Adult-Like Phase and Multi-scale Assistance for Isointense Infant Brain Tissue Segmentation

Jiameng Liu[1], Feihong Liu[2,1], Kaicong Sun[1], Mianxin Liu[3], Yuhang Sun[1,4],
Yuyan Ge[1,5], and Dinggang Shen[1,6,7(✉)]

[1] School of Biomedical Engineering, ShanghaiTech University, Shanghai, China
dgshen@shanghaitech.edu.cn
[2] School of Information Science and Technology, Northwest University, Xi'an, China
[3] Shanghai Artificial Intelligence Laboratory, Shanghai, China
[4] School of Biomedical Engineering, Southern Medical University, Guangzhou, China
[5] Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an,
China
[6] Shanghai United Imaging Intelligence Co. Ltd., Shanghai, China
[7] Shanghai Clinical Research and Trial Center, Shanghai, China

Abstract. Precise brain tissue segmentation is crucial for infant development tracking and early brain disorder diagnosis. However, it remains challenging to automatically segment the brain tissues of a 6-month-old infant (isointense phase), even for manual labeling, due to inherent ongoing myelination during the first postnatal year. The intensity contrast between gray matter and white matter is extremely low in isointense MRI data. To resolve this problem, in this study, we propose a novel network with multi-phase data and multi-scale assistance to accurately segment the brain tissues of the isointense phase. Specifically, our framework consists of two main modules, *i.e.*, semantics-preserved generative adversarial network (SPGAN) and Transformer-based multi-scale segmentation network (TMSN). SPAGN bi-directionally transfers the brain appearance between the isointense phase and the adult-like phase. On the one hand, the synthesized isointense phase data augments the isointense dataset. On the other hand, the synthesized adult-like images provide prior knowledge to the ambiguous tissue boundaries in the paired isointense phase data. TMSN integrates features of multi-phase image pairs in a multi-scale manner, which exploits both the adult-like phase data, with much clearer boundaries as structural prior, and the surrounding tissues, with a larger receptive field to assist the isointense data tissue segmentation. Extensive experiments on the public dataset show that our proposed framework achieves significant improvement over the state-of-the-art methods quantitatively and qualitatively.

Keywords: Isointense infant tissue segmentation · Semantic-preserved GAN · Anatomical guidance segmentation

---

J. Liu and F. Liu—These authors contributed equally to this work.

1   Introduction

Charting brain development during the first postnatal year is crucial for identifying typical and atypical changes of brain tissues, *i.e.*, grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF), which can be utilized for diagnosis of autism and other brain disorders [16]. Early identification of autism is very important for effective intervention, and currently, an accurate diagnosis can be achieved as early as 12 months old [21]. Advancing diagnosis at an earlier stage could provide more time for greatly improving intervention [17]. Unfortunately, it remains challenging in obtaining a precise tissue map from the early infancy structural MRI (sMRI) data, such as T1 or (and) T2.

The acquisition of a mass of tissue maps relies on automatic segmentation techniques. However, the accuracy cannot be guaranteed, and the most difficult case is the segmentation of the isointense phase (6-9 months) data [18], as shown in Fig. 1, where the intensity distributions of GM and WM are highly overlapped compared with the adult-like phase ($\geq$ 9 months) [18] (clear boundaries between GM and WM). Such a peculiarity incurs two challenges to conventional segmentation methods: *i*) lack of well-annotated data for training automatic segmentation algorithms; and *ii*) GM and WM demonstrate a low contrast result in limited anatomical information for accurately distinguishing GM and WM, even with enough high-quality annotations as training data.

Conventional methods can mainly be classified into two categories: 1) the registration-based methods and 2) the learning-based methods, as introduced below. The registration-based methods usually utilize a single previous-defined atlas, for cross-sectional data, or a sequence of atlases, for longitudinal data-guided methods, to indirectly obtain the tissue maps [13,14,19,20]. Those methods require a substantial number of atlases and follow-up adult-like images to guide the segmentation process, which is known for low accuracy in segmenting infant tissues due to the rapid developmental changes in early life. The learning-based methods have gained significant prominence for individualized segmentation [8,22], which is also exploited to segment isointense brain tissues. For example, Nie et al. [9] used fully convolutional networks (FCNs) to directly segment tissue from MR images, and Wang et al. [23] employed an attention mechanism for better isointense infant brain segmentation. Also, Bui et al. proposed a 3D CycleGAN [2] to utilize the isointense data to synthesize adult-like data, which can be employed to make the segmentation of isointense data more accurate.

In this work, we propose a novel Transformer-based framework for isointense tissue segmentation via T1-weighted MR images, which is composed of two stages: *i.e.*, *i*) the semantics-preserved GAN (SPGAN), and *ii*) Transformer-based multi-scale segmentation network (TMSN). Specifically, SPGAN is a bidirectional synthesis model that enables both the isointense data synthesis using adult-like data and vice verse. The isointense structural MRI data is paired with the segmented tissue maps for extending the training dataset. Additionally, the synthesized adult-like data from isointense infant data are adopted to assist segmentation in TMSN. TMSN incorporates a Transformer-based cross-branch
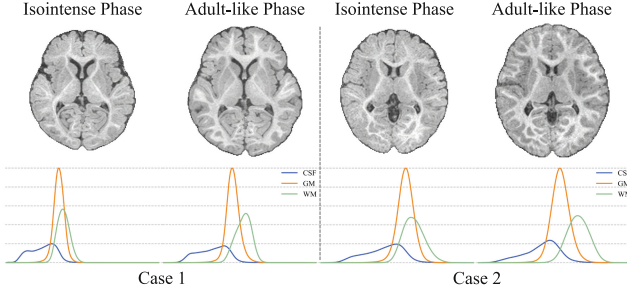
Fig. 1. T1-weighted images (the first row) and different tissue intensity distributions (the second row) of two typical infant brains scanned in isointense and adult-like phases, respectively

fusing (TCF) module which exploits supplementary tissue information from a patch with a larger receptive field to guide the local segmentation. Extensive experiments are conducted on the public dataset, National Database for Autism Research (NDAR) [5,12], and the results demonstrate that our proposed framework outperforms other state-of-the-art methods, particularly in accurately segmenting ambiguous tissue boundaries.

2   Methods

We propose a Transformer-based framework for isointense tissue segmentation, which includes two main modules: i) the semantic-preserved GAN (SPGAN), and ii) the Transformer-based multi-scale segmentation network (TMSN). An overview of the framework is provided in Fig. 2, SPGAN is designed to synthesize isointense or adult-like data to augment the training dataset and guide segmentation with anatomical information (Sect. 2.1). TMSN is designed for isointense tissue segmentation by utilizing the tissue information from the synthesized adult-like data (Sect. 2.2).

2.1   Semantics-Preserved Multi-phase Synthesis

Motivated by the effectiveness of classical unpaired image-to-image translation models [24,25], we propose a novel bi-directional generative model that synthesizes both isointense and adult-like data while introducing essential anatomical constraints to maintain structural consistency between the input images and the corresponding synthesized images. As shown in Fig. 2 (b.1), the SPGAN consists of two generators, G_{A2I} and G_{I2A}, along with their corresponding discriminators D_{A2I} and D_{I2A}. The adult-like infant images are aligned to the isointense infant images using affine registration methods [1], before training SPGAN.
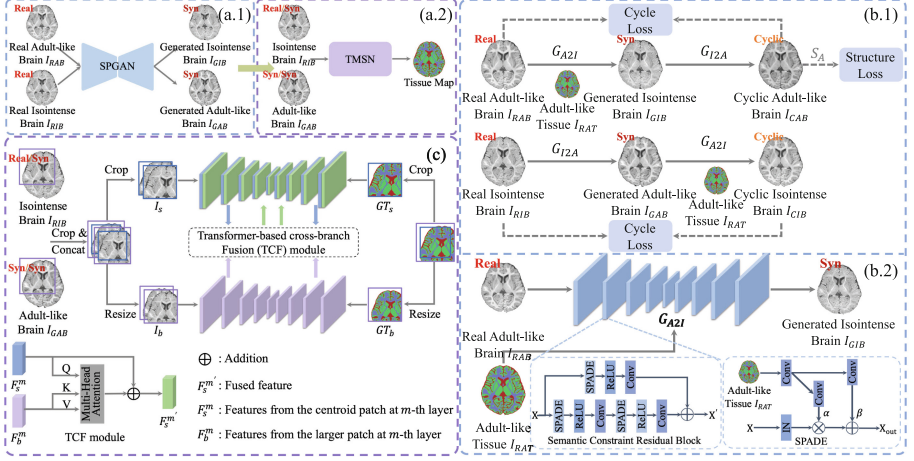
Fig. 2. (a) An overview of our proposed framework including: *i*) the semantics-preserved GAN (SPGAN), and *ii*) Transformer-based multi-scale segmentation (TMSN) network; (b) Detailed SPGAN architecture with the discriminator omitted for making the figure concise; (c) Detailed architecture of TMSN.

Isointense Phase Synthesis. The generator G_{A2I} is designed to synthesize isointense data from adult-like ones, which is constructed based on the 3D UNet [3] model, consisting of 4 encoder layers and 4 decoder layers. Inspired by GauGAN [10], we incorporate the spatial adaptive denormalization (SPADE) into G_{A2I} to maintain structural consistency between the synthesized isointense brain images and corresponding adult-like brain images, as shown in Fig. 2 (b.2). SPADE module takes the adult-like tissue maps to perform feature normalization during synthesis stage of isointense phase data. Thus, the generated isointense brain images can obtain consistent tissue structures with adult-like brain images. Specifically, X_{in} with a spatial resolution of $(H \times W \times D)$ represents the input feature map to each layer, consisting of a batch of N samples and C channels. For an arbitrary n-th sample at the c-th channel, the activation value associated with location (i,j,k) is denoted as $x_{n,c,i,j,k}$. The SPADE operator adopts two convolution blocks to learn the modulation maps $\alpha(M)$ and $\beta(M)$ for a given tissue map M, where the first convolution block makes the number of channels in the tissue map consistent with the number of channels in the feature map. The modulated value at (i,j,k) in each feature map is formulated as:

$$x'_{n,c,i,j,k} = \alpha(M)_{n,c,i,j,k} \frac{x_{n,c,i,j,k} - \mu_c}{\sigma_c} + \beta(M)_{n,c,i,j,k} \qquad (1)$$

where μ_c and σ_c denote the mean and standard deviation, respectively. Additionally, we adopt PatchGAN [25] as the discriminator D_{A2I} to provide adversary loss to train the generator G_{A2I}.

Adult-Like Phase Synthesis. The generator G_{I2A} is designed to synthesize adult-like infant brain images from isointense infant brain images, which is employed to provide clear structural information for identifying the ambiguous tissue boundaries in isointense brain images. To ensure the synthesized adult-like infant brain images can provide tissue information as realistic and accurate as the real images, we utilize a pre-trained adult-like brain tissue segmentation network S_A (3D UNet) to preserve the structural similarity between the synthesized adult-like data and the real adult-like data and promote more reasonable anatomical structures in the synthesized images. To achieve that goal, during the training of SPGAN, we freeze the parameters of S_A and adopt the mean square error (MSE) to penalize the dissimilarity between the tissue probability maps of the real and synthesized brain images (extracted by the pre-trained segmentation model S_A). The MSE loss is formulated as $\mathcal{L}_{MSE} = MSE(S_A(I_{RAB}), S_A(I_{CAB}))$ where I_{RAB} and I_{CAB} denote the real adult-like brain images and synthesized adult-like brain images synthesized by G_{I2A}. The overall loss function of SPGAN is defined as:

$$\mathcal{L}_{SPGAN} = \gamma_1 \mathcal{L}_{G_{I2A}} + \gamma_2 \mathcal{L}_{G_{A2I}} + \gamma_3 \mathcal{L}_{MSE} + \gamma_4 \mathcal{L}_{cycle} \tag{2}$$

where $\mathcal{L}_{cycle}$ denotes the cycle consistency loss between two generators.

2.2   Transformer-Based Multi-scale Segmentation

The overview of our proposed segmentation network (TMSN) is shown in Fig. 2 (c). In order to guide the network with anatomical prior, TMSN takes the pair of isointense and the corresponding adult-like images as input. The isointense and adult-like images are concatenated and cropped to image patches and then, they are fed into the top branch of TMSN. Due to the fact that neighboring tissues can provide additional information and greatly improve segmentation performance, hence, we also employ a multi-scale strategy by taking an image pair of the larger receptive field as the input of the bottom branch. It is important to note that the image pairs of the top and bottom branches should have the same sizes.

Moreover, we fuse the features of the two branches at the same block using a novel Transformer-based cross-branch fusion (TCF) module, where the Transformer can better capture relationships across two different branches (with local and global features). Specifically, F_s^m and F_b^m denote the learned features of the two branches at the m-th stage, respectively. The TCF module treats the F_s^m and F_b^m as the input of the encoder and decoder in the multi-head Transformer to learn the relationships between the centroid tissues in the top branch and the corresponding surrounding tissues in the bottom branch and then fuses the bottom branch features with the ones of the top branch by utilizing the learned relationships. The proposed TCF module can be formulated as:

$$F_s^{m'} = F_s^m + [Q(F_s^m).(K(F_b^m))^\top] * V(F_b^m) \tag{3}$$

where Q, K, and V denote the query, key, and value in the Transformer.

Table 1. Ablation analysis of our proposed framework. DA: data augmentation; SE: structural enhancement; SegNet: single-branch segmentation network; SegNetMS: multi-branch segmentation network based on channel concatenation; SegNetMSAtt: proposed Transformer-based multi-branch segmentation network.

Method	DA	SE	Dice (%) ↑			HD (mm) ↓			ASD ×10 (mm) ↓		
			WM	GM	CSF	WM	GM	CSF	WM	GM	CSF
SegNet			93.68 ± 0.55	92.45 ± 0.62	93.31 ± 0.76	5.44 ± 1.56	11.19 ± 1.86	8.69 ± 1.19	1.44 ± 0.17	1.17 ± 0.10	0.88 ± 0.09
SegNet	✓		94.31 ± 0.54	93.35 ± 0.60	94.50 ± 0.64	5.34 ± 1.27	11.26 ± 2.01	10.60 ± 2.40	1.30 ± 0.16	1.01 ± 0.09	0.70 ± 0.07
SegNet		✓	94.01 ± 0.61	93.28 ± 0.61	94.92 ± 0.58	7.40 ± 2.11	8.19 ± 2.62	10.15 ± 2.29	1.43 ± 0.19	1.00 ± 0.10	0.65 ± 0.06
SegNet	✓	✓	94.40 ± 0.56	93.83 ± 0.57	94.74 ± 0.62	5.44 ± 1.05	8.38 ± 2.51	10.04 ± 2.36	1.25 ± 0.15	0.93 ± 0.09	0.67 ± 0.06
SegNetMS			95.02 ± 0.41	94.50 ± 0.46	95.44 ± 0.48	6.51 ± 1.94	8.29 ± 2.36	7.95 ± 1.16	1.12 ± 0.13	0.84 ± 0.09	0.57 ± 0.05
SegNetMS	✓		95.38 ± 0.37	94.99 ± 0.44	95.79 ± 0.51	4.96 ± 1.17	8.87 ± 2.32	8.20 ± 1.75	1.04 ± 0.11	0.77 ± 0.08	0.53 ± 0.05
SegNetMS		✓	95.88 ± 0.41	95.40 ± 0.42	96.18 ± 0.43	4.86 ± 0.94	8.11 ± 2.79	7.87 ± 1.72	0.94 ± 0.12	0.70 ± 0.08	0.48 ± 0.04
SegNetMS	✓	✓	96.06 ± 0.40	95.62 ± 0.45	96.35 ± 0.48	5.14 ± 1.18	7.61 ± 2.37	8.22 ± 1.99	0.90 ± 0.11	0.66 ± 0.07	0.45 ± 0.05
SegNetMSAtt			95.68 ± 0.36	95.10 ± 0.47	95.85 ± 0.55	4.90 ± 0.97	8.44 ± 2.77	8.33 ± 1.35	1.00 ± 0.12	0.74 ± 0.07	0.53 ± 0.06
SegNetMSAtt	✓		95.82 ± 0.36	95.40 ± 0.46	96.23 ± 0.48	4.82 ± 0.78	$\mathbf{7.59 \pm 2.53}$	8.32 ± 1.63	0.95 ± 0.12	0.68 ± 0.07	0.47 ± 0.05
SegNetMSAtt		✓	96.21 ± 0.38	95.79 ± 0.42	96.50 ± 0.46	5.24 ± 0.95	8.39 ± 2.87	7.44 ± 1.22	0.86 ± 0.11	0.64 ± 0.07	0.43 ± 0.05
SegNetMSAtt	✓	✓	$\mathbf{96.31 \pm 0.36}$	$\mathbf{95.89 \pm 0.41}$	$\mathbf{96.60 \pm 0.42}$	$\mathbf{4.71 \pm 0.99}$	7.89 ± 2.65	$\mathbf{7.30 \pm 1.24}$	$\mathbf{0.84 \pm 0.11}$	$\mathbf{0.62 \pm 0.07}$	$\mathbf{0.42 \pm 0.04}$

We take a hybrid Dice and focal loss to supervise the two segmentation branches as follows:

$$\mathcal{L}_{Seg} = \lambda_1 \mathcal{L}_s(I_s, GT_s) + (1 - \lambda_1)\mathcal{L}_b(I_b, GT_b) \tag{4}$$

where I_s and GT_s, respectively, denote the input and ground truth (GT) of the top branch, and I_b and GT_b represent the ones of the bottom branch. The final tissue segmentation results are obtained from the top branch.

3   Experiments and Results

3.1   Dataset and Evaluation Metrics

We evaluated our proposed framework for isointense infant brain tissue segmentation on the public dataset NDAR [5,12]. The NDAR dataset comprises T1-weighted brain images of 331 cases at the isointense phase and 368 cases at the adult-like phase (12-month-old), where 180 cases contain both time points. The GT annotations for GM, WM, and CSF were manually labeled by experienced radiologists. The data was aligned to MNI space and normalized to standard distribution using z-score normalization with dimensions of 182, 218, and 182 in the x, y, and z axis, respectively. The dataset was randomly divided into three subsets, *i.e.*, 70% for training, 10% for validation, and 20% for testing. For quantitative evaluation, we employed three assessment metrics including Dice score, Hausdorff distance (HD), and average surface distance (ASD) [15].

3.2   Implementation Details

Our proposed framework is implemented based on PyTorch 1.7.1 [11] and trained on a workstation equipped with two NVIDIA V100s GPUs. We employ the Adam optimizer [7] with momentum of 0.99 and weight decay of 0.001. The learning

Table 2. Quantitative comparison with the state-of-the-art (SOTA) methods.

Method	Dice (%) ↑			HD (mm) ↓			ASD ×10 (mm) ↓		
	WM	GM	CSF	WM	GM	CSF	WM	GM	CSF
3D UNet [3]	93.68 ± 0.55	92.45 ± 0.62	93.32 ± 0.76	5.44 ± 1.57	11.19 ± 1.86	8.69 ± 1.19	1.44 ± 0.17	1.17 ± 0.10	0.88 ± 0.09
DualAtt [4]	95.21 ± 0.34	94.61 ± 0.41	95.51 ± 0.47	5.24 ± 1.22	10.96 ± 1.89	7.94 ± 1.42	1.07 ± 0.11	0.85 ± 0.08	0.56 ± 0.04
SwinUNETR [6]	94.99 ± 0.40	94.17 ± 0.48	94.88 ± 0.54	4.88 ± 1.14	$\mathbf{7.60 \pm 2.58}$	8.47 ± 1.82	1.13 ± 0.14	0.86 ± 0.08	0.64 ± 0.05
Ours	$\mathbf{96.31 \pm 0.36}$	$\mathbf{95.89 \pm 0.41}$	$\mathbf{96.60 \pm 0.42}$	$\mathbf{4.71 \pm 0.99}$	7.89 ± 2.65	$\mathbf{7.30 \pm 1.23}$	$\mathbf{0.84 \pm 0.11}$	$\mathbf{0.62 \pm 0.07}$	$\mathbf{0.42 \pm 0.04}$

rate is set as 0.001 and reduced to 90% of the current learning rate every 50 epochs for both SPGAN and TMSN. For SPGAN the hyperparameters γ_1, γ_2, γ_2, γ_4 are set to 10, 10, 1 and 1 in our experiments, respectively. For TMSN, the patch sizes of the top and bottom branches are set as $128 \times 128 \times 128$ and $160 \times 160 \times 160$, respectively. The total parameter number of SPGAN and TMSN is 44.51M. The hyperparameter λ_1 in TMSN is set to 0.6. The inference for one single image took about 8.82 s. Code is available at this link.

3.3   Evaluation and Discussion

We conduct extensive experiments to evaluate the effectiveness of our infant brain tissue segmentation framework. Table 1 summarizes the results of the ablation study and demonstrates the contribution of each component in our framework. Particularly, SegNet represents the baseline model using a single scale, based on which, SegNetMS involves an additional branch that captures a larger receptive field and employs channel concatenation to fuse the features of these two branches. SegNetMSAtt replaces feature concatenation with the Transformer-based TCF module. Each of the above three configurations contains four different settings, as listed in Table 1. The results demonstrate the benefits of using both isointense phase data augmentation and adult-like phase structural enhancement.

Effectiveness of Multi-phase Synthesis. In Table 1, we demonstrate the effectiveness of the synthesis of isointense and adult-like phase images by the proposed synthesis network SPGAN. When comparing with the methods based on limited samples and augmented samples, we witness that data augmentation (DA) effectively improves segmentation performance. Furthermore, by comparing with the methods with or without structural enhancement (SE), we observe that the generated adult-like infant brain images provide richer and more reliable structural information to guide the identification of ambiguous tissue boundaries at the isointense phase. When combining DA and SE, an improvement of the segmentation performance on all evaluating metrics emerge, which demonstrates that the DA and SE can be effectively integrated to bring significant benefits for improving tissue segmentation performance.

Effectiveness of Multi-scale Assistance. In order to demonstrate the effectiveness of the additional branch in TMSN, which contains a larger receptive
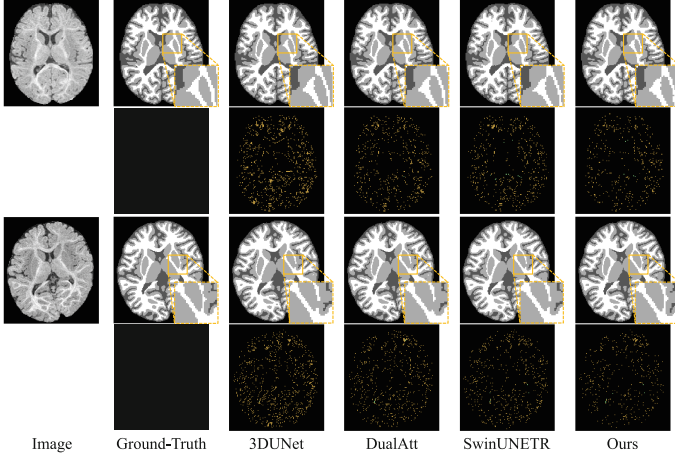
| | | | | | |
| Image | Ground-Truth | 3DUNet | DualAtt | SwinUNETR | Ours |

Fig. 3. Qualitative comparison with SOTA methods: two representative cases with corresponding error maps.

field, we show the corresponding results as one part of the ablation study in Table 1. Comparing the multi-scale SegNetMS with the single-scale variant Seg-Net, we can see that SegNetMS improves the Dice score from 94.32% to 96.01% and achieves a decrease in HD from 7.95 mm to 6.99 mm on average. This indicates that the multi-scale strategy enables the segmentation model to make use of the intrinsic relations between the centroid tissues and the corresponding counterparts with a larger receptive field, which leads to great improvement in the segmentation performance. In our proposed model SegNetMSAtt, we replace the channel concatenation in SegNetMS with TCF. We can find that the TCF module can more effectively build up the correspondence between top and bottom branches such that the surrounding tissues can well guide the segmentation.

Comparison with State-of-the-Arts (SOTAs). We conduct a comparative evaluation of our framework against several SOTA learning-based tissue segmentation networks, including 1) the 3D UNet segmentation network [3], 2) the DualAtt network [4] utilizing both spatial and channel attention, and 3) the SwinUNETR network [6] employing a 3D Swin Transformer. From the results listed in Table 2, we can find that our framework achieves the highest average Dice score and lower HD and ASD compared with the SOTA methods.

To further illustrate the advanced performance of our framework, we provide a visual comparison of two typical cases in Fig. 3. It can be seen that the predicted tissue maps (the last column) obtained by our method are much closer to the GT (the second column), especially in the subcortical regions. We can draw the same conclusion from the error maps (the second and the fourth rows in Fig. 3) between the segmented tissue maps obtained by each method and GT. Both the quantitative improvement and the superiority from the qualitative results

performance consistently show the effectiveness of our proposed segmentation framework, as well as the two main components, *i.e.*, SPGAN, and TMSN.

4   Conclusion

In this study, we have presented a novel Transformer-based framework to segment tissues from isointense T1-weighted MR images. We designed two modules, *i.e.*, *i*) semantic-preserved GAN (SPGAN), and *ii*) Transformer-based multi-scale segmentation network (TMSN). SPGAN is designed to synthesize both isointense and adult-like data, which augments the dataset and provides supplementary tissue constraint for assisting isointense tissue segmentation. TMSN is used to segment tissues from isointense data under the guidance of the synthesized adult-like data. Their advantages are distinctive. For example, SPGAN overcomes the lack of training samples and synthesizes adult-like infant brain images with clear structural information for enhancing ambiguous tissue boundaries. TMSN exploits the pair of isointense and adult-like phase images, as well as the multi-scale scheme, to provide more information for achieving accurate segmentation performance. Extensive experiments demonstrate that our proposed framework outperforms the SOTA approaches, which shows the potential in early brain development or abnormal brain development studies.

References

1. Avants, B.B., Tustison, N., Song, G., et al.: Advanced normalization tools (ants). Insight J. **2**(365), 1–35 (2009)
2. Bui, T.D., Wang, L., Lin, W., Li, G., Shen, D.: 6-month infant brain MRI segmentation guided by 24-month data using cycle-consistent adversarial networks. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), pp. 359–362. IEEE (2020)
3. 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
4. Fu, J., et al.: Dual attention network for scene segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3146–3154 (2019)
5. Hall, D., Huerta, M.F., McAuliffe, M.J., Farber, G.K.: Sharing heterogeneous data: the national database for autism research. Neuroinformatics **10**, 331–339 (2012)

6. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin UNETR: swin transformers for semantic segmentation of brain tumors in MRI images. In: Crimi, A., Bakas, S. (eds.) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2021. LNCS, vol. 12962. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-08999-2_22

7. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

8. Liu, J., et al.: Multi-scale segmentation network for Rib fracture classification from CT images. In: Lian, C., Cao, X., Rekik, I., Xu, X., Yan, P. (eds.) MLMI 2021. LNCS, vol. 12966, pp. 546–554. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87589-3_56

9. Nie, D., Wang, L., Gao, Y., Shen, D.: Fully convolutional networks for multi-modality isointense infant brain image segmentation. In: 2016 IEEE 13Th International Symposium on Biomedical Imaging (ISBI), pp. 1342–1345. IEEE (2016)

10. Park, T., Liu, M.Y., Wang, T.C., Zhu, J.Y.: GauGAN: semantic image synthesis with spatially adaptive normalization. In: ACM SIGGRAPH 2019 Real-Time Live!, p. 1 (2019)

11. Paszke, A., et al.: PyTorch: an imperative style, high-performance deep learning library. In: Advances in Neural Information Processing Systems 32 (2019)

12. Payakachat, N., Tilford, J.M., Ungar, W.J.: National Database for Autism Research (NDAR): big data opportunities for health services research and health technology assessment. Pharmacoeconomics **34**(2), 127–138 (2016)

13. Prastawa, M., Gilmore, J.H., Lin, W., Gerig, G.: Automatic segmentation of MR images of the developing newborn brain. Med. Image Anal. **9**(5), 457–466 (2005)

14. Shi, F., Yap, P.T., Fan, Y., Gilmore, J.H., Lin, W., Shen, D.: Construction of multi-region-multi-reference atlases for neonatal brain MRI segmentation. Neuroimage **51**(2), 684–693 (2010)

15. Taha, A.A., Hanbury, A.: Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med. Imaging **15**(1), 1–28 (2015)

16. Tierney, A.L., Nelson, C.A., III.: Brain development and the role of experience in the early years. Zero Three **30**(2), 9 (2009)

17. Wang, L., et al.: Anatomy-guided joint tissue segmentation and topological correction for 6-month infant brain MRI with risk of autism. Hum. Brain Mapp. **39**(6), 2609–2623 (2018)

18. Wang, L., et al.: Benchmark on automatic six-month-old infant brain segmentation algorithms: the iSeg-2017 challenge. IEEE Trans. Med. Imaging **38**(9), 2219–2230 (2019)

19. Wang, L., et al.: Integration of sparse multi-modality representation and anatomical constraint for isointense infant brain MR image segmentation. Neuroimage **89**, 152–164 (2014)

20. Wang, L., et al.: Segmentation of neonatal brain MR images using patch-driven level sets. Neuroimage **84**, 141–158 (2014)

21. Wang, L., Shi, F., Yap, P.T., Gilmore, J.H., Lin, W., Shen, D.: 4D multi-modality tissue segmentation of serial infant images (2012)

22. Wang, R., Lei, T., Cui, R., Zhang, B., Meng, H., Nandi, A.K.: Medical image segmentation using deep learning: a survey. IET Image Proc. **16**(5), 1243–1267 (2022)

23. Wang, Z., Zou, N., Shen, D., Ji, S.: Non-local U-Nets for biomedical image segmentation. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 6315–6322 (2020)

24. Yang, H., et al.: Unpaired brain MR-to-CT synthesis using a structure-constrained CycleGAN. In: Stoyanov, D., et al. (eds.) DLMIA/ML-CDS -2018. LNCS, vol. 11045, pp. 174–182. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00889-5_20

25. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2223–2232 (2017)