Comparing deep learning architectures on gamma-spectroscopy analysis for nuclear waste characterization

Otero, A. G. L. [a] , Potiens Júnior, A. J. [a] , Marumo, J. T. [a]

[a] Instituto de Pesquisas Energéticas e Nucleares, IPEN–CNEN/SP, Av. Prof. Lineu Prestes, 2242 – Cidade Universitária – CEP 05508-000 São Paulo – SP – Brasil

aglotero@usp.br

ABSTRACT

Neural networks, particularly deep neural networks, are used nowadays with great success in several tasks, such as image classification, image segmentation, translation, text to speech, speech to text, achieving super-human performance. In this study, the capabilities of deep learning are explored on a new field: gamma-spectroscopy analysis, comparing the classification performance of different deep neural network architectures. The following architectures where tested: VGG-16, VGG-19, Xception, ResNet, InceptionV3, and MobileNet, which are available through the Keras Deep Learning framework to identify several different radionuclides (Am-241, Ba133, Cd-109, Co-60, Cs-137, Eu-152, Mn-54, Na-24, and Pb-210). Using an HPGe detector to acquire several gamma spectra from different sealed sources to create a dataset used for the training and validation of the neural network's comparison. This study demonstrates the strengths and weaknesses of applying deep learning on gamma-spectroscopy analysis for nuclear waste characterization.
Keywords: gamma-spectroscopy analysis; deep learning;

1. INTRODUCTION

Neural networks are one artificial intelligence technique that mimics the biological brain [1], Figure 1 [2], through software, Figure 2, although the method was described long before, their use only gained traction by the last decade due to hardware and software improvements.
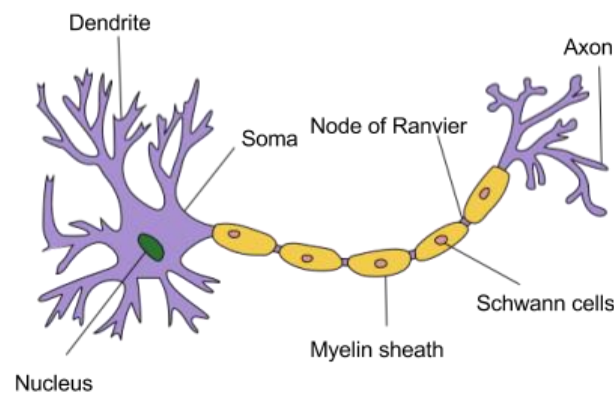

Figure 1: *The biological neuron.*
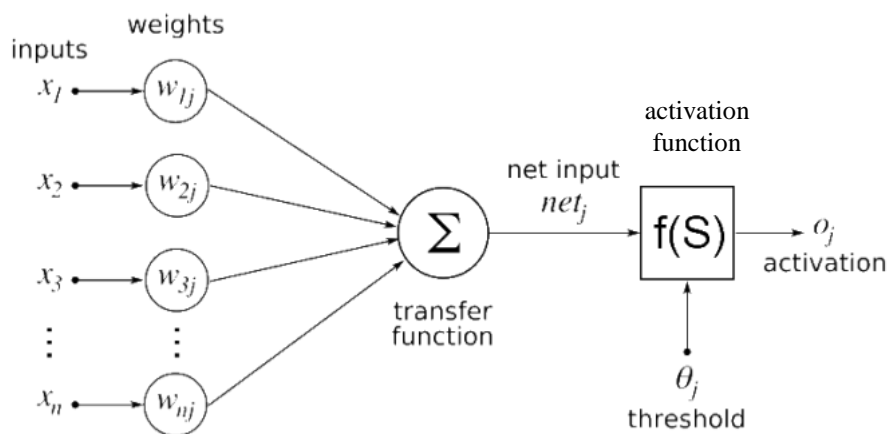Source: adapted after [2]

Figure 2: *One artificial neuron from a perception neural network.*
Source: adapted after [8]

Nowadays deep neural networks are applied in several new tasks, giving the computer power provided by modern devices, combined with the abundant availability of data, new architectures for new models are being created, achieving performance levels that surpasses humans in several tasks [3 ,4, 5].

These successes are the drivers for this study that evaluates the performance of well-known deep neural networks architectures in a new task: gamma-spectroscopy analysis for nuclear waste characterization.

Previous works evaluate the use of neural networks on such task, although such studies applied the perceptron neural network architecture [6, 7]. The use of perception [8] neural network architecture has some drawbacks such as small network size with only three layers (Input, Hidden, and Output); train data set must fit into the main computer memory; there is only one activation function.

Deep neural networks [9], in contrast, allows the use of training sets larger than the main memory; an arbitrary number of layers and several different activation functions. Especially the ability to train deeper networks allowed new levels of performance, as shown in Figure 3.
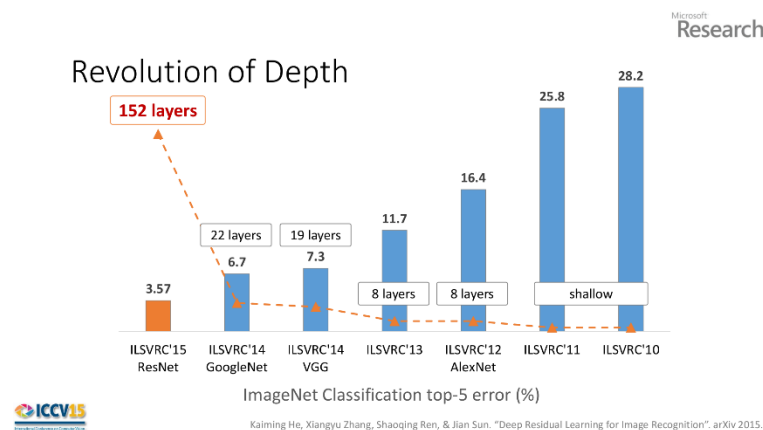

Figure 3: *The impact of larger networks on classification performance.*
Source: adapted after [12]

At the IPEN Radioactive Waste Management Department, final product from the waste treatment must be characterized. Nowadays, this work is manual. This study presents the results of different deep neural network architectures on gamma-spectroscopy classification, which is: given a gamma spectrum the deep neural network must identify which radionuclide composes the spectra (which can be more than one). Aiming future automation of this waste characterization process step.

2. MATERIALS AND METHODS

Measurements were taken from different sealed sources, Figure 4, to create the base dataset. From this base dataset several new spectra where generated containing random radionuclide combinations, up to three radionuclides, resulting in a synthetic dataset.

The synthetic dataset was split into training and testing sets. Measurement from a triple-sealed source, containing Am-241, Cs-137, and Co-60 was added to the testing set. All architectures where trained using the training set and the performance was evaluated using the testing set.
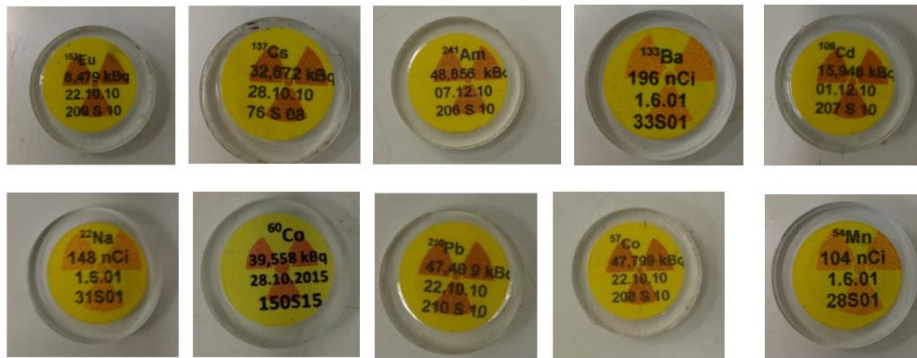
Figure 4: *The sealed sources used on this study.*

The eleven sealed sources: Am-241, Pb-210, Cs-137, Eu-152, Ba-133, Na-22, Co-60, Co-57, Mn-54, Cd-109, Triple calibration source with Am-241, Cs-137 and Co-60.

The geometry of the measurements is reported at Figure 5. The sealed source is 5 cm far from the HPGe detector that is inside the collimator.
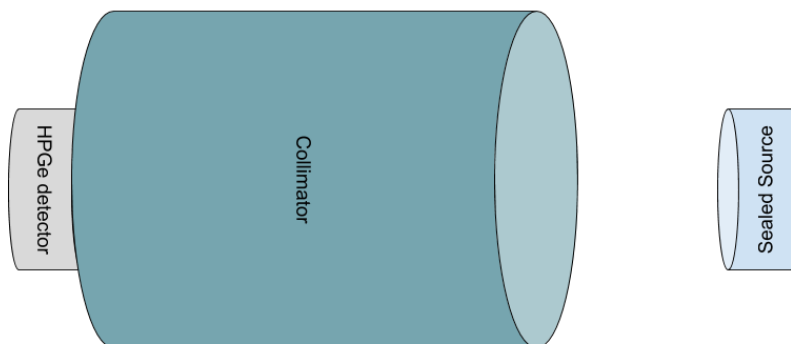

Figure 5: *Measurements geometry.*

The experiment loop is composed by these steps: model creation, one model per network architecture; adjust the output layer to a multi class output; train; and validate the trained model.

The code used in this experiment was developed using Ubuntu 18.04 (kernel version 4.15) as operational system, python (version 3.6.6) as a programming language, and the deep neural network models were built using Keras (version 2.2.4), a deep learning framework created on top of Tensorflow (version 1.13.1), deep learning accelerating library for high-performance computing.

The training and inference were performed on an Intel I7 personal computer equipped with one Nvidia GTX 1060 GPU, using Ubuntu 16.04 as the operating system and Nvidia CUDA (version 9.2.148) and cuDNN (version 7.1.4) libraries for deep learning computations acceleration.

During the training phase, all experiments used Stochastic Gradient Descendant as the optimizer, binary cross entropy as loss function, learning rate of 0.001 and 250 epochs of training, and 0.5 (from 0.0 to 1.0) as the class threshold for accuracy score.

3. RESULTS AND DISCUSSION

Table 1 summarizes the results after training all models using the same training and testing set.

Table 1: Results for each network architecture

Network architecture	Training time (seconds)	Number of layers	Number of neurons	Validation loss	Validation accuracy
Xception [10]: use of residual connections with depth-wise convolutions layers instead of standard convolution layers	1750	36	21,822,698	0.50	18.12%
VGG-19 [11]: use of several consecutive convolution layers grouped 4 by 4	2250	19	70,404,042	0.22	80.62%
VGG-16 [11]: use of several consecutive convolution layers grouped 3 by 3	2000	16	65,094,346	0.21	80.62%
ResNet50 [12]: use of several consecutive convolution layers with some of the convolution blocks with residual connections (shortcuts) between the input and output of the block	1500	50	23,601,930	0.55	18.75%
MobileNet [13]: use of depth-wise convolution layers instead of standard convolution layers	500	25	3,238,538	0.39	21.87%
InceptionV3 [14]: use of several blocks consisted by the concatenation of different convolution layers with different filter sizes	1250	48	21,822,698	0.28	22.50%

As the performance results of VGG-16 and VGG-19 seems to be equal, the top_k_categorical_accuracy was the tiebreaker metric. This metric is measured on the accuracy of the correct prediction being in the top-k predictions, as the data mixes up to three radionuclides in each spectrum the k parameter, in this case, is equal to three. The results are reported in Table 2.

Table 2: VGG-16 validation results

Network architecture	top_k_categorical_accuracy (k=3)
VGG-19	95.62%
VGG-16	95.00%

The tiebreaker metric indicates the VGG-19 architecture as the architecture which the best performance for the classification task.

4. CONCLUSION

The experiments demonstrated the feasibility of deep neural networks for gamma-spectroscopy analysis, correctly identifying several radionuclides. This approach is innovative due to the use of the raw spectra data without any pre/post-processing.

REFERENCES

[1] W. S. MCCULLOCH; W. PITTS, "A logical calculus of the ideas immanent in nervous activity," **Bull. Math. Biophys**., 1943.

[2] W. C. Contributors, "Neuron Hand-tuned." [Online]. Available: https://commons.wikimedia.org/w/index.php?title=File:Neuron_Hand-tuned.svg&oldid=347708343.

[3] R. RAINA; A. MADHAVAN; A. Y. NG, "Large-scale deep unsupervised learning using graphics processors," 2009.

[4] D. CIREȘAN; A. GIUSTI; L. M. GAMBARDELLA; J. SCHMIDHUBER, "Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images," **NIPS**, 2012.

[5] O. RONNEBERGER: P. FISCHER; T. BROX, "U-net: Convolutional networks for biomedical image segmentation," **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, 2015.

[6] P. E. KELLER; R. T. KOUZES, "Gamma spectral analysis via neural networks," **Proc. 1994 IEEE Nucl. Sci. Symp**. - NSS'94, 1995.

[7] P. E. KELLER; L. J. KANGAS; G. L. TROYER; S. HASHEM; R. T. KOUZES, "Nuclear Spectral Analysis via Artificial Neural Networks for Waste Handling," **IEEE Trans. Nucl. Sci.**, vol. 42, no. 4, pp. 709–715, 1995.

[8] F. ROSENBLATT, "The perceptron: A probabilistic model for information storage and organization in the brain," **Psychol. Rev.**, 1958.

[9] 9. G. H. Yann LeCun, Yoshua Bengio, "Deep learning," Nat. Methods, 2015.

[10] F. CHOLLET, "Xception: Deep learning with depthwise separable convolutions," **Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017**, vol. 2017-Janua, pp. 1800–1807, 2017.

[11] K. SIMONYAN; A. ZISSERMAN, "Very Deep Convolutional Networks for Large-Scale Image Recognition," pp. 1–14, 2014.

[12]　K. HE; X. ZHANG; S. REN; AND J. SUN, "Deep residual learning for image recognition," **Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.**, vol. 2016-Decem, pp. 770–778, 2016.

[13]　A. G. HOWARD ET AL., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," 2017.

[14]　C. SZEGEDY, V. VANHOUCKE, S. IOFFE, J. SHLENS, AND Z. WOJNA, "Rethinking the Inception Architecture for Computer Vision," **Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.**, vol. 2016-Decem, pp. 2818–2826, 2016.