Learning residual motion correction for fast and robust 3D multiparametric MRI

Carolin M. Pirkl [a,b,*], Matteo Cencini [c,d], Jan W. Kurzawski [e], Diana Waldmannstetter [a,f], Hongwei Li [a,f], Anjany Sekuboyina [a,f,g], Sebastian Endt [a,b], Luca Peretti [h,c,d], Graziella Donatelli [i,d], Rosa Pasquariello [c], Mauro Costagli [c,j], Guido Buonincontri [c,d], Michela Tosetti [c,d], Marion I. Menzel [k,b,l,1], Bjoern H. Menze [f,a,1]

[a] *Department of Computer Science, Technical University of Munich, Garching, Germany*
[b] *GE Healthcare, Munich, Germany*
[c] *IRCCS Fondazione Stella Maris, Pisa, Italy*
[d] *Fondazione Imago7, Pisa, Italy*
[e] *Pisa Division, National Institute for Nuclear Physics (INFN), Pisa, Italy*
[f] *Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland*
[g] *Department of Neuroradiology, Klinikum rechts der Isar, Munich, Germany*
[h] *Department of Computer Science, University of Pisa, Pisa, Italy*
[i] *Azienda Ospedaliero-Universitaria Pisana, Pisa Italy*
[j] *Department of Neuroscience, Rehabilitation, Ophtalmology, Genetics, Maternal and Child Sciences (DINOGMI), University of Genova, Genova, Italy*
[k] *Almotion Bavaria, Faculty of Electrical Engineering and Information Technology, Technische Hochschule Ingolstadt, Ingolstadt, Germany*
[l] *Department of Physics, Technical University of Munich, Garching, Germany*

ARTICLE INFO

ABSTRACT

Voluntary and involuntary patient motion is a major problem for data quality in clinical routine of Magnetic Resonance Imaging (MRI). It has been thoroughly investigated and, yet it still remains unresolved. In quantitative MRI, motion artifacts impair the entire temporal evolution of the magnetization and cause errors in parameter estimation. Here, we present a novel strategy based on residual learning for retrospective motion correction in fast 3D whole-brain multiparametric MRI. We propose a 3D multiscale convolutional neural network (CNN) that learns the non-linear relationship between the motion-affected quantitative parameter maps and the residual error to their motion-free reference. For supervised model training, despite limited data availability, we propose a physics-informed simulation to generate self-contained paired datasets from *a priori* motion-free data. We evaluate motion-correction performance of the proposed method for the example of 3D Quantitative Transient-state Imaging at 1.5T and 3T. We show the robustness of the motion correction for various motion regimes and demonstrate the generalization capabilities of the residual CNN in terms of real-motion *in vivo* data of healthy volunteers and clinical patient cases, including pediatric and adult patients with large brain lesions. Our study demonstrates that the proposed motion correction outperforms current state of the art, reliably providing a high, clinically relevant image quality for mild to pronounced patient movements. This has important implications in clinical setups where large amounts of motion affected data must be discarded as they are rendered diagnostically unusable.

1. Introduction

The capability to non-invasively provide high-resolution volumetric imaging with versatile soft-tissue contrast makes Magnetic Resonance Imaging (MRI) the modality of choice for the detection, characterization and monitoring of a wide variety of medical con-

* Corresponding author at: Department of Computer Science, Technical University of Munich, Garching, Germany.
E-mail address: carolin.pirkl@tum.de (C.M. Pirkl).
[1] Contributed equally.

ditions in modern radiology. Despite the various advantages of MRI over other medical imaging modalities, its main drawback for clinical routine is that image encoding and acquisition is inherently slow. Due to the prolonged scan times, MRI acquisitions are extremely susceptible to voluntary and involuntary patient motion. The arising motion artifacts impair diagnostic information and decrease the validity of obtained images, making diagnosis inaccurate or ambiguous. Images become diagnostically unusable and scans have to be repeated, potentially under sedation. Motion robustness is therefore a key requirement for data quality in routine imaging and is especially crucial for pediatric and elderly patients and those who are affected by diseases that prevent them from maintaining a still position throughout the acquisition. It is therefore of high clinical relevance to develop techniques that effectively prevent or resolve motion artifacts.

Subject motion causes erroneous signal allocation in *k*-space. How this mismatch propagates to image-space and in what kind of artifact appearance it manifests highly depends on:

- the *acquisition and readout schemes* (e.g. Cartesian vs. non-Cartesian),
- the targeted *diagnostic question* (e.g. static vs. time-resolved acquisition),
- the *body region* to be imaged (e.g. rigid vs. non-rigid organ motion) and / or
- the *condition of the patient* and his / her ability to cooperate.

Motion therefore remains a complex problem without a universal solution. A number of conceptually different motion mitigation, reduction and correction methods have been previously suggested (Zaitsev et al., 2015; Lee et al., 2021). In general, all these strategies can be classified into one of the following categories (Godenschweger et al., 2016):

- *Prevention* to avoid, suppress or minimize motion events in the first place, e.g., through fixation, sedation or breath-holding.
- *Sequence design* to increase motion-robustness of the acquisition, e.g., via accelerated scans, less sensitive (non-Cartesian) readout schemes or triggering / gating techniques.
- *Correction* to prospectively or retrospectively detect and account for motion, e.g., via external tracking, data-driven (self-) navigator information or incorporation in the reconstruction model.

For quantitative MRI, body motion is particularly problematic. This is because multiple acquisitions are required to encode the parameter information together with the spatial information. Despite its great potential for quantifying tissue parameters and hence for characterizing diseases, the resulting long scan times of quantitative MRI techniques together with the correspondingly high susceptibility to motion are the main barrier for establishing MRI-based tissue parameter mapping in routine clinical practice (Seiberlich et al., 2020).

With the overarching goal of accelerating MRI, significant methodological advances in terms of acquisition and reconstruction - with a great impulse due to the recent advances of deep learning methods - paved the way towards shorter scan times and potentially higher robustness to motion. From the initial incentive in the context of qualitative MRI, these methods are now also entering the field of quantitative MRI. In addition, there is great clinical interest to progress from state-of-the-art single-parameter to more efficient multiparameter mapping techniques, to overcome the practical limitations of quantitative MRI. In this context, the seminal work on MR Fingerprinting (Ma et al., 2013) for the joint mapping T_1, T_2 and proton density (*PD*) certainly constituted a major step forward and aroused the development of other efficient multiparameter encoding schemes in the transient state (e.g. Jiang et al., 2015; Christodoulou et al., 2018;

Gomez et al., 2019), demonstrating high quantification accuracy together with high repeatability and reproducibility (Jiang et al., 2017; Panda et al., 2019; Buonincontri et al., 2021). These attributes together with the improved motion robustness compared to conventional quantitative MRI with lengthy scanning protocols make transient-state techniques promising candidates for the clinical practice.

Although motion artifacts are generally reduced in these highly accelerated acquisition schemes, they are not entirely insensitive to patient motion. Previous work on motion correction for transient-state techniques almost exclusively concentrated on 2D acquisition schemes (Mehta et al., 2018; Cruz et al., 2019; Xu et al., 2019). To the best of our knowledge, for 3D variants thereof (e.g. Ma et al., 2018; Gómez et al., 2020), there is only the recently proposed navigator-based retrospective rigid motion correction of Kurzawski et al. (2020). For motion estimation and subsequent correction in 3D Quantitative Transient-state Imaging (QTI, Gómez et al., 2020), they rely upon intrinsic self-navigator information that is reconstructed from the segmented acquisition scheme. Although this realignment strategy was demonstrated to be capable of retaining a critical amount of the underlying multiparameter information, the temporal resolution of the self-navigators limits the resolvable motion scales so that quantitative maps of T_1, T_2 and *PD* showed remaining motion artifacts.

With the aim of improving on the navigator-based motion correction of Kurzawski et al. (2020), we developed a residual learning strategy for retrospective motion correction in fast 3D multiparametric MRI. We take their previously developed method and improve by adding five main contributions:

1. We propose a residual learning model to learn the non-linear relationship between the motion-affected T_1, T_2 and *PD* maps and the residual error maps, i.e., the deviation from the motion-free counterpart. In contrast to a direct mapping to the motion-free domain, this allows an efficient, decoupled representation and hence identification of the motion artifacts in the sparse residual space.
2. We rely on a 3D multiscale convolutional neural network (CNN) architecture to capture the intrinsic 3D nature of spatial correlations that inevitably arise (1) from spatially correlated image artifacts due to subject movements in the 3D space, and (2) from the 3D acquisition scheme with spatial undersampling and multi-coil imaging, additionally causing a mixing of signal components in the 3D space.
3. We present a physics-informed simulation of motion artifacts for generating self-contained, paired training data, enabling supervised model training without the need for large amounts of paired acquisitions or fully sampled data.
4. We evaluate the performance of the proposed method on real-motion data of healthy volunteers acquired with the 3D QTI scheme at 1.5T and 3T field strengths.
5. We analyze the generalization capabilities of our method with clinical cases of pediatric and adult patients.

In the present study we build upon our previous work (Pirkl et al., 2021). Based on the initial methodological concepts, we extended the recent paper by a more detailed overview of related works. We also added a comprehensive formalisation of the encoding and reconstruction problem and a detailed description of the proposed physics-informed motion simulation together with the underlying motion model. From a methodological perspective, we refined the training strategy, evaluated the performance of the proposed motion correction for different motion regimes in a dedicated sensitivity analysis and performed

an extensive tissue-specific evaluation for the real-motion *in vivo* data.

2. Related works

Recent advances of deep learning demonstrated improved robustness of MRI acquisitions towards subject motion, achieving similar performance as conventional methods or even outperforming the current state of the art in terms of image quality, reconstruction speed or computational efficiency.

In this section, we review these deep learning-based motion reduction and correction approaches for single-contrast qualitative MRI, and then give an overview of how motion and motion-induced artifacts are handled in quantitative MRI, particularly focusing on fast transient-state methods, such as MR Fingerprinting (Ma et al., 2013; Jiang et al., 2015).

2.1. Deep learning-based motion correction

Most of the recently proposed deep learning-based algorithms for reducing and correcting motion artifacts are retrospective methods and purely data-driven, i.e., without the need for any *a priori* information or additional (navigator) acquisitions.

In general, deep learning-based motion encoding and resolving methods almost exclusively take place in image space and mainly rely on deep CNN architectures trained in a supervised fashion (Lee et al., 2021). Depending on the specific scope, i.e., the region of interest and the respective type of motion, previously presented learning strategies either target motion artifacts caused by *non-rigid* patterns, e.g., due to involuntary respiratory, cardiac or intestinal motion (Lee et al., 2018; Tamada et al., 2019; Kromrey et al., 2020), or *rigid* patterns, e.g., due to head movements (Duffy et al., 2018; Johnson and Drangova, 2018; Pawar et al., 2018; Pawar et al., 2019; Johnson and Drangova, 2019; Dou et al., 2019; Bydder et al., 2019; Haskell et al., 2019; Sommer et al., 2020; Usman et al., 2020; Duffy et al., 2021).

Although being initially designed for different types of body motion, consequently resulting in different artifacts, these techniques can in general be differentiated into two main strategies: *Motion correction*, i.e., to infer an artifact-free MRI image from the motion-corrupted input, where all aforementioned works belong to, and *motion detection/parameter estimation*, e.g., to be used for subsequent realignment-based correction (Miao et al., 2016; Hou et al., 2018).

As stated before, the propagation of the movement that occurs during the *k*-space acquisition, to the final appearance of the artifact in image space is not only dependent on the type of motion, but also on the acquisition scheme, e.g., Cartesian vs. non-Cartesian readouts. The majority of deep learning-based motion correction approaches focused on Cartesian sampling patterns, which are predominant in conventional state-of-the-art MRI. There is only little work on non-Cartesian techniques so far (Dou et al., 2019). This is the case even though non-Cartesian gradient waveforms are superior to Cartesian readouts (not only) in terms of motion robustness. As they do not have unique frequency- and phase-encoding directions, artifacts manifest as diffuse image blurring. In contrast, subject motion translates as coherent ghosting or geometric distortions in case of Cartesian sampling, which is easier to correct for.

Another conceptually different approach to motion correction is to integrate motion encoding into the (deep learning) reconstruction model (Christodoulou et al., 2018; Schlemper et al., 2018; Huang et al., 2021). Modeling temporal correlations such as in dynamic cardiac MRI acquisitions allows to better capture spatially redundant information and to identify and correct for motion-induced artifacts.

2.2. Motion correction for fast multiparametric MRI

Quantitative MRI opens the possibility to obtain quantitative measures of the actual tissue parameter(s) by acquiring multiple contrast-weighted images within one scan. The signal evolution, however, that encodes the underlying (multiple) parameter information, is vulnerable to various sources of artifacts, including motion. This additionally complicates the translation of these techniques into routine clinical practice. To overcome this barrier, numerous strategies have been proposed to avoid or mitigate motion artifacts in quantitative MRI. Of these, accelerated acquisition schemes combined with advanced reconstruction techniques, e.g., compressed sensing, or motion prevention methods, such as breath-holding or triggering, have been shown to improve motion robustness of both the quantitative and qualitative MRI schemes. Also, retrospective correction approaches have been successfully transferred from conventional contrast-weighted scans to quantitative MRI techniques. More recently, deep learning strategies have entered the field. Gong et al. (2021) proposed a CNN-based correction for diffusion-weighted MRI acquisitions, while Tamada et al. (2019) and Kromrey et al. (2020) successfully improved image qualities of dynamic contrast-enhanced MRI exams by learning to resolve motion-induced artifacts in quantitative perfusion maps.

For both qualitative and quantitative MRI, fast acquisition and encoding schemes were shown to be powerful means to prevent motion artifacts. One advantage of highly accelerated parameter mapping techniques, e.g., based on transient-state readouts, is the higher motion robustness due to a shorter acquisition compared to lengthy conventional quantitative MRI methods. Despite the previously demonstrated high immunity to subject movements, motion artifacts are attenuated but not completely suppressed. Retrospective methods based on a realignment of the individual time frames are commonly used to correct for motion artifacts in MR Fingerprinting and other transient-state schemes, such as the QTI scheme of Gomez et al. (2020). The high temporal resolution that is achieved by these techniques is *a priori* beneficial for retrospective detection and correction of spatial mismatches between consecutive time frames. However, the massive *k*-space undersampling renders a precise realignment on these native time scales impossible. That is, to collect a sufficient amount of spatial information for image-based registration, the acquired image time-series are either reconstructed using sliding-window techniques or consecutive time-points are collapsed. Although these techniques have proved valuable for reducing motion-induced artifacts, the underlying motion patterns cannot be resolved in the high, natively acquired temporal resolution. This inevitably causes residual artifacts that cannot be fully resolved. This particularly complicates motion correction for 3D acquisition schemes. To achieve the same high temporal resolution as 2D readouts, 3D acquisition schemes (Ma et al., 2018; Gómez et al., 2020) are commonly characterized by segmented readouts with the 3D *kt*-space being filled iteratively, e.g., as illustrated in Fig. 1a for the example of 3D QTI. In contrast to 2D *kt*-sampling schemes, motion does not only cause spatial misalignment between subsequent time frames, but also results in discrepancies between iteratively sampled *k*-space locations of the same time frame. This is the main reason why previous work on motion correction for transient-state imaging has almost exclusively concentrated on 2D acquisition schemes (Mehta et al., 2018; Cruz et al., 2019; Xu et al., 2019).

3. QTI acquisition and reconstruction model

In this section, we aim to give a concise summary of the 3D QTI framework that this work builds upon. We start with a review of the acquisition and encoding scheme, with a particular
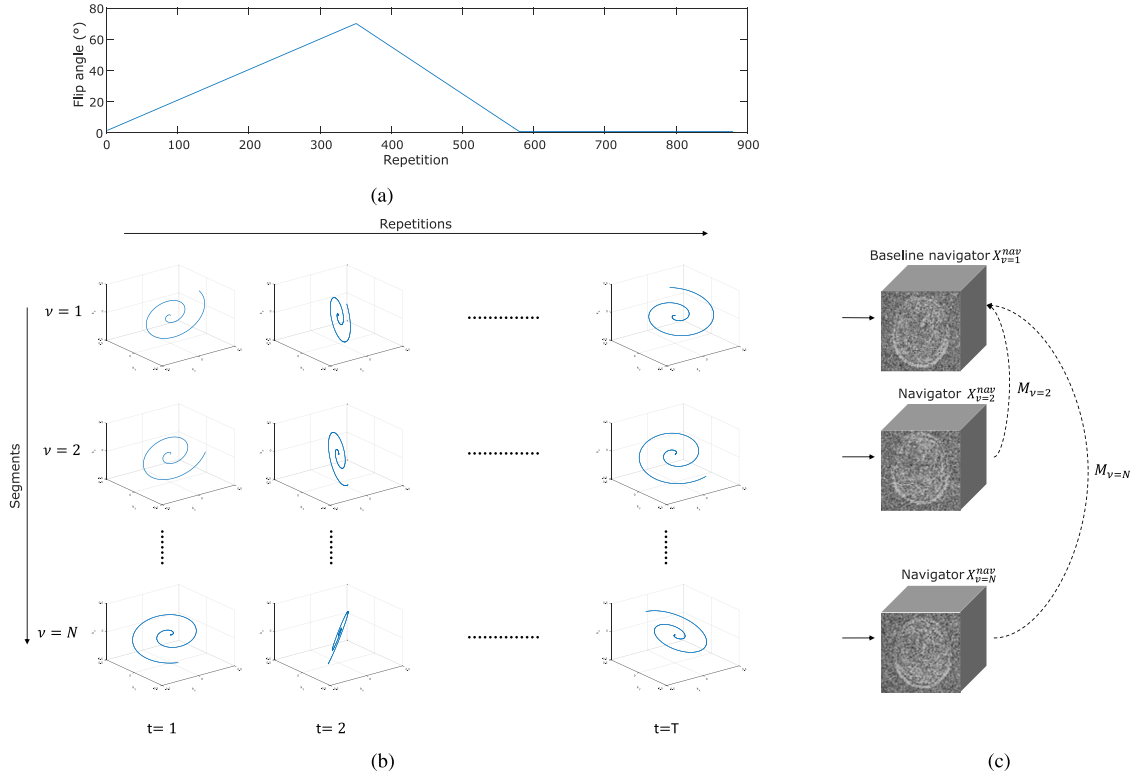
Fig. 1. Illustration of the 3D Quantitative Transient-state Imaging (QTI) scheme and the navigator-based retrospective rigid motion correction. (a) Excitation scheme that is repeated segment by segment. (b) Acquisition scheme based on segmented (multishot) readouts. (c) Realignment of the self-navigators $X^{nav}_{\nu>1}$ to the first baseline navigator $X^{nav}_{\nu=1}$ to estimate and subsequently correct for inter-segment motion M_ν.

focus on the segmented signal readout (Section 3.1). We then give a formal description of the underlying image and parameter reconstruction model (Section 3.2) and how rigid motion is integrated into the encoding operator (Section 3.3). Finally, we briefly recap the navigator-based motion correction as proposed by Kurzawski et al. (2020) (Section 3.4).

3.1. Acquisition and encoding scheme

The QTI acquisition scheme employed in this work follows a Steady-state Free Precession (SSFP) scheme to jointly encode T_1 and T_2 relaxation time information as the magnetization evolves in the transient state. An initial adiabatic inversion pulse precedes a train of RF excitations with flip angles following a ramp-up/ramp-down pattern as illustrated in Fig. 1a. The 3D kt-space is acquired using spiral projections combined with a multishot readout based on consecutive segments ν ($\nu = 1, \ldots, N = 56$) of the same RF excitation scheme. To iteratively fill the 3D kt-space, in-plane and spherical rotation angles, φ and θ, of the gradient waveform are randomly permuted from one repetition to the next (Fig. 1b).

3.2. Reconstruction model

Taking both the temporal signal evolution and the spatial Fourier relationship into account, the 3D kt-space data are then reconstructed by formulating the image reconstruction problem as a linear spatio-temporal model (Ma et al., 2013)

$$Y = \mathcal{E}(X) + \xi. \tag{1}$$

Here, $Y \in \mathbb{C}^{T \times m \times C}$ is the multi-coil raw kt-space acquired at time points $t = 1, \ldots, T$, m are the sampled k-space locations, ξ is a noise term and $X \in \mathbb{C}^{T \times n}$ is the image time-series to be reconstructed with n spatial voxels.

The forward encoding operator $\mathcal{E}$ with

$$\mathcal{E} := U^\Omega FS \tag{2}$$

is composed of the (under-) sampling operator U^Ω, with Ω describing the temporally varying kt-space sampling locations, the Fourier transform F and the multi-coil sensitivity profiles S (Golbabaee et al., 2021).

Further, a temporal projection $V \in \mathbb{C}^{T \times s}$ with $s << T$ is applied to transfer the high T-dimensional problem to a low-rank subspace, Range$(V) \subset \mathbb{C}^T$. With temporal subspace compression, the reconstruction model from Eq. (1) is given by

$$Y \approx \mathcal{E}(VX') + \xi, \tag{3}$$

where X' is the subspace-projected image time-series.

To retrieve the quantitative tissue properties $q \in \mathbb{R}^{Q \times n}$, including but not limited to T_1, T_2 and PD, that are encoded in the temporal transient-state signal responses, the parameter inference is formulated as a voxel-wise optimization problem

$$\hat{q}(r_i) = \operatorname{argmin}_q \left|\left| X'(r_i) - PD(r_i)V^H \mathcal{B}(q(r_i), \eta) \right|\right| \\ \forall i \in 1, \ldots, n, \tag{4}$$

where $\mathcal{B}(q, \eta)$ is the Bloch signal response as a function of the tissue parameters q and the time-dependent sequence parameters η, including flip angle FA, repetition time TR and echo time TE.

To solve the optimization problem in Eq. (4), it is approximated by a dictionary matching pursuit (Mallat and Zhang, 1993), i.e., by voxel-wise grid searching a precomputed dictionary $\mathcal{D}$ of synthetic signal responses to retrieve the most correlated dictionary entry $D_j := \mathcal{B}(q_j, \eta)$ for each reconstructed subspace signal evolution $X'(r_i) \; \forall i \in 1, \ldots, n$.
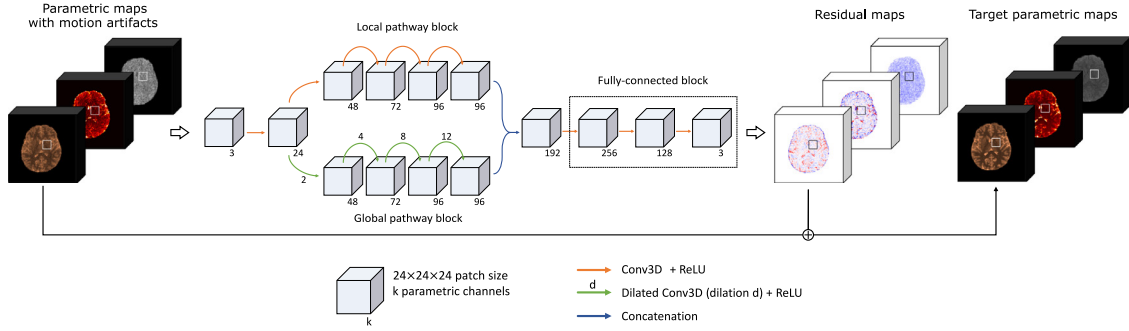
Fig. 2. Schematic diagram of the proposed residual learning model. The 3D patch-based multiscale CNN receives the motion-affected parametric maps as input and outputs the residual maps, i.e., the deviation from the high-quality maps. The learned residuum is then used to retrieve the final, motion-corrected parameter maps.

3.3. Rigid motion model

When integrating a rigid motion operator $M := TR$ with

$$T : \tilde{X}_t(r) \mapsto \tilde{X}_t(r \pm \delta r), \quad \text{for translations along } x, y, z,$$
$$R : \tilde{X}_t(r) \mapsto \tilde{X}_t(R(\delta\alpha)r), \quad \text{for rotations around } x, y, z, \tag{5}$$

into the image encoding formalism, Eq. (1) becomes

$$\tilde{Y} = \mathcal{E}(MX) + \xi, \tag{6}$$

where X is the motion-free image time-series. Spatial translations are reflected by $\delta r = (\delta x, \delta y, \delta z)^\top$ and the rotation matrix $R(\delta\alpha) = R_x(\delta\alpha_x)R_y(\delta\alpha_y)R_z(\delta\alpha_z)$ accounts for spatial rotations.

For the segmented, i.e., multishot, acquisition scheme (as illustrated in Fig. 1b), where the k-space of a time point t is filled iteratively, Eq. (1) can be written as

$$Y_t = \sum_{\nu=1}^N \mathcal{E}_{\nu,t}(X_{\nu,t}) + \xi_{\nu,t}$$
$$= \sum_{\nu=1}^N U_{\nu,t}^\Omega FS(X_{\nu,t}) + \xi_{\nu,t}, \tag{7}$$

where the subsequently acquired segments of the same time point t are indicated by ν.

In this case, Eq. (6) becomes

$$\tilde{Y}_t = \sum_{\nu=1}^N U_{\nu,t}^\Omega FS(M_{\nu,t}X_{\nu,t}) + \xi_{\nu,t}. \tag{8}$$

3.4. Navigator-based rigid motion correction

The navigator-based correction (Kurzawski et al., 2020) identifies motion-induced misalignment in the acquired image-time series $\tilde{X} = MX$. To do so, the full kt-space $\tilde{Y}$ is binned into subsequently acquired segments $\tilde{Y}_\nu$,

$$\tilde{Y}_\nu = \sum_{t=1}^T \mathcal{E}_{\nu,t}(\tilde{X}_{\nu,t}) + \xi_t$$
$$= \sum_{t=1}^T U_{\nu,t}^\Omega FS(\tilde{X}_{\nu,t}) + \xi_{\nu,t}, \tag{9}$$

from which full 3D equal-contrast navigator images $X_\nu^{nav} \in \mathbb{C}^{1 \times n}$ are reconstructed $\forall \nu = 1, \ldots, N$ as illustrated in Fig. 1c. To achieve sufficiently high spatial information despite the high degree of spatial undersampling, the first singular value coefficient of the subspace-projected image time-series $X_\nu^{nav} = V^H \tilde{X}_\nu$ with $V \in \mathbb{C}^{T \times s=1}$ is used for the self-navigators X_ν^{nav}. The self-navigators $X_{\nu>1}^{nav}$ are then aligned to the first baseline navigator $X_{\nu=1}^{nav}$ based on the normalized correlation coefficient (Penny et al., 2006) to estimate the effective inter-segment movement M_ν, i.e., δr_ν and $\delta\alpha_\nu$, and to subsequently realign the kt-space data $\tilde{Y}_\nu$ accordingly. The corrected raw kt-space series are then fed into the reconstruction pipeline as described in Section 3.2 to yield the motion-corrected parametric maps.

4. Residual learning for retrospective 3D motion correction

Inspired by the recently demonstrated efficacy of residual learning techniques (Zhang et al., 2017; Jin et al., 2017; Ulas et al., 2018), we adopt this approach to resolve motion artifacts in fast 3D multiparametric MRI. Considering the motion correction as an inference task between the motion-affected image input and the residual error, i.e., the deviation of the motion-corrupted image data from the high-quality, motion-free reference, has been shown to be more effective than a direct mapping (Tamada et al., 2019; Haskell et al., 2019; Kromrey et al., 2020; Sommer et al., 2020; Liu et al., 2020). In this work, we propose a residual CNN model to efficiently learn the non-linear relationship of a motion-corrupted multiparameter space $\tilde{q}$ and its high-quality, motion-free reference q. We demonstrate our deep learning method by taking the 3D QTI framework for joint T_1, T_2 and PD mapping as an exemplary acquisition and parameter encoding scheme. The residual learning strategy allows us to disentangle the propagation of motion (from the kt-domain to the parameter domain in image space), from the *a priori* high-dimensional problem of joint spatial image and multiparameter encoding (Eq. (6)). This way, we can efficiently identify and resolve the motion-induced corruptions in the parameter maps in a more sparse representation.

We show that the potential application scenario of the proposed residual learning strategy is twofold: First, our deep learning framework is able to resolve residual motion artifacts that could not be sufficiently corrected by the navigator-based method of Kurzawski et al. (2020). Second, the trained algorithm can also be applied as a stand-alone method to mitigate mild image corruptions without prior navigator-based correction.

In this section, we introduce the proposed deep learning strategy with its key components: the residual CNN model (Section 4.1, Section 4.2 and Section 4.3) and the physics-informed generation of artificial motion artifacts (Section 4.4).

4.1. Residual CNN model

We propose a patch-based 3D multiscale CNN architecture as illustrated in Fig. 2. The model receives a 3D patch of $24 \times 24 \times 24$ voxels of the quantitative maps degraded by motion artifacts as input. The T_1, T_2 and PD information is captured along the channel dimension with $Q = 3$ parametric channels. The model is built by a local and a global pathway to account for the inherent overlay of encoded spatial information on multiple scales. These dual pathway structures were recently shown to efficiently capture spatial relationships that naturally arise from the Fourier imaging principle (Kamnitsas et al., 2017; Kim et al., 2017; Ulas et al., 2018). 3D QTI relies on multi-coil readouts and the massive spiral undersampling causes spatial aliasing across the entire imaging volume. With the proposed multiscale implementation we capture the

full dimensionality of the encoded information - the encoded primary parameter and spatial information and also the interfering secondary motion effects. The local pathway is built by four 3D convolutional layers, each with a kernel size of $3 \times 3 \times 3$ and subsequent rectified linear unit (ReLU) non-linearity. This way, the local pathway extracts localized, spatially adjacent features, while in parallel the global pathway is constructed to account for longer-range contextual information. It comprises four ReLU-activated 3D dilated convolutional layers with dilation factors of $2, 4, 8$ and 12, and kernel sizes of $3 \times 3 \times 3$. To efficiently combine the decoded multiscale information and to eventually yield the targeted output, i.e., the residual error maps for each of the $Q = 3$ parameter maps, local and global features are concatenated along the channel dimension and fed into a block of fully-connected layers. To maintain the spatial dimension throughout the CNN, zero-padding is applied prior to each convolutional operation and the fully-connected layers are implemented such that they only act on the channel dimension.

4.2. CNN training

The residual CNN was trained based on *in vivo* 3D QTI data from ten healthy volunteers who underwent the acquisition twice. The first *no motion reference* dataset was acquired with the instruction to hold still, while for the second *motion* scan the volunteers were asked to rapidly move the head as detailed by Kurzawski et al. (2020). The *no motion reference* dataset was used to generate paired training sets by retrospectively applying rigid motion to the acquired *kt*-space data as described in Section 4.4. The *motion* dataset was used for testing only.

Following a supervised training strategy, the residual CNN model was supposed to learn the non-linear mapping from the parametric maps with simulated motion artifacts $\tilde{q}$ to the residual error $\epsilon = q - \tilde{q}$, with q being the motion-free counterpart. Following a leave-one-out strategy, 10,000 training patches were randomly sampled from the retrospectively corrupted datasets of eight subjects. For validation, 3,000 patches were sampled from another subject's data. To ensure sufficient amount of training/validation information, only patches with less than 70% background were considered. The one remaining subject's data was held back for testing. The residual CNN was trained with a learning rate of $1e^{-4}$ for a maximum of 100 epochs and a mini-batch size of 20, using Adam optimization to minimize the $\mathcal{L}_1$ loss function

$$\mathcal{L}_1 = \sum_J |\epsilon - \hat{\epsilon}|, \tag{10}$$

between the predicted output $\hat{\epsilon}$ and the corresponding reference of the residual error ϵ over all samples in the training set J. As final model state we kept the epoch with the best validation loss.

4.3. CNN testing and application

To feed motion-affected parametric maps into the trained residual CNN model, the full 3D field of view was divided into overlapping patches of $24 \times 24 \times 24$ voxels, sampled by a step size of 2 voxels along x, y, z. The predicted residual error patches were then added to the motion-affected input patches. All motion-corrected patches of a neighborhood were mapped back to the original field of view and averaged to yield the full 3D motion-corrected T_1, T_2 and *PD* maps.

4.4. Physics-informed simulation of motion

For supervised model training, we created self-contained pairs of motion-corrupted and motion-free data. To do so, the motion-free raw *kt*-space acquisitions were retrospectively distorted by applying artificial motion patterns. We then made implicit use of the
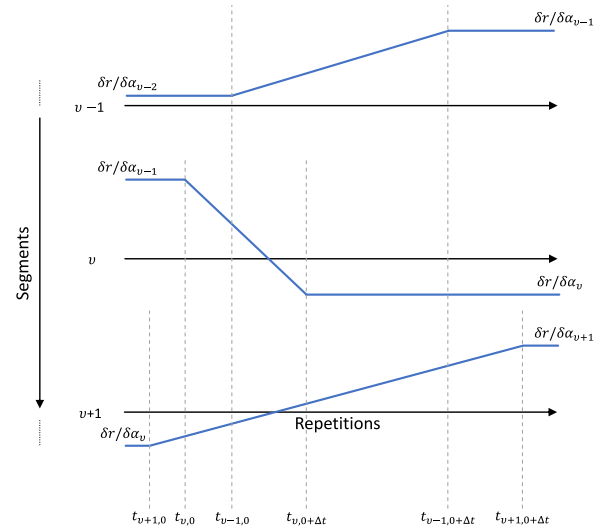

Fig. 3. Schematic diagram of the proposed simulation of continuous rigid motion patterns.
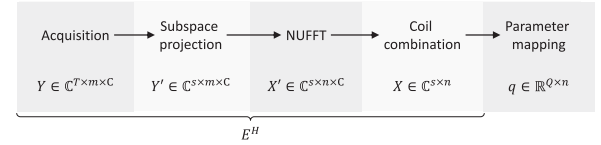

Fig. 4. Schematic diagram of the reconstruction pipeline.

forward encoding operator $\mathcal{E}^H$ to propagate the simulated movements from the *kt*-space through the reconstruction pipeline to the parameter maps.

The brain is less affected by non-rigid deformations than abdominal or thorax regions, where respiratory, cardiac or intestines motion is prevalent. Neglecting the effect of brain pulsation and blood flow, rigid head movements, i.e., global translations δr or rotations $\delta \alpha$, are the prevailing cause of motion artifacts in head and neck MRI. Therefore we applied continuous translations and rotations to the motion-free time frames $Y_{v,t}$ of the acquired *kt*-space. Translations were performed by global shifts of $Y_{v,t}$ along x, y, z, while rotations around x, y, z-axes were applied to the gradient waveform.

Aiming for a realistic artifact appearance, translation and rotation patterns were created to continuously vary the misalignment of consecutive *kt*-space time points, considering translations $\delta r = (\delta x, \delta y, \delta z)^\top$ and rotations $\delta \alpha = (\delta \alpha_x, \delta \alpha_y, \delta \alpha_z)^\top$ with a maximum displacement from the initial position of $-20mm \leq \Delta r \leq 20mm$ and $-20° \leq \Delta \alpha \leq 20°$, respectively (Kurzawski et al., 2020). To simulate continuous head movements with varying displacements and time scales throughout the acquisition, we randomly sampled effective displacements $\delta r_v, \delta \alpha_v$ for each segment $v > 1$. For each of these motion events, an onset $t_{v,0}$ and an offset $t_{v,0+\Delta t}$ time point were randomly sampled. Between $t_{v,0}$ and $t_{v,0+\Delta t}$, translation and rotation steps were continuously applied to the consecutive time-points $t_{v,0} \leq t \leq t_{v,0+\Delta t}$ to move from one displacement $(\delta r_{v-1}, \delta \alpha_{v-1})$ to the next $(\delta r_v, \delta \alpha_v)$ (Fig. 3).

For model training, we performed a navigator-based correction to mitigate artifacts due to inter-segment movements M_v in the first place as schematically shown in Fig. 5. The obtained parametric maps with residual artifacts due to intra-segment movements M_t were the input to the CNN.
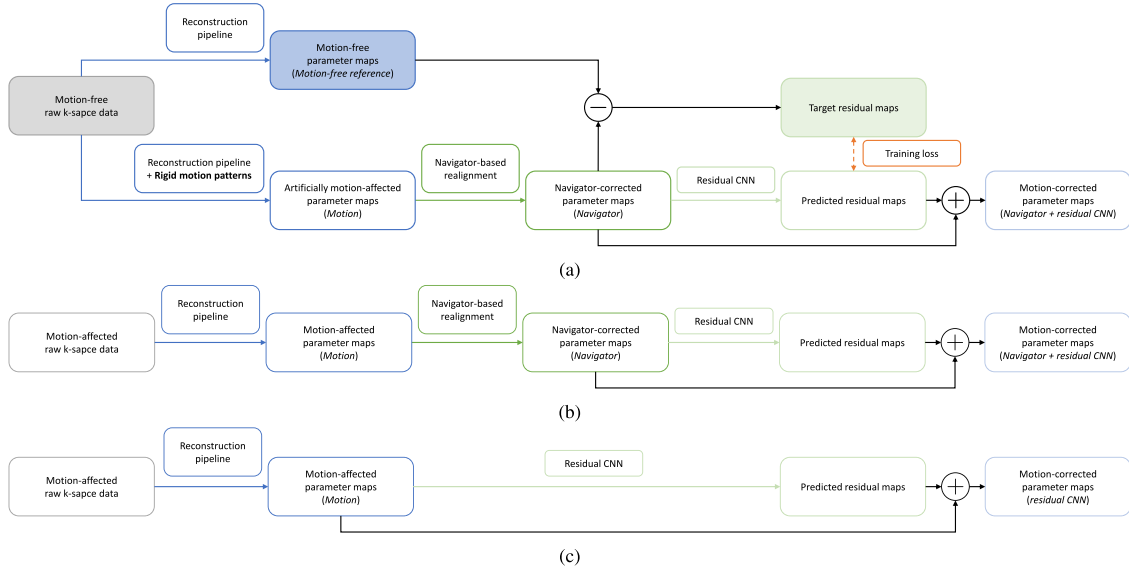
Fig. 5. Data flow diagrams. (a) Proposed simulation of motion-affected data + residual CNN training setup. (b+c) Data processing setup in case of residual CNN motion correction with and without prior navigator-based realignment, respectively.

5. Methodology and experimental setup

5.1. In vivo data

All *in vivo* data presented in this study were acquired in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Approval was granted by the local ethics boards.

5.1.1. Acquisition and reconstruction

Ten healthy volunteers, one pediatric and one adult patient were scanned with the 3D QTI acquisition scheme as described in Section 3.1. All healthy subjects were scanned on both a 1.5T HDxt and a 3T MR750 scanner (both GE Healthcare, Milwaukee, WI). The pediatric and the adult patients were examined at 1.5T and 3T, respectively. For both magnetic field strengths, the sequence parameters were identical, i.e., $TI = 18ms$, $TE = 0.5ms$, $TR = 8.5ms$ and $0.7° \leq FA \leq 70°$. Data were acquired with a field of view of $225 \times 225 \times 225$ mm^3 with $1.125 \times 1.125 \times 1.125$ mm^3 isotropic voxel size and $T = 880$ repetitions.

The formulation of the image reconstruction problem in Eq. (1) translates into the reconstruction pipeline as schematically shown in Fig. 4: The spatially undersampled 3D kt-space time-series are first zero-filled and then projected into a lower dimensional subspace in the time domain. In this work, we applied the singular value decomposition (SVD) variant for temporal subspace compression (McGivney et al., 2014). In this case, V projects the full temporal time-series to its first $s = 10$ singular values. Subspace-projection is followed by a Non-Uniform Fast Fourier Transform (NUFFT, F, Knoll et al., 2014) and subsequent adaptive coil sensitivity estimation and combination (S, Walsh et al., 2000). Parameter inference was performed using a pattern matching algorithm. Using the Extended Phase Graphs (EPG) formalism (Weigel et al., 2010), a dictionary $\mathcal{D}$ of synthetic signal evolutions $\mathcal{B}(q, \eta)$ was generated for $T_1 = [20ms, 5000ms]$ and $T_2 = [10ms, 500ms]$ with steps of 20ms for $T_1 < 3000ms$, 200ms for $T_1 \geq 3000ms$, 5ms for $T_2 < 300ms$ and 50ms for $T_2 \geq 300ms$, i.e., the same granularity and parameter ranges of the T_1 and T_2 spaces as in the work of Jiang et al. (2015).

5.1.2. Processing

For quantitative evaluation, we transformed the quantitative maps from the motion-affected and motion-free reference acquisition to a common space. For each subject, we registered the motion-free parameter space to its counterpart with motion artifacts. To ensure sufficient image quality, FSL's FLIRT-based rigid alignment (with 6 degrees of freedom) was performed based on the brain-extracted T_1 maps of the motion-free and motion-affected dataset, after initial navigator-based motion correction for the latter one (Jenkinson and Smith, 2001; Jenkinson et al., 2002). Based on the motion-free T_1 and T_2 maps, we obtained white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF) tissue segmentation maps using FAST (Zhang et al., 2001).

5.2. Quantitative performance evaluation

Quantitative analysis was based on the following voxel-wise performance measures:

– Mean absolute percentage error (MAPE)

$$\text{MAPE}(q_{pred}, q_{ref}) = \frac{1}{I} \sum_{i=0}^{I-1} \frac{|q_{pred,i} - q_{ref,i}|}{\max(\lambda, q_{ref,i})}, \quad (11)$$

where $\lambda = 1e^{-5}$ to avoid undefined results for $q_{ref,i} = 0$.

– Root mean square error (RMSE)

$$\text{RMSE}(q_{pred}, q_{ref}) = \sqrt{\frac{1}{I} \sum_{i=0}^{I-1} (q_{pred,i} - q_{ref,i})^2}, \quad (12)$$

– Peak signal-to-noise ratio (PSNR)

$$\text{PSNR}(q_{pred}, q_{ref}) = 20 \log_{10} \left(\frac{q_{max}}{\sqrt{\text{MSE}(q_{pred}, q_{ref})}} \right), \quad (13)$$

where $\text{MSE} = \frac{1}{I} \sum_{i=0}^{I-1} (q_{pred,i} - q_{ref,i})^2$ is the mean-squared error (MSE). For $q = T_1, T_2, PD$, q_{max} was set as the maximum possible parameter value in the dictionary $\mathcal{D}$, i.e., $q_{max} = 5000ms, 500ms, 1$.

For all the above measures, the motion-free parameter maps were used as the reference q_{ref}. As an additional measure, we calculated the signal to noise ratio (SNR)

$$\text{SNR}(q) = \frac{\bar{q}}{\sigma}. \quad (14)$$

6. Experiments and results
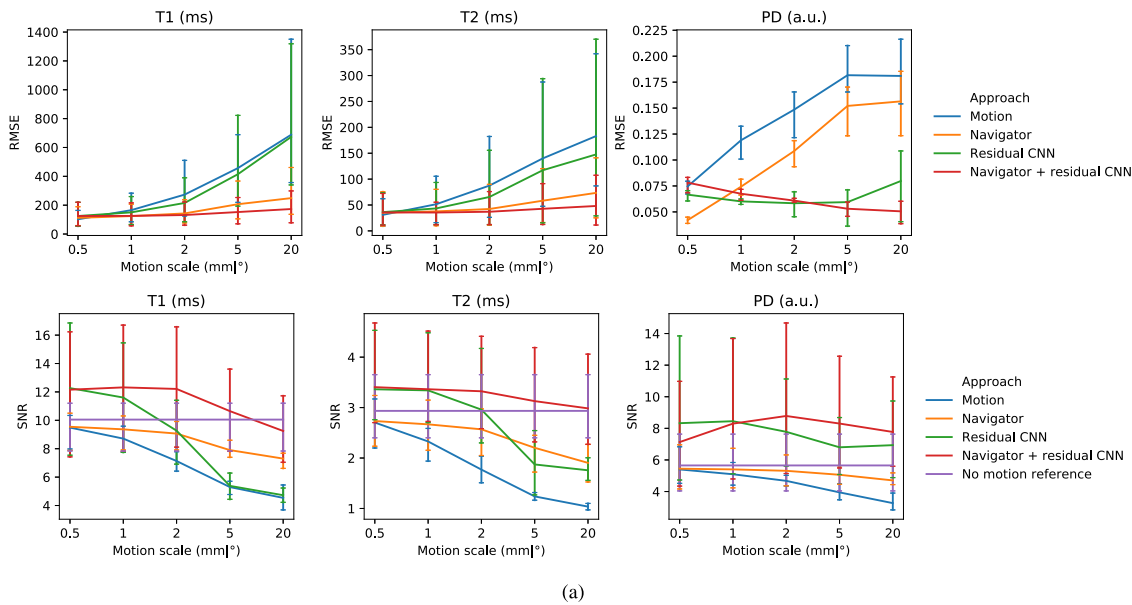
6.1. Sensitivity analysis

To evaluate the efficacy of the proposed residual learning method and to understand how it depends on the respective motion scale, we performed a sensitivity analysis as follows. We applied the trained residual CNN model to the held-back test set of the 3T acquisition. To stress-test how the intensity of the movement influences model performance, we retrospectively applied the same randomly sampled pattern of translations and rotations to the initially motion-free kt-space data with varied amplitudes of the motion to simulate scenarios from mild ($\Delta r = \pm 0.5mm$, $\Delta \alpha = \pm 0.5°$) up to pronounced movements ($\Delta r = \pm 20mm$, $\Delta \alpha = \pm 20°$). We then investigated robustness of the proposed residual CNN model (*Residual CNN*), the navigator-based correction (*Navigator*) and when applied in combination (*Navigator + residual CNN*).

The qualitative assessment (Fig. 6) shows that the navigator-based approach by design identifies and corrects for movements on a coarser time scale. That is, very small motion scales, e.g., $\Delta r = \pm 0.5mm$, $\Delta \alpha = \pm 0.5°$ (Fig. 6a) could not be resolved due to the limited temporal resolution ($\backsim 7s = T \cdot TR$) of the navigator-images X_ν^{nav}. Theoretically, the navigator-based approach could use temporally higher resolved self-navigators. However, this would result in lower SNR of the navigator-images and would be insufficient for reliable realignment. For milder motion scales, intra-segment movements and the thereby induced phase inconsistencies are the primary cause of image artifacts. In these domains the residual CNN outperformed the navigator correction. It effectively resolved image artifacts and achieved high image quality of the quantitative maps when performed with and without initial navigator-based correction. The purpose of the residual CNN was to correct for intra-segment motion that cannot be resolved by the navigator approach. As the corrective effect of the preceding navigator-based approach is minor, actual amplitudes and time scales of the uncorrected and navigator-corrected data are similar. That is, we observe that the residual CNN applied as a stand-alone tool was as effective as in combination with the navigator approach. This is also seen from the quantitative analysis (Fig. 7): For all quantitative maps, i.e., T_1, T_2 and PD, residual CNN correction with and without prior navigator-correction achieved higher SNR than the navigator-correction, even exceeding the high quality of the reference maps. We attribute this to the fact that although being acquired with the volunteer holding as still as possible, very minor movements are not entirely avoidable, causing a slight degradation of the image quality and hence lower SNR as the corrected maps. This fits the observation that RMSE between corrected and reference maps for PD indicates better agreement with the reference for the navigator-corrected maps than for the residual CNN-corrected counterparts. While the residual CNN reliably corrected motion artifacts for $\Delta r = \pm 1mm$, $\Delta \alpha = \pm 1°$, achieving high image qualities of the parameter maps (Fig. 6b), the impact of larger-scale inter-segment motion M_ν becomes the more and more dominant cause of image artifacts for increasing motion amplitudes. As visually observed from Fig. 6d to e, image artifacts that arise from movements with $\Delta r \geq \pm 5mm$, $\Delta \alpha \geq \pm 5°$ could not be resolved by the residual CNN when applied directly to the motion-corrupted maps. The navigator-based correction clearly improved image qualities of the motion-affected parameter maps in these motion regimes, where inter-segment movements becomes the prevailing effect. However, it could not achieve high-enough image qualities to suffice for clinical diagnosis. Only the combination of the navigator-based realignment and residual CNN processing demonstrated high motion-correction performance throughout the entire range of examined motion regimes, achieving clearly improved image qualities of the quantitative maps of T_1, T_2 and PD
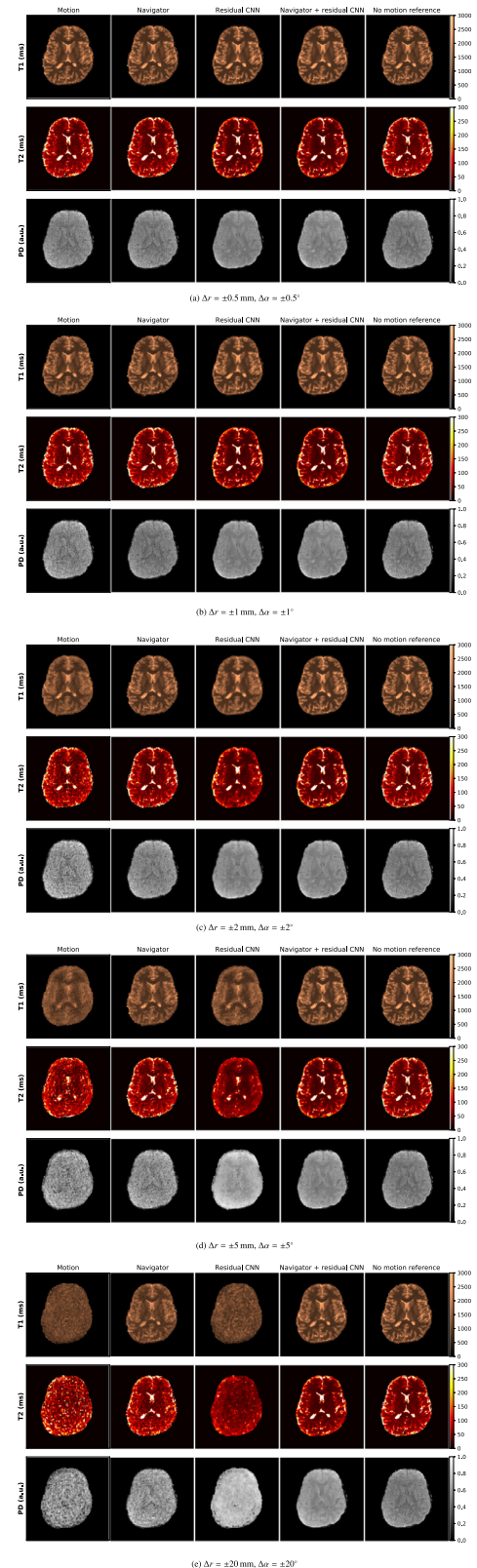
Fig. 6. Sensitivity analysis - qualitative results. Axial views of a representative volunteer dataset acquired at 3T with retrospectively applied rigid motion patterns in different regimes, from mild ($\Delta r = \pm 0.5mm$, $\Delta \alpha = \pm 0.5°$) to pronounced ($\Delta r = \pm 20mm$, $\Delta \alpha = \pm 20°$) movements. From left to right, quantitative maps are shown for the case without any correction (*Motion*), after navigator-based correction (*Navigator*), after residual CNN correction (*residual CNN*), after residual CNN correction with prior navigator-based correction (*Navigator + residual CNN*) and as obtained from the motion-free reference scan (*No motion reference*).

Fig. 7. Sensitivity analysis - quantitative results. Quantitative comparison with respect to the motion-free reference based on RMSE and SNR measures. The vertical bars reflect the individual tissue types, i.e., WM, GM and CSF.

even in case of very pronounced head movements $\Delta r = \pm 20mm$, $\Delta \alpha = \pm 20°$, which would be clinically useless without the correction (Fig. 6e).

As such, the combination of the residual CNN with the navigator-based correction effectively mitigated rigid motion patterns on two scales:

1. The navigator-based motion estimation and its subsequent correction in the *kt*-space reliably corrected for inter-segment movements M_ν and the thereby induced image artifacts.
2. The residual CNN identified and resolved motion artifacts and phase inconsistencies due to continuous intra-segment movements M_t, including residual artifacts that are unresolved by the limited temporal resolution of the self-navigators.

Importantly, the proposed two-stage motion correction preserves brain tissue boundaries, e.g., in adjacent healthy WM, GM and CSF structures, for mild to intermediate motion scales ($\Delta r < \pm 5mm$, $\Delta \alpha < \pm 5°$, Fig. 6a–c). Brain tissue structure is not smoothed out but is well recovered and clearly visible in the quantitative maps after motion correction. Also in pronounced motion cases, e.g., $\Delta r = \pm 20mm$, $\Delta \alpha = \pm 20°$ (Fig. 6e), the proposed correction recovers a major portion of the underlying tissue structure, previously corrupted by massive artifacts. Despite the clear improvement in image quality, the high image quality as obtained for small to intermediate motion scales is not reached in these excessive motion regimes. We would like to emphasize that such extreme scenarios, however, exceed the range of head movements that are observed in day-to-day clinical routine (cf. Fig. 10).

6.2. Performance analysis on real-motion in vivo data

The proposed residual CNN was solely trained on artifacts obtained from artificial motion patterns. To show that this training strategy allows for reliable translation to real-motion scenarios, we tested the motion-correction performance based on the motion-affected volunteer datasets for both the 1.5T and 3T acquisitions. To to so, we performed a ten-fold cross-validation experiment by repeating the training setup from Section 4.2. Each of the resulting ten model instances was then tested on the held-back volunteer dataset with real motion. Before being fed into the model, we first corrected for inter-segment movements using navigator-based realignment scheme. The cross-validation was performed independently for the 1.5T and the 3T dataset. In both cases, the corrected maps of T_1, T_2 and *PD* were then compared to the co-registered motion-free reference.

The proposed motion-correction strategy, i.e., the combination of initial navigator-based correction followed by the residual CNN provided T_1, T_2 and *PD* maps with high image qualities and good agreement with the motion-free reference maps. This is qualitatively shown for two representative test cases of the cross-validation experiments for the 1.5T (Fig. 8a) and the 3T acquisition (Fig. 8b), respectively. This is the case even though the residual CNN model was trained on purely artificially corrupted data. The residual CNN (*Navigator + residual CNN*) demonstrated a clear improvement compared to the results obtained without the deep learning-based correction (*Navigator*), providing parameter maps with image qualities that meet or even outperform the motion-free reference in terms of SNR (Fig. 9a, b bottom). Quantitative evaluation of the cross-validation experiment, i.e., by comparing the motion-affected (*Motion*) and motion-corrected maps (*Navigator, Navigator + residual CNN*) to the reference maps (*No motion reference*) by means of RMSE, MAPE and PSNR substantiated the previous qualitative findings. Overall, quantitative measures (RMSE↓, MAPE↓, PSNR↑) indicate that the residual CNN effectively resolved remaining motion artifacts, pushing the outcome of the navigator-based correction towards higher image qualities for all parameters (T_1, T_2, PD) and all tissue classes (WM, GM, CSF), with biggest improvements compared to the *Navigator* approach in case of T_2 parameters in WM and GM. Despite the same general tendency for both field strengths, quantitative analysis suggests better correspondence between *Navigator + residual CNN* and *No motion reference* for the 3T than for the 1.5T dataset. We see this observation arising from a mixture of different effects that overlay each other. First, the cross-validation experiments rely on real-motion scans. That is, although the same group of healthy volunteers was scanned at 1.5T and 3T, the actual head movements in terms of intensity and timing are not identical. Second, the 3T data has intrinsically higher SNR than the 1.5T data (Fig. 9a, b bottom). This results in higher image qualities of the self-navigators X_ν and even-

(a)

(b)
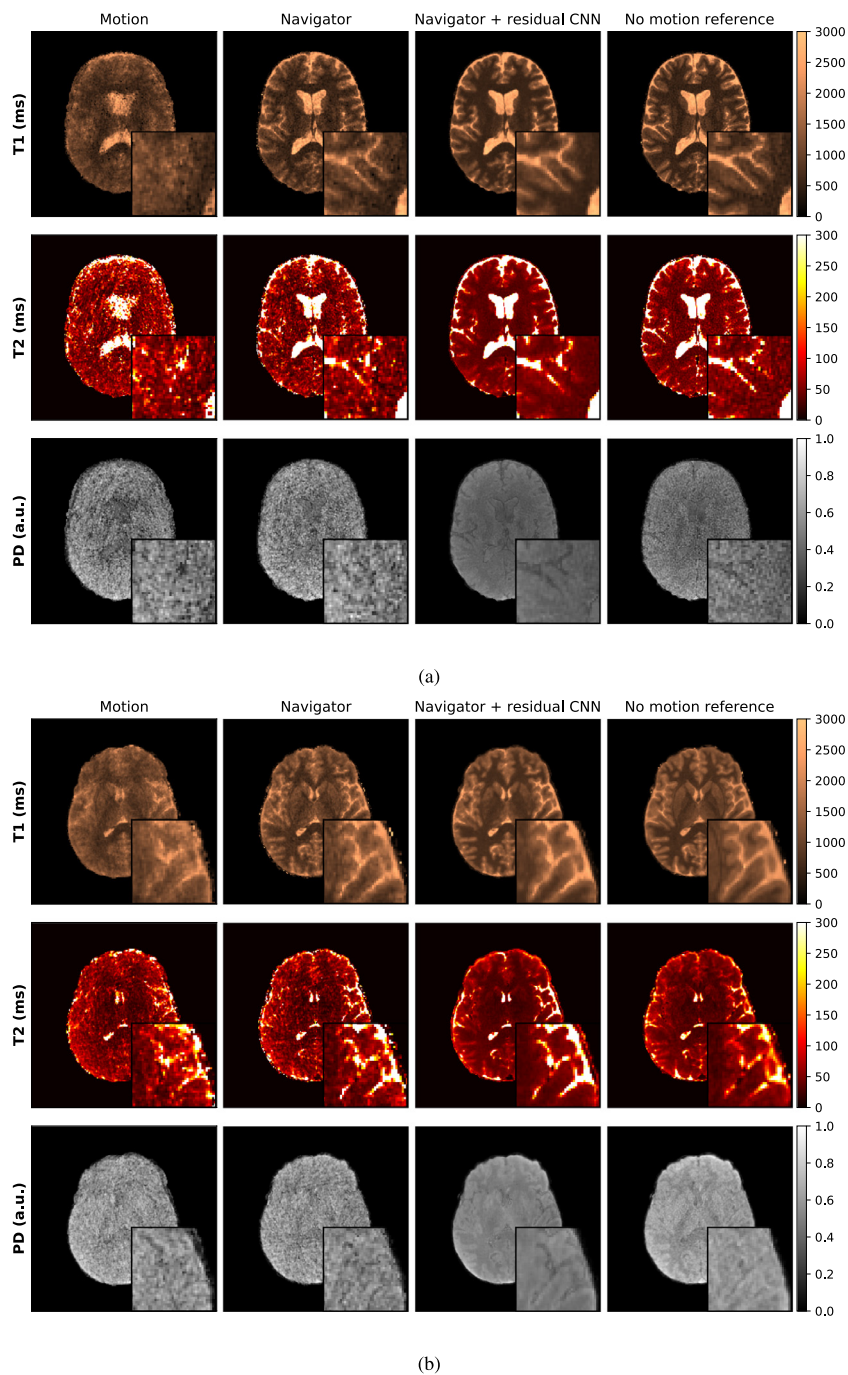
Fig. 8. Axial views of representative real-motion volunteer test dataset acquired at (a) 1.5T and (b) 3T. T_1, T_2 and *PD* maps show pronounced motion-induced artifacts (*Motion*) compared to the *No motion reference*. Remaining artifacts after *Navigator*-based correction are resolved by the residual CNN (*Navigator + residual CNN*), providing high-quality, artifact-free parameter maps.
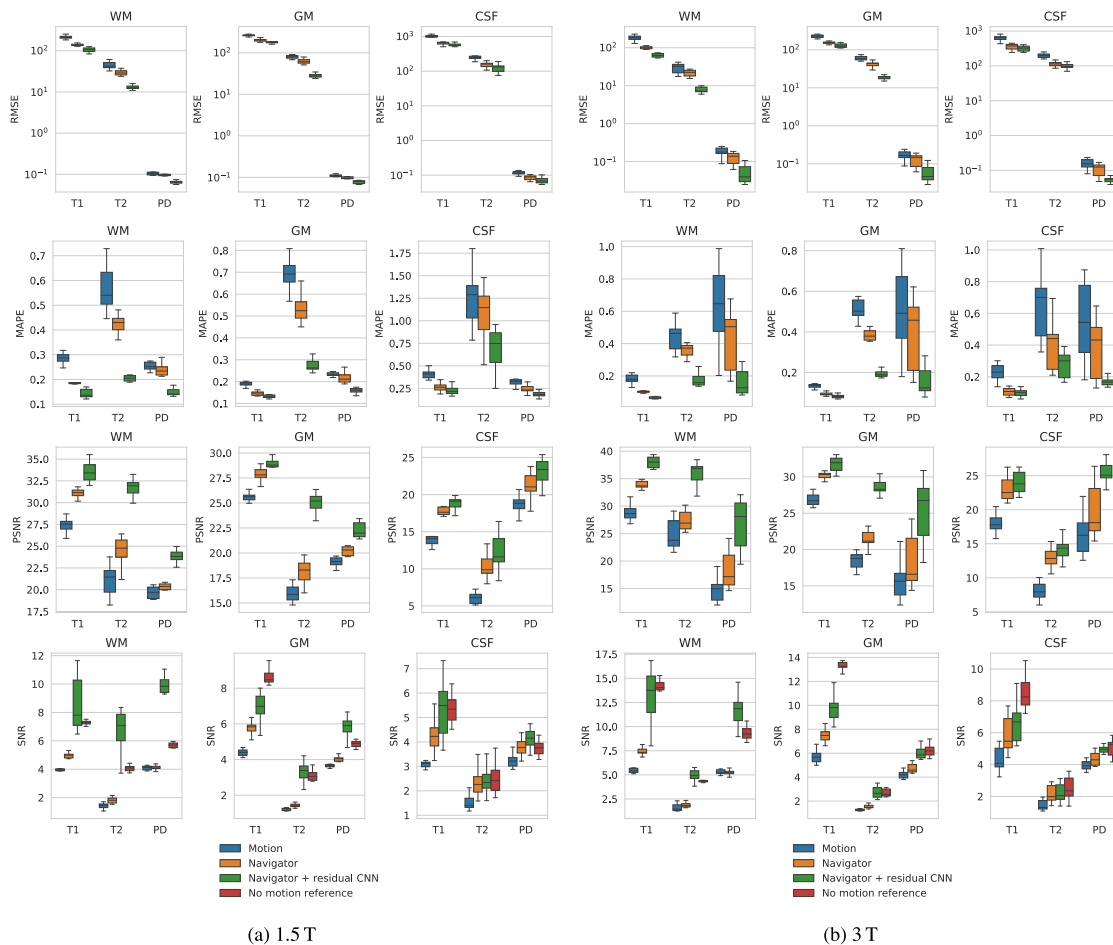
Fig. 9. Quantitative, tissue-specific analysis of the cross-validation experiment based on RMSE, MAPE, PSNR and SNR as performance measures. The boxes extend from the first to the third quartile, with the line indicating the median. The whiskers reflect the 1.5-times interquartile range. Outliers in the boxplots were omitted for clarity.

tually leads to a more accurate realignment of the navigators. The better agreement observed for 3T scans is also in line with the previous hypothesis that high-quality scans might still be affected by minor, noise-like motion and aliasing artifacts and are therefore rather considered a reference than a ground truth. Third, as observed in Section 6.1, the residual CNN resolved motion-induced artifacts even to an extent beyond the image qualities of the reference maps and additionally suppresses noise-like aliasing. The latter effect is more pronounced for the 1.5T data with intrinsically lower SNR, also explaining the generally lower overall agreement with the reference.

Overall, the observed generalizability to real-motion scenarios proved the efficacy of the proposed supervised training scheme based on the physics-informed generation of motion and motion-free data pairs.

6.3. Generalization analysis to clinical pediatric and oncological data

To demonstrate and validate the generalizability of our approach beyond the healthy population, we applied the trained model to the clinical 3D QTI scans of one pediatric (8-year old) patient with subtotal agenesis of the corpus callosum, and one adult patient with glioblastoma.

For the clinical data, representative axial views are shown in Fig. 10a and b. The proposed two-stage motion-correction strategy (*Navigator + residual CNN*) yielded high-quality, artifact-free para-

metric maps. The improvement in image quality was substantiated by initial radiological evaluation. In both cases, motion artifacts in the uncorrected parameter maps (*Motion*) were less pronounced than for the volunteer data (Fig. 8). Based on the findings of the sensitivity analysis in Section 6.1, we attribute this to the fact that the head movements of the patients were rather mild compared to the very pronounced (and advised) movements of the volunteers. This hypothesis is in line with the negligible, barely visible effect of the navigator-based correction (*Navigator*). Therefore, to enhance the quality of the maps, we also applied the residual CNN correction without prior navigator-based realignment (*Residual CNN*) and achieved similarly high image qualities in the parametric maps as suggested by the sensitivity analysis before. These findings also substantiate the generalization capabilities of the residual CNN beyond the healthy volunteer data. In the proposed deep learning strategy with its key features - the patch-based data perception and the sparse representation in the residual domain, we decoupled the primary parameter information and the anatomical context from the secondary artifact component. These initial results in the evaluated patient data give confidence that the proposed deep learning model is agnostic to the respective brain anatomy in terms of pathologies (healthy vs. diseased) and development (adult vs. pediatric). To gain further insights into model performance and robustness, we are in the process of rolling out this initial proof-of-concept study to an extended clinical evaluation covering diverse disease patterns and pathologies.

(a) Pediatric patient (1.5 T)
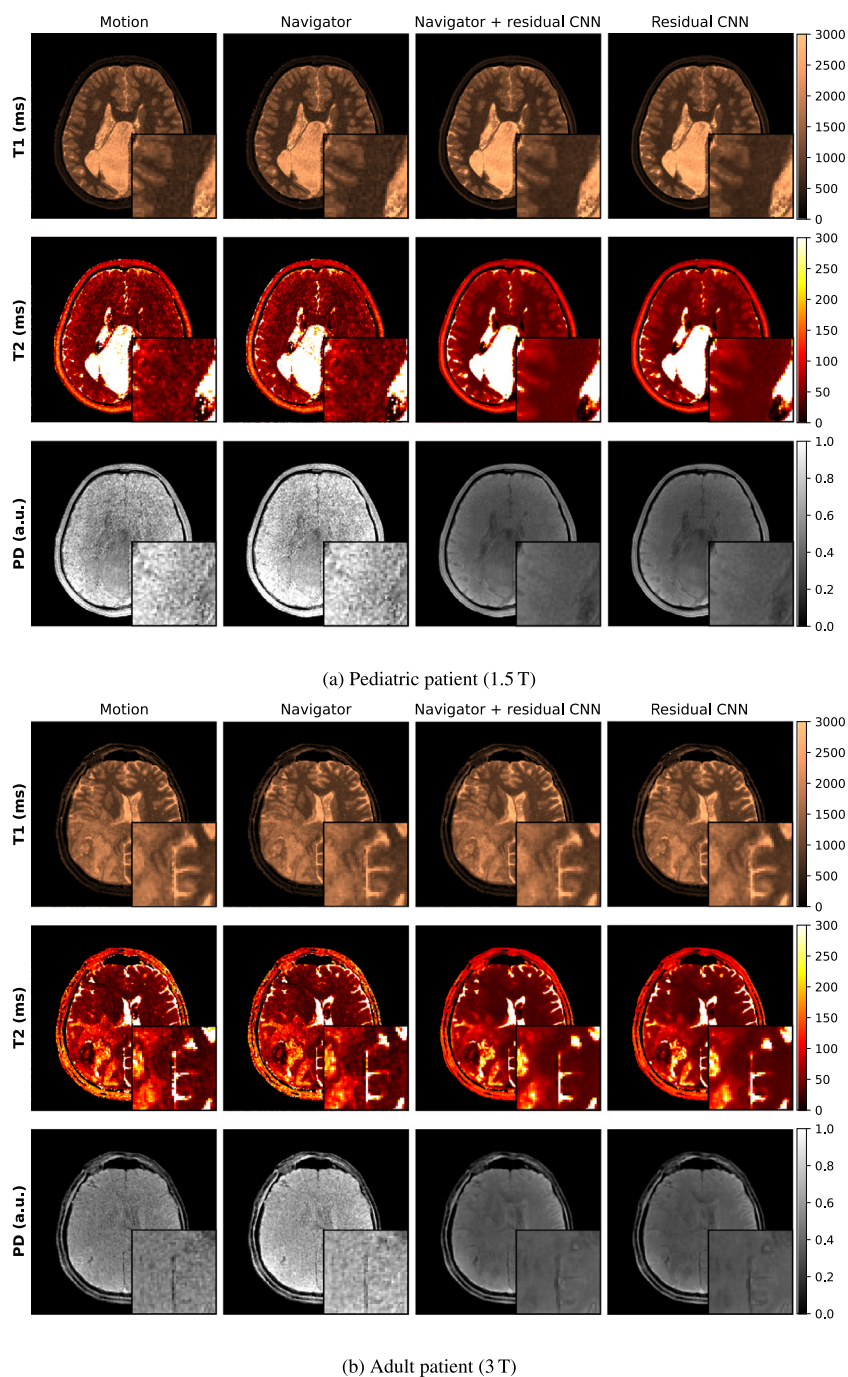
(b) Adult patient (3 T)

Fig. 10. Axial views of two representative clinical patient cases. (a) Pediatric patient with subtotal agenesis of the corpus callosum and inter-hemispheric cyst, scanned at 1.5T (b) Adult patient with glioblastoma in the temporo-parietal region with cystic-necrotic and hemorrhagic components, and marked perilesional edema, scanned at 3T. For both patients, the residual CNN improved image quality of all parametric maps (*Navigator + residual CNN*), mitigating image artifacts attributed to head movements during scan sessions.

7. Conclusion

In this work, we presented a deep learning model for retrospective motion correction in fast 3D whole-brain multiparametric MRI. With the proposed physics-informed motion simulation, we generated self-contained data pairs for supervised model training, circumventing the need for large amounts of paired acquisitions. The 3D multiscale CNN architecture captures the inherent 3D nature of the motion artifacts. With a residual learning strategy, we took advantage of the sparsity in the residual maps, effectively disen-

tangling primary anatomical and parameter information from the secondary motion artifacts. Combining the proposed residual CNN with a navigator-based algorithm demonstrated high efficacy, providing high quality maps of T_1, T_2 and *PD* in scenarios from mild to pronounced motion artifacts. We substantially improved the quality of motion-affected quantitative maps in case of healthy volunteers but also for pediatric and adult patients with pathological findings and different brain development stages. This is particularly important in clinical setups where motion corrupted scans frequently have to be repeated, possibly under sedation, because of

motion artifacts. With short scanning time and the higher motion-immunity, quantitative MRI may become a standard for clinical practices.

Declaration of Competing Interest

Carolin M. Pirkl and Marion I. Menzel are employees at GE Healthcare, Munich, Germany. Matteo Cencini and Michela Tosetti receive research funding from GE Healthcare. All other authors declare that they do not have any financial or non-financial conflict of interests.

CRediT authorship contribution statement

Carolin M. Pirkl: Conceptualization, Methodology, Software, Data curation, Validation, Writing – original draft, Project administration. **Matteo Cencini:** Conceptualization, Methodology, Software, Data curation, Validation, Writing – review & editing. **Jan W. Kurzawski:** Conceptualization, Methodology, Software, Data curation, Writing – review & editing. **Diana Waldmannstetter:** Methodology, Software, Writing – review & editing. **Hongwei Li:** Methodology, Software, Writing – review & editing. **Anjany Sekuboyina:** Methodology, Software, Writing – review & editing. **Sebastian Endt:** Methodology, Software, Writing – review & editing. **Luca Peretti:** Methodology, Software. **Graziella Donatelli:** Data curation, Validation, Writing – review & editing. **Rosa Pasquariello:** Validation. **Mauro Costagli:** Conceptualization, Data curation, Validation, Supervision, Funding acquisition, Writing – review & editing. **Guido Buonincontri:** Conceptualization, Data curation, Validation, Supervision, Funding acquisition. **Michela Tosetti:** Conceptualization, Supervision, Funding acquisition. **Marion I. Menzel:** Conceptualization, Supervision, Funding acquisition. **Bjoern H. Menze:** Conceptualization, Supervision, Funding acquisition.

Acknowledgements

References

Buonincontri, G., Kurzawski, J.W., Kaggie, J.D., Matys, T., Gallagher, F.A., Cencini, M., Donatelli, G., Cecchi, P., Cosottini, M., Martini, N., Frijia, F., Montanaro, D., Gomez, P.A., Schulte, R.F., Retico, A., Tosetti, M., 2021. Three dimensional MRF obtains highly repeatable and reproducible multi-parametric estimations in the healthy human brain at 1.5T and 3T. Neuroimage 226, 117573. doi:10.1016/j.neuroimage.2020.117573.

Bydder, M., Ghodrati, V.K., Ali, F.A., Hu, P., 2019. Deep CNN for outlier detection: a complementary tool to low-rank based methods for reducing motion artefacts. Montral, QC, Canada.

Christodoulou, A.G., Shaw, J.L., Nguyen, C., Yang, Q., Xie, Y., Wang, N., Li, D., 2018. Magnetic resonance multitasking for motion-resolved quantitative cardiovascular imaging. Nat. Biomed. Eng. 2 (4), 215–226. doi:10.1038/s41551-018-0217-y.

Cruz, G., Jaubert, O., Schneider, T., Botnar, R.M., Prieto, C., 2019. Rigid motion-corrected magnetic resonance fingerprinting. Magn. Reson. Med. 81 (2), 947–961. doi:10.1002/mrm.27448.

Dou, Q., Feng, X., Wang, Z., Weller, D., Meyer, C., 2019. Deep learning motion compensation for Cartesian and spiral trajectories. Montral, QC, Canada.

Duffy, B.A., Zhang, W., Tang, H., Zhao, L., Law, M., Toga, A.W., Kim, H., 2018. Retrospective correction of motion artifact affected structural MRI images using deep learning of simulated motion. https://openreview.net/forum?id=H1hWfZnjM.

Duffy, B.A., Zhao, L., Sepehrband, F., Min, J., Wang, D.J., Shi, Y., Toga, A.W., Kim, H., 2021. Retrospective motion artifact correction of structural MRI images using deep learning improves the quality of cortical surface reconstructions. Neuroimage 230, 117756. doi:10.1016/j.neuroimage.2021.117756.

Godenschweger, F., Kgebein, U., Stucht, D., Yarach, U., Sciarra, A., Yakupov, R., Lõsebrink, F., Schulze, P., Speck, O., 2016. Motion correction in MRI of the brain. Phys. Med. Biol. 61 (5), R32–56. doi:10.1088/0031-9155/61/5/R32.

Golbabaee, M., Buonincontri, G., Pirkl, C.M., Menzel, M.I., Menze, B.H., Davies, M., Gomez, P.A., 2021. Compressive MRI quantification using convex spatiotemporal priors and deep encoder-decoder networks. Med. Image Anal. 69, 101945. doi:10.1016/j.media.2020.101945.

Gómez, P.A., Cencini, M., Golbabaee, M., Schulte, R.F., Pirkl, C., Horvath, I., Fallo, G., Peretti, L., Tosetti, M., Menze, B.H., Buonincontri, G., 2020. Rapid three-dimensional multiparametric MRI with quantitative transient-state imaging. Sci. Rep. 10 (1), 13769. doi:10.1038/s41598-020-70789-2.

Gong, T., Tong, Q., Li, Z., He, H., Zhang, H., Zhong, J., 2021. Deep learning-based method for reducing residual motion effects in diffusion parameter estimation. Magn. Reson. Med. 85 (4), 2278–2293. doi:10.1002/mrm.28544.

Gomez, P.A., Molina-Romero, M., Buonincontri, G., Menzel, M.I., Menze, B.H., 2019. Designing contrasts for rapid, simultaneous parameter quantification and flow visualization with quantitative transient-state imaging. Sci. Rep. 9 (1), 8468. doi:10.1038/s41598-019-44832-w. Number: 1 Publisher: Nature Publishing Group.

Haskell, M.W., Cauley, S.F., Bilgic, B., Hossbach, J., Splitthoff, D.N., Pfeuffer, J., Setsompop, K., Wald, L.L., 2019. Network accelerated motion estimation and reduction (NAMER): convolutional neural network guided retrospective motion correction using a separable motion model. Magn. Reson. Med. 82 (4), 1452–1461. doi:10.1002/mrm.27771.

Hou, B., Miolane, N., Khanal, B., Lee, M.C.H., Alansary, A., McDonagh, S., Hajnal, J.V., Rueckert, D., Glocker, B., Kainz, B., 2018. Computing CNN loss and gradients for pose estimation with Riemannian geometry. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-Lpez, C., Fichtinger, G. (Eds.), Medical Image Computing and Computer Assisted Intervention - MICCAI 2018. Springer International Publishing, Cham, pp. 756–764. doi:10.1007/978-3-030-00928-1_85.

Huang, Q., Xian, Y., Yang, D., Qu, H., Yi, J., Wu, P., Metaxas, D.N., 2021. Dynamic MRI reconstruction with end-to-end motion-guided network. Med. Image Anal. 68, 101901. doi:10.1016/j.media.2020.101901.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17 (2), 825–841.

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5 (2), 143–156.

Jiang, Y., Ma, D., Keenan, K.E., Stupic, K.F., Gulani, V., Griswold, M.A., 2017. Repeatability of magnetic resonance fingerprinting T1 and T2 estimates assessed using the ISMRM/NIST MRI system phantom. Magn. Reson. Med. 78 (4), 1452–1457. doi:10.1002/mrm.26509.

Jiang, Y., Ma, D., Seiberlich, N., Gulani, V., Griswold, M.A., 2015. MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout. Magn. Reson. Med. 74 (6), 1621–1631. doi:10.1002/mrm.25559.

Jin, K.H., McCann, M.T., Froustey, E., Unser, M., 2017. Deep convolutional neural network for inverse problems in imaging. IEEE Trans. Image Process. 26 (9), 4509–4522. doi:10.1109/TIP.2017.2713099. Conference Name: IEEE Transactions on Image Processing.

Johnson, P.M., Drangova, M., 2018. Motion correction in MRI using deep learning. In: Proceedings of the 27th Annual Meeting of International Society for Magnetic Resonance in Medicine (ISMRM), Paris, France. Paris, France.

Johnson, P.M., Drangova, M., 2019. Conditional generative adversarial network for 3D rigid-body motion correction in MRI. Magn. Reson. Med. 82 (3), 901–910. doi:10.1002/mrm.27772.

Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B., 2017. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. 36, 61–78. doi:10.1016/j.media.2016.10.004.

Kim, K.H., Choi, S.H., Park, S.-H., 2017. Improving arterial spin labeling by using deep learning. Radiology 287 (2), 658–666. doi:10.1148/radiol.2017171154. Publisher: Radiological Society of North America.

Knoll, F., Schwarzl, A., Diwoky, C., Sodickson, D.K., 2014. gpuNUFFT - an open-source GPU library for 3D gridding with direct Matlab interface. In: Proceedings of the 22th Annual Meeting of International Society for Magnetic Resonance in Medicine (ISMRM), Milan, Italy.

Kromrey, M.-L., Tamada, D., Johno, H., Funayama, S., Nagata, N., Ichikawa, S., Kõhn, J.-P., Onishi, H., Motosugi, U., 2020. Reduction of respiratory motion artifacts in gadoxetate-enhanced MR with a deep learning-based filter using convolutional neural network. Eur. Radiol. 30 (11), 5923–5932. doi:10.1007/s00330-020-07006-1.

Kurzawski, J.W., Cencini, M., Peretti, L., Gomez, P.A., Schulte, R.F., Donatelli, G., Cosottini, M., Cecchi, P., Costagli, M., Retico, A., Tosetti, M., Buonincontri, G., 2020. Retrospective rigid motion correction of three-dimensional magnetic resonance fingerprinting of the human brain. Magn. Reson. Med. 84 (5), 2606–2615. doi:10.1002/mrm.28301.

Lee, H., Ryu, K., Nam, Y., Lee, J., Kim, D.-H., 2018. Reduction of respiratory motion artifact in c-spine imaging using deep learning: is substitution of navigator possible?. Paris, France.

Lee, S., Jung, S., Jung, K.J., Kim, D.H., 2021. Deep learning in MR motion correction: a brief review and a new motion simulation tool (view2Dmotion). Invest. Magn. Reson. Imaging 24 (4), 196–206. http://www.koreamed.org/SearchBasic.php?RID=2510449.

Liu, J., Kocak, M., Supanich, M., Deng, J., 2020. Motion artifacts reduction in brain MRI by means of a deep residual network with densely connected multi-resolution blocks (DRN-DCMB). Magn. Reson. Imaging 71, 69–79. doi:10.1016/j.mri.2020.05.002.

Ma, D., Gulani, V., Seiberlich, N., Liu, K., Sunshine, J.L., Duerk, J.L., Griswold, M.A., 2013. Magnetic resonance fingerprinting. Nature 495 (7440), 187–192. doi:10.1038/nature11971.

Ma, D., Jiang, Y., Chen, Y., McGivney, D., Mehta, B., Gulani, V., Griswold, M., 2018. Fast 3D magnetic resonance fingerprinting for a whole-brain coverage. Magn. Reson. Med. 79 (4), 2190–2197. doi:10.1002/mrm.26886.

Mallat, S., Zhang, Z., 1993. Matching pursuits with time-frequency dictionaries. IEEE Trans. Signal Process. 41 (12), 3397–3415. doi:10.1109/78.258082. Conference Name: IEEE Transactions on Signal Processing.

McGivney, D.F., Pierre, E., Ma, D., et al., 2014. SVD compression for magnetic resonance fingerprinting in the time domain. IEEE TMI 33, 2311–2322. doi:10.1109/TMI.2014.2337321.

Mehta, B.B., Ma, D., Pierre, E.Y., Jiang, Y., Coppo, S., Griswold, M.A., 2018. Image reconstruction algorithm for motion insensitive MR Fingerprinting (MRF): MORF. Magn. Reson. Med. 80 (6), 2485–2500. doi:10.1002/mrm.27227.

Miao, S., Wang, Z.J., Liao, R., 2016. A CNN regression approach for real-time 2D/3D registration. IEEE Trans. Med. Imaging 35 (5), 1352–1363. doi:10.1109/TMI.2016.2521800. Conference Name: IEEE Transactions on Medical Imaging.

Panda, A., Chen, Y., Ropella-Panagis, K., Ghodasara, S., Stopchinski, M., Seyfried, N., Wright, K., Seiberlich, N., Griswold, M., Gulani, V., 2019. Repeatability and reproducibility of 3D MR fingerprinting relaxometry measurements in normal breast tissue. J. Magn. Reson. Imaging 50 (4), 1133–1143. doi:10.1002/jmri.26717.

Pawar, K., Chen, Z., Shah, N., Egan, G., 2019. Suppressing motion artefacts in MRI using an Inception-ResNet network with motion simulation augmentation. NMR Biomed doi:10.1002/nbm.4225.

Pawar, K., Chen, Z., Shah, N.J., Egan, G.F., 2018. Motion Correction in MRI using deep convolutional neural network. In: Proceedings of the 26th Annual Meeting of International Society for Magnetic Resonance in Medicine (ISMRM), Paris, France, p. 3.

Penny, W., Friston, K., Kiebel, S., Nichols, T., 2006. Statistical Parametric Mapping: The Analysis of Functional Brain Images. Academic Press, Massachusetts doi:10.1016/B978-0-12-372560-8.X5000-1.

Pirkl, C.M., Cencini, M., Kurzawski, J.W., Waldmannstetter, D., Li, H., Sekuboyina, A., Endt, S., Peretti, L., Donatelli, G., Pasquariello, R., Costagli, M., Buonincontri, G., Tosetti, M., Menzel, M.I., Menze, B.H., 2021. Residual learning for 3D motion corrected quantitative MRI: robust clinical T1, T2 and proton density mapping. In: Proceedings of the Forth Conference on Medical Imaging with Deep Learning. PMLR 143, pp. 618–632.

Schlemper, J., Caballero, J., Hajnal, J.V., Price, A.N., Rueckert, D., 2018. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Trans. Med. Imaging 37 (2), 491–503. doi:10.1109/TMI.2017.2760978. Conference Name: IEEE Transactions on Medical Imaging.

Seiberlich, N., Gulani, V., Calamante, F., Campbell-Washburn, A., Doneva, M., Hu, H.H., Sourbron, S. (Eds.), 2020. Quantitative Magnetic Resonance Imaging. volume 1 of Advances in Magnetic Resonance Technology and Applications. Academic Press. URL: https://www.sciencedirect.com/science/article/pii/B978012817057100007X, doi:10.1016/B978-0-12-817057-1.00007-X ISSN: 2666-9099.

Sommer, K., Saalbach, A., Brosch, T., Hall, C., Cross, N.M., Andre, J.B., 2020. Correction of motion artifacts using a multiscale fully convolutional neural network. AJNR Am. J. Neuroradiol. 41 (3), 416–423. doi:10.3174/ajnr.A6436.

Tamada, D., Kromrey, M.-L., Ichikawa, S., Onishi, H., Motosugi, U., 2019. Motion artifact reduction using a convolutional neural network for dynamic contrast enhanced MR imaging of the liver. Magn. Reson. Med. Sci. 19 (1), 64–76. doi:10.2463/mrms.mp.2018-0156.

Ulas, C., Tetteh, G., Thrippleton, M.J., Armitage, P.A., Makin, S.D., Wardlaw, J.M., Davies, M.E., Menze, B.H., 2018. Direct estimation of pharmacokinetic parameters from DCE-MRI using deep CNN with forward physical model loss. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-Lpez, C., Fichtinger, G. (Eds.), Medical Image Computing and Computer Assisted Intervention - MICCAI 2018. Springer International Publishing, Cham, pp. 39–47. doi:10.1007/978-3-030-00928-1_5.

Usman, M., Latif, S., Asim, M., Lee, B.-D., Qadir, J., 2020. Retrospective motion correction in multishot MRI using generative adversarial network. Sci. Rep. 10 (1), 4786. doi:10.1038/s41598-020-61705-9. Publisher: Nature Publishing Group.

Walsh, D.O., Gmitro, A.F., Marcellin, M.W., 2000. Adaptive reconstruction of phased array MR imagery. Magn. Reson. Med. 43 (5), 682–690.

Weigel, M., Schwenk, S., Kiselev, V.G., Scheffler, K., Hennig, J., 2010. Extended phase graphs with anisotropic diffusion. J. Magn. Reson. 205 (2), 276–285. doi:10.1016/j.jmr.2010.05.011.

Xu, Z., Ye, H., Lyu, M., He, H., Zhong, J., Mei, Y., Chen, Z., Wu, E.X., Chen, W., Feng, Q., Feng, Y., 2019. Rigid motion correction for magnetic resonance fingerprinting with sliding-window reconstruction and image registration. Magn. Reson. Imaging 57, 303–312. doi:10.1016/j.mri.2018.11.001.

Zaitsev, M., Maclaren, J., Herbst, M., 2015. Motion artefacts in MRI: a complex problem with many partial solutions. J. Magn. Reson. Imaging 42 (4), 887–901. doi:10.1002/jmri.24850.

Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L., 2017. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Trans. Image Process. 26 (7), 3142–3155. doi:10.1109/TIP.2017.2662206. Conference Name: IEEE Transactions on Image Processing.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20 (1), 45–57. doi:10.1109/42.906424.