# A Multi-Step Machine Learning Approach to Directional Gamma Ray Detection

Matthew Durbin, *Student Member, IEEE,* Ryan Sheatlsey, *Student Member, IEEE,* Patrick McDaniel, *Fellow, IEEE,* and Azaree Lintereur, *Member, IEEE*

2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC) | 978-1-7281-7693-2/20/$31.00 ©2020 IEEE | DOI: 10.1109/NSS/MIC42677.2020.9507918

***Abstract*—Directional detection predicts the angular component of a gamma ray source's location by analyzing the distribution of counts received across an array of stationary detectors. The array's response to the source is a function of angle, as well as other factors such as distance, energy, and obstructions. The effectiveness of an angular prediction in a real-world environment is therefore dependent on the inclusion of these phenomena when processing the detector array data. With sufficiently representative training data that captures these variables, it is hypothesized that machine learning algorithms can aid in this angular prediction process due to their success in other complex data processing applications. A multi-step approach is introduced, in which machine learning algorithms are tasked with addressing specific complexities of the overall analysis. Initial results indicate that this multi-step method is a viable option which can be used to analyze different components of the array response. Presented here is a proof-of-concept with simulated datasets of three different isotopes, and measured datasets of two different isotopes. Preliminary results indicate that a multi-step machine learning approach improves the overall angle-prediction accuracy compared to a single phase machine learning algorithm and a least-squares comparison to a reference table.**

***Index Terms*—Machine Learning, Source Search, Spectroscopic Analysis**

## I. Introduction

THE localization of gamma ray sources is an important component of many homeland and nuclear security applications. By analyzing the distribution of counts across detectors within an array during a stationary acquisition, the angular component of a source's location can be predicted. Current analysis methods include reference table based least squares algorithms (RTLS) with stationary arrays [1], and maximum-likelihood based methods with mobile arrays [2]. The ability of these two methods to determine source direction is made possible by two main effects: differences in each detector's solid angle with respect to the source, and self-occlusion of the array. This yields a unique array response for each source angle at a fixed stand-off distance, allowing for an angular prediction by analysis of the detector outputs. A simulated angular response curve for the detector array used in this work is shown in Fig. 1.

Manuscript submitted December 20th, 2020. This work was supported in part by the U.S. Nuclear Regulatory Commission's Nuclear Education Program Fellowship Grant Program.

M. Durbin and A. Lintereur are with the Ken and Mary Alice Lindquist Department of Nuclear Engineering Department, Pennsylvania State University, University Park, PA 16802 USA (email: {mdurbin,atl21}@psu.edu).

R. Sheatsley and P. McDaniel are with the Department of Computer Science and Engineering, Pennsylvania State University, University Park, PA 16802 USA (email: sheatsley@psu.edu, mcdaniel@cse.psu.edu).

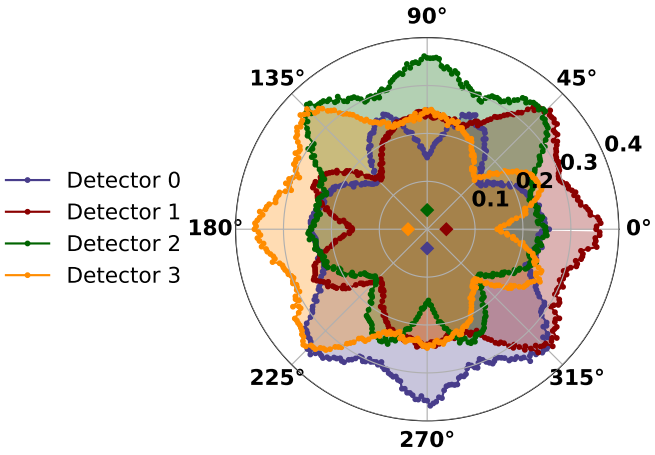


Fig. 1. Simulated angular response of the detector array used in this work for a 300 cm fixed source stand-off distance. Radial values represent the fraction of counts received in each detector, and diamonds represent detector positions.

In addition to angle, the array response to measurements in real-world scenarios will also be a function of geometric factors, gamma ray energy, and varying background. Thus, the data analysis methods must be robust to the phenomena that affect the array response.

Machine learning (ML) algorithms have shown promise in various radiation detection applications, and incorporating a multi-step ML analysis method has proven beneficial for some applications, such as pulse shape discrimination [3]. ML has not, however, been applied to stationary directional detection. To address the many factors relevant to this domain, a multi-step ML analysis is proposed. In this approach, dedicated steps are assigned to classify the raw data, and based on that classification, the data is ultimately processed by a specific angular prediction algorithm. As a proof-of-concept, two algorithmic phases are implemented in this analysis. The first phase classifies the isotope to address the energy dependence of the array response, and the second predicts the angular component of the source's location using isotope-specific training data. Preliminary results are presented, where performance is assessed in terms of accuracy, and compared against both a single-step ML analysis and an RTLS based approach.

## II. Methods

The array used in this work consists of four 5.08 cm by 10.16 cm by 40.64 cm (2"x4"x16") NaI detectors arranged in

978-1-7281-7693-2/20/$31.00 ©2020 IEEE

a 30 cm square with respect to the inner faces, with the 40.64 cm axis orthogonal to the source plane, as shown in Fig. 2.

Fig. 2. Detector array and associated electronics. Source holder (brown) can be seen behind the array.

Simulated datasets were acquired utilizing Monte Carlo n-Particle (MCNP6) [4], in which the detector array, point source, and concrete floor were modeled and energy binned pulse height tallies were recorded. Separate datasets of 10,000 individual trials each for $^{60}$Co, $^{137}$Cs, and $^{192}$Ir were simulated, with different random source locations 0.5-5 m away from the array center. Each trial was simulated with 10 million particles with energies corresponding to each isotope. Measured datasets included 225 $^{60}$Co (7 $\mu Ci$) trials and 150 $^{137}$Cs (30 $\mu Ci$) trials with count times five and two minutes respectively. Distances ranged from 0.5-5 m, and a CAEN digitizer was used for signal collection and pulse processing.

For the proof of concept approach, a k-Nearest Neighbor (KNN) algorithm was applied in both phases of the analysis, first for isotope identification and then for source location. A KNN algorithm takes the mode output value of the user specified number of nearest neighbors in the input feature space, as defined by the euclidean distance [5]. The simplified input features for the isotope classification phase were taken as the counts in each bin of a reduced ten-bin spectra, where each bin had an energy width of 0.18 MeV over the range of 0.2-2 MeV. As developing an isotope identification method is not the focus of this work, the reduced number of bins were used for simplicity, and the input feature scheme was not optimized. The inputs for the angular predictors were the fraction of counts received in each detector. Due to the need for large quantities of training data, simulated data was used for the training phase for both the simulated and measured tests.

To test the utility of the multi-step approach, datasets were trained in three ways: using single isotope training sets, using a training set that combines isotopes, and using the multi-step approach in which the training set used for the angular predictions is decided by the results of the isotope classifier. Additionally, an RTLS analog was tested in each case. Reference tables were generated from results produced with the source located at a distance of 3 m from the detector at each angle. The 3 m distance was chosen as it gave better results than other tested distances. All three isotopes were tested for the simulations, while due to availability only $^{60}$Co and $^{137}$Cs were tested for the measurements. A k-fold cross validation scheme[5] was implemented to iteratively split the datasets for training and testing, to mitigate the effects of a specific split. Results in the next section are therefore presented as a mean and standard deviation across the k-folds.

## III. Preliminary Results

The results of the simulated datasets are shown in Fig. 3, where a yellower shade indicates higher accuracy.

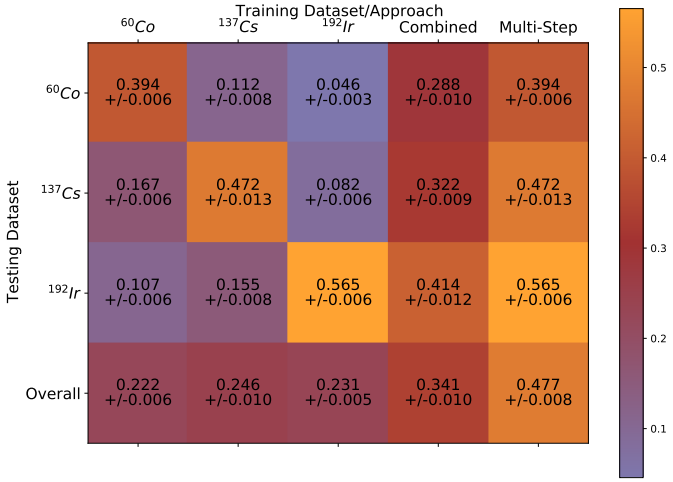


Fig. 3. Results of simulated tests

As expected, testing and training on the same isotope yielded the best results. The last row in Fig. 3 gives the overall results across all tested datasets. While there was improvement using a combined training dataset, the multi-step approach yielded the best results. Additionally, ML performance was greater than that of the RTLS method across all tests, for example, the overall multi-step approach of the LSRT yielded 0.3 % accuracy. Measured results are shown in Fig. 4.

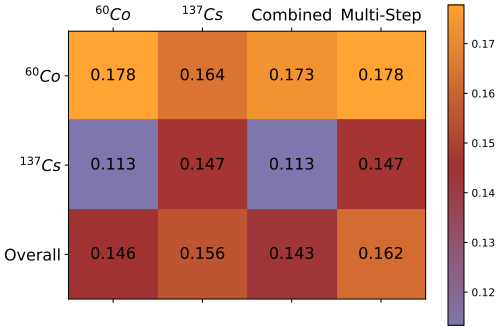


Fig. 4. Results of measured tests

Similar trends were seen with the measured tests: the best performance was achieved when the testing and training datasets were of the same isotope, there was an improvement in the overall results with the combined dataset, and the highest accuracy (0.162) was seen with the multi-step approach. Here again, ML out performed the RTLS approach. While results of a ML algorithm are largely dependent on the agreement of

testing and training data, the outcomes of this preliminary study demonstrate that the complexities associated with directional detection type measurements could benefit from a multi-step algorithmic approach.

## IV. Conclusion

To address the complexities associated with processing directional detection data, a multi-step ML approach is proposed with algorithms applied to dedicated phases of the analysis. Due to the energy dependence of the array's angular response, the preliminary scheme involved an isotope classifier, followed by an isotope-specific angular predictor. Results of experimentally measured and simulated datasets benefited from combined datasets, but saw the highest accuracy when the multi-step approach was used. The simulated tests, for example, saw an overall improvement of approximately 20% with the multi-step approach compared to using any single isotope in training. Future work will include investigating the allowable discrepancies between simulated and measured data when attempting to train on simulations, as well as expanding and optimizing the multi-step approach to address additional complexities in directional detection.

## References


[1] D. Hanna, L. Sagnières, P. J. Boyle, and A. M. MacLeod, "A directional gamma-ray detector based on scintillator plates," *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 797, pp. 13–18, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.nima.2015.06.019

[2] J. C. Curtis, R. J. Cooper, T. H. Joshi, B. Cosofret, T. Schmit, J. Wright, J. Rameau, D. Konno, D. Brown, F. Otsuka, E. Rappeport, M. Marshall, and J. Speicher, "Simulation and validation of the Mobile Urban Radiation Search (MURS) gamma-ray detector response," *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 954, no. August 2018, p. 161128, 2020. [Online]. Available: https://doi.org/10.1016/j.nima.2018.08.087

[3] C. Fu, A. Di Fulvio, S. D. Clarke, D. Wentzloff, S. A. Pozzi, and H. S. Kim, "Artificial neural network algorithms for pulse shape discrimination and recovery of piled-up pulses in organic scintillators," *Annals of Nuclear Energy*, vol. 120, pp. 410–421, 2018.

[4] T. Goorley, M. James, T. Booth, F. Brown, J. Bull, L. J. Cox, J. Durkee, J. Elson, M. Fensin, R. A. Forster, J. Hendricks, H. G. Hughes, R. Johns, B. Kiedrowski, R. Martz, S. Mashnik, G. McKinney, D. Pelowitz, R. Prael, J. Sweezy, L. Waters, T. Wilcox, and T. Zukaitis, "Initial MCNP6 release overview," *Nuclear Technology*, vol. 180, no. 3, pp. 298–315, 2012.

[5] F. Pedregosa, V. Gael, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and D. Edouard, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 1, pp. 2825–2830, 2011.