

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

Sophia Lopez*

*Department of Mathematics and Statistics
University of North Carolina at Greensboro, USA*

Abby Nam*

*Department of Psychology
Lafayette College, USA*

Ayan Kabaria*

*River Hill High School
Howard County, MD, USA*

Muhammad Khalid*

*Department of Mathematics
Baruch College, CUNY, USA*

Ertan Dogan*

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

Sidhya Pathak*

*Department of Computer Science
University of Virginia, USA*

Victor Sandrin*

*Department of Neuroscience
University of Arizona, USA*

Xueying Sun*

*Department of IT and Management
Illinois Institute of Technology, USA*

Ehsan Shakeri

*Dept. of Mathematics and Statistics
U. of Maryland, Baltimore County, USA*

Harrison Lewis

*Dept. of Mathematics and Statistics
U. of Maryland, Baltimore County, USA*

Hussam Fateen

*Dept. of Mathematics and Statistics
U. of Maryland, Baltimore County, USA*

Matthias K. Gobbert

*Dept. of Mathematics and Statistics
U. of Maryland, Baltimore County, USA*

Farshad Safavi

*Department of Radiation Oncology
U. of Maryland School of Medicine, USA*

Ananta Chalise

*Department of Radiation Oncology
U. of Maryland School of Medicine, USA*

Lei Ren

*Department of Radiation Oncology
Northwestern University, USA*

Stephen W. Peterson

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

Jerimy C. Polf

*M3D, Inc.
USA*

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.

Index Terms—Proton beam therapy, Compton cameras, machine learning, image reconstruction, deep learning models

I. INTRODUCTION

Proton beam radiotherapy is a cancer treatment that uses high-energy protons to destroy tumors. Compared to other conventional treatments, such as X-ray radiotherapy, protons deposit most of their energy at a specific depth known as the Bragg peak [11]. As such, in an ideal setting, the tumor receives the majority of the radiation, mitigating the irradiation of surrounding healthy tissue. To maximize the benefits of proton beam radiotherapy, clinicians must accurately determine the location of the Bragg peak to ensure healthy tissue is not damaged by radiation.

To achieve this task, clinicians may use detectors known as Compton cameras to capture real-time information. When protons interact with patient matter, they emit prompt gamma rays that the Compton camera can detect, providing valuable information about the beam's path and the location of the

*These authors contributed equally to this work.

Bragg peak. This information allows the clinicians to adjust the gantry angle and proton beam energy in real time such that the tumor can be correctly targeted.

However, Compton cameras are limited by their finite time resolution. As a result, distinct gamma rays may be incorrectly classified as a single event, or the prompt gamma interactions may be misordered. The resulting reconstructed images may then contain distortion and noise, making it more difficult to accurately locate the proton beam in the body [11]. To address these deficits, machine learning models can be trained to classify each prompt gamma interaction, and then repair or remove unusable events. The current study utilized a deep Fully Connected Network (FCN) implemented through the Big Data REU Integrated Development and Experimentation (BRIDE) platform to classify 24 experimental setups utilizing real-world triple-scatter phantom data.

To improve the denoising of scattering data for Compton Camera image reconstruction in proton beam radiotherapy, this work makes the following contributions:

- Generated 1-dimensional and 2-dimensional image reconstructions of the proton beam profile for experimental triple scatter data.
- Compared and evaluated the effectiveness of Simple Back Projection (SBP) and Kernel-Expectation Maximization (KEM) for reconstructions.
- Created and implemented a Python script to reorient the experimental triples data to improve image reconstruction quality and prevent erroneous denoising of detectors.

The code of this work is publicly available in a GitHub repository [7].

This paper is structured as follows: Section II provides background on proton beam radiotherapy, Compton camera imaging, and the motivation for machine learning-based scatter classification. Section III describes the experimental methodology, including machine learning prediction, experimental data processing, beam reconstruction, and data reorientation. Section IV presents the experimental reconstruction results and evaluates the impact of scatter correction and detector reorientation. Finally, Section V concludes the paper and discusses future work.

II. APPLICATION BACKGROUND

A. Proton Beam Radiotherapy

Proton beam therapy is a type of external beam radiation therapy that uses high-energy proton beams to deposit radiation at tumor sites, destroying cancerous tissue. Unlike many other radiation therapies, such as X-ray therapy, that have a gradual dose deposition pattern resulting in irradiation of healthy tissue sites around the tumor, proton beam therapy deposits the majority of the radiation at a localized point known as the Bragg peak. During treatment, the Bragg peak must be located directly on the tumor to deposit radiation precisely at the correct location to avoid damaging healthy tissue; it cannot undershoot or overshoot the tumor, as seen in Fig. 1. This is only possible if clinicians are able to determine the exact location of the Bragg peak.

B. Compton Camera Image Reconstruction

Prompt gamma rays are emitted when protons in the beam interact with patient matter and can be captured by detectors called Compton cameras [10]. When prompt gammas interact with the camera, the device captures the scattering's spatial (x_i, y_i, z_i) coordinates and energy level (e_i) . From the captured scattering data, the Compton cone of emission with multiple detectors inside the camera can be used to statistically reconstruct the location of the source of the prompt gammas [9]. The prompt gamma origin locations can then be used to visually determine the path of the proton beam and the Bragg peak location through image reconstruction algorithms. However, due to the Compton camera's finite time resolution, gamma scatterings are detected almost simultaneously, and thus, in the incorrect order, leading to excessive noise in the resulting image reconstructions. This renders the reconstructions blurry and wholly unusable for determining whether the Bragg peak is accurately placed on the tumor location.

1) *Scattering Types*: There are multiple types of scattering events detected by the Compton camera; in order to separate out false events that create image noise due to mischaracterized scattering sequences, the following groupings are used:

- 1) True Triples: these consist of three sequential interactions with the Compton camera and occur in six distinct orderings (123, 132, 213, 231, 312, 321). Only the 123 ordering can be used for image reconstruction.
- 2) Doubles to Triples (DtoT): these consist of a double interaction and an independent single interaction that are detected as the same scattering. There are six distinct orderings of DtoTs (124, 134, 214, 234, 324, 314). The "4" signifies to the second prompt gamma interaction during misdetection.
- 3) False Triples: these consist of three independent events that are falsely detected as a single scattering. They result in noisy images and must be discarded during image reconstruction.

C. The Need for Machine Learning

The limited efficiency and resolution of Compton cameras introduces several classes of misidentified events. Applying machine learning to each case can recover misordered triples, improve event identification, and remove spurious events. The removal and repair of scatter events by the machine learning model results in a denoised and subsequently sharper image that is more useful for locating the Bragg peak.

D. Related Works

Neural networks have been studied as a mechanism to denoise prompt gamma scatter data for the reconstruction of radiotherapy beam images. In [4], [3], and [2] neural networks were trained to correct GEANT4-simulated constant-density water phantom data as well as voxelized patient phantom data. The ability to repair triple scatter events was demonstrated in [1]. Most recently, the Compton cameras capturing these scatters were relocated to a nozzle-mounted position in [12],

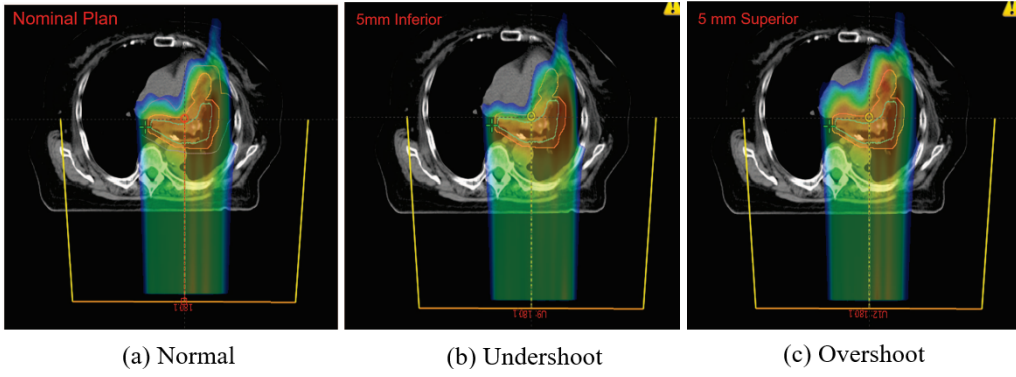


Fig. 1: Range and location uncertainties of the proton beam and Bragg peak. (a) shows the Bragg peak placed precisely on the tumor, (b) shows the beam undershooting the tumor, and (c) shows the beam overshooting the tumor [8].

motivating a new approach to reorient the neural networks to the data.

III. METHODOLOGY

A. 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 reconstructions. 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.

B. 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.

C. 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.

D. 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 [4]. 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 2. 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].

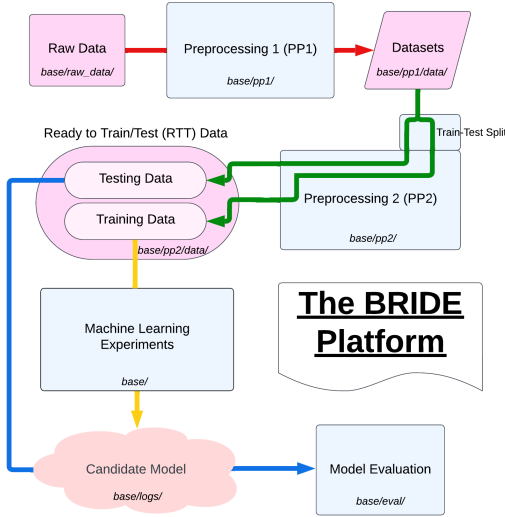


Fig. 2: Flowchart of the structure of the BRIDE platform [5].

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.

E. 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 [12]. 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.

F. 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 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

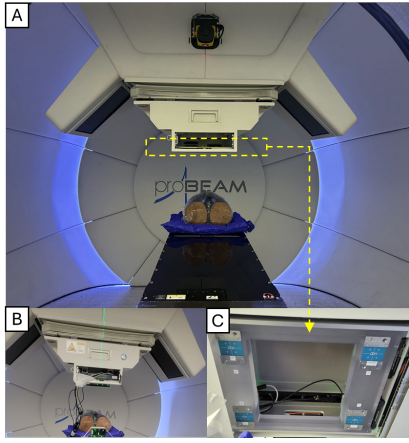


Fig. 3: Experimental proton beam imaging setup used to acquire the prompt gamma datasets. (a) Clinical proton therapy gantry with the nozzle-mounted quad Compton camera prompt gamma imaging (PGI) system. (b) Detector mounted beneath the treatment nozzle during phantom irradiation. (c) Bottom view of the detector assembly [12].

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-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.

G. 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.

H. 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 [6].

IV. RESULTS

A. Overall Results

SBP 1-D profiles and 2-D image reconstructions generally included a sharper and more distinct Bragg peaks as well as cleaner, more centered beams than KEM image reconstructions, though all reconstructions retained a significant amount of noise. In particular, SBP tended to filter out two of the four Compton detectors in the 2-D yz -beam profiles. However, after a second run of SBP and KEM image algorithms with reoriented experimental data, KEM saw 1-D profile improvements in the clarity and peak gamma emission, but no significant differences were observed for SBP [6].

B. Experimental Triples Data Reconstructions

From the experimental triples data provided by gantry setup experiments, a total of 96 image reconstructions were produced from 24 datasets. For each dataset, the results of control reconstructions and repaired reconstructions were compared for each algorithm. On the 1-dimensional x -axis profile, a single clear, sharp Bragg peak with a near 100% gamma emission was expected, and on the 1-dimensional y - and z -axis profiles, a centered, symmetric parabolic profile was expected. The 2-dimensional yz beam profiles were examined as well, expecting a centered, circular beam slice.

C. Initial Beam Reconstructions

1) *SBP Image Algorithm Reconstructions*: Reconstructions created with the SBP algorithm displayed visible measured Bragg peaks in the x -axis profiles, although centering on the edge of the field of view, at the position of the detectors. In the y -axis and z -axis profiles for control data reconstructions, there were double peaks symmetrical about the center. It was signified that these double peaks represented the two detectors on each axis, totaling to four detectors in the field of view. 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 phantom on the patient field.

For dataset 11, for which reconstructions are displayed, 9,831 events out of 16,707 total events were kept by the SBP algorithm. In the repaired reconstructions, 5,593 events out of 9,049 total filtered events were kept.

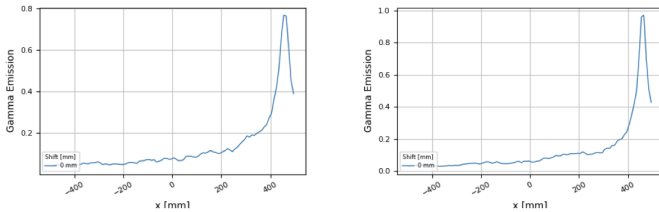


Fig. 4: Dataset 11, SBP 90G, x -profiles, control vs. repaired.

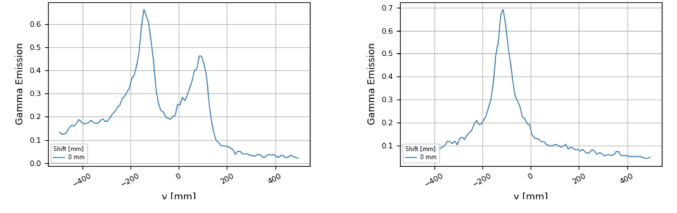


Fig. 5: Dataset 11, SBP 90G, y -profiles, control vs. repaired.

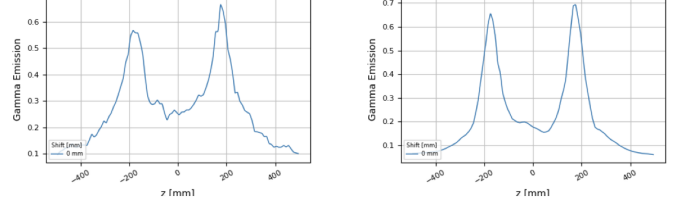


Fig. 6: Dataset 11, SBP 90G, z -profiles, control vs. repaired.

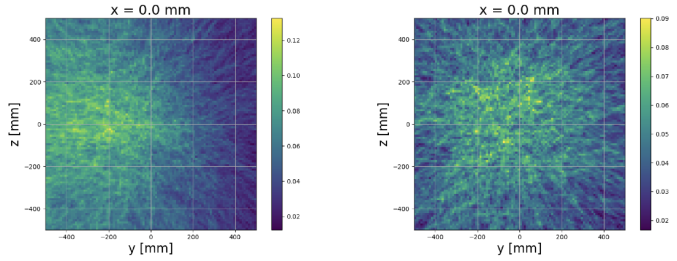


Fig. 7: Dataset 11, SBP 90G, yz -slices, control vs. repaired.

2) *KEM Image Algorithm Reconstructions*: Reconstructions with the KEM algorithm, on both control and repaired data, included highly unstructured and mostly noisy 1-D profiles and 2-D images. Although some reconstructions on repaired data included Bragg peaks in the x -axis profile, the recorded gamma emission values for those Bragg peaks were often negligible, which led to the assumption of these peaks being just noise.

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

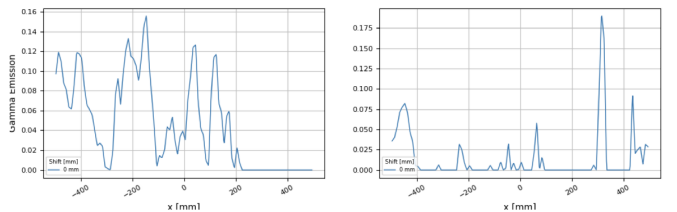


Fig. 8: Dataset 17, KEM 90G, x -profiles, control vs. repaired.

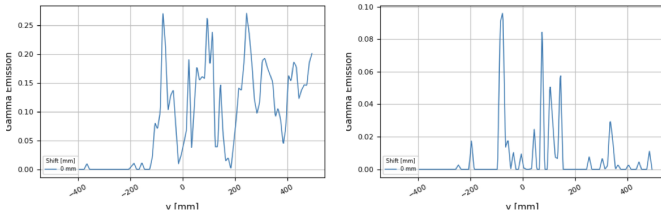


Fig. 9: Dataset 17, KEM 90G, y-profiles, control vs. repaired.

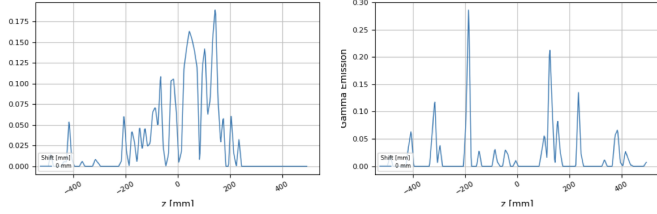


Fig. 10: Dataset 17, KEM 90G, z-profiles, control vs. repaired.

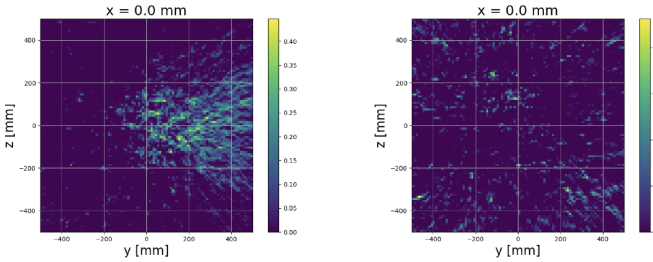


Fig. 11: Dataset 17, KEM 90G, yz-slices, control vs. repaired.

D. Detector Anomalies

In most repaired reconstructions using the SBP algorithm, the y-axis profile of reconstructions included one of the double peaks filtered out, leaving a singular peak in the slight negative direction. It was hypothesized that the filtering out of two detectors in the repaired reconstructions was due to the orientation of the training data versus the experimental data. While the simulated data that the model was trained on included the detectors underneath the patient, on the positive y-axis, the experimental setup detectors were installed on the snout, beside the patient on the positive and negative x-axis. The rotational difference in these two setups revealed that the classification model was not generalizable to different environments, and thus a constant setup of orientation across all cases was required.

E. Reorientation Experimental Triples Data Reconstructions

The CORE pipeline was re-run on each of the 24 cases of reoriented triples experimental data. Based on the gantry angle of the original experiments, the data was reoriented by flipping the positive x-coordinates to either the positive or negative y-coordinates. The reoriented data was then run through the FCN model and converted back to the original orientations to run reconstructions. The field of view of all reconstructions was also narrowed as to omit the noise from the Compton

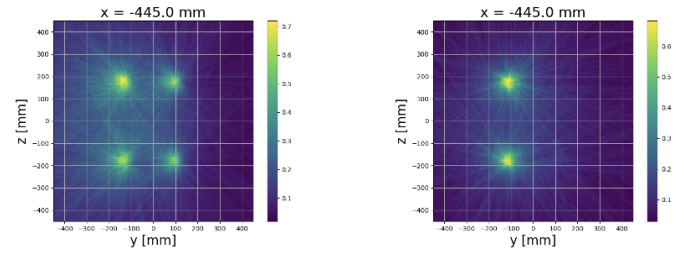


Fig. 12: Dataset 14, SBP 270G, yz-profiles, control vs. repaired. Four detectors displayed in the controlled reconstruction while two were filtered out in the repaired.

detectors. As with the first round of reconstructions, the control reconstructions from the original runs were compared to the new repaired reconstructions that used the reoriented data for each algorithm. Once again the same criteria was examined for each x-, y-, and z-profile as well as the 2-D yz beam slice.

1) *SBP Image Algorithm Reconstructions*: Reoriented reconstructions using the SBP algorithm resulted in no significant change on the data as before reorientations 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 2-D yz-profiles additionally became shifted in the far negative y-direction. The y- and z-profiles displayed slight improvement in gamma emission values and no longer displayed the double peak of the four detectors.

Interestingly enough, for dataset 11 shown, 16,707 total events were recorded in total. However, in the reoriented and repaired dataset, still 16,706 total events remained, 9,078 of which were kept by the SBP algorithm.

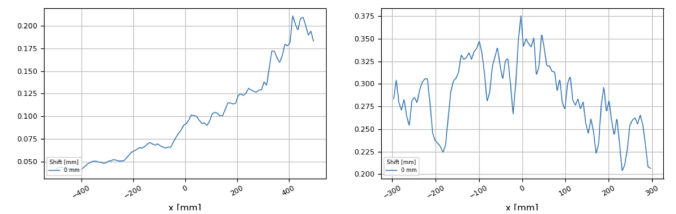


Fig. 13: Dataset 11, SBP 90G, x-profiles, control vs. reoriented repaired.

2) *KEM Image Algorithm Reconstructions*: Interestingly, the 1-D profiles of the repaired and reoriented data reconstructed with the KEM algorithm displayed better results than original repaired reconstructions, including sharper Bragg peaks and higher gamma emissions of up to 80%. 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 in the patient field, suggesting

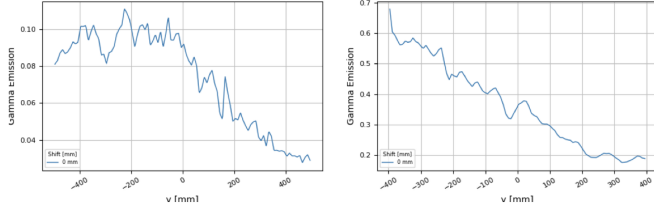


Fig. 14: Dataset 11, SBP 90G, y-profiles, control vs. reoriented repaired.

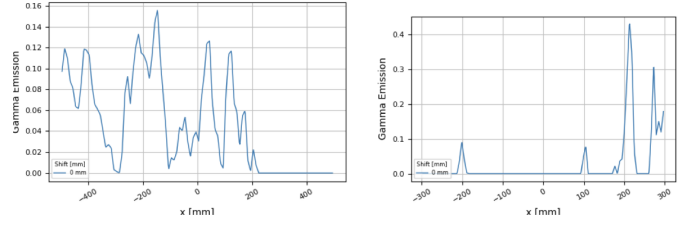


Fig. 17: Dataset 17, SBP 90G, x-profiles, control vs. reoriented repaired.

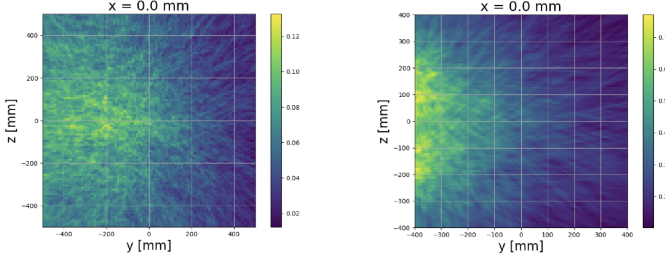


Fig. 15: Dataset 11, SBP 90G, z-profiles, control vs. reoriented repaired.

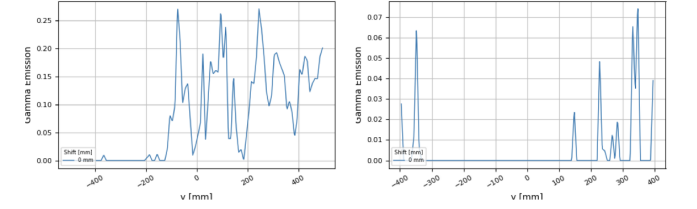


Fig. 18: Dataset 17, SBP 90G, y-profiles, control vs. reoriented repaired.

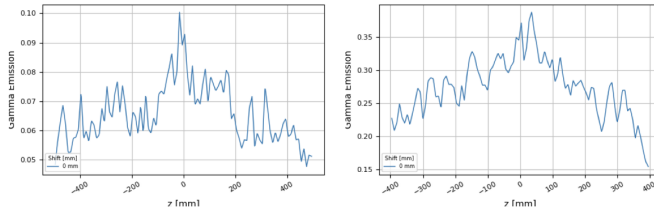


Fig. 16: Dataset 11, SBP 90G, yz-slice, control vs. reoriented repaired.

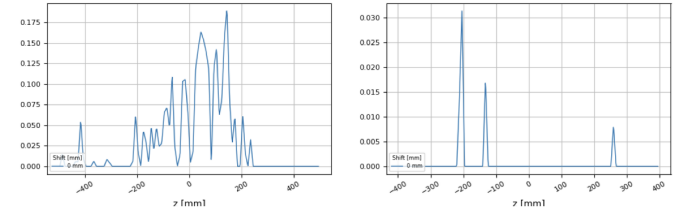


Fig. 19: Dataset 17, SBP 90G, z-profiles, control vs. reoriented repaired.

that reorientation had allowed the algorithm to filter out noise more accurately. 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. Although the y- and z-profiles saw less improvement and the 2-D yz-beam slices still showed scattered noise, the reorientations provided a significant development to the KEM reconstructions.

Similar to the case in dataset 11 for SBP reconstructions, in the original dataset 17, 42,373 total events were recorded, however there were still 42,370 events after the classification, 23,776 of which were kept by the KEM algorithm.

V. 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

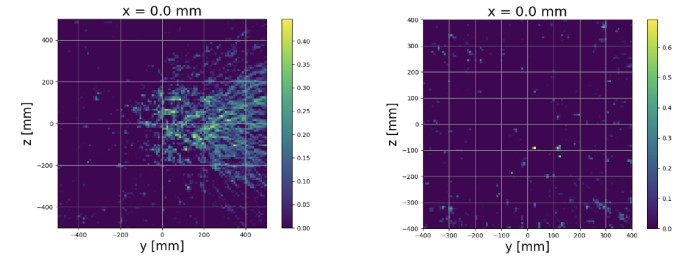


Fig. 20: Dataset 17, SBP 90G, yz-profiles, control vs. reoriented repaired.

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.

VI. FUTURE WORK

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 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] Carlos A. Barajas. *Neural Networks for the Sanitization of Compton Camera Based Prompt Gamma Imaging Data for Proton Radiotherapy*. Ph.D. Thesis, Department of Mathematics and Statistics, University of Maryland, Baltimore County, 2022.
- [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 Jeremy 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] 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 Jeremy C. Polf. Large-scale optimizations in proton beam radiotherapy by neural network denoising of robust simulated patient data. In *2025 IEEE International Conference on Data Mining Workshops (ICDMW 2025)*, pages 2807–2815, 2025.
- [4] 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 Jeremy 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.
- [5] Michael O. Chen, Julian Hodge, Peter L. Jin, Ella Protz, Elizabeth Wong, Ruth Obe, Ehsan Shakeri, Mostafa Cham, Matthias K. Gobbert, Carlos A. Barajas, Vijay R. Sharma, Sina Mossahebi, Lei Ren, Stephen W. Peterson, and Jeremy C. Polf. Improving gamma imaging in proton therapy by sanitizing compton camera simulated patient data using neural networks through the bride pipeline. In *2024 IEEE International Conference on Big Data (Big Data 2024)*, pages 7463–7470, 2024.
- [6] Ayan Kabaria, Muhammad Khalid, Sophia Lopez, Abby Nam, Ertan Dogan, Sidhya Pathak, Victor Sandrin, Xueying Sun, Ehsan Shakeri, Harrison Lewis, Hussam Fateen, Matthias K. Gobbert, Farshad Safavi, Ananta Chalise, Vijay R. Sharma, Lei Ren, Stephen W. Peterson, and Jeremy C. Polf. Experimental prompt gamma scatter classification and reorientation for proton beam image reconstruction. Technical Report HPCF–2026–5, UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2026.
- [7] Ayan Kabaria, Muhammad Khalid, Sophia Lopez, Abby Nam, Ertan Dogan, Sidhya Pathak, Victor Sandrin, Xueying Sun, Ehsan Shakeri, Harrison Lewis, Hussam Fateen, Matthias K. Gobbert, Farshad Safavi, Ananta Chalise, Vijay R. Sharma, Lei Ren, Stephen W. Peterson, and Jeremy C. Polf. Github repository. <https://github.com/big-data-lab-umbc/big-data-reu>, 2026.
- [8] Ruth Obe, Brandt Kaufmann, Kaelen Baird, Sam Kadel, Yasmin Soltani, Mostafa Cham, Matthias K. Gobbert, Carlos A. Barajas, Zhuoran Jiang, Vijay R. Sharma, Lei Ren, Stephen W. Peterson, and Jeremy C. Polf. Accelerating real-time imaging for radiotherapy: Leveraging multi-GPU training with PyTorch. In *2023 International Conference on Machine Learning and Applications (ICMLA 2023)*, pages 1735–1742, 2023.
- [9] Jeremy C Polf, Stephen Avery, Dennis S Mackin, and Sam Beddar. Imaging of prompt gamma rays emitted during delivery of clinical proton beams with a compton camera: feasibility studies for range verification. *Physics in Medicine & Biology*, 60(18):7085, 2015.
- [10] Jeremy C. Polf, Carlos A. Barajas, Stephen W. Peterson, Dennis S. Mackin, Sam Beddar, Lei Ren, and Matthias K. Gobbert. Applications of machine learning to improve the clinical viability of Compton camera based in vivo range verification in proton radiotherapy. *Front. Phys.*, 10:838273, 2022.
- [11] Jeremy C. Polf and Katia Parodi. Imaging particle beams for cancer treatment. *Phys. Today*, 68(10):28–33, 2015.
- [12] Farshad Safavi, Stephen W. Peterson, Sina Mossahebi, Ananta Chalise, Vijay R. Sharma, Matthias K. Gobbert, Jeremy C. Polf, and Lei Ren. The first nozzle-mounted compton camera prompt gamma imaging system for in vivo proton therapy dose verification. <https://arxiv.org/abs/2606.03978>, 2026. arXiv 2606.03978.