



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Fetal MRI Reconstruction by Global Diffusion and Consistent Implicit Representation

Junpeng Tan¹, Xin Zhang^{1,3}(✉), Chunmei Qing^{1,3}, Chaoxiang Yang⁴, He Zhang⁵, Gang Li⁶, and Xiangmin Xu^{2,3}(✉)

- ¹ The School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China.
tjeeepscut@gmail.com, eexinzhang@scut.edu.cn
- ² School of Future Technology, South China University of Technology, Guangzhou 511442, China.
- ³ Pazhou Lab, Guangzhou 510640, China.
- ⁴ Department of Radiology, Guangdong Women and Children Hospital, Guangzhou 510010, China.
- ⁵ Department of Radiology, Obstetrics, and Gynecology Hospital, Fudan University, Shanghai, 200011, China.
- ⁶ Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, 27599, The United States of America.

Abstract. Although the utilization of multi-stacks can solve fetal MRI motion correction and artifact removal problems, there are still problems of regional intensity heterogeneity, and global consistency discrimination in 3D space. To this end, we propose a novel coarse-to-fine self-supervised fetal brain MRI Radiation Diffusion Generation Model (RDGM). Firstly, we propose a novel self-supervised regionally Consistent Implicit Neural Representation (CINR) network with a double-spatial voxel association consistency mechanism to solve regional intensity heterogeneity. CINR enhances regional 3D voxel association and complementarity by two-voxel mapping spaces to generate coarse MRI. We also fine-tune the weighted slice reconstruction loss to improve the network reconstruction performance. Moreover, we propose the Global Diffusion Discriminative Generation (GDDG) fine module to enhance volume global consistency and discrimination. The noise diffusion is used to transform the global intensity discriminant information in 3D volume. The experiments on two real-world fetal MRI datasets demonstrate that RDGM achieves state-of-the-art results.

Keywords: Fetal MRI · Slice-to-Volume Reconstruction · Neural Radiation Field · Diffusion Model.

1 Introduction

Fetal Magnetic Resonance Imaging (MRI) is susceptible to inter-slice motion artifacts [1] [2] [3]. Although many methods can reduce the motion artifacts within slices, there are still some problems with 3D space discontinuity such

as insufficient inter-slice motion correction and noise. 3D volume reconstruction using multiple fetal MRI stacks acquired from different directions has received increasing attention. In the early work, Ali et al. proposed a robust M-estimation solution that minimizes the robust error norm function between slices generated by the model and real slices [4]. Maria et al. proposed a method to reconstruct fetal volume MRI from 2D slices [5], by combining intensity matching and robust statistics from 2D slices and excluding the misregistered or damaged voxels and slices. These Slice-to-Volume (SVR) methods are widely used for fetal brain motion correction after rigid transformation. Rigid registration cannot solve the misregistration problem caused by the deformable structures. Amir et al. presented Patch-to-Volume Reconstruction (PVR) [6], which could reconstruct the large field of view of non-rigid deformable structures. Further, Alena et al. proposed non-rigid motion correction deformable SVR (DSVR) [7]. Compared with SVR and PVR, the reconstruction performance of DSVR is improved. Notably, these SVR methods are all refactoring methods based on machine learning, and the processing process is tedious.

Recently, SVR methods can realize registration and reconstruction directly with deep learning. Xu et al. proposed Spatio-Temporal Resolution Enhancement with Simulated Scans (STRESS) [8]. Xu et al. proposed Iterative Transformer for Slice-to-Volume Registration (SVoRT) [9]. Shi et al. proposed an Affinity Fusion-based Framework for Iteratively Random Motion correction (AFFIRM) [10]. However, the existing SVR deep learning methods trained on simulated data are time-consuming, especially when high-resolution MRI volume is required. Xu et al. proposed implicit Neural Representation for SVR (NeSVoR) [11], which was a self-supervised implicit network [12] and overcame the high storage costs of dense discretized voxel grids [13]. Although the existing fetal 3D volume reconstruction methods have good effects, there are still defects in fetal MRI with severe motion artifacts and noise. Inter-slice regional consistency, as well as the overall integrity of the volume, are not considered.

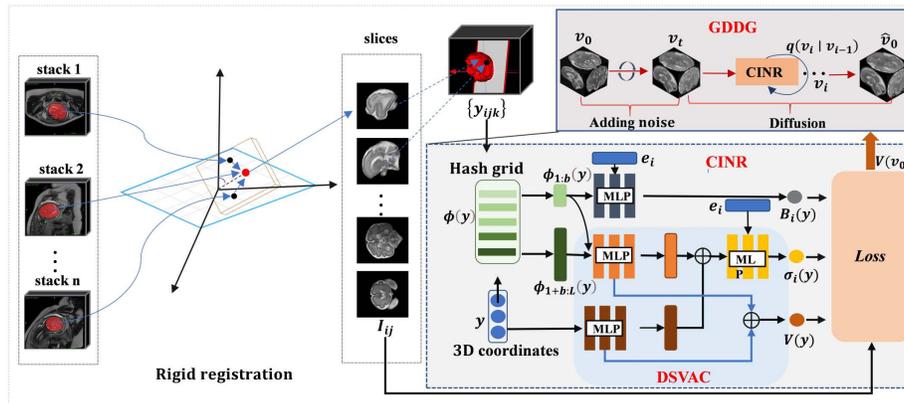


Fig. 1. The general block diagram of our method RDGM.

To address the above issues, we propose a novel coarse-to-fine self-supervised fetal brain 3D MRI reconstruction method, called Radiation Diffusion Generation Model (RDGM). As shown in Fig. 1, multi-stacks of fetal MRI and brain masks are used as input. Initially, the pre-trained SVoRT model is employed for rigid registration of the slices to accommodate varying datasets. Subsequently, to address regional intensity heterogeneity, a coarse SVR 3D training strategy is conducted using the proposed Consistency Implicit Neural Representation (CINR). CINR involves proposing a 3D voxel batch association map from two distinct voxel mapping spaces, which is the Double-Spatial Voxel Association Consistency (DSVAC) mechanism. Moreover, to make the CINR network more robust and improve the network reconstruction performance, we fine-tune the weighted slice reconstruction loss. Finally, to enhance volume global consistency and discrimination, the Global Diffusion Discriminative Generation (GDDG) mechanism is proposed for generating fine high-fidelity fetal brain MRI 3D volume. GDDG uses the trained CINR network as the bias volume generation network in the diffusion model to optimize the diffusion of noise voxels and generate discriminative high-quality 3D volume.

2 Methods

Multiple stacks of fetal MRI data are donated as $X = [x_1, x_2, \dots, x_n]$, and n is the number of stacks. First, we use the masks of each stack $M = [m_1, m_1, \dots, m_n]$ by NiftyMIC [14] to segment X . Then, we use a pre-trained model to achieve rigid registration of each slice in X . The corresponding slice transformations can be expressed as $T = [T_1, T_2, \dots, T_r]$ and the slices of all stacks denote as $Y = [y_1, y_2, \dots, y_r]$, where r is the number of all slices. In the CINR sub-module, let $I \in \mathbb{R}^{r \times N_p}$ become the data of the acquired slices, I_{ij} denotes the intensity of the j -th pixel in the i -th slice, and N_p is the number of pixels in each slice. V is an unknown reconstruction volume. All input slices can be represented as an array of all voxel coordinates as $Y \in \mathbb{R}^{N_v \times 3}$ in the CINR intensity generation network, and N_v is the number of reconstruction volume voxels.

2.1 Consistency Implicit Neural Network

The Implicit Neural Representation network (INR) has been widely used in 3D rendering [15] [16] [17]. However, the spatial mapping of existing methods will lose the region-related consistency of the original spatial voxels [11]. To this end, we propose a new regional intensity learning mechanism for 3D batch voxels in two different voxel mapping spaces, DSVAC. In this way, by combining the input voxel regions of a 3D batch of the CINR network, the network can strengthen the connection between points in different voxel mapping spaces.

$$[V_1(y), Z_1(y)] = MLP_{V_1}(y); [V_2(y), Z_2(y)] = MLP_{V_2}(\phi(y)) \quad (1)$$

$$V(y) = V_1(y) + V_2(y); Z(y) = Z_1(y) + Z_2(y) \quad (2)$$

where $MLP_{V_1}(\cdot)$ and $MLP_{V_2}(\cdot)$ are the original regional intensity learning network and hash grid encoding intensity learning network, respectively. All MLP s have one hidden layer with 64 units and $ReLU$ activation. $\phi(\cdot)$ is the hash grid encoding, $\phi(y) = [\phi_1(y), \phi_2(y), \dots, \phi_L(y)]$ and L is the length of hash grid encoding. $V(y)$ and $Z(y)$ are the coordinate feature vector and intensity feature vector at position y , respectively. Further, we adopt a continuous slices acquisition model to generate the intensity of j -th pixel in the i -th slice I_{ij} .

$$I_{ij} = C_i \int_{\Omega} M_{ij}(y) B_i(y) [V(y) + \epsilon_i(y)] dy \quad (3)$$

$$M_{ij}(y) = g(T_i^{-1} \circ y - p_{ij}; \Sigma); g(u; \Sigma) = \frac{1}{\sqrt{(2\pi)^3 \det(\Sigma)}} \exp\left(-\frac{1}{2} u^T \Sigma^{-1} u\right) \quad (4)$$

$$B_i(y) = MLP_B(\phi_{1:b}(y), e_i) \quad (5)$$

where Ω is the 3D mask region, C_i is scaling factor, $M_{ij}(y)$ is the coefficient of spatially aligned. T_i is the rigid transformation of the i -th slice from transformer-based pre-trained SVoRT. p_{ij} is the location of pixel I_{ij} in the slice coordinates. Σ is the covariance matrix of the Gaussian PSF. $B_i(y)$ is the bias field, which uses the low-level encoding $\phi_{1:b}(y) = [\phi_1(y), \phi_2(y), \dots, \phi_b(y)]$ and the slice embedding e_i to train $B_i(y)$. $\epsilon_i(y)$ is white Gaussian noise with $\mathbb{E}[\epsilon_i(y)] = 0$ and $\mathbb{E}[\epsilon_i(y)\epsilon_i(z)] = \sigma_i^2(y)\delta(y-z)$, $\delta(\cdot)$ is the Dirac delta function. According to Eq. (4), the mean and variance of pixel I_{ij} can be denoted as:

$$\mathbb{E}[I_{ij}] = C_i \int_{\Omega} M_{ij}(y) B_i(y) V(y) dy; \sigma_{ij}^2 = C_i^2 \int_{\Omega} M_{ij}^2(y) B_i^2(y) \sigma_i^2(y) dy \quad (6)$$

$$\sigma_i^2(y) = MLP_{\sigma}(Z(y), e_i) \quad (7)$$

According to the continuous iterations of the training network $MLP_{V_1}(\cdot)$, $MLP_{V_2}(\cdot)$, $MLP_B(\cdot)$ and $MLP_{\sigma}(\cdot)$, variables $V(y)$, $M_{ij}(y)$, $B_i(y)$ and σ_{ij}^2 can be obtained. Combined with Eq. (4), the intensity of each pixel is calculated.

2.2 Global Diffusion Discriminative Generation

Here, we adopt the idea of diffusion [18] [19] [20][21] generation for the volume-to-volume reconstruction of fetal brain MRI 3D volumes. Combined with the CINR, we propose a GDDG module. The global feature intensity can be learned and the uniform distribution of the overall intensity can be strengthened by the global discriminative optimization of the noise voxels with continuous diffusion. Let the CINR output V as GDDG v_0 , we define a forward diffusion process q that gradually adds Gaussian noise to the initial MRI v_0 over t iterations:

$$q(\mathbf{v}_{1:t} | \mathbf{v}_0) = \prod_{i=1}^t q(\mathbf{v}_i | \mathbf{v}_{i-1}); q(\mathbf{v}_i | \mathbf{v}_{i-1}) = \mathcal{N}(\mathbf{v}_i | \sqrt{\alpha_i} \mathbf{v}_{i-1}, (1 - \alpha_i) \mathbf{I}) \quad (8)$$

where the scalar parameters $\alpha_{1:t}$ are hyper-parameters, \mathbf{I} is unit tensor and $0 < \alpha_i < 1$. The distribution of v_i given v_0 by marginalizing out the intermediate steps can be as

$$q(\mathbf{v}_i | \mathbf{v}_0) = \mathcal{N}(\mathbf{v}_i | \sqrt{\gamma_i} \mathbf{v}_0, (1 - \gamma_i) \mathbf{I}) \quad (9)$$

where $\gamma_i = \prod_{m=1}^i \alpha_m$, recall that the generation model is trained to estimate noise. Here, we use CINR as the score generation network model. v_{i-1} can derive the posterior distribution as

$$\mathbf{v}_{i-1} = \frac{1}{\sqrt{\alpha_i}} \left(\mathbf{v}_i - \frac{1 - \alpha_i}{\sqrt{1 - \gamma_i}} \text{CINR}(\mathbf{v}_0, \mathbf{v}_i, \gamma_i) \right) \quad (10)$$

Thus, given v_i , through the continuous iterative generation of Eq. (10), we approximate v_0 as

$$\hat{\mathbf{v}}_0 = \frac{1}{\sqrt{\gamma_i}} \left(\mathbf{v}_i - \sqrt{1 - \gamma_i} \text{CINR}(\mathbf{v}_0, \mathbf{v}_i, \gamma_i) \right) \quad (11)$$

2.3 Loss Functions

1) Slice Reconstruction Loss: We use the generated pixel intensity mean \bar{I}_{ij} and variance σ_{ij}^2 to reconstruct the underlying volume by minimizing the negative log-likelihood Gaussian distribution.

$$\mathcal{L}_I = \frac{1}{|\mathcal{B}|} \sum_{(i,j) \in \mathcal{B}} \lambda \frac{(I_{ij} - \bar{I}_{ij})^2}{2\sigma_{ij}^2} + \frac{1}{2} \log(\sigma_{ij}^2) \quad (12)$$

where $\mathcal{B} \subset \{1, \dots, N_s\} \times \{1, \dots, N_p\}$ is a batch of data, I_{ij} is initial pixel intensity. λ is the trade-off parameter.

2) Image Regulation Loss: We adopt several regularization methods $r_m(\cdot)$ (isotropic total variation, first-order Tikhonov, and edge-preserving) to improve image quality and suppress noise.

$$\mathcal{R}_V = \sum_{m=1}^3 \frac{2}{K|\mathcal{B}|} \sum_{(i,j) \in \mathcal{B}} \sum_{k=1}^{K/2} r_m \left(\frac{|V(x_{ijk}) - V(x_{ijl})|}{\|x_{ijk} - x_{ijl}\|_2} \right) \quad (13)$$

Here, K is the number of subjects, we split K into $K/2$ pairs. $|V(x_{ijk}) - V(x_{ijl})| / \|x_{ijk} - x_{ijl}\|_2$ is the directional derivative for each pair, and $l = k + K/2$.

3) Bias Field Loss: According to [11], the bias field B_i and the volume V is unique up to a constant factor. To disambiguate (B_i, V) and $(cB_i, \frac{1}{c}V)$, we should force the mean log bias field to be zero, where c is any constant. Therefore, we can get the bias field loss function as follows:

$$\mathcal{R}_B = \left(\frac{1}{K|\mathcal{B}|} \sum_{(i,j) \in \mathcal{B}} \sum_{k=1}^K \log B_i(x_{ijk}) \right)^2 \quad (14)$$

Finally, the training optimization problem can be denoted as

$$\underset{\theta}{\operatorname{argmin}} \mathcal{L}(\theta), \quad \mathcal{L} = \mathcal{L}_I + \lambda_V \mathcal{R}_V + \lambda_B \mathcal{R}_B \quad (15)$$

where parameters λ_V and λ_B are the trade-off weights for the regularization terms, θ is the set of trained network parameters. The learning rate is 0.001 and λ is set as 20, $\lambda_V = 2$, $\lambda_B = 100$, $\alpha_0 = 0.001$, $t = 10$. The experimental device is a GeForce RTX 3090 Ti 24G.

3 Experiments and Results

3.1 Datasets

We use two real-world datasets to validate the proposed model. Dataset A contains 132 stacks from 44 subjects, with the gestational age ranging from 24 weeks to 37 weeks. These stacks are acquired using a 3T Philips Ingenia scanner and the resolution is $0.71 \times 0.71 \times 3.0 \text{ mm}^3$, $TR/TE = 15000/177 \text{ ms}$. This study is approved by the institutional review board (Guangdong Maternal and Child Health Hospital, China). Dataset B contains 1106 stacks from 284 subjects. These stacks are acquired using a 1.5T Siemens Avanto scanner and the resolution is $0.54 \times 0.54 \times 4.4 \text{ mm}^3$, $TR/TE = 1350/92 \text{ ms}$. The gestational ages are from 21 weeks to 36 weeks. This study is approved by the institutional review board (Obstetrics and Gynecology Hospital, Fudan University, China).

3.2 Compared with Results of Different Methods

We adopt some state-of-the-art SVR methods as the baselines, which have open-source implementation: NiftyMic [14], SVRTK [1], DSVR [7], SVoRT [9] and NeSVoR [11]. As shown in Fig. 2, our method achieves the best reconstruction volumes. However, the experimental results of subjects 2 and 3 are slightly less accurate than those of subject 1 and have a little artifact. Dataset B has more severe motion artifacts. Due to severe motion in subjects 2 and 3, the experimental results of CVRTK, NiftyMic, and DSVR are not shown in Fig. 2.

3.3 Simulated Fetal Brain Data

The Fetal Brain Atlas sample (26 weeks) [22] is evaluated by slices to simulate real-world fetal stacks [23]. The maximum translation and rotation in the motion trajectory is 21.4 mm/s and 59.7 degree/s, respectively. We compare quantitative metrics, including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM), Root Mean Square Error (RMSE), and Normalized Cross-Correlation (NCC). In Table 1, our proposed method achieves the best results in all quantitative metrics.

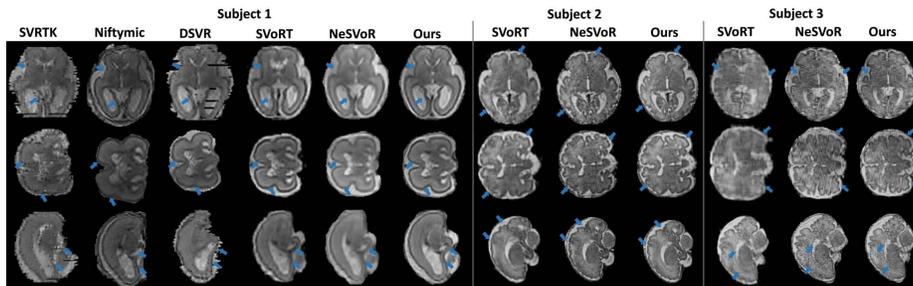


Fig. 2. The reconstruction volumes on different methods. Subject 1 (25 weeks) is from dataset A, and subject 2 (34 weeks) and subject 3 (35 weeks) are from dataset B.

Table 1. Mean (Standard Deviation) of Quantitative Metrics on Simulated Data.

Methods	PSNR(dB) \uparrow	SSIM \uparrow	RMSE \downarrow	NCC \uparrow
SVRTK	23.7439(0.1234)	0.3991(0.0012)	0.8640(0.0012)	0.2785(0.0029)
SVoRT	24.8237(0.2865)	0.3906(0.0052)	0.9581(0.0079)	0.1816(0.0555)
DSVR	25.5134(2.2824)	0.3959(0.0012)	0.8693(0.0013)	0.2728(0.0030)
NeSVoR	26.0515(0.0554)	0.3782(0.0005)	0.8846(0.0009)	0.4567(0.0041)
Ours	26.3754(0.0126)	0.4200(0.0001)	0.7807(0.0001)	0.5452(0.0029)

3.4 Reconstructing Volumes at Different Gestational Ages

To verify the effectiveness of RDGM with more samples, we performed reconstruction experiments on fetuses of different gestational ages using samples from dataset A. In Fig. 3, RDGM reconstructs great results at different gestational ages. We can see a tendency for fetal brains to grow with gestational age.

3.5 Ablation Study

Fetal brain atlas samples (26-31 weeks) are evaluated by volume to simulate ablation studies. we evaluate the model component by ablating the baseline (NeSVoR), only improving the GDDG fine module (OD), the proposed method CINR without GDDG fine module (ND), only improving network (ON), the proposed method GDDG without CINR (NN), only improving loss functions (OL), the proposed method without our improved loss function (NL), and RDGM. In Table 2, compared with the baseline, the reconstruction results of all our improved modules are better than NeSVoR.

3.6 Downstream Task Analysis

We quantify and analyze the gestational age estimation of the downstream task for the reconstruction results in Table 3 and Fig. 4, respectively. In Table 3, we show the reconstruction results of different algorithms for the downstream task of fetal brain age estimation quantified experimentally through the literature [24]. Several common evaluation criteria are used, including the mean absolute error

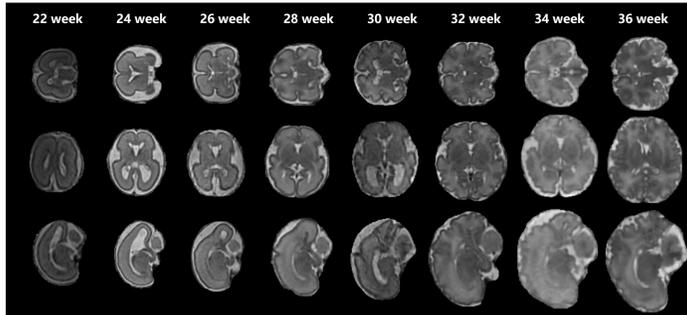


Fig. 3. The reconstruction results at different gestational ages.

Table 2. Ablation Study of Sub-module Quantitative on Simulated data.

Methods	PSNR(dB) \uparrow	SSIM \uparrow	NRMSE \downarrow	NCC \uparrow
w/o baseline	19.6583(0.6119)	0.6716(0.0646)	0.1096(0.0079)	0.5131(0.0415)
w/o OL	19.7958(0.6327)	0.6742(0.0574)	0.1018(0.0059)	0.5275(0.0409)
w/o NL	20.2403(0.6271)	0.6770(0.0646)	0.0961(0.0063)	0.5194(0.0445)
w/o ON	20.2088(0.4302)	0.6991(0.0118)	0.1084(0.0073)	0.5108(0.0207)
w/o NN	20.0470(0.7552)	0.7106(0.0646)	0.0947(0.0054)	0.5246(0.0403)
w/o OD	20.0296(0.2147)	0.6716(0.0106)	0.1081(0.0040)	0.5153(0.0164)
w/o ND	19.8502(0.6600)	0.6756(0.0615)	0.1031(0.0062)	0.5261(0.0417)
w Full A	20.8536(0.7686)	0.7191(0.0490)	0.0949(0.0083)	0.5331(0.0374)

(MAE), the standard deviation (STD), Pearson’s correlation coefficient (PCC), and Spearman’s rank correlation coefficient (SRCC). As shown in Fig. 4, the distribution of gestational age prediction with linear regression is visualized for the two datasets. In the top row of Fig. 4, the reconstructed MRI volumes by our proposed method have a more clustered scatter diagram in the downstream fetal brain age estimation task. In the lower row of Fig. 4, between 21-25 weeks, the reconstructed MRI volumes of the three methods could not correctly predict brain age. Most of the correct points fall in the 26-36 weeks range. This indicates that the dataset has clear high-quality data acquisition in this range.

4 Conclusion

In this work, we propose a new real-world multi-stacks fetal brain MRI reconstruction model, called RDGM. This method incorporates slices rigid registration (SVoRT), the novel slice-to-volume regional consistency self-supervise generation (CINR), and volume-to-volume global associative high-quality diffusion reconstruction (GDDG). RDGM also reconstructs motion artifact data obtained from various real scans by self-supervised learning. Especially, our proposed method can achieve optimal reconstruction results in both high and low-quality data. gestational age estimation experimental demonstrates the effectiveness of the proposed method. In the future, we will investigate reconstruction algorithms for fetal MRI with fewer stacks and lower-quality acquisitions.

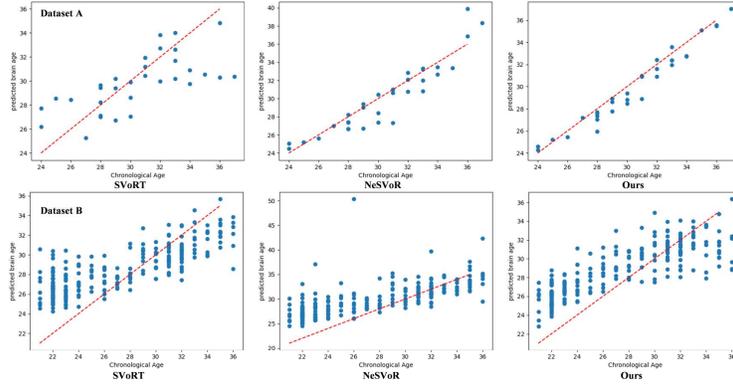


Fig. 4. Results of gestational age estimation.

Table 3. Ablation Study of Sub-module Quantitative on Simulated data.

Datasets	Methods	Evaluation Metrics				
		MAE	STD	PCCage	SRCCage	SRCCgap
Dataset A	SVoRT	2.003	2.47	0.683	0.766	-0.659
	NeSVoR	1.311	1.399	0.907	0.706	-0.634
	Ours	0.681	0.684	0.979	0.979	-0.157
Dataset B	SVoRT	2.871	3.113	0.799	0.774	-0.895
	NeSVoR	3.243	3.293	0.705	0.794	-0.793
	Ours	2.851	2.885	0.821	0.835	-0.784

Acknowledgment. This work is supported by the following grants: Key-Area Research and Development Program of Guangdong Province (2023B0303040001); Guangdong Basic and Applied Basic Research Foundation(2024A1515010180); Guangdong Provincial Key Laboratory of Human Digital Twin (2022B1212010004) and GDHDT01.

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

References

- [1] Maria Kuklisova-Murgasova, Quaghebeur, et al. Reconstruction of fetal brain mri with intensity matching and complete outlier removal. *Medical image analysis*, 16(8):1550–1564, 2012.
- [2] Lucilio Cordero-Grande, Ortuño-Fisac, et al. Fetal mri by robust deep generative prior reconstruction and diffeomorphic registration: application to gestational age prediction. *arXiv preprint arXiv:2111.00102*, 2021.
- [3] Priscille de Dumast, Thomas Sanchez, H el ene Lajous, et al. Simulation-based parameter optimization for fetal brain mri super-resolution reconstruction. *arXiv preprint arXiv:2211.14274*, 2022.
- [4] Ali Gholipour, Judy A Estroff, and Simon K Warfield. Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain mri. *IEEE transactions on medical imaging*, 29(10):1739–1758, 2010.
- [5] Bahram Marami, Benoit Scherrer, Onur Afacan, et al. Motion-robust diffusion-weighted brain mri reconstruction through slice-level registration-based motion tracking. *IEEE transactions on medical imaging*, 35(10):2258–2269, 2016.
- [6] Amir Alansary, Martin Rajchl, Steven G McDonagh, et al. Pvr: patch-to-volume reconstruction for large area motion correction of fetal mri. *IEEE transactions on medical imaging*, 36(10):2031–2044, 2017.
- [7] Alena Uus, Tong Zhang, Laurence H Jackson, et al. Deformable slice-to-volume registration for motion correction of fetal body and placenta mri. *IEEE transactions on medical imaging*, 39(9):2750–2759, 2020.
- [8] Junshen Xu, Esra Abaci Turk, P Ellen Grant, et al. Stress: Super-resolution for dynamic fetal mri using self-supervised learning. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VII 24*, pages 197–206. Springer, 2021.
- [9] Junshen Xu, Daniel Moyer, P Ellen Grant, et al. Svort: Iterative transformer for slice-to-volume registration in fetal brain mri. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VI*, pages 3–13. Springer, 2022.
- [10] Wen Shi, Haoan Xu, Cong Sun, et al. Affirm: Affinity fusion-based framework for iteratively random motion correction of multi-slice fetal brain mri. *IEEE Transactions on Medical Imaging*, 2022.
- [11] Junshen Xu, Daniel Moyer, Borjan Gagoski, et al. Nesvor: Implicit neural representation for slice-to-volume reconstruction in mri. *IEEE Transactions on Medical Imaging*, 2023.
- [12] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, et al. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.

- [13] Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, et al. Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems*, 33:7537–7547, 2020.
- [14] Michael Ebner, Guotai Wang, Wenqi Li, et al. An automated framework for localization, segmentation and super-resolution reconstruction of fetal brain mri. *NeuroImage*, 206:116324, 2020.
- [15] Qing Wu, Yuwei Li, Lan Xu, et al. Irem: high-resolution magnetic resonance image reconstruction via implicit neural representation. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VI 24*, pages 65–74. Springer, 2021.
- [16] Ricardo Martin-Brualla, Noha Radwan, Mehdi SM Sajjadi, et al. Nerf in the wild: Neural radiance fields for unconstrained photo collections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7210–7219, 2021.
- [17] Thomas Müller, Alex Evans, Christoph Schied, et al. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics (ToG)*, 41(4):1–15, 2022.
- [18] Hyungjin Chung, Jeongsol Kim, Sehui Kim, et al. Parallel diffusion models of operator and image for blind inverse problems. *arXiv preprint arXiv:2211.10656*, 2022.
- [19] Jiaming Song, Arash Vahdat, Morteza Mardani, et al. Pseudoinverse-guided diffusion models for inverse problems. In *International Conference on Learning Representations*, 2023.
- [20] Gwanghyun Kim, Taesung Kwon, and Jong Chul Ye. Diffusionclip: Text-guided diffusion models for robust image manipulation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2426–2435, 2022.
- [21] Nan Liu, Shuang Li, Yilun Du, et al. Compositional visual generation with composable diffusion models. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XVII*, pages 423–439. Springer, 2022.
- [22] Ali Gholipour, Caitlin K Rollins, Clemente Velasco-Annis, et al. A normative spatiotemporal mri atlas of the fetal brain for automatic segmentation and analysis of early brain growth. *Scientific reports*, 7(1):476, 2017.
- [23] Junshen Xu, Sayeri Lala, Borjan Gagoski, et al. Semi-supervised learning for fetal brain mri quality assessment with roi consistency. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VI 23*, pages 386–395. Springer, 2020.
- [24] Jian Cheng, Ziyang Liu, Hao Guan, Zhenzhou Wu, Haogang Zhu, Jiyang Jiang, Wei Wen, Dacheng Tao, and Tao Liu. Brain age estimation from mri using cascade networks with ranking loss. *IEEE Transactions on Medical Imaging*, 40(12):3400–3412, 2021.