2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 979-8-3503-2447-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/EMBC40787.2023.10340902

Content-Noise Feature Fusion Neural Network for Image Denoising in Magnetic Particle Imaging*

Tan Wang, Liwen Zhang, Zechen Wei, Yusong Shen, Jie Tian, *Fellow, IEEE* and Hui Hui, *Member, IEEE*

***Abstract*— Magnetic particle imaging (MPI) is a tomographic imaging method that quantitatively determines the distribution of magnetic nanoparticles (MNPs). However, the performance of MPI is primarily limited by the noise in the receive coil and electronic devices, which causes quantification errors for MPI images. Existing methods cannot efficiently eliminate noise while preserve structural details in MPI images. To address this problem, we propose a Content-Noise Feature Fusion Neural Network equipped with tailored modules of noise learning and content learning. It can simultaneously learn content and noise features of raw MPI images. Experimental results show that the proposed method outperforms the state-of-the-art methods on structural details preservation and image noise reduction of different levels.**

I. Introduction

Magnetic particle imaging (MPI) is a tomographic imaging method with the advantages of high sensitivity and linear quantitative ability [1]. Due to these characteristics, MPI has been implemented in various realms of biomedical research, like the detection of atherosclerotic plaque [2] and the presence of hepatocellular carcinomas [3]. Under combined static and dynamic magnetic fields, MPI obtains a spatially resolved magnetic nanoparticles (MNPs) imaging signal. This signal is further processed to obtain the spatial distribution of MNPs after performing an image reconstruction. However, there are noise and background signals in the measured raw MPI signals, which leads to distortion in the reconstructed image, thus making errors in the quantitative analysis [4].

Previous studies based on background subtraction aim to improve the signal-to-noise ratio (SNR) in the received signal, thus improving the quality of MPI images [5]. Due to the difficulty of modeling the statistical characteristics of MPI signals with complex noise, the existing methods for directly processing signals cannot simultaneously eliminate image noise while keeping structural details. Therefore, more effective methods to improve the quality of MPI images are required to be developed.

Recently, deep learning (DL) has proved to be effective in noise suppression for low-dose computed tomography (CT) [6], stripe elimination of light-sheet microscopic images [7]. Besides, Shang et al. proposed a deep-learning approach to improve the spatial resolution of MPI by fusing a dual-sampling convolutional neural network (FDS-MPI) [8]. These above methods have achieved good performance for medical image denoising, and they either learn content images (i.e., full-dose or high-resolution images) directly or learn noise images first and get the content images by subtracting. Both learning strategies have their inherent advantages: noise learning performs well in structural and contrast details preservation, while content learning shows more stable noise cancellation performance [9]. To apply the features of both the content and the noise to denoising MPI images, we proposed a Content-Noise Feature Fusion Neural Network (CNFFNet) equipped with tailored modules of noise learning and content learning in this work. Moreover, we verified the proposed model using a self-designed MPI image denoising dataset and authentic MPI images, the main contributions of this work are summarized as follows:

1) We proposed a CNFFNet to improve the quality of MPI images that simultaneously learns content and noise features and utilize image information.
2) We comprehensively investigated CNFFNet's denoising performance over the MPI image denoising dataset. Experiment results show that CNFFNet outperforms the state-of-the-art methods in structure preservation and noise reduction of different noise levels.
3) The proposed CNFFNet is validated on the authentic MPI image dataset. The results show that our method can deal with a larger MNPs concentration range than other methods.

II. METHOD

A.MPI Image Dataset

We used the simulated MPI images and authentic noise images to build our dataset. For MPI process simulation, we took simulated phantom images as the distributions of MNPs. The magnetization of MNPs in response to an excited magnetic field was obtained by the Langevin equation as follows [8]:

$$M(r,t) = M_0(r) \times \left[coth\left(\frac{m|B(r,t)|}{k_B T}\right) - \frac{k_B T}{m|B(r,t)|}\right] \quad (1)$$

*This work was supported in part by the National Natural Science Foundation of China under Grant: 62027901.

Tan Wang, Liwen Zhang, Zechen Wei, Jie Tian, and Hui Hui are with CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China (Corresponding author: Hui Hui, email: hui.hui@ia.ac.cn).

Tan Wang, Zechen Wei and Hui Hui are with University of Chinese Academy of Sciences, Beijing, 100080, China.

Tan Wang, Zechen Wei, Jie Tian, and Hui Hui are with Beijing Key Laboratory of Molecular Imaging, Beijing, 100190, China.

Yusong Shen is with School of Computer Science and Engineering, Southeast University, Nanjing, 211189, China.

Jie Tian is with Key Laboratory of Big Data-Based Precision Medicine (Beihang University), Ministry of Industry and Information Technology of the People's Republic of China, Beijing, 100191, People's Republic of China.

$$M_0(r) = \frac{N(r)m}{\Delta V}, \tag{2}$$

where $B(r,t)$ denotes the total magnetic field at position r and time t; m refers to the magnetic moment; $N(r)$ indicates the number of particles each contributing a magnetic moment $m = \frac{1}{6}\pi D^3 M_{sat}$; M_{sat} is the saturation magnetization. We used the X-space reconstruction method to process the imaging signals and got MPI images [10]. All the MPI simulation parameters are listed in Table I.

The workflow of the dataset construction is shown in Fig. 1. Firstly, we designed four kinds of phantoms, each phantom has four parameter vectors, including position, size, spacing, and intensity gradient change coefficient. The parameter vectors are dynamically changing in a certain range by a program based on OpenCV. We generated 4,000 phantoms of each kind through the program. Secondly, we took simulated phantoms as the input of the simulated X-space reconstruction algorithm [10]. The simulated reconstruction is carried out without noise, so its output was chosen as the ground truth label. Thirdly, we use the preclinical MPI scanner (MOMENTUM, Magnetic Insight Inc. Alameda, CA, USA) to scan phantoms without MNPs to obtain authentic MPI noise images. Fourthly, we randomly cropped the real noise images to a size of 128×128. Finally, noisy images were generated by randomly adding one of the cropped noise images to the clean MPI image. σ is a coefficient that represents the intensity of the added noise images. We constructed three datasets with the noise of different levels: (1) σ = 0.3, (2) σ = 0.25, (3) σ = 0.2. Each dataset contains 16,000 pairs of images.

TABLE I
PARAMETERES FOR MPI SIMULATION

Symbol	Parameters	Value	Unit
μ_0	Permeability of vacuum	$4\pi e^{-7}$	N A^{-2}
D	Nanoparticle diameter	20	nm
M_s	Saturation magnetization	$4.77e^{-5}$	Am^{-1}
m	Magnetic moment	$6.75e^{-18}$	Am^2
T	Kelvin temperature	293.15	K
K_B	Boltzmann constant	$1.38e^{-23}$	JK^{-1}
G	Gradient	5.7	Tm^{-1}
FOV	Field of view	20×20	mm^2

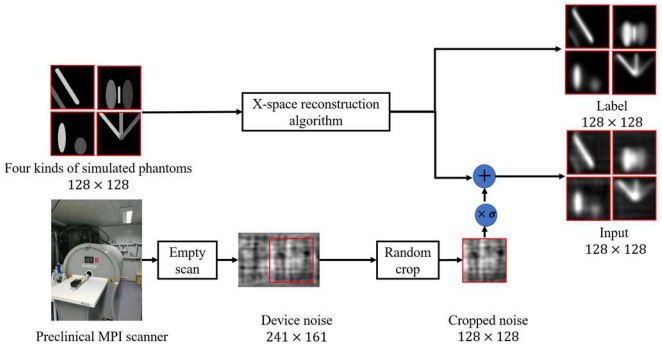

Fig. 1. Illustration of the workflow of dataset construction

B. Feature Fusion Learning

In the image domain, the degradation process of MPI can be expressed as:

$$I_{corrupted} = I_{content} + I_{noise} \tag{3}$$

Here, $I_{content}$ is the clean image that needs to be recovered, which precisely reveals the concentration distribution of MNPs. $I_{corrupted}$ is the noisy image. There are two learning strategies for existing image denoising methods based on deep learning. One learns a function to map the corrupted image to the content image directly, called content learning (CL):

$$I'_{content} = P_c(I_{corrupted}) \tag{4}$$

$$P_c(I_{corrupted}) = F^c_{fi}\left(F^c_{if}(I_{corrupted})\right) \tag{5}$$

$I'_{content}$ denotes the predicted content image, and P_c refers to the content prediction model composed of two parts: F^c_{if} extracts image features from $I_{corrupted}$, and F^c_{fi} recovers the content image from these image features. The other learns a residual mapping and gets the content image by subtracting the inputted image from the learned noise image, called noise learning (NL):

$$I'_{content} = I_{corrupted} - I'_{noise} \tag{6}$$

$$I'_{noise} = P_n(I_{corrupted}) = F^n_{fi}\left(F^n_{if}(I_{corrupted})\right), \tag{7}$$

where I'_{noise} refers to the predicted noise and P_n denotes the noise prediction model. F^n_{fi} and F^n_{if} are the corresponding image regressor and image feature extractor.

In order to make full use of the features from $I_{corrupted}$. We proposed a content-noise feature fusion neural network based on feature fusing learning for MPI image denoising. And the whole denoising process of CNFFNet can be expressed with a function H:

$$\begin{aligned} I'_{content} &= H(I_{corrupted}) \\ &= F^f_{if}(f(f'_{content}, f'_{noise})) \\ &= F^f_{if}\left(f\left(F^c_{if}(I_{corrupted}), F^n_{if}(I_{corrupted})\right)\right) \end{aligned} \tag{8}$$

$I_{corrupted}$ is used as the input of F^c_{if} and F^n_{if} to get content features $f'_{content}$ and noise features f'_{noise} respectively. Then, the fusion function f combines the $f'_{content}, f'_{noise}$, and translate the fused features to F^f_{if} to obtain the final output $I'_{content}$.

C. Content-Noise Feature Fusion Neural Network

The proposed CNFFNet, as shown in Fig. 2, mainly consists of three parts: feature extract module, feature fusion module, and image regress module.

The feature extraction module contains two identical feature extractors, which extract the input image's content features and noise features, respectively. Fig. 2(b) shows the detailed structure of our feature extract sub-network with the following components. Firstly, the extractor includes three down-sampling and up-sampling layers. Each layer consists of a 3 × 3 convolution operation, a batch normalization operation, and a dropout operation in case of overfitting. Secondly, down-sampling is conducted by a convolution operation to extract more features. Finally, instead of directly concatenating low-level and high-level features, we add a convolutional block attention module (CBAM) for each layer.

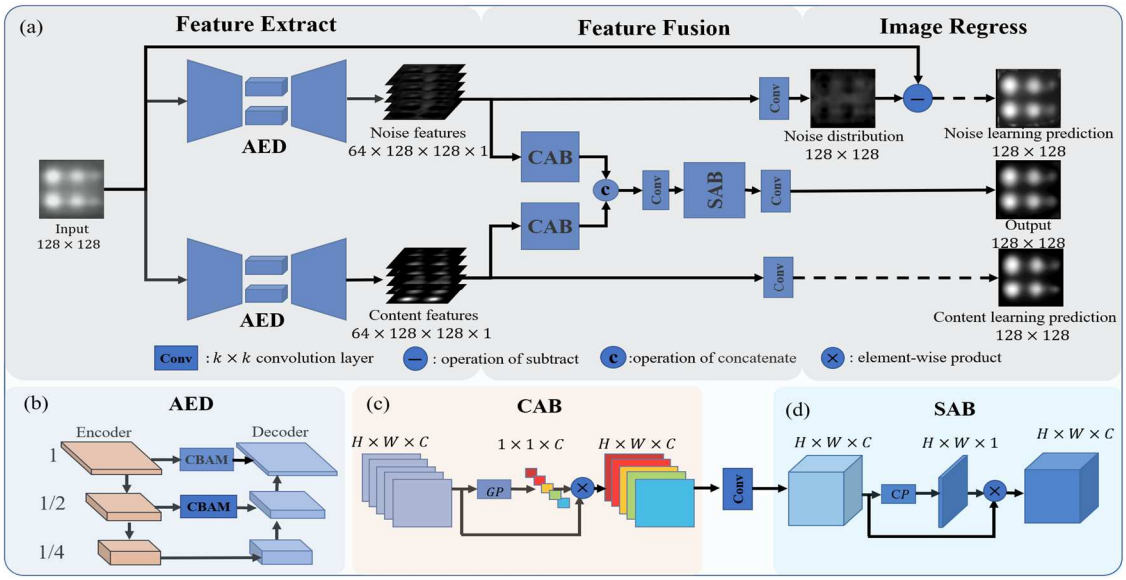

Fig. 2. Architecture and tailored modules of our proposed Content-Noise Feature Fusion Neural Network (CNFFNet). (a) Architecture of CNFFNet; (b) Attention Encoder-Decoder Feature Extractor architecture, CBAM denotes connect CAB and SAB in series; (c) Channel Attention Block architecture, GP denotes global average pooling. (d) Spatial Attention Block architecture, CP denotes channel average pooling.

Since the feature extraction module extracts two different kinds of features from the input image, we need to fuse them efficiently through the attention mechanism. As shown in Fig. 2(a), our feature fusion module consists of two channel attention blocks, two convolution layers and one spatial attention block. Firstly, we use the channel attention block to process the output of two feature extractors, respectively, and filter out important feature map channels. Secondly, we concatenate the output features directly and use a convolution layer to fuse the extracted features. Thirdly, a spatial attention block is conducted to determine the most critical regions in the feature maps. The fused feature maps will be the input of the image regress module and get the output. Furthermore, two convolution layers are added behind the two extractors as the corresponding regressors, and we trained our CNFFNet end-to-end with the following optimization problem:

$$\underset{H}{arg\,min}\left(\left\|\boldsymbol{I}_{NL\,prediction}-\boldsymbol{I}_{label}\right\|_2^2+\left\|\boldsymbol{I}_{Output}-\boldsymbol{I}_{label}\right\|_2^2+\left\|\boldsymbol{I}_{CL\,prediction}-\boldsymbol{I}_{label}\right\|_2^2\right) \quad (9)$$

III. Experiments & Results

A. Implementation details

In experiments, we randomly divided the three datasets in 7:1:2 to form the training, verification, and test sets. All the networks were trained on the training set (σ = 0.25). All the networks were tested on three test sets. All the networks were trained for 200 epochs using Adam optimizer with a batch size of 3. The initial learning rate was set at 0.003. For a quantitative assessment, three evaluation metrics were determined: root mean square error (RMSE), peak signal-to-noise ratio (PSNR), and structural similarity index measure (SSIM).

B. Ablation experiments

To verify the superiority of our proposed network, a series of ablation experiments were conducted over the simulated dataset (σ = 0.25). 1) the effectiveness of learning content and noise features simultaneously through the CNFFNet. 2) the effectiveness of the feature fusion module.

In detail, we removed the upper branch of CNFFNet to form Content Subnet (CSNet), which only extracted content features for MPI image denoising. We removed the lower branch of CNFFNet to form Noise Subnet (NSNet), which only extracted noise features for MPI image denoising. We removed the feature fusion module from the CNFFNet and formed No Fusion Net (NFNet). We compared CNFFNet with the networks above, and the experiment results are summarized in Table II. CNFFNet has achieved better performance than the CSNet and the NSNet in terms of RMSE, PSNR, and SSIM. Furthermore, the experiment results indicate that the feature fusion module is essential for CNFFNet. Without the feature fusion module, the performance of our proposed network decreased from 40.07 dB to 37.56 dB.

TABLE II
Evaluation Results of Ablation Study

	RMSE(× 10^{-2})	PSNR	SSIM
Input	12.61±1.60	18.05±1.05	0.24±0.12
CSNet	1.48±1.08	37.20±2.94	0.92±0.11
NSNet	1.16±0.48	39.24±2.96	0.92±0.06
NFNet	1.53±1.44	37.56±3.90	0.90±0.14
CNFFNet	**1.13±0.98**	**40.07±3.76**	**0.98±0.02**

C. Comparisons with existing denoising methods.

To evaluate the performance of CNFFNet, we compared it with three methods: fusing dual-sampling convolutional neural network (FDS-MPI) [8], residual encoder-decoder convolutional neural network (RCNN) [6], and residual channel attention network (RCAN) [11]. All the networks were trained on the training set (σ = 0.25) and tested on the three test sets. The experimental results are summarized in Table III, our network achieves better performance than other methods on three datasets with different noise levels, which shows that our method has good generalization. Furthermore, CNFFNet outperforms on structural details preservation. As shown in the red box in Fig. 3, the shape of the low-

concentration sample changes from a circle to a triangle after being processed by other methods.

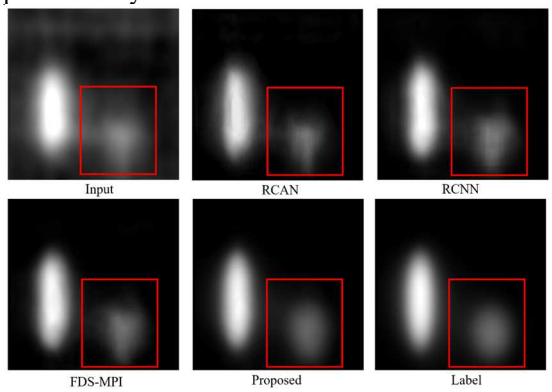

Fig. 3. Illustrations of the results from different methods

TABLE III
PERFORMANCE COMPARISON BASED ON DIFFERENT METHODS ON THREE TESTING DATASETS.

Noise level	Methods	RMSE($\times 10^{-2}$)	PSNR	SSIM
σ=0.3	Input	15.90 ± 2.09	16.04 ± 1.09	0.22 ± 0.08
	RCNN	2.88 ± 1.10	31.33 ± 2.93	0.84 ± 0.06
	RCAN	3.66 ± 1.24	29.14 ± 2.63	0.83 ± 0.07
	FDS-MPI	2.45 ± 1.09	32.96 ± 3.50	0.93 ± 0.05
	CNFFNet	$\mathbf{2.00 \pm 1.24}$	$\mathbf{35.17 \pm 4.38}$	$\mathbf{0.96 \pm 0.04}$
σ=0.25	Input	12.61±1.60	18.05±1.05	0.24±0.12
	RCNN	1.99±0.84	34.62±3.15	0.89±0.17
	RCAN	2.41±1.01	32.97±3.19	0.90±0.07
	FDS-MPI	1.89±0.88	35.21±3.54	0.95±0.04
	CNFFNet	**1.13±0.98**	**40.07±3.76**	**0.98±0.02**
σ=0.2	Input	9.56 ± 2.20	20.60 ± 1.85	0.28 ± 0.12
	RCNN	2.06 ± 0.85	34.33 ± 3.13	0.88 ± 0.07
	RCAN	2.53 ± 0.97	32.44 ± 2.80	0.91 ± 0.06
	FDS-MPI	1.87 ± 0.91	35.43 ± 3.75	0.95 ± 0.04
	CNFFNet	$\mathbf{1.52 \pm 1.40}$	$\mathbf{38.32 \pm 5.25}$	$\mathbf{0.97 \pm 0.04}$

D. *Phantom verification*

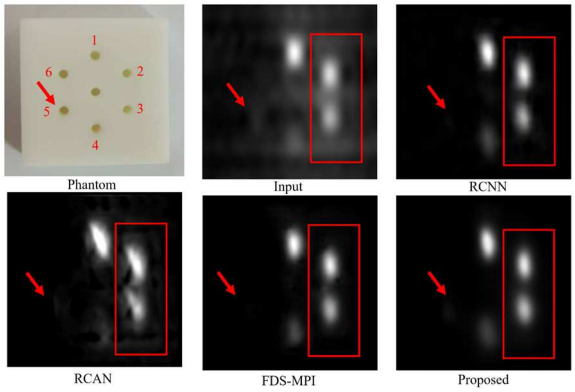

Fig. 4. Illustrations of the results of phantom verification.

TABLE IV
EVALUATION RESULTS OF PHANTOM VERIFICATION

Metric	Input	RCNN	RCAN	FDS-MPI	Proposed
RMSE($\times 10^{-2}$)	14.18	7.66	9.87	7.38	7.01

To evaluate the performance of CNFFNet on the real MPI images, we supplemented the experiment with a phantom that consists of six cylinders of equal volume of $25\mu L$, as shown in Fig. 4. The concentration of MNPs in the first cylinder is 10 $mg \cdot mL^{-1}$, each of the remaining cylinders has a concentration half of the previous one. The experimental results are shown in the Fig. 4. There is some noise between samples of different concentrations in the images processed by other methods, as shown by the red box in the figures. We simulate the ground truth image according to the shape of phantom, and the evaluation results are summarized in Table IV. CNFFNet performs better than other methods on the real MPI image. Moreover, CNFFNet has a larger concentration processing range than other methods. It can roughly recover the shape of the fifth particle sample, while other methods remove it as noise, as shown by the red arrow in the figures.

IV. CONCLUSION

In this study, we proposed a content-noise feature fusion neural network for MPI image denoising. CNFFNet is equipped with tailored modules of noise learning and content learning to utilize image information from corrupted MPI images. The advantage of CNFFNet is that it performs well in structures preservation and can handle noise of different levels. Furthermore, it is validated over the authentic MPI image dataset.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant: 62027901.

REFERENCES

[1] B. Gleich and J. Weizenecker, "Tomographic imaging using the nonlinear response of magnetic particles," Nature, vol. 435, no. 7046, pp. 1214-7, Jun 30 2005.

[2] W. Tong et al., "Highly sensitive magnetic particle imaging of vulnerable atherosclerotic plaque with active myeloperoxidase-targeted nanoparticles," Theranostics, vol. 11, no. 2, p. 506, 2021.

[3] J. Zhuo et al., "Enhanced glypican-3-targeted identification of hepatocellular carcinoma with liver fibrosis by pre-degrading excess fibrotic collagen," Acta Biomaterialia, 2023.

[4] H. Paysen, O. Kosch, J. Wells, N. Loewa, and F. Wiekhorst, "Characterization of noise and background signals in a magnetic particle imaging system," Phys Med Biol, vol. 65, no. 23, Nov 27 2020.

[5] T. Knopp, M. Grosser, M. Graeser, T. Gerkmann, and M. Moeddel, "Efficient Joint Estimation of Tracer Distribution and Background Signals in Magnetic Particle Imaging Using a Dictionary Approach," Ieee Transactions on Medical Imaging, Article vol. 40, no. 12, pp. 3568-3579, Dec 2021.

[6] H. Chen et al., "Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network," IEEE Trans Med Imaging, vol. 36, no. 12, pp. 2524-2535, Dec 2017.

[7] Z. Wei et al., "Elimination of stripe artifacts in light sheet fluorescence microscopy using an attention-based residual neural network," Biomedical Optics Express, vol. 13, no. 3, pp. 1292-1311, 2022.

[8] Y. Shang et al., "Deep learning for improving the spatial resolution of magnetic particle imaging," Phys Med Biol, vol. 67, no. 12, Jun 10 2022.

[9] M. Geng et al., "Content-Noise Complementary Learning for Medical Image Denoising," IEEE Trans Med Imaging, vol. 41, no. 2, pp. 407-419, Feb 2022.

[10] Y. Shen, C. Hu, P. Zhang, J. Tian, and H. Hui, "A novel software framework for magnetic particle imaging reconstruction," International Journal of Imaging Systems and Technology, vol. 32, no. 4, pp. 1119-1132, 2022.

[11] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in Proceedings of the European conference on computer vision (ECCV), 2018, pp. 286-301.