KiMoCo-Net: Dual-Domain Learning for MRI Motion Artifact Correction

Marina Maher Fekry[1], Mohamed A. El-Beltagy[1], Meena M. Makary[2,3], and Mohammed A. Al-masni[4,*]

[1] Engineering Mathematics and Physics Department, Engineering Faculty, Cairo University, Cairo 12613, Egypt
[2] Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University, Cairo 12613, Egypt
[3] Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA, USA
[4] Department of Bioengineering, King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, Saudi Arabia

marinamaher513@eng.cu.edu.eg, zbeltagy@eng.cu.edu.eg, MMAKARY@mgh.harvard.edu, m.almasani@sejong.ac.kr

Abstract. Magnetic Resonance Imaging (MRI) is a vital non-invasive imaging modality that is highly sensitive to patient motion, which disrupts k-space consistency and produces structured artifacts in reconstructed images. Most existing deep learning–based motion correction methods operate exclusively in the image domain, implicitly discarding phase information and limiting their ability to correct motion-induced frequency distortions. Although several recent approaches incorporate k-space awareness, they primarily use k-space to detect corrupted lines or impose data-consistency constraints, while motion correction itself remains confined to the image domain. As a result, phase-inconsistent frequency components introduced during acquisition are not explicitly corrected. To address this limitation, we propose KiMoCo-Net, a dual-domain motion correction framework that aligns learning-based correction with the physics of MRI acquisition. KiMoCo-Net consists of two sequential stages: (i) a frequency-domain correction network that operates directly on complex-valued k-space to restore phase and magnitude consistency prior to reconstruction, and (ii) a spatial-domain refinement network that enhances anatomical structures after inverse Fourier transformation. The model is trained end-to-end using a hybrid objective that jointly enforces k-space fidelity and image-domain structural similarity. Extensive experiments on synthetic and real motion-corrupted MRI data demonstrate that KiMoCo-Net outperforms state-of-the-art image-domain and hybrid methods under severe motion while exhibiting robust performance across motion severities, achieving substantial improvements in structural fidelity and image sharpness (18.2% SSIM and 7.61% PSNR gains). Moreover, KiMoCo-Net maintains low inference time, supporting its practical applicability in clinical settings.

Keywords: MRI motion correction, K-space learning, Image reconstruction, Dual-domain networks.

1 Introduction

Magnetic Resonance Imaging (MRI) acquires images by sampling spatial frequency information in k-space and reconstructing the image through an inverse Fourier transform. This acquisition process is inherently sensitive to patient motion, which disrupts k-space consistency and leads to structured artifacts such as ghosting and blurring in the reconstructed image [1]. Motion artifacts remain a major practical challenge in MRI, degrading diagnostic quality and often requiring costly rescans.

Existing motion artifact correction strategies can be broadly categorized into prospective and retrospective approaches. Prospective methods compensate for motion during acquisition but require additional hardware and complex system integration, limiting their applicability in routine clinical settings [2]. Retrospective methods correct motion after data acquisition; therefore, they are more practical and widely adopted. However, classical retrospective approaches rely on explicit motion estimation and often fail under complex or severe motion patterns [3].

Recent learning-based methods have demonstrated promising performance in MRI motion artifact suppression by directly learning a mapping from corrupted to clean images. Most existing deep learning approaches [4-6] operate exclusively in the image domain, applying CNNs to reconstructed magnitude images. While effective in reducing visible artifacts, such methods overlook a critical fact: motion corrupts MRI data at the k-space acquisition level, introducing phase inconsistencies and frequency-domain distortions that cannot be fully recovered once the image is reconstructed.

Several recent works incorporate k-space awareness into learning-based motion correction, for example, by detecting corrupted k-space lines or enforcing data-consistency losses. Safari et al. [7] proposed PI-MoCoNet, which combines a U-Net for identifying corrupted k-space lines with a Swin U-Net for image-space correction. Corrupted lines are used to guide a k-space data-consistency loss, selectively constraining the reconstruction. More recently, Hemidi et al. introduced IM-MoCo [8], a two-stage framework that separates motion detection and correction. In IM-MoCo, k-space lines are classified as corrupted or clean, and correction is performed through an implicit forward reconstruction model operating in image space. However, in these approaches, motion correction itself is still performed in the image domain, with k-space used primarily as a supervisory or regularization signal. This indirect use of k-space limits the ability to explicitly correct phase and frequency-domain errors introduced during acquisition. Other recent works [9-11] estimate motion parameters directly from k-space and apply them to compensate for motion-induced corruption.

Hybrid frequency–image domain architectures have been explored in related contexts, such as accelerated. For example, a hybrid frequency-domain and image-domain network combines k-space and image-space processing for under-sampled data [12].

Inspired by these perspectives, we propose KiMoCo-Net, an end-to-end dual-domain k-space and image-space motion correction network composed of two cascaded stages. The first network operates in k-space, correcting motion-induced corruption by jointly modeling phase and magnitude information. The corrected k-space is then transformed into the image domain, where a second network refines the reconstruction and suppresses residual artifacts. By aligning the correction process with the physical origin of motion artifacts, KiMoCo-Net enables more faithful signal recovery and improved image fidelity compared to image-domain-only or k-space-guided methods. The effectiveness of the proposed framework is validated through extensive experiments on both synthetically simulated and real motion-corrupted MRI data, demonstrating robust performance under realistic acquisition conditions. Code is available at https://github.com/marina513/Prior-Mixer-KiMoCoNet.

2 Materials and Methods

2.1 Problem Formulation and Motion Simulation

MRI acquisition is performed in k-space, where phase-encoding lines are sequentially sampled over time. Motion occurring during acquisition disrupts k-space consistency and leads to structured artifacts in the reconstructed image. This process can be expressed as:

$$\widetilde{K} = \sum_{n=1}^{N} S_n \mathcal{F}(M_n I_{GT}) , \tag{1}$$

where $I_{GT} \in \mathbb{C}^{H \times W}$ denotes the motion-free ground-truth image, M_n represents the rigid motion operator (i.e., translation or rotation), $\mathcal{F}(\cdot)$ is the Fourier transform, and $\widetilde{K}$ denotes the motion-corrupted k-space. The binary mask S_n selects the k-space lines acquired under the nth motion state, and N denotes the number of distinct motion events. Acquisition of phase-encoding lines under varying motion states leads to k-space inconsistencies across spatial configurations, which are globally propagated by the inverse Fourier transform.

To simulate motion artifacts, we use data from the publicly available Calgary–Campinas dataset [13], consisting of 102 volumes with a spatial size of 100×256×256. In total, 9,000 2D slices are used for training and 1,200 for testing. Synthetic rigid motion is generated using Eq. (1) by controlling three parameters: rotation (±5° to ±25°), translation (±5 mm to ±25 mm), and the number of motion events (5 to 25). Based on these parameters, we define three motion severity levels: mild motion (±5), moderate motion (±15), and severe motion (±25).

To evaluate generalization under real acquisition conditions, we further test the model on the PMoC3D dataset [14], which contains real motion-corrupted MRI scans.

2.2 KiMoCo-Net: Dual-Domain Motion Correction Framework

We introduce KiMoCo-Net, a sequential dual-domain network with complex-valued k-space correction followed by image-domain refinement. While hybrid dual-domain architectures have previously been investigated for under-sampled MRI reconstruction [12], motion corruption represents a fundamentally different inverse problem. KiMoCo-Net decomposes correction into two stages: *k-space correction* that restores frequency consistency, and *image-domain refinement* that removes residual artifacts and enhances anatomical details. Restoring frequency coherence first allows the image-

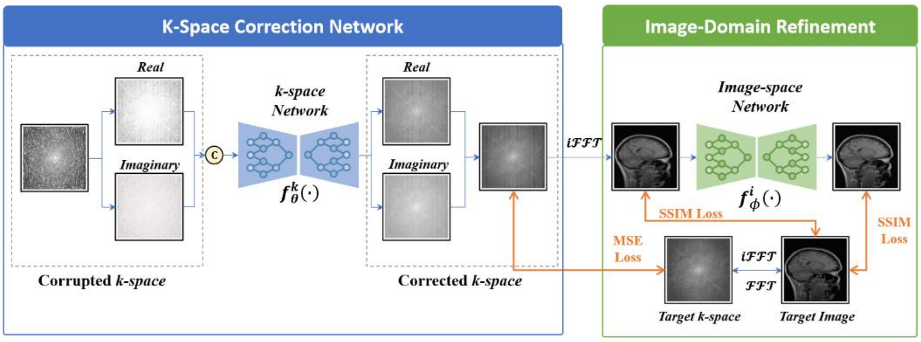

Fig. 1. KiMoCo-Net dual-domain architecture. The network consists of two sequential stages. Left: frequency-domain correction operates on the real and imaginary components of motion-corrupted k-space, restoring phase and magnitude consistency via a residual network. Right: image-domain refinement operates on the inverse Fourier-reconstructed image to suppress residual artifacts and enhance structural details.

stage network to operate on a better-conditioned reconstruction, focusing on structural refinement rather than compensating for globally mixed Fourier artifacts. An overview of the architecture is shown in **Fig. 1**. Mathematically, the motion corrected k-space $\widehat{K}$ and image-space $\hat{I}_{final}$ through the proposed KiMoCo-Net can be expressed as:

$$\widehat{K} = f_\theta^k(\widetilde{K}), \qquad \hat{I}_{final} = f_\phi^i\left(\mathcal{F}^{-1}(\widehat{K})\right), \tag{2}$$

where f_θ^k and f_ϕ^i represent the k-space and image-space networks, respectively.

K-Space Correction Network: The goal of the k-space stage is to restore frequency-domain consistency before image reconstruction. Given corrupted k-space $\widetilde{K}$, its real and imaginary components are concatenated along the channel dimension, enabling joint modeling of magnitude and phase information while remaining compatible with standard real-valued convolutional operations.

To stabilize optimization and balance the dynamic range across channels, per-sample normalization is applied before processing. The normalized k-space is then passed through an encoder-decoder correction network with residual learning, which encourages the model to focus on correcting motion-induced inconsistencies rather than reconstructing the entire signal. The k-space correction is formulated as:

$$\widehat{K} = f_\theta^k(\widetilde{K}) + \widetilde{K}. \tag{3}$$

The corrected k-space $\widehat{K}$ is subsequently denormalized and transformed into the image domain via inverse Fourier transform. The magnitude of the resulting complex-valued image is then computed to obtain the intermediate reconstruction $\hat{I}$. The k-space correction stage is supervised using a hybrid objective that combines a frequency-domain fidelity term with an image-domain structural constraint:

$$\mathcal{L}_k = \left\|\widehat{K} - K_{GT}\right\|_2^2 + \left(1 - \mathrm{SSIM}(\hat{I}, I_{GT})\right), \tag{4}$$

where K_{GT} and I_{GT} denote the motion-free k-space and image, respectively. This objective encourages restoration of frequency coherence while ensuring that corrections lead to structurally consistent reconstructions.

Image-Domain Refinement Networks: Although k-space correction significantly mitigates motion-induced inconsistencies, residual artifacts and local distortions may persist after reconstruction. The image-space refinement stage is designed to suppress these residual artifacts and enhance fine anatomical structures.

The image-domain network operates on the intermediate reconstruction produced by the k-space stage, where the reconstructed corrected image from the corrected k-space is passed into another encoder-decoder refinement network that focuses on spatial continuity and structural detail preservation. This stage is supervised using a structural similarity objective:

$$\mathcal{L}_i = 1 - \mathrm{SSIM}\big(\hat{I}_{final}, I_{GT}\big), \tag{5}$$

where $\hat{I}_{final}$ denotes the final refined reconstruction.

2.3 KiMoCo-Net Training and Optimization

By decomposing motion correction into frequency-domain correction and image-space refinement, KiMoCo-Net reduces the complexity of the overall inverse problem and allows each stage to address complementary aspects of motion-induced degradation.

The final training objective jointly optimizes both stages in an end-to-end manner using a combination of k-space and image-domain loss functions:

$$\mathcal{L} = \mathcal{L}_k + \mathcal{L}_i\ . \tag{6}$$

This joint optimization encourages coordinated learning across domains

Both the k-space and image-space networks follow a standard real-valued U-Net–based encoder–decoder architecture with identical depth. Each encoder consists of four resolution levels, progressively increasing the number of feature channels while reducing spatial resolution, followed by a symmetric decoder that restores the original resolution. Convolutional blocks are interleaved with instance normalization and ReLU activations, and residual connections are employed to stabilize training and focus learning on motion-induced distortions rather than full signal reconstruction.

The network is trained for 30 epochs using the Adam optimizer (β_1 = 0.9, β_2= 0.999) with a learning rate of 10^{-4} and batch size of 8. We used NVIDIA GeForce GTX 1080 Ti, requiring approximately 6 seconds to process each test volume.

3 Results

KiMoCo-Net consistently improves reconstruction quality across quantitative metrics and visual assessment. Compared to motion-corrupted inputs, it increases SSIM by 18.2% and PSNR by 7.6%, while reducing blurriness (BLUR) [15] by 3.9%, indicating enhanced structural fidelity and sharper anatomical details.

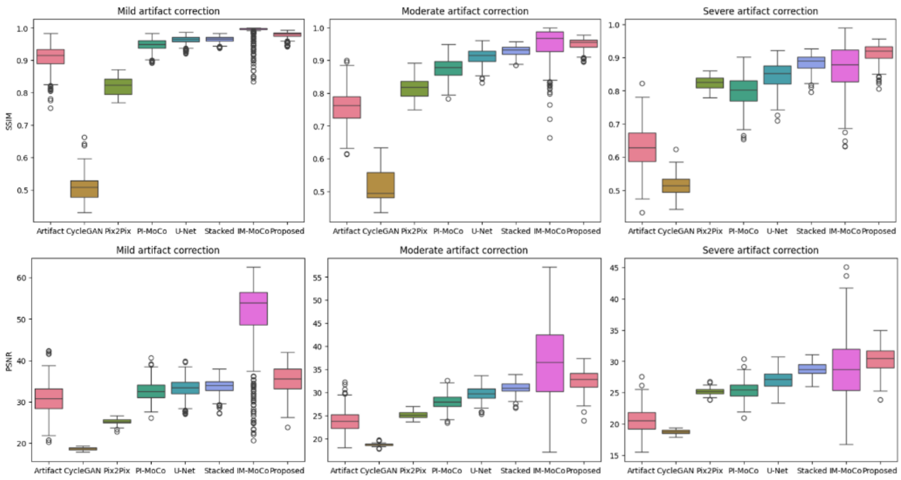

Fig. 2. Boxplots of SSIM (top) and PSNR (bottom) across three motion severity levels (mild, moderate, severe). KiMoCo-Net achieves higher median scores and fewer extreme outliers.

Table 1. Quantitative comparison of KiMoCo-Net with state-of-the-art motion correction methods in terms of SSIM, PSNR, blurriness (BLUR), and inference time (seconds). * indicates a statistically significant difference from KiMoCo-Net according to the Wilcoxon signed-rank test ($p < 0.05$); † indicates no statistically significant difference ($p \geq 0.05$).

Method	**SSIM ↑**	**PSNR ↑**	**BLUR ↓**	**Time**
Corrupted Simulated Data	*0.765±0.124**	*25.07±5.02**	*0.449±0.0213**	-
CycleGAN [16] (ICCV 2017)	0.685±0.169*	23.05±4.73*	0.420±0.0277*	0.046
Pix2Pix [17] (CVPR 2017)	0.820±0.144*	25.20±3.49*	0.378±0.0327*	0.003
PI-MoCo [7] (Med. Phys. 2025)	0.873±0.069*	28.73±3.50*	0.418±0.0253*	0.033
U-Net [18] (MICCAI 2015)	0.906±0.055*	30.07±3.05*	0.430±0.0332*	0.031
Stacked U-Net [6] (NeuroImage 2022)	0.926±0.036*	31.19±2.42*	0.423±0.0272*	0.145
IM-MoCo [8] (MICCAI 2024)	0.935±0.073†	38.80±11.8*	0.416±0.0259*	50.03
KiMoCo-Net (Ours)	0.947±0.032	32.68±3.28	0.410±0.0276	0.061

3.1 Comparison with State-of-the-Art

We compared KiMoCo-Net with CycleGAN [16], Pix2Pix [17], U-Net [18], PI-MoCo [7], Stacked U-Nets [6], and IM-MoCo [8] using identical training and testing configurations. Quantitative results averaged across different motion severities are summarized in **Table 1**. Overall, KiMoCo-Net achieves the highest SSIM among all competing methods, outperforming the baselines by margins ranging from 1.2%–26.2%.

To further analyze performance across different motion severities, we examined the boxplots shown in **Fig. 2**. KiMoCo-Net maintains consistently strong performance with fewer extreme outliers, indicating improved reconstruction stability. Although IM-MoCo achieves higher SSIM under mild motion (99.02% vs. 97.82%) and comparable performance under moderate motion (94.92% vs. 95.08%), its performance deteriorates substantially under severe corruption (86.80±6.8% vs. 91.35±2.7%). In contrast, KiMoCo-Net exhibits substantially lower variance and maintains superior performance under severe motion, indicating improved robustness to challenging motion corruption.

This observation is supported by Wilcoxon signed-rank tests, which reveals clear severity-dependent behavior. IM-MoCo achieves significantly higher SSIM under mild motion (p=2.15×10^{-37}), whereas KiMoCo-Net achieves significantly higher SSIM under severe motion (p=5.7×10^{-40}). Consequently, the overall SSIM difference between the two methods is not statistically significant (p=0.06), indicating that averaging across motion severities obscures important severity-dependent performance differences.

Similar severity-dependent trends are observed for PSNR (p<0.05). While IM-MoCo achieves the highest mean PSNR (38.80±11.8 dB), this result is largely driven by its superior performance under mild motion (51.13±8.66 dB vs. 35.29±3.30 dB for KiMoCo-Net). Under severe motion, however, KiMoCo-Net achieves higher PSNR (30.25±1.87 dB vs. 28.82±4.61 dB) while exhibiting substantially lower variance, indicating more reliable reconstructions under challenging motion conditions.

In addition, KiMoCo-Net offers a substantial computational advantage. The proposed method requires approximately 0.06 s per sample, compared with approximately 50 s for IM-MoCo.

3.2 Qualitative Evaluation on Simulated and Real Data

Fig. 3 presents qualitative comparisons under severe simulated motion. KiMoCo-Net more effectively suppresses ghosting and blurring artifacts, producing reconstructions with reduced and spatially localized error maps. Competing methods often exhibit residual distortions or over-smoothing, particularly along cortical boundaries. The proposed method preserves fine anatomical structures, consistent with the improved SSIM, PSNR, and blur metrics. **Fig. 4** shows reconstructions on real motion-corrupted MRI data, where motion-free ground truth is unavailable. Visual assessment indicates improved artifact removal, clearer anatomical delineation, and more homogeneous background regions across methods, particularly in the cerebellum. Residual streaking artifacts observed in competing methods are substantially reduced.

3.3 Ablation Study

To evaluate the contribution of each domain and determine whether the observed performance gains arise from increased architectural depth or cross-domain complementarity, we conducted an ablation study (**Table 2**) including: (1) a k-space U-Net trained exclusively in the frequency domain, (2) an image-domain U-Net, (3) cascaded image-domain U-Nets, and (4) the proposed dual-domain KiMoCo-Net. Single-domain models achieve comparable SSIM (0.908 and 0.906, respectively), and increasing image-domain depth alone does not yield further improvement (0.909). This indicates that additional spatial depth alone is insufficient for effective motion artifact mitigation.

In contrast, KiMoCo-Net achieves a substantial performance gain, reaching 0.947±0.032 SSIM, alongside improved PSNR and reduced blurriness. Notably, the

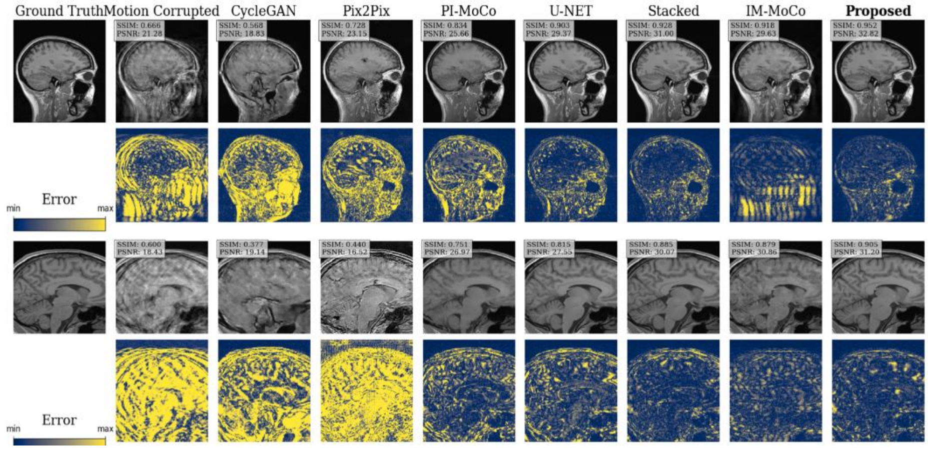

Fig. 3. Qualitative results under severe simulated motion artifacts. Reconstructed images (rows 1,3) and corresponding absolute error maps (rows 2,4; blue–yellow: low–high).

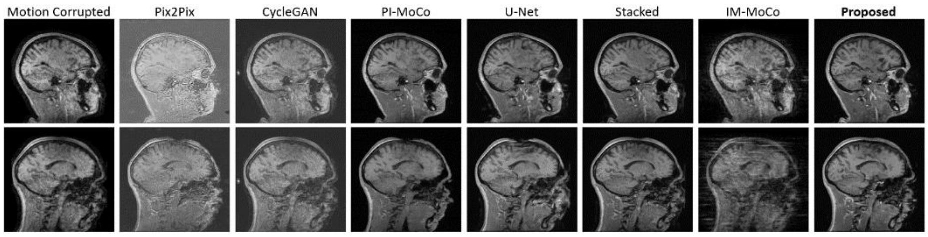

Fig. 4. Qualitative results on real motion-corrupted MRI scans, where motion-free ground truth is unavailable.

proposed model also exhibits lower variance compared to single-domain approaches, indicating enhanced reconstruction stability.

These results demonstrate that the observed improvements arise from complementary frequency- and image-domain correction, rather than increased network capacity.

Table 2. Ablation study evaluating the contribution of k-space correction and image-space refinement components in KiMoCo-Net.

Method	SSIM ↑	PSNR ↑	BLUR ↓
Corrupted Simulated Data	*0.765±0.124*	*25.07±5.02*	*0.449±0.0213*
K-space Only (Complex U-Net)	0.908±0.059	32.00±4.82	0.446±0.0263
Image-space Only (U-Net)	0.906±0.055	30.07±3.05	0.430±0.0332
Image-space Only Cascaded U-Nets	0.909±0.050	30.12±2.77	0.424±0.0313
KiMoCo-Net (K-space + Image-space)	0.947±0.032	32.68±3.28	0.410±0.0276

4 Conclusion

We present KiMoCo-Net, a sequential dual-domain framework for MRI motion correction. By explicitly restoring phase and magnitude consistency in k-space prior to image reconstruction, the proposed approach aligns the correction process with the physical origin of motion artifacts. This decomposition reduces the ill-posedness of the inverse problem, enabling the frequency-domain stage to address global inconsistencies while the image-domain stage focuses on refining fine anatomical details. Experiments on both synthetic and real motion-corrupted MRI data demonstrate that KiMoCo-Net achieves superior performance under severe motion while maintaining stable performance across different motion severities. The proposed framework maintains low inference time, supporting its practicality for clinical deployment.

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

References

1. Friston, K., Williams, S., Howard, R., Frackowiak, R., Turner, R.: Movement-related effects in fMRI time-series. Magn. Reson. Med. 35(3), 346–355 (1996). https://doi.org/10.1002/mrm.1910350312
2. Maclaren, J., Herbst, M., Speck, O., Zaitsev, M.: Prospective motion correction in brain imaging: a review. Magn. Reson. Med. 69(3), 621–636 (2013). https://doi.org/10.1002/mrm.24314
3. Atkinson, D., Hill, D., Stoyle, P., Summers, P., Clare, S., Bowtell, R., et al.: Automatic compensation of motion artifacts in MRI. Magn. Reson. Med. 41(1), 163–170 (1999). https://doi.org/10.1002/(SICI)1522-2594(199901)41:1%3C163::AID-MRM23%3E3.0.CO;2-9
4. Küstner, T., Armanious, K., Yang, J., Yang, B., Schick, F., Gatidis, S.: Retrospective correction of motion-affected MR images using deep learning frameworks. Magn. Reson. Med. 82(4), 1527–1540 (2019). https://doi.org/10.1002/mrm.27783
5. Duffy, B., Zhang, W., Tang, H., Zhao, L., Law, M., Toga, A., et al.: Retrospective correction of motion artifact affected structural MRI images using deep learning of simulated motion. NeuroImage 230, 117756 (2021). https://doi.org/10.1016/j.neuroimage.2021.117756
6. Al-masni, M., Lee, S., Yi, J., Kim, S., Gho, S., Choi, Y., et al.: Stacked U-Nets with self-assisted priors towards robust correction of rigid motion artifact in brain MRI. NeuroImage 259, 119411 (2022). https://doi.org/10.1016/j.neuroimage.2022.119411
7. Safari, M., Wang, S., Eidex, Z., Qiu, R. L., Chang, C. W., Yu, D. S., et al.: A physics-informed deep learning model for MRI brain motion correction. Med. Phys. 52(12), e70197 (2025). https://doi.org/10.1002/mp.70197
8. Al-Haj Hemidi, Z., Weihsbach, C., Heinrich, M.: IM-MoCo: Self-supervised MRI motion correction using motion-guided implicit neural representations. In: Int. Conf. Med. Image Comput. Comput.-Assist. Interv. (MICCAI), Lect. Notes Comput. Sci., vol. 15007, pp. 382–392 (2024). https://doi.org/10.1007/978-3-031-72104-5_37
9. Dabrowski, O., Falcone, J.-L., Klauser, A., Songeon, J., Kocher, M., Chopard, B., et al.: SISMIK for Brain MRI: Deep-Learning-Based Motion Estimation and Model-Based Motion

Correction in k-Space. IEEE Trans. Med. Imaging 44(1), 396–408 (2025). https://doi.org/10.1109/TMI.2024.3446450
10. Hossbach, J., Splitthoff, D. N., Cauley, S., Clifford, B., Polak, D., Lo, W. C., et al.: Deep learning-based motion quantification from k-space for fast model-based magnetic resonance imaging motion correction. Med. Phys. 50(4), 2148–2161 (2023). https://doi.org/10.1002/mp.16119
11. Singh, N. M., Dey, N., Hoffmann, M., Fischl, B., Adalsteinsson, E., Frost, R., et al.: Data consistent deep rigid MRI motion correction. Proc. Mach. Learn. Res. 227, 368–381 (2024).
12. Souza, R., Frayne, R.: A hybrid frequency-domain/image-domain deep network for magnetic resonance image reconstruction. In: Proc. SIBGRAPI, pp. 257–264 (2019).
13. Souza, R., Lucena, O., Garrafa, J., Gobbi, D., Saluzzi, M., Appenzeller, S., et al.: An open, multi-vendor, multi-field-strength brain MR dataset and analysis of publicly available skull stripping methods agreement. NeuroImage 170, 482–494 (2018). https://doi.org/10.1016/j.neuroimage.2017.08.021
14. Wang, K., Klug, T., Ruschke, S., Kirschke, J. S., Heckel, R.: Reliable evaluation of MRI motion correction: dataset and insights. arXiv:2506.05975 (2025). https://doi.org/10.48550/arXiv.2506.05975
15. Crété-Roffet, F., Dolmière, T., Ladret, P., Nicolas, M.: The blur effect: Perception and estimation with a new no-reference perceptual blur metric. In: Human Vision and Electronic Imaging XII, Proc. SPIE, vol. 6492 (2007). https://doi.org/10.1117/12.702790
16. Zhu, J., Park, T., Isola, P., Efros, A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp. 2223–2232 (2017). https://doi.org/10.1109/ICCV.2017.244
17. Isola, P., Zhu, J. Y., Zhou, T., Efros, A. A.: Image-to-Image Translation with Conditional Adversarial Networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1125–1134 (2017). https://doi.org/10.1109/CVPR.2017.632
18. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: Int. Conf. Med. Image Comput. Comput.-Assist. Interv. (MICCAI), Lect. Notes Comput. Sci., vol. 9351, pp. 234–241 (2015). https://doi.org/10.1007/978-3-319-24574-4_28