ASSOCIAÇÃO BRASILEIRA DE ENERGIA NUCLEAR – ABEN
2019 International Nuclear Atlantic Conference - INAC 2019
Santos, SP, Brazil, October 21-25, 2019

Agência Brasileira do ISBN
ISBN 978-85-99141-08-3
9 788599 141083

COMPARING DEEP LEARNING ARCHITECTURES ON GAMMA-SPECTROSCOPY ANALYSIS FOR NUCLEAR WASTE CHARACTERIZATION

André Gomes Lamas Otero[1], Ademar José Potiens Júnior, Eduardo Petrolini Calzeta, Julio Takehiro Marumo
Instituto de Pesquisas Energéticas e Nucleares (IPEN / CNEN - SP)
Av. Professor Lineu Prestes 2242
05508-000 São Paulo, SP
aglotero@usp.br 1

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, we explore the capabilities of deep learning on a new field: gamma-spectroscopy analysis, comparing the classification performance of different deep neural networks architectures. We choose VGG-16, VGG-19, Xception, ResNet, InceptionV3 and MobileNet architectures which are available through the Keras Deep Learning framework to identify several different radionuclides (Am-241, Ba-133, 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 created a dataset that was used for the training and validation of the neural networks comparison. This study demonstrates the strengths and weakness of applying deep learning on gamma-spectroscopy analysis for nuclear waste characterization.

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.

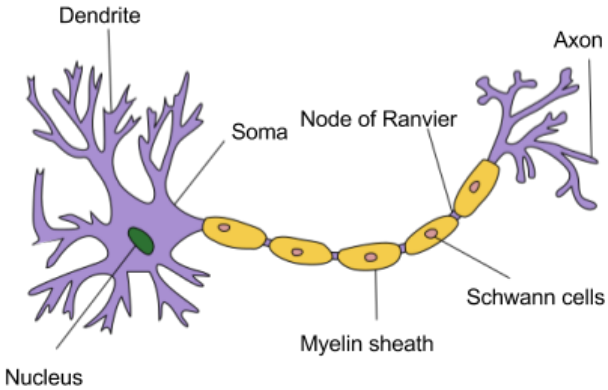

Figure 1: The biological neuron.

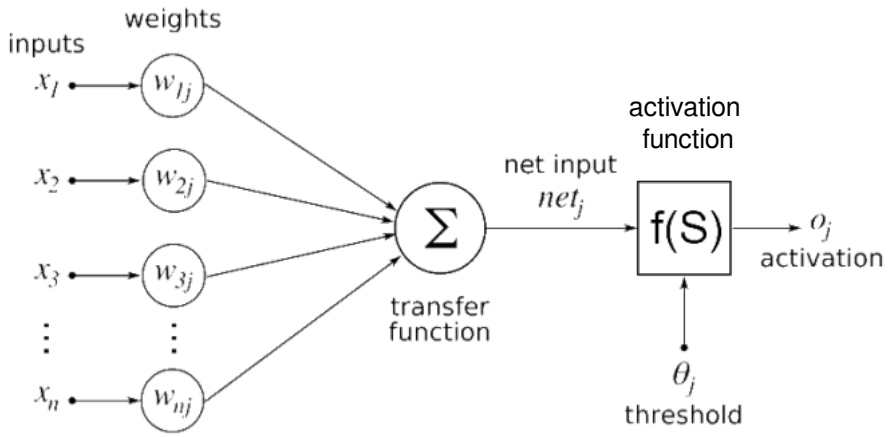

Figure 2: One artificial neuron from a perception neural network.

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, researchers are capable of trying new architectures for new models, achieving a performance level that surpasses humans in several tasks [3,4,5].

These successes instigated our team to evaluate the performance of well known deep neural networks architectures in a new task: gamma-spectroscopy analysis for nuclear waste characterization.

Past 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 set must fit into the main computer memory; there is only one activation function.

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

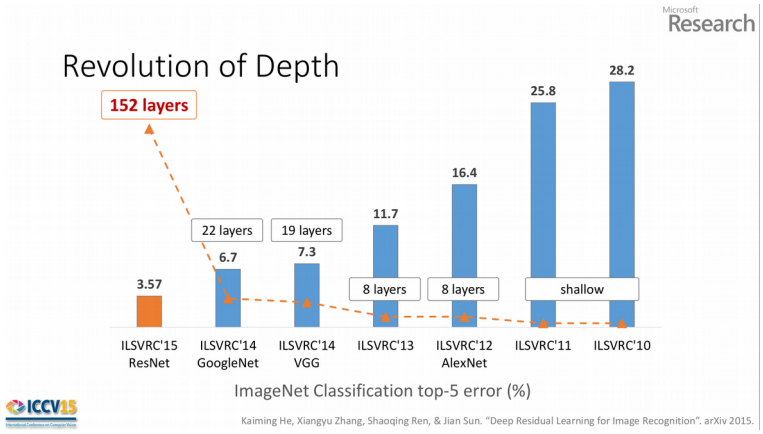

Figure 3: The impact of larger networks on classification performance.

At the IPEN Radioactive Waste Management Department, we must characterize the final product from the wast treatment. Nowadays, this work is manual. In this study, we present the results of different deep neural networks architectures on gamma-spectroscopy classification, which is: given a spectrum the deep neural network must say the radionuclide that composes the spectra, which can be more than one — aiming future automation of this step of the waste characterization process.

2. MATERIALS AND METHODS

We performed measurements from different sealed sources, Figure 4, to create the base dataset, from this base dataset we generated several new spectra that contains random radionuclide combinations, up to three radionuclide, then we split this base dataset into train and test sets. We create a validation set using the original measurements, plus one measurement from a triple sealed source, containing Am-241, Cs-137, and Co-60.

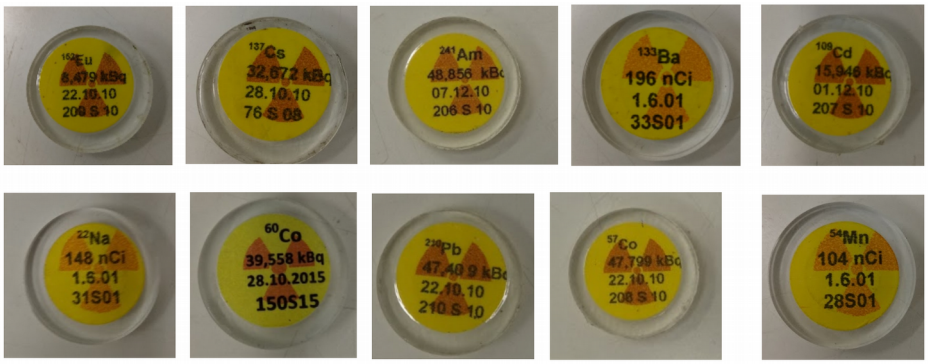

Figure 4: The sealed sources used on this study.

We have 11 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.

At Figure 5, we report the geometry of the measurements. The sealed source is 5 cm from the collimator witch an HPGe detector.

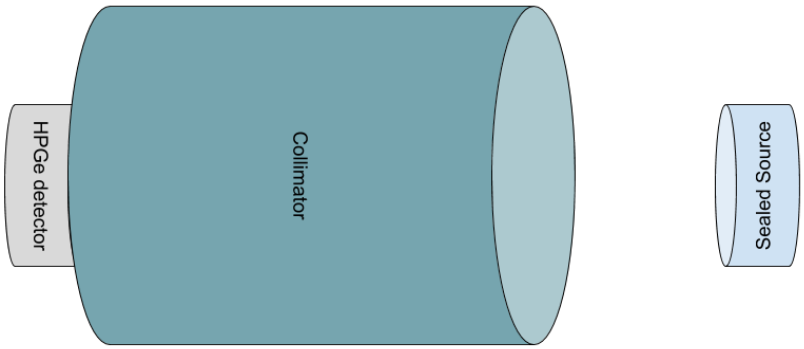

Figure 5: Measurements geometry.

The experiment loop consists in: instantiate the model, adjust the output layer to a multi class output, train, and validate the trained model.

The code developed used Ubuntu 18.04 (kernel version 4.15) as host operational system, python (version 3.6.6) as a programing 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.

We performed training and inference 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 Descendent as the optimizer, binary cross entropy as loss function, learning rate of 0.001 and 250 epochs of training, our accuracy score use 0.5 (from 0.0 to 1.0) as the class threshold.

The model is feed with the raw spectra data.

3. RESULTS

After training all models using the same data, we end up with the results reported in Table 1.

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%

Network architecture	Training time (seconds)	Number of layers	Number of neurons	Validation loss	Validation accuracy
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, we compute one more metric to be the tiebreaker, the top_k_categorical_accuracy, with is measured on the accuracy of the correct prediction being in the top-k predictions, as we are mixing up to three radionuclides in each spectrum our k = 3. Results reported in Table 2.

Table 2: VGG-16 validation results – part 1

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

Using as tiebreaker the metric that computes the accuracy on top-k predictions we choose the VGG-19 as the best model for our task.

4. CONCLUSIONS

Our experiments demonstrated the feasibility of deep neural networks for gamma-spectroscopy analysis, correctly identifying several radionuclide. Our approach is innovative as we use the raw spectra data without any pre/post processing.

ACKNOWLEDGMENTS

We thank IPEN-CNEN/SP and CAPES for financial support.

REFERENCES

1. W. S. McCulloch and 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, and A. Y. Ng, “Large-scale deep unsupervised learning using graphics processors,” 2009.

4. D. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, “Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images,” in *NIPS*, 2012.

5. O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015.

6. P. E. Keller and 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, and 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. 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 and 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.