Mamba-driven MRI-to-CT Synthesis for MRI-only Radiotherapy Planning

Konstantinos Barmpounakis[1,2], Theodoros P. Vagenas[1], Maria Vakalopoulou[2,3], and George K. Matsopoulos[1]

[1] School of Electrical and Computer Engineering, National Technical University of Athens
kbarbounakis@biomed.ntua.gr
[2] Archimedes, Athena Research Center
[3] CentraleSupélec, University of Paris-Saclay

Abstract. Radiotherapy workflows for oncological patients increasingly rely on multi-modal medical imaging, commonly involving both Magnetic Resonance Imaging (MRI) and Computed Tomography (CT). MRI-only treatment planning has emerged as an attractive alternative, as it reduces patient exposure to ionizing radiation and avoids errors introduced by inter-modality registration. While nnU-Net-based frameworks are predominantly used for MRI-to-CT synthesis, we explore Mamba-based architectures for this task, aiming to showcase the advantages of state-space modeling for cross-modality translation compared to standard convolutional neural networks. Specifically, we adapt both the U-Mamba and the SegMamba architecture, originally proposed for segmentation, to perform cross-modality image generation. Our 3D Mamba architecture effectively captures complex volumetric features and long-range dependencies, thus allowing accurate CT synthesis while maintaining fast inference times. Experiments were conducted on a subset of SynthRAD2025 dataset, comprising registered single-channel MRI–CT volume pairs across three anatomical regions. Quantitative evaluation is performed via a combination of image similarity metrics computed in Hounsefield Units (HU) and segmentation-based metrics obtained from TotalSegmentator to ensure geometric consistency is preserved. The findings pave the way for the integration of state-space models into radiotherapy workflows.

Keywords: synthetic CT · MRI-only based radiotherapy · Mamba · SynthRAD2025.

1 Introduction

Accurate imaging forms the foundation of diagnoses and treatment protocols in oncology, enabling effective clinical decisions. In the context of radiotherapy, the complementary strengths of Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) are leveraged to attain precise tumor targeting with optimal treatment outcomes while minimizing exposure to healthy tissues [1]. The former

modality allows for superior soft-tissue contrast and tumor outlining accuracy, whereas the latter contributes to reliable dose calculation due to the inherent relationship between Hounsefield Units (HU) and electron density.

Since radiotherapy is typically delivered in weekday fractions over several weeks, it would ideally require MRI and CT scans prior to each treatment session. However, repeated CT acquisition increases patient exposure to ionizing radiation beyond the therapeutic dose and is not always technically feasible or economically viable in certain clinical settings. Conventional workflows require registration approaches in order to transfer the anatomical delineations from MRI onto the CT scan used for dose calculation. The accumulation of registration errors may further degrade the treatment performance. Therefore, a simplified MRI-only workflow emerges as a promising alternative, eliminating the need for multiple CT scans [2,3]. Recent advances in medical imaging computation, including cross-modality translation using deep learning techniques, enable the accurate synthesis of CT images directly from MRI data [4].

A major limitation in the development of MRI-only workflows is the scarcity of large-scale, heterogeneous and well-curated paired MRI-CT datasets. Constructing such datasets is a tedious and resource-intensive process, encompassing multi-institutional coordination, standardized acquisition protocols, quality control, ethical approval as well as carefully designed pre-processing pipelines and registration procedures. Despite these constraints, several studies have explored MRI-only workflows leveraging unpaired data in an unsupervised or semi-supervised manner. For instance, Brou Boni et al. proposed an unpaired Cycle-GAN framework to synthesize CT from MRI eliminating the need for paired data [5], while MaskGAN enforces anatomical consistency through automatically extracted structural masks during unpaired MR–CT translation [6].

This paper introduces the use of Mamba-based architectures for MRI-to-CT synthesis across multiple anatomical regions. To the best of our knowledge, this work represents the first integration of the U-Mamba and SegMamba frameworks into a generative medical imaging task [7,8]. We further propose a compound loss scheme specifically designed for this task and ultimately demonstrate the robustness and the efficiency of the Mamba modules in CT synthesis, establishing them as a viable alternative to conventional nnU-Net and transformer-based architectures.

2 Methodology

In this study, synthetic CT generation is formulated as a supervised image-to-image translation task, where the network learns to predict CT intensity distributions from spatially aligned MRI volumes. Given paired MRI–CT data, we investigate five encoder–decoder deep learning frameworks extending across multiple state-of-the-art architectural backbones, involving a conventional U-Net [9], the self-configuring nnU-Net pipeline [10,11], a SwinUNETR [12] as a transformer-based alternative and the two currently leading state-space-driven

models in medical imaging, namely SegMamba and U-Mamba [7,8]. The methodological approach is displayed in Figure 1.

Architectures Convolutional Neural Networks (CNNs) have long been the dominant paradigm for 3D medical image tasks due to their strong inductive bias toward local spatial context, enabling precise voxel-wise predictions and stable training even with limited data. Building upon the standard U-Net design, nnU-Net introduces a self-configuring framework that automatically adapts preprocessing, normalization, network topology and training settings to a given dataset, providing a robust baseline by eliminating manual hyper-parameter tuning. Despite their effectiveness, purely convolutional architectures exhibit limited receptive fields and struggle to capture long-range dependencies inherent in diverse applications. Transformer-based hybrids such as SwinUNETR serve as a prominent alternative by integrating hierarchical self-attention through shifted window mechanisms and drawing global dependencies between input and output [13]. However, transformers exhibit quadratic computational complexity with respect to the number of tokens and typically require large-scale training data to achieve competitive performance.

State-space models (SSMs) perform sequential data processing through latent dynamical systems combining long-range modeling capabilities with favorable computational scaling [14]. Mamba introduces selective state-space blocks that integrate SSMs productively into deep learning frameworks. In terms of network design, U-Mamba retains the self-configuring characterisitcs of nnU-Net and replaces conventional convolutional blocks with Mamba modules, whereas SegMamba employs a hybrid hierarchical architecture that fuses state-space components with convolutional layers, demonstrating strong performance in volumetric segmentation tasks.

Loss function A compound loss strategy is adopted to balance global intensity reconstruction and anatomically meaningful refinement during training. In reconstruction tasks, the Mean Absolute Error (MAE) is commonly used as a baseline objective due to its robustness to outliers and its ability to preserve overall intensity fidelity. However, conventional MAE treats all voxels equally and does not account for the heterogeneous clinical relevance of different regions. To address this limitation, we introduce a weighted MAE in which voxel-wise errors are masked based on the HU intensity of the target CT. The weighting function that generates a 3D map $\boldsymbol{\omega} \in \mathbf{R}^{H \times W \times D}$ based on the target CT intensities in HU is expressed as:

$$\omega(x, y, z) = \begin{cases} 3, & t(x, y, z) > 300, \\ 1.5, & -700 < t(x, y, z) \leq 300, \\ 0.5, & t(x, y, z) \leq -700, \end{cases} \tag{1}$$

where (x, y, z) denotes the spatial voxel coordinates. The values have been selected to address domain-specific requirements. For instance, voxels with $y > 300$

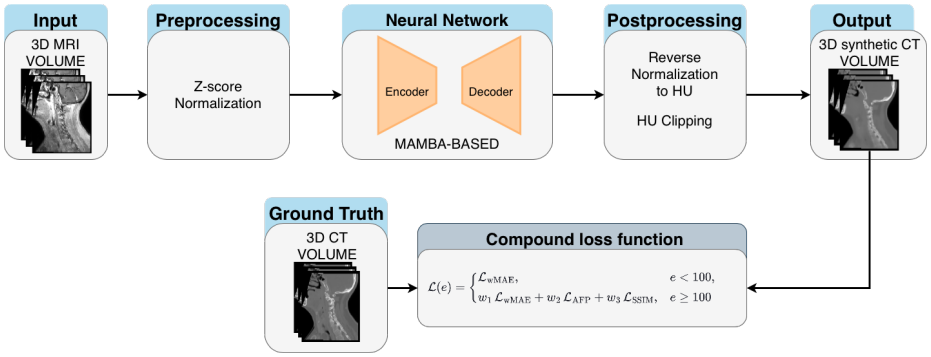

Fig. 1. Illustration of the overall proposed framework. Input MRI volumes undergo Z-score normalization and feed a Mamba-based encoder–decoder architecture. The predicted output is transformed back to HU via reverse normalization on dataset-level and appropriate clipping. During training, the network is supervised using paired 3D CT volumes with a staged compound loss function

correspond to bone structures and receive the highest penalty to promote accurate reconstruction of dense regions that are important for downstream tasks such as dose calculation [16]. The range $-700 < y \leq 300$ represents soft tissue, including most organs, muscle, fat or water-equivalent regions and is assigned an intermediate weight to preserve structural consistency and intensity contrast. Finally, voxels with $y < -700$ correspond to air or background and receive the lowest weight due to their minimal diagnostic information. The weighted MAE is thus formulated below:

$$\mathcal{L}_{\mathrm{wMAE}} = \frac{\sum_{x,y,z} \omega(x,y,z)\ \left|\hat{t}(x,y,z) - t(x,y,z)\right|}{\frac{1}{HWD}\sum_{x,y,z} \omega(x,y,z)}, \tag{2}$$

where $\hat{t}(x,y,z)$ and $t(x,y,z)$ denote correspondingly the predicted and ground-truth CT intensities at voxel location (x,y,z). The normalization by the average value of $\boldsymbol{\omega}$ ensures numerical stability across batches.

To promote structural consistency during training, we incorporate the Structural Similarity Index Measure (SSIM) [17], which evaluates local luminance, contrast as well as structural agreement. The SSIM loss is thereby defined as:

$$\mathcal{L}_{\mathrm{SSIM}} = 1 - SSIM(\hat{t}, t). \tag{3}$$

Arthur et al. recently introduced the Anatomical Feature-Prioritized (AFP) loss, a novel objective function designed to enhance the reconstruction of localized anatomical details in medical image synthesis [18]. Unlike conventional pixel-wise losses that primarily optimize global intensity similarity, AFP operates in feature space by comparing multi-scale embeddings extracted from a pre-trained foundational model. Inspired by perceptual loss formulations, the method leverages hierarchical representations learned for anatomical delineation,

encouraging the conformation of clinically relevant structures rather than solely minimizing voxel-wise differences. In our framework, AFP is incorporated into the compound loss through TotalSegmentator embeddings [19], guiding the synthesis process toward improved geometric consistency and anatomically meaningful reconstructions. The overall training objective has the following structure:

$$\mathcal{L}(e) = \begin{cases} \mathcal{L}_{\text{wMAE}}, & e < 100, \\ w_1\,\mathcal{L}_{\text{wMAE}} + w_2\,\mathcal{L}_{\text{AFP}} + w_3\,\mathcal{L}_{\text{SSIM}}, & e \geq 100, \end{cases} \tag{4}$$

where e denotes the current epoch. The staged formulation of the loss function reflects the progressive learning strategy adopted in this work. During the initial training phase, models are optimized using the intensity-aware $\mathcal{L}_{\text{wMAE}}$, which encourages stable convergence and promotes accurate reconstruction of global CT intensity distributions, thereby enabling the network to first learn realistic medical volume representations. After the 100th epoch, the structural similarity loss $\mathcal{L}_{\text{SSIM}}$ and the AFP term $\mathcal{L}_{\text{AFP}}$ are appended to further guide optimization toward improved geometric consistency and anatomically meaningful details. This progressive loss scheduling allows the model to transition from intensity-driven learning to structure-aware refinement.

Quantitative Evaluation Quantitative evaluation is performed using standard image similarity metrics, including MAE, Peak Signal-to-Noise Ratio (PSNR) and Multi-Scale SSIM (MS-SSIM), which extends SSIM by averaging structural similarity across multiple resolution scales [20]. In addition, anatomical consistency is assessed through downstream segmentation analysis by computing Dice Similarity Coefficient (DSC) and the 95th percentile of Hausdorff Distance (HD95) between several masks obtained with TotalSegmentator on the generated and target CT volumes [19,21].

3 Experiments

3.1 Dataset & Preprocessing

The SynthRAD2023 Grand Challenge dataset was developed to overcome the scarcity of co-registered modalities, providing 540 rigidly registered MRI-CT pairs obtained from three Dutch medical centers across two different anatomical regions: brain and pelvis [22]. Building on this foundation, the SynthRAD2025 Grand Challenge dataset substantially expands both the scale and diversity of paired data, comprising 890 MRI–CT pairs spanning head-and-neck, thoracic and abdominal sites across five European medical centers [23].

The current work utilizes a curated subset of the SynthRAD2025 dataset. In particular, we focus on the MRI–CT synthesis subtask, which comprises paired data from four European medical centers across three anatomical regions: abdomen (AB), head-and-neck (HN), and thorax (TH). Data from center D were excluded due to usage restrictions. In total, 890 MRI–CT pairs are provided

together with an outline mask defining the patient's body region. The dataset follows an official 65/10/25 split into training, validation and test sets, respectively, where only the training percentage is fully accessible. Consequently, we used the training set and further divided it into 90%/10% subsets, employing the larger portion for model development and keeping the remaining 10% fully unseen until the evaluation procedure. The resulting subset includes 461 training cases and 52 test cases, distributed across the three anatomical regions as follows: 157 training and 18 test cases for abdomen (AB), 141 and 15 for head-and-neck (HN), and 163 and 19 for thorax (TH), respectively.

The provided SynthRAD2025 data have undergone an initial preprocessing pipeline including rigid registration of MRI to CT using the Elastix framework to ensure spatial alignment [24]. Facial structures were removed for anonymization and then all volumes were resampled to a common voxel spacing of $1 \times 1 \times 3\,\text{mm}$ before being cropped with a 10-pixel margin around automatically generated patient outline masks. In addition, we apply supplementary preprocessing performing per-scan z-score normalization of MRI volumes, whereas CT intensities were clipped to $[-1024, 1500]$ HU followed by dataset-level z-score normalization, without further modifying the shared voxel spacing.

3.2 Implementation Details

All models were trained from scratch on an NVIDIA A100-SXM4-80GB GPU. Baseline U-Net and SwinUNETR were implemented within the SegMamba training pipeline using a patch size of $[64, 192, 192]$ and a batch size of 2. Patch sampling was guided by the accompanying outline masks to ensure a minimum body coverage of 70% per patch, without additional data augmentation. Default self-configuration procedures were retained for nnU-Net and U-Mamba, adapted to the generative setting. Optimization was performed using AdamW with an initial learning rate of 5×10^{-4} and a polynomial decay schedule [25], whereas nnU-Net followed its standard stochastic gradient descent configuration. All models were trained for 500 epochs except nnU-Net which was trained for 1000 epochs following its default training schedule. The proposed compound loss function was used consistently across all experiments. The source code will be made publicly available upon acceptance of the manuscript.

3.3 Results

The quantitative results are gathered in Table 1, while qualitative comparisons are illustrated in Figure 2. SegMamba achieves the lowest MAE among all compared methods, indicating improved voxel-wise intensity fidelity in the synthesized CT volumes. In terms of image similarity, nnU-Net attains the highest PSNR and MS-SSIM scores, reflecting strong global reconstruction quality and multi-scale structural preservation.

Within the identical training pipeline, using the same data split and random seed as well, SegMamba consistently outperforms both U-Net and SwinUNETR across all evaluation metrics. Although SegMamba achieves a slightly lower MAE

compared to nnU-Net, it fails to surpass nnU-Net in the remaining similarity and geometric consistency metrics. This performance gap may be attributed to the fully self-configuring design of the nnU-Net framework, its extended training schedule (1000 vs. 500 epochs) and its larger model capacity (88M parameters for nnU-Net vs. 67M for SegMamba).

Similarly, U-Mamba also outperforms the baseline architectures and remains competitive with nnU-Net, despite being trained for fewer epochs (500 vs. 1000) and having a lower model capacity (73M vs. 88M parameters). These findings further support the effectiveness of Mamba-based designs in this cross-modality translation setting.

The highest DSC and lowest HD95 imply correspondingly improved volumetric overlap and more accurate boundary delineation. While nnU-Net excels quantitavely regarding geometric consistency, the Mamba-based models demonstrate more stable visual coherence across different anatomical regions, with fewer localized artifacts and improved intensity homogeneity. These qualitative observations align with the quantitative improvements in MAE and competitive structural metrics in comparison with nnU-Net.

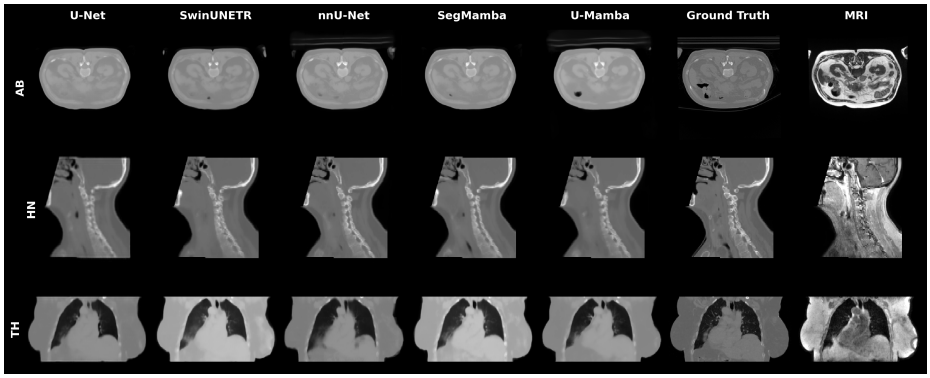

Fig. 2. Visual comparison of U-Net, SwinUNETR, nnU-Net, SegMamba, and U-Mamba on representative cases for each anatomical site. The target CT images and corresponding MRI images are shown in the last columns

4 Conclusion

The current study probes the suitability of Mamba-based architectures for MRI-to-CT synthesis within the context of radiotherapy planning. Experiments were conducted on a curated subset of the SynthRAD2025 dataset across three anatomical regions, where the U-Mamba and SegMamba frameworks were adapted to image-to-image translation task and evaluated under a controlled experimental protocol. In addition, we explored a weighted HU intensity-aware loss designed

Table 1. Quantitative comparison of image similarity and geometric consistency metrics. Arrows indicate whether higher (↑) or lower (↓) values are preferred. For each model, the value reported in parentheses denotes its number of parameters (in millions).

Model	**MAE ↓**	**PSNR ↑**	**MS-SSIM ↑**	**DSC ↑**	**HD95 ↓**
U-Net	109.86 ± 24.81	25.52 ± 1.56	0.8376 ± .0719	0.564 ± .142	13.18 ± 6.79
SwinUNETR	106.97 ± 25.46	25.62 ± 1.60	0.8395 ± .0714	0.522 ± .153	18.07 ± 10.65
nnU-Net	103.72 ± 23.20	**25.90 ± 1.59**	**0.8492 ± .0687**	**0.630 ± .138**	**8.55 ± 4.90**
Seg-Mamba	**103.33 ± 25.29**	25.82 ± 1.61	0.8481 ± .0685	0.594 ± .146	11.60 ± 6.75
U-Mamba	103.37 ± 23.05	25.80 ± 1.52	0.8465 ± .0653	0.580 ± .143	13.25 ± 7.34

to emphasize clinically relevant tissue ranges, thereby boosting radiotherapy accuracy and supporting optimized MRI-only treatment planning workflows. The assessed models outperformed conventional convolutional and transformer-based baselines and remained competitive with the widely adopted nnU-Net framework, despite reduced training epochs and lower model capacity. The results demonstrate that selective state-space models constitute a viable alternative in cross-modality translation domain. Future work will investigate scaling the capacity of Mamba-based architectures and further optimization of the current training framework and anatomy-specific model specialization to fully exploit their representational capabilities.

Acknowledgments. This work has been partially supported by project MIS 5154714 of the National Recovery and Resilience Plan Greece 2.0 funded by the European Union under the NextGenerationEU Program. The authors gratefully acknowledge NVIDIA Corporation for the GPU hardware grant that facilitated the conducted computational experiments.

Disclosure of Interests. The authors declare that they have no competing interests.

References

1. Schmidt, M.A., Payne, G.S.: Radiotherapy planning using MRI. Physics in Medicine and Biology **60**(22), R323–R361 (2015). https://doi.org/10.1088/0031-9155/60/22/R323
2. Edmund, J.M., Nyholm, T.: A review of substitute CT generation for MRI-only radiation therapy. Radiotherapy and Oncology **12**, 28 (2017). https://doi.org/10.1186/s13014-016-0747-y
3. Owrangi, A.M., Greer, P.B., Glide-Hurst, C.K.: MRI-only treatment planning: benefits and challenges. Physics in Medicine and Biology **63**(5), 05TR01 (2018). https://doi.org/10.1088/1361-6560/aaaca4
4. Acquah, I.K., et al.: Deep learning-based synthetic CT generation from MRI for enhanced precision in MRI-only radiotherapy dose planning. Polish Journal of Medical Physics and Engineering **31**(3), 219–226 (2025). https://doi.org/10.2478/pjmpe-2025-0025

5. Brou Boni, K.N.D., Klein, J., Gulyban, A., Reynaert, N., Pasquier, D.: Improving generalization in MR-to-CT synthesis in radiotherapy by using an augmented cycle generative adversarial network with unpaired data. Medical Physics **48**(6), 3003–3010 (2021). https://doi.org/10.1002/mp.14866
6. Phan, V.M.H., Liao, Z., Verjans, J.W., To, M.S.: Structure-preserving synthesis: MaskGAN for unpaired MR–CT translation. In: Greenspan, H., et al. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2023, LNCS, vol. 14229, pp. 63–73. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-43999-5_6
7. Ma, J., Li, F., Wang, B.: U-Mamba: Enhancing Long-range Dependency for Biomedical Image Segmentation. *arXiv preprint arXiv:2401.04722* (2024).
8. Xing, Z., Ye, T., Yang, Y., Liu, G., Zhu, L.: SegMamba: Long-range sequential modeling Mamba for 3D medical image segmentation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2024, LNCS, vol. 15008, pp. 578–588. Springer Nature Switzerland (2024).
9. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
10. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. Nature Methods **18**(2), 203–211 (2021). https://doi.org/10.1038/s41592-020-01008-z
11. Isensee, F., Wald, T., Ulrich, C., Baumgartner, M., Roy, S., Maier-Hein, K., Jäger, P.F.: nnU-Net revisited: A call for rigorous validation in 3D medical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2024. Lecture Notes in Computer Science, vol. 15009, pp. 488–498. Springer Nature Switzerland (2024)
12. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin UNETR: Swin transformers for semantic segmentation of brain tumors in MRI images. In: Crimi, A., Bakas, S. (eds.) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2021. Lecture Notes in Computer Science, vol. 12962, pp. 272–284. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-08999-2_22
13. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. In: Guyon, I., von Luxburg, U., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (eds.) Advances in Neural Information Processing Systems, vol. 30. Curran Associates, Inc. (2017)
14. Gu, A., Goel, K., Ré, C.: Efficiently modeling long sequences with structured state spaces. In: International Conference on Learning Representations (ICLR) (2022)
15. Gu, A., Dao, T.: Mamba: Linear-Time Sequence Modeling with Selective State Spaces. In: International Conference on Learning Representations (ICLR) (2024)
16. Nie, D., Trullo, R., Lian, J., Wang, L., Petitjean, C., Ruan, S., Wang, Q., Shen, D.: Medical Image Synthesis with Context-Aware Generative Adversarial Networks. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2017. LNCS, vol. 10435, pp. 417–425. Springer, Cham (2017).
17. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4), 600–612 (2004). https://doi.org/10.1109/TIP.2003.819861

18. Longuefosse, A., Denis de Senneville, B., Dournes, G., Benlala, I., Baldacci, F., Desbarats, P.: Anatomical feature-prioritized loss for enhanced MR to CT translation. Physics in Medicine & Biology **70**(14) (2025). https://doi.org/10.1088/1361-6560/adea07
19. Wasserthal, J., Meyer, M., Breit, H.-C., Cyriac, J., Yang, S., Segeroth, M.: TotalSegmentator: Robust segmentation of 104 anatomic structures in CT images. Radiology: Artificial Intelligence **5**(5), e230024 (2023). https://doi.org/10.1148/ryai.230024
20. Dohmen, M., Klemens, M.A., Baltruschat, I.M., et al.: Similarity and quality metrics for MR image-to-image translation. Scientific Reports **15**, 3853 (2025). https://doi.org/10.1038/s41598-025-87358-0
21. Celaya, A., Rivière, B.M., Fuentes, D.T.: A generalized surface loss for reducing the Hausdorff distance in medical imaging segmentation. In: Medical Imaging with Deep Learning (MIDL) (2023)
22. Thummerer, A., van der Bijl, E., Galapon Jr, A., Verhoeff, J.J.C., Langendijk, J.A., Both, S., van den Berg, C.N.A.T., Maspero, M.: SynthRAD2023 Grand Challenge dataset: Generating synthetic CT for radiotherapy. Medical Physics **50**(7), 4664–4674 (2023). https://doi.org/10.1002/mp.16529
23. Thummerer, A., van der Bijl, E., Galapon Jr, A., Kamp, F., Savenije, M., Muijs, C., Aluwini, S., Steenbakkers, R.J.H.M., Beuel, S., Intven, M.P.W., Langendijk, J.A., Both, S., Corradini, S., Rogowski, V., Terpstra, M., Wahl, N., Kurz, C., Landry, G., Maspero, M.: SynthRAD2025 Grand Challenge dataset: Generating synthetic CTs for radiotherapy from head to abdomen. Medical Physics **52**(7), e17981 (2025). https://doi.org/10.1002/mp.17981
24. Shamonin, D.P., Bron, E.E., Lelieveldt, B.P.F., Smits, M., Klein, S., Staring, M.: Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease. Frontiers in Neuroinformatics 7(50), 1–15 (2014)
25. Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. In: International Conference on Learning Representations (ICLR) (2019)