

An Open-Source Benchmark for Modular PET Reconstruction: Establishing a Baseline for STIR and CASToR

S. Mahmoud Mousawi^{1*} and Aram Teymurazyan²

^{1*}Department of Computer Science, University of Regina, 3737 Wascana Parkway, Regina, S4S 0A2, SK, Canada.

²Department of Physics, University of Regina, 3737 Wascana Parkway, Regina, S4S 0A2, SK, Canada.

*Corresponding author(s). E-mail(s): seyyed.mousavi@uregina.ca;
Contributing authors: aram.teymurazyan@uregina.ca;

Abstract

Objective: This study establishes and validates a complete data processing pipeline for comparing open-source PET reconstruction platforms. Using a single Geant4 simulation as a common source of list-mode data, we provide a direct performance comparison between STIR and CASToR, two major software packages, for a modular, small-bore scanner.

Methods: Our Geant4 simulation generated list-mode data by modeling the spatial distribution and energy spectrum of $F-18$ positrons, their annihilation, and the subsequent recording of crystal interaction positions for gamma-ray pairs. This single dataset was used as input for both STIR (v6.2.0) and CASToR (v3.2). We employed the NEMA NU 4-2008 standard as a quantitative blueprint for evaluation, measuring spatial resolution, uniformity, recovery coefficients, and spill-over ratios.

Results: The pipeline successfully generated compatible list-mode data for both STIR and CASToR. Quantitative analysis revealed a consistent trend of STIR achieving higher recovery coefficients than CASToR, though the magnitude of this advantage varied depending on the feature size. In contrast, both platforms showed comparable performance in spatial resolution, uniformity, and spill-over ratios.

Conclusion: We demonstrate a robust methodology for a fair, cross-platform comparison of PET reconstruction software from a unified simulation source. The results confirm that both STIR and CASToR are capable and reliable tools, with STIR showing a distinct trend towards better quantitative recovery. This work

provides a practical guide for researchers selecting software for specialized scanners, highlighting that the optimal choice may depend on the specific quantitative metrics prioritized for their application.

Keywords: Image Reconstruction, NEMA NU 4-2008, Positron Emission Tomography (PET), Scanner, STIR

1 Introduction

Open-source software platforms have become cornerstones of innovation and reproducibility in positron emission tomography (PET) image reconstruction. They differ in their philosophy, purpose, method, and programming language framework [1–7]. However, they provide accessible, transparent, and customizable tools for researchers to develop and validate new algorithms, ultimately accelerating advancements in quantitative imaging. Among the most prominent of these platforms are STIR (Software for Tomographic Image Reconstruction) and CASToR (Customizable and Advanced Software for Tomographic Reconstruction) [1, 2]. Both are powerful, community-driven projects capable of handling complex reconstruction scenarios. There are numerous works that have used STIR or CASToR for their image reconstruction and performance analysis [8–10]. However, for research groups and facilities utilizing custom, modular scanner geometries—such as those designed for plant imaging or specialized preclinical studies [11–15]—a critical question remains: what are the practical performance differences between these two tools when applied to a unified dataset from a non-standard scanner?

A direct, controlled comparison is essential. While both STIR and CASToR are based on well-established statistical reconstruction methods, they differ significantly in their underlying architecture, projectors, and user implementation. These differences can lead to variations in quantitative accuracy, spatial resolution, and noise properties that are not well-documented for specialized imaging systems. Without a rigorous baseline comparison, researchers lack the evidence needed to make an informed choice between these platforms for their specific applications, potentially leading to suboptimal image quality or misinterpretation of results.

To address this gap, this work establishes a comprehensive, fair-comparison framework. We present a complete data processing pipeline that generates synthetic list-mode data from a single, unified Geant4 [16–18] simulation of a modular, small-bore scanner, ensuring that any performance differences observed are attributable to the software themselves and not to variations in the input data. This dataset, designed around the NEMA NU 4-2008 standard [19] as a quantitative blueprint, is then reconstructed using both STIR and CASToR, with a focus on key metrics such as spatial resolution, recovery coefficients, and noise uniformity. The use of a Monte Carlo simulation, mostly through GATE [20, 21] or Geant4 [16–18] toolkit and connecting it with a reconstruction software to study the performance of real world scanners as well as the performance of reconstruction software is well established [8, 12, 22–28]. This

approach eliminates the cost, time, and other factors associated with the real world experiments while closely approximating actual real world cases.

The primary goal of this study is to provide a definitive performance benchmark, guiding researchers in selecting and configuring the most appropriate software for their specialized scanners. Furthermore, by establishing this rigorous baseline and making the simulation and analysis pipeline publicly available, this work deliberately constructs the essential foundation for a subsequent, critical research direction: the integration of modern artificial intelligence techniques. The use of deep learning and generative AI has established a new line of research in PET for denoising and enhancing the quality and contrast of the reconstructed images [29–35]. The controlled dataset and known ground truth provided here are the fundamental prerequisites for developing and validating deep learning-based super-resolution and denoising networks, marking a purposeful first step toward enhancing the capabilities of modular PET imaging through AI.

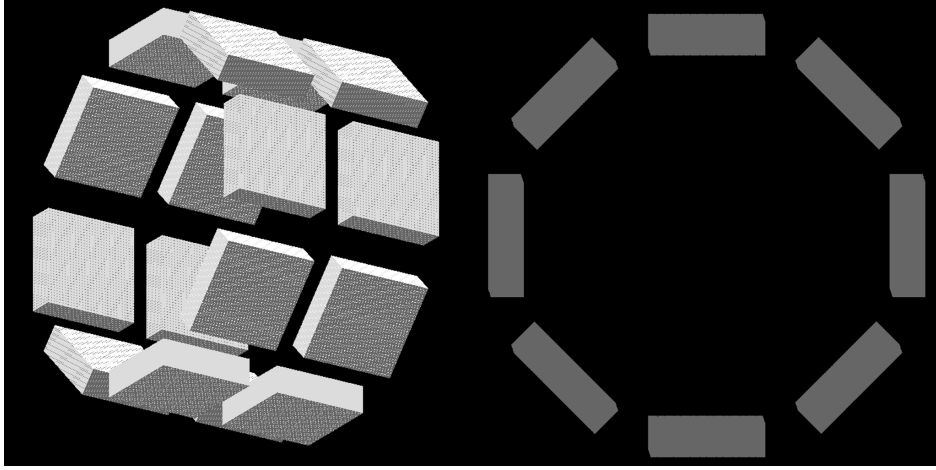
2 Setup and Methods

Our custom-built scanner consists of 16 detector modules arranged in two modules axially and eight modules transaxially forming an octagonal prism (see Fig. 1). The inner radius is 73 mm. Each module is an array of 35×35 pixelated LYSO scintillators. Each crystal is of dimensionality $1.2 \times 1.2 \times 13 \text{ mm}^3$ and there is a gap of 0.1 mm between each two crystals within a module both axially and transaxially which is filled with G4_TEFロン. The distance between the centers of two detector modules is 60 mm in axial direction. The middle ring gap between modules is filled with Tungsten (G4_W) and any photon entering it is killed in order to ensure the photon detection only occurs at the surface of each crystal and not by its side. To put them in other words, the scanner comprises 70 rings of crystals such that each ring consists of 8 modules equally distributed around the ring and each module containing 35 crystals with a gap of 0.1 mm between those crystals. Axially, all rings are equally distributed, 0.1 mm, except that there is a gap of 14.6 mm between the ring number 35th and the ring number 36. The shape of the scanner is provided in Fig. 1.

2.1 Data

The data was obtained by simulation using Geant4 v10.4.7 [16–18]. We simulated NEMA NU 4-2008 image quality phantom [19] (depicted in Fig. 2) as a quantitative blueprint to measure uniformity, recovery coefficients, and spill-over ratios. We also simulated four source points at distances 5, 10, 15, 25 mm from the center of the scanner for the spatial resolution measurement as described in [19, Section 3]. Our simulation generated list-mode data by modeling the spatial distribution and energy spectrum of $F-18$ positrons [36], their annihilation, and the subsequent recording of crystal interaction positions for gamma-ray pairs. Particle gun were used to instantiate the positrons. For simplicity and faster simulation, we also filled the world volume with vacuum, i.e. *G4-Galactic*. Due to the nature of our simulation, we minimized the effect of random and scattered events. As a result, the effect of those events are

Fig. 1 Geometry of the Scanner



negligible and ignored. The list-mode data was then fed to the reconstruction software as described in the following subsection.

Fig. 2 Image Quality phantom



2.2 The Reconstruction Software

In this study, we used CASToR (v3.2) [2] and STIR (v6.2.0) [1]. Both software were fed the same list-mode data. In STIR, the list mode data was first converted to projection data using *lm_to_projdata* from the official STIR library. The list mode data for STIR was prepared in SAFIR format [23, 24, 37]. On the other hand, the list mode

data was fed to CASToR directly with the appropriate header file. The geometry of the scanner was described to STIR using a crystal map. The crystal map contains crystal coordinates along with the id of their corresponding ring and layer. The names of the crystal map, the list mode data, and a template for the projection metadata along with an OPTIONAL value for the LOR randomization sigma value are listed in a parameter file for CListModeDataSAFIR class. Then this parameter file is fed to *lm_to_projdata* using another parameter file for *lm_to_projdata*. On the other hand, we used a binary Look-Up-Table (LUT) along with its header file to describe the geometry of the scanner to CASToR. For more details and other ways to describe the geometry of scanners to each software, consult each software’s documentation.

In STIR, the images are reconstructed using *OSMAPOS*, which is an implementation of OSEM-One Step Late algorithm, on projection data. However, in our parameter file, we assume that there is no prior to make *OSMAPOS* more close to OSEM to be a better match for CASToR algorithm. In CASToR, *OSEM* algorithm is obtained by using *MLEM* with subsets. In other words, “-opti *MLEM* -it 6:4” actually means that CASToR is using OSEM with 6 iterations and 4 subsets. In STIR, we need to specify the number of subiterations instead of the number of iterations. In this work, the number of subiterations in STIR is set to the number of iterations in CASToR multiplied by the number of subsets.

The output images of both software were of the same dimension $241 \times 241 \times 160$ where each voxel size is as $0.25 \text{ mm} \times 0.25 \text{ mm}$ transaxially and 0.65 mm in axial direction. This choice of voxel size is an arbitrary trade off between the accuracy and reconstruction time. However, odd numbers for transaxial FOV dimensions are chosen to make the centre of FOV to be in the middle of the central pixel. In addition, we did not apply any post filters to the images. The sensitivity images were calculated by their respective software.

To correct for the PSF effect, in CASToR we can pass a filter to the software using the flag “-conv filter:filter.conf::psf” or “-conv filter, x_1,x_2,x_3 ::psf” where filter can be Gaussian or any other appropriate filter depending on the scenario, x_1 is the transaxial FWHM, x_2 is the axial FWHM, and x_3 is the number of sigmas included in the kernel (recommendations: at least 3 and maximum 5). Unlike CASToR, STIR does not implement a single projector for PSF correction. Instead, we need to apply image-based filters by using separate projectors to pre-process in forward projection and post-process in backward projection.

3 Results

To compare the two image reconstruction software, we measured factors like spatial resolution, uniformity, recovery coefficients (RCs), and spillover ratio (SOR) as described in NEMA NU 4-2008 standard [19]. The results are reported in the following subsections.

3.1 Spatial Resolution

As mentioned in NEMA NU 4-2008 standard [19], this measurement highlights the ability of a reconstruction toolkit to distinguish between two source points after reconstruction. This is done by measuring the FWHM and FWTM of the point source response function in radial and tangential directions by forming one-dimensional response functions through the peak. To obtain the response function, we considered all the one dimensional profiles that are parallel to the direction of the measurement and that lie within at least twice the (roughly estimated) FWHM of the directions perpendicular to the measurement direction. Then, we selected one of the orthogonal directions and sum projected all the profiles (pixel-wise) in this selected direction and then we again sum projected the resulting profiles in the other orthogonal direction. This gave us a profile in the direction of the measurement. To calculate FWHM and FWTM, we employed linear interpolation between the adjacent pixel at half and one tenth of the maximum peak value, respectively. To determine the peak value and location, we used parabolic fit to the pixel with the greatest value and its two immediate neighbors. For example, to calculate FWHM in X direction, we sum projected all the pixels falling between the range in y direction. Then we sum projected all the pixels with the range in z (axial) direction. Having this one dimensional profile, we can calculate peak, FWHM, and FWTM values as described.

As mentioned earlier, we placed the sources at distances 5, 10, 15, and 25 in two planes each perpendicular to the axial axis of the scanner where one at the center of the axial FOV, and the other one at one-fourth of the FOV length from the center of the scanner. Although it is mentioned in [19] to use $Na-22$ as the radionuclide, we used $F-18$ for this measurement. For more details, see [19].

For STIR, we employed the well-known ray tracing matrix projector which uses an optimisation of Siddon's algorithm [38]. To find the best match from CASToR, we employed Incremental Siddon [39], Joseph [40], and Multi Siddon (based on Incremental Siddon) projectors¹ from CASToR denoted as CASToR-I, CASToR-J, and CASToR-M, respectively. Additionally, in this case we did not use any point spread correction as per [19].

Tables 1 and 2 show the number of recorded pairs of crystals for each point source. According to NEMA NU 4-2008 standard, there should be at least 10^5 prompt counts per source. Tables 3 and 4 report the resolution values at different distances for the cases where the point sources are at the center of the axial FOV and where the point sources are at 0.25 of the length of axial FOV. In addition, Tables 5 and 6 report peak location of each point source for the two software and their projectors.

Tables 3 and 4 shows that among the chosen projectors from CASToR, Joseph demonstrates a better spatial resolution and we use it as the single projector for CASToR along with ray tracing for STIR in the next section.

3.2 Image Quality

As we have described in section 2.1, the image quality phantom Fig. 2 was simulated to study the quality of images reconstructed by each software. This phantom aims to

¹See CASToR help.

Table 1 Number of Recorded Events (Center of the Axial FOV)

Location	5 mm	10 mm	15 mm	25 mm
Number of Events	2,874,797	2,514,360	2,315,041	2,152,626

Table 2 Number of Recorded Events (off Center)

Location	5 mm	10 mm	15 mm	25 mm
Num of Events	1,808,530	1,598,365	1,493,480	1,504,390

Table 3 Resolution Values Reported at the Center of the Axial FOV

	CASToR-I							
	5 mm		10 mm		15 mm		25 mm	
	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM
Radial	1.02	1.79	1.02	1.83	0.87	1.83	0.88	1.78
Tangential	0.88	1.75	0.89	1.78	0.91	1.78	1.04	1.8
	CASToR-J							
	5 mm		10 mm		15 mm		25 mm	
	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM
Radial	0.89	1.73	0.94	1.68	0.85	1.71	0.86	1.75
Tangential	0.86	1.7	0.87	1.73	0.89	1.75	0.95	1.71
	CASToR-M							
	5 mm		10 mm		15 mm		25 mm	
	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM
Radial	0.8	1.46	0.93	1.69	0.66	2.2	3.46	4.79
Tangential	0.74	1.44	0.73	1.43	0.74	1.45	0.84	1.9
	STIR							
	5 mm		10 mm		15 mm		25 mm	
	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM
Radial	0.78	1.51	0.83	1.48	0.75	1.57	0.65	1.45
Tangential	0.73	1.47	0.72	1.47	0.76	1.49	0.85	1.61

simulate and capture all details of a total body small animal scan like a small rodent with hot lesions along with uniform hot and some cold regions. Although it is not a perfect modeling of the case due to the complex nature of a small animal scan, it provides a standardized way of measuring the image quality in some particular cases.

Table 4 0.25 the Length of the Scanner from the Center of the Axial FOV

	CASToR-I							
	5 mm		10 mm		15 mm		25 mm	
	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM
Radial	1.04	1.8	1.04	1.85	0.82	1.82	0.85	1.9
Tangential	0.88	1.75	0.88	1.78	0.9	1.77	1.04	1.83
	CASToR-J							
	5 mm		10 mm		15 mm		25 mm	
	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM
Radial	0.96	1.8	1.01	1.77	0.79	1.69	0.84	1.73
Tangential	0.87	1.71	0.87	1.73	0.9	1.75	0.97	1.76
	CASToR-M							
	5 mm		10 mm		15 mm		25 mm	
	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM
Radial	0.74	1.4	0.82	1.6	0.63	3.77	1.8	4.43
Tangential	0.7	1.37	0.65	1.35	0.68	1.36	0.77	1.6
	STIR							
	5 mm		10 mm		15 mm		25 mm	
	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM	FWHM	FWTM
Radial	0.8	1.53	0.84	1.48	0.76	1.6	0.69	1.48
Tangential	0.74	1.48	0.72	1.49	0.77	1.5	0.84	1.59

While we did not employ PSF correction for spatial resolution part, we applied positron range correction for this measurement. The average (mean) positron range in water is reported to be 0.44 mm (using a GATE simulation) [41], 0.46 mm (analytically) [42], 0.48 mm [42, 43], 0.51 mm [44], 0.57 mm (PeneloPET Simulation) [43], and 0.6 mm (theoretically) [45, 46], 0.64 mm (semiempirical estimates) [43]. For simplicity of computation and a good approximation of the *F-18* positron range, we assumed that the positron range for *F-18* in water is 0.5 mm. In CASToR we simply passed it as `-conv gaussian,0.5,0.5,3::psf` to the software. For STIR, we used separable Gaussian model to deconvolve for the positron range in forward and backward projections separately².

Overall, the two reconstructed images are very similar in terms of reconstruction quality. As you can see in Fig. 4 we have considered three regions for each image as follows: the rods region, the uniform middle region, and the region containing the cold air and water areas. For each reconstructed volumetric 3D image, we obtained three

²The implementation in the OSMAPOS� parameter file can be found at <https://github.com/Recursion-Q19/CASTiR>

Table 5 Peak Location in mm at the Center of the FOV

	Radial			
	5 mm	10 mm	15 mm	25 mm
CASToR-I	5.0807	9.9639	14.9585	24.9478
CASToR-J	5.0274	9.9954	14.9371	24.9575
CASToR-M	5.0907	9.9477	15.0125	27.5415
STIR	5.033	9.9944	14.9753	24.9794
	Tangential			
	5 mm	10 mm	15 mm	25 mm
CASToR-I	0.001404	0.002034	0.0006079	0.0007351
CASToR-J	0.001581	0.001331	0.0004623	0.000961
CASToR-M	0.00007753	0.002336	-0.001953	-0.0001656
STIR	-0.0006223	0.003575	-0.007296	0.02025

Table 6 Peak Location in mm at 0.25 off Center of the FOV

	Radial			
	5 mm	10 mm	15 mm	25 mm
CASToR-I	5.1448	9.8721	14.9787	24.9378
CASToR-J	5.0818	9.9357	14.952	24.9351
CASToR-M	5.0367	9.9932	15.002	27.563
STIR	5.0476	9.9823	14.9082	24.9186
	Tangential			
	5 mm	10 mm	15 mm	25 mm
CASToR-I	-0.003091	-0.002149	-0.00007502	0.002656
CASToR-J	-0.002012	-0.002598	-0.0001472	0.001406
CASToR-M	-0.003776	-0.0009253	0.003026	0.006676
STIR	-0.002269	0.002799	-0.007988	0.013

2D images where each of which was obtained by Z-projecting 15 consecutive slices in each region (roughly illustrated on the phantom in Fig. 3). The Z-projection calculates and stores the mean intensity across pixels axially to yield the 2D image. For each 2D image, we measured intensity per pixel and calculated mean square error along with χ^2 by drawing a diagonal line through the images and obtaining the profiles. We used the same line to obtain the profiles for all 2D images. This way, we can compare the reconstructed images by each software against one another. Fig. 5 shows the profiles of the three selected regions for the images reconstructed by each software.

In subsections below, we report uniformity, recovery coefficients (RCs), and spill-over ratio (SOR).

Fig. 3 Illustrating 3 main regions. Region A: Rod region, Region B: Uniformity region, Region C: Air and water region. Regions C.1 and C.2 are used to calculate Spill-Over Ratio (SOR). Visualized using DAWN.

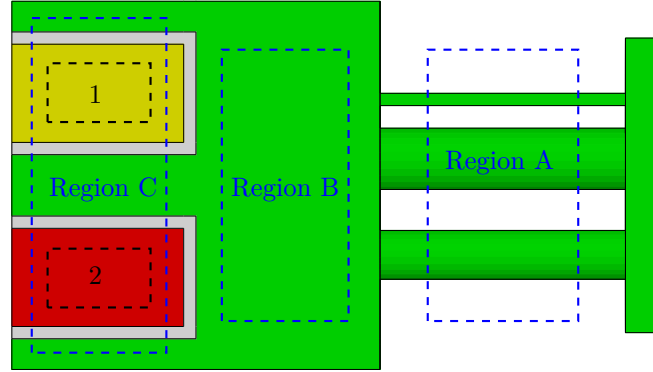
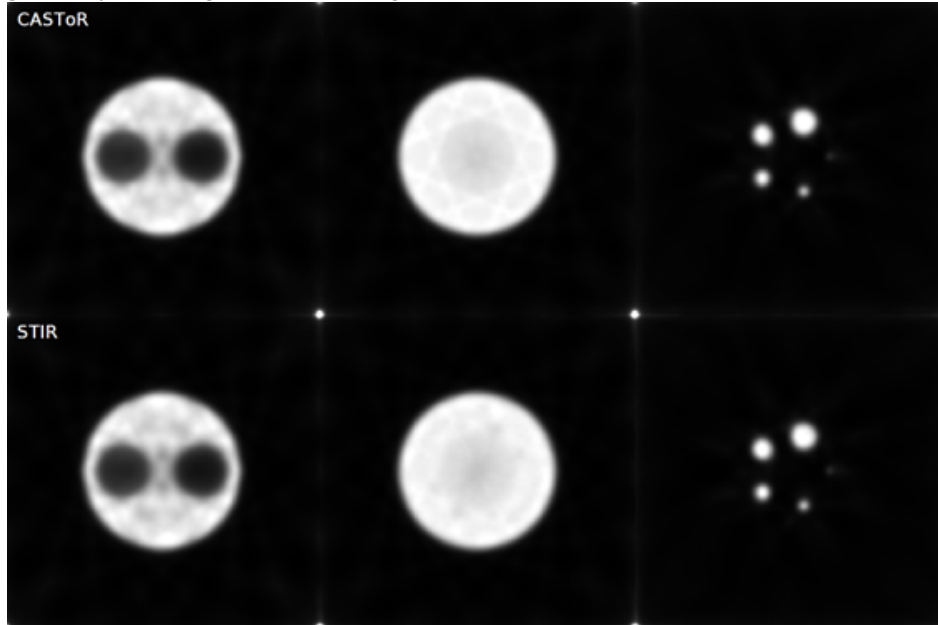


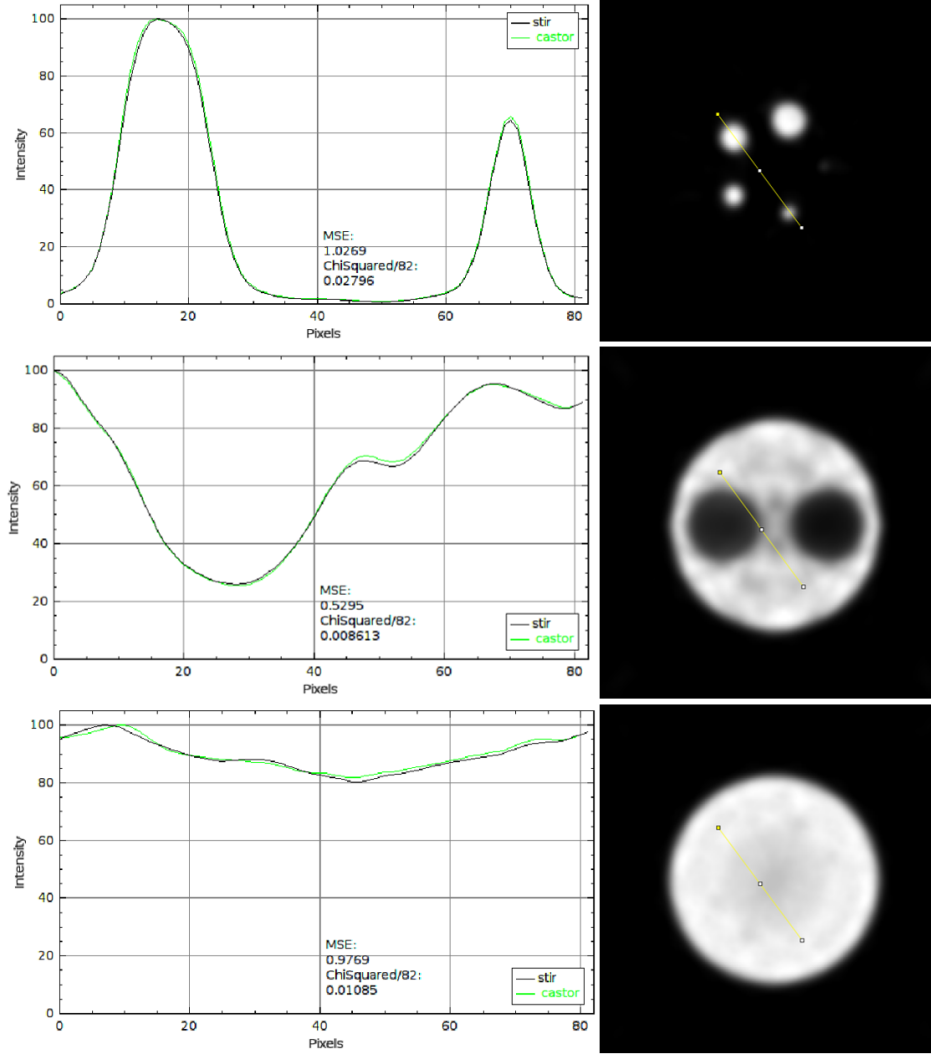
Fig. 4 Projected Images at Different Regions From CASToR and STIR Reconstructions



3.2.1 Uniformity

As per [19], in this section we report the mean, maximum, and minimum pixel intensities for a cylindrical volume of diameter 22.5 mm and length of 10 mm drawn over the center of the uniform active region of the phantom (Region B in Fig. 3).

Fig. 5 Profiles at Different Regions From CASToR and STIR Reconstructions



The mean value here is used to calculate the recovery coefficients and spill-over ratio in the subsequent subsections. The measurement results for these metrics are presented in Table 7.

3.2.2 Recovery Coefficients

In this section, we report the recovery coefficients with their corresponding percentage STDs for the five rods. To calculate this, we considered the slices covering the central 10 mm of the rods region (Region A in Fig. 3). We then averaged all of these slices to obtain a single image of lower noise. See Table 8 for the results of this measurement. To compute these results, we first found the coordinates of the pixel of maximum

Table 7 Report for Uniformity

	Average	Maximum	Minimum	% <i>STD</i>
CASToR-J	0.1802	0.2345	0.1318	1.1175
STIR	0.2702	0.3316	0.2002	1.1749

value for each rod by considering a circle of radius twice the rod’s actual radius in the image. Next, we drew line profiles at recorded coordinates through the five rods across the axial direction. We then find the mean across each line profile and finally divide these mean values for each rod by the mean activity of the uniform region calculated in the previous subsection to achieve the recovery coefficients.

We used the following formula to calculate the standard deviation of the recovery coefficients:

$$\%STD_{RC} = 100 \times \sqrt{\left(\frac{STD_{lineprofile}}{mean_{lineprofile}}\right)^2 + \left(\frac{STD_{uniform}}{mean_{uniform}}\right)^2} \quad (1)$$

Table 8 Report for Recovery Coefficients

	1mm	% <i>STD</i>	2mm	% <i>STD</i>	3mm	% <i>STD</i>	4mm	% <i>STD</i>	5mm	% <i>STD</i>
CASToR-J	0.2782	18.7783	0.8184	5.3679	0.7659	4.4299	0.7254	6.2648	0.71	7.0724
STIR	0.281	17.9745	0.9006	4.3881	0.7627	4.1723	0.7632	7.7279 A	0.791	5.8062

3.2.3 Accuracy of Corrections

The accuracy of correction measurement accounts for the Spill-Over Ratio (SOR) and its associated standard deviation. To calculate the spill-over ratio, we considered two cylindrical volumes, covering the central parts of the two cold regions (water-filled and air-filled cylinders), each of which of diameter 4 *mm* and length 7.5 *mm* (Regions C.1 and C.2 in Fig. 3). The spill-over ratio for each cold region is then computed as the mean intensity in each volume of interest divided by the mean intensity of the background which is calculated in the uniformity section. The standard deviation is calculated in a similar manner to equation 1. Table 9 summarizes the measurement results for accuracy of correction as described in section 6.4.3 of NEMA NU-4 2008 standard [19].

4 Discussion

Our study presents a comprehensive methodological pipeline and comparative analysis of the STIR and CASToR reconstruction software for a custom modular PET scanner. Utilizing a Geant4 simulation to generate positron annihilations and record detector

Table 9 Report for Accuracy of Corrections

Software	Water-filled Cylinder		Air-filled Cylinder	
	SOR	% <i>STD</i>	SOR	% <i>STD</i>
CASToR-J	0.1064	9.1996	0.1938	6.8613
STIR	0.1052	8.5627	0.1916	6.461

interactions, we created a single, consistent source of list-mode data, which was then processed through both software packages. The performance was evaluated using a standardized blueprint from the NEMA NU 4-2008 protocol, quantifying key image quality factors such as spatial resolution, noise uniformity, and the accuracy of activity recovery in features of various sizes. In the spirit of open science, the full source code for the Geant4 simulation, the data conversion pipeline, and the scripts for metric evaluation will be shared via a GitHub repository. The results confirm the general reliability of both tools while identifying a distinct, though variable, advantage for STIR in terms of recovery coefficients.

Having established this performance baseline, our future work will focus on employing deep learning super-resolution networks to enhance the quantitative recovery of features, using the images from this study as the training data. This direct comparison not only guides researchers in software selection but also deliberately constructs a controlled and publicly accessible experimental framework. The framework is specifically designed to validate and benchmark the next generation of AI-driven image enhancement algorithms, marking a deliberate first step in a larger research program to integrate modern computational methods with proven reconstruction techniques.

Acknowledgements. We would like to thank STIR and CASToR communities for the fruitful discussions.

Possible To Do:

- Adding some graphics using Dawn if possible.
- Mentioning that we did not do any normalization, depth of interaction modeling. These could be as the future work.

References

- [1] Thielemans, K., Tsoumpas, C., Mustafovic, S., Beisel, T., Aguiar, P., Dikaïos, N., Jacobson, M.W.: Stir: software for tomographic image reconstruction release 2. *Physics in Medicine & Biology* **57**(4), 867 (2012) <https://doi.org/10.1088/0031-9155/57/4/867>
- [2] Merlin, T., Stute, S., Benoit, D., Bert, J., Carlier, T., Comtat, C., Filipovic, M., Lamare, F., Visvikis, D.: Castor: A generic data organization and processing code framework for multi-modal and multi-dimensional tomographic reconstruction. *Physics in Medicine and Biology* **63** (2018) <https://doi.org/10.1088/1361-6560/aadac1>

- [3] Polson, L.A., Fedrigo, R., Li, C., Sabouri, M., Dzikunu, O., Ahamed, S., Karakatsanis, N., Kurkowska, S., Sheikhzadeh, P., Esquinas, P., Rahmim, A., Uribe, C.: Pytomography: A python library for medical image reconstruction. *SoftwareX* **29**, 102020 (2025) <https://doi.org/10.1016/j.softx.2024.102020>
- [4] Markiewicz, P.J., Thielemans, K., Schott, J.M., Atkinson, D., Arridge, S.R., Hutton, B.F., Ourselin, S.: Rapid processing of pet list-mode data for efficient uncertainty estimation and data analysis. *Physics in Medicine & Biology* **61**(13), 322 (2016) <https://doi.org/10.1088/0031-9155/61/13/N322>
- [5] Markiewicz, P.J., Ehrhardt, M.J., Erlandsson, K., Noonan, P.J., Barnes, A., Schott, J.M., Atkinson, D., Arridge, S.R., Hutton, B.F., Ourselin, S.: Niftypet: a high-throughput software platform for high quantitative accuracy and precision pet imaging and analysis. *Neuroinformatics* **16**(1), 95–115 (2018)
- [6] Pedemonte, S., Catana, C., Van Leemput, K.: An inference language for imaging. In: Cardoso, M.J., Simpson, I., Arbel, T., Precup, D., Ribbens, A. (eds.) *Bayesian and graphical Models for Biomedical Imaging*, pp. 61–72. Springer, Cham (2014)
- [7] Najmaoui, Y., Chemli, Y., Toussaint, M., Petibon, Y., Marty, B., Fontaine, K., Gallezot, J.-D., Razdevšek, G., Orehar, M., Dhaynaut, M., Guehl, N., Dolenec, R., Pestotnik, R., Johnson, K., Ouyang, J., Normandin, M., Tétrault, M.-A., Lecomte, R., Fakhri, G.E., Marin, T.: Yrt-pet: An open-source gpu-accelerated image reconstruction engine for positron emission tomography. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 1–1 (2025) <https://doi.org/10.1109/TRPMS.2025.3619872>
- [8] Dao, V., Mikhaylova, E., Ahnen, M., Fischer, J., Thielemans, K., Tsoumpas, C.: Evaluation of stir library adapted for pet scanners with non-cylindrical geometry. *Journal of Imaging* **8**(6) (2022) <https://doi.org/10.3390/jimaging8060172>
- [9] Razdevšek, G., Fakhri, G.E., Marin, T., Dolenec, R., Orehar, M., Chemli, Y., Gola, A.G., Gascon, D., Majewski, S., Pestotnik, R.: Flexible and modular pet: Evaluating the potential of tof-doi panel detectors. *Medical Physics* **52**(5), 2845–2860 (2025) <https://doi.org/10.1002/mp.17741> <https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1002/mp.17741>
- [10] Wadhwa, P., Thielemans, K., Efthimiou, N., Wangerin, K., Keat, N., Emond, E., Deller, T., Bertolli, O., Deidda, D., Delso, G., Tohme, M., Jansen, F., Gunn, R.N., Hallett, W., Tsoumpas, C.: Pet image reconstruction using physical and mathematical modelling for time of flight pet-mr scanners in the stir library. *Methods* **185**, 110–119 (2021) <https://doi.org/10.1016/j.ymeth.2020.01.005> . *Methods on simulation in biomedicine*
- [11] Pajak, M.Z., Volgyes, D., Pimlott, S.L., Salvador, C.C., Asensi, A.S., McKeown, C., Waldeck, J., Anderson, K.I.: Nema nu4-2008 performance evaluation of albira: A two-ring small-animal pet system using continuous lyso crystals. *Open*

- [12] Chang, Y., Talebitaher, A., Thompson, K., Papandreou, Z., Teymurazyan, A., Siciliano, S.: Plant-specific modular pet: Data processing with castor and performance evaluation. In: 2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC), pp. 1–3 (2018). IEEE
- [13] Goertzen, A.L., Bao, Q., Bergeron, M., Blankemeyer, E., Blinder, S., Cañadas, M., Chatzioannou, A.F., Dinelle, K., Elhami, E., Jans, H.-S., Lage, E., Lecomte, R., Sossi, V., Surti, S., Tai, Y.-C., Vaquero, J.J., Vicente, E., Williams, D.A., Laforest, R.: Nema nu 4-2008 comparison of preclinical pet imaging systems. *Journal of Nuclear Medicine* **53**(8), 1300–1309 (2012) <https://doi.org/10.2967/jnumed.111.099382> <https://jnm.snmjournals.org/content/53/8/1300.full.pdf>
- [14] Becker, R., Buck, A., Casella, C., Dissertori, G., Fischer, J., Howard, A., Ito, M., Khateri, P., Lustermann, W., Oliver, J.F., *et al.*: The safir experiment: Concept, status and perspectives. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **845**, 648–651 (2017)
- [15] Bebié, P., Becker, R., Commichau, V., Debus, J., Dissertori, G., Djambazov, L., Eleftheriou, A., Fischer, J., Fischer, P., Ito, M., Khateri, P., Lustermann, W., Ritzer, C., Ritzert, M., Röser, U., Tsoumpas, C., Warnock, G., Weber, B., Wyss, M.T., Zagodzinska-Bochenek, A.: Safir-i: Design and performance of a high-rate preclinical pet insert for mri. *Sensors* **21**(21) (2021) <https://doi.org/10.3390/s21217037>
- [16] Agostinelli, S., Allison, J., Amako, K., Apostolakis, J., Araujo, H., Arce, P., Asai, M., Axen, D., Banerjee, S., Barrand, G., Behner, F., Bellagamba, L., Boudreau, J., Broglia, L., Brunengo, A., Burkhardt, H., Chauvie, S., Chuma, J., Chytrcek, R., Cooperman, G., Cosmo, G., Degtyarenko, P., Dell’Acqua, A., Depaola, G., Dietrich, D., Enami, R., Feliciello, A., Ferguson, C., Fesefeldt, H., Folger, G., Foppiano, F., Forti, A., Garelli, S., Giani, S., Giannitrapani, R., Gibin, D., Gómez Cadenas, J.J., González, I., Gracia Abril, G., Greeniaus, G., Greiner, W., Grichine, V., Grossheim, A., Guatelli, S., Gumplinger, P., Hamatsu, R., Hashimoto, K., Hasui, H., Heikkinen, A., Howard, A., Ivanchenko, V., Johnson, A., Jones, F.W., Kallenbach, J., Kanaya, N., Kawabata, M., Kawabata, Y., Kawaguti, M., Kelner, S., Kent, P., Kimura, A., Kodama, T., Kokoulin, R., Kossov, M., Kurashige, H., Lamanna, E., Lampén, T., Lara, V., Lefebvre, V., Lei, F., Liendl, M., Lockman, W., Longo, F., Magni, S., Maire, M., Medernach, E., Minamimoto, K., Mora de Freitas, P., Morita, Y., Murakami, K., Nagamatsu, M., Nartallo, R., Nieminen, P., Nishimura, T., Ohtsubo, K., Okamura, M., O’Neale, S., Oohata, Y., Paech, K., Perl, J., Pfeiffer, A., Pia, M.G., Ranjard, F., Rybin, A., Sadilov, S., Di Salvo, E., Santin, G., Sasaki, T., Savvas, N., Sawada, Y., Scherer, S., Sei, S., Sirotenko, V., Smith, D., Starkov,

- N., Stoecker, H., Sulkimo, J., Takahata, M., Tanaka, S., Tcherniaev, E., Safai Tehrani, E., Tropeano, M., Truscott, P., Uno, H., Urban, L., Urban, P., Verderi, M., Walkden, A., Wander, W., Weber, H., Wellisch, J.P., Wenaus, T., Williams, D.C., Wright, D., Yamada, T., Yoshida, H., Zschesche, D.: Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506**(3), 250–303 (2003) [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8)
- [17] Allison, J., Amako, K., Apostolakis, J., Araujo, H., Arce Dubois, P., Asai, M., Barrand, G., Capra, R., Chauvie, S., Chytrcek, R., Cirrone, G.A.P., Cooperman, G., Cosmo, G., Cuttone, G., Daquino, G.G., Donszelmann, M., Dressel, M., Folger, G., Foppiano, F., Generowicz, J., Grichine, V., Guatelli, S., Gumplinger, P., Heikkinen, A., Hrivnacova, I., Howard, A., Incerti, S., Ivanchenko, V., Johnson, T., Jones, F., Koi, T., Kokoulin, R., Kossov, M., Kurashige, H., Lara, V., Larsson, S., Lei, F., Link, O., Longo, F., Maire, M., Mantero, A., Mascialino, B., McLaren, I., Mendez Lorenzo, P., Minamimoto, K., Murakami, K., Nieminen, P., Pandola, L., Parlati, S., Peralta, L., Perl, J., Pfeiffer, A., Pia, M.G., Ribon, A., Rodrigues, P., Russo, G., Sadilov, S., Santin, G., Sasaki, T., Smith, D., Starkov, N., Tanaka, S., Tcherniaev, E., Tome, B., Trindade, A., Truscott, P., Urban, L., Verderi, M., Walkden, A., Wellisch, J.P., Williams, D.C., Wright, D., Yoshida, H.: Geant4 developments and applications. *IEEE Transactions on Nuclear Science* **53**(1), 270–278 (2006) <https://doi.org/10.1109/TNS.2006.869826>
- [18] Allison, J., Amako, K., Apostolakis, J., Arce, P., Asai, M., Aso, T., Bagli, E., Bagulya, A., Banerjee, S., Barrand, G., Beck, B.R., Bogdanov, A.G., Brandt, D., Brown, J.M.C., Burkhardt, H., Canal, P., Cano-Ott, D., Chauvie, S., Cho, K., Cirrone, G.A.P., Cooperman, G., Cortés-Giraldo, M.A., Cosmo, G., Cuttone, G., Depaola, G., Desorgher, L., Dong, X., Dotti, A., Elvira, V.D., Folger, G., Francis, Z., Galoyan, A., Garnier, L., Gayer, M., Genser, K.L., Grichine, V.M., Guatelli, S., Guèye, P., Gumplinger, P., Howard, A.S., Hrivnáčová, I., Hwang, S., Incerti, S., Ivanchenko, A., Ivanchenko, V.N., Jones, F.W., Jun, S.Y., Kaitaniemi, P., Karakatsanis, N., Karamitros, M., Kelsey, M., Kimura, A., Koi, T., Kurashige, H., Lechner, A., Lee, S.B., Longo, F., Maire, M., Mancusi, D., Mantero, A., Mendoza, E., Morgan, B., Murakami, K., Nikitina, T., Pandola, L., Paprocki, P., Perl, J., Petrović, I., Pia, M.G., Pokorski, W., Quesada, J.M., Raine, M., Reis, M.A., Ribon, A., Ristić Fira, A., Romano, F., Russo, G., Santin, G., Sasaki, T., Sawkey, D., Shin, J.I., Strakovsky, I.I., Taborda, A., Tanaka, S., Tomé, B., Toshito, T., Tran, H.N., Truscott, P.R., Urban, L., Uzhinsky, V., Verbeke, J.M., Verderi, M., Wendt, B.L., Wenzel, H., Wright, D.H., Wright, D.M., Yamashita, T., Yarba, J., Yoshida, H.: Recent developments in geant4. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **835**, 186–225 (2016) <https://doi.org/10.1016/j.nima.2016.06.125>
- [19] NEMA Standard Publication NU 4-2008: Performance Measurements of Small

- [20] Jan, S., Santin, G., Strul, D., Staelens, S., Assié, K., Autret, D., Avner, S., Barbier, R., Bardiès, M., Bloomfield, P.M., Brasse, D., Breton, V., Bruyndonckx, P., Buvat, I., Chatziioannou, A.F., Choi, Y., Chung, Y.H., Comtat, C., Donnarieix, D., Ferrer, L., Glick, S.J., Groiselle, C.J., Guez, D., Honore, P.-F., Kerhoas-Cavata, S., Kirov, A.S., Kohli, V., Koole, M., Krieguer, M., Laan, D.J.v.d., Lamare, F., Largeron, G., Lartizien, C., Lazaro, D., Maas, M.C., Maigne, L., Mayet, F., Melot, F., Merheb, C., Pennacchio, E., Perez, J., Pietrzyk, U., Rannou, F.R., Rey, M., Schaart, D.R., Schmidtlein, C.R., Simon, L., Song, T.Y., Vieira, J.-M., Visvikis, D., Walle, R.V.d., Wieërs, E., Morel, C.: Gate: a simulation toolkit for pet and spect. *Physics in Medicine and Biology* **49**(19), 4543–4561 (2004) <https://doi.org/10.1088/0031-9155/49/19/007>
- [21] Jan, S., Benoit, D., Becheva, E., Carlier, T., Cassol, F., Descourt, P., Frisson, T., Grevillot, L., Guigues, L., Maigne, L., Morel, C., Perrot, Y., Rehfeld, N., Sarrut, D., Schaart, D.R., Stute, S., Pietrzyk, U., Visvikis, D., Zahra, N., Buvat, I.: Gate v6: a major enhancement of the gate simulation platform enabling modelling of ct and radiotherapy. *Physics in Medicine & Biology* **56**(4), 881 (2011) <https://doi.org/10.1088/0031-9155/56/4/001>
- [22] Kalaitzidis, P.: Modelling and simulations of a clinical pet-system using the gate monte carlo software. Master of science thesis, Lund University (2018). <https://lup.lub.lu.se/luur/download?func=downloadFile&recordId=8958979&fileId=8958982>
- [23] Khateri, P.: SAFIR PROTOTYPE PET INSERT: IMAGE RECONSTRUCTION AND NEMA CHARACTERIZATION. Theses, ETH Zurich (2021). <https://doi.org/10.3929/ethz-b-000485496>
- [24] Khateri, P., Fischer, J., Lustermann, W., Tsoumpas, C., Dissertori, G.: Implementation of cylindrical pet scanners with block detector geometry in stir. *EJNMMI physics* **6**(1), 15 (2019)
- [25] Baran, J., Borys, D., Brzeziński, K., Gajewski, J., Silarski, M., Chug, N., Coussat, A., Czerwiński, E., Dadgar, M., Dulski, K., Eliyan, K.V., Gajos, A., Kacprzak, K., Kapłon, Klimaszewski, K., Konieczka, P., Kopeć, R., Korcyl, G., Kozik, T., Krzemień, W., Kumar, D., Lomax, A.J., McNamara, K., Niedźwiecki, S., Olko, P., Panek, D., Parzych, S., Perez del Rio, E., Raczyński, L., Simbarashe, M., Sharma, S., Shivani, Shopa, R.Y., Skóra, T., Skurzok, M., Stasica, P., Stępień, E.Ł., Tayefi, K., Tayefi, F., Weber, D.C., Winterhalter, C., Wiślicki, W., Moskal, P., Ruciński, A.: Feasibility of the j-pet to monitor the range of therapeutic proton beams. *Physica Medica* **118**, 103301 (2024) <https://doi.org/10.1016/j.ejmp.2024.103301>
- [26] Das, M., Sharma, S., Beyene, E.Y., Bilewicz, A., Choinski, J., Chug, N., Curceanu, C., Czerwinski, E., Eliyan, K.V., Hajduga, J., Jalali, S., Kacprzak,

- K., Kaplanoglu, T., Kaplon, L., Kasperska, K., Khreptak, A., Korcyl, G., Kozik, T., Kubat, K., Kumar, D., Venadan, A.K., Lisowski, E., Lisowski, F., Medrala-Sowa, J., Moyo, S., Mryka, W., Niedźwiecki, S., Pandey, P., Parzych, S., Porcelli, A., Rachwal, B., Rio, E.P., Rädler, M., Rominger, A., Shi, K., Skurzok, M., Stolarz, A., Szumlak, T., Tanty, P., Ardebili, K.T., Tiwari, S., Walczak, R., Stepien, E.L., Moskal, P.: First positronium imaging using 44sc with the j-pet scanner: a case study on the nema-image quality phantom. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 1–1 (2025) <https://doi.org/10.1109/TRPMS.2025.3621554>
- [27] Giovagnoli, D.: Image reconstruction for three-gamma PET imaging. Theses, Ecole nationale supérieure Mines-Télécom Atlantique (November 2020). <https://theses.hal.science/tel-03123557>
- [28] Aggarwal, K.: Impact of preclinical PET scanner characteristics on the overall image quality. Theses, Université de Strasbourg (October 2019). <https://theses.hal.science/tel-02490686>
- [29] Webber, G., Reader, A.J.: Diffusion models for medical image reconstruction. *BJR|Artificial Intelligence* **1**(1), 013 (2024) <https://doi.org/10.1093/bjrai/ubae013> <https://academic.oup.com/bjrai/article-pdf/1/1/ubae013/59189039/ubae013.pdf>
- [30] Singh, I.R., Denker, A., Barbano, R., Kereta, v., Jin, B., Thielemans, K., Maass, P., Arridge, S.: Score-based generative models for pet image reconstruction. *Machine Learning for Biomedical Imaging 2*(Generative Models), 547–585 (2024) <https://doi.org/10.59275/j.melba.2024-5d51>
- [31] Reader, A.J., Corda, G., Mehranian, A., Costa-Luis, C.d., Ellis, S., Schnabel, J.A.: Deep learning for pet image reconstruction. *IEEE Transactions on Radiation and Plasma Medical Sciences* **5**(1), 1–25 (2021) <https://doi.org/10.1109/TRPMS.2020.3014786>
- [32] Mehranian, A., Reader, A.J.: Model-based deep learning pet image reconstruction using forward–backward splitting expectation–maximization. *IEEE Transactions on Radiation and Plasma Medical Sciences* **5**(1), 54–64 (2021) <https://doi.org/10.1109/TRPMS.2020.3004408>
- [33] Webber, G., Mizuno, Y., Howes, O.D., Hammers, A., King, A.P., Reader, A.J.: Generative-model-based fully 3d pet image reconstruction by conditional diffusion sampling. In: 2024 IEEE Nuclear Science Symposium (NSS), Medical Imaging Conference (MIC) and Room Temperature Semiconductor Detector Conference (RTSD), pp. 1–2 (2024). <https://doi.org/10.1109/NSS/MIC/RTSD57108.2024.10657861> . <https://arxiv.org/abs/2412.04319>
- [34] Webber, G., Mizuno, Y., Howes, O.D., Hammers, A., King, A.P., Reader, A.J.: Likelihood-scheduled score-based generative modeling for fully 3d pet image

- p reconstruction.
- IEEE Transactions on Medical Imaging*
- 44**
- (11), 4445–4456 (2025)
- <https://doi.org/10.1109/TMI.2025.3576483>
- [35] Hatvani, J.: Medical image enhancement using deep learning and tensor factorization techniques. Theses, Université Paul Sabatier - Toulouse III ; Pázmány Péter katolikus egyetem (Budapest ; 1992-....) (May 2021). <https://theses.hal.science/tel-04164508>
 - [36] Blanco, A.: Positron range effects on the spatial resolution of rpc-pet. 2006 IEEE Nuclear Science Symposium Conference Record **4**, 2570–2573 (2006)
 - [37] Fischer, J.: SAFIR input file format for STIR. Institute for Particle Physics, ETH Zurich. (2015). <http://doi.org/10.5905/ethz-1007-22>
 - [38] Siddon, R.L.: Fast calculation of the exact radiological path for a three-dimensional ct array. *Medical physics* **12**(2), 252–255 (1985)
 - [39] Sundermann, E., Jacobs, F., Christiaens, M., De Sutter, B., Lemahieu, I.: A fast algorithm to calculate the exact radiological path through a pixel or voxel space. *Journal of Computing and Information Technology* **6** (1998)
 - [40] Joseph, P.M.: An improved algorithm for reprojecting rays through pixel images. *IEEE Transactions on Medical Imaging* **1**(3), 192–196 (1982) <https://doi.org/10.1109/TMI.1982.4307572>
 - [41] Emond, E.C., Groves, A.M., Hutton, B.F., Thielemans, K.: Effect of positron range on pet quantification in diseased and normal lungs. *Physics in Medicine & Biology* **64**(20), 205010 (2019) <https://doi.org/10.1088/1361-6560/ab469d>
 - [42] Lehnert, W., Gregoire, M.-C., Reilhac, A., Meikle, S.R.: Analytical positron range modelling in heterogeneous media for pet monte carlo simulation. *Physics in Medicine & Biology* **56**(11), 3313 (2011) <https://doi.org/10.1088/0031-9155/56/11/009>
 - [43] Cal-González, J., Herraiz, J.L., España, S., Corzo, P.M.G., Vaquero, J.J., Desco, M., Udias, J.M.: Positron range estimations with penelopet. *Physics in Medicine & Biology* **58**(15), 5127 (2013) <https://doi.org/10.1088/0031-9155/58/15/5127>
 - [44] Bai, B., Ruangma, A., Laforest, R., Tai, Y.-C., Leahy, R.M.: Positron range modeling for statistical pet image reconstruction. 2003 IEEE Nuclear Science Symposium. Conference Record (IEEE Cat. No.03CH37515) **4**, 2501–25054 (2003)
 - [45] Jødal, L., Le Loirec, C., Champion, C.: Positron range in pet imaging: an alternative approach for assessing and correcting the blurring. *Physics in Medicine & Biology* **57**(12), 3931 (2012) <https://doi.org/10.1088/0031-9155/57/12/3931>

- [46] Bailey, D.L., Karp, J.S., Surti, S.: Physics and instrumentation in pet. In: Positron Emission Tomography: Basic Sciences, pp. 13–39. Springer, London (2005)