Sparse-view Cone-Beam CT Image Reconstruction Based on Conditional Generative Adversarial Networks for Sinogram Restoration

Xuan Xian
Liuzhou City Vocational College
No.1 Wenyuan Road,Guantang Avenue, Yufeng, Liuzhou, Guangxi, China
xxuanxuan@163.com

Guohang He
School of Computer and Electronic Information
Guangxi University
Nanning, Guangxi, China
2313301014@st.gxu.edu.cn

Yuan Linshan
School of Computer Science and Technology
Harbin Engineering University
Harbin, Heilongjiang, China
shanyuanlin@foxmail.com

Haole Guo
School of Computer Science and Technology
Harbin Engineering University
Harbin, Heilongjiang, China
323067022@hrbeu.edu.cn

Zihan Deng
College of Instrument Science and Engineering
Harbin Institute of Technology
Harbin, Heilongjiang, China
306398570@qq.com

2025 IEEE 5th International Conference on Software Engineering and Artificial Intelligence (SEAI) | 979-8-3315-1361-0/25/$31.00 ©2025 IEEE | DOI: 10.1109/SEAI65851.2025.11108841

***Abstract*—Sparse - view projection in cone - beam CT (CBCT) cuts dose and time but causes streak artifacts. Sinogram inpainting can generate missing data for better image quality. A residual encoder-decoder generative adversarial network (Pix2pixGAN) is proposed for sparse - view CBCT reconstruction. It replaces the U - Net generator with a residual encoder - decoder module and uses a PatchGAN - based discriminator. Trained with real CBCT data, Pix2pixGAN was tested at 1/2, 1/3, and 1/4 sparse sampling, compared with linear interpolation and the original Pix2pixGAN. Results show it performs best, especially at 1/4 sampling. In the sinogram domain, RMSE reduces by 7.2%, PSNR increases by 1.5%, and SSIM improves by 1.4%. In the reconstruction image domain, RMSE drops by 5.4%, PSNR rises by 1.6%, and SSIM improves by 1.0%. In conclusion, Pix2pixGAN is suitable for high - quality sparse - angle CBCT reconstruction and has potential for fast, low - dose scanning.**

Keywords—Brain imaging, CBCT, Modal transfer, Self-supervised learning, Image synthesis

I. Introduction

Cone Beam Computed Tomography (CBCT) with flat-panel detectors is widely used in multiple fields. Excessive CBCT scanning doses can harm scanned objects. Sparse projection and reducing exposure parameters are common low-dose scanning methods. Sparse projection not only cuts the dose but also speeds up scanning and avoids motion artifacts from long scanning times [1-2]. Using sparse projection data to generate data for missing angles can enhance the completeness and image quality in sparse projection CBCT, which is significant for fast-scanning CBCT. Currently, the FDK algorithm is the standard for commercial CBCT scanners[3-4].

To solve sparse projection reconstruction problems in Fan-beam Computed Tomography (FBCT), sinogram inpainting fills missing sine info to improve projection data completeness [5]. Traditional methods (LI, directional interpolation, etc.) reduce streak artifacts in reconstructed images but cause issues like continuous structure loss and secondary artifacts. Deep learning, especially CNN, has been applied in FBCT's sinogram inpainting and artifact removal with good results [6]. In [7], U-Net repaired sinograms and FBP enhanced image quality. In [8], two improved U-Nets removed artifacts effectively, with one better than TV iterative reconstruction. In [9], a multi-residual U-Net reduced artifacts while preserving image details.

CBCT and FBCT have different projection data types. CBCT's single-angle projection data lacks angle correlation. Some studies used directional interpolation methods: stacking 3D in to generate missing angle data directly, re-extracting as 2D sinograms then repairing, but with issues of structure loss and secondary artifacts. However, the converted CBCT sinogram has SSIM with FBCT's, so FBCT's deep learning inpainting methods can be applied to CBCT to overcome traditional interpolation flaws, improve reconstruction, and fill the research gap.

Since 2014, GAN with generator and discriminator has many variants used widely in computer vision tasks. In medical imaging. They show GAN's potential for CBCT reconstruction. This paper combines residual module and GAN to propose PIX2PIXGAN for CBCT sinogram inpainting. Replacing U-Net in Pix2pixGAN with residual module retains structure and continuity, prevents gradient issues, and a PatchGAN-based discriminator boosts generator performance. The reconstruction method has 3 steps: convert sparse data to 3D sinogram, upsample, train/test with PIX2PIXGAN, convert results back, and use FBP for reconstruction and evaluate in projection and image domains.

II. Data Processing

This section mainly introduces the methods of data collection and processing to better test the performance of the proposed method in this study.

Collected by the X-ray scanner developed by XRE nv Company, with walnuts scanned. The scanner consists of a cone-beam micro-focus X-ray source, a CMOS flat-panel detector, and an object turntable. The X-ray source has a rated 90 kV voltage and 90 W power, emitting multi-energy rays. The detector has a 1536×1944 pixel matrix with 74.8 μm pixels. For walnut scanning, the tube voltage was set at 40 kV, power at 12 W, exposure time at 80 ms, and the detector used 2×2 binning mode. The distances from the X-ray source focus to the rotation axis and detector center were 66 mm and 199 mm respectively, resulting in a 3.016 magnification ratio. The turntable with foam fixation could rotate 360° with a 0.3° step and had 1201 projections (first and last positions coincided). Only data from the medium height position of the X-ray source relative to walnuts was chosen.

Here, V represents detector pixels in the rotation direction, U is for those perpendicular to the rotation plane, N is the projection number, and η is the sparse sampling factor. In scanning, the detector and X-ray source rotate synchronously. The detector pixel matrix is U×V, and full-sampling has N projections. To get paired sinograms for full-sampling and repair, a linear interpolation upsampling method was used:

Stack the N full-sampled CBCT projection images by angle to form a U×V×N matrix.

Use η to extract data from the matrix to get a U×V×N/η matrix.

Transform the two matrices to get U×N×V (full-sampling) and U×N/η×V (sparse-sampling) sinograms, which can be split into slices of U count with V×N and V×N/η dimensions respectively.

Upsample sparse-sampling sinogram slices by linear interpolation to get a U×N×V matrix as the sinogram to be repaired, which pairs with the full-sampling sinogram for training and testing. Given dataset parameters (U = 972, V = 768, N = 1201, η = 2, 3, 4), 10 walnuts' data were for training and 2 for testing. After sinogram conversion and data cleaning, 8460 slices formed the training set and 1944 slices the test set.

III. Method

A. *Sine graph repair process based on PIX2PIXGAN*

In the realm of sparse projection Cone Beam Computed Tomography (CBCT) reconstruction, the training phase of PIX2PIXGAN plays a crucial role. The full-sampling sinogram slices, which are meticulously prepared by following a series of precise data acquisition and transformation steps, along with the sinogram slices that require repair due to sparse sampling-induced information gaps, are introduced into the PIX2PIXGAN framework as the fundamental training data constituents. Here, the full-sampling sinogram slices, being the benchmark or the ground truth, hold the key to guiding the learning process of the network. They possess the complete and accurate information that the network strives to approximate during the training iterations.

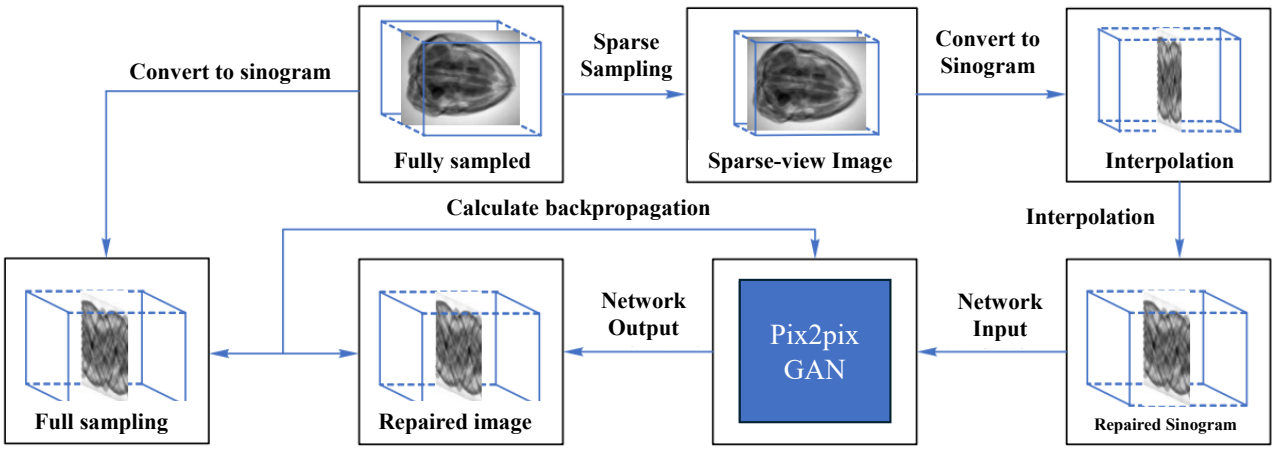

Figure 1. Pix2pixGAN performance evaluation process

As the training unfolds within the Pix2pixGAN's complex architecture, which is composed of multiple layers of residual encoder-decoder modules and a well-designed generative adversarial setup, a meticulous loss calculation mechanism comes into play. The network painstakingly computes the disparity, quantified as the loss, between the full-sampling sinogram slices and the sinogram slices that it generates during the repair attempts. This loss is not merely a numerical value but a reflection of how far the network's current output deviates from the ideal, complete data. Through the powerful technique of backpropagation, this loss is then channeled back through the layers of the network, triggering a cascade of adjustments to the network's parameters. These adjustments, small yet cumulative, work towards reducing the loss over successive training epochs, gradually establishing a highly accurate mapping relationship. This mapping relationship is what enables the network to take the incomplete, to-be-repaired sinogram slices and transform them into something that closely resembles the full-sampling

sinogram slices, mimicking the complete information they contain.

Finally, the Filtered Back Projection (FBP) algorithm, a time-tested and widely used reconstruction technique in the field of CT imaging, is deployed to transform the refined projection data into the final repaired reconstructed image. This image, along with the repaired sinogram itself, becomes the yardstick for evaluating the performance of Pix2pixGAN. By analyzing various metrics such as the reduction in artifacts, the preservation of structural details, and the overall visual quality in both the sinogram domain and the reconstructed image domain, one can comprehensively assess how effectively Pix2pixGAN has managed to repair the sparse projection data and improve the quality of the CBCT reconstruction process.

B. Residual Encoding and Decoding Generator

A sinogram is the result of stacking projection data in the order of sampling angles. Compared with the projection data, it exhibits a high degree of continuity in image structure, meaning that each part is highly correlated with its neighboring regions. The original generator of Pix2pixGAN is U-Net, and the multiple downsampling structures adopted by U-Net can lead to the loss of information in the original image, thus impairing this continuity. To maintain the continuity of the sinogram as much as possible, this paper replaces the original generator of Pix2pixGAN with the residual module, which retains the continuous structural information in the sinogram through symmetric convolutional and deconvolutional layers with residual connections. The structure of the residual generator is shown in Figure 2. In the residual module, each convolutional layer and deconvolutional layer have the same structure, with 96 channels, a convolution kernel size of 4×4, and a convolution stride of 1. Each convolutional and deconvolutional layer is followed by a Leaky ReLU (Rectified Linear Unit) layer as the activation function layer of the network, with the parameter of Leaky ReLU set to 0.2.

To boost network performance, a PatchGAN-based conditional discriminator differentiates the consistency between fully sampled and repaired images. Unlike traditional GAN discriminators that give a single "true" or "false" for generator outputs, this one yields an N×N evaluation matrix, with each value assessing a specific input region. It has fewer params, runs faster, suits large images well, and works great even when the matrix is much smaller than the generated image. Its 5-layer structure has a 4×4 convolution kernel, strides of 2 for the first 3 layers and 1 for the last 2. The first 4 layers have 192 channels, the last has 32, each followed by a Leaky ReLU. Feeding (xf, y) and (x, y) in simultaneously, after 5 convolutions, they turn into two 192×50 feature matrices. The discriminator pixel-wise assesses similarity, outputting a 192×50 matrix, where each element reflects the likeness of 70×70 regions in (xf, y) and (x, y). More similarity means values closer to 1, greater difference to 0.

C. Conditional Adversarial Generative Networks

As depicted in Figure 2, the architecture of Pix2pixGAN consists primarily of two core components: the generator G and the conditional discriminator D. The generator G is tasked with transforming the input sinogram slices X that require repair into the repaired sinogram slices Y', aiming to produce an output that closely approximates the characteristics and information content of the target image Y. Its objective is to fill in the gaps and rectify the deficiencies present in the original, damaged slices.

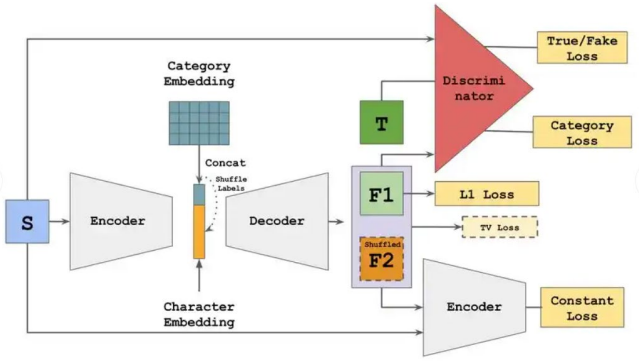

Figure 2. Network architecture diagram of Pix2pixGAN

IV. Experiment and Results

In the context of this research endeavor, the deep learning framework of choice is PyTorch, specifically version 1.12.0, renowned for its robustness and versatility in handling complex neural network architectures. The computational powerhouse facilitating our experiments is a state-of-the-art workstation, outfitted with not one, but two high-performance NVIDIA GeForce 2080Ti GPU cards. These GPUs, with their advanced processing capabilities, ensure rapid and efficient computation, enabling the training and testing of our proposed Pix2pixGAN model with remarkable speed. The workstation operates under the Ubuntu 20.04 operating system, which, known for its stability and compatibility with scientific computing applications, provides an ideal environment for our deep learning tasks.

Figure 3-5 shows the relevant results. For the projection data with varying sparse sampling factors η (specifically, values of 2, 3, and 4), the learning rates are meticulously tuned to 0.00005, 0.00008, and 0.00010 respectively. This fine-tuning of the learning rate is essential to accommodate the different levels of sparsity in the data, ensuring that the model converges optimally without getting trapped in local minima or experiencing divergence during training. During both the training and testing phases, a normalization step is implemented to map the gray values of the input images to the range of [0, 1]. This normalization not only simplifies the numerical computations within the network but also helps in enhancing the overall stability and performance of the model by ensuring that all input data is on a standardized scale, thereby facilitating better learning and generalization.

Compared with repaired sinograms, the reconstructed images after repair are the key for clinical diagnosis and material composition analysis, more suitable for subjectively evaluating sinogram repair. We selected slices of repaired reconstructed images under three sparse sampling conditions (η = 2, 3, 4) and compared them with corresponding fully sampled ones at the same window settings.

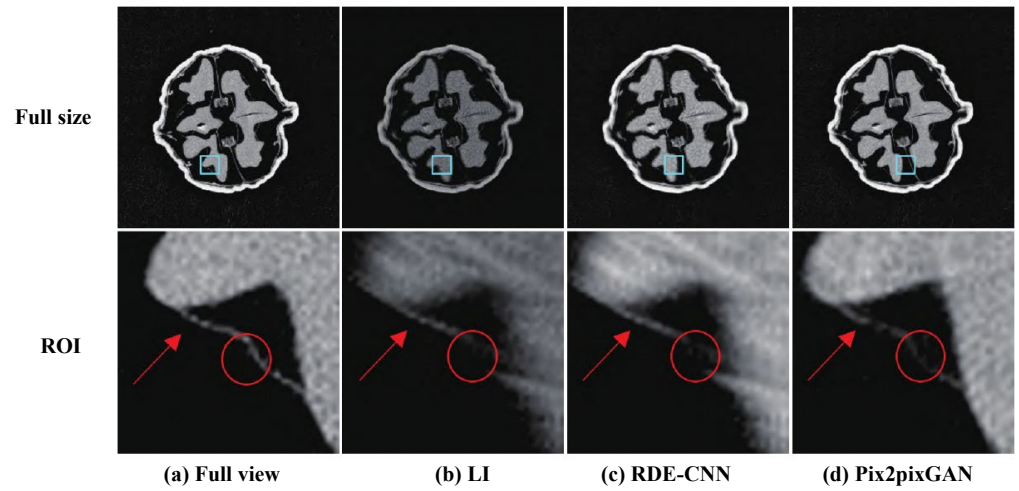

Figure 3. Comparison of restored and reconstructed images under 1/4 sparse sampling conditions

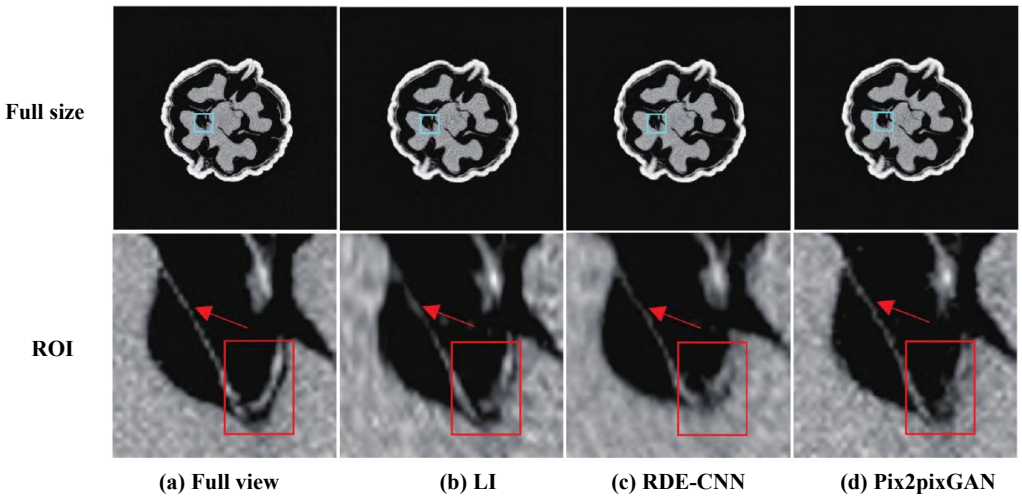

Figure 4. Comparison of restored and reconstructed images under 1/3 sparse sampling conditions

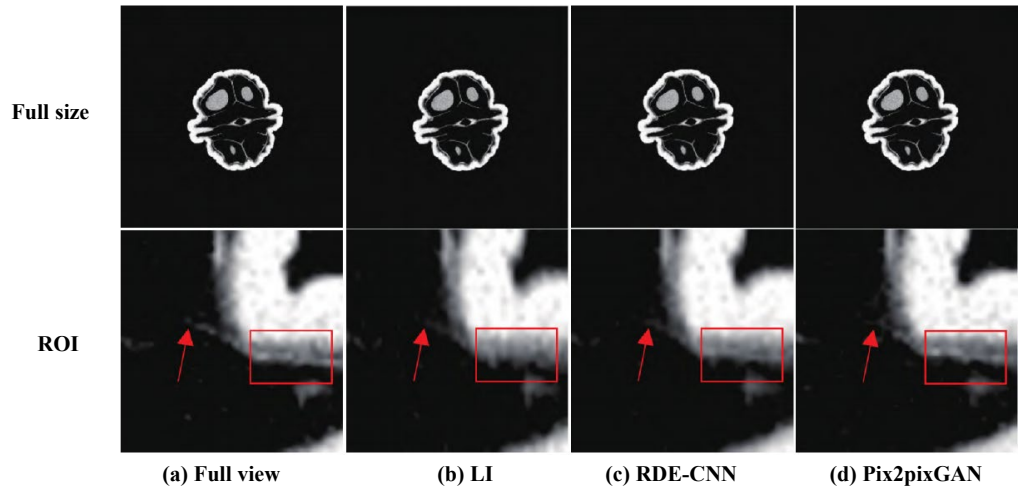

Figure 5. Comparison of restored and reconstructed images under 1/2 sparse sampling conditions

V. Conclusion

In the domain of medical imaging, enhancing the quality of Cone Beam CT (CBCT) scans is crucial. This research presents Pix2pixGAN, a novel solution for sinogram inpainting in sparse-sampled CBCT. Sinograms, vital for CT imaging, frequently face incomplete sampling issues. Pix2pixGAN combines the merits of residual module and Pix2pixGAN. By substituting Pix2pixGAN's generator with the residual module, it excels at retaining essential structural details. The RED module's residual learning capacity captures fine anatomical features, strengthening the overall repair prowess. A PatchGAN-based discriminator is also integrated. It differentiates repaired and full-sampled sinograms. Through adversarial training, the generator and discriminator vie with each other, spurring the network to improve. This process hones the inpainting results, making them closer to the originals.

Experiments under 1/2, 1/3, and 1/4 sparse sampling were carried out. In both sinogram and reconstructed image domains, Pix2pixGAN surpassed LI, and Pix2pixGAN. Metrics like RMSE, PSNR, and SSIM affirmed its superiority. Significantly, as data loss grew more severe with lower sampling ratios, Pix2pixGAN 's edge became even more pronounced.

To sum up, Pix2pixGAN proves to be an outstanding sinogram inpainting scheme for CBCT. It offers great potential to augment the diagnostic utility of CBCT scans, especially when dealing with limited data acquisition challenges in clinical practice.

References

[1] MOHAN K A, VENKATAKRISHNAN S V, GIBBS J W, et al. TIMBIR: a method for time-space reconstruction from interlaced views[J]. IEEE Transactions on Computational Imaging, 2015, 1(2): 96-111.DOI: 10.1109/TCI.2015.2431913

[2] KALKE M, SILTANEN S. Sinogram interpolation method for sparse-angle tomography[J]. Applied Mathematics, 2014, 5(3): 423-441. DOI:10.4236/am.2014.53043

[3] KIM H G, YOO H. Image enhancement for computed tomography using directional interpolation for sparsely-sampled sinogram[J]. Optik, 2018, 166: 227-235. DOI:10.1016/j.ijleo.2018.03.139

[4] KÖSTLER H, PRÜMMER M, RÜDE U, et al. Adaptive variational sinogram interpolation of sparsely sampled CT data[C]//Proceedings of the 18th International Conference on Pattern Recognition. Piscataway: IEEE, 2006, 3: 778-781. DOI: 10.1109/ICPR.2006.225

[5] HERRAIZ J L, ESPANA S, VICENTE E, et al. Frequency selective signal extrapolation for compensation of missing data in sinograms[C]//Proceedings of the 2008 IEEE Nuclear Science Symposium/Medical Imaging Conference/16th International Workshop on Room-Temperature Semiconductor X- and Gamma-ray Detectors' Conference Record. Piscataway: IEEE, 2008: 4299-4302. DOI: 10.1109/NSSMIC.2008.4774234

[6] Ohnesorge B, Flohr T, Schwarz K, et al. Efficient correction for CT image artifacts caused by objects extending outside the scan field of view[J]. Medical physics, 2000, 27(1): 39-46.

[7] LEE H, LEE J, KIM H, et al. Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction[J]. IEEE Transactions on Radiation and Plasma Medical Sciences, 2019, 3(2): 109-119. DOI:10.1109/TRPMS.2018.2867611

[8] HAN Y, YE J C. Framing U-Net via deep convolutional framelets: application to sparse-view CT[J]. IEEE Transactions on Medical Imaging, 2018, 37(6): 1418-1429. DOI: 10.1109/TMI.2018.2823768

[9] Fang Y, Wei X, Ma J. High-precision DOA estimation based on synthetic aperture and sparse reconstruction[J]. Sensors, 2023, 23(21): 8690.