Geant4-Based Simulation and Image Reconstruction of Tc-99m for Radiopharmaceutical Imaging

Ehsan Shakeri
Department of Mathematics and Statistics
University of Maryland, Baltimore County
Baltimore, MD, USA
ehsans1@umbc.edu

Matthias K. Gobbert
Department of Mathematics and Statistics
University of Maryland, Baltimore County
Baltimore, MD, USA
gobbert@umbc.edu

Vijay R. Sharma
Department of Radiation Oncology
University of Maryland School of Medicine
Baltimore, MD, USA
vsharma@som.umaryland.edu

Lei Ren
Department of Radiation Oncology
University of Maryland School of Medicine
Baltimore, MD, USA
lren@som.umaryland.edu

Ananta Chalise
Department of Radiation Oncology
University of Maryland School of Medicine
Baltimore, MD, USA
ananta.chalise@umm.edu

Stephen W. Peterson
Department of Physics
University of Cape Town
Rondebosch, South Africa
steve.peterson@uct.ac.za

Jerimy C. Polf
M3D Inc
Ann Arbor, MI, USA
jpolf@m3dimaging.com

***Abstract*—This paper presents a framework for simulating and reconstructing radiopharmaceutical images using a Compton camera (CC) for potential use in personalized dosimetry. We developed a Geant4-based Monte Carlo simulation model of a two-stage CC and a Technetium-99m (Tc-99m) source, simulating gamma-ray emissions and their interactions across multiple camera orientations. This synthetic data was used to evaluate two 3D image reconstruction algorithms: simple back projection (SBP) and a kernel weighted back projection (KWBP) method. While KWBP demonstrated superiority over SBP by producing reconstructions with significantly less noise and better-preserved source geometry, both methods were limited by artifacts and blurring due to restricted angular sampling. To address these limitations, we integrated a deep learning-based image enhancement step into the reconstruction pipeline. We trained and evaluated two convolutional neural networks, a RED-CNN and a U-Net, to denoise the reconstructed images. Our results show that both networks effectively suppressed noise, but the U-Net, trained with a hybrid mean squared error (MSE) and structural similarity index measure (SSIM) loss function, delivered superior performance. Quantitative analysis on a held-out test set showed that the U-Net achieved a lower average MSE of 0.0041, compared to the RED-CNN's 0.0103. Furthermore, U-Net qualitatively outperformed the RED-CNN by more accurately preserving the structural integrity and shape of the sources. These findings establish a two-step strategy—combining physics-based reconstruction with data-driven denoising—as a promising pathway for refining Compton camera images for clinical applications.**

***Index Terms*—Compton camera, image reconstruction, radiopharmaceutical imaging, deep learning, gamma-ray**

I. INTRODUCTION

Radiopharmaceutical therapy is an emerging cancer treatment modality that leverages radioactive isotopes to deliver targeted radiation to malignant cells while minimizing damage to surrounding healthy tissues. This therapy is widely used for treating neuroendocrine tumors, prostate cancer, and other malignancies, with growing significance in personalized medicine due to the increasing demand for patient-specific organ dosimetry. However, a key challenge in Radiopharmaceutical therapy is achieving precise dosimetry, which is essential for optimizing treatment efficacy and minimizing adverse effects.

Currently, dosimetry protocols rely on multiple imaging sessions using single-photon emission computed tomography (SPECT) and positron emission tomography (PET) scanners to assess the distribution of radiopharmaceuticals such as Technetium-99m (Tc-99m) and Gallium-67 (Ga-67) over time. This process is resource-intensive, requiring access to expensive scanners and forcing patients to undergo repeated imaging sessions over several days. These logistical challenges have limited the widespread adoption of accurate personalized dosimetry in clinical settings.

To address this limitation, M3D, Inc. is developing a compact, cost-effective imaging system that can be placed at the patient bedside to capture real-time images of radiopharmaceutical distribution. This system aims to replace the need for full SPECT scans by acquiring multiple 2D images from different angles, which are then reconstructed into a 3D representation of the tracer uptake in the body. However, due to limitations in data acquisition and imaging system resolution, these reconstructed images often contain noise and artifacts. A DL model can be integrated into the reconstruction pipeline

DOI 10.1109/ICDMW69685.2025.00242

to clean and refine the images, ensuring that the final 3D representation provides an accurate quantitative analysis of radiopharmaceutical uptake.

This research builds on experience gained from a Geant4 model of the Polaris-J CC (H3D Inc., Ann Arbor, MI) used for prompt gamma imaging [5], where images from multiple camera positions were reconstructed and then improved using machine learning techniques. In this work, we adapt Monte Carlo simulation methods to model the imaging system, define a radiotracer source within a patient phantom, and reconstruct high-quality images that can be used for dosimetry analysis. To the best of our knowledge, this work represents the first demonstration (using Geant4) of a Compton camera's capability for radiopharmaceutical imaging in a therapy context.
The main contributions of this work include:

- **Monte Carlo simulation of a Compton camera system:** We developed a Geant4-based simulation model of a two-stage Compton camera and a radiopharmaceutical source. The model simulates gamma emissions from a Tc-99m source, models their interactions with the proposed camera design (including scatter and absorber detector stages), and records the resulting energy deposition events.
- **Implementation of 3D image reconstruction algorithms:** We implemented and evaluated Compton image reconstruction algorithms (SBP and a KWBP method) to produce 3D source images of the gamma emissions. We assessed the quality and accuracy of the reconstructed images against the known ground truth distribution.
- **Deep learning-based image enhancement:** We trained RED-CNN and U-Net models to suppress noise and artifacts in the reconstructed images. The denoised outputs were quantitatively evaluated against Monte Carlo ground truth to confirm improvements in accuracy for dosimetry analysis.
- **Framework for future experimental validation:** We refined the simulation and reconstruction pipeline to align with the expected design of the M3D imaging system, laying the groundwork for transitioning from simulation studies to experimental data collection in the future.

The rest of the paper is organized as follows. Section II provides background on Compton camera imaging and Tc-99m radiopharmaceuticals. Section III discusses related work. Section IV describes our methodology, including the simulation setup, reconstruction algorithms, and deep learning (DL) models. Section V presents experimental results. Section VI offers discussion and future directions. Finally, Section VII concludes the paper.

II. BACKGROUND

A. Compton Camera Principles

A schematic of a conventional gamma camera that relies on mechanical collimators is shown in Figure 1. In contrast, a Compton camera (CC) is a gamma-ray imaging device that exploits Compton scattering to determine the direction of incoming gamma rays. Unlike traditional gamma cameras that suffer from low detection efficiency due to the use of collimators, CCs offer significantly higher detection efficiency and the potential for improved spatial resolution. The underlying principle of CC operation is Compton scattering, a process in which an incident gamma-ray interacts with an electron in the detector material, resulting in a loss of energy and a change in the gamma-ray's trajectory [1].

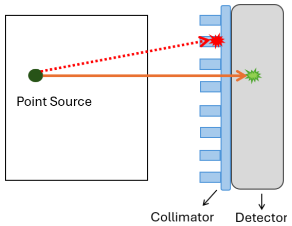


Fig. 1: Illustration of collimation principles: A point source emits rays, filtered by a collimator, with the straight ray reaching the detector.

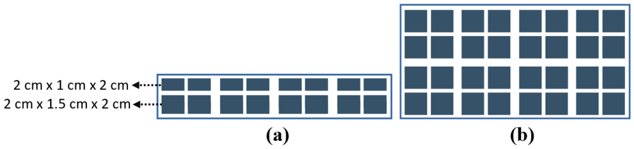
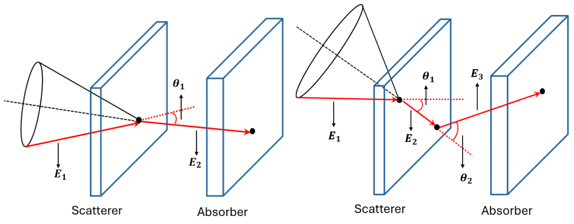


Fig. 2: Views of the Polaris-J CC: (a) Side view, (b) Bottom view.

The Compton scattering process is governed by the well-known Compton equation, which relates the energies of the incident and scattered photons to the scattering angle [2]:

$$\cos\theta = 1 + \frac{m_e c^2}{E_2} - \frac{m_e c^2}{E_1}, \tag{1}$$

where E_1 is the energy of the incoming gamma-ray, E_2 is the remaining energy after the scatter (so that $\Delta E_1 = E_1 - E_2$ is deposited in the first detector), $m_e c^2 \approx 511$ keV is the electron rest energy, and θ is the Compton scatter angle.



(a) (b)

Fig. 3: Schematic illustration of a two-stage CC. (a) A double-scatter event showing the Compton scatter angle θ_1 and associated energies E_1 and E_2. (b) A triple-scatter event with scatter angles θ_1 and θ_2, and corresponding energies E_1, E_2, and E_3. The diagrams also depict the projected Compton cones used in image reconstruction.

The Polaris-J CC is a two-stage system. Figures 2a and 2b illustrate the side and bottom views of the Polaris-J CC, respectively. In a two-stage CC, the first stage (scatter detector) measures the position and energy of an initial Compton interaction, and the second stage (absorber detector) measures the position and energy of the subsequent interaction. By combining the information from both detectors, the scattering angle θ can be calculated and the possible direction of the incident gamma-ray can be constrained to a surface of a cone (see Figure 3a and 3b). This surface is known as the *Compton cone* [3].

A key advantage of CCs over conventional gamma cameras is their higher detection efficiency. Collimators used in conventional cameras block most of the emitted photons to achieve directional information, greatly reducing sensitivity (illustrated in Figure 1). CCs, on the other hand, require no mechanical collimation and thus can detect a much larger fraction of emitted photons [4]. By accurately measuring the interaction positions and energies in both stages and applying the Compton kinematic relationships, the incident photon's origin direction can be reconstructed.

The use of CCs in radiopharmaceutical imaging offers significant potential for enhancing diagnostic accuracy and improving patient outcomes. By providing higher detection efficiency and improved spatial resolution, CCs can facilitate more precise localization of radiotracers within the body, leading to more accurate dosimetry and more effective treatment planning.

Monte Carlo simulations are a valuable tool for evaluating CC performance in various scenarios. In this study, we implemented a two-stage CC design in Geant4 to simulate photon interactions and generate synthetic projection data for image reconstruction. Panthi et al. [5] previously developed a Geant4 Monte Carlo model of the Polaris-J CC for proton therapy range verification. We adapted and modified their model to meet the specific requirements of radiopharmaceutical imaging.

B. Radiopharmaceutical Imaging with Tc-99m

Radiopharmaceuticals are radioactive compounds used for diagnostic and therapeutic purposes in nuclear medicine. These agents consist of a radioisotope (which emits detectable radiation) attached to a pharmaceutical that targets specific biological processes or organs. In imaging, the emitted radiation is detected by gamma cameras or SPECT systems to visualize the distribution of the radiopharmaceutical in the body.

Tc-99m is the most widely utilized radioisotope in diagnostic nuclear medicine due to its favorable properties. It has a short half-life of about 6 hours (minimizing patient radiation exposure) and emits a 140.5 keV gamma photon suitable for imaging with standard gamma cameras [6]. The chemistry of Tc-99m allows it to be bound to various carrier molecules, enabling imaging of a wide range of physiological processes. For example, Tc-99m MDP (methylene diphosphonate) is commonly used for bone scans to highlight areas of high bone metabolism, and Tc-99m sestamibi is used for myocardial perfusion imaging to assess cardiac blood flow [7].

In our simulations, we model an isotropic Tc-99m source emitting 140 keV gamma-rays to mimic a typical radiopharmaceutical distribution. This involves accurately representing the photon energy spectrum and spatial distribution of the source within a phantom. We use Geant4 to model the decay of Tc-99m and the subsequent gamma emission, and we place the source within a defined volume (e.g., a 3 cm radius sphere) corresponding to the region of interest (such as a tumor).

III. RELATED WORK

Compton cameras have been investigated in various medical imaging contexts. Hirasawa et al. [3] developed a compact two-stage semiconductor Compton camera and demonstrated its potential for gamma-ray imaging, highlighting improved efficiency over collimated systems. Polaris-series Compton cameras (like Polaris-J) have been explored for prompt gamma imaging in proton therapy; for example, Panthi et al. [5] simulated a Polaris-J based CC for verifying proton range via prompt gamma detection. Wu et al. [8] proposed a GPU-accelerated 3D reconstruction method for Compton camera-based radioisotope imaging, emphasizing the importance of computational efficiency for iterative algorithms. Kim and Lee [9] provide a comprehensive review of Compton camera image reconstruction techniques, including recent advances in AI-driven methods. To our knowledge, the use of Compton cameras for directly imaging radiopharmaceutical therapy distributions (as we propose) has not been previously reported, making our study novel in this domain.

Deep learning has shown great promise for improving image reconstruction and denoising in medical imaging. For low-dose CT, Chen et al. [10] introduced a residual encoder-decoder convolutional neural network (RED-CNN) to reduce noise, achieving significant improvement in image quality. In situations of limited-angle or sparse-view imaging, deep networks have been used to enhance reconstructions: e.g., Jiang et al. [11] demonstrated that a deep CNN can augment under-sampled cone-beam computed tomography (CT) images, and Jiang et al. [12] developed a multi-scale U-Net to improve digital tomosynthesis for lung imaging. Lang et al. [13] employed a hybrid DL approach (combining a transformer and 3D U-Net) for protoacoustic imaging in proton therapy, achieving accurate real-time reconstruction. In the context of Compton imaging, Polf et al. [14] applied neural networks to filter and reorder Compton event data for prompt gamma imaging, illustrating the potential of machine learning to improve CC data quality. Our approach builds on this prior work by applying DL at the image reconstruction stage to specifically address noise and artifacts in Compton camera images.

IV. METHODOLOGY

A. Simulation and Data Generation

We used Geant4 (version 10.7) [15] to simulate the proposed Compton camera and radiotracer source. The detector geometry replicates the Polaris-J design, consisting of a scatter detector followed by an absorber detector. Each detector stage is an array of pixellated cadmium zinc telluride (CZT) crystals (eight modules per stage, as shown in Figure 2a and 2b, each module containing four CZT blocks of dimensions 2.0 cm × 1.0 cm × 2.0 cm in Stage 1 and 2.0 cm × 1.5 cm × 2.0 cm in Stage 2). The centers of the two detector stages are separated by 7.5 cm. A spherical tumor phantom of radius 3 cm is placed at the origin, representing the radiopharmaceutical distribution.

The radioactive source is configured to emit 140.5 keV gamma photons isotropically within the spherical volume, corresponding to Tc-99m. Standard electromagnetic and radioactive decay physics models are employed in Geant4 to handle gamma-ray interactions (Compton scattering, photoelectric absorption) and particle transport. We simulate 10^8 photons for each acquisition. As photons travel through the phantom and detectors, we record every Compton scatter event that deposits energy in both the scatter and absorber detectors. For each detected event, the positions (x_1, y_1, z_1) and (x_2, y_2, z_2) and energies E_1, E_2 of the first and second interactions are logged (for triple-interaction events, a third interaction is recorded as well).

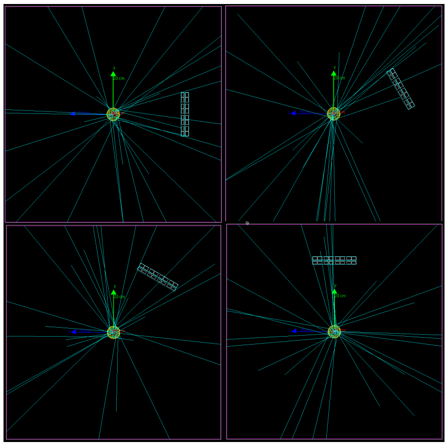


Fig. 4: Geant4 simulation geometries showing the CC at four of twelve orientations (0°, 30°, 60°, 90°) about the x-axis, 30 cm from the Stage 1 detector to the source. Thirty gamma-ray trajectories (yellow) are shown for illustration.

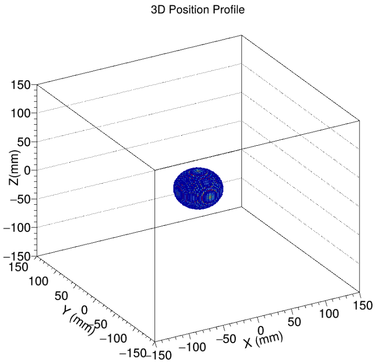


Fig. 5: Isotropic gamma source position profile.

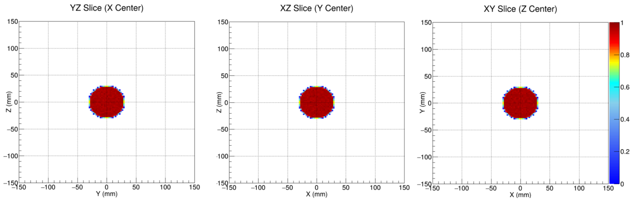


Fig. 6: Central 2D slices of the 3D gamma source distribution along the YZ, XZ, and XY planes. These slices reveal the uniform spherical structure of the source.

To emulate a tomographic scan, the camera is virtually rotated around the source in 12 orientations, each separated by 30°, covering a full 360° about the x-axis. At each orientation, a 10^8-photon simulation is performed and the detected events are recorded. Figure 4 illustrates the geometry for four representative angles. The data from all 12 positions are then combined into a single event list, serving as input to the reconstruction algorithms. The ground-truth isotropic source distribution is shown in Figure 5, with corresponding central slices (YZ, XZ, XY planes) in Figure 6.

B. Image Reconstruction Algorithms

The recorded Compton events, combined from all orientations, are processed to reconstruct a 3D image of the source activity. We implemented two reconstruction methods: simple back-projection (SBP) and a kernel weighted back projection (KWBP) algorithm.

1) Simple Back-Projection (SBP): In Compton imaging, each detected event defines a *Compton cone*, the surface of possible source locations consistent with that event's geometry. The SBP method back-projects all such cones into a 3D voxel volume and identifies regions of high overlap as likely source locations. Specifically, for each event with interactions at (x_1, y_1, z_1) and (x_2, y_2, z_2) and measured energies E_1, E_2, we compute the Compton scatter angle θ via (1). The cone axis is taken along the line $\vec{A} = (x_2 - x_1,\, y_2 - y_1,\, z_2 - z_1)$ connecting the two interaction points. We then increment all voxels lying on the surface of this cone (apex at (x_1, y_1, z_1), opening half-angle θ) in a 3D grid of $128 \times 128 \times 128$ voxels covering a $15 \times 15 \times 15$ cm^3 volume around the source. After processing all events, areas with a higher number of cone overlaps correspond to higher estimated source intensity.

While SBP is straightforward and fast, it produces images with significant blurring, low contrast, and high background noise, especially when only a limited number of projection angles is available. In our simulations, the raw SBP reconstructions were dominated by noise and did not clearly resolve the spherical source, motivating the use of a more sophisticated statistical reconstruction approach.

2) Kernel Weighted Back Projection (KWBP): We implemented a KWBP reconstruction algorithm that utilizes a cone-based kernel model, inspired by approaches in [2] and [5]. In this method, the image is reconstructed by calculating a gamma emission probability for each voxel. This is done by back-projecting all Compton cones into the imaging space, with each voxel contributing to the image based on its distance to the cone surface.

We define a gamma emission probability for voxel $\mathbf{x}$ given a dataset of M Compton cones as:

$$p(\mathbf{x}) = \frac{1}{Mh} \sum_{m=1}^{M} K\Big(\frac{d(\mathbf{x}, c_m)}{h}\Big), \tag{2}$$

where the sum is over the M Compton cones $(c_1, c_2, ..., c_M)$ from double- or triple-scatter events. $d(\mathbf{x}, c_m)$ is the minimum distance from the voxel $\mathbf{x}$ to the surface of the m-th Compton cone, which is also referred to as the point of closest approach (PCA). $K(z)$ is an Epanechnikov kernel defined as $K(z) = 0.75(1 - z^2)$ for $|z| < 1$ and 0 otherwise, and h is a bandwidth parameter chosen to be greater than or equal to the voxel size. This approach assigns higher values to voxels near the cone surface of an event. Finally, noise is suppressed by randomly shuffling all cones, back-projecting them, and subtracting the resulting image from the calculated image. For more information about the KWBP algorithm, readers are referred to [5].

C. Deep Learning Denoising

To further enhance the reconstructed images, we employed DL models to post-process the volume produced by KWBP. Because directly tackling the 3D volume is challenging with limited training data, we focused on 2D cross-sectional slices. Specifically, we extracted the central YZ-plane slice from each 3D reconstruction—which corresponds to a true cross-section of the source—and its corresponding slice from the ground truth activity distribution.

Dataset: We generated a dataset of paired noisy and clean slices by simulating multiple source configurations. We simulated 16 spherical sources with radii ranging from 2.5 mm to 40 mm and 23 ellipsoidal or irregular sources, including elongated shapes, to provide variety. Each case was simulated with 12 camera angles and reconstructed with KWBP as above. This yielded a total of 1180 slice pairs. We extracted multiple slices per volume in most cases. The dataset was split 80%/20% into training and testing sets (944 and 236 slices, respectively) and all images were normalized to $[0, 1]$ intensity.

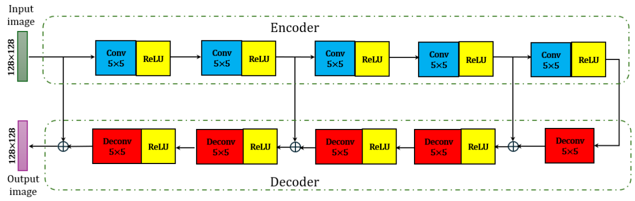


Fig. 7: RED-CNN architecture with symmetric encoder–decoder and residual skips for edge preservation and noise reduction.

Networks: We explored two convolutional neural network architectures for denoising: a RED-CNN [10] and a U-Net [17]. Both are encoder-decoder designs, but they differ in how they implement skip connections to preserve high-frequency details.

The RED-CNN employs a symmetric encoder-decoder structure with *residual skip connections*, as illustrated in Figure 7. The encoder uses five convolutional layers with down-sampling to extract features from the noisy slice. The decoder mirrors this path with five deconvolutional layers to restore resolution. Residual shortcuts add the input of each layer to the output of its corresponding decoder layer, which helps preserve fine details while suppressing noise. We used ReLU activations throughout the network except for the final output layer.

Figure 8 illustrates the U-Net architecture implemented in our work. The network follows an encoder-decoder architecture with *skip concatenations*. It comprises four downsampling levels, with two 3×3 convolutional layers per level. The encoder layers use $\{64, 128, 256, 512\}$ filters and 2×2 max-pooling for down-sampling. The decoder upsamples the feature maps using transposed convolutions and concatenates them with the corresponding high-resolution features from the encoder, enabling precise localization with contextual information. The network uses batch normalization and ReLU activations, and a final 1×1 convolution with a sigmoid activation produces the denoised slice.

1) Loss Functions and Training Dynamics: Let I be the ground-truth slice and $\hat{I}$ the network output, both normalized to $[0, 1]$. The MSE loss is

$$L_{\text{MSE}}(I, \hat{I}) = \frac{1}{N} \sum_{i=1}^{N} \big(I_i - \hat{I}_i\big)^2, \tag{3}$$

which measures the average squared pixel difference. While effective for reducing pixel-wise errors, MSE often produces overly smooth, blurry images by suppressing fine details.

To preserve perceptual quality, we also use the structural similarity index measure (SSIM) loss, defined as $L_{SSIM}(I, \hat{I}) = 1 - SSIM(I, \hat{I})$, where

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \tag{4}$$

Here μ, σ^2, and σ_{xy} denote local means, variances, and covariance, with C_1, C_2 preventing division by zero. SSIM evaluates similarity in luminance, contrast, and structure, aligning more closely with human visual perception.

Our hybrid loss combines pixel-level accuracy and perceptual fidelity:

$$L_{hybrid}(I, \hat{I}) = \alpha \cdot L_{MSE}(I, \hat{I}) + \beta \cdot L_{SSIM}(I, \hat{I}), \tag{5}$$

where constants α and β balance the contributions. This ensures quantitatively accurate yet visually realistic reconstructions, critical in medical imaging.

V. RESULTS

A. Image Reconstruction Results

The KWBP reconstruction of the primary source, a 3 cm-radius sphere (Figure 5), is shown as a 3D isosurface in

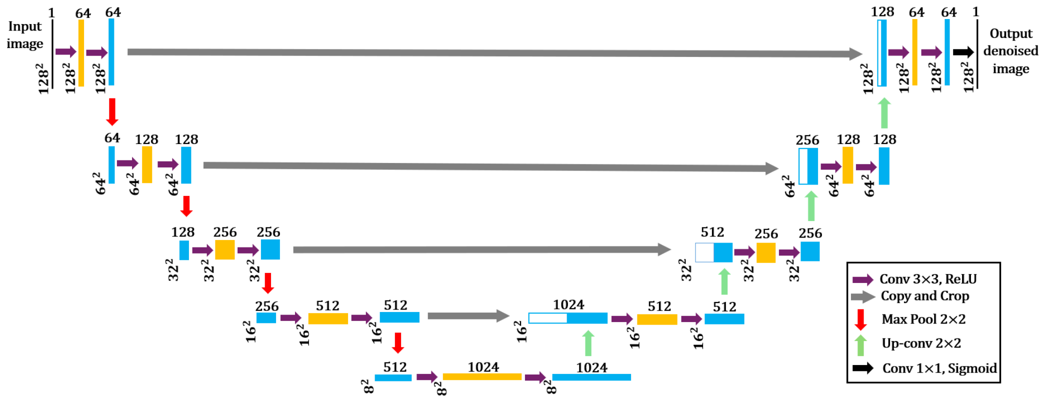


Fig. 8: U-Net architecture with encoder–decoder and skip concatenations for detail preservation during denoising.

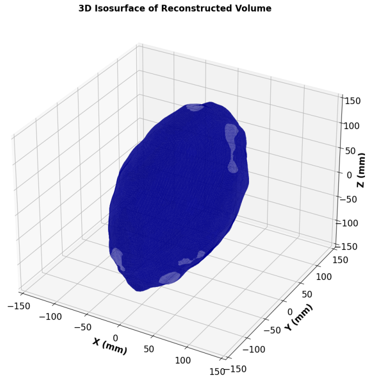


Fig. 9: 3D isosurface visualization of the reconstructed source using the KWBP method. The image was reconstructed on a 15 cm × 15 cm × 15 cm volume discretized into 128^3 voxels.

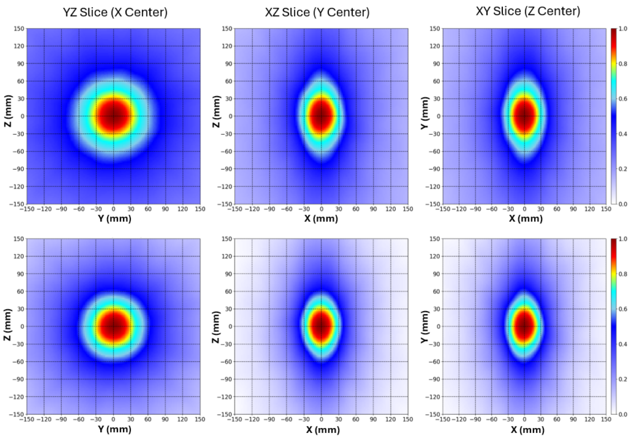


Fig. 10: Central orthogonal slices of the 3D source image reconstructed using SBP (top) and KWBP (bottom). Each row shows the YZ, XZ, and XY planes through the volume.

Figure 9. The reconstruction was performed within a 15.0 cm× 15.0 cm × 15.0 cm imaging volume, discretized into 128 × 128 × 128 voxels, with a kernel bandwidth of 15. Although the source is expected to appear spherical, the reconstruction is blurred and the recovered radius exceeds 3 cm, reflecting the effects of limited angular coverage and noise. Figure 10 compares central slices of the KWBP and SBP reconstructions. In the YZ slice, which corresponds to the true cross-section of the sphere because the camera was rotated about the x-axis, the KWBP result more closely resembles the expected 3 cm circle, whereas the SBP result appears blurred and severely noisy. The XZ and XY slices reveal that both methods exhibit elongation and artifacts caused by the limited 12-view acquisition; however, these distortions are substantially more pronounced in the SBP image. Overall, KWBP better preserves source geometry and provides higher contrast than SBP, yielding a sharper and more localized estimate. Some blurring and artifacts remain due to noise and the limited 12-view angular sampling.

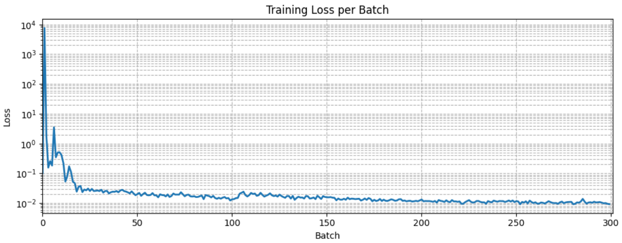


Fig. 11: Training loss curve of RED-CNN over 20 epochs.

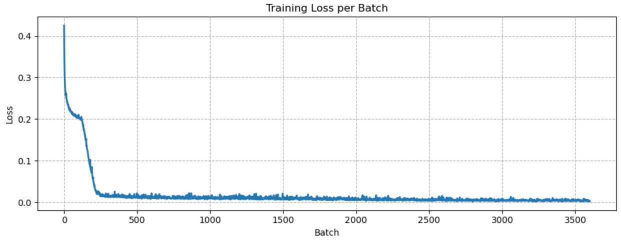


Fig. 12: Training loss curve of U-Net over 120 epochs.

B. Deep Learning Denoising Results

Both RED-CNN and U-Net networks were trained on 944 paired slices. The RED-CNN was optimized with MSE loss using the Adam optimizer at a learning rate of 10^{-3} for 20 epochs with a batch size of 16. The U-Net was trained for 120 epochs with the same optimizer and learning rate, but employed the hybrid loss defined in (5), where $\alpha = 0.8$ and $\beta = 0.2$. SSIM was computed over 11×11 windows. This hybrid formulation encouraged the U-Net to preserve structural details while also minimizing pixel-wise errors.

All experiments were conducted on a workstation equipped with an NVIDIA L40S GPU (48 GB memory). Training the RED-CNN model required approximately 105 seconds for 20 epochs, while the U-Net model required 299 seconds for 120 epochs on the same dataset. The longer runtime of U-Net is attributable both to its deeper encoder–decoder architecture with more trainable parameters and to the greater number of training epochs used for convergence.

Figures 11 and 12 present the training loss curves for RED-CNN and U-Net, respectively. RED-CNN converged within approximately 20 epochs, while U-Net converged by 120 epochs and achieved a lower final error under the hybrid loss.

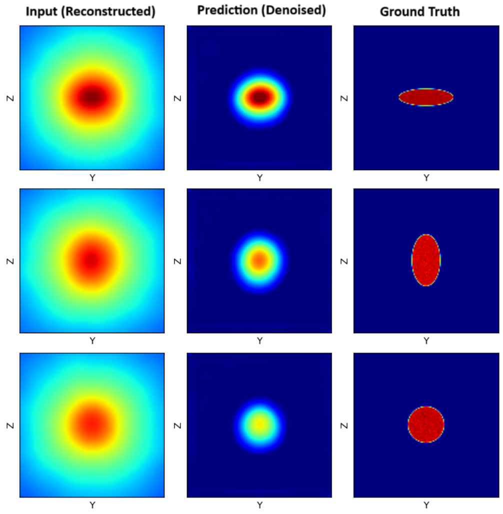


Fig. 13: Example of RED-CNN denoising. Left: Noisy KWBP slice. Middle: RED-CNN output. Right: Ground truth. Noise is largely removed, but some structural elongation is lost, and intensity is slightly underestimated.

The DL models were evaluated on a held-out test set of 236 slices. Quantitatively, the U-Net achieved a lower average MSE (0.0041) compared to RED-CNN (0.0103), indicating improved fidelity to the ground truth. In terms of perceptual quality, both models substantially enhanced the noisy KWBP reconstructions, but U-Net consistently preserved structural details more faithfully. We also evaluated the peak signal-to-noise ratio (PSNR), where RED-CNN achieved 20.95 ± 3.04 dB and U-Net achieved 28.04 ± 4.74 dB. Together, the MSE and PSNR metrics confirm the superior performance of U-Net over RED-CNN under limited-view Compton camera data.

The RED-CNN effectively suppressed the reconstruction noise present in the input slices. Figure 13 shows three representative examples. The denoised outputs are smoother than the noisy inputs and enhance the visibility of the source region. However, RED-CNN tends to over-smooth structural details; for example, in the first slice, the true source is elongated, whereas the output appears more isotropic, indicating a loss of structural accuracy. In addition, the peak activity is underestimated in the second and third slices compared to the ground truth, suggesting that the model reduces intensity as part of the denoising process.

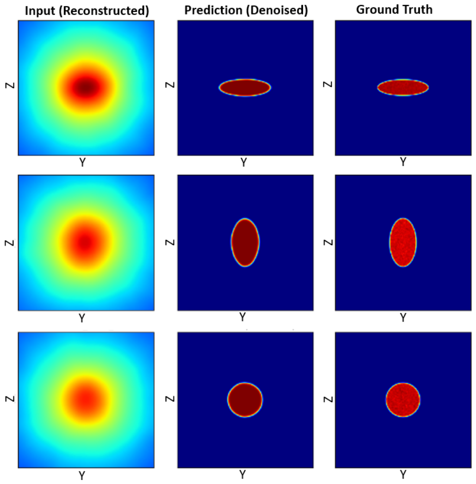


Fig. 14: Example of U-Net denoising with hybrid loss. Left: Noisy slice. Middle: U-Net output. Right: Ground truth. Noise is greatly reduced while the elongated shape and size of the source are preserved.

TABLE I: Per-slice PSNR values (dB) for the three examples shown in Figures 13 and 14.

Method	Slice #0	Slice #1	Slice #2
RED-CNN	17.5	19.6	19.8
U-Net	22.6	25.1	26.3

Figure 14 illustrates the U-Net results on the same slices. Trained with the hybrid loss, U-Net preserved structural details far better than RED-CNN, yielding outputs that closely matched the ground truth in both shape and extent. Across the test set, U-Net consistently maintained source size and geometry, which we attribute to its skip connections and the SSIM component of the loss function. The quantitative PSNR values for the three representative slices are summarized in Table I, where U-Net outperformed RED-CNN by approximately 5–7 dB in each case, consistent with the qualitative improvements observed in Figures 13 and 14.

Overall, DL post-processing substantially improved image quality. The denoised images exhibit higher contrast and clearer source boundaries while retaining the fidelity of the original distribution. Such improvements are critical for clinical translation, as they may enable more reliable quantification of radiotracer uptake in organs and tumors.

VI. Discussion and Future Works

Our results demonstrate the feasibility of using a Compton camera to image radiopharmaceutical distributions and highlight the benefit of integrating data-driven denoising into the reconstruction process. The KWBP algorithm provided a reasonable initial reconstruction of the source distribution from highly noisy projection data, and subsequent deep learning enhancement was able to recover something very close to the ground truth image.

There are several directions for future work. First, while our DL approach was applied slice-by-slice, an obvious extension is to operate in 3D. A 3D CNN or U-Net could take the entire 3D volume as input and potentially learn to enforce consistency across slices, which might further improve artifact reduction. However, training a 3D model would require a larger dataset or patch-based training due to memory constraints. A related idea is to train a network to directly reconstruct the 3D activity distribution from the set of 2D projection images. Emerging techniques like neural radiance fields (NeRF) [18], [19] have shown that neural networks can represent volumetric scenes and synthesize novel views; adapting such approaches to the Compton imaging geometry is an intriguing possibility [20].

The translation of this work to practical applications will require accounting for more realistic conditions. Our simulations were performed under an idealized scenario with a known source geometry embedded in a uniform medium. In reality, patient anatomy is highly heterogeneous and introduces both attenuation and scattering of gamma rays. A compelling extension involves simulating imaging within the anatomical context of a patient's CT scan, following the approach demonstrated by [21]. Geant4 supports the import of DICOM-based CT data as a voxelized phantom [22], enabling evaluation of how a CC could image a radiotracer distribution, for example within a liver tumor, against a realistic anatomical background. In such settings, our reconstruction and denoising pipeline would face additional challenges, including non-uniform attenuation, and would likely require further corrections or retraining of the deep learning models. If successful, this technique could provide a practical, bedside imaging solution for treatment dose verification in radiopharmaceutical therapy, improving patient-specific dosimetry and treatment outcomes.

Another important direction is to perform a systematic sensitivity analysis with respect to the number of projection views. In this study, reconstructions were limited to 12 projection angles; future work will investigate how varying the number of projections influences reconstruction quality and the subsequent performance of deep learning denoising. Such an analysis will help establish the robustness of the proposed pipeline under practical acquisition constraints.

VII. Conclusion

This study successfully demonstrated a comprehensive framework for simulating and reconstructing radiopharmaceutical images using a CC and a Geant4-based model. We developed a pipeline that simulates gamma-ray interactions from a Tc-99m source within a phantom and then reconstructs the 3D source distribution from multi-view acquisitions. The 3D images of the source distribution were reconstructed using the SBP and the KWBP methods. Our comparative analysis of these reconstruction algorithms confirmed that the KWBP algorithm is superior to SBP, as it produced reconstructions with significantly less noise and better-preserved source geometry.

Despite the improvements offered by KWBP, the reconstructed images still contained notable noise and artifacts due to limited angular sampling. To address these limitations, we implemented and evaluated two DL models for post-reconstruction image enhancement: a RED-CNN and a U-Net.

Our results show that both networks effectively suppressed noise, but the U-Net, trained with a hybrid MSE and SSIM loss function, delivered superior performance. Quantitative evaluation on the test set showed that the U-Net achieved a lower average MSE of 0.0041 compared to the RED-CNN's 0.0103, and a higher PSNR of 28.04 ± 4.74 dB compared to 20.95 ± 3.04 dB. Qualitatively, the U-Net outputs more accurately preserved the elongated shapes and structural integrity of the sources, which is critical for dosimetry analysis. Together, the MSE and PSNR results consistently confirmed the superiority of U-Net over RED-CNN, reinforcing the robustness of our proposed two-step reconstruction and denoising pipeline.

The proposed two-step strategy—using physics-based reconstruction followed by deep learning denoising—combined the accuracy of model-driven methods with the noise-reduction capability of data-driven approaches, resulting in clearer and more reliable images of the radiotracer distribution.

Acknowledgments

This work is supported by the grant "REU Site: Online Interdisciplinary Big Data Analytics in Science and Engineering" from the National Science Foundation (grant no. OAC–2348755). Co-authors Sharma and Ren additionally acknowledge support by NIH. We acknowledge the UMBC High Performance Computing Facility and the financial contributions from NIH, NSF, CIRC, and UMBC for this work.

References

[1] G. F. Knoll, *Radiation Detection and Measurement*. John Wiley & Sons, 2010.

[2] S. W. Peterson, D. Robertson, and J. J. Polf, "Optimizing a three-stage CC for measuring prompt gamma rays emitted during proton radiotherapy," *Phys. Med. Biol.*, vol. 55, no. 22, p. 6841, 2010.

[3] M. Hirasawa, S. Takeda, S. Watanabe, and S. Ohsuka, "Development of a compact CC using Si/CdTe semiconductor detectors," *Nucl. Instrum. Methods Phys. Res. A*, vol. 692, pp. 163–167, 2012.

[4] J. F. Ziegler, *The Stopping and Ranges of Ions in Matter*. Pergamon Press, 1999.

[5] R. Panthi, P. Maggi, S. Peterson, D. Mackin, J. Polf, and S. Beddar, "Secondary particle interactions in a Compton camera designed for *in vivo* range verification of proton therapy," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 5, no. 3, pp. 383–391, May 2021.

[6] G. B. Saha, *Physics and Radiobiology of Nuclear Medicine*, 4th ed. New York: Springer, 2010.

[7] S. Vallabhajosula, *Molecular Imaging: Radiopharmaceuticals for PET and SPECT*. New York: Springer, 2009.

[8] R.-Y. Wu, C.-R. Geng, F. Tian, Z.-Y. Yao, C.-H. Gong, H.-N. Han, J.-F. Xu, Y.-S. Xiao, and X.-B. Tang, "GPU-accelerated three-dimensional reconstruction method of the Compton camera and its application in radionuclide imaging," *Nucl. Sci. Tech.*, vol. 34, no. 4, p. 52, 2023.

[9] S. M. Kim and J. S. Lee, "A comprehensive review on Compton camera image reconstruction: From principles to AI innovations," *Biomed. Eng. Lett.*, vol. 14, no. 6, pp. 1175–1193, 2024.

[10] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imaging*, vol. 36, no. 12, pp. 2524–2535, Dec. 2017.

[11] Z. Jiang *et al.*, "Augmentation of CBCT reconstructed from under-sampled projections using deep learning," *IEEE Trans. Med. Imaging*, vol. 38, no. 11, pp. 2705–2715, 2019.

[12] Z. Jiang, F. F. Yin, Y. Ge, and L. Ren, "Enhancing digital tomosynthesis (DTS) for lung radiotherapy guidance using patient-specific deep learning," *Phys. Med. Biol.*, vol. 66, no. 3, p. 035009, 2021.

[13] Y. Lang, Z. Jiang, L. Sun, L. Xiang, and L. Ren, "Hybrid-supervised deep learning for domain transfer 3D protoacoustic image reconstruction," *Phys. Med. Biol.*, vol. 69, no. 8, p. 085007, 2024.

[14] J. C. Polf, C. A. Barajas, S. W. Peterson, D. S. Mackin, S. Beddar, L. Ren, and M. K. Gobbert, "Applications of machine learning to improve the clinical viability of Compton camera based *in vivo* range verification in proton radiotherapy," *Front. Phys.*, vol. 10, p. 838273, 2022.

[15] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, *et al.*, "Geant4—a simulation toolkit," *Nucl. Instrum. Methods Phys. Res. A*, vol. 506, no. 3, pp. 250–303, 2003.

[16] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imaging*, vol. 1, no. 2, pp. 113–122, 1982.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. MICCAI*, 2015, pp. 234–241.

[18] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," *Commun. ACM*, vol. 65, no. 1, pp. 99–106, 2021.

[19] F. Zhu, S. Guo, L. Song, K. Xu, J. Hu, *et al.*, "Deep review and analysis of recent NeRFs," *APSIPA Trans. Signal Inf. Process.*, vol. 12, no. 1, 2023.

[20] X. Wang, S. Hu, H. Fan, H. Zhu, and X. Li, "Neural radiance fields in medical imaging: Challenges and next steps," *arXiv preprint* arXiv:2402.17797, 2024.

[21] V. R. Sharma, Z. Jiang, S. Mossahebi, E. Shakeri, A. Chalise, M. K. Gobbert, S. W. Peterson, J. C. Polf, and L. Ren, "Modeling prompt gamma (PG) emission, detection and imaging in real patient anatomy using a novel Compton camera for dose verification in proton therapy," *Phys. Med. Biol.*, vol. 70, no. 12, p. 125004, 2025.

[22] CERN, "Geant4 DICOM example," [Online]. Available: https://geant4-userdoc.web.cern.ch/Doxygen/examples_doc/html/ExampleDICOM.html