LAWRENCE LIVERMORE NATIONAL LABORATORY

LLNL-PROC-825095

Feature Extraction and Design from Gamma-ray Spectra for Radionuclide Identification

K. T. Nelson, J. R. Romo, M. Monterial, S. E. Labov, K. E. Nelson

July 29, 2021

INMM & ESDARSA Joint Virtual Annual Meeting
Vienna, Austria
August 23, 2021 through September 1, 2021

Proceedings of the INMM & ESDARSA Joint Virtual Annual Meeting
August 23-26 & August 30-September 1, 2021

Feature Extraction and Design from Gamma-ray Spectra for Radionuclide Identification

Kai Tyrus Nelson[1,2,*], John-Ryan Romo[3], Mateusz Monterial[1], Karl E. Nelson[1], Simon E. Labov[1]

[1]*Lawerence Livermore National Laboratory, Livermore, CA, USA.*
[2]*Las Positas College, Livermore, CA, USA.*
[3]*University of New Mexico, Albuquerque, NM, USA.*
**nelson268@llnl.gov*

Abstract

In this study we compare the performance of various strategies in extracting features from gamma-ray spectra for radionuclide identification. The primary objective of feature design is to reduce the number of dimensions for the classifier, therefore improving performance while avoiding overfitting. We used two feature extraction methods, principal component analysis (PCA) and peak integration, and also used the raw spectra. Multilayer Perceptron (MLP) classifier was used to compare the performance between the different feature extraction methods. Training and testing samples were generated with a 3"x3" NaI detector model with a source library of 33 radionuclides with a spanning set of shielding configurations. The drawn samples included variable background and mixtures of SNM (Special Nuclear Material) with masking sources. The overall performance of each feature design was assessed using the F1-score. Individual radionuclides that performed best and worst in each feature design were compared as well.

[1] This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. LLNL-PROC-825095

Introduction

Interpreting gamma-ray spectra to identify source radionuclides requires subject matter experts which are fewer in number than the demand from the various deployed detectors. Algorithms can fill this gap by allowing unskilled operators to identify the measured radionuclides. Machine learning classifiers can perform such classification, given a representative training data set, and have been studied in recent years as an alternative to custom-built expert systems. Neural networks are a popular classifier, because of their ability to ingest a variety of data without the need of extensive pre-processing.

There have been previous studies conducted on using neural networks to identify radio nuclides from their gamma-ray spectra. Many of these studies use raw gamma-ray spectra as features to feed into the neural network [2][5][6]. However, spectra distributed across many channels create a very high dimensionality space in which noise can exist. This can potentially increase the difficulty of the problem for the neural network as it has to be able to recognize a nuclide across any variation of noise by itself.

Feature extraction, and pre-calculations of features are often employed in machine learning problems. They can have benefits such as reducing noise, reducing dimensionality, and removing irrelevant information. This work will focus on comparing different feature transforms when applied to the spectra, and how it affects identification performance.

Neural Networks

Neural Networks

A simple neural network is a nonlinear mathematical model which tries to find an arbitrary relation mapping R^a to R^1. It achieves this by using gradient-descent based methods. This data is given a number of samples which allows it to train and optimize towards the relation which best fits the data that is given. The model updates its parameters to find a nonlinear complex function to model the relation.

Neural networks are often used when machine learning is applied to identification of radionuclides [2][3][5][6]. They are a good baseline classifier as they have a robustness to perform on many kinds of features without many constraints on the data it is fed. Therefore, neural networks are capable of handling both raw spectra, and extracted features used in this work.

Optimization

In the implementation of our neural network, we used the ADAM optimizer due to the robustness of the method in application across machine learning problems [1]. We chose to use commonly applied ADAM parameters: $\alpha = 0.001$, which changes the rate of learning in the model, and $\beta = 0.9$, ϐ = 0.99, ε = 10^-8 which all affect the momentum in the gradient descent method. For implementation of this method, we used the Scikit-learn implementation of the Multi-Layer Perceptron classifier [4].

Features

In order to train a network to be robust to all variations of noise, background, and gain would take a very large amount of data and would likely have a large training time. Pre-calculations to account for noise, background, and reduce dimensionality help reduce the variations within the data allowing for the classifier to be more robust without needing as much training data. Two feature designs, Principal Component Analysis (PCA) and peak extraction, will be compared alongside "raw" features (counts in all channels) from the full gamma-ray spectrum.

<u>Sample Generation</u>

Training and testing sets were constructed by sampling from a pre-computed spectra template library of 33 radionuclides with a spanning set of shielding configurations. The PCA transform and peak extraction are then performed on the generated spectra. This ensures that all the sets contain data from the same spectrums for fair comparison.

<u>Raw Spectra</u>

The raw gamma-ray spectra are created via sampling from templates generated from GADRAS for a NaI 3x3 detector model and bucketed into 2000 channels from 0 to 6000 keV [7]. The background samples are composed of world average background and variable sub-components which include the primordial terrestrial radionuclides. In any given background sample, the proportion of the average background varies between 20% to 40% and sub-components comprise the remaining 60% to 80%.

The source samples for each radionuclide include a spanning set of shielding configurations. The background samples are combined with source samples in order to satisfy a specified signal-to-noise ratio (SNR). In this case, the SNR is the proportion of the source counts to counts from the background sample. In addition to the background, we include a Cs-137 intrinsic source which is used in the real world for gain stabilization. The intrinsic source rate is assumed to be 100 counts per second (cps) and the background rate varies between 150 to 400 cps. Each sample is drawn between 30 to 60 seconds, and the chosen SNR along with aforementioned rates implies the appropriate source sample counts to draw from a chosen template. For mixture sets, the samples are a mixture of 60% to 80% of an arbitrary masking nuclide and 20% to 40% of Special Nuclear Material (U-235 or Pu-239).

PCA Features

The PCA transform is a well-known transform which reduces the dimensionality of the feature space. By creating new axes along the vectors through the feature space which have the most variance, the features can be reordered into corresponding eigenvectors and eigenvalues representing the data. Then the features that account for little variance within the features space can be removed. This allows the most information regarding the variance to be preserved in the fewest number of features. The spectra were re-binned to a square root binning taking 256 bins from 20 - 3000 KeV. This increases the volume of the information stored in each bin so that meaningful data can be extracted from its variance.

Peak Extracted Features

The Benchmark for Radionuclide Identification Algorithm (BARNI) was used to extract peak features [8]. BARNI uses a continuum estimator followed by a derivative method for finding the peaks. If the peak is found within a predefined region-of-interest (ROI) then the counts in that peak and the square-root of the continuum counts under that peak are added to that ROI features. The regions are defined ahead of time by sampling each radionuclide with a high number of counts (50k+) and automatically selecting the regions around a high density of peak locations. A set of ROIs are defined for each radionuclide, which are normalized to the total counts in those sets of regions, with the total counts representing their own feature.

Procedures

Performance Metric

The F1-score metric was used for the measurement of performance of general classification. This ensures that for general classification a model has to be able to perform well identifying nuclides by demonstrating both high precision and recall.

SNM masking cases are also evaluated for total accuracy of identifying concealed SNM. For these masking samples, samples are counted correctly only if the model identifies the present SNM in the samples. The accuracy is the ratio of the number of correctly identified samples over the total number of samples.

Training and Testing Data Sets

In this study there were two main groups that were tested: single nuclide and masked SNM. Each group was trained and tested over a range of SNRs, with masked SNM necessitating a higher SNR in order to achieve reasonable performance.

Sets of "low" SNR ranges (1-5, 5-10, 10-15, and 15-20) were generated for the single nuclide cases. Masking was not used as the SNR range is too low to get meaningful data to identify the masked nuclide. Models were trained on each of these, then tested against all low SNR sets. They were evaluated by F1-score performance.

Sets of “high” SNR ranges (50-100, 100-15, 150-200) were generated for both single nuclide and masking cases. Models were trained on each of these, then tested against all other high SNR sets. They were compared on metrics of F1-score for identification of single nuclides, F1-score for identification of masking cases, and total accuracy of SNM identification for masking cases.

Results

F1-scores of Models Trained and Tested on Single Nuclide Samples Set

The performance of single nuclide trained models tested on single nuclide cases at low SNR ranges is shown in Figure 1. At the lowest SNR bucket (1 to 5), identification was difficult with the best model being a PCA feature oriented model trained on the same SNR bucket performing with an F1-score of 0.4. The majority of models tended to perform best on sets that have the same SNR ranges to the sets that they were trained on, and often outperformed models from the same feature design that were trained on other SNR ranges. This excludes the raw gamma-ray spectra featured model trained on 1 to 5 SNR range which failed to identify the majority of nuclides. Overall identification tended to become easier at the higher end of the SNR ranges for the majority of models. PCA features performance was comparable to the raw gamma-ray spectra performance; meanwhile the peak extracted featured models tended to perform worse overall in comparison to other featured models.

The poor performance of the raw gamma-ray spectrum trained model may be attributed to the fact that the raw spectra was trained off of a very high noise on an unprocessed spectrum, thus preventing it from finding the meaningful ranges necessary for identification. This demonstrates that feature extraction and selection can improve performance if the training set is noisy and has relatively low amount of relevant signal.

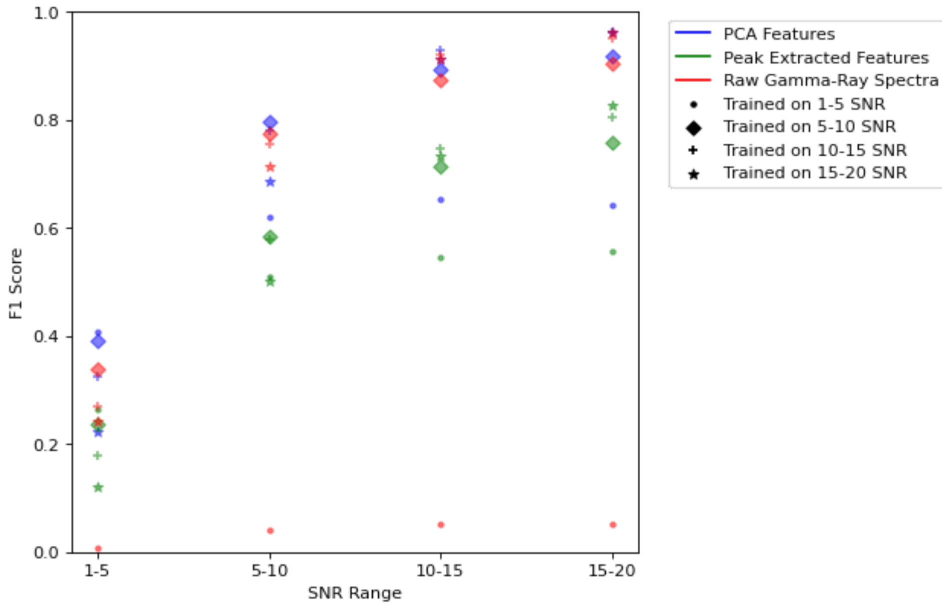

Figure 1 F1-Score Performance for Identification of Single Nuclides trained and tested at low SNR ranges.

F1-scores of Models Trained and Tested on Single Nuclide and SNM Masking Sets

For this set of evaluations, we compared the performance of training on only the single nuclide set against training with both single and SNM masking sets. The training models on the set that included singles and masked SNM performed far better on SNM masking test sets than the models trained purely on single nuclide sets, as shown in Figure 2. This came with a miniscule detriment to overall performance in the single nuclide identification when trained on the single nuclide and SNM masking set, as shown in Figure 3. Overall, the raw spectra tended to have the highest robustness to the different kinds of sets, with raw spectra models being consistently the top contender for most arbitrary nuclide as well as SNM identification tasks.

Interestingly, the models trained on the lowest SNR ranges tended to have the best performance for identifying singles; the models trained on the 50 to 100 SNR bracket performed close to the top for single nuclide sets for the single nuclide set trained models per their respective feature transforms and the models trained on the 50 to 100 SNR bracket set tended to perform the best on single nuclide identification for the single nuclide and SNM masking trained models per their respective feature transforms. In the case of the single nuclide and SNM masking set trained models, this could be due to the models being trained on a high amount of Pu-239, and U-235. This would cause the SNM to have more weight in the total proportion of the cost function, causing classification of those nuclides to be more important to the classifier. Training on the lower SNR range would have the MLP have to identify the other nuclide from less source, potentially making the network be more sensitive to those nuclides as the source is increased.

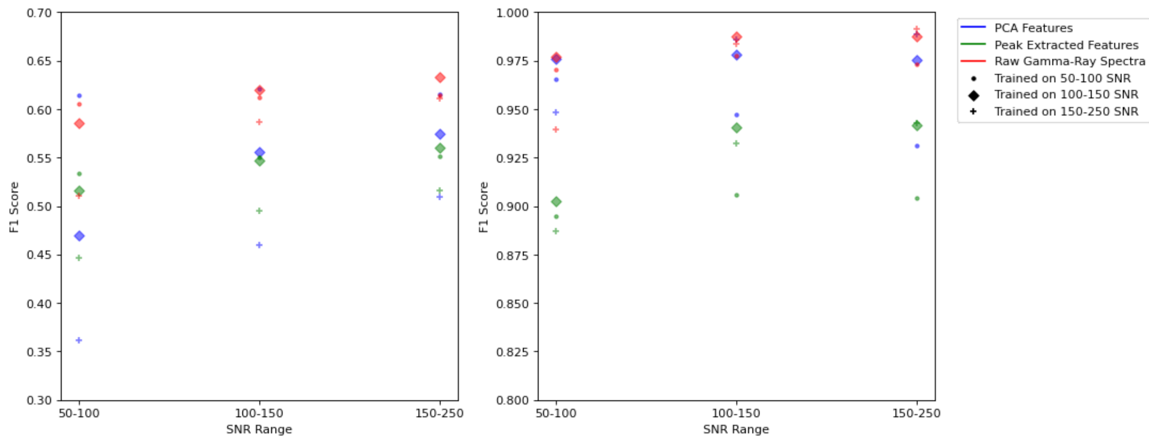

Figure 2 Comparison of model performance (F1 score) trained on Single Nuclide Set (left) and trained on both Single Nuclide and SNM Masking Set (right), tested against the SNM Masking Set.

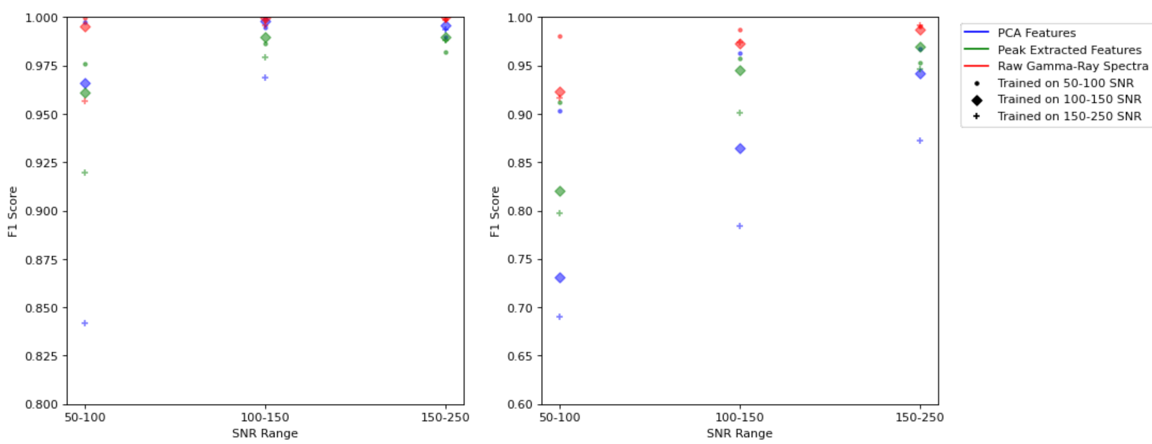

Figure 3 Comparison of model performance (F1 score) trained on Single Nuclide Set (left) and trained on both Single Nuclide and SNM Masking Set (right), tested against the Single Nuclides Set.

<u>Accuracy of SNM Indemnification of Models Trained on Single Nuclide and SNM Masking Sets</u>

For identification of SNM sources masked by other nuclides, training on the set including singles as well as masked SNM sources far outperformed training purely on single nuclides, as shown in Figure 4. This is likely due to the fact that the single nuclide trained models were not trained on Pu-239 or U-235 in quantities as low as the ones seen in the SNM masking set. For masked sources, the peak extracted feature-trained models performed the best out of models trained on only the single nuclide set. For overall performance on identifying masked SNM, the PCA features trained on the single nuclide and masking sets performed the best, with each SNR bucket for the SNM masking set being best identified by the PCA featured model trained on the same set.

For the models trained on only the single nuclide set, the best performance came from the peak extracted features. This may be due to the extracted features pre-defining its regions of interest to feed the model. The feature extracted model may perform better than the PCA or raw gamma-ray spectra models as it is fed features specific to peaks from Pu-239 and U-235. This may lead it to be easier to identify the SNM when it is present in a mixture compared to the other models as the peak information and noise reduction allows for a more obvious distinction to the presence of SNM in the sample.

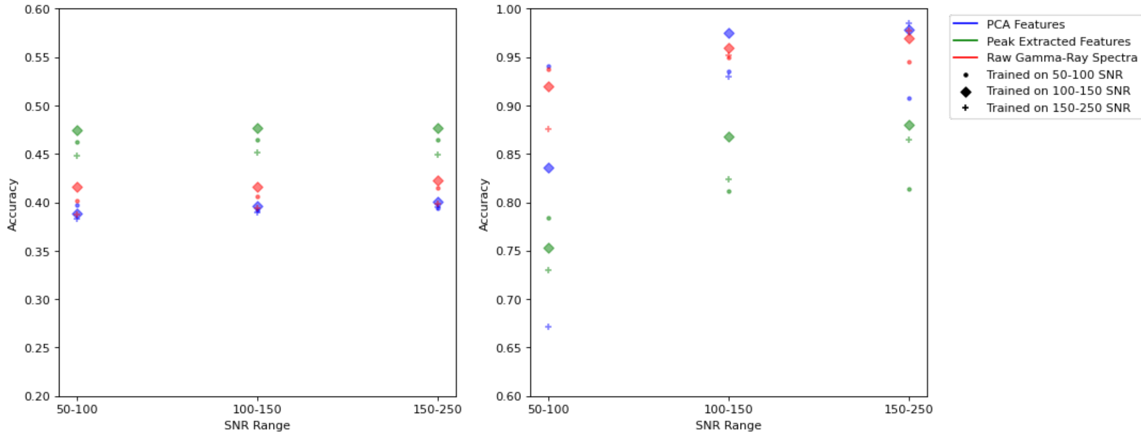

Figure 4 Accuracy of SNM identification for models trained on only Single Nuclide Set (left) and trained on both Single Nuclide and SNM Masking sets (right).

<u>Models Training Time</u>

The comparison of model training times for the different feature extraction methods and raw spectra is shown in Figure 5. The raw gamma-ray spectra models had the longest overall training time across all models, often taking over double the duration of the other models to train. In general, the preprocessed features took significantly less time to process.

The peak extracted features took the least time, despite having more feature dimensions than the PCA feature extraction. This may be due to the feature extraction features pre-defining regions of interest for classification. This would make the model not have to learn the many of peaks associated with each nuclide from scratch allowing it to more quickly converge towards its optimum.

The lower amount of training times for pre-processed features will be useful as the variation in the distribution of training data is increased and more samples will be required for training.

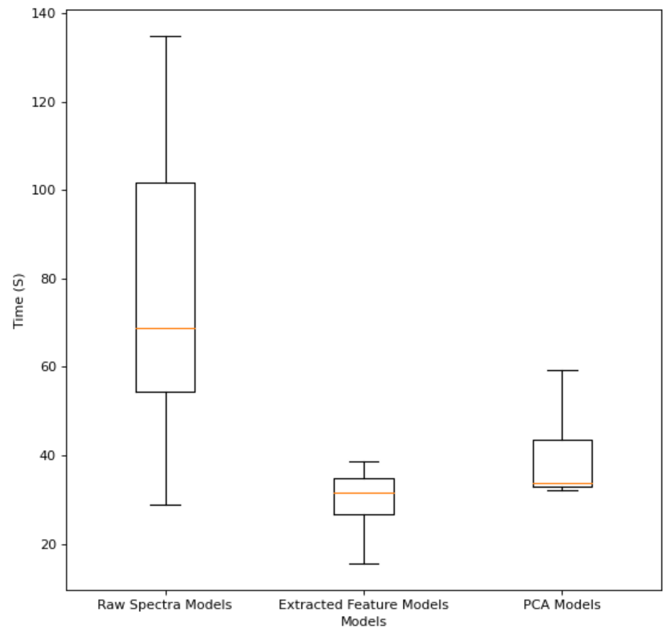

Figure 5 Comparison of model training time for different feature extraction methods.

Conclusion

Overall, raw gamma-ray spectra, as well as PCA extracted features had the best overall performance when it came to identification of the source for gamma-ray spectrum. Though performing well in most cases, raw spectra appeared to be more sensitive to the data it was trained on then either of the other feature transforms, entirely failing to form meaningful identifications when trained on some sets with high noise.

In addition, sometimes PCA featured models performed better than raw gamma-ray features by a notable margin. This may be attributed to the square root binning of the spectra allowing for better identification of nuclides at higher energies. Although the peak extracted features tended to perform worse on most cases, there were cases where it performed better than the other models when trained on a less complex set than what it was tested on. This may be useful when applying the extracted features to real world data, as the more concise peak data may serve useful for identification when the model has to look at data from distributions it has not seen. Feature extraction appears promising in reducing the amount of training time as well as attaining adequate levels of performance. This will be useful as the size of training sets scales up when the models have to be applied to real-world data.

Neural Networks are capable of performing classification tasks for nuclides from NaI gamma-ray spectra using various feature transforms and are capable of doing this across various SNR ranges. The best feature design is dependent on the classification task and the data which it is trained on. Future work will pertain to how calibration error affects each of the performances of the feature designs.

References

[1] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *arXiv:1412.6980 [cs]*, Jan. 2017, Accessed: Jul. 20, 2021. [Online]. Available: http://arxiv.org/abs/1412.6980

[2] M. Kamuda, J. Stinnett, and C. Sullivan, “Automated isotope identification algorithm using artificial neural networks,” *IEEE Transactions on Nuclear Science*, vol. 64, no. 7, Apr. 2017, doi: 10.1109/TNS.2017.2693152.

[3] J. He *et al.*, “Rapid radionuclide identification algorithm based on the discrete cosine transform and BP neural network,” *Annals of Nuclear Energy (Oxford)*, vol. 112, pp. 1–8, 2018, doi: 10.1016/j.anucene.2017.09.032.

[4] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.

[5] P. Olmos *et al.*, “A new approach to automatic radiation spectrum analysis,” *IEEE Transactions on Nuclear Science*, vol. 38, no. 4, pp. 971–975, Aug. 1991, doi: 10.1109/23.83860.

[6] E. Yoshida, K. Shizuma, S. Endo, and T. Oka, “Application of neural networks for the analysis of gamma-ray spectra measured with a Ge spectrometer,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 484, no. 1, pp. 557–563, May 2002, doi: 10.1016/S0168-9002(01)01962-3.

[7] Mitchell, Dean, and & Mattingly, John. *Gamma Detector Response and Analysis Software (GADRAS) v. 16.0*. Computer software. Vers. 01. USDOE. 24 Dec. 2009. Web.

[8] M. Monterial, K.E. Nelson, S.E. Labov, S. Sangiorgio Benchmarking Algorithm for Radio Nuclide Identification (BARNI) Literature Review (No. LLNL-SR-767242) Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States) (2019)