

ARTICLE TYPE

DEEP-DISORDER: Motion Correction in 3D MRI via Segment Reconstruction and Registration

Laurens Beljaards¹ | Martijn Nagtegaal¹ | Chinmay Rao¹ | Yiming Dong¹ | Matthias J.P. van Osch¹ | Nicola Pezzotti^{2,3} | Mariya Doneva⁴ | Marius Staring¹

¹Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands

²Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands

³Cardiologs, Paris, France

⁴Philips Innovative Technologies, Hamburg, Germany

Correspondence

Corresponding author Laurens Beljaards.
Email: l.r.beljaards@lumc.nl

Present address

Albinusdreef 2, 2333 ZA Leiden

Abstract

3D MR image acquisition is inherently time-intensive, rendering it susceptible to patient motion during scanning. This may introduce significant blurring and artifacts, potentially necessitating re-acquisition.

We propose a modular framework to retrospectively correct for intra-scan motion in 3D brain MRI, without active motion tracking. Serving as the backbone of our approach is an existing distributed and incoherent sampling scheme (DISORDER), combined with a fast network trained for highly undersampled reconstruction. This enables approximate reconstructions of anatomy after every few seconds, using only a tiny fraction of k-space data (< 2%). While these reconstructions are only approximate, we postulate they are sufficient to estimate motion patterns at said temporal resolution. Groupwise registration, notable for its elimination of registration bias, is utilized for estimating rigid motion parameters, which are leveraged to reconstruct the measured data with reduced motion artifacts.

The approach was evaluated on 94 retrospectively and 3 prospectively motion-corrupted *in vivo* 3D T1-weighted brain MRI acquisitions. The estimated motion parameters matched the known retrospective motion with 0.06 mm and 0.13° accuracy, resulting in an improvement in reconstruction quality from 0.942 ± 0.026 to 0.992 ± 0.003 SSIM for the retrospective scans. The prospective scans improved from 0.915 ± 0.024 to 0.936 ± 0.014 SSIM after correction in the case of gradual motion and from 0.764 ± 0.008 to 0.923 ± 0.011 SSIM for extreme motion.

In conclusion, the proposed approach, that is free of external tracking devices or navigators, successfully estimated and corrected 3D motion between small sub-portions of a scan. This resulted in vastly improved image quality, making volumetric MRI substantially more tolerant to motion.

KEY WORDS

Retrospective Motion Correction, Image Reconstruction, Deep Learning, Groupwise Registration, DISORDER

Word Count: 5521 (Including appendix)

1 | INTRODUCTION

Multi-shot MRI is an imaging technique used to obtain high-resolution anatomical images by acquiring k-space data over multiple excitation pulses, which is a time-intensive process that requires the subject to lie still for several minutes. This can be especially challenging for young or vulnerable patients. If the patient does move during MRI acquisition, this can introduce artifacts and distortions into the image with their severity depending on the amount of motion, and may in the worst case require the scan to be reacquired. One study estimated the costs of motion at \$115,000 per scanner per year in 2015¹, while another study

Abbreviations: CG, Conjugate Gradient; SENSE, Sensitivity Encoding; DISORDER, Distributed and Incoherent Sample Orders for Reconstruction Deblurring using Encoding Redundancy; DL-Recon, Deep Learning based Reconstruction model; CNN, Convolutional Neural Network; GAN, Generative Adversarial Network; TFE, Turbo Field Echo; SSIM, Structural Similarity Index Measure; PSNR, Peak Signal-to-Noise Ratio;

in 2020 estimated the annual cost of patient head motion at \$45,066 per scanner without pediatric examinations and \$364,242 with pediatric examinations². Thus, development of approaches that make MRI more robust to motion without increasing scan time or decreasing image quality is paramount. Retrospective intra-scan motion correction aims to reduce motion artifacts in the reconstruction of an MRI image, often requiring an estimate of the motion of the subject to do so^{3,4,5,6,7,8,9}. One way of obtaining motion parameters is via tracking devices, though depending on the approach there may be additional effort that complicates the workflow⁴, increased cost, and also the accuracy of the external tracking devices may be insufficient. On the other hand, data-driven approaches are preferred as they do not require external hardware.

Retrospective motion correction is deeply intertwined with image reconstruction, thus addressing these two problems jointly confers synergistic benefits. After all, a better motion estimate can enable better reconstruction and vice versa. Such an approach may involve incorporating a motion term in the forward model and iteratively estimating the motion parameters and motion-corrected image¹⁰. An example of such a work is the DISORDER¹¹ framework, which additionally introduces a specialized sampling scheme. This distributed and incoherent sampling scheme maximizes the sensitivity to motion, which improves the ability to detect motion and solve the aligned reconstruction problem. While these approaches have the benefit of being data consistent, the iterative process can be computationally expensive and difficult to converge. One work proposed integrating a CNN trained to remove motion artifacts to improve the conditioning and convergence of the joint search¹². Some approaches, on the other hand, avoid the repeated image updates in the joint optimization, for example using a scout scan to replace the image estimate, thereby speeding up the reconstruction but slowing down the acquisition¹³, or a deep-learning based reconstruction to replace the image estimate¹⁴, optimizing only the motion parameters using a data consistency loss.

AI-based reconstruction end-to-end models have fast inference times and can be trained to reconstruct images that appear free of motion artifacts, often using a physics-informed motion-corruption simulation to generate paired training data⁵. Approaches that do not explicitly model the motion parameters have been proposed using 3D CNNs¹⁵, foundation models¹⁶, Fourier aggregation¹⁷ and GANs¹⁸. However, if the underlying motion is not modeled explicitly, deviating from the measured k-space data (data inconsistency) is required to correct the motion artifacts. In the case of severe motion, such a model may be prone to hallucinations, potentially resulting in missing anatomical details or the generation of non-existent anatomy in the reconstruction^{19,14}.

We therefore propose a framework that retrospectively corrects for motion in 3D MRI requiring no external tracking device, reference scan, or fully sampled center of k-space. The proposed approach generates reconstructions of small scan segments that are registered in a groupwise fashion, thereby estimating a time series of rigid motion parameters to be used in a final motion-corrected reconstruction. Only the estimated motion parameters are used in the final motion-corrected reconstruction, enabling the enforcement of data consistency and avoiding hallucinations. The estimation of motion parameters by registering some form of low quality reconstructions for each scan segment has been applied before in self-navigated approaches, for example by repeatedly sampling rotated blocks of k-space and using the overlapping central data from each to generate low-resolution images²⁰, or by apodization of the repeatedly sampled central k-space data with a Fermi filter²¹. However, these approaches rely on a fully sampled center of k-space for each segment, requiring a slower sequence⁴ or a sequence that already oversamples the k-space center²¹. We instead postulate that deep learning enables the reconstruction of small scan segments without a fully sampled k-space center at a sufficient quality for motion estimation. To ensure each scan segment contains a few lines from the center of k-space, we use the distributed incoherent sampling introduced for the DISORDER¹¹ method. We have therefore named our framework DEEP-DISORDER.

In the following, we detail our approach, and compare it to the original DISORDER reconstruction on retrospectively motion-corrupted data, as well as on prospectively motion-corrupted and motion-free 3D T1-weighted *in vivo* brain scans.

2 | METHODS

2.1 | Problem Definition

The data acquisition process is modeled by the forward equation $y = Ax$, with y denoting the measured k-space data, x denoting the underlying image, and the forward operator $A = MFSC_\theta$ describing the acquisition process. Here, C_θ is the segment-wise motion corruption operator that projects the image into different motion states with different motion parameters θ . In the proposed pipeline, trilinear interpolation is used in the operator C_θ . S represents the coil sensitivity encoding operator, F is the Fourier transform, and M is the masking operator that selects k-space lines by applying different sampling masks to the different

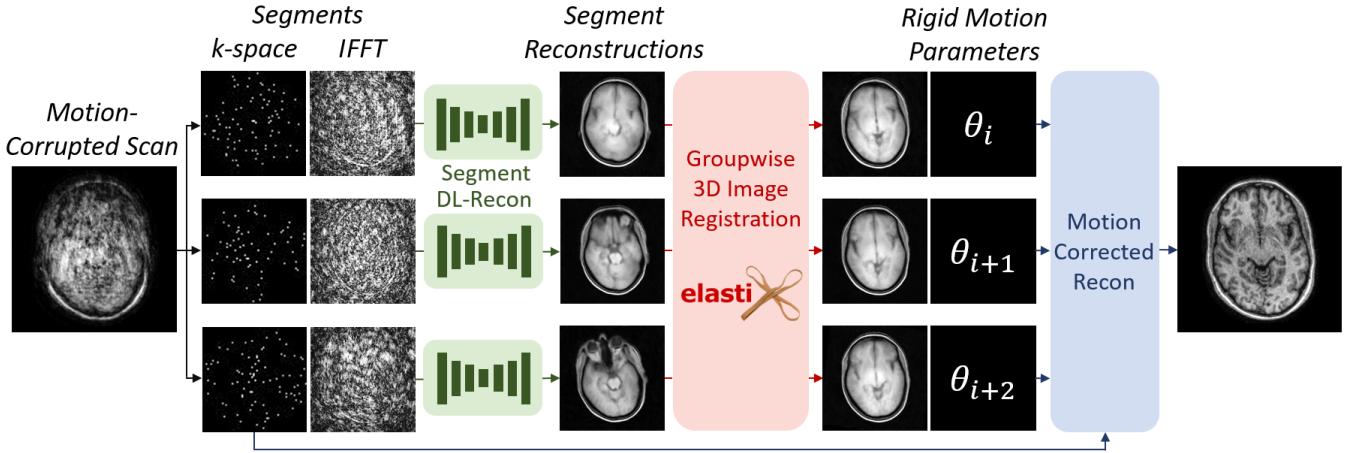


FIGURE 1 Overview of the proposed motion correction approach. The acquired motion-corrupted k-space is split into segments based on acquisition time. A zero-filled fast Fourier transform is applied and an unrolled 3D U-Net performs reconstructions for each extremely undersampled segment. These reconstructions are aligned using groupwise registration, yielding one set of estimated motion parameters per segment. Finally, these motion parameters and the measured data are used in the forward operator of a classical reconstruction approach to produce a motion-corrected image. The coil sensitivities are omitted from the figure for clarity.

segments. We define a segment as a fraction of k-space that represents one estimated motion state. As our acquisition process (a 3D-T1-weighted TFE sequence) comprises multiple shots, it follows naturally that we partition the k-space at shot boundaries with each segment consisting of a whole number of shots. Estimated motion parameters $\hat{\theta}$ can be used to invert the displacements and rotations that occurred during the acquisition for each segment, yielding the motion-corrected reconstruction problem

$$\hat{x} = \underset{x}{\operatorname{argmin}} \|MFSC_{\hat{\theta}}x - y\|_2^2, \quad (1)$$

which can be solved using CG-SENSE²² adapted for motion compensation^{18,23}. Given the estimated forward operator $A = MFSC_{\hat{\theta}}$, the least-squares solution for a given A can be found by solving the equation:

$$A^H Ax = A^H y. \quad (2)$$

Since the matrix $A^H A$ is too large to be inverted or stored explicitly, the system is solved iteratively using the Conjugate Gradient (CG) algorithm²⁴.

2.2 | DISORDER Reconstruction

The original DISORDER reconstruction approach¹¹ solves for x by iteratively updating the estimated motion parameters θ for a fixed image, and updating the image using fixed motion parameters:

$$\hat{x}^{(i+1)} = \underset{x}{\operatorname{argmin}} \|MFSC_{\hat{\theta}^{(i)}}x - y\|_2^2, \quad (3)$$

$$\hat{\theta}^{(i+1)} = \underset{\theta}{\operatorname{argmin}} \|MFSC_{\theta}\hat{x}^{(i+1)} - y\|_2^2. \quad (4)$$

DISORDER introduced a distributed incoherent sampling scheme that for each shot acquires a broad variety of frequencies from all over k-space, making the motion easier to detect and making the problem easier to solve. We feed the final motion parameters estimated by DISORDER to the same motion-corrected CG-SENSE algorithm described in Equation 1 to obtain a motion-corrected reconstruction comparable to the proposed approach.

2.3 | DEEP-DISORDER Reconstruction

An overview of the proposed retrospectively motion-corrected MRI reconstruction approach is given in Figure 1. The framework estimates 3D motion by reconstructing and registering segments of the acquired MRI data. We use the DISORDER sampling scheme to ensure a balanced distribution of low and high frequencies in every shot. As a single shot may contain very little anatomical information, we group multiple shots into a segment to reach an image quality that is sufficient for the subsequent registration process. Reconstructing the segments by zero-filling does not yield sufficient image quality for registration as these reconstructions lack clear anatomical features due to the extremely high acceleration factor, and each segment has different undersampling artifacts, as seen in Figure 1. Moreover, the segment reconstruction problem is too ill-posed for classical reconstruction algorithms. Therefore, we pass the segments through a deep learning reconstruction network that was trained to reconstruct extremely highly undersampled data. The segment reconstruction network f_{recon} reconstructs the n^{th} k-space segment as

$$\hat{x}_n = f_{\text{recon}}(S^H F^{-1} M_n^H y; \phi). \quad (5)$$

During training, a random segment index r and random rigid transformation C_r are generated to obtain a synthetic motion-corrupted k-space segment $y_r = M_r F S C_r x$ from uncorrupted image x . The network parameters ϕ are learned according to the optimization objective

$$\phi^* = \underset{\phi}{\operatorname{argmin}} \mathcal{L}(C_r x, f_{\text{recon}}(S^H F^{-1} M_r^H y_r; \phi)) \quad (6)$$

with loss function \mathcal{L} . These segment reconstructions are limited in quality due to the high undersampling factor and resulting lack of information, but we postulate that they are sufficiently accurate for registration purposes. The magnitude images of the segment reconstructions are registered with groupwise registration²⁵ using the medical image registration toolbox Elastix^{26,27}. By employing groupwise registration, we avoid having to select an arbitrary fixed image, eliminating bias in the selection by optimizing the group objective across all images via a consensus template representative of the entire group. This could be especially relevant if the quality of the segment reconstructions varies, which could potentially happen in the case of intra-segment motion, or variations in coil sensitivity due to the position of the subject. Additionally, groupwise registration avoids the accumulation of errors inherent in sequential pairwise registrations since transformations are jointly estimated.

A registration mask was used to ignore the neck region, which is often affected by non-rigid motion, and instead focus on head registration. During inference, the segment reconstructions are registered to estimate the rigid motion parameters

$$\{\hat{\theta}_1, \dots, \hat{\theta}_N\} = f_{\text{reg}}(\{\hat{x}_1, \dots, \hat{x}_N\}). \quad (7)$$

The groupwise registration objective function of f_{reg} is described by

$$\{\hat{\theta}, \hat{T}\} = \underset{\theta, T}{\operatorname{argmin}} \sum_{s=1}^N \left\{ \sum_v \left\| \theta_s(\hat{x}_s(v)) - T(v) \right\|_2^2 + \text{dist}(e, \theta_s) \right\}. \quad (8)$$

Here, θ_s is the transformation for segment s , $T(v)$ is the common template for voxel v , and $\text{dist}(e, \theta_s)$ is a regularization term enforcing closeness to the identity transformation e ²⁸.

Finally, as described in Equation 1, the resulting registration parameters are used as motion parameters in the motion-corrected CG-SENSE reconstruction algorithm to reconstruct the measured data with greatly reduced motion artifacts.

2.4 | Training and Implementation Details

For segment reconstruction we trained an unrolled 3D U-Net^{29,30} with four resolution levels with 16, 32, 64, and 128 feature maps per level. Each layer of the model consists of group normalization, 3D convolution and ReLU operators, respectively. The model received a highly undersampled zero-filled 3D image as input and was tasked to estimate the corresponding uncorrupted fully sampled 3D image. A blend of SSIM³¹ loss and smooth-L1 loss was used during training. Both the input and target were randomly rotated and translated during training to account for different patient positions.

Unrolled or iterative networks are commonly used in MRI reconstruction^{32,33}, as they enable iterative reconstruction in which data consistency is continually enforced. Since each segment has few k-space lines, the benefit of data consistency that an iterative approach provides is small. Nonetheless, we observed a minor improvement in performance over a single-stage approach from 0.043 to 0.033 NMSE on 90 retrospectively corrupted scans, albeit at the cost of an increased GPU memory footprint. We thus use a segment reconstruction approach with six consecutive residual U-Nets that were trained end-to-end. The reconstruction models were pretrained for 16 segments on an NVIDIA RTX A6000 GPU for 1×10^6 iterations with a batch size of 2 and a group normalization size of 8, and fine-tuned on the desired number of segments for 4×10^5 iterations, which was 64 for most experiments. We used a leave-one-out approach to reduce out-of-distribution issues on the test set, by further fine-tuning multiple models on a subset of motion-free volunteer scans and evaluating each on the unseen volunteer.

For groupwise registration we used Elastix with a multi-resolution pyramid of 4 levels and 2500 iterations per resolution. The schedule consists of three pyramids, where the result of the previous pyramid serves as an initialization of the next. The used registration pipeline and settings differ from DISORDER, as the proposed approach performs only a final groupwise registration while DISORDER performs many quick motion update steps alternating with reconstruction steps. We selected a high number of iterations and resolution pyramids to ensure convergence. We did not optimize these settings for runtime performance. The registration took about 10 minutes for $R=2$ on the CPU.

2.5 | Number of Segments

Besides the hyperparameters of the segment reconstruction and registration, there are only a few factors of influence for our method, most notably the number of segments. There is a natural tradeoff between the achievable image quality for each segment and the temporal resolution for the motion correction. To find a balanced configuration for which the proposed approach performs best, we trained and evaluated our approach on acquisition patterns with different numbers of segments, which were constructed using random-checkered sampling¹¹. Figure 2 shows the reconstruction performance for a given number of segments on 90 retrospectively sampled motion-corrupted scans with shaking or nodding-like motion with 320 motion states, demonstrating that the ideal number of segments is around 64 for our current setup.

2.6 | Generalizability

An experiment is included in the appendix that investigates generalizability to different retrospectively motion-corrupted data, specifically two spoiled gradient echo 3D-scans with different echo-times, with and without fine-tuning. The segment reconstruction model can handle unseen data that is moderately similar to the training data, but the performance lowers as the contrast gets further from T1w data which the model was trained on. It is demonstrated that this issue can be solved by including a number of scans of the unseen contrast in the training set. Additionally, successful correction of motion in higher resolution scans (384^3 voxels) was demonstrated when the motion estimation is run on a lower resolution (192^3 voxels) to reduce the need for compute resources.

2.7 | Data

Retrospectively motion-corrupted data: We focus our experiments on T1-weighted 3D MRI scans. We used an in-house dataset of 361 3D T1-weighted turbo field echo scans from our radiological clinic without intentional motion for training and validation, and 94 for the synthetic test set. 416 of the scans were acquired on Philips Ingenia 3.0T scanners, and 39 on an Ingenia Elition 3.0T X scanner. All participants provided informed consent and the study was approved by the institutional review board. The scans were performed in a transverse orientation. The flip angle was 8° , the repetition time (TR) around 9.9 ms, and the echo time (TE) around 4.6 ms. The scans were acquired with $1.8 \times$ acceleration. Coil-combined images were obtained using the scanner's built-in compressed sensing reconstruction without post-processing. These scans were retrospectively motion-corrupted by simulating the k-space acquisition process, giving us paired motion-corrupted and ground-truth data. For each motion state, we applied a translation and rotation to the coil-combined image and we sampled k-space lines from this transformed image according to the acquisition pattern. The acquisition matrices of the original data had different sizes, which we cropped or zero-padded to 240^3 . For inter-segment motion, we used one motion state per segment, meaning 64 motion states with multiple

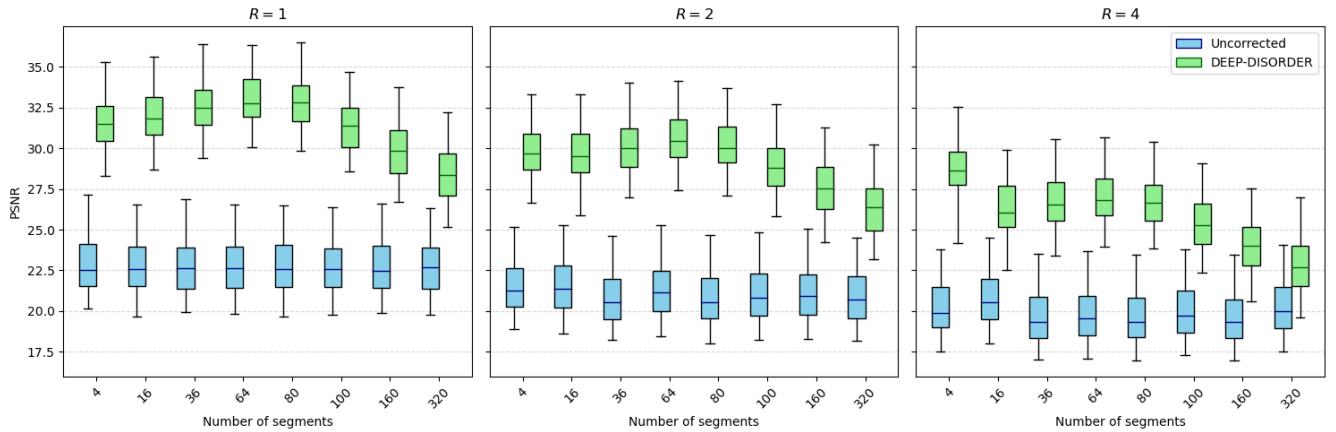


FIGURE 2 Motion-corrected reconstruction performance on synthetic data with different numbers of shots, with one shot per segment.

shots sharing one motion state. This is less realistic, but matches the number of motion states our approach can estimate, allowing us to see how well it resolves correctable motion. To simulate intra-segment motion, we used 320 motion states, giving each shot its own motion state. We used a TFE factor of 180 in our simulated acquisitions, meaning 180 k-space lines per shot and 320 shots, resulting in 900 k-space lines per segment for 64 segments. We obtained coil sensitivity maps from the SENSE reference scan that is automatically acquired for sequences that require coil sensitivity maps for image reconstruction. The coil sensitivity maps were compressed to 8 coils by a singular value decomposition³⁴ and normalized using root-sum-of-squares normalization. The scans were retrospectively undersampled by discarding a number of shots corresponding to the desired acceleration factor. Different motion patterns were used in the different experiments. To create random inter-segment motion, we sampled 64 translations and rotations from a normal distribution with a standard deviation of 2 mm and 1° for each axis ('spatially medium motion') or 12 mm and 6° ('spatially extreme motion'). This resulted in an average distance between the centers of two segments of 4.5 mm or 27.1 mm respectively. We optionally added intra-segment motion on top of this motion, randomly generated in a similar fashion but with 320 motion states and only 1/16th as severe, as we expected such motion to be less severe due to the shorter timeframe. Besides the random motion patterns, we also used more realistic rigid head motion patterns from a previous work³⁵ where a scanner with an optical tracking system was used to track participants who were instructed to perform shaking or nodding-like motion.

Prospectively motion-corrupted data: We acquired additional 3D T1-weighted prospectively motion-corrupted scans and still reference scans from 3 healthy volunteers for testing and 1 for validation, who all provided informed consent. The volunteers were instructed to move according to a specified motion pattern in each scan. For the gradual motion pattern, the volunteers were asked to occasionally slightly rotate their head as if slowly looking around, mimicking the drift often seen in clinical settings by patients who gradually relax or tire. For the jittery motion pattern, the volunteers were asked to abruptly rotate their head in a random direction approximately once every 15 seconds, such that their nose moves by approximately 2 cm each time. For the spatially extreme motion pattern, the volunteers were asked to look as far as possible to the bottom left at the start of the acquisition and rotate their head until they faced the top right at the end of the acquisition. The volunteers were informed in advance of the length of the scan and were asked to perform this motion in a stepwise fashion by moving every 15 seconds as it can be straining to slowly move continuously. This results in a reproducible large motion pattern with visible progression, which enables visually checking the reconstructed segments and registration for correctness while no ground truth motion is available. The first, middle and last segment reconstructions of the volunteer used for validation can be seen in Figure 1. We did not apply padding in the head coil, to enable more severe motion and test the limits of what motion can be corrected. The acquisitions used the MP-RAGE³⁶ sequence and were acquired on a 1.5T scanner (Ingenia, Philips Healthcare, Best, The Netherlands). The acquisitions were of size 191 × 191 × 384 and consisted of 133 or 68 shots for prospective accelerations $R=1$ and $R=2$ with a TFE factor of 216. Using 3 shots per segment, this leads to 648 k-space lines per segment. We used flip angles of 8°, TR of 4.6 ms and TE of 2.1 ms. The acquisitions lasted 6 minutes and 44 seconds for $R=1$ and 3 minutes and 29 seconds for $R=2$, and were performed in a sagittal orientation. Coil sensitivity maps were again obtained from the automatically acquired SENSE reference scan, compressed to 8 coils and normalized.

3 | RESULTS

For evaluation, we compare our method against DISORDER, which is closest to our work and an excellent example of another navigator-free and retrospective approach. In Section 3.1, we evaluate both the DISORDER and proposed DEEP-DISORDER approach on retrospectively motion-corrupted data. Section 3.2 shows the performance of both approaches on deliberately prospectively motion-corrupted MRI acquisitions of healthy volunteers. We selected the segment size to be 3 shots for both approaches on the retrospectively corrupted data. The final reconstructions were obtained by using the estimated motion parameters in the exact same motion-corrected CG-SENSE reconstruction algorithm with the same settings for both DISORDER and DEEP-DISORDER. Thus, the only difference in reconstruction quality between the approaches comes from the estimated motion. As is well known³⁷, running the CG algorithm for too many iterations can lead to amplified noise in the final reconstruction, so we selected 10 and 3 iterations for the retrospectively and prospectively corrupted data, respectively. The reconstruction quality was evaluated using two measures: the Structural Similarity Index (SSIM)³¹, which quantifies perceptual similarity by comparing luminance, contrast and structural information; and the Peak Signal-to-Noise Ratio (PSNR), which quantifies reconstruction fidelity via the ratio of the maximum possible signal power to the error power.

3.1 | Evaluation on Retrospectively Motion-Corrupted Data

Table 1 shows a comparison on the retrospectively motion-corrupted test set between a baseline uncorrected approach that uses a zero-motion estimate, DISORDER and DEEP-DISORDER. Each approach estimated 64 motion states. The first row evaluates DISORDER and DEEP-DISORDER on motion-free data. DISORDER estimates motion closer to zero due to the nature of its optimization loop, but neither approach deteriorates image quality. For medium random motion and nodding motion, both approaches correct the motion similarly well enough to produce near-perfect reconstructions. Of these, the pattern with the lowest SSIM after correction ($R=2$ intra-segment nodding motion) is shown in the top row of Figure 3. The motion-corrected image is almost indistinguishable from the motion-free image, bar a slight increase in noise. The remaining images show the performance for extreme motion. Once again a major improvement by both approaches over the uncorrected image can be seen, but unlike before, DEEP-DISORDER corrects it more accurately, both in quantitative measures and in visual appearance. For the extreme intra-segment motion, neither approach sufficiently corrects for the motion. After all, the motion that happens within a segment is not estimated and therefore uncorrected. While this extreme motion pattern effectively illustrates the breaking point of both approaches, the displacements are too large to typically occur in practice. In practice, the breaking point may be at less severe motion since not all effects of motion can be accounted for in the forward model.

3.2 | Evaluation on Prospectively Motion-Corrupted Data

Table 2 evaluates the baseline motion-agnostic reconstruction approach, the DISORDER approach and DEEP-DISORDER on prospectively motion-corrupted in-vivo data. The motion-corrected reconstructions are registered to a motion-free reference acquisition for evaluation. The lower neck area was ignored during this registration via a mask as there may be physical deformations between the reconstructed scans and the reference. The mouth and lower neck area were masked out for quantitative evaluation since the brain is the region of interest. The first row of Figure 4 shows the reconstructions of both approaches on volunteer acquisitions without intentional motion. The higher noise level compared to the $R=1$ reference scan can be attributed in large part to the fact that the data is two times undersampled, as the uncorrected reconstruction also shows increased noise. The results demonstrate that both approaches do not meaningfully deteriorate image quality even in cases without motion. On scans where the volunteers performed gradual motions, we see major blurring in the uncorrected reconstructions, especially at the front of the head, as highlighted by the arrows. The values in Table 2 do not show a large increase in SSIM for the gradual motion, but the reconstructions of DISORDER and DEEP-DISORDER are greatly deblurred as seen in the figure. Some amplified noise and minor artifacts remain visible, particularly in the $R=2$ case. The remaining images show examples of severe jittery and extreme motion. The uncorrected reconstructions have lost all detail, while DISORDER and DEEP-DISORDER produce clearly improved reconstructions. The $R=1$ jittery and $R=1$ extreme motion-corrected reconstructions no longer contain structural artifacts that hinder diagnostic ability, but the noise is highly amplified relative to the reference scan. Finally, the bottom example with $R=2$ extreme motion could not be sufficiently corrected and contains both structural artifacts and severe

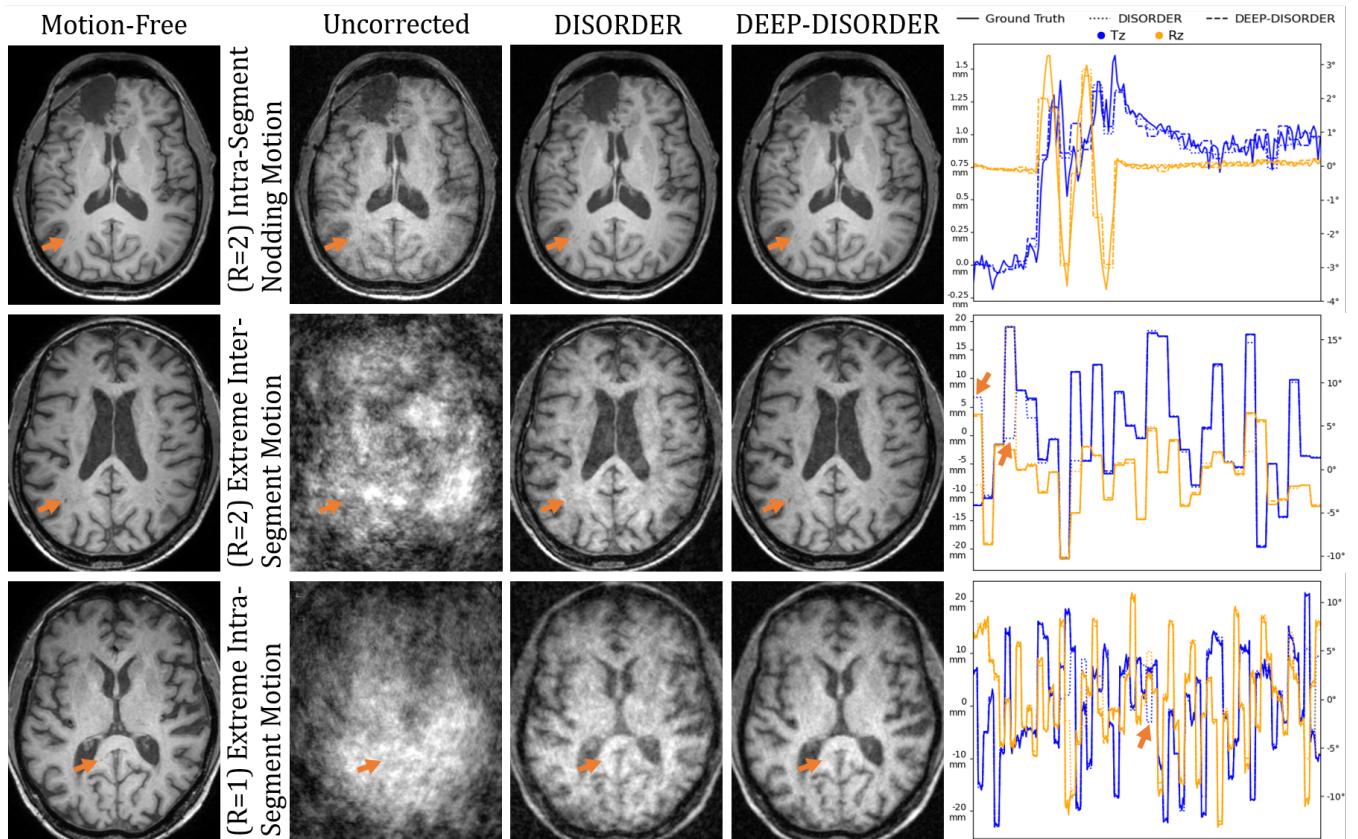


FIGURE 3 Reconstructions of retrospectively motion-corrupted and retrospectively undersampled data. The arrows highlight tiny structures that are rendered imperceptible by motion artifacts. DISORDER and DEEP-DISORDER correct the majority of motion artifacts, making the structures visible again for all but the most extreme motion patterns. The estimated and ground truth motion parameters are shown for a selected axis of translation (T) and rotation (R), with arrows highlighting outlying estimations by DISORDER.

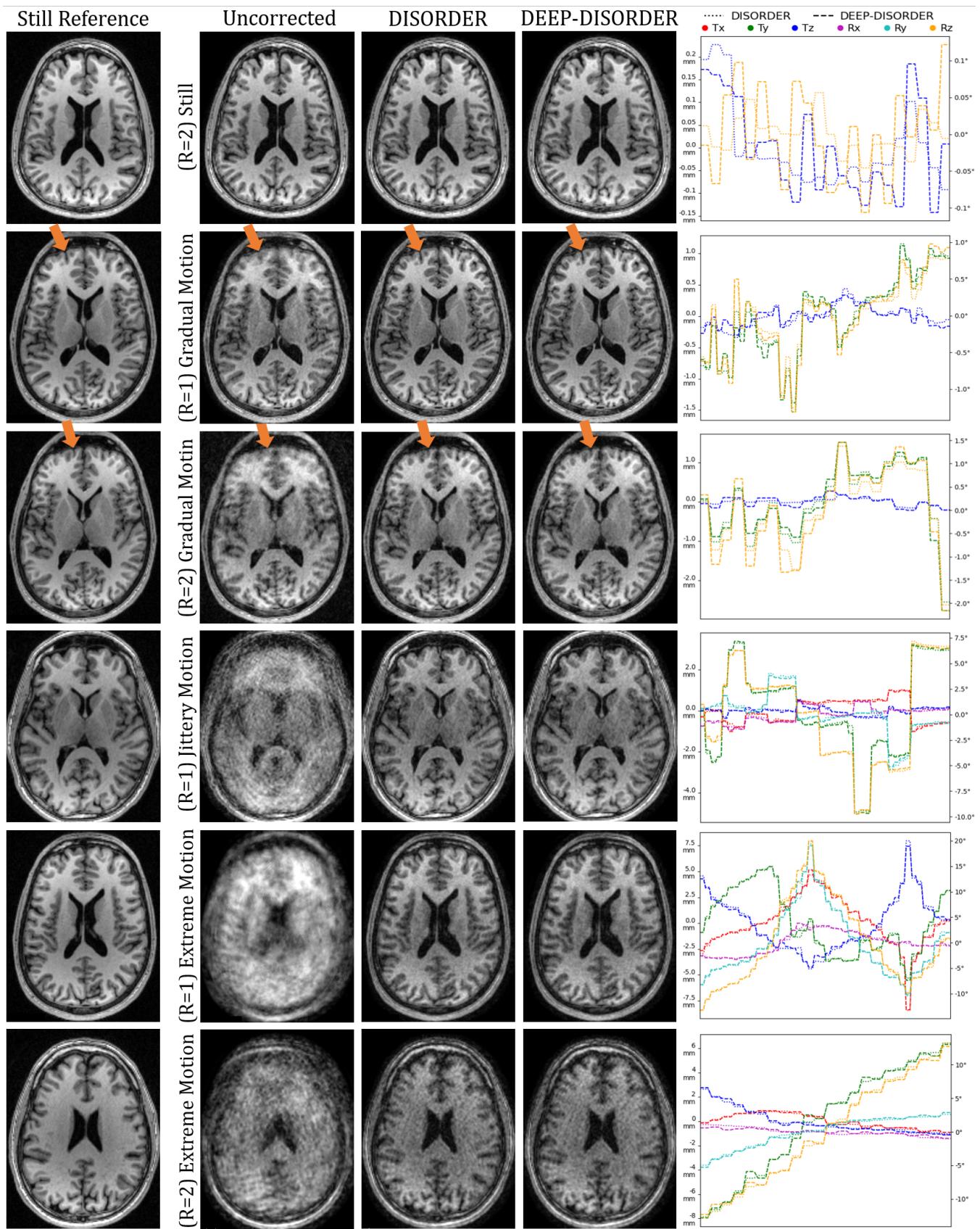


FIGURE 4 Reconstructions of prospectively motion-corrupted data, and corresponding $R=1$ still reference scans. The arrows highlight blurring caused by volunteer motion, most of which is corrected by DISORDER and DEEP-DISORDER. Arbitrarily selected translations (T) and rotations (R) are plotted on the right.

4 | DISCUSSION AND CONCLUSION

We presented a framework for motion estimation and motion-corrected reconstruction. The proposed approach reconstructs fractions (segments) of k-space and registers them to estimate the motion of the subject. The estimated parameters can be used in a motion-corrected reconstruction, such as a CG-SENSE reconstruction, to remove the majority of motion artifacts.

On the retrospectively motion-corrupted dataset, both the proposed DEEP-DISORDER and the reference approach (DISORDER) produced visually flawless images in the case of mild inter-segment motion. Neither approach perceptually degraded the image quality when no motion was present, indicating there is no downside to applying the motion correction when patients lie still. For spatially extreme inter-segment motion, DISORDER saw a drop in performance while DEEP-DISORDER successfully corrected for it with 0.16 mm accuracy. While DISORDER implicitly aligns the segments by optimizing only for the k-space fit, DEEP-DISORDER performs explicit image-domain registration. The latter enables a Gaussian multi-resolution pyramid strategy, which may increase the capture range and accuracy. A multi-resolution pipeline for DISORDER on the other hand would entail cropping k-space, reducing data to 4 to 5 k-space lines per shot for the lowest resolution, and it is not clear if DISORDER would still work with this little information. DEEP-DISORDER's explicit registration also enables the use of similarity metrics in the spatial domain, which potentially enriches or regularizes the problem definition, thereby making the optimization easier and more flexible. The experiments with retrospective intra-segment motion demonstrated that rapid motion occurring within a segment will elude correction, introducing mild blurring artifacts. However, in a realistic scenario, the brevity of each segment's acquisition window probably limits the severity of such intra-segment motion. Nevertheless, future research could investigate the possibility of intra-segment correction, potentially in combination with lower resolution registrations or deep-learning based registration to lower the registration time. Another direction for future work could be to explore hybrid solutions where DEEP-DISORDER is used to initialize a DISORDER reconstruction, for example the estimation of large motion by DEEP-DISORDER followed by refinement of the motion parameters at finer temporal scales by DISORDER.

For the prospectively motion-corrupted volunteer data, DISORDER and the proposed approach vastly reduced motion artifacts in the images, particularly for realistic gradual motion. The artifacts and noise amplification from the most extreme motion could not be completely corrected, though these voluntary motions were more severe than can be expected in practice as a result of our instructions to the volunteers since we wanted to find the breaking point of the correction approaches, and the fact that we deliberately scanned without padding in the head coil. The fact that the motion-corrected reconstructions of the prospectively motion-corrupted data still contain some motion artifacts and noise compared to the motion-free reference scans can be attributed to multiple factors. Firstly, not all effects of motion can be accounted for in the forward model, such as the intricate effects of motion on the coil sensitivities, B0 inhomogeneities and spin history effects. This directly hampers the final motion-corrected reconstruction quality, and potentially indirectly the segment reconstruction quality through less realistic training data. Mismatches between the coil sensitivities and the undersampled data due to motion are known to lead to aliasing and artifacts³⁸. Motion also results in suboptimal k-space sampling, thereby increasing the effective undersampling factor. Secondly, we used the basic conjugate gradient solver in our motion-corrected iterative reconstruction. While an optimized reconstruction algorithm with post-processing or denoising may produce better images, this was not studied by us as the focus was on the motion estimation. The approach performed slightly worse on the volunteer dataset when the segment reconstruction model was not fine-tuned on volunteer scans, mostly as a result of different acquisition settings between the test set that was sampled sagittally and the training set that was sampled transversely, negatively impacting the motion estimation. This seems to highlight a downside of our approach, as it involves training a 3D U-Net, which requires enough training data to be available and may limit generalizability if the data is not varied. However, deep learning-based products that can reconstruct a wide variety of contrasts (T1w, T2w, FLAIR, etc) with different acquisition settings and k-space coverage are in the portfolio of every major scanner vendor nowadays. It stands to reason that a segment reconstruction model can similarly learn to reconstruct segments from a wide range of contrasts and different k-space coverages if supplied with enough varied data during training. While we used 64 segments in the experiments, the optimal number of segments could vary depending on factors like the spatio-temporal scale of the acquisition, acquisition geometry, coil profiles and levels of motion. This optimal number is however unknown at scan time, as for example subject motion is not known in advance. Another point of attention is that for inference the motion-corrupted scans should be sampled with DISORDER-like sampling, which samples the same k-space lines as conventional sampling but in a different order to optimize the detectability of the motion.

Registered segments could be potentially checked automatically for quality issues as all segments show the same anatomy. For example, image difference metrics between the registered segment reconstructions could be used to estimate potential issues by corrupted shots within a segment. Additionally, corrupted shots at the border between two adjacent segments could be detected

by checking if the absolute distance between the corresponding motion parameters in the estimated motion trajectory exceeds a certain threshold. During the final motion-corrected reconstruction, individual shots or segments could be rejected or reduced in weight. Furthermore, data consistency metrics during the final reconstruction may help to identify outliers among the shots to reduce their influence on the final image quality. Moreover, when estimated continuously, feedback of estimated motion could help improve the efficiency of MRI examination. For example, the first two segments could already be used to detect motion early in the scan in less than 4% of total scan-time, enabling the MR technician to provide feedback to the patient, or acquiring additional data in an on-line manner when motion is detected during scanning.

In conclusion, this work outlined a framework for retrospectively correcting motion artifacts in 3D MR images. We introduced a deep learning-based segment reconstruction model that can estimate the anatomy of *in vivo* subjects on a fine temporal resolution (1/64th of k-space). These approximate reconstructions enable estimating subject motion using conventional groupwise registration. Finally, the estimated motion parameters can be leveraged in a motion-corrected reconstruction algorithm. In an evaluation on scans from volunteers, the approach did not deteriorate motion-free scans while removing the vast majority of motion artifacts from motion-corrupted scans, making volumetric MRI substantially more tolerant to motion. The proposed approach is modular and flexible in nature. For example, the rigid registration component could be replaced with deformable registration to enable non-rigid motion correction, which can be challenging for conventional iterative approaches. This may be explored in future work.

5 | AUTHOR CONTRIBUTION STATEMENT

Laurens Beljaards: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing. **Martijn Nagtegaal:** Data Curation, Investigation, Methodology, Writing – Review & Editing. **Chinmay Rao:** Data Curation, Software, Writing – Review & Editing. **Yiming Dong:** Data Curation, Software, Writing – Review & Editing. **Matthias J.P. van Osch:** Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Review & Editing. **Nicola Pezzotti:** Conceptualization, Funding Acquisition, Methodology, Software, Supervision, Writing – Review & Editing. **Mariya Doneva:** Conceptualization, Methodology, Supervision, Writing – Review & Editing. **Marius Staring:** Conceptualization, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Writing – Review & Editing.

6 | FUNDING AND ACKNOWLEDGEMENTS

This publication is part of the project ROBUST: Trustworthy AI-based Systems for Sustainable Growth with project number KICH3.LTP.20.006, which is (partly) financed by the Dutch Research Council (NWO), Philips, and the Dutch Ministry of Economic Affairs and Climate Policy (EZK) under the program LTP KIC 2020-2023. Laurens Beljaards and Chinmay Rao receive support from this project. Martijn Nagtegaal and Yiming Dong are supported by a PPS Allowance made available by Health~Holland Top Sector Life Sciences & Health. The authors thank Jochen Keupp, Jakob Meineke, Elwin de Weerdt and Kirsten Koolstra for their guidance and constructive discussions.

7 | CONFLICT OF INTEREST STATEMENT

The authors declare the following financial interests or personal relationships which may be considered as potential competing interests: Laurens Beljaards and Chinmay Rao declare the grant and research support they receive from Philips. Mariya Doneva declares employment at Philips. Nicola Pezzotti declares former employment at Philips. The C.J. Gorter MRI center receives research support from Philips. The other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

8 | DATA AVAILABILITY STATEMENT

The authors do not have permission to share data. Code is available at <https://github.com/MoCoMRI/DEEP-DISORDER>.

9 | REFERENCES

REFERENCES

1. Andre JB, Bresnahan BW, Mossa-Basha M, et al. Toward Quantifying the Prevalence, Severity, and Cost Associated With Patient Motion During Clinical MR Examinations. *Journal of the American College of Radiology*. 2015;12(7):689-695. doi: <https://doi.org/10.1016/j.jacr.2015.03.007>
2. Slipsager J, Glimberg S, Søgaard J, et al. Quantifying the Financial Savings of Motion Correction in Brain MRI: A Model-Based Estimate of the Costs Arising From Patient Head Motion and Potential Savings From Implementation of Motion Correction. *Journal of Magnetic Resonance Imaging*. 2020;52(3):731-738. doi: [10.1002/jmri.27112](https://doi.org/10.1002/jmri.27112)
3. Godenschweger F, Kägelein U, Stucht D, et al. Motion correction in MRI of the brain. *Physics in Medicine and Biology*. 2016;61(5):R32-R56. doi: [10.1088/0031-9155/61/5/R32](https://doi.org/10.1088/0031-9155/61/5/R32)
4. Zaitsev M, Maclarens J, Herbst M. Motion artifacts in MRI: A complex problem with many partial solutions. *Journal of Magnetic Resonance Imaging*. 2015;42(4):887-901. Epub 2015 Jan 28. PMID: 25630632; PMCID: PMC4517972. doi: [10.1002/jmri.24850](https://doi.org/10.1002/jmri.24850)
5. Chang Y, Li Z, Saju G, Mao H, Liu T. Deep learning-based rigid motion correction for magnetic resonance imaging: A survey. *Meta-Radiology*. 2023;1(1):100001. doi: <https://doi.org/10.1016/j.metrad.2023.100001>
6. Spieker V, Eichhorn H, Hammerink K, et al. Deep Learning for Retrospective Motion Correction in MRI: A Comprehensive Review. *IEEE Transactions on Medical Imaging*. 2024;43(2):846-859. doi: [10.1109/tmi.2023.3323215](https://doi.org/10.1109/tmi.2023.3323215)
7. Atkinson D, Hill D, Stoyle P, Summers P, Keevil S. Automatic Correction of Motion Artifacts in Magnetic Resonance Images Using an Entropy Focus Criterion.. *IEEE transactions on medical imaging*. 1997;16:903-10. doi: [10.1109/42.650886](https://doi.org/10.1109/42.650886)
8. Loktyushin A, Nickisch H, Pohmann R, Schölkopf B. Blind Retrospective Motion Correction of MR Images. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*. 2014;73. doi: [10.1002/mrm.24615](https://doi.org/10.1002/mrm.24615)
9. Haskell MW, Cauley SF, Wald LL. TArgeted Motion Estimation and Reduction (TAMER): Data Consistency Based Motion Mitigation for MRI Using a Reduced Model Joint Optimization. *IEEE Transactions on Medical Imaging*. 2018;37(5):1253-1265. doi: [10.1109/TMI.2018.2791482](https://doi.org/10.1109/TMI.2018.2791482)
10. Cordero-Grande L, Teixeira RP, Hughes E, Hutter J, Price A, Hajnal J. Sensitivity Encoding for Aligned Multishot Magnetic Resonance Reconstruction. *IEEE Transactions on Computational Imaging*. 2016;2:1-1. doi: [10.1109/TCI.2016.2557069](https://doi.org/10.1109/TCI.2016.2557069)
11. Cordero-Grande L, Ferrazzi G, Teixeira RPAG, O'Muircheartaigh J, Price AN, Hajnal JV. Motion-corrected MRI with DISORDER: Distributed and incoherent sample orders for reconstruction deblurring using encoding redundancy. *Magnetic Resonance in Medicine*. 2020;84(2):713-726. doi: <https://doi.org/10.1002/mrm.28157>
12. Haskell MW, Cauley SF, Bilgic B, et al. Network Accelerated Motion Estimation and Reduction (NAMER): Convolutional neural network guided retrospective motion correction using a separable motion model. *Magnetic Resonance in Medicine*. 2019;82(4):1452-1461. doi: <https://doi.org/10.1002/mrm.27771>
13. Polak D, Splitthoff DN, Clifford B, et al. Scout accelerated motion estimation and reduction (SAMER). *Magnetic Resonance in Medicine*. 2022;87(1):163-178. doi: <https://doi.org/10.1002/mrm.28971>
14. Singh NM, Dey N, Hoffmann M, et al. Data Consistent Deep Rigid MRI Motion Correction. *Proceedings of Machine Learning Research*. 2024;227:368-381. Proc Mach Learn Res. 2024;227:368-381.
15. Duffy BA, Zhao L, Sepehrband F, et al. Retrospective motion artifact correction of structural MRI images using deep learning improves the quality of cortical surface reconstructions. *NeuroImage*. 2021;230:117756. doi: <https://doi.org/10.1016/j.neuroimage.2021.117756>
16. Sun Y, Wang L, Li G, Lin W, Wang L. A foundation model for enhancing magnetic resonance images and downstream segmentation, registration and diagnostic tasks. *Nature Biomedical Engineering*. 2025;9(4):521-538. doi: [10.1038/s41551-024-01283-7](https://doi.org/10.1038/s41551-024-01283-7)
17. Solomon O, Patriat R, Braun H, et al. Motion robust magnetic resonance imaging via efficient Fourier aggregation. *Medical Image Analysis*. 2023;83:102638. doi: <https://doi.org/10.1016/j.media.2022.102638>
18. Usman M, Latif S, Asim M, Lee BD, Qadir J. Retrospective Motion Correction in Multishot MRI using Generative Adversarial Network. *Scientific Reports*. 2020;10(1):4786. doi: [10.1038/s41598-020-61705-9](https://doi.org/10.1038/s41598-020-61705-9)
19. Beljaards L, Pezzotti N, Rao C, Doneva M, Osch vMJP, Staring M. AI-based motion artifact severity estimation in undersampled MRI allowing for selection of appropriate reconstruction models. *Medical Physics*. 2024;51(5):3555-3565. doi: <https://doi.org/10.1002/mp.16918>
20. Pipe JG. Motion correction with PROPELLER MRI: application to head motion and free-breathing cardiac imaging. *Magnetic Resonance in Medicine*. 1999;42(5):963-969. doi: [10.1002/\(sici\)1522-2594\(199911\)42:5<963::aid-mrm17>3.0.co;2-1](https://doi.org/10.1002/(sici)1522-2594(199911)42:5<963::aid-mrm17>3.0.co;2-1)
21. Kecskemeti S, Samsonov A, Velikina J, et al. Robust Motion Correction Strategy for Structural MRI in Unsedated Children Demonstrated with Three-dimensional Radial MPnRAGE. *Radiology*. 2018;289:180180. doi: [10.1148/radiol.2018180180](https://doi.org/10.1148/radiol.2018180180)
22. Pruessmann KP, Weiger M, Börnert P, Boesiger P. Advances in sensitivity encoding with arbitrary k-space trajectories. *Magnetic Resonance in Medicine*. 2001;46(4):638-651. doi: [10.1002/mrm.1241](https://doi.org/10.1002/mrm.1241)
23. Guhaniyogi S, Chu ML, Chang HC, Song AW, Chen Nk. Motion immune diffusion imaging using augmented MUSE for high-resolution multi-shot EPI. *Magnetic Resonance in Medicine*. 2016;75(2):639-652. doi: <https://doi.org/10.1002/mrm.25624>
24. Hestenes MR, Stiefel E. Methods of conjugate gradients for solving linear systems. *Journal of research of the National Bureau of Standards*. 1952;49:409-435.
25. Metz C, Klein S, Schaap M, Walsum T, Niessen W. Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach. *Medical image analysis*. 2010;15:238-49. doi: [10.1016/j.media.2010.10.003](https://doi.org/10.1016/j.media.2010.10.003)
26. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging*. 2010;29(1):196-205. doi: [10.1109/TMI.2009.2035616](https://doi.org/10.1109/TMI.2009.2035616)
27. Shamonin DP, Bron EE, Lelieveldt BP, Smits M, Klein S, Staring M. Fast Parallel Image Registration on CPU and GPU for Diagnostic Classification of Alzheimer's Disease. *Frontiers in Neuroinformatics*. 2014;7. doi: [10.3389/fninf.2013.00050](https://doi.org/10.3389/fninf.2013.00050)
28. Wu G, Jia H, Wang Q, Shen D. Groupwise Registration with Sharp Mean. *Medical Image Computing and Computer-Assisted Intervention*. 2010;13(Pt 2):570-577. doi: [10.1007/978-3-642-15745-5_70](https://doi.org/10.1007/978-3-642-15745-5_70)
29. Wolny A, Cerrone L, Vijayan A, et al. Accurate and versatile 3D segmentation of plant tissues at cellular resolution. *eLife*. 2020;9:e57613. doi: [10.7554/eLife.57613](https://doi.org/10.7554/eLife.57613)
30. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. 2016:424-432.

31. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*. 2004;13(4):600–612.
32. Zeng G, Guo Y, Zhan J, et al. A Review on Deep Learning MRI Reconstruction without Fully Sampled k-space. *BMC Medical Imaging*. 2021;21(1):195. doi: 10.1186/s12880-021-00727-9
33. Pezzotti N, Yousefi S, Elmahdy MS, et al. An Adaptive Intelligence Algorithm for Undersampled Knee MRI Reconstruction. *IEEE Access*. 2020;8:204825-204838.
34. Buehrer M, Pruessmann KP, Boesiger P, Kozerke S. Array compression for MRI with large coil arrays. *Magnetic Resonance in Medicine*. 2007;57(6):1131–1139. doi: 10.1002/mrm.21237
35. Eichhorn H, Frost R, Kongsgaard A, et al. Evaluating the performance of markerless prospective motion correction and selective reacquisition in a general clinical protocol for brain MRI. *PsyArXiv*. 2022. doi: 10.31234/osf.io/vzh4g
36. Brant-Zawadzki M, Gillan GD, Nitz WR. MP RAGE: a three-dimensional, T1-weighted, gradient-echo sequence—initial experience in the brain. *Radiology*. 1992;182(3):769–775. doi: 10.1148/radiology.182.3.1535892
37. Maier O, Baete SH, Fyrdahl A, et al. CG-SENSE revisited: Results from the first ISMRM reproducibility challenge. *Magnetic Resonance in Medicine*. 2021;85(4):1821–1839. Epub 2020-11-12doi: 10.1002/mrm.28569
38. Hamilton J, Franson D, Seiberlich N. Recent advances in parallel imaging for MRI. *Progress in Nuclear Magnetic Resonance Spectroscopy*. 2017;101:71–95. doi: <https://doi.org/10.1016/j.pnmrs.2017.04.002>