CDMAF-CEST: Conditional Diffusion model for multi-acceleration-factor CEST-MRI Reconstruction

Pei Cai[1], Lequan Yu[2], Ziyan Wang[1], Jiawen Wang[1], Kyongtae Ty Bae[1], Henry KF Mak[1], Jianpan Huang[1*]

*[1] Department of Diagnostic Radiology, The University of Hong Kong, Hong Kong, China; [2] Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong, China; * Corresponding Author (jphuang@hku.hk).*

INTRODUCTION: CEST MRI is a promising molecular imaging technique that provides molecular-level information within the body[1]. However, the prolonged scan time required for acquiring images at multiple saturation frequency offsets hinders its application in clinical settings[2]. To accelerate CEST-MRI acquisition, advanced deep learning techniques have been explored for CEST MRI reconstruction[2-4]. Currently, diffusion-based generative models have demonstrated competitive ability in image reconstruction tasks[5-8]. In this study, we leverage a conditional diffusion model to reconstruct the high-resolution (HR) CEST image conditioned on its low-resolution (LR) counterpart that is undersampled in k-space along with the M0 image at multiple acceleration factors (CDMAF-CEST). To the best of our knowledge, this is the first diffusion-based CEST-MRI reconstruction work.

METHODS: (1) CDMAF-CEST: As illustrated in Fig. 1a, the approach achieves multi-acceleration-factor CEST-MRI SR through forward and reverse diffusion processes. Given the HR image $x_0 \sim q(x_0)$, the forward diffusion process gradually adds Gaussian noise to x_0 over T diffusion steps. On the other hand, the reverse diffusion process p aims to denoise the image from x_T step by step, with the conditioning part of its associated LR image y and M0 image. This process can be expressed as: $p_\theta(x_{0:T}) = \prod_{t=1}^{T} p_\theta(x_{t-1}|x_t)$, $p_\theta(x_{t-1}|x_t, y, z) = N(x_{t-1}; \mu_\theta(x_t, y, z, t), \sigma^2 \boldsymbol{I})$, where p_θ denotes a parameterized model with trainable parameters θ. To estimate the reverse distribution by learning latent representations from various input, we adopt a multi-stream U-Net model with disentanglement loss and charbonnier loss[8], as shown in Fig. 1b. (2) Dataset: The human CEST dataset was acquired using a GE Signa 3T MRI scanner on 24 healthy volunteers. For each subject, CEST images from 12 slices and 40 frequency offsets (one M0 image at −300 ppm and 39 CEST images ranging from −20 ppm to 20 ppm) were acquired. Each image was normalized and center cropped to a size of 224×224. We split 24 healthy brain CEST-MRIs into 19 for training (9120 images), 2 for validation (960 images), and 3 for testing (1440 images). The LR images were obtained by undersampling varied numbers of central k-space lines. (3) Implementation details: CDMAF-CEST was implemented using PyTorch with the following settings: diffusion steps T = 1000, batch size = 4. The model was trained for 100,000 iterations on four NVIDIA-SMI Tesla V100 16GB GPUs using the AdamW optimizer with a learning rate of 10^{-4}.

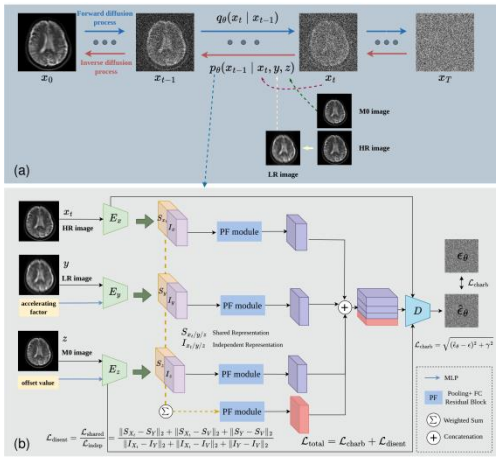

Fig. 1. Illustration of CDMAF model. (a) The forward and inverse diffusion process of the proposed conditional diffusion model. (b) The network architecture.

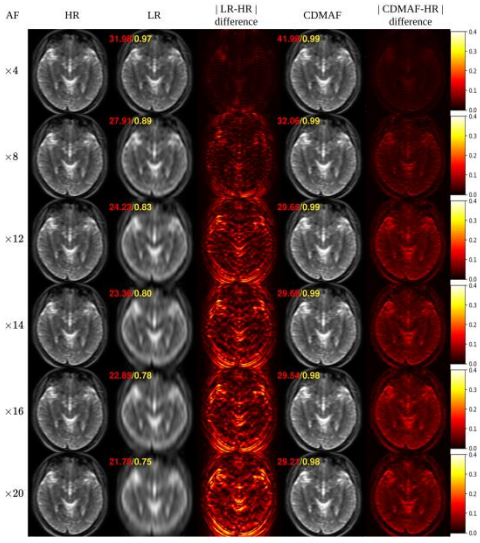

Fig. 2. The reconstruction results of the CDMAF-CEST at various AFs. The red and yellow numbers are the PSNR and SSIM, respectively.

RESULTS & DISCUSSION: Fig. 2 shows examplary reconstruction results of CDMAF-CEST with acceleration factors (AFs) of ×4, ×8, ×12, ×14, ×16, ×20. Compared with the LR images, the proposed method can generate HR images comparable to the original images with a high PSNR and SSIM, and a low difference at all AFs. Notably, our method can reconstruct spatial details even at a large AF of ×20. This could be owing to diffusion models' ability to improve image resolution by sharpening and recover image details through learning distributions. Despite of its promise, the inference of the CDMAF-CEST is time-consuming, which can be addressed by accelerating the diffusion sampling process. Ongoing work is being made to address this problem and further improve the network performance by leveraging attention mechanisms.

CONCLUSION: This study demonstrates the feasibility of using diffusion models for CEST-MRI reconstruction. The results revealed great adaptability of CDMAF-CEST to a wide range of AFs, effectively restoring image details with high SSIM and PSNR values. CDMAF-CEST has potential to promote the clinical application of CEST MRI by enhancing the image quality within the limited acquisition time.

ACKNOWLEDGMENTS: The University of Hong Kong: 109000487, 204610401 and 204610519.

REFERENCES: **1.** P. C. van Zijl, *et al*. Magn Reson Med 2011;65:927. **2.** Xu J, *et al*. Magn Reson Med. 2023;91:583. **3.** Yang Z, *et al*. IEEE J Biomed Health Inform 2024. **4.** Pemmasani Prabakaran RS, *et al*. NMR Biomed. 2024;e5130. **5.** Korkmaz Y, *et al*. MICCAI. 2023;491-501. **6.** Xie Y, *et al*. MICCAI. 2022; 655-664. 2019;32:e4139. **7.** Peng C, *et al*. MICCAI. 2022;623-633. **8.** Mao Y, *et al*. LNCS, vol 14229.2023.