# Set-Based Groupwise Registration for Variable-Length, Variable-Contrast Cardiac MRI

Yi Zhang[1], Yidong Zhao[1], Tijmen Toxopeus[1], Maša Božić-Iven[1], Sebastian Weingärtner[1], and Qian Tao[1]

Department of Imaging Physics, Delft University of Technology, The Netherlands

**Abstract.** Quantitative cardiac magnetic resonance imaging (MRI) enables non-invasive myocardial tissue characterization but relies on robust motion correction within these variable-length, variable-contrast image sequences. Groupwise registration, which simultaneously aligns all images, has shown greater robustness than pairwise registration for motion correction. However, current deep-learning-based groupwise registration methods cannot generalize across MRI sequences: the architecture typically encodes input data as a fixed-length channel stack, which rigidly couples network design to protocol-specific sequence length, input ordering, and contrast dynamics. At inference time, any change in imaging protocols will render the network unusable. In this work, we introduce *Any$^2$Reg*, a new set-based groupwise registration framework that takes a quantitative MRI sequence as an unordered set. This set formulation fundamentally decouples network design from sequence length and input ordering. By utilizing a shared encoder and correlation-guided feature aggregation, *Any$^2$Reg* constructs a permutation-invariant canonical reference for registration, and learns a permutation-equivariant mapping from images to deformation fields. Additionally, we extract contrast-insensitive image features from an existing foundation model to handle extreme contrast variations. Trained exclusively on a single public $T_1$ mapping dataset (STONE, sequence length $L = 11$), Any$^2$Reg generalizes to two unseen quantitative MRI datasets (MOLLI, ASL) with variable lengths ($L \in [11, 60]$) and different contrast dynamics. It achieves strong cross-protocol generalization in a zero-shot manner, and consistently improves downstream quantitative mapping quality. Notably, while designed for quantitative MRI sequences, our framework is directly applicable to Cine MRI sequences for inter-cardiac-phase registration.

**Keywords:** Cardiac MRI · Image Registration · Set-based Methods

## 1 Introduction

Cardiac magnetic resonance imaging (MRI) is essential for diagnosing and monitoring cardiovascular diseases. Modern examinations typically include quantitative mapping, such as the widely used $T_1$ mapping [6,4,22,5], for non-invasive myocardial tissue assessment. Quantitative sequences estimate parameters by fitting signal models voxel-wise, assuming strict anatomical alignment across images.

However, cardiac and respiratory motion cause misalignment, introducing spatial blurring, biased parameter estimates, and increased map uncertainty [14,15]. Motion correction is a prerequisite for reliable quantitative cardiac MRI [25,27].

Quantitative cardiac MRI sequences are challenging to register. Their imaging protocols have variable sequence lengths $L$, specific sampling orders, and different contrast evolution governed by different physical processes [19,8]. Inter-frame differences reflect an entangled combination of true motion and physiological signal evolution. This poses significant challenges for traditional pairwise image registration, which requires a predefined reference image and fails for poor-contrast images [2,18]. Therefore, groupwise registration is methodologically preferred, which aligns all images simultaneously to an implicit reference [12,29,28]. Classical groupwise registration performs optimization per sequence [12], and is computationally expensive. Importantly, the optimization scales poorly to sequence length and can take minutes to hours for long sequences.

Deep learning provides an amortized alternative to per-sequence optimization, significantly accelerating inference [3]. Recent works leverage deep learning to speed up groupwise registration for practical use [3,20,29,18]. However, existing works predominantly use U-Net-style backbones [24], which hardcode the input image sequence as a fixed-length, ordered channel stack [17,28,9,10]. Such hard-coding of sequence length and order imposes a rigid prior, precluding generalization to other quantitative sequences with different lengths or contrasts. Though some recent works have explored variable-size group modelling for inter-subject brain MRI atlas construction [11,1], they primarily focus on homogeneous image contrasts and cannot address the dramatic contrast changes in cardiac MRI (e.g. contrast nulling, flipping, etc. as shown in Fig. 1).

In principle, a quantitative MRI acquisition can be viewed as an observation set, each element being an image, sampled from a physical process (e.g., $T_1$ relaxation). Quantitative mapping operates on the set, where standard fitting (least-squares minimization) is invariant to permutations in the set. Ideally, motion correction models should also respect this set symmetry, but most groupwise registration methods use fixed-channel models [17,28,9,10], where the sequence length and frame order critically determine learning, contradicting set symmetry.

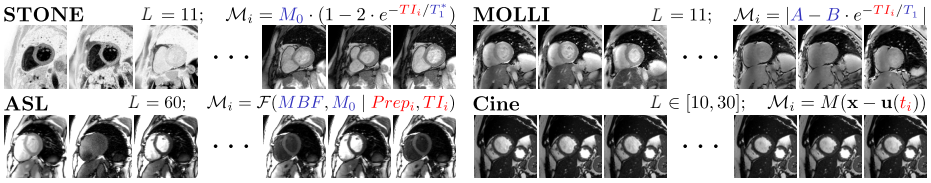


**Fig. 1.** Cardiac MRI sequences exhibit high heterogeneity across imaging protocols with different mechanisms ($\mathcal{M}_i$) and sequence lengths ($L$). Analytical models are expressed by external parameters (red; e.g., inversion times $TI_i$, preparation $Prep_i$, or phases $t_i$) and internal properties (dark blue; e.g., $T_1, T_1^*$, $MBF$, $M_0$, $A$, or $B$).

Motivated by this critical gap, we propose Any$^2$Reg, a set-based groupwise registration framework that treats cardiac MRI sequences as unordered sets. Instead of fixed-length channel stacking, Any$^2$Reg uses a shared encoder to extract features per image. It uses correlation-guided aggregation to construct a canonical reference, which is strictly permutation-invariant. Any$^2$Reg enables a new way of groupwise registration, generalizable to variable-length, variable contrast cardiac MRI sequences. Our contributions are threefold:

- We introduce Any$^2$Reg, a set-based groupwise registration architecture that models quantitative MRI sequences as unordered sets. Any$^2$Reg fundamentally decouples network design from sequence length $L$ and acquisition order, enabling highly efficient and scalable registration of variable-length sequences with near-linear time complexity *w.r.t.* $L$ in practice.
- We propose a correlation-guided aggregation module to construct a canonical reference, and an optional module for contrast-insensitive feature fusion. This makes Any$^2$Reg applicable to variable-contrast sequences.
- We demonstrate that Any$^2$Reg, trained exclusively on a public $T_1$ database, achieves strong zero-shot generalization on three unseen cardiac MRI datasets with varying sequence lengths and image contrasts. It consistently outperforms state-of-the-art baselines in registration and downstream mapping.

## 2 Method

### 2.1 Problem formulation and set symmetry for registration

Let $\Omega \subset \mathbb{R}^2$ denote the image domain. A quantitative cardiac MRI baseline sequence consists of $L$ frames acquired under different settings, represented as a set $\mathcal{I} = \{I_1, \dots, I_L\}$ with $I_i : \Omega \to \mathbb{R}$. Groupwise registration estimates dense displacement fields $\{\mathbf{u}_i\}_{i=1}^{L}$ and transformations $\phi_i : \Omega \to \Omega$ defined by $\phi_i(x) = x + \mathbf{u}_i(x)$, $x \in \Omega$, mapping all frames to a shared canonical coordinate system. Let $\Phi = \{\phi_1, \dots, \phi_L\}$ and $\hat{I}_i = I_i \circ \phi_i$ denote the transformation set and aligned images, respectively.

The alignment facilitates downstream voxel-wise quantitative mapping. Under ideal alignment, tissue properties $\psi(x)$ at location $x$ are estimated by

$$\hat{\psi}(x) = \arg\min_{\psi} \sum_{i=1}^{L} \ell\big(\hat{I}_i(x),\, \mathcal{M}_i(\psi(x))\big), \quad x \in \Omega, \tag{1}$$

where $\mathcal{M}_i$ is the signal model for the $i$-th frame under its acquisition settings (Fig. 1), and $\ell$ is a pointwise loss. Eq. 1 suggests two structural requirements for registration: it should generalize across varying $L$, and respect permutation symmetry over frame indices $\{1, \dots, L\}$.

### 2.2 Permutation-equivariant registration network

To handle varying $L$ and frame-index symmetry, the registration operator $\mathcal{R} : \mathcal{I} \mapsto \Phi$ should be permutation-equivariant:

$$\mathcal{R}(\pi \cdot \mathcal{I}) = \pi \cdot \mathcal{R}(\mathcal{I}), \qquad \forall \pi \in S_L. \tag{2}$$

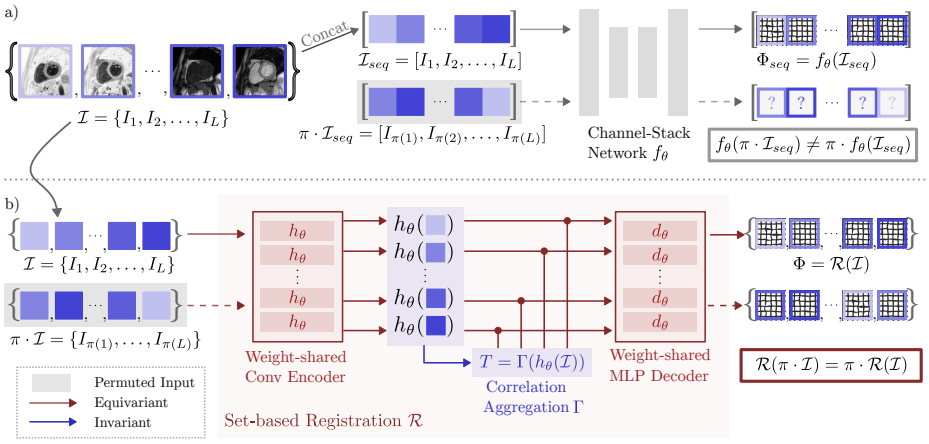


**Fig. 2.** Comparison of two learning-based groupwise registration designs: (a) The conventional channel stacking design concatenates frames; it is order-sensitive ($f_\theta(\pi \cdot \mathcal{I}_{seq}) \neq \pi \cdot f_\theta(\mathcal{I}_{seq})$) and tied to a fixed length $L$. (b) Set-based design (Any$^2$Reg) uses shared encoders and a canonical reference $T = \Gamma(\mathcal{I})$ broadcasted to all frames; it enables permutation-equivariant registration ($\mathcal{R}(\pi \cdot \mathcal{I}) = \pi \cdot \mathcal{R}(\mathcal{I})$) for an arbitrary $L$.

That is, permuting the input frames should only permute the predicted transformations. In groupwise registration, this calls for an order-invariant canonical context to which all frames align. We enforce this with an invariant-to-equivariant design: a set operator $\Gamma$ constructs a reference, $\Gamma(\pi \cdot \mathcal{I}) = \Gamma(\mathcal{I})$, which is broadcast to all framewise weight-shared branches, thus permutation-equivariant.

Based on this principle, we propose Any$^2$Reg (Fig. 2b), a set-based multi-scale registration network with $K$ scales ($k = 1$ the finest). At scale $k$, a frame-shared encoder extracts per-frame features $\{z_i^{(k)}\}_{i=1}^{L}$. An invariant operator $\Gamma^{(k)}$ aggregates a canonical feature $T^{(k)}$ from $\{z_i^{(k)}\}$ and broadcasts it to all frames. The concatenated features $[z_i^{(k)}, T^{(k)}]$ update the next-scale features and drive a shared CorrMLP-style decoder [21] to predict incremental transformations $\phi_i^{(k)}$. The final transformation composes predictions from coarse to fine:

$$\phi_i = \phi_i^{(1)} \circ \uparrow\big(\phi_i^{(2)}\big) \circ \cdots \circ \uparrow\big(\phi_i^{(K)}\big), \tag{3}$$

where $\uparrow(\cdot)$ upsamples to the finest grid. Sec. 2.3 details our instantiation of $\Gamma^{(k)}$.

### 2.3 Correlation-guided set aggregation

We instantiate the scale-specific invariant operator $\Gamma^{(k)}$ using a correlation-guided principal component (PC) aggregation as in [12,23].

Given scale-$k$ features $z_i^{(k)} \in \mathbb{R}^{C_k \times H_k \times W_k}$, we form vectorized $\ell_2$-normalized representations $\tilde{z}_i^{(k)}$ and compute the frame-wise correlation matrix $\mathbf{C}^{(k)} \in$

$\mathbb{R}^{L\times L}$. Let $v^{(k)}$ be the leading eigenvector (first PC) of $\mathbf{C}^{(k)}$, which captures the principal mode of variation across the sequence. We then aggregate the canonical feature using the entries $v_i^{(k)}$ of the $v^{(k)}$:

$$T^{(k)} = \sum_{i=1}^{L} w_i^{(k)} z_i^{(k)}, \qquad w_i^{(k)} = \frac{v_i^{(k)}}{\sum_{j=1}^{L} v_j^{(k)}}. \tag{4}$$

Under a frame-wise index permutation, $v_i^{(k)}$ (and thus $w_i^{(k)}$) permutes accordingly. Therefore, with a deterministic eigenvector sign convention of $\sum_{j=1}^{L} v_j^{(k)} > 0$ is permutation-invariant, leading to $\Gamma^{(k)}\Big(\{z_i^{(k)}\}_{i=1}^{L}\Big)$.

To inject this canonical context back to each frame, we broadcast $T^{(k)}$ and apply a shared update:

$$z_i^{(k+1)} = \sigma\Big(\mathrm{Conv}^{(k)}\Big([z_i^{(k)}, T^{(k)}]\Big)\Big), \tag{5}$$

which preserves permutation equivariance since $T^{(k)}$ is shared across frames and $\mathrm{Conv}^{(k)}$ uses shared weights. The set aggregation operates on a relatively small $L \times L$ matrix, and therefore adds negligible overhead compared with per-frame convolutions. This is further empirically validated in Sec. 4.

### 2.4 Auxiliary anatomy-driven feature stream

Anatomy-driven, contrast-agnostic features have been shown to improve learning-based registration [3,9]; when available from a pretrained model (e.g., a segmentation foundation model), they can be optionally incorporated as an auxiliary stream. Let $g(\cdot)$ be a frozen feature extractor. For each frame $I_i$, we extract

$$F_i = g(I_i), \quad F_i \in \mathbb{R}^{C_f \times H_f \times W_f}. \tag{6}$$

Applying $g$ independently to each frame yields a feature set $\mathcal{F} = \{F_1, \ldots, F_L\}$ that preserves the set structure and permutation equivariance.

The auxiliary stream is encoded in parallel by a frame-shared multi-scale encoder $h_\theta'^{(k)}$. At each scale, we fuse pretrained features with image-branch features via a lightweight convolution before applying $\Gamma^{(k)}$. This leads to set aggregation without altering the permutation-equivariant formulation.

### 2.5 Loss functions

Using the final warped images $\hat{I}_i = I_i \circ \phi_i$, a correlation-weighted template is constructed as $T_I = \sum_{i=1}^{L} w_i \hat{I}_i$ where weights $\{w_i\}$ are computed via the correlation-guided weighting scheme (Sec. 2.3). We use conditional template entropy (CTE) [23] for multi-contrast registration as the dissimilarity function:

$$\mathcal{L}_{\mathrm{CTE}}(\hat{\mathcal{I}}, T_I) = \frac{1}{L}\sum_{i=1}^{L} H\Big(T_I \mid \hat{I}_i\Big), \qquad H(T \mid I) = H(T, I) - H(I). \tag{7}$$

When the auxiliary feature stream is utilized, the warped feature maps $\hat{F}_i = F_i \circ \phi_i$ yield the feature template $T_F = \sum_i w_{f,i} \hat{F}_i$. The weights $\{w_{f,i}\}$ are independently computed via correlation in the feature space. The total loss is:

$$\mathcal{L} = \mathcal{L}_{\text{CTE}}(\hat{\mathcal{I}}, T_I) + \lambda_f \mathcal{L}_{\text{CTE}}(\hat{\mathcal{F}}, T_F) + \lambda_s \mathcal{L}_{\text{smooth}}(\Phi) + \lambda_c \mathcal{L}_{\text{cyclic}}(\Phi), \quad (8)$$

where $\mathcal{L}_{\text{smooth}}$ penalizes displacement gradients to ensure deformation regularity and $\mathcal{L}_{\text{cyclic}}$ [29] suppresses non-zero group motion to prevent spatial drift.

## 3 Experiments

**Data.** We evaluate four heterogeneous cardiac MRI datasets comprising 2D short-axis slices (Fig. 1). In-domain training uses the public STONE dataset (free-breathing slice-interleaved $T_1$ mapping; 210 subjects; 1050 slices, 150/30/30 train/val/test subject-wise split, $L = 11$) [26,7]. For zero-shot evaluation, we use an in-house MOLLI dataset (modified Look-Locker $T_1$ mapping; 10 subjects, 60 slices, pre-/post-contrast, $L = 11$) [22] and an in-house ASL dataset (myocardial arterial spin labeling; 6 subjects; $L = 60$ via 12 images$\times$5 repetitions) [5]. Additionally, we include the public ACDC dataset (bSSFP Cine; test split; 25 subjects, 256 slices; $L \in [10, 30]$) [4], which is not a quantitative sequence and contains regular cardiac motion, for an out-of-distribution stress test.

**Evaluation metrics.** We report slice-wise all-pair Dice for left ventricle (LV) and myocardium (Myo) on STONE (expert annotations), and for LV, right ventricle (RV), and Myo on MOLLI, ASL, and Cine using ACDC-trained nnU-Net predictions [13] independent of the feature stream. We manually annotate RV insertion points on STONE and MOLLI to measure target registration error (TRE). On STONE, we further assess $T_1$ mapping quality via per-pixel curve-fitting $R^2$ and deformation regularity via std($\log \det J$).

**Baselines and ablations.** We compare with Elastix [16] (optimization-based; PCA2 similarity [12]; B-spline; 4-pixel spacing), GroupRegNet [29] (channel-stacked, one-shot), PCA-Relax [28] (channel-stacked learning with PCA2), and MultiMorph [1] adapted to groupwise registration, including its segmentation-guided variant (MultiMorph-Seg). We also evaluate Any$^2$Reg without foundation model features, replace correlation with mean aggregation for ablation. We include an instance-optimization (IO) [28] variant that refines test cases for 30 inference steps.

**Implementation.** Images are center-cropped to $192 \times 192$. The encoder utilizes $K = 4$ scales with 16 channels. Auxiliary anatomical features are derived from a pretrained Reverse Imaging segmentation foundation model [30] trained on the ACDC train split. For MultiMorph-Seg, official STONE annotations are used. Models are trained for 50 epochs via AdamW (lr $10^{-4}$) with $\lambda_f = 0.5$, $\lambda_s = 10$, and $\lambda_c = 0.05$. For fair comparison, MultiMorph optimizes the same CTE loss; all hyperparameters are tuned on the STONE validation with comparable std($\log \det J$) and folding ratios $< 10^{-5}$. One-shot GroupRegNet optimizes from random initialization for 250 steps. Any$^2$Reg IO refines for 30 steps (lr $3 \times 10^{-5}$). Implementation used PyTorch 2.8.0 on an NVIDIA RTX5090. Code and weights are available at `https://github.com/YiZhang025/Any2Reg`.

**Table 1.** Main quantitative results for in-domain (STONE) and three cardiac MRI datasets. $\sigma_{\log|J|} \equiv \text{std}(\log\det J) \times 10^{-2}$. **-** indicates not applicable due to sequence length constraints; **Best**/Second mark the best and second-best results.

| Method | STONE | | | MOLLI | | ASL | Cine |
|---|---|---|---|---|---|---|---|
| | Dice ↑ | TRE ↓ | $\sigma_{\log\|J\|}$ ↓ | Dice ↑ | TRE ↓ | Dice ↑ | Dice ↑ |
| Raw | 78.8±7.1 | 4.19±1.26 | / | 78.6±14.8 | 3.16±1.50 | 76.7±7.8 | 79.1±4.8 |
| Elastix | 85.0±5.4 | 3.46±1.08 | 12.7±4.5 | 79.2±14.3 | 3.10±1.32 | 79.4±8.1 | 86.8±3.8 |
| GroupRegNet | 83.0±6.2 | 3.60±1.04 | 9.3±3.5 | 79.7±14.6 | 3.09±1.36 | 78.6±8.5 | 80.5±4.7 |
| PCA-Relax | 80.6±6.6 | 4.01±1.16 | 14.0±4.7 | 78.2±14.8 | 3.13±1.49 | - | - |
| MultiMorph | 83.1±6.4 | 3.77±1.15 | 14.7±1.9 | 78.4±14.3 | 3.11±1.39 | 75.2±5.7 | 84.8±4.3 |
| MultiMorph-Seg | 85.0±5.9 | 3.71±1.17 | 15.2±2.2 | 78.8±14.1 | 3.18±1.36 | 76.9±6.4 | 86.1±3.9 |
| Any$^2$Reg w/o FM | 84.2±5.9 | 3.65±1.11 | **8.3±1.4** | 78.9±14.5 | 3.17±1.43 | 79.0±7.2 | 89.5±2.7 |
| Any$^2$Reg mean | 86.4±4.8 | 3.23±0.86 | 14.2±1.6 | 84.7±14.7 | 3.02±0.97 | 80.3±7.7 | 89.9±2.5 |
| Any$^2$Reg | 86.9±4.7 | 3.18±0.88 | 12.3±3.0 | 84.8±14.8 | **2.96±0.96** | 80.5±7.9 | 89.8±2.5 |
| Any$^2$Reg IO | **86.9±4.4** | **3.14±0.83** | 13.7±3.1 | **84.9±15.0** | 2.98±0.95 | **80.6±7.6** | **90.9±2.1** |

## 4 Results and Discussion

**Main performance.** Table 1 presents quantitative results for in-domain and zero-shot cross-protocol scenarios. On STONE, Any$^2$Reg+IO achieves the highest Dice and lowest TRE. The default Any$^2$Reg consistently outperforms U-Net baselines across all datasets. Ablation on aggregation shows Any$^2$Reg surpasses mean aggregation (Any$^2$Reg mean) in quantitative MRI datasets, validating that correlation yields a more informative canonical representation under contrast variation. Any$^2$Reg w/o FM also outperforms baselines in zero-shot cases; this confirms that set-based aggregation improves generalization independently of the anatomy-driven features. Strong zero-shot performance on Cine shows robust generalization to a completely different spatiotemporal sequence. In contrast, fixed-length channel-stacked method PCA-Relax fails on ASL and Cine due to sequence-length constraints.

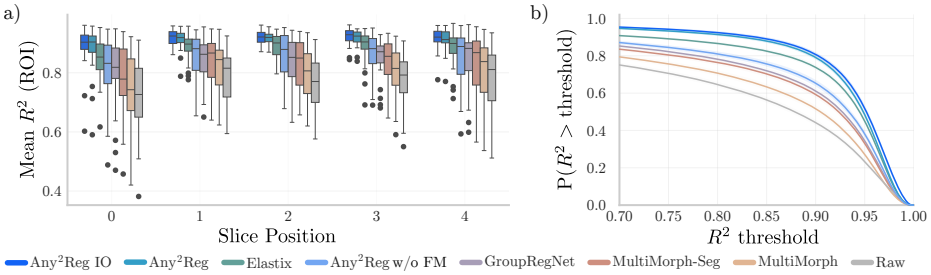


**Fig. 3.** $T_1$ mapping quality on STONE (LV+Myo). (a) Slice-wise mean $R^2$. Slice position: base (0) to apex (4). Any$^2$Reg IO demonstrates the best alignment at all slice positions. (b) $R^2$ survival curves show Any$^2$Reg has consistently higher fitting quality as a result of superior alignment.

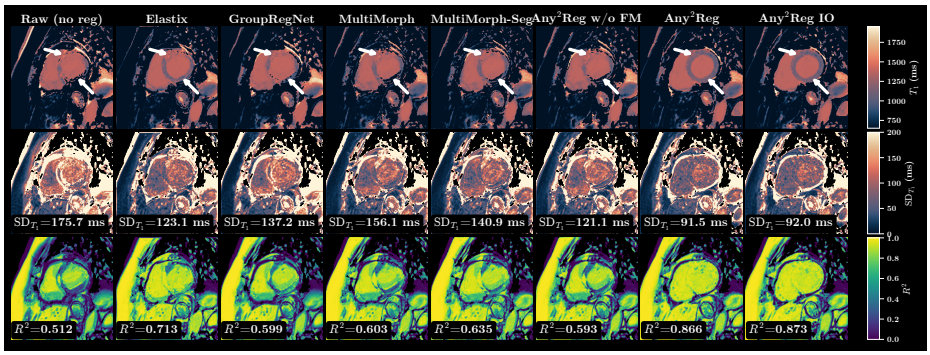


**Fig. 4.** Qualitative result on a representative STONE sequence with $T_1$ map, fitting uncertainty $\mathrm{SD}_{T_1}$ (↓), and pixelwise fitting $R^2$ (↑). White arrows indicate the location of large motion. Both Any$^2$Reg and Any$^2$Reg IO yield clear $T_1$ maps with high precision.

**Scalability analysis.** We evaluated Any$^2$Reg's scalability up to $L = 512$. Any$^2$Reg inference scales linearly (∼2.4 ms per additional image), taking ∼0.2 s at $L = 64$ and ∼1.2 s at $L = 512$. The correlation aggregation GPU runtime overhead remains negligible ($< 4\%$). The 30-step IO variant also scales linearly (∼12 s at $L = 64$). Conversely, conventional method (Elastix) grows superlinearly (∼70 s at $L = 64$; ∼260 s at $L = 128$).

**Quantitative Mapping.** Figure 3 illustrates the benefits for downstream $T_1$ mapping, wherein Any$^2$Reg maintains the highest $R^2$ survival rates. The most substantial margin is observed in the base segment, where motion degrades alignment. Qualitatively, as shown in Figure 4, imprecise registration mixes the signal contributions from adjacent tissues across frames. This degrades the quality of fit, resulting in reduced $R^2$ values and high $T_1$ fitting uncertainty ($\mathrm{SD}_{T_1}$) [14]. Any$^2$Reg effectively resolves these motion artifacts, achieving a superior fit ($R^2 = 0.866$, $\mathrm{SD}_{T_1} = 91.5\,\mathrm{ms}$) in high precision (low $\mathrm{SD}_{T_1}$) $T_1$ maps.

## 5 Conclusion

We propose a new way of groupwise registration, Any$^2$Reg, which can operate on variable-length, variable-contrast image sequences of cardiac MRI. Any$^2$Reg reformulates groupwise registration as a set-to-set prediction problem, and completely eliminates the rigid coupling of sequence length and image order in conventional channel-stacked architectures. Trained solely on one quantitative MRI sequence, Any$^2$Reg achieves robust zero-shot generalization across heterogeneous MRI protocols, including unseen quantitative sequences and spatiotemporal Cine sequences. It improves spatial alignment and downstream quantitative mapping over state-of-the-art baselines while maintaining speed advantage. Any$^2$Reg provides a highly scalable and robust solution for registering cardiac MRI, and its generic setup enables extension to other medical image registration applications.

## References


1. Abulnaga, S.M., Hoopes, A., Dey, N., Hoffmann, M., Fischl, B., Guttag, J., Dalca, A.: Multimorph: On-demand atlas construction. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 30906–30917 (2025)
2. Arava, D., Masarwy, M., Khawaled, S., Freiman, M.: Deep-learning based motion correction for myocardial t1 mapping. In: IEEE International Conference on Microwaves, Antennas, Communications and Electronic Systems. pp. 55–59 (2021)
3. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: Voxelmorph: a learning framework for deformable medical image registration. IEEE transactions on medical imaging **38**(8), 1788–1800 (2019)
4. Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P.A., Cetin, I., Lekadir, K., Camara, O., Ballester, M.A.G., et al.: Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE transactions on medical imaging **37**(11), 2514–2525 (2018)
5. Božić-Iven, M., Rapacchi, S., Tao, Q., Pierce, I., Thornton, G., Nitsche, C., Treibel, T.A., Schad, L.R., Weingärtner, S.: Improved reproducibility for myocardial asl: Impact of physiological and acquisition parameters. Magnetic Resonance in Medicine **91**(1), 118–132 (2024)
6. Campello, V.M., Gkontra, P., Izquierdo, C., Martin-Isla, C., Sojoudi, A., Full, P.M., Maier-Hein, K., Zhang, Y., He, Z., Ma, J., et al.: Multi-centre, multi-vendor and multi-disease cardiac segmentation: the m&ms challenge. IEEE Transactions on Medical Imaging **40**(12), 3543–3554 (2021)
7. El-Rewaidy, H., Nezafat, M., Jang, J., Nakamori, S., Fahmy, A.S., Nezafat, R.: Nonrigid active shape model–based registration framework for motion correction of cardiac t1 mapping. Magnetic resonance in medicine **80**(2), 780–791 (2018)
8. Haaf, P., Garg, P., Messroghli, D.R., Broadbent, D.A., Greenwood, J.P., Plein, S.: Cardiac t1 mapping and extracellular volume (ecv) in clinical practice: a comprehensive review. Journal of Cardiovascular Magnetic Resonance **18**(1), 89 (2016)
9. Hanania, E., Volovik, I., Barkat, L., Cohen, I., Freiman, M.: Pcmc-t1: Free-breathing myocardial t1 mapping with physically-constrained motion correction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 226–235. Springer (2023)
10. Hanania, E., Zehavi-Lenz, A., Volovik, I., Link-Sourani, D., Cohen, I., Freiman, M.: Mbss-t1: Model-based subject-specific self-supervised motion correction for robust cardiac t1 mapping. Medical Image Analysis **102**, 103495 (2025)
11. He, Z., Chung, A.C.: Instantgroup: Instant template generation for scalable group of brain mri registration. IEEE Transactions on Image Processing (2025)
12. Huizinga, W., Poot, D.H., Guyader, J.M., Klaassen, R., Coolen, B.F., van Kranenburg, M., Van Geuns, R., Uitterdijk, A., Polfliet, M., Vandemeulebroucke, J., et al.: Pca-based groupwise image registration for quantitative mri. Medical image analysis **29**, 65–78 (2016)
13. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)
14. Kellman, P., Arai, A.E., Xue, H.: T1 and extracellular volume mapping in the heart: estimation of error maps and the influence of noise on precision. Journal of Cardiovascular Magnetic Resonance **15**(1), 1–12 (2013)
15. Kellman, P., Hansen, M.S.: T1-mapping in the heart: accuracy and precision. Journal of cardiovascular magnetic resonance **16**, 1–20 (2014)

16. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.: Elastix: a toolbox for intensity-based medical image registration. IEEE transactions on medical imaging **29**(1), 196–205 (2009)
17. Li, X., Zhang, Y., Zhao, Y., van Gemert, J., Tao, Q.: Contrast-agnostic groupwise registration by robust pca for quantitative cardiac mri. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 77–87 (2023)
18. Li, Y., Wu, C., Qi, H., Si, D., Ding, H., Chen, H.: Motion correction for native myocardial t1 mapping using self-supervised deep learning registration with contrast separation. NMR in Biomedicine **35**(10), e4775 (2022)
19. Makela, T., Clarysse, P., Sipila, O., Pauna, N., Pham, Q.C., Katila, T., Magnin, I.E.: A review of cardiac image registration methods. IEEE Transactions on medical imaging **21**(9), 1011–1021 (2002)
20. Martín-González, E., Sevilla, T., Revilla-Orodea, A., Casaseca-de-la Higuera, P., Alberola-López, C.: Groupwise non-rigid registration with deep learning: an affordable solution applied to 2d cardiac cine mri reconstruction. Entropy **22**(6), 687 (2020)
21. Meng, M., Feng, D., Bi, L., Kim, J.: Correlation-aware coarse-to-fine mlps for deformable medical image registration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9645–9654 (2024)
22. Messroghli, D.R., Radjenovic, A., Kozerke, S., Higgins, D.M., Sivananthan, M.U., Ridgway, J.P.: Modified look-locker inversion recovery (molli) for high-resolution t1 mapping of the heart. Magnetic Resonance in Medicine **52**(1), 141–146 (2004)
23. Polfliet, M., Klein, S., Huizinga, W., Paulides, M.M., Niessen, W.J., Vandemeulebroucke, J.: Intrasubject multimodal groupwise registration with the conditional template entropy. Medical image analysis **46**, 15–25 (2018)
24. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention, October 5-9, 2015, Part III 18. pp. 234–241. Springer (2015)
25. Tao, Q., van der Tol, P., Berendsen, F.F., Paiman, E.H., Lamb, H.J., van der Geest, R.J.: Robust motion correction for myocardial t1 and extracellular volume mapping by principle component analysis-based groupwise image registration. Journal of Magnetic Resonance Imaging **47**(5), 1397–1405 (2018)
26. Weingärtner, S., Roujol, S., Akçakaya, M., Basha, T.A., Nezafat, R.: Free-breathing multislice native myocardial t1 mapping using the slice-interleaved t1 (stone) sequence. Magnetic resonance in medicine **74**(1), 115–124 (2015)
27. Xue, H., Shah, S., Greiser, A., Guetter, C., Littmann, A., Jolly, M.P., Arai, A.E., Zuehlsdorff, S., Guehring, J., Kellman, P.: Motion correction for myocardial t1 mapping using image registration with synthetic image estimation. Magnetic resonance in medicine **67**(6), 1644–1655 (2012)
28. Zhang, Y., Zhao, Y., Huang, L., Xia, L., Tao, Q.: Deep-learning-based groupwise registration for motion correction of cardiac t 1 mapping. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 586–596. Springer (2024)
29. Zhang, Y., Wu, X., Gach, H.M., Li, H., Yang, D.: Groupregnet: a groupwise one-shot deep learning-based 4d image registration method. Physics in Medicine & Biology **66**(4), 045030 (2021)
30. Zhao, Y., Zhang, Y., Simonetti, O., Han, Y., Tao, Q.: Reverse imaging for wide-spectrum generalization of cardiac mri segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer (2025)