Noise-Aware Complementary GAN for Low-Dose CT Image Denoising

Paturi Siva Prakash[1], Karthikeyan A[2]

School of Computer Science and Engineering
Vellore Institute of Technology, Vellore, India

Course: SWE1010 - Digital Image Processing | **Semester:** Fall 2025-26 | **Slot:** G1+TG1 | **Faculty:** Dr. S. Geetha
[1]20MIA1107 | [2]23MIA1123

Abstract

Computed Tomography (CT) scans are essential for modern medical diagnosis, but the radiation involved poses long-term health risks. While reducing radiation dose is desirable, it considerably increases noise in reconstructed CT images, affecting diagnostic accuracy. This project proposes a novel Noise-Aware Complementary Generative Adversarial Network (NAC-GAN) designed to address these limitations. The model uses a dual-head generator that predicts both the clean CT image and corresponding noise residual. Combined with two discriminators—one focusing on local textures and another on global structural realism—the system encourages high-quality denoising while retaining clinically relevant details. Preliminary results demonstrate strong potential for enhancing low-dose CT image quality while maintaining diagnostic integrity.

Keywords—*CT denoising, Generative Adversarial Networks, Low-dose imaging, Medical image processing, Dual-head architecture, Noise modeling*

I. INTRODUCTION

CT imaging plays a vital role in medical diagnostics due to its ability to provide clear and detailed cross-sectional views of internal structures. However, CT scans rely on ionizing radiation, which poses serious health risks for patients undergoing repeated imaging. To address this, medical practitioners often reduce radiation dose, but this significantly increases noise in the final image. A noisy CT scan makes it difficult for radiologists to identify abnormalities, as noise can obscure subtle structures or mimic pathological patterns.

Over the past decade, deep learning has emerged as a powerful tool for medical image denoising. Convolutional Neural Networks (CNNs) have shown great promise due to their ability to learn complex noise patterns directly from data. However, traditional CNN approaches often struggle to maintain textural and anatomical details present in original CT images. Generative Adversarial Networks (GANs) introduced a new direction by focusing on producing perceptually realistic outputs, but most GAN-based approaches rely on a single-path generator that directly maps noisy images to clean ones, often compromising between noise reduction and detail preservation.

This project addresses these limitations by explicitly modeling both the noise and underlying clean signal. By encouraging the generator to learn these two components complementarily, we create a stronger foundation for high-quality denoising. The proposed Noise-Aware Complementary GAN represents a novel approach to low-dose CT enhancement.

II. RELATED WORK

Early approaches to CT image denoising relied heavily on classical image processing filters. Methods such as Non-Local Means and BM3D attempted to use patch similarity and transform-domain filtering to remove noise. Although these methods provided moderate improvements, they often resulted in excessive smoothing, especially in low-contrast regions, and were computationally expensive.

The advent of deep learning revolutionized medical image denoising. Models like DnCNN introduced residual learning, where networks predict noise rather than clean output, allowing faster convergence. U-Net architectures enhanced performance through skip connections that transmitted high-frequency information from encoder to decoder. Later models such as RED-CNN specifically optimized CNN structures for CT restoration.

GAN-based approaches brought a major shift. By introducing a discriminator that distinguishes between real and generated images, GANs encourage generators to create more natural and realistic textures. Several studies applied GANs to CT denoising with considerable success. However, most works depended on single output pathways and patch-level discriminators, sometimes leading to locally correct but globally inconsistent outputs.

The gap in existing literature is the lack of models that simultaneously predict clean images and noise residuals with both local and global structural supervision. This motivated the present work.

III. PROPOSED METHODOLOGY

A. System Architecture

The proposed NAC-GAN is designed around the principle that explicitly modeling noise helps the generator learn better representations for denoising. The generator has two separate decoding pathways: one dedicated to reconstructing the clean image and another to predicting underlying noise. These complementary outputs enable more accurate restoration.

B. Generator Design

At the core is a U-Net-based generator with a shared encoder. The encoder extracts features from noisy images and gradually compresses them into latent representation. From this latent space, two independent decoders operate in parallel. The reconstruction decoder recovers the clean CT image, while the noise decoder predicts the noise component. Since noisy images mathematically equal the sum of clean images and noise, this dual-head mechanism encourages balanced representation that captures details precisely.

C. Discriminator Architecture

The system includes two discriminators. The first is a PatchGAN-inspired local discriminator that examines small patches across denoised images to ensure realistic local textures. The second is a global discriminator that evaluates entire CT slices to maintain overall anatomical plausibility. Together, these provide supervision at different scales, creating more realistic results.

D. Loss Function

The loss function combines reconstruction, adversarial, and noise modeling components, with a planned orthogonality constraint to enforce independence between noise and reconstruction errors. This holistic combination ensures the generator is guided by both pixel-level correctness and perceptual realism.

IV. EXPERIMENTAL SETUP

A. Implementation Details

The entire system is implemented in Python using PyTorch deep learning framework. Supporting libraries include NumPy, h5py, OpenCV, and scikit-image for preprocessing, data loading, and metric computation. The codebase follows modular structure separating data handling, model definitions, training scripts, and utilities.

During training, both discriminators are updated before updating the generator to ensure meaningful gradients and prevent early collapse. The system includes validation routines computing PSNR and SSIM after each epoch. Cosine-annealing learning rate scheduling gradually reduces learning rate to stabilize later training epochs.

B. Dataset Description

The raw dataset consists of clinical CT volumes stored in HDF5 format. Each volume contains multiple slices varying in contrast, resolution, and anatomical coverage. Intensity windowing is applied to highlight soft tissues before normalization. Slices are resized to 256×256 pixels for consistency.

Clean slices extracted from volumes are used to generate synthetic noisy versions. Initial dataset versions use Gaussian noise for simplicity and computational speed. Although this doesn't perfectly represent real CT noise—which follows combined Poisson and Gaussian distributions—it allows initial model training and debugging. More realistic Poisson-based simulation is planned for future iterations.

After preprocessing, approximately 3,684 clean images and 1,040 noisy-clean pairs were generated. The dataset is divided into training, validation, and testing splits (70:15:15 ratio), ensuring slices from the same patient don't appear in multiple splits to prevent data leakage.

C. Training Strategy

Training NAC-GAN requires careful balancing of multiple loss components and stable GAN optimization. Adam optimizer is used with parameters commonly recommended for GANs. Several stabilization techniques are incorporated: label smoothing, spectral normalization, and gradient clipping. TensorBoard logging allows real-time visualization of loss curves and sample outputs.

Due to hardware limitations, initial experiments run on CPU, resulting in longer training times. Despite constraints, the training loop is optimized to reduce overhead through parallel data loading and memory-efficient processing.

V. RESULTS AND DISCUSSION

A. Evaluation Metrics

Primary evaluation metrics are Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). PSNR quantifies pixel-level accuracy compared to ground truth. Higher PSNR values indicate lower reconstruction error. However, PSNR alone doesn't always reflect perceptual quality.

SSIM addresses this by measuring luminance, contrast, and structural similarity. Since CT images contain subtle diagnostic textures, SSIM provides meaningful indication of whether denoised images retain important structures. Qualitative evaluation through visual comparison complements numerical metrics.

B. Preliminary Results

Initial experiments with limited training epochs show encouraging trends. Even after one or two epochs, the model demonstrates noticeable noise reduction without severely compromising anatomical structures. Although slight smoothing is visible, this is typical in early-stage training and expected to improve with additional iterations.

The dual-head architecture performs as intended, producing both clean reconstructions and noise predictions without implementation issues. Training logs confirm stable loss behavior with no signs of mode collapse. Visualization outputs indicate the network learns meaningful representations despite hardware limitations.

These preliminary results validate the feasibility of the proposed architecture. Complete results will be established after full dataset preprocessing and longer training runs. Baseline comparisons with BM3D, DnCNN, and RED-CNN are planned to contextualize NAC-GAN performance.

C. Challenges Encountered

The primary challenge is hardware limitation. The Intel Arc GPU currently available is not fully compatible with PyTorch CUDA drivers, forcing CPU-only operation. This dramatically reduces training speed and restricts experimentation. Completing dataset preprocessing is computationally expensive and time-consuming.

Achieving realistic noise modeling remains challenging. Real CT noise includes complex Poisson and Gaussian component mixtures. Replicating this accurately without actual low-dose CT pairs is difficult. Balancing multiple GAN loss components also requires substantial experimentation.

VI. CONCLUSION AND FUTURE WORK

This project introduces a novel dual-head GAN architecture for low-dose CT image denoising. By explicitly modeling both clean reconstruction and noise components, the proposed approach addresses major limitations of existing reconstruction-only models. Adding both local and global discriminators ensures generated images preserve fine textures and overall structural realism.

Despite hardware limitations, development and initial testing indicate strong potential. Once trained on full datasets with realistic noise simulation and optimized loss functions, the model is expected to provide substantial improvements in CT image quality at reduced radiation doses.

Future enhancements include completing dataset preprocessing, implementing orthogonality loss, and conducting longer training sessions with suitable hardware. Medium-term goals involve architectural enhancements such as attention mechanisms, multi-scale processing, and perceptual loss integration. Long-term plans include extending to 3D volume denoising, clinically realistic noise modeling, uncertainty estimation, and validation on real low-dose clinical data. This work contributes toward safer and more reliable diagnostic imaging, ultimately supporting better patient outcomes.

REFERENCES

[1] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3142-3155, 2017.

[2] H. Chen et al., "Low-dose CT via convolutional neural network," Biomedical Optics Express, vol. 8, no. 2, pp. 679-694, 2017.

[3] I. Goodfellow et al., "Generative adversarial nets," Advances in Neural Information Processing Systems, pp. 2672-2680, 2014.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," Medical Image Computing and Computer-Assisted Intervention, pp. 234-241, 2015.

[5] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," IEEE Transactions on Image Processing, vol. 16, no. 8, pp. 2080-2095, 2007.

[6] Q. Yang et al., "Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," IEEE Transactions on Medical Imaging, vol. 37, no. 6, pp. 1348-1357, 2018.

[7] P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," IEEE Conference on Computer Vision and Pattern Recognition, pp. 1125-1134, 2017.

[8] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, 2004.