



**Advanced
Reactors &
Fuel
Cycles**

ADVANCED REACTORS AND FUEL CYCLES GROUP
Technical Report Series
ARFC-NPRE-17-00

Inception Neural Networks for Isotope Identification

Prepared for:
AMERICAN NUCLEAR
SOCIETY
CONTRACT NN-NNNN

Prepared by:
Samuel DOTSON
Mark KAMUDA
Prof. Kathryn HUFF

June 5, 2019

Urbana, IL
Department of Nuclear, Plasma, and Radiological Engineering
University of Illinois at Urbana-Champaign

1 Introduction

Developing algorithms that accurately identify the isotope sources of low-resolution gamma-ray spectra will be an important advance over the current isotope identification workflow (Sullivan, 2010). Previous work has shown that artificial neural networks can perform isotope identification using low-resolution gamma-ray spectrometers (Cite: [3], [4], [5]). The purpose of this paper is to introduce a new feature to the existing architecture of the Artificial Neural Network for Spectroscopic Analysis (ANNSA) package and to improve the training data simulations to better emulate background radiation found in real measurements. The new feature is known as an Inception Neural Network (INN) that implements wide convolutional layers with several filters, rather than the typical single-filter layer in a simple convolutional neural network (CNN). The features of a gamma-ray spectrum vary depending on the full width at half max (FWHM) of the photopeaks. Simultaneously applying multiple filters of different sizes allows an INN to capture more features during a single layer than a CNN. We hypothesize that an INN will also be robust to changes in background radiation thereby generalizing the ANNSA framework to more scenarios. We will compare the INN to a simple CNN to determine if the improvement in accuracy is enough to warrant the increased computational complexity. Finally, new training data will be obtained through simulations with GADRAS-DRF (Mitchell and Harding, 2014) software.

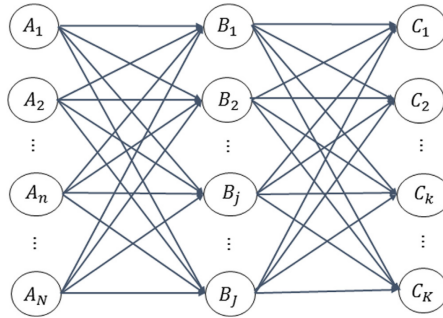


Figure 1: An arbitrary neural network that maps values with weights (arrows).

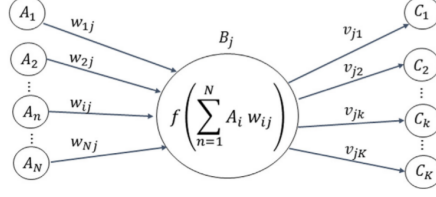


Figure 2: A single neuron being passed through an activation function, f .

2 Theory – Artificial Neural Networks

An artificial neural network (ANN) is a function that maps values from \mathbb{R}_N to \mathbb{R}_K by mimicking biological neurons. Examples of an arbitrary neural net and a single neuron are shown in Fig. 1 and Fig. 2. The sum of the inputs times the weights pointing to a neuron are passed through an activation function, here it is a rectified linear unit,

$$Relu = \operatorname{argmax}(0, x), \quad (1)$$

and then used as the input for the next layer, as shown in Fig. 2. An ANN may be trained by iteratively updating the weights of a network that minimizes an error function E . The weights are updated through back propagation by taking the derivative of E with respect to the weights. The error function minimized during training of the INN is cross-entropy,

$$E = - \sum_c^M y_{o,c} \ln p_{o,c}. \quad (2)$$

Eq. 2 is the cross-entropy for a multiclass classification, where there are more than two possible labels for a given input. M is the total number of labels for a given model, in this case it corresponds to 29 radionuclides (ANSI, 2015). Variable $y_{o,c}$ is binary, indicating whether observation, o , has the correct label, c . Variable $p_{o,c}$ is the probability that o is a member of c . The complete INN model is shown in Fig. 3 and Fig. 4. The input for the INN is a 2"x2" NaI spectrum of 1024 channels and the final output is a softmax given by,

$$\operatorname{softmax}(z_j) = \frac{\exp(z_j)}{\sum_{k=1}^k}. \quad (3)$$

The input data is passed through three ‘inception’ layers and, after flattening, the output is passed to a dense layer which gives the final softmax output. Each ‘inception’ layer has a bottleneck, a convolution, and a concatenation (Szegedy, et. al, 2014). The bottleneck is performed to reduce the computational complexity before large convolution filters are applied. A large filter can be factorized into several smaller filters to reduce the number of required multiplications. The convolution layer uses filters to pick features of a spectrum that have local spatial significance, but no long-range relationships. Once the convolution has been done with several filters, in parallel, the outputs of each of those convolutions are concatenated into a single tensor that is passed to the next ‘inception’ layer, or dense layer if the last layer has been reached. Just like a typical CNN (Fig. [5]) the final step after feature identification, is classification. Classification is performed by using a dense, or fully connected, layer that learns weights that correspond to a probability for a certain label. Here, the corresponding labels are radioactive isotopes.

3 Methods

3.1 Training set creation

It is infeasible to get enough real measurements to properly train a neural network. Thus all of the data used to train the neural network here has been simulated using GADRAS-DRF (Mitchell and Harding, 2014). The 29 isotopes in the dataset are based on the American National Standards Institute performance criteria for handheld instruments for the detection and identification of radionuclides, ANSI N42-34-2015 (ANSI, 2015). Previous work (Kamuda, et. al 2018) used a uniform distribution of background isotopes and a constant average count of 65 counts per second (cps) for the background. Simulating background radiation in this manner is insufficient for training a neural network to be robust against background variations and thus cannot be generalized to real conditions. We are updating the simulation protocol to increase the variability in background conditions. New simulated data will include random noise in a range of 40 to 200 cps. We believe this range is realistic for background radiation.

3.2 Network Structure and Hyperparameters

A neural network can memorize training sets resulting in overtraining and a misidentification of novel data. This is especially true for an INN, whose

weights are difficult to tune with simple back propagation. To solve this problem during training, we include an intermediate softmax output that allows for error corrections before the entire forward pass is complete. This branch is ignored during prediction, but offers a way to prevent overtraining. In the fully connected layer, dropout regularization is used to force the neural network to learn new pathways. Another method to prevent overfitting is the use of hyperparameters to optimize performance. There is no way to know which hyperparameters will influence the model before training, so a random hyperparameter search is performed to identify which are important (Bergstra and Bengio, 2012). For CNN structures, like the INN, hyperparameters include the number and sizes of convolutional filters, and the number of nodes and dropout rate for the fully connected layer.

3.3 Benchmark Techniques

In order to compare the efficiency of these two neural networks and the effectiveness of updated datasets, we will train each network twice. Once reusing data from previous work (Kamuda, et. al, 2018) and again with the improved datasets. This will allow us to identify if accuracy improved more due to a sophisticated algorithm or the advance is attributable to better training data, or some combination thereof. The INN will demonstrate a satisfactory improvement over the CNN if the amount of time required to train the network results in an equivalent reduction in variance for predictions. For concreteness, a 10 increase in training time should correspond to a 10, or greater, decrease in the variance for the INN when compared with the training time and variance of the CNN.

4 Conclusion

In this study, we compare the accuracy of two neural networks, a convolutional neural network with two convolutional layers and an inception neural network with three inception layers, for identifying radioactive isotopes in low-resolution gamma-ray spectra. We hypothesize that the INN would exhibit an increase in robustness commensurate to its computational complexity and that training the neural networks with larger and varied datasets will also improve its robustness. Improvements in the ability to identify isotopes present in a sample of radioactive material will have implications for *[add something here...]*