

Experimental Prompt Gamma Scatter Classification and Reorientation for Proton Beam Image Reconstruction

REU Site: Online Interdisciplinary Big Data Analytics in Science and Engineering

Ayan Kabaria¹, Muhammad Khalid², Sophia Lopez³, Abby Nam⁴,
Ertan Dogan⁵, Sidhya Pathak⁶, Victor Sandrin⁷, Xueying Sun⁸,
Ehsan Shakeri⁹, Harrison Lewis⁹, Hussam Fateen⁹, Matthias K. Gobbert⁹,
Farshad Safavi¹⁰, Ananta Chalise¹⁰, Lei Ren¹¹,
Stephen W. Peterson¹², and Jerimy C. Polf¹³

¹River Hill High School, Howard County, MD

²Department of Mathematics, Baruch College, City University of New York

³Department of Mathematics and Statistics, University of North Carolina at Greensboro

⁴Department of Psychology, Lafayette College

⁵A. James Clark School of Engineering, University of Maryland, College Park

⁶Department of Computer Science, University of Virginia

⁷Department of Neuroscience, University of Arizona

⁸Department of Information Technology and Management, Illinois Institute of Technology

⁹Department of Mathematics and Statistics, UMBC

¹⁰Department of Radiation Oncology, University of Maryland School of Medicine

¹¹Department of Radiation Oncology, Northwestern University

¹²Department of Physics, University of Cape Town, South Africa

¹³M3D, Inc.

Technical Report HPCF-2026-5, hpcf.umbc.edu > Publications

Abstract

Prompt gamma rays emitted by cells during proton beam radiotherapy can be detected by Compton cameras and used to reconstruct images to verify accurate dose delivery and Bragg peak location. However, image reconstruction is complicated by misclassified scattering events because of the limited temporal resolution of the cameras. To address this, machine learning models were trained on simulated datasets generated using Geant4 and MCDE and implemented through the Big Data REU Integrated Development and Experimentation (BRIDE) platform to classify and filter scatter events for more accurate image reconstruction. After identifying the best-performing models, predictions using a deep Fully Connected Network (FCN) were created for 24 cases of experimental phantom triples data. Images were reconstructed on CORE using Kernel Expectation-Maximization (KEM) and Simple Back Projection (SBP) algorithms on both raw and repaired data, then compared to assess the accuracy of the Bragg peak estimation and noise reduction. SBP reconstructions had more complete beam reconstructions and Bragg peak profiles than KEM reconstructions; however, they still included noise in 2-D beam profiles. KEM 1-D profile reconstructions improved after a reorientation script was implemented to transform the scatter data before classification, which reoriented the data such that it was consistent with training, and reversed before reconstruction.

Key words. Proton beam therapy, Compton camera, machine learning, image reconstruction, deep learning models

1 Introduction

Proton beam radiotherapy is a form of external radiation beam therapy that is used for cancer treatment. It is typically favored over other forms of treatment, such as X-ray therapy, for specific

complex tumors or tumors in specific locations because it can be calibrated such that a majority of the radiation dose is deposited at a specific depth known as the Bragg peak, leaving the healthy tissue around the tumor mostly unaffected [7].

However, due to range uncertainty, it is extremely difficult to accurately align the Bragg peak with the precise location of the tumor. To mitigate these uncertainties, creating a safety margin that encompasses an area slightly larger than the tumor can ensure that the tumor is maximally irradiated, at the risk of potentially delivering radiation to healthy tissue in the surrounding area [7].

To fully exploit the unique advantages of proton therapy, it is therefore desirable to tighten the safety margins. Previous research has focused on real-time imaging methods to improve beam positioning accuracy [8]. The distance the proton travels in the patient’s body is largely affected by a variety of factors such as tissue density, patient movement, imaging limitations, etc. Thus, a promising approach to continuously track proton beams using prompt gammas is necessary to correctly target the Bragg peaks. When proton beams traveling through the body collide with atomic nuclei, they release excess energy in the form of prompt gamma rays. Multistage detectors known as Compton cameras can capture these prompt gamma interactions in real time as single, double, and triple scatters. However, Compton cameras are limited by their finite time resolution. Consequently, distinct gamma rays can be incorrectly classified as a single event, or the recorded order of the prompt gamma interactions can be incorrect. The resulting reconstructed images may then contain distortion and noise, making it more difficult to accurately deduce the location of the proton beam in the body [7].

Machine learning techniques have previously been used to remedy the deficits in the reconstructed images. Machine learning models can be trained to classify each prompt gamma interaction, and then repair or remove unusable events. The models implemented in the machine learning framework of the present study include deep Fully Connected Networks (FCNs), Long Short-Term Memory Neural Networks (LSTMs), 1-D Convolutional Neural Networks (CNNs), and an Event Classifier Transformer [2]. The current study utilized an FCN to classify the 24 experimental setups utilizing real-world triple-scatter phantom data.

The remainder of this report is organized as follows: Section 2 expands upon proton beam radiotherapy, Compton camera imaging, scatter types, and the impetus for employing machine learning models. Section 3 introduces the technical background for machine learning and further discusses the machine learning models of interest, as well as the specific hardware and software used to run experiments. Section 4 explains each stage of the data generation pipeline, using the Polaris-J Monte-Carlo (PJMC) framework, Geant4, and Monte-Carlo Detector Effects (MCDE) to simulate data. The Big Data REU Integrated Development and Experimentation (BRIDE) platform [3] is also discussed for its use in training machine learning models to classify prompt gamma scatters. Section 5 presents the results of using the SBP or KEM reconstruction algorithms on unrepaired and repaired data for each of the 24 experimental triple-scatter datasets, and examines the efficacy of reorienting the test data to prevent SBP from erroneously removing detectors in reconstructions. Section 6 recapitulates the most significant findings, compares them to prior studies, and considers future directions for the research.

2 Background

2.1 Proton Beam Radiotherapy

Radiotherapy functions as a cancer treatment by depositing energy through nuclear interactions into bodily tissue, damaging tumor cells. Ideally, the radiation treatment dose should be administered to the entire tumor while minimizing damage to the surrounding healthy tissue and organs.

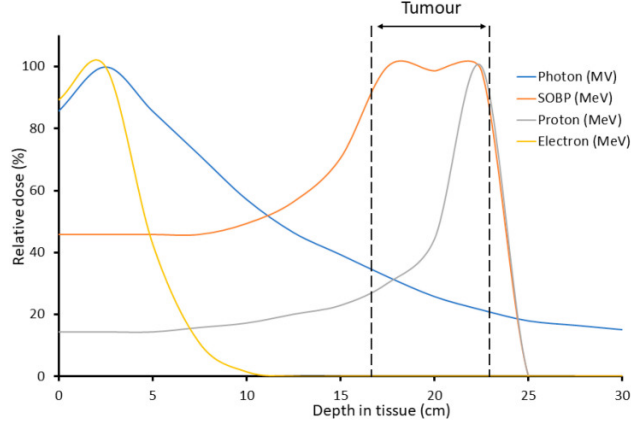


Figure 2.1: Energy deposition pattern of the spread-out Bragg peak (SOBP) compared with other treatment energy types [5].

In the past decade, radiotherapy has increasingly employed proton and carbon-ion beams for treatment. Conventional X-ray therapy utilizes high-energy X-rays, which deposit the maximum radiation dose shortly after entering the body and continuously release energy as they travel through tissue. Consequently, healthy tissue may be exposed to radiation [7].

In contrast, proton beam radiotherapy utilizes charged particles, protons, which deposit less energy passing through healthy tissue. When the protons reach the tumor location, they release the majority of their radiation in a sudden burst at the site known as the beam range, or Bragg peak. The radiation dose then drops sharply as the protons exit the treatment area beyond the Bragg peak, reducing the irradiation of healthy tissue around the tumor. Using this pattern of energy deposition, more of the radiation dose can be delivered to the tumor and less to the surrounding critical organs and tissue. As a result, patients may experience fewer complications due to radiation therapy and, moreover, are more likely to be cured of their cancer [7]. The dose deposition pattern of the Bragg peak compared to other forms of treatment is further illustrated in Figure 2.1 [5].

Determining the exact location of the Bragg peak in proton beam radiotherapy would allow clinicians to exert greater control and precision in radiation dose delivery when trying to target tumors. While this is particularly useful for treating cancers near critical organs, clinicians are limited by the degree of error that exists when predicting the location of the Bragg peak; even a minor misalignment may lead to the Bragg peak overshooting or undershooting adjacent to the tumor site, damaging healthy tissue instead.

Despite these difficulties in locating the Bragg peak, clinicians can still exploit the advantages of proton beam radiotherapy by creating a safety margin that both accounts for surrounding healthy tissue that may be irradiated due to range uncertainty and ensures that the entire tumor receives the dose. Figure 2.2 demonstrates the importance of accurately determining the safety margin for a breast cancer patient. The purple margin indicates the area to be treated. In (a), the yellow outline represents an estimated treatment volume that accounts for uncertainty in determining the exact location of the Bragg peak. As indicated by the overlap of the treatment area with the heart and left anterior descending coronary artery (LAD), administering the dose as pictured could unnecessarily damage adjacent healthy heart tissue and render the patient more susceptible to radiation-induced heart disease. In (b), an improved safety margin significantly reduces the risk of such complications by eliminating the overlap between the heart and LAD with the intended treatment volume [7].

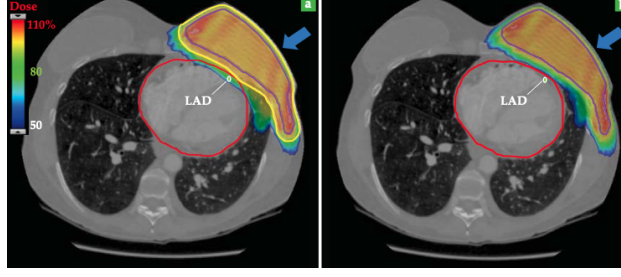


Figure 2.2: (a) The proton beam (blue) intended to target the tumor (purple) surrounded by the safety margin (yellow) overlaps with healthy heart tissue (red) and the left anterior descending (LAD) coronary artery (white circle). (b) The optimized safety margin successfully minimizes overlap of the intended treatment volume with the heart tissue and LAD [7].

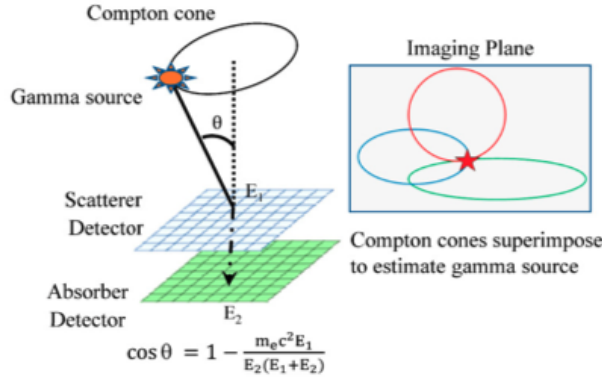


Figure 2.3: An example of a Compton cone of emission [6].

2.2 Compton Camera and Image Reconstruction

To optimize safety margins such that the tumor is completely irradiated and healthy tissue remains intact, clinicians can use detectors known as Compton cameras to capture real-time information. When the protons interact with the atoms in patient tissue, they emit prompt gamma rays that the Compton camera can detect, providing valuable information regarding the path of the beam and the possible location of the Bragg peak. This information allows the clinicians to adjust the gantry angle and proton beam energy real-time such that the tumor can be correctly targeted.

As seen in Figure 2.3, a Compton camera consists of three detection layers — two scatterer layers and one absorption layer. In the first two interactions in the scatterer layers, an incoming gamma ray deposits a portion of its energy via Compton scattering and is deflected at a scatter angle θ . The twice-scattered gamma then usually travels to the absorption layer, where its remaining energy is fully deposited. Although less common, a prompt gamma may also scatter backwards (i.e., contacting the second scatterer or absorber followed by the first).

By applying the Compton kinematic equation to the three chronological measured energy deposits, θ can be reconstructed via

$$\cos \theta_1 = 1 - m_e c^2 \left(\frac{1}{E_2 + E_3} - \frac{1}{E_1 + E_2 + E_3} \right), \quad (2.1)$$

where $m_e c^2$ is the electron rest mass energy, E_1 is the energy deposited in the first interaction,

Class	Interaction 1				Interaction 2				Interaction 3			
123	e_1	x_1	y_1	z_1	e_2	x_2	y_2	z_2	e_3	x_3	y_3	z_3
132	e_1	x_1	y_1	z_1	e_3	x_3	y_3	z_3	e_2	x_2	y_2	z_2
213	e_2	x_2	y_2	z_2	e_1	x_1	y_1	z_1	e_3	x_3	y_3	z_3
231	e_2	x_2	y_2	z_2	e_3	x_3	y_3	z_3	e_1	x_1	y_1	z_1
312	e_3	x_3	y_3	z_3	e_1	x_1	y_1	z_1	e_2	x_2	y_2	z_2
321	e_3	x_3	y_3	z_3	e_2	x_2	y_2	z_2	e_1	x_1	y_1	z_1
124	e_1	x_1	y_1	z_1	e_2	x_2	y_2	z_2	single			
214	e_2	x_2	y_2	z_2	e_1	x_1	y_1	z_1	single			
134	e_1	x_1	y_1	z_1	single				e_2	x_2	y_2	z_2
314	e_2	x_2	y_2	z_2	single				e_1	x_1	y_1	z_1
234	single				e_1	x_1	y_1	z_1	e_2	x_2	y_2	z_2
324	single				e_2	x_2	y_2	z_2	e_1	x_1	y_1	z_1
444	single				single				single			

Figure 2.4: Scatter event classifications [1].

E_2 is the energy deposited in the second interaction, and E_3 is the energy deposited in the third interaction. Although the true source of the prompt gamma cannot be reconstructed with complete accuracy, this equation generates a conical surface on which the origin likely lies. When many of these cones are reconstructed from multiple detected events, an image reconstruction algorithm can identify the most likely source of the gamma emission [8].

2.2.1 Scatter Types

As collisions of prompt gammas inside of a Compton camera happen at the speed of light, there are often errors in triple scatter classifications and introduced background noise. Due to this, there are 13 scatter order classifications for the correct and incorrect detections of triple scatters, as shown in Figure 2.5, generalized into four main cases: True Triples, Misordered Triples, Doubles-to-Triples, and False Triples.

True Triple Scatter: The most useful detection, however least common, is the true triple event, including class 123, which occurs when the Compton camera detects the triple scatter of a single event in the correct order. In this case, the event can be left as-is during data repair.

Misordered Triple Scatter: The misordered triple scatter event, including classes 123, 213, 231, 312, 321, occurs when three scatter interactions are correctly detected from a single event, however are labeled out of order. In this case, the data must be repaired by reordering the triple in the correct sequence.

Double-to-Triple Scatter (DtoT): The double-to-triple scatter event, including classes 124, 214, 134, 314, 234, 324, is when two of the three scatters are from one event and the remaining scatter is either from another scatter or extraneous noise. The "4" in the class label indicates an outside or unrelated event. Here, the model throws away the extraneous event and reorders the double scatter.

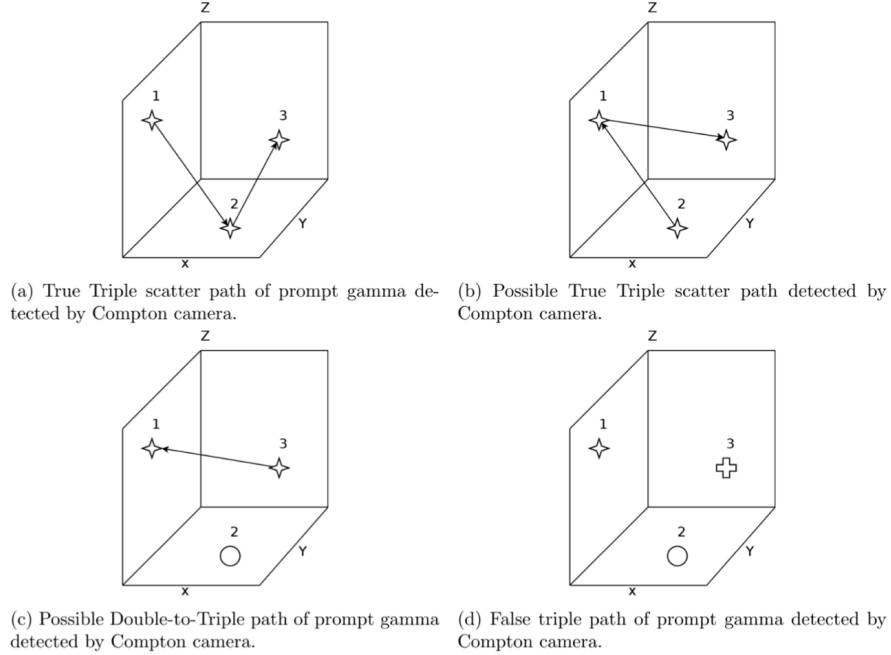


Figure 2.5: Scatter events [4].

False Triple Scatter: The final case is the false triple scatter event, class 444, where three independent scatters are incorrectly classified into one event, in which case the data must be discarded.

2.2.2 The Need for Machine Learning

Beyond these, the efficiency and resolution of the Compton camera introduce several classes of misidentified events. Machine learning can be applied across all three misclassification cases to recover misordered triples, improve event identification, and remove spurious events. The removal and repair of scatter events by the machine learning model can result in a less noisy, sharper image that can be more useful in determining the location of the Bragg peak.

3 Methodology

3.1 Hardware and Software

This research primarily utilized Graphics Processing Units (GPUs) in the chip cluster managed by the UMBC High Performance Computing Facility (hpcf.umbc.edu). The chip cluster's GPU portion includes four nodes with eight RTX 2080Ti GPUs (11 GB of GDDR6 memory each), seven nodes with eight RTX 6000 GPUs (24 GB of GDDR6 memory each), two nodes of eight RTX 8000 (48 GB of GDDR6 memory each), two nodes with two NVIDIA H100 GPUs (100 GB of HBM3 memory each), and ten nodes with four L40S GPUs (4 GB of GDDR6 memory each).

The machine learning models were built and implemented using PyTorch v2.5.1 (<https://PyTorch.org>). To pre-process and manipulate data, scikit-learn v1.7.2 (<https://scikit-learn.org/stable/>), pandas v2.3.3 (<https://pandas.pydata.org/>), and NumPy v2.2.6 (<https://numpy.org/>) were used. SciPy v1.15.2 (<https://scipy.org>) and Matplotlib v3.10.9 (<https://matplotlib.org>) were used to build re-

constructions. The Python environment Anaconda3 (<https://www.anaconda.com/>) was used to manage the aforementioned project dependencies and execute Python scripts throughout the data generation and experimental pipeline.

The project pipeline was organized into several subdirectories within the Team 2 directory: 1-PJMC, 2-MCDE, 3-BRIDE, and 4-CORE. The first three subdirectories were involved in the simulation of patient data and the classification of prompt gamma scatters. The 4-CORE directory and associated workflow facilitated the application of trained machine learning models to real-world experimental data. PJMC, MCDE, BRIDE, and CORE required the submission of slurm scripts as interactive jobs to the UMBC High Performance Computing chip cluster.

3.2 Polaris-J Monte-Carlo (PJMC)

The Polaris-J Monte-Carlo (PJMC) framework is a customized Geant4-based simulation tool used to model proton beam interactions with patient tissue and generate realistic prompt gamma emission data for machine learning and image reconstruction studies. PJMC utilizes the Geant4 particle transport toolkit to model the interactions of high-energy protons with matter and the resulting prompt gamma emissions detected by a Compton camera. Within a given run, the `parameters.mac` file could be edited to modify the proton beam position, beam energy, and number of simulated events. The result of submitting the `compiled.root` file was a `compiled.root` file that contained the simulated detector events. For use with the Monte-Carlo Detector Effects workflow, `compiled.root` was converted into a CSV by running the `converter_core.py` script. A successful PJMC result for a given run contained a "511" CSV representing high-energy collisions and annihilation data, as well as a "PG" (prompt gamma) CSV containing data for lower-energy collisions.

3.3 Monte-Carlo Detector Effects (MCDE)

The CSVs containing the simulated Geant4 data are subsequently passed through the Monte-Carlo Detector Effects (MCDE) model. MCDE processes the simulated data such that it is more similar to the actual patient data that a Compton camera would produce. This allows the simulated datasets to more accurately represent detector behavior observed during experimental proton beam imaging. To achieve this, MCDE adds detector timing and trigger effects to the Geant4 data. MCDE separates the detected interactions into single, double, and triple scatter events, allowing triple scatter data to be used in subsequent machine learning pre-processing within BRIDE. Although MCDE produces files for singles, doubles, and triples, only the triple-scatter events are used in the subsequent machine learning pipeline. To process the 511 and PG CSV files, the `run_chip.slurm` script was submitted as a SLURM job. This executed `Script_ErgInit_Peter_3e9.py`, which created a new `saveMCDE` directory containing text files for the separated single-, double-, and triple-scatter events.

3.4 BRIDE Platform and Machine Learning Prediction

The Big Data REU Integrated Development and Experimentation (BRIDE) platform is a machine learning development platform that allows for rigorous testing and flexible model development. Furthermore, BRIDE is advantageous because it is a high-level modular code workflow and leverages PyTorch Lightning's Distributed Data Parallel to use multiple GPUs during training [3]. Currently, the available models for training include the Fully Connected Neural Network (FCN), the Long Short-term Memory Neural Network (LSTM), the 1-D Convolutional Neural Network (CNN) and the Event Classifier Transformer. Hyperparameter tuning is automated using Grid Search [2]. In

the present research, BRIDE was used to apply an FCN used by the 2025 Big Data REU team, `impr_fcn` [2], to classify different scattering events.

The overall structure of BRIDE is visualized in Figure 3.1. BRIDE contains multiple distinct components: first, directories containing CSVs of raw prompt gamma scattering data are stored in the **raw_data** directory of BRIDE. During the first round of pre-processing, through **pp1**, the raw data is filtered to only include triples. The result is a CSV file containing class labels for true triples, double-to-triples, and false triples events. This CSV is then used for the second round of processing through **pp2**, creating a training, validation, and test set split. After pre-processing, the data can be used to train and test a specified machine learning model. The **logs** directory stores model checkpoints and training and validation metrics in output logs. Additionally, the metrics resulting from the testing the model on the test set are kept in the **eval** directory [2].

Prediction jobs within the BRIDE platform are created by specifying a unique **run_id** and creating an associated `.yaml` configuration file. The configuration file specifies the pre-trained FCN checkpoint, the experimental dataset to be evaluated, and the output locations for the prediction results. The 3-BRIDE directory we used contained various subdirectories, which are outlined in more detail in [2].

For each of the 24 experimental datasets, a `predict.py` script was executed through a slurm job, loading the pre-trained model and generating predicted classifications for each prompt gamma triple-scatter event. Unlike typical BRIDE runs, when we applied the FCN model to our experimental triples data, we modified the `predict.sh` script to skip the evaluation and only specified the test set for the train-test split parameters in the config files. Since experimental, real-world data was used instead of simulated data, we were more interested in using a machine learning model shown to be effective to classify and repair the data before creating reconstructions than testing the validation accuracy of the model itself. The resulting predictions were saved as NumPy arrays in the **eval** directory and converted back into CSV format for use in the CORE reconstruction pipeline. The preliminary results of the reconstruction indicated that the performance of the model varied between different gantry angles. To investigate whether detector orientation influenced prediction accuracy, subsequent experiments incorporated coordinate reorientation of the experimental datasets prior to machine learning prediction while maintaining the same pretrained FCN model.

3.5 Experimental Data Source

The experimental datasets used in this study were provided by Farshad Safavi and collected using a nozzle-mounted quad Compton camera prompt gamma imaging (PGI) system integrated with a clinical proton therapy gantry. During proton beam irradiation, prompt gamma events were acquired from both water and High-Density Polyethylene (HDPE) phantoms under multiple beam delivery conditions, including varying beam energies and gantry angles. The detector system recorded prompt gamma interactions throughout irradiation, producing the experimental triple-scatter datasets used in this study [8]. The datasets consisted of experimental Compton camera triple-scatter events collected under multiple experimental conditions. These datasets were provided to the research team for evaluations and served as the input for the machine learning prediction and beam reconstruction workflow.

3.6 Experimental Triples Data Reconstructions

This experiment utilized data from an experimental proton beam radiotherapy system using both water phantoms and High-Density Polyethylene (HDPE) phantoms. The objective of this study was to apply a previously trained machine learning model [2] to experimental Compton camera data

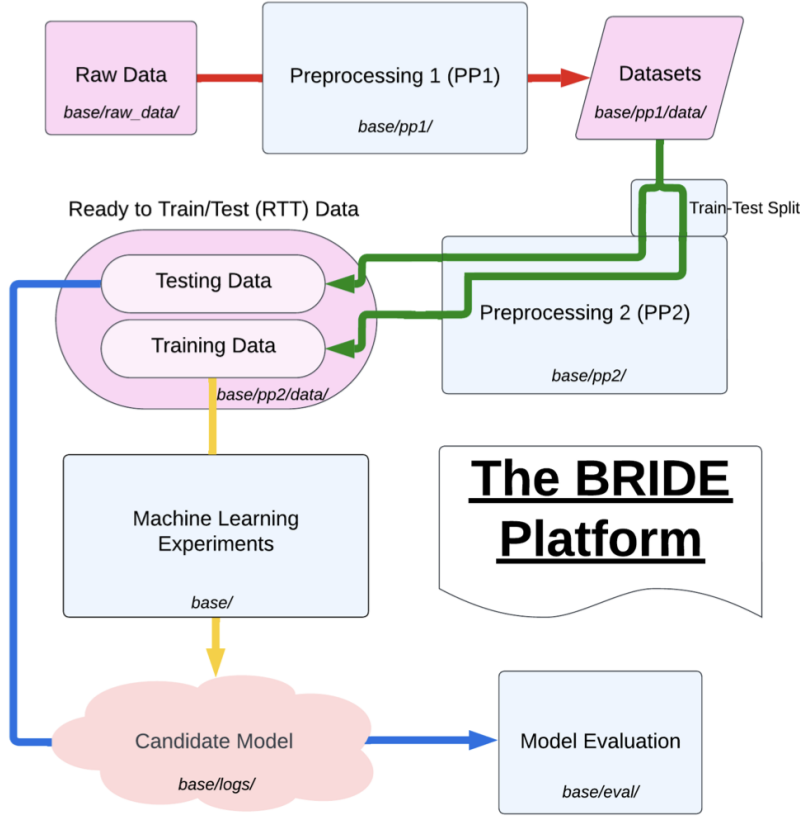


Figure 3.1: Overview of the BRIDE machine learning workflow. Raw post-MCDE data undergoes pre-processing, train-test splitting, machine learning experimentation, and model evaluation. For this study, the pretrained FCN model was used only for prediction on experimental datasets [3].

and evaluate whether the repaired data improved proton beam image reconstruction compared to the original control data. The phantoms in the proton beam radiotherapy system were placed to simulate human tissue without a live patient: water phantoms to simulate softer tissue, and HDPE phantoms to simulate higher-density tissue such as bone. Although phantoms cannot precisely replicate the variation of electron density inside an actual human body, they are the standard method when patient data is inaccessible. 24 different snout and beam orientations were used for data capture, including different gantry angles, different beam energies, and other factors.

Raw Compton camera triple scatter data from the experiment was first converted to NumPy arrays and saved. The datasets were then repaired using the BRIDE platform with a trained checkpoint of a Fully Connected Network (FCN) model from the 2025 Big Data REU team. A slurm script in the `runs` directory executed `predict.py` for each of the 24 datasets. During prediction, the pretrained FCN checkpoint was loaded and used to classify each prompt gamma triple-scatter event. The resulting predicted classifications were saved as NumPy arrays in the `eval` directory. The predicted classifications were then converted back into CSV format, producing repaired datasets compatible with the CORE reconstruction pipeline. To reconstruct the data, symbolic links for the repaired and control (original) data files were created in the `4-CORE/2_DATA` directory. Two different image reconstruction algorithms were run on the control and repaired data: Kernel Expectation-

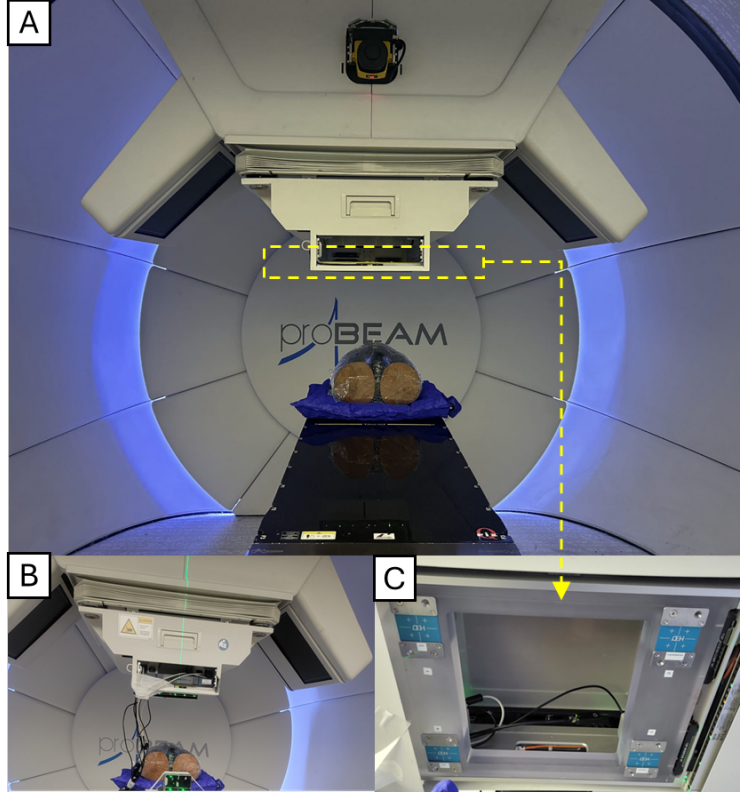


Figure 3.2: Experimental proton beam imaging setup used to acquire the prompt gamma datasets. (a) Clinical proton therapy gantry with the nozzle-mounted quad Compton camera prompt gamma imaging (PGI) system. (b) Detector mounted beneath the treatment nozzle during phantom irradiation. (c) Bottom view of the detector assembly [8].

Maximization (KEM) and Simple Back-Projection (SBP). This allowed for qualitative comparison of proton beam reconstruction before and after machine learning correction. A full field of view was used for the image reconstruction outputs, including a range of -500 mm to 500 mm for each axis. A slurm script and Python script were run for each trial which produced 1-D Bragg peak profile predictions and 2-D beam slices for each axis.

3.7 Beam Reconstruction and Evaluation

The purpose of the `4-CORE` directory is to reconstruct prompt gamma emission images from Compton camera data and estimate the location of the proton beam Bragg peak. These reconstructions provide a visual representation of the beam path and allow image quality to be evaluated after machine learning based scatter correction.

Reconstruction parameters were specified using the `triples.cfg` configuration file, which allowed selection of the reconstruction algorithm and imaging parameters. The reconstruction algorithms currently support Simple Back Projection (SBP), Octane, Kernel Expectation-Maximization (KEM), and Stochastic Origin Ensembles (SOE). Multiple reconstruction algorithms were evaluated to compare their ability to locate the proton beam and suppress background noise after machine learning correction. The resulting reconstruction `output.dat` after submitting the SLURM job will populate in the `6_OUTPUT/1-CORE_out` directory. Running `recon_plotter_ummc.py` subsequently

generates visualizations of one-dimensional proton beam profiles and two-dimensional proton beam slices.

3.8 Data Reorientation

The preliminary results of the reconstruction indicated that the pretrained FCN model performed inconsistently across different gantry angles. Specifically, two-dimensional yz -profiles for the SBP image reconstructions that displayed the position and energy of the four detectors at approximately $x = -445.0$ mm were consistently missing two detectors on the positive- y side.

An explanation proposed was that the orientation of the experimental Compton camera differed from the orientations represented in the training data. During training, the detectors were placed under the patient’s bench, and, therefore the resulting cones were oriented facing the negative- y direction. As such, the machine learning model would have been attempting to generalize to test data that utilized a setup wherein the detectors were flanking the sides of the bench instead, at the 90-degree and 270-degree gantry angles. To remedy this, experimental datasets were reoriented before machine learning prediction by applying coordinate transformations corresponding to the geometry of the detector, such that the coordinate plane for the test data was transformed to match the coordinate plane that was expected from the training data. Before creating the image reconstructions, the reorientation was then reversed to evaluate the efficacy of the new reconstructions with the previous reconstructions. A Python script, `reorientData.py`, was run on each data set to automate the coordinate transformations. The script inputs a CSV containing the 12-column raw triple-scatter data and asks the user to pass the argument "90" or "270" to specify the gantry angle at which the data were collected. For data collected with the cameras oriented at the 90-degree gantry angle, the original positive- x values were changed to positive- y values, and the original positive- y values were changed to the negative- x values. For data captured at the 270-degree gantry angle, the original positive- x values were changed to the negative- y values, and the original positive- y values were changed to the positive- x values. The reoriented datasets were then processed through the same aforementioned BRIDE prediction and CORE reconstruction pipeline to evaluate whether reconstruction quality improved.

4 Results

SBP image reconstructions were clearer and less noisy than KEM image reconstructions, though all reconstructions retained a significant amount of noise. In particular, SBP tended to filter out the detectors in the 2-D yz -beam profiles. After re-running SBP and KEM with the reoriented data, KEM saw some improvements, but no differences were observed for SBP.

4.1 Experimental Triples Data Reconstructions

From the experimental triples data provided by Farshad Safavi, we produced a total of 96 image reconstructions from 24 datasets. For each dataset, we compared the results of control reconstructions and repaired reconstructions for each algorithm. We focused on the 1-dimensional x -axis profile for an ideal single clear, sharp Bragg peak, and on the 1-dimensional y - and z -axis profiles for an expected centered, symmetric parabolic profile. We additionally analyzed the 2-dimensional yz beam profiles, expecting a centered, circular beam slice.

4.1.1 Initial SBP Reconstructions

SBP reconstructions had visible measured Bragg peaks in the x -axis profiles. In the y -axis and z -axis profiles for control data reconstructions, there were double peaks symmetrical about the center.

For dataset 11, in the control reconstructions, 9,831 events out of 16,707 total events were kept. In the repaired reconstructions, 5,593 events out of 9,049 total events were kept.

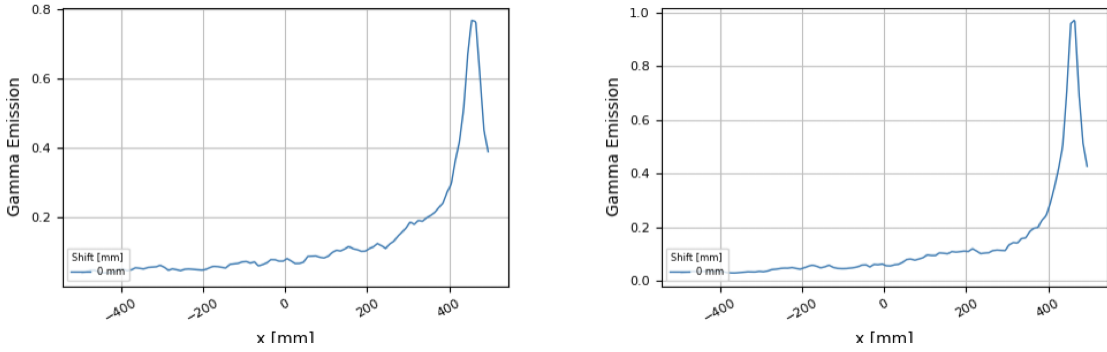


Figure 4.1: Dataset 11, SBP 90G, x -profiles, control vs. repaired.

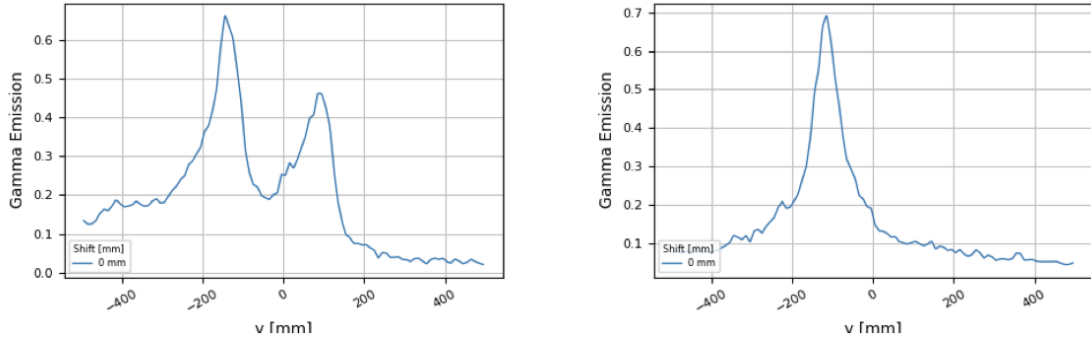


Figure 4.2: Dataset 11, SBP 90G, y -profiles, control vs. repaired.

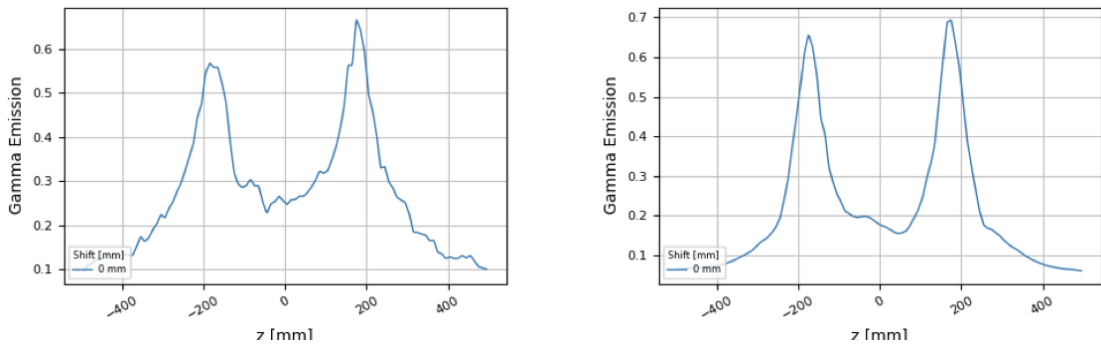


Figure 4.3: Dataset 11, SBP 90G, z -profiles, control vs. repaired.

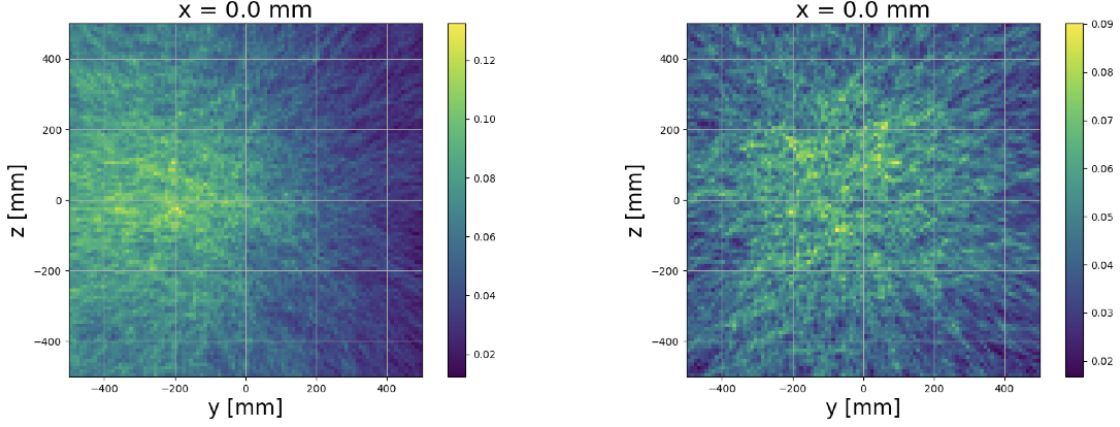


Figure 4.4: Dataset 11, SBP 90G, yz -profiles, control vs. repaired.

4.1.2 Initial KEM Reconstructions

For the KEM reconstructions, on both control and repaired data, the 1-D profiles and 2-D images were highly unstructured and contained mostly noise. Although some reconstructions on repaired data included Bragg peaks in the x -axis profile, the gamma emission values for those Bragg peaks were often negligible.

Note that datasets 17 and 18 are identical; the reconstructions below are labeled as dataset 17 for consistency. In the control reconstructions, 26,917 events out of 42,373 total events were kept. In the repaired reconstructions, 15,090 events out of 23,198 total events were kept.

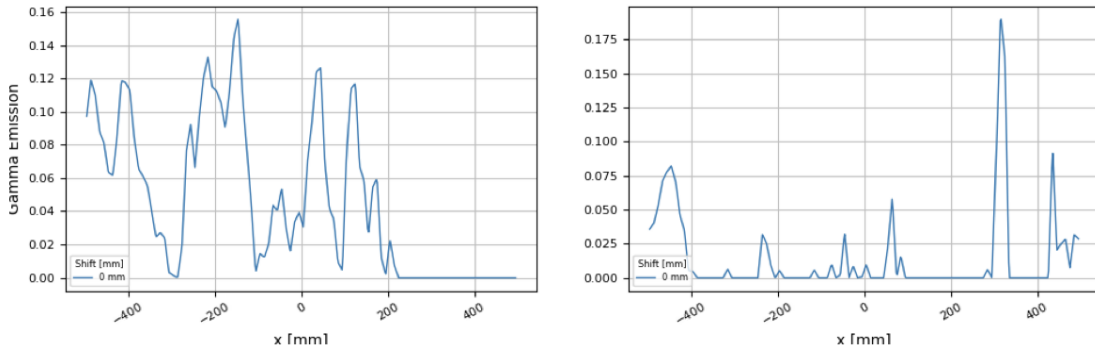


Figure 4.5: Dataset 17, KEM 90G, x -profiles, control vs. repaired.

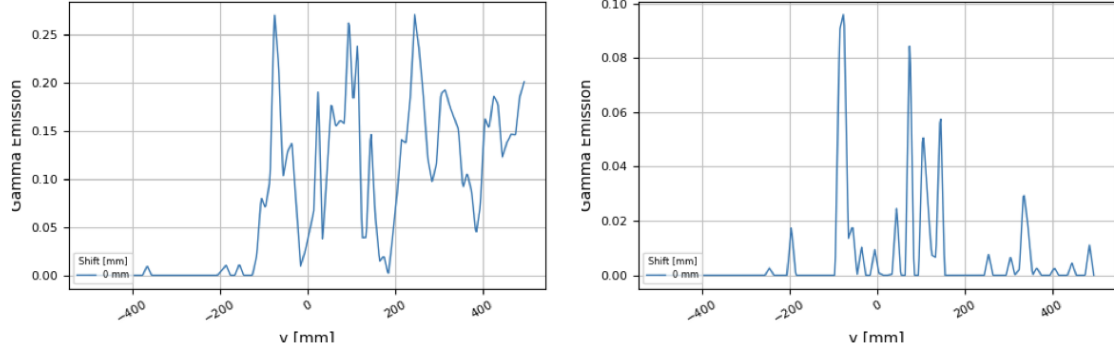


Figure 4.6: Dataset 17, KEM 90G, y -profiles, control vs. repaired.

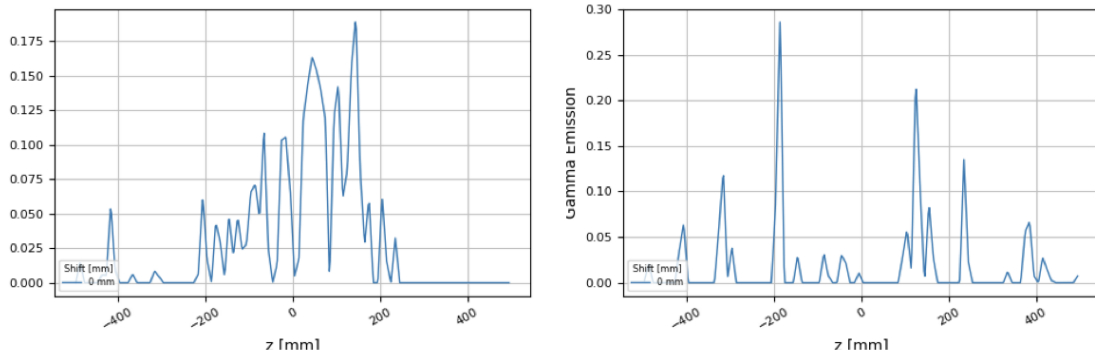


Figure 4.7: Dataset 17, KEM 90G, z -profiles, control vs. repaired.

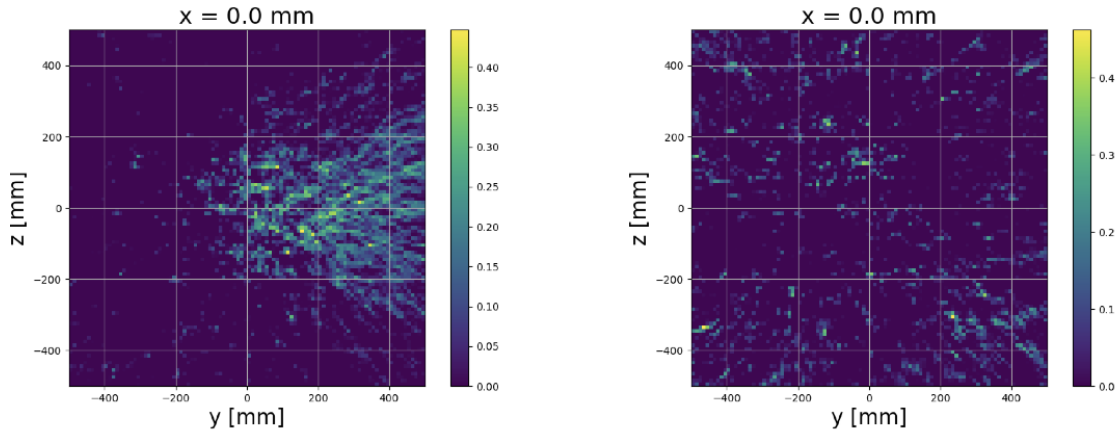


Figure 4.8: Dataset 17, KEM 90G, yz -profiles, control vs. repaired.

4.1.3 Detector Anomalies

In most repaired reconstructions, the y -axis profile had one of the peaks filtered out, leaving a singular peak in the slight negative direction. However, after further analysis, we found that the Bragg peaks measured in each SBP image reconstruction were actually the noise of the detectors

at the edge of the frame. This is best illustrated in the missing peak in the repaired SBP y -profile reconstruction in Figure 4.2. No real Bragg peak from the proton beam was measured, as the number of cones converging at the four Compton camera detectors at $x = \pm 420$ mm overrode the detection of any other prompt gammas coming from the water phantom. The double peaks in the 1-D profiles for the y - and z -axes were revealed to be the concentration of cones at the four detectors.

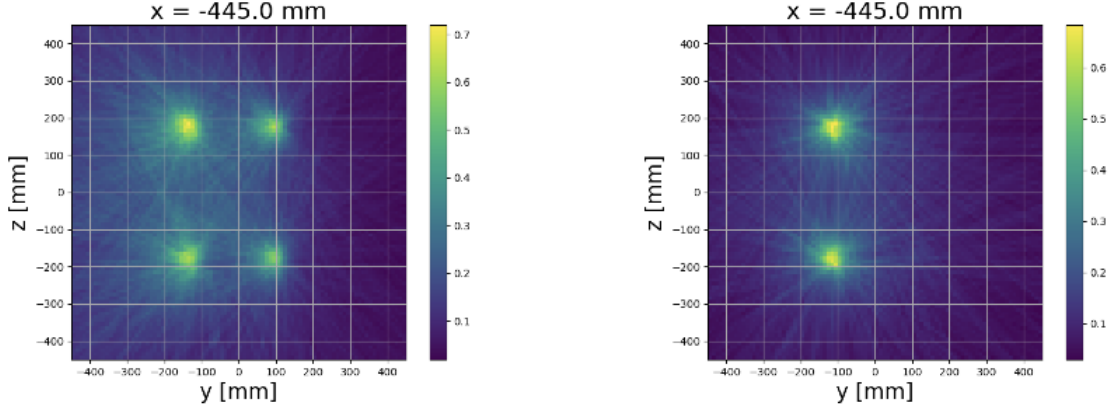


Figure 4.9: Dataset 14, SBP 270G, yz -profiles, control vs. repaired.

Interestingly, for every SBP reconstruction with repaired data, two of the four detectors at positive y -values were filtered out by the FCN model as noise, as illustrated in Figure 4.9. It was hypothesized that the detectors were denoised due to different placements of the detectors during training compared to testing. For the simulated training data inputted into the model, the detectors were placed below the patient, on the positive y -axis. However, in the experimental setup, the detectors were located on the positive and negative x -axis. This suggests the training setup cannot be generalized to other detector angles, necessitating subsequent reorientation of the data.

4.2 Reoriented Experimental Triples Data Reconstructions

We re-ran CORE on each of the 24 cases of reoriented triples experimental data. As with the first round of reconstructions, we compared the control reconstructions from the original runs to the new repaired reconstructions that used the reoriented data for each algorithm. Once again, we examined the same criteria for each x -, y -, and z -profile as well as the 2-D yz beam profile.

4.2.1 Reoriented SBP Reconstructions

When analyzing SBP reoriented reconstruction results, we found that the reorientations had no significant effect on the data except displaying all four Compton camera detectors. However, because of the narrowed field of view of the reconstruction area, the sharp points of the detectors were mostly filtered out. The Bragg peaks in the x -profiles were not improved from the first repaired reconstructions, and in most cases resembled a linear shape instead of a distinct Bragg peak. The y - and z -profiles displayed slight improvement in gamma emission values.

In the control reconstructions, 9,831 events out of 16,707 total events were kept. In the repaired reconstructions, 9,078 events out of 16,706 total events were kept.

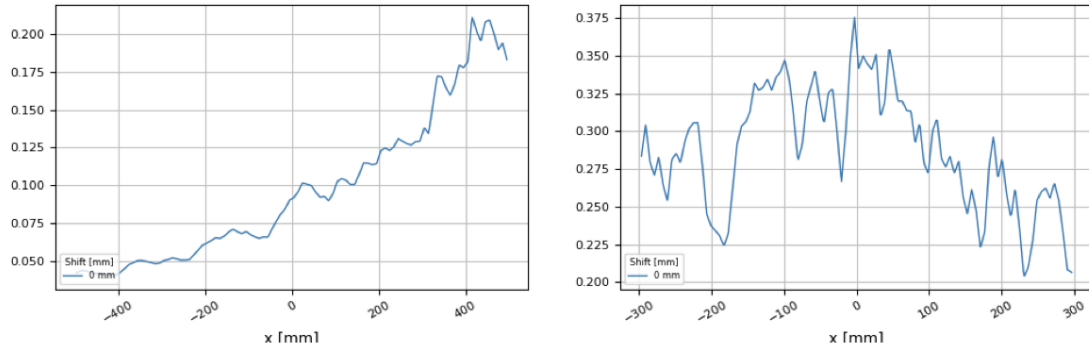


Figure 4.10: Dataset 11, SBP 90G, x -profiles, control vs. repaired.

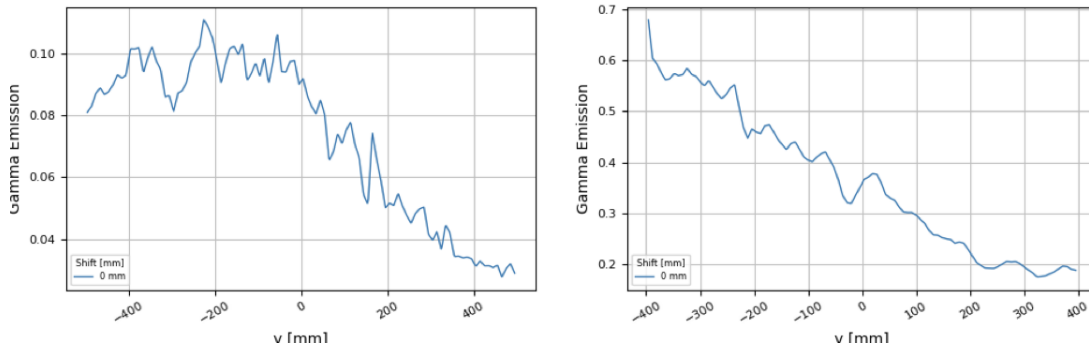


Figure 4.11: Dataset 11, SBP 90G, y -profiles, control vs. repaired.

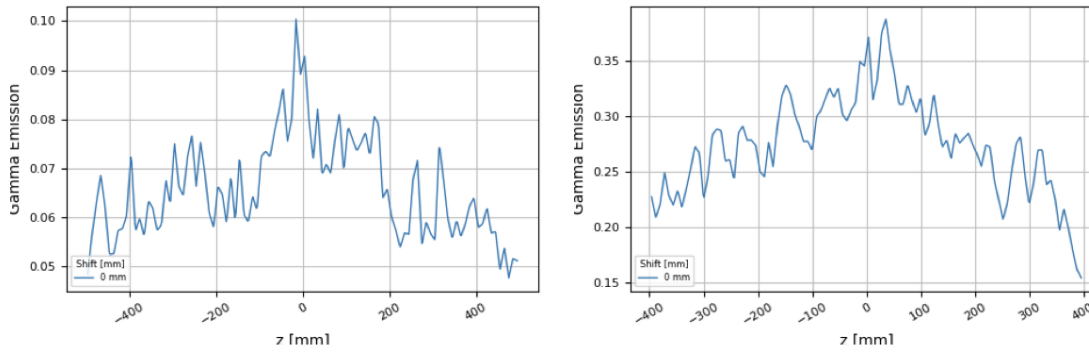


Figure 4.12: Dataset 11, SBP 90G, z -profiles, control vs. repaired.

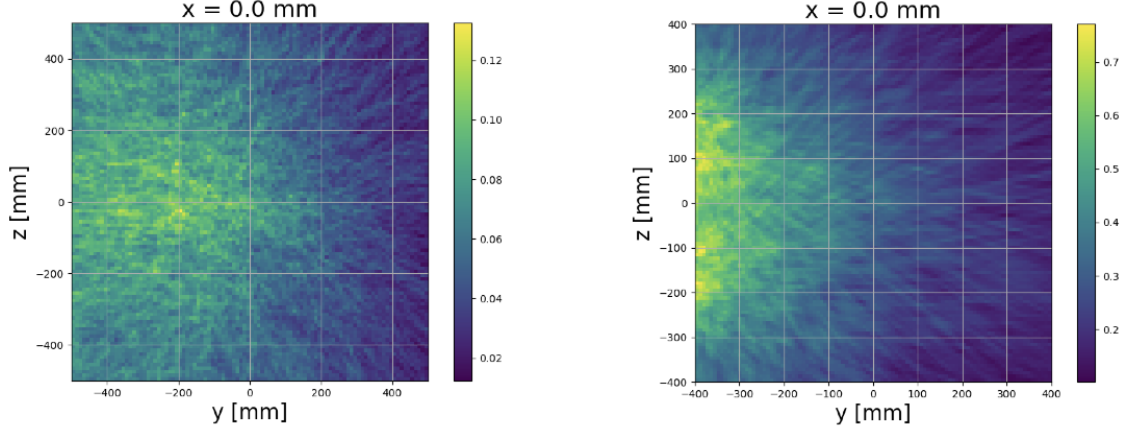


Figure 4.13: Dataset 11, SBP 90G, yz -profiles, control vs. repaired.

4.2.2 Reoriented KEM Reconstructions

Interestingly, the 1-D profiles of the repaired and reoriented data reconstructed with KEM were better than they were in the original repaired reconstructions. This was true for both 90-degree and 270-degree gantry angles. In the x -profiles, although there was still a significant amount of noise, a clearer, smoother shape resembling a Bragg peak in the graph could be observed, suggesting that reorientation had allowed the algorithm to filter out noise more accurately.

Additionally, the peak gamma emission recorded in the x -profiles increased to upwards of 80%, demonstrating a measured, concentrated Bragg peak. Similarly, for the y - and z - profiles, the gamma emissions increased significantly although retaining approximately the same profile shape. Some spikes of gamma emission present in control reconstructions appeared to be filtered out for the y - and z -profiles, but did not meaningfully contribute to improved symmetry of the profiles. Despite the substantial improvement in the 1-D beam profiles, similar to the original repaired reconstructions, the reoriented KEM 2-D beam slices contained almost entirely noise.

In the control reconstructions, 26,917 events out of 42,373 total events were kept. In the repaired reconstructions, 23,776 events out of 42,370 total events were kept.

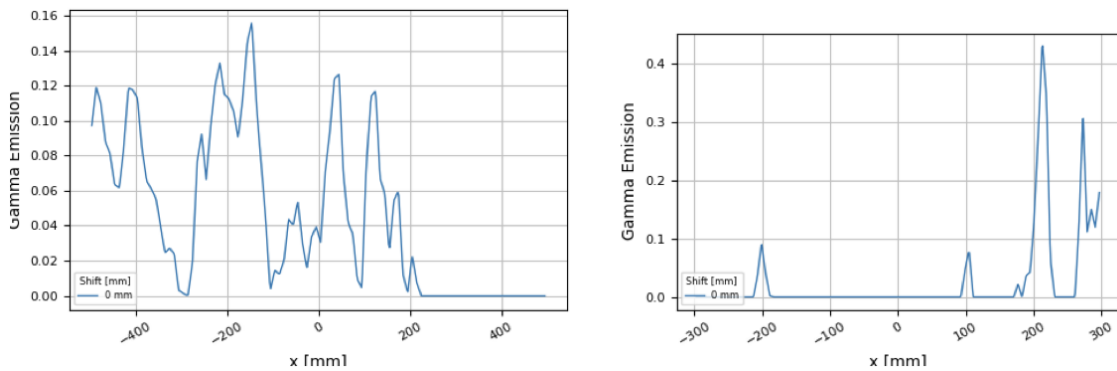


Figure 4.14: Dataset 17, SBP 90G, x -profiles, control vs. repaired.

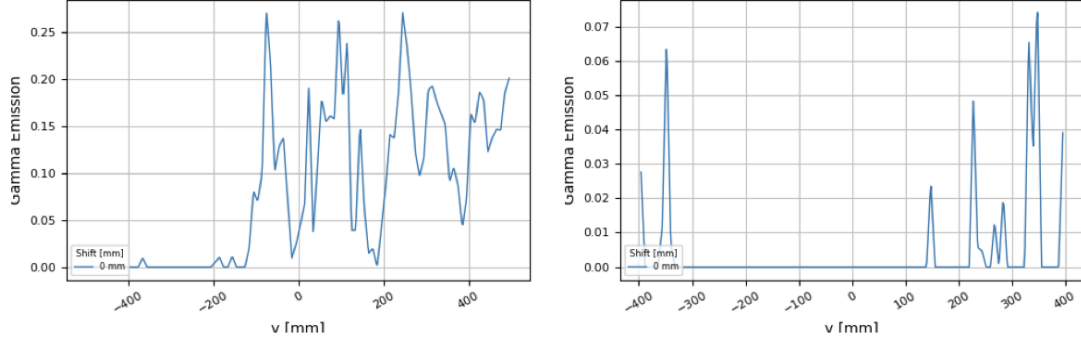


Figure 4.15: Dataset 17, SBP 90G y -profiles, control vs. repaired.

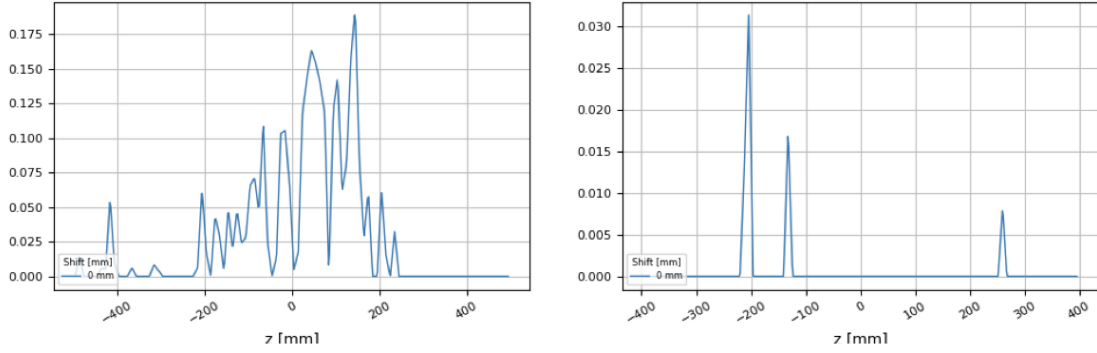


Figure 4.16: Dataset 17, SBP 90G, z -profiles, control vs. repaired.

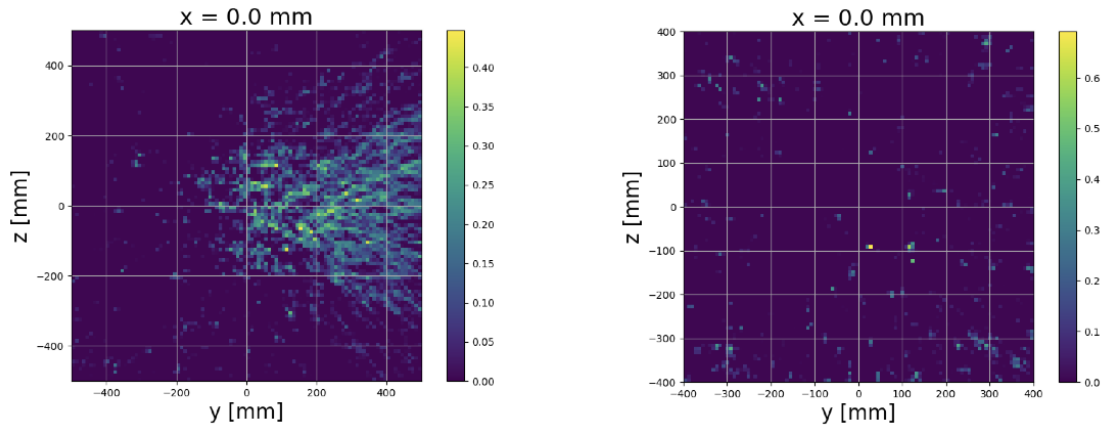


Figure 4.17: Dataset 17, SBP 90G, yz -profiles, control vs. repaired.

5 Conclusions

In proton beam radiation therapy, the proton beam is administered to the patient and the depth of the radiation dose can be verified in real-time via the prompt gamma rays released when the beam

collides with tissue. Compton cameras are used to reconstruct the source of the prompt gamma rays that it detects. In this experiment, we sought to improve the ability of Compton cameras to identify that source. By doing so, we know what tissue the proton beam interacts with in radiation therapy, allowing doctors to monitor the location of the proton beam’s Bragg peak in the patient and make any treatment adjustments.

Because of the inherent noisiness of thousands of prompt gammas entering a detector at the speed of light every second, Compton Cameras often misclassify the triple scatter events of the prompt gamma. Events can be misclassified as misordered triples detected in the wrong order, Double-to-Triples with two scatters from a single prompt gamma and the third scatter as noise, or most commonly, false triples, three independent events classified as a triple scatter. We posited that machine learning algorithms can be used to extract value from the misclassified events, potentially improving image reconstructions by reducing inaccuracies. While false triples cannot be useful in reconstructions, Doubles-to-Triples can be restored to double events and misordered triples can be restored to true triples.

The first step in the pipeline was training deep learning models on simulated triple scatter data, where the detectors were located on the positive y-axis. We used a fully trained Fully Connected Network (FCN) model to classify experimental data where the detectors were located on the x-axis, with 90 degree and 270 degree gantry angles. To reconstruct the control and repaired data, we used two image reconstruction – Kernel Expectation-Maximization (KEM) and Simple Back Projection (SBP). While SBP demonstrated some potential with sharp Bragg peaks and visible beams, KEM did not perform as well, and displayed mostly external noise in results. The controlled image reconstruction performed the best, followed closely by SBP, and then KEM, a distant third.

We noticed that because there was a difference between the detector axes between the simulated data used for training and the experimental data, further experiments were needed. We created a script to flip the X and Y axes of the experimental data to align the training data on the same axis as the experimental data, re-ran the FCN model on the reoriented data, and then changed the data back for image reconstruction.

After constructing the new images, we noticed a stark improvement in KEM. Though KEM was not as proficient as SBP or the control state at reconstructing the images, there were clear improvements demonstrated from KEM with much higher gamma emission at the Bragg peak and fairly centered y - and z -profile graphs.

In the future, we can explore better aligning the axes across both sets of data, including developing more generalizable machine learning models to classify multiple data with multiple varying factors, including different orientations, positionings, densities, and energy levels of beams. One tangible step that can be done is to see if the deep learning models respond to training the detectors on a different axis. The ultimate goal for this research is to conduct both classification and image reconstruction in real-time such that physicians can correct the position of the proton beam in the patient without delay.

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 Safavi and Ren additionally acknowledge support by NIH. We acknowledge the computational resources in the UMBC High Performance Computing Facility (hpcf.umbc.edu) and the financial contributions from NIH, NSF, CIRC, and UMBC for this work.

References

- [1] Carlos A. Barajas, Matthias K. Gobbert, and Jerimy C. Polf. Deep residual fully connected neural network classification of Compton camera based prompt gamma imaging for proton radiotherapy. *Front. Phys.*, 11:903929, 2023.
- [2] Angelo Calingo, Bikash Gautam, Peter L. Jin, Sidhya Pathak, Michelle Zhao, Hussam Fateen, Harrison Lewis, Matthias K. Gobbert, Vijay R. Sharma, Lei Ren, Ananta Chalise, Stephen W. Peterson, and Jerimy C. Polf. Improving proton beam radiotherapy by classifying simulated patient data in Compton camera imaging with neural networks. Technical Report HPCF–2025–5, UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2025.
- [3] Michael O. Chen, Julian Hodge, Peter L. Jin, Ella Protz, Elizabeth Wong, Ruth Obe, Ehsan Shakeri, Mostafa Cham, Matthias K. Gobbert, Carlos A. Barajas, Zhuoran Jiang, Vijay R. Sharma, Lei Ren, Sina Mossahebi, Stephen W. Peterson, and Jerimy C. Polf. Improving gamma imaging in proton therapy by sanitizing compton camera simulated patient data using neural networks through the bride pipeline. Technical Report HPCF–2024–5, UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2024.
- [4] Joseph Clark, Anaise Gaillard, Justin Koe, Nithya Navarathna, Daniel J. Kelly, Matthias K. Gobbert, Carlos A. Barajas, and Jerimy C. Polf. Sequence-based models for the classification of Compton camera prompt gamma imaging data for proton radiotherapy on the GPU clusters taki and ada. Technical Report HPCF–2022–12, UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2022.
- [5] Jonathan R. Hughes and Jason L. Parsons. FLASH radiotherapy: Current knowledge and future insights using proton-beam therapy. *Int. J. Mol. Sci.*, 21(18):6492, 2020.
- [6] Raj Kumar Parajuli, Makoto Sakai, Ramila Parajuli, and Mutsumi Tashiro. Development and applications of Compton camera — a review. *Sensors*, 22:7374, 2022.
- [7] Jerimy C. Polf and Katia Parodi. Imaging particle beams for cancer treatment. *Phys. Today*, 68(10):28–33, 2015.
- [8] Farshad Safavi, Stephen W. Peterson, Sina W. Mossahebi, Ananta K. Chalise, Vijay R. Sharma, Matthias K. Gobbert, Jerimy C. Polf, and Lei Ren. The first nozzle-mounted Compton camera prompt gamma imaging system for in vivo proton therapy dose verification. *IEEE Transactions on Instrumentation and Measurement*, submitted (2026).