MORE: Multi-Organ Tomographic REconstruction Dataset
Appendix

A Project Page

We provide a webpage for our **MORE** dataset.

B Code for Reproducibility

We provide the code for the methods in our benchmark in https://huggingface.co/datasets/WSKINGDOM/MORE/blob/main/code.zip. Now these codes are under the process of organizing and will be publicly available on GitHub soon.

C Dataset Details

Scan Parameters The CT scans were acquired using a Siemens SOMATOM Definition AS+ scanner. The detailed scan parameters are included in the attachment ScanParameters.csv.

Metadata We include the metadata of the dataset in the attachment 'Metadata.csv'. The metadata includes the patient ID, anatomy, lesion, and the number of scans for each instance.

D Experiment Details

Implementation Details RED-CNN [2] is a denoising method that uses a residual encoder-decoder network to remove noise from the input image. Therefore, the input of RED-CNN is the noisy image reconstructed by FBP (for CT), and then outputs a denoised image. On the other hand, the training process of AUTOMAP involves implementing a deep neural network with a feed-forward architecture composed of fully connected layers followed by sparse convolutional autoencoders. The network is trained to map sparse measurements to high-quality images.

The diffusion-based methods MCG [4], DiffusionMBIR [3], and SWORD [7] learn the diffusion process, and the reconstruction iteratively updates the image by the diffusion process. For the NeRF-based method NeRP [6], which does not need training data, directly uses a neural network to implicitly model the 3D volume and update the parameters under the supervision of the sparse measurements.

Our GIFT also does not need training data. Compared to the implicit modeling of NeRF, GIFT explicitly models the 3D volume as a set of Gaussians and updates the parameters of the Gaussians under the supervision of the sparse measurements. The Gaussians are Initialized with random means and standard deviations, and the intensity is initialized with the average intensity of the sparse measurements. Every 100 iterations, we adaptively densify the Gaussians to ensure the reconstruction quality inspired by 3DGS [5].

More details about the implementation of the benchmarks are provided in the code repository.

E Full Benchmark Results

From Table 5 to Table 19, we provide the full benchmark results on the 15 types of les in our dataset. The results are reported in terms of PSNR and SSIM following the main paper. The best results are highlighted in **bold**.

Table 5: Benchmark on Emphysema.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	29.58	0.714	28.44	0.644	27.06	0.623	27.27	0.588
MCG [4]	✓	32.72	0.820	32.84	0.821	34.47	0.843	32.90	0.820
DiffusionMBIR [3]	✓	32.58	0.933	32.64	0.936	32.45	0.932	32.24	0.932
SWORD [7]	✓	35.38	0.879	34.52	0.864	33.78	0.849	32.30	0.827
FBP [1]	×	18.55	0.365	16.29	0.293	14.77	0.248	12.03	0.193
NeRP [6]	×	25.41	0.744	25.21	0.735	25.40	0.745	25.39	0.745
R^2-Gaussian [8]	×	38.88	0.943	38.52	0.939	37.95	0.932	37.69	0.928
GIFT (Ours)	×	**39.47**	**0.950**	**39.04**	**0.946**	**38.42**	**0.941**	**38.04**	**0.937**

Table 6: Benchmark on Ureteral Calculi.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	37.04	0.901	35.63	0.913	32.07	0.759	31.46	0.844
MCG [4]	✓	37.94	0.901	37.99	0.901	38.04	0.902	38.05	0.902
DiffusionMBIR [3]	✓	38.37	0.968	38.24	0.967	38.13	0.967	38.90	0.966
SWORD [7]	✓	42.35	0.973	40.93	0.967	39.42	0.960	37.63	0.947
FBP [1]	×	23.09	0.515	19.42	0.462	16.89	0.416	14.02	0.355
NeRP [6]	×	26.91	0.801	26.68	0.789	26.95	0.802	26.66	0.785
R^2-Gaussian [8]	×	41.37	0.971	41.09	0.966	40.05	0.962	39.45	0.956
GIFT (Ours)	×	**43.43**	**0.982**	**42.24**	**0.980**	**40.82**	**0.976**	**40.11**	**0.975**

Table 7: Benchmark on Rib Fracture.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	29.61	0.707	28.97	0.682	27.94	0.658	27.97	0.585
MCG [4]	✓	34.81	0.851	34.94	0.852	34.96	0.853	35.07	0.854
DiffusionMBIR [3]	✓	34.64	0.950	34.64	0.952	34.54	0,951	34.35	**0.950**
SWORD [7]	✓	36.51	0.877	35.90	0.864	35.53	0.855	34.76	0.838
FBP [1]	×	19.33	0.388	16.64	0.324	14.76	0.280	12.69	0.230
NeRP [6]	×	25.77	0.778	25.10	0.744	25.63	0.771	25.60	0.769
R^2-Gaussian [8]	×	40.72	0.960	39.48	0.952	38.65	0.948	37.68	0.943
GIFT (Ours)	×	**42.43**	**0.972**	**41.05**	**0.962**	**40.01**	**0.953**	**39.43**	0.948

Table 8: Benchmark on Appendicitis.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	36.96	0.904	35.54	0.906	31.30	0.838	32.59	0.854
MCG [4]	✓	38.76	0.908	38.96	0.909	38.97	0.897	38.36	0.899
DiffusionMBIR [3]	✓	38.34	0.960	38.28	0.959	38.24	0.966	38.00	0.967
SWORD [7]	✓	**44.18**	0.976	**42.62**	0.971	40.85	0.964	37.79	0.949
FBP [1]	×	23.37	0.516	19.63	0.462	18.17	0.427	14.67	0.366
NeRP [6]	×	27.15	0.821	27.25	0.817	27.38	0.819	27.28	0.817
R^2-Gaussian [8]	×	41.47	0.964	40.78	0.959	40.16	0.954	39.79	0.948
GIFT (Ours)	×	42.03	**0.981**	41.63	**0.981**	**41.15**	**0.979**	**40.22**	**0.976**

ACM MM, 2025, Dublin, Ireland
2025. ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Table 9: Benchmark on Pneumonia.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	31.78	0.733	30.43	0.672	29.22	0.680	27.82	0.578
MCG [4]	✓	32.87	0.810	33.05	0.813	33.19	0.814	33.33	0.815
DiffusionMBIR [3]	✓	33.34	0.954	33.26	0.953	33.10	0.952	32.86	**0.951**
SWORD [7]	✓	39.69	0.901	38.75	0.887	38.02	0.875	36.40	0.850
FBP [1]	×	17.57	0.323	15.73	0.264	14.66	0.229	12.73	0.182
NeRP [6]	×	25.52	0.694	26.16	0.733	25.93	0.722	25.64	0.701
R^2-Gaussian [8]	×	39.73	0.953	38.96	0.949	38.40	0.944	37.92	0.938
GIFT (Ours)	×	**41.77**	**0.967**	**40.96**	**0.962**	**40.31**	**0.956**	**39.11**	0.946

Table 10: Benchmark on Cerebral Hemorrhage.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	35.47	0.895	33.29	0.864	30.46	0.786	29.26	0.766
MCG [4]	✓	39.14	0.898	39.23	0.899	39.32	0.899	39.31	0.899
DiffusionMBIR [3]	✓	39.04	0.969	39.29	0.973	39.05	0.971	38.53	0.969
SWORD [7]	✓	34.90	0.742	33.50	0.740	31.86	0.737	29.57	0.732
FBP [1]	×	24.13	0.526	21.54	0.490	19.70	0.460	17.52	0.413
NeRP [6]	×	25.38	0.789	25.98	0.804	25.02	0.760	24.23	0.764
R^2-Gaussian [8]	×	40.97	0.968	40.54	0.964	39.88	0.960	38.79	0.955
GIFT (Ours)	×	**43.71**	**0.984**	**42.94**	**0.981**	**41.68**	**0.978**	**40.56**	**0.974**

Table 11: Benchmark on Kidney Stones.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	36.65	0.882	34.70	0.909	31.98	0.802	30.89	0.798
MCG [4]	✓	38.16	0.909	38.43	0.911	38.49	0.912	38.67	0.913
DiffusionMBIR [3]	✓	28.84	0.964	38.92	0.966	38.79	0.964	38.54	0.964
SWORD [7]	✓	43.58	0.980	42.27	0.976	40.95	0.971	39.51	0.961
FBP [1]	×	22.88	0.483	19.39	0.439	16.27	0.398	13.52	0.341
NeRP [6]	×	26.17	0.767	26.25	0.773	26.11	0.772	26.16	0.776
R^2-Gaussian [8]	×	44.00	0.983	42.94	0.978	41.97	0.975	40.26	0.969
GIFT (Ours)	×	**44.37**	**0.988**	**43.45**	**0.987**	**42.99**	**0.986**	**41.20**	**0.982**

Table 12: Benchmark on Fatty Liver.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	36.60	0.857	35.64	0.876	32.48	0.743	32.73	0.836
MCG [4]	✓	37.97	0.897	38.07	0.897	38.12	0.898	38.14	0.898
DiffusionMBIR [3]	✓	38.04	0.961	37.95	0.960	37.86	0.960	37.68	0.959
SWORD [7]	✓	43.47	0.973	42.21	0.968	40.76	0.961	38.45	0.948
FBP [1]	×	22.29	0.482	18.10	0.431	16.54	0.395	13.87	0.342
NeRP [6]	×	26.89	0.785	27.27	0.808	26.81	0.784	26.93	0.792
R^2-Gaussian [8]	×	42.58	0.977	42.03	0.974	41.36	0.969	40.80	0.965
GIFT (Ours)	×	**44.46**	**0.987**	**43.96**	**0.986**	**43.47**	**0.985**	**42.54**	**0.983**

Table 13: Benchmark on Gallbladder Stones.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	36.15	0.892	35.59	0.913	32.41	0.797	31.80	0.868
MCG [4]	✓	38.13	0.897	38.47	0.901	38.01	0.897	37.95	0.899
DiffusionMBIR [3]	✓	38.20	0.966	38.22	0.966	38.19	0.967	37.86	0.965
SWORD [7]	✓	43.66	0.974	42.34	0.969	40.56	0.961	37.64	0.943
FBP [1]	×	23.94	0.548	20.27	0.494	17.46	0.445	14.68	0.380
NeRP [6]	×	27.03	0.809	27.12	0.814	26.81	0.799	26.86	0.806
R^2-Gaussian [8]	×	42.58	0.982	41.93	0.979	41.56	0.975	**41.03**	0.972
GIFT (Ours)	×	**43.73**	**0.985**	**42.91**	**0.984**	**42.15**	**0.982**	40.55	**0.977**

Table 14: Benchmark on Hepatic Cyst.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	36.74	0.930	35.52	0.905	31.33	0.791	33.36	0.854
MCG [4]	✓	37.87	0.891	37.91	0.891	37.94	0.891	37.94	0.891
DiffusionMBIR [3]	✓	37.92	0.955	38.02	0.957	37.91	0.956	37.50	0.952
SWORD [7]	✓	42.84	0.973	41.42	0.967	39.81	0.960	37.12	0.946
FBP [1]	×	25.26	0.603	19.94	0.525	17.27	0.475	14.26	0.416
NeRP [6]	×	26.65	0.808	26.57	0.804	26.65	0.808	26.39	0.799
R^2-Gaussian [8]	×	42.05	0.976	41.57	0.972	40.83	0.967	**40.02**	0.966
GIFT (Ours)	×	**42.96**	**0.981**	**42.29**	**0.980**	**41.47**	**0.977**	39.12	**0.971**

Table 15: Benchmark on Elbow Fracture.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	34.41	0.847	34.05	0.789	27.42	0.777	29.82	0.732
MCG [4]	✓	37.20	0.857	37.13	0.858	37.08	0.856	36.80	0.852
DiffusionMBIR [3]	✓	37.06	0.932	36.93	0.930	36.89	0.931	36.75	0.930
SWORD [7]	✓	**42.83**	0.959	38.67	0.917	37.39	0.901	34.71	0.865
FBP [1]	×	26.15	0.459	22.32	0.382	19.93	0.337	16.95	0.279
NeRP [6]	×	28.14	0.826	28.31	0.827	28.06	0.823	28.18	0.835
R^2-Gaussian [8]	×	41.99	0.950	41.37	0.946	40.63	0.942	**40.12**	0.939
GIFT (Ours)	×	42.82	**0.961**	**41.94**	**0.954**	**41.19**	**0.949**	39.97	**0.948**

Table 16: Benchmark on Spinal Fracture.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	23.86	0.866	23.94	0.841	23.92	0.832	23.70	0.810
MCG [4]	✓	38.52	0.913	38.52	0.913	38.48	0.912	38.40	0.911
DiffusionMBIR [3]	✓	39.34	0.973	39.27	0.973	**39.08**	**0.972**	**38.49**	**0.969**
SWORD [7]	✓	40.94	0.946	38.02	0.930	34.68	0.901	28.85	0.834
FBP [1]	×	16.41	0.793	15.20	0.766	14.73	0.741	13.96	0.698
NeRP [6]	×	28.10	0.847	26.24	0.779	27.95	0.840	26.48	0.790
R^2-Gaussian [8]	×	40.51	0.970	39.06	0.961	38.33	0.957	37.98	0.953
GIFT (Ours)	×	**41.23**	**0.981**	**39.70**	**0.977**	38.41	0.971	37.68	0.968

Table 17: Benchmark on Foot Fracture.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	37.53	0.860	35.61	0.783	32.52	0.817	32.46	0.837
MCG [4]	✓	39.40	0.891	39.62	0.895	39.43	0.894	39.45	0.894
DiffusionMBIR [3]	✓	40.45	0.956	40.31	0.955	40.22	0.954	**40.26**	0.957
SWORD [7]	✓	34.92	0.927	36.40	0.905	31.95	0.866	28.33	0.783
FBP [1]	×	23.45	0.235	18.80	0.181	17.23	0.160	14.46	0.132
NeRP [6]	×	30.69	0.921	30.82	0.926	30.76	0.932	30.56	0.927
R^2-Gaussian [8]	×	39.97	0.960	39.12	0.956	38.46	0.951	37.90	0.949
GIFT (Ours)	×	**41.81**	**0.981**	**41.21**	**0.980**	**40.51**	**0.974**	39.60	**0.977**

Table 18: Benchmark on Wrist Fracture.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	36.61	0.810	34.73	0.825	31.73	0.870	30.78	0.744
MCG [4]	✓	37.14	0.887	37.53	0.889	37.65	0.890	37.64	0.889
DiffusionMBIR [3]	✓	36.91	0.953	36.94	0.954	36.73	0.952	36.31	0.950
SWORD [7]	✓	36.74	0.903	33.91	0.874	31.57	0.832	28.91	0.766
FBP [1]	×	21.35	0.231	17.95	0.197	15.69	0.174	12.96	0.146
NeRP [6]	×	29.55	0.893	28.77	0.891	29.62	0.897	29.56	0.893
R^2-Gaussian [8]	×	38.42	0.974	38.04	0.973	37.66	0.970	36.98	0.967
GIFT (Ours)	×	**40.28**	**0.984**	**39.71**	**0.983**	**38.89**	**0.976**	**37.78**	**0.973**

Table 19: Benchmark on Subarachnoid Hemorrhage.

Method	Pretrain	180-view		120-view		90-view		60-view	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RED-CNN [2]	✓	29.52	0.874	29.65	0.877	29.55	0.877	29.42	0.863
MCG [4]	✓	38.78	0.908	38.87	0.909	38.81	0.908	38.79	0.908
DiffusionMBIR [3]	✓	39.46	0.975	39.38	0.975	39.20	0.974	38.70	0.973
SWORD [7]	✓	42.54	0.965	39.60	0.955	36.71	0.938	31.84	0.895
FBP [1]	×	21.31	0.440	20.84	0.423	19.22	0.404	17.93	0.361
NeRP [6]	×	23.72	0.760	23.34	0.760	23.84	0.800	24.04	0.791
R^2-Gaussian [8]	×	41.89	0.973	41.11	0.968	40.35	0.964	39.77	0.956
GIFT (Ours)	×	**43.29**	**0.993**	**42.74**	**0.993**	**41.98**	**0.992**	**41.02**	**0.992**

F Datasheet

Motivation

Q1. For what purpose was the dataset created?
Answer: Mainly for two purposes: (1) to provide a comprehensive dataset of multiple anatomys and lesions for medical image reconstruction research instead of focusing on a single organ or disease in the existing datasets; (2) to evaluate the performance of different reconstruction methods on a diverse dataset to ensure the robustness of the methods.

Q2. Who created this dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)?
Answer: The dataset was created by the BCMI Lab at Shanghai Jiao Tong University and the Radiology Department at Suzhou Xiangcheng People's Hospital.

Q3. Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number.
Answer: The dataset was created without any specific funding or associated grants.

Q4. Any other comments?
Answer: Examined by three experienced radiologists during the data collection process, we ensure the distribution of the dataset is consistent with the real-world distribution of medical images and thus suitable for evaluating the performance of medical image reconstruction methods in clinical practice.

Composition

Q5. What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)?
Answer: Each patient can be viewed as an instance in the dataset, and each instance contains a 3D volume of CT scans. Each scan is a DICOM file that represents a 2D image slice of the 3D volume.

Q6. How many instances are there in total (of each type, if appropriate)?
Answer: There are 135 CT instances in total, with 65,575 CT slices.

Q7. Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set?
Answer: This dataset contains all possible instances of anatomys and lesions that are commonly seen in clinical practice and have been carefully selected by three experienced radiologists to ensure the diversity and representativeness of the dataset.

Q8. What data does each instance consist of? "Raw" data (e.g., unprocessed text or images) or features?
Answer: Yes, we provide both the unprocessed DICOM files and the reconstructed 3D volumes for each instance. However, the private information of patients in the raw data has been removed to protect privacy.

Q9. Is there a label or target associated with each instance?
Answer: Yes, each instance has been labeled with the specific anatomy and lesion it represents.

Q10. Is any information missing from individual instances?
Answer: No.

Q11. Are relationships between individual instances made explicit (e.g., users' movie ratings, social network links)?
Answer: Yes, each scan is associated with a specific patient and the corresponding anatomy and lesion.

Q12. Are there recommended data splits (e.g., training, development/validation, testing)?
Answer: Yes, we provide a training set and a testing set for each type of scan. The training set is used to train the reconstruction models (for those training-based methods), and the testing set is used to evaluate the performance of the models.

Q13. Are there any errors, sources of noise, or redundancies in the dataset?
Answer: No. We aimed to provide a high-quality dataset for medical image reconstruction research, and the dataset has been carefully examined to ensure the quality and reliability of the data.

Q14. Is the dataset self-contained, or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)?
Answer: This dataset is self-contained since all the data was collected from Suzhou Xiangcheng People's Hospital.

Q15. Does the dataset contain data that might be considered confidential (e.g., data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals non-public communications)?
Answer: No. All the confidential information of patients has been removed to protect privacy, i.e., one cannot identify the patients from the dataset.

Q16. Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety?
Answer: No.

Q17 Does the dataset identify any subpopulations (e.g., by age, gender)?
Answer: Yes, the dataset contains scans from patients and thus the subpopulations can be recognized. For example, the pelvic structures of males and females are different. However, these data are essential for medical image reconstruction research and do not contain any private information that can be used to identify individuals.

Q18 Is it possible to identify individuals (i.e., one or more natural persons), either directly or indirectly (i.e., in combination with other data) from the dataset?
Answer: No. The private information of patients has been removed to protect privacy, and the dataset cannot be used to identify the patients.

Q19 Does the dataset contain data that might be considered sensitive in any way (e.g., data that reveals race or ethnic origins, sexual orientations, religious beliefs, political opinions or union memberships, or locations; financial or health data; biometric or genetic data; forms of government identification, such as social security numbers; criminal history)?
Answer: No, and the dataset has been approved by the ethics committee of Suzhou Xiangcheng People's Hospital.

Q20 Any other comments?
Answer: This dataset is intended for research purposes only. All DICOM data has been anonymized to protect patient privacy and comply with the Helsinki Declaration.

Collection Process

Q21. How was the data associated with each instance acquired?
Answer: First, the patients underwent CT scans at Suzhou Xiangcheng People's Hospital. Then, the radiologists examined the scans and selected the representative scans for each anatomy and lesion. Finally, the selected scans were collected and stored in the dataset.
Q22. What mechanisms or procedures were used to collect the data (e.g., hardware apparatus or sensor, manual human curation, software program, software API)?
Answer: The data was collected using CT scanners at Suzhou Xiangcheng People's Hospital. The DICOM files were then extracted from the scanners and stored in the dataset.
Q23. If the dataset is a sample from a larger set, what was the sampling strategy?
Answer: This dataset is newly collected and does not sample from a larger set. However, during the data collection process, the radiologists selected the representative scans for each anatomy and lesion.
Q24. Who was involved in data collection process (e.g., students, crowd-workers, contractors) and how were they compensated (e.g., how much were crowd-workers paid)?
Answer: The collection process of the raw data was conducted by the radiologists at Suzhou Xiangcheng People's Hospital, and then the data was processed and stored by the BCMI Lab at Shanghai Jiao Tong University. No compensation was involved in the data collection process.
Q25. Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (e.g., recent crawl of old news articles)?
Answer: The data was collected over a period of 2 months, from April 2024 to June 2024. The creation timeframe of the data associated with the instances matches the data collection timeframe.
Q26. Were any ethical review processes conducted (e.g., by an institutional review board)?
Answer: Yes, this dataset has been reviewed and approved by the ethics committee of Suzhou Xiangcheng People's Hospital. The approval number is 2024-KY-03.
Q27. Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (e.g., websites)?
Answer: No. All data was collected directly from the patients who underwent CT scans at Suzhou Xiangcheng People's Hospital.
Q28 Were the individuals in question notified about the data collection?
Answer: This research is a retrospective study, for which a Waiver of Informed Consent Application has been signed and approved by the Ethics Committee.
Q29 Did the individuals in question consent to the collection and use of their data?
Answer: The same as the answer to the previous question.
Q30 If consent was obtained, were the consenting individuals provided with a mechanism to revoke their consent in the future or for certain uses?
Answer: N/A; This dataset is intended for research purposes only, and should not be used for any other purposes.
Q31 Has an analysis of the potential impact of the dataset and its use on data subjects (e.g., a data protection impact analysis) been conducted?
Answer: We have anonymized the data to protect patient privacy and comply with the Helsinki Declaration.
Q32 Any other comments?
Answer: This dataset has been collected in compliance with the Helsinki Declaration and the regulations of the ethics committee of Suzhou Xiangcheng People's Hospital.

Preprocessing, Cleaning, and/or Labeling

Q33. Was any preprocessing/cleaning/labeling of the data done (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)?
Answer: Yes, and we provide both the raw DICOM files and the processed image files in the dataset. In the preprocessing step, the raw DICOM files were converted to 2D image slices following the standard practice in SimpleITK.
Q34. Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (e.g., to support unanticipated future uses)?
Answer: Yes, both are provided to facilitate future research and unanticipated uses.
Q35. Is the software used to preprocess/clean/label the instances available?
Answer: Yes, all the programs are open-source and free to use.
Q36. Any other comments?
Answer: The raw data is stored in the DICOM format, and the processed image files are stored in the PNG format. Notably, the raw data is unique but there can be multiple ways to preprocess the data. For reproducibility and future research, we provide the processed image files in the dataset.

Uses

Q37. Has the dataset been used for any tasks already?
Answer: No, this dataset is novel and has not been used for any tasks yet.
Q38. Is there a repository that links to any or all papers or systems that use the dataset?
Answer: No, this dataset is novel and has not been used in any papers or systems yet.
Q39. What (other) tasks could the dataset be used for?
Answer: Besides medical image reconstruction, this dataset can also be used for the classification task of different lesions because each instance is labeled with the specific anatomy and lesion it represents.
Q40. Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses?
Answer: No, since both the raw data and the processed image files are provided in the dataset, the dataset will be easy to use for future research.
Q41. Are there any tasks for which the dataset should not be used?

Answer: No. However, this dataset is intended for research purposes only.

Q42. Any other comments?

Answer: The license of the dataset is CC BY-NC 4.0, which means the dataset can be used for non-commercial purposes with proper attribution.

Distribution

Q43. Will the dataset be distributed to third parties outside of the entity (e.g., company, institution, organization) on behalf of which the dataset was created?

Answer: Yes, our dataset will be made publicly available for research purposes.

Q44. How will the dataset will be distributed (e.g., tarball on website, API, GitHub)

Answer: The dataset is hosted at Huggingface, and may also be uploaded to other platforms later. The latest news will be updated on our https://more-med.github.io/webpage.

Q45. When will the dataset be distributed?

Answer: It will be distributed after the publication of the paper.

Q46. Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)?

Answer: Yes, the dataset will be distributed under the CC BY-NC 4.0 license.

Q47 Have any third parties imposed IP-based or other restrictions on the data associated with the instances?

Answer: Since this dataset is intended for research purposes only, there are no restrictions on the data associated with the instances.

Q48. Do any export controls or other regulatory restrictions apply to the dataset or to individual instances?

Answer: No.

Q49. Any other comments?

Answer: The dataset is intended for research purposes only and should not be used for any other purposes.

Maintenance

Q50. Who will be supporting/hosting/maintaining the dataset?

Answer: Currently, this dataset is hosted on the Hugging Face Datasets platform. To avoid data loss, we will also upload the dataset to other platforms such as Google Drive. The latest news will be updated in our webpage.

Q51. How can the owner/curator/manager of the dataset be contacted (e.g., email address)?

Answer: Please contact us via email or the issue page on Hugging Face or GitHub.

Q52. Is there an erratum?

Answer: The erratum will be maintained on the webpage of this dataset.

Q53. Will the dataset be updated (e.g., to correct labeling errors, add new instances, delete instances)?

Answer: Yes, we will update the dataset if there are any labeling errors or new instances to be added.

Q54. If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (e.g., were individuals in question told that their data would be retained for a fixed period of time and then deleted)?

Answer: The dataset is anonymized and does not contain any private information of patients. As long as the dataset is used for research purposes only, there are no limits on the retention of the data.

Q55. Will older versions of the dataset continue to be supported/hosted/maintained?

Answer: Yes, we will maintain the older versions of the dataset to ensure reproducibility and traceability of the research results.

Q56. If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so?

Answer: If so, they should also follow the CC BY-NC 4.0 license and provide proper attribution.

Q57. Any other comments?

Answer: We will keep maintaining the dataset and provide the latest information on the webpage.

References

[1] Ronald Newbold Bracewell and ACf Riddle. 1967. Inversion of fan-beam scans in radio astronomy. *Astrophysical Journal, vol. 150, p. 427* 150 (1967), 427.

[2] Hu Chen, Yi Zhang, Mannudeep K Kalra, Feng Lin, Yang Chen, Peixi Liao, Jiliu Zhou, and Ge Wang. 2017. Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE transactions on medical imaging* 36, 12 (2017), 2524–2535.

[3] Hyungjin Chung, Dohoon Ryu, Michael T Mccann, Marc L Klasky, and Jong Chul Ye. 2023. Solving 3d inverse problems using pre-trained 2d diffusion models. IEEE. In *CVF Conference on Computer Vision and Pattern Recognition*, Vol. 1. 6.

[4] Hyungjin Chung, Byeongsu Sim, Dohoon Ryu, and Jong Chul Ye. 2022. Improving diffusion models for inverse problems using manifold constraints. *Advances in Neural Information Processing Systems* 35 (2022), 25683–25696.

[5] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 2023. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics* 42, 4 (2023), 1–14.

[6] Liyue Shen, John Pauly, and Lei Xing. 2022. NeRP: implicit neural representation learning with prior embedding for sparsely sampled image reconstruction. *IEEE Transactions on Neural Networks and Learning Systems* (2022).

[7] Kai Xu, Shiyu Lu, Bin Huang, Weiwen Wu, and Qiegen Liu. 2024. Stage-by-stage Wavelet Optimization Refinement Diffusion Model for Sparse-View CT Reconstruction. *IEEE Transactions on Medical Imaging* (2024).

[8] Ruyi Zha, Tao Jun Lin, Yuanhao Cai, Jiwen Cao, Yanhao Zhang, and Hongdong Li. 2024. R^2-Gaussian: Rectifying Radiative Gaussian Splatting for Tomographic Reconstruction. In *Advances in Neural Information Processing Systems (NeurIPS)*.