Tumor Segmentation with Heterogeneity Clustering in Non-contrast Breast MRI

Xinyu Xie[1,2], Luyi Han[3,6], Yonghao Li[2], Yaofei Duan[1], Yue Sun[1], Muzhen He[7], Tao Tan[1(✉)], Dinggang Shen[2,4,5(✉)]

[1] Faculty of Applied Sciences, Macao Polytechnic University, Macao 999078, China
taotan@mpu.edu.mo
[2] School of Biomedical Engineering, ShanghaiTech University & State Key Laboratory of Advanced Medical Materials and Devices, Shanghai 201210, China
dgshen@shanghaitech.edu.cn
[3] Department of Radiology and Nuclear Medicine, Radboud University Medical Centre, Nijmegen 6525GA, The Netherlands
[4] Shanghai Clinical Research and Trial Center, Shanghai 201210, China
[5] United Imaging Intelligence Co., Ltd., Shanghai 200230, China
[6] Department of Radiology, Netherlands Cancer Institute, Amsterdam 1066CX, The Netherlands
[7] Fuzhou University Affiliated Provincial Hospital, Fuzhou 350001, China

Abstract. Breast tumor segmentation in dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) achieves precise delineation of tumor boundaries and subregions by capturing rich tissue heterogeneity information. However, its reliance on contrast agents may cause adverse effects, and the acquisition of complete time-series data involves a complex process. In contrast, current non-contrast image segmentation methods suffer from insufficient accuracy due to the lack of explicit tissue heterogeneity information. To address these limitations, we propose an approach for tumor heterogeneity estimation and segmentation in non-contrast images. The core idea is to extract tissue heterogeneity information from DCE-MRI and transfer it to a non-contrast image segmentation network, achieving tumor segmentation accuracy comparable to DCE-MRI-based methods. Our approach uses a vector quantized-variational autoencoder (VQ-VAE)-based clustering model to transform images into heterogeneity maps, capturing structural features of tumor subregions. These maps serve as the ground truth for training. Then, a heterogeneity information prediction model (HIPM) estimates heterogeneity maps from non-contrast images. These features are utilized as prior information to guide the segmentation network, further improving segmentation accuracy. Experimental results demonstrate that the cluster compactness (CPN) and Davies-Bouldin index (DBN) of the clustering reach approximately 0.05 and 0.001, respectively, indicating high clustering accuracy. Our method provides intuitive visualization of tumor heterogeneity without the need for contrast agents and significantly improves segmentation accuracy, with Dice Similarity Coefficient (DSC), Positive Predictive Value (PPV), and Sensitivity (SEN) increased by 20% compared to other non-contrast image segmentation networks.

Keywords: Intra-tumor heterogeneity · Tumor segmentation · Non-contrast breast imaging.

1 Introduction

Breast cancer is one of the most common malignant tumors in women, and its heterogeneity is reflected in aspects such as tumor size, shape, composition, and biological behavior, which presents significant challenges for early diagnosis and precision treatment [9]. In breast imaging diagnosis, dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) and non-contrast magnetic resonance imaging (non-contrast MRI) are two commonly used tools [4]. DCE-MRI enhances tissue perfusion and permeability using gadolinium (Gad) contrast agents, highlighting high-intensity signals in tumor regions, revealing dynamic blood flow characteristics, and providing detailed information on tumor microvasculature and tissue structure [21]. This technique includes images acquired at multiple time points, providing dynamic data on how the tumor evolves over time, with the time-intensity curve (TIC) being one of the most important analysis tools in DCE-MRI. TIC [13] describes changes in signal intensity over time in the tumor region and reflects heterogeneity characteristics of the tissue, such as blood flow velocity and vascular permeability, which helps differentiate between benign and malignant tumors.

However, gadolinium contrast agents can cause adverse reactions such as nephrogenic systemic fibrosis and allergic reactions, limiting their widespread use [3]. In contrast, non-contrast MRI eliminates these side effects by requiring no contrast agents, while still providing essential anatomical information about the tumor, including tumor boundaries, size, shape, and relationship with surrounding tissues [2, 17]. This is crucial for the early detection and accurate localization of breast cancer [11]. Nonetheless, non-contrast MRI faces challenges in accurately assessing tumor malignancy, vascularity, and invasiveness. This is particularly evident in early breast cancer screening, where it shows limited sensitivity to small or low-vascular tumors [7].

In recent years, deep learning has made significant progress in breast cancer MRI segmentation research. The widely used Unet has been extensively applied to breast tumor segmentation [14, 23]. These methods primarily focus on extracting structural features from non-contrast images while neglecting the exploration of tumor heterogeneity in non-contrast enhanced images. Some studies have employed automatic segmentation of DCE-MRI to assist in tumor region identification [7, 21, 24, 25]. However, these methods typically rely on complete DCE-MRI sequences [5], while the high dimensionality of the data poses challenges in sequence feature modeling, with tumor segmentation models consuming substantial computational resources [18]. Although some studies have attempted to generate post-contrast images from non-contrast images to reconstruct missing time-point images in DCE-MRI [11, 3, 22], using methods such as generative adversarial networks (GANs) or denoising diffusion probabilistic models (DDPMs) [16], pixel-level image restoration is highly challenging and unstable.

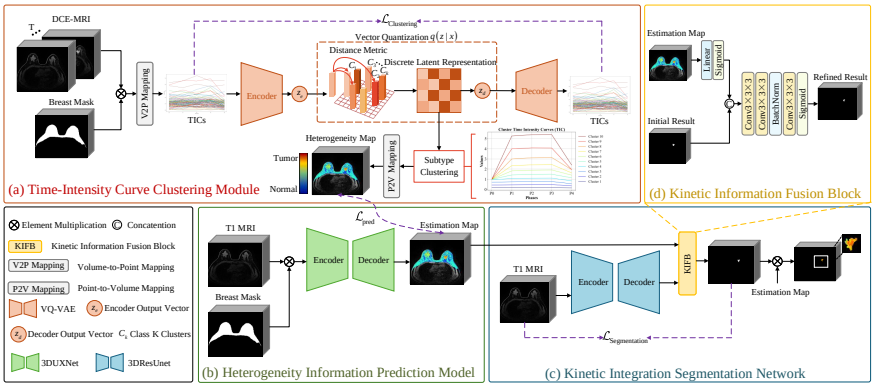


Fig. 1. Overview of proposed tumor segmentation with heterogeneity clustering models. The network operates in three phases: (a) generating a tumor heterogeneity map via time-intensity curve clustering; (b) estimating tumor heterogeneity from non-contrast images using a predictive network; and (c) combining structural details with the heterogeneity map for precise tumor segmentation and visualization. The first phase is excluded during testing.

We propose a method combining heterogeneity estimation and tumor segmentation to enhance breast cancer diagnosis accuracy and visualize tumor heterogeneity using non-contrast images. First, a TIC clustering method based on VQ-VAE is designed to generate heterogeneity maps, capturing and visualizing tumor heterogeneity from DCE-MRI. Next, a prediction network estimates heterogeneity information directly from non-contrast MRI, eliminating the need for full DCE-MRI sequences. Finally, a kinetic integration segmentation network merges structural features from non-contrast images with synthesized heterogeneity features, leveraging spatial-kinetic complementarity to improve segmentation accuracy. This approach avoids contrast agents, simplifies key information extraction, and improves clinical feasibility. Experimental results show improved segmentation performance and clear tumor heterogeneity visualization, aiding clinical diagnosis and personalized treatment.

2 Method

The proposed framework consists of a time-intensity curve clustering module (TIC-Clustering), a heterogeneity information prediction model (HIPM), and a Kinetic Integration Segmentation Network (KIS-Net). First, we propose an unsupervised clustering model based on VQ-VAE, which utilizes TIC information from DCE-MRI to divide the tumor into multiple biologically distinct subregions, revealing its internal heterogeneity. Next, we employ a 3D U-Net variant (3DUXNet) [8] to predict TIC information from the no-enhanced images, supplementing heterogeneity features. Finally, we design a kinetic integration seg-

mentation network that combines TIC predictions with structural information from non-contrast MRI to achieve more precise tumor segmentation.

2.1 Time-Intensity Curve Clustering Module

To reduce the difficulty of obtaining heterogeneity features of lesions from non-contrast images, we propose a clustering method based on TIC from DCE-MRI. In the first step of this module, we refer to [10] and construct a point set of input points based on the sequence I_t through Volume-to-Point mapping. As shown in Fig. 1(a), for each voxel v, its TIC is constructed based on the signal intensity S_t over time, represented as $TIC_v = \{S_0, S_1, \ldots, S_T\}$. To eliminate signal discrepancies among individuals, the TIC is normalized as $\hat{S}_t = \frac{S_t - \min(S)}{\max(S) - \min(S)}$. The normalized TIC is then fed into the VQ-VAE encoder, which outputs the low-dimensional feature representation z_e. Subsequently, VQ-VAE quantizes z_e using a fixed-size learnable codebook $C = \{c_1, c_2, \ldots, c_K\}$, mapping it to the closest cluster center $z_q = c_k$. Here, $k = \arg\min_j \|z_e - c_j\|_2$ is the cluster index, and c_k is the cluster center in the codebook. The training objective of VQ-VAE is to minimize both the reconstruction loss and the quantization loss, ensuring that the TIC after encoding, quantization, and decoding is accurately reconstructed, and that the quantized features align with the codebook centers. The clustering loss function $\mathcal{L}_{\text{Clustering}}$ is defined as follows:

$$\mathcal{L}_{\text{Clustering}} = \mathcal{L}_{rec} + \beta \mathcal{L}_{vq} = \|TIC - \hat{TIC}\|_2^2 + \beta \|\text{sg}[z] - c_k\|_2^2 \tag{1}$$

$\mathcal{L}_{rec}$ means the reconstruction loss, and $\mathcal{L}_{vq}$ means the quantization loss. $\text{sg}[\cdot]$ represents the stop-gradient operation, which prevents gradient updates from affecting the codebook. Finally, the TIC of each voxel is assigned a clustering index,which generates multiple heterogeneity maps from the original breast image. β is a balance hyperparameter that controls $\mathcal{L}_{\text{rec}}$ and $\mathcal{L}_{\text{vq}}$ to $\mathcal{L}_{\text{Clustering}}$.

2.2 Heterogeneity Information Prediction Model

In the second step of our framework, the goal is to predict heterogeneity information solely from the T1 map, similar to that obtained from DCE-MRI. To achieve this, we employ the well-established 3DUXNet [8], a network known for its exceptional performance in image prediction and generation tasks. During training, we use the clustering results from the previous step as supervision labels. Specifically, the prediction loss consists of Dice loss and Weighted Cross-Entropy loss, defined as:

$$\mathcal{L}_{\text{pred}} = \mathcal{L}_{\text{Dice}} + \mathcal{L}_{\text{WCE}} \tag{2}$$

where $\mathcal{L}_{\text{Dice}}$ represents the Dice loss, which measures the overlap between the predicted and ground truth segmentation, and $\mathcal{L}_{\text{WCE}}$ denotes the Weighted Cross-Entropy loss, which addresses class imbalance by assigning higher weights to

underrepresented classes. The $\mathcal{L}_{\text{Dice}}$ is formulated as:

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2\sum_i \hat{y}_i y_i}{\sum_i \hat{y}_i + \sum_i y_i} \tag{3}$$

and the $\mathcal{L}_{\text{WCE}}$ is defined as:

$$\mathcal{L}_{\text{WCE}} = -\sum_i w_i \left(y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)\right) \tag{4}$$

where $\hat{y}_i$ and y_i represent the predicted probability and ground truth label for the i-th voxel, respectively, and w_i is the weight assigned to the i-th voxel, which is inversely proportional to the frequency of the class to address class imbalance.

2.3 Kinetic Integration Segmentation Network

In the third stage, we propose a segmentation network that integrates the predicted heterogeneity maps and non-contrast T1-weighted breast MRI images. The multiple heterogeneity maps contain values quantifying tumor malignancy and provide heterogeneity prior. To effectively combine the heterogeneity information and spatial information of tumors, we designed a Kinetic Information Fusion Block (KIFB), as shown in Fig. 1(d). Specifically, the non-contrast T1 images are processed through a convolutional network to extract structural features. Meanwhile, this module embeds the TIC heterogeneity prior derived from non-contrast images into high-dimensional vectors and processes them through convolutional layers and a Sigmoid activation function. The two information streams are fused via concatenation, followed by multiple layers of 3×3×3 convolutions, a Batch Normalization (BN) layer, and a final Sigmoid activation function to output the segmentation result. The entire framework is based on ResUnet [19] as the baseline network, optimizing information flow and gradient propagation. Segmentation loss $\mathcal{L}_{\text{seg}}$ is defined as:

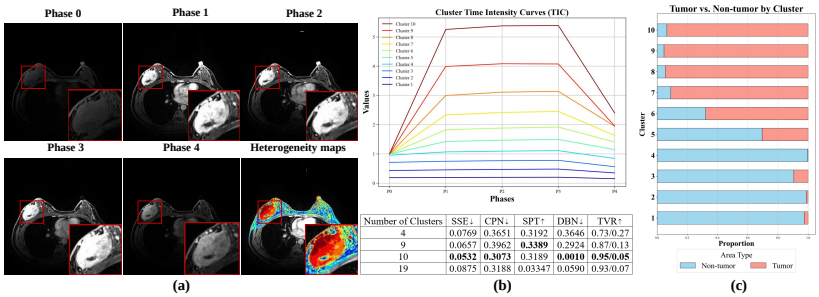
$$\begin{aligned} \mathcal{L}_{\text{seg}} &= \mathcal{L}_{\text{Dice}} + \lambda \mathcal{L}_{\text{BCE}} \\ &= \left(1 - \frac{2\sum_i \hat{y}_i y_i}{\sum_i \hat{y}_i + \sum_i y_i}\right) - \lambda \frac{1}{N} \sum_i \left(y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)\right) \end{aligned} \tag{5}$$

Here, $\hat{y}_i$ represents the predicted probability for the i-th pixel, y_i is the ground truth label (0 or 1), and N is the total number of pixels. λ is a balancing hyperparameter used to control contributions of $\mathcal{L}_{\text{Dice}}$ and $\mathcal{L}_{\text{BCE}}$ to $\mathcal{L}_{\text{segmentation}}$.

3 Experiments

Dataset The comprehensive dataset includes 163 cases from the ISPY1 dataset , 922 cases from the DUKE dataset , and 300 cases from Fuzhou Hospital dataset.

`https://www.cancerimagingarchive.net/collection/ispy1/`
`https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=70226903`



Number of Clusters	SSE↓	CPN↓	SPT↑	DBN↓	TVR↑
4	0.0769	0.3651	0.3192	0.3646	0.73/0.27
9	0.0657	0.3962	**0.3389**	0.2924	0.87/0.13
10	**0.0532**	**0.3073**	0.3189	**0.0010**	**0.95/0.05**
19	0.0875	0.3188	0.03347	0.0590	0.93/0.07

Fig. 2. (a) Visualization of clustering results obtained by the proposed method. (b) Average curves of different clusters. (c) Distribution of voxel ratios.

Each case contains multi-phase DCE-MRI images and T1-weighted images. The dataset is split into training and testing samples in a 8:2 ratio. The cases from ISPY1 and In-house have been meticulously voxel-labeled by radiology experts, and the manual segmentation results, carefully reviewed, serve as ground truth. In contrast, we use "Box-to-Mask" on the DUKE dataset to generate segmentation labels, which inevitably include non-tumor regions. Therefore, we only use this dataset in the first two stages of the network. Additionally, we use nnU-Net [6] to generate full-breast masks for all datasets.

3.1 Experimental Setup

Implementation To conduct experiments and train the networks, we utilize the PyTorch platform and run all experiments on a single NVIDIA RTX A40 GPU with 48 GB memory. For the clustering model network, the initial learning rate is set to 0.005, which decays by 50% every 50 epochs, with a total of 300 epochs. For the TIC prediction network, we adopt the 3DUXNet model as the baseline, with an initial learning rate of 0.001, and the total number of epochs as 1000. For the TIC-guided segmentation network, the initial learning rate is set to 0.005 and the total number of epochs is set to 300. We conduct extensive experiments to evaluate the performance of the proposed method and compare it with state-of-the-art (SOTA) methods. For the hyperparameters, the value of β is set to 0.25, and the value of λ is set to 5, as adopted in [12, 24]. The code will be published at `https://github.com/millieXie/HCNet`.

Comparative Methods and Evaluation Metrics We evaluate clustering performance using four key metrics [10]; intra-cluster compactness (CPN), inter-cluster separability (SPT), Davies-Bouldin index (DBN) and the max of TIC cluster voxel Ratio (TVR). To evaluate the second-stage TIC prediction, we use Mean Accuracy, and Mean Intersection over Union (Mean IoU) [8]. For a comprehensive evaluation of the proposed segmentation method, we compare it with baseline models using the following metrics [15]: Dice Similarity Coefficient

(DSC) to quantify the overlap between manually and automatically segmented label maps; Positive Predictive Value (PPV) and Sensitivity (SEN) to measure segmentation precision and recall; and Average Surface Distance (ASD).

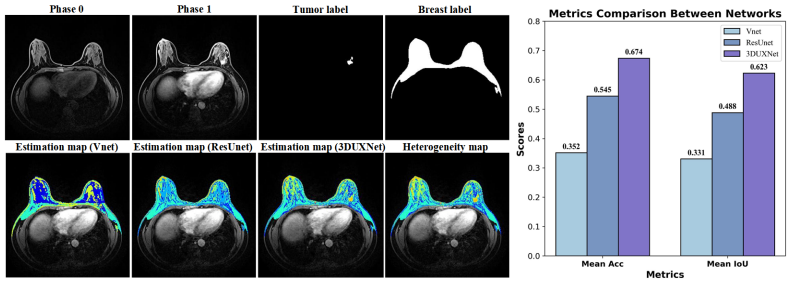


Fig. 3. The left image displays breast tumor images and their masks at different phases (top row), along with a qualitative comparison of tumor heterogeneity maps from various prediction networks (bottom row). The right image presents a quantitative comparison of these networks.

3.2 Experimental Results

TIC Clustering performance analysis In the first stage, we extract TICs from the complete DCE-MRI time series and perform cluster analysis on the tumor regions. As shown in Fig. 2(a), the clustering visualization effectively distinguishes between tumor and non-tumor regions. Fig. 2(b) displays the average curves of different clusters, and combined with Fig. 2(a), the heterogeneity characteristics of the tumor are clearly observed. Additionally, Fig. 2(c) presents the voxel ratio distribution of each cluster, significantly reflecting the differences in cluster category proportions between tumor and non-tumor regions, further validating the effectiveness of the proposed clustering method. To determine the optimal number of clusters (K), we conduct an ablation experiment. The results show that when K=10, metrics such as SSE (0.0532), CPN (0.3073), DBN (0.0010), and TVR (0.95/0.05) achieve optimal values, leading to set K to 10. Clinical expert evaluation further confirms that the clustering results effectively distinguish tumor regions of different malignancy levels.

Heterogeneity prediction performance analysis As shown in Fig. 3, we present the pre-image (Phase 0), post-image (Phase 1), tumor label (Tumor label), breast label (Breast label), tumor heterogeneity map, and prediction result from various networks (including VNet [12], ResUnet [19], and 3DUXNet [8]). It can be observed that VNet's prediction performance is poor, with some regions, especially tumor edges, appearing blurred. In contrast, ResUnet shows improved prediction results, particularly in the tumor edge regions, demonstrating higher

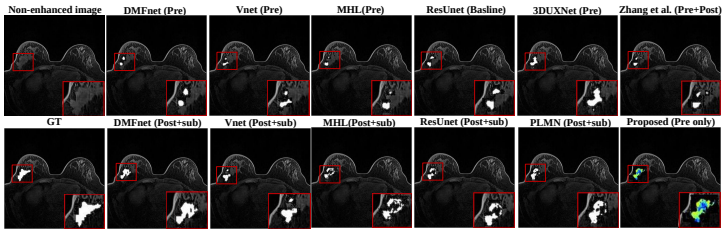


Fig. 4. Qualitative analysis of segmentation results. The first row shows results using only pre-contrast images as input. The second row presents results with post-contrast and subtraction-enhanced images as input, except our network.

Table 1. Visualization of the final segmentation results compared with other methods.

Method	Pre-C	Post-C	Sub	DSC↑	PPV↑	SEN↑	ASD↓
Zhang et al. [7]	✓	✓		0.5616	0.6237	0.6454	4.8751
ALMN [24]		✓	✓	0.5518	0.5803	0.6148	4.7572
PLMN [25]		✓	✓	0.6265	0.6131	0.7315	4.6446
MHL [20]		✓	✓	0.6295	0.6385	0.7104	4.7541
VNet [12]		✓	✓	0.5789	0.5991	0.7204	4.9872
3DUXNet [20]		✓	✓	0.5815	0.5383	0.7185	8.8854
DMFNet [1]		✓	✓	**0.6389**	0.6469	**0.7421**	4.2885
ResUnet (Baseline) [19]		✓	✓	0.6372	**0.6469**	0.7235	**2.7667**
MHL [20]	✓			0.3624	0.5454	0.2991	53.1351
Vnet [12]	✓			0.4182	0.4725	0.5243	60.4957
3DUXNet [8]	✓			0.3985	0.3935	0.4525	20.0793
DMFNet [1]	✓			0.4452	0.4649	0.4913	17.6689
ResUnet (Baseline) [19]	✓			0.4629	0.5813	0.5232	39.0312
Proposed (pre+TIC prediction)	✓			**0.6070**	**0.7005**	**0.7101**	**2.2941**

accuracy, though significant differences still exist in some areas. 3DUXNet excels in detail processing, and despite the highly challenging nature of the task, it achieves optimal values in metrics such as mean accuracy (0.6740) and mean IoU (0.6228). Given the low resolution of non-contrast images, noise interference, and the subtle heterogeneity of tumor regions, accurately predicting TIC information and achieving high-precision segmentation is undoubtedly an extremely challenging task.

Kinetic integration segmentation performance analysis Table 1 presents a comparison of the proposed method with state-of-the-art approaches. Our method (Ours) performs excellently across multiple metrics, achieving the best results in PPV (0.7005) and ASD (2.2941), demonstrating significant advantages in prediction accuracy and boundary fitting. Our method outperforms the baseline network using only non-contrast images in all segmentation metrics. Compared to networks that use both post-contrast images, the proposed method also shows improvement. Notably, our approach relies solely on non-contrast images, offering effective tumor heterogeneity visualization, particularly in resource-limited clinical settings.

4 Conclusion and Discussion

In this work, we propose a method for non-contrast breast MRI that integrates heterogeneity parameter estimation and tumor segmentation. Our approach predicts heterogeneity parameter distributions from non-contrast images to guide segmentation, offering clinicians intuitive insights into tumor heterogeneity. Evaluated on generalized datasets and compared with SOTA methods, our method achieves superior performance, obtaining results comparable to DCE-MRI while being safer and more cost-effective. However, challenges remain in handling complex tumor regions due to significant heterogeneity and individual variability, indicating potential for further refinement. Key advantages of our method include: 1) leveraging potential heterogeneity information in non-contrast images to enhance segmentation, and 2) improving segmentation accuracy through predicted heterogeneity maps.

Acknowledgments

This work was supported in part by National Natural Science Foundation of China (grant numbers 82441023, U23A20295, 62131015, 82394432), the China Ministry of Science and Technology (S20240085, STI2030-Major Projects-2022 ZD0209000, STI2030-Major Projects-2022ZD0213100), Shanghai Municipal Central Guided Local Science and Technology Development Fund (No. YDZX2023 3100001001), The Key R&D Program of Guangdong Province, China (grant number 2023B0303040001), HPC Platform of ShanghaiTech University, Macao Polytechnic University Grant (RP/FCA-08/2024), and the Science and Technology Development Fund of Macao (0105/2022/A).

Disclosure of Interests

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Chen, C., Liu, X., Ding, M., Zheng, J., Li, J.: 3d dilated multi-fiber network for real-time brain tumor segmentation in mri. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part III 22. pp. 184–192. Springer (2019)
2. Fan, J., Cao, X., Wang, Q., Yap, P.T., Shen, D.: Adversarial learning for mono-or multi-modal registration. Medical image analysis **58**, 101545 (2019)
3. Han, L., Tan, T., Zhang, T., Huang, Y., Wang, X., Gao, Y., Teuwen, J., Mann, R.: Synthesis-based imaging-differentiation representation learning for multi-sequence 3d/4d mri. Medical Image Analysis **92**, 103044 (2024)

4. Han, L., Tan, T., Zhang, T., Wang, X., Gao, Y., Lu, C., Liang, X., Dou, H., Huang, Y., Mann, R.: Non-adversarial learning: Vector-quantized common latent space for multi-sequence mri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 481–491. Springer (2024)
5. He, J., Zhao, X., Luo, Z., Su, S., Li, S., Zhang, G.: Tsesnet: Temporal-spatial enhanced breast tumor segmentation in dce-mri using feature perception and separability. In: Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence. pp. 803–811 (2024)
6. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)
7. Jiadong, Z., Cui, Z., Shi, Z., Jiang, Y., Zhang, Z., Dai, X., Yang, Z., et al.: A robust and efficient ai assistant for breast tumor segmentation from dce-mri via a spatial-temporal framework (2023)
8. Lee, H.H., Bao, S., Huo, Y., Landman, B.A.: 3d ux-net: A large kernel volumetric convnet modernizing hierarchical transformer for medical image segmentation. arXiv preprint arXiv:2209.15076 (2022)
9. Levy-Jurgenson, A., Tekpli, X., Kristensen, V.N., Yakhini, Z.: Spatial transcriptomics inferred from pathology whole-slide images links tumor heterogeneity to survival in breast and lung cancer. Scientific reports **10**(1), 18802 (2020)
10. Lv, T., Hong, X., Liu, Y., Miao, K., Sun, H., Li, L., Deng, C., Jiang, C., Pan, X.: Ai-powered interpretable imaging phenotypes noninvasively characterize tumor microenvironment associated with diverse molecular signatures and survival in breast cancer. Computer Methods and Programs in Biomedicine **243**, 107857 (2024)
11. Lv, T., Liu, Y., Miao, K., Li, L., Pan, X.: Diffusion kinetic model for breast cancer segmentation in incomplete dce-mri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 100–109. Springer (2023)
12. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D vision (3DV). pp. 565–571. Ieee (2016)
13. Nie, K., Baltzer, P., Preim, B., Mistelbauer, G.: Knowledge-assisted comparative assessment of breast cancer using dynamic contrast-enhanced magnetic resonance imaging. In: Computer Graphics Forum. vol. 39, pp. 13–23. Wiley Online Library (2020)
14. Piantadosi, G., Sansone, M., Sansone, C.: Breast segmentation in mri via u-net deep convolutional neural networks. In: 2018 24th international conference on pattern recognition (ICPR). pp. 3917–3922. IEEE (2018)
15. Qian, X., Pei, J., Han, C., Liang, Z., Zhang, G., Chen, N., Zheng, W., Meng, F., Yu, D., Chen, Y., et al.: A multimodal machine learning model for the stratification of breast cancer risk. Nature Biomedical Engineering pp. 1–15 (2024)
16. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695 (2022)
17. Xiang, L., Chen, Y., Chang, W., Zhan, Y., Lin, W., Wang, Q., Shen, D.: Deep-learning-based multi-modal fusion for fast mr reconstruction. IEEE Transactions on Biomedical Engineering **66**(7), 2105–2114 (2018)
18. Yeung, M., Sala, E., Schönlieb, C.B., Rundo, L.: Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. Computerized Medical Imaging and Graphics **95**, 102026 (2022)

19. Yu, L., Yang, X., Chen, H., Qin, J., Heng, P.A.: Volumetric convnets with mixed residual connections for automated prostate segmentation from 3d mr images. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 31 (2017)
20. Zhang, J., Saha, A., Zhu, Z., Mazurowski, M.A.: Hierarchical convolutional neural networks for segmentation of breast tumors in mri with application to radiogenomics. IEEE transactions on medical imaging **38**(2), 435–447 (2018)
21. Zhang, T., Tan, T., Han, L., Wang, X., Gao, Y., van Dijk, J., Portaluri, A., Gonzalez-Huete, A., D'Angelo, A., Lu, C., et al.: Important-net: integrated mri multi-parametric increment fusion generator with attention network for synthesizing absent data. Information Fusion **108**, 102381 (2024)
22. Zhang, Z., Han, L., Zhang, T., Lin, Z., Gao, Q., Tong, T., Sun, Y., Tan, T.: Unimrisegnet: Universal 3d network for various organs and cancers segmentation on multi-sequence mri. IEEE Journal of Biomedical and Health Informatics (2024)
23. Zhao, L., Wang, T., Chen, Y., Zhang, X., Tang, H., Zong, R., Tan, T., Chen, S., Tong, T.: Msabynet: A multiscale subtraction attention network framework based on bayesian loss for medical image segmentation. Biomedical Signal Processing and Control **103**, 107393 (2025)
24. 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 Recognition **129**, 108723 (2022)
25. Zhou, L., Zhang, Y., Zhang, J., Qian, X., Gong, C., Sun, K., Ding, Z., Wang, X., Li, Z., Liu, Z., et al.: Prototype learning guided hybrid network for breast tumor segmentation in dce-mri. IEEE Transactions on Medical Imaging (2024)