MRI Motion Correction Through Disentangled CycleGAN Based on Multi-Mask K-Space Subsampling

Gang Chen, Han Xie, Xinglong Rao, Xinjie Liu, Martins Otikovs, Lucio Frydman, Peng Sun, Zhi Zhang, Feng Pan, Lian Yang, Xin Zhou, Maili Liu, Qingjia Bao, and Chaoyang Liu

***Abstract*— This work proposes a new retrospective motion correction method, termed DCGAN-MS, which employs disentangled CycleGAN based on multi-mask k-space subsampling (DCGAN-MS) to address the image domain translation challenge. The multi-mask k-space subsampling operator is utilized to decrease the complexity of motion artifacts by randomly discarding motion-affected k-space lines. The network then disentangles the subsampled, motion-corrupted images into content and artifact features using specialized encoders, and generates motion-corrected images by decoding the content features. By utilizing multi-mask k-space subsampling, motion artifact features become more sparse compared to the original image domain, enhancing the efficiency of the DCGAN-MS network. This method effectively corrects motion artifacts in clinical gadoxetic acid-enhanced human liver MRI, human brain MRI from fastMRI, and preclinical rodent brain MRI. Quantitative improvements are demonstrated with SSIM values increasing from 0.75 to 0.86 for human liver MRI with simulated motion artifacts, and from 0.72 to 0.82 for rodent brain MRI with simulated motion artifacts. Correspondingly, PSNR values increased from 26.09 to 31.09 and from 25.10 to 31.77. The method's performance was further validated on clinical and preclinical motion-corrupted MRI using the Kernel Inception Distance (KID) and Fréchet Inception Distance (FID) metrics. Additionally, ablation experiments were conducted to confirm the effectiveness of the multi-mask k-space subsampling approach.**

***Index Terms*— Gadoxetic acid-enhanced human liver MRI, motion artifact, unpaired learning, multi-mask k-space subsampling, CycleGAN.**

Received 29 September 2024; revised 14 December 2024; accepted 20 December 2024. Date of publication 30 December 2024; date of current version 3 April 2025. This work was supported in part by the National Key Research and Development Program under Grant 2023YFE0113300 and Grant 2022YFF0707000, in part by the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant XDB0540300, in part by the National Natural Science Foundation of China under Grant 22327901, and in part by the Science and Technology Innovation Talent and Service Special Project of Hubei Province under Grant 2023EHA003. *(Corresponding authors: Lian Yang; Qingjia Bao.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board at Union Hospital, Tongji Medical College, Huazhong University of Science and Technology.

Gang Chen, Xinjie Liu, Zhi Zhang, and Qingjia Bao are with the State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China, and also with the University of Chinese Academy of Sciences, Beijing 100864, China (e-mail: chengang23@mails.ucas.ac.cn; liuxinjie@apm.ac.cn; zhangzhi@apm.ac.cn; qingjia.bao@apm.ac.cn).

Han Xie is with Wuhan Railway Vocational College of Technology, Wuhan 430071, China (e-mail: xiehan@wru.edu.cn).

Xinglong Rao is with the State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China (e-mail: xinglongrao@163.com).

Martins Otikovs and Lucio Frydman are with the Department of Chemical and Biological Physics, Weizmann Institute of Science, Rehovot 76100, Israel (e-mail: otikovs@gmail.com; lucio.frydman@weizmann.ac.il).

Peng Sun is with the Clinical & Technical Support, Philips Healthcare, Beijing 100600, China (e-mail: peng.sun@philips.com).

Feng Pan and Lian Yang are with the Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430071, China (e-mail: uh_fengpan@hust.edu.cn; yanglian@hust.edu.cn).

Xin Zhou, Maili Liu, and Chaoyang Liu are with the State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China, also with the University of Chinese Academy of Sciences, Beijing 100864, China, and also with the Optics Valley Laboratory, Hubei 430074, China (e-mail: xinzhou@wipm.ac.cn; ml.liu@apm.ac.cn; chyliu@apm.ac.cn).

Digital Object Identifier 10.1109/TMI.2024.3523949

I. Introduction

MAGNETIC resonance imaging (MRI) is a widely used imaging modality for both clinical [1] and preclinical [2] applications due to its ability to visualize both the anatomy and function of tissues and organs as well as pathological processes [3], [4]. However, since its inception, MRI has been hampered by motion artifacts due to subject motion [5], [6]. An example is liver MRI with gadoxetic acid, a hepatocyte-specific MR contrast agent [7], [8]. Previous studies [9], [10], [11] have reported a high incidence (5%-18%) of severe image quality degradation due to acute transient dyspnea or transient severe motion during gadoxetic acid-enhanced MRI. In addition, in pediatric or stroke patients in clinical [12] or rodents in preclinical studies [13], images are always degraded due to subject motion between consecutive acquisitions. Therefore, it is essential to develop an effective motion artifact correction technique to improve the quality of motion-corrupted images.

Although various prospective and retrospective strategies have been proposed to remove motion artifacts [6], [14],

[15], [16], non-rigid motion correction remains challenging, especially for abdominal regions such as liver MRI [17]. Prospective techniques perform a real-time update of the image acquisition by utilizing either optical tracking of target markers [18] or continuously reacquired images from dedicated navigator scans [19]. However, these prospective solutions are suboptimal in liver MRI as nonrigid motion is difficult to track. On the other hand, the retrospective motion correction methods modify the k-space or image data after the acquisition [20]. These methods can be classified into the primary information-based methods that incorporate prior information and the data-driven autofocusing methods that do not depend on any device or navigator. The data-driven autofocusing motion correction approaches are appealing, as they can be easily applied to all scanners. Unfortunately, these approaches often result in a poorly conditioned and non-convex optimization problem.

With the rapid development of deep learning, many researchers have demonstrated great potential for MRI motion correction [21], especially for dramatically reducing computation time [22] or improving the data-driven autofocusing motion correction convergence [23]. Many methods are based on end-to-end supervised learning networks. The classical networks are trained by pairs of motion-corrupted and motion-free images by comparing the network's prediction with the ground truth motion-free images, using a voxel intensity-based cost function, i.e., L1 or L2 loss. Within this framework, Tamada et al. [24] proposed a deep learning-based network based on multi-channel convolutional neural networks (MARC) to remove the motion artifacts in DCE-MRI images of the liver. A method [25] combining dense convolutional neural networks and residual U-Net was proposed to detect and correct motion artifacts in brain MRI. Yang et al. [26] proposed a model-based synthetic data-driven learning (MOST-DL) method that combines Bloch simulation and general MRI models, which can significantly reduce ghosting and motion artifacts in T2 images. Shi et al. [27] proposed an affinity fusion-based framework for iteratively random motion (AFFIRM), which uses affinity fusion to integrate features between 2D slices and reconstructed 3D volumes to accurately estimate motion in multi-slice fetal brain MRI.

In addition to the classical supervised networks, several generative end-to-end adversarial network (GAN) [28] based methods have also been proposed. These GAN-based methods introduce a generator network that maps the motion-corrupted image to a motion-free one, and a discriminator network that aims to distinguish the predicted image from a ground truth image. The GAN-based conjugate gradient SENSE (CG-SENSE) reconstruction method [29] uses the CG-SENSE algorithm to reconstruct images from multishot motion-corrupted k-space data. These images are further processed by GAN for motion artifact correction. Bao et al. [30] proposed a motion correction method based on the conditional generative adversarial network with minimum entropy (cGANME), and demonstrated that the additional entropy loss can improve the final MR image quality. Gao et al. [31] proposed a hierarchical perception adversarial learning framework (HP-ALF), which employs a hierarchical mechanism to provide structural information of MRI images for adversarial learning from global and regional perspectives.

The supervised motion correction methods always require paired data to train the network. The acquisition of paired data is often impractical and time-consuming. A common approach to generating the paired data is through simulations. However, the simulations cannot accurately and sufficiently reflect all possible forms of real-world artifacts. To cope with the lack of paired motion-free and corrupted data, many unsupervised generative models that aim to correct motion artifacts with only unpaired data have been proposed. CycleGAN [32] architecture consisting of two GANs is one of the most important unsupervised networks. Two generators, one corrupting a motion-free image and one correcting an unpaired corrupted image, are trained to invert each other (cycle transform). Typical CycleGAN-based motion correction methods are Cycle-MedGAN [33] and Cycle-MedGAN V2.0 [34], which capitalize on the adversarial loss to capture high-frequency texture information, and the perceptual loss to enhance the sharpness of the translated images. Recently, Liu et al. [35] suggested a disentangled unsupervised cycle-consistent adversarial network (DUNCAN), which demonstrated that artifact-corrupted images can be disentangled into an anatomical content component and an artifact component. Pan et al. [36] proposed an unsupervised network (DR-CycleGAN) that introduces a novel content consistency loss to eliminate entanglement. In a different setting, Oh et al. [37] attempted to correct motion-corrupted measurements by combining repeated randomly subsampled reconstructions. The motion artifacts were reduced in probability by random subsampling to reject the k-space outliers, and then an optimal transport-driven cycleGAN (OT-cycleGAN) was used to reconstruct the subsampled images.

In this work, we introduce a new retrospective motion correction method, DCGAN-MS (Disentangled CycleGAN based on multi-mask k-space subsampling), designed to address motion artifacts in MRI. First, we apply multi-mask k-space subsampling to reduce these artifacts, leveraging prior knowledge that motion-induced artifacts predominantly arise from outliers in the k-space phase-encoding direction. This multi-mask subsampling approach will generate a domain with fewer motion artifacts, facilitating the correction process while preserving image content information. Second, we design a network based on a disentangled CycleGAN architecture to address the lack of paired ground truth data for training. By reformulating motion correction as an image domain translation task, the network reconstructs motion-free images from multi-mask subsampled, motion-corrupted inputs. We evaluate the method's performance using Kernel Inception Distance (KID) [38] and Frèchet Inception Distance (FID) [39] metrics, which measure the similarity between the distributions of corrected and motion-free images. Experiments on both simulated and in vivo preclinical/clinical MRI data demonstrate the effectiveness of the proposed approach. Additionally, ablation studies confirm the impact of the multi-mask subsampling method on enhancing motion correction.

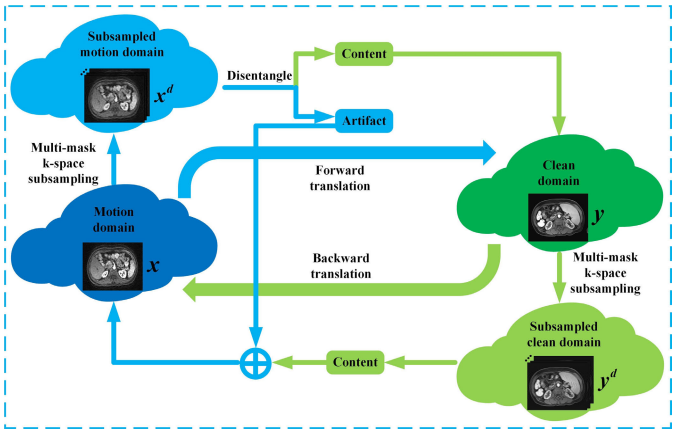

Fig. 1. Overview of the proposed DCGAN-MS. First, multi-mask k-space subsampling is applied to reduce motion artifacts and create a subsampled motion domain. Then, forward translation corrects motion-corrupted images into motion-free images, while backward translation regenerates motion-corrupted images from motion-free ones.

II. Methods

A. The Architecture of DCGAN-MS

The main idea of DCGAN-MS is illustrated in Fig. 1. Given unpaired motion-corrupted images x (from the motion domain) and motion-free images y (from the clean domain), our goal is to translate these images into their corresponding counterparts in the opposite domains. First, multi-mask k-space subsampling is applied to reduce motion artifacts and obtain a subsampled motion domain. Second, forward translation corrects motion-corrupted images into motion-free images, while backward translation regenerates motion-corrupted images from motion-free ones. In the forward translation, the motion-corrupted images x are encoded and disentangled into content and artifact features, while the motion-free images y are encoded into content features. The content features from the motion-corrupted images x are then decoded to reconstruct images in the clean domain. Conversely, the artifact features of x are combined with the content features of the motion-free images y to generate new motion-corrupted images in the motion domain

The multi-mask k-space subsampling operator is a crucial component in reducing the complexity of motion artifacts in motion-corrupted images, allowing the network to more effectively disentangle content and artifact features. By randomly discarding motion-affected k-space lines, this operator sparsifies the motion artifact features, making it easier for the disentangled CycleGAN framework to isolate these artifacts. As a result, the network can focus more efficiently on reconstructing the underlying content, even in cases with severe motion artifacts. This sparsification process, as demonstrated in the ablation studies (discussed further below), leads to a significant improvement in the model's performance. Fig. 2 provides a detailed visualization of the proposed DCGAN-MS architecture. The method is designed around three core elements: multi-mask subsampling, cross-domain translation, and within-domain translation. These elements work synergistically to address the challenge of motion correction in MRI. Multi-mask k-space subsampling simplifies artifact complexity; cross-domain translation enables the transformation between motion-corrupted and motion-free image domains; and within-domain translation ensures reconstruction consistency.

1) Multi-Mask Subsampling: The transient motion of the subject introduces phase errors in the k-space data along the phase-encoding direction. In most cases, these motion-corrupted k-space outliers can be assumed to be sparse, leading to the following motion-corrupted k-space data: [40]:

$$k_x(r,p) = \begin{cases} k_y(r,p)e^{-j\Phi(p)}, & p \in \mathbb{K} \\ k_y(r,p) \end{cases} \tag{1}$$

where $j = \sqrt{-1}$, $k_x(r,p)$ and $k_y(r,p)$ refer to the motion-corrupted and motion-free k-space data, respectively. r and p indicate the read-out and phase-encoding directions, respectively. $\phi(p)$ is the displacement (in radian) phase caused by motion along the phase-encoding direction. The variable κ denotes the phase-encoding indices where the motion occurs.

Therefore, by applying a k-space random subsampling operator to the motion-corrupted image x in the phase-encoding direction, we can mitigate the impact of motion artifacts by randomly discarding some motion-affected k-space lines. However, subsampling will also introduce aliasing artifacts while reducing motion-related artifacts. Therefore, the multi-mask k-space subsampling will be utilized to reduce the effect of aliasing. The particular process is as follows:

(1) Input the raw k-space or image data. If the input is image data, the Fourier transform $\mathcal{F}$ needs to be applied for the input motion-corrupted images x and the motion-free images y to transform the image data into the k-space domain, resulting in $k_x = \mathcal{F}(x), k_y = \mathcal{F}(y)$.

(2) The k-space data k_x and k_y are subsampled using multi-mask to obtain subsampled k-space data $k_x^s = \mathcal{T} \odot k_x, k_y^s = \mathcal{T} \odot k_y$, where $\odot$ represents element-by-element multiplication, and $\mathcal{T}$ is the multiple subsampling matrix composed of 0 and 1.

(3) The subsampled k-space data k_x^s and k_y^s are converted back to image space data through an inverse Fourier transform $\mathcal{F}^{-1}$ to obtain the multi-mask subsampled image $x^d = \mathcal{F}^{-1}(k_x^s), y^d = \mathcal{F}^{-1}(k_y^s)$.

2) Cross-Domain Translation: Fig. 2(a) shows the architecture for cross-domain translation, which is designed to obtain the reconstructed motion-corrected images. Firstly, the subsampled motion-corrupted images x^d are encoded and disentangled to the content features $Z_c^x = E_c(x^d)$ and artifact features $Z_a^x = E_a(x^d)$. The subsampled clean images y^d are encoded to the content features $Z_c^y = E_c(y^d)$. Finally, the artifact features Z_a^x from the motion-corrupted images and the content features Z_c^y from the clean images are decoded to generate the new motion-corrupted image $\tilde{y} = G_m(Z_c^y, Z_a^x)$. Similarly, the content features Z_c^x from the corrupted images are decoded to obtain the motion-corrected images $\tilde{x} = G_c(Z_c^x)$.

One key advantage of DCGAN-MS is its ability to train the network without requiring a paired dataset. This is achieved by performing cross-domain translation twice, resulting in a cycle transformation of the input, as shown in Fig. 2(a). This

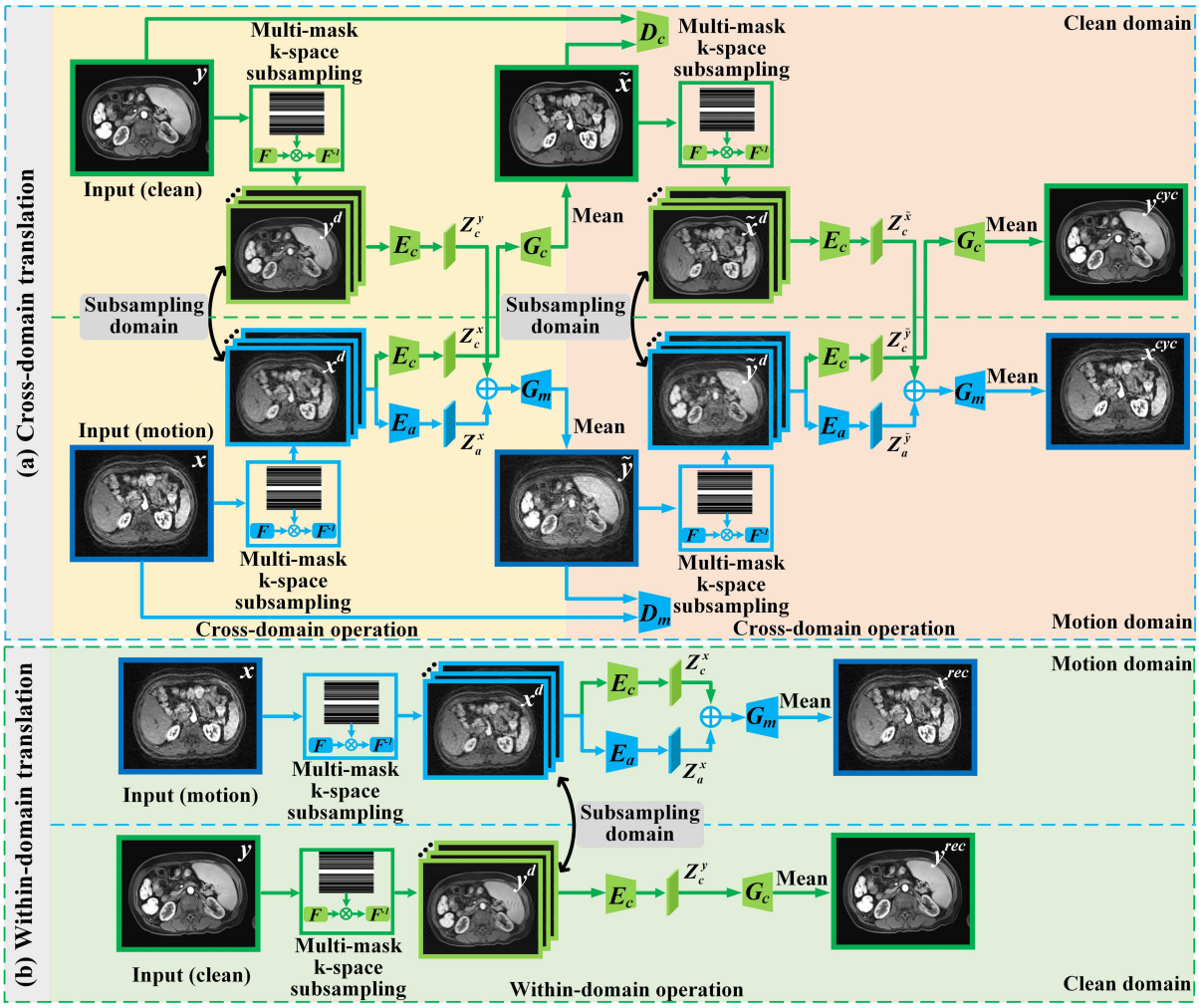

Fig. 2. The architecture of the proposed method (DCGAN-MS). The upper panel (a) illustrates the cross-domain translation process, which utilizes disentangled CycleGAN based on multi-mask k-space subsampling to address motion correction. The lower panel (b) shows the within-domain translation, implemented as a self-supervised constraint to enhance the training efficiency of the encoder and decoder. Key components include: E_c (content feature encoder), E_a (artifact feature encoder), G_c (clean domain generator), G_m (motion domain generator acting as the decoder), D_c (clean domain discriminator), and D_m (motion domain discriminator).

cycle transformation imposes a cycle consistency constraint, ensuring that the learned mappings between domains remain consistent. DCGAN-MS trains the generator and discriminator adversarially. As illustrated in Fig. 2, the model includes two encoders, E_c and E_a along with two generator-decoders, G_c and G_m. The discriminator D_c is used to distinguish whether the generated motion-corrected images $\tilde{x} = G_c(Z_c^x)$ belong to the motion-clean domain. And the discriminators D_m distinguish whether the generated motion-corrupted images $\tilde{y} = G_m(Z_c^y, Z_a^x)$ belong to the motion domain. The generators aim to produce realistic motion-corrected images to deceive the discriminators, while the discriminators work to improve their classification accuracy and avoid being fooled. The detailed architectures of the encoders, generators, and discriminators are presented in Fig. 3.

3) Within-Domain Translation: Fig. 2(b) shows the architecture for within-domain translation. The goal of the within-domain translation is to train the encoder and decoder more effectively by imposing a self-reconstruction loss. Similar to the cross-domain translation, the motion-corrupted images x and motion-free images y are multi-mask k-space subsampled along the phase-encoding direction to obtain the subsampled images x^d and y^d. Then, the motion-corrupted images x^d are also disentangled into the content features Z_c^x and artifact features Z_a^x. The clean images y^d are encoded to the content features Z_c^y. However, unlike cross-domain translation, we do not swap the content features. Thus, the reconstructed motion-corrupted $x^{rec} = G_m(Z_a^x, Z_c^x)$ and motion-corrected $y^{rec} = G_c(Z_c^y)$ should be identical to x and y, respectively.

B. The Loss Function

The total loss function of our method contains four parts (see Fig. 4): the adversarial loss L_{adv}, the cycle-consistency loss L_{cyc}, the reconstruction-consistency loss L_{rec}, and the content-consistency loss L_{sum}:

$$L_{total} = \lambda_{adv} L_{adv} + \lambda_{cyc} L_{cyc} + \lambda_{rec} L_{rec} + \lambda_{con} L_{con} \quad (2)$$

where λ_{adv}, λ_{cyc}, λ_{rec} and λ_{con} are the balance factors to ensure similar contributions among all these losses.

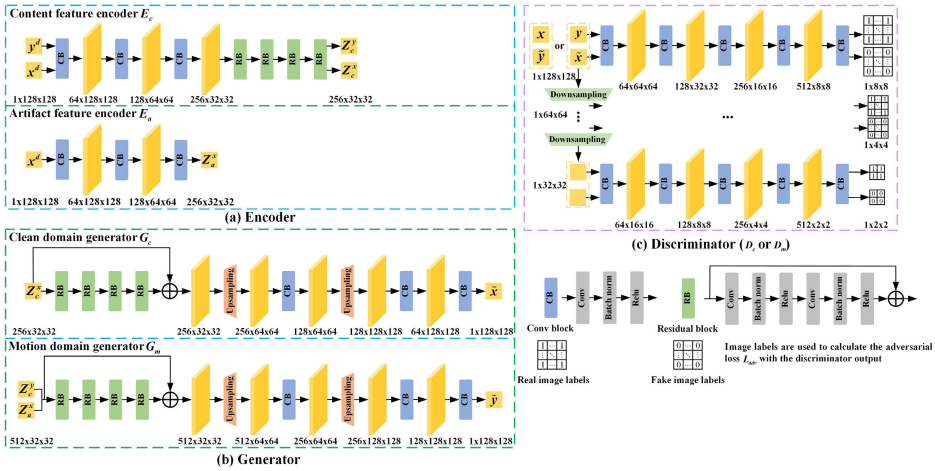

Fig. 3. The detailed structure of the proposed model, including encoders E_c and E_a, generators G_c and G_m and the discriminator. (a) Two kinds of encoders: the content feature encoder E_c and the artifact feature encoder E_a. (b) Two kinds of generators: the clean domain generator G_c whose input is only content features, and the motion domain generator G_m whose input contains both content features and artifact features. (c) The discriminators (D_c, D_m), which evaluate the quality and realism of the generated images.

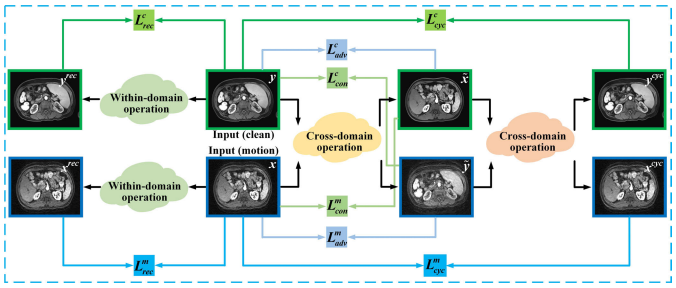

Fig. 4. Description of loss functions in the proposed method.

1) Adversarial Loss: The adversarial loss L_{adv} is a critical component of GAN-based networks, measuring the discrepancy between the distribution of real motion-free data and the distribution of generated motion-corrected data. This loss encourages the generator to produce realistic images that can deceive the discriminator, while pushing the discriminator to correctly classify input images as real or fake. The adversarial loss functions like a minimax game between the generator and discriminator, where the generator aims to minimize the loss, and the discriminator attempts to maximize it.

In the proposed method, there are two types of adversarial losses: L^C_{adv}, which encourages the generation of more realistic motion-corrected images. L^M_{adv}, which encourages the generation of more realistic motion-corrupted images.

$$L_{adv} = L^C_{adv} + L^M_{adv} \tag{3}$$

$$L^C_{adv} = E[\|D_c(y)\|_2] + \mathbb{E}[\|D_c(\tilde{x}) - 1\|_2],$$
$$L^M_{adv} = E[\|D_m(x)\|_2] + \mathbb{E}[\|D_m(\tilde{y}) - 1\|_2] \tag{4}$$

$$\tilde{x} = G_c\big(E_c(x^d)\big),\ \tilde{y} = G_m\big(E_c(y^d), E_a(x^d)\big) \tag{5}$$

where $\tilde{x}$ are the generated motion-corrected images, $\tilde{y}$ are the generated motion-corrupted images, $\|\cdot\|_2$ is the L_2 norm, and $\mathbb{E}$ denotes the expectation operator.

2) Cycle-Consistency Loss: The cycle-consistency loss L_{cyc} is a crucial component of DCGAN-MS, ensuring that the generated images retain the content of the original images after cross-domain translation. This loss operates on the principle that if an image is translated from the motion domain to the clean domain, and then back to the motion domain, the result should closely resemble the original image. Similarly, if an image is translated from the clean domain to the motion domain, and then back to the clean domain, the original image should be recovered. The cycle-consistency loss measures the difference between the original and reconstructed images using the L_1 norm and works to minimize this difference. This constraint helps maintain the content integrity of the input images throughout the translation process.

$$L_{cyc} = L^C_{cyc} + L^M_{cyc} \tag{6}$$

$$L^C_{cyc} = \mathrm{E}[\|y - y^{cyc}\|_1],\ L^M_{cyc} = \mathrm{E}[\|x - x^{cyc}\|_1] \tag{7}$$

$$x^{cyc} = G_m\left(E_c\big(\Phi(\tilde{x})\big), E_a\big(\Phi(\tilde{y})\big)\right),$$
$$y^{cyc} = G_c\left(E_c(\Phi(\tilde{y}))\right) \tag{8}$$

where x^{cyc} is the cycle reconstructed motion-corrupted images, y^{cyc} is the cycle reconstructed motion-corrected images, and $\|\cdot\|_1$ is the L_1 norm.

3) Within-Domain Reconstruction-Consistency Loss: L_{rec} is used to ensure that the encoder and decoder can preserve the content of the input images after encoding and decoding, which measures the pixel-wise discrepancy between the input image and its reconstructed version in the within-domain translation. This loss is based on the idea that if we encode an image from the motion/clean domain and then decode it back to the corresponding domain, we should obtain an image that is identical or very similar to the original image. Reconstruction loss can effectively avoid the problem of losing information or introducing noise during the encoding and decoding process.

$$L_{rec} = L^C_{rec} + L^M_{rec} \tag{9}$$

$$L^C_{rec} = \mathbb{E}[\|y - y^{rec}\|_1],\ L^M_{rec} = \mathbb{E}[\|x - x^{rec}\|_1] \tag{10}$$

$$x^{rec} = G_m\big(E_c(x^d), E_a(x^d)\big),\ y^{rec} = G_c\big(E_c(y^d)\big) \tag{11}$$

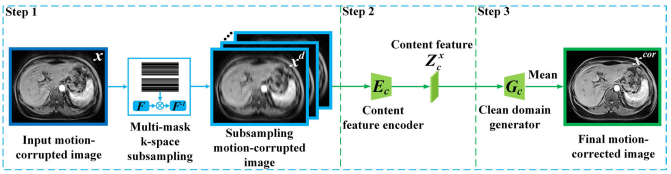

Fig. 5. The complete inference process of motion correction in DCGAN-MS.

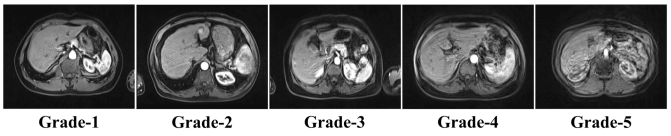

Fig. 6. Five-point Likert scale based on artifact significance. Grade 1: No artifacts; Grade 2: Mild artifacts; Grade 3: Moderate artifacts; Grade 4: Severe artifacts; Grade 5: Non-diagnostic.

where x^{rec} are the reconstructed motion-corrupted images, and y^{rec} are the reconstructed motion-corrected images.

4) Content Consistency Loss: Due to the lack of sufficient supervised information for the translated images, relying only on the constraints of the feature domain discriminator may lead to "hallucinations" in the generated images. As we know, motion artifacts always occur in the phase-encoding direction and have a negligible effect on the calculation of the sum of pixel values in each column along the phase-encoding direction compared to the original content of the images [36]. With content consistency loss L_{con} constraints for the final motion-corrected images, we can effectively reduce hallucinations.

$$L_{con} = L_{con}^{C} + L_{con}^{M} \tag{12}$$

$$L_{con}^{C} = \sum_{i}^{n} \left(\left\| \frac{sum_i(y)}{max_i(y)} - \frac{sum_i(\tilde{y})}{max_i(\tilde{y})} \right\|_1 \right),$$

$$L_{con}^{M} = \sum_{i}^{n} \left(\left\| \frac{sum_i(x)}{max_i(x)} - \frac{sum_i(\tilde{x})}{max_i(\tilde{x})} \right\|_1 \right) \tag{13}$$

where n is the number of the column along the phase-encoding direction, sum_i is the summation of the pixel value in the i_{th} column, and max_i is the maximum value of the pixel value in the i_{th} column.

C. The Pipeline of Training and Inference

Training procedure (1) Subsampling of the input unpaired images with different mask N times ($N = 15$ in this study).

(2) Disentangle the multi-mask subsampled images into content and artifact features.

(3) With extracted content and artifact features, a cross-domain operator is employed to obtain generated motion-free or motion-corrupted images. Subsequently, the cross-domain operator is executed once more to ascertain the cycle-consistency loss. At the same time, a within-domain operator is utilized to obtain within-domain reconstruction-consistency loss.

(4) Train the network (DCGAN-MS) with hybrid loss.

Inference procedure (1) Subsampling of the input motion-corrupted images x with different mask N times ($N = 15$ in this study) to obtain x^d (see Fig. 5).

(2) Disentangle the subsampled motion-corrupted images x^d to the content features $Z_c^x = E_c(x^d)$ with the trained network E_c.

(3) The content features Z_c^x are decoded using the trained network G_c, and the final motion-corrected image $x^c or$ is obtained through average aggregation processing.

III. Experiments

A. Dataset

1) Clinical Gadoxetic Acid-Enhanced Human Liver MRI: The clinical dataset comprised gadoxetic acid-enhanced MRI scans of the liver collected from Union Hospital, Huazhong University of Science and Technology, after privacy-preserving preprocessing. The study was approved by the institutional review board (IRB No. 2019-09-021), and written informed consent was waived. All MR image data (in DICOM file format) were exported from the institutional Vue PACS system and underwent deidentification preprocessing. A total of 176 examinations were acquired on a commercial 1.5T MR scanner (MAGNETOM Avanto, Siemens Healthineers, Germany), and 132 examinations were acquired on a 3T MR scanner (MAGNETOM Skyra, Siemens Healthineers, Germany). The study followed standard scanning protocols, images were acquired using the 3Dimensional Volumetric Interpolated Breath-hold Examination (3D-VIBE) pulse sequence [42] with the following scan parameters: FOV 380 mm × 380 mm, acquisition matrix size 320 × 320, TE = 1.29 ms, TR = 4.5 ms, and slice thickness = 3 mm. Two radiologists assessed the grades of motion artifacts. For subjective evaluation, they rated the motion artifacts of arterial phase MR images before and after motion correction using a five-point Likert scale based on the significance of the artifacts [9], [10] (refer to Fig. 6). Of the 308 examinations containing artifacts, we utilized 50 with grade 2 to 5 motion artifacts and 58 without artifacts as a training dataset. The remaining 258 examinations with artifacts were used as a testset.

2) Preclinical Brain Rodent MRI: We also evaluated our algorithm in the preclinical study for rodent brain MRI. The preclinical MRI experiments were performed on a Bruker Biospec 7.0 T/20 cm MRI scanner. A 72-mm-diameter volume coil was used for radio-frequency (RF) transmission and a quadrature surface coil for signal detection. The body temperature was kept at 37 °C with water circulation. Images were acquired using the Rapid Acquisition with Relaxation Enhancement (RARE) pulse sequence and the following scan parameters: FOV 28 mm × 28 mm, acquisition matrix size 256 × 256, ETL = 8, TE = 20 ms, TR = 2500 ms, slice thickness = 0.6 mm, number of slices = 30. A total of 20 rats were scanned. To obtain realistic motion-affected images, we conducted in vivo experiments with rodents by gradually reducing the isoflurane concentration to zero. We maintained the respiratory rate around 120 bpm for acquiring the motion-affected imaging data.

3) Motion Simulation for Quantitative Analysis: To evaluate the performance of motion correction methods in terms of

TABLE I
QUANTITATIVE EVALUATION OF MOTION CORRECTION RESULTS USING VARIOUS METHODS FOR CLINICAL, PRECLINICAL, AND PUBLIC DATASETS WITH SIMULATED MOTION ARTIFACTS. (BEST PERFORMANCE IS INDICATED IN BOLD BLACK)

Method	Human liver MRI		Rodent brain MRI		Human brain MRI from fastMRI	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
Motion-corrupted	0.75 ± 0.08	26.09 ± 5.03	0.72 ± 0.10	25.10 ± 3.43	0.79 ± 0.04	27.27 ± 1.97
MARC [24]	0.79 ± 0.05	29.73 ± 2.63	0.77 ± 0.07	30.12 ± 3.44	0.82 ± 0.04	**29.28 ± 2.45**
DUNCAN [35]	0.80 ± 0.05	27.68 ± 1.66	0.76 ± 0.08	28.45 ± 3.65	0.82 ± 0.05	28.19 ± 2.41
Bootstrap [40]	0.82 ± 0.06	28.93 ± 3.04	0.77 ± 0.08	28.67 ± 3.63	0.79 ± 0.03	28.26 ± 1.72
UDDN [41]	0.85 ± 0.05	28.77 ± 4.13	0.82 ± 0.07	31.09 ± 4.32	0.83 ± 0.05	27.95 ± 2.69
Ours	**0.86 ± 0.04**	**31.09 ± 1.89**	**0.82 ± 0.06**	**31.77 ± 4.49**	**0.84 ± 0.04**	28.63 ± 1.60

SSIM and PSNR, we simulated motion artifacts on both preclinical and clinical datasets. For the clinical human liver MRI, 58 motion-free subjects were selected as the ground truth. Artifact simulation was performed by introducing phase errors to the k-space data, following the methods in [24] and [40]. For the preclinical rodent brain MRI, motion simulation accounted for both inter-shot motion (within the repetition time) and intra-shot motion (within the echo train), as described in [30]. Additionally, human brain T2-weighted MRI from the FastMRI public dataset [43] was utilized to further validate the generalizability of the proposed method. The fastMRI dataset scanning parameters are: field of view (FOV) ranging from 220 mm × 220 mm to 240 mm × 240 mm, acquisition matrix size of 320 × 320, echo time (TE) of 95-112 ms, and repetition time (TR) of 3112-8400 ms.

B. Implementation of Methods

The experiments were conducted using a system equipped with an NVIDIA Geforce RTX 3090 with 24GB of GPU memory and an Intel Xeon Gold 6230 CPU at 2.10 GHz. For the training process, we implemented our network on the PyTorch platform and utilized Adam [44] for optimization. The learning rate was initially set to 0.0001 with exponential decay over 10K iterations. To save memory and speed up training, we randomly cropped the original images to obtain 128 × 128 pixel images prior to training.

For comparison, we also implemented three state-of-the-art deep learning-based methods: one supervised method (MARC) [24] and three unsupervised methods (DUNCAN [35], Bootstrap [40] and UDDN [41]). MARC is an end-to-end supervised deep learning approach for motion correction. We trained the MARC network with the simulated data and tested it with both the simulated and real in vivo data.

Evaluating the quality of the model can be challenging when motion correction attempts to generate high-dimensional data (motion-corrected data) with complex structures, especially for real in vivo images. The Kernel Inception Distance (KID) [38] and Frèchet Inception Distance (FID) [39] are common metrics used to compare the feature distributions of the generated and real images to assess how well the generative model preserves the diversity and quality of the original domain. In this study, we used the Kernel Inception Distance (KID) and Frèchet Inception Distance (FID) in conjunction with commonly used evaluation metrics such as the Peak Signal to Noise Ratio (PSNR) and the Structural Similarity Index (SSIM) to assess the quality of the generated motion-corrected images. The KID and FID scores decrease as the distribution of the generated images approaches that of the real motion-free images. We used t-distributed stochastic neighborhood embedding (t-SNE) [45] to reduce the dimensionality of the feature vectors and assess the dissimilarity between the distribution of the generated motion-corrected data and the real motion-free data. The source code and example data are available on https://github.com/baoqingjia/DCGAN-MS.

IV. EXPERIMENTAL RESULTS

A. Clinical MRI Data

1) Simulated Motion Artifacts: Fig. 7 provides both a quantitative and qualitative comparison of motion correction results for simulated motion artifacts in gadoxetic acid-enhanced MRI. Each image includes a zoomed-in region to highlight details, with the corresponding PSNR and SSIM values shown in the upper left corner. The MARC method [24] shows a reduction in motion artifacts; however, it often introduces noticeable blurring, leading to a loss of fine image details. Both the DUNCAN [35] and Bootstrap [40] methods manage to avoid significant blurring, but their overall effectiveness in correcting motion artifacts remains limited. In contrast, our proposed method successfully eliminates motion artifacts while preserving the finer structural details of the image. This is especially evident in the zoomed regions, where our method's results appear much closer to the ground truth compared to other methods. To further validate the generalizability of our method, we compared the motion correction results of various methods applied to human brain MRI from fastMRI, utilizing simulated motion artifacts for evaluation. As illustrated in Fig. 8, each image is accompanied by a difference map that highlights the discrepancies between the corrected images and the ground truth. These difference maps provide a visual representation of the effectiveness of each method in reducing motion artifacts and preserving the integrity of the original images. By analyzing these results, we can assess how well our approach performs compared to other state-of-the-art methods, reinforcing its robustness and versatility across different datasets and imaging conditions.

In addition to visual improvements, our method significantly boosts quantitative performance. As shown in Table I, both PSNR and SSIM values are markedly enhanced after applying our motion correction technique, indicating more accurate and clearer reconstructions. To further validate these findings, Fig. 9 presents SSIM histograms comparing different methods.

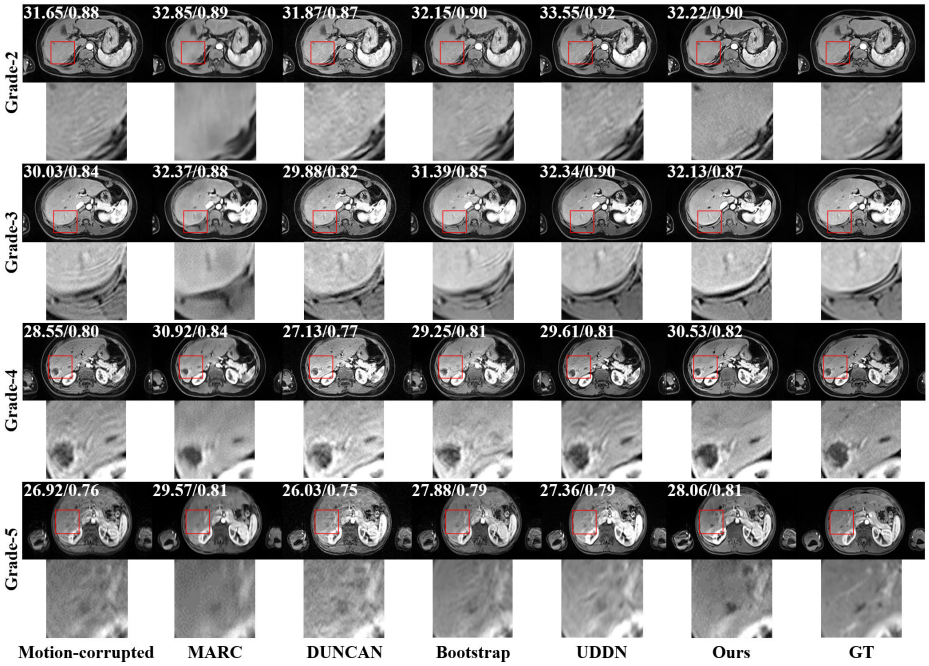

Fig. 7. Quantitative and qualitative comparison of motion correction results using various methods for human liver MRI with simulated motion artifacts. The top-to-bottom rows show motion correction for varying degrees of motion artifacts. Our method is compared with four state-of-the-art techniques: MARC [24], DUNCAN [35], Bootstrap [40] and UDDN [41]. The PSNR and SSIM values for each image are displayed in the corners.

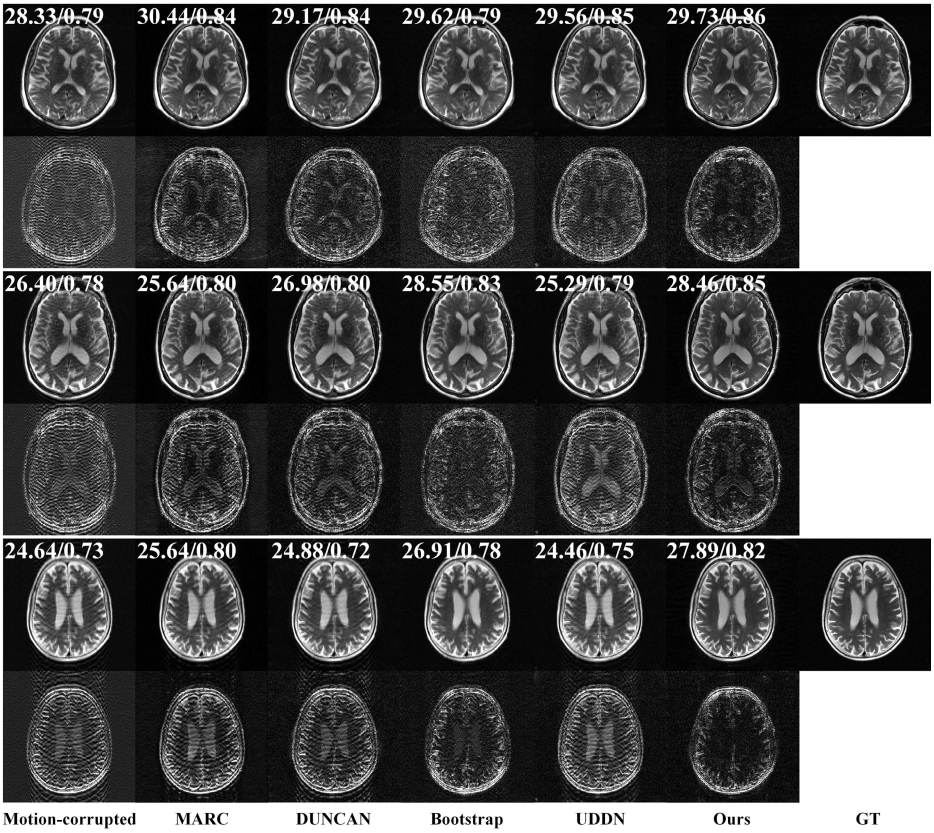

Fig. 8. Quantitative and qualitative comparison of motion correction results using various methods for human brain MRI from fastMRI with simulated motion artifacts.

The SSIM values for the motion-corrupted images primarily fall between 0.5 and 0.9, highlighting the degradation caused by motion. However, after applying our method, most images achieve SSIM values above 0.8, significantly outperforming other state-of-the-art techniques. This demonstrates not only the robustness of our approach in addressing motion artifacts

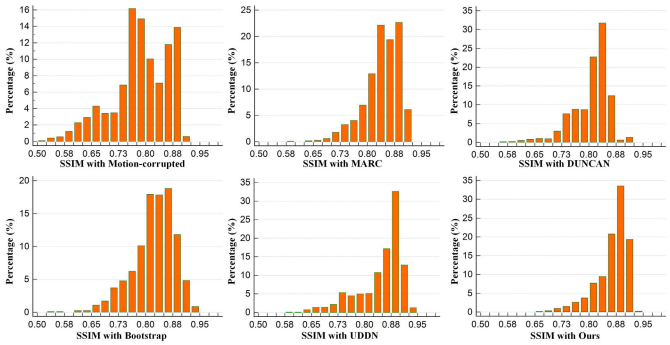

Fig. 9. SSIM histograms of motion correction results using various methods for human liver MRI with simulated motion artifacts.

TABLE II
RADIOGRAPHIC ASSESSMENT OF MOTION CORRECTION RESULTS USING VARIOUS METHODS FOR CLINICAL GADOXETIC ACID-ENHANCED HUMAN LIVER MRI

Method	Artifact	Cohen's kappa	Blurring	Cohen's kappa
Motion-corrupted	2.9 ± 1.3	0.92	1.2 ± 0.4	1.0
MARC [24]	**1.6 ± 0.5**	0.89	2.5 ± 0.5	0.71
DUNCAN [35]	2.3 ± 0.7	0.86	1.3 ± 0.3	0.80
Bootstrap [40]	2.1 ± 0.8	0.84	1.2 ± 0.4	0.88
UDDN [41]	2.0 ± 0.5	0.90	1.1 ± 0.5	1.0
Ours	1.8 ± 0.6	0.89	**1.1 ± 0.4**	1.0

but also its ability to maintain and even enhance image quality across a wide range of conditions.

2) Real Motion Artifacts: Fig. 10 presents the motion correction results for clinical gadoxetic acid-enhanced MRI images that exhibit real in vivo motion artifacts. Our proposed method demonstrates superior performance compared to other techniques in motion correction, effectively avoiding blurring and preserving fine details in the images. This preservation of detail is crucial for accurate diagnosis and interpretation, especially in complex abdominal imaging scenarios.

To gain a more comprehensive understanding of the clinical relevance of these results, we invited two experienced radiologists to conduct a thorough radiologic assessment of the corrected images. The two physicians, one with 25 years of experience (Lian Yang) and the other with 13 years (Feng Pan) in abdominal radiology, independently evaluated the images for the presence of artifacts and blurring. The findings from their evaluations are summarized in Table II, which details their assessments regarding the effectiveness of the different methods. Image quality was evaluated using a standardized 5-point visual scoring system, where: 1 = excellent image quality with no visible artifacts; 2 = mild artifacts but satisfactory diagnostic confidence; 3 = moderate artifacts with limited diagnostic confidence; 4 = poor image quality with severe artifacts; and 5 = non-diagnostic image quality due to pronounced artifacts. Additionally, image blurring was assessed using a 4-point scoring system: 1 = no blurring; 2 = mild blurring; 3 = moderate blurring; and 4 = severe blurring. To ensure the reliability of the evaluations, the consistency of the radiologists' scores was further analyzed using Cohen's kappa measure.

Our results demonstrate the efficacy of the proposed method in mitigating motion artifacts, achieving a significantly lower mean motion artifact score of 1.8 ± 0.6 (indicating mild artifacts with satisfactory diagnostic confidence) compared to 2.9 ± 1.3 for motion-corrupted images (indicating moderate artifacts with limited diagnostic confidence). Notably, the model does not introduce additional image blurring. Furthermore, radiographic assessments of simulated motion correction (see Table III) validate the improved visual quality of the images. Expert evaluations confirm that our approach significantly enhances diagnostic confidence and accuracy. This validation by experienced radiologists highlights the clinical applicability of our method, underscoring its potential to improve diagnostic outcomes and patient care in real-world scenarios.

TABLE III
RADIOGRAPHIC ASSESSMENT OF MOTION CORRECTION RESULTS USING VARIOUS METHODS FOR HUMAN LIVER MRI WITH SIMULATED MOTION ARTIFACTS

Method	Artifact	Cohen's kappa	Blurring	Cohen's kappa
Motion-corrupted	3.4 ± 1.2	0.84	1.5 ± 0.4	1.0
MARC [24]	**1.4 ± 0.4**	0.84	2.2 ± 0.6	0.84
DUNCAN [35]	1.9 ± 0.7	0.86	1.5 ± 0.3	0.77
Bootstrap [40]	1.7 ± 0.5	0.89	1.3 ± 0.4	1.0
UDDN [41]	1.5 ± 0.6	0.84	1.5 ± 0.4	0.90
Ours	1.5 ± 0.7	0.83	**1.3 ± 0.3**	1.0

B. Preclinical MRI Data

We also applied the method to the preclinical simulated and in vivo data. Fig. 11 shows the comparison for simulated motion-corrupted rodent brain MRI images. The motion artifact correction performance of our method is better than the other methods in terms of PSNR and SSIM (see Table I). Fig. 12 shows the motion-correction results for preclinical rodent brain MRI with real motion artifact using various methods. These results illustrate that our method can significantly reduce motion artifacts and outperform the other methods.

C. The KID and FID Results

As mentioned above, the evaluation of the performance of the motion correction method for in vivo data is not trivial. In this study, we further analyzed the Kernel Inception Distance (KID) and Frèchet Inception Distance (FID) for different methods. As shown in Table IV, our method has significantly lower KID and FID scores than the other methods. This indicates that the image distribution of corrected images using our method is closer to the distribution of images without motion artifacts.

We also use t-distributed stochastic neighbor embedding (t-SNE) plots to visually analyze the performance of different motion correction methods for the clinical gadoxetic acid-enhanced human liver MRI with real motion artifact. The corrected and clean images are processed by the pre-trained Inception v3 network [46] to obtain high-level image features (1×2048), and Principal Component Analysis (PCA) [47] and t-SNE further reduce the feature dimension to 1×2. By plotting the dimensionality reduction features (1×2) of the two images, we can compare the distributions of the corrected and clean images (see Fig. 13). The MARC-corrected images have a distribution that is significantly different from that

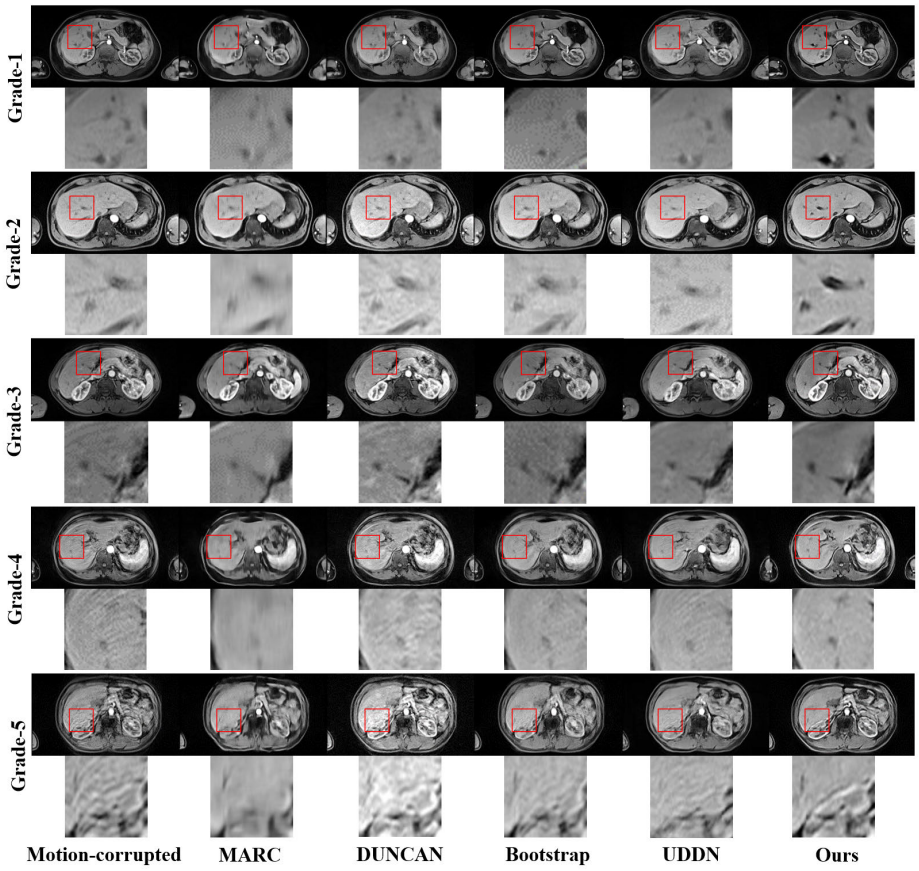

Fig. 10. Motion correction results of clinical gadoxetic acid-enhanced human liver MRI using various methods.

TABLE IV
COMPARISON OF KID AND FID SCORES FOR DIFFERENT MOTION CORRECTION METHODS APPLIED TO BOTH CLINICAL AND PRECLINICAL MRI DATA

Method	Clinical gadoxetic acid-enhanced human liver MRI		Preclinical rodent brain MRI	
	KID	FID	KID	FID
DUNCAN [35]	0.080	22.76	0.053	87.82
Bootstrap [40]	0.080	22.74	0.019	49.00
UDDN [41]	0.073	20.89	0.049	84.80
Ours	**0.066**	**18.69**	**0.015**	**44.91**

of the clean images, with a noticeable distance between them. The DUNCAN-corrected images have a distribution that exhibits a clear boundary from that of the clean images. The images corrected by Bootstrap and our method are intertwined with the distribution of the clean images and do not show clear boundaries. However, the images corrected by our method are more tightly intertwined, indicating that they are closer to the clean images.

V. DISCUSSION

This work presents a motion correction method that utilizes a disentangled CycleGAN framework based on multi-mask subsampled motion-corrupted images. The efficacy of this method was evaluated using quantitative metrics SSIM and PSNR on three datasets with simulated motion artifacts: human liver MRI, human brain MRI from fastMRI, and rodent brain MRI. Furthermore, the efficacy of this method was validated on clinical gadoxetic acid-enhanced human liver MRI and preclinical rodent brain MRI, with real motion artifacts. For the preclinical dataset, the analysis was conducted using raw k-space data, while for clinical studies, we relied on magnitude image data. This distinction arises from the fact that raw k-space data is rarely archived in clinical settings due to storage constraints. Notably, our proposed method effectively addresses motion artifacts in both raw k-space data and reconstructed magnitude images. It is well established that parallel imaging and compressed sensing are widely adopted techniques in clinical MRI, and applying our method to real raw k-space data in conjunction with these techniques presents an exciting opportunity for future research. We hypothesize that integrating our method with parallel imaging and compressed sensing will yield even better image quality. Firstly, accelerated and parallel imaging techniques can significantly reduce acquisition times, thereby minimizing the likelihood of motion events occurring during scanning. Secondly, our retrospective motion correction method is designed to manage residual motion artifacts that may still be present in the images. This combination could enhance the overall efficacy of motion correction, leading to improved diagnostic capabilities in clinical practice.

In contrast to supervised motion correction methods, unsupervised approaches do not require large amounts of

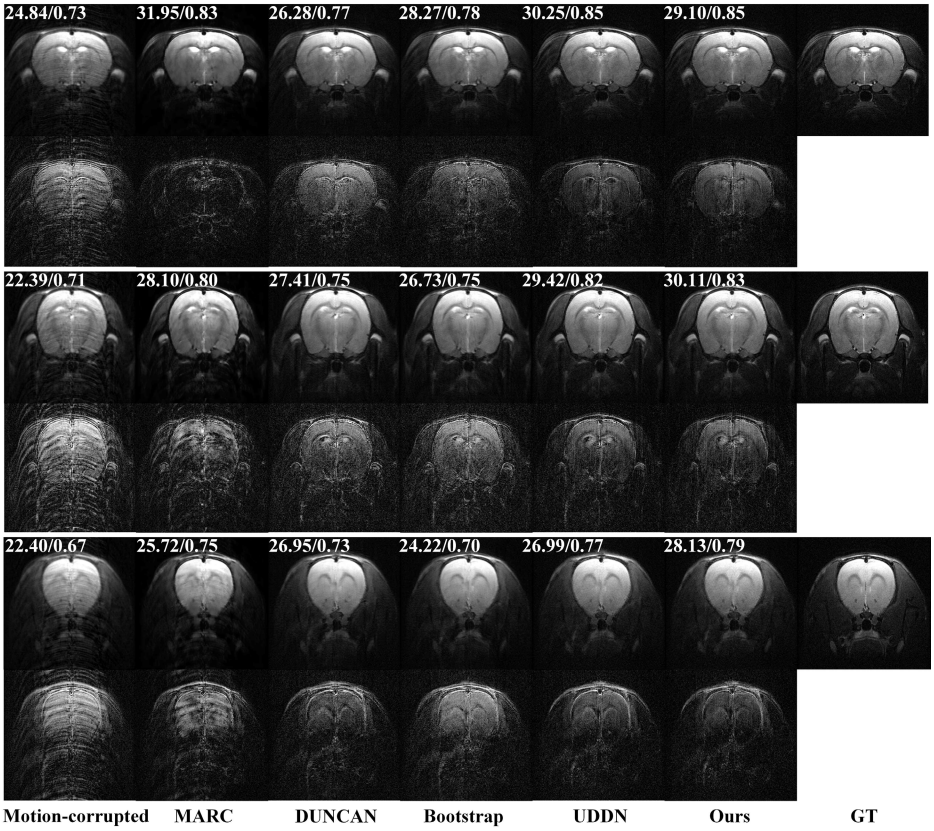

Fig. 11. Quantitative and qualitative comparison of motion correction results using various methods for rodent brain MRI with simulated motion artifacts.

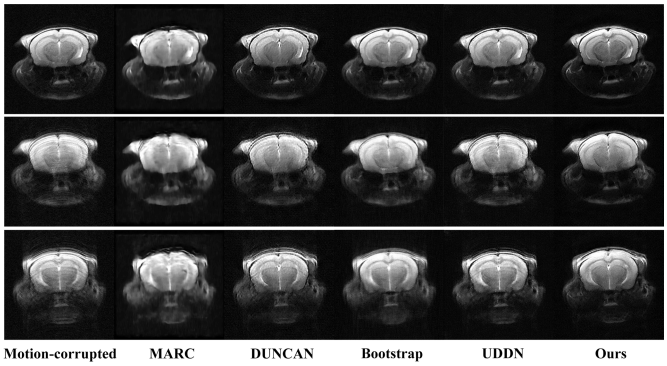

Fig. 12. Motion correction results of preclinical rodent brain MRI using various methods.

paired data for training, making them more adaptable in real-world scenarios where such data is often scarce. Additionally, unsupervised networks [48], [49] are generally more robust to distribution shifts and better equipped to handle out-of-distribution samples compared to their supervised counterparts. While methods like MARC can achieve higher quantitative metrics such as PSNR and SSIM by minimizing the L1 loss between clean and motion-corrupted images, they tend to underperform on real in vivo images with complex motion artifacts. For instance, as shown in Fig. 7, the supervised MARC method [24] often generates blurry images. Generative AI-based methods are better suited for preserving fine structural details in images than end-to-end deep learning motion correction techniques. Motivated by these advantages, we developed a novel motion correction method leveraging CycleGAN, one of the most advanced frameworks in generative AI. To further enhance image quality, we also introduced a multi-scale discriminator, which evaluates the differences between motion-free and motion-corrected images at various scales. This multi-scale approach ensures that finer details are preserved while reducing motion artifacts, ultimately leading to sharper, more realistic images across a range of scales.

The occurrence of hallucinations remains a persistent challenge in generative AI-based motion correction for MRI. Hallucinations occur when the generative model introduces unrealistic or spurious structures into the image, which can mislead clinical interpretation. To tackle this issue, some researchers have incorporated physics-based motion models into the MRI reconstruction process [17], [50], [51], [52]. For example, Hossbach et al. [17] proposed a novel approach to rigid-body motion correction that combines traditional model-driven techniques with deep learning algorithms. Similarly, Levac et al. [53] developed a diffusion model-based method for accelerated MRI motion correction, which jointly estimates both the motion-free image and rigid motion parameters from subsampled, motion-corrupted 2D k-space data. However, these methods are mostly tailored to rigid motion, where the physical model is relatively simple compared to the complexities involved in non-rigid motion, such as that seen in liver MRI. The goal of our study is to develop a motion correction

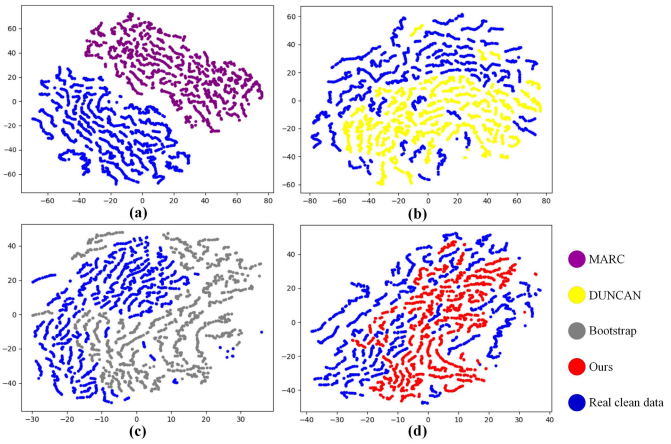

Fig. 13. t-SNE plots between motion correction results of various methods and clinical gadoxetic acid-enhanced human liver MRI. (a) DUNCAN [35], (b) Bootstrap [40], (c) UDDN [41], and (d) Ours method.

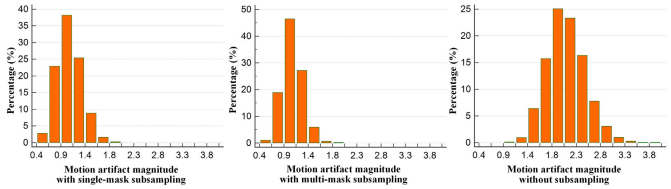

Fig. 14. Histograms of motion artifact magnitude for 1,000 clinical gadoxetic acid-enhanced human liver MRI, measured by the L2 norm of the disentangled artifact module E_a. From left to right: single-mask subsampling, multi-mask subsampling, and without subsampling.

method that is effective in handling non-rigid motion, particularly for clinical gadoxetic acid-enhanced human liver MRI. To achieve this, we have devised a generative AI-based motion correction framework utilizing image-to-image translation. A key focus of this approach is minimizing hallucinations by leveraging three carefully crafted loss functions.

The first of these is the cycle-consistency loss L_{cyc}, which ensures that the content of the input image is maintained after it undergoes cross-domain translation between motion-corrupted and motion-free domains. This prevents the network from losing crucial anatomical information during the generative process. The second loss function, within-domain reconstruction consistency loss L_{rec}, ensures that the network can faithfully reconstruct the original content of the input image after encoding and decoding within the same domain. This is critical for preserving image fidelity throughout the correction process. The third loss function, content consistency loss L_{con}, is designed specifically to address the challenge of hallucinations. Since motion artifacts typically manifest in the phase-encoding direction of MRI, they have little impact on the overall sum of pixel values in each column along that direction. By incorporating this knowledge, L_{con} provides a form of supervised information that guides the network to retain the original content of the image while removing motion artifacts. This not only helps mitigate hallucinations but also ensures that the corrected images remain faithful to the true underlying anatomical structures. However, we acknowledge that the proposed method is not entirely free from hallucination artifacts. To assess this, two experienced radiologists analyzed a total of 300 real in vivo motion-corrupted images (motion grades 4-5). Among these, hallucinations were observed in 4 images when using the proposed method, compared to 16 images when content consistency loss was not employed. This indicates that while our approach reduces the occurrence of hallucinations, there is still room for further improvement.

We also compared our method to two state-of-the-art unsupervised deep-learning-based methods, namely, DUNCAN [35] and Bootstrap [40]. DUNCAN is conceptually similar to the CycleGAN architecture, which consists of two generators, to overcome the need for paired datasets. Moreover, DUNCAN also disentangles images by separating image content features from motion artifact features. Although DUNCAN significantly improves upon MARC and CycleGAN in terms of motion correction, its performance can degrade when dealing with highly complex or severe motion artifacts, as it primarily focuses on generating fully sampled motion-free images from fully sampled motion-corrupted images.

The Bootstrap method approaches motion correction from a probabilistic subsampling perspective. It trains only one generator to remove subsampling artifacts rather than directly addressing motion artifacts. The method does not learn to differentiate between motion-free and motion-corrupted images but instead generates motion-free images by "cheating" the discriminator—replacing randomly dropped k-space lines with denoised or reconstructed lines based on the assumption that these missing lines should resemble uncorrupted data. However, when the subsampled motion-corrupted images significantly differ from their motion-free counterparts, the Bootstrap method may struggle to produce satisfactory results.

In contrast, the primary focus of our method is motion correction through disentangled CycleGAN architecture, augmented with multi-mask subsampling of motion-corrupted k-space data. The subsampling operator randomly discards motion-affected k-space lines, mitigating the impact of motion artifacts [54], [55], [56] and aiding in the disentanglement of image content from motion artifact features. This operator helps to sparsify motion artifact features, improving the efficiency of the CycleGAN model, particularly when motion artifacts are severe. However, if the subsampling mask primarily retains most of the motion-corrupted k-space data, the motion correction results may not be optimal. While subsampling can introduce aliasing artifacts, our multi-mask subsampling strategy helps minimize aliasing effects, much like the Bootstrap method. However, our method diverges from Bootstrap in both reconstruction and correction. Unlike Bootstrap, which only uses motion-free images for training, our method leverages both motion-free and motion-corrupted datasets. This enables our model to better handle complex and severe motion artifacts, resulting in more accurate and effective motion correction, especially in challenging cases.

In the following section, we evaluate the role of multi-mask k-space subsampling from two critical perspectives: simplifying the motion artifacts and preserving the consistency of image content information.

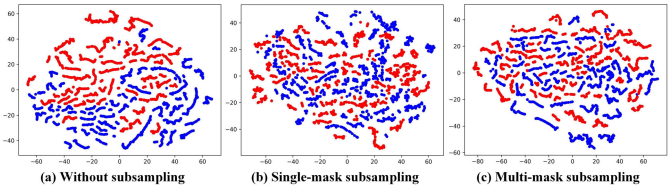

Fig. 15. t-SNE plots for visualizing three scenarios: (a) motion-corrupted versus motion-free images, (b) single-mask subsampling of motion-corrupted and motion-free images, and (c) multi-mask subsampling of motion-corrupted and motion-free images.

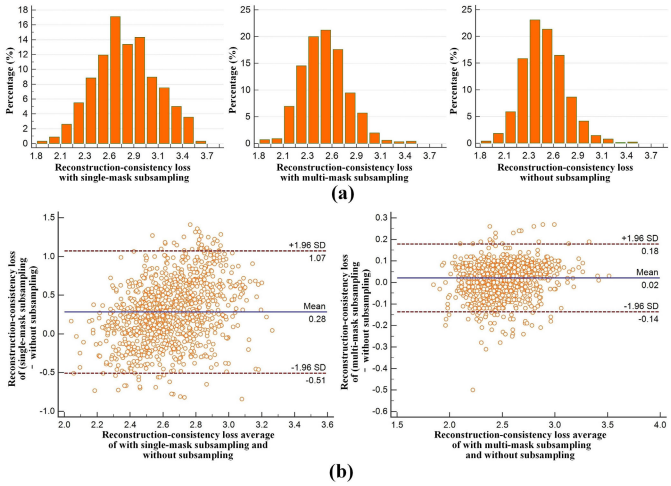

Fig. 16. The histogram and Blad-Altman plots of within-domain reconstruction-consistency loss between the motion-free images y and the reconstructed images y^{rec}. The up row shows the comparison of the histogram. From left to right: single-mask subsampling, multi-mask subsampling, and without subsampling. The second row shows the Blad-Altman plots.

A. Subsampling to Simplify the Motion Artifacts

We quantify the motion artifacts by computing the L2 norm of the disentangled artifact modules E_a. Fig. 14 shows the histograms of artifact amplitudes for 1000 clinical gadoxetic acid-enhanced MRI images (including with single-mask subsampling, multi-mask subsampling, and without subsampling. Without subsampling, artifact amplitudes are distributed around 1.4-3.3. After subsampling, the artifact feature amplitudes are reduced to 0.4-1.8. The results suggest that subsampling can significantly reduce the motion artifact complexity, which may facilitate the network disentanglement artifact features. Moreover, we can observe that the artifacts with multi-mask subsampling are very similar to the single-mask subsampling.

Similarly, we used t-SNE to visualize the following three scenarios: motion-corrupted and motion-free images, motion-corrupted and motion-free images with single-mask subsampling, and motion-corrupted and motion-free images with multi-mask subsampling. As shown in the Fig. 15, the distributions of motion-corrupted and motion-free images become increasingly similar after subsampling. This suggests that subsampling effectively reduces the complexity of motion artifacts, helping to align the feature distribution of motion-corrupted images with that of motion-free images.

TABLE V
QUANTITATIVE EVALUATION OF MOTION CORRECTION RESULTS USING DIFFERENT SUBSAMPLING MASK NUMBERS N FOR HUMAN LIVER MRI WITH SIMULATED MOTION ARTIFACTS

numbers of subsampled masks (N)	SSIM	PSNR
Input	0.75 ± 0.08	26.09 ± 5.03
$N = 1$	0.84 ± 0.05	29.54 ± 2.21
$N = 5$	0.84 ± 0.05	29.70 ± 2.33
$N = 10$	0.84 ± 0.04	29.72 ± 2.15
$N = 15$	**0.86 ± 0.04**	**31.09 ± 1.89**
$N = 20$	0.85 ± 0.05	29.74 ± 2.09

TABLE VI
QUANTITATIVE EVALUATION OF MOTION CORRECTION RESULTS USING DIFFERENT ACCELERATION FACTORS R FOR HUMAN LIVER MRI WITH SIMULATED MOTION ARTIFACTS

Factor R	SSIM	PSNR
Input	0.75 ± 0.08	26.09 ± 5.03
$R = 1$	0.82 ± 0.05	29.04 ± 2.75
$R = 2$	0.83 ± 0.05	29.33 ± 2.03
$R = 3$	**0.86 ± 0.04**	**31.09 ± 1.89**
$R = 4$	0.84 ± 0.05	29.41 ± 2.51

B. Multi-Mask Subsampling That Preserves the Consistency of Image Content Information

The proposed method with subsampling aims to reduce artifacts effectively without losing the content information. To evaluate the impact of the subsampling operator on the content information, we calculate the within-domain reconstruction-consistency loss L_{rec} between the reconstructed images y^{rec} and the input motion-free images y (see details in the loss function), and compare the loss with or without subsampling. The histograms and Bland-Altman plot in Fig. 16 indicate that the reconstruction-consistency loss tends to increase after applying single-mask subsampling (with an average reconstruction loss increase of 0.28) compared to the case without subsampling. This observation implies that the subsampling operator leads to content information loss in motion-corrupted images. To address this issue, a concept similar to bootstrap is introduced, involving the use of multi-mask k-space subsampling to recover content information. The reconstruction-consistency loss histograms exhibit remarkable similarity between multi-mask subsampling and without subsampling conditions, with an average reconstruction-consistency loss difference of only 0.02.

We experimented with the effect of different numbers of subsampled masks (denoted as N) on the motion correction results. Table V shows that both PSNR and SSIM improve as N increases. The performance stabilizes when N reaches 15. Therefore, we choose $N = 15$ as a compromise between performance improvement and computational overhead. Throughout these ablation experiments, we keep the acceleration factor R fixed at 3 to streamline the adjustment of the hyperparameter N. Similarly, to validate the impact of different subsampling rates, we also examine different acceleration factors, denoted as R. Fig. 17 illustrates an example of correction results with different motion artifacts and different acceleration factors R. Table VI provides a quantitative

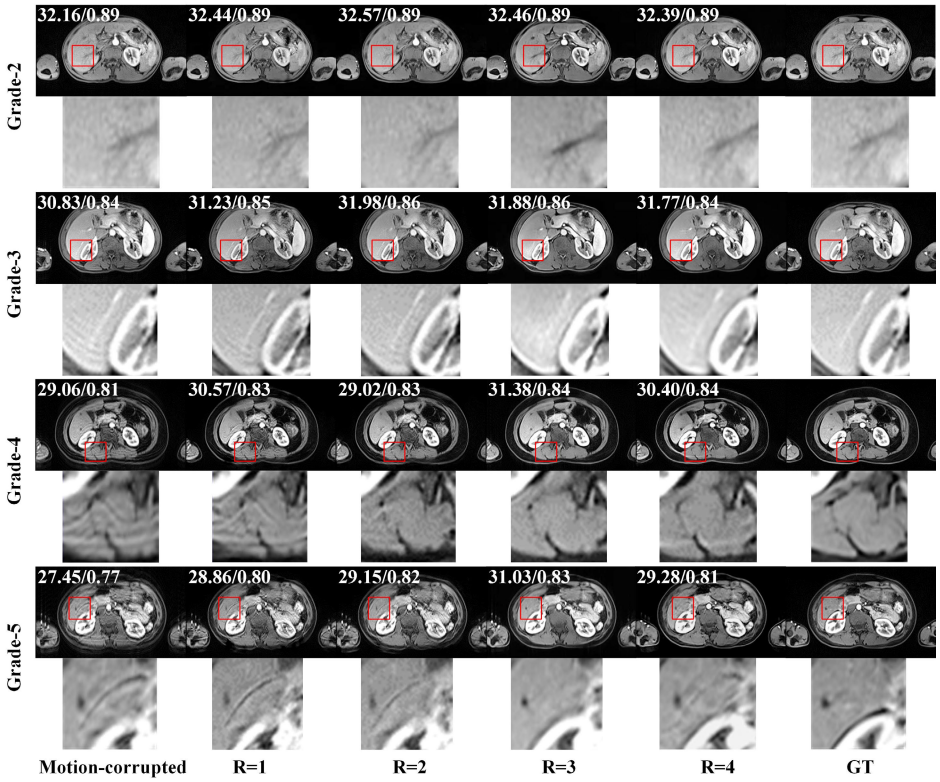

Fig. 17. Quantitative and qualitative comparison of motion correction using different acceleration factors R for human liver MRI with simulated motion artifacts. From left to right: motion corrupted, without subsampling ($R = 1$), with subsampling factor $R = 2$, $R = 3$, $R = 4$, and the ground truth (GT).

comparison across different acceleration factors. Based on these results, motion artifacts are most effectively removed when an acceleration factor $R = 3$ is used. Throughout these ablation experiments, we kept the number of sub-sampled masks at $N = 15$ to facilitate the adjustment of the acceleration factor hyperparameter.

VI. Conclusion

We propose a retrospective motion correction method that used disentangled CycleGAN based on multi-mask k-space subsampling (DCGAN-MS). The main idea is to convert motion correction to the image domain transfer problem, which can be solved by disentangled CycleGAN. Moreover, the mutli-mask subsampling strategy is introduced to reduce features of the motion artifacts and simplify the domain transfer problem. The network can disentangle the motion-corrupted images into the content and artifact features by corresponding trained encoders and obtain the corresponding motion-free images by applying the decoder with the content domain. DCGAN-MS can correct the motion artifact for clinical gadoxetic acid-enhanced human liver MRI and preclinical rodent brain MRI without paired datasets. It also outperforms other unsupervised methods in quantitative metrics, including SSIM and PSNR for datasets with simulated motion artifacts, and KID, FID for the real in vivo datasets.

References

[1] C. Wu et al., "Clinical applications of magnetic resonance imaging based functional and structural connectivity," *NeuroImage*, vol. 244, Dec. 2021, Art. no. 118649. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811921009228

[2] C. Albanese et al., "Preclinical magnetic resonance imaging and systems biology in cancer research: Current applications and challenges," *Amer. J. Pathol.*, vol. 182, no. 2, pp. 312–318, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S000294401200819X

[3] K. Sano et al., "Imaging study of early hepatocellular carcinoma: Usefulness of gadoxetic acid–enhanced MR imaging," *Radiology*, vol. 261, no. 3, pp. 834–844, Dec. 2011.

[4] M. J. Park et al., "Small hepatocellular carcinomas: Improved sensitivity by combining gadoxetic acid–enhanced and diffusion-weighted MR imaging patterns," *Radiology*, vol. 264, no. 3, pp. 761–770, Sep. 2012.

[5] M. Zaitsev, J. Maclaren, and M. Herbst, "Motion artifacts in MRI: A complex problem with many partial solutions," *J. Magn. Reson. Imag.*, vol. 42, no. 4, pp. 887–901, Oct. 2015.

[6] V. Spieker et al., "Deep learning for retrospective motion correction in MRI: A comprehensive review," *IEEE Trans. Med. Imag.*, vol. 43, no. 2, pp. 846–859, Feb. 2024.

[7] D. A. Bluemke et al., "Efficacy and safety of MR imaging with liver-specific contrast agent: U.S. multicenter phase III study," *Radiology*, vol. 237, no. 1, pp. 89–98, Oct. 2005.

[8] O. Döhr, R. Hofmeister, M. Treher, and H. Schweinfurth, "Preclinical safety evaluation of Gd-EOB-DTPA (Primovist)," *Investigative Radiol.*, vol. 42, no. 12, pp. 830–841, Dec. 2007.

[9] M. S. Davenport et al., "Comparison of acute transient dyspnea after intravenous administration of gadoxetate disodium and gadobenate dimeglumine: Effect on arterial phase image quality," *Radiology*, vol. 266, no. 2, pp. 452–461, Feb. 2013.

[10] J. A. Pietryga, L. M. B. Burke, D. Marin, T. A. Jaffe, and M. R. Bashir, "Respiratory motion artifact affecting hepatic arterial phase imaging with gadoxetate disodium: Examination recovery with a multiple arterial phase acquisition," *Radiology*, vol. 271, no. 2, pp. 426–434, May 2014.

[11] U. Motosugi, P. Bannas, C. A. Bookwalter, K. Sano, and S. B. Reeder, "An investigation of transient severe motion related to gadoxetic acid–enhanced MR imaging," *Radiology*, vol. 279, no. 1, pp. 93–102, Apr. 2016.

[12] M. Wilke et al., "Clinical application of advanced MR methods in children: Points to consider," *Ann. Clin. Translational Neurol.*, vol. 5, no. 11, pp. 1434–1455, Nov. 2018.

[13] G. Russo, X. Helluy, M. Behroozi, and D. Manahan-Vaughan, "Gradual restraint habituation for awake functional magnetic resonance imaging combined with a sparse imaging paradigm reduces motion artifacts and stress levels in rodents," *Frontiers Neurosci.*, vol. 15, Dec. 2021, Art. no. 805679.

[14] J. Maclaren, M. Herbst, O. Speck, and M. Zaitsev, "Prospective motion correction in brain imaging: A review," *Magn. Reson. Med.*, vol. 69, no. 3, pp. 621–636, Mar. 2013.

[15] Y. Chang, Z. Li, G. Saju, H. Mao, and T. Liu, "Deep learning-based rigid motion correction for magnetic resonance imaging: A survey," *Meta-Radiol.*, vol. 1, no. 1, Jun. 2023, Art. no. 100001.

[16] H. Chen et al., "High-resolution multi-shot diffusion-weighted MRI combining markerless prospective motion correction and locally low-rank constrained reconstruction," *Magn. Reson. Med.*, vol. 89, no. 2, pp. 605–619, Feb. 2023.

[17] J. Hossbach et al., "Deep learning-based motion quantification from k-space for fast model-based magnetic resonance imaging motion correction," *Med. Phys.*, vol. 50, no. 4, pp. 2148–2161, Apr. 2023.

[18] M. Ippoliti, M. Lukas, W. Brenner, T. Schaeffter, M. R. Makowski, and C. Kolbitsch, "3D nonrigid motion correction for quantitative assessment of hepatic lesions in DCE-MRI," *Magn. Reson. Med.*, vol. 82, no. 5, pp. 1753–1766, Nov. 2019.

[19] A. Singh et al., "Optical tracking with two markers for robust prospective motion correction for brain imaging," *Magn. Reson. Mater. Phys., Biol. Med.*, vol. 28, no. 6, pp. 523–534, Dec. 2015.

[20] K. Vecchiato et al., "Evaluation of DISORDER: Retrospective image motion correction for volumetric brain MRI in a pediatric setting," *Amer. J. Neuroradiol.*, vol. 42, no. 4, pp. 774–781, Apr. 2021.

[21] Z. Zhou, P. Hu, and H. Qi, "Stop moving: MR motion correction as an opportunity for artificial intelligence," *Magn. Reson. Mater. Phys., Biol. Med.*, vol. 37, no. 3, pp. 397–409, Feb. 2024.

[22] M. Moche, R. Trampel, T. Kahn, and H. Busse, "Navigation concepts for MR image-guided interventions," *J. Magn. Reson. Imag.*, vol. 27, no. 2, pp. 276–291, Feb. 2008.

[23] E. Kuzmina, A. Razumov, O. Y. Rogov, E. Adalsteinsson, J. White, and D. V. Dylov, "Autofocusing+: Noise-resilient motion correction in magnetic resonance imaging," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, Jan. 2022, pp. 365–375.

[24] D. Tamada, M.-L. Kromrey, S. Ichikawa, H. Onishi, and U. Motosugi, "Motion artifact reduction using a convolutional neural network for dynamic contrast enhanced MR imaging of the liver," *Magn. Reson. Med. Sci.*, vol. 19, no. 1, pp. 64–76, 2020.

[25] I. Oksuz, "Brain MRI artefact detection and correction using convolutional neural networks," *Comput. Methods Programs Biomed.*, vol. 199, Feb. 2021, Art. no. 105909.

[26] Q. Yang et al., "MOdel-based SyntheTic data-driven learning (MOST-DL): Application in single-shot T_2 mapping with severe head motion using overlapping-echo acquisition," *IEEE Trans. Med. Imag.*, vol. 41, no. 11, pp. 3167–3181, Nov. 2022.

[27] W. Shi et al., "AFFIRM: Affinity fusion-based framework for iteratively random motion correction of multi-slice fetal brain MRI," *IEEE Trans. Med. Imag.*, vol. 42, no. 1, pp. 209–219, Jan. 2023.

[28] I. Goodfellow et al., "Generative adversarial nets," in *Proc. Adv. neural Inf. Process. Syst.*, vol. 27, Jan. 2023, pp. 73–76.

[29] M. Usman, S. Latif, M. Asim, B.-D. Lee, and J. Qadir, "Retrospective motion correction in multishot MRI using generative adversarial network," *Sci. Rep.*, vol. 10, no. 1, p. 4786, Mar. 2020.

[30] Q. Bao et al., "Retrospective motion correction for preclinical/clinical magnetic resonance imaging based on a conditional generative adversarial network with entropy loss," *NMR Biomed.*, vol. 35, no. 12, p. 4809, Dec. 2022.

[31] Z. Gao, Y. Guo, J. Zhang, T. Zeng, and G. Yang, "Hierarchical perception adversarial learning framework for compressed sensing MRI," *IEEE Trans. Med. Imag.*, vol. 42, no. 6, pp. 1859–1874, Jan. 2023.

[32] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2242–2251.

[33] K. Armanious, C. Jiang, S. Abdulatif, T. Küstner, S. Gatidis, and B. Yang, "Unsupervised medical image translation using cycle-MedGAN," in *Proc. 27th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2019, pp. 1–5.

[34] K. Armanious, A. Tanwar, S. Abdulatif, T. Küstner, S. Gatidis, and B. Yang, "Unsupervised adversarial correction of rigid MR motion artifacts," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1494–1498.

[35] S. Liu, K.-H. Thung, L. Qu, W. Lin, D. Shen, and P.-T. Yap, "Learning MRI artefact removal with unpaired data," *Nature Mach. Intell.*, vol. 3, no. 1, pp. 60–67, Jan. 2021.

[36] F. Pan et al., "Correction of arterial-phase motion artifacts in gadoxetic acid-enhanced liver MRI using an innovative unsupervised network," *Bioengineering*, vol. 10, no. 10, p. 1192, Oct. 2023.

[37] G. Oh, B. Sim, H. Chung, L. Sunwoo, and J. C. Ye, "Unpaired deep learning for accelerated MRI using optimal transport driven CycleGAN," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1285–1296, 2020.

[38] M. Bińkowski, D. J. Sutherland, M. Arbel, and A. Gretton, "Demystifying MMD GANs," 2018, *arXiv:1801.01401*.

[39] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local Nash equilibrium," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2017, pp. 1–12.

[40] G. Oh, J. E. Lee, and J. C. Ye, "Unpaired MR motion artifact deep learning using outlier-rejecting bootstrap aggregation," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3125–3139, Nov. 2021.

[41] B. Wu, C. Li, J. Zhang, H. Lai, Q. Feng, and M. Huang, "Unsupervised dual-domain disentangled network for removal of rigid motion artifacts in MRI," *Comput. Biol. Med.*, vol. 165, Oct. 2023, Art. no. 107373.

[42] T. C. Lauenstein et al., "Three-dimensional volumetric interpolated breath-hold MR imaging for whole-body tumor staging in less than 15 minutes: A feasibility study," *Amer. J. Roentgenol.*, vol. 179, no. 2, pp. 445–449, Aug. 2002.

[43] J. Zbontar et al., "FastMRI: An open dataset and benchmarks for accelerated MRI," 2018, *arXiv:1811.08839*.

[44] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[45] L. V. D. Maaten and G. E. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 86, pp. 2579–2605, Jan. 2008.

[46] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.

[47] A. Maćkiewicz and W. Ratajczak, "Principal components analysis (PCA)," *Comput. Geosci.*, vol. 19, no. 3, pp. 303–342, 1993.

[48] Q. Bao et al., "Unsupervised cycle-consistent network using restricted subspace field map for removing susceptibility artifacts in EPI," *Magn. Reson. Med.*, vol. 90, no. 2, pp. 458–472, Aug. 2023.

[49] Q. Bao et al., "Unsupervised deep learning model for correcting Nyquist ghosts of single-shot spatiotemporal encoding," *Magn. Reson. Med.*, vol. 91, no. 4, pp. 1368–1383, Apr. 2024.

[50] İ. Öksüz et al., "Detection and correction of cardiac MRI motion artefacts during reconstruction from k-space," in *Proc. 22nd Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, Shenzhen, China. Cham, Switzerland: Springer, Jan. 2019, pp. 695–703.

[51] H. Eichhorn et al., "Physics-aware motion simulation for T2*-weighted brain MRI," in *Proc. Int. Workshop Simulation Synth. Med. Imag.* Cham, Switzerland: Springer, Jan. 2023, pp. 42–52.

[52] H. Eichhorn et al., "Physics-informed deep learning for motion-corrected reconstruction of quantitative brain MRI," 2024, *arXiv:2403.08298*.

[53] B. Levac, S. Kumar, A. Jalal, and J. I. Tamir, "Accelerated motion correction with deep generative diffusion models," *Magn. Reson. Med.*, vol. 92, no. 2, pp. 853–868, Aug. 2024.

[54] P. Guo, Y. Mei, J. Zhou, S. Jiang, and V. M. Patel, "ReconFormer: Accelerated MRI reconstruction using recurrent transformer," *IEEE Trans. Med. Imag.*, vol. 43, no. 1, pp. 582–593, Jan. 2024.

[55] Z. Li et al., "Encoding enhanced complex CNN for accurate and highly accelerated MRI," *IEEE Trans. Med. Imag.*, vol. 43, no. 5, pp. 1828–1840, May 2024.

[56] Z. Li et al., "Complementation-reinforced network for integrated reconstruction and segmentation of pulmonary gas MRI with high acceleration," *Med. Phys.*, vol. 51, no. 1, pp. 378–393, Jan. 2024.