

Biophysics-Enhanced Neural Representations for Patient-Specific Respiratory Motion Modeling

Jan Boysen ¹, Hristina Uzunova ³, Heinz Handels ^{1,2}, Jan Ehrhardt ^{1,2}

¹ German Research Center for Artificial Intelligence (DFKI), Lübeck, DE

² Institute of Medical Informatics, University of Lübeck, Lübeck, DE

³ Clinic for Orthopedics and Orthopedic Surgery, University Medicine Greifswald, Greifswald, DE

Abstract

A precise spatial delivery of the radiation dose is crucial for the treatment success in radiotherapy. In the lung and upper abdominal region, respiratory motion introduces significant treatment uncertainties, requiring special motion management techniques. To address this, respiratory motion models are commonly used to infer the patient-specific respiratory motion and target the dose more efficiently. In this work, we investigate the possibility of using implicit neural representations (INR) for surrogate-based motion modeling. Therefore, we propose physics-regularized implicit surrogate-based modeling for respiratory motion (PRISM-RM). Our new integrated respiratory motion model is free of a fixed reference breathing state. Unlike conventional pairwise registration techniques, our approach provides a trajectory-aware spatio-temporally continuous and diffeomorphic motion representation, improving generalization to extrapolation scenarios. We introduce biophysical constraints, ensuring physiologically plausible motion estimation across time beyond the training data. Our results show that our trajectory-aware approach performs on par in interpolation and improves the extrapolation ability compared to our initially proposed INR-based approach. Compared to sequential registration-based approaches both our approaches perform equally well in interpolation, but underperform in extrapolation scenarios. However, the methodical features of INRs make them particularly effective for respiratory motion modeling, and with their performance steadily improving, they demonstrate strong potential for advancing this field.

Keywords

Physics-Informed Machine Learning, Implicit Neural Representations, Registration, Respiratory Motion Modeling

Article informations

<https://doi.org/10.59275/j.melba.2026-1ba1>

©2026 Boysen et al.. License: CC-BY 4.0

Volume 2026, Received: 2025-07, Published 2026-03

Corresponding author: jan.boysen@dfki.de

Special issue: MELBA-BVM 2025 Special Issue

Guest editors: Andreas Maier, Thomas Deserno, Heinz Handels, Klaus Maier-Hein, Christoph Palm, Thomas Tolxdorff, Katharina Breininger



1. Introduction

The aim of radiotherapy is to accurately deliver a high dose of radiation to the tumor while sparing as much healthy tissue as possible. In tumor treatment in the lung and upper abdomen, respiratory motion is one of the main sources of error, and respiratory motion management such as respiratory gating or tracking is required to compensate for tumor motion during treatment (Dhont et al., 2020). The ideal scenario for managing respiratory motion is to directly observe and track the tumor and organs at risk (OARs) during radiation therapy. However, achieving this requires suitable imaging technologies integrated into the linear accelerator that either provide

functional information, such as PET (Fiorina et al., 2021), or high soft tissue contrast, such as MRI (Cerviño et al., 2011). The clinical application of techniques like MRI-guided radiation therapy (MRIgRT) is currently limited by the need for additional infrastructure, increased treatment and personnel costs, and the lack of medium- to long-term clinical outcome data (Keall et al., 2022). Less complex and more cost-effective linac-integrated imaging techniques such as MV/kV imaging enable sufficiently reliable and robust marker-less tracking only for specific cases and tumor locations (Zhang et al., 2018). An alternative to direct imaging and tracking of tumor and OARs is indirect monitoring using internal or external surrogate signals. Examples of

internal surrogate signals include implanted fiducials tracked by kV imaging or electromagnetic transponders (Cho et al., 2009; Bertholet et al., 2019), while external surrogate signals include spirometers, pressure belts, or optical tracking systems. In this work, we rely on external surrogate signals, although the proposed methodology is equally applicable to internal surrogate signals.

Then, based on a surrogate signal that can be acquired during treatment, deformable motion models allow for the estimation of respiratory motion in an entire anatomical region, including the tumor and nearby OAR (McClelland et al., 2013). Such models are usually generated based on previously acquired 4D image data using a sequential process including two-stages: first, the motion is estimated from the imaging data using deformable image registration and second, a correspondence model approximates the relationship between the surrogate signal and the motion using regression modeling (McClelland et al., 2013; Wilms et al., 2014). Later work (McClelland et al., 2017; Huang et al., 2024) proposed an integrated model that jointly optimizes image registration and correspondence model fitting, improving the reconstruction of 4D image data. The majority of approaches assume simple linear or polynomial relationships between the surrogate signals and local deformations.

The estimation of motion from 4D image data is usually done by pairwise registration of the dynamic 3D image frames to a fixed image, acquired at a reference breathing state (Ehrhardt and Lorenz, 2013). A variety of optimization- and learning-based registration approaches has been applied specifically to thorax images, see Xiao et al. (2023) for an overview. Recently, implicit neural representations (INRs) have gained attention in the field of medical image registration (Sun et al., 2024), performing on par with the state-of-the-art for deformable lung registration (Wolterink et al., 2022; Zimmer et al., 2024). INRs offer a continuous representation of deformation, and their capacity for analytical differentiation provides significant benefits for implementing PDE-based regularizers and loss functions (Chen et al., 2023). Closely related to physics-informed neural networks (PINNs) (Raissi et al., 2019), biophysical constraints can be introduced into the registration model as proposed by Arratia López et al. (2023) and, moreover, INRs naturally fit into a diffeomorphic registration setting (Sun et al., 2022; Tian et al., 2025; Han et al., 2022; Zhang and Liang, 2023; Sun et al., 2024). Another line of research uses INRs for motion-compensated image reconstruction and benefits in particular from the implicit continuity regularization of INRs and the ability to reconstruct from sparsely and irregularly sampled data (Feng et al., 2025; Shen et al., 2024; Zhang et al., 2023). Although some of these approaches include the cardiac or respiratory phase as a temporal dimension, the motion is only implicitly modeled. Despite the suc-

cess of the approaches in the context of image registration and motion-compensated image reconstruction, there is rarely any work utilizing INRs for surrogate-based motion modeling.

In this work, we investigate implicit neural representations for creating patient-specific motion models that represent the lung deformation over an entire breathing cycle. Specifically, based on a 4D image dataset with associated surrogate signals, we aim to learn a model that is able to estimate the internal motion state from unseen surrogate signals. This work builds upon our previous conference paper (Boysen et al., 2025), in which we developed an integrated INR-based motion model that uses spatial position and surrogate signals as input dimensions to predict local motion. The network is capable of learning complex functions to model the relationship between surrogate signals and internal breathing motion. It learns the respiration-induced motion deformations from sparsely sampled image data by registering images from different breathing states to a reference state.

This paper extends our previous work in several aspects: First, a much more detailed mathematical derivation and description of the approach is presented in the context of surrogate based motion modeling. More importantly, based on the transport equation, we extend our INR-based approach to a diffeomorphic, time-continuous respiratory motion model that no longer requires a fixed reference image for pair-wise registration. Instead, we introduce a group-wise diffeomorphic INR-based registration, enabling motion estimation between arbitrary respiratory phases. To determine the motion from one breathing phase to another, our model traces the trajectories along the breathing curve. Model training is, therefore, performed by sampling along the respiratory trajectories between arbitrary image pairs. Besides the physics-based spatial regularization introduced in Boysen et al. (2025), additional temporal constraints are employed to ensure smooth and plausible motion trajectories. The motivation for these substantial adaptations of our approach is twofold: First, we aim to improve the robustness by increasing the number of available image pairs for training from $n - 1$ when using a fixed reference image to $n(n - 1)$ with arbitrary image pairs. Second, we address the extrapolation limitations commonly observed in implicit neural representations and also reported in Boysen et al. (2025). Since surrogate signals are recorded continuously while images are only available at a discrete amount of time points, our trajectory-based sampling scheme together with the temporal smoothness constraint forces our model to learn plausible motion information, even for surrogate signals lying outside of acquired image information.

We call our adapted and improved approach a physics-regularized implicit surrogate-based modeling for respiratory motion (PRISM-RM), and provide evaluation experiments

on 3D+t (4D) CT images of lung cancer patients with corresponding surrogate signals to show the advantages of these adaptations, especially in extrapolation scenarios. Further, we investigate the introduced spatial and temporal regularizations in an ablation study.

2. Methods

For surrogate-based modeling we consider the following scenario: given is a 4D image dataset $\mathcal{I} = \{I_j\}_{j=0}^{n-1}$ consisting of n 3D images $I_j : \Omega \rightarrow \mathbb{R}$ ($\Omega \subset \mathbb{R}^3$) acquired at time-points $t_j \in [0, T]$ and a time-continuous surrogate signal $\mathbf{s} : [0, T] \rightarrow \mathcal{S}$ ($\mathcal{S} \subset \mathbb{R}^{n_s}$) of dimension n_s , where n_s is the surrogate signal dimensionality, such that $\mathbf{s}(t_j)$ corresponds to I_j .

2.1 Background on Surrogate-Based Motion Modeling

Surrogate-based motion models aim to predict internal organ motion based on externally measured surrogate signals \mathbf{s} . More formally, the function to be determined is

$$\varphi : \Omega \times \mathcal{S} \rightarrow \Omega \quad \hat{\mathbf{x}} = \varphi^\theta(\mathbf{x}, \mathbf{s}). \quad (1)$$

The function φ^θ is parametrized by θ and predicts the position of a particle $\hat{\mathbf{x}}$ initially positioned at $\mathbf{x} \in \Omega$ given the measured signal $\mathbf{s} \in \mathcal{S}$. The parameters θ are the coefficients of a regression model that relates the surrogate signals to the internal spatial motion. They are learned from 4D image data acquired together with surrogate signals. Usually, patient-specific images acquired before radiotherapy are used to build a correspondence model; however, population-based models can also be created using multiple 4D image datasets (McClelland et al., 2013; Wilms et al., 2014, 2017).

2.1.1 Sequential Respiratory Motion Modeling

Sequential respiratory motion models describe a group of motion modeling approaches that feature two sequential stages (McClelland et al., 2013). First, the respiratory motion is estimated by image registration from the images I_j . Second, a correspondence model is fitted mapping the motion with the corresponding surrogate signals $\mathbf{s}(t_j)$ (see Fig. 1 for a schematic overview).

In the first stage, all images I_j are registered to a fixed reference state image, e.g. $I_{ref} = I_0$, to obtain the transformations $\phi_j : \Omega \rightarrow \Omega, j = 1, \dots, n-1$. In the case of pair-wise deformable image registration, the following objective is optimized:

$$\mathcal{L}_j(\phi_j) = \mathcal{D}(I_{ref}, I_j \circ \phi_j) + \alpha \mathcal{R}(\phi_j), \quad (2)$$

where \mathcal{D} denotes a similarity measure and \mathcal{R} is a regularization term weighted by α .

In the second stage, the relationship between the acquired surrogate signals $\mathbf{s}(t_j)$ and the $n-1$ estimated deformation fields ϕ_j is learned by fitting a regression model to approximate the deformations, i.e. $\phi_j \approx \varphi^\theta(\mathbf{s}(t_j))$. Most approaches use simple linear or polynomial models combined with standard fitting techniques such as ordinary least squares (McClelland et al., 2013). Following the example of Wilms et al. (2014), the deformations are represented by the group exponential map of a stationary velocity field $\phi_j(\mathbf{x}) = \exp(v_j)(\mathbf{x})$ (Arsigny et al., 2006). Let $\mathbf{V}_j \in \mathbb{R}^{3m}$ be the matrix of vectorized representations of the velocity fields v_j , where m denotes the number of voxels. The parameters θ of a linear correspondence model with regression coefficients $\mathbf{B} \in \mathbb{R}^{3m \times n_s}$ and intercept $\mathbf{b}_0 \in \mathbb{R}^{3m}$ are given by:

$$\theta = \{\mathbf{B}, \mathbf{b}_0\} = \arg \min_{\{\mathbf{B}, \mathbf{b}_0\}} \sum_{j=1}^{n-1} \|\mathbf{V}_j - \mathbf{B}\mathbf{s}(t_j) - \mathbf{b}_0\|_2^2. \quad (3)$$

During inference, a newly measured surrogate signal \mathbf{s}' is applied to predict velocities $\mathbf{V} = \mathbf{B}\mathbf{s}' + \mathbf{b}_0$ followed by exponentiation to compute the deformation. To model the variations in motion between inspiration and expiration (hysteresis), many approaches use separate models for inhalation and exhalation due to the limited expressivity of correspondence models.

2.1.2 Integrated Respiratory Motion Modeling

In the integrated respiratory motion modeling approach, the idea is to directly estimate the parameters θ of the correspondence model from the images I_j and corresponding surrogate signals $\mathbf{s}(t_j)$. Here, an optimization over the parameters θ of the correspondence model simultaneously adapts the reference image $I_{ref} = I_0$ to all remaining images $I_j, j = 1, \dots, n-1$. The objective can be defined analogously to Eq. (2):

$$\mathcal{L}(\theta) = \sum_{j=1}^{n-1} \mathcal{D}(I_{ref}, I_j \circ \varphi^\theta(\mathbf{s}(t_j))) + \alpha \mathcal{R}(\varphi^\theta(\mathbf{s}(t_j))). \quad (4)$$

In contrast to the sequential approach (Sec. 2.1.1) the parameters θ are optimized directly by a gradient-based algorithm, and the registration-based estimation of intermediate deformations ϕ_j is avoided; thus, the optimization is carried out simultaneously over all images and surrogates. The integrated modeling approach is shown in Fig. 1.

One integrated respiratory motion model was introduced by McClelland et al. (2017). The authors define a linear correspondence model in order to map the surrogate signals to the parameters of the deformable transformation, such as velocities, displacements or B-spline functions. Theoretically, the correspondence model can be defined in any differentiable fashion, e.g. with polynomials or spline-based

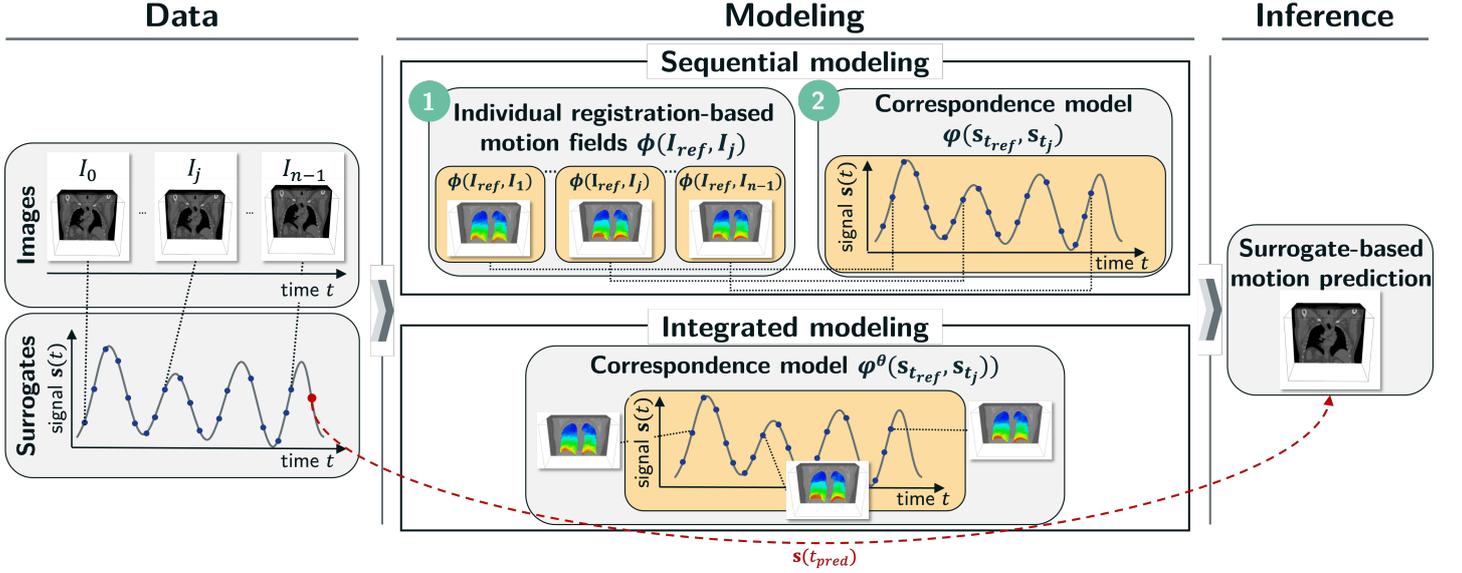


Figure 1: Sequential and integrated respiratory motion modeling using image data with corresponding surrogate signals over time.

regression, however, this approach requires a strict definition of the relation between surrogate signal and motion parameters. Omitting the frame-wise estimation of deformations increases robustness and has been applied successfully to, e.g., 4D image reconstruction when only partial data was acquired per time point (McClelland et al., 2017; Huang et al., 2024).

In our previously published conference paper (Boysen et al., 2025), we aimed to lift the restrictions of an explicitly defined correspondence model and explore the capabilities of the recently emerged INRs (Sitzmann et al., 2020). INRs map spatial coordinates \mathbf{x} (and surrogate values) to their corresponding image intensity by training a dense neural network as a mapping function $f(\mathbf{x})$, enabling continuous spatio-temporal representations, making them invariant to image resolution and size. Especially, for patient-specific applications, INRs have shown great advantages in terms of memory consumption and speed and, thus, pose a promising alternative to conventional deep-learning approaches.

Inspired by Wolterink et al. (2022), we propose to use surrogate-conditioned INRs to implicitly represent the displacement-based warping function between a globally fixed reference image I_{ref} and some image I_j with the corresponding surrogate signal $s(t_j)$:

$$\varphi^\theta(\mathbf{x}, \mathbf{s}) = \mathbf{x} + f^\theta(\mathbf{x}, \mathbf{s}), \quad (5)$$

with θ being the parameters of the INR. Notably, this approach still requires a fixed reference image I_{ref} and also to learn separate models for the inhalation and exhalation phases. Additionally, to address the overfitting risk of deep neural networks, we include a physics-inspired spatial regularization inspired by PINNs (Raissi et al., 2019) (Sec. 2.3).

2.2 Trajectory-Aware Respiratory Motion Modeling

Following the idea of integrated respiratory motion models (Sec. 2.1.2), we propose to extend the INR-based integrated respiratory modeling approach by re-formulating it as a reference-free diffeomorphic model. We assume that for a particle at position $\mathbf{x} \in \Omega$ at time $t_0 = 0$ and surrogate signal $s(t_0 = 0)$, the function $\varphi(\mathbf{x}, s(t_j)) : \Omega \times \mathcal{S} \rightarrow \Omega$ defines the position of that particle at time t_j for a given surrogate signal $s(t_j)$. More generally and for ease of notation regarding the particle's position, let $\phi_{r,t}(\mathbf{x})$ be the position of a particle at time t with starting position \mathbf{x} at time r , in our example $\varphi(\mathbf{x}, s(t_j)) = \phi_{t_0,t_j}(\mathbf{x})$. To ensure diffeomorphism, the motion $\phi_{r,t}$ is generated as a flow of the time-dependent velocity field v defined by the transport equation:

$$\frac{\partial \phi_{r,\cdot}(\mathbf{x})}{\partial t} = v^\theta(\phi_{r,t}(\mathbf{x}), \mathbf{s}(t)) \quad \text{s.t.} \quad \phi_{r,r}(\mathbf{x}) = \mathbf{x}. \quad (6)$$

The time-dependent velocity field v^θ is represented by an INR parametrized by network weights θ , and using the spatial coordinate and surrogate as inputs. To now estimate the motion between two time frames t_j and t_k , starting from the given velocity field, an integration over time is required:

$$\phi_{t_j,t_k}^\theta(\mathbf{x}) = \mathbf{x} + \int_{t_j}^{t_k} v^\theta(\phi_{t_j,\tau}(\mathbf{x}), \mathbf{s}(\tau)) d\tau. \quad (7)$$

The resulting deformations ϕ_{t_j,t_k}^θ are diffeomorphic at all times $t_j, t_k \in [0, T]$ (Younes, 2019). Here, we propose to realize integration using an Euler-based scheme with variable step-size and a fixed number of evaluations. Note

that this formulation diverges from the widely used scaling-and-squaring-based approaches (Arsigny et al., 2006; Dalca et al., 2019) in two ways: we employ variable integration intervals rather than $[0, 1]$, and we employ time-dependent velocity fields, similarly to LDDMM (Beg et al., 2005).

The time-continuous formulation in Eqs. (6) and (7) now allows us to reformulate the integrated motion model in Eq. (4) to:

$$\mathcal{L}(\theta) = \sum_{\substack{(j,k) \in [1, \dots, n]^2 \\ j \neq k}} \mathcal{D}(I_j, I_k \circ \phi_{t_j, t_k}^\theta) + \alpha \mathcal{R}(\phi_{t_j, t_k}^\theta, v^\theta), \quad (8)$$

with notation analogous to Eq. (2). This reformulated approach has a number of advantages. First, by omitting a fixed reference frame, we can use $n(n-1)$ registration pairs to learn the network parameters instead of only $n-1$ as in the approach in Sec. 2.1.2. Second, for the calculation of the motion deformations ϕ_{t_j, t_k} using the integral in Eq. (7) the recorded surrogate signal is sampled densely and overlapping, even at time-points without acquired image information. Because all samples contribute to the loss function, the network learns that there is a continuous relationship between the surrogate signal and internal motion, i.e. the velocities. Furthermore, provided that the surrogate signals allow to distinguish between inhalation and exhalation, we can model the respiratory motion continuously across all respiratory phases. Therefore, we call our model trajectory-aware.

Commonly used approaches such as normalized cross-correlation (NCC) and bending energy can be chosen as distance measures and regularizers (Wolterink et al., 2022). To enforce a plausible, smooth and continuous respiratory motion model, we propose a more complex bio-physically motivated spatial regularizer and an additional temporal regularization that exploits the Euler-based temporal sampling in Eq. (7).

2.3 Physics-Enhanced Spatial and Temporal Regularization

Inspired by PINNs (Raissi et al., 2019), we aim to tackle the data shortage by utilizing physics-inspired regularization. We assume the lung and tumor tissue to behave like a neohookean hyperelastic material, which closely mimics its natural properties (Concha et al., 2018) and can be formulated as (Arratia López et al., 2023):

$$\mathcal{R}_{ph}(\varphi^\theta) = \text{trace}(\mathbf{C}) - 3 - \log(|\mathbf{J}|^2) + \lambda(|\mathbf{J}| - 1)^2 \quad (9)$$

where $|\mathbf{J}|$ is the determinant of the jacobian $\mathbf{J} = d\phi^\theta/d\mathbf{x}$, \mathbf{C} is defined as $\mathbf{C} = \mathbf{J}^T \mathbf{J}$, and λ is a hyperparameter. Applying \mathcal{R}_{ph} as regularization in Eq. 4 as proposed in our previous work (Sec. 2.1.2) creates a physics-enhanced INR, which is subjected to physically valid displacements in the space domain Ω at times t_j available in the dataset. Notably,

only utilizing the \mathcal{R}_{ph} regularization does not guarantee any smoothness in the time domain, which neglects the naturally smooth respiratory motion characteristics. Thus, we propose to extend the regularizer in our new trajectory-aware motion model (Sec. 2.2) by a temporal constraint that guarantees temporal smoothness of the velocity field and physical plausibility of the resulting deformations. Specifically, we propose to add a total variation regularization in the temporal domain to the velocities:

$$\mathcal{R}_t(v^\theta) = \frac{1}{2} \int_{t_r}^{t_k} \sqrt{\|\nabla_t v^\theta\|_2^2 + \epsilon} dt \quad (10)$$

where ∇_t is the temporal gradient and ϵ a small constant. Eq. 10 allows for sharp edges along time, which is particularly important for the respiratory motion at the end-inhale and end-exhale breathing states. The complete spatio-temporal regularizer is then defined as follows with α_t and α_{ph} being hyperparameters:

$$\mathcal{R}_{st}(\varphi^\theta, v^\theta) = \alpha_{ph} \mathcal{R}_{ph}(\varphi^\theta) + \alpha_t \mathcal{R}_t(v^\theta). \quad (11)$$

3. Experimental Setup

The purpose of our experimental setup is to investigate the interpolation and extrapolation capabilities of the trajectory-aware motion model. Specifically, the interpolation setup is intended to investigate how well the motion model performs for breathing states within the range of the surrogate signal in the training data, but without including the corresponding image in the training of the model. The extrapolation setup aims to investigate the motion model’s performance for breathing states outside the range of the surrogate signal in the training data. The corresponding images are also not included in the training data.

3.1 Data

For our experiments, we use a subset of the in-house CT dataset from Wilms et al. (2014), that contains 4D upper-body CT images of eleven lung cancer patients. For every patient either 10 or 14 images are available at equidistant breathing stages between the inhalation and exhalation state. Specifically, one image is recorded at the end-inhalation (EI) state, one image is recorded at the end-exhalation (EE) state, and the remaining images are evenly split among the inhalation and exhalation phases and sorted based on the amplitude. The original dataset contains spirometry measurements of air volume and air flow that were recorded simultaneously with the images (see Yamamoto et al. (2012) and Ehrhardt et al. (2007) for details on image acquisition and image reconstruction). However, to maintain comparability in our experiments, we use simulated spirometry measurements that were calculated retrospectively from the image data and provided

by Wilms et al. (2014): $s(t) = (V(t), \partial V(t)/\partial t)^T$. The lung volumes are extracted and normalized per patient, then bootstrapped by the lung volume at maximum exhalation, which is defined as $V = 0$. Furthermore, 70 to 90 corresponding landmarks are manually set per patient at the end-inhale (EI), end-exhale (EE), mid-inhale (MI) and mid-exhale (ME) states (see Fig. 2 with a corresponding surrogate signal), where the mid breathing states are located at the center between the end-inhale and end-exhale states in the corresponding breathing phases. All images entail visible cancerous tissue.

3.2 Experiments

Our experimental setup simulates a patient-specific scenario to learn the patient’s breathing pattern during a training phase, enabling the prediction of respiratory motion based on a given surrogate signal during inference. In order to evaluate the performance of our method for accurate respiratory motion modeling, we design a leave-out-based setup to simulate interpolation and extrapolation scenarios. Using this particular leave-out setup allows for ground truth evaluation of the predicted respiratory motion, since landmarks are given for the left-out images. This scenario is like the real-world case of predicting breathing movements outside the known patient-specific range of variability (i.e. extrapolation). By excluding multiple breathing states, we can estimate the model’s performance in more extreme cases and assess its robustness. Moreover, utilizing this experiment design allows for a direct performance comparison to the available baseline for the used data (Wilms et al., 2014; Boysen et al., 2025). The extrapolation scenario is particularly important for the clinical application of a respiratory motion model, because respiratory motion is highly variable (e.g. day-to-day, breath-to-breath variability) and the motion model must perform well during inference time on unseen surrogate signals to ensure the radiotherapy effectiveness. PRISM-RM is specifically designed to improve performance in this extrapolation scenario, where the spatio-temporal regularization combined with a diffeo-

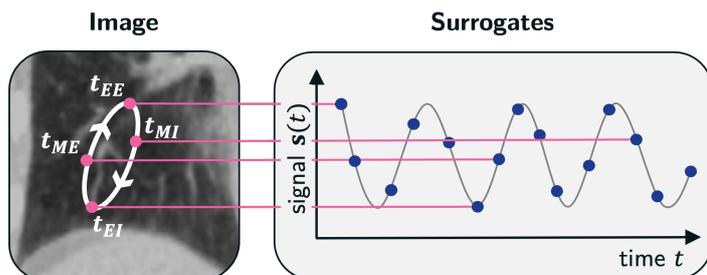


Figure 2: Exemplary motion trajectory of a single voxel tracked over time (left) with a corresponding surrogate signal (right).

morphic setting targets improved smoothness and physical plausibility over time.

Three experiments were performed in which correspondence models were applied to estimate patient-specific motion between end-inspiration and mid-inspiration (EI→MI), mid-expiration (EI→ME), and end-expiration (EI→EE). Notably, the trajectory-aware approach allows to use any image as a reference for evaluation due to its reference-free nature. However, to ensure comparability, we choose the EI breathing state (Wilms et al., 2014; Boysen et al., 2025). In all three experiments we use the same dataset per patient but with different training/test data splits. The two interpolation experiments (EI→MI and EI→ME) are evaluated by leaving out the phases at mid-expiration (ME) and mid-inspiration (MI) during training. The extrapolation experiment (EI→EE) is designed to evaluate the extrapolation ability of the approaches, i.e. in cases where the respiratory motion lies outside the range of the acquired training images. To make the extrapolation scenario more severe, the neighboring breathing states are omitted during training in addition to the EE breathing state. Specifically, if 10 images are available for a patient one additional breathing state in each breathing phase is left out, in case 14 are available two are left out. A visual explanation of our experiments is represented in Fig. 3 showing the respiratory states included in the training data of the respective experiment in white and the excluded one(s) in pink. Here, the trajectory of a single landmark over the entire breathing phase is shown. Where, specifically, for the extrapolation experiment (EI→EE) the breathing states MI and ME are part of the training set and EE is used for testing.

For performance evaluation in validation and test time, the target registration error (TRE) between the corresponding landmarks is used. Specifically, for evaluation during testing, the TRE between the landmarks of the warped EI breathing state and the leftout breathing state is used.

As the baseline method we use the sequential two-stage model proposed by Wilms et al. (2014), which yields state-of-the-art performance for respiratory motion modeling. We then compare our INR-based integrated motion model

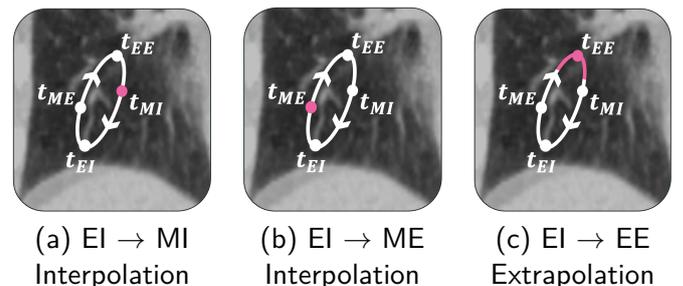


Figure 3: Visualization of the three experiments showing the respiratory states included in the training data in white and the excluded one(s) in pink.

(Sec. 2.1.2) and the trajectory-aware approach (Sec. 2.2) to it.

3.3 Implementation Details

The velocity fields are represented by a sinusoidal activated INR (Sitzmann et al., 2020), which is composed of three hidden layers with 256 neurons each. During training at every epoch we sample 10.000 points in $\Omega \times \mathcal{S}$. Coordinates are sampled using a weighted image-gradient based strategy in the lung region of the image at the end-inhale breathing state of each patient. The surrogate signals of the 10.000 points per epoch are sampled using two different strategies. 40% are taken from the surrogate signals \mathcal{S} available for images I_j in the dataset (see the blue points in Fig. 4 and pink points for a specific registration pair). For the other 60% the surrogate signal is uniformly sampled over the entire patients breathing cycle (see the yellow points in Fig. 4). Thus, for this 60% split no image data is available. Specifically, the volumes $V(t)$ sampled uniformly over the entire breathing cycle and the corresponding derivatives are interpolated from the real-world surrogate $s_{\partial V(t)/\partial t}$. The interpolation is done in a piece-wise linear fashion between the known derivatives at the breathing states available in the dataset. We use the normalized cross-correlation $\mathcal{D}(\cdot) = \text{NCC}(\cdot)$ as a similarity measure for the inputs where fixed and moving image are available in dataset. For the 60% of points at times without available images, no image-based similarity measure is calculated and used for optimization. However, the regularizer is applied to all 10.000 points, because it does not require any image information. During training we validate on the TRE between the EI breathing state and all wrapped images with available landmarks. The regularizer as described above is applied to all points and its weighting hyperparameters (Eq. (11)) are chosen to $\alpha_{ph} = 0.001$ and $\alpha_t = 0.1$. The λ of the physics loss (Eq. (9)) is chosen to $\lambda = 1$ as in (Arratia López et al., 2023). The learning rate is set to 1×10^{-5} .

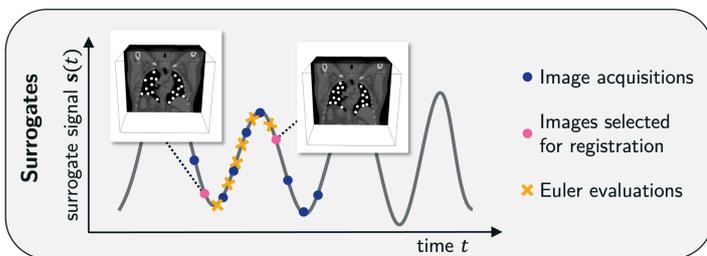


Figure 4: Registration of two images during one optimization step using an Euler scheme. Euler evaluations (yellow dots) can be made in a continuous manner between the actual image acquisitions (blue dots), allowing for the registration of two arbitrary images.

4. Results

Tab. 1 presents the results of PRISM-RM, our previously introduced integrated approach and the sequential approach compared for the three leave-out scenarios all on the same dataset. The integrated INR-based respiratory motion model shows on-par performance with the sequential baseline model for interpolation (ME, MI), but significantly underperforms in extrapolation (EE). PRISM-RM performs comparable to the integrated INR-based respiratory motion model, but outperforms it in extrapolation, although it still performs worse than the sequential two-stage approach. Indeed, a paired T-test reveals no statistically significant differences between PRISM-RM and our previous INR-based integrated approach (Sec. 2.1.2) in the interpolation scenarios EI \rightarrow MI and EI \rightarrow ME, but significantly improved results in terms of extrapolation EI \rightarrow EE ($p < 0.05$).

Fig. 5 shows the magnitude of the resulting displacement fields for patient six in the ME experiment, in other words, the displacement magnitude between breathing states EI and ME. We show patient six as an example, because this patient shows approximately the same TRE improvement between the unregistered images (Init) and the PRISM-RM registered images as the dataset average. The magnitude increases from blue to red. Note that only intra-patient registration uses ME image information and can therefore be considered a pseudo ground truth. The displacement fields of both models, the sequential and trajectory-aware model, infer the largest displacement at the bottom of the lung. Considering the largest physiological difference between the two breathing states is located at the bottom of the lung (see the red area in Fig. 5(a)), both models infer a plausible displacement field. However, the displacement field magnitude of our trajectory-aware model (Fig. 5(c)) is much smoother than the one of the sequential approach (Fig. 5(b)).

4.1 Ablation Study

We conduct an ablation study to investigate the impact of the two different regularization types (spatial and temporal) on the interpolation and extrapolation performance of PRISM-RM. Tab. 2 shows an ablation study on the different regularization types of the trajectory-aware model. We compare the overall mean and standard deviation of the TRE resulting from registration using both the spatial and temporal regularizers simultaneously (PRISM-RM) or each regularizer independently. In the interpolation experiment EI \rightarrow MI only using the spatial regularizer performs best by a small margin. For the EI \rightarrow ME interpolation experiment PRISM-RM performs comparable to our previously introduced only spatially regularized integrated INR approach and also compared to only using the temporal regularizer. Thus, the integration of both a temporal and spatial reg-

Table 1: Results of three registration scenarios measured as TRE in millimeters. In bold we show the best performance among the two INR-based methods. Presented methods: ‘Init’ initial alignment before registration; ‘Seq’ registration by a sequential two-stage approach (Wilms et al., 2014); ‘Integ’ registration by our proposed integrated INR approach (Sec. 2.1.2); ‘PRISM-RM’ registration with the new trajectory-aware approach (Sec. 2.2).

ID	EI → MI				EI → ME				EI → EE			
	Init	Seq	Integ	PRISM-RM	Init	Seq	Integ	PRISM-RM	Init	Seq	Integ	PRISM-RM
01	3.98	1.68	1.96	2.15	2.08	1.48	1.54	1.86	4.25	1.29	2.54	2.10
02	5.08	1.65	2.10	1.92	2.51	1.32	1.97	1.82	6.29	1.56	2.85	2.17
03	4.43	1.83	1.92	1.72	2.41	1.83	1.99	2.68	5.45	1.74	2.81	2.41
04	4.95	2.24	2.08	2.27	2.52	1.77	2.14	2.34	6.19	1.76	3.34	2.13
05	3.76	2.42	2.04	1.70	2.41	1.78	2.20	2.02	6.79	2.34	4.26	3.98
06	4.56	2.15	1.84	1.73	1.98	1.87	1.72	1.65	6.44	1.46	4.61	2.51
07	3.80	1.78	1.74	1.81	1.92	1.97	1.82	1.96	4.31	1.61	2.87	2.51
08	6.68	2.59	2.11	2.22	3.84	2.08	2.16	3.46	10.76	2.66	5.83	4.38
09	4.11	2.02	1.77	1.60	2.76	1.70	1.77	2.86	6.40	1.74	3.64	2.85
10	4.03	1.68	1.90	1.81	2.07	1.43	1.90	1.62	6.06	1.60	4.41	2.63
11	6.71	2.24	2.50	2.19	2.94	1.77	2.00	2.50	8.31	2.65	3.07	3.45
Mean	4.74	2.03	2.00	1.92	2.49	1.73	1.93	2.25	6.48	1.86	3.66	2.83
Std	1.06	0.33	0.21	0.24	0.55	0.23	0.20	0.58	1.81	0.47	1.01	0.77

Table 2: Ablation study results are measured in TRE in millimeters. Compared methods: ‘Init’ before registration; ‘PRISM-RM’ trajectory-aware approach; ‘S’ trajectory-aware approach with only spatial regularization; ‘T’ trajectory-aware approach with only temporal regularization.

	EI → MI				EI → ME				EI → EE			
	Init	PRISM-RM	S	T	Init	PRISM-RM	S	T	Init	PRISM-RM	S	T
Mean	4.74	1.92	1.88	1.95	2.49	2.25	2.25	2.26	6.48	2.83	2.98	2.89
Std	1.06	0.24	0.23	0.24	0.55	0.58	0.44	0.42	1.81	0.77	1.01	0.9

Table 3: The average compute time per epoch in seconds during training of the different model types. Compared methods: ‘PRISM-RM’ trajectory-aware approach; ‘S’ trajectory-aware approach with only spatial regularization; ‘T’ trajectory-aware approach with only temporal regularization.

	PRISM-RM	S	T
MI	40.1s	27.8s	21.2s
ME	40.3s	29.7s	25.7s
EE	18.9s	13.6s	11.2s

ularization overall does not necessarily improve the model performance significantly in interpolation scenarios. However, for the extrapolation experiment (EI → EE) Tab. 2 shows that the integration of regularization both in space and time improves the performance significantly. Also, the decreasing standard deviation in the extrapolation suggests that the combination also increases PRISM-RM’s certainty about the motion. Only using spatial regularization yields worse results compared to solely using temporal regularization in the extrapolation case. This follows the intuition, since the spatial regularization only induces constraints at

the individual time steps and not over time, which is the additional constraint that is necessary for the extrapolation scenario, since the model cannot derive information from the data in this case.

The models were trained on a NVIDIA A30. The training times for training and inference of the different experiments in the ablation study are presented in Tab. 3. The epoch duration in the extrapolation experiment is the shortest as the number of breathing states in the dataset is smaller compared to the interpolation experiments (Sec. 3.2). Further, the trajectory-aware approach with only the temporal regularization is faster than the other methods, because fewer derivatives must be calculated due to the missing physics-inspired regularization. Hence, the inference duration for a single image is similar in all experiments, as no additional backpropagation is required for the regularizer. The inference of a model takes around 2.2s for a single image.

5. Discussion and Conclusion

This work explores the potential of INRs in integrated surrogate-based respiratory motion modeling. In our initial

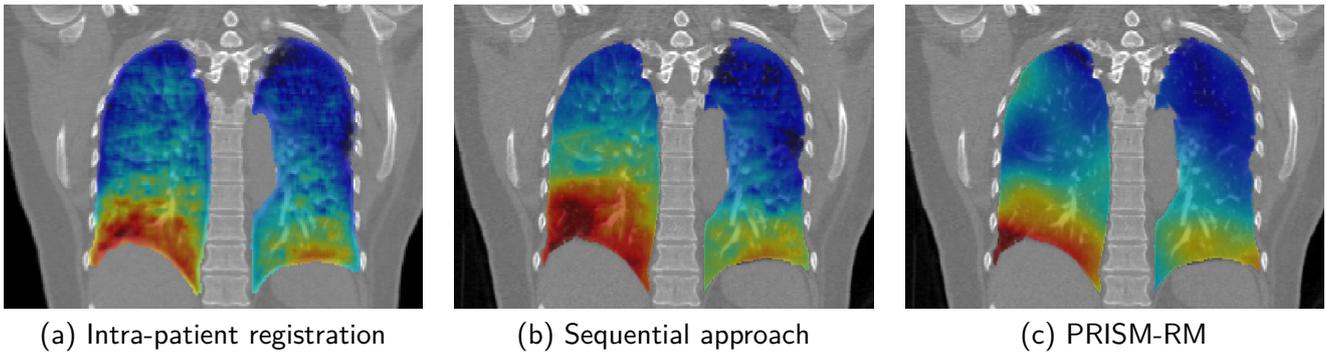


Figure 5: Color-coded magnitudes of the motion (increasing from blue to red) estimated by an image-based intra-patient registration in the ME setting (a); the motion estimation using a sequential surrogate-based registration followed by an ordinary least squares regression (Wilms et al., 2014) (b); and the motion estimation by PRISM-RM (c).

work (Boysen et al., 2025) INRs were used to enable the learning of complex spatial and temporal relationships between surrogate and internal motion in an integrated model. This allows for advancing beyond sequential two-stage models and the assumptions of common linear or polynomial correspondence models. The evaluation results show this INR-based model to be equivalent to established two-stage models for interpolation scenarios, but its ability to generalize to unseen breathing states during extrapolation remains limited and requires further methodological improvements.

As a continuation of this work, we present PRISM-RM, a trajectory-aware INR-based framework for integrated respiratory motion modeling that addresses key limitations of our previous work, especially in extrapolation to unseen breathing states. Unlike previous methods constrained by fixed reference states, our proposed PRISM-RM framework represents a time-continuous representation of the entire respiratory cycle and thus allows motion estimation between arbitrary respiratory phases, significantly extending the training basis by using arbitrary image pairs. This is further enhanced by the applied Euler integration along continuous surrogate signals, whereby a consistent and smooth motion is learned even outside of available image information. Since the lung tissue is modeled as a neohookean hyperelastic material, the smooth deformation is desirable and also corresponds to its natural behavior. By leveraging spatio-temporal regularization within the diffeomorphic framework, our method enables smooth, physiologically consistent motion fields while maintaining adaptability to patient-specific variations.

Our results show, that PRISM-RM maintains a comparable performance in interpolation tasks while significantly improving extrapolation accuracy compared to the integrated INR-based respiratory motion model, which is critical for unseen breathing patterns during treatment delivery. Our ablation study finds the integration of spatio-temporal regularization, in particular the temporal component, crucial to stabilize predictions across the breathing cycle. This

greatly aligns with the physiological continuity of respiratory motion. Notably, the decreasing variance in the extrapolation suggests increased confidence in the trajectory-aware model's predictions, despite the data limitations inherent to surrogate-based approaches.

Despite the advances of PRISM-RM in INR-based modeling, limitations remain. Although the performance in extrapolation improved, it still lags behind sequential approaches. This indicates that INRs are not yet fully overcoming the well-known generalization barriers especially in the extrapolation experiment for unseen respiratory states. Thus, further methodological improvements are needed, e.g. we plan to apply neo-Hookean regularization also along the Euler steps, similar to the temporal regularization, to improve the spatial regularization at times where no images are available. Another line of future work will be the exploration of global motion atlases for initialization which afterwards can be conditioned to individual patients and we plan to investigate how the scaling-and-squaring algorithm can be applied to a sparse velocity field to take advantage of its computational effectiveness. Overall, PRISM-RM reveals a promising line of research and is a meaningful step towards clinically applicable and patient-specific respiratory motion modeling, with the potential to improve precision in radiotherapy delivery.

Acknowledgments

This work was funded by grant 16IS24006C from the German Federal Ministry of Research, Technology and Space (BMFTR) and by the state of Schleswig-Holstein, Germany, through grant 22024003.

Ethical Standards

The work follows appropriate ethical standards in conducting research and writing the manuscript, following all applicable

laws and regulations regarding treatment of animals or human subjects.

Conflicts of Interest

We declare we don't have conflicts of interest.

Data availability

The used dataset was introduced by Wilms et al. (2014) and kindly made available for this work. For details on the data availability for third-parties the interested reader is kindly referred to the corresponding author.

References

- Pablo Arratia López, Hernán Mella, Sergio Uribe, Daniel E. Hurtado, and Francisco Sahli Costabal. WarpPINN: Cine-MR image registration with physics-informed neural networks. *Medical Image Analysis*, 89:102925, October 2023. ISSN 13618415. .
- Vincent Arsigny, Olivier Commowick, Xavier Pennec, and Nicholas Ayache. A Log-Euclidean Framework for Statistics on Diffeomorphisms. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Dough Tygar, Moshe Y. Vardi, Gerhard Weikum, Rasmus Larsen, Mads Nielsen, and Jon Sporring, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, volume 4190, pages 924–931. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-44707-8 978-3-540-44708-5. . Series Title: Lecture Notes in Computer Science.
- M. Faisal Beg, Michael I. Miller, Alain Trouvé, and Laurent Younes. Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, February 2005. ISSN 1573-1405. .
- Jenny Bertholet, Antje Knopf, Björn Eiben, Jamie McClelland, Alexander Grimwood, Emma Harris, Martin Menten, Per Poulsen, Doan Trang Nguyen, Paul Keall, and Uwe Oelfke. Real-time intrafraction motion monitoring in external beam radiotherapy. *Physics in Medicine and Biology*, 64(15):15TR01, August 2019. ISSN 0031-9155. .
- Jan Boysen, Hristina Uzunova, Jan Ehrhardt, and Heinz Handels. Surrogate-based Respiratory Motion Estimation using Physics-enhanced Implicit Neural Representations. In Christoph Palm, Katharina Breininger, Thomas Derseno, Heinz Handels, Andreas Maier, Klaus H. Maier-Hein, and Thomas M. Tolxdorff, editors, *Bildverarbeitung für die Medizin 2025*, pages 13–18. Springer Fachmedien Wiesbaden, Wiesbaden, 2025. ISBN 978-3-658-47421-8 978-3-658-47422-5. . Series Title: Informatik aktuell.
- Laura I Cerviño, Jiang Du, and Steve B Jiang. MRI-guided tumor tracking in lung cancer radiotherapy. *Physics in Medicine and Biology*, 56(13):3773–3785, July 2011. ISSN 1361-6560. .
- Honglin Chen, Rundi Wu, Eitan Grinspun, Changxi Zheng, and Peter Yichen Chen. Implicit Neural Spatial Representations for Time-dependent PDEs. In *Proceedings of the 40th International Conference on Machine Learning*, pages 5162–5177. PMLR, July 2023. ISSN: 2640-3498.
- Byungchul Cho, Per R. Poulsen, Alex Sloutsky, Amit Sawant, and Paul J. Keall. First Demonstration of Combined kV/MV Image-Guided Real-Time Dynamic Multileaf-Collimator Target Tracking. *International Journal of Radiation Oncology*Biophysics*Physics*, 74(3):859–867, July 2009. ISSN 0360-3016. .
- Felipe Concha, Mauricio Sarabia-Vallejos, and Daniel E. Hurtado. Micromechanical model of lung parenchyma hyperelasticity. *Journal of the Mechanics and Physics of Solids*, 112:126–144, March 2018. ISSN 00225096. .
- Adrian V. Dalca, Guha Balakrishnan, John Guttag, and Mert R. Sabuncu. Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. *Medical Image Analysis*, 57:226–236, October 2019. ISSN 1361-8415. .
- Jennifer Dhont, Susan V. Harden, Lessandra Y. S. Chee, Katharine L. Aitken, Gerard G. Hanna, and Jenny Bertholet. Image-guided Radiotherapy to Manage Respiratory Motion: Lung and Liver. *Clinical Oncology*, 32(12):792–804, December 2020. ISSN 0936-6555. .
- Jan Ehrhardt and Cristian Lorenz. *4D Modeling and Estimation of Respiratory Motion for Radiation Therapy*. Biological and Medical Physics, Biomedical Engineering. Springer, Berlin, Heidelberg, 2013. ISBN 978-3-642-36440-2 978-3-642-36441-9. .
- Jan Ehrhardt, René Werner, Dennis Säring, Thorsten Frenzel, Wei Lu, Daniel Low, and Heinz Handels. An optical flow based method for improved reconstruction of 4D CT data sets acquired during free breathing. *Medical Physics*, 34(2):711–721, 2007. ISSN 2473-4209. .

- Jie Feng, Ruimin Feng, Qing Wu, Xin Shen, Lixuan Chen, Xin Li, Li Feng, Jingjia Chen, Zhiyong Zhang, Chunlei Liu, Yuyao Zhang, and Hongjiang Wei. Spatiotemporal Implicit Neural Representation for Unsupervised Dynamic MRI Reconstruction. *IEEE Transactions on Medical Imaging*, 44(5):2143–2156, May 2025. ISSN 0278-0062, 1558-254X. . Publisher: Institute of Electrical and Electronics Engineers (IEEE).
- Elisa Fiorina, Veronica Ferrero, Guido Baroni, Giuseppe Battistoni, Nicola Belcari, Niccolo Camarlinghi, Piergiorgio Cerello, Mario Ciocca, Micol De Simoni, Marco Donetti, Yunsheng Dong, Alessia Embriaco, Marta Fischetti, Gaia Franciosini, Giuseppe Giraudo, Aafke Kraan, Francesco Laruina, Carmela Luongo, Davide Maestri, Marco Magi, Giuseppe Magro, Etesam Malekzadeh, Carlo Mancini Terracciano, Michela Marafini, Ilaria Mattei, Enrico Mazzoni, Paolo Mereu, Riccardo Mirabelli, Alfredo Mirandola, Matteo Morrocchi, Silvia Muraro, Alessandra Patera, Vincenzo Patera, Francesco Pennazio, Alessandra Retico, Angelo Rivetti, Manuel Dionisio Da Rocha Rolo, Valeria Rosso, Alessio Sarti, Angelo Schiavi, Adalberto Sciubba, Elena Solfaroli Camillocci, Giancarlo Sportelli, Sara Tampellini, Marco Toppi, Giacomo Traini, Serena Marta Valle, Francesca Valvo, Barbara Vischioni, Viviana Vitolo, Richard Wheadon, and Maria Giuseppina Bisogni. Detection of Interfractional Morphological Changes in Proton Therapy: A Simulation and In Vivo Study With the INSIDE In-Beam PET. *Frontiers in Physics*, 8:578388, January 2021. ISSN 2296-424X. .
- Kun Han, Shanlin sun, Xiangyi Yan, Chenyu You, Hao Tang, Junayed Naushad, Haoyu Ma, Deying Kong, and Xiaohui Xie. Diffeomorphic Image Registration with Neural Velocity Field, November 2022. arXiv:2202.12498 [cs].
- Yuliang Huang, Bjoern Eiben, Kris Thielemans, and Jamie R. McClelland. Resolving Variable Respiratory Motion From Unsorted 4D Computed Tomography. In Marius George Linguraru, Qi Dou, Aasa Feragen, Stamatia Giannarou, Ben Glocker, Karim Lekadir, and Julia A. Schnabel, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*, pages 588–597, Cham, 2024. Springer Nature Switzerland. ISBN 978-3-031-72378-0. .
- Paul J. Keall, Caterina Brighi, Carri Glide-Hurst, Gary Liney, Paul Z. Y. Liu, Suzanne Lydiard, Chiara Paganelli, Trang Pham, Shanshan Shan, Alison C. Tree, Uulke A. Van Der Heide, David E. J. Waddington, and Brendan Whelan. Integrated MRI-guided radiotherapy — opportunities and challenges. *Nature Reviews Clinical Oncology*, 19(7): 458–470, July 2022. ISSN 1759-4782. .
- Jamie R. McClelland, David J. Hawkes, Tobias R. Schaeffter, and Andrew P. King. Respiratory motion models: A review. *Medical Image Analysis*, 17(1):19–42, January 2013. ISSN 1361-8415. .
- Jamie R. McClelland, Marc Modat, Simon Arridge, Helen Grimes, Derek D’Souza, David Thomas, Dylan O’Connell, Daniel A Low, Evangelia Kaza, David J Collins, Martin O Leach, and David J Hawkes. A generalized framework unifying image registration and respiratory motion models and incorporating image reconstruction, for partial image data or full images. *Physics in Medicine & Biology*, 62(11):4273, May 2017. ISSN 0031-9155. . Publisher: IOP Publishing.
- Maziar Raissi, Paris G. Perdikaris, and George E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, February 2019. ISSN 00219991. .
- Liyue Shen, John Pauly, and Lei Xing. NeRP: Implicit Neural Representation Learning With Prior Embedding for Sparsely Sampled Image Reconstruction. *IEEE Transactions on Neural Networks and Learning Systems*, 35(1):770–782, January 2024. ISSN 2162-237X, 2162-2388. . Publisher: Institute of Electrical and Electronics Engineers (IEEE).
- Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit Neural Representations with Periodic Activation Functions. In *Advances in Neural Information Processing Systems*, volume 33, pages 7462–7473. Curran Associates, Inc., 2020.
- Shanlin Sun, Kun Han, Deying Kong, Hao Tang, Xiangyi Yan, and Xiaohui Xie. Topology-Preserving Shape Reconstruction and Registration via Neural Diffeomorphic Flow. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20845–20855, 2022.
- Shanlin Sun, Kun Han, Chenyu You, Hao Tang, Deying Kong, Junayed Naushad, Xiangyi Yan, Haoyu Ma, Pooya Khosravi, James S. Duncan, and Xiaohui Xie. Medical image registration via neural fields. *Medical Image Analysis*, 97:103249, October 2024. ISSN 1361-8415. .
- Lin Tian, Hastings Greer, Raúl San José Estépar, Roni Sengupta, and Marc Niethammer. NePhi: Neural Deformation Fields for Approximately Diffeomorphic Medical Image Registration. In Aleš Leonardis, Elisa Ricci, Stefan Roth, Olga Russakovsky, Torsten Sattler, and Gül Varol, editors, *Computer Vision – ECCV 2024*, pages 213–237, Cham, 2025. Springer Nature Switzerland. ISBN 978-3-031-73223-2. .

- Matthias Wilms, René Werner, Jan Ehrhardt, Alexander Schmidt-Richberg, Heinz-Peter Schlemmer, and Heinz Handels. Multivariate regression approaches for surrogate-based diffeomorphic estimation of respiratory motion in radiation therapy. *Physics in Medicine and Biology*, 59(5):1147–1164, March 2014. ISSN 1361-6560. .
- Matthias Wilms, René Werner, Tokihiro Yamamoto, Heinz Handels, and Jan Ehrhardt. Subpopulation-based correspondence modelling for improved respiratory motion estimation in the presence of inter-fraction motion variations. *Physics in Medicine & Biology*, 62(14):5823, June 2017. ISSN 0031-9155. . Publisher: IOP Publishing.
- Jelmer M. Wolterink, Jesse C. Zwienenberg, and Christoph Brune. Implicit Neural Representations for Deformable Image Registration. In *Proceedings of The 5th International Conference on Medical Imaging with Deep Learning*, pages 1349–1359. PMLR, December 2022.
- Hanguang Xiao, Xufeng Xue, Mi Zhu, Xin Jiang, Qingling Xia, Kai Chen, Huanqi Li, Li Long, and Ke Peng. Deep learning-based lung image registration: A review. *Computers in Biology and Medicine*, 165:107434, October 2023. ISSN 0010-4825. .
- Tokihiro Yamamoto, Sven Kabus, Jens Von Berg, Cristian Lorenz, Melody P. Chung, Julian C. Hong, Billy W. Loo, and Paul J. Keall. Reproducibility of Four-dimensional Computed Tomography-based Lung Ventilation Imaging. *Academic Radiology*, 19(12):1554–1565, December 2012. ISSN 10766332. .
- Laurent Younes. *Shapes and Diffeomorphisms*, volume 171 of *Applied Mathematical Sciences*. Springer, Berlin, Heidelberg, 2019. ISBN 978-3-662-58495-8 978-3-662-58496-5. .
- Chulong Zhang and Xiaokun Liang. INR-LDDMM: Fluid-based Medical Image Registration Integrating Implicit Neural Representation and Large Deformation Diffeomorphic Metric Mapping, November 2023. arXiv:2308.09473 [eess].
- Pengpeng Zhang, Margie Hunt, Arina B. Telles, Hai Pham, Michael Lovelock, Ellen Yorke, Guang Li, Laura Happersett, Andreas Rimner, and Gig Mageras. Design and validation of a MV/kV imaging-based markerless tracking system for assessing real-time lung tumor motion. *Medical Physics*, 45(12):5555–5563, December 2018. ISSN 2473-4209. .
- You Zhang, Hua-Chieh Shao, Tinsu Pan, and Tielige Mengke. Dynamic cone-beam CT reconstruction using spatial and temporal implicit neural representation learning (STINR). *Physics in Medicine & Biology*, 68(4):045005, February 2023. ISSN 0031-9155, 1361-6560. . Publisher: IOP Publishing.
- Veronika A. Zimmer, Kerstin Hammernik, Vasiliki Sideri-Lampretsa, Wenqi Huang, Anna Reithmeir, Daniel Rueckert, and Julia A. Schnabel. Towards Generalised Neural Implicit Representations for Image Registration. In Anirban Mukhopadhyay, Ilkay Oksuz, Sandy Engelhardt, Dajiang Zhu, and Yixuan Yuan, editors, *Deep Generative Models*, pages 45–55, Cham, 2024. Springer Nature Switzerland. ISBN 978-3-031-53767-7. .