineation using statistical properties of the breast US images and vector quantization based clustering algorithms, Int. J. Image Graph. Signal Process 5 (11) (2013) 1–12.

[20] W.K. Moon, et al., Tumor detection in automated breast ultrasound images using quantitative tissue clustering, Med. Phys. 41 (4) (2014).

[21] J. Shan, H.D. Cheng, Y.X. Wang, A novel segmentation method for breast ultrasound images based on neutrosophic l-means clustering, Med. Phys. 39 (9) (2012) 5669–5682.

[22] A. Madabhushi, D.N. Metaxas, Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions, IEEE Trans. Med. Imaging 22 (2) (2003) 155–169.

[23] H.D. Cheng, et al., A novel Markov random field segmentation algorithm and its application to breast ultrasound image analysis, Proceedings of the 8th Joint Conference on Information Sciences, 2005, pp. 644–647.

[24] P. Gu, et al., Automated 3D ultrasound image segmentation to aid breast cancer image interpretation, Ultrasonics 65 (2016) 51–58.

[25] Q.H. Huang, et al., A robust graph-based segmentation method for breast tumors in ultrasound images, Ultrasonics 52 (2) (2012) 266–275.

[26] A. Belaid, et al., Phase-based level set segmentation of ultrasound images, IEEE Trans. Inf Technol. Biomed. 15 (1) (2011) 138–147.

[27] Q.H. Huang, Y.Z. Luo, Q.Z. Zhang, Breast ultrasound image segmentation: a survey, Int. J. Comput. Assist. Radiol. Surg. 12 (3) (2017) 493–507.

[28] P.N.T. Wells, M. Halliwell, Speckle in ultrasonic-imaging, Ultrasonics 19 (5) (1981) 225–229.

[29] B. Liu, et al., Probability density difference-based active contour for ultrasound image segmentation, Pattern Recogn. 43 (6) (2010) 2028–2042.

[30] W.L. Zhang, et al., Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation, Neuroimage 108 (2015) 214–224.

[31] E. Gibson et al., Deep residual networks for automatic segmentation of laparoscopic videos of the liver, in: Medical Imaging 2017: Image-Guided Procedures, Robotic

Interventions, and Modeling, 2017, 10135.

[32] K. Kamnitsas, et al., Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Med. Image Anal. 36 (2017) 61–78.

[33] R. Cardenes, R. de Luis-Garcia, M. Bach-Cuadra, A multidimensional segmentation evaluation for medical image data, Comput. Methods Programs Biomed. 96 (2) (2009) 108–124.

[34] J. Long E. Shelhamer T. Darrell, Fully convolutional networks for semantic segmentation, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (Cvpr), 2015, pp. 3431–3440.