MRI Motion Artifact Correction via Frequency-Assisted Artifact Disentanglement and Confidence-Guided Knowledge Distillation

Jiazhen Wang[1], Heran Yang[1], Yizhe Yang[1], and Jian Sun[1,2(✉)]

[1] Xi'an Jiaotong University, Xi'an, China
{jzwang,yyz0022}@stu.xjtu.edu.cn, {hryang,jiansun}@xjtu.edu.cn
[2] State Industry-Education Integration Center for Medical Innovations at Xi'an Jiaotong University, Xi'an, China

Abstract. Motion artifacts degrade MR image quality affecting clinical diagnoses. Although deep learning-based motion artifact correction (MAC) methods show promise, they are limited by the lack of real paired motion-corrupted and motion-free images. We propose a novel frequency-assisted artifact disentanglement learning framework for MAC of MR images. Our approach integrates a frequency-decomposed motion correction network (FDMC-Net) for content-artifact disentanglement over the real unpaired data, coupled with confidence-guided knowledge distillation using simulated paired data. Specifically, considering that motion artifacts are primarily caused by high-frequency k-space misalignment, FDMC-Net decomposes motion-corrupted MR images into low-frequency and high-frequency components and then employs dedicated encoders to disentangle content and artifact features. FDMC-Net is trained by unsupervised cycle-consistent adversarial loss over realistic unpaired data, and confidence-guided knowledge distillation loss by distilling a teacher model trained on simulated paired data. Experiments demonstrate its state-of-the-art performance, with ablation studies confirming the effectiveness of frequency-assisted disentanglement and confidence-guided distillation.

Keywords: Motion artifact correction · Frequency-assisted artifact disentanglement · Confidence-guided knowledge distillation.

1 Introduction

Magnetic resonance imaging (MRI) is a non-invasive, radiation-free technique providing high-resolution images of soft tissues. However, the prolonged acquisition time and involuntary patient movement during scanning often introduce various motion artifacts, such as ghosting and blurring [15]. These artifacts significantly degrade image quality, potentially obscuring critical diagnostic information and negatively impacting clinical decision-making for radiologists.

To mitigate motion artifacts, various prospective and retrospective methods have been proposed. Prospective methods utilize external tracking devices or

navigators for real-time motion estimation, but they require additional hardware and scanner modifications [8]. In contrast, retrospective artifact correction (RAC) applies motion correction during image reconstruction after data acquisition, eliminating the need for extra hardware [15]. While CNN-based RAC methods [2, 5, 6, 9, 10, 14, 16–18] have shown promise in removing motion artifact, their supervised training requires paired motion-corrupted and motion-free images, which are difficult to obtain in practice. Simulated artifacts provide an alternative but may not fully replicate real-world motion distortions.

To address the lack of real-paired data, many unsupervised generative networks have been proposed using unpaired data [4, 13, 20–22]. The CycleGAN [22] architecture consists of two generators: one corrupts a motion-free image, while the other corrects an unpaired motion-corrupted image using cycle consistency loss, along with adversarial loss distinguishing between synthetic and real images. Recently, several methods have been proposed to enhance CycleGAN for motion correction [4, 11, 20]. Liu et al. [11] introduced DUNCAN, a disentangled unsupervised cycle-consistent adversarial network, to separate content and artifacts. Multi-Net [7] is a multi-task framework for motion artifact correction (MAC) and tissue segmentation. DCGAN-MS [4] uses multi-mask k-space subsampling with CycleGAN to reduce the complexity of motion artifact. Despite promise, they still struggle with handling complex real-world artifacts and preserving anatomical structures.

In this paper, we propose a frequency-assisted artifact disentanglement framework for MR image motion artifact correction (MAC). Our approach integrates a frequency-decomposed motion correction network (FDMC-Net) for content-artifact disentanglement in low-frequency (LF) and high-frequency (HF) components and employs confidence-guided knowledge distillation for synthetic-to-real knowledge transfer. Specifically, FDMC-Net employs a learnable frequency decoupling module to decompose motion-corrupted MR images into HF and LF components, leveraging the prior knowledge that motion artifacts predominantly arise from high-frequency k-space misalignments. These LF and HF components are then processed by specialized encoders to disentangle anatomical content from motion artifacts, enabling targeted suppression of artifacts while preserving structural fidelity. Finally, the LF and HF content features are fused to reconstruct motion-free MR images. To overcome the lack of real paired data, we combine cycle-consistent adversarial learning on real unpaired data with confidence-guided knowledge distillation (CGKD) using simulated pseudo-pairs. In CGKD, a teacher model trained on simulated data generates pseudo-labels, while a confidence map prioritizes reliable regions to guide the FDMC-Net during distillation. Experiments conducted on simulated and real motion-corrupted MRI datasets show that our method outperforms state-of-the-art approaches.

2 General Approach and Network Architecture

Given an unpaired motion-free MR image I_{mf} and a motion-corrupted MR image I_{mc}, we propose a frequency-decomposed motion correction network(FDMC-

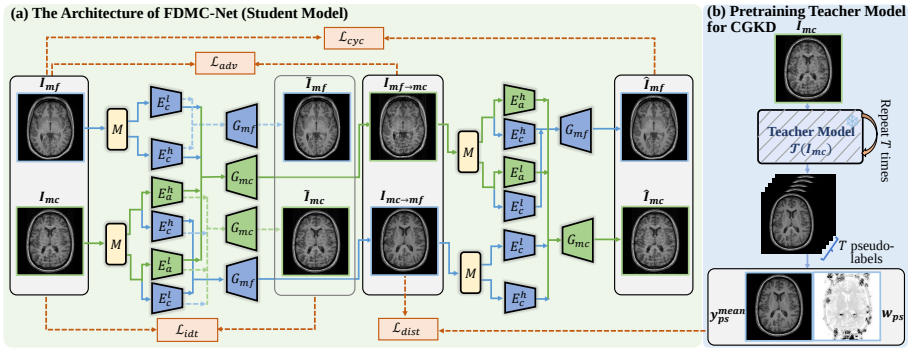

Fig. 1: (a) The architecture of the frequency-decomposed motion correction network (FDMC-Net). (b) The teacher model pre-trained on simulated paired data for confidence-guided knowledge distillation (CGKD).

Net) for correcting motion artifact in MRI. As shown in Fig. 1(a), FDMC-Net decomposes motion-corrupted image I_{mc} into low-frequency (LF) and high-frequency (HF) components, then disentangles artifact and content features separately within both LF and HF components using specialized encoders. The LF and HF content features are fused to reconstruct motion-free images, while content features from the motion-free image I_{mf} are combined with artifact features from I_{mc} to generate motion-corrupted images.

2.1 Network Architecture

The architecture of FDMC-Net is shown in Fig. 1(a), comprising key components such as: frequency decomposed module M, content encoders (i.e., E_c^l, E_c^h), artifact encoders (i.e., E_a^l, E_a^h), and decoders (i.e., G_{mf}, G_{mc}). Specifically, given an unpaired motion-corrupted image I_{mc} and a motion-free image I_{mf}, FDMC-Net first decomposes both images into their corresponding low-frequency (LF) and high-frequency (HF) components using the frequency decomposed module M:

$$F_{mc}^l, F_{mc}^h = M(I_{mc}), \quad F_{mf}^l, F_{mf}^h = M(I_{mf}), \tag{1}$$

where F_{mc}^l and F_{mc}^h are the LF and HF components of the features from I_{mc}, and F_{mf}^l and F_{mf}^h are the LF and HF components of the features from I_{mf}.

For the LF and HF components of the motion-corrupted MR image I_{mc}, we use the LF and HF content encoders (i.e., E_c^l, E_c^h), and LF and HF artifact encoders (i.e., E_a^l, E_a^h) to encode and disentangle the corresponding content features (i.e., C_{mc}^l, C_{mc}^h) and artifact features (i.e., A_{mc}^l, A_{mc}^h), respectively:

$$C_{mc}^l = E_c^l(F_{mc}^l), \quad A_{mc}^l = E_a^l(F_{mc}^l), \quad C_{mc}^h = E_c^h(F_{mc}^h), \quad A_{mc}^h = E_a^h(F_{mc}^h). \tag{2}$$

Similarly, for the LF and HF components of motion-free image I_{mf}, we only apply the LF and HF content encoders E_c^l and E_c^h to extract the content features:

$$C_{mf}^l = E_c^l(F_{mf}^l), \quad C_{mf}^h = E_c^h(F_{mf}^h). \tag{3}$$

Next, the LF and HF content features of both the motion-corrupted image I_{mc} and the motion-free image I_{mf} are passed to the motion-free decoder G_{mf}, which reconstructs the corresponding motion-free MR images:

$$I_{mc\to mf} = G_{mf}(C_{mc}^l, C_{mc}^h), \quad \tilde{I}_{mf} = G_{mf}(C_{mf}^l, C_{mf}^h), \tag{4}$$

where $I_{mc\to mf}$ is the reconstructed motion-free version of I_{mc}, and $\tilde{I}_{mf}$ is the self-reconstructed motion-free image of I_{mf}. Conversely, the LF and HF artifact features are combined with the LF and HF content features to reconstruct the motion-corrupted MR images:

$$\tilde{I}_{mc} = G_{mc}(C_{mc}^l, C_{mc}^h, A_{mc}^l, A_{mc}^h), \quad I_{mf\to mc} = G_{mc}(C_{mf}^l, C_{mf}^h, A_{mc}^l, A_{mc}^h), \tag{5}$$

where $I_{mf\to mc}$ is a synthesized motion-corrupted image, combining artifacts from I_{mc} and content from I_{mf}. $\tilde{I}_{mc}$ is the self-reconstructed image of I_{mc}.

Furthermore, FDMC-Net incorporates a cycle consistency structure to ensure reversible transformations between motion-free and motion-corrupted images. As shown in the right side of Fig. 1(a), we apply the frequency decomposition and content-artifact disentanglement to the generated motion-corrupted image $I_{mf\to mc}$ and motion-free image $I_{mc\to mf}$, and then decode them to reconstruct motion-free image $\hat{I}_{mf}$ and motion-corrupted image $\hat{I}_{mc}$.

Frequency Decomposed Module M. This module M is designed to decompose both motion-free and motion-corrupted MR images into LF components and HF components by using dynamically gated learnable filters for each spatial location. Specifically, a convolution block B, consisting of two convolutional layers with ReLU activation, is applied to extract shallow features $F_{mc} = B(I_{mc}) \in \mathbb{R}^{H\times W\times C}$. Then, a gated low-pass filter K_{lp} is learned by

$$K_{lp} = \varphi(\hat{F}_{mc} \odot \sigma(\mathrm{Conv}_{1\times1}(\hat{F}_{mc}))), \ \hat{F}_{mc} = \mathrm{Conv}_{1\times1}(F_{mc}), \tag{6}$$

where $K_{lp} \in \mathbb{R}^{H\times W\times k^2}$, with $k = 3$ representing the kernel size of the learned filter. $\mathrm{Conv}_{1\times1}$ is a 1×1 convolution layer, and σ is a Sigmoid operation. φ is a Softmax operation that ensures the generated filter is a low-pass filter [23]. The low-pass filter is then applied to the shallow features F_{mc} to obtain the low-frequency components F_{mc}^l and high-frequency components F_{mc}^h:

$$F_{mc}^l(i,j) = \sum_u \sum_v \bar{K}_{lp}(i,j,u,v) F_{mc}(i+u, j+v), \ F_{mc}^h = F_{mc} - F_{mc}^l, \tag{7}$$

where $\bar{K}_{lp}$ is the reshaped filter of K_{lp}, i and j represent the spatial coordinates, and $u, v \in \{-1, 0, 1\}$ denote the surrounding locations.

Encoder. The encoder consists of four subnetworks: a LF content encoder E_c^l, a LF artifact encoder E_a^l, a HF content encoder E_c^h, and a HF artifact encoder

E_a^h. These encoders share the same architecture and network parameters settings used in the SOTA unsupervised MAC method, DCGAN-MS [4].
Decoder. The decoder comprises two subnetworks: a motion-free decoder G_{mf} and a motion-corrupted decoder G_{mc}. Except using 1×1 convolutional layers with varying channel numbers to match the channel dimensions, the remaining architecture and network parameter settings follow DCGAN-MS [4].

3 Training Method

To ensure that the outputs of FDMC-Net align with the desired MR images for effective content-artifact disentanglement, we design two loss functions to train the FDMC-Net, as shown in Fig. 1. First, a cycle-consistency adversarial loss $\mathcal{L}_{disc}$ is applied using real unpaired data. To improve reconstruction accuracy, we introduce a confidence-guided knowledge distillation loss $\mathcal{L}_{dist}$, which uses confidence-weighted pseudo-labels from a teacher model trained on simulated paired data to guide FDMC-Net in reconstructing the motion-free MR image $I_{mc\to mf}$. The total loss function is formulated as:

$$\mathcal{L}_{total} = \mathcal{L}_{disc} + \lambda_{dist}\mathcal{L}_{dist}, \tag{8}$$

where λ_{dist} controls the weight of the confidence-guided distillation loss, set to 10 empirically. The influence of different λ_{dist} values is evaluated in Sec. 4.5.

3.1 Cycle-consistency Adversarial Training Loss

The cycle-consistency adversarial training loss consists of three components including adversarial loss $\mathcal{L}_{adv}$, identity loss $\mathcal{L}_{idt}$, and cycle-consistency loss $\mathcal{L}_{cyc}$:

$$\mathcal{L}_{disc} = \mathcal{L}_{adv} + \lambda_{idt}\mathcal{L}_{idt} + \lambda_{cyc}\mathcal{L}_{cyc}, \tag{9}$$

where λ_{idt} and λ_{cyc} are set to 20 as in paper [4].
Adversarial Loss. Since no real paired motion-free image I_{mf} and motion-corrupted image I_{mc} are provided, we employ discriminator D_{mc} and D_{mf} to distinguish real motion-corrupted images and motion-free images from generated ones. The adversarial loss $\mathcal{L}_{adv}$ is defined as follows:

$$\begin{aligned}\mathcal{L}_{adv} = {} & \mathbb{E}[\log D_{mc}(I_{mc})] + \mathbb{E}[1 - \log D_{mc}(I_{mf\to mc})] \\ & + \mathbb{E}[\log D_{mf}(I_{mf})] + \mathbb{E}[1 - \log D_{mf}(I_{mc\to mf})].\end{aligned} \tag{10}$$

Identity Loss. The identity loss $\mathcal{L}_{idt}$ is designed to enforce consistency between the self-reconstructed image and input image: $\mathcal{L}_{idt} = \|\tilde{I}_{mf} - I_{mf}\|_1 + \|\tilde{I}_{mc} - I_{mc}\|_1$.
Cycle-consistency Loss. The cycle consistency loss is introduced to ensure the consistency between the MR image reconstructed by cyclic transformation and the input MR image: $\mathcal{L}_{cyc} = \|\hat{I}_{mf} - I_{mf}\|_1 + \|\hat{I}_{mc} - I_{mc}\|_1$.

3.2 Confidence-guided Knowledge Distillation Training Loss

We further outline the teacher model training on synthetic paired data, and present the derived confidence-guided knowledge distillation loss.

Supervised Training on Synthetic Paired Data. The teacher model is trained with supervised learning on synthetic paired data, where motion-corrupted images are generated using the forward model $I_{sim} = M\mathcal{F}T_\theta I_{mf}$ used in [1, 3, 18]. The teacher learns to map these motion-corrupted images I_{sim} to their corresponding motion-free versions I_{mf}. The L1 loss minimizes the difference between predictions and ground truth, allowing the teacher model to generate pseudo-labels for real unpaired motion-corrupted data.

Architecture of Teacher Network. The teacher model $\mathcal{T}$ consist of an encoder $\mathcal{E}$, a dynamic mixture-of-experts module $\mathcal{M}$, and a decoder $\mathcal{D}$. The encoder has three convolution blocks and four residual blocks with two downsampling layers. The decoder comprises four residual blocks and three convolution blocks with two upsampling layers. The dynamic mixture-of-experts module $\mathcal{M}$ leverages Gumbel-Softmax for stochastic expert selection, enabling differentiable probabilistic routing based on input features:

$$\begin{aligned} \mathcal{M}(F) &= \sum\nolimits_{i=1}^{N} \pi_i(F)\mathcal{R}_i(F), \\ \pi \triangleq [\pi_1(F), \cdots, \pi_N(F)] &= \text{GumbelSoftmax}(\text{FC}(\text{GMP}(F)+\text{GAP}(F))), \end{aligned} \tag{11}$$

where GMP and GAP denote global max and average poolings. FC is a fully connected layer, and π_i is the weight of expert i. $\mathcal{R}_i$ is the expert network i, which consists of two convolutional layers and a ReLU activation. The teacher model is summarized as: $y_{ps} = \mathcal{D} \circ \mathcal{M} \circ \mathcal{E}(I_{mc})$.

Confidence-guided Knowledge Distillation. Due to the domain gap between synthetic and real data, pseudo-labels generated by the teacher model may be inaccurate, affecting FDMC-Net training. To address this, we introduce confidence-guided knowledge distillation loss. Specifically, $\mathcal{M}$ uses stochastic expert selection in Eqn. (11), which can generate T pseudo-labels (e.g., y_{ps}^1, $\cdots$, y_{ps}^T) for each real unpaired motion-corrupted image I_{mc} by running the teacher model T times. We then compute the pixel-wise mean and variance of these pseudo-labels: $y_{ps}^{mean} = \frac{1}{T}\sum_{i=1}^{T} y_{ps}^i$, $y_{ps}^{var} = \frac{1}{T}\sum_{i=1}^{T}\left(y_{ps}^i - y_{ps}^{mean}\right)^2$. Next, we define the pixel-wise confidence map w_{ps} as $w_{ps} = 1 - \text{Sigmoid}\left(y_{ps}^{var}/\rho\right)$, where ρ is a hyper-parameter set empirically as 0.0001. This confidence map reflects the uncertainty in the pseudo-labels, allowing the model to focus more on reliable regions during training. Finally, we define the confidence-guided distillation loss between the generated motion-free MR image $I_{mc\to mf}$ of FDMC-Net and the averaged pseudo-labels y_{ps}^{mean}, weighted by the confidence map w_{ps}:

$$\mathcal{L}_{dist} = \left\| w_{ps} \odot I_{mc\to mf} - w_{ps} \odot y_{ps}^{mean} \right\|_1. \tag{12}$$

4 Experiments

Dataset. We evaluate our method on the MR-ART dataset, which contains T1-weighted 3D structural MRI images from 148 healthy adults, provided by

Table 1: Quantitative comparison of different methods on MR-ART dataset under simulated motion(SM1 and SM2) and real motion(RM1 and RM2).

Methods	SM1		SM2		RM1		RM2	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Corrupted	30.85	0.8523	25.76	0.8795	30.80	0.9185	30.02	0.8756
CycleGAN	29.32	0.9295	25.83	0.8729	29.58	0.9041	28.99	0.8634
DUNCAN	31.89	0.9542	25.96	0.8745	31.17	0.9526	30.12	0.9390
UDDN	32.02	0.9547	26.05	0.8746	31.43	0.9533	30.57	0.9426
UNAEN	31.65	0.9536	25.86	0.8731	31.03	0.9522	30.16	0.9391
DCGAN-MS	32.54	0.9578	26.44	0.8762	31.99	0.9546	31.03	0.9457
Ours	**33.98**	**0.9713**	**27.63**	**0.8814**	**33.02**	**0.9632**	**32.26**	**0.9569**

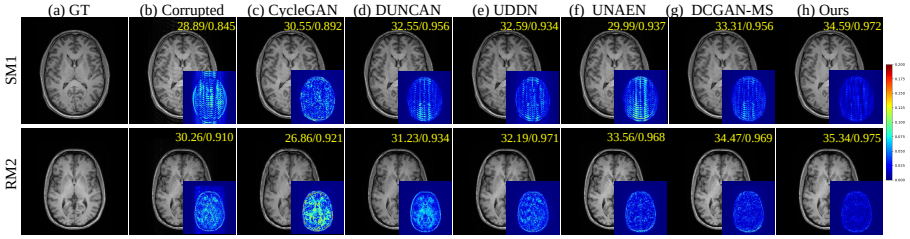

Fig. 2: Qualitative comparison of different methods on SM1 and RM2.

OpenNeuro [12]. The data is acquired using a 3T MRI scanner with the following parameters: TE = 3 ms, TR = 2300 ms, flip angle = 9°, and FOV = 256 × 256 mm. The dataset is split into a training set with 108 subjects and a test set with 40 subjects. Each subject has a motion-free data, two real motion-corrupted data (moderate real motion (RM1) and heavy real motion (RM2)), and two simulated motion-corrupted data (SM1 and SM2). Specifically, SM1 ("pitch15dur2p5nnods10") simulates 10 nods with fixed 15° pitch rotation. SM2 ("rot0to15nnods5") simulates 5 nods with random 0-15° rotation across the pitch, yaw, and roll axis. For training, we use motion-corrupted data from the first 50 subjects and motion-free data from the last 50 subjects to ensure unpaired data. In addition, SyN registration by ANTs is applied to roughly align each pair of motion-free and real motion-corrupted data.

Implementation Details. We implement our model by PyTorch, and train it on an NVIDIA RTX 3090 GPU for 50 epochs using the Adam with a learning rate of 1×10^{-4} and a batch size of 1. We take PSNR and SSIM [19] as metrics.

Performance Evaluation. We compare the FDMC-Net with state-of-the-art methods, including CycleGAN [22], DUNCAN [11], UDDN [20], UNAEN [21], and DCGAN-MS [4]. The motion-corrupted image is denoted as "Corrupted". In Table 1, we show the quantitative results of different methods on simulated data (SM1, SM2) and real data (RM1, RM2) with different motion severities. While most methods improve reconstruction over "Corrupted", our approach consistently achieves the best performance, attributed to precise content-artifact disentanglement and knowledge transfer from the teacher model trained on simulated

Table 2: The ablation study of the key components (i.e., FDM and CGKD) of our FDMC-Net on SM1.

Methods	PSNR	SSIM
Baseline	31.78	0.9534
w/o CGKD	33.03	0.9632
Ours ($w_{ps} = \mathbf{1}$)	33.36	0.9650
Ours	**33.98**	**0.9713**

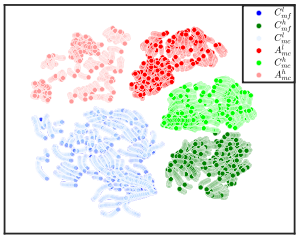

Fig. 3: T-SNE visualization of features in Eqn.(2) and Eqn.(3).

Table 3: Effect of CGKD on SM1.

Methods	PSNR	SSIM
$\lambda_{dist} = 1$	32.85	0.9620
$\lambda_{dist} = 5$	33.57	0.9672
$\lambda_{dist} = 10$	**33.98**	**0.9713**
$\lambda_{dist} = 20$	33.68	0.9679

Table 4: Effect of T on SM1.

Methods	PSNR	SSIM
$T = 2$	33.59	0.9683
$T = 4$	**33.98**	**0.9713**
$T = 6$	33.73	0.9690
$T = 8$	33.37	0.9661

data. Figure 2 further validates the superiority of our method, showing the lowest reconstruction error and the best recovery of fine image details. Additionally, the t-SNE visualization in Fig. 3 confirms effective content-artifact disentanglement, with well-separated content and artifact features. Notably, the LF content features of motion-free and motion-corrupted images exhibit substantial overlap, indicating that LF content remains stable despite motion artifacts.

Effectiveness of Key Components. We evaluate the effectiveness of key components, including frequency decomposed module (FDM) and confidence-guided knowledge distillation (CGKD), on SM1. In Table 2, (A) "Baseline" refers to FDMC-Net without FDM and CGKD; (B) "w/o CGKD" indicates FDMC-Net without CGKD; (C) "Ours ($w_{ps} = \mathbf{1}$)" represents FDMC-Net with a single pseudo-label generated by the teacher model; (D) "Ours" is our full FDMC-Net with both FDM and CGKD. As shown in Table 2, our method (i.e., "Ours") outperforms all variants, demonstrating the effectiveness of both FDM and CGKD. Additionally, by using the confidence map in Eqn. (12), our method improves the reconstruction accuracy compared to "Ours ($w_{ps} = \mathbf{1}$)".

Effect of CGKD. Table 3 shows the effect of the confidence-guided knowledge distillation loss in Eqn. (8) by varying λ_{dist}. Our method achieves the best when $\lambda_{dist} = 10$, and consistently better than competitors when $\lambda_{dist} = 1, 5, 10, 20$.

Effect of Number of Generated Pseudo-labels. Table 4 evaluates the effect of the number of pseudo-labels T generated by the teacher model on FDMC-Net training on SM1. From these results, our method achieves the best performance at $T = 4$, and consistently outperforms the competitors shown in Table 4.

5 Conclusion

This paper introduced a novel frequency-assisted artifact disentanglement learning framework to correct the MR image motion artifacts. It is based on a frequency decomposed motion correction network (FDMC-Net) for content-artifact disentanglement on real unpaired data and additionally trained by confidence-guided knowledge distillation on simulated paired data. Experiments on simulated and real datasets show the superiority of the proposed FDMC-Net. In future, we will apply the method to more real datasets, e.g., abdominal dataset.

Acknowledgments. This work is supported by National Natural Science Foundation of China (12426313, 12125104, 12090021, 62206216), Key-Area Research and Development Program of Guangdong Province (2022B0303020003).

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

References

1. Alexander, L., Hannes, N., Rolf, Pohmannand Bernhard, S.: Blind retrospective motion correction of mr images. Magn. Reson. Med. **70**(6), 1608–1618 (2013)
2. Armanious, K., Jiang, C., Fischer, M., Küstner, T., Hepp, T., Nikolaou, K., Gatidis, S., Yang, B.: Medgan: Medical image translation using gans. Comput. Med. Imaging. Graph. **79**, 101684 (2020)
3. Ben, A.D., et al: Retrospective motion artifact correction of structural mri images using deep learning improves the quality of cortical surface reconstructions. Neuroimage **230**, 117756 (2021)
4. Chen, G., Xie, H., Rao, X., Liu, X., Otikovs, M., Frydman, L., Sun, P., Zhang, Z., Pan, F., Yang, L., et al.: Mri motion correction through disentangled cyclegan based on multi-mask k-space subsampling. IEEE Trans. Med. Imaging (2024)
5. Haskell, M.W., et al: Network accelerated motion estimation and reduction (namer): Convolutional neural network guided retrospective motion correction using a separable motion model. Magn. Reson. Med. **82**(4), 1452–1461 (2019)
6. Junchi, L., Mehmet, K., Mark, S., Jie, D.: Motion artifacts reduction in brain mri by means of a deep residual network with densely connected multi-resolution blocks (drn-dcmb). Magn. Reson. Imaging. **71**, 69–79 (2020)
7. Jung, S., Choi, Y., Al-masni, M.A., Jung, M., Kim, D.H.: Deformation-aware segmentation network robust to motion artifacts for brain tissue segmentation using disentanglement learning. In: MICCAI. pp. 213–222 (2024)
8. Kay, N., Peter, B.: Prospective correction of affine motion for arbitrary mr sequences on a clinical scanner. Magn. Reson. Med. **54**(5), 1130–1138 (2005)
9. Kuzmina, E., Razumov, A., Rogov, O.Y., Adalsteinsson, E., White, J., Dylov, D.V.: Autofocusing+: Noise-resilient motion correction in magnetic resonance imaging. In: MICCAI. pp. 365–375 (2022)
10. Li, F., Zhou, Z., Fang, Y., Cai, J., Wang, Q.: Moco-diff: Adaptive conditional prior on diffusion network for mri motion correction. In: MICCAI. pp. 411–421 (2024)
11. Liu, S., Thung, K.H., Qu, L., Lin, W., Shen, D., Yap, P.T.: Learning mri artefact removal with unpaired data. Nature Machine Intelligence **3**(1), 60–67 (2021)

12. Nárai, Á., Hermann, P., Auer, T., Kemenczky, P., Szalma, J., Homolya, I., Somogyi, E., Vakli, P., Weiss, B., Vidnyánszky, Z.: Movement-related artefacts (mr-art) dataset of matched motion-corrupted and clean structural mri brain scans. Scientific data **9**(1), 630 (2022)
13. Oh, G., Lee, J.E., Ye, J.C.: Unpaired mr motion artifact deep learning using outlier-rejecting bootstrap aggregation. IEEE Trans. Med. Imaging **40**(11), 3125–3139 (2021)
14. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: MICCAI. pp. 234–241 (2015)
15. Spieker, V., Eichhorn, H., Hammernik, K., Rueckert, D., et al.: Deep learning for retrospective motion correction in mri: A comprehensive review. IEEE Trans. Med. Imaging **43**(2), 846–859 (2024)
16. Thomas, K., Karim, A., Jiahuan, Y., Bin, Y., Fritz, S., Sergios, G.: Retrospective correction of motion-affected mr images using deep learning frameworks. Magn. Reson. Med. **82**(4), 1527–1540 (2019)
17. Upadhyay, U., Chen, Y., Hepp, T., Gatidis, S., Akata, Z.: Uncertainty-guided progressive gans for medical image translation. In: MICCAI. pp. 614–624 (2021)
18. Wang, J., Yang, y., Yang, y., Sun, J.: Dual domain motion artifacts correction for mr imaging under guidance of k-space uncertainty. In: MICCAI (2023)
19. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process **13**(4), 600–612 (2004)
20. Wu, B., Li, C., Zhang, J., Lai, H., Feng, Q., Huang, M.: Unsupervised dual-domain disentangled network for removal of rigid motion artifacts in mri. Comput. Biol. Med. **165**, 107373 (2023)
21. Zhou, Y., Li, H., Liu, J., Kong, Z., Huang, T., Ahn, E., Lv, Z., Kim, J., Feng, D.D.: Explicit abnormality extraction for unsupervised motion artifact reduction in magnetic resonance imaging. IEEE J. Biomed. Health Inform. (2024)
22. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networkss. In: ICCV (2017)
23. Zou, X., Xiao, F., Yu, Z., Li, Y., Lee, Y.J.: Delving deeper into anti-aliasing in convnets. IJCV. **131**(1), 67–81 (2023)