USC Stevens Neuroimaging and Informatics Institute

SynthDiffusion: Diffusion-Based Modeling On Computed Tomography for MR-Only Radiotherapy

Lujia Zhong, Zhiwei Deng

Laboratory of Neuro Imaging (LONI),

USC Stevens Neuroimaging and Informatics Institute, Keck School of Medicine of USC, Los Angeles, CA

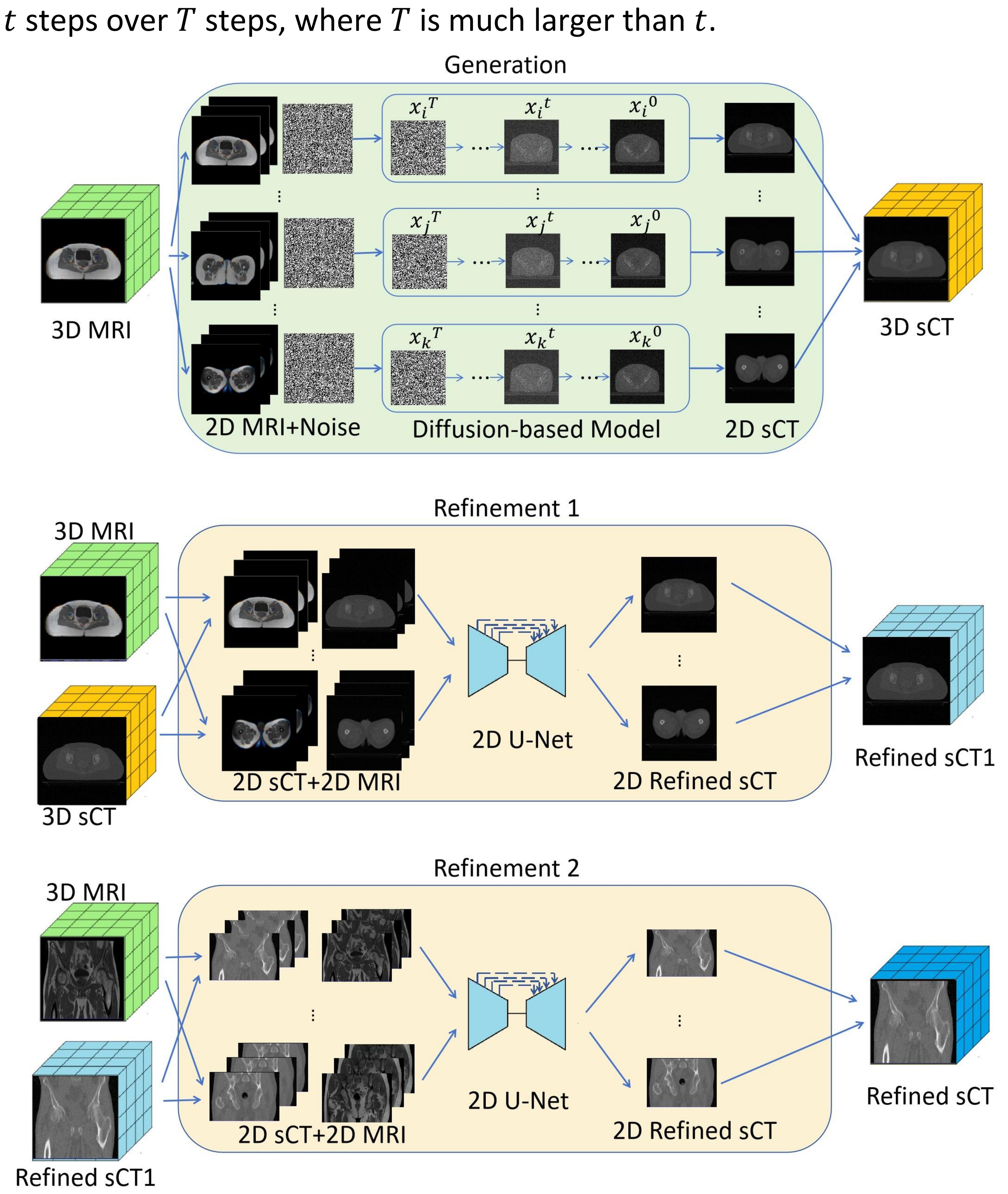
lujiazho@usc.edu

Introduction

- To obtain the CT images and corresponding tissue attenuation information for dose calculation and treatment planning, additional radiation to the patients are usually inevitable.
- The key of utilizing CT for dose calculation is the tissue attenuation information, which is not directly available on MR data and the superb soft-tissue contrast of MR is not enough.
- Developing MRI-only RT can not only reduce the radiation to patients, but also help to simplify and speed up the workflow for disease cure, which requires highly effective modeling for generating CT given MR data.
- Our work integrates the diffusion-based model for the synthesized CT data generation, which can generate highly accurate and realistic CT based on MR.
- We equip the diffusion model with additional two refinement networks for better smoothness and less artifacts due to the limitation of computation in this challenge.

Methods

- **Variational Diffusion Model.** The Variational Diffusion Model can be simply viewed as a Markovian Hierarchical Variational Autoencoder with additional constraints: (1) The latent space is not learned but instead it's pre-defined as a perturbed image space with Gaussian noise, which means each latent space has the same dimension as the initial image space; (2) The final latent space after T timesteps is theoretically a standard Gaussian space.
- **Workflow.** As shown below, each MRI slice is padded and re-sampled to the size of 256 × 256 firstly, and then the diffusion model with slice-wise consistency constraint is employed to generate the sCT images. After that, the generated sCT images are re-sampled to the same shape as the original MRI images. Finally, two vanilla U-Net models are employed to remove the noise and artifacts.
- **Refinement Networks**. The refinement networks are not required for generating high-quality synthesized CT data. Without computation and time limitations, we can follow the whole reverse process of adding Gaussian noise and generate no-noise and no-artifact data. But to accommodate the diffusion model, which is relatively more expensive to GAN, we only sample a few steps and deploy refinement networks for further denoising.
- **DDIM Sampling**. Even though a diffusion model is trained with T timesteps and hyperparameters $\overline{\alpha}$ and $\overline{\beta}$, we can consider it being trained with fewer timesteps t with another set of hyperparameters $\overline{\alpha}'$ and $\overline{\beta}'$. Therefore, with a well-trained diffusion model, we can expedite its sampling process to only t steps over T steps, where T is much larger than t.

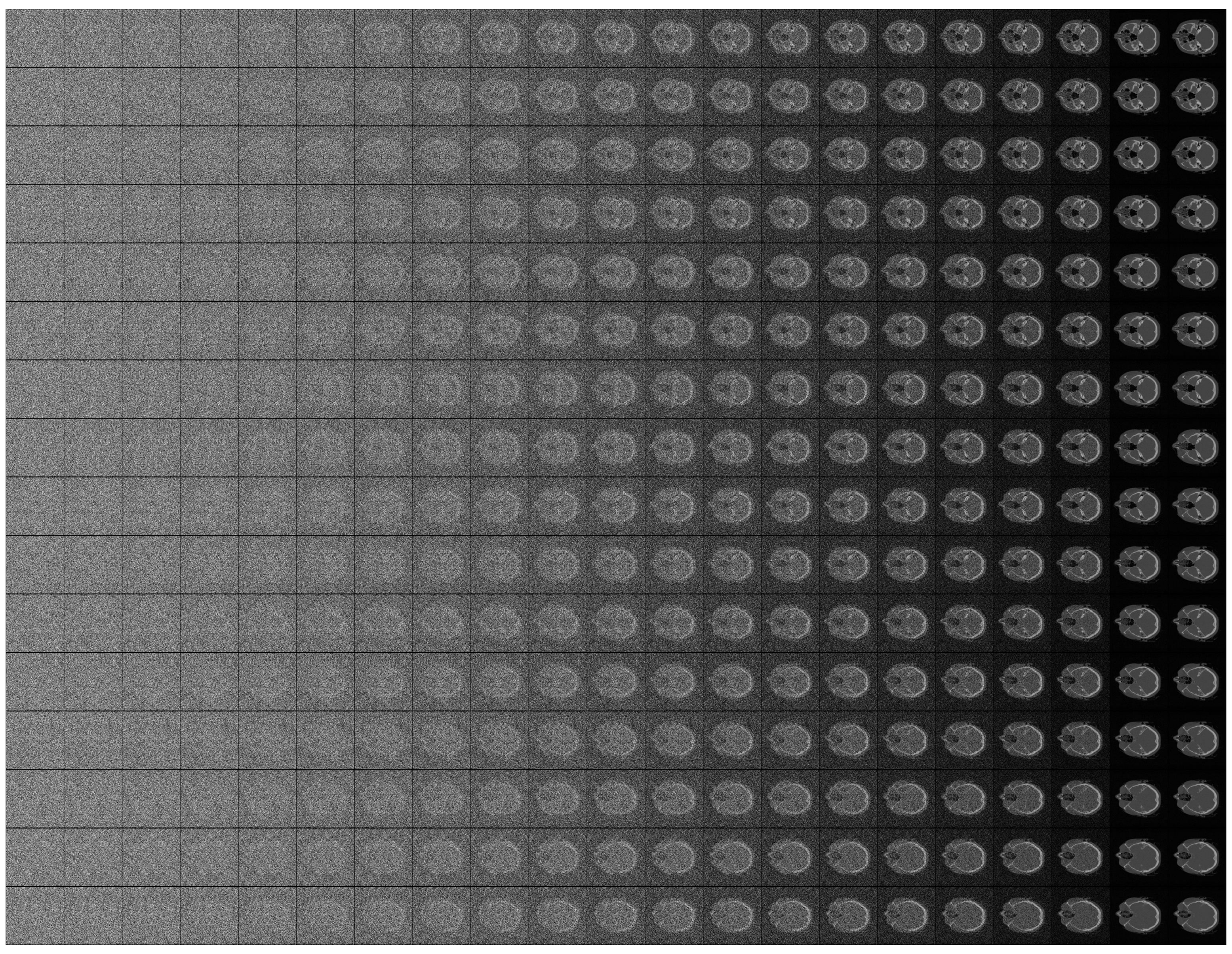


Results & Analysis

- **Diffusion Model vs. GAN and U-net variants.** Diffusion models are known to suffer less from mode collapse, which is very common in GAN models, which means Diffusion model can generate a large diversity of samples but this can also lead to unappreciated randomness and larger variance, especially in medical image processing. Compared to various U-net variants, the Diffusion models also have larger diversity on generation. This can be seen in the Challenge results as shown below, where our best prediction on CT data surpass all other teams' best prediction, but also generally larger standard deviation on predictions.

Method		MAE↓	PSNR↑	DDIM↑
MSEP (Rank 1)	Max	107.73	33.41	0.9472
	Min	36.67	24.63	0.7852
	Std	13.40	1.78	0.0288
Conditional GAN (Rank 4)	Max	108.64	35.52	0.9743
	Min	29.59	24.55	0.7841
	Std	13.06	1.60	0.0297
Ours (Rank 8)	Max	113.2	**37.57**	**0.9849**
	Min	**27.10**	23.95	0.7715
	Std	14.40	1.74	0.0339

- **Sampling strategy.** In our case, limited by Diffusion model's nature of much longer inference time, sampling phase are every bit as important as the training phase. Original DDPM won't survive in any challenge that has strict running time limit. DDIM is chosen to accelerate the sampling process, but which still cannot completely denoise the image and shift the mean of samples from random Gaussian zero mean to real CT mean. Better sampling strategies could lead to better generation, we leave which for future exploration. An example of DDIM sampling with only 20 steps is shown below.



Conclusions

In this paper, we developed a Diffusion-based modeling for generating high-quality synthetic CT data given provided MR data. For the balance between computation and time consumption and data quality, we proposed two following refinement networks that can largely reduce inference time of Diffusion model and improve data quality by further denoising and mean shifting. For future work, we will explore more power of sampling methods and conditional Diffusion model structures that better utilizes MR information for faster and better generation.

References

- Chernak, E.S., Rodriguez-Antunez, A., Jelden, G.L., Dhaliwal, R.S. and Lavik, P.S., 1975. The use of computed tomography for radiation therapy treatment planning. Radiology, 117(3), pp.613-614.
- Luo, C., 2022. Understanding diffusion models: A unified perspective. arXiv preprint arXiv:2208.11970.
- Ronneberger, O., Fischer, P. and Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 (pp. 234-241). Springer International Publishing.
- Song, J., Meng, C. and Ermon, S., 2020. Denoising diffusion implicit models. arXiv preprint arXiv:2010.02502.
- Yang, L., Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., Shao, Y., Zhang, W., Cui, B. and Yang, M.H., 2022. Diffusion models: A comprehensive survey of methods and applications. arXiv preprint arXiv:2209.00796.

LONI

NICR Neuro Image Computing Research

Acknowledgement: This work is supported by the National Institute of Health (NIH) under grants R01EB022744, RF1AG077578, RF1AG056573, RF1AG064584, R21AG064776, U19AG078109, and P41EB015922.