3978

Diffusion Probabilistic Generative Models for Accelerated in-NICU, Permanent Magnet Neonatal MRI Reconstruction

Yamin Ishraq Arefeen[1,2], Brett Levac[1], and Jonathan Tamir[1]
[1]Chandra Family Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, United States, [2]Department of Imaging Physics, MD Anderson Cancer Center, Houston, TX, United States

Synopsis

Keywords: AI Diffusion Models, AI/ML Image Reconstruction, Lower-field, Neonatal, Clinical Translation

Motivation: Accelerating lower-field MRI in the neonatal-intensive-care-unit may reduce motion artifacts and increase accessibility to a wider range of patients. However, parallel imaging and machine learning reconstruction models trained on adult MRI do not apply in the neonatal setting.

Goal(s): This work accelerates single-channel neonatal MRI with diffusion-probabilistic-generative models trained from limited and noisy data collected on a permanent magnet system.

Approach: The proposed training method combines datasets from multiple contrasts and orientations with class embeddings and applies a self-supervised denoiser before training. Diffusion posterior sampling reconstructs images from under-sampled k-space.

Results: Our method enables 1.5x reduction in scan-time using a single-channel.

Impact: The improvement in acquisition speed of T_1 and T_2 weighted lower field neonatal MRI protocols using diffusion-probabilistic-generative models, trained with methods designed to handle the noisy, limited data, improves accessibility of MRI to patients in the neonatal-intensive-care-unit.

Introduction

Lower-field (below 1.5T) MRI sited directly in the neonatal-intensive-care-unit (NICU) [1] increases accessibility of neonatal brain abnormality assessment [2-4] but suffers from frequent motion artifacts [5]. Parallel imaging [6] and machine learning [7] reduce motion sensitivity in adults by shortening acquisition times. However, in-NICU permanent magnet MR systems with one receive coil cannot employ parallel imaging [1]. The inherently noisier images and limited public data repositories also make training machine learning algorithms for neonatal MRI challenging.

This work proposes accelerating neonatal MRI acquired with the in-NICU 1T permanent magnet Embrace System (Aspect Imaging) with diffusion-probabilistic-generative models. First, we gather a training dataset of clinical images in collaboration with Shaare-Zedek-Medical-Center. Then to address the low quantity and SNR of the dataset, we **(1)** modify the existing diffusion network architecture to support inputs with varying matrix sizes to expand training size, **(2)** train a single model on all orientations and contrasts with class embeddings, **(3)** apply self-supervised denoising to boost training dataset SNR.

Methods

Neonatal Dataset: With Aspect Imaging and Shaare-Zedek-Medical-Center under IRB approval and informed consent/assent, we gathered a dataset acquired with the 1T Embrace System from 128 neonatal subjects scanned with coronal, axial, and sagittal T_2 Fast-Spin-Echo (FSE) and axial T_1 Spin-Echo (SE). We randomly split this dataset into 108 training and 20 test subjects. We re-sized all slices to 200×200 matrix size in k-space and removed the first and last four slices from each volume, resulting in 8659 FSE and 3224 SE training slices. B_0/B_1 correction, available on Embrace [8], was not applied.

Proposed Generative Model Training: We trained a U-Net [9] style network with 65 million parameters and used the "EDM" hyper-parameters, loss function, and diffusion noise schedules [10]. For our neonatal setting, we first modified the network architecture so that the model can take arbitrary matrix size inputs by re-sizing all potential input dimensions to integer values. Thus, we re-size and use all heterogeneous data for training, and the model handles matrix size discrepancies at inference. Second, we trained a single model on all the coronal, sagittal, and axial FSE and axial SE data simultaneously. Our method one-hot encodes each contrast/orientation and trains a multi-layer perceptron that takes the encoding vector as input and outputs an embedding that the main U-Net incorporates as an input. Finally, we train another U-Net for denoising in a self-supervised fashion on our (inherently) noisy training dataset with Noisier2Noise [11]. This denoising model is applied to the dataset as a pre-training step before training our "EDM" based generative model. For inference, we averaged 5 posterior samples using diffusion-posterior-sampling [12]. Figure 1 summarizes our proposed training and inference pipeline.

Experiment, Training with all data with class embeddings versus using a separate model for each contrast/orientation: We trained a separate generative models for classes: axial (3599 slices), coronal (2687 slices), sagittal (3599 slices) FSE and axial SE (3224 slices). The test FSE and SE data were retrospectively under-sampled by $R = 2.0$ by removing data associated with each echo train to capture realistic signal decay [13,14]. We compared reconstructions using the multiple separate models for each class to a single model trained using all data with and without the proposed embedding. No models used denoising pre-training.

Experiment, Ablation on embedding and denoising: We trained four generative models on all of the data: **(1)** without embeddings, without denoising; **(2)** without embeddings, with denoising; **(3)** with embeddings, without denoising; and **(4)** with embeddings, with denoising. Test FSE and SE data were retrospectively under-sampled at average $R \in \{1.6, 2.0\}$, and we compared reconstructions using the four models.

Results

Figure 2 shows normalized-root-mean-squared-error (NRMSE) violin-plots ($\times 100$) of reconstructions on each contrast/orientation. Models trained on each class separately perform worse quantitatively in comparison to training on all data. Incorporating class embedding further improves performance with all data. Figure 3 displays example reconstructions.

Figure 4 shows NRMSE violin-plots exploring the effects of embedding and denoising when using all data. Embedding mostly seems to improve performance, but denoising does not yield clear quantitative improvements. In Figure 5 with example reconstructions, notice in the zoomed SE axial results row that embedding without denoising yields lower NRMSE than with denoising, but suffers from increased hallucinations.

Figures

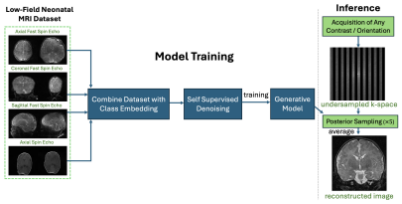

Figure 1: The proposed training and inference pipeline. Images from different orientations and contrasts are combined into a single dataset using class embeddings. A denoiser, trained in a self-supervised fashion, denoises the dataset before training. Diffusion posterior sampling reconstructs an image from under-sampled k-space by averaging 5 posterior samples. The proposed methods enable training of diffusion models for accelerated reconstruction on noisy and limited neonatal MRI data.

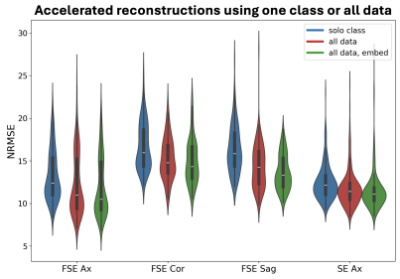

Figure 2: Violin plots comparing NRMSE across the test dataset of posterior sampling reconstructions on $2\times$ under-sampled data using diffusion models trained on just a single image class, trained on all data, and trained on all data with class embeddings. Results on all contrasts and orientations suggest that a diffusion model trained by combining all data with class embeddings yields the best quantitative performance.

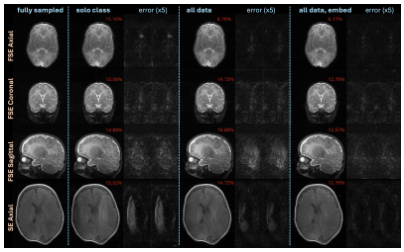

Figure 3: Example images from the experiment comparing posterior sampling reconstructions on $2\times$ under-sampled data using diffusion models trained on just a single image class, trained on all data, and trained on all data with class embeddings. This is a specific illustration of the quantitative conclusion that using all data combined with class embeddings to train the diffusion model yields best performance.

Discussion and Conclusion

Denoising might not always yield quantitative improvements because we compute NRMSE with respect to the original, inherently noisy fully-sampled image [15]. While this work explored accelerated reconstruction, the flexibility of generative models also enable application in other settings like motion correction [16]. Future work will perform a clinical reader study to evaluate the diagnostic utility of our images.

Acknowledgements

This work was supported in part by Aspect Imaging, NSF CCF-2239687 (CAREER), NSF IFML 2019844, and a JCCO fellowship.

References

[1] Kirsten Thiim, Elizabeth Singh, Srinivasan Mukundan, Ellen Grant, Edward Yang, Mohamed El-Dib, and Terrie Inder. "Clinical experience with an in-nicu magnetic resonance imaging system". Journal Of Perinatology, 2022.

[2] Lianne Woodward, Peter Anderson, Nicola Austin, Kelly Howard, and Terrie Inder. "Neonatal mri to predict neurodevelopmental outcomes in preterm infants". The New England Journal of Medicine, 2006.

[3] H. Kidokoro, JJ. Neil, and T.E. Inder. "New mr imaging assessment tool to define brain abnormalities in very preterm infants at term". American Journal of Neuroradiology, 2013.

[4] Shamik Trivedi, Zachary Vesoulis Rakesh Rao, Steve Liao, Joshua Shimony, Robert McKinstry, and Amit Mathur. "A validated clinical mri injury scoring system in neonatal hypoxic-ischemic encephalopathy". Pediatric Radiology, 2017.

[5] Jessica Dubois, Marianne Alison, Serena Counsell, Lucie Hertz-Pannier, Petra Huppi, and Manon Benders. "Mri of the neonatal brain: A review of methodological challenges and neuroscientific advances". Journal of Magnetic Resonance Imaging, 2020.

[6] Anagha Deshmane, Vikas Gulani, Mark Griswold, and Nicole Seiberlich. "Parallel mr imaging". Journal of Magnetic Resonance Imaging, 2012.

[7] Reinhard Heckel, Mathews Jacob, Akshay Chaudhari, Or Perlman, and Efrat Shimron. "Deep learning for accelerated and robust mri reconstruction". Magnetic Resonance Materials in Physics, Biology and Medicine, 2024.

[8] Kensuke Sekihara, Shigeru Matsui, and Hideki Kohno. Nmr imaging for magnets with large nonuniformities. IEEE Transactions on Medical Imaging, 1986.

[9] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. "Score-based generative modeling through stochastic differential equations". International Conference on Learning Representations, 2021.

[10] Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. "Elucidating the design space of diffusion-based generative models". Advances in Neural Information Processing Systems, 2022.

[11] Nick Moran, Dan Schmidt, Yu Zhong, and Patrick Coady. "Noisier2noise: Learning to denoise from unpaired noisy data". Conference on Computer Vision and Pattern Recognition, 2020.

[12] Hyungjin Chung, Jeongsol Kim, Michael T. Mccann, Marc L. Klasky, and Jong Chul Ye. "Diffusion posterior sampling for general noisy inverse problems". International Conference on Learning Representations, 2023.

[13] Junaid Rajput, Simon Weinmueller, Jonathan Endres, Peter Dawood, Florian Knoll, Andreas Maier, and Moritz Zaiss. "Death by retrospective undersampling caveats and solutions for learning-based mri reconstructions". International Conference on Medical Image Computing and Computer-Assisted Intervention, 2024.

[14] Brett Levac, Sidharth Kumar, Sofia Kardonik, and Jonathan Tamir. "Fse compensated motion correction for mri using data driven methods". International Conference on Medical Image Computing and Computer-Assisted Intervention, 2022.

[15] Jiayang Wang, Di An, and Justin Haldar. "The “hidden noise” problem in mr image reconstruction". Magnetic Resonance in Medicine, 2024.

[16] Brett Levac, Sidharth Kumar, Ajil Jalal, and Jonathan Tamir. "Accelerated motion correction with deep generative diffusion models". Magnetic Resonance in Medicine, 2024.

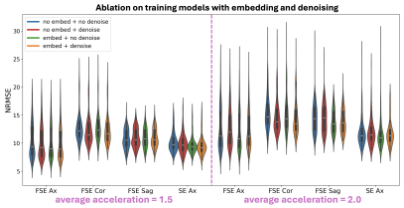

Figure 4: Violin plots comparing NRMSE across the dataset of posterior sampling reconstructions on $1.5\times$ and $2\times$ under-sampled data using diffusion models trained on all data with an ablation on using class embeddings and denoising. While incorporating class embeddings improves performance, denoising does not always yield quantitative improvement. This may be because quantitative metrics are computed with respect to the original fully-sampled, inherently noisy image.

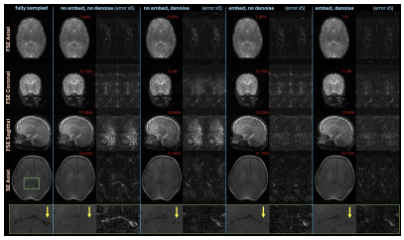

Figure 5: Example reconstructions from the experiment comparing posterior sampling reconstructions on $2\times$ under-sampled data using diffusion models trained on all data with an ablation on using class embeddings and denoising. In the SE axial results row with zoomed in images, notice how no embedding without denoising yields lower NRMSE than with denoising, but suffers from increased hallucinations.