Neural Cardiac Representations for Motion-Resolved Super-Resolution and Segmentation

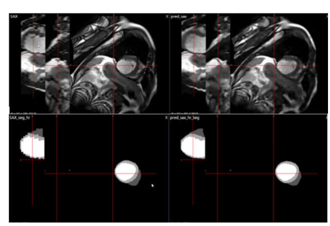
Phanie D Negho, Nora Vogt, Ziad Al-Haj Hemidi, Mattias P Heinrich, Julien Oster

IADI U1254, Inserm, Université de Lorraine, Nancy, France.

Aims: This work aims to jointly optimize super-resolution, inter-slice motion correction, and segmentation to reconstruct coherent high-resolution 3D cardiac volumes from sparsely sampled, anisotropic, and misaligned 2D slices. Building on our previous work demonstrating accurate motion correction and reconstructions for phantom studies, we extend the approach to in-vivo MRI and cardiac segmentation.

Methods: The proposed implicit neural representation (INR) model learns a continuous volumetric representation by aligning input coordinates (x, y, z) of short-axis (SAX) and long-axis (LAX) slices using learnable transformations (rotation: (r_x, r_y, r_z); translation: (t_x, t_y, t_z)). The transformed coordinates are processed by a SIREN network with two prediction heads, yielding image reconstructions and segmentation maps. The model consists of four hidden layers, each containing 128 neurons, and was trained using the AdamW optimizer for 5000 iterations using the following objective function: $L_{\text{Total}} = L_{\text{MSE}}^{\text{SAX}} + L_{\text{MSE}}^{\text{LAX1}} + L_{\text{MSE}}^{\text{LAX2}} + L_{\text{MSE}}^{\text{LAX3}} + \lambda L_{\text{CE}}^{\text{SAX}}$. We tested the method on real data from a single patient consisting of 13 SAX slices and three LAX slices from different views, with 1mm in-plane and 8mm through-plane resolution. While all slices contribute to optimizing transformation parameters and the reconstruction head, only SAX view segmentation masks are used for supervision of the segmentation head. At inference time, predictions are generated for an isotropic, high-resolution 3D coordinate grid (1mm) using the SIREN network.

Results: Qualitative results demonstrate high-quality isotropic 3D reconstructions with effective inter-slice motion correction, outperforming simple interpolation methods in visual smoothness.



Left: interpolation, right: model prediction.

Conclusion: The proposed multi-task framework achieves promising 3D spatial consistency and anatomical structure preservation. Future work will explore temporal (4D) modeling and quantitative assessment of reconstruction quality (e.g., Sliced Wasserstein Distance).