Contrast-insensitive motion correction for MRI cardiac T1 mapping

Yue, Chengyu; Huang, Lu; Huang, Lihong; Guo, Yi; Tao, Qian; Xia, Liming; Wang, Yuanyuan

DOI
10.1016/j.bspc.2024.107330

Publication date
2025

Document Version
Final published version

Published in
Biomedical Signal Processing and Control

Biomedical Signal Processing and Control 102 (2025) 107330

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Contrast-insensitive motion correction for MRI cardiac T1 mapping

Chengyu Yue [a], Lu Huang [b], Lihong Huang [a], Yi Guo [a], Qian Tao [c], Liming Xia [b,*], Yuanyuan Wang [a,*]

[a] School of Information Science and Technology, Fudan University, 200438, China
[b] Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, 430030, China
[c] Department of Imaging Physics, Delft University of Technology, 2628 CJ, The Netherlands

ARTICLE INFO

Keywords:
MRI
Cardiac T1 mapping
Image registration
Motion correction
Augmentation

ABSTRACT

Cardiac T1 mapping by magnetic resonance imaging (MRI) is an important clinical tool for the diagnosis and treatment of cardiovascular diseases. In practice, involuntary cardiac and respiratory motion often results in reduced accuracy and precision in T1 estimation. Motion correction is an essential preprocessing step, however, with intensive contrast changes among baseline images, both optimization-based and deep-learning (DL)-based registration methods still struggle to estimate structural similarity between images, especially when image contrast is poor and displacement is large. In this work, we propose a novel registration metric that is highly insensitive to large contrast changes, based on modified modality independent neighborhood descriptor (mo-MIND). To accommodate severe motions, we further propose pre-deformation as an augmentation strategy at the training stage. We combine the proposed mo-MIND-based metric and the augmentation strategy in a U-Net architecture to tackle the challenges of motion correction for cardiac T1 mapping. Experimental results and ablation studies demonstrated that our method achieved improved registration performance compared to several established baselines, leading to significantly reduced T1 mapping error and improved landmark stability.

1. Introduction

Quantitative evaluation of diseased myocardium in cardiac magnetic resonance imaging (MRI) plays an important role in the diagnosis and prognosis of cardiovascular disease [1]. Cardiac T1 mapping is a novel quantitative tool in cardiac MRI for myocardial tissue characterization [2] and has been used in the assessment of iron overload, edema, infarction, diffuse fibrosis, and other myocardial pathologies [3,4]. In T1 mapping, the map of T1 longitudinal magnetization relaxation time (T1 map) is obtained from pixel-wise T1 exponential relaxation curve fitting in a sequence of T1-weighted (T1w) images, also called baseline images, which can be acquired by several advanced T1 mapping techniques [5–8].

Modified Look-Locker inversion recovery (MOLLI) is one of the most widely used myocardial T1 mapping techniques, where 2D baseline images are acquired after two or three inversion pulses during multiple heart cycles within a single breath-hold [9]. By using electrocardiographic gating (ECG-gating) for intermittent sampling, MOLLI minimizes the influence of motion [5]. In recent years, variations of MOLLI have been developed allowing shortened breath-hold durations and reduced sensitivity to heart rate. The shortened MOLLI (ShMOLLI) scheme uses sequential inversion recovery measurements with a single breath hold of only nine successive heartbeats [6]. Other pulse sequences including saturation recovery single-shot acquisition (SASHA) [7] and saturation pulse prepared heart-rate-independent inversion recovery (SAPPHIRE) [8] are also used in clinical practice. However, all the above techniques still require good breath-holding during acquisition, while the inability to hold breath, especially for patients, will cause severe fitting artifacts and thus compromise the accuracy and precision of T1 estimation. Despite employing techniques such as ECG-gating and acquiring images in mid- or end-diastole, artifacts still exist because of triggering errors or drifts [10]. More recently, free-breathing T1 mapping techniques have been proposed, yet they take longer to acquire baseline images and are more susceptible to respiratory motions, with reduced clinical applicability [11]. Besides, cardiac motion due to inter-beat or inter-respiration variability persists, e.g. septal shifts or diaphragmatic drift [9]. Therefore, retrospective non-rigid registration to correct motion between images remains an essential step for cardiac T1 mapping.

Different from conventional registration methods that predict the deformation field between a pair of fixed and moving images, registration for cardiac T1 mapping aims to align a series of images,

* Corresponding authors.
E-mail addresses: lmxia@tjh.tjmu.edu.cn (L. Xia), yywang@fudan.edu.cn (Y. Wang).

https://doi.org/10.1016/j.bspc.2024.107330
Received 10 April 2024; Received in revised form 27 September 2024; Accepted 8 December 2024
Available online 20 December 2024

thus demanding a tailored approach. Moreover, besides addressing the geometric deformations induced by motion, registration for cardiac T1 mapping data must deal with drastic contrast variations among images, including contrast inversion, partial volume effects, and signal nulling [12].

Registration models for cardiac T1 mapping can be divided into two categories: traditional models and DL-based models. Traditional models demonstrate promising performance by iteratively optimizing an energy function for each unseen image pair or sequence, which is generally applicable but computationally expensive. Recent learning-based registration models exhibit enhanced efficiency by minimizing a global function over the training set and utilizing optimal parameters for direct registration at inference.

Optimization-based registration methods have been developed specifically for cardiac T1 mapping data. In [12], a simulated approach based on an initial T1 map was proposed to estimate synthetic images that are motion-free and have similar contrast to the original images. The synthetic images act as references for the alignment of baseline images and will be updated through iteration. In [13], a two-step registration approach was developed, which involves a global affine registration module combined with criterion selection to handle contrast variations and a subsequent refined local non-rigid motion estimation module based on automatic feature tracking. All images are registered to a reference image, which is chosen to be the baseline image with the largest or second-largest inversion time (TI). In [11,14], a group-wise image registration method for quantitative MRI was proposed, which aligns all images simultaneously by minimizing an energy function based on principal component analysis (PCA). In [15], motion correction for cardiac T1 mapping is performed by maximizing the similarity of normalized gradient fields between the reference image and other images. In [9], a model-based registration framework combining a data-driven initialization with a non-rigid registration approach was proposed, which incorporated the T1 exponential relaxation model in the optimization process as an information prior.

While traditional methods have achieved promising performance on the registration of cardiac T1 mapping data through iterative optimization, DL-based frameworks for image registration recently gained popularity for more efficient computation and fast registration. In [16], a convolutional neural network (CNN) called Voxelmorph was proposed to parameterize the registration function, and the parameters of the neural network were optimized using a set of training images. Once the network is trained, the deformation fields can be predicted rapidly from new pairs of fixed and moving images. The accuracy and efficiency of Voxelmorph have been demonstrated on brain MRI and have motivated DL-based registration models for cardiac T1 mapping recently [17]. Arava et al. [18] proposed a learning-based model and a cascading scheme to enable a faster registration for cardiac T1 mapping, where the negative mutual information between the registered baseline images and the reference ones was used as the similarity loss. However, the mutual information metric which is commonly used in multi-modal image registration proved to be insufficient in the face of contrast variations of T1 mapping data such as signal nulling [12]. Li et al. [10] evaluated a motion correction method for cardiac T1 mapping using self-supervised DL-based registration with contrast separation, which can distinguish contrast agent from myocardium and blood pool in the case of just slight motions among baseline images. This method may introduce uncertain noise in other regions except the myocardium and blood pool. Kim et al. [19] first introduced a diffusion model [20] for image registration, which leveraged the latent feature from score function of the diffusion model to predict the deformation field and has demonstrated a notable performance on the cardiac MRI dataset. However, it is impractical to learn the common distribution of the baseline images under the completely different contrast. In a similar spirit as Xue et al. [12], Hanania et al. [21] combined a T1 exponential relaxation model fitting module to generate parametric maps and motion-free synthetic images, with an image registration module that predicts deformation fields for baseline images. Registration is carried out by minimizing the similarity loss between the registered baseline images and the synthetic images. However, this method can only be applied to T1 mapping sequences with an identical number of images, while severe motions in the T1 mapping data may pose significant challenges for generating synthetic images.

In summary, DL-based models have achieved tremendous progress for cardiac T1 mapping registration, but the challenges of intensive contrast variations coupled with cardiac and respiratory motion persist. As mentioned above, Li et al. [10] and Hanania et al. [21] attempted to leverage the T1 magnetization relaxation model to alleviate the impact of contrast variations, but their approaches are limited to specific data and unsubstantial motions among images. Besides, the commonly used mutual information metric is shown to be insufficient to guide the registration for T1 mapping data of dynamic contrasts [12]. The situation is further exacerbated in the presence of severe motions.

In this work, we primarily focus on these two challenges, namely, contrast variation and unpredictable motion that coexist in cardiac T1 mapping. Our contributions are two-fold: First, we proposed a generic contrast-insensitive similarity metric based on modified modality independent neighborhood descriptor (mo-MIND), which reliably represents the anatomical structures under arbitrary variations of image contrast. Second, to specifically cope with the severe motions typical for cardiac T1 mapping, we proposed a pre-deformation strategy, which showed to be a highly effective way to augment data in training the registration network.

2. Method

2.1. Modified modality independent neighborhood descriptor

One of the most significant challenges of registration for cardiac T1 mapping is to deal with drastic contrast variations such as signal nulling (poor contrast) and contrast inversion. These variations make the calculation of similarity loss difficult and inconsistent, complicating the identification of corresponding anatomical structures across images. While the mutual information metric works in general for multi-modal image registration when both images have sufficient contrast [22], it yields limited performance when the contrast in images is extremely poor, as in the case of nulling point in cardiac T1 mapping [12]. In this work, we modify and propose a contrast-insensitive image descriptor, called mo-MIND, tailored to the cardiac T1 mapping registration problem.

In [23], Heinrich et al. introduced a modality independent neighborhood descriptor (MIND), a multi-dimensional image descriptor defined as a feature-based self-similarity to describe the local patterns. The convincing performance of MIND in multi-modal medical image registration has been demonstrated in [24–26]. By computing in a dense fashion over the whole image, MIND provides a high-dimensional anatomical representation of image structure. As shown in Fig. 1, for any location x in a two-dimensional (2D) image I, the MIND feature can be defined by a Gaussian weighted patch distance D_p and a variance estimate V in the spatial search region R:

$$\mathrm{MIND}\left(I,x,x_r\right)=exp\left(-\frac{D_p\left(I,x,x_r\right)}{V\left(I,x\right)}\right),\quad x_r\in R \tag{1}$$

where x_r located at one of the pixels in its search region R which are bounded with red box. The size of R corresponds to the number of dimensions of MIND. The specific features of the region near the right ventricular insertion are shown at the bottom.

$D_p\left(I,x,x_r\right)$ denotes the point-wise squared difference between two Gaussian weighted patches $p_G\left(x\right)$ and $p_G\left(x_r\right)$ as shown in Fig. 1, which center at x and x_r:

$$D_p\left(I,x,x_r\right)=\sum\left(p_G\left(x\right)-p_G\left(x_r\right)\right)^2 \tag{2}$$

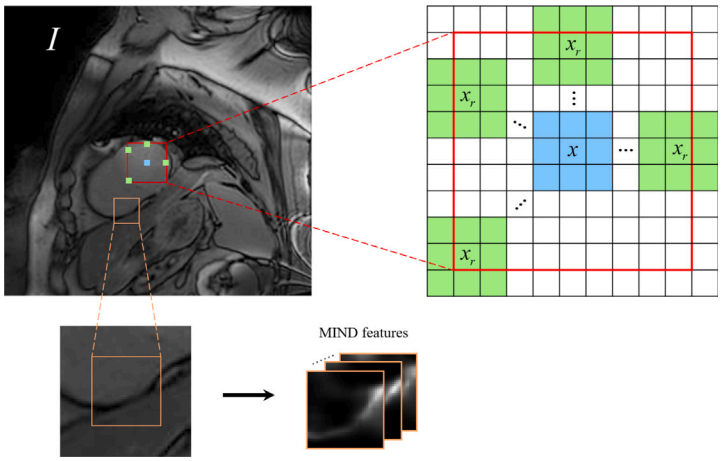

Fig. 1. The concept of extracting the MIND features on cardiac T1 mapping data. The edge features of the right ventricle represented by MIND are illustrated.

where Gaussian weighted patch p_G is obtained by using Gaussian kernel convolution on the image patch p, and Σ represents element summation.

The denominator $V(I,x)$ in Eq. (1) is an estimation of the local variance, Heinrich et al. [23] use the mean of Gaussian weighted patch distances D_p within the neighborhood of x to measure the variance:

$$V(I,x) = \frac{1}{N}\sum_{n\in N} D_p\left(I,x,x_n\right) \tag{3}$$

where N is neighborhood region of x.

However, dividing D_p by the result of Eq. (3) is just like a local normalization operation, which indeed reduces the sensitivity of the features to spatially varying noise, yet decreases the distinctiveness of extracted anatomical representation. Specifically, the features extracted by MIND only represent restricted anatomical information because of the violent estimation of V. Instead, we assume that the calculation of local variance should not depend on the average of patch distances themselves, but needs to take the actual distribution of image intensities near the location x into account.

Incorporated with the image descriptor in Eq. (1), the local variance should be related to the amount of noise, we propose to compute the variance of neighboring patches and denote it as modified local variance V_m:

$$V_m(I,x) = \underset{n\in N}{\mathrm{Var}}\left(\sum p(x_n)\right) \tag{4}$$

where $\sum p\left(x_n\right)$ is the sum of all elements of the patch $p\left(x_n\right)$ centered at x_n. In this way, the mean of Gaussian weighted patch distances V is replaced by the variance of neighboring patches V_m. A higher value for V_m represents there are larger differences in image intensities within the neighborhood N, indicating sharper and stronger responses, while a smaller value means similar local intensities, leading to smooth results. According to Eqs. (1), (2), (4), our modified modality independent neighborhood descriptor can be written as follows:

$$mo\text{-MIND}\left(I,x,x_r\right) = \exp\left(-\frac{\sum\left(p_G(x) - p_G\left(x_r\right)\right)^2}{\underset{n\in N}{\mathrm{Var}}\left(\sum p(x_n)\right)}\right), \quad x_r \in R \tag{5}$$

The patch distances are calculated through square of the pixel-wise difference between Gaussian weighted patches $p_G(x)$ and $p_G\left(x_r\right)$, just as that in [23]. Based on concept of image self-similarity [27], the image features including regions, edges, corners and textures are all treated in a unified manner, which makes it possible to extract anatomical representation through similarity of image patches independent of contrast variation. Associated with the patch distances $D_p = \sum(p_G(x) - p_G(x_r))^2$ in Eq. (5), mo-MIND will obtain a high response at x for patches $p_G(x)$ that are similar to the patch $p_G\left(x_r\right)$, and vice versa.

The local variance $V_m = \mathrm{Var}_{n\in N}\left(\sum p(x_n)\right)$ refers to the estimation of neighboring noise, which reflects the texture complexity and the intensity of change in region N. While extracting the structural representation with mo-MIND, the V_m are used to control the effect of local intensity differences on the patch distances D_p. Specifically, large value of V_m means greater variation of local structure, which allows more significant intensity differences between patches to still be considered similar. Instead, small values of V_m indicate less structural variation. In this case of smoothed regions, only patches with high similarity can achieve a strong response in the resulting representation. To sum up, by introducing local variance, mo-MIND is able to dynamically adjust the similarity criteria according to the complexity of the local structure in different regions, while enhancing the robustness in the scenarios of poor contrast or contrast inversion.

Ultimately, the modified modality independent image descriptor is formulated through the relative intensity distributions of the local neighborhoods. For each pixel x in image I, a vector of shape $|R|$ is derived corresponds to all x_r in the search region R, which is denoted as mo-MIND. According to Eq. (5), the overall size of mo-MIND is $|R| \times H \times W$, where H and W are the height and width of the image respectively.

The overview of the MIND and our mo-MIND are schematically reported in Fig. 2. we use both of them to extract the features on two intentionally selected baseline images from one sequence in which the contrast inversion is conspicuous. For simplicity and representativity, only the features within the subject-specific rectangular region of interest covering both ventricles are shown in Fig. 2. It can be seen that the intensity of the blood pool in image a is higher than that of the myocardium, whereas in image b it is significantly lower. The second column displays the features extracted by MIND, the anatomical structures of MIND are somewhat disordered, and the broad outlines of ventricles are incomplete and unclear. On the contrary, mo-MIND provides a more robust representation for both images as depicted in the rightmost column in Fig. 2, even in the presence of severe contrast inversion.[1]

Once the mo-MIND of reference and registered images are obtained, the similarity metric between them can be defined as the mean of absolute differences between the mo-MIND features:

$$\ell_{\mathrm{sim}}(I,J) = \frac{1}{|R|\times H\times W}\sum\left|mo\text{-MIND}\left(I,x,x_r\right) - mo\text{-MIND}\left(J,x,x_r\right)\right| \tag{6}$$

[1] The baseline images in Figs. 2–6 are enhanced by Gamma transformation of varying degrees for better visualization.

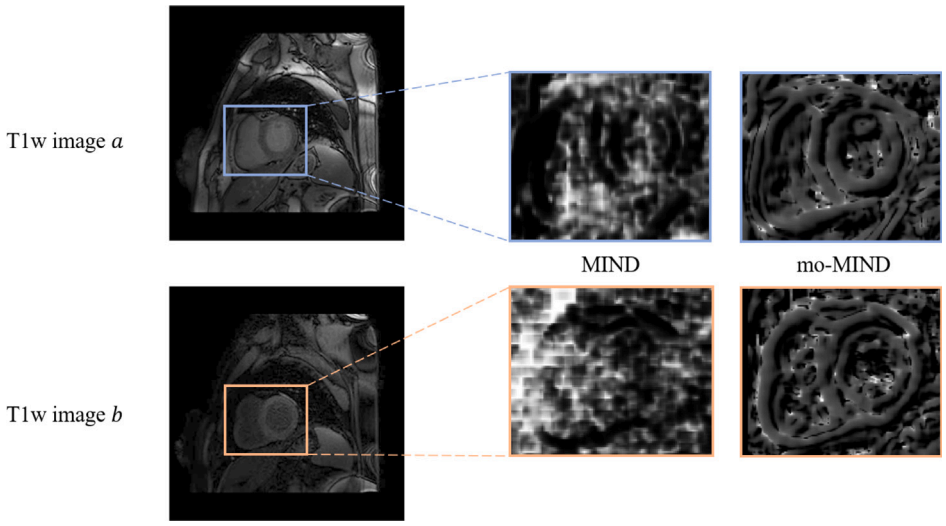

Fig. 2. The comparison of features from MIND and mo-MIND. Two images with significant contrast differences are selected as examples, where the features of the ventricular region are displayed at the right. It can be seen that our proposed mo-MIND can provide a clearer anatomical representation than MIND, robust to the variations in image contrast.

where I and J are reference image and registered image respectively, and $\Sigma|\cdot|$ represents the sum of all elements. During registration process, pixels are considered to be matched if they have a similar neighborhood structure (i.e., similar local structure) in different images.

2.2. Pre-deformation augmentation

Depending on the patients, involuntary respiratory and cardiac motion often occur when acquiring T1 mapping sequences. Such motions are unpredictable and may not exist in the training data. When it occurs at inference time, it will make a well-trained registration model dysfunctional. To address this problem, we propose a pre-deformation strategy for the registration network. As severe motions occur mostly in the ventricular region, we apply this deformation augmentation focusing on this region.

Similar to the data augmentation in [28,29]. Our pre-deformation strategy is carried out through B-spline-based free-form deformation (FFD) [30]. We denote the domain of the 2D baseline image as $\Omega = \{(x, y)\,|0 \le x \le X, 0 \le y \le Y\}$. Let Ψ be a grid of control points $\psi_{i,j}$ with uniform space $n_x \times n_y$ overlaid on the region Ω. Then, the FFD can be written as the 2D tensor product of the familiar one-dimensional (1D) cubic B-spline:

$$v(x, y) = \sum_{l=0}^{3}\sum_{m=0}^{3} B_l(u)\, B_m(v)\, \psi_{i+l,j+m} \tag{7}$$

where $i = \left\lfloor \frac{x}{n_x} \right\rfloor - 1$, $j = \left\lfloor \frac{y}{n_y} \right\rfloor - 1$, $u = \frac{x}{n_x} - \left\lfloor \frac{x}{n_x} \right\rfloor$, $v = \frac{y}{n_y} - \left\lfloor \frac{y}{n_y} \right\rfloor$ and B_l, B_m are the uniform cubic B-spline basis functions. The locations of control points $\psi_{i,j}$ serve as parameters of the B-spline-based FFD, which form the displacement vector $v(x, y)$. The spaces of control points n_x, n_y influence the degree of FFD and determine the resolution of the grid Ψ. In particular, more significant deviations of control points from their initial positions bring about more substantial deformations. Large spaces of control points yield smooth deformations, while small spaces allow fine and sharp ones.

We apply a grid of control points Ψ with uniform space 7×7, as the motions in T1 mapping data are typically elastic and smooth. Since the motions primarily manifest around the region of ventricles, we specifically identify and adjust a subset of control points near the ventricles to simulate these motions, keeping the remaining control points fixed. Fig. 3 gives an example of the pre-deformation strategy. The red spots represent the fixed points and the yellow spots are the control points we specified to generate deformations near the ventricles. As the ventricular region mostly lies in the center of cardiac MRI, we suggest adjusting the central 3×3 point set to simulate the motions. The deviations of control points along height and width are denoted as d_x and d_y respectively, and they follow a uniform distribution:

$$(d_x, d_y) \sim U([-d, d], [-d, d]) \tag{8}$$

where d is the distance threshold at pixel level. In Eq. (8), $-d$ represents shifting the control points to the left or downward, while d denotes moving them to the right or upward. Through random deformation augmentation of the training data, the registration model accesses and learns anatomical representation across a broad range of scales, thereby improving the robustness of registration for cardiac T1 mapping.

2.3. Registration model for contrast-insensitive motion correction

Registration is usually formulated as an optimization problem aiming to search for a deformation field $\phi \in \mathbb{R}^{H\times W\times 2}$ that maps coordinates of the fixed image $I_f \in \mathbb{R}^{H\times W}$ to those of the moving image $I_m \in \mathbb{R}^{H\times W}$. Typically, ϕ is often characterized by a displacement vector field u that specifies the vector offset from I_f to I_m for each pixel: $\phi = Id + u$, where Id is the identity deformation field. The energy function of the optimization problem is defined as:

$$\arg\min_{\phi} \ell(I_f, I_m, \phi) = \arg\min_{\phi} \ell_{\text{sim}}(I_f, I_m \circ \phi) + \lambda \text{Reg}(\phi) \tag{9}$$

where $I_m \circ \phi \in \mathbb{R}^{H\times W}$ represents the registered image obtained by warping the moving image I_m with the deformation field ϕ. $\ell_{\text{sim}}(\cdot,\cdot)$ measures the dissimilarity of the fixed image and the registered image, $\text{Reg}(\phi)$ acts as a regularization term for the deformation field ϕ and scalar λ is the trade-off parameter.

Traditional registration models for cardiac T1 mapping minimize the energy function for each pair of reference and baseline images through iterative optimization. Instead, we propose a DL-based registration model, as shown in Fig. 4, for motion correction of cardiac T1 mapping. The network parameters are optimized by minimizing the energy function in Eq. (9) over a set of T1 mapping data.

The architecture of our proposed model for contrast-insensitive motion correction is depicted in Fig. 4(a). Given one certain T1 mapping sequence, the model takes original baseline images I^i, $i = 1, 2, \ldots, N$ as input and generates random-warped images I^i_W, $i = 1, 2, \ldots, N$ through our pre-deformation module firstly. The warped images along with their templates I_t are then fed into a U-Network to predict the

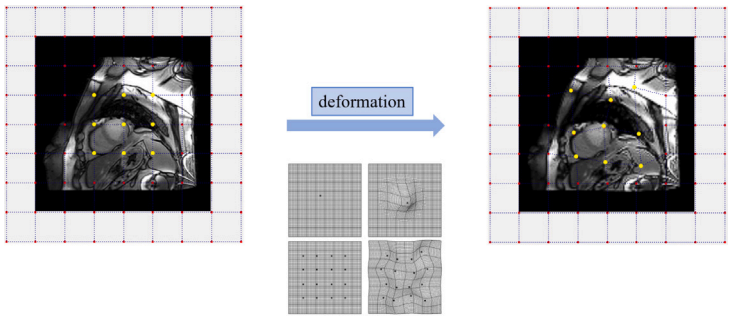

Fig. 3. The schematic diagram of pre-deformation strategy by B-spline-based FFD. Specific control points (yellow spots) near the ventricles are adjusted to warp the baseline images and simulate motions in T1 mapping data [31].

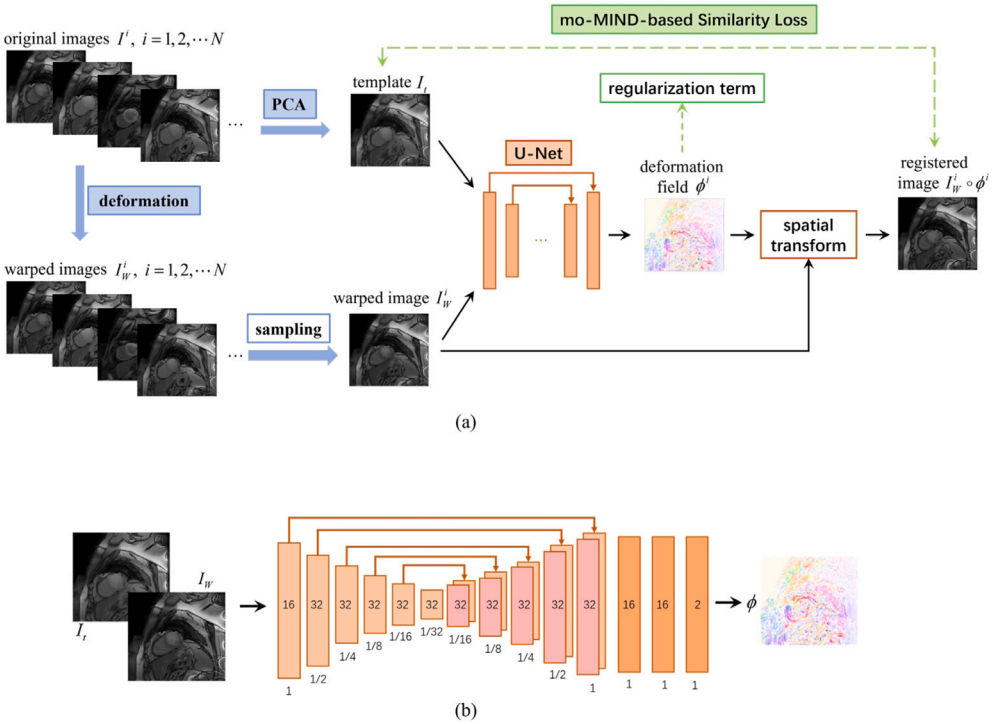

Fig. 4. The overall architecture of our proposed registration model for cardiac T1 mapping. (a) Contrast-insensitive motion correction scheme with mo-MIND-based similarity loss and pre-deformation augmentation. During training, batches of warped baseline images and the corresponding templates are fed into the network. (b) Illustration of the U-Network. Each rectangle represents a 2D feature, generated from the preceding feature through a 2D convolutional layer. The spatial resolution of each feature with respect to the input images is marked underneath.

deformation fields ϕ^i for each pair of I_W^i and I_t. Finally, the registered images $I_W^i \circ \phi^i$ are obtained through a spatial transformation function which warps I_W^i using the deformation fields ϕ^i. The parameters of our registration model are optimized by minimizing the energy function in Eq. (9), where the $\ell_{\text{sim}}(\cdot,\cdot)$ is the mo-MIND-based loss in Eq. (6). We implement regularization term $\lambda \text{Reg}(\phi)$ as two loss functions:

$$\lambda \text{Reg}(\phi) = \lambda_1 \ell_{\text{smooth}}(\phi) + \lambda_2 \ell_{\text{small}}(\phi) \tag{10}$$

where $\ell_{\text{smooth}}(\phi) = \frac{1}{\mathcal{P}} \sum_{p \in \mathcal{P}} \|\nabla u(p)\|^2$ enforces the smoothness of the deformation fields and $\ell_{small}(\phi) = \sum_{p \in \mathcal{P}} \|u(p)\|^2$ is used to penalize large deformations. In our experiments, the hyperparameters λ_1 and λ_2 are empirically set to 1.0 and 0.1 to balance the gradient values and deformations to a small range for topological consistency. During inference, given the original baseline images I^i and their templates I_t, the corresponding deformation field ϕ^i can be rapidly predicted through the trained registration network.

The U-Network structure is illustrated in Fig. 4(b), both the encoder and decoder layers consist of convolutional layers with kernel size of 3 and stride of 1 followed by LeakyReLU activation function. In the encoding stage, coarse representations of the input images are captured by successive convolutional layers and max pooling. In the decoding stage, convolutions, upsampling, and skip connection are applied alternatively to propagate learned features from the encoder to the decoder layers.

In our proposed registration model for cardiac T1 mapping, a practical template image of each sequence should be formulated, which represents the center of all baseline images (i.e., a common space). Considering the superior performance of PCA in data dimensionality reduction [32], we use PCA to generate the PCA-based template, which contains the principal component information of all baseline images from one T1 mapping sequence. Specifically, we treat different baseline images as different variables and each pixel as an observation to compute the first principal component that serves as the PCA-based template. The eigenvector $V = (w_1, w_2, w_3 \cdots w_N)$ associated with the largest eigenvalue is used to construct the template through $I_t = \sum_{i=1}^{N} w_i \cdot I^i$, where I^i is the original baseline images from one sequence and N is the number of images.

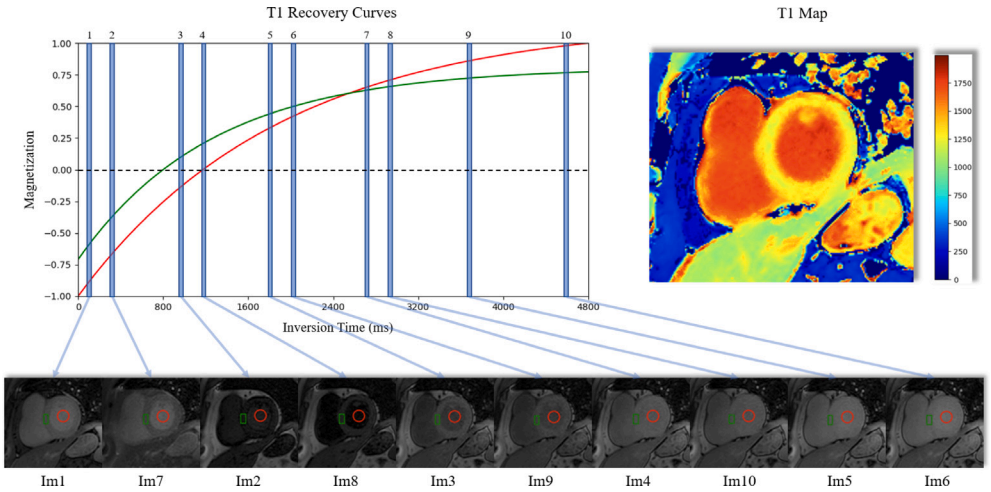

Fig. 5. A sequence of baseline images is illustrated at the bottom along the inversion time, with the T1 map at the top right and the T1 relaxation curves for myocardium (green) and blood pool (red) at the top left.

2.4. Cardiac T1 mapping

Once a sequence of motion-free baseline images is acquired, the T1 map can be estimated through pixel-wise curve fitting using the three-parameter signal model:

$$S(TI) = A - B \cdot \exp\left(-\frac{TI}{T_1^*}\right) \tag{11}$$

where TI is the inversion time and $S(TI)$ denotes corresponding signal intensity at each pixel. The T1 longitudinal magnetization relaxation time (T1 value) is calculated through the resulting parameters A, B and apparent T1 value T_1^* using the equation: $T_1 = \left(\frac{B}{A} - 1\right) \cdot T_1^*$[5]. Fig. 5 shows a sequence of baseline images from our dataset and the corresponding T1 map, the T1 relaxation curves of two regions of interest along inversion time are also reported. It can be seen that the tissue characterization of the myocardium (green), blood pool (red), etc. are much different from each other with differing values of T1. Wholesale signal nulling that occurs near zero crossing in T1 exponential relaxation curves is also expressed in $Im2$ and $Im8$.

3. Experiments and results

3.1. Dataset

We used an in-house cardiac MRI dataset including 36 consecutive patients (age 54 ± 16, 15 female). The acquisitions were performed on a 3.0T Ingenia MR-scanner (Philips Healthcare, Best, The Netherlands). Each subject has both pre-contrast and post-contrast MOLLI sequences. There are 108 sequences in total, with lengths ranging from 9 to 14. Since pairwise registration is used, the number of baseline images does not affect the pipeline. All the baseline images in our dataset cover the chest cavity with the ventricles located approximately at the central region. We adjust all images to the same size 448×448 before they are fed into the registration model, and the registered baseline images will be converted back to their original sizes. The pre-deformation augmentation is carried out during training.

In our implementation, we performed a 5-fold cross-validation on our dataset for all the comparison methods and ablation studies. Specifically, we identified a test set of 18 sequences which containing labeled right ventricular insertion points and then randomly divided the rest of the dataset into the training set and validation set on a proportion of 4:1 for each fold.

3.2. Evaluation metrics

We validated our method using the mean of T1 mapping error and the stability of right ventricle insertion points (SRVI). We followed Kellman et al. [33] to transform the standard deviation (SD) of the residual fitting error into the SD of the estimated parameters. The resulting SD of the T1 value estimation refers to the T1 mapping error. By calculating the T1 mapping error pixel by pixel, we can obtain the T1 error map i.e., the SD map of the estimated T1 value. The T1 mapping error reflects the accuracy of the T1 exponential relaxation curve fitting in Eq. (11). According to the T1 longitudinal magnetization relaxation principle, image intensities at the same positions along inversion time should precisely match an exponential function. In other words, better alignment on the baseline images results in reduced fitting error and subsequently small T1 mapping error. To this end, we used the mean of T1 mapping error across all T1 mapping sequences in the dataset to evaluate the registration performance. In our implementation, only the T1 mapping error in the manually annotated myocardial region of interest (ROI) was taken into account.

In addition, we also introduced a novel evaluation criterion based on the right ventricle insertion points. These points have been demonstrated clinically important in several cardiac diseases such as dilated cardiomyopathy [34]. As shown in Fig. 6, the right ventricle insertion points are located at the junction of the free wall of the right ventricle with the interventricular septum, where the green spots are the anterior insertion points and the blue ones denote the posterior insertion points. We marked these points in the original baseline images and evaluated the registration performance by measuring their stability in the registered images. For our implementation, given one certain T1 mapping sequence, the SRVI of two kinds of insertion points can be formulated respectively as the SD of their positions across all baseline images. A smaller SRVI value indicates a better alignment, particularly in the region of the ventricles. The mean SRVI value for all T1 mapping sequences serves as another metric for comparing different registration methods.

3.3. Experimental setup

Elastix [35], a classical registration toolbox that packed a set of methods for medical image registration, acts as the first baseline. This model was deployed on an Intel Xeon CPU E5-2687 W v4, using the following settings on our dataset: transform: B-spline, final spaces of grid: 64, metric: mutual information with 32 histogram bins, optimizer: adaptive stochastic gradient descent, number of resolutions: 3, number of iterations per resolution: 500. We additionally use VoxelMorph [16]

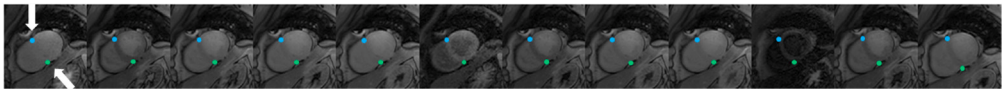

Fig. 6. Illustration of the right ventricle insertion points in the baseline images from one T1 mapping sequence. The SD of their locations well indicates the registration performance.

Table 1
The mean of T1 mapping error and the mean SRVI on the test set for all the models in comparison.

Method	T1 mapping error (ms)	SRVI (mm)	
		Anterior insertion	Posterior insertion
Original	53.97 ± 2.89 ($p < 0.05$)	1.36 ± 0.17 ($p < 0.05$)	1.16 ± 0.15 ($p < 0.05$)
B-spline	46.58 ± 2.87 ($p < 0.05$)	1.75 ± 0.49 ($p < 0.05$)	1.56 ± 0.28 ($p < 0.05$)
VoxelMorph	40.68 ± 2.84 ($p < 0.05$)	1.22 ± 0.18 ($p < 0.05$)	1.08 ± 0.12 ($p < 0.05$)
DiffuseMorph	43.20 ± 3.08 ($p < 0.05$)	1.11 ± 0.11 ($p < 0.05$)	1.20 ± 0.10 ($p < 0.05$)
PCMC-T1	41.61 ± 3.69 ($p < 0.05$)	1.07 ± 0.13 ($p < 0.05$)	1.03 ± 0.11 ($p < 0.05$)
Ours	36.92 ± 3.15 ($p < 0.05$)	1.02 ± 0.13 ($p < 0.05$)	1.02 ± 0.13 ($p < 0.05$)

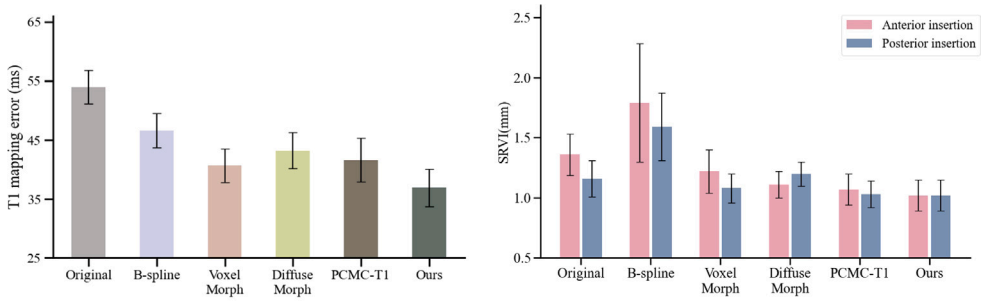

Fig. 7. The mean of T1 mapping error (a) and the mean SRVI (b) for original and registered T1 mapping data.

with the default settings as another baseline, the similarity metric of which is the local normalized cross-correlation (LNCC). The registration manner of Voxelmorph is the same as our proposed model, i.e., aligning all baseline images to their templates.

The DL-based model containing Voxelmorph, DiffuseMorph, PCMC-T1, our proposed model and their variants in following experiments are all implemented in PyTorch using the PyTorch framework on a single Nvidia RTX V100 GPU. The Adam optimization [36] is used to minimize the loss function for all these methods. According to VoxelMorph, our proposed model and their variants, we trained them for 1500 epochs with batch size 4 and learning rate 8×10^{-4}. Besides, we choose the search region $R = 11$ in MIND and mo-MIND and the threshold $d = 20$ in our pre-deformation strategy. The implementation details of DiffuseMorph are the same as its default setting for cardiac MR data. As for PCMC-T1, we first crop the T1 mapping sequences to length of 9, and then trained the model for 1000 epochs with the default settings in [21]. All methods used for our experiments have been adjusted to optimal performance.

3.4. Comparison of different methods

In this section, we evaluated the performance of our registration model and compared it with one classical registration method and three DL-based deformable registration models.

Quantitative results across six different scenarios: the original baseline images, and the registered baseline images acquired by B-spline, Voxelmorph, DiffuseMorph, PCMC-T1 and our proposed mode are reported in Table 1, with the histograms of the mean of T1 mapping error (a) and the mean of SRVI (b) shown in Fig. 7. It can be observed that our proposed registration model outperforms all other methods, with the lowest T1 mapping error of 36.92 ± 3.15 ms ($p < 0.05$) and the same lowest SRVI values of 1.02 ± 0.13 mm ($p < 0.05$) for both anterior and posterior insertion points. As depicted in Table 1, four different methods in comparison also demonstrate notable registration performance with significantly lower T1 mapping errors than that of original baseline images. Our proposed model further reduces the T1 mapping error, which is 3.76 ms lower than that of the best comparison model VoxelMorph, and 4.69 ms lower than that of the most recent model PCMC-T1. The results of SRVI outcomes, shown in Fig. 7(b), reveal that our proposed method achieved the best SRVI values, indicating the most stable alignment, particularly within the ventricular regions. The DL-based models including Voxelmorph, DiffuseMorph and PCMC-T1 also achieved fine performances on the SRVI, with much lower SRVI values than that of original baseline images. However, B-spline got an even worse SRVI value than that of original baseline images, which are 0.39 mm increase for anterior insertion points and 0.40 mm increase for posterior insertion points, respectively. The unexpected performance on SRVI is caused by the occurrence of the liquefaction in B-spline transform, indicating that B-spline struggled to correct the motions among T1 mapping data.

The visual results of different scenarios are displayed in Fig. 8, including the T1 maps and T1 error maps of one given sequence before and after registration. The T1 maps and T1 error maps corresponding to the original baseline images are illustrated in the first column, and the results after motion correction with different methods are shown in the subsequent columns. The motions among original baseline images introduced severe artifacts in the corresponding T1 map and significant mapping errors in the error map. After applying motion correction, the quality of T1 maps improved, with fewer artifacts and more homogeneous mapping. The overall T1 mapping error of five registered cases also reduced compared to that of the original images. As depicted in the T1 error maps, the fitting failure (marked by white arrow in Fig. 8) became fewer after registration. However, all the comparison methods struggled with the motions among T1 mapping data, leading to residual artifacts in the T1 maps and increased T1 mapping error. Our proposed method well aligned both slight and severe motions under the impact of drastic contrast variations and demonstrated the best registration performance with the lowest T1 mapping error.

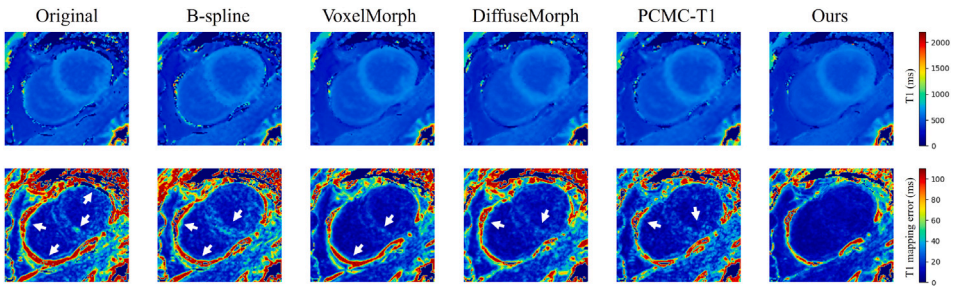

Fig. 8. Visualization of the registration results for comparison. The T1 maps and T1 error maps of one sequence are displayed, where the white arrows in T1 error maps denote significant fitting errors.

Table 2
Effect of different similarity losses on registration performance. The mean values of T1 mapping error and the mean SRVI are reported, where "mo-MIND" and "MIND" represent the similarity loss based on them, and NMI denotes the normalized mutual information loss.

Method	T1 mapping error (ms)	SRVI (mm)	
		Anterior insertion	Posterior insertion
Ours w mo-MIND	36.92 ± 3.15 ($p < 0.05$)	1.02 ± 0.13 ($p < 0.05$)	1.02 ± 0.13 ($p < 0.05$)
Ours w MIND	39.16 ± 2.94 ($p < 0.05$)	1.08 ± 0.15 ($p < 0.05$)	1.03 ± 0.10 ($p < 0.05$)
Ours w NMI	40.09 ± 3.55 ($p < 0.05$)	1.30 ± 0.15 ($p < 0.05$)	1.16 ± 0.12 ($p < 0.05$)

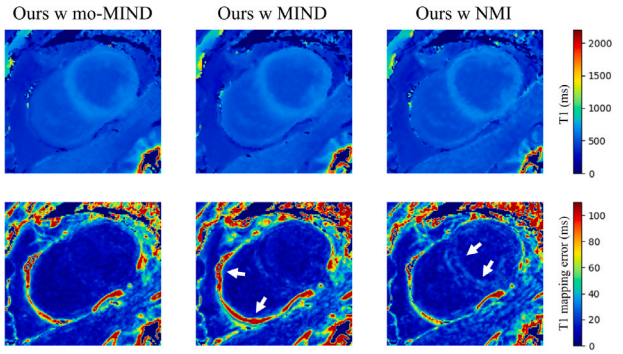

Fig. 9. Visual comparison of registration results with different similarity losses. The T1 maps and T1 error maps of one T1 sequence are displayed.

3.5. Ablation studies

In this section, we investigated the effectiveness of the contrast-insensitive loss based on mo-MIND and the pre-deformation augmentation module.

We evaluated the registration performance of our proposed model with different similarity losses, including mo-MIND-based loss, MIND-based loss, and normalized mutual information (NMI) loss. The quantitative results of the T1 mapping error in Table 2 show that our proposed model with the similarity loss based on mo-MIND achieved much better registration performance than the other two cases. The mean of T1 mapping error was 36.92 ± 3.15 ms ($p < 0.05$) for our proposed model with mo-MIND-based loss, significantly lower than that of the model with MIND-based loss 39.16 ± 2.94 ms ($p < 0.05$) and the model with NMI loss 40.09 ± 3.55 ms ($p < 0.05$). The corresponding T1 maps and T1 error maps of different cases are displayed in Fig. 9 for visualization. In the T1 error maps, it is evident that the overall fitting errors of our proposed model using similarity loss based on mo-MIND are much lower than those of the models with the other two losses. The registration model employing MIND-based loss also exhibited fine performance in the region of myocardium and blood pool, but the edges of the ventricles were not well aligned with higher values. In contrast, the model utilizing negative NMI as the similarity loss achieved the lowest T1 mapping error near the edges, yet it struggled in the region of myocardium and blood pool.

According to the pre-deformation strategy, we evaluated the registration performance of our proposed model and Voxelmorph, with and without this augmentation. As reported in Table 3, the mean of T1 mapping error significantly rises after removing the pre-deformation augmentation, from the mean of T1 mapping error of 36.92 ± 3.15 ms ($p < 0.05$) to 39.17 ± 2.86 ms ($p < 0.05$). Besides, without the pre-deformation module, our proposed model still outperformed Voxelmorph, with a reduction of 1.51 ms in the T1 mapping error. The reason is attributed to that the anatomical representation is well extracted and utilized in ours. Interestingly, it can be observed that Voxelmorph performed even worse after employing the pre-deformation strategy, with 5.27 ms of T1 mapping error increase. That means the LNCC loss used in Voxelmorph is struggling with capturing the spatial correspondence between images, let alone in the presence of severe motions. Fig. 10 showed the T1 maps and T1 error maps of one certain T1 mapping sequence, which reflect consistent behavior with the quantitative results in Table 3, where more significant T1 mapping error corresponds to more serious artifacts in the T1 maps and much higher values in the error maps.

4. Discussion

In this study, we proposed a robust registration model aimed at motion correction for cardiac T1 mapping data. Our work incorporates two key components: Firstly, we proposed a contrast-insensitive image descriptor called mo-MIND, which demonstrated robustness to the drastic variations in image contrast. The mo-MIND yields consistent anatomical representation for images independent of differing variations, enabling our model to accurately identify the corresponding structures among baseline images. Secondly, a pre-deformation augmentation strategy was introduced to augment the training data. By integrating this random deformation augmentation, the registration performance was further improved, especially in the presence of severe motions. Experimental results showed that our proposed method outperformed both B-spline and the baseline Voxelmorph, in terms of the T1 mapping error and the SRVI value.

The proposed mo-MIND demonstrates a strong capability of extracting anatomical representation under the drastic contrast variations in T1 mapping data. This modified descriptor takes the actual distribution of image intensities to estimate the local variance, producing much clearer and more explicit anatomical information compared to MIND. This structural information is utilized to compute the similarity loss for

Table 3
Effect of pre-deformation augmentation on registration performance. The mean values of T1 mapping error and the mean SRVI are reported, where "pre" refers to our pre-deformation strategy.

Method	T1 mapping error (ms)	SRVI (mm)	
		Anterior insertion	Posterior insertion
Ours w pre	$36.92 \pm 3.15\,(p < 0.05)$	$1.02 \pm 0.13\,(p < 0.05)$	$1.02 \pm 0.13\,(p < 0.05)$
Ours w/o pre	$39.17 \pm 2.86\,(p < 0.05)$	$1.17 \pm 0.15\,(p < 0.05)$	$1.13 \pm 0.13\,(p < 0.05)$
Voxelmorph w/o pre	$40.68 \pm 2.84\,(p < 0.05)$	$1.22 \pm 0.18\,(p < 0.05)$	$1.08 \pm 0.12\,(p < 0.05)$
Voxelmorph w pre	$45.95 \pm 2.83\,(p < 0.05)$	$1.27 \pm 0.17\,(p < 0.05)$	$1.10 \pm 0.13\,(p < 0.05)$

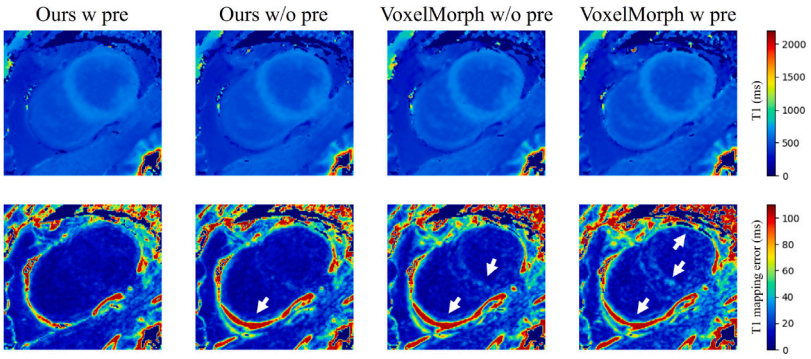

Fig. 10. Visual comparison of registration results with and without pre-deformation augmentation. The T1 maps and T1 error maps of one T1 mapping sequence are displayed.

model optimization, resulting in a notably improved registration performance. Experimental results in Section 3.4 indicated the efficacy of the mo-MIND-based similarity loss in enhancing registration performance.

The pre-deformation strategy in our registration model augments the T1 mapping data, emulating realistic motions in practice. Its effectiveness was empirically demonstrated by our ablation study which compared the registration performance of our proposed model with and without this deformation augmentation in Section 3.4.

5. Limitation and future work

Our proposed contrast-insensitive motion correction is carried out in a pairwise manner. As such, the registration pipeline is flexible to quantitative sequences of different length and our model can be adapted to other registration tasks. Nonetheless, one limitation of pairwise registration is that the overall interrelation between baseline images and the information of the T1 exponential relaxation model is not taken into account. In future work, we will focus on how to integrate the interrelationships among baseline images and the extensions of other quantitative cardiac MRI mapping. Besides, we are expecting to leverage diffusion models to generate baseline images that are motion-free and evaluate the registration performance using simulated motions.

6. Conclusion

In this paper, we present a robust registration method for the motion correction of cardiac T1 mapping data. The proposed contrast-insensitive image descriptor mo-MIND produces robust anatomical representation and facilitates reliable loss calculation, even for those baseline images with extremely poor contrast. By incorporating the pre-deformation augmentation into training, we further enhance the registration accuracy in situations of large motion. Experimental results demonstrated significantly improved registration performance compared with current baseline methods, with reduced T1 mapping error and SRVI value. The proposed methodologies tackle the major challenges of cardiac T1 mapping motion correction in an effective and generic manner, with robust structure representation and enhanced motion accommodation. Therefore, our method can be applied to the expanding family of quantitative cardiac MRI sequences, to facilitate clinical practice and medical research.

CRediT authorship contribution statement

Chengyu Yue: Writing – original draft, Software, Methodology, Conceptualization. **Lu Huang:** Validation, Resources, Formal analysis. **Lihong Huang:** Visualization, Validation. **Yi Guo:** Validation, Methodology, Investigation. **Qian Tao:** Writing – review & editing, Methodology, Data curation. **Liming Xia:** Resources, Investigation. **Yuanyuan Wang:** Supervision, Conceptualization.

Research data

Due to the sensitive nature of the questions asked in this study, survey respondents were assured raw data would remain confidential and would not be shared.

Declaration of competing interest

The 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.

Acknowledgment

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix

The drastic variations in image contrast introduce significant noise in the calculation of the similarity loss for T1 mapping registration. In other words, The varying contrast may lead to more substantial losses compared to the anatomical misalignment. Consequently, several similarity metrics exhibit heightened sensitivity to contrast variations rather than anatomical deformations in the T1 mapping data, making it challenging to identify corresponding structures. In this section, we evaluated the sensitivity of several similarity metrics on changes in deformation and contrast. The results demonstrated that our mo-MIND-based similarity loss is less sensitive to contrast variations compared to other metrics.

Given one certain similarity loss $\ell_{\text{sim}}(\cdot,\cdot)$, we take image I and its transformed image $Tr_{d,c}(I)$ as input to calculate the loss, where

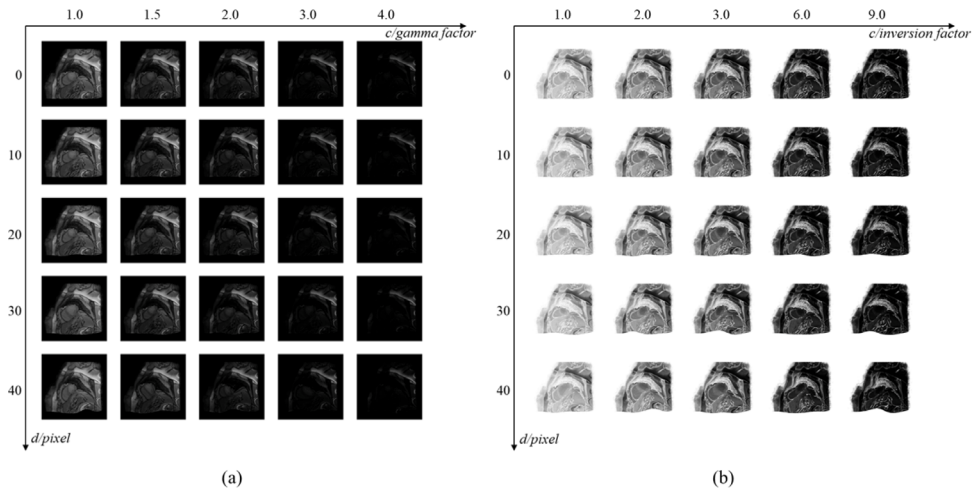

Fig. A.1. The schematic diagram of transformations on deformation and contrast. (a) The Gamma transform combined with the deformation transform. (b) The inversion transform combined with the deformation transform.

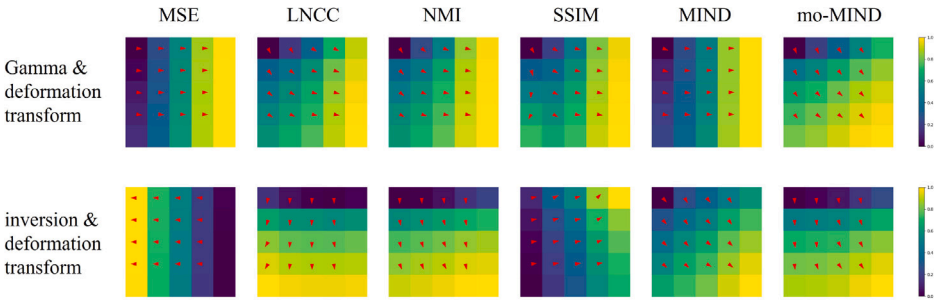

Fig. A.2. Illustration of loss maps of different similarity losses with different transformations. The results of the Gamma transform combined with the deformation transform are displayed in the first row while the results of the inversion transform combined with the deformation transform are displayed in the second row. Each column corresponds to the results of a specific similarity loss.

$Tr_{d,c}(I)$ denotes deformation and contrast transform applied to image I with parameters d and c. The deformation transform is the same as that in our pre-deformation strategy of Section 2.2 and d is the distance threshold in Eq. (8). For the contrast transform, we respectively utilize Gamma transform: $y = x^c$ and inversion transform: $y = (1-x)^c$. In the Gamma transform, c serves as the Gamma factor regulating the overall intensities of the image, while in the inversion transform, c controls the overall intensities of the inverted image. We then use these two contrast transformations, combined with the deformation transform to translate baseline images. The parameters d and c are several fixed values sampled within a suitable range to simulate the challenges of deformation and contrast in our dataset. An illustration of the simulation scheme is shown in Fig. A.1.

For our implementation, five often-used metrics of registration were evaluated, including Mean Squared Error (MSE), Local Normalized Cross Correlation (LNCC), Normalized Mutual Information (NMI), Structural Similarity Index Measure (SSIM), MIND-based loss, and our proposed similarity loss based on mo-MIND. According to each metric, we computed the mean of similarity losses $\ell_{\text{sim}}\left(I, Tr_{d,c}(I)\right)$ for each specific transformation across all baseline images in the dataset. Finally, we visualized the results in image intensities following the layout of Fig. A.1 and obtained twelve loss maps, as shown in Fig. A.2. The intensities at each position in the loss maps represent the mean of similarity loss using particular transformations.

To further investigate the sensitivity of the metrics on deformation and contrast, we calculated the normalized gradients over the loss maps and plotted them with red arrows. Specifically, we approximated the spatial gradients using differences between the results of neighboring transformations and displayed them on their corresponding location in the loss maps. The directions of the normalized gradients represent the fastest-growing directions and indicate the sensitivity of different similarity losses on changes in deformation and contrast. For instance, the contrast-insensitive similarity metrics should be more robust to contrast variations, with their gradient directions closer to the direction of increasing deformation. In Fig. A.2, we observed the MSE varied almost entirely with changes in contrast, the LNCC and NMI both performed insensitive to contrast changes in the inversion transform but got worse in the Gamma transform. The SSIM behaved similarly to LNCC and NMI according to the inversion transform, yet it failed in the Gamma transform. The MIND-based similarity metric was also highly sensitive to contrast changes, making it hard to guide the registration model to identify corresponding structures. At last, our proposed mo-MIND-based similarity metric demonstrated contrast-insensitive properties for both Gamma and inversion transform, i.e., much more robust against drastic variations in image contrast.

The assessment method about the sensitivity of similarity losses on deformation and contrast can be explored for other data and may guide the registration tasks for other multi-modal images.

References

[1] Y.-C. Kim, K.R. Kim, K. Choi, M. Kim, Y. Chung, Y.H. Choe, EVCMR: A tool for the quantitative evaluation and visualization of cardiac MRI data, Comput. Biol. Med. 111 (2019) 103334, http://dx.doi.org/10.1016/j.compbiomed.2019.103334, URL: https://www.sciencedirect.com/science/article/pii/S0010482519302033.

[2] T.D. Karamitsos, A. Arvanitaki, H. Karvounis, S. Neubauer, V.M. Ferreira, Myocardial tissue characterization and fibrosis by imaging, JACC: Cardiovasc. Imaging 13 (5) (2020) 1221–1234, http://dx.doi.org/10.1016/j.jcmg.2019.06.030, URL: https://www.sciencedirect.com/science/article/pii/S1936878X19307119.

[3] D.R. Messroghli, J.C. Moon, V.M. Ferreira, L. Grosse-Wortmann, T. He, P. Kellman, J. Mascherbauer, R. Nezafat, M. Salerno, E.B. Schelbert, A.J. Taylor, R. Thompson, M. Ugander, R.B. van Heeswijk, M.G. Friedrich, Clinical recommendations for cardiovascular magnetic resonance mapping of T1, T2, T2* and extracellular volume: A consensus statement by the society for cardiovascular magnetic resonance (SCMR) endorsed by the European association for cardiovascular imaging (EACVI), J. Cardiovasc. Magn. Reson. 19 (1) (2017) 75, http://dx.doi.org/10.1186/s12968-017-0389-8.

[4] P. Haaf, P. Garg, D.R. Messroghli, D.A. Broadbent, J.P. Greenwood, S. Plein, Cardiac T1 mapping and extracellular volume (ECV) in clinical practice: a comprehensive review, J. Cardiovasc. Magn. Reson.: Off. J. Soc. Cardiovasc. Magn. Reson. 18 (1) (2016) 89, http://dx.doi.org/10.1186/s12968-016-0308-4.

[5] D.R. Messroghli, A. Radjenovic, S. Kozerke, D.M. Higgins, M.U. Sivananthan, J.P. Ridgway, Modified Look-Locker inversion recovery (MOLLI) for high-resolution T1 mapping of the heart, Magn. Reson. Med. 52 (1) (2004) 141–146, http://dx.doi.org/10.1002/mrm.20110.

[6] S.K. Piechnik, V.M. Ferreira, E. Dall'Armellina, L.E. Cochlin, A. Greiser, S. Neubauer, M.D. Robson, Shortened modified look-locker inversion recovery (ShMOLLI) for clinical myocardial T1-mapping at 1.5 and 3 T within a 9 heartbeat breathhold, J. Cardiovasc. Magn. Reson.: Off. J. Soc. Cardiovasc. Magn. Reson. 12 (1) (2010) 69, http://dx.doi.org/10.1186/1532-429X-12-69.

[7] K. Chow, J.A. Flewitt, J.D. Green, J.J. Pagano, M.G. Friedrich, R.B. Thompson, Saturation recovery single-shot acquisition (SASHA) for myocardial t(1) mapping, Magn. Reson. Med. 71 (6) (2014) 2082–2095, http://dx.doi.org/10.1002/mrm.24878.

[8] S. Weingärtner, M. Akçakaya, T. Basha, K.V. Kissinger, B. Goddu, S. Berg, W.J. Manning, R. Nezafat, Combined saturation/inversion recovery sequences for improved evaluation of scar and diffuse fibrosis in patients with arrhythmia or heart rate variability, Magn. Reson. Med. 71 (3) (2014) 1024–1034, http://dx.doi.org/10.1002/mrm.24761.

[9] S. Tilborghs, T. Dresselaers, P. Claus, G. Claessen, J. Bogaert, F. Maes, P. Suetens, Robust motion correction for cardiac T1 and ECV mapping using a T1 relaxation model approach, Med. Image Anal. 52 (2019) 212–227, http://dx.doi.org/10.1016/j.media.2018.12.004, URL: https://www.sciencedirect.com/science/article/pii/S136184151830344X.

[10] Y. Li, C. Wu, H. Qi, D. Si, H. Ding, H. Chen, Motion correction for native myocardial T1 mapping using self-supervised deep learning registration with contrast separation, NMR Biomed. 35 (10) (2022) e4775, http://dx.doi.org/10.1002/nbm.4775, URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/nbm.4775. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nbm.4775.

[11] Q. Tao, P. van der Tol, F.F. Berendsen, E.H.M. Paiman, H.J. Lamb, R.J. van der Geest, Robust motion correction for myocardial T1 and extracellular volume mapping by principle component analysis-based groupwise image registration, J. Magn. Reson. Imaging: JMRI 47 (5) (2018) 1397–1405, http://dx.doi.org/10.1002/jmri.25863.

[12] H. Xue, S. Shah, A. Greiser, C. Guetter, A. Littmann, M.-P. Jolly, A.E. Arai, S. Zuehlsdorff, J. Guehring, P. Kellman, Motion correction for myocardial T1 mapping using image registration with synthetic image estimation, Magn. Reson. Med. 67 (6) (2012) 1644–1655, http://dx.doi.org/10.1002/mrm.23153, URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.23153. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.23153.

[13] S. Roujol, M. Foppa, S. Weingärtner, W.J. Manning, R. Nezafat, Adaptive registration of varying contrast-weighted images for improved tissue characterization (ARCTIC): Application to T1 mapping, Magn. Reson. Med. 73 (4) (2015) 1469–1482, http://dx.doi.org/10.1002/mrm.25270, URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.25270. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.25270.

[14] W. Huizinga, D.H.J. Poot, J.M. Guyader, R. Klaassen, B.F. Coolen, M. van Kranenburg, R.J.M. van Geuns, A. Uitterdijk, M. Polfliet, J. Vandemeulebroucke, A. Leemans, W.J. Niessen, S. Klein, PCA-based groupwise image registration for quantitative MRI, Med. Image Anal. 29 (2016) 65–78, http://dx.doi.org/10.1016/j.media.2015.12.004, URL: https://www.sciencedirect.com/science/article/pii/S1361841515001851.

[15] S. Zhang, T.T. Le, S. Kabus, B. Su, D.J. Hausenloy, S.A. Cook, C.W.L. Chin, R.S. Tan, Cardiac magnetic resonance T1 and extracellular volume mapping with motion correction and co-registration based on fast elastic image registration, Magn. Reson. Mater. Phys. Biol. Med. 31 (1) (2018) 115–129, http://dx.doi.org/10.1007/s10334-017-0668-2.

[16] G. Balakrishnan, A. Zhao, M.R. Sabuncu, J. Guttag, A.V. Dalca, VoxelMorph: A learning framework for deformable medical image registration, IEEE Trans. Med. Imaging (2019) http://dx.doi.org/10.1109/TMI.2019.2897538.

[17] R.A. Gonzales, Q. Zhang, B.W. Papież, K. Werys, E. Lukaschuk, I.A. Popescu, M.K. Burrage, M. Shanmuganathan, V.M. Ferreira, S.K. Piechnik, MOCOnet: Robust motion correction of cardiovascular magnetic resonance T1 mapping using convolutional neural networks, Front. Cardiovasc. Med. 8 (2021) http://dx.doi.org/10.3389/fcvm.2021.768245, URL: https://www.frontiersin.org/articles/10.3389/fcvm.2021.768245.

[18] D. Arava, M. Masarwy, S. Khawaled, M. Freiman, Deep-learning based motion correction for myocardial T1 mapping, in: 2021 IEEE International Conference on Microwaves, Antennas, Communications and Electronic Systems, COMCAS, 2021, pp. 55–59, http://dx.doi.org/10.1109/COMCAS52219.2021.9629048, URL: https://ieeexplore.ieee.org/abstract/document/9629048.

[19] B. Kim, I. Han, J.C. Ye, DiffuseMorph: Unsupervised deformable image registration using diffusion model, in: S. Avidan, G. Brostow, M. Cissé, G.M. Farinella, T. Hassner (Eds.), Computer Vision – ECCV 2022, Springer Nature Switzerland, Cham, 2022, pp. 347–364, http://dx.doi.org/10.1007/978-3-031-19821-2_20.

[20] J. Ho, A. Jain, P. Abbeel, Denoising diffusion probabilistic models, 2020, http://dx.doi.org/10.48550/arXiv.2006.11239, URL: http://arxiv.org/abs/2006.11239. arXiv:2006.11239 [cs, stat].

[21] E. Hanania, I. Volovik, L. Barkat, I. Cohen, M. Freiman, PCMC-T1: Free-breathing myocardial T1 mapping with physically-constrained motion correction, in: H. Greenspan, A. Madabhushi, P. Mousavi, S. Salcudean, J. Duncan, T. Syeda-Mahmood, R. Taylor (Eds.), Medical Image Computing and Computer Assisted Intervention – MICCAI 2023, in: Lecture Notes in Computer Science, Springer Nature Switzerland, Cham, 2023, pp. 226–235, http://dx.doi.org/10.1007/978-3-031-43990-2_22.

[22] F. Maes, D. Loeckx, D. Vandermeulen, P. Suetens, Image registration using mutual information, in: N. Paragios, J. Duncan, N. Ayache (Eds.), Handbook of Biomedical Imaging: Methodologies and Clinical Research, Springer US, Boston, MA, 2015, pp. 295–308, http://dx.doi.org/10.1007/978-0-387-09749-7_16.

[23] M.P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F.V. Gleeson, S.M. Brady, J.A. Schnabel, MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration, Med. Image Anal. 16 (7) (2012) 1423–1435, http://dx.doi.org/10.1016/j.media.2012.05.008, URL: https://www.sciencedirect.com/science/article/pii/S1361841512000643.

[24] Y. Wang, T. Fu, C. Wu, J. Xiao, J. Fan, H. Song, P. Liang, J. Yang, Multimodal registration of ultrasound and MR images using weighted self-similarity structure vector, Comput. Biol. Med. 155 (2023) 106661, http://dx.doi.org/10.1016/j.compbiomed.2023.106661, URL: https://www.sciencedirect.com/science/article/pii/S0010482523001269.

[25] Y. Chi, Y. Xu, H. Liu, X. Wu, Z. Liu, J. Mao, G. Xu, W. Huang, A two-step deep learning method for 3DCT-2DUS kidney registration during breathing, Sci. Rep. 13 (1) (2023) 12846, http://dx.doi.org/10.1038/s41598-023-40133-5, URL: https://www.nature.com/articles/s41598-023-40133-5. Number: 1 Publisher: Nature Publishing Group.

[26] Z. Zhou, S. Wang, J. Hu, A. Liu, X. Qian, C. Geng, J. Zheng, G. Chen, J. Ji, Y. Dai, Unsupervised registration for liver CT-MR images based on the multiscale integrated spatial-weight module and dual similarity guidance, Comput. Med. Imaging Graph. 108 (2023) 102260, http://dx.doi.org/10.1016/j.compmedimag.2023.102260, URL: https://www.sciencedirect.com/science/article/pii/S0895611123000782.

[27] A. Buades, B. Coll, J.-M. Morel, A non-local algorithm for image denoising, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'05, Vol. 2, 2005, pp. 60–65, http://dx.doi.org/10.1109/CVPR.2005.38, URL: https://ieeexplore.ieee.org/document/1467423. ISSN: 1063-6919. vol. 2.

[28] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in: N. Navab, J. Hornegger, W.M. Wells, A.F. Frangi (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Springer International Publishing, Cham, 2015, pp. 234–241, http://dx.doi.org/10.1007/978-3-319-24574-4_28.

[29] F. Milletari, N. Navab, S.-A. Ahmadi, V-Net: Fully convolutional neural networks for volumetric medical image segmentation, in: 2016 Fourth International Conference on 3D Vision, 3DV, 2016, pp. 565–571, http://dx.doi.org/10.1109/3DV.2016.79, URL: https://ieeexplore.ieee.org/document/7785132.

[30] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, D. Hawkes, Nonrigid registration using free-form deformations: application to breast MR images, IEEE Trans. Med. Imaging 18 (8) (1999) 712–721, http://dx.doi.org/10.1109/42.796284, URL: https://ieeexplore.ieee.org/document/796284. Conference Name: IEEE Transactions on Medical Imaging.

[31] S.-Y. Lee, K.-Y. Chwa, S.Y. Shin, Image metamorphosis using snakes and free-form deformations, in: Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '95, Association for Computing Machinery, New York, NY, USA, 1995, pp. 439–448, http://dx.doi.org/10.1145/218380.218501, URL: https://dl.acm.org/doi/10.1145/218380.218501.

[32] T. Che, Y. Zheng, X. Sui, Y. Jiang, J. Cong, W. Jiao, B. Zhao, DGR-Net: Deep groupwise registration of multispectral images, in: A.C.S. Chung, J.C. Gee, P.A. Yushkevich, S. Bao (Eds.), Information Processing in Medical Imaging, in: Lecture Notes in Computer Science, Springer International Publishing, Cham, 2019, pp. 706–717, http://dx.doi.org/10.1007/978-3-030-20351-1_55.

[33] P. Kellman, A.E. Arai, H. Xue, T1 and extracellular volume mapping in the heart: estimation of error maps and the influence of noise on precision, J. Cardiovasc. Magn. Reson. 15 (1) (2013) 56, http://dx.doi.org/10.1186/1532-429X-15-56.

[34] Y. Mikami, A. Cornhill, S. Dykstra, A. Satriano, R. Hansen, J. Flewitt, M. Seib, S. Rivest, R. Sandonato, C.P. Lydell, A.G. Howarth, B. Heydari, N. Merchant, N. Fine, J.A. White, Right ventricular insertion site fibrosis in a dilated cardiomyopathy referral population: phenotypic associations and value for the prediction of heart failure admission or death, J. Cardiovasc. Magn. Reson. 23 (1) (2021) 79, http://dx.doi.org/10.1186/s12968-021-00761-0.

[35] S. Klein, M. Staring, K. Murphy, M.A. Viergever, J.P.W. Pluim, Elastix: A toolbox for intensity-based medical image registration, IEEE Trans. Med. Imaging 29 (1) (2010) 196–205, http://dx.doi.org/10.1109/TMI.2009.2035616, URL: https://ieeexplore.ieee.org/document/5338015. Conference Name: IEEE Transactions on Medical Imaging.

[36] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, 2014, http://dx.doi.org/10.48550/arXiv.1412.6980, CoRR URL: https://www.semanticscholar.org/paper/Adam%3A-A-Method-for-Stochastic-Optimization-Kingma-Ba/a6cb366736791bcccc5c8639de5a8f9636bf87e8.