

A-IDE : Agent-Integrated Denoising Experts

Uihyun Cho*¹

CUIHYUN12@YONSEI.AC.KR

Namhun Kim*¹

KSOUTH0413@YONSEI.AC.KR

¹ *Yonsei University, Seoul, South Korea*

Editors: Under Review for MIDL 2025

Abstract

Recent advances in deep learning-based denoising methods have improved the quality of Low-Dose CT (LDCT) images. However, due to anatomical variability and limited data availability, a single model often struggles to generalize effectively across multiple anatomical regions. To address this limitation, we propose the Agent-Integrated Denoising Experts (A-IDE) framework. A-IDE integrates three region-specialized RED-CNN models under the control of a decision-making large language model (LLM) agent. This agent analyzes anatomical priors extracted from BiomedCLIP and dynamically routes incoming LDCT scans to the most suitable specialized model. We highlight three major advantages. First, A-IDE shows robust performance in heterogeneous and data-scarce environments. Second, the framework reduces risk of overfitting by distributing tasks among multiple experts. Finally, the fully automated agent-driven routing eliminates the need for manual intervention. Experimental results in the Mayo-2016 dataset confirm that A-IDE achieves superior performance in RMSE, PSNR, and SSIM compared to a single unified denoiser.

Keywords: Low-Dose CT, Image Denoising, LLM, Agents

1. Introduction

LDCT significantly reduces patient radiation at the cost of increased noise and artifacts, which can obscure diagnostic details and adversely affect clinical decision making (Zhang et al., 2024a). Although deep learning-based methods like RED-CNN (Chen et al., 2017), GANs (You et al., 2019), Transformers (Zhang et al., 2021), and Diffusion models (Zhao et al., 2023) have shown notable success in denoising LDCT images, they typically rely on single and generalized architecture. However, anatomical regions differ greatly in terms of Hounsfield Unit (HU) intensity distributions, scanner protocols, and artifact characteristics. As a result, not a single method consistently performs well across all regions (Deng and Campbell, 2025). Furthermore, data scarcity and privacy constraints often lead models to overfit on a limited range of anatomies or characteristics, particularly in underrepresented or imbalanced categories (Won et al., 2021). These regional variations and limited data highlight the need for context-aware, region-specific denoising frameworks (Chen et al., 2023) (Yang et al., 2025).

To overcome these limitations, we propose A-IDE, which employs a decision-making LLM agent to evaluate the anatomical context of each incoming LDCT image and orchestrate the optimal denoising experts. First, the input CT image is processed by BiomedCLIP (Zhang et al., 2024b) to generate an anatomical prior distribution over anatomical structures. This distribution, along with textual descriptions of three pretrained RED-CNN

* Contributed equally

denoising models—each specialized for a distinct anatomical region—is passed to a LLM agent (gpt-4o) (OpenAI, 2024). Guided by prompts detailing each model’s anatomical focus, the agent dynamically selects the most appropriate expert model for the given input. The selected model then reconstructs the denoised image patches and reports quantitative metrics. Our experimental results show that A-IDE achieves superior performance in terms of PSNR, SSIM, and RMSE compared to baseline approach.

2. Methods

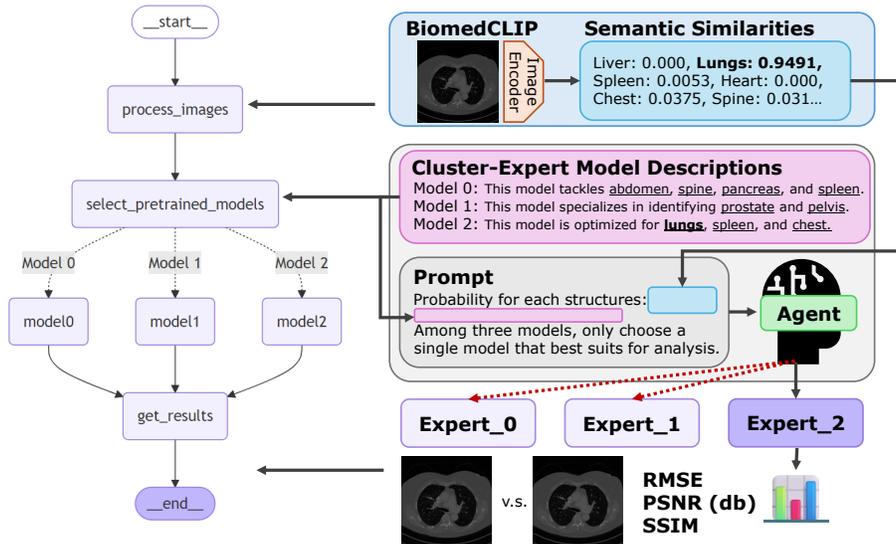


Figure 1: A-IDE Architecture.

We use the 2016 Low-Dose X-ray CT Grand Challenge dataset (Mayo-2016) (McCollough et al., 2017), which includes paired LDCT and NDCT images from 10 anonymized patient scans. Each 512×512 image is normalized and divided into non-overlapping 55×55 patches to yield 81 patches per image. To simulate anatomy-specific denoising, we cluster the dataset into three groups based on semantic embeddings extracted from BiomedCLIP. These high-dimensional vectors are reduced to two dimensions using Principal Component Analysis (PCA) and clustered using K-means ($k=3$). We characterize each cluster by averaging the anatomical structure probabilities over a randomly selected set of 100 images, thereby capturing the dominant anatomical features within each group.

We train a baseline model and three cluster-specific expert models using the RED-CNN architecture. The models are optimized using the Adam optimizer with an initial learning rate of $1e-5$, a mean squared error (MSE) loss, and a learning rate scheduler that decays the rate to with a minimum threshold $1e-10$. We then introduce an LLM-driven agent that autonomously selects the most suitable expert model based on the anatomical content of each CT image. Each image is first processed by BiomedCLIP to extract semantic embeddings and generate a probability distribution over twenty anatomical structures. This distribution, along with concise descriptions of the three specialized RED-CNN models, is

passed as a prompt to a GPT-4o agent. Guided by this contextual information, the agent selects the optimal model to perform patch-wise denoising.

3. Results and Discussion

Methods	RMSE	PSNR	SSIM
Baseline	0.097 \pm 0.00164	43.06 \pm 1.73	0.9557 \pm 0.0125
Expert 0	0.097 \pm 0.00245	42.15 \pm 2.28	0.9483 \pm 0.0200
Expert 1	0.086 \pm 0.00165	43.33 \pm 1.82	0.9576 \pm 0.0107
Expert 2	0.107 \pm 0.00196	40.89 \pm 1.39	0.9435 \pm 0.0147
A-IDE	0.094 \pm 0.00169	43.42 \pm 1.72	0.9583 \pm 0.0110

Table 1: Evaluation metrics (mean \pm standard deviation) for the baseline, cluster-specific expert models, and the proposed A-IDE framework. A-IDE consistently outperforms the baseline in RMSE, PSNR, and SSIM, and achieves the highest PSNR and SSIM scores among all models.

Table 1 presents the quantitative performance of all models in terms of RMSE, PSNR, and SSIM. The single, generalized baseline model demonstrates competitive performance across all metrics. However, the cluster-specific expert models, especially Expert 1, exhibit superior performance due to training tailored to specific anatomical regions. The proposed A-IDE framework effectively leverages the strengths of these individual anatomical experts to deliver superior overall performance. Notably, A-IDE outperforms all individual expert models in both PSNR and SSIM, demonstrating the effectiveness of the proposed framework’s adaptive integration strategy in harnessing the complementary strengths of multiple specialized denoisers to achieve more accurate and reliable reconstruction—particularly in the presence of complex, overlapping, or anatomically ambiguous regions. As a result, A-IDE achieves enhanced visual fidelity and structural preservation compared to both the baseline and individual expert models.

4. Conclusion

Our results demonstrate the effectiveness of employing an intelligent agent to dynamically route inputs to the most appropriate expert model. By harnessing the complementary strengths of multiple specialized RED-CNN experts, the agent-driven A-IDE framework achieves a balanced trade-off between low reconstruction error, high signal-to-noise ratio, and accurate structural preservation. As a result, A-IDE delivers robust performance across diverse imaging scenarios, positioning it as a promising solution for enhancing LDCT reconstruction in real-world clinical applications.

References

- Hu Chen, Yi Zhang, Mannudeep K Kalra, et al. Low-dose ct with a residual encoder-decoder convolutional neural network. *IEEE Trans. Med. Imaging*, 36(12):2524–2535, 2017.
- Zhihao Chen, Qi Gao, Yi Zhang, and Hongming Shan. *ASCON: Anatomy-Aware Supervised Contrastive Learning Framework for Low-Dose CT Denoising*, page 355–365. Springer Nature Switzerland, 2023. ISBN 9783031439995. doi: 10.1007/978-3-031-43999-5_34. URL http://dx.doi.org/10.1007/978-3-031-43999-5_34.
- Z. Deng and J. Campbell. Sparse mixture-of-experts for non-uniform noise reduction in mri images. *arXiv preprint arXiv:2501.14198*, 2025.
- C.H. McCollough, A. Bartley, R.E. Carter, B. Chen, T.A. Drees, P. Edwards, D.R. Holmes, A.E. Huang, F. Khan, S. Leng, K. McMillan, G.J. Michalak, K.M. Nunez, L. Yu, and J.G. Fletcher. Low-dose ct for the detection and classification of metastatic liver lesions: Results of the 2016 low dose ct grand challenge. *Medical Physics*, 44:e339–e352, 2017.
- OpenAI. Gpt-4o system card, 2024. URL <https://arxiv.org/abs/2410.21276>.
- D.K. Won, E. Jung, S. An, P. Chikontwe, and S.H. Park. Self-supervised learning based ct denoising using pseudo-ct image pairs. *arXiv preprint arXiv:2104.02326*, 2021.
- Z. Yang, Y. Chen, Z. Wang, H. Shan, Y. Chen, and Y. Zhang. Patient-level anatomy meets scanning-level physics: Personalized federated low-dose ct denoising empowered by large language model. 2025. Unpublished manuscript.
- Chenyu You, Guang Li, Yi Zhang, et al. Ct super-resolution gan constrained by the identical, residual, and cycle learning ensemble (gan-circle). *IEEE Trans. Med. Imaging*, 39(1):188–203, 2019.
- J. Zhang, W. Gong, L. Ye, F. Wang, Z. Shangguan, and Y. Cheng. A review of deep learning methods for denoising of medical low-dose ct images. *Computers in Biology and Medicine*, 2024a. doi: 10.1016/j.compbiomed.2024.108112.
- Sheng Zhang, Zhihong Chen, Yi Zhang, and Hongming Shan. A multimodal biomedical foundation model trained from fifteen million image–text pairs. *NEJM AI*, 2(1), 2024b. doi: 10.1056/AIoa2400640. URL <https://ai.nejm.org/doi/full/10.1056/AIoa2400640>.
- Zhicheng Zhang, Lequan Yu, Xiaokun Liang, Wei Zhao, and Lei Xing. Transct: Dual-path transformer for low dose computed tomography. In *Medical Image Computing and Computer Assisted Intervention*, pages 55–64. Springer, 2021.
- Mingyuan Zhao, Haoran Wang, Wenxiang Xu, and Ge Wang. Denoising diffusion probabilistic model for low-dose ct restoration. *IEEE Transactions on Medical Imaging*, 42(12):3211–3222, 2023.