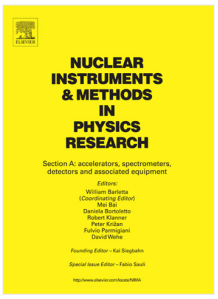
# Accepted Manuscript




An automated isotope identification and quantification algorithm for isotope mixtures in low-resolution gamma-ray spectra

Mark Kamuda, Jifu Zhao, Kathryn Huff

PII: S0168-9002(18)31377-9
DOI: https://doi.org/10.1016/j.nima.2018.10.063
Reference: NIMA 61385

To appear in: *Nuclear Inst. and Methods in Physics Research, A*

Received date : 31 July 2018
Accepted date : 9 October 2018

Please cite this article as: M. Kamuda, et al., An automated isotope identification and quantification algorithm for isotope mixtures in low-resolution gamma-ray spectra, *Nuclear Inst. and Methods in Physics Research, A* (2018), https://doi.org/10.1016/j.nima.2018.10.063

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Manuscript
Click here to view linked References

ACCEPTED MANUSCRIPT

# An Automated Isotope Identification and Quantification Algorithm for Isotope Mixtures in Low-Resolution Gamma-ray Spectra

Mark Kamuda[a,*], Jifu Zhao[a], Kathryn Huff[a]

[a] *University of Illinois at Urbana-Champaign*

**Abstract**

Pattern recognition algorithms such as artificial neural networks (NNs) and convolution neural networks (CNNs) are prime candidates to perform automated gamma-ray spectroscopy. The way these models train and operate mimics how trained spectroscopists identify spectra. These models have shown promise in identifying gamma-ray spectra with large calibration drift and unknown background radiation fields. In this work, two algorithms for mixtures of radioisotopes based on NN and CNN are presented and evaluated.

*Keywords:* Automated Isotope Identification, Neural Networks, Gamma-ray Spectroscopy

## 1. Introduction

Traditionally, isotope identification is conducted by trained spectroscopist. Rawool-Sullivan et al. identified a common workflow performed by a group of gamma-ray spectroscopists [1]. This workflow included discriminating background and source photopeaks, adjusting the calibration using background photopeaks and checking for shielding effects in the low-energy photopeaks. Once photopeaks are identified, the spectroscopist would use their prior knowledge of isotope emissions (or consult a database of these emissions) to match isotopes to the spectrum. The researchers also noted that while spectroscopists used this book knowledge, they often would use intuition developed from analyzing tens or hundreds of gamma-ray spectra. The researchers also noted the difficulty in incorporating this subjective analysis into an automated algorithm.

Neural networks (NNs) can mimic the intuition a trained spectroscopist uses when identifying spectra. It has been previously shown that NNs can be trained to identify and quantify isotopes in gamma-ray spectra [2]. There have also been a number of published papers which apply NNs to automated isotope identification. NNs have been applied to peak fitting [3], isotope identification [4, 5], and activity estimation [4, 6, 7].

Previous work applying NNs to spectroscopy have focused on fully-connected architectures. Fully-connected NNs do not assume the input channels have local spatial structure, while convolutional neural networks (CNN) do. Because gamma-ray spectra have local spatial structure in the form of photopeaks and Compton continua, it may be better to use a convolutional NN over a fully connected NN. This work will focus on comparing the performance of a fully-connected NN and a CNN for automated gamma-ray spectroscopy.

## 2. Artificial Neural Networks

An NN is a mathematical model that attempts to map an arbitrary function from $\mathbb{R}^M$ to $\mathbb{R}^N$, where $M$ and $N$ are any integers. An NN accomplishes this by mimicking biological neurons. An example of an ANN is shown in Figure 1. This network has N neurons in input layer A, J neurons in hidden layer B, and K neurons in output layer C. Neurons in adjacent layers are connected by weights, represented in Figure 1 by arrows connecting nodes.

The operation of a neuron is summarized in Figure 2. Each neuron operates by summing the products of the previous layers values ($A_1$, $A_2$, ... $A_N$) and each individual weight ($w_{1j}$, $w_{2j}$, ... $w_{nj}$) connecting nodes. This summation is then operated on by an activation function, typically rectified linear unit (ReLU), which is passed onto the next layer of the network.

An ANN may be trained by setting the network weights, **W**, connecting the neurons in such a way that

*Corresponding author
*Email addresses:* kamuda1@illinois.edu (Mark Kamuda), jzhao59@illinois.edu (Jifu Zhao), kdhuff@illinois.edu (Kathryn Huff)

ACCEPTED MANUSCRIPT

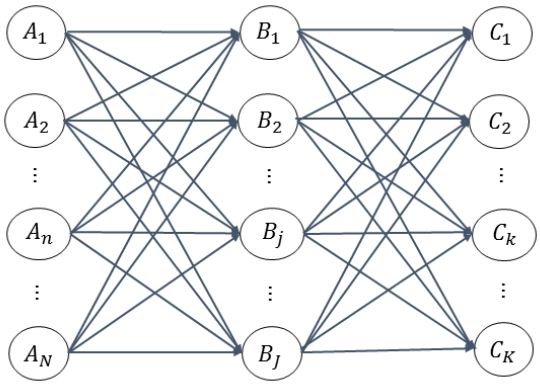



Figure 1: Example ANN with input layer A, hidden layer B, and output layer C.

they minimize some error metric, E, between target values, $\mathbf{T}$, in a training dataset, $\mathbf{Y}$, and the ANN output given that training dataset, $f(\mathbf{Y};\mathbf{W})$,

$$\underset{\mathbf{W}}{\operatorname{argmin}}\ \mathrm{E}(f(\mathbf{Y};\mathbf{W}),\mathbf{T}). \tag{1}$$

A popular numerical method to solving Equation 1 is gradient descent through the back-propagation of errors [8]. This method changes the ANN weights using the derivative of the error function with respect to the weights.

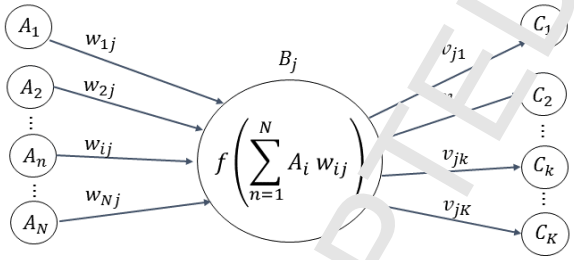



Figure 2: Summary of the operation of a single neuron.

The complete FC-NN model used in this work is shown in Figure 3. The FC-NN uses 1024 channels of a 2" x 2" NaI spectrum as input, and has a softmax output,

$$softmax(z_j) = \frac{\exp(z_j)}{\sum_{k=1}^{K}\exp(z_k)}, \tag{2}$$

after two fully connected layers. The softmax output normalizes the output vector to sum to unity and guarantees all elements are positive. This ensures the outputs can be mapped to mixing coefficients for each isotope in the library. An isotopes mixing coefficient describes how many counts in a given spectrum are attributable to that isotope.

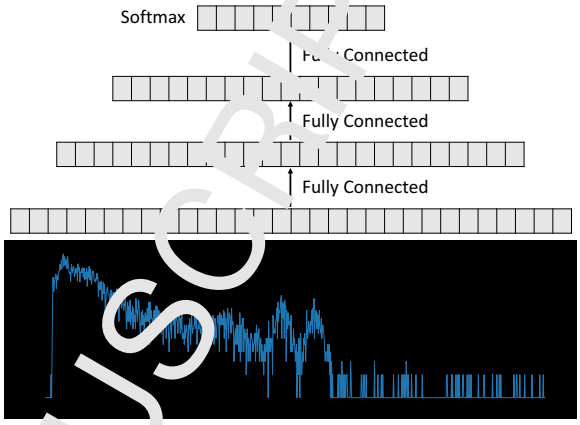



Figure 3: Illustration of Fully Connected Neural Network Structure (the input spectrum contains $^{60}Co$).

The CNN model used in this work is shown in Figure 4. The input and output of the CNN are the same as the FC-NN. The difference between the CNN and FC-NN are the convolution and max pooling layers. Convolution layers activations are created by convolving 1-D filters with the previous layer's signal. Max pooling is a sub-sampling operation that attempt to combine low-level features and to encourage spatial invariance by reducing the resolution of the previous layers [9]. After the convolution and pooling layers, the features are flattened into a vector and fed into a fully-connected architecture. The weights of the fully-connected network and the 1-D convolution filters are learned through training.

## 3. Methods

### 3.1. Training Set Creation

In order to train an ANN, a training set must be provided. The 29 isotopes in the dataset are based on the American National Standards Institute performance criteria for hand-held instruments for the detection and identification of radionuclides, ANSI N42-34-2006 [10]. From this set of isotopes, template spectra for each isotope were simulated using GADRAS [11]. These sources are simulated using a 30 cm standoff distance. The default calibration for the templates is set for a maximum gamma-ray energy of 3 MeV. To teach the model a range of different detector calibrations, each spectrum's channels were linearly rebinned by some percent. After rebinning, the resulting spectrum was reconstructed using third order spline interpolation with

ACCEPTED MANUSCRIPT

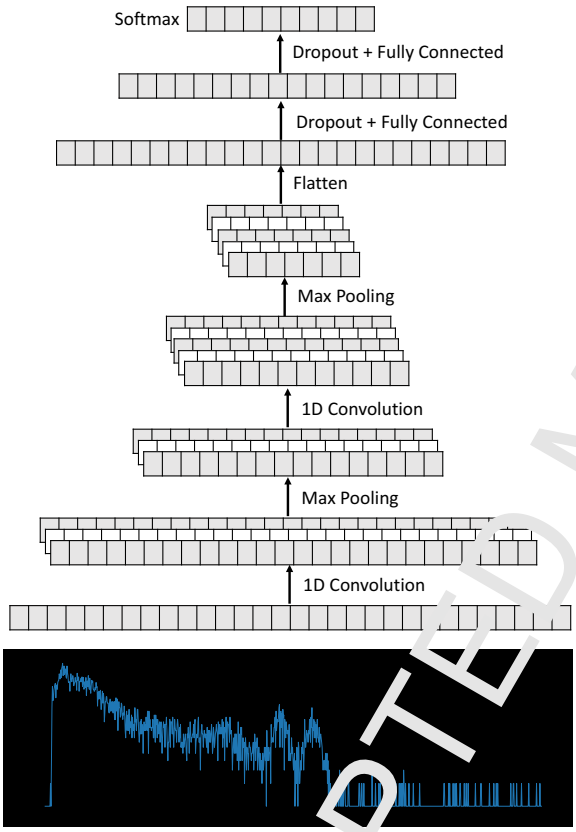



Figure 4: Illustration of Convolutional Neural Network Structure (the input spectrum contains $^{60}Co$).

the new bin positions. The magnitude of each rebinning was uniformly chosen between a ± 25%. Based on lab observations using an Ortec 905-3 NaI spectrometer, the average background count rate was set at 65 counts per second (cps). Each spectrum had a total number of count logarithmically distributed between $6.5 \times 10^2$ and $6.5 \times 10^4$. Each isotope, excluding those in background, had an equal probability of being included in each spectrum. The counts from background were distributed uniformly between background thorium, background uranium, and background $^{40}$K.

### 3.2. Network Structure and Hyperparameter Search

In general, ANNs have a tendency to memorize their training set in a process called overtraining. An overtrained ANN will tend to incorrectly identify novel data. To prevent this, hyperparameters were used in the present work to prevent overfitting and optimize performance. Unfortunately, there is currently no known method to know which hyperparameters have an impact on model performance before training. Because of this, a number of popular hyperparameters are typically added to a model and a random hyperparameter search is used to identify those which are important [12]. For the FC-NN, a network structure and hyperparameter choices are based on a previously trained FC-NN [2]. Due to a large observed variance in output from different FC-NN's with different weight initialization, the output from 10 unique FC-NN's were averaged using a technique called bagging [13].

For CNN structures, shown in Figure 4, hyperparameters include the number of convolutional filter, kernel size for each convolution filter, number of nodes for the fully connected layers, and the dropout rate. A random search was performed over the predefined hyperparameter space, shown in Table 1, and the best hyperparameter combination is chosen based on the performance of CNN on a separate validation dataset.

Table 1: Hyperparameter Space for CNNs

| Hyperparameter | Values |
|---|---|
| # filters | 3, 5, 7, 9, 11 |
| kernel size | 5, 10, 15, 20, 25, 30<br>35, 40, 45, 50, 55, 60 |
| dropout rate | 0.0, 0.1, 0.2, 0.3, 0.4, 0.5 |
| dense layer size | 32, 64, 128, 256 |

ACCEPTED MANUSCRIPT

## 4. Results and Discussion

To investigate how the FC-NN and CNN compared to each other, four simulated datasets were considered. Each dataset includes the 29 isotopes in the training library. For each isotope, 100 single-isotope spectra of 9 different mixing coefficients were simulated. The mixing coefficients were evenly spaced between [0.1, 0.9]. In addition to this, each spectrum had its calibration randomly changed and a random background generated in the same way as the training set. The first dataset investigated how well each algorithm worked in the best-case scenario, when the data are generated from the same distribution used to train the algorithms. The three additional datasets are adversarial examples, which are meant to represent real-world issues that could affect algorithm performance. Performance in these datasets explores the generalization limits of the presented training methodology for automated isotope quantification. For machine learning algorithms, generalization describes the ability for an algorithm to fit data outside the training dataset.

The first adversarial dataset used the same spectral templates as the training dataset for the sources, but not the background. The background was sampled from a 2" x 2" NaI detector with a 10 hour integration time. The second adversarial dataset was simulated using template spectra simulated with standoff distances of 15 cm and 60 cm. Changing the standoff distance changes the shape of the spectrum by affecting the peak-to-total ratio and shape of the Compton continuum. The third adversarial dataset was created using template spectra simulated with a full-width-at-half-maximum (FWHM) of 10% at 662 keV. The training set used a FWHM of 7.44% at 662 keV. The wider FWHM distorts the spectrum by making photopeaks closer to each other more difficult to differentiate.

### 4.1. Algorithm Quantification Performance on Simulated Spectra

Spectra with a total of $10^3$ counts were simulated in the manner described above. Using the outputs, a box-and-whisker plot was created comparing the outputs of the FC-NN and CNN. This plot is shown in Figure 5. In general, the CNN outperforms the FC-NN in terms of variance. In addition to this, the median outputs of the CNN were closer to the true mixing coefficient for mixing coefficients between 0.1 and 0.5. The median outputs are comparable between the two models for coefficients between 0.6 and 0.8.

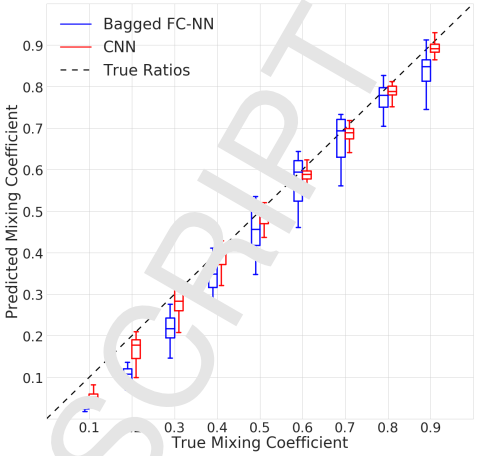



Figure 5: Predicted mixing coefficient for the FC-NN (blue) and CNN (red), averaged over all library isotopes. The total counts in each spectrum is $10^3$. Each spectrum is simulated using the same template as the training dataset.

### 4.2. Algorithm Quantification Performance on Adversarial Datasets

The three adversarial datasets presented represent performance in spectra with a real background radiation field, varying standoff distance, and a detector response with a wider FWHM than the models were trained with.

#### 4.2.1. Simulated Sources Injected Into Measured Background

Figure 6 shows significant performance degradation from both models when quantifying spectra simulated with a measured background instead of the simulated background used in training. The CNN performed particularly poorly, underestimating the true mixing coefficient with a large variance. This may indicate the CNN is particularly sensitive to background in count-starved spectra. The FC-NN also underestimated the true count contribution, but compared to the CNN the median of the output distributions were closer to the true mixing coefficient, especially as the mixing coefficient increased. Because each channel of the spectrum is weighted by the coefficients of the network, previously described in Figure 2, the FC-NN may exploit the spectrum's photopeaks, like a region of interest (ROI) algorithm. Because the ratio of counts in photopeaks will not change significantly with varying background, it may be well suited to quantify isotopes in spectra with unknown background radiation patterns.

In contrast to the FC-NN, the CNN weighs responses to convolution filters, explained in Figure 4, in a way that may be able to use the entire gamma-ray spectrum. The potential drawback is that the convolution fil-

ACCEPTED MANUSCRIPT

ters learned by the CNN that worked well for simulated background may not work well for real background radiation. This was seen in the CNNs poor performance in Figure 6.

CNN performance improved with additional counts, as seen in Figure 7. For mixing coefficients between 0.4 and 0.7 The CNN had a smaller variance in it's outputs and modes closer to true mixing coefficients compared to the FC-NN. The increase in counts lowered the variance between channels in the simulated sources. Low channel-to-channel variance may help the CNN recognize the pattern of the source templates, similar to how a template matching algorithm might improve with a lower channel-to-channel variance signal.

The FC-NN operates similarly with $10^3$ and $10^4$ counts, keeping a large variance while underestimating the true mixing coefficient. The slight improvement in the FC-NN's performance are explainable if the FC-NN acted primarily like a region of interest algorithm. The increase in counts may not change the regions explored by the FC-NN. Because the FC-NN was comparing the same regions with a background outside the training set, the relationship (for example the ratio of counts between photopeaks) between these regions may be outside those learned by the FC-NN. This implies that g...eralizing to backgrounds outside those used in the training set will be more difficult for the FC-NN compared to the CNN.

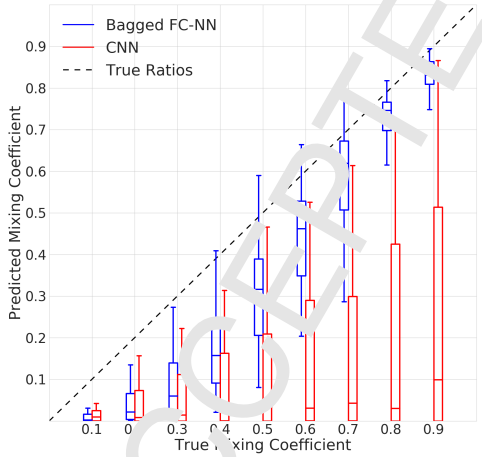



Figure 6: Predicted mix... co...icient for the FC-NN (blue) and CNN (red), averaged o... ...l library isotopes. The total counts in each spectrum is $10^3$. In ea... sp...trum, the source is simulated using the same template as the train... dataset and the background is sampled from a measured spectrum.

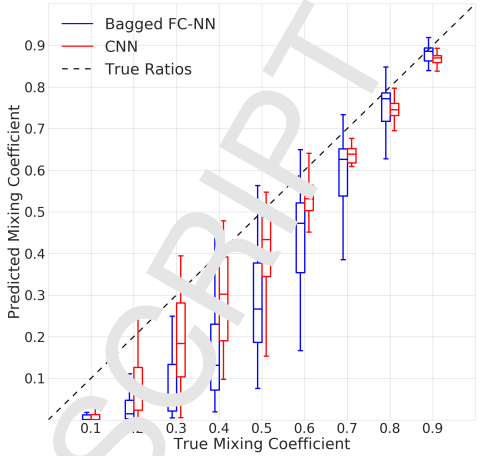



Figure 7: Pr...icted ...ixing coefficient for the FC-NN (blue) and CNN (red), averaged o... r all library isotopes. The total counts in each spectrum i... $10^4$. In ea... spectrum, the source is simulated using the same tem...late as ...e t...ining dataset and the background is sampled from a measur... spectrum.

### *4.2.2. Effects of Standoff Distance on Quantification Performance*

In general, both the FC-NN and CNN had a larger variance in output when quantifying spectra with a standoff distance of 15 cm, Figure 8, compared to spectra with a standoff distance of 30 cm, Figure 5. The increase in variance was more severe for the FC-NN compared to the CNN. The median of the predictions was also consistently higher for the CNN than for the FC-NN. This is likely due to the difference in the peak-to-total ratios of the different simulation templates. Templates with a 15cm standoff have a peak-to-total ratio of 0.29 at 662 keV, whereas the training templates used a 30cm standoff and have a peak-to-total ratio of 0.25 at 662 keV. Because of the increased peak-to-total ratio, a model that employs photopeaks (instead of the full spectrum) to quantify isotopes will overpredict an isotopes true mixing coefficient. The large variance in output combined with underprediction from the FC-NN indicates poor generalization with respect to changes in standoff distance in count starved spectra.

In spectra with more counts FC-NN performance improves, shown by an increase in the median outputs seen in Figure 9. Despite this improvement the FC-NN still has a larger output variance than the CNN, indicating that the CNN generalizes better to changes in standoff distance.

Increasing the standoff distance to 60 cm decreases the peak-to-total ratio to 0.21 at 662 keV. This decreased ratio leads to underpredictions, seen in Figures 10 and 11. The trends for the larger standoff distance are sim-

ACCEPTED MANUSCRIPT

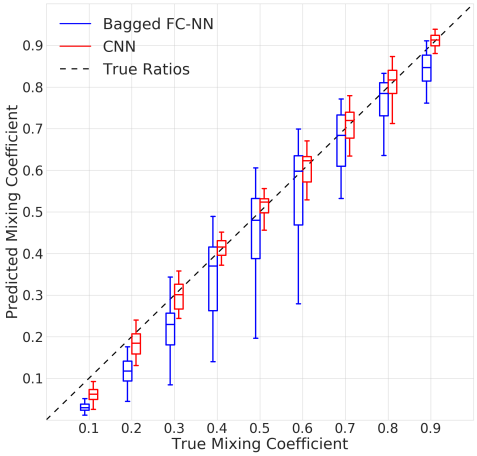



Figure 8: Predicted mixing coefficient for the FC-NN (blue) and CNN (red), averaged over all library isotopes. The total counts in each spectrum is $10^3$. In each spectrum, the source is simulated using a standoff distance of 15cm.

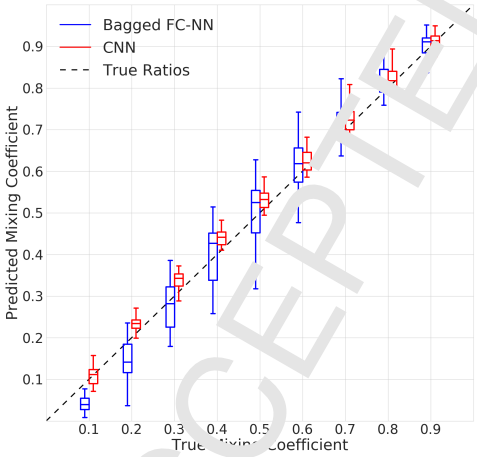



Figure 9: Predicted mixing coefficient for the FC-NN (blue) and CNN (red), averaged over all library isotopes. The total counts in each spectrum is $10^4$. In each spectrum, the source is simulated using a standoff distance of 15cm.

ilar to those for a smaller standoff distance. In general the variance in the outputs of the CNN are smaller than the outputs from the FC-NN.

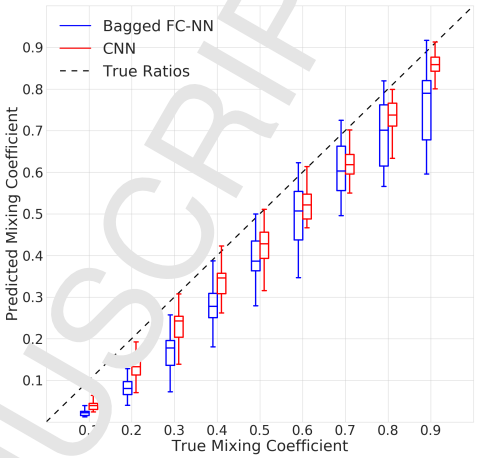



Figure 10: Predicted mixing coefficient for the FC-NN (blue) and CNN (red), averaged over all library isotopes. The total counts in each spectrum is $10^3$. In each spectrum, the source is simulated using a standoff distance of 60cm.

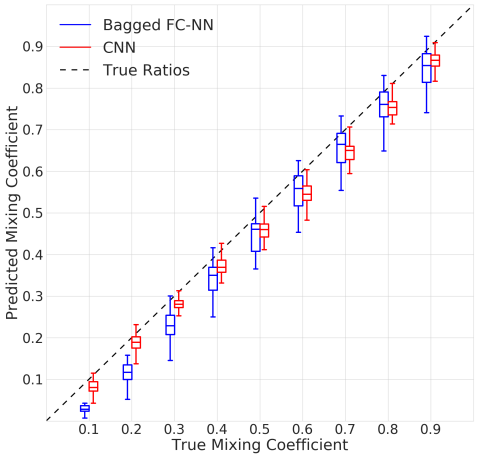



Figure 11: Predicted mixing coefficient for the FC-NN (blue) and CNN (red), averaged over all library isotopes. The total counts in each spectrum is $10^4$. In each spectrum, the source is simulated using a standoff distance of 60cm.

#### *4.2.3. Effects of Detector Full-Width-at-Half-Maximum on Quantification Performance*

Spectra in the following dataset have a FWHM of 10% at 662 keV. This is larger than spectra in the training dataset, which have a FWHM of 7.44% at 662 keV. A FWHM of 10% at 662 keV represents a worst-case scenario for a typical NaI detector. The larger FWHM

ACCEPTED MANUSCRIPT

results in wider spectral peaks, which may overlap to form spectral templates different from those used in the training dataset. Good performance indicates that the model is insensitive to changes in FWHM and has generalized with respect to changes in different radiation detector's FWHM.

Seen in Figure 12, the modes of each dataset are consistently below their respective true mixing coefficients. This could be evidence that each model is close to generalizing to changes in FWHM. This may be due to each model identifying and quantifying spectra using ROI's around photopeaks in each spectrum. Because the photopeaks in these spectra are wider then those used during training, consistently fewer counts will be in the ROI's found by the algorithms trained with narrower ROI's. Despite this, the model performance was worse on this dataset than the dataset in Figure 5, indicating generalization performance can be improved with respect to changes in FWHM.

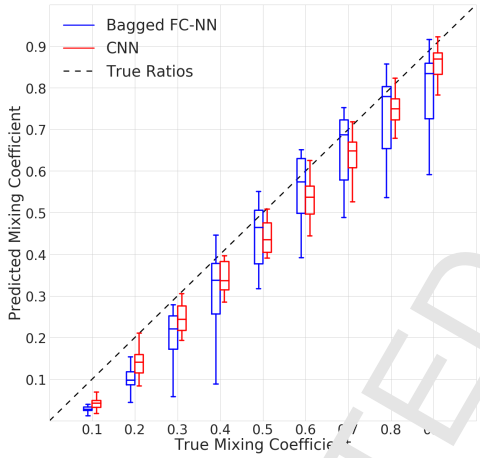



Figure 12: Predicted mixing coefficient for the FC-NN (blue) and CNN (red), averaged over all library isotopes. The total counts in each spectrum is $10^3$. Each spectrum is generated using templates with a wider FWHM than the training dataset.

## 5. Conclusion

In general, the CNN had a smaller variance on outputs for each dataset explored compared to the FC-NN. This may be because the FC-NN focuses primarily on photopeaks while the CNN is able to use both the photopeaks and other features like the Compton continuum. For this reason, future work applying machine learning to spectroscopy may benefit by focusing on convolutional models in the place of fully connected architectures.

As discussed in section 4.2.1 both models are very sensitive to changes in the background radiation field. This implies that the current method of simulating background is insufficient for either model to generalize to real background radiation conditions. This may be improved by adding measured background templates to the simulations or by varying the simulated background templates so they represent more realistic background.

As seen in section 4.2.2 the FC-NN and CNN tend to overpredict mixing coefficients when standoff distances are smaller than those used in the training dataset and underpredict mixing coefficients when standoff distances are larger. This shows that both algorithms primarily use photopeaks to quantify isotopes in spectra. This also shows some generalization to spectra with standoff distances outside the training set. Improvements can be made by adding additional templates with different standoff distances.

As seen in section 4.2.3, both models have generalized to changes in FWHM. This motivates future gamma-ray spectroscopy training datasets incorporate different FWHM's on a coarse grid.

## Acknowledgement

This work was funded by the Consortium for Verification Technology under Department of Energy National Nuclear Security Administration award number DE-NA0002534.

## References

[1] M. Rawool-Sullivan, J. Bounds, S. Brumby, L. Prasad, J. Sullivan, Steps toward automated gamma ray spectroscopy steps toward automated gamma ray spectroscopy: How a spectroscopist deciphers an unknown spectrum to reveal the radioactive source.

[2] M. Kamuda, C. J. Sullivan, An automated isotope identification and quantification algorithm for isotope mixtures in low-resolution gamma-ray spectra, Radiation Physics and Chemistry.

[3] R. Abdel-Aal, Comparison of algorithmic and machine learning approaches for the automatic fitting of gaussian peaks, Neural Computing and Applications 11 (1) (2002) 17–29.

[4] R. Abdel-Aal, M. Al-Haddad, Determination of radioisotopes in gamma-ray spectroscopy using abductive machine learning, Nuclear Instruments and Methods in Physics Research A 391 (1996) 275–288.

[5] M. Medhat, Artificial intelligence methods applied for quantitative analysis of natural radioactive sources, Annals of Nuclear Energy 45 (2012) 73 – 79.

[6] V. Vigneron, J. Morel, M. Lepy, J. Martinez, Statistical modelling of neural networks in y-spectrometry, Nuclear Instruments and Methods in Physics Research A 396 (1996) 642647.

[7] M. Kamuda, J. Stinnett, C. J. Sullivan, Automated isotope identification algorithm using artificial neural networks, IEEE Transactions on Nuclear Science 64 (2017) 1858 – 1864.

ACCEPTED MANUSCRIPT

[8] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning representations by back-propagating errors, Nature 323 (1986) 533–536.

[9] D. Scherer, A. Müller, S. Behnke, Evaluation of pooling operations in convolutional architectures for object recognition, in: K. Diamantaras, W. Duch, L. S. Iliadis (Eds.), Artificial Neural Networks – ICANN 2010, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 92–101.

[10] American National Standard Performance Criteria for Hand-Held Instruments for the Detection and Identification of Radionuclides, ANSI N42.34-2006 (2007).

[11] D. J. Mitchell, L. T. Harding, GADRAS isotope ID user's manual for analysis of gamma-ray measurements and api for linux and android, SAND2014-3933.

[12] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, Journal of Machine Learning Research 13 (2012) 281–305.

[13] L. Breiman, Bagging predictors, Machine Learning 24 (2) (1996) 123–140. doi:10.1007/BF00058655.
URL https://doi.org/10.1007/BF00058655

ACCEPTED MANUSCRIPT