Received: 7 February 2023 | Revised: 4 July 2023 | Accepted: 6 July 2023

DOI: 10.1002/mp.16731

RESEARCH ARTICLE

MEDICAL PHYSICS

Anatomy-aware computed tomography-to-ultrasound spine registration

Mohammad Farid Azampour[1,2] | **Maria Tirindelli**[1,3] | **Jane Lameski**[1] | **Miruna Gafencu**[1] | **Eleonora Tagliabue**[4] | **Emad Fatemizadeh**[2,4] | **Ilker Hacihaliloglu**[5] | **Nassir Navab**[1]

[1]Chair for Computer Aided Medical Procedures & Augmented Reality, Technical University of Munich, Munich, Bavaria, Germany

[2]Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran

[3]ImFusion GmbH, Munich, Bavaria, Germany

[4]Department of Computer Science, University of Verona, Verona VR, Italy

[5]Department of Radiology, Department of Medicine, University of British Columbia, Vancouver, Canada

Correspondence
Mohammad Farid Azampour, Chair for Computer Aided Medical Procedures, and Augmented Reality, Technical University of Munich, Boltzmannstraße 15, 85748, Garching bei Muenchen, Bavaria, Germany.
Email: mf.azampour@tum.de

Funding information
Qatar National Research Fund, Grant/Award Number: NPRP-11S-1219-170106; Bayerische Forschungsstiftung, Grant/Award Number: DOK-180-19

Abstract

Background: Ultrasound (US) has demonstrated to be an effective guidance technique for lumbar spine injections, enabling precise needle placement without exposing the surgeon or the patient to ionizing radiation. However, noise and acoustic shadowing artifacts make US data interpretation challenging. To mitigate these problems, many authors suggested using computed tomography (CT)-to-US registration to align the spine in pre-operative CT to intra-operative US data, thus providing localization of spinal landmarks.

Purpose: In this paper, we propose a deep learning (DL) pipeline for CT-to-US registration and address the problem of a need for annotated medical data for network training. Firstly, we design a data generation method to generate paired CT-US data where the spine is deformed in a physically consistent manner. Secondly, we train a point cloud (PC) registration network using anatomy-aware losses to enforce anatomically consistent predictions.

Methods: Our proposed pipeline relies on training the network on realistic generated data. In our data generation method, we model the properties of the joints and disks between vertebrae based on biomechanical measurements in previous studies. We simulate the supine and prone position deformation by applying forces on the spine models. We choose the spine models from 35 patients in VerSe dataset. Each spine is deformed 10 times to create a noise-free data with ground-truth segmentation at hand. In our experiments, we use one-leave-out cross-validation strategy to measure the performance and the stability of the proposed method. For each experiment, we choose generated PCs from three spines as the test set. From the remaining, data from 3 spines act as the validation set and we use the rest of the data for training the algorithm.

To train our network, we introduce anatomy-aware losses and constraints on the movement to match the physics of the spine, namely, rigidity loss and bio-mechanical loss. We define rigidity loss based on the fact that each vertebra can only transform rigidly while the disks and the surrounding tissue are deformable. Second, by using bio-mechanical loss we stop the network from inferring extreme movements by penalizing the force needed to get to a certain pose.

Results: To validate the effectiveness of our fully automated data generation pipeline, we qualitatively assess the fidelity of the generated data. This

assessment involves verifying the realism of the spinal deformation and subsequently confirming the plausibility of the simulated ultrasound images. Next, we demonstrate that the introduction of the anatomy-aware losses brings us closer to state-of-the-art (SOTA) and yields a reduction of 0.25 mm in terms of target registration error (TRE) compared to using only mean squared error (MSE) loss on the generated dataset. Furthermore, by using the proposed losses, the rigidity loss in inference decreases which shows that the inferred deformation respects the rigidity of the vertebrae and only introduces deformations in the soft tissue area to compensate the difference to the target PC. We also show that our results are close to the SOTA for the simulated US dataset with TRE of 3.89 mm and 3.63 mm for the proposed method and SOTA respectively. In addition, we show that our method is more robust against errors in the initialization in comparison to SOTA and significantly achieves better results (TRE of 4.88 mm compared to 5.66 mm) in this experiment.

Conclusions: In conclusion, we present a pipeline for spine CT-to-US registration and explore the potential benefits of utilizing anatomy-aware losses to enhance registration results. Additionally, we propose a fully automatic method to synthesize paired CT-US data with physically consistent deformations, which offers the opportunity to generate extensive datasets for network training.

The generated dataset and the source code for data generation and registration pipeline can be accessed via https://github.com/mfazampour/medphys_ct_us_registration.

KEYWORDS

anatomy-aware deep learning, physics-based data generation, point cloud, registration, spine, ultrasound

1 | INTRODUCTION

Lumbar spine injections are commonly performed in clinical practice to treat chronic back pain via the injection of steroids or anesthetics at the level of the inflamed regions.[1] Ultrasound (US) guidance has demonstrated to be an effective guidance technique for providing visual feedback during the procedure, increasing its effectiveness compared to blind palpation. Unlike Fluoroscopy, US can provide 3D information, while preventing both surgeon and patient from exposure to ionizing radiation.[2,3] However, interpretation of spine US data is challenging due to the presence of artifacts caused by the spine curvature and the shadowing effects produced by bony structures. To address these issues, several authors have proposed to use 3D-to-3D multimodal registration to align pre-operative CT or spine atlases to intra-operative, 3D US volumes. Volumetric US scans are obtained by compounding 2D B-mode US images acquired with a tracked transducer.[4,5] Spine alignment in CT-to-US data is relevant, as spine deformations are likely to occur between these two modalities due to the different patient positioning used to acquire the data (prone for CT and supine for US)

Spine CT-to-US registration for spinal injections are typically formulated as PC registration between bone surfaces extracted on the CT and US data. PC alignment can be solved using standard techniques as iterative closest point (ICP), coherent point drift (CPD)[6–10] or Kalman filtering (KF).[11–13] Notably, Nagpal et al.[14] proposed a variation of the standard CPD method, where a biomechanical constraint is added as a regularization term in the iterative optimization, so to account for the relative motions between vertebrae. Instead of applying the standard CPD on the whole lumbar spine, the authors suggest splitting vertebrae in the CT volume and considering one vertebra at a time in the CPD optimization. Specifically, at each step of the iterative CPD registration, they allow the motion of one vertebra at a time in the CT point cloud. Moreover, to avoid unnatural deformations of the spine, they propose to constrain the relative motions between vertebrae in a way to avoid unrealistic deformations, such as deformations where consecutive vertebrae compenetrate each other or excessively drift from each other. To do so, they add virtual springs between vertebrae and use the spring energy generated by the vertebrae displacement as a regularization term in the registration optimization. The points connecting the springs are selected manually on the CT point cloud. The spring energy, that is, the regularization term, is therefore given by $\alpha|(l_{new} - l_{initial})|$ where l_{new} is the length of the spring at the given optimization step and $l_{initial})|$ is the initial length of the spring. In their work, they prove the advantages of using a biomechanical-constrained multi-body registration for CT-to-US spine registration, showing an improvement

over state of art methods. In this paper, we denote the work of Nagpal and colleagues as multi-body CPD (MB-CPD) and compare against it in the experiments and results section.

Other recent approaches tried to solve this registration problem in the image domain using intensity based optimization. Masoumi et al.[15] suggested using a patch-wise normalized cross correlation (NCC) or correlation ratio (CR). They introduced a dataset acquired on an ex-vivo phantom and measured the efficacy of their proposed algorithm for rigidly registering the US and CT in the image domain. Furthermore, Gueziri et al.[16] introduced a fast rigid registration approach that works close to real time by moving the time consuming optimization part to GPU. In their method a single caudo-cranial axial sweep procedure is used to acquire intra-operative ultrasound images. This scan trajectory is then utilized to initiate the registration transform. Subsequent refinement of this transform is achieved by identifying the posterior vertebra surface locations, which are then used to compute gradient-based alignment in the image domain.

With the advent of DL, several authors have proposed using convolutional networks for solving the PC registration problem.[17,18] Although not yet employed for CT-US spine registration, deep learning methods have shown improved performance compared to non-deep learning techniques with increased robustness against convergence to local minima.[17] However, Deep Learning methods require a large amount of data to be trained, which is not straightforward in the medical context. To tackle the data scarcity problem, recently many works focused on exploiting data simulation for other tasks such as image segmentation for ultrasound imaging.[19,20] Furthermore, DL-based point cloud registration methods are typically designed for computer vision dataset rather than medical ones. Therefore, they typically do not account for the anatomical prior and mechanical behavior of the target anatomy. However, it has been shown in non DL methods that regularizing the registration via the introduction of biomechanical constraints yields improve performance.[7,13] DL approaches also show improvement in the performance and feasibility of the inferred deformation field when biomechanical constraints are added to the training phase, for example, for magnetic resonance imaging (MRI) to CT images registration[21] or registration of images of the prostate.[18]

In this paper, we propose a deep-learning-based pipeline for CT-to-US spine registration and a novel fully automatic data generation method, to effectively generate paired CT-US images for network training. To the best of our knowledge, this is the first DL-based method for CT-to-US spine registration, accounting for anatomical priors of the spine. Both the code and the generated dataset will be made publicly available upon publication. The contribution of this work is twofold:

1. We propose a pipeline for CT-to-US registration. Furthermore, we introduce two anatomy-aware loss functions for network training, accounting for the anatomical priors of the spine biomechanics.
2. We propose a method for generating synthetic paired CT-US of the same patient spines with different deformations, ensuring: *i*. Realistic deformations of the spine *ii*. Realistic Point Cloud distribution.

2 | METHOD AND MATERIALS

2.1 | Dataset generation

One of the major bottlenecks in training deep networks for medical applications is the ubiquitous problem of lack of data. For our application, we need paired CT-US images from the same patient. Gathering such data can be challenging due to the time-consuming process of getting ethical approval and the often limited pool of patients who provide their consent. To overcome this issue, we propose a novel fully automatic data generation/data augmentation method, allowing to generate intra-patient paired CT-US data, starting from CT data and accounting for deformation arising from different patient positioning. Unlike spinal US or paired CT-US datasets, there exists several publicly available abdominal CT datasets, containing large amount of data.[1,22] In this work, we selected 35 volumes from the VerSe Dataset,[1] where the full lumbar spine is visible and no fixation screw is present. The dataset consists of the CT volumes together with the vertebra segmentation and classification. The resolution of the images differ from one patient to another. However, for all the of the images the voxel size is smaller than 1 mm^3 . The intra-slice spacing ranges between 0.2 to 0.9 mm and the inter-slice spacing ranges between 0.6 to 1.0mm.

To generate data using only the CT images, two challenges have to be addressed. First, we need to generate physically consistent spine deformations in the CT data. Second, we need to simulate US data of the deformed CT. In Figure 1a flowchart of the data generation pipeline is reported. In what follows we describe the two steps Deformation modeling and US simulation in more detail.

2.1.1 | Deformation modeling

The position in which a spinal injection is performed can vary depending on the specific technique used and the preference of the healthcare professional. Three common positions are supine position, sitting position and lateral decubitus position which is lying on the side while holding the knees. Sitting position is used mostly for epidural injections while the lateral and supine is used for thoracic, cervical or facet joint injections. These

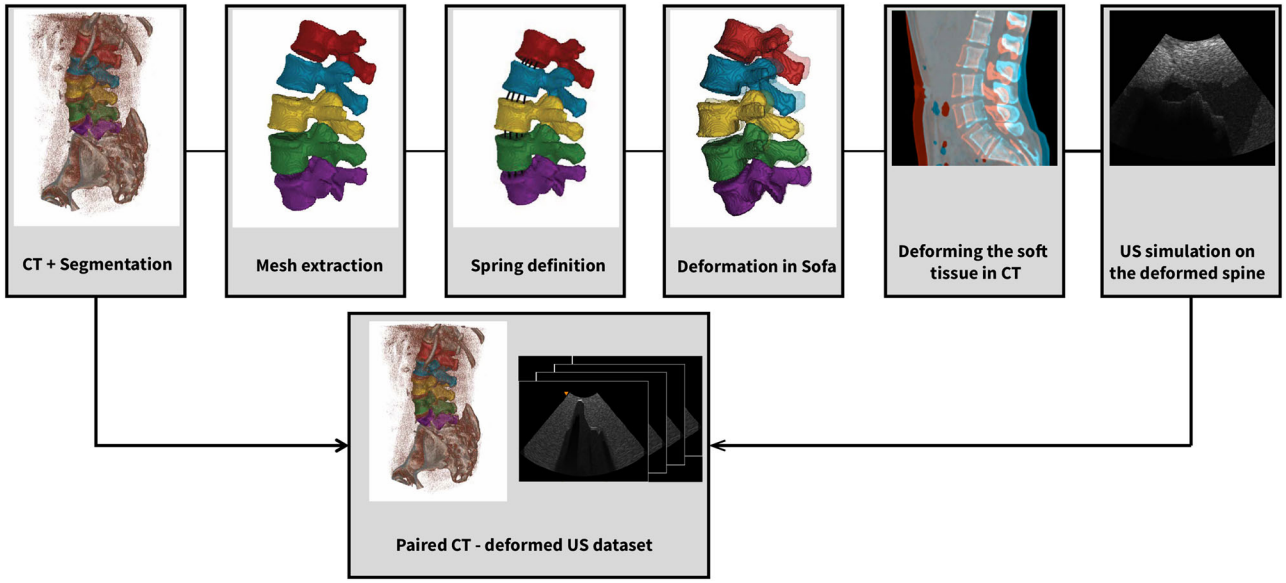


FIGURE 1 The paired CT-US generation pipeline.

positions allow for the spine to curve naturally and provide easier access to the injection site. In what follows, we show how we can realistically simulate these positions in CT domain starting from a CT image taken in the prone position.

To this end, we define a simplified physical model of the spine. We model bones as rigid tissues. The intervertebral fluid at the level of facet joints and disks is modeled by a set of springs in parallel. The process of spring definition occurs automatically by determining points on each vertebrae body and facets that connect neighboring vertebrae. The process is as follows:

Disk springs

For each vertebra mesh $\mathcal{V}_i \in \{L_1, L_2, L_3, L_4, L_5\}$ we first determine the points that belong to each of the vertebral body faces. We compute the centroid c_i, which is usually situated in the spinal canal, right above the vertebral body. Therefore, we consider all points below this point to belong to the vertebral body $B(\mathcal{V}_i) = \{p_i \in \mathcal{V}_i \mid y_{p_i} < y_{c_i}, p_i \in \mathbb{R}^3\}$. Because we are only interested in points that are on the surface of the two vertebral body sides denoted as F_1 and F_2, we divide the points from $B(\mathcal{V}_i)$ based on the z_{c_i}, the z coordinate of the centroid c_i and additionally enforce constraints regarding the normal vector n_{p_i} of point p_i: $F_1 = \{p_i \in B(\mathcal{V}_i) \mid z_{p_i} < z_{c_i},\ n_{p_i} = (0, 0, -1),\ p_i \in \mathbb{R}^3\}$ and $F_2 = \{p_i \in B(\mathcal{V}_i) \mid z_{p_i} > z_{c_i},\ n_{p_i} = (0, 0, 1),\ p_i \in \mathbb{R}^3\}$. Finally, we shuffle all points in F_1 and F_2 and define the inter-body (IB) springs between neighboring vertebrae $\{\mathcal{V}_i, \mathcal{V}_j\} \in \{\{L_1, L_2\}, \{L_2, L_3\}, \{L_3, L_4\}, \{L_4, L_5\}\}$ by the two end points: $IB = \{\{e_1, e_2\} \mid e_1 \in F_1(\mathcal{V}_j), e_2 \in F_2(\mathcal{V}_i)\}$. $F_1(\mathcal{V}_j)$ and $F_2(\mathcal{V}_i)\}$. These can be visualized in Figure 2.

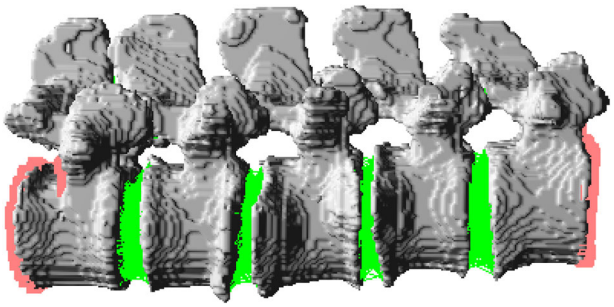


FIGURE 2 Automatically generated springs in between vertebral bodies colored in green.

Facet springs

We first determine the facet points. For this purpose, we create a set of pairwise neighboring vertebrae $\{\mathcal{V}_i, \mathcal{V}_j\} \in \{\{L_1, L_2\}, \{L_2, L_3\}, \{L_3, L_4\}, \{L_4, L_5\}\}$. Since the facets from two neighboring vertebrae are usually situated very close to each other, we define the facet points for $\mathcal{V}_i$ as the points having the least distance from the neighboring vertebra. In practical terms, every point with distance less than 1 *mm* is considered to be a facet point: $Facet_{Left}(\mathcal{V}_i) = \{p_i \in \mathcal{V}_i \mid \exists p_j \in \mathcal{V}_j \text{ if } d(p_i, p_j) < 1mm,\ x_{p_i} < x_{c_i}\}$ and $Facet_{Right}(\mathcal{V}_i) = \{p_i \in \mathcal{V}_i \mid \exists p_j \in \mathcal{V}_j \text{ if } d(p_i, p_j) < 1mm,\ x_{p_i} > x_{c_i}\}$, where c_i refers to the centroid of $\mathcal{V}_i$ and $d(p_i, p_j)$ is the Euclidean distance function. Afterwards, we shuffle all points in $Facet_{Left}(\mathcal{V}_i)$, $Facet_{Left}(\mathcal{V}_j)$, $Facet_{Right}(\mathcal{V}_i)$, $Facet_{Right}(\mathcal{V}_j)$ and define facet joint springs for each side by pairing two end points: $Joint_{Left} = \{e_1, e_2 \mid e_1 \in Facet_{Left}(\mathcal{V}_i),\ e_2 \in Facet_{Left}(\mathcal{V}_j)\}$ and $Joint_{Right} = \{e_1, e_2 \mid e_1 \in Facet_{Right}(\mathcal{V}_i),\ e_2 \in Facet_{Right}(\mathcal{V}_j)\}$. In Figure 3, we visualized these springs for a sample spine.

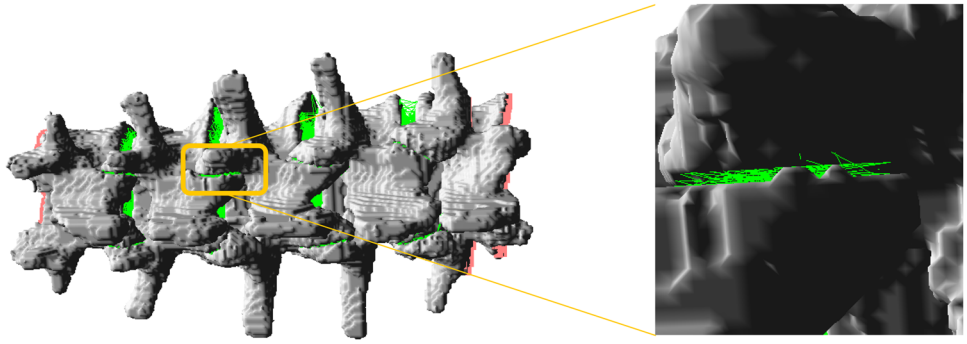


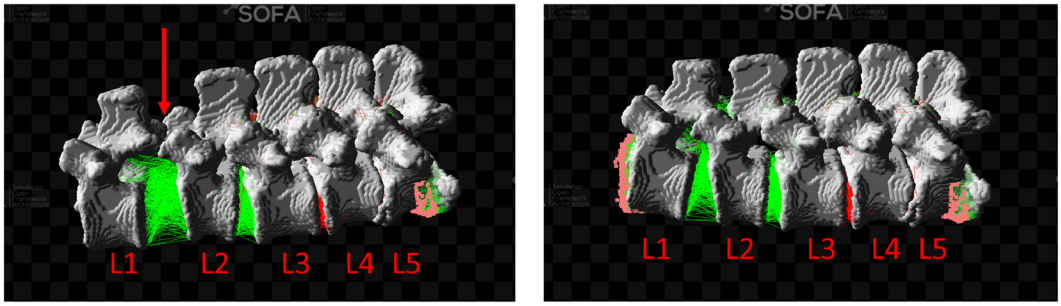
FIGURE 3 Automatically generated springs in between facets of neighboring vertebrae.

TABLE 1 Parameters used in SOFA framework.

	Spring stiffness (N/m)	Springs number (#)	Equivalent stiffness (MN/m)
Inter-vertebral fluid	500-1000	400-800	0.2-0.4
Facet joint	8000	200-500	0.16-4

As reported in Table 1, the number of springs at the level of the facet joints varies from 200 to 500 springs, where each spring has a stiffness of 8000 N/m, for an overall equivalent stiffness ranging from 0.6 and 2 MN/m. Each intervertebral disk is connected through a number of vertebrae ranging from 400 to 800 springs, with a stiffness of 500 N/m, making the overall stiffness of the springs varies between 0.16 and 4 MN/m, consistent with the literature.[23–28] The damping coefficients of the springs were set to 3 N/s for the inter-body fluid springs and 500 N/s for the facet joint springs. The applied forces varied up to 0.7 N. The selection of the correct stiffness (i.e., the selection of the number of springs and springs stiffness) is crucial to obtain realistic deformations, as depicted in Figure 4, where it is shown how an incorrect selection of the stiffness at the level of the facet joint results in an unrealistic deformation.

To control the overall motion of the spine, we connect L1 and L5 to a still constraint through a set of springs in parallel, simulating the vertebrae connections to T12 and S1. The physical simulation of the spine deformation is performed using the SOFA framework.[29] We apply forces on the individual vertebrae along the anterior-posterior axis. On L3 we apply the largest force and reduce it gradually towards L1 and L5. To take into consideration multiple possible spine curvatures, for each spine we generate one deformation along the anterior axis and one in the posterior direction. While an anterior deformation has the effect of increasing the distances between the spinous processes, the posterior deformation has the opposite effect. To expand the deformation space and cover multiple scenarios, we define force intervals for each vertebra as shown in the second row of Table 2. We mention that forces along lateral axes account for the slight shifts in patient's positioning. Moreover, to further allow for multi-tissue US simulation, we automatically also segment the soft tissue, muscle, and fat in the CT images using the TotalSegmentator.[31] We extrapolate the deformation field calculated on the spine to the surrounding soft tissue using a radial basis function (RBF) approximator with thin plate spline as the kernel. The resulting dense deformation field is then applied to the CT image. We thus obtain a complete, multi-tissue deformed CT image which



(a) (b)

FIGURE 4 Comparison between resulting spine deformation when (a) The stiffness is not correctly selected (i.e., missing springs at the level of L1 facet), yielding unrealistic spine deformation, where the vertebrae are detached from one another (highlighted with the red arrow) (b) The stiffness is correctly selected, yielding realistic spine deformations.

TABLE 2 Force intervals applied on the diverse lumbar vertebrae along the three main axes that consider different spine curvature during injection. F_i is the force applied on L_i, $i \in \{1, 2, 3, 4, 5\}$.

Axes	F1(N/m)	F2(N/m)	F3(N/m)	F4(N/m)	F5(N/m)
Lateral (X)	[−10,9]	[−10,9]	[−10,9]	[−10,9]	[−10,9]
Anterior/posterior (Y)	[±10, ±15]	[±15, ±20]	[±30, ±50]	[±15, ±20]	[±10, ±15]
Superior/inferior (Z)	0	0	0	0	0

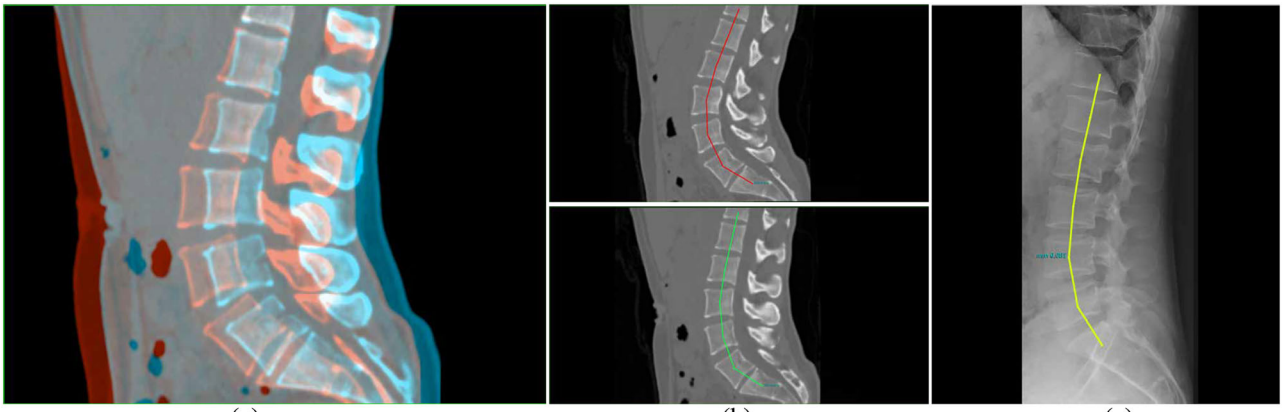


FIGURE 5 Overlay and side-by-side view of a sample CT of a patient, before and after deforming and a comparison with a x-ray from lateral decubitus position. (a) shows the overlay of the image before (red) and after (blue) applying the deformation. In (b), the spine curvature between the two states is visualized. We can see how the spine and the soft tissue around it deformed to account for the shape deformation happening between supine and prone position. We can see that although the deformation is quite large (around 20 mm), the generated physics-aware deformation respects the boundaries of the vertebrae and there is no collision between the vertebrae after deformation. Furthermore, (c) shows an example of an x-ray during a spinal injection in pre-injection state (Case courtesy of Ian Bickle, Radiopaedia.org, rID: 47355) with its spinal curvature annotated. By comparing the spinal curves, we can see that the spinal curvature after deformation (green) is closer than the original one (red) to the pre-injection x-ray shape (shown in yellow).

results from changes in spine curvature as shown in Figure 5.

2.1.2 | US simulation

For US simulation we use the simulation pipeline described in [32], implemented in ImFusion[1]. This simulation pipeline involves a ray-tracing engine at its core. The engine builds the data for each radio frequency scanline separately, defining the recorded echo for each scanline as the sum of two main terms: the reflected energy from tissue boundaries and the backscattered energy from scattering points throughout the scanline. Reflection is modeled as a function of the angle of incidence, ultrasound wave amplitude, and the acoustic impedance difference between two adjacent tissues. As ultrasound wave traverses through tissue, it attenuates and can cause artifacts like shadowing and speckles. The remaining energy of the sound beam is modeled using the Beer-Lambert Law, where the initial energy, wave frequency, and attenuation coefficient of the medium are taken into account. The other term in the returned echo, the back-scattered energy, is the product of the remaining wave amplitude and the convolution of the point spread function with random scatterers. The scatterers are generated from a generative model based on two random textures combined with tissue-specific parameters μ_0, μ_1, and σ_0. The model generates scatterers of various spatial and acoustic densities for each tissue type.

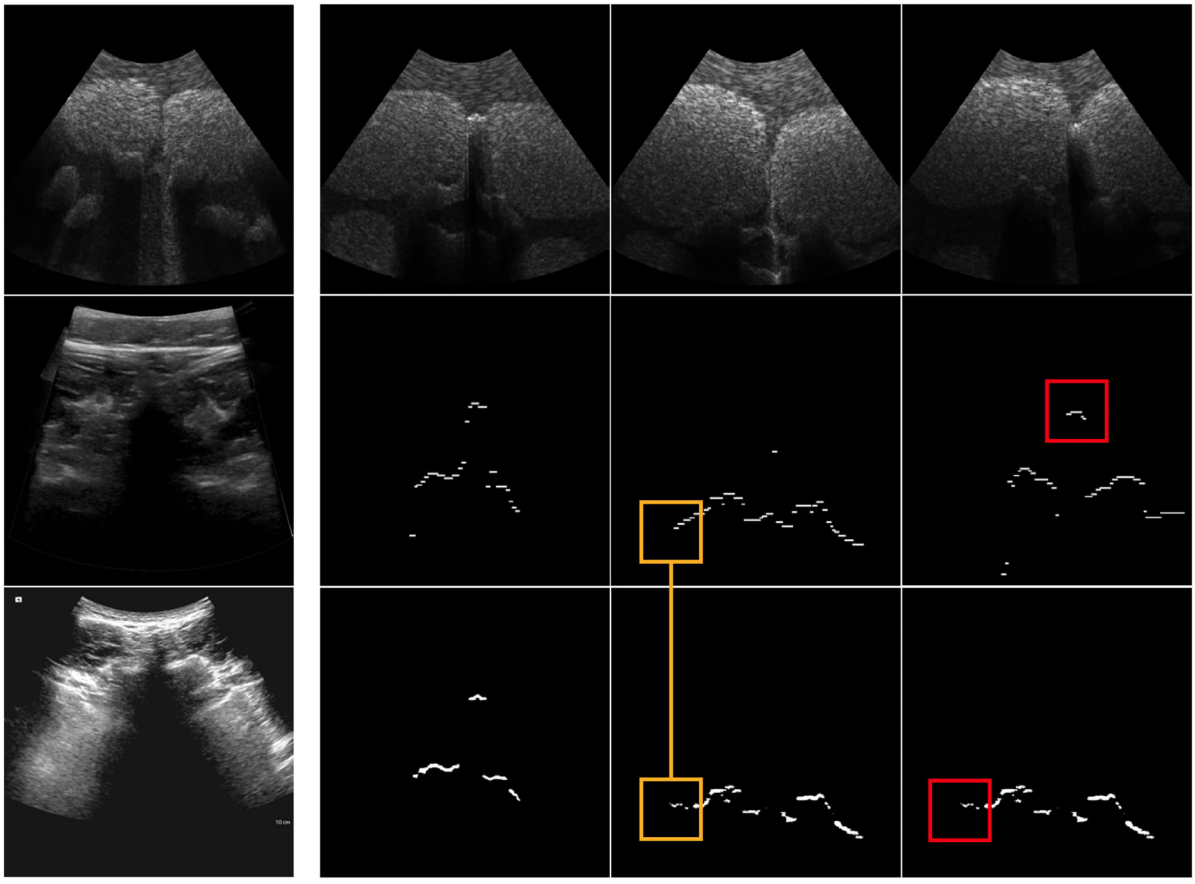
We use a 40 mm curvilinear probe with 100 mm depth, 55 mm focus depth, 2 MHz transmit frequency, and 512 scan lines. We set the time gain compensation (TGC) to 0.3. To achieve more realistic simulated ultrasound, we need to consider the soft tissue surrounding the spine as well. We use the multi-tissue deformed CT image as a starting point for the US simulation. In Table 3 we report the selected tissue-specific parameters. Z, c and a are impedance, speed of sound and attenuation coefficients and μ_0, μ_1, and σ are additional parameters to control scatterer properties in the simulation.

The trajectory of the probe is defined as a straight line moving from the highest point in the sacrum to the lowest point of T11 along the spine direction, with the probe in a transverse orientation. Figure 6 shows the result for 4 slices of the generated ultrasound and show a qualitative comparison between a simulated frame

[1] ImFusion GmbH, Munich, Germany

TABLE 3 Parameters used for US simulation.

	$Z\left(\frac{g}{cm^2}\right)$	c (m/s)	$a\left(\frac{dB}{MHz}\times cm\right)$	$_{0}$	$_{1}$	α_0
Soft Tissue	163000	1540	0.54	0.52	0.5	0
Bone	612000	1800	3	0.77	0.56	0.10
Muscle	163000	1568	1.09	0.52	0.50	0
Fat	102000	1530	0.69	0.80	0.5	0



(a) (b)

FIGURE 6 Qualitative comparison of simulated and real ultrasound images and segmentation results. (a) represents three images, first is the simulated ultrasound, second row is an ultrasound image acquired by a Zonare z.one ultra sp Convertible Ultrasound System (ZONARE Medical Systems, Inc., Mountain View, California, United States) and the third row is recorded using an ACUSON Juniper ultrasound system (Siemens Industry, Erlangen, Germany). Although there is a domain gap between real and simulated images we can see that the prominent effects are preserved. Furthermore there is a large intra-domain distance in real images depending on the probe and the acquisition parameter. (b) Three simulated ultrasounds from the validation set, the ground truth label and the output of the network is visualized here. We can see that the network output mostly matches the ground truth label with minor deviation of the shape (indicated by orange) or small parts missing or added (in red).

and two real spinal ultrasound frames acquired using different devices.

2.2 | Registration pipeline

In Figure 7, we provide a flowchart of the proposed registration pipeline. Firstly, the CT is segmented, and the posterior points of the spine are extracted to discard points that are not visible in the US data, thus improving the registration result according to.[7,13] The posterior bone surface extraction is performed by letting a set of parallel rays travel in the Anterior-Posterior direction in the CT-labelmap, starting from the posterior side. Whenever one of these rays encounters a non-zero voxel, the position of that voxel is stored as a surface point in the resulting posterior surface CT point cloud. Secondly, bones are automatically segmented on US data, using a U-net network architecture[33] trained on the generated US data. Thirdly, the extracted labelmaps and

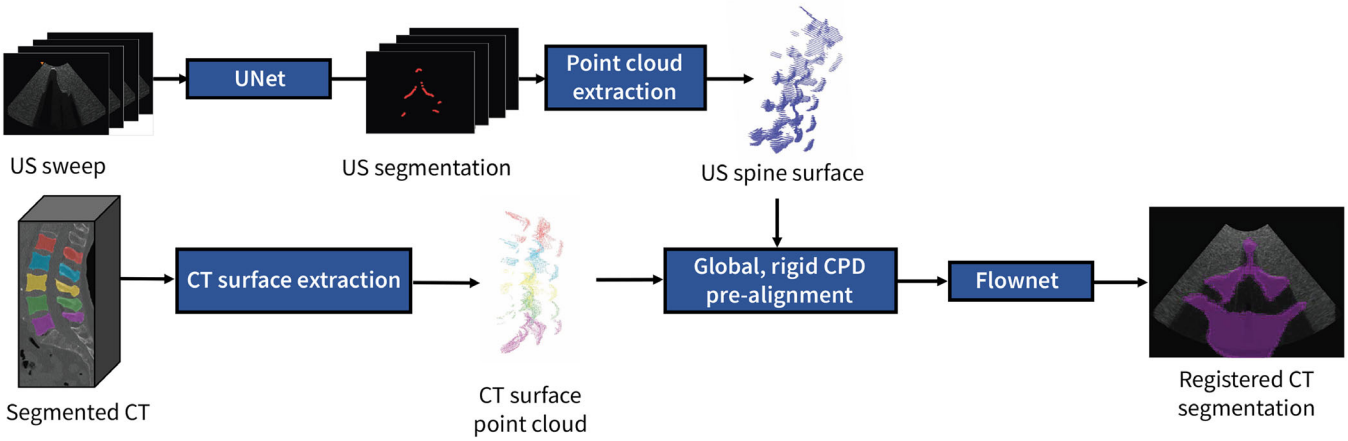


FIGURE 7 Registration Pipeline: On the one side, the US image is segmented using a U-net network. The segmentation is combined with the tracking stream of the US probe to generate point clouds. On the other side, the CT segmentation is raycasted and converted to a point cloud. The extracted US and CT point clouds are then fed into the registration network.

the US tracking information are used to generate point clouds of the bones data. Finally, the CT and US point clouds are registered together using a registration network, where anatomical consistent transformations are enforced via the introduction of anatomy-aware losses. The CT point cloud is considered the moving set and the Ultrasound point cloud the target set. A detailed description of the anatomy-aware losses as well as of the network architecture is provided in sections 2.2.1 and 2.2.2.

2.2.1 | Anatomy-aware losses

Deep deformable registration usually suffers from estimating unrealistic deformation fields.[34] To alleviate this problem, we introduce anatomy-aware losses that penalize the network based on the anatomical constraints governing spinal movement. These constraints are twofold. First, the vertebral body can only move rigidly. Second, the difference between rigid transformations of neighboring vertebrae is controlled and limited by the joints.

Considering these two constraints, we define two losses to regularize the deformable registration, the rigidity loss, and the bio-mechanical loss. We define rigidity loss ($\mathcal{L}_{rigid}$) as the difference between the sum of the $l2$-norm of all the edges (e_j) belonging to a vertebra i (M_{v_i}) before and after transformation. M_{v_i} is the set of all edges between each two point belonging to vertebra i. The rigidity loss encourages deformations that are locally Euclidean.

$$\mathcal{L}_{rigid} = \sum_{i=1}^{L} \left(\sum_{e_j \in M_{v_i}} \frac{1}{N_i} \left(\|e_j\| - \|\hat{\phi}(e_j)\| \right)^2 \right)^{0.5} \quad (1)$$

where L is the total number of vertebrae we want to register, N_i is the number of edges defined between points of the given vertebra v_i, $\hat{\phi}$ is the estimated transformation by the network for all of points in the point cloud and $\|\hat{\phi}(e_j)\|$ is the norm of edge e_j after applying the estimated transformation $\hat{\phi}$.

The bio-mechanical loss ($\mathcal{L}_{bio}$) considers the force required to achieve a certain pose on the spine. We model the joints ($J_{i,l}$) between two neighboring vertebra (v_i and v_l) by springs with different stiffness $k_{J_{i,l}}$. We set the rest length for each spring (x_{rest}) equal to the initial distance between the points at the two ends of the corresponding joint. $\mathcal{L}_{bio}$ is defined as the sum of the potential energy of the springs after the transformation.

$$\mathcal{L}_{bio} = \sum_{i=1}^{L} \sum_{l=i}^{L} \frac{1}{2} k_{J_{i,l}} \left\| x_{\hat{\phi}} - x_{rest} \right\|^2 \quad (2)$$

where $x_{\hat{\phi}}$ is the distance between the same points considered for calculating the resting distance after applying $\hat{\phi}$.

2.2.2 | Network architecture

For the network architecture, we use the FlowNet3D[17] as the base architecture, which has shown promising results for medical data registration.[18] The network architecture consists of a feature extraction layer, inspired by the PointNet++ architecture[35] followed by flow embedding, downsampling, and upsampling layers. The feature extraction layer is composed of a set of hierarchical abstraction layers, that extract local features in the point cloud. Compared to another feature extraction architecture, the capability of PointNet++ to encode local features in the point cloud is highly

relevant for the task of spine registration. For this task, it is crucial that the network not only captures the global appearance of the whole spine, but also the appearance and relative position of local features as vertebrae or vertebral landmarks. Once a set of hierarchical features is extracted from both source and target point clouds, the flow embedding layer encodes the motion of the points based on both their features and relative positions in space. Finally, downsampling and upsampling convolutional layers ensure the smoothness of the extracted flow. In practical terms, the structure includes four sets of convolutional layers, one layer for flow embedding, and four sets of up-convolutional layers (matching the four sets of convolutional layers). Additionally, there is a final linear flow regression layer that generates the predicted scene flow in three-dimensional space. The set of up-convolutional layers are equipped with skip connections. Each trainable layer utilizes multi-layer perceptrons, employing Linear-BatchNorm-ReLU layers with adjustable widths for its linear layer.

2.3 | Experiments

2.3.1 | Segmentation network

We trained the U-net segmentation network for 50 epochs using Adam optimizer and binary cross-entropy (BCE) loss with an 80%–20 % split between training and validation set and achieved a dice score of 0.58. The labeling for the input US data was automatically generated by interpolating the posterior bone surface extracted from the CT labels onto the US data. In Figure 6 , qualitative results on segmenting simulated ultrasound are depicted. We can see from the results that the network output is of adequate quality to be used in the pipeline.

2.3.2 | Registration metrics

To evaluate the network performance, we considered three different metrics: The TRE, the translation and absolute quaternion distance between the ground truth, and estimated transformation matrices for each vertebra. For computing the TRE, we first manually selected points on the facet joints in the source point clouds. We then compute the TRE as the distance between these points transformed with the ground truth and predicted transformations. As the registration network provides a displacement for each point in the source point cloud, we used a Least Square Fitting method to estimate the best rigid transformation $\hat{T}_i$ between each vertebra in the source point cloud and its deformed target. For a given spine, the translation and absolute quaternion distance (AQD) are given by

$$d_{translation} = \frac{1}{L}\sum_i \left\| t^i_{gt} - t^i_{pred} \right\|$$
$$d_{AQD} = \frac{1}{L}\sum_i \left(1 - \left\langle q^i_{gt}, q^i_{pred} \right\rangle\right) \quad (3)$$

where t^i_{gt}, t^i_{pred} are the translation terms in the ground truth and predicted transformation matrices of the i-th vertebra $\hat{T}_i$, respectively, and q^i_{gt}, q^i_{pred} are the quaternions associated with the rotation term in the ground truth and predicted transformation matrices of the i-th vertebra, respectively.

TABLE 4 Computation cost of each step of data generation pipeline and inference time of the network. The reported time for ultrasound simulation is for generating an US sweep with 50 frames. All of the experiments were done a machine with 128 GB of RAM, a AMD Ryzen Threadripper 1950X 16-Core Processor and an Nvidia RTX 3090.

Step	Time (ms)	Step	Time (ms)
Spring generation	1200	Ultrasound simulation	10000
Applying force on spine	500	Ground truth label generation	800
Deformation extrapolation	3000	Inference time	50

2.3.3 | Computation cost

We first talk about the time needed for generating the data and then about the inference time. In Table 4, we report the time needed for deforming one spine with one force field and finally generating one simulated ultrasound sweep with its corresponding point cloud.

2.3.4 | Ablation study and comparison with SOTA

Purpose: The purpose of this experiment is to evaluate the proposed losses and to compare the network results to non-DL SOTA methods in ideal conditions, that is, with ground-truth noise-free surface point clouds. To obtain these surface point clouds, we extract the posterior spine surface from the deformed spine labelmap (from which US simulations are generated) using the same procedure described in Section 2.2. By doing so, we bypass any noise that might be added during the segmentation process.

Implementation: We train the registration network for 60 epochs using Adam optimizer. The learning rate is set

to 0.001, and we choose the same model parameters as[17] except for the number of points in the abstraction layers, which we double based on our hyperparameter tuning. The model expects a set of 4096 points in each input point cloud which we randomly pick from the input point cloud. To better assess and represent the stability of the model against different splits of the data, we perform leave-one-out cross-validation experiments. We choose one spine for each experiment as the test set, and we randomly divide the rest of the data to have 18 and 3 spines in training and validation sets. During training, we select the best model based on the TRE score on the validation set.

We choose a combination of losses for training our network. First loss is the mean squared error (MSE) between the known ground truth deformation field and the prediction of the network. Second is the rigidity loss defined in Equation 1 and the third is the biomechanical loss as defined in Equation 2. The last loss is the Chamfer loss between the target PC and the warped PC. The Chamfer Loss is a metric used to measure the dissimilarity between two sets of points in a point cloud. It is commonly used in computer vision tasks such as shape matching, point cloud registration, and object detection.

Given two sets of points P and Q, each consisting of N and M points respectively, the Chamfer Loss $L_{\text{Chamfer}}(P, Q)$ is calculated as follows:

$$L_{\text{Chamfer}}(P, Q) = \sum_{p \in P} \min_{q \in Q} \|p - q\|_2^2 + \sum_{q \in Q} \min_{p \in P} \|q - p\|_2^2, \tag{4}$$

where $\| \cdot \|_2$ represents the Euclidean distance between two points.

During hyperparameter tuning, we use TRE for evaluating the network performance. We notice that the choice of the coefficients for Chamfer loss does not significantly affect the network performance. However, performance varies with the coefficient range for rigidity and biomechanical losses. The acceptable coefficient range is [1, 20] and [0.2 1.5] for rigidity and biomechanical loss, respectively. Based on our experiments, we set the coefficients to 10.0, 5.0, and 0.5 for Chamfer, rigidity, and biomechanical loss, respectively.

We perform an ablation study on a set of training losses: MSE only, MSE with Chamfer distance, MSE with rigidity loss, MSE with biomechanical loss and MSE with rigidity and biomechanical loss. We further compare the model results to the multi-body CPD (MB-CPD) method proposed in [7]. We analyze the effect of the springs on the results of the non-DL method and report the different results of the MB-CPD method when using intervertebral springs and in the absence of springs. Compared to the work proposed by [7], we introduce two additional springs connecting L1 and L5 to fixed constraints. This addition yields improved performance as it allows regularizing the displacement from their initial position, thus avoiding unrealistic transformation of L1 and L5. The average TRE before registering the data is 5.44 mm with standard deviation 2.31..

2.3.5 | Robustness against errors in initialization

Purpose: The purpose of this experiment is to analyze the effect of errors in initialization for both the MB-CPD and the proposed method.

Implementation: In our generated dataset, an initial alignment of high quality is inherently available as a default setting. To measure the robustness, we conducted an experiment involving the deliberate introduction of errors to this initial alignment. Specifically, we manually introduced rotational and translational errors around at least one of the axes, in the range of ± 10 degrees. Subsequently, we evaluated and compared the performance of both the proposed method and the MB-CPD on two distinct scenarios: correctly initialized data without errors, and data containing the introduced rotational errors.

2.3.6 | Dependency on the dataset size

Purpose: With this experiment we want to evaluate the dependency of the registration network results on the size of the training set.

Implementation: We evaluate the results of the model trained with "MSE + Bio" when using 50% and 66% of the training data for training.

2.3.7 | Network performance on US segmentation

Purpose: The purpose of this experiment is to evaluate the robustness of the network against partial occlusions and artifacts in the point clouds, arising from acoustic shadowing, as well as inaccuracies in the US segmentation.

Implementation: We evaluate the network performance, when trained with the best set of hyperparameters and losses (MSE+Biomechanical loss) and we compare against the best non-DL SOTA method.

3 | RESULTS

3.1 | Ablation Study and comparison with SOTA

In Table 5 we report the results of the ablation and the comparison study. We can see that the results with rigidity loss and the biomechanical loss tend to improve (although not significantly) compared to using MSE. For

TABLE 5 Ablation study on the registration network result when using the proposed losses compared to using the standard loss (MSE) on ground-truth noise-free data. Values after ± show the standard deviation.

Method	TRE (mm)	Translation (mm)	AQD	Rigidity loss
MSE	3.82 ± 0.60	4.65 ± 0.87	0.019 ± 0.032	1.15 ± 0.11
MSE + chamfer	3.88 ± 0.89	4.98 ± 1.06	0.023 ± 0.035	1.45 ± 0.15
MSE + rigidity	3.79 ± 0.78	4.65 ± 0.95	0.021 ± 0.028	0.83 ± 0.21
MSE + bio	3.67 ± 0.63	4.51 ± 0.90	0.018 ± 0.020	0.91 ± 0.18
MSE + bio + rigidity	3.76 ± 0.64	4.84 ± 0.87	0.023 ± 0.025	0.96 ± 0.17
MB-CPD	3.78 ± 2.37	7.53 ± 4.78	0.035 ± 0.032	-
MB-CPD + springs	3.46 ± 2.12	4.93 ± 3.38	0.037 ± 0.025	-

Abbreviations: MB-CPD, multi-body CPD; MSE, mean squared error.

TABLE 6 Robustness of the proposed method and the SOTA method to errors in initialization. The word *uninit* marks the rows presenting result on data with erroneous initialization. Values after ± show the standard deviation.

Method	TRE (mm)	Translation (mm)	AQD
MB-CPD (uninit)	5.66 ± 3.55	6.47 ± 3.24	0.044 ± 0.034
MB-CPD	3.46 ± 2.12	4.93 ± 3.38	0.037 ± 0.025
proposed method (uninit)	4.88 ± 1.65	5.79 ± 1.24	0.036 ± 0.028
proposed method	3.67 ± 0.63	4.51 ± 0.90	0.018 ± 0.020

Abbreviation: MB-CPD, multi-body CPD.

the rigidity loss and the biomechanical loss, the TRE is reduced by 0.03 and 0.15 and rigidity loss is reduced by 28% and 20%, respectively, when compared to MSE (Table 5). To assess the statistical significance of the proposed method, we perform a paired *t*-test and compare the variation of losses against "MSE + Bio" which shows that although the result improves, this improvement is not significant ($p \geq 0.05$). From the MB-CPD results we can notice that, as for the network, the presence of springs affects the overall performance of the method. The rigid loss is not given for the MB-CPD method, as it is inherently 0, given that the method provides a rigid transformation matrix for each vertebra. In the tables, we report the mean value and the standard deviation for each value. We calculate the *p*-values to compare "MB-CPD + Springs" and "MSE + Bio" which results in a non-significant superiority of the SOTA method over the proposed method. However it is worth noting that while for our proposed method, it only takes 50ms to register, for MB-CPD, this number varies between 10 to 15 s depending on the number of iterations. As can be seen from the results, our proposed method generates results close the SOTA while being orders of magnitude faster. In what follows, we compare the robustness of the two methods against errors in the initialization.

3.2 | Robustness against errors in initialization

Table 6 presents a comparative analysis of the impact of initialization on the results. The findings demonstrate that in contrast to MB-CPD, the proposed method exhibits greater resilience and is less susceptible to the adverse effects caused by errors in initialization. To support this claim, we employ a paired *t*-test and the result ($p = 0.02$) shows that the proposed method is significantly more robust compared to MB-CPD (TRE of 4.88 mm in comparison to 5.66 mm for MB-CPD without initialization, Table 6). This is due to the fact that during training the method, we can expose our network to different configurations of misalignment in the initialization.

3.3 | Dependency on the dataset size

When training only using 50% and 66% of the training set, we obtained a TRE of 4.78 mm ± 1.55 and 4.34 mm ± 1.27, respectively. This result indicates that the network performance benefits from increasing training set, as expected.

3.4 | Network performance on US segmentation

In Table 7 we report the results of our study on the simulated US dataset. It can be seen that the results are consistent with the ones obtained for ground truth segmentation data, suggesting that the network is robust against partial occlusions that can occur in the US data due to the probe orientation. The error is in an acceptable range for lumbar facet joint injections according

TABLE 7 Comparison of registration results of the proposed method and the non-DL counterpart on simulated US point clouds. Values after ± show the standard deviation.

Method	TRE (mm)	Translation	AQD	Rigidity loss
MSE + Bio	3.89 ± 0.91	4.89 ± 1.31	0.019 ± 0.023	1.01 ± 0.18
MB-CPD + springs	3.63 ± 2.27	5.12 ± 4.20	0.043 ± 0.030	-

Abbreviations: MB-CPD, multi-body CPD; MSE, mean squared error.

to,[7] proving the potential of DL methods for spine registration, and the efficacy of introducing anatomy-aware losses when training DL models.

4 | DISCUSSION

This section examines two main processes: the data generation pipeline and the registration pipeline. The data generation pipeline involves the creation of paired CT-US data, where our proposed method successfully generates accurate deformations to convert a spinal CT image into an image captured from a prone or lateral position (Figure 5). This pipeline enables the generation of CT-US data with known ground truth transformations, which demonstrates a value for network pre-training when transitioning to real data (Figure 1). To enhance efficiency and scalability, we introduce automated feature extraction, trajectory planning, and a streamlined data generation process.

In the registration pipeline, the reported results show that the proposed losses improve the registration outcome (Table 5). The method trained on the current data set does not outperform the standard (non-DL) methods, but while such standard non-DL methods are solely relying on relative point positions in source and target PC, the proposed network architecture is designed to encode both global and local features of the input point clouds. Therefore, with a larger training population, the network could learn global and local features of the spine and better capture the underlying mechanical behavior. We have supported this claim by presenting the network performance's dependency on the size of the training dataset (section 2.3.7), underlying the major potential of the method over standard registration techniques, which do not offer room for improvements with an increasing number of training data. Moreover, the MB-CPD is more dependent on the initial alignment, as shown by the results in the absence of a good pre-alignment (Table 6).

Although the generalizability of the method to real data is highly relevant, the acquisition of real paired CT-US dataset presents several challenges in terms of ethical approval and CT and 3D US data gathering and the definition of ground-truth deformations. We strongly believe that our generated data can help in pre-training the network and result in better convergence once access to real data for training is made possible.

Overall, our proposed approach demonstrates the feasibility and potential of employing DL methods for US-to-CT spine registration. It also introduces a novel method for generating a substantial amount of data to train and evaluate such models in a controlled environment, serving as an initial step before transitioning to real data.

5 | CONCLUSIONS

This study introduces a novel deep learning pipeline for CT-to-US spine registration, incorporating anatomical losses to facilitate learning of spinal biomechanics. Although biomechanical constraints were employed for DL approaches trained for other modalities (MRI-CT)[21] or other organs (prostate),[18] application to CT-to-US spine registration was not investigated. Additionally, we propose a data generation method to overcome the challenge of procuring sufficient paired CT-US data, marking a novel approach in this domain.

However, while promising, our method did not surpass standard (non-DL) techniques in the existing dataset, indicating a need for a larger training population to enhance performance. Furthermore, despite the potential utility of our artificially generated data for preliminary network training, the procurement of authentic paired CT-US datasets continues to be a formidable challenge due to ethical concerns and practical difficulties.

Overall, our study represents a pioneering step towards applying deep learning to CT-to-US spine registration. Yet, further research and more comprehensive real-world data collection are essential to fully realize this method's potential.

ACKNOWLEDGMENTS

This work was partially supported by the Bayerische Forschungsstiftung (Bavarian Research Foundation), under Grant DOK-180-19. Also, the authors were partially supported by the grant NPRP-11S-1219-170106 from the Qatar National Research Fund (a member of the Qatar Foundation). The findings herein are however solely the responsibility of the authors.

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts to disclose.

REFERENCES

1. Sekuboyina A, Husseini ME, Bayat A, et al. VerSe: a vertebrae labelling and segmentation benchmark for multi-detector CT images. *Med Image Anal*. 2021;73:102166.
2. Wu T, Zhao W-H, Dong Y, Song HX, Li JH. Effectiveness of ultrasound-guided versus fluoroscopy or computed tomography scanning guidance in lumbar facet joint injections in adults with facet joint syndrome: a meta-analysis of controlled trials. *Arch Phys Med Rehabil*. 2016;97:1558-1563.
3. Evansa I, Logina I, Vanags I, Borgeat A. Ultrasound versus fluoroscopic-guided epidural steroid injections in patients with degenerative spinal diseases: a randomised study. *Eur J Anaesthesiol*. 2015;32:262-268.
4. Gueziri HE, Santaguida C, Collins DL. The state-of-the-art in ultrasound-guided spine interventions. *Med Image Anal*. 2020;65:101769.
5. Seitel A, Sojoudi S, Osborn J, et al. Ultrasound-guided spine anesthesia: feasibility study of a guidance system. *Ultrasound Med Biol*. 2016;42:3043-3049.
6. Myronenko A, Song X. Point set registration: coherent point drift. *IEEE Trans Pattern Anal Mach Intell*. 2010;32:2262-2275.
7. Nagpal S, Hacihaliloglu I, Ungi T, et al. A global CT to US registration of the lumbar spine. *Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling*. Vol. 9036. International Society for Optics and Photonics; 2014:90362O.
8. Behnami D, Seitel A, Rasoulian A, et al. Joint registration of ultrasound, CT and a shape+ pose statistical model of the lumbar spine for guiding anesthesia. *Int J Comp Assist Radiol Surg*. 2016;11:937-945.
9. Anas EMA, Seitel A, Rasoulian A, et al. Registration of a statistical model to intraoperative ultrasound for scaphoid screw fixation. *Int J Comput Assist Radiol Surg*. 2016;11:957-965.
10. Ciganovic M, Ozdemir F, Pean F, Fuernstahl P, Tanner C, Goksel O. Registration of 3D freehand ultrasound to a bone model for orthopedic procedures of the forearm. *Int J Comput Assist Radiol Surg*. 2018;13:827-836.
11. Moghari MH, Abolmaesumi P. Comparing unscented and extended Kalman filter algorithms in the rigid-body point-based registration. In: *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE; 2006:497-500.
12. Moghari MH, Abolmaesumi P. Point-based rigid-body registration using an unscented kalman filter. *IEEE Trans Med Imaging*. 2007;26:1708-1728.
13. Rasoulian A, Abolmaesumi P, Mousavi P. Feature-based multibody rigid registration of CT and ultrasound images of lumbar spine. *Med Phys*. 2012;39:3154-3166.
14. Nagpal S, Abolmaesumi P, Rasoulian A, et al. A multi-vertebrae CT to US registration of the lumbar spine in clinical data. *Int J Comput Assist Radiol Surg*. 2015;10:1371-1381.
15. Masoumi N, Belasso CJ, Ahmad MO, Benali H, Xiao Y, Rivaz H. Multimodal 3D ultrasound and CT in image-guided spinal surgery: public database and new registration algorithms. *Int J Comput Assist Radiol Surg*. 2021;16:555-565.
16. Gueziri HE, Drouin S, Yan CX, Collins DL. Toward real-time rigid registration of intra-operative ultrasound with preoperative CT images for lumbar spinal fusion surgery. *Int J Comput Assist Radiol Surg*. 2019;14:1933-1943.
17. Liu X, Qi CR, Guibas LJ. Flownet3D: Learning scene flow in 3D point clouds. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE; 2019:529-537.
18. Fu Y, Lei Y, Wang T, et al. Biomechanically constrained non-rigid MR-TRUS prostate registration using deep learning based 3D point cloud matching. *Med Image Anal*. 2021;67:101845.
19. Velikova Y, Simson W, Salehi M, Azampour MF, Paprottka P, Navab N. Cactuss: common anatomical CT-US space for US examinations. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer; 2022:492-501.
20. Velikova Y, Azampour MF, Simson W, Duque VG, Navab N. LOTUS: learning to optimize task-based US representations. 2023. arXiv preprint arXiv:2307.16021.
21. Jian B, Azampour MF, De Benetti F, et al. Weakly-supervised biomechanically-constrained CT/MRI Registration of the spine. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer; 2022:227-236.
22. Deng Y, Wang C, Hui Y, et al. CTSpine1K: a Large-Scale Dataset for spinal vertebrae segmentation in computed tomography, arXiv preprint arXiv:2105.14711. 2021.
23. Izambert O, Mitton D, Thourot M, Lavaste F. Dynamic stiffness and damping of human intervertebral disc using axial oscillatory displacement under a free mass system. *Eur Spine J*. 2003;12:562-566.
24. Schmidt H, Häußler K, Wilke H-J, Wolfram U. Structural behavior of human lumbar intervertebral disc under direct shear. *J Appl Biomater Funct Mater*. 2015;13:66-71.
25. Kemper A, McNally C, Manoogian S, McNeely D, Duma S. Stiffness properties of human lumbar intervertebral discs in compression and the influence of strain rate. *Lumbar Injury Biomechanics*. 2013;1-127.
26. Kurutz M. In vivo age-and sex-related creep of human lumbar motion segments and discs in pure centric tension. *J Biomech*. 2006;39:1180-1190.
27. Keller TS, Colloca CJ, Béliveau J-G. Force-deformation response of the lumbar spine: a sagittal plane model of posteroanterior manipulation and mobilization. *Clin Biomech*. 2002;17:185-196.
28. Valdevit A, Noonan E, Sidoti H, Chung R, Ritter A, Errico T. Elevated frequency loading of facet joints. *J Biomed Eng Biosci*. 2015;2:12-17.
29. Faure F, Duriez C, Delingette H, et al. SOFA: a Multi-model framework for interactive physical simulation. In: Payan Y, ed. *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*. Vol. 11 of Studies in Mechanobiology, Tissue Engineering and Biomaterials. Springer; 2012:283-321.
30. TotalSegmentator. Accessed on 05/23/2023. https://totalsegmentator.com/
31. Wasserthal J, Meyer M, Breit H-C, Cyriac J, Yang S, Segeroth M. TotalSegmentator: robust segmentation of 104 anatomical structures in CT images. 2022.
32. Salehi M, Ahmadi S-A, Prevost R, Navab N, Wein W. Patient-specific 3D ultrasound simulation based on convolutional ray-tracing and appearance optimization. In: Navab N, Hornegger J, Wells WM, Frangi A, eds. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer International Publishing; 2015:510-518.
33. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: *International Conference on Medical image Computing and Computer-Assisted Intervention*. Springer; 2015:234-241.
34. Chen X, Diaz-Pinto A, Ravikumar N, Frangi AF. Deep learning in medical image registration. *Prog Biomed Eng*. 2021;3:012003.
35. Qi CR, Yi L, Su H, Guibas LJ. Pointnet++: deep hierarchical feature learning on point sets in a metric space. Advances in neural information processing systems (NeurIPS). 2017;30.

How to cite this article: Azampour MF, Tirindelli M, Lameski J, et al. Anatomy-aware computed tomography-to-ultrasound spine registration. *Med Phys*. 2024;51:2044–2056. https://doi.org/10.1002/mp.16731