


ORIGINAL ARTICLE

Open Access



Denoising diffusion-based MRI to CT image translation enables automated spinal segmentation

Robert Graf^{1*} , Joachim Schmitt¹, Sarah Schlaeger¹, Hendrik Kristian Möller¹, Vasiliki Sideri-Lampretsa², Anjany Sekuboyina^{1,3}, Sandro Manuel Krieg⁴, Benedikt Wiestler¹, Bjoern Menze³, Daniel Rueckert^{2,5} and Jan Stefan Kirschke¹

Abstract

Background Automated segmentation of spinal magnetic resonance imaging (MRI) plays a vital role both scientifically and clinically. However, accurately delineating posterior spine structures is challenging.

Methods This retrospective study, approved by the ethical committee, involved translating T1-weighted and T2-weighted images into computed tomography (CT) images in a total of 263 pairs of CT/MR series. Landmark-based registration was performed to align image pairs. We compared two-dimensional (2D) paired — Pix2Pix, denoising diffusion implicit models (DDIM) image mode, DDIM noise mode — and unpaired (SynDiff, contrastive unpaired translation) image-to-image translation using “peak signal-to-noise ratio” as quality measure. A publicly available segmentation network segmented the synthesized CT datasets, and Dice similarity coefficients (DSC) were evaluated on in-house test sets and the “MRSpineSeg Challenge” volumes. The 2D findings were extended to three-dimensional (3D) Pix2Pix and DDIM.

Results 2D paired methods and SynDiff exhibited similar translation performance and DCS on paired data. DDIM image mode achieved the highest image quality. SynDiff, Pix2Pix, and DDIM image mode demonstrated similar DSC (0.77). For craniocaudal axis rotations, at least two landmarks per vertebra were required for registration. The 3D translation outperformed the 2D approach, resulting in improved DSC (0.80) and anatomically accurate segmentations with higher spatial resolution than that of the original MRI series.

Conclusions Two landmarks per vertebra registration enabled paired image-to-image translation from MRI to CT and outperformed all unpaired approaches. The 3D techniques provided anatomically correct segmentations, avoiding underprediction of small structures like the spinous process.

Relevance statement This study addresses the unresolved issue of translating spinal MRI to CT, making CT-based tools usable for MRI data. It generates whole spine segmentation, previously unavailable in MRI, a prerequisite for biomechanical modeling and feature extraction for clinical applications.

Key points

- Unpaired image translation lacks in converting spine MRI to CT effectively.
- Paired translation needs registration with two landmarks per vertebra at least.
- Paired image-to-image enables segmentation transfer to other domains.

*Correspondence:

Robert Graf
robert.graf@tum.de

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

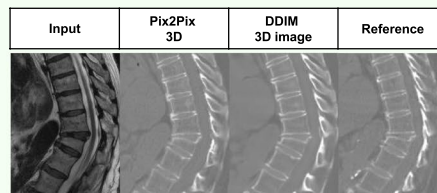
- 3D translation enables super resolution from MRI to CT.
- 3D translation prevents underprediction of small structures.

Keywords Deep learning, Image processing (computer assisted), Magnetic resonance imaging, Spine, Vertebral body

Graphical Abstract

Denoising diffusion-based MRI to CT image translation enables automated spinal segmentation

- Whole spine segmentation was unavailable in MRI, a prerequisite for biomechanical modeling and feature extraction for clinical applications.
- Two landmarks per vertebra registration enabled paired image-to-image translation from MRI to CT. The 3D techniques provided anatomically correct segmentations, avoiding underprediction of small structures like the spinous process.
- 3D translation allows super-resolution CT images from sagittal MRI



Our translation from T2-weighted MRI to CT with Pix2Pix 3D and Denoising diffusion implicit model (DDIM) 3D. The reference is a registered real CT.



3D visualization of the generated segmentations. With the rescaling to 1-mm isotropic, we receive a super-resolution of our mask in the thick slice direction.

This study addresses the unresolved issue of translating spinal MRI to CT, making CT-based tools usable for MRI data. It generates whole spine segmentation, previously unavailable in MRI, a prerequisite for biomechanical modeling and feature extraction for clinical applications.



**Eur Radiol Exp (2023) Graf R, Schmitt J, Schlaeger S et al.
DOI: 10.1186/s41747-023-00385-2**

Background

The different image contrast of computed tomography (CT) and magnetic resonance imaging (MRI) offer distinct clinical utilities. Segmentation is a prerequisite to automatically extract biomarkers, especially in large cohorts like the German National Cohort [1] or the UK Biobank [2]. While the extraction of the precise bone structure of the spine from CT is publicly available [3, 4], neither a segmentation nor an annotated ground truth dataset for the whole spine including the posterior elements is currently available for MRI.

Accurate segmentations are not only vital for scientific studies but also enable the exact localization of abnormalities in clinical routine. Unlike CT, MRI provides additional information about bone marrow edema-like changes, intervertebral disc degeneration, degenerative endplate changes, ligaments, joint effusions, and the spinal cord. Robust and precise segmentation and quantification of such spinal structures are a prerequisite, *e.g.*, to evaluate large epidemiologic

studies or to enable automated reporting. An alternative to labor-intensive manual annotations is the potential use of image-to-image translation to extract bony structures. This approach may overcome challenges like partial volume effects (*e.g.*, at the spinous process) and subtle signal differences (*e.g.*, of vertebral end plates and ligaments in MRI), which are easily distinguishable in high-resolution CT but not in MRI.

Image-to-image translation involves transforming images from one domain to another, and several deep learning methods have been employed for this purpose, including Pix2Pix [5], CycleGAN [6], and contrastive unpaired translation (CUT) [7]. These methods have been used in various studies to generate missing sequences, translate to different domains, enhance image quality, and improve resolution [8]. In the medical domain, these methods have shown success in rigid structures like the brain, head, and pelvis, where registration guarantees that both domains have similar tissue distributions and anomalies [8]. However, if biases

are not accounted for, the model may hallucinate new structures to fit both distributions [9]. Due to this difficulty, translating warpable structures like the spine is less explored in the literature. Some successful implementations have shown that translated images can be similar to the target images and might mislead medical experts [10–14]. However, none of these works has focused on using translations for downstream tasks, such as segmentations in the output domain.

This study aimed to develop and compare different image translation networks for pretrained CT-based segmentation models when applied to MRI datasets (Fig. 1). The primary focus was on segmenting the entire spine, with special attention to accurately translating the posterior spine structures, as they pose challenges in MRI delineation. We compared generative adversarial network (GAN)-based approaches [5, 7] with new denoising diffusion models [15–17]. Denoising diffusion functions are fundamentally different from GANs, as they add and remove noise to an image instead of relying on the discriminator and generator zero-sum game in GANs. In the computer vision domain, denoising diffusion models have outperformed GANs in various tasks, including upscaling, inpainting, image restoration, and paired image-to-image translation [18]. While diffusion has been applied to medical image translation tasks in a limited number of papers [17, 19–22], we adapted the conditional denoising

diffusion for paired image-to-image two-dimensional (2D) and three-dimensional (3D) translation.

The purposes of this study were as follows: (1) to improve existing image-to-image translation for spine MRI to CT translation by improving all steps of the process, from data alignment, implementation of new denoising diffusion translations and comparison to GANs, and finally extension of our findings to 3D translation; (2) to utilize the translated CT images for automatic segmentation of the entire spine, eliminating the need for a manually labeled segmentation mask in the original MRI domain; and (3) to develop the ability to generate full spine segmentations on MRI, which are currently not available.

Methods

In brief, we aligned CT and MR spine images through rigid landmark registration [23]. With this paired data, we trained various image-to-image models to generate synthetic CT images. We used an available CT segmentation algorithm [3, 4] to generate vertebral masks in these synthesized CTs for the original MRI. These resulting segmentations were subsequently used to generate new landmarks for new training data (Fig. 1). During inference, the MRI is sufficient to generate a segmentation by translating the MRI to a synthetic CT

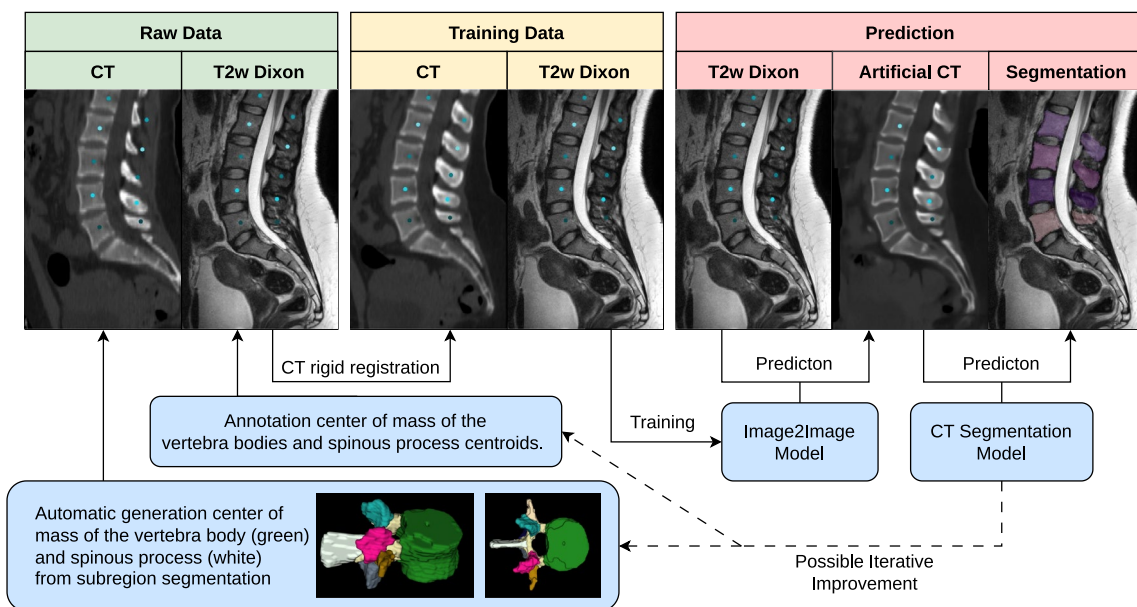


Fig. 1 Training pipeline. In our datasets, we identified the center of the vertebral body and spinous process (green box; raw data). Based on the center points, we rigidly registered CT onto MRI to align the bone structures between the two images (yellow box; training data). Aligned images were used to train our image-to-image models. Finally, the MRIs of validation and test sets were translated to CT images. Segmentation was performed on synthesized CT images and, consequently, was perfectly aligned with the original MRIs (blue box from left to right; prediction). The generated segmentations can be used for generating additional and new center points to iteratively optimize the registration

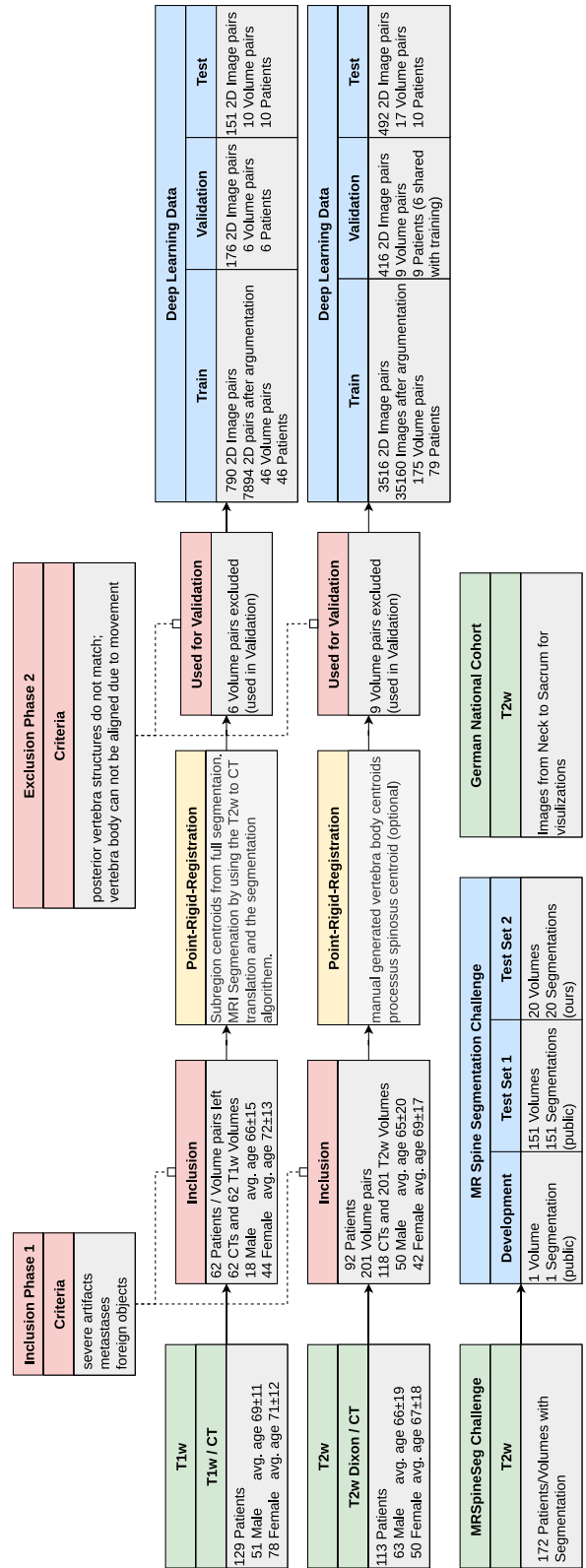


Fig. 2 Datasets, preparation, exclusion, and split. MRI data were acquired with 12 different scanners from 3 different vendors. Additionally, we used the MRSegCtg for external testing. For the 2D training, we only consider 2D slices containing a spine. We demonstrated generalizability using a full-body MRI from the German National Cohort dataset for the figures in this paper

and subsequently applying an existing CT segmentation algorithm. We compared different landmark registrations and 2D models. Finally, we adapted the results into 3D models and assessed the accuracy of the resulting segmentations.

Data

In this study, we retrospectively collected sagittal T1-weighted and T2-weighted MRI and corresponding CT images of the spine from the same patient within a week. Approval from the local ethics committee was obtained, and informed consent was waived. Figure 2 illustrates our data selection process. Sixty-two T1-weighted image series (18 males, aged 66 ± 15 years [mean \pm standard deviation]; 44 females, aged 72 ± 13 years) were used from another unpublished in-house study, including five thoracic and 57 lumbar volumes. Additionally, a new dataset was collected of 201 T2-weighted image series (50 males, aged 65 ± 20 years; 42 females, aged 69 ± 17 years) from 92 patients, including 38 cervical, 99 thoracic, and 70 lumbar volumes. Patients with fractures and degenerative changes were included, while those with motion artifacts, metastases, and foreign objects were excluded, because for segmentation models, it would benefit when the translation suppresses these anomalies. We performed rigid registration of the matching MRIs and CTs based on the center of mass of the vertebral body and the spinous process (Fig. 1, bottom left). In-house test set, training, and validation set were split patient-wise for different MRI acquisitions of other spine regions. For validation, six T1-weighted and nine T2-weighted MRIs were used as they could not be aligned with the CTs due to substantially different patient positioning.

We used 172 lumbar MRI and segmentation volumes from the MRSpineSeg Challenge (MRSSegClg) [24, 25] for external evaluation of Dice similarity coefficient (DSC). This dataset focuses on the lumbar region, but the segmentation exceeds the bony borders, questioning its validity. One subject was used for pipeline development and validation. Validation sets were used to find optimal inference parameters and to avoid overfitting. Since the labels in MRSSegClg encompass not only the bony spine but also adjacent ligaments and soft tissue, we manually adjusted the labels for a subset of 20 volumes to restrict them solely to the bone. We analyzed these subsets as two distinct datasets.

Image preprocessing

CT and MR datasets were rigidly registered [23] by using landmarks to facilitate paired image translation. For the single-landmark approach, we selected the center of mass (CM) of the vertebral bodies. To

address rotational misalignment around the cranio-caudal axis, the CM of the spinal processes was added for the two-landmark approach, as such rotational misalignment was frequently observed. Landmarks for CT were automatically determined based on vertebral and subregion segmentations (Fig. 1). For the T2-weighted images, we manually identified the CM points for both the vertebral bodies and the spinous processes. The manual centroid selection and ground truth segmentation corrections in the test sets were performed by J. S., a radiologist with 3 years of experience. To obtain the points for the T1-weighted images, we synthesized CTs by adapting the T2 weighted to CT translation, generating segmentation from synthetic images, and extracting the CMs. Roughly 10 to 20% of the failure cases were first excluded and then translated with models that were trained on the other T1-weighted images. This proved sufficient to generate all CM points. To assess the impact of additional landmarks on registration, we computed the DSC using our pipeline on the T2-weighted dataset using the manual ground truth as a reference.

CT images were transformed to the range of $[-1, 1]$ by dividing the values by 1,000 HU and clamping outliers to retain air, soft tissue, and bone while suppressing extreme intensities. Linear rescaling was applied to the MRI data, converting the range from $[0, \text{max}]$ to $[-1, 1]$. To account for varying intensities, MRIs were augmented with a random color jitter (brightness, contrast randomization: 0.2). Image pairs were resampled to a uniform spatial resolution of 1×1 mm in the sagittal plane and a slice thickness of 2.5–3.5 mm, as acquired in the MRI. To enhance the training data by a factor of 10 and simulate weak scoliosis and unaligned acquisition, we introduced 3D image deformations using the elastic deformation Python plug-in [26]. Subsequently, the volumes were sliced into 2D sagittal images, and slices without segmentation were removed. Random cropping was performed to adjust the image size to 256×256 pixels.

Models for image-to-image translation

To compare various image-to-image translation methods, we implemented two unpaired methods, namely CUT [7] and SynDiff [17], along with three paired methods, Pix2Pix [5], DDIM noise, and DDIM image. The training process involved unregistered and registered data using both single- and two-landmark approaches. For DDIM, we employed a UNet architecture [26] with convolutional self-attention and embeddings for the timesteps, which we refer to as self-attention U-network (SA-UNet) [18, 27, 28]. The diffusion mechanism predicted either noise

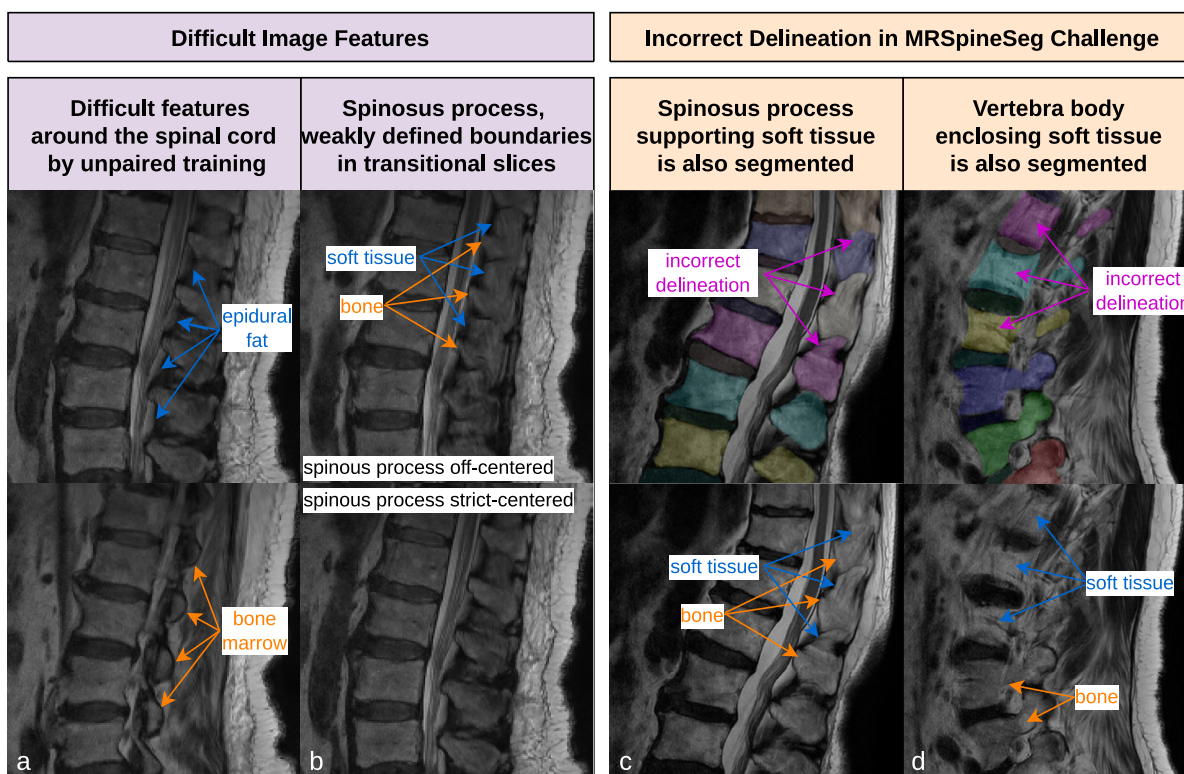


Fig. 3 Difficulties of the MRI data for unpaired training and issues with the MRSSegClg segmentation. **a** The bone marrow of the posterior elements and the epidural fat were not easily differentiated. Unpaired learning has issues translating the arcus as bone and the epidural fat as soft tissue in the CT domain. **b** In posterior elements, bone and soft tissue boundaries are weakly defined due to partial volume effects in and around the spinous process. **c** The segmentations of the MRSSegClg include soft tissues around the spinous process, caused by difficulties of the original annotators as described in B. **d** The soft tissues around the vertebrae are also segmented in the MRSSegClg. **c** and **d** show the reasons why we manually improved the segmentation in a small subset

or the image, with the other computed during inference. A learning rate of 0.00002 was used, and we set the timestep to $t=20$ for the DDIM inference parameter. The value of $\eta = 1$ (noise generation is fully random) was determined by optimizing on the validation set. We compared our approach to CUT [7], Pix2Pix [5], and SynDiff [17]. During our experiments, we performed a hyperparameter search for the reference ResNet and UNet. Additionally, we introduced a weighted structural similarity index metric (SSIM) loss from a recent paper [29] to update the loss formulation. To further explore the impact of different models and methods, we also tested CUT and Pix2Pix with the SA-UNet. All models were randomly initialized. In our analysis of DDIM, we ablated three inference parameters [16, 30]. However, the results did not show substantial effects, and we have included them in the [Supplementary material](#) along with brief descriptions of the tested methods.

Image quality

The evaluation of image quality involved comparing actual and synthesized CT images. To quantify this, we

used the “peak signal-to-noise ratio” (PSNR) metric. In this context, the reference image serves as the signal, while the divergence between the two images is considered the noise. A PSNR value above 30 dB indicates that the difference between the two images is imperceptible to the human eye [10]. It is important to note that we did not control the correspondence of soft tissue, as it fell outside the scope of our downstream task. To handle this in our evaluation, we masked pixels that were further than 10 pixels away from a segmented spine structure, setting them to zero. We also computed the absolute difference (L1) mean squared error (MSE), SSIM, and visual information fidelity (VIFp).

Downstream task: segmentation

We utilized a publicly available segmentation algorithm [3, 4] on the synthesized CT images. We then compared the DSC globally and on a vertebral level between the synthesized and ground truth segmentations in four datasets. The segmentation ground truth of the in-house datasets was derived from the aligned CT image and was

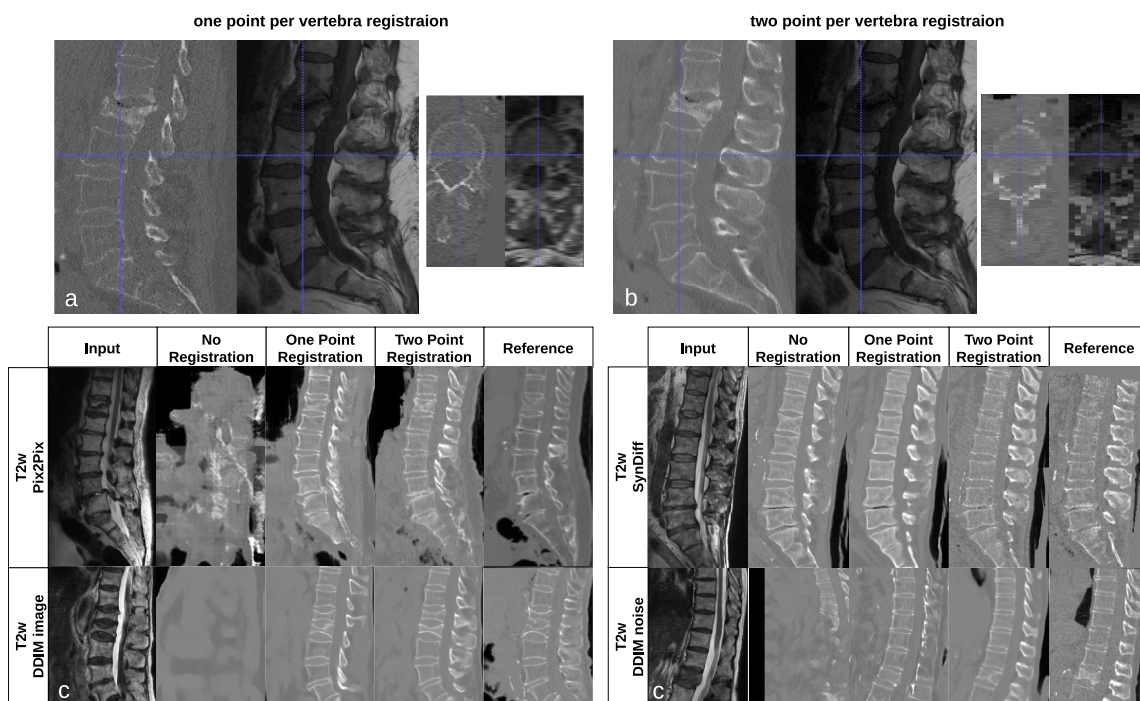


Fig. 4 Comparison of one and two registration points per vertebra *versus* real data. **a** We registered with a single point in the center of the vertebral body. The vertebral body could rotate along the spine axis. This caused the posterior vertebra structures to be misaligned. **b** When we registered the images with an additional point on the spinous process, we avoided this rotation around the spine itself. The blue dashed lines are for locating the relation between axial and sagittal slices. **c** Translation with networks trained on registrations with 0, 1, or 2 points per vertebra. Images are from the in-house T2-weighted test dataset. Posterior structures are only reconstructed correctly with 2-point registration. *DDIM* Denoising diffusion implicit model

manually corrected (datasets 1 and 2). The segmentation of the MRSSegClg that is known to exceed the bony structures (dataset 3) and a manually corrected subset of MRSSegClg (dataset 4) [24, 25]. In Fig. 3c and d, the segmentation reaching beyond the bony structures of MRSSegClg is highlighted. For analysis purposes, we excluded structures that the CT segmentation algorithm could not segment, such as the sacrum and partially visualized vertebrae.

3D image translation with diffusion

The first implementations of both DDIM and Pix2Pix in 3D, similar to the 2D approach, did not converge. We thus implemented changes according to recommendations of Bieder et al. [31]. To optimize graphics processing unit storage, we eliminated attention layers and replaced concatenation skip connections with addition operations. Additionally, we introduced a position embedding by concatenating ramps ranging from zero to one of the original images’ full dimensions into the input. The training was done on 3D patches, and our approach

used a patch size of (128 × 128 × 32), where the left/right side was limited to 32 pixels due to the image shape. This setup is “fully convolutional,” which means that during inference, an image of any size can be computed by the network as long the sides are divisible by 8. To the best of our knowledge, this represents the first 3D image-to-image translation with diffusion. Since 3D translations require to include the left/right direction, we resampled all images to 1 mm isotropic.

Statistical analysis and software

We employed a paired *t*-test to assess the significance of PSNR and DSC between different models. To achieve a fixed size of 256 × 256 pixels for assessing image quality, we used one crop per image slice. When reporting differences in multiple experiments, we present the worst (*i.e.*, highest) *p*-value. We skip significance calculations other image quality metrics because the results are redundant. For 3D data, we pad the test data, and the 3D models generate 1-mm isotropic volumes, which are later resampled to the original MRI size.

Table 1 Image quality for T1-weighted and T2-weighted MRI to CT translation

From T1-weighted MRI	L1↓	MSE↓	PSNR↑	SSIM↑	VIFp↑
CUT ResNN (unpaired)	0.0224	0.0050	23.50	0.835	0.295
CUT SA-UNet (unpaired)	0.0295	0.0083	21.76	0.819	0.269
Pix2Pix UNet	0.0143	0.0023	27.37	0.881	0.392
Pix2Pix SA-UNet	0.0135	^a 0.0020	27.82	0.883	0.394
SynDiff (unpaired)	0.0150	0.0024	27.01	0.865	0.373
DDIM noise $\eta = 1, t = 20, w = 0$	0.0136	0.0021	27.60	0.879	0.396
DDIM image $\eta = 1, t = 20, w = 0$	^a 0.0131	^a 0.0020	^a 27.89	^a 0.887	^a 0.411
From T2-weighted MRI	L1↓	MSE↓	PSNR↑	SSIM↑	VIFp↑
CUT ResNN (unpaired)	0.0213	0.0046	23.72	0.848	0.312
CUT SA-UNet (unpaired)	0.0215	0.0046	23.75	0.850	0.311
Pix2Pix UNet	0.0142	0.0023	26.95	0.895	0.392
Pix2Pix SA-UNet	0.0142	0.0023	26.87	0.890	0.384
SynDiff (unpaired)	0.0140	0.0022	27.12	0.885	0.385
DDIM noise $\eta = 1, t = 20, w = 0$	0.0139	0.0023	26.92	0.894	0.391
DDIM image $\eta = 1, t = 20, w = 0$	^a 0.0131	^a 0.0021	^a 27.36	^a 0.898	0.401
Pix2Pix 3D	0.0188	0.0039	26.38	0.889	0.428
DDIM 3D noise $\eta = 1, t = 25$	0.0194	0.0041	26.22	0.894	^a 0.444
DDIM 3D image $\eta = 1, t = 25$	0.0189	0.0040	26.22	0.892	0.434

Arrows indicate if smaller or bigger is better. As a visual aid, we marked the best values with ^a. We marked multiple values if they were below the rounding threshold. The ground truth is registered real CTs. The image pairs are from the test set of our in-house data

CUT Contrastive unpaired translation, CT Computed tomography, DDIM Denoising diffusion implicit model, MRI Magnetic resonance imaging, MSE Mean squared error, PSNR Peak signal-to-noise ratio, SA-UNet Self-attention U-network, SSIM Structural similarity index metric, VIFp Visual information fidelity

Results

Influence of rigid registration

Networks trained on unregistered data were incapable of learning the difference between soft tissue and bone. During our early testing, we noticed that most methods could correctly identify the vertebral body, but translating the posterior structures was impossible. Especially, the spinous process was often omitted in the translation, as shown in Fig. 4. “One point per vertebra” registration was sufficient for the vertebral body translation, but the spine could rotate around the craniocaudal axis. This caused the spinous process to disappear in translated images (Fig. 4a, b). Additionally, confusion between epidural fat and bone shifted the entire posterior elements towards the spinal cord. Overcoming this issue required accounting for rotation by adding additional points to the rigid registration (Fig. 4). Next to visual findings, we observed a significant increase in DSC from 1 to 2 points per vertebra registration: Pix2Pix 0.68 to 0.73 ($p < 0.003$); SynDiff 0.74 to 0.77 ($p < 0.001$); DDIM noise 0.55 to 0.72 ($p < 0.011$); and

DDIM image 0.70 to 0.75 ($p < 0.001$). Notably, the best unpaired method, SynDiff, could not learn posterior structure translation without registration (DSC without registration 0.75).

Image quality

The unpaired CUT models performed worse than all others ($p < 0.001$), while all other models performed on a similar level (Table 1 for PSNR and other common metrics). Example outputs from the test sets can be seen in Fig. 5. The Pix2Pix with the SA-UNet performed better on T1-weighted images and worse on T2-weighted images than the smaller UNet (T1 weighted, $p < 0.001$; T2 weighted, $p = 0.041$). Even though SynDiff had an unpaired formulation, it had similar results compared to our paired Pix2Pix and DDIM noise (slightly worse in T1 weighted and better in T2 weighted, all $p < 0.003$). The DDIM image mode performed slightly better than the DDIM noise mode ($p < 0.001$), SynDiff ($p < 0.001$), and Pix2Pix ($p < 0.001$). DDIM image mode produces images with less noise than the original data. Less noise should make the segmentation easier. Overall, the DDIM image mode was our best-performing 2D model.

Downstream task: segmentation

Three 2D models shared the best DSC: Pix2Pix SA-UNet, SynDiff, and DDIM image mode (Table 2): Pix2Pix SA-UNet versus SynDiff, $p = 0.019$; Pix2Pix SA-UNet versus DDIM image mode, $p < 0.001$; and DDIM image mode versus SynDiff, $p = 0.455$. DDIM in noise mode and Pix2Pix UNet (DDIM noise versus Pix2Pix UNet, $p = 0.972$) were worse than the three best models ($p < 0.001$). The CUT reconstruction was unsuited for segmentation and was the worst model (CUT versus all $p < 0.001$). An example of the segmentation from different translations for a full spine can be found in Fig. 6 in an example dataset from the German National Cohort [1].

We observed comparable rankings in the MRSSegClg [24, 25] and T1-weighted datasets when excluding the vertebral body (Table 3). In the in-house T2-weighted test set, SynDiff has a considerably higher DSC than Pix2Pix SA-UNet and DDIM image mode ($p < 0.001$), indicating a better performance in the “more complicated” anatomical structures for this data set only.

The correction of the MRSSegClg segmentations resulted in an increased DSC of up to 0.02. The rankings of all methods on the original versus the corrected MRSSegClg dataset were mostly consistent, indicating that no method had exploited the false delineation by overpredicting the segmentation.

Overall, Pix2Pix SA-UNet, DDIM image mode, and SynDiff were equally capable of producing CT images

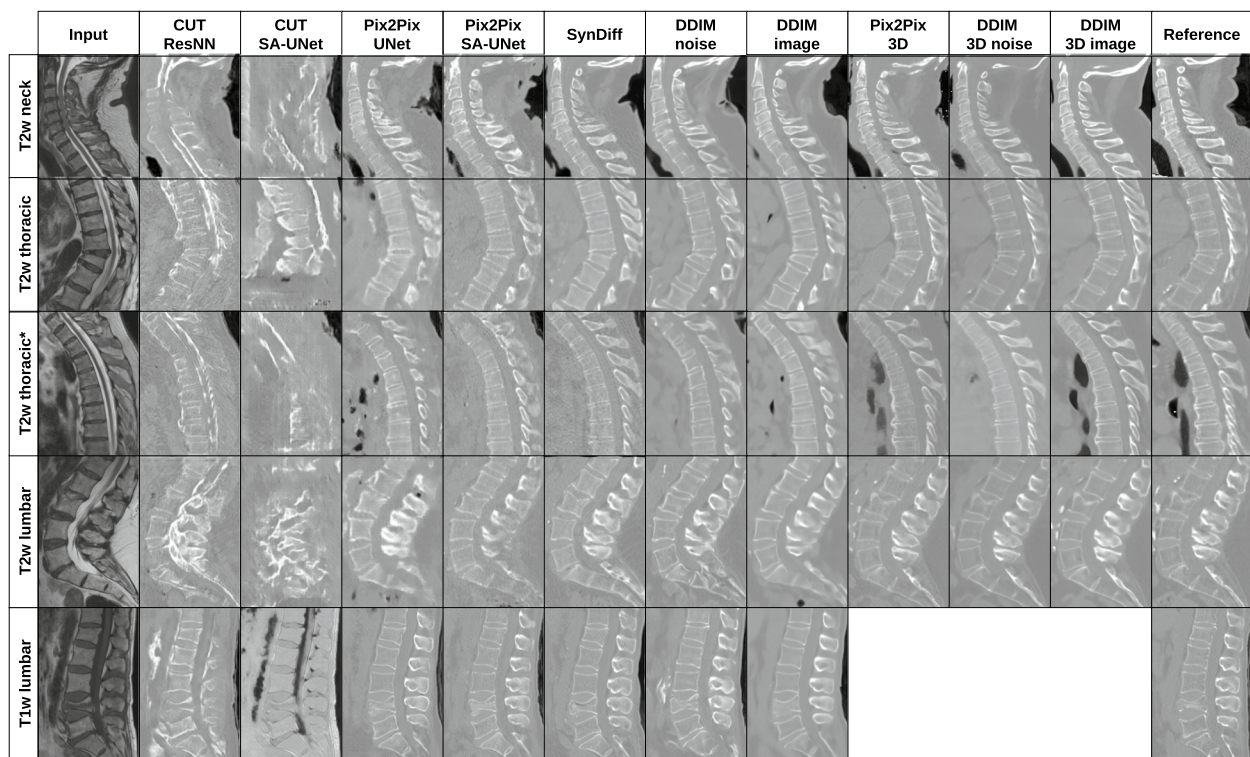


Fig. 5 Translation from test sets T1-weighted/T2-weighted MRI to CT from the neck to the lumbar vertebra. We did not control the type of reconstruction of the CT. Therefore, the noise level and appearance could differ from the reference and were still considered correct. The 3D variances were trained on an improved training set, which was only done for T2 weighted. The reference is a registered real CT. * is an off-angle acquisition with strong partial volume effects. The dataset contains a high number of broken vertebral bodies, which causes them to be also translated correctly. *CUT* Contrastive unpaired translation, *DDIM* Denoising diffusion implicit model, *SA-UNet* Self-attention U-network

Table 2 Average Dice similarity coefficient† per volume and per vertebra on the T1 weighted, T2-weighted MRI, and the MRSSegClg

Dataset	Per volume	Per vertebra	Per volume	Per vertebra	Per volume	Per vertebra	Per volume	Per vertebra
	T1 weighted	T1 weighted	T2 weighted	T2 weighted	MRSSegClg	MRSSegClg	MRSSegClg (our)	MRSSegClg (our)
CUT ResNN (unpaired)	0.30	0.28	0.49	0.46	0.54	0.49	0.54	0.50
CUT SA-UNet (unpaired)	0.09	0.08	0.26	0.23	0.02	0.01	0.03	0.02
Pix2Pix UNet	0.79	0.80	0.73	0.69	0.75	0.74	0.76	0.76
Pix2Pix SA-UNet	^a 0.82	0.82	0.75	0.72	^a 0.77	^a 0.76	0.77	0.77
SynDiff (unpaired)	0.80	0.81	^a 0.77	^a 0.74	^a 0.77	^a 0.76	0.77	0.76
DDIM noise $\eta = 1, t = 20, w = 0$	0.78	0.77	0.72	0.69	0.75	0.73	0.77	0.78
DDIM image $\eta = 1, t = 20, w = 0$	^a 0.82	^a 0.83	0.75	0.72	^a 0.77	^a 0.76	^a 0.78	0.78
Pix2Pix 3D			0.79	0.78	0.78	^b 0.78	0.79	^b 0.80
DDIM 3D noise $\eta = 1, t = 25$			0.79	0.78	0.78	^b 0.78	^b 0.80	^b 0.80
DDIM 3D image $\eta = 1, t = 25$			^b 0.80	^b 0.79	^b 0.79	^b 0.78	^b 0.80	^b 0.80

MRSSegClg (ours) is a split where we improved the segmentation better to align the segmentation with the actual bone structure. T1-weighted and T2-weighted ground truths are corrected segmentations of the registered CTs. We marked the best values for the 2D cases with ^a and the overall best with ^b

CUT Contrastive unpaired translation, *DDIM* Denoising diffusion implicit model, *MRSSegClg* MRSpineSeg Challenge, *SA-UNet* Self-attention U-network

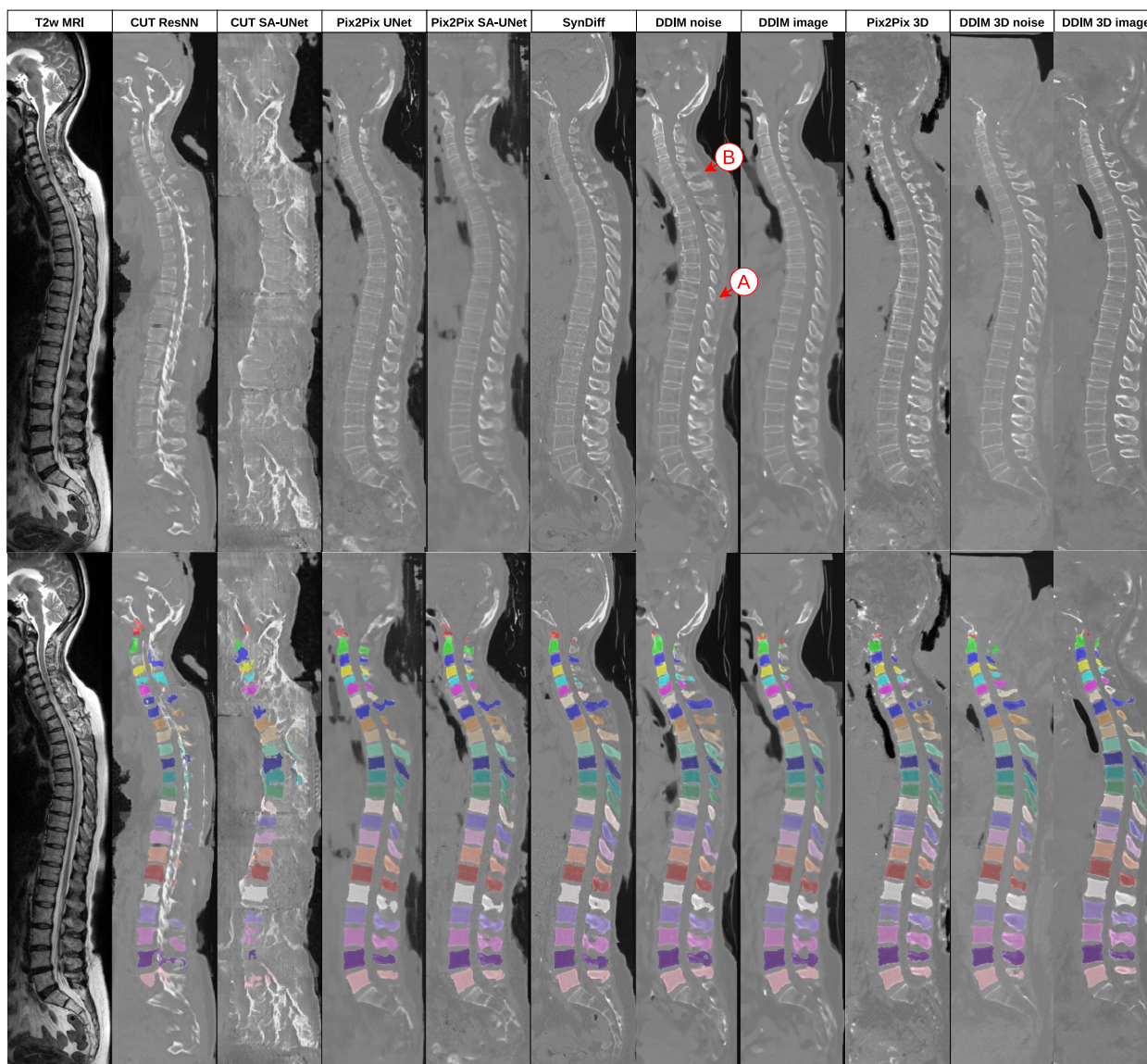


Fig. 6 Translation from T2-weighted MR to CT and the segmentation results in an external full spine scan. The MRI shown is a random image from the German National Cohort dataset. The CT translation is stitched. The 2D networks only work on a fixed size of 256×256, and the 3D models run out of memory for the entire image. The 2D networks needed classifier-free guidance ($w=1$) for these out-of-distribution images or else the neck regions would not form correctly because the frontal area has a drop in magnetic resonance signal. The 3D networks do not delineate the background and soft tissue when we use a small number of steps ($t=25$). **A** We observed underpredictions in the thorax process spinous. **B** The neck has higher variability between different translations. Moving to 3D translation resolves these issues. *CUT* Contrastive unpaired translation, *DDIM* Denoising diffusion implicit model, *SA-UNet* Self-attention U-network

for the segmentation algorithm, closely followed by DDIM noise mode and the Pix2Pix UNet.

3D image translation with diffusion

All 3D models increased the DSC compared to our 2D models ($p<0.006$). Pix2Pix 3D and DDIM 3D noise performed on a similar level, while DDIM 3D image

performances were consistently a bit better close to the rounding threshold ($p<0.001$). PSNR showed a drop compared to the 2D variants. The 3D models outperform all 2D models on posterior structures (Fig. 7: T2 weighted, $p<0.024$; MRSegClg (ours), $p<0.005$ for DDIM 3D image, $p<0.062$ for DDIM 3D noise; $p<0.462$ for Pix2Pix 3D; posterior structures are unavailable in the original MRSegClg). With the rescaling

Table 3 Average posterior structures Dice similarity coefficient[†] per volume and per vertebra

Dataset	Per volume T1 weighted	Per vertebra T1 weighted	Per volume T2 weighted	Per vertebra T2 weighted	Per volume MRSegClg (our)	Per vertebra MRSegClg (our)
CUT ResNN (unpaired)	0.09	0.09	0.17	0.15	0.16	0.13
CUT SA-UNet (unpaired)	0.01	0.01	0.07	0.05	0.00	0.00
Pix2Pix UNet	0.64	0.62	0.55	0.50	0.56	0.55
Pix2Pix SA-UNet	^a 0.68	^a 0.67	0.59	0.54	^a 0.58	0.56
SynDiff (unpaired)	0.67	^a 0.67	^a 0.63	^a 0.58	^a 0.58	^a 0.57
DDIM noise $\eta = 1, t = 20, w = 0$	0.61	0.59	0.50	0.46	0.57	0.56
DDIM image $\eta = 1, t = 20, w = 0$	^a 0.68	^a 0.67	0.58	0.53	^a 0.58	^a 0.57
Pix2Pix 3D			0.69	0.67	0.59	0.58
DDIM 3D noise $\eta = 1, t = 25$			^b 0.70	^b 0.68	0.60	^b 0.60
DDIM 3D image $\eta = 1, t = 25$			^b 0.70	^b 0.68	^b 0.61	^b 0.60

The vertebral body is removed from the calculation by an automatic subregion segmentation on the T1 weighted, T2 weighted, and MRSegClg (ours). The unchanged MRSegClg could not be subregion segmented. We marked the best values from 2D cases with ^a and the overall best with ^b

CUT Contrastive unpaired translation, DDIM Denoising diffusion implicit model, MRSegClg MRSpineSeg Challenge, SA-UNet Self-attention U-network

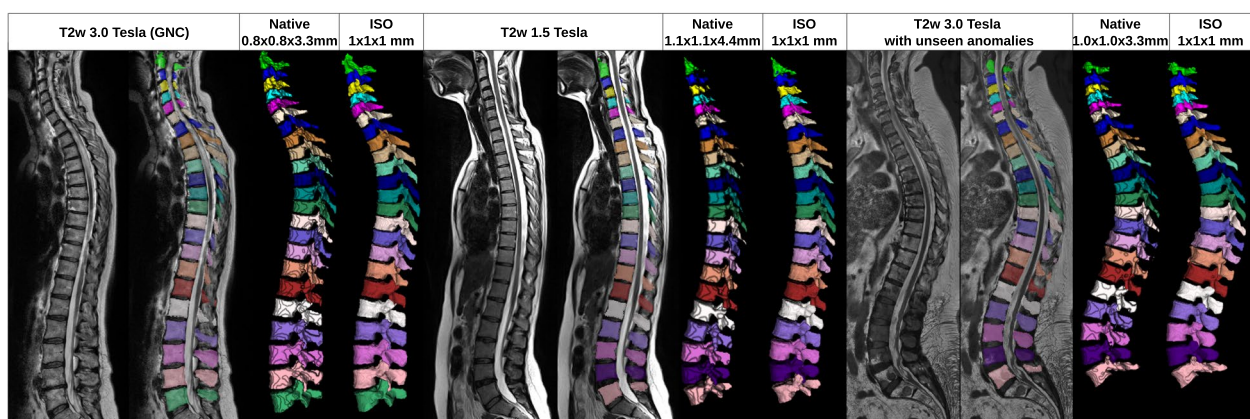


Fig. 7 3D visualization of the generated segmentations out of the German National Cohort and in-house datasets. The 3D translation models produce isometric segmentation (iso) that looks biologically correct. After downscaling to the native resolution (native), we observe that the spinous process gets deformed by reducing the slice thickness because the spinous process is thinner than two to three slices. The examples are translated by the DDIM image mode model. We observe no noticeable drop in translation quality for MRIs from other scanners. Degenerative changes that are not in the training set are often repaired during translation. While it can partially reproduce when vertebral bodies grow together, which is present in rare cases in the training set. This can be observed by the over-segmentation in the right image from vertebra 7 to 10 counted from the bottom. DDIM Denoising diffusion implicit model, ISO Isometric segmentation, Native Native resolution segmentation

to 1-mm isotropic, we receive a super-resolution of our mask in the thick slice direction that resembles a more realistic 3D shape than the native resolution (Fig. 7).

Discussion

This study successfully demonstrated the feasibility of translating standard sagittal spine MRI into the CT domain, enabling subsequent CT-based image processing. Specifically, the registration process, with a

minimum of 2 points per vertebra, enables accurately translating posterior structures, which are typically challenging for image translation and segmentation. To achieve this, a low-data registration technique was introduced for pairing CT and MRI images, which can be automated by our translation and segmentation pipeline. In our low-data domain, paired translation methods performed on a similar level, with DDIM in image mode being the single best model. The spinous process was not

always correctly translated in our 2D approaches. We resolved this issue by changing the process to 3D. Our 3D methods had a drop in image quality compared to the 2D translation. We believe this is due to the required resampling from the 1-mm isotropic output to the native resolution of the test data. Ultimately, the image-to-image translation facilitated MRI segmentation using a pre-trained CT segmentation algorithm for all spine regions.

Our results extend prior works that have been limited to high-resolution gradient-echo Dixon T1-weighted sequences to CT translations [14, 32, 33] as well as to intra-modality MR translations for different contrasts from standard T1-weighted and T2-weighted TSE sequences to short tau inversion recovery [34] or T2-weighted fat-saturated images [35], frequently used in spinal MRI. Commercial products are available for MRI to CT translation [36, 37]. However, in contrast to our approach, they require a dedicated, isotropic gradient-echo sequence. They are unavailable for standard T1-weighted or even T2-weighted TSE sequences. Acquiring an additional, dedicated image only for segmentation is resource and time demanding in everyday medical practice and not possible at all in existing data like in available large epidemiological studies like the German National Cohort.

Mature preprocessing pipelines enable image translation in other body regions [8]. For example, in brain MRI, every sample can rigidly be registered to an atlas, and the non-brain tissue is removed. However, in the spine, where vertebrae may be moving between acquisitions, such a simple, rigid preprocessing is impossible. Additionally, the mapping of intensities from the MR to the CT domain is highly dependent on the anatomy, *e.g.*, fat and water would have similar signals in T2-weighted MRI but have substantially different density values in CT, despite being in close anatomical location with a high intersubject variability. Consequently, a network cannot learn the relationship between anatomy and intensity translation based on unpaired images; the tested unpaired method CUT [7] would require additional constraints to learn an anatomically correct translation. SynDiff [17] has an unpaired CycleGAN [6] in its formulation and worked on paired datasets similar to paired methods. Still, it could not correctly translate the posterior structures on unmatched data. We demonstrated that our rigid registration is a required preprocessing for a correct translation, even for SynDiff, and we believe that better processing, such as deformable registration, can lead to better results. However, to account for inter-vertebra movement between two acquisitions due to different patient lying positions between CT and MR acquisitions would require whole vertebral segmentation. Other papers combat this issue by using axial slices, which only

need a local vertebra registration [10–12] or only focusing on the lumbar spine [5–9], where acquisitions can be performed in a more standardized patient positioning than the cervical spine. Oulbacha and Kadourys's et al. [38] also use sagittal slices like our study. However, they face similar challenges with incorrectly translating posterior structures, as observed in their figures. To address these issues, we employed dedicated preprocessing techniques and transitioned to a 3D approach.

Our study has limitations. Our pipeline enables us to generate segmentations that are available in other modalities. This method cannot produce segmentations of structures that are not segmented but visible in the input domain. We observed weaknesses in translating neck and thoracic regions when using external images, especially for the 2D methods. The posterior elements in the thoracic region were still the most difficult, and the segmentation and the translation showed more errors compared to other regions. Classifier-free guidance showed substantial improvement in language-based DDIM generation [30] and had a visible impact in 2D translation on an out-of-training distribution like the German National Cohort images. Still, the difference in image quality and the DSC are too small to measure. Therefore, we excluded classifier-free guidance [30] from our analysis, as the effect was too small to be investigated in available test sets. The same is true for testing a different number of time steps and the determinism parameter η . We go in more detail about these inference parameters in the [Supplemental materials](#).

In conclusion, we were able to show that image segmentations can be generated in a novel target domain without manual annotations if segmentations exist for another image domain, and paired data for both domains can be obtained. For the spine, we identified minimum registration requirements for paired image-to-image translations. With this approach, SynDiff, Pix2Pix, and DDIM enabled translation of 2D images resulting in similarly good downstream segmentations. We introduced a 3D variant of conditional diffusion for image-to-image translation that improved the segmentation of posterior spinal elements compared to 2D translation. The synthesized segmentations represent a novel ground truth for MRI-based spine segmentations that are prerequisites for spine studies involving large cohorts.

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
CM	Center of mass
CUT	Contrastive unpaired translation
DDIM	Denosing diffusion implicit model
DSC	Dice similarity coefficient

GAN	Generative adversarial network
MRSegClg	MRSpineSeg Challenge
PSNR	Peak signal-to-noise ratio
SA-UNet	Self-attention U-network
SSIM	Structural similarity index metric

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s41747-023-00385-2>.

Additional file 1.

Authors' contributions

The conceptualization of the DL pipeline was done by RG, SS, BW, DR, HM, JSK, DR, and AS. The registration conceptualization and development were done by RG, VSL, BM, DR, and BW. The center of mass was annotated, and segmentation correction was done by JS. The data registration and selection were done by JS, SMK, and RG. The formal analysis was done by RG and HM. The investigation of the registration issue was supported by RG, JS, HM, VSL, BM, JSK, AS, and SMK. Project supervision was done by SMK, BW, BM, DR, AS, and JSK. The original draft was done by RG. The review and editing were done by RG, JS, SS, HM, VSL, SMK, and JSK. The visualization was done by RG. All authors read and approved the final manuscript. The corresponding author and some co-authors (RG, HM, JSK) are members of the ERC grant iBack-epic. The focus of this group is to investigate the causes of lower back pain with epidemiological studies. Lower back pain is one of the most common age-related diseases. The segmentation is a prerequisite for our research, and other groups have already noted their interest in our generated segmentations.

Funding

Open Access funding enabled and organized by Projekt DEAL. The research for this article received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (101045128—iBack-epic—ERC2021-COG).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. The MRSegClg dataset is available under <https://www.spinesegmentation-challenge.com/>. The used segmentation algorithm can be accessed by <https://anduin.bonescreen.de/>. Our code for registration and our deep learning methods are available under point registration, URL: <https://github.com/robert-graf/Pointregistration>, <https://doi.org/10.5281/zenodo.8198697>; platform independent, Python 3.10 or higher with packages simpleitk nibabel jupyter simpleitk pillow pyparsing matplotlib; license: MIT License; readable conditional denoising diffusion — URL <https://github.com/robert-graf/Readable-Conditional-Denoising-Diffusion> — <https://doi.org/10.5281/zenodo.8221159>; platform independent — Python 3.10 or higher with packages pytorch pytorch-lightning numpy configargparse einops ipykernel ipython joblib nibabel pandas scikit-image scikit-learn scipy tqdm ema-pytorch; and license: MIT License. Other used publicly available algorithms are as follows: SynDiff: <https://github.com/icon-lab/SynDiff> (reference [17]); platform independent — Python > = 3.6.9 with packages torch > = 1.7.1 torchvision > = 0.8.2 cuda > = 11.2 ninja; deformable data argumentation: <https://pypi.org/project/elasticdeform/> (reference [26]); <https://doi.org/10.5281/zenodo.4563333>; and platform independent, Python package.

Declarations

Ethics approval and consent to participate

The ethics committee of the Technical University Munich approved this retrospective, German-law-compliant study and waived the need for informed consent. We asked the local ethics committee if they saw any issues with our general research, and they found no complaints (593/21 S-NP). We use existing MRI and CT images from the Clinic "Rechts der Isar." We pseudonymized the MRI and CT. According to German law, we can do research with existing images without written consent internally if we follow the guidelines of patient privacy and ethics. We adhere to the German ethics standards for medical research on patient data generated by normal clinical routine.

Consent for publication

Not applicable.

Competing interests

JSK and AS are cofounders and shareholders of Bonescreen GmbH. See <https://bonescreen.de/>. The authors who analyzed and controlled the data are not employees, cofounders, or shareholders of Bonescreen GmbH. The other authors declare that they have no competing interests.

Author details

¹Department of Diagnostic and Interventional Neuroradiology, School of Medicine, Technical University of Munich, Munich, Germany. ²Institut Für KI Und Informatik in Der Medizin, Klinikum Rechts Der Isar, Technical University of Munich, Munich, Germany. ³Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland. ⁴Department of Neurosurgery, Klinikum Rechts Der Isar, School of Medicine, Technical University of Munich, Munich, Germany. ⁵Visual Information Processing, Imperial College London, London, UK.

Received: 7 August 2023 Accepted: 12 September 2023

Published online: 14 November 2023

References

- Bamberg F, Kauczor H-U, Weckbach S et al (2015) Whole-body MR imaging in the German National Cohort: rationale, design, and technical background. *Radiology* 277:206–220. <https://doi.org/10.1148/radiol.2015142272>
- Allen N, Sudlow C, Downey P et al (2012) UK Biobank: current status and what it means for epidemiology. *Health Policy Technol* 1:123–126. <https://doi.org/10.1016/j.hlpt.2012.07.003>
- Sekuboyina A, Hussein ME, Bayat A et al (2021) VerSe: A vertebrae labeling and segmentation benchmark for multi-detector CT images. *Med Image Anal* 73:102166. <https://doi.org/10.1016/j.media.2021.102166>
- Sekuboyina A, Hussein ME, Bayat A, et al (2021) Anduin is a freely available research tool to segment vertebrae in a CT scan, uploaded as NIFTI data. In: *bonescreen anduin*. <https://anduin.bonescreen.de/>. Accessed 12 Oct 2022
- Isola P, Zhu J-Y, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit (CVPR)*. 2017:1125–1134. <https://doi.org/10.1109/CVPR.2017.632>
- Zhu J-Y, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proc. IEEE Int. Conf. Comput. Vis.* pp 2223–2232. <https://doi.org/10.1109/ICCV.2017.244>
- Park T, Efros AA, Zhang R, Zhu J-Y (2020) Contrastive learning for unpaired image-to-image translation. In: *Proceedings of the IEEE International Conference on Computer Vision (ECCV)*, 2020. Springer, pp 319–345. https://doi.org/10.1007/978-3-030-58545-7_19
- Wang T, Lei Y, Fu Y et al (2021) A review on medical imaging synthesis using deep learning and its clinical applications. *J Appl Clin Medical Phys* 22:1–36. <https://doi.org/10.1002/acm2.13121>
- Cohen JP, Luck M, Honari S (2018) Distribution matching losses can hallucinate features in medical image translation. In: *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018: 21st International Conference, Granada, Spain, September 16–20, 2018, Proceedings, Part I*. Springer, pp 529–536. https://doi.org/10.1007/978-3-030-00928-1_60
- Lee JH, Han IH, Kim DH et al (2020) Spine computed tomography to magnetic resonance image synthesis using generative adversarial networks: a preliminary study. *J Korean Neurosurg Soc* 63:386–396. <https://doi.org/10.3340/jkns.2019.0084>
- Hong K-T, Cho Y, Kang CH et al (2022) Lumbar spine computed tomography to magnetic resonance imaging synthesis using generative adversarial network: visual turing test. *Diagnostics* 12:530. <https://doi.org/10.3390/diagnostics12020530>
- Jin C-B, Kim H, Liu M et al (2019) DC2Anet: generating lumbar spine MR images from CT scan data based on semi-supervised learning. *Appl Sci* 9:2521. <https://doi.org/10.3390/app9122521>

13. Galbusera F, Bassani T, Casaroli G et al (2018) Generative models: an upcoming innovation in musculoskeletal radiology? A preliminary test in spine imaging. *Eur Radiol Exp* 2:1–13. <https://doi.org/10.1186/s41747-018-0060-7>
14. Jans LB, Chen M, Elewaut D et al (2021) MRI-based synthetic CT in the detection of structural lesions in patients with suspected sacroiliitis: comparison with MRI. *Radiology* 298:343–349. <https://doi.org/10.1148/radiol.2020201537>
15. Ho J, Jain A, Abbeel P (2020) Denoising diffusion probabilistic models. In: Larochelle H, Ranzato M, Hadsell R, et al (eds) *Proceedings of the 34th International Conference on Neural Information Processing Systems (NeurIPS 2020)*. Curran Associates, Inc., pp 6840–6851, <https://doi.org/10.48550/arXiv.2006.11239>
16. Song J, Meng C, Ermon S (2021) Denoising diffusion implicit models. In: *International Conference on Learning Representations (ICLR)*. <https://doi.org/10.48550/arXiv.2010.02502>
17. Özbey M, Dalmaz O, Dar SU et al (2023) Unsupervised medical image translation with adversarial diffusion models. *IEEE Trans Med Imaging*. <https://doi.org/10.1109/TMI.2023.3290149>
18. Saharia C, Chan W, Chang H, et al (2022) Palette: image-to-image diffusion models. In: *ACM SIGGRAPH 2022 Conference Proceedings*, pp 1–10. <https://doi.org/10.1145/3528233.3530757>
19. Wolleb J, Sandkühler R, Bieder F, Cattin PC (2022) The Swiss Army knife for image-to-image translation: multi-task diffusion models. *arXiv preprint arXiv:220402641*. <https://doi.org/10.48550/arXiv.2204.02641>
20. Kim B, Oh Y, Ye JC (2022) Diffusion adversarial representation learning for self-supervised vessel segmentation. In: *The Eleventh International Conference on Learning Representations (ICLR)*, 2021. <https://doi.org/10.48550/arXiv.2209.14566>
21. Lyu Q, Wang G (2022) Conversion between CT and MRI images using diffusion and score-matching models. *arXiv preprint arXiv:220912104*. <https://doi.org/10.48550/arXiv.2209.12104>
22. La Barbera G, Boussaid H, Maso F, et al (2022) Anatomically constrained CT image translation for heterogeneous blood vessel segmentation. In: *BMVC 2022 - The 33rd British Machine Vision Conference*. London. <https://doi.org/10.48550/arXiv.2210.01713>
23. Beare R, Lowekamp B, Yaniv Z (2018) Image segmentation, registration and characterization in R with SimpleITK. *J Stat Softw* 86:8. <https://doi.org/10.18637/jss.v086.i08>
24. Pang S, Pang C, Zhao L et al (2020) SpineParseNet: spine parsing for volumetric MR image by a two-stage segmentation framework with semantic image representation. *IEEE Trans Med Imaging* 40:262–273. <https://doi.org/10.1109/TMI.2020.3025087>
25. Pang S, Pang C, Zhao L, et al (2020) MRSpineSeg challenge. In: *spinesegmentation challenge*. <https://www.spinesegmentation-challenge.com/>. Accessed 19 Oct 2022
26. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference*. Springer, pp 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
27. Nichol AQ, Dhariwal P (2021) Improved denoising diffusion probabilistic models. In: *Proceedings of the 38th International Conference on Machine Learning*. PMLR, pp 8162–8171. <https://doi.org/10.48550/arXiv.2102.09672>
28. Dhariwal P, Nichol A (2021) Diffusion models beat gans on image synthesis. In: *Adv. Neural Inf. Process. Syst.* 34 (NeurIPS 2021), pp 8780–8794. <https://doi.org/10.48550/arXiv.2105.05233>
29. Finck T, Li H, Schlaefer S et al (2022) Uncertainty-aware and lesion-specific image synthesis in multiple sclerosis magnetic resonance imaging: a multicentric validation study. *Front Neurosci*. <https://doi.org/10.3389/fnins.2022.889808>
30. Ho J, Salimans T (2021) Classifier-free diffusion guidance. In: *NeurIPS 2021 Workshop on deep generative models and downstream applications*. <https://doi.org/10.48550/arXiv.2207.12598>
31. Bieder F, Wolleb J, Durrer A, et al (2023) Diffusion models for memory-efficient processing of 3D medical images. *arXiv preprint arXiv:230315288* <https://doi.org/10.48550/arXiv.2303.15288>
32. Morbée L, Chen M, Herregods N et al (2021) MRI-based synthetic CT of the lumbar spine: geometric measurements for surgery planning in comparison with CT. *Eur J Radiol* 144:109999. https://doi.org/10.1007/978-3-030-58545-7_19
33. van der Kolk BBY, Slotman DJJ, Nijholt IM et al (2022) Bone visualization of the cervical spine with deep learning-based synthetic CT compared to conventional CT: a single-center noninferiority study on image quality. *Eur J Radiol* 154:110414. <https://doi.org/10.1016/j.ejrad.2022.110414>
34. Haubold J, Demircioglu A, Theysohn JM et al (2021) Generating virtual short tau inversion recovery (STIR) images from T1- and T2-weighted images using a conditional generative adversarial network in spine imaging. *Diagnostics* 11:1542. <https://doi.org/10.3390/diagnostics11091542>
35. Schlaefer S, Drummer K, Hussein ME et al (2023) Implementation of GAN-based, synthetic T2-weighted fat saturated images in the routine radiological workflow improves spinal pathology detection. *Diagnostics* 13:974. <https://doi.org/10.3390/diagnostics13050974>
36. Florkow MC, Zijlstra F, Willemsen K et al (2020) Deep learning-based MR-to-CT synthesis: the influence of varying gradient echo-based MR images as input channels. *Magn Reson Med* 83:1429–1441. <https://doi.org/10.1002/mrm.28008>
37. Hoesl M, Corral NE, Mistry N (2022) White paper: MR-based synthetic CT reimaged - an AI-based algorithm for continuous Hounsfield units in the pelvis and brain - with syngo.via RT image suite (VB60). https://marketing.webassets.siemens-healthineers.com/4db6e75384fa9081/5832ca0e472/siemens-healthineers_syngo-via_white-paper-MR-based-Synthetic-CT.PDF. Accessed 16 Jun 2023
38. Oulbacha R, Kadoury S (2020) MRI to CT synthesis of the lumbar spine from a pseudo-3D cycle GAN. In: *IEEE 17th international symposium on biomedical imaging (ISBI) 2020*. IEEE; pp 1784–1787. <https://doi.org/10.1109/ISBI45749.2020.9098421>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)