



SimPLe: Similarity-Aware Propagation Learning for Weakly-Supervised Breast Cancer Segmentation in DCE-MRI

Yuming Zhong and Yi Wang^(✉)

Smart Medical Imaging, Learning and Engineering (SMILE) Lab, Medical UltraSound Image Computing (MUSIC) Lab, School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen University, Shenzhen, China
onewang@szu.edu.cn

Abstract. Breast dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) plays an important role in the screening and prognosis assessment of high-risk breast cancer. The segmentation of cancerous regions is essential useful for the subsequent analysis of breast MRI. To alleviate the annotation effort to train the segmentation networks, we propose a weakly-supervised strategy using extreme points as annotations for breast cancer segmentation. Without using any bells and whistles, our strategy focuses on fully exploiting the learning capability of the routine training procedure, i.e., the *train - fine-tune - retrain* process. The network first utilizes the pseudo-masks generated using the extreme points to *train* itself, by minimizing a contrastive loss, which encourages the network to learn more representative features for cancerous voxels. Then the trained network *fine-tunes* itself by using a similarity-aware propagation learning (SimPLe) strategy, which leverages feature similarity between unlabeled and positive voxels to propagate labels. Finally the network *retrains* itself by employing the pseudo-masks generated using previous fine-tuned network. The proposed method is evaluated on our collected DCE-MRI dataset containing 206 patients with biopsy-proven breast cancers. Experimental results demonstrate our method effectively fine-tunes the network by using the SimPLe strategy, and achieves a mean Dice value of 81%. *Our code is publicly available at <https://github.com/Abner228/SmileCode>.*

Keywords: Breast cancer · Weakly-supervised learning · Medical image segmentation · Contrastive learning · DCE-MRI

1 Introduction

Breast cancer is the most common cause of cancer-related deaths among women all around the world [8]. Early diagnosis and treatment is beneficial to improve the survival rate and prognosis of breast cancer patients. Mammography, ultrasonography, and magnetic resonance imaging (MRI) are routine imaging modalities for breast examinations [15]. Recent clinical studies have proven that

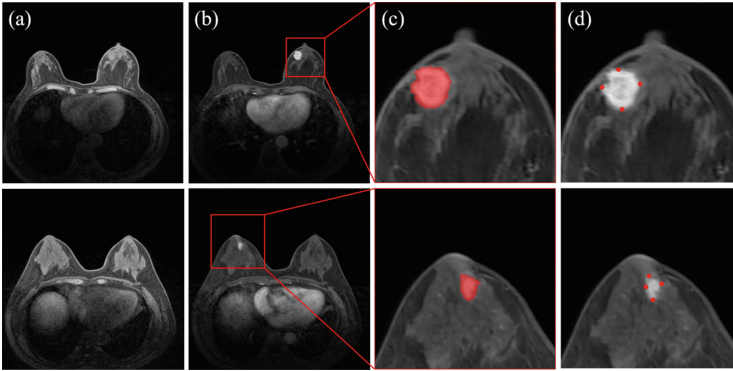


Fig. 1. Breast MRI and different annotations: (a) T1-weighted images, (b) corresponding contrast-enhanced images, (c) the cancer annotation with full segmentation masks, and (d) the cancer annotation using extreme points (*note that to facilitate the visualization, here we show the extreme points in 2D images, our method is based on 3D*).

dynamic contrast-enhanced (DCE)-MRI has the capability to reflect tumor morphology, texture, and kinetic heterogeneity [14], and is with the highest sensitivity for breast cancer screening and diagnosis among current clinical imaging modalities [17]. The basis for DCE-MRI is a dynamic T1-weighted contrast enhanced sequence (Fig. 1). T1-weighted acquisition depicts enhancing abnormalities after contrast material administration, that is, the cancer screening is performed by using the post-contrast images. Radiologists will analyze features such as texture, morphology, and then make the treatment plan or prognosis assessment. Computer-aided feature quantification and diagnosis algorithms have recently been exploited to facilitate radiologists analyze breast DCE-MRI [12, 22], in which automatic cancer segmentation is the very first and important step.

To better support the radiologists with breast cancer diagnosis, various segmentation algorithms have been developed [20]. Early studies focused on image processing based approaches by conducting graph-cut segmentation [29] or analyzing low-level hand-crafted features [1, 11, 19]. These methods may encounter the issue of high computational complexity when analyzing volumetric data, and most of them require manual interactions. Recently, deep-learning-based methods have been applied to analyze breast MRI. Zhang *et al.* [28] proposed a mask-guided hierarchical learning framework for breast tumor segmentation via convolutional neural networks (CNNs), in which breast masks were also required to train one of CNNs. This framework achieved a mean Dice value of 72% on 48 testing T1-weighted scans. Li *et al.* [16] developed a multi-stream fusion mechanism to analyze T1/T2-weighted scans, and obtained a Dice result of 77% on 313 subjects. Gao *et al.* [7] proposed a 2D CNN architecture with designed attention modules, and got a Dice result of 81% on 87 testing samples. Zhou *et al.* [30] employed a 3D affinity learning based multi-branch ensemble network for the

segmentation refinement and generated 78% Dice on 90 testing subjects. Wang *et al.* [24] integrated a combined 2D and 3D CNN and a contextual pyramid into U-net to obtain a Dice result of 76% on 90 subjects. Wang *et al.* [25] proposed a tumor-sensitive synthesis module to reduce false segmentation and obtained 78% Dice value. To reduce the huge annotation burden for the segmentation task, Zeng *et al.* [27] presented a semi-supervised strategy to segment the manually cropped DCE-MRI scans, and attained a Dice value of 78%.

Although [27] has been proposed to alleviate the annotation effort, to acquire the voxel-level segmentation masks is still time-consuming and laborious, see Fig. 1(c). Weakly-supervised learning strategies such as extreme points [5, 21], bounding box [6] and scribbles [4] can be promising solutions. Roth *et al.* [21] utilized extreme points to generate scribbles to supervise the training of the segmentation network. Based on [21], Dorent *et al.* [5] introduced a regularized loss [4] derived from a Conditional Random Field (CRF) formulation to encourage the prediction consistency over homogeneous regions. Du *et al.* [6] employed bounding boxes to train the segmentation network for organs. However, the geometric prior used in [6] can not be an appropriate strategy for the segmentation of lesions with various shapes. To our knowledge, currently only one weakly-supervised work [18] has been proposed for breast mass segmentation in DCE-MRI. This method employed three partial annotation methods including single-slice, orthogonal-slice (i.e., 3 slices) and interval-slice (~ 6 slices) to alleviate the annotation cost, and then constrained segmentation by estimated volume using the partial annotation. The method obtained a Dice value of 83% using the interval-slice annotation, on a testing dataset containing only 28 patients.

In this study, we propose a simple yet effective weakly-supervised strategy, by using extreme points as annotations (see Fig. 1(d)) to segment breast cancer. Specifically, we attempt to optimize the segmentation network via the conventional *train - fine-tune - retrain* process. The initial training is supervised by a contrastive loss to pull close positive voxels in feature space. The fine-tune is conducted by using a similarity-aware propagation learning (SimPLe) strategy to update the pseudo-masks for the subsequent retrain. We evaluate our method on a collected DCE-MRI dataset containing 206 subjects. Experimental results show our method achieves competitive performance compared with fully supervision, demonstrating the efficacy of the proposed SimPLe strategy.

2 Method

The proposed SimPLe strategy and the *train - fine-tune - retrain* procedure is illustrated in Fig. 2. The extreme points are defined as the left-, right-, anterior-, posterior-, inferior-, and superior-most points of the cancerous region in 3D. The initial pseudo-masks are generated according to the extreme points by using the random walker algorithm. The segmentation network is firstly trained based on the initial pseudo-masks. Then SimPLe is employed to fine-tune the network and update the pseudo-masks. At last, the network is retrained from random initialization using the updated pseudo-masks.

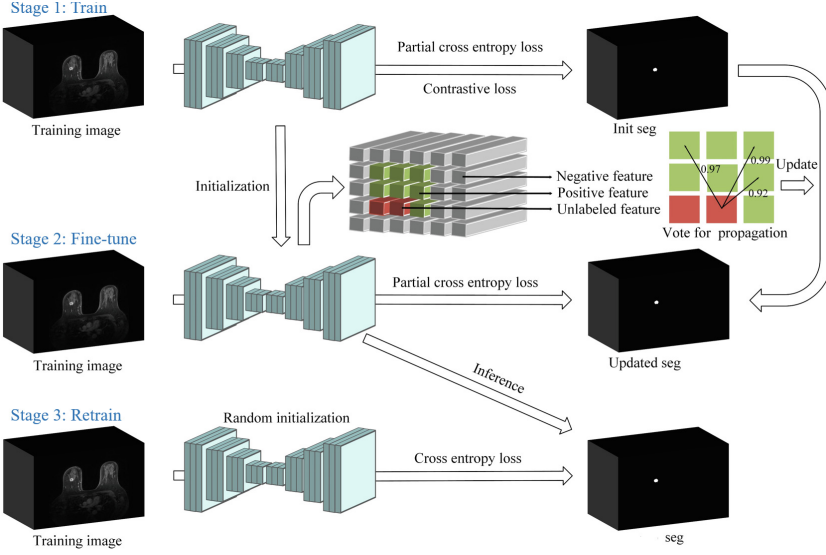


Fig. 2. The schematic illustration of the proposed similarity-aware propagation learning (SimPLE) and the *train - fine-tune - retrain* procedure for the breast cancer segmentation in DCE-MRI.

2.1 Generate Initial Pseudo-masks

We use the extreme points to generate pseudo-masks based on random walker algorithm [9]. To improve the performance of random walker, according to [21], we first generate scribbles by searching the shortest path on gradient magnitude map between each extreme point pair via the Dijkstra algorithm [3]. After generating the scribbles, we propose to dilate them to increase foreground seeds for random walker. Voxels outside the bounding box (note that once we have the six extreme points, we have the 3D bounding box of the cancer) are expected to be the background seeds. Next, the random walker algorithm is used to produce a foreground probability map $\hat{Y} : \Omega \subset \mathbb{R}^3 \rightarrow [0, 1]$, where Ω is the spatial domain. To further increase the area of foreground, the voxel at location \mathbf{k} is considered as new foreground seed if $\hat{Y}(\mathbf{k})$ is greater than 0.8 and new background seed if $\hat{Y}(\mathbf{k})$ is less than 0.1. Then we run the random walker algorithm repeatedly. After seven times iterations, we set foreground in the same way via the last output probability map. Voxels outside the bounding box are considered as background. The rest of voxels remain unlabeled. This is the way initial pseudo-masks $Y_{init} : \Omega \subset \mathbb{R}^3 \rightarrow \{0, 1, 2\}$ generated, where 0, 1 and 2 represent negative, positive and unlabeled.

2.2 Train Network with Initial Pseudo-masks

Let $X : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ denotes a training volume. Let f and θ be network and its parameters, respectively. A simple training approach is to minimize the partial cross entropy loss \mathcal{L}_{pce} , which is formulated as:

$$\mathcal{L}_{pce} = - \sum_{Y_{init}(\mathbf{k})=0} \log(1 - f(X; \theta)(\mathbf{k})) - \sum_{Y_{init}(\mathbf{k})=1} \log(f(X; \theta)(\mathbf{k})). \quad (1)$$

Moreover, supervised contrastive learning is employed to encourage voxels of the same label to gather around in feature space. It ensures the network to learn discriminative features for each category. Specifically, features corresponding to N negative voxels and N positive voxels are randomly sampled, then the contrastive loss \mathcal{L}_{ctr} is minimized:

$$\begin{aligned} \mathcal{L}_{ctr} = & -\frac{1}{2N-1} \sum_{\mathbf{k}_n \in \mathcal{N}(\mathbf{k})} \log(1 - \sigma(\text{sim}(\mathcal{Z}(\mathbf{k}), \mathcal{Z}(\mathbf{k}_n))/\tau)) \\ & - \frac{1}{2N-1} \sum_{\mathbf{k}_p \in \mathcal{P}(\mathbf{k})} \log(\sigma(\text{sim}(\mathcal{Z}(\mathbf{k}), \mathcal{Z}(\mathbf{k}_p))/\tau)), \end{aligned} \quad (2)$$

where $\mathcal{P}(\mathbf{k})$ denotes the set of points with the same label as the voxel \mathbf{k} and $\mathcal{N}(\mathbf{k})$ denotes the set of points with the different label. $\mathcal{Z}(\mathbf{k})$ denotes the feature vector of the voxel at location \mathbf{k} . $\text{sim}(\cdot, \cdot)$ is the cosine similarity function. σ denotes sigmoid function. τ is a temperature parameter.

To summarize, we employ the sum of the partial cross entropy loss \mathcal{L}_{pce} and the contrastive loss \mathcal{L}_{ctr} to train the network with initial pseudo-masks:

$$\mathcal{L}_{train} = \mathcal{L}_{pce} + \mathcal{L}_{ctr}. \quad (3)$$

2.3 SimPLe-Based Fine-Tune and Retrain

The performance of the network trained by the incomplete initial pseudo-masks is still limited. We propose to fine-tune the entire network using the pre-trained weights as initialization. The fine-tune follows the SimPLe strategy which evaluates the similarity between unlabeled voxels and positive voxels to propagate labels to unlabeled voxels. Specifically, N positive voxels are randomly sampled as the referring voxel. For each unlabeled voxel \mathbf{k} , we evaluate its similarity with all referring voxels:

$$\mathcal{S}(\mathbf{k}) = \sum_{i=1}^N \mathbb{I}\{\text{sim}(\mathcal{Z}_{\mathbf{k}}, \mathcal{Z}_i) > \lambda\}, \quad (4)$$

where $\mathbb{I}(\cdot)$ is the indicator function, which is equal to 1 if the cosine similarity is greater than λ and 0 if less. If $\mathcal{S}(\mathbf{k})$ is greater than αN , the voxel at location \mathbf{k} is considered as positive. Then the network is fine-tuned using the partial cross entropy loss same as in the initial train stage. The loss function $\mathcal{L}_{finetune}$ is formulated as:

$$\mathcal{L}_{finetune} = \mathcal{L}_{pce} - w \cdot \sum_{\mathcal{S}(\mathbf{k}) > \alpha N} \log(f(X; \theta)(\mathbf{k})), \quad (5)$$

where w is the weighting coefficient that controls the influence of the pseudo labels. To reduce the influence of possible incorrect label propagation, pseudo labels for unlabeled voxels are valid only for the current iteration when they are generated.

After the fine-tune completed, the network generates binary pseudo-masks for every training data, which are expected to be similar to the ground-truths provided by radiologists. Finally the network is retrained from random initialization by minimizing the cross entropy loss with the binary pseudo-masks.

3 Experiments

Dataset. We evaluated our method on an in-house breast DCE-MRI dataset collected from the Cancer Center of Sun Yat-Sen University. In total, we collected 206 DCE-MRI scans with biopsy-proven breast cancers. All MRI scans were examined with 1.5T MRI scanner. The DCE-MRI sequences (TR/TE = 4.43 ms/1.50 ms, and flip angle = 10°) using gadolinium-based contrast agent were performed with the T1-weighted gradient echo technique, and injected 0.2 ml/kg intravenously at 2.0 ml/s followed by 20 ml saline. The DCE-MRI volumes have two kinds of resolution, $0.379 \times 0.379 \times 1.700 \text{ mm}^3$ and $0.511 \times 0.511 \times 1.000 \text{ mm}^3$.

All cancerous regions and extreme points were manually annotated by an experienced radiologist via ITK-SNAP [26] and further confirmed by another radiologist. We randomly divided the dataset into 21 scans for training and the remaining scans for testing¹. Before training, we resampled all volumes into the same target spacing $0.600 \times 0.600 \times 1.000 \text{ mm}^3$ and normalized all volumes as zero mean and unit variance.

Implementation Details. The framework was implemented in PyTorch, using a NVIDIA GeForce GTX 1080 Ti with 11GB of memory. We employed 3D U-net [2] as our network backbone.

- *Train*: The network was trained by stochastic gradient descent (SGD) for 200 epochs, with an initial learning rate $\eta = 0.01$. The ploy learning policy was used to adjust the learning rate, $(1 - \text{epoch}/200)^{0.9}$. The batch size was 2, consisting of a random foreground patch and a random background patch located via initial segmentation Y_{init} . Such setting can help alleviate class imbalance issue. The patch size was $128 \times 128 \times 96$. For the contrastive loss, we set $N = 100$, temperature parameter $\tau = 0.1$.
- *Fine-tune*: We initialized the network with the trained weights. We trained it by SGD for 100 iterations, with $\eta = 0.0001$. The ploy learning policy was also used. For the SIMPLe strategy, we set $N = 100$, $\lambda = 0.96$, $\alpha = 0.96$, $w = 0.1$.

¹ We have tried different amount of training data to investigate the segmentation performance of the fully-supervised network. The results showed that when using 21, 42, 63 scans for training, the Dice results changed very little, within 0.3%. Therefore, to include more testing data, we chose to use 21 (10%) out of 206 scans for training.

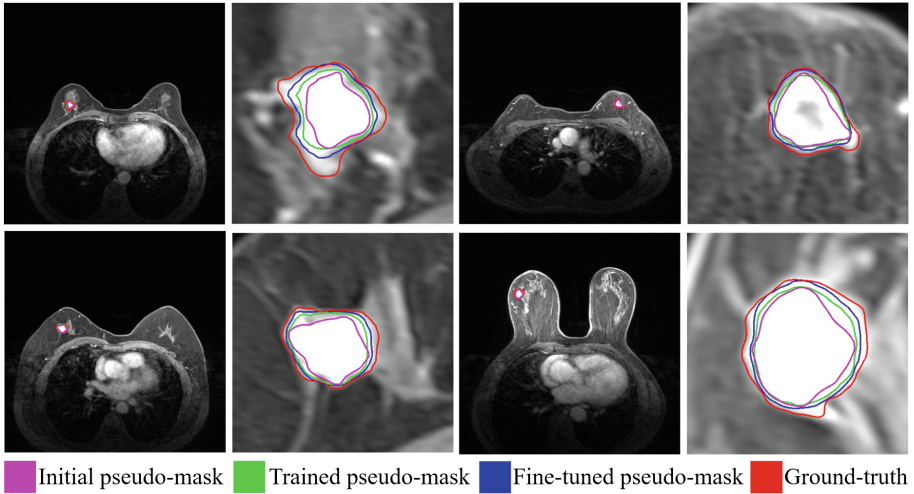


Fig. 3. The pseudo-masks (shown as boundaries) at different training stages, including the initial pseudo-mask generated by the random walker (purple), the trained network’s output (green), the fine-tuned pseudo-mask using SimPLe (blue) and the ground-truth (red). Note that all images here are the training images. (Color figure online)

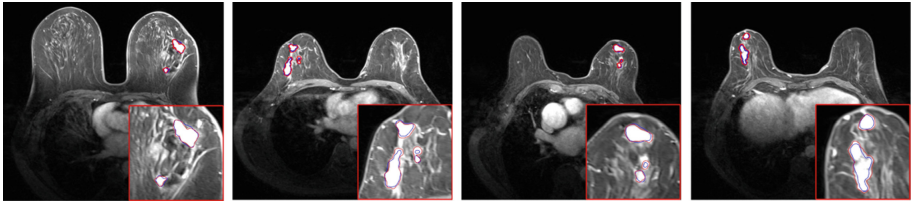


Fig. 4. The segmentation visualization in transversal slices. The blue and red contours are the segmented boundaries and the ground-truths, respectively. (Color figure online)

After fine-tuned, the last weights were used to generate the binary pseudo-masks.

- *Retrain*: The training strategy was the same as the initial train stage.
- *Inference*: A sliding window approach was used. The window size equaled the patch size used during training. Stride was set as half of a patch.

Quantitative and Qualitative Analysis. We first verified the efficacy of our SimPLe in the training stage. Figure 3 illustrates the pseudo-masks at different training stages. It is obvious that our SimPLe effectively updated the pseudo-masks to make them approaching the ground-truths. Therefore, such fine-tuned pseudo-masks could be used to retrain the network for better performance.

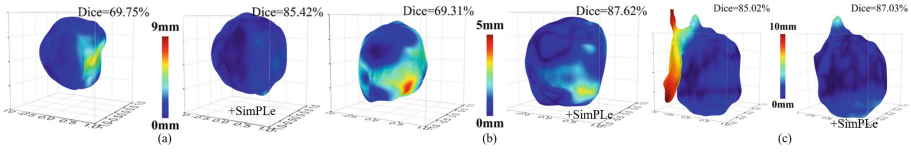


Fig. 5. Three cases of 3D visualization of the surface distance between segmented surface and ground-truth. Each case shows the $\mathcal{L}_{pce} + \mathcal{L}_{ctr}$ result and the SimPLe fine-tuned result. The proposed SimPLe consistently enhances the segmentation.

Table 1. The numerical results of different methods for breast cancer segmentation

Methods	Dice [%]	Jaccard [%]	ASD [mm]	95HD [mm]
$\mathcal{L}_{pce} + \mathcal{L}_{crf}$ [4]	64.97 ± 28.66	53.83 ± 27.26	1.01 ± 0.81	3.90 ± 3.91
$\mathcal{L}_{pce} + \mathcal{L}_{ctr}$	69.39 ± 24.09	57.36 ± 23.31	0.95 ± 0.66	3.60 ± 3.49
Entropy Min [10]	71.94 ± 17.86	58.74 ± 18.69	0.88 ± 0.58	3.16 ± 3.77
Mean Teacher [23]	65.92 ± 25.59	53.82 ± 24.74	1.02 ± 0.64	3.74 ± 3.29
Bounding Box [13]	77.02 ± 17.15	65.08 ± 17.92	0.89 ± 0.57	2.54 ± 5.20
$\mathcal{L}_{pce} + \mathcal{L}_{crf} + \text{SimPLe}$	79.71 ± 17.72	68.99 ± 18.84	0.74 ± 0.55	2.48 ± 1.93
$\mathcal{L}_{pce} + \mathcal{L}_{ctr} + \text{SimPLe}$	81.20 ± 13.28	70.01 ± 15.02	0.69 ± 0.44	2.40 ± 1.69
Fully Supervision	81.52 ± 19.40	72.10 ± 20.45	0.68 ± 0.63	2.40 ± 2.76

Figure 4 visualizes our cancer segmentation results on the testing data. Table 1 reports the quantitative Dice, Jaccard, average surface distance (ASD), and Hausdorff distance (95HD) results of different methods. We compared our method with an end-to-end approach [4] that proposed to optimize network via CRF-regularized loss \mathcal{L}_{crf} . Although our \mathcal{L}_{ctr} supervised method outperformed \mathcal{L}_{crf} [4], the networks trained only using the initial pseudo-masks could not achieve enough high accuracy (Dice values < 70%). In contrast, the proposed SimPLe largely boosted the performance of the basically trained networks, by +14.74% Dice and +15.16% Jaccard (*v.s.* \mathcal{L}_{crf}), +11.81% Dice and +12.65% Jaccard (*v.s.* \mathcal{L}_{ctr}). Table 1 also shows the comparison results of three general weakly-supervised strategies, including entropy minimization [10], mean teacher [23], and bounding box [13]. Our method consistently outperformed these strategies with respect to all evaluation metrics. Furthermore, our method achieved competitive Dice results compared with fully supervision, which again proves the efficacy of the proposed SimPLe strategy. Note that the average annotation time for extreme points and full masks were 31 s and 95 s per scan, respectively. Figure 5 visualizes the 3D distance map between the segmented surface and ground-truth. It can be observed that our SimPLe consistently enhanced the segmentation.

4 Conclusion

We introduce a simple yet effective weakly-supervised learning method for breast cancer segmentation in DCE-MRI. The primary attribute is to fully exploit the simple *train - fine-tune - retrain* process to optimize the segmentation network via only extreme point annotations. This is achieved by employing a similarity-aware propagation learning (SimPLe) strategy to update the pseudo-masks. Experimental results demonstrate the efficacy of the proposed SimPLe strategy for weakly-supervised segmentation.

Acknowledgements. This work was supported in part by the National Natural Science Foundation of China under Grants 62071305, 61701312 and 81971631, and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2022A1515011241.

References

1. Ashraf, A.B., Gavenonis, S.C., Daye, D., Mies, C., Rosen, M.A., Kontos, D.: A multichannel markov random field framework for tumor segmentation with an application to classification of gene expression-based breast cancer recurrence risk. *IEEE Trans. Med. Imaging* **32**(4), 637–648 (2012)
2. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) *MICCAI 2016. LNCS*, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49
3. Dijkstra, E.: A note on two problems in connexion with graphs. *Numerische Mathematik* **1**, 269–271 (1959)
4. Dorent, R., et al.: Scribble-based domain adaptation via co-segmentation. In: Martel, A.L., et al. (eds.) *MICCAI 2020. LNCS*, vol. 12261, pp. 479–489. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59710-8_47
5. Dorent, R., et al.: Inter extreme points geodesics for end-to-end weakly supervised image segmentation. In: de Bruijne, M., et al. (eds.) *MICCAI 2021. LNCS*, vol. 12902, pp. 615–624. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87196-3_57
6. Du, H., Dong, Q., Xu, Y., Liao, J.: Weakly-supervised 3D medical image segmentation using geometric prior and contrastive similarity. *arXiv preprint arXiv:2302.02125* (2023)
7. Gao, Y., Zhao, Y., Luo, X., Hu, X., Liang, C.: Dense encoder-decoder network based on two-level context enhanced residual attention mechanism for segmentation of breast tumors in magnetic resonance imaging. In: 2019 IEEE International Conference on Bioinformatics and Biomedicine, pp. 1123–1129. IEEE (2019)
8. Giaquinto, A.N., et al.: Breast cancer statistics, 2022. *CA Cancer J. Clin.* **72**(6), 524–541 (2022)
9. Grady, L.: Random walks for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(11), 1768–1783 (2006)
10. Grandvalet, Y., Bengio, Y.: Semi-supervised learning by entropy minimization. *Adv. Neural Inf. Process. Syst.* **17**, 1–8 (2004)

11. Gubern-Mérida, A., et al.: Automated localization of breast cancer in DCE-MRI. *Med. Image Anal.* **20**(1), 265–274 (2015)
12. Jiang, Y., Edwards, A.V., Newstead, G.M.: Artificial intelligence applied to breast MRI for improved diagnosis. *Radiology* **298**(1), 38–46 (2021)
13. Kervadec, H., Dolz, J., Wang, S., Granger, E., Ayed, I.B.: Bounding boxes for weakly supervised segmentation: global constraints get close to full supervision. In: *Medical Imaging with Deep Learning*, pp. 365–381 (2020)
14. Kim, J.Y., et al.: Kinetic heterogeneity of breast cancer determined using computer-aided diagnosis of preoperative MRI scans: relationship to distant metastasis-free survival. *Radiology* **295**(3), 517–526 (2020)
15. Lee, C.H., et al.: Breast cancer screening with imaging: recommendations from the society of breast imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer. *J. Am. Coll. Radiol.* **7**(1), 18–27 (2010)
16. Li, C., Sun, H., Liu, Z., Wang, M., Zheng, H., Wang, S.: Learning cross-modal deep representations for multi-modal MR image segmentation. In: Shen, D., et al. (eds.) *MICCAI 2019. LNCS*, vol. 11765, pp. 57–65. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32245-8_7
17. Mann, R.M., Cho, N., Moy, L.: Breast MRI: state of the art. *Radiology* **292**(3), 520–536 (2019)
18. Meng, X., et al.: Volume-awareness and outlier-suppression co-training for weakly-supervised MRI breast mass segmentation with partial annotations. *Knowl.-Based Syst.* **258**, 109988 (2022)
19. Militello, C., et al.: Semi-automated and interactive segmentation of contrast-enhancing masses on breast DCE-MRI using spatial fuzzy clustering. *Biomed. Signal Process. Control* **71**, 103113 (2022)
20. Rezaei, Z.: A review on image-based approaches for breast cancer detection, segmentation, and classification. *Expert Syst. Appl.* **182**, 115204 (2021)
21. Roth, H.R., Yang, D., Xu, Z., Wang, X., Xu, D.: Going to extremes: weakly supervised medical image segmentation. *Mach. Learn. Knowl. Extract.* **3**(2), 507–524 (2021)
22. Sheth, D., Giger, M.L.: Artificial intelligence in the interpretation of breast cancer on MRI. *J. Magn. Reson. Imaging* **51**(5), 1310–1324 (2020)
23. Tarvainen, A., Valpola, H.: Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results. *Adv. Neural Inf. Process. Syst.* **30**, 1–10 (2017)
24. Wang, H., Cao, J., Feng, J., Xie, Y., Yang, D., Chen, B.: Mixed 2D and 3D convolutional network with multi-scale context for lesion segmentation in breast DCE-MRI. *Biomed. Signal Process. Control* **68**, 102607 (2021)
25. Wang, S., et al.: Breast tumor segmentation in DCE-MRI with tumor sensitive synthesis. *IEEE Trans. Neural Netw. Learn. Syst.* **34**, 4990–5001 (2021)
26. Yushkevich, P.A., Piven, J., Cody Hazlett, H., Gimpel Smith, R., Ho, S., Gee, J.C., Gerig, G.: User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* **31**(3), 1116–1128 (2006)
27. Zeng, X., Huang, R., Zhong, Y., Xu, Z., Liu, Z., Wang, Y.: A reciprocal learning strategy for semisupervised medical image segmentation. *Med. Phys.* **50**(1), 163–177 (2023)
28. Zhang, J., Saha, A., Zhu, Z., Mazurowski, M.A.: Hierarchical convolutional neural networks for segmentation of breast tumors in MRI with application to radio-genomics. *IEEE Trans. Med. Imaging* **38**(2), 435–447 (2018)

29. Zheng, Y., Baloch, S., Englander, S., Schnall, M.D., Shen, D.: Segmentation and classification of breast tumor using dynamic contrast-enhanced MR images. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007. LNCS, vol. 4792, pp. 393–401. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-75759-7_48
30. Zhou, L., Wang, S., Sun, K., Zhou, T., Yan, F., Shen, D.: Three-dimensional affinity learning based multi-branch ensemble network for breast tumor segmentation in MRI. *Pattern Recogn.* **129**, 108723 (2022)