

Classification of Compton Camera Based Prompt Gamma Imaging for Proton Radiotherapy by Random Forests

Carlos A. Barajas

*Department of Mathematics and Statistics
University of Maryland, Baltimore County
Baltimore, Maryland, USA*

Gerson C. Kroiz

*Department of Mathematics and Statistics
University of Maryland, Baltimore County
Baltimore, Maryland, USA*

Matthias K. Gobbert

*Department of Mathematics and Statistics
University of Maryland, Baltimore County
Baltimore, Maryland, USA*

Jerimy C. Polf

*Department of Radiation Oncology
University of Maryland School of Medicine
Baltimore, Maryland, USA*

Abstract—Proton beam radiotherapy is a method of cancer treatment that uses proton beams to irradiate cancerous tissue, while simultaneously sparing healthy tissue. One promising method of real-time imaging during treatment is the use of a Compton camera, which can image prompt gamma rays that are emitted along the beam's path through the patient. However, because of limitations in the Compton camera's ability to detect prompt gammas, the reconstructed images are often noisy and unusable for verifying proton treatment delivery. Machine learning ensemble methods like random forests are able to automatically learn patterns that exist in numerical data, making them a promising method to analyze Compton camera data for the purpose of reducing noise in the reconstructed images. We conduct a hyperparameter search to find an optimal random forest model. We then present the results of the best performing random forest model, which demonstrate that this ensemble method is less effective than competing machine learning techniques for this application.

Index Terms—Proton beam therapy, Prompt gamma imaging, Compton camera, Machine learning, Random forest.

I. INTRODUCTION

Proton beams' primary advantage in cancer treatment as compared to other forms of radiation therapy, such as x-rays, is their finite range. The radiation delivered by the beam reaches its maximum, known as the Bragg peak, at the very end of the beam's range [1], [2]. Little to no radiation is delivered beyond this point. By exploiting the properties of the Bragg peak it is possible to only irradiate cancerous tissues, avoiding any damage to the surrounding healthy tissues [3]. However, due to uncertainties in the range of the beam, relative to important organs in the body, it is difficult to make optimal use of the Bragg peak during treatment.

The Compton camera is one method for real time imaging, which works by detecting prompt gamma rays emitted along

the path of the beam [4]. By analyzing how prompt gamma rays scatter through the Compton camera, it is possible to reconstruct their origin. It has been suggested that the range of the proton beam in the patient could be verified by using a Compton camera to image the prompt gamma rays emitted during proton treatment delivery. There are a couple hurdles in achieving this goal. The Compton camera does not explicitly record the sequential order of the prompt gammas that interact with the camera's internals. The physical limitations of the camera cause the interactions of any detected true triple, which consists of three interactions, to be output in a random order. It also often creates false couplings where multiple prompt gamma rays appear as originating from a single prompt gamma ray [5]. These problems make the reconstructed images of the beam's shape and depth based on Compton camera data noisy and unusable for practical purposes [6], [7].

We intend on using various machine learning technologies to try and correct the bad data produced by the Compton camera. If any one of the various possible algorithms or techniques can adequately clean the data then the usage of Compton cameras in proton beam therapy could become a reality. This would enhance a doctor's ability to make same-day treatment adjustments which inevitably leads to better treatment. In this work specifically we use random forests as they are well known and commonly used ensemble method.

The remaining sections of this work are organized as follows: Section II-A introduces the Compton camera and how it is used in proton beam therapy for cancer treatment and its current limitations. Section II-B gives a brief overview of random forests. Section II-C details how the data has to be handled, changed, and labelled for deep learning viability. Section III catalogues the hardware and software used for all research activities. Section IV details the performance of random forests used for prompt gamma classification. Section V presents our conclusions from this work.

II. BACKGROUND

A. Compton Cameras in Proton Beam Therapy

Proton beam therapy was first proposed as a cancer treatment in [8]. The beam's particles lose kinetic energy as they traverse the patient, the amount of radiation delivered by the beam is low at its entry point, gradually rising until the beam nears the end of its range, at which point the delivered dose rapidly reaches its maximum [2]. This point of maximum dose is called the Bragg peak and the discovery and additional details associated with it can be seen in [1], [2]. One of the most important things about the Bragg peak is that little to no radiation is delivered beyond the Bragg peak. These characteristics of proton beam therapy give it a distinct advantage over x-rays. Exploiting its finite range, medical practitioners can confine the radiation of the beam to areas solely affected by cancerous tumors allowing vital organs beyond the tumor to be spared [3].

In order to exploit the full advantages of proton therapy, many researchers are investigating methods to image the beam in real time as it passes through the patient's body [3], [6], [9]. One proposed method for real time imaging is by detecting prompt gamma rays that are emitted along the path of the beam using a Compton camera. As the proton beam enters the body, protons in the beam interact with atoms in the body, emitting prompt gamma rays which eventually collide with special modules inside the Compton camera. These modules have a non-zero time-resolution during which all interactions are recorded as occurring simultaneously. This means that the output ordering of the interactions is random and arbitrary. The collection of all internal interaction data that a camera module collects during a single readout cycle is referred to as an event [10]. Should two prompt gamma rays enter the same module of the Compton camera during the same readout cycle, the camera would record the resulting interactions as a part of the same event. This presents a notable problem because image reconstruction methods assume that all interactions in an event correspond to the same prompt gamma ray *and* are correctly ordered. For each interaction (also called a Compton scatter) an (x, y, z) location and the energy e deposited are recorded. At the higher dosage rates typically used in treatment, proton beams emit a larger number of prompt gamma rays per unit time, increasing the likelihood of bad events [11].

B. Ensemble Methods

Random forests are a commonly used ensemble method that average the results from decision trees for classification, regression, and other forms of machine learning. The studies in this paper use random forest models trained for classification accuracy. Within the random forest, each individual decision tree in itself is another form of machine learning classification, where based on characteristics of a sample, the sample is classified. Typically, random forests outperform single decision trees as they can average the results of many decision trees. The averaging of the individual decision trees directly tackles any overfitting that occurs when training individual

decision trees. A more in-depth explanation of the ensemble method is described in [12]. All hyperparameter studies use hyperparameter search methods defined in the `sklearn` library. Additionally the random forests themselves are the random forest classifier from `sklearn`. The results of the hyperparameter studies are shown in Section IV.

C. Data Preprocessing

Let each interaction be represented by 1, 2, or 3 for any given event. When a triple is correctly ordered we call that a 123 event. When a triple is misordered it is represented as 132, 213, 231, 312, or 321. Table I shows all possible orderings and the arrangement of the data for each class in their respective interactions. To explain the labelling system more, consider the 312 event from Table I. In the 312 event the data which should be interaction 1, $[e_1, x_1, y_1, z_1]$, shows up as interaction 2. Similarly, $[e_2, x_2, y_2, z_2]$ shows up as interaction 3 but should be interaction 2. Lastly, $[e_3, x_3, y_3, z_3]$ shows up as interaction 1 but should be interaction 3. To improve the performance of our random forests, we find it useful to append the Euclidean distances between interactions to the data output by the Compton camera. Each class has the same number of events with 140K events per class. Recall that only the 123 events are usable for reconstruction.

In order to normalize our data we treat our energy and spatial data differently. We fold all the energies into a single column for preprocessing and then normalize energy with the `PowerTransformer` (Yeo-Johnson) from `sklearn.Preprocessing`. For our spatial data we fold all of the data into x , y , and z ; then we normalize the columns with `MaxAbsScaler` from `sklearn.Preprocessing`. After normalization we defold the data back into our interaction formats as seen in Table I.

III. HARDWARE USED

The studies in this work use a distributed-memory cluster of 42 compute nodes with each with two 18-core Intel Xeon Gold 6140 Skylake CPUs (2.3 GHz clock speed, 24.75 MB L3 cache, 6 memory channels). Each node has 384 GB of memory (12×32 GB DDR4 at 2666 MT/s). The nodes are connected by a network of four 36-port EDR (Enhanced Data Rate) InfiniBand switches (100 Gb/s bandwidth, 90 ns latency). These nodes are contained in the cluster tank of the UMBC High Performance Computing Facility (HPCF), whose webpage at hpcf.umbc.edu can provide more details.

All studies and preprocessing used one or more of the following python packages with the respective version:

- Python 3.7.6,
- Numpy 1.18.1,
- Scipy 1.4.1,
- Pandas 1.1.0.dev0+690.g690e382 (configured for `icc 19.0.1.144 20181018`),

TABLE I

LIST OF THE INPUT CLASSES AND THEIR ACTUAL INTERACTION DATA IN THE FOLLOWING COLUMNS. THE INPUT CLASS ENCODES HOW THE ENERGY AND SPATIAL DATA ARE ORDERED IN THE DATA FILE FOR TRAINING. FOR EXAMPLE FOR A 312 EVENT $[e_1, x_1, y_1, z_1]$ SHOWS UP AS INTERACTION 3 WHEN IT SHOULD SHOW UP AS INTERACTION 1. SIMILARLY, $[e_2, x_2, y_2, z_2]$ SHOWS UP AS INTERACTION 1 BUT SHOULD BE INTERACTION 2. LASTLY, $[e_3, x_3, y_3, z_3]$ SHOWS UP AS INTERACTION 2 BUT SHOULD BE INTERACTION 3.

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

TABLE II

ALL POSSIBLE HYPERPARAMETER VALUES USED IN SEARCHING FOR THE OPTIMAL RANDOM FOREST.

Hyperparameter	Values
min_samples_split	2, 5, 10
min_samples_leaf	1, 2, 4
max_depth	10, 20, 30, 40, 50
max_features	auto, sqrt, log2
bootstrap	true and false

TABLE III

HYPERPARAMETER VALUES USED IN THE BEST RANDOM FOREST.

Hyperparameter	Value
min_samples_split	5
min_samples_leaf	2
max_depth	50
max_features	auto
bootstrap	false

IV. RESULTS

We train over a collection of hyperparameters using sklearn's RandomForestClassifier for classification and sklearn's RandomizedCVSearch for hyperparameter tuning given the collection of hyperparameters mentioned in Table II. We train our random forests on a MCDE 150MeV 0kMU beam dataset. Only the random forest with the best accuracy will be considered and evaluated.

Figure 1 is a confusion matrix which is generated by a random forest, whose parameters are listed in Table III, classifying the Monte-Carlo plus Detector Effects (MCDE) 150MeV 20kMU beam data. The first column of the table is the input class and the every proceeding column is the percentage of input which was assigned to class at the top of column. Each cell is colored by accuracy relative to the largest percentage present in entire confusion matrix.

We see that the dominant classification for each row is the input class itself. For the 123 class we see that the second and third highest classifications result in around 30% of the data and the remaining 10% is spread among the remaining classes. The observation that the 2nd and 3rd highest classifications soak around 30% of the data is consistent for each input class. There is no answer which can consistently explain what the input class, say 123, has in connection to a dominant misclassification like, say, 213. These behaviors are the same for Figures 2 and 3.

Lastly we notice that the random forest has a minimum correct classification accuracy of 52.9% and a maximum correct classification accuracy of 64.8% across all three testing sets. We know that of the 6 classes only one of them is usable for reconstruction. If we take a data file and do not use the random forest classifications we would expect that approximately 16% (1/6) would be usable for reconstruction. If we used the random forest classification to reorder the triples then we might expect that, on average, 58% of the data file would be usable for reconstruction.

	123	132	213	231	312	321
123	60.1	13.5	14.6	4.5	2.9	4.4
132	14.1	53.3	3.6	5.1	20.2	3.7
213	13.9	4.3	57.5	17.1	4.9	2.3
231	4.3	8.1	10.6	60.0	3.7	13.4
312	4.8	8.9	7.8	3.0	63.7	11.9
321	8.8	3.6	4.7	13.3	12.7	57.0

Fig. 1. Confusion matrix for a random forest trained on true triples data from a 150MeV 0kMU beam and tested on the MCDE 150MeV 20kMU beam.

	123	132	213	231	312	321
123	60.6	14.6	15.1	3.7	2.6	3.5
132	12.8	52.9	3.5	4.7	22.7	3.5
213	14.6	3.8	54.6	19.1	5.3	2.6
231	2.8	7.3	10.7	61.3	3.6	14.4
312	4.5	7.8	7.4	2.9	63.5	14.0
321	8.7	3.8	4.8	12.4	13.4	57.0

Fig. 2. Confusion matrix for a random forest trained on true triples data from a 150MeV 0kMU beam and tested on the MCDE 150MeV 100kMU beam.

	123	132	213	231	312	321
123	62.7	14.9	11.8	3.6	1.9	5.0
132	11.5	58.7	2.2	4.6	19.0	4.1
213	15.2	4.3	55.4	18.8	3.9	2.4
231	4.8	8.7	9.2	56.1	4.8	16.4
312	4.8	8.2	6.5	2.2	64.8	13.5
321	8.2	4.6	6.3	10.6	13.5	56.9

Fig. 3. Confusion matrix for a random forest trained on true triples data from a 150MeV 0kMU beam and tested on the MCDE 150MeV 180kMU beam.

We also have four additional MCDE testing tests at each of the previously used dose rates. All of the conclusions made about Figures 1, 2, and 3 all hold for the unlisted results mentioned. There are differences in the exact percentages but the general relationships discussed are identical. The results from the random forest are far below the minimum classification accuracy 80% for each class.

V. CONCLUSIONS

Compton cameras are promising tool which could enable real-time same-day adjustments to a patient's proton radiotherapy cancer treatment. This is because the Compton camera can capture the prompt gamma rays that are emitted by the patient's tissues when the proton beam enters the body. When the prompt gamma rays collide with the Compton camera internals the camera can record the spatial position and energy deposited. These values can then be used to reconstruct pictures of the original beam inside the patient. The are two major drawbacks which prevent the usage of Compton cameras in this manner are that the Compton cameras cannot correctly determine the order of the internal collisions and it cannot determine if two prompt gamma rays entered at the same time. Prompt gamma image reconstruction algorithms expect the interaction to be ordered for proper and accurate reconstruction. Any incorrect ordering will cause noise and render the resulting image unusable. There are 6 possible orderings for a prompt gamma ray which scatters inside the Compton camera three times but only one of the orderings is usable for reconstruction.

In this work we attempted to use random forests from `sklearn` to determine the correct ordering of our simulated Compton camera data. Rather than just using a single configuration we did a hyperparameter study to try and find the optimal hyperparameters for random forests on our data. We tested the best performing random forest on three simulated test data sets which closely resemble a real-world proton beam captured by a Compton camera. The best random forest performed rather poorly with a minimum classification accuracy of $\approx 53\%$ and a maximum classification accuracy of $\approx 65\%$ between our test sets. If we were to use the best random forest classification to correct a data file of balanced true triples then the amount of reconstruction viable data goes from 16% to 58%. We need a minimum classification accuracy of 80% for each class to be competitive with other methods. In addition to this, we prefer a classification accuracy greater than 95% for real-world usage. Based on these findings we do not have high hopes that other ensemble methods will perform well enough. We have already have great success using neural networks in [13]. Nevertheless, we intend to test other ensemble methods like K-means clustering and support vector machines to see if they could do better at understanding our data.

ACKNOWLEDGMENTS

This work is supported by the grant "CyberTraining: DSE: Cross-Training of Researchers in Computing, Applied Mathematics and Atmospheric Sciences using Advanced Cyberin-

frastructure Resources" from the National Science Foundation (grant no. OAC-1730250). The research reported in this publication was also supported by the National Institutes of Health National Cancer Institute under award number R01CA187416. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The hardware used in the computational studies is part of the UMBC High Performance Computing Facility (HPCF). The facility is supported by the U.S. National Science Foundation through the MRI program (grant nos. CNS-0821258, CNS-1228778, and OAC-1726023) and the SCREMS program (grant no. DMS-0821311), with additional substantial support from the University of Maryland, Baltimore County (UMBC). See hpcf.umbc.edu for more information on HPCF and the projects using its resources. Co-author Carlos Barajas additionally acknowledges support as HPCF RA.

REFERENCES

- [1] W. H. Bragg, "On the absorption of α rays, and on the classification of α rays from the radium," *London Edinburgh Dublin Philos. Mag. J. Sci.*, vol. 8, no. 48, pp. 719–725, 1904. [Online]. Available: <https://doi.org/10.1080/14786440409463245>
- [2] W. H. Bragg and R. Kleeman, "On the ionization curves of radium," *London Edinburgh Dublin Philos. Mag. J. Sci.*, vol. 8, no. 48, pp. 726–738, 1904. [Online]. Available: <https://doi.org/10.1080/14786440409463246>
- [3] J. C. Polf and K. Parodi, "Imaging particle beams for cancer treatment," *Phys. Today*, vol. 68, no. 10, pp. 28–33, 2015.
- [4] A. H. Compton, "A quantum theory of the scattering of x-rays by light elements," *Phys. Rev.*, vol. 21, no. 5, pp. 483–502, 1923.
- [5] R. Panthi, P. Maggi, S. Peterson, D. Mackin, J. Polf, and S. Beddar, "Secondary particle interactions in a Compton camera designed for *in vivo* range verification of proton therapy," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 3, pp. 383–391, 2021.
- [6] H. Rohling, M. Prieegnitz, S. Schoene, A. Schumann, W. Enghardt, F. Hueso-González, G. Pausch, and Fiedler, "Requirements for a Compton camera for *in vivo* range verifications of proton therapy," *Phys. Med. Biol.*, vol. 62, pp. 2795–2811, 2017.
- [7] P. G. Ortega, I. Torres-Espallardo, F. Cerutti, A. Ferrari, J. E. Gillam, C. Lacasta, G. Llosá, J. F. Oliver, P. R. Sala, P. Solevi, and M. Rafecas, "Noise evaluation of Compton camera imaging for proton therapy," *Phys. Med. Biol.*, vol. 60, pp. 1845–1863, 2015. [Online]. Available: <https://cds.cern.ch/record/2162923>
- [8] R. R. Wilson, "Radiological use of fast protons," *Radiology*, vol. 47, no. 5, pp. 487–491, 1946.
- [9] M. Fontana, J. Ley, D. Dauvergne, N. Freud, J. Krimmer, J. M. Létang, V. Maxim, M. Richard, I. Rinalida, and É. Testa, "Monitoring ion beam therapy with a Compton camera: Simulation studies of the clinical feasibility," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 4, no. 2, pp. 218–232, 2020.
- [10] J. C. Polf, P. Maggi, R. Panthi, S. Peterson, D. Mackin, and S. Beddar, "The effects of Compton camera data acquisition and readout timing on PG imaging for proton range verification," *IEEE Trans. Radiat. Plasma Med. Sci.*, pp. 1–1, 2021.
- [11] P. Maggi, S. Peterson, R. Panthi, D. Mackin, H. Yang, Z. He, S. Beddar, and J. Polf, "Computational model for detector timing effects in Compton-camera based prompt-gamma imaging for proton radiotherapy," *Phys. Med. Biol.*, vol. 65, no. 12, p. 125004, 2020.
- [12] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, 2001.
- [13] C. A. Barajas, G. C. Kroiz, M. K. Gobbart, and J. C. Polf, "Deep learning based classification methods of Compton camera based prompt gamma imaging for proton radiotherapy," UMBC High Performance Computing Facility, University of Maryland, Baltimore County, Tech. Rep. HPCF-2021-1, 2021. [Online]. Available: <http://hpcf.umbc.edu>