Real-time radionuclides detection using artificial intelligence

Filipe André Carvalho Mendes*

Gamma-ray spectroscopy is the usual method to identify detected radioactive hot-spots. Classical Gamma spectroscopy involves many phases, longer analysis and usually an expert to reach an identification result. Thus, developing and upgrading the identification systems available has been a challenge for security and defence organisations such as Departments of Homeland Security, Emergency Response Teams, Customs and Border Control.

This work proposes an approach using machine learning techniques that is intended to be implemented as an easy to use identification system, meaning that it can be used by anyone without experience in the field. The proposed solution makes use of artificial neural networks to produce a classification to a given spectrum obtained with a CZT sensor. The system is trained using simulated data and is then tested with real acquisition spectra. Single and multiple isotope identification on each sample is explored, highlighting the benefits of an implementation of this kind as well as possible improvements.

Additionally, an example of a portable application is suggested using a Raspberry Pi. It is noteworthy that the artificial neural networks developed could be implemented in other devices such as a mobile phone with a connection to a detector. This kind of standalone and portable system could be used on site by humans or even by unmanned vehicles such as drones.

I. INTRODUCTION

The majority of the people around the world is constantly exposed to radiation arising from their phones, microwaves, routers, communication antennas, high exposure to the sun. More often than not, this kind of exposure is not harmful to living beings, however not every type of radiation is equal. Radiation with higher energy is usually referred to as ionising radiation. These energetic beams represent a real threat to our well being since they are energetic enough to modify the DNA structure in human cells, causing damage to tissues and organs. For that matter, it is of high importance to detect and identify possible radioactive sources.

There has been a growing concern on this matter of radiation control especially due to the raise of awareness among people, fuelled by disasters like Chernobyl and Fukushima. Nuclear activated components can be found in a variety of different scenarios such as reactor outages, laboratory test facilities, storage areas for contaminated fusion and fission remains and even particular equipment used, for instance, in health care. Criminal and unauthorised acts related to nuclear and other radioactive material such as illegal dropout of activated substances in remote locations and terrorism are also to be taken into consideration. Naturally-Occurring Radioactive Materials (NORM) [1] is a term used to specify all naturally occurring radioactive materials which occur naturally or where human activities have increased the potential for exposure compared with an unaltered situation. These materials potentially include all radioactive elements found in the environment such as uranium, thorium and potassium. Any of their decay products also represent a serious threat. Moreover, with the current rates of using nuclear reactions, namely for electricity production, the possibility of nuclear accidents cannot be neglected, thus control and monitoring must be enhanced and improved. Not only detecting but also identifying the detected radioactive sources is important. Identifying an unknown source is usually relevant since this information might hint the cause for such radiation detection and also define how to mitigate its risks.

Artificial Intelligence (AI) applications have been increasing in the scientific and even commercial areas in the past few years. Following this huge dissemination of AI methods such as artificial neural networks (ANN) and Deep Learning, it was figured out that there could be some room to explore this recent developments in this area and establish a connection between AI and gamma-ray spectroscopy, contributing to the improvement of current Gamma-ray analysis applications.

II. RADIOACTIVITY

Atoms found in nature are either stable or unstable. The instability of an atom's nucleus may result from an excess of either neutrons or protons, or even excessive energy. A radioactive atom will attempt to reach stability by ejecting nucleons (protons or neutrons), as well as other particles, or by releasing energy in other forms [2] .

* filipe.c.mendes@tecnico.ulisboa.pt

This process of attempting to reach a more stable form for the atom is denominated radioactive decay. Depending on the type of process itself, this decay can be sorted out in different categories: Alpha, Beta and Gamma.

An Alpha decay consists in the transformation of the atom nucleus (parent) into a new nucleus (daughter) by the emission of an alpha particle. Similarly, a beta decay consists in the emission of a beta particle (electron or positron) from the parent nucleus.

In a gamma decay the nucleus simply transitions from a higher energy state to a lower energy state releasing energy by means of the emission of electromagnetic radiation (gamma ray). Following a α or β decay, the daughter nucleus may be left in one of these high energetic states, progressing to a more stable state via a supplementary gamma decay. Since gamma-ray carries no charge nor does it have an associated mass, there is no change in the element as a result of emission of a gamma rays.

A. Interaction of Gamma Radiation with matter

The majority of gamma ray interactions with matter can be described by 3 main processes: Photoelectric effect, Compton scattering and Pair production [3] .

In the photoelectric effect the incident photon gives all of its energy to a bound electron in an atom, leading to the ejection of the specific electron with kinetic energy (E_e) equal to the difference between the incident photon energy (E_γ) and the binding energy (E_b) of such electron.

$$E_e = E_\gamma - E_b \tag{1}$$

This process allows the detector to accurately measure the energy that the incident photon transferred to the electron, corresponding to a well defined energy peak in the spectrum. These peaks are denominated Full-Energy Photopeaks (FEP).

Compton scattering occurs when an incident photon transfers part of its energy to a free or loosely bound electron via collision process. The amount of energy transferred is dependant on the angle between the direction of the incident photon and the direction of the scattered photon. The scattered photon leaves the detector and the detected energy is the kinetic energy of the electron. The energies of the scattered photon and electron are given by:

$$\begin{aligned} E_{\gamma'} &= \frac{E_\gamma}{1+E_0(1-\cos\theta)} \\ E_e = E_\gamma - E_{\gamma'} &= E_\gamma - \frac{E_\gamma}{1+E_0(1-\cos\theta)} \end{aligned} \tag{2}$$

where E_γ is the incident gamma ray energy, $E_{\gamma'}$ is the scattered photon energy, E_e is the electron energy, E_0 is $\frac{E_\gamma}{m_e c^2}$ and θ represents the scattering angle for the scattered photon.

Pair production consists in the creation of a positron-electron pair when the gamma ray is travelling through matter, usually in the vicinity of an atomic nucleus. To make this process possible, the incident gamma ray must have at least 1.022 MeV of energy which corresponds to the combined rest mass of those two particles. The positron is unstable causing it to lose its kinetic energy and find an available electron to annihilate. During this annihilation, two gamma photons with the energy of 511 KeV are created in opposite directions.

III. GAMMA-RAY SPECTROSCOPY

Gamma-ray spectroscopy is an analytical technique that can be used to identify various properties of the radioactive isotopes present in a sample of a specific radioactive substance. The energy of incident gamma-rays produced by the sample are acquired and measured by a detector, being then compared to the known energy of gamma-rays produced by radioisotopes and determining the identity of the emitter. This technique has many applications, such as in material analysis, geological exploration or even computer tomography.

In order to analyse the sources of radiation, the detected gamma emissions are measured and used to produce an energy spectrum. A detailed analysis of this spectrum is useful to determine the identity and quantity of gamma emitters present in a sample.

A. Gamma-ray Spectrum Components

Gamma ray spectrum is basically an histogram that represents the number of occurrences detected by the spectrometer system for a certain energy. Each individual radionuclide has a different decay scheme and consequently a distinct spectrum that is used to identify them. Gamma rays entering the detector can undergo any of the possible interaction processes described in Section II. These processes are responsible for the several features that can be encountered in the spectrum. Relevant spectrum features include: Full Energy Photopeaks (FEP), Compton Edge and Continuum, Backscaterring peaks, X-ray peaks, Annihilations peaks and escape peaks (see Fig. 1).

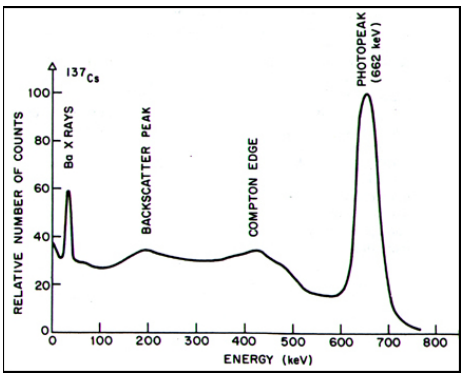


Figure 1: Example of Cesium-137 spectrum [4].

IV. STATE OF THE ART METHODOLOGY

After understanding the several features that may be present in a gamma spectrum it is important to understand the methodology used in order to apply that knowledge in a way that enables the user to identify the radionuclide in question. Approaches to radionuclide identification can be divided in two wide-ranging categories: peak search and template matching.

A. Peak search identification methods

The most commonly used algorithm to perform a peak search is called derivative search [5]. This procedure consists in applying the first and second derivatives to the entire spectrum and identify possible peaks by the signal variation of such derivatives.

If all spectra were a smooth and continuous curve, a derivative peak search could be immediately applied to the raw spectrum. Unfortunately, this is not the case of real acquisitions which contain undesired noise that must be filtered/smoothed. Several different smoothing techniques may be applied for gamma-ray spectroscopy [6]. From simple moving average to far more intricate and complex polynomial approaches such as B-splines [7] or Savitzky-Golay filter [8]. Fourier transforms [9] have been utilised for this matter as well as some more recent procedures such as wavelet analysis [10][11].

Another possible alternative to derivative peak search is to use "deconvolution" methods. Deconvolution is an algorithmic approach that tries to reverse the effects that a physical detection system has on the original theoretical source spectrum. Some examples of deconvolution techniques include Maximum Likelihood Expectation Maximization (ML-EM), Maximum Entropy Method (MEM) and linear regularization [12].

The final classification procedure consists in gathering all the data obtained in the previous steps and matching it against a known nuclide library in order to possibly identify the radionuclides present in the sample. Classifier algorithms that have been applied to radionuclide identification include [6]: Expert systems (essentially hard coded decision trees) [13], Naïve Bayes [14][15][16], Nearest neighbour [17][18] and also Support vector machines [19][20].

B. Template Matching identification methods

Template matching requires a separate type of algorithms that can search the space of possible radionuclides in order to find the correct mixture that contributes to the a certain spectrum. These types of algorithms fall into two broad categories [6]: Heuristic and Systematic.

Heuristic algorithms involve a strategy based approach that compares the sampled spectrum with a huge number of possible spectra. One solution is to start from the full set of possible radionuclides and sequentially eliminate them on the way to a solution (strip down) [21][22].

Differently from heuristic ones, systematic algorithms consider multiple possible solutions at each decision node, therefore decreasing the problem of path dependence in heuristic methods [6].

C. Artificial Neural Networks previous implementations

The majority of the aforementioned algorithms require specific data pre-processing in order to be properly utilised. Some of them inclusively require special tuning of parameters, making those types of approaches unsuitable for being used by non expert users. This kind of problems can be overcame by making use of artificial intelligence algorithms such as ANNs. Although ANN implementation can be extremely consuming both time and resource wise due to training and the usual need for high performance computers, all of these difficulties are condensed in the development stage. After development, the algorithm requires a relatively low computational power to be used (this heavily depends on the purpose of the ANN) and more importantly, presents a far more user friendly experience since the user provides the input and receives an output without any additional parameters to tune. Additionally, the successful implementation of ANNs in other complex areas such as computer vision and image classification [23], further enhances the significance of exploring such approach.

ANN approaches have already been applied in several previous works using low to medium energy resolution detector systems (scintillator detectors) [24][25][26][27]. Since the input for the ANN is the entire spectrum, there is no need for more complex pre-processing procedures, such as transformations or filters, making this applications easy to operate by the end user. To the best of my knowledge, no work was found on ANN development for radionuclide identification using high resolution detectors such as Cadmium zinc telluride (CZT) ones. Taking this into consideration, the goal of this thesis is to further explore the possibilities of using ANNs to analyse gamma-ray spectra from a CZT sensor and identify radioactive isotopes.

V. PROPOSED SOLUTION

In this work an ANN approach is proposed as a possible solution to perform isotope identification on samples acquired with a CZT detector.

Artificial Neural Networks are one of the most widely used tools of Machine Learning today. As the "neural" part of their name suggests, the way they work is inspired in the human brain as they are intended to replicate the way that we (humans) learn.

A simple ANN is composed by an input layer with as many neurons as there are input signals, at least one hidden layer and an output layer with as many neurons as output signals.

Each neuron-neuron interaction can be described by receiving the input signals, multiplying them by each respective weight and sending these values to an activation function that delivers a final output signal. These weights are the constants that we get from the training/learning procedures. These functions, also known as transfer functions, act as thresholds for the signal usually outputting values between 0 and 1 or -1 and 1.

During the training process, the weights of the ANN are updated with the intention of minimising the error in relation to the expected output for each of the training data examples. This error can be evaluated and measured by what is called a loss function. Commonly used loss functions include Mean Squared Error (MSE), Mean Absolute Error (MAE) normally used for regression problems or even Hinge Loss or Cross Entropy that are usually applied to classification problems [28]. This process of minimisation of a function with respect to a set of parameters (in this case the loss function in respect to the weights) is at the root of many computer science issues [29] and can be carried out by optimisation algorithms like Gradient-based learning. Popular algorithms include Stochastic Gradient Descend (SGD), Momentum based GD, Nesterov Accelerated Gradient Descent (NAG), RMSprop and Adaptive Moment Estimation (ADAM), every single one of them being some variant of the classical Gradient Descent Algorithm. ADAM which is a combination of RMSprop and Momentum is considered to be current state of the art [30].

VI. SOLUTION IMPLEMENTATION

A. Training Data Generation

Since obtaining a data-set composed of real spectra was not possible, the selected approach was to resort to simulated data and use them as training samples. In order to obtain such material, a simulator was used: GADRAS-DRF from Nuclear Energy Agency [31].

In this work, the idea is to develop an ANN that can identify the data acquired with a specific CZT detector. Thus, the detector dimensions were given a fixed set of values that were obtained by looking at the data-sheets [32] of the detector in use.

Both energy resolution and energy calibration parameters of the simulated detector were defined using data provided by [33] where they performed an analysis of the CZT detector used in this work. The channel number was set to 1024, since this analysis used this configuration.

A range of values were chosen for each relevant parameter of the simulator in order to obtain a more complete data-set. A more detailed description of each of the parameters used for the simulation of the data-set can be found in the complete thesis document.

The final list of selected radionuclides used for learning is:

- ^{22}Na, ^{60}Co, ^{137}Cs, ^{152}Eu, ^{241}Am, ^{226}Ra, ^{228}Ac, ^{235}U, ^{40}K, ^{133}Ba, ^{222}Rn, ^{57}Co, ^{54}Mn, ^{204}Tl, ^{7}Be - 15 Isotopes

The following step was to define the simulation parameters namely: acquisition time, distance from source to detector, activity of the source and background noise. A combination of an activity of 1 μC and an acquisition time of 3600 real seconds was chosen, this choice being based on characteristics of previous experimental acquisitions. Distance to the source was set to 10 cm, height was set to 100 cm and Poisson statistics were applied.

Taking into consideration that the ANN should be able to identify the radionuclides regardless of the location and background noise, a set of possible background values were picked for the simulations.

An independent spectrum was produced for each isotope for all the different detector configurations and simulation parameters previously described, scaling the number of single isotope spectrum samples to a total of **10 935 samples**. The same process was performed for multiple radionuclide simulations, where a combination of up to 3 isotopes were simulated in a single sample. In this case all radinuclides were simulated as having equal activity in the combination, since having different combinations of activity would lead to an enviable huge number of samples. A total of **76 545** and **331 695 samples** were simulated for dual and triple isotope combinations respectively.

B. ANN Structure Development

The idea for this implementation is to receive the full acquired spectrum after a maximum normalisation, hence the size of the input layer must be equal to the number of channels of the detector itself (1024).

Regarding the output of the ANN, the chosen approach was to have an output neuron for each of the possible identifiable radionuclides, totalling 15 neurons. For the case of single-isotope identification (multi-class classification) the activation of choice was Softmax since we want to predict the most probable radionuclide in the sample. When trying to identify more than one radionuclide in the same sample (multi-label classification) the approach must be distinct. For this case the activation function chosen was Sigmoid.

In this work, three different hidden layers structures were tested for each of the desired classification options (single and multi radionuclide identification). Each independent approach used one, two or three hidden layers. The base model is shown in Fig. 2.

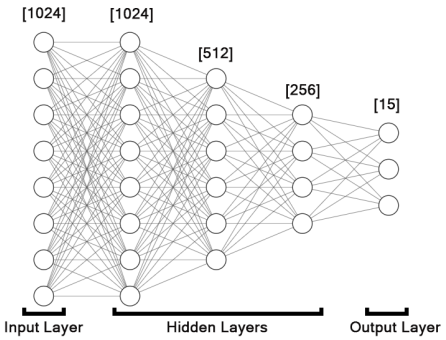


Figure 2: Base model for the artificial neural network structure.

ReLu activation was chosen for every neuron belonging to the hidden layers.

Tensorflow [34] and Keras API [35] were used for the implementation of the mentioned ANNs.

C. ANN Training

For this particular implementation, binary cross-entropy loss function was used for Sigmoid activation and categorical cross-entropy for Softmax activation. The learning algorithm chosen was ADAM optimiser.

The regularisation mechanisms that were analysed consisted in L1 and L2 regularisation. In addition, features such as batch size and number of epochs were studied in order to obtain better results.

The starting point for stopping mechanism implemented consisted in evaluating the training loss value and stop the learning process when this value did not reduce at least 10^{-4} over 10 consecutive iterations. Regarding the batch size, common values include 32, 64, 128, 256 [36]. The value of 128 was chosen as a first approach.

VII. PERFORMANCE ANALYSIS CRITERIA

The performance of the resulting ANNs was evaluated using a set of real spectra acquired in 3 different scenarios with the CZT detector previously introduced. The 3 groups of samples can be categorised as Single or Multi radionuclide acquisitions and include the radionuclide samples mentioned in Table I.

Set Name	Radionuclides	Sample Name	Aquisition Time	Background sample?
Single_Set #1 (SS1)	^{60}Co	*SS1_60Co*	30 min	Yes
	^{137}Cs	*SS1_137Cs#1*	30 min	Yes
	^{137}Cs	*SS1_137Cs#2*	30 min	Yes
	^{152}Eu	*SS1_152Eu*	30 min	Yes
	^{22}Na	*SS1_22Na*	30 min	Yes
Single_Set #2 (SS2)	^{60}Co	*SS2_60Co#1*	90 min	Yes
	^{60}Co	*SS2_60Co#2*	90 min	Yes
	^{241}Am	*SS2_241Am*	90 min	Yes
Multi_Set #1 (MS1)	^{226}Ra, ^{214}Pb, ^{214}Bi *	*MS1_#1*	30 min	Yes
	^{226}Ra, ^{214}Pb, ^{214}Bi *	*MS1_#2*	30 min	Yes
	^{226}Ra, ^{214}Pb, ^{214}Bi *	*MS1_#3*	30 min	Yes
	^{226}Ra, ^{214}Pb, ^{214}Bi *	*MS1_#4*	30 min	No

Table I: Data-sets used for evaluating the performance of the ANNs. *** Expected radionuclides.**

SS1 and *SS2* were acquired at the Lab using sources with well-known identities, meaning that the radionuclides attributed to each of the samples are correct. For *MS1* data-set the spectra were acquired in a real scenario of an old ore mine, so the attributed radionuclides to each sample correspond to the expected identities obtained with other analysis.

Before classification, each spectrum was normalised and subject to a background subtraction procedure consisting in taking the original spectrum and subtracting the entirety of the background noise acquisition (proportional to acquisition time of each sample).

When evaluating the performance of a resulting ANN trained for classification, the usual approach is to run the algorithm to analyse a specific set of "problems" and attribute an evaluation of correct (if the classification was correct) or incorrect (otherwise) to each of the processed examples. At the end, the *Accuracy* of a given algorithm is obtained based on the ratio of right guesses against the total number of "problems" that were analysed from that specific test set. Additionally to this accuracy metric, *Precision* and *Recall* [37] are commonly used to analyse the performance of multi-label classification algorithms.

Finally, another interesting metric to analyse is the F1-score which essentially combines precision and recall values in order to obtain a final score.

VIII. SINGLE RADIONUCLIDE CLASSIFICATION RESULTS

In order to obtain an optimal identification system, several steps of optimisation procedures must be taken. Only the final ANN results are shown in the extended abstract. A more detailed description of the relevant optimisation phases is presented in the complete thesis document. The following results are obtained using *SoMx1* configuration which uses only one hidden layer fo 1024 neurons, for a batch size of 128 and L1 and L2 regularization of 0.001.

A. Background Subtraction effect

Background subtraction is a mechanism that intends to filter some of the noise present in the spectra of the samples. This process consist basically in acquiring a sample spectrum of the area where little to none radioactivity from the source under analysis is detected and subtracting it from the acquired spectra. Background noise sample acquisition is not always possible, making it relevant to analyse the performance of the ANN when classifying samples without background subtraction.

Table II reveals the classification output for each sample with and without background subtraction. Note that not having background subtraction affects especially the isotopes with lower activity counts or with acquisitions made farther from the source, where the ratio between

source and background counts is lower. This might be the case of samples *SS1_60Co* and *SS1_137Cs#2* which show significantly better results when background subtraction is applied. Unexpectedly, the confidence value for *SS1_22Na* increased without background subtraction, despite it being the sample with lower activity.

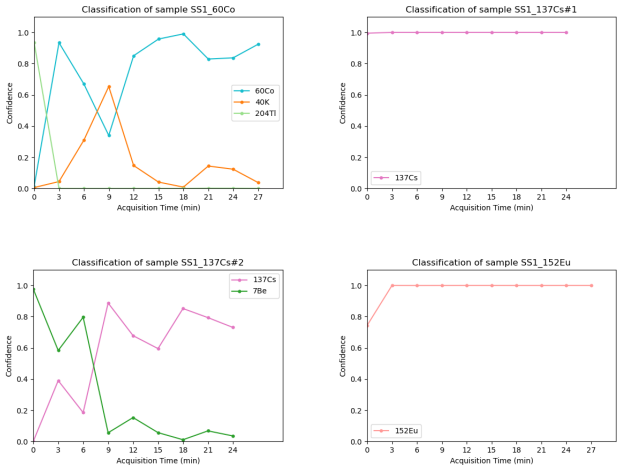
Background Subtraction	**Yes**	**No**
SS1_60Co	$^{60}Co - 91\%$	$^{22}Na - 53\%$ $^{40}K - 41\%$
SS1_137Cs#1	$^{137}Cs - 100\%$	$^{137}Cs - 99\%$
SS1_137Cs#2	$^{137}Cs - 72\%$	$^{22}Na - 18\%$ $^{204}Tl - 53\%$
SS1_152Eu	$^{152}Eu - 99\%$	$^{152}Eu - 99\%$
SS1_22Na	$^{22}Na - 42\%$	$^{22}Na - 96\%$
SS2_60Co#1	$^{60}Co - 99\%$	$^{60}Co - 91\%$
SS2_60Co#2	$^{60}Co - 55\%$	$^{60}Co - 56\%$
SS2_241Am	$^{22}Na - 15\%$ $^{241}Am - 65\%$	$^{22}Na - 15\%$ $^{241}Am - 65\%$
Precision **Recall**	0.88 1.00	0.75 0.55
F1-score	**0.94**	**0.63**

Table II: SoMx1 classification results with and without background subtraction. Regularisation of $L1 : 0.001$ and $L2 : 0.001$

Based on these results, background subtraction should be performed whenever possible since it led to an overall better classification.

B. Acquisition Time effect

The Acquisition time for a given sample has an impact on the form of the spectrum, since some features might take longer to be defined. It is then relevant to examine how the acquisition time of each sample spectrum would affect the classification output provided by the ANN. For this matter, the spectrum of each sample was classified every 3 minutes, creating a graph that represents the ANN output in function of acquisition time.



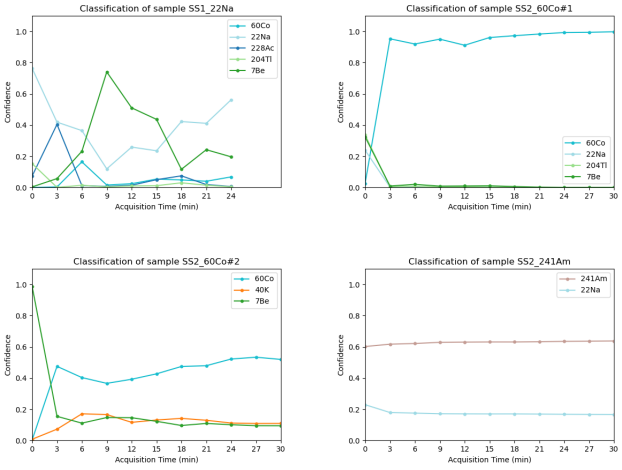


Figure 3: Classification of each sample along acquisition time.

Fig. 3 reveals that the majority of the samples take between 3 and 9 minutes of acquisition time to reach their final classification confidence levels. Sample *SS1_60Co*, *SS1_137#2* and *SS1_22Na* are the ones that exhibit noticeable fluctuations until reaching the final classification. This can possibly be explained by the fact that these spectra present higher background noise than the other samples, hindering the classification of the ANN.

Overall, this analysis suggest that the designed ANN is suitable for quick isotope identification with subjectively sufficient confidence levels.

C. False Positives Analysis

In order to test the possibility of obtaining False Positive outputs when using the resulting ANN, a data-set containing 4 new isotopes was created. These isotopes were: ^{65}Zn, ^{243}Am, ^{238}U and ^{109}Cd. The new data-set was generated equally to the training set, totalling 729 samples for each isotope. None of these samples was fed to the ANN during training, so the expected classification output would be to get no identification at all.

Ideally there should be a class that represents "none of the above". Unfortunately, a class like this does not always make sense or is even achievable. In this particular case, creating a class like this would require the simulation of a large number of samples consisting of all the other possible radionuclides. If we were to simulate a data-set like this one, it would make more sense to label all these extra samples and define new classes for them. Since a class of this kind is not possible, the ideal way of solving this problem is to establish some thresholds for the confidence levels output by the ANN.

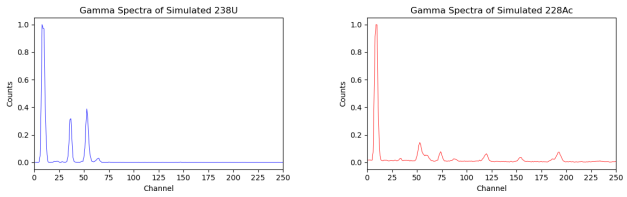
The perfect output for "none of the above" cases would be to have the probability distributed over the total classes. In this case, each output neuron should give a value of 6.66%. A usual threshold for classification is the 50% value.

Isotopes	False Positives		Total
^{65}Zn	571 114	^{40}K ^{60}Co	685
^{109}Cd	415	^{241}Am	415
^{238}U	729	^{228}Ac	729
^{243}Am	27 216	^{22}Na ^{222}Rn	243

Table III: False positives results for *SoMx1* using a threshold of 50%

Looking at Table III it is possible to notice that the results for each of the tested isotopes present a high number of false positives relatively to the total 729 samples. This can be elucidated by the fact that an ANN with Softmax activation function in the output layer will always try to provide a classification result.

These results can actually be explained by the form of each spectrum. For example, the ANN classifies all ^{238}U samples as being ^{228}Ac since they have two highly relevant peaks in the same energy range Fig.4. All the remaining results could be explained similarly.



(a) Spectrum of ^{238}U. (b) Spectrum of ^{228}Ac.

Figure 4: Comparison between Simulated spectra from ^{238}U and ^{228}Ac ([0,250] Channels). The number of counts is normalised to the maximum value.

In order to avoid this kind of problem, a possible solution would be to either train the ANN with more data or develop different ANN that can identify specific radionuclides. The ANN should always be trained with all the isotopes that we want to identify and should only be used when radioactivity has been previously detected for example by a Geiger-Muller counter, in order to avoid false positives when analysing background noise only.

D. Final ANN Results

The ANN that provided better results for single radionuclide identification was *SoMx1*, obtaining very interesting F1-score values for the utilised evaluating data-set. Additionally, it is interesting to analyse more closely the classification of each of the 8 samples that are part of the evaluating data-set. Each final output produces a graph containing the final acquisition spectra and highlighting the expected peaks from each of the identified isotopes.

SS1_60Co sample in particular is difficult to classify even for an expert since the resulting spectrum contains a large amount of background noise (see Fig. 5).

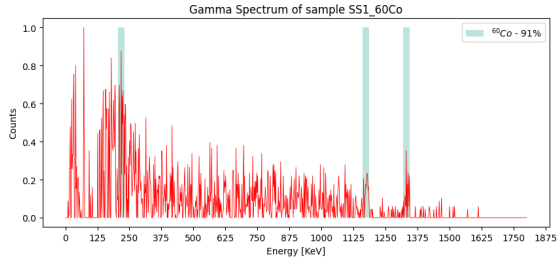


Figure 5: *SS1_Co60* spectra with classification peaks.

It is difficult to describe exactly how the ANN classifies a specific sample, since it simply receives an input and gives an output. However, in this case the correct classification of the isotope is possibly due to the network being able to identify the photo-peaks present in the 1173 keV and 1332 keV areas. If the photo-peaks were from a lower energetic area, they could be more difficult do distinct from the noise.

Signalling relevant peaks from the identified radionuclides in the spectra provides relevant information that may help to visually confirm the classification obtained by the ANN.

The spectrum of *SS1_137Cs#1* and *SS1_137Cs#2* samples were acquired in similar conditions, so a discrepancy in the classification output is something that should be explained. The correct classification of sample *SS1_137Cs#1* is observed among almost all the analysed ANN structures. In fact, the only relevant difference in the measurements of these two sources was the distance between the detector and the source.

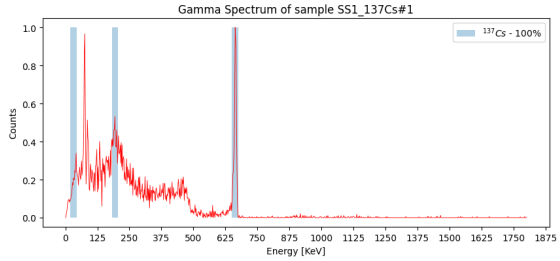


Figure 6: *SS1_137Cs#1* spectra with classification peaks.

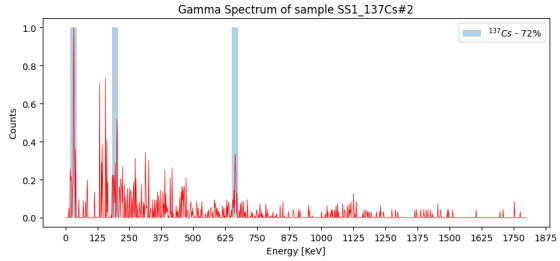


Figure 7: *SS1_137Cs#2* spectra with classification peaks.

Sample *SS1_137Cs#1* was acquired much closer to the source, revealing a much more consistent classification result. Being further form the source causes the ratio between actual source counts and background noise counts

to be lower, making it more difficult for the ANN to correctly classify the spectrum. Although a good background subtraction procedure can attenuate this problem, we can clearly observe that sample *SS1_137Cs#2* (7) is much noisier and consequently has a lower confidence level output than sample *SS1_137Cs#1* (6) .

Similarly to *SS1_137Cs#1*, the *SS1_152Eu* sample was correctly identified by almost every ANN configuration. This can be explained by the fact that the acquired spectrum is quite similar to the training ^{152}Eu spectrum, having well defined photopeaks.

The final ANN struggles to correctly identify the ^{22}Na radionuclide present in the *SS1_22Na* sample. From Fig. 8 it is possible to observe that the identifying peaks are quite difficult to distinct from the noise.

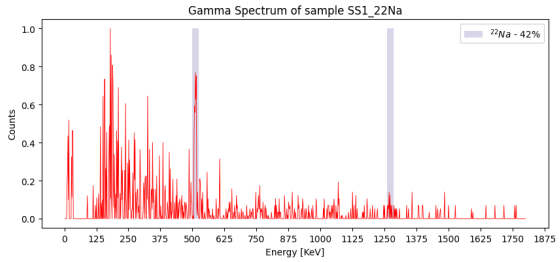


Figure 8: *SS1_22Na* spectra with classification peaks.

This difficulty is probably the reason for the lower confidence level of only 42%. Although the 1274 keV photo peak is extremely difficult to recognise, the annihilation peak of 511 keV is quite visible.

Regarding the samples of *SS2*, all of the acquisitions were made equally, meaning that the disparity between the classification results for samples *SS2_60Co#1* and *SS2_60Co#2* must be analysed.

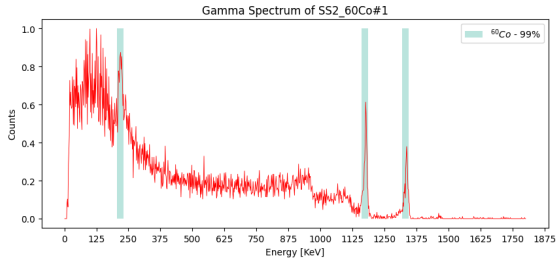


Figure 9: *SS2_60Co#1* spectra with classification peaks.

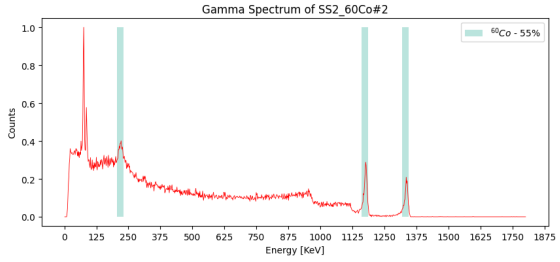


Figure 10: *SS2_60Co#2* spectra with classification peaks.

From Fig. 9 and Fig. 10 we can see that the spectrum from sample *SS2_60Co#2* contains a peak in the 70-80 keV region. This is the characteristic X-Ray peak from the lead casing that is covering the source. It is possible that this extra peak can cause some confusion to the ANN, leading to a worst classification confidence in the sample that is encased in lead.

Finally, the classification result for sample *SS2_251Am* is correct but with a relatively low confidence.

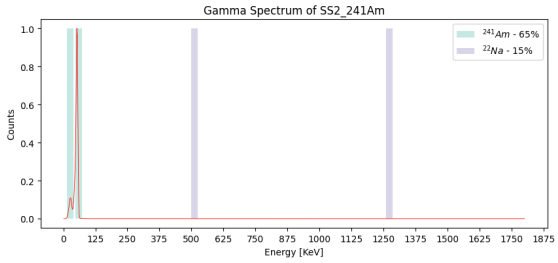


Figure 11: *SS2_241Am* spectra with classification peaks.

This spectrum is really well defined so a possible reason for this confidence value is related with the fact that the photo peaks of ^{241}Am are encountered in the lower energetic region of the spectra. Peaks in this area can be mistakenly considered as noise by the ANN. Consequently, the ANN might try to identify another isotope in the sample. This is likely the case of misidentification of ^{22}Na in this sample, since its photo peaks are far away from any relevant structure in the spectrum as seen in Fig. 11.

In this case, signalling relevant peaks enables us to quickly discard ^{22}Na from the possible isotopes by visually understanding that it is most likely not present.

IX. MULTI RADIONUCLIDE CLASSIFICATION RESULTS

For multi radionuclide classification the same principles used in the previous section can be applied. One of the major difficulties in this case is dealing with such a huge number of data samples (more than 400 thousand). With the intention of reducing the implementation and optimisation process, it was assumed that the same ANN structure analysis made for single isotope could be applied to the multi radionuclide ANN.

A. ANN Results and Discussion

The parameters chosen for regularisation and early stopping criteria were the same as the final ANN for the previous section.

Based on the results of the previous section, the preferred ANN structure was using only one or two hidden layers. Softmax activation function can no longer be used for this classification since more than one isotope can be

identified, meaning that the output values of the ANN do not sum up to 1.

Both *Sig1* and *Sig2* were trained using the whole single and multi data-set for batch sizes of 32, 64 and 128.

Similarly to single isotope identification, evaluating the resulting ANN using the training data set resulted in a perfect score, obtaining an F1-score value equal to 1. However, this was not expected and hints that the ANN might be over-fitting.

The performance analysis on the evaluation data-set started with only the single isotope samples, since a multi isotope classification solution should perform adequately with single identification as well.

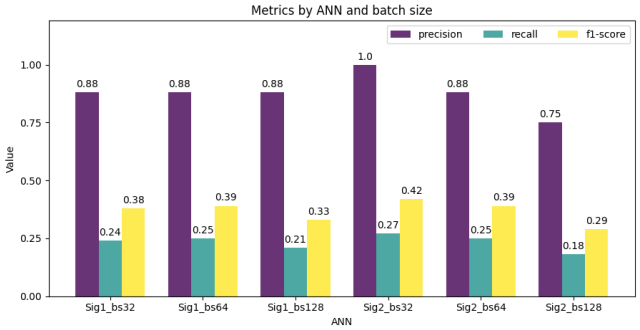


Figure 12: Precision, Recall and F1-score values for each ANN configuration using single isotope evaluation data-set.

The results shown on Fig. 12 present a low F1-score value for every ANN structure tested. Note that the precision metric shows an overall high number, suggesting that the ANN is capable of identifying the present radionuclides. However, the recall values are low due to the huge number of FPs that the networks output. These results might be related to how the expected output is fed to the ANN during training.

The expected output fed to the network is an array containing 0s and 1s. The 0 value indicates that the isotope is not present in the spectra and the 1 value indicates that the isotope is present. While in single identification this is an adequate approach, it was concluded that for multi isotope identification, representing the existing radionuclides with a simple 1 value might not be ideal. The problem arises when one of the present isotopes contributes with many more counts than the other, making the second isotope almost unnoticeable. When this happens, a spectrum with only the more active radionuclide might look nearly identical to a spectrum containing both radionuclides.

Defining the expected output containing equal confidence values for both isotopes can possibly lead the ANN to output more false positives, since the ANN is supposed to learn something that is not as noticeable. This makes the network look for smaller variations in the spectra, possibly mistaking noise for a radionuclide and leading to incorrect classifications.

Additionally, for multi isotope identification, the training data labels should take into account the relative contribution of each of the present radionuclides in the sample. A similar approach can be found in [25], where the number of counts that each radionuclide contributes to the spectra is controlled and taken into consideration. Another possibility would be to normalise each individual spectra and then sum the normalised counts of each one, creating a new spectra with well defined features from each of the desired radionuclides.

X. FINAL REMARKS

The work presented in this thesis provide a preliminary indication that ANN might be a promising solution for radionuclide identification. Although there is a need for further analysis and testing, especially due to the dimensions of the relatively small evaluation data-set available, the results provide a positive insight on this subject revealing some potential. Both the flexibility, as well as the accessibility of use by non expert users of the ANN makes this type of solution particularly interesting. All in all, radionuclide identification based on machine learning algorithms such as artificial neural networks should be further explored.

The following contributions were provided by this work:

- Creation of an ANN capable of identifying single isotope spectrum samples with relative high degree of confidence. This ANN was tested using real data acquisitions. The results show the ANN provides fine classification outputs even for small acquisitions times and could be possibly applied to real-time identification systems.

- It was shown that using Softmax activation function in the output layer resulted in a better classification as opposed to Sigmoid activation which presented worse results. The best performing network was the one containing a single hidden layer, since fewer hidden layers also resulted in better classification outputs.

- Despite providing relevant outputs for single isotope samples, the results for multi isotope classification were not so conclusive and additional works is required. It was possible to understand the importance of data-set labelling since this is the possible reason for such results. More samples need to be acquired since even if results were better, it would be difficult to correctly validate the performance of the ANN due to the lack of real data samples for multi isotopes.

- A proof-of-concept was developed using a Raspberry Pi and CZT detector. This real world application presented high portability and flexibility, being perfect for real-time acquisitions either on foot or using unmanned vehicles.

- Development of detailed data-set generation explanation using GADRAS-DRF software. This description can be used to create different data-sets for other detectors, making this approach easily scalable.

XI. FUTURE WORK

Several possible enhancements and upgrades came up while making this thesis. Some future ideas include:

- Explore the labelling of the training samples for multi radionuclide identification applications. Training data labels should take into account the relative contribution of each of the present radionuclides in the sample. The same could be applied to generate more samples with various types of background contribution.
- Merge existing algorithms with this application. Using filtering techniques to filter some of the noise of the acquired spectrum could enhance the classification capabilities of the developed ANN. Nonetheless, it is worth mentioning that for a real time analysis, the pre-processing of the data should always be simple enough to be implemented in real time.
- Acquiring more real data would be ideal for upgrading the performance evaluation of the ANN and even contribute to possible optimisations. If big enough, some of the real data-set could be used to train the ANN.
- False positive occurrences should also be further analysed. Possibility of more complex threshold systems or even other ANN structures can be explored.

[1] "Naturally-Occurring Radioactive Materials (NORM)," https://www.world-nuclear.org, Accessed: 2019-12-29.
[2] Arpansa, Australian Government, "Radioactivity," https://www.arpansa.gov.au, Accessed: 2020-12-08.
[3] *Identification of Unknown Radionuclides*, TECHNICAL UNIVERSITY DRESDEN - Institute of Power Engineering Training Reactor.
[4] "137Cs Spectrum image," http://www.people.vcu.edu, Accessed: 2020-01-10.
[5] G. R. Gilmore, *Practical Gamma-ray Spectrometry* (2008).
[6] M. Monterial, K. E. Nelson, S. E. Labov, and S. Sangiorgio, (2019), 10.2172/1544518.
[7] M.-H. Zhu, L. Liu, Z. You, and A. Xu (2008) pp. 691 – 695.
[8] M. Deriche and M. Raad, Instruments and Experimental Techniques **54**, 30 (2011).
[9] A. Abdel-Hafiez, Pramana **67**, 457 (2006).
[10] C. J. Sullivan, S. E. Garner, and K. B. Butterfield, in *IEEE Symposium Conference Record Nuclear Science 2004.*, Vol. 1 (2004) pp. 281–286 Vol. 1.
[11] G. Yu, J. Gu, L. Hou, Z. Li, Y. Wang, and Y. Zhang, Science China Physics, Mechanics and Astronomy **56** (2013), 10.1007/s11433-013-5185-3.
[12] L. Meng and D. Ramsden, Nuclear Science, IEEE Transactions on **47**, 1329 (2000).
[13] H. T. T. Aarnio, P A and J. T. Routti, "Expert system for nuclide identification and interpretation of gamma-spectrum analysis," (1992).
[14] J. V. Candy, (2010), 10.2172/1116930.
[15] C. Sullivan and J. Stinnett, Nuclear Instruments and Methods in Physics Research Section A Accelerators Spectrometers Detectors and Associated Equipment **784** (2014), 10.1016/j.nima.2014.11.113.
[16] J. B. Stinnett, *Automated isotope identification algorithms for low-resolution gamma spectrometers*, Ph.D. thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois (2016).
[17] D. Boardman, M. Reinhard, and A. Flynn, IEEE Transactions on Nuclear Science **59**, 154 (2012).
[18] R. C. Runkle, M. F. Tardiff, K. K. Anderson, D. K. Carlson, and L. E. Smith, IEEE Transactions on Nuclear Science **53**, 1418 (2006).
[19] M. Alamaniotis, S. Lee, and T. Jevremovic, Nuclear technology **191** (2015), 10.13182/NT14-75.
[20] H. Hata and K. e. a. Yokoyama, Applied radiation and isotopes: including data, instrumentation and methods for use in agriculture, industry and medicine **104**, 143 (2015).
[21] e. a. T. B. Gosnell, J. M. Hall, *Gamma-Ray Identification of Nuclear Weapon Materials*, Tech. Rep. W-7405-ENG-48 (Lawrence Livermore National Laboratory, 1997).
[22] M. Hogan, S. Yamamoto, and D. Covell, Nuclear Instruments and Methods **80**, 61 (1970).
[23] "YOLO: Real-Time Object Detection," https://pjreddie.com/darknet/yolo/, Accessed: 2020-01-10.
[24] D. Liang and P. G. et al., Annals of Nuclear Energy **133**, 483 (2019).
[25] M. Kamuda and C. J. Sullivan, Radiation Physics and Chemistry **155**, 281 (2019), iRRMA-10.
[26] L. Chen and Y.-X. Wei, Nuclear Instruments Methods in Physics Research Section A-accelerators Spectrometers Detectors and Associated Equipment - NUCL INSTRUM METH PHYS RES A **598**, 450 (2009).
[27] M. Kamuda, J. Stinnett, and C. Sullivan, IEEE Transactions on Nuclear Science **PP**, 1 (2017).
[28] "Common Loss functions in machine learning," https://towardsdatascience.com, Accessed: 2020-01-10.
[29] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, Proceedings of the IEEE **86**, 2278 (1998).
[30] "A Deeper Look into Gradient Based Learning for Neural Networks," https://towardsdatascience.com, Accessed: 2020-01-10.
[31] "NEA - Nuclear Energy Agency," Accessed 06-June-2020.
[32] *Gamma-Radiation CdZnTe Microspectrometer with exchangeable detector modules μSPEC*, ZRF RITEC SIA, Gustava Zemgala gatve 71A.
[33] J. Borbinha and Y. e. a. Romanets, Sensors **20**, 1538 (2020).
[34] "TensorFlow," https://www.tensorflow.org/, Accessed: 2020-01-10.
[35] "Keras," https://keras.io/, Accessed: 2020-01-10.
[36] J. Brownlee, "A Gentle Introduction to Mini-Batch Gradient Descent and How to Configure Batch Size," https://machinelearningmastery.com, Accessed: 2020-01-10.
[37] "Precision and Recall," https://en.wikipedia.org/wiki/ Precision_and_recall, Accessed: 2020-01-10.