DLMA - Final Presentation

Wavelet and Diffusion Models

Presenter: Leonhard Zirus
Director: Prof. Dr. Nassir Navab
Supervisor: Mohammad Farid Azampour
Date: 20nd June of 2024

Technische Universität München
JOHNS HOPKINS WHITING SCHOOL of ENGINEERING

Outline

MOTIVATION

Motivation

Optimal network performance in computer vision requires **extensive** datasets of **clean**, **high-quality** images ⇒ especially in medical imaging

Challenge for medical context:

→ High cost and resource intensity in obtaining detailed CT or MRI scans

→ often compromised by noise and artifacts.

Motivation

Optimal network performance in computer vision requires **extensive** datasets of **clean**, **high-quality** images ⇒ especially in medical imaging

Challenge for medical context:

→ High cost and resource intensity in obtaining detailed CT or MRI scans

→ often compromised by noise and artifacts.

Generative AI Models: can produce high-quality synthetic images supplementing existing datasets

⇒ improved robustness and accuracy

⇒ **But:** → requires large datasets and computational intensity for training

→ synthetic images often lacking intricate details

⇒ **Wavelet Transformations** ?

Background

Generative Models

designed to generate new data samples resembling a given dataset by learning its underlying distribution.

⇒ used for: generating images, text, and audio.

Common Implementations for medical imaging:

→ Variational Autoencoders (VAEs)

→ Generative Adversarial Networks (GANs)

Generative Models

designed to generate new data samples resembling a given dataset by learning its underlying distribution.

⇒ used for: generating images, text, and audio.

Common Implementations for medical imaging:

→ Variational Autoencoders (VAEs)

→ Generative Adversarial Networks (GANs)

Score-Based Generative Models (SGMs)

SGMs estimate the score function of data distribution to generate new samples by following gradients

→ Denoising Score Matching (DSM)

→ Stochastic Differential Equations (SDE)

Diffusion Models

use a two-stage process:

→ forward diffusion adds Gaussian noise

→ reverse diffusion recovers the original data

⇒ known for high-quality but slow sample generation.

Wavelet Transform

- Mathematical transform to decompose data into frequency components
- Captures BOTH time and frequency localization (≠ Fourier)
- > analysis of signals/data with non-stationary properties
- > multi-resolution analysis (at various levels of detail)

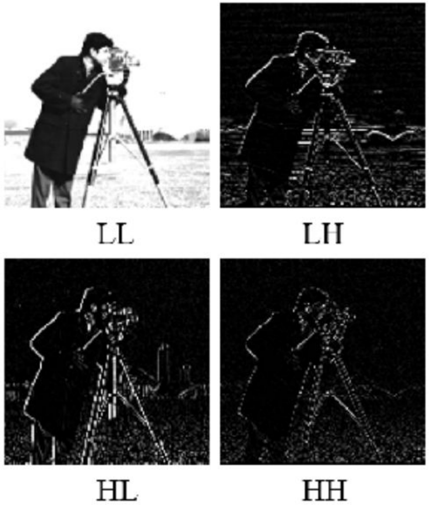
Used for: signal processing, data compression, data smoothing, image denoising, etc.

[7] Shin, Y. H., Park, M. J., Lee, O. Y., & Kim, J. O. (2020). Deep orthogonal transform feature for image denoising. IEEE Access, 8, 66898-66909.

Wavelet Transform

- Mathematical transform to decompose data into frequency components
- Captures BOTH time and frequency localization (≠ Fourier)
- > analysis of signals/data with non-stationary properties
- > multi-resolution analysis (at various levels of detail)

Used for: signal processing, data compression, data smoothing, image denoising, etc.



discrete wavelet transform [9]

LL (Low-Low): Approximation of original image

LH (Low-High): Horizontal edge details

HL (High-Low): Vertical edge details

HH (High-High): diagonal details and fine textures in the image

[9] Shin, Y. H., Park, M. J., Lee, O. Y., & Kim, J. O. (2020). Deep orthogonal transform feature for image denoising. IEEE Access, 8, 66898-66909.

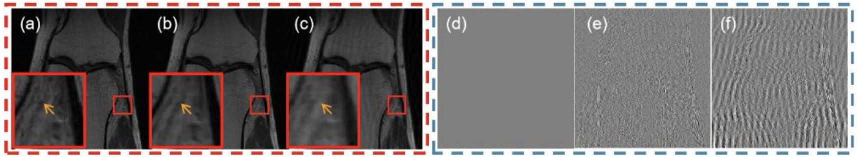
Advancements in Medical Imaging

Wavelet-Improved Score-Based Generative Model for Medical Imaging

Technische Universität München + Johns Hopkins Whiting School of Engineering

SGMs in Medical Imaging

- SGMs need high-quality datasets for best performance
- > medical imaging modalities are inherently noisy and ridden by artifacts
- >> low-dose CT, sparse-view CT, and fast MRI



(a) ground truth (b), (c) reconstruction result of SGM trained with clean and noisy images (d), (e), (f) differences between reconstructed images and the ground truth respectively

[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.

Contributions

Adaptive wavelet sub-network
⇒ Novel and effective denoising mechanism
⇒ enhancing the robustness of SGM → higher quality reconstruction results

[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.

Contributions

Adaptive wavelet sub-network
⇒ Novel and effective denoising mechanism
⇒ enhancing the robustness of SGM → higher quality reconstruction results

Unified Framework adaptive wavelet integrated with SGM
⇒ wavelet provides cleaner images → improve SGM reconstruction
⇒ SGMs provides better reconstruction results → refines wavelet sub-network

[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.

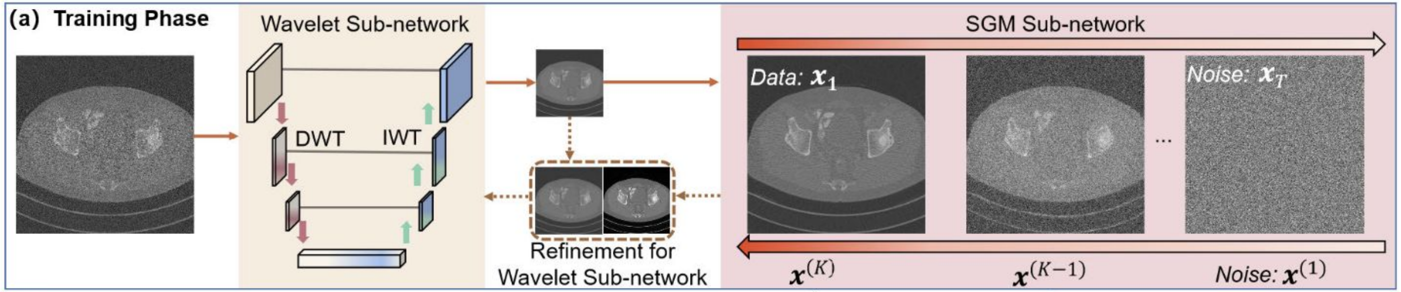
Contributions

Adaptive wavelet sub-network
⇒ Novel and effective denoising mechanism
⇒ enhancing the robustness of SGM → higher quality reconstruction results

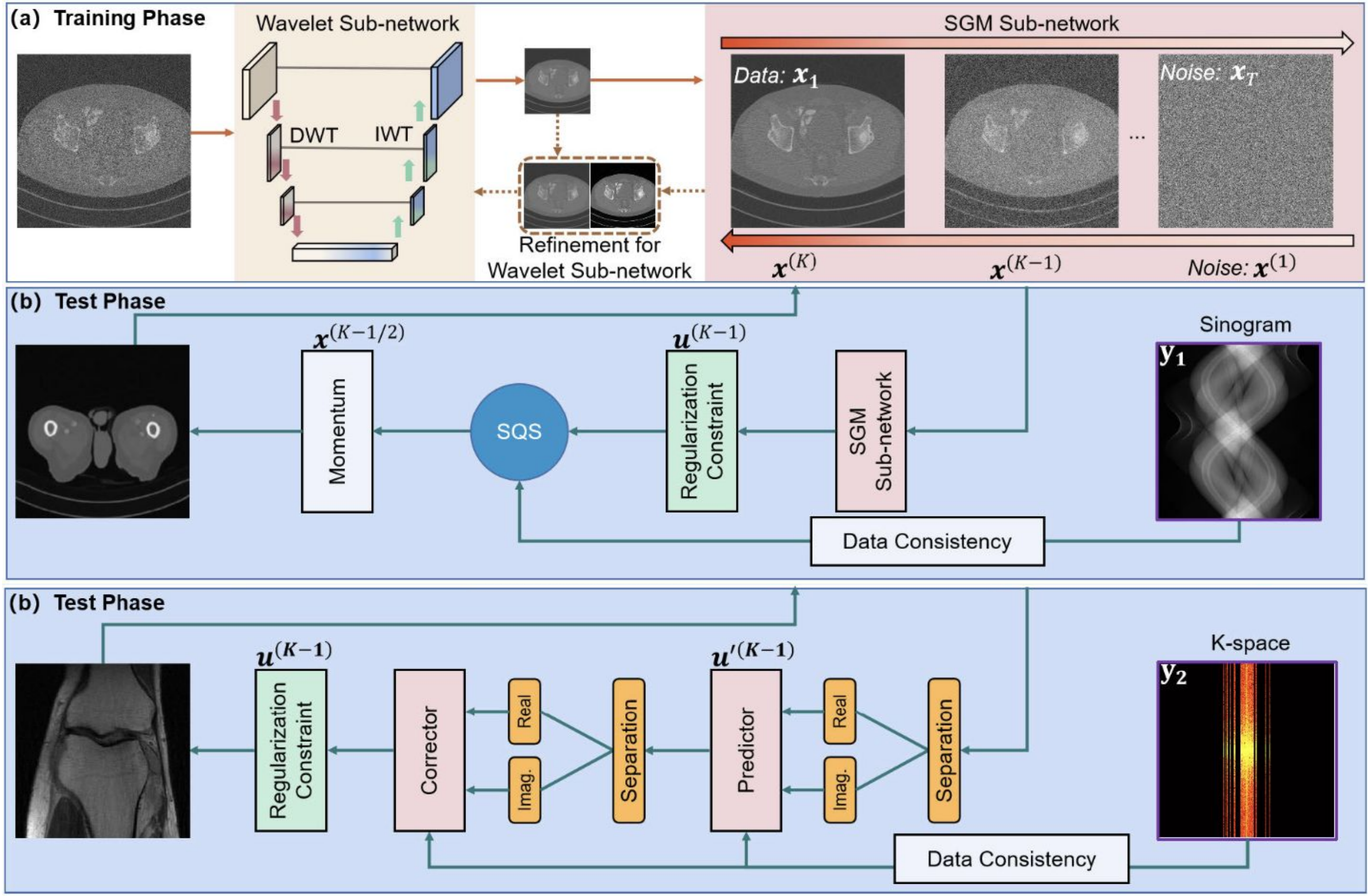
Unified Framework adaptive wavelet integrated with SGM
⇒ wavelet provides cleaner images → improve SGM reconstruction
⇒ SGMs provides better reconstruction results → refines wavelet sub-network

Multi-perspective regularization
⇒ integrate prior knowledge in diffusion process: regularized sparsity, data consistency
⇒ handle under-sampled and noisy input → accelerate search process

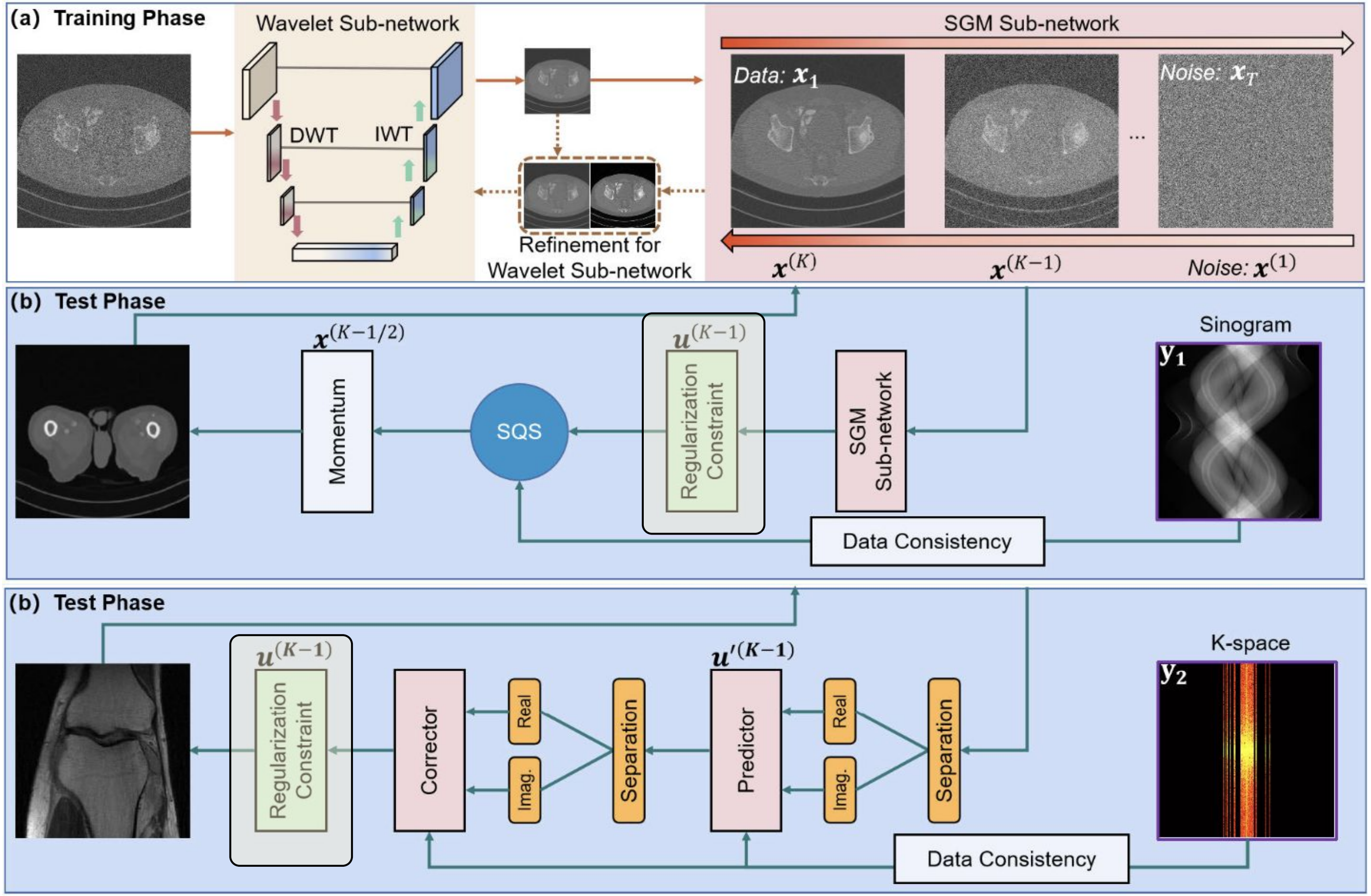
[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.



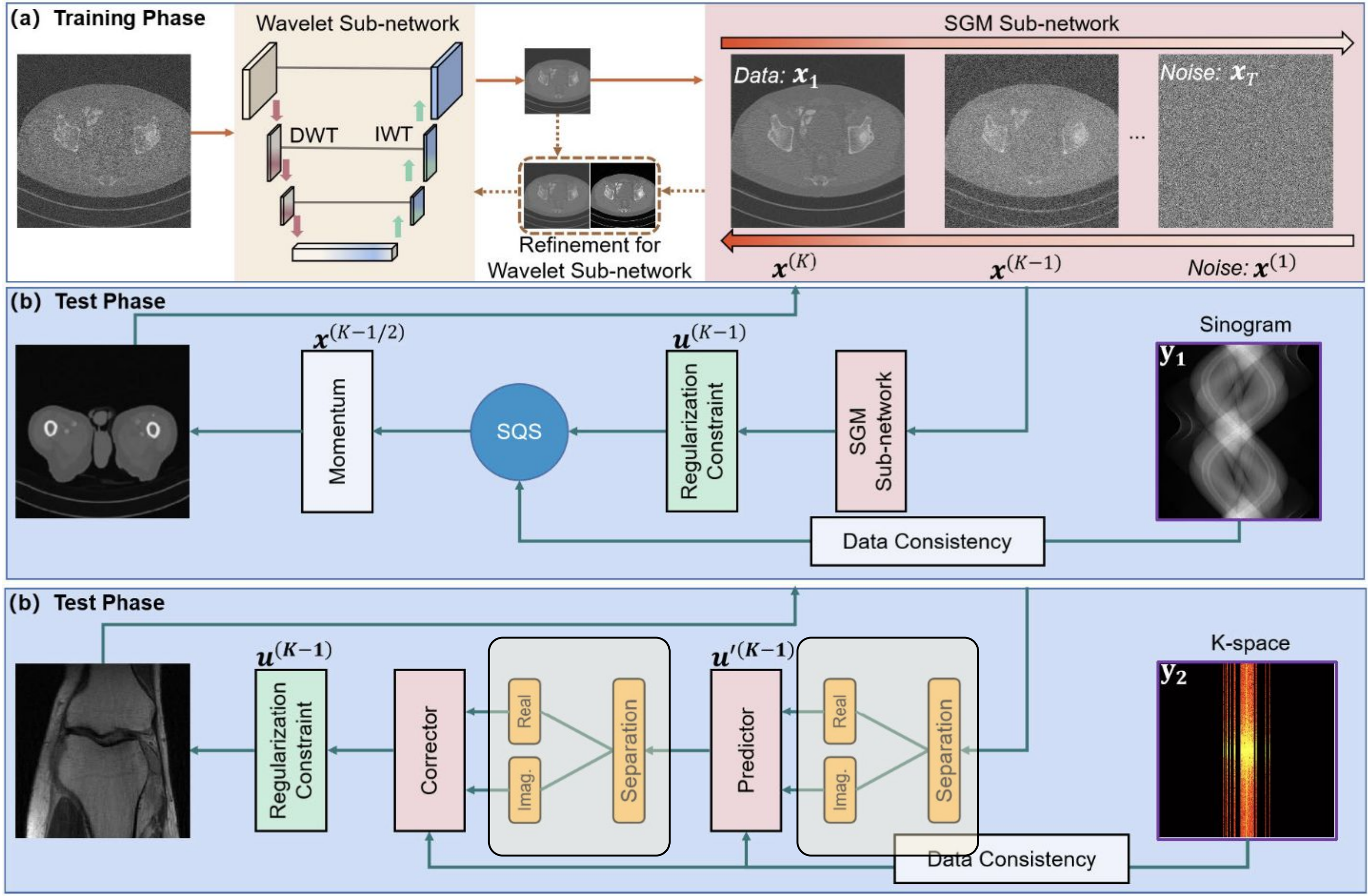
[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.



[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.



[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.



[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.

Results

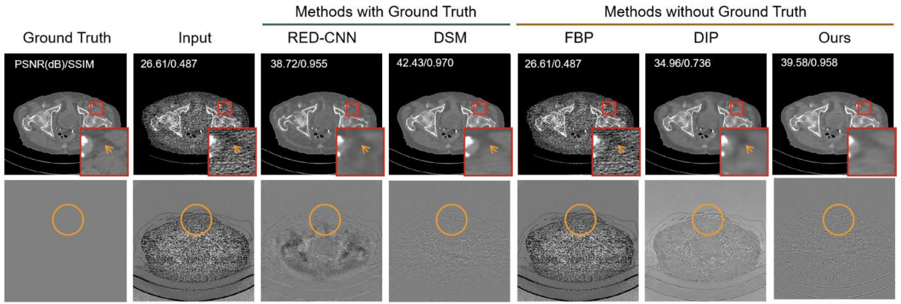


Fig. 7. Low-dose CT reconstruction results. From left to right Ground Truth, Input, RED-CNN, DSM, FBP, DIP, and our method. The 1^{st} row displays the reconstructed images, while the 2^{nd} row shows the difference images relative to the ground truth. Display windows [−50 1200] HU are used.

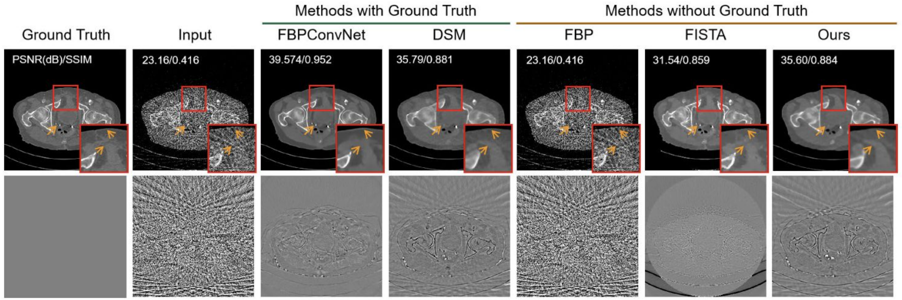


Fig. 8. The reconstruction results of 60 views from AAPM CT datasets. The columns from left to right: the Ground Truth, Input, FBPConvNet, DSM, FBP, FISTA, and our method. The 1^{st} row displays the reconstructed images, while the 2^{nd} row shows the difference images relative to the ground truth. Display windows [100 1300] HU are used.

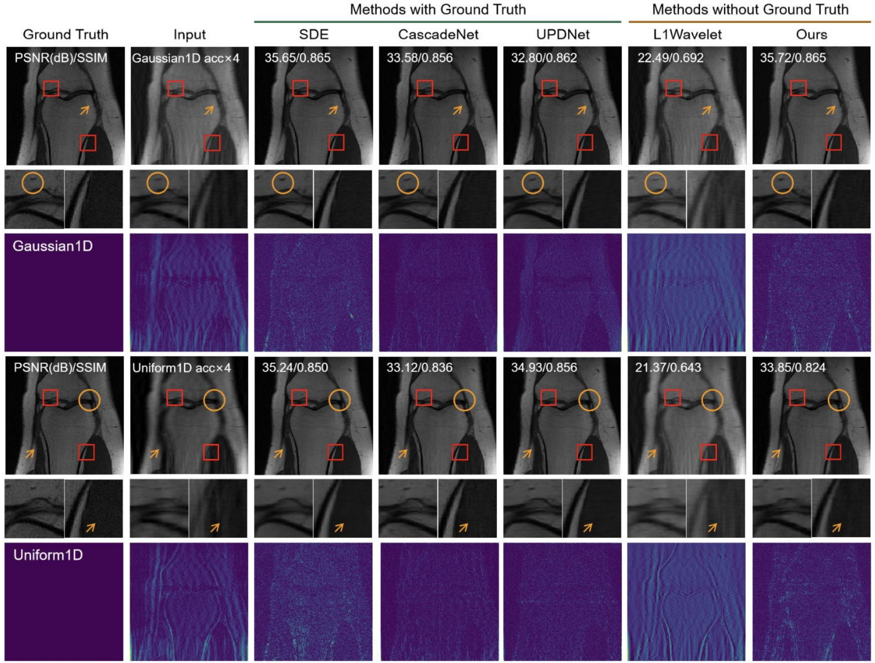


Fig. 9. ×4 acceleration MRI reconstruction results with Gaussian 1D and Uniform 1D masks. From left to right Ground Truth, Input, L1Wavelet, SDE, CascadeNet, UPDNet, and our method. The 1^{st} to 3^{rd} rows display one case of the image, two enlarged ROIs (knee joint ROI and tissue ROI), and the difference images in comparison to the ground truth, respectively. Similarly, the 4^{th} to 6^{th} rows display another image case, enlarged ROIs, and the different images in comparison to the ground truth, respectively.

[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.

Results - metric evaluation

Peak Signal-to-Noise Ratio (PSNR): indicating overall image fidelity and quality

→ ratio between the maximum possible power of an image signal and the power of corrupting noise

Structural Similarity Index Measure (SSIM): similarity between two images

→ based on luminance, contrast, and structure

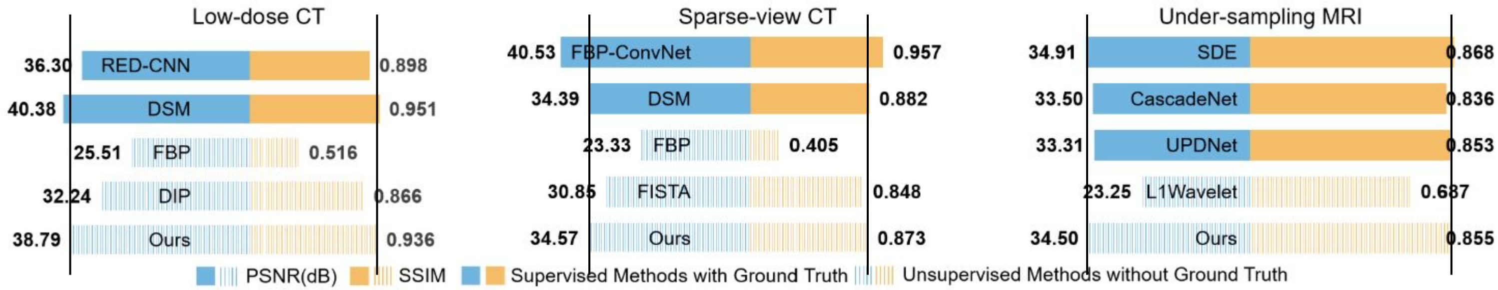


Fig. 10. Statistical results of different methods in terms of PSNR and SSIM in three reconstruction tasks, including low-dose CT reconstruction, sparse-view CT reconstruction, and under-sampled MRI reconstruction with Gaussian 1D mask.

[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.

Evaluation

Pro

→ robustness to noise

→ high reconstruction quality without ground truth images

→ preserves fine details and structural information

→ effective across various imaging modalities (Ultrasound ?)

⇒ beneficial unsupervised learning capability in clinical settings

Limitations

→ substantial amounts of computation time due to SGMs

⇒ increased computational complexity and time due to wavelet

⇒ limitation for real-time applications

→ may struggle to recover very fine details under extreme imaging conditions

⇒ requires further enhancements through integration with other reconstruction techniques

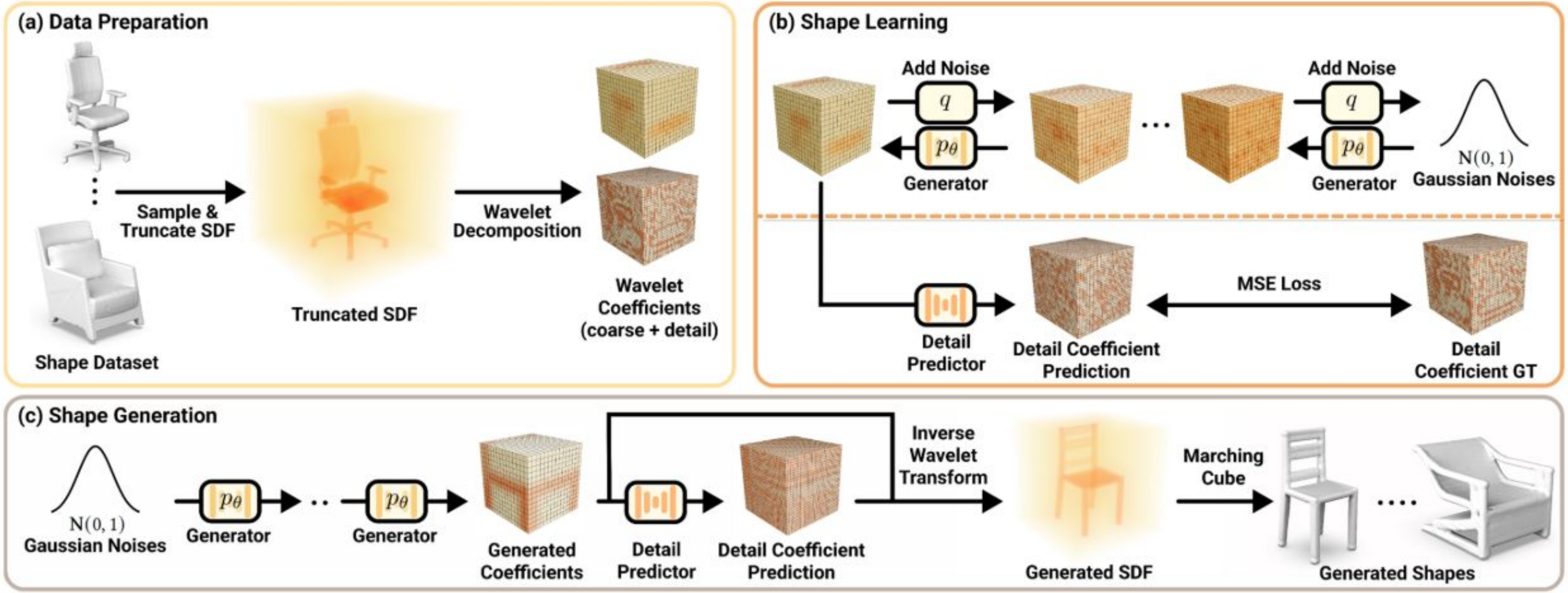
[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.

3D Shape Generation

Neural Wavelet-domain Diffusion for 3D Shape Generation

Technische Universität München + Johns Hopkins Whiting School of Engineering

Overview



(a) Data Preparation
Sample & Truncate SDF
Wavelet Decomposition
Truncated SDF
Wavelet Coefficients (coarse + detail)
Shape Dataset
(b) Shape Learning
Add Noise
q
p_θ
Generator
Add Noise
q
p_θ
Generator
N(0, 1)
Gaussian Noises
Detail Predictor
Detail Coefficient Prediction
MSE Loss
Detail Coefficient GT
(c) Shape Generation
N(0, 1)
Gaussian Noises
p_θ
Generator
p_θ
Generator
Generated Coefficients
Detail Predictor
Detail Coefficient Prediction
Inverse Wavelet Transform
Generated SDF
Marching Cube
Generated Shapes

Data Preparation

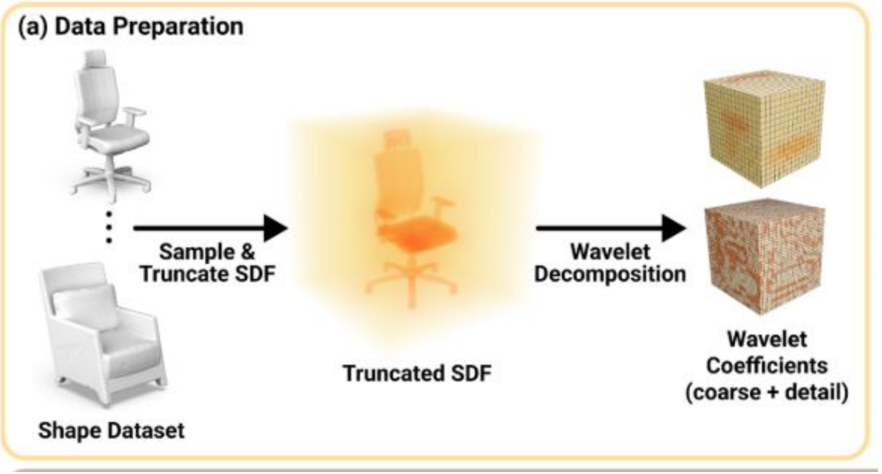
Sample a signed distance field (SDF)

→ truncate its distance values to avoid redundant information

⇒ TSDF (truncated SDF)

> Reduce the shape representation redundancy

> Focus the shape learning process on the shape's structures and fine details



Transform TSDF to the wavelet domain

[5] Hui, K. H., Li, R., Hu, J., & Fu, C. W. (2022, November). Neural wavelet-domain diffusion for 3d shape generation. In SIGGRAPH Asia 2022 Conference Papers (pp. 1-9).

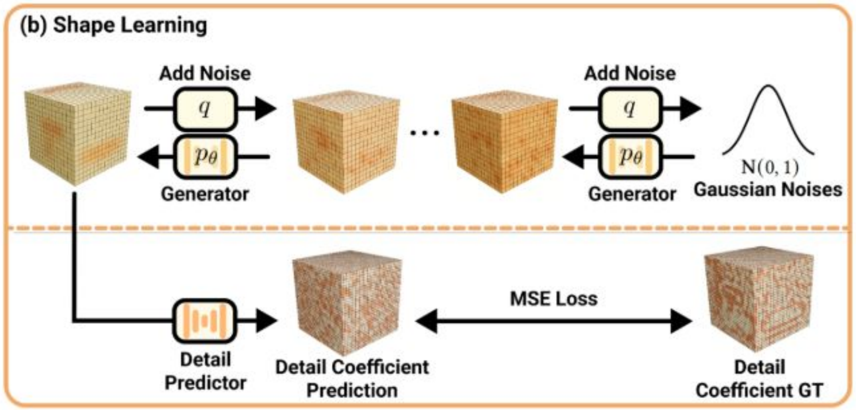
Shape Learning

Generator Network → generation of coarse image through diffusion

⇒ denoising diffusion probabilistic model

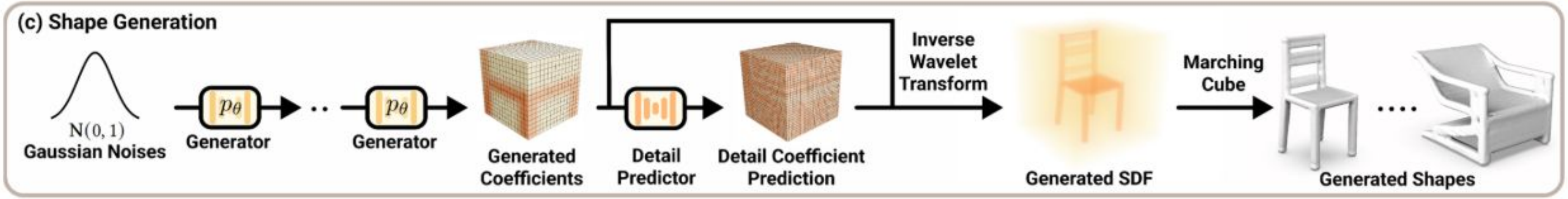
Detail Predictor Network → learning to predict the details for the generated shapes

⇒ adding fine details → more realistic and intricate 3D Shapes



[5] Hui, K. H., Li, R., Hu, J., & Fu, C. W. (2022, November). Neural wavelet-domain diffusion for 3d shape generation. In SIGGRAPH Asia 2022 Conference Papers (pp. 1-9).

Shape Generation (Inference)



→ trained generator network creates a coarse coefficient volume from a random noise sample

→ detail predictor then adds fine details to this coarse volume

⇒ inverse wavelet transform followed by the marching cube algorithm → reconstruct the final 3D shape

[5] Hui, K. H., Li, R., Hu, J., & Fu, C. W. (2022, November). Neural wavelet-domain diffusion for 3d shape generation. In SIGGRAPH Asia 2022 Conference Papers (pp. 1-9).

Results

Method	Chair						Airplane					
	COV		MMD		1-NNA		COV		MMD		1-NNA	
	CD	EMD	CD	EMD	CD	EMD	CD	EMD	CD	EMD	CD	EMD
IM-GAN [Chen and Zhang 2019]	56.49	54.50	11.79	14.52	61.98	63.45	61.55	62.79	3.320	8.371	76.21	76.08
Voxel-GAN [Kleineberg et al. 2020]	43.95	39.45	15.18	17.32	80.27	81.16	38.44	39.18	5.937	11.69	93.14	92.77
Point-Diff [Luo and Hu 2021]	51.47	**55.97**	12.79	16.12	61.76	63.72	60.19	62.30	3.543	9.519	74.60	72.31
SPAGHETTI [Hertz et al. 2022]	49.19	51.92	14.90	15.90	70.72	68.95	58.34	58.38	4.062	8.887	78.24	77.01
Ours	**58.19**	55.46	**11.70**	**14.31**	**61.47**	**61.62**	**64.78**	**64.40**	**3.230**	**7.756**	**71.69**	**66.74**

coverage (COV): generated shapes coverage of the shapes in the given 3D repository

minimum matching distance (MMD): fidelity of the generated shapes

1-NN classifier accuracy (1-NNA): how well a classifier differentiates the generated shapes from those in the repository

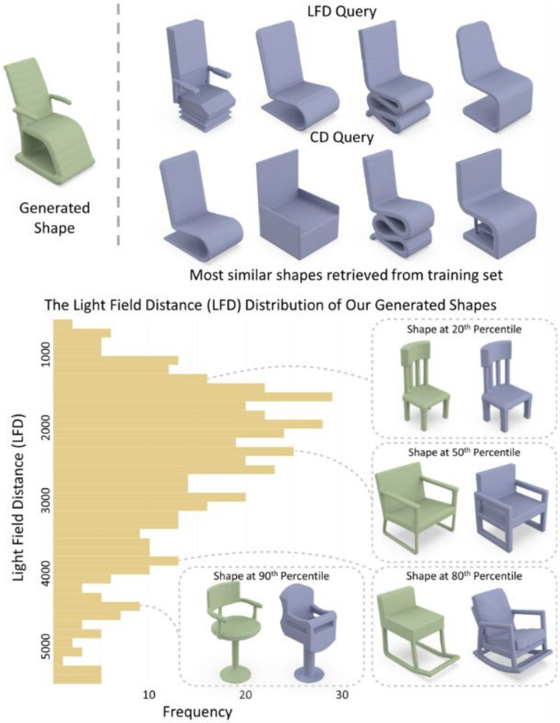
⇒ Chamfer Distance (CD) & Earth Mover's Distance (EMD)

[5] Hui, K. H., Li, R., Hu, J., & Fu, C. W. (2022, November). Neural wavelet-domain diffusion for 3d shape generation. In SIGGRAPH Asia 2022 Conference Papers (pp. 1-9).

Results - Qualitative

[5] Hui, K. H., Li, R., Hu, J., & Fu, C. W. (2022, November). Neural wavelet-domain diffusion for 3d shape generation. In SIGGRAPH Asia 2022 Conference Papers (pp. 1-9).

Results - Novelty



[5] Hui, K. H., Li, R., Hu, J., & Fu, C. W. (2022, November). Neural wavelet-domain diffusion for 3d shape generation. In SIGGRAPH Asia 2022 Conference Papers (pp. 1-9).

Evaluation

Pro

→ High fidelity in generated shapes

→ Detailed and realistic structures

→ Effective coverage of a broad range of shapes

→ Clean surfaces free from artifacts

→ Realistic generation of novel shapes

Limitations

→ high computational cost due to iterative process

⇒ Long computing times despite subsampling

[5] Hui, K. H., Li, R., Hu, J., & Fu, C. W. (2022, November). Neural wavelet-domain diffusion for 3d shape generation. In SIGGRAPH Asia 2022 Conference Papers (pp. 1-9).

Optimization of Generative Modeling

Wavelet Score-Based Generative Modeling

Contributions

Analysis of the computational effort for new data generation

Wavelet Score-based Generative Model (WSGM)

→ improve computation efficiency of SGMs

⇒ time complexity growing INDEPENDENTLY with image size

[4] Guth, F., Coste, S., De Bortoli, V., & Mallat, S. (2022). Wavelet score-based generative modeling. Advances in Neural Information Processing Systems, 35, 478-491.

Contributions

Analysis of the computational effort for new data generation

Wavelet Score-based Generative Model (WSGM)

→ improve computation efficiency of SGMs

⇒ time complexity growing INDEPENDENTLY with image size

Theorems for controlling errors of time discretizations of SGMs

⇒ proving accelerations obtained by scale separation with wavelets

⇒ empirically verified by showing that WSGM provides an acceleration for the synthesis of physical processes at phase transition and natural image datasets.

[4] Guth, F., Coste, S., De Bortoli, V., & Mallat, S. (2022). Wavelet score-based generative modeling. Advances in Neural Information Processing Systems, 35, 478-491.

SGMs: discretization and score regularity

Theorem 1: for gaussian distribution $N(0,\Sigma)$

→ provides an upper bound on the Kullback-Leibler (KL) divergence error ε between a Gaussian distribution and its discretized version

⇒ dependant on the number of time steps N to reach ε as a function of the condition number κ of covariance matrix Σ.

⇒ indicates that the number of time steps **should increase** with the condition number of the covariance matrix → in typical cases the number of time steps increases with the image size

[4] Guth, F., Coste, S., De Bortoli, V., & Mallat, S. (2022). Wavelet score-based generative modeling. Advances in Neural Information Processing Systems, 35, 478-491.

SGMs: discretization and score regularity

Theorem 1: for gaussian distribution N(0,Σ)

→ provides an upper bound on the Kullback-Leibler (KL) divergence error ε between a Gaussian distribution and its discretized version

⇒ dependant on the number of time steps N to reach ε as a function of the condition number κ of covariance matrix Σ.

⇒ indicates that the number of time steps **should increase** with the condition number of the covariance matrix → in typical cases the number of time steps increases with the image size

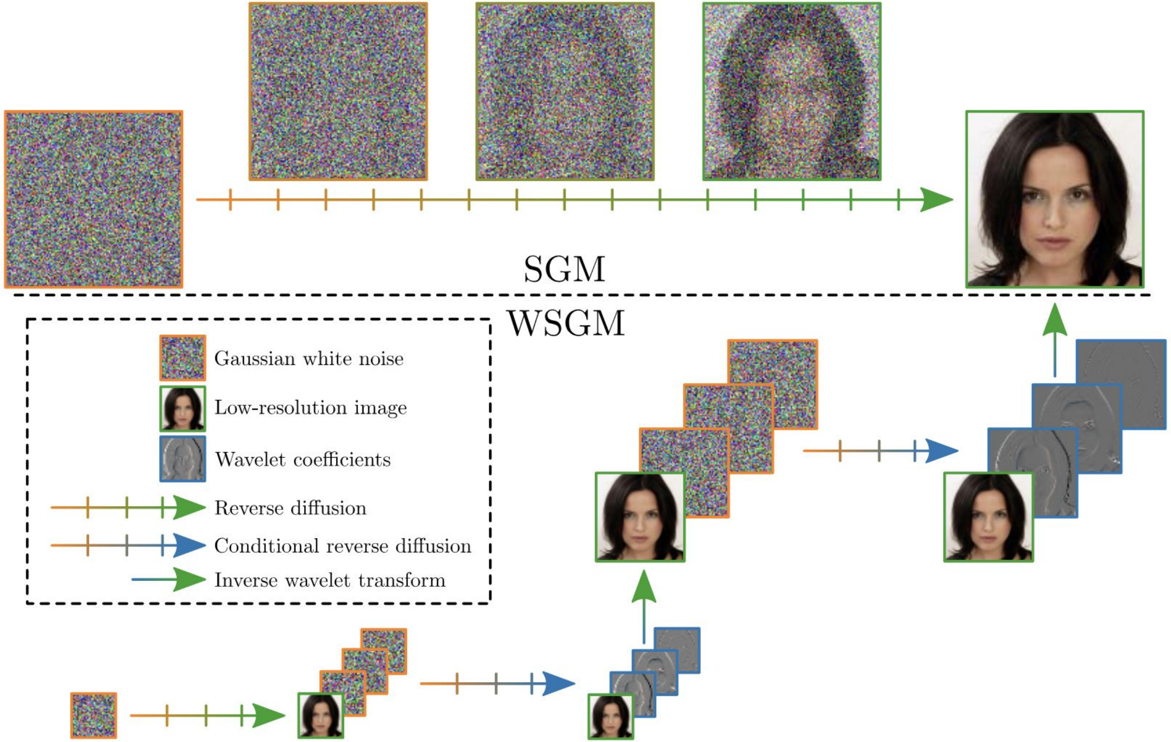
Theorem 2: extension of theorem 1 to non-gaussian processes

⇒ Well-conditioned covariance matrix is crucial to minimize error.

⇒ Non-Gaussian processes with ill-conditioned matrices may need more discretization steps to achieve small errors.

[4] Guth, F., Coste, S., De Bortoli, V., & Mallat, S. (2022). Wavelet score-based generative modeling. Advances in Neural Information Processing Systems, 35, 478-491.

Wavelet Score-Based Generative Model (WSGM)



[4] Guth, F., Coste, S., De Bortoli, V., & Mallat, S. (2022). Wavelet score-based generative modeling. Advances in Neural Information Processing Systems, 35, 478-491.

Discretization and Accuracy for Gaussian Processes

Theorem 3: for Gaussian multiscale processes using the WSGM method

→ Establishes bounds for convergence and error rates.

⇒ This bound is not dependent on the conditioning number of Σ

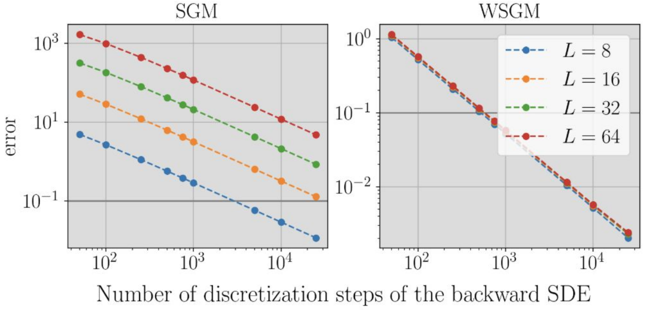
⇒ The number of diffusion steps required to reach a fixed error is independent of the input data size (e.g., image size).

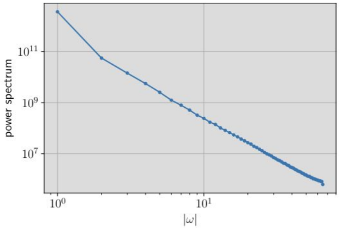
Demonstrates general applicability beyond just Gaussian processes, suggesting broader potential use.

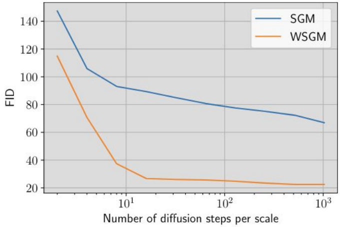
[4] Guth, F., Coste, S., De Bortoli, V., & Mallat, S. (2022). Wavelet score-based generative modeling. Advances in Neural Information Processing Systems, 35, 478-491.

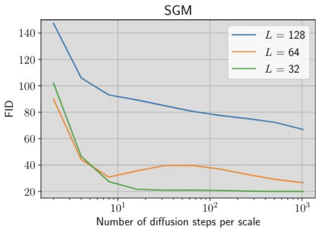
Results

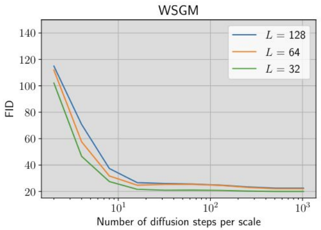
Fréchet Inception Distance (FID): evaluates the quality of generated images by comparing them to a set of real images











Evaluation

Pro

→ Efficiency with fewer iterations

→ Scalable for high-resolution images

→ Superior perceptual quality with lower FID scores

Limitations

→ High complexity due to wavelet transforms and conditional distributions

→ Mainly effective for near-Gaussian multiscale processes

⇒ Extending to non-Gaussian processes may require further techniques

[4] Guth, F., Coste, S., De Bortoli, V., & Mallat, S. (2022). Wavelet score-based generative modeling. Advances in Neural Information Processing Systems, 35, 478-491.

Discussion

Wavelet Transforms in Generative Models

→ Enhance generative models by handling multiscale data and capturing fine details.

→ Ensure high fidelity, preserving intricate structures and improving image quality.

→ Robust in various applications:

⇒ Handling noisy training data.

⇒ Maintaining performance across different scales.

⇒ Achieving better perceptual quality.

Challenges:

→ Increased system complexity.

→ Potentially longer computation times.

→ Limited suitability for all situations.

Promise in Medical Imaging

→ WSGMs' independence from generated image size ⇒ potential for high resolution datasets

→ Unsupervised operation with fine detail preservation and less sensitivity to artifacts

→ Potential for diverse 3D dataset generation of smooth anatomical structures

Future Research Directions:

→ Integrate strengths while mitigating limitations

→ Hybrid approaches combining wavelet transforms with other techniques

⇒ E.g. extend WSGMs to 3D for improved computational efficiency and shape generation

⇒ Investigate performance on ultrasound images, the noisiest and blurriest modality

References

[1] Celard P, Iglesias EL, Sorribes-Fdez JM, Romero R, Vieira AS, Borrajo L. A survey on deep learning applied to medical images: from simple artificial neural networks to generative models. Neural Comput Appl. 2023;35(3):2291-2323. doi:10.1007/s00521-022-07953-4

[2] Croitoru, F. A., Hondru, V., Ionescu, R. T., & Shah, M. (2023). Diffusion models in vision: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[3] Donoho, D. L. (1995). De-noising by soft-thresholding. IEEE transactions on information theory, 41(3), 613-627.

[4] Guth, F., Coste, S., De Bortoli, V., & Mallat, S. (2022). Wavelet score-based generative modeling. Advances in Neural Information Processing Systems, 35, 478-491.

[5] Hui, K. H., Li, R., Hu, J., & Fu, C. W. (2022, November). Neural wavelet-domain diffusion for 3d shape generation. In SIGGRAPH Asia 2022 Conference Papers (pp. 1-9).

[6] K. Kim, G. El Fakhri, and Q. Li, “Low-dose CT reconstruction using spatially encoded nonlocal penalty,” Med. Phys., vol. 44, no. 10, pp. e376–e390, Oct. 2017.

[7] Pinto-Coelho L. How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications. Bioengineering. 2023; 10(12):1435. https://doi.org/10.3390/bioengineering10121435

[8] S. Niu et al., “Sparse-view X-ray CT reconstruction via total generalized variation regularization,” Phys. Med. Biol., vol. 59, no. 12, pp. 2997–3017, Jun. 2014.

[9] Shin, Y. H., Park, M. J., Lee, O. Y., & Kim, J. O. (2020). Deep orthogonal transform feature for image denoising. IEEE Access, 8, 66898-66909.

[10] Sifuzzaman, M., Islam, M. R., & Ali, M. Z. (2009). Application of wavelet transform and its advantages compared to Fourier transform.

[11] Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., & Poole, B. (2020). Score-based generative modeling through stochastic differential equations. arXiv preprint arXiv:2011.13456.

[12] Walker, J. S., & Nguyen, T. Q. (2001). Wavelet-based image compression. Sub-chapter of CRC Press book: Transforms and Data Compression, 267-312.

[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.

QUESTIONS?

Backup Slides

Wavelet enhanced SGM in medical imaging - Dataset & Implementation details

Dataset:

- CT: provided by the AAPM challenge → 5410 images with 512 × 512 pixels and 1 mm thickness ⇒ low-dose and sparse-view is simulated
- MRI: fastMRI knee joint dataset

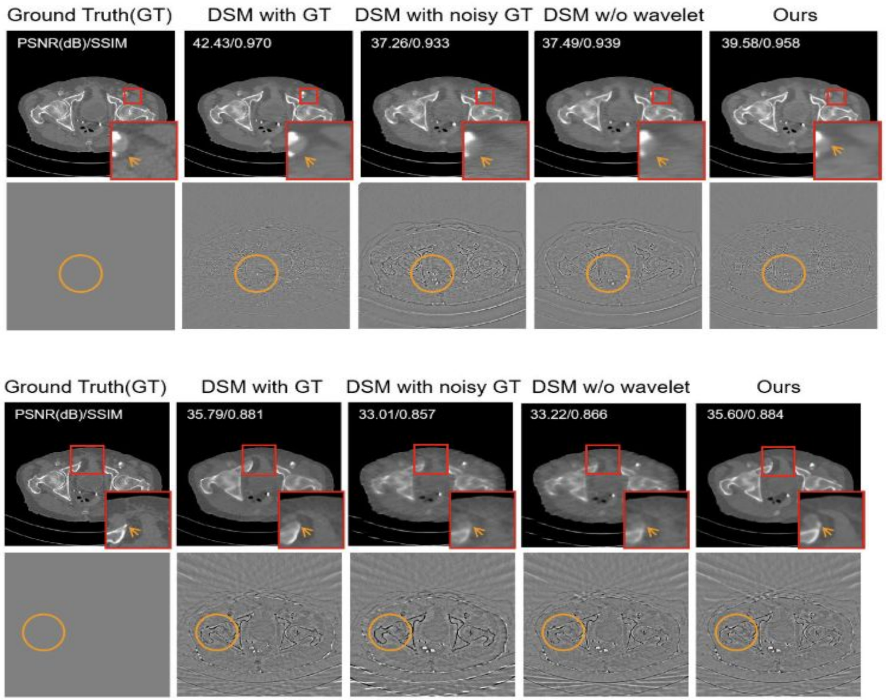
Implementation Details:

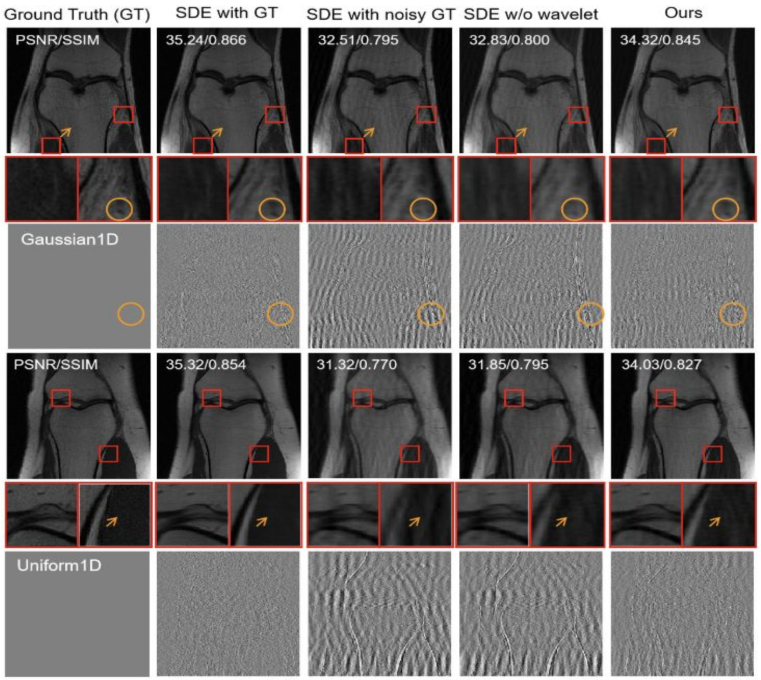
- Haar wavelet
- Network implementation following P. Liu, H. Zhang, W. Lian, and W. Zuo, "Multi-level wavelet convolutional neural networks," IEEE Access, vol. 7, pp. 74973–74985, 2019
- ⇒ modified UNet-architecture

Hardware: Not specified

Code: https://zenodo.org/records/8266123

Results - Ablation Study





[13] Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. IEEE transactions on medical imaging.

Regularization Constraint

transformer img → sino/k-space
balance weight
SGM

$$\boldsymbol{x} = \arg\min_{\boldsymbol{x}} \left[\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \lambda_1 \Phi_2(\Phi_1(\boldsymbol{x})) \right], \quad (12)$$

sinogram/ k-space
reconstructed image
regularization constraint

noise? → $y = Ax + \delta \Leftrightarrow \delta = y - Ax$

$\boxed{\text{data consistency}}$

$$(\boldsymbol{x}^{(k+1)}, \boldsymbol{u}^{(k+1)}) = \arg\min_{\boldsymbol{x},\boldsymbol{u}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \lambda_1 \Phi_2(\Phi_1(\boldsymbol{u}))$$

$$s.t., \boldsymbol{u} = \boldsymbol{x}, \quad (13)$$

iteration number in the reverse diffusion process (up to k)
t ≙ iteration number in forward diffusion process (up to T)
k = T-t

weighting factor > 0

$$(\boldsymbol{x}^{(k+1)}, \boldsymbol{u}^{(k+1)}) = \arg\min_{\boldsymbol{x},\boldsymbol{v}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \lambda_1 \Phi_2(\Phi_1(\boldsymbol{u})) + \lambda_2 \|\boldsymbol{u} - \boldsymbol{x}\|^2. \quad (14)$$

$$\boldsymbol{x}^{(k+1)} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \lambda_2 \left\| \boldsymbol{x} - \boldsymbol{u}^{(k)} \right\|^2, \quad (15)$$

$$\boldsymbol{u}^{(k+1)} = \arg\min_{\boldsymbol{u}} \left\| \boldsymbol{u} - \boldsymbol{x}^{(k+1)} \right\|^2 + \lambda_1 \Phi_2(\Phi_1(\boldsymbol{u})). \quad (16)$$

CT:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k+1/2)} + \gamma \left(\boldsymbol{x}^{(k+1/2)} - \boldsymbol{x}^{(k)} \right). \quad (18)$$

$$\boldsymbol{u}^{(k+1)} = \Phi_2(\Phi_1(\boldsymbol{x}^{(k+1)})), \quad (19)$$

MRI:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \eta \boldsymbol{A}_2^*(\boldsymbol{A}_2 \boldsymbol{x}^{(k)} - \boldsymbol{y}_2) - \lambda_2(\boldsymbol{x}^{(k)} - \boldsymbol{u}^{(k)}), \quad (20)$$

$$\begin{cases} \boldsymbol{u}_{real}^{(k+1)} = \Phi_2(\Phi_1(\boldsymbol{x}_{real}^{(k+1)})) \\ \boldsymbol{u}_{imag}^{(k+1)} = \Phi_2(\Phi_1(\boldsymbol{x}_{imag}^{(k+1)})), \end{cases} \quad (21)$$

3D Shape Gen - Dataset & Implementation details

Dataset: ShapeNet dataset

Implementation Details:

- modified 3D version of the U-Net architecture → same structure for both
- Generator: 800,000 iterations - detail predictor: 60,000 iterations
- Learning rate: $1e^{-4}$
- Training takes three days (generator) and 12 hours (detail predictor)
- The inference takes around six seconds per shape on an RTX 3090 GPU

Hardware:

- pyTorch on a GPU cluster with four RTX3090 GPUs

Code: https://github.com/edward1997104/Wavelet-Generation

Results - Ablation Study

Method	COV ↑		MMD ↓		1-NNA ↓	
	CD	EMD	CD	EMD	CD	EMD
Full Model	**58.19**	**55.46**	**11.70**	**14.31**	**61.47**	**61.62**
W/o detail predictor	54.20	50.96	12.32	14.54	62.46	62.57
VAD Generator	21.83	26.77	21.83	26.77	95.20	93.62
Direct predict TSDF	50.51	50.67	12.83	15.24	68.69	68.29

WSGM - Dataset & Implementation details

Dataset: CelebA-HQ image dataset - 128×128 images

Implementation Details:

- Haar wavelets
- UNet architecture

Hardware: Not specified

Code: Pseudo Code in Paper