Enhancing X-Ray Fluorescence Computed Tomography Image Quality via Deep Image Prior-Based Denoising

We are developing a 3D X-ray fluorescence computed tomography (XFCT) system using non-radioactive-labeled compounds for biofunctional imaging. To enhance image quality and detection limits, we applied deep image prior (DIP), a convolutional neural network, as a pre-denoising method to projection images before reconstruction. XFCT imaging experiments were conducted using synchrotron radiation on phantoms and mouse brain samples. The DIP pre-denoising showed superior denoising effectiveness compared to post-denoising and other denoising methods, improving the contrast-to-noise ratio by 3.7–4.6 times and the detection limit without degrading the spatial resolution. The image quality of mouse brain significantly improved.

Nuclear medicine imaging technologies, including positron emission tomography (PET) and single-photon emission computed tomography (SPECT), are indispensable for visualizing physiological functions. These imaging modalities contribute not only to clinical diagnosis but also to preclinical animal studies such as drug development, regenerative medicine, and disease modeling. However, PET and SPECT rely on radioisotopes, which require strict radiation control, limiting their routine usability. In contrast, X-ray fluorescence computed tomography (XFCT) offers an emerging alternative for molecular imaging without using radioactive materials.

XFCT combines X-ray fluorescence analysis and tomographic imaging to visualize the spatial distribution of non-radioactive-labeled compounds and nanoparticles inside biological samples. Our research group developed a 3D XFCT system with a volume monochromatic X-ray beam, multi-pinhole collimator, and a 2D detector. This setup enabled us to image iodine compounds and gold nanoparticles with high sensitivity and spatial resolution [1, 2]. However, practical imaging is limited by object size and low photon count due to radiation dose constraints. For instance, mouse brains are approximately 10 mm in diameter, but the mouse head measures around 15 mm, reducing the amount of detectable fluorescence X-rays. This creates the need for software-based image enhancement to improve image quality under low-count conditions.

Denoising is commonly implemented using regularized reconstruction methods like total variation (TV), but parameter tuning is often heuristic. Deep learning has emerged as a powerful tool for medical image enhancement; however, it typically requires large training datasets, which are difficult to obtain in biomedical imaging. Therefore, we applied a training-free deep learning method called deep image prior (DIP) [3]. DIP uses an untrained convolutional neural network (CNN) to denoise images by leveraging the network's structural bias towards natural image statistics. DIP has shown effectiveness in other modalities such as PET, CT, and MRI.

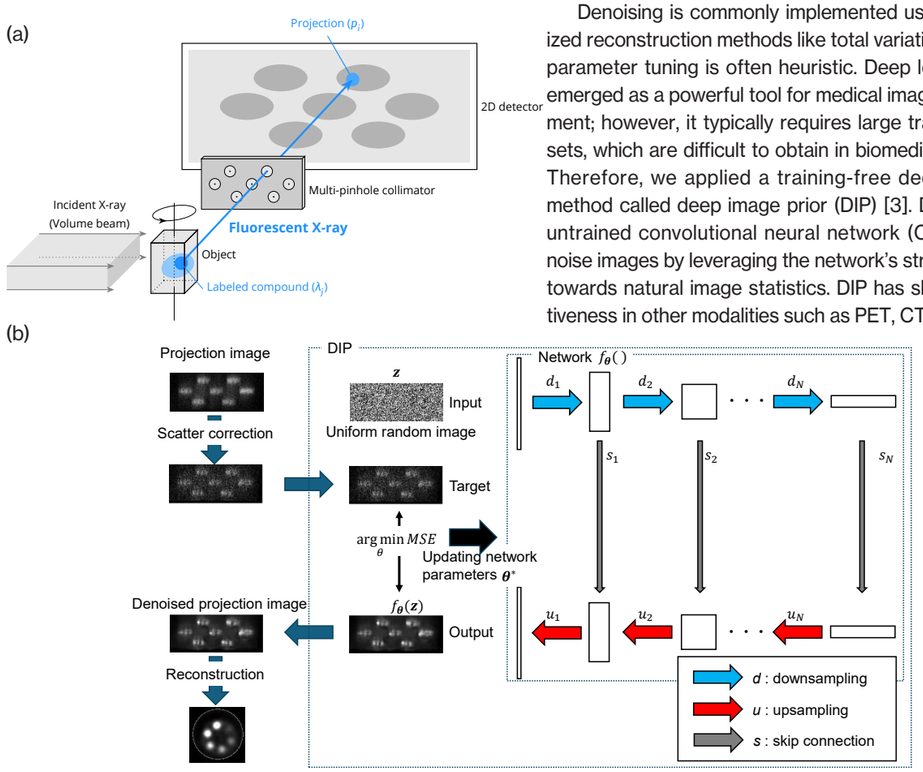

Figure 1: (a) Schematic of 3D XFCT using multi-pinhole collimator. (a) is reused from reference [4]. Copyright Springer Nature. **(b)** Overview of DIP and procedure of XFCT image generation with DIP pre-denosing.

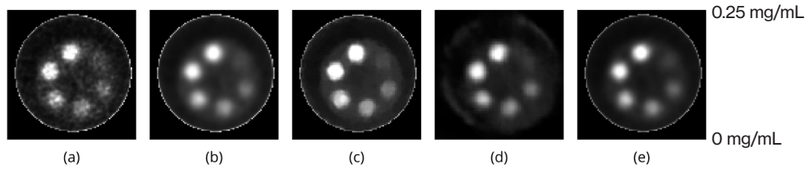

Figure 2: XFCT-reconstructed images of a six-hole phantom filled with iodine solution: **(a)** OS-EM, **(b)** GF + OS-EM, **(c)** EM-TV, **(d)** Post-DIP and **(e)** Pre-DIP. Figure 2 is reused from reference [4]. Copyright Springer Nature.

In this study, we evaluated DIP for noise reduction in multi-pinhole XFCT (mp-XFCT) and compared two applications [4]: DIP as a post-processing method after image reconstruction (Post-DIP) and DIP applied to the projection images before reconstruction (Pre-DIP). To our knowledge, this is the first direct comparison of pre- versus post-denoising using DIP in XFCT.

The mp-XFCT system consisted of a rotation stage, a multi-pinhole collimator, and a 2D detector (**Fig. 1(a)**). The system geometry was designed to reduce Compton scattering by placing the detector at a 90° angle relative to the incident X-ray beam. Iodine was used as the imaging element, detected via K-shell fluorescence at 28.3 keV. Reconstructed images were generated from projection images using the ordered subsets expectation maximization (OS-EM) algorithm, which is well-suited to low-count data.

DIP was implemented using an encoder–decoder CNN with skip connections (**Fig. 1(b)**). DIP minimized the mean square error between the network output and the noisy target image using stochastic gradient descent. A naturally denoised image can be obtained by stopping the update of the parameters before the noise component recovers. The scatter-corrected projection images were subjected to DIP processing, and reconstructed images were generated from the denoised projection images using OS-EM.

For validation, we performed experiments using a physical phantom and an extracted mouse brain. The phantom had six holes filled with iodine solutions at concentrations from 0.05 to 0.30 mg/mL. Denoising was evaluated across five methods: OS-EM alone (**Fig. 2(a)**), OS-EM with Gaussian filtering (GF+OS-EM) (**Fig. 2(b)**), EM-TV reconstruction (**Fig. 2(c)**), Post-DIP (**Fig. 2(d)**), and Pre-DIP (**Fig. 2(e)**). Performance was measured by contrast-to-noise ratio (CNR), spatial resolution (FWHM), and detection limit (LOD).

Pre-DIP demonstrated the best CNR improvement, increasing CNR by 3.4 to 4.6 times over OS-EM. While GF+OS-EM and EM-TV also reduced noise, they introduced blurring or artifacts. Pre-DIP retained spatial resolution with only minimal degradation and improved the detection limit from 0.069 to 0.035 mg/mL – approaching the required threshold for *in vivo* imaging. The quantification of iodine concentrations showed strong linearity ($R^2 \approx 0.98$–0.99) for both Pre-DIP and Post-DIP.

Mouse brain imaging further validated the effectiveness of Pre-DIP. Images reconstructed using OS-EM alone were noisy (**Fig. 3(a)**), obscuring anatomical details. In contrast, Pre-DIP significantly improved image clarity (**Fig. 3(b)**), allowing visualization of iodine-labeled compounds in the cortex and thalamus.

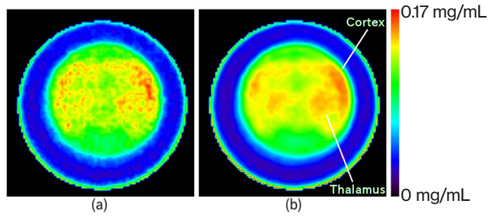

Figure 3: XFCT-reconstructed images of the extracted mouse brain where ^{127}I-iofetamine has accumulated: **(a)** OS-EM, **(b)** Pre-DIP. Figure 3 is reused from reference [4]. Copyright Springer Nature.

In conclusion, applying DIP to projection data prior to reconstruction proved more effective than traditional denoising methods and even DIP post-denoising. Pre-DIP preserved spatial resolution, improved contrast, and reduced the detection limit, all without requiring training data. This approach provides a powerful tool for high-quality XFCT imaging under low-photon conditions and represents a significant step toward practical *in vivo* 3D molecular imaging. Future work will combine this software approach with hardware improvements to enhance detection efficiency.

REFERENCES

[1] T. Sasaya, N. Sunaguchi, K. Hyodo, T. Zeniya and T. Yuasa, *Sci. Rep.* **7**, 5742 (2017).
[2] T. Sasaya. N. Sunaguchi, S.-J. Seo, K. Hyodo, T. Zeniya, J.-K. Kim and T. Yuasa, *Nucl. Instrum. Methods Phys. Res. A* **886**, 71 (2018).
[3] D. Ulyanov. A. Vedaldi and V. Lempitsky, *Int. J. Comput. Vis.* **128**, 1867 (2020).
[4] S. Kusakari, K. Sato, Y. Tsushima, M. Matsuoka, T. Sasaya, N. Sunaguchi, K. Matsubara, H. Kawashima, K. Hyodo, T. Yuasa and T. Zeniya, *Int. J. Comput. Assist. Radiol. Surg.* **20**, 665 (2025).

BEAMLINE

AR-NE7A

S. Kusakari[1], K. Sato[1], Y. Tsushima[1], M. Matsuoka[2], T. Sasaya[2], N. Sunaguchi[3], K. Matsubara[4], H. Kawashima[5], K. Hyodo[6], T. Yuasa[2] and T. Zeniya[1] ([1]Hirosaki Univ., [2]Yamagata Univ., [3]Nagoya Univ., [4]Akita Pref. Univ., [5]Kyoto Pharm. Univ., [6]KEK-IMSS-PF)