



Collaborative Modality Generation and Tissue Segmentation for Early-Developing Macaque Brain MR Images

Xueyang Wu^{1,2,3}, Tao Zhong^{1,2,3}, Shujun Liang^{1,2,3}, Li Wang⁴, Gang Li⁴,
and Yu Zhang^{1,2,3}(✉)

¹ School of Biomedical Engineering, Southern Medical University,
Guangzhou 510515, China

² Guangdong Provincial Key Laboratory of Medical Image Processing,
Southern Medical University, Guangzhou 510515, China

³ Guangdong Province Engineering Laboratory for Medical Imaging and Diagnostic
Technology, Southern Medical University, Guangzhou 510515, China
yuzhang@smu.edu.cn

⁴ Department of Radiology and BRIC, University of North Carolina at Chapel Hill,
Chapel Hill, USA

Abstract. In neuroscience research, automatic segmentation of macaque brain tissues in magnetic resonance imaging (MRI) is crucial for understanding brain structure and function during development and evolution. Acquisition of multimodal information is a key enabler of accurate tissue segmentation, especially in early-developing macaques with extremely low contrast and dynamic myelination. However, many MRI scans of early-developing macaques are acquired only in a single modality. While various generative adversarial networks (GAN) have been developed to impute missing modality data, current solutions treat modality generation and image segmentation as two independent tasks, neglecting their inherent relationship and mutual benefits. To address these issues, this study proposes a novel Collaborative Segmentation-Generation Framework (CSGF) that enables joint missing modality generation and tissue segmentation of macaque brain MR images. Specifically, the CSGF consists of a modality generation module (MGM) and a tissue segmentation module (TSM) that are trained jointly by a cross-module feature sharing (CFS) and transferring generated modality. The training of the MGM under the supervision of the TSM enforces anatomical feature consistency, while the TSM learns multi-modality information related to anatomical structures from both real and synthetic multi-modality MR images. Experiments show that the CSGF outperforms the conventional independent-task mode on an early-developing macaque MRI dataset with 155 scans, achieving superior quality in both missing modality generation and tissue segmentation.

Keywords: Tissue segmentation · Missing-modality generation · Macaque · Early brain development · Multi-task collaboration

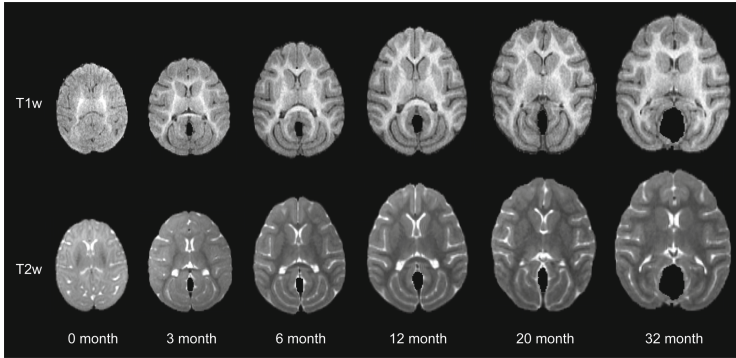


Fig. 1. T1w and T2w brain MR image of macaques at different ages.

1 Introduction

The precise processing and analysis of brain MR images are critical in advancing neuroscience research. As a widely used animal model with high systematic similarity to humans in various aspects, macaques play an indispensable role in understanding brain mechanisms in development, aging and evolution, exploring the pathogenesis of neurological diseases, and validating the effectiveness of clinical techniques and drugs [1]. In both humans and macaques, early brain development is a complex, dynamic, and regionally heterogeneous process that plays a crucial role in shaping later brain structural, functional, and cognitive outcomes [19]. While the human neuroimaging research is advancing rapidly, the macaque neuroimaging research lags behind, partly due to challenges in data acquisition, lack of tailored processing tools, and the inapplicability of human neuroimaging analysis tools to macaques [2]. Therefore, the macaque neuroimaging research is a rapidly evolving field that demands specialized tools and expertise. Of particular importance is brain tissue segmentation, a crucial prerequisite for quantitative volumetric analysis and surface-based studies, which aim to partition the brain into distinct regions such as white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF) [3]. As illustrated in Fig. 1, the anatomical structures, age-related contrasts, and size of macaque brains undergo significant changes during early development [18], posing challenges for accurate tissue segmentation. Especially during the early postnatal months, the WM and GM show extremely low contrast, making their boundaries difficult to detect. The development of automatic and robust tissue segmentation algorithms for macaque brains during early developing stages is thus of great importance.

Deep learning has emerged as a powerful tool in medical image processing and analysis, with Convolutional Neural Networks (CNNs) showing remarkable performance in various applications, e.g., image segmentation [4], cross-modality generation [5, 6] and multi-task learning [7, 9] with state-of-the-art results. However, there is currently a lack of dedicated deep learning-based tools for brain tissue segmentation of macaque MRI data during early development, which can be

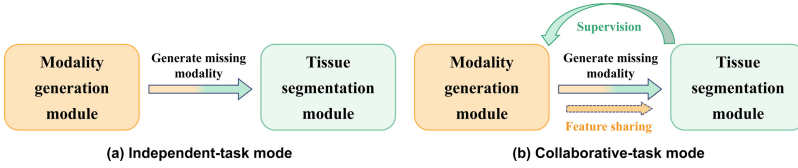


Fig. 2. A brief comparison of two modes: (a) Independent-task mode conducts image generation and tissue segmentation separately and step by step; (b) the proposed collaborative-task mode combines them through feature sharing and supervised learning.

attributed to several reasons. First, collecting macaque brain MRI data is challenging, and there is a dearth of publicly available large-scale labeled datasets for deep learning-based tissue segmentation. Second, brain tissue segmentation in early postnatal stages usually requires multi-modality images to provide more useful information than a single modality [8]. However, macaque datasets that are publicly available often only contain 3D T1-weighted (T1w) images and lack corresponding tissue labels and 3D T2-weighted (T2w) information. For example, the UW-Madison Rhesus MRI dataset [24] includes 592 macaque samples with T1w images only. To compensate for the missing modality information, generative adversarial networks (GAN) [13] have been applied to impute the missing modality in neuroimages, due to their potential for image-to-image translation [14–16]. Among them, pix2pix algorithm first proposes a solution for style transfer from image to image, which can be applied in the modality translation of medical images. Recently PTNet3D transforms the CNN-based frame into a Transformer-based frame, which reduces the computation amount required when processing high-resolution images. In these solutions, modality generation and downstream tasks are treated as two cascading independent tasks, as shown in Fig. 2(a). However, it is difficult to determine whether the generated missing modality has a natural positive effect on the downstream tasks [11]. In fact, both modality generation and tissue segmentation tasks require the feature extraction and learning of brain anatomical structures to achieve voxel-wise translation, showing a high task similarity and correlation. Essentially, the tissue map can guide the missing modality generation with anatomical information, while the generated missing modality can also provide useful complementary information for refining tissue segmentation [12]. To effectively capture and utilize the features and task correlation between two tasks, it is desirable to integrate tissue segmentation and modality generation into a unified framework. This can be achieved by imputing the missing modality image in a tissue-oriented manner, as illustrated in Fig. 2(b).

In this study, we present a 3D Collaborative Segmentation-Generation Framework (CSGF) for early-developing macaque MR images. The CSGF is designed in the form of multi-task collaboration and feature sharing, which enables it to complete the missing modality generation and tissue segmentation simultaneously. As depicted in Fig. 3(a), the CSGF comprises two modules:

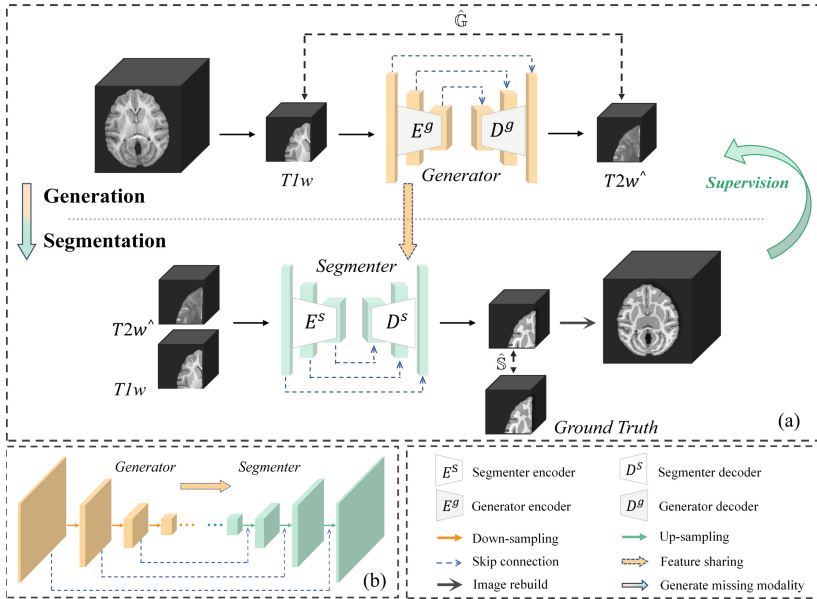


Fig. 3. Overview of the proposed 3D Collaborative Segmentation-Generation Framework (CSGF) (a). We use the T1w as the input of the modality generation module (MGM) to generate the missing modality T2w ($T2w^{\wedge}$). Meanwhile, the generated T2w ($T2w^{\wedge}$) and T1w are used as the input to the tissue segmentation module (TSM), and the results will be rebuilt into complete 3D MR images and represent the MGM loss and the TSM loss, respectively. In the structure diagram of cross-module feature sharing (CFS) (b), green blocks represent the decoder of the generator, and orange blocks represent the encoder of the segmenter. (Color figure online)

a modality generation module (MGM) and a tissue segmentation module (TSM), which are trained collaboratively through two forward information flows. Specifically, MGM and TSM will be linked by transferring generated missing-modality and cross-module feature sharing (CFS), ensuring that the MGM is trained under the supervision of the TSM, thereby imposing a constraint on anatomical feature consistency and providing multi-modality information.

The proposed CSGF offers several advantages over existing methods: (1) The collaborative learning of both MGM and TSM enables the missing modality to be imputed in a tissue-oriented manner, and hence the generated neuroimages are more consistent with real neuroimages from an anatomical point of view; (2) The CFS mechanism between the two modules provides more prosperous feature guidance for the TSM, thus also imposing constraints on the anatomical feature consistency of the MGM encoder; (3) The CSGF achieves improvements in both missing modality generation and brain tissue segmentation, especially for infant macaques with exceptionally low contrast between different tissues.

2 Method

2.1 Dataset

The experimental dataset was from the public UNC-Wisconsin Neurodevelopment Rhesus Database [10], acquired by a GE MR750 3.0T MRI scanner and covered by animal research protocols approved by relevant Institutional Animal Care and Use committees. A total of 155 developing macaque structural MRI samples between 0 and 36 months of age were used for experiments. Each sample contains T1w image and T2w image with following parameters: T1w, matrix = 256×256 , resolution = $0.5469 \times 0.5469 \times 0.8 \text{ mm}^3$; T2w, matrix = 256×256 , resolution = $0.6016 \times 0.6016 \times 0.6 \text{ mm}^3$. For image preprocessing, FMRIB's Linear Image Registration Tool (FLIRT) in FSL (version 5.0) [20] was used to rigidly align each T2w image to its corresponding T1w image, followed by resampling all images to 0.5469 mm mm isotropic resolution. Brain skulls were removed using the method in [2] and intensity inhomogeneity correction was performed using N4 bias correction [25]. The initial tissue segmentation of each scan was obtained by LINKS [17], which was then manually corrected significantly by experienced experts using ITK-SNAP software [21].

2.2 Problem Formulation

We construct this Collaborative Segmentation-Generation Framework (CSGF) based on the primary modality T1w (denoted as \mathbf{T}_i^1) and the auxiliary modality T2w (denoted as \mathbf{T}_i^2) data. Let $\{\mathbf{T}_i^1, \mathbf{T}_i^2, \mathbf{M}_i\}_{i=1 \sim N}$ be the dataset consisting of N scans, where \mathbf{T}_i^1 and \mathbf{T}_i^2 represent the T1w and T2w of i^{th} scan, respectively; \mathbf{M}_i represents the tissue label of i^{th} scan. This framework can be formulated as

$$\hat{\mathbf{M}}_i = \mathbb{S}(\mathbf{T}_i^1, \mathbf{T}_i^2), \quad (1)$$

where $\hat{\mathbf{M}}_i$ is the estimated tissue label of the i^{th} scan. However, in practical application, it is often the case that some scans only contain the T1w modality. Therefore we construct a modality generation module (MGM) by a mapping function \mathbb{G} to generate the missing T2w modality as

$$\hat{\mathbf{T}}_i^2 = \mathbb{G}(\mathbf{T}_i^1), \quad (2)$$

where $\hat{\mathbf{T}}_i^2$ is the generated T2w based on T1w modality for the i^{th} scan. Then, the tissue segmentation module (TSM) constructs a mapping function \mathbb{S} as

$$\hat{\mathbf{M}}_i = \mathbb{S}(\mathbf{T}_i^1, \mathbb{G}(\mathbf{T}_i^1)) = \mathbb{S}(\mathbf{T}_i^1, \hat{\mathbf{T}}_i^2) \approx \mathbb{S}(\mathbf{T}_i^1, \mathbf{T}_i^2). \quad (3)$$

2.3 Framework Overview and Collaborative Learning

The proposed CSGF consists of two key modules: the MGM and the TSM. These two modules are linked through the cross-module feature sharing (CFS),

thus being trained collaboratively, as illustrated in Fig. 3(a). The CFS was created by a skip connection between the generator’s down-sampling and the segmenter’s up-sampling, as depicted in Fig. 3(b). This connection enables greater interaction and anatomy constraints between the two modules by additional back-propagating the gradient of the loss function.

The proposed MGM and TSM are trained collaboratively. Given a dataset $\{\mathbf{T}_i^1, \mathbf{T}_i^2, \mathbf{M}_i\}_{i=1 \sim N}$, the generator in MGM can be trained as

$$\hat{\mathbb{G}} = \underset{i=1}{\operatorname{argmin}} \sum \{ \|\mathbb{G}(\mathbf{T}_i^1) - \mathbf{T}_i^2\|_{GAN} + \|\mathbb{G}(\mathbf{T}_i^1) - \mathbf{T}_i^2\|_{MSE} \}, \quad (4)$$

where $\|*\|_{GAN}$ represents the GAN loss and $\|*\|_{MSE}$ represents mean square error (MSE) loss. Meanwhile, the segmenter in TSM can be trained as

$$\hat{\mathbb{S}} = \underset{i=1}{\operatorname{argmin}} \sum \{ \|\mathbb{S}(\mathbf{T}_i^1, \mathbb{G}(\mathbf{T}_i^1)) - \mathbf{M}_i\|_{CE} + \|\mathbb{S}(\mathbf{T}_i^1, \mathbb{G}(\mathbf{T}_i^1)) - \mathbf{M}_i\|_{DICE} \}, \quad (5)$$

where $\|*\|_{CE}$ represents the cross-entropy loss and $\|*\|_{DICE}$ represents dice coefficient loss. During testing, for each input scan, if it has both T1w and T2w, we predict its tissue label as $\hat{\mathbf{M}}_i = \mathbb{S}(\mathbf{T}_i^1, \mathbf{T}_i^2)$. If it only has T1w modality, we first generate its T2w modality by $\hat{\mathbf{T}}_i^2 = \mathbb{G}(\mathbf{T}_i^1)$, and then predict its tissue label as $\hat{\mathbf{M}}_i = \mathbb{S}(\mathbf{T}_i^1, \mathbb{G}(\mathbf{T}_i^1))$.

3 Results and Discussion

3.1 Experimental Details

The total scans of MRI data used in this study was 155, divided into training and test sets in a 4:1 ratio. Before training, the data had the redundant background removed and cropped into a size of $160 \times 160 \times 160$, and then overlapping patches of size 64 were generated. The intensity values of all patches were normalized to $[-1, 1]$. Both MGM and TSM were optimized simultaneously during training. We employed the Adam optimizer with a learning rate of 0.0002 for the first 50 epochs, after which the learning rate gradually decreased to 0 over the following 50 epochs. In the testing phase, overlapping patches with stride 32 generated from testing volumes were input into the model, and the obtained results were rebuilt to a volume, during which the overlaps between adjacent patches were averaged. All experiments were implemented based on Pytorch 1.13.0 and conducted on NVIDIA RTX 4090 GPUs with 24GB VRAM in Ubuntu 18.04.

3.2 Evaluation of Framework

Totally 31 scans of early-developing macaques were used for framework evaluation. To provide a comparison baseline, we used PTNet3D (PTNet), one state-of-the-art algorithm for medical image translation, and U-Net [23], the most

widely used segmentation model, to construct the CSGF. First, we trained two U-Net models as baseline models, one of which used only the T1w modality as input, while the other used both T1w and T2w modalities. Next, we trained a PTNet model, which generates corresponding T2w images based on the input T1w images. Finally, we constructed the CSGF framework using the PTNet and U-Net models described in Sect. 2. It is important to note that we constructed two versions of CSGF, one with the embedding of CFS to reflect the role of feature sharing and one without. According to the developmental stage of macaques [22], we divided the evaluation scans into three groups, including 7 in early infancy (age range of 0 to 6 months), 9 in infancy (age range of 7 to 12 months) and 15 in yearlings and juveniles (age range of 13 to 36 months). Compared with the other two stages, the brain in the early infancy stage from 0 to 6 months is in a state of vigorous development, especially the extremely low contrast between tissues from 0 to 2 months, which brings great challenges to this study.

Table 1. Quantitative comparison of different ablation settings. And ‘ind.’ means independent-task mode.

Method	Cerebrospinal fluid		Gray matter		White matter	
	Dice (%)	ASD (mm)	Dice (%)	ASD (mm)	Dice (%)	ASD (mm)
Early infancy (0–6 months)						
U-Net	74.82 ± 11.51	0.31 ± 0.19	90.59 ± 5.60	0.30 ± 0.21	79.57 ± 14.37	0.36 ± 0.28
PTNet+U-Net(ind.)	73.30 ± 11.23	0.49 ± 0.23	90.02 ± 5.46	0.35 ± 0.22	77.34 ± 14.43	0.41 ± 0.27
CSGF w/o CFS	76.90 ± 11.80	0.28 ± 0.19	90.97 ± 5.54	0.27 ± 0.19	80.21 ± 14.20	0.35 ± 0.28
CSGF with CFS	77.54 ± 11.44	0.26 ± 0.18	90.99 ± 5.73	0.27 ± 0.19	80.57 ± 14.40	0.33 ± 0.27
Infancy (7–12 months)						
U-Net	92.07 ± 3.21	0.06 ± 0.09	96.65 ± 0.73	0.06 ± 0.02	94.46 ± 1.21	0.08 ± 0.03
PTNet+U-Net(ind.)	91.11 ± 3.85	0.13 ± 0.09	96.30 ± 0.83	0.09 ± 0.05	93.82 ± 1.36	0.09 ± 0.03
CSGF w/o CFS	93.12 ± 3.10	0.06 ± 0.09	96.89 ± 0.75	0.05 ± 0.02	94.58 ± 1.19	0.07 ± 0.02
CSGF with CFS	93.44 ± 3.01	0.05 ± 0.09	97.01 ± 0.76	0.05 ± 0.02	94.80 ± 1.27	0.07 ± 0.03
Yearlings and juveniles (13–36 months)						
U-Net	92.71 ± 0.99	0.04 ± 0.01	96.69 ± 0.41	0.05 ± 0.01	94.70 ± 0.89	0.07 ± 0.02
PTNet+U-Net(ind.)	91.30 ± 2.01	0.11 ± 0.09	96.36 ± 0.49	0.07 ± 0.02	94.04 ± 1.01	0.08 ± 0.03
CSGF w/o CFS	93.65 ± 1.00	0.03 ± 0.01	96.95 ± 0.43	0.05 ± 0.01	94.81 ± 0.87	0.07 ± 0.02
CSGF with CFS	93.79 ± 0.46	0.03 ± 0.01	97.08 ± 0.43	0.04 ± 0.01	95.03 ± 0.88	0.06 ± 0.02

In the evaluation of tissue segmentation, we assumed the practical scenario that the test scans contain only T1w images. We compared four results, including U-Net with T1w input, U-Net with T1w and generated T2w input from PTNet (i.e. independent-task mode), CSGF without CFS embedding and the full CSGF framework. We quantitatively evaluate the results by using the average surface distance (ASD) and the dice coefficient, as reported in Table 1. The results show that compared to U-Net with a single T1w images input, U-Net with independent PTNet-generated T2w images as an additional modality input has worse results, especially for data during infancy. This may be due to the fact that the

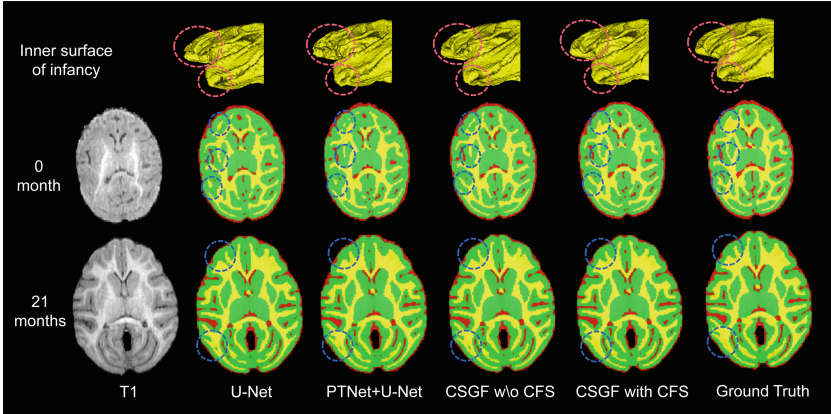


Fig. 4. Representative results of tissue segmentation. Red, green, and yellow colors represent cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM), respectively. Blue dotted circles indicate significantly different results across different methods. Typical results of the reconstructed inner surface during infancy are also displayed on the top row, where red dotted circles indicate apparent differences. (Color figure online)

T2w images generated in the independent-task mode do not accurately express the anatomical structures of different tissues based on the real T1w images, thus introducing unnecessary noise for subsequent tissue segmentation. In comparison, the results of the two CSGF methods are superior to those of the above two, which may be due to the fact that under task-oriented supervised training, the T2w images generated by PTNet tend to retain anatomical structures that are conducive to subsequent tissue segmentation. Notably, the method in which CFS is embedded achieves the best results, showing that tissue segmentation also benefits from the deconstruction of features by the MGM during the encoding stage. Figure 4 presents representative visual results of tissue segmentation at two age stages. It can be seen that the result of CSGF is the closest to the ground truth, retaining more anatomical details.

In the evaluation of the modality generation, we compared generated T2w images from T1w images based on three methods, including PTNet, CSGF w/o CFS (i.e. PTNet under the supervision of U-Net) and the final CSGF framework. We quantitatively evaluated the results by using the peak signal to noise ratio (PSNR) and structural similarity index (SSIM), as reported in Table 2. It demonstrates that co-training supervised by subsequent tissue segmentation leads to PTNet in CSGF with improved generative results, as compared to a single PTNet. Especially in the infancy stage, the results of CSGF have been significantly improved. This improvement can be attributed to the fact that the generation module is constrained by task-oriented anatomical features, which encourage the preservation of complete tissue structures. This helps to address the challenge of low tissue contrast to some extent. The stable performance of

Table 2. Modality generation results on early infancy (0–6 months), infancy (7–12 months), yearlings and juveniles (13–36 months) macaques.

Method	Early infancy		Infancy		Yearlings and juveniles	
	PSNR(dB)	SSIM(%)	PSNR(dB)	SSIM(%)	PSNR(dB)	SSIM(%)
PTNet	25.74 ± 2.12	84.51 ± 6.19	28.07 ± 1.54	92.22 ± 1.92	27.44 ± 3.71	88.74 ± 12.53
CSGF w/o CFS	26.64 ± 1.90	84.97 ± 5.62	27.95 ± 1.39	92.02 ± 1.63	26.99 ± 3.46	88.72 ± 11.41
CSGF with CFS	27.74 ± 1.56	85.92 ± 5.38	28.17 ± 1.66	92.03 ± 1.60	27.37 ± 3.53	88.97 ± 11.57

the generative module suggests that anatomical feature consistency and task relevance can coexist and be co-optimized between MGM loss and TSM loss, leading to the preservation of image quality in the generative model. Furthermore, the incorporation of CFS enhances the generation results, possibly due to the added anatomical feature constraints that encourage the MGM to prioritize the encoding of anatomical information.

4 Conclusion

We propose the novel Collaborative Segmentation-Generation Framework (CSGF) to deal with the missing-modality brain MR images for tissue segmentation in early-developing macaques. Under this framework, the modality generation module (MGM) and tissue segmentation module (TSM) are jointly trained through cross-module feature sharing (CFS). The MGM is trained under the supervision of the TSM, while the TSM is trained with real and generated neuroimages. Comparative experiments on 155 scans of developing macaque data show that our CSGF outperforms conventional independent-task mode in both modality generation and tissue segmentation, showing its great potential in neuroimage research. Furthermore, as the proposed CSGF is a general framework, it can be easily extended to other types of modality generation, such as CT and PET, combined with other image segmentation tasks.

Acknowledgements. This work was supported in part by the National Natural Science Foundation of China #61971213, and #U22A20350, the National Natural Science Foundation of Youth Science Foundation project of China #62201246, and #62001206, and the Guangzhou Basic and Applied Basic Research Project #2023A04J2262.

References

1. Roelfsema, P.R., et al.: Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research. *Neuron* **82**(6), 1200–1204 (2014)
2. Zhong, T., et al.: DIKA-Nets: Domain-invariant knowledge-guided attention networks for brain skull stripping of early developing macaques. *Neuroimage* **227**, 117649 (2021)
3. Li, G., et al.: Computational neuroanatomy of baby brains: a review. *Neuroimage* **185**, 906–925 (2019)

4. Isensee, F., et al.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**(2), 203–211 (2021)
5. Yu, B., et al.: Medical image synthesis via deep learning. *Deep Learn. Med. Image Anal.* 23–44 (2020)
6. Zhou, B., Liu, C., Duncan, J.S.: Anatomy-constrained contrastive learning for synthetic segmentation without ground-truth. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12901, pp. 47–56. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87193-2_5
7. Zheng, H., et al.: Phase Collaborative Network for Two-Phase Medical Imaging Segmentation. [arXiv:1811.11814](https://arxiv.org/abs/1811.11814) (2018)
8. Zhou, T., et al.: A review: deep learning for medical image segmentation using multi-modality fusion. *Array* **3**, 100004 (2019)
9. Pan, Y., Chen, Y., Shen, D., Xia, Y.: Collaborative image synthesis and disease diagnosis for classification of neurodegenerative disorders with incomplete multi-modal neuroimages. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12905, pp. 480–489. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87240-3_46
10. Young, J.T., et al.: The UNC-Wisconsin rhesus macaque neurodevelopment database: a structural MRI and DTI database of early postnatal development. *Front. Neurosci.* **11**, 29 (2017). <https://pubmed.ncbi.nlm.nih.gov/28210206/>
11. Xie, G., et al.: A Survey of Cross-Modality Brain Image Synthesis. *arXiv preprint arXiv:2202.06997* (2022)
12. Yu, Z., Zhai, Y., Han, X., Peng, T., Zhang, X.-Y.: MouseGAN: GAN-based multiple MRI modalities synthesis and segmentation for mouse brain structures. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12901, pp. 442–450. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87193-2_42
13. Goodfellow, I., et al.: Generative adversarial networks. *Commun. ACM* **63**(11), 139–144 (2020)
14. Isola, P., et al.: Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*, pp. 1125–1134. IEEE(2017)
15. Zhu, J., et al.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*, pp. 2223–2232. IEEE (2017)
16. Zhang, X., et al.: PTNet3D: a 3D high-resolution longitudinal infant brain MRI synthesizer based on transformers. *IEEE Trans. Med. Imaging* **41**(10), 2925–2940 (2022)
17. Wang, L., et al.: Links: learning-based multi-source integration framework for segmentation of infant brain images. *NeruoImage* **108**, 160–172 (2015)
18. Zhong, T., et al.: Longitudinal brain atlases of early developing cynomolgus macaques from birth to 48 months of age. *NeruoImage* **247**, 118799 (2022)
19. Wang, F., et al.: Developmental topography of cortical thickness during infancy. *Proc. Natl. Acad. Sci.* **116**(32), 15855–15860 (2019)
20. Jenkinson, M., et al.: FSL. *Neuroimage* **62**(2), 782–790 (2012)
21. Yushkevich, P.A., et al.: User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* **31**(3), 1116–1128 (2006)
22. National Centre for the Replacement, Reduction and Refinement of Animals in Research. <https://macaques.nc3rs.org.uk/about-macaques/life-history>

23. 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
24. UW-Madison Rhesus MRI dataset. https://fcon_1000.projects.nitrc.org/indi/PRIME/uwmadison.html
25. Tustison, N.J., et al.: N4ITK: improved N3 bias correction. *IEEE Trans. Med. Imaging* **29**(6), 1310–1320 (2010)