Self-Supervised Low-Field MRI Denoising with Multi-Contrast Structural Guidance

Yaogong Zhang, MS[1]; Qiwen Ye, BS[2]; Zhengzheng Liu, PhD[2] ; Fang-Fang Yin, PhD[1] ; Lei Zhang, PhD[1]

[1]Medical Physics Graduate Program, Duke Kunshan University, Kunshan, China

[2]School of Electrical Engineering, Huazhong University of Science and Technology, Wuhan, China

2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

ABSTRACT

Purpose:
Low-field MRI is cost-efficient and is a recent booming direction in MRI field, while suffers intrinsically low SNR. Deep learning–based supervised denoising requires paired high-quality data that are rarely available for low-field MRI. We develop a self-supervised denoising framework that leverages co-registered multi-contrast MRI to enable denoising without needs of high-quality paired data.

Methods:
The public 0.3T M4Raw MRI dataset was used with 158 subjects for training and 25 for testing. Co-registered T1/T2-w images were jointly used as complementary anatomical guidance. Each noisy image was decomposed into two paired views via complementary sub-sampling, yielding interleaved observations with nearly identical signal content but different noise realizations. The network was trained to predict one view from its paired counterpart. With approximately zero-mean, signal-independent noise, noisy–to-noisy regression drives the network toward the expected signal. T1 and T2-w MRI were jointly-denoised using a dual-stream encoder–decoder with a shared bottleneck, while keeping modality-specific branches and intra-modality skips to prevent cross-contrast feature leakage. Training was stabilized by two consistency regularizers: cross-contrast structural consistency (implemented with MIND descriptors) to align anatomical patterns, and shared-sampling consistency to reduce sampling-induced variance by enforcing agreement at shared sampling locations.

Results:
Our method achieved the highest performance among multiple state-of-the-art (SOTA) self-supervised methods (T1-w: PSNR 32.77 dB, SSIM 0.915; T2-w: PSNR 32.42 dB, SSIM 0.891), improving PSNR by ~0.9–1.1 dB and SSIM by ~0.01 over the best-performing SOTA method. Qualitatively, zoom-in views show denoising without too much smoothing or unnatural texture patterns, compared with other self-supervised methods.

Conclusion:
A self-supervised multi-contrast MRI denoising network that learns from noisy data without high-quality references was developed for low-field MRI. By jointly denoising co-registered T1/T2-w MRI with a leakage-aware dual-stream architecture and consistency regularizers that preserve shared anatomy, the method achieved improved SSIM/PSNR over SOTA self-supervised methods.

CONTACT

Lei Zhang, Ph.D.
Medical Physics Graduate Program, Duke Kunshan University, Kunshan, Suzhou, Jiangsu, 215316, China
lei.zhang@dukekunshan.edu.cn
(314) 323-6469

INTRODUCTION

Low-field MRI has received renewed clinical and research interest because of its lower cost, improved accessibility, and potential for portable point- of-care imaging. Nevertheless, the intrinsically low signal-to-noise ratio (SNR) of low-field acquisitions remains a major barrier. Recently, in Science journal, Zhao et al. has showed that via supervised deep learning, image quality of 0.05T MRI could be enhanced to a level similar to 3T MRI scanners.

However, such methods were focusing on supervised learning where paired high-quality images were required for training. In practice, due to the accessibility and clinical constraints, paired MRI of both low and high fields were scarce. In addition, the supervised learning focused more on the signal learning, where hallucination might pose challenge for clinical applications.

Therefore, the aim of this study was to propose a self-supervised learning framework that intrinsically does not require paired high-quality data for training. In addition, instead of signal learning, this study focus on noise reduction, where hallucination issue might be reduced potentially.

METHODS AND MATERIALS

The materials used in this study was a public dataset containing multi-contrast brain MRI acquired on a 0.3T low-field system. T1-w and T2-w contrasts were used. 158 volunteer image sets were used for training and 25 used for testing. Reference or the high SNR images were generated by averaging 6 repeated acquisitions for the 25 datasets.

Figure 1 shows the overview of the proposed self-supervised learning framework. The noisy T1 and T2-w images were first separated into non-overlapping pseudo-input and pseudo-target sub-images via a spatial subsampling operator.

Then the T1 and T2-w pseudo-input $g_1(y_{T1})$ and $g_1(y_{T2})$ were fed into T1 and T2-w arm respectively. In each arm, the input image went through encoding, disentangled into shared and distinct heads (s_1, d_1, s_2, d_2), where the shared heads (s_1, s_2) went through modality-independent neighborhood descriptor (MIND) to extract structure feature and maximize the similarity between the two shared heads.

After the cross-contrast fusion of (s_1, d_1, s_2, d_2) in each arm, decoding were performed and resulted in output images $\hat{o}_{T1}$ and $\hat{o}_{T2}$. (**Figure 2**)

The loss function of the network is composed of 3 losses. $L_{total} = L_1 + \lambda_1 L_2 + \lambda_2 L_3$, where $\lambda_{1,2}$ are coefficients. $L_1 = \| \hat{o}_{T1} - g_2(y_{T1}) \|_2^2 + \| \hat{o}_{T2} - g_2(y_{T2}) \|_2^2$, where $g_2(y_{T1})$ and $g_2(y_{T2})$ are the pseudo-target images. This is the main part of the L_{total}. L_2 is calculating the difference between MIND descriptor of the two arms. L_3 is calculating the strength of the distinct head d_1 and d_2, to reduce cross-contrast leakage.

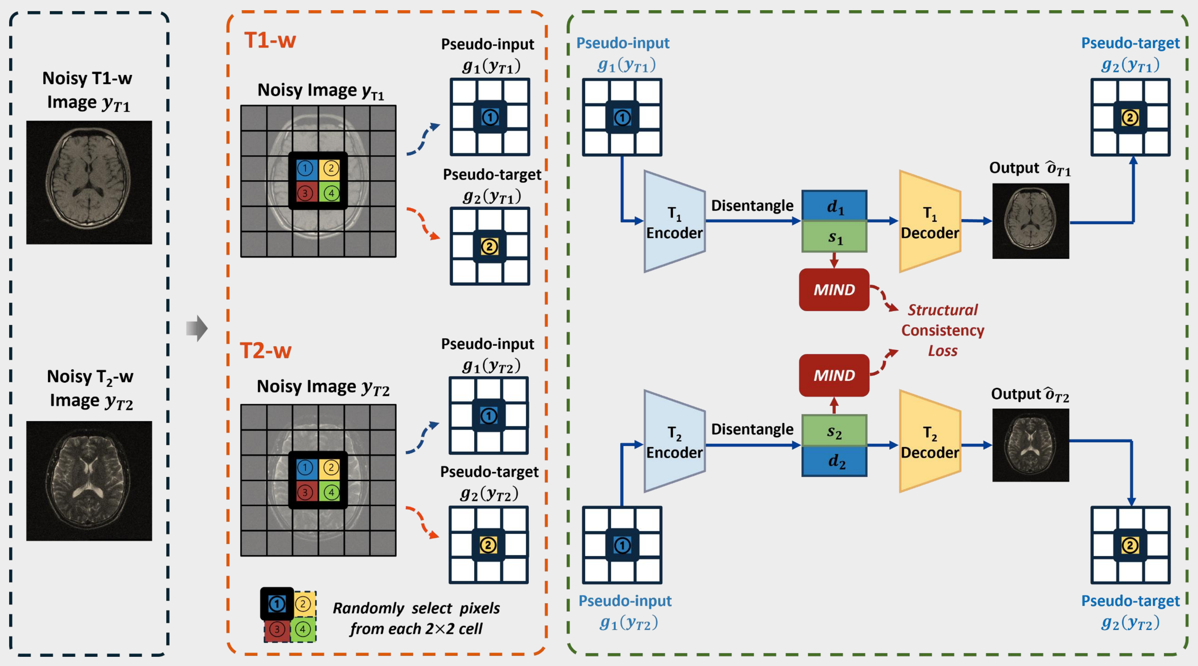

Figure 1. Overview of the self-supervised learning framework

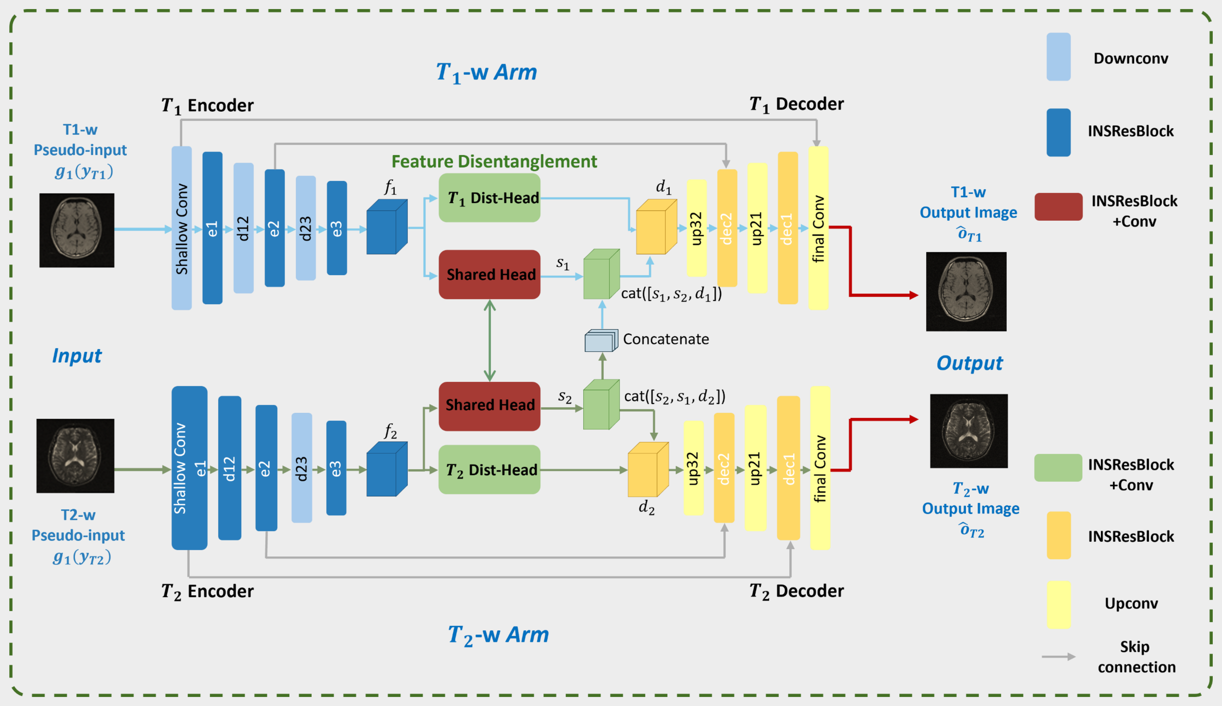

Figure 2. Architecture of the proposed dual-stream feature-disentanglement network

The model was trained using the Adam optimizer with an initial learning rate set to 1×10^{-3}, which decayed to 1×10^{-4} at the 100th epoch. Eight-fold data augmentation was performed. All experiments were conducted on NVIDIA GeForce RTX 5090, with CUDA 13.0 and the PyTorch environment.

Visual comparison between the proposed method and multiple state-of-the-art (SOTA) methods were performed. PSNR and SSIM metrics were used for the quantitative evaluation.

RESULTS

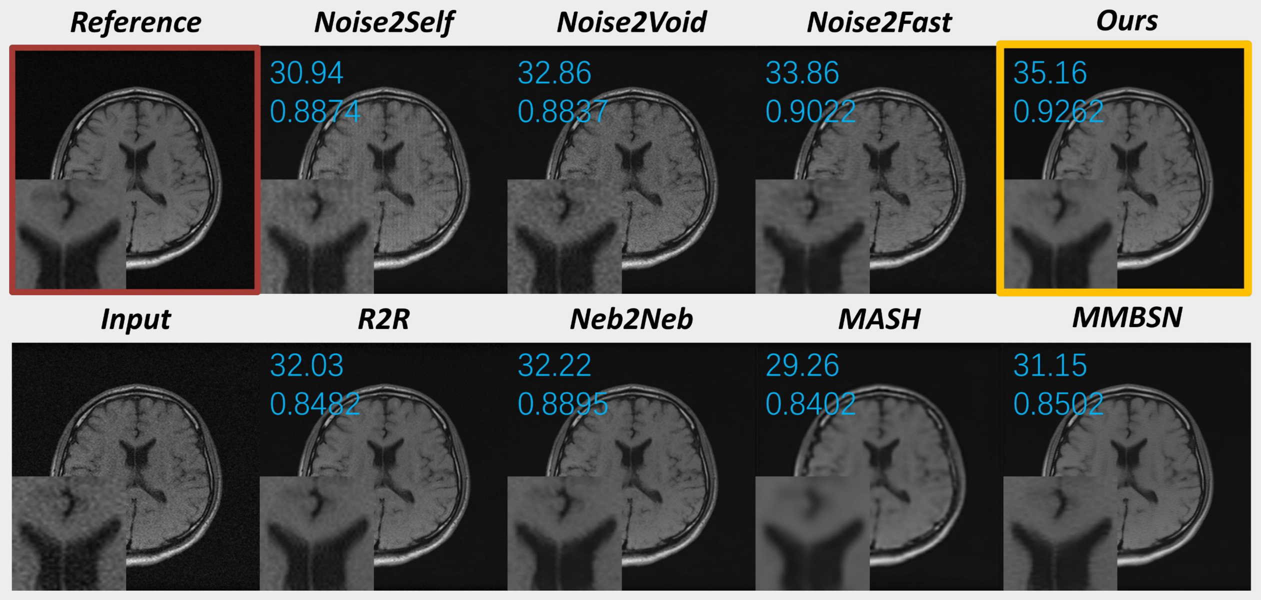

Figure 3. Visual comparison of de-noising performance of multiple methods (T1-w)

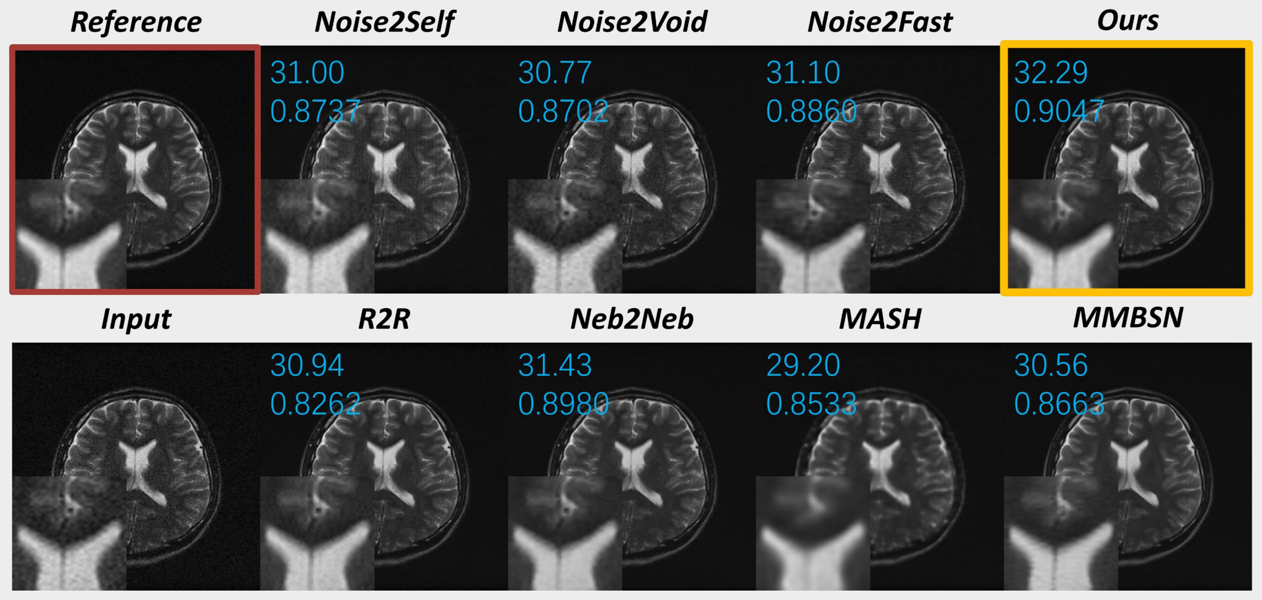

Figure 4. Visual comparison of de-noising performance of multiple methods (T2-w)

Visual comparison (**Figure 3 and 4**) shows that the proposed method was able to reconstruct images closer to the reference, without too much smoothing or unnatural texture patterns like appeared in some other methods.

Quantitative results (**Table 1**) demonstrates that the proposed method was able to achieve higher mean PSNR and SSIM, as compared to original low SNR images, as well as 7 SOTA methods, improving PSNR by ~0.9 – 1.1 dB, and SSIM by ~0.01, over the best performing SOTA method, for T1-w and T2-w MRI, respectively.

Table 1. Quantitative comparison (PSNR and SSIM) on the M4Raw Dataset (0.3T)

Method	T1-w	T2-w
Original	30.26 ± 2.25 / 0.792 ± 0.038	29.27 ± 1.39 / 0.746 ± 0.029
N2F	31.46 ± 2.50 / 0.885 ± 0.026	30.81 ± 1.62 / 0.862 ± 0.023
MASH	30.72 ± 2.09 / 0.875 ± 0.023	30.06 ± 1.28 / 0.841 ± 0.023
N2S	30.71 ± 2.25 / 0.876 ± 0.027	29.93 ± 1.55 / 0.853 ± 0.024
N2V	31.02 ± 2.46 / 0.865 ± 0.029	30.37 ± 1.58 / 0.844 ± 0.025
MMBSN	31.20 ± 2.29 / 0.873 ± 0.026	30.32 ± 1.36 / 0.842 ± 0.023
R2R	30.83 ± 2.14 / 0.795 ± 0.035	30.69 ± 1.63 / 0.830 ± 0.023
Neb2Neb	31.92 ± 2.72 / 0.903 ± 0.024	31.34 ± 1.74 / 0.880 ± 0.023
Ours	$\mathbf{32.77 \pm 2.96}$ / $\mathbf{0.915 \pm 0.025}$	$\mathbf{32.42 \pm 2.11}$ / $\mathbf{0.891 \pm 0.024}$

DISCUSSION

The main innovations of this study were as follows. A self-supervised learning framework was proposed where paired data is not a prerequisite for learning. Simultaneous multi-contrast MRI (e.g., T1-w, T2-w) SNR improvement were achieved which showed advantage over SOTA methods. Dual-arm feature disentanglement network and spatial structural consistency loss were embedded to better utilize the complementary and shared information between multi-contrast MRIs.

Future work will include expanding the current method to more image contrasts, and exploring other MRI image datasets.

CONCLUSIONS

Low-field MRI has received renewed clinical and research interest because of its lower cost, improved accessibility and assist by deep learning methods. In this study, a simultaneous multi-contrast guided self-supervised denoising framework for low-field MRI was proposed and investigated.

A higher PSNR and SSIM were achieved as compared to SOTA methods, without over-smoothing or unnatural noise textures. It suggests the potential of using self-supervised learning and multi-contrast MRI for the improvement of MRI image quality.

REFERENCES

1. Zhao, Y., et al. “Whole-body magenetic resonance imaging at 0.05 Tesla”, Science **384**, (2024)
2. Kimberly, W. T., et al. "Brain imaging with portable low-field MRI." Nature Reviews Bioengineering **1**, 617-630 (2023)
3. Lyu, M., et al. "M4Raw: A multi-contrast, multi-repetition, multi-channel MRI k-space dataset for low-field MRI research." Scientific Data **10**, 264 (2023)