

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/327629592>

Densely Deep Supervised Networks with Threshold Loss for Cancer Detection in Automated Breast Ultrasound: 21st International Conference, Granada, Spain, September 16–20, 2018, Proc...

Chapter in Lecture Notes in Computer Science · September 2018

DOI: 10.1007/978-3-030-00937-3_73

CITATIONS

2

READS

525

10 authors, including:



[Yi Wang](#)

The Chinese University of Hong Kong

23 PUBLICATIONS 59 CITATIONS

[SEE PROFILE](#)



[Qin Chenchen](#)

Shenzhen University

5 PUBLICATIONS 5 CITATIONS

[SEE PROFILE](#)



[Xin Yang](#)

The Chinese University of Hong Kong

47 PUBLICATIONS 607 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Deep learning for 3D medical image analysis [View project](#)



MICCAI Workshop on STACOM [View project](#)



Densely Deep Supervised Networks with Threshold Loss for Cancer Detection in Automated Breast Ultrasound

Na Wang¹, Cheng Bian¹, Yi Wang^{1,2}, Min Xu³, Chenchen Qin¹, Xin Yang⁴,
Tianfu Wang¹, Anhua Li³, Dinggang Shen⁵, and Dong Ni^{1,2}✉

¹ National-Regional Key Technology Engineering Laboratory for Medical
Ultrasound, School of Biomedical Engineering, Health Science Center,
Shenzhen University, Shenzhen, China

nidong@szu.edu.cn

² Medical UltraSound Image Computing (MUSIC) Lab, Shenzhen, China

³ Department of Ultrasound, Sun Yat-Sen University Cancer Center,
State Key Laboratory of Oncology in South China,
Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

⁴ Department of Computer Science and Engineering,
The Chinese University of Hong Kong, Hong Kong, China

⁵ Department of Radiology and BRIC, University of North Carolina at Chapel Hill,
Chapel Hill, NC 27599, USA

Abstract. Automated breast ultrasound (ABUS) is a new and promising tool for diagnosing breast cancer. However, reviewing ABUS images is extremely time-consuming and oversight errors could happen. We propose a novel 3D convolutional network for automatic cancer detection in ABUS. Our contribution is twofold. First, we propose a threshold loss function to provide voxel-level adaptive threshold for discriminating cancer and non-cancer, thus achieving high sensitivity with low FPs. Second, we propose a densely deep supervision (DDS) mechanism to improve the sensitivity significantly by utilizing multi-scale discriminative features of all layers. Both class-balanced cross entropy loss and overlap loss are employed to enhance DDS performance. The efficacy of the proposed network is validated on a dataset of 196 patients with 661 cancer regions. The 4-fold cross-validation experiments show our network obtains a sensitivity of 93% with 2.2 FPs per ABUS volume. Our proposed novel network can provide an accurate and automatic cancer detection tool for breast cancer screening by maintaining high sensitivity with low FPs.

1 Introduction

Recently, automated breast ultrasound (ABUS) has been developed as a new and promising tool for diagnosing breast cancer. ABUS can provide 3D views of the breast by automatically scanning the whole breast. It also has a number of advantages over traditional 2D handheld ultrasound: higher reproducibility, less operator dependence, and less image acquisition time [8]. However, reviewing ABUS images is extremely time-consuming, because a typical exam often

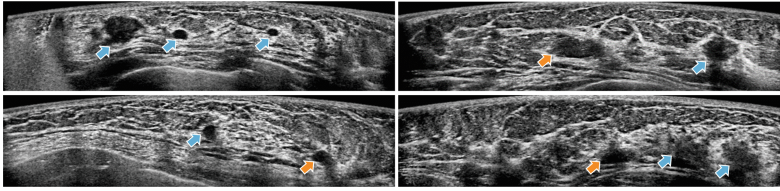


Fig. 1. Example ABUS images. Blue arrows indicate biopsy-proven cancer regions, and orange arrows indicate cancer mimicry with very similar appearance to cancer regions.

consists of three volumes of each breast in order to cover the whole breast. Furthermore, the large size of the ABUS volumes may cause oversight errors in some malignancies. Therefore, automated cancer detection in ABUS is highly expected to assist clinicians in facilitating the identification of breast cancer.

As shown in Fig. 1, computer-aided detection (CAdE) of cancer from ABUS images remains very challenging. First, cancers often possess high intraclass appearance variation caused by various factors like imaging artifacts of acoustic shadows and speckles, deformation of soft tissues and large difference in lesion size. Second, malignant lesions may possibly appear similar to other structures, such as benign lesions and normal hypoechoic structures. Third, small cancers are mostly likely to be missed even by clinical experts because of the relatively low quality of ultrasound imaging, the existence of similar normal structures and the oversight errors. Finally, the severe class imbalance between cancer and non-cancer voxels is another challenge because of the extremely small size of the cancer relative to the large ABUS volume. In this situation, the predictive model developed based on machine learning methods could be biased and inaccurate.

To better assist clinician with cancer screening, a number of CAdE approaches have been developed. Moon *et al.* developed a CAdE system based on a two-stage multi-scale blob analysis method [5]. The system showed sensitivities of 100%, 90% and 70% with false positives (FPs) per volume of 17.4, 8.8, and 2.7, respectively. Tan *et al.* proposed a multi-stage system using an ensemble of neural networks to classify cancers [9]. Although the FPs per volume could be controlled at 1, the sensitivity was only 64%. Lo *et al.* employed watershed segmentation to extract potential abnormalities in ABUS and reduced FPs using various quantitative features [3]. The sensitivities were 100%, 90% and 80% with FPs per volume of 9.44, 5.42, and 3.33, respectively. Generally, maintaining high sensitivity with low FPs remains a vital problem in ABUS CAdE.

Recently, a surge of deep learning is becoming dominant over traditional CAdE methods [7]. We propose a novel 3D convolutional neural network (CNN) for automatic cancer detection in ABUS. We believe we are the first to employ deep learning based techniques for this problem. Our contribution is twofold. First, we propose a threshold loss function by adding a threshold map (TM) layer in the CNN. The proposed method provides voxel-level adaptive threshold to classify voxels into cancer or non-cancer, thus achieving high sensitivity with low FPs. Second, we propose a densely deep supervision (DDS) mechanism

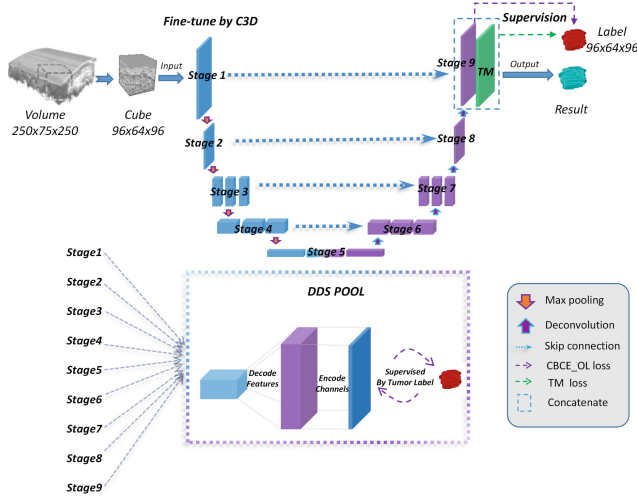


Fig. 2. The illustration of the proposed network for cancer detection in ABUS.

to improve the sensitivity significantly by utilizing multi-scale discriminative features of all layers [2]. We employ two loss functions to enhance DDS performance. Specifically, class-balanced cross entropy is employed to tackle the issue of limited positive training samples; overlap loss is for selecting discriminative cancer representations. The proposed network was extensively evaluated on a 196-patient dataset, with 661 cancer regions.

2 Methods

Figure 2 illustrates the proposed network, which leverages the DDS to learn more discriminative cancer representations for the improvement of detection sensitivity, and utilizes the proposed TM to adaptively refine the probability map for the FPs reduction while maintaining high sensitivity.

2.1 Network Architecture

We choose the most successful segmentation net 3D U-net [6], [11] as our backbone architecture but make the following modifications (Fig. 2): (a) we employ the pre-trained C3D model [10] to fine-tune our network parameters to restrain the over-fitting issue induced by limited ABUS training samples; (b) we design a DDS mechanism to effectively learn discriminative features for cancer identification and meanwhile boost the gradient flow of the whole network; (c) we add a TM layer to provide voxel-level adaptive threshold for optimizing the probability map, thus achieving high sensitivity with low FPs. Specifically, the TM is automatically learned from the complementary information of learned features, label information and predicted probability map. (d) other customizations: each

convolutional layer is connected with a batch normalization (BN) layer and a rectified linear unit (ReLU); each of the stages 3–7 uses 3 convolutional layers to increase receptive fields for using more global information.

2.2 Densely Deep Supervision

Deep neural networks are powerful to generate abundant multi-scale features for object detection in natural images. Nevertheless, this is challenging in ABUS images, because breast cancers possess high intraclass appearance variation and some are relatively subtle. Furthermore, due to gradient vanishing issue, the parameter tuning processes of the 3D CNN may encounter low efficiency and overfitting problems. Taking advantages of deeply supervised nets (DSN) [2], we implement the DDS into our 3D U-net to alleviate above problems by fully exploiting the multi-scale features from all stages. Specifically, we input each of the stages 1-9 and the concatenation of all stages into DDS pool (thus totally 10 DSNs), and introduce a DDS loss function to supervise the generation of cancer probability map. The DDS loss function is defined as

$$\mathcal{L}_{dds}(X, Y; \mathcal{W}, \omega) = \sum_{t=1}^{T-1} \left(\theta_t * (\mathcal{L}_{cbce}^{(t)}(X, Y; \mathcal{W}, \omega^{(t)}) + \mathcal{L}_{ol}^{(t)}(X, Y; \mathcal{W}, \omega^{(t)})) \right), \quad (1)$$

where X is training image and Y is corresponding label image, \mathcal{W} is the weight of main network, $\omega = (\omega^{(1)}, \omega^{(2)}, \dots, \omega^{(T)})$ where $(\omega^{(1)}, \omega^{(2)}, \dots, \omega^{(T-1)})$ are the weights of each DSN and $\omega^{(T=11)}$ is the weight of TM layer, $\theta = (\theta_1, \theta_2, \dots, \theta_T)$ are the coefficients to weight each DSN loss and threshold loss in total loss, respectively. The class-balanced cross entropy (CBCE) loss \mathcal{L}_{cbce} and overlap (OL) loss \mathcal{L}_{ol} in Eq. (1) will be explained in the following paragraphs.

Class-Balanced Cross Entropy Loss. Considering lots of breast cancers are relatively subtle, the distribution of cancer/non-cancer regions in a large ABUS volume is heavily biased. We employ a CBCE loss to tackle the issue of limited positive training samples. Specifically, we introduce a class-balancing weight α to offset the imbalance between cancer and non-cancer voxels and define the CBCE loss function as

$$\begin{aligned} \mathcal{L}_{cbce}^{(t)}(\mathcal{W}, \omega^{(t)}) = & -\alpha \sum_{i \in Y_+} \log Ps(y_i = 1 | X; \mathcal{W}, \omega^{(t)}) \\ & -(1 - \alpha) \sum_{i \in Y_-} \log Ps(y_i = 0 | X; \mathcal{W}, \omega^{(t)}), \end{aligned} \quad (2)$$

where $y_i \in \{0, 1\}$ represents the label of Y at location i . $\alpha = \text{sum}(Y_-)/\text{sum}(Y)$ and $1 - \alpha = \text{sum}(Y_+)/\text{sum}(Y)$, where Y_- and Y_+ are non-cancer and cancer label sets. $Ps(y_i | X; \mathcal{W}, \omega^{(t)}) = e^{z_j} / \sum_k e^{z_k} \in (0, 1)$ denotes the feature map obtained by *softmax* function, where z_j is the score of class j and $k = 2$.

Overlap Loss. To further improve the detection sensitivity, especially for better identification of subtle cancers, we design a new loss function, i.e. overlap loss,

to learn more discriminative features for cancer representations. The OL loss function is defined as

$$\mathcal{L}_{ol}^{(t)}(\mathcal{W}, \omega^{(t)}) = \sum_{i=1}^{|X|} (Ps(y_i = 1|X; \mathcal{W}, \omega^{(t)}) * Ps(y_i = 0|X; \mathcal{W}, \omega^{(t)})). \quad (3)$$

The purpose of this loss is to ensure that cancer regions should have little overlap with non-cancer regions, thus the optimal overlapped region should be zero. By minimizing the overlap loss, we attempt to learn more discriminative features to distinguish cancer and non-cancer regions.

2.3 Threshold Map

Although the probability map generated by the proposed DDS mechanism can indicate cancer locations with high sensitivity, it may still have some high probability regions which are actually normal tissues. Thus the further post-processing of the probability map is essential to obtain a better detection. However, conventional methods often fail in simultaneously achieving high sensitivity and low FPs, for example, fixed threshold is sensitive to the selected value; softmax method often achieves high FPs when coping with challenging problems; directly performing conditional random field tends to lower the sensitivity.

To address the above issues, we concatenate and train a threshold map (TM) layer in our network to adaptively refine the probability map for better detection. The proposed TM can provide voxel-level adaptive threshold to classify voxels into cancer or non-cancer by making use of all the information from learned features, label information and probability map, thus achieving a good balance between high sensitivity and low FPs. To train the TM, we design a new loss function, i.e. threshold loss, which can be calculated as follow:

$$\mathcal{L}_{threshold}(\mathcal{W}, \omega^{(t)}) = 1 - \frac{2 * |Mask(y_i = 1|X; \mathcal{W}, \omega^{(t)}) * Y|}{|Mask(y_i; \mathcal{W}, \omega^{(t)})| + |Y|}, \quad (4)$$

$$Mask(y_i; \mathcal{W}, \omega^{(t)}) = 1/(1 + e^{-tmp}), \quad (5)$$

$$tmp = \begin{cases} Ps(y_i = 1|X; \mathcal{W}, \omega^{(t)}), & Ps(y_i) > Threshold_Map(y_i) \\ -e^{10}, & else \end{cases}. \quad (6)$$

The objective of this threshold loss is to learn a voxel-wise threshold map, which can be further employed to adaptively refine the probability map by suppressing non-cancer regions and meanwhile maintaining cancer regions. To the best of our knowledge, we are the first to design a threshold map to adaptively optimize the probability map. The efficacy of the proposed TM is shown in our experiments.

We summarize the total loss function for our cancer detection as

$$\mathcal{L}_{total} = \mathcal{L}_{dds} + \theta_T * \mathcal{L}_{threshold}. \quad (7)$$

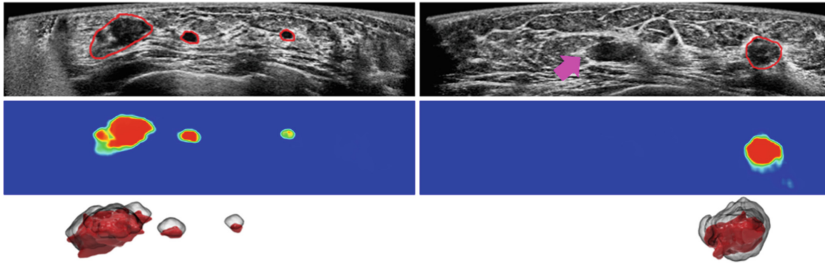


Fig. 3. Example results of cancer detection in ABUS. Top row: ABUS images with annotated cancer locations (red contours) and purple arrow indicates cancer mimicry. Middle row: refined cancer probability maps by the proposed threshold map. Bottom row: 3D visualization of our detected cancers (gray) and ground truth (red).

3 Experiments

Materials. Experiments were carried on the dataset obtained using Invenia ABUS system (GE, USA) in Sun Yat-Sen University Cancer Center. Informed consent for this retrospective study was obtained from our institutional review board. To cover the whole breast, three volumes including anterior-posterior, medial and lateral passes were obtained for each breast. Thus for each patient, six ABUS volumes were acquired. The voxel resolutions of the acquired 3D ABUS volumes were 0.511 mm, 0.082 mm and 0.200 mm in the transverse, sagittal and coronal direction, respectively.

In this study, ABUS data from 196 women (age range: 30–75 years, mean 49 years) with biopsy-proven breast cancers were collected. From these ABUS data, 559 volumes were annotated by an experienced clinician, which included 661 cancer regions (volume: 0.01–86.54 cm³, mean: 2.84 cm³). Four-fold cross-validation was conducted to evaluate the detection performance. As a control, 119 ABUS volumes with no abnormal findings were also included for evaluations.

Implementation Details. Our proposed framework was implemented with the popular library Keras for Tensorflow. To tackle the issues of limited cancer samples and demanding 3D computational cost, we divided ABUS volume into multiple $96 \times 64 \times 96$ cubes and adopted data augmentation (i.e., translation, rotation, cropping, flipping) for training, and further combine the predicted cube into a volume as detection result. The framework was trained on a 8x NVIDIA Tesla GPU. Adaptive moment estimation was used to train the whole framework. We set the learning rate as $1e-4$ and learning stopped after 30000 iterations.

Detection Performance. We extensively compared our method with state of the art, including SegNet [1], FCN [4], U-net [11]. To illustrate the efficacy of the proposed triplet loss, we further evaluated the proposed network with different loss functions, including dice loss (DL), cross entropy (CE) loss, CBCE loss, CBCE-OL loss, and the triplet CBCE-OL-TM loss.

Figure 3 visualizes the cancer detection results by our network. By utilizing the proposed DDS and threshold map, our network can output accurate cancer probability maps even when cancers were subtle or cancer mimicry existed.

Table 1. Sensitivities and corresponding FPs per volume for different methods

Method	Sensitivity (%)	FPs per volume		
		Cancer	Normal	All
3D-SegNet [1]	63.75	1.69	2.42	1.82
3D-FCN [4]	66.47	1.84	2.77	2.00
3D-Unet [11]	79.61	0.56	1.04	0.65
3D-Unet-DDS-DL	84.21	3.56	3.06	3.47
3D-Unet-DDS-CE	76.16	4.77	4.90	4.79
3D-Unet-DDS-CBCE	88.69	3.60	4.48	3.75
3D-Unet-DDS-CBCE-OL	100.00	9.81	8.62	9.60
3D-Unet-DDS-CBCE-OL-TM	93.04	2.23	2.19	2.22

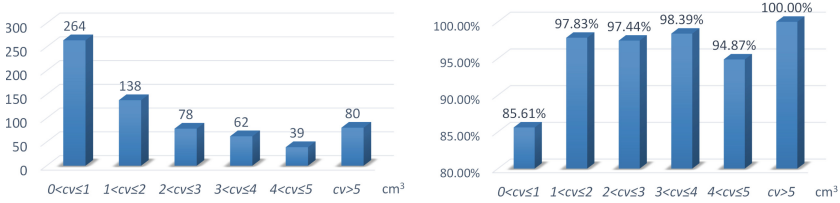

Fig. 4. Left: the cancer volume (cv) distribution of all 661 cancer regions. Right: the detection sensitivities for different ranges of cancer volume.

Table 1 lists the sensitivities and corresponding FPs per volume for different methods. Our network obtained a sensitivity of 93% with 2.2 FPs per ABUS volume. Compared to SegNet and FCN, our network significantly improved detection sensitivity meanwhile still controlled FPs at about 2. Although U-net obtained a FP value less than 1, its sensitivity is less than 80%. By observing the results of our network with different loss functions in Table 1, the designed DDS and TM contributed to the improvement of detection performance. Specifically, the DDS with CBCE-OL loss contributed to select discriminative cancer representations and the TM loss helped to adaptively optimize probability map for the FPs reduction while maintaining high sensitivity. Table 1 also records the difference between FPs in cancer and normal volumes. Our network got the number of FPs for normal volumes slightly lower than that for abnormal volumes.

Figure 4 further illustrates the volume distribution of all 661 cancer regions, as well as the corresponding detection sensitivities for different ranges of cancer volume. It can be observed that our network achieved a sensitivity above 85% even when cancer volume was smaller than 1 cm³, and when cancer volume was larger than 5 cm³, we got a sensitivity of 100%.

4 Conclusion

In this paper, we propose a novel 3D convolutional network for automatic cancer detection in ABUS. To the best of our knowledge, we are the first to employ deep

learning techniques for this problem. In the proposed network, a novel threshold map is designed to provide voxel-level adaptive threshold to classify voxels into cancer or non-cancer, thus achieving high sensitivity with low FPs. Furthermore, a densely deep supervision mechanism is employed to improve the sensitivity greatly by utilizing multi-scale discriminative features of all layers. Experiments show our network obtains a sensitivity of 93% with 2.2 FPs per ABUS volume. Our method can provide an accurate and automatic cancer detection tool for breast cancer screening by maintaining high sensitivity and low FPs.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grant 61571304, Grant 61701312, and in part by the Shenzhen Peacock Plan under Grant KQTD2016053112051497.

References

1. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(12), 2481–2495 (2017)
2. Lee, C.Y., Xie, S., Gallagher, P., Zhang, Z., Tu, Z.: Deeply-supervised nets. In: *Artificial Intelligence and Statistics*, pp. 562–570 (2015)
3. Lo, C.M., et al.: Multi-dimensional tumor detection in automated whole breast ultrasound using topographic watershed. *IEEE Trans. Med. Imaging* **33**(7), 1503–1511 (2014)
4. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: *CVPR*, pp. 3431–3440 (2015)
5. Moon, W.K., Shen, Y.W., Bae, M.S., Huang, C.S., Chen, J.H., Chang, R.F.: Computer-aided tumor detection based on multi-scale blob detection algorithm in automated breast ultrasound images. *IEEE Trans. Med. Imaging* **32**(7), 1191–1200 (2013)
6. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
7. Shen, D., Wu, G., Suk, H.I.: Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* **19**, 221–248 (2017)
8. Shin, H.J., Kim, H.H., Cha, J.H.: Current status of automated breast ultrasonography. *Ultrasonography* **34**(3), 165 (2015)
9. Tan, T., Platel, B., Mus, R., Tabar, L., Mann, R.M., Karssemeijer, N.: Computer-aided detection of cancer in automated 3-D breast ultrasound. *IEEE Trans. Med. Imaging* **32**(9), 1698–1706 (2013)
10. Tran, D., Bourdev, L., Fergus, R., Torresani, L., Paluri, M.: Learning spatiotemporal features with 3D convolutional networks. In: *ICCV*, pp. 4489–4497. IEEE (2015)
11. Yang, X.: Towards automatic semantic segmentation in volumetric ultrasound. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) *MICCAI 2017*. LNCS, vol. 10433, pp. 711–719. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66182-7_81