The Application of Machine Learning to Event Classification in Radiation Detectors

The Application of Machine Learning to Event Classification in Radiation Detectors

By Richard Garnett,

A Thesis Submitted to the School of Graduate Studies in the Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University

Doctor of Philosophy (2022)

Hamilton, Ontario (Department of Physics and Astronomy)

TITLE: The Application of Machine Learning to Event Classification in Radiation Detectors

AUTHOR: Richard Garnett (McMaster University)

SUPERVISOR: Dr. Soo Hyun Byun

NUMBER OF PAGES: xxvii, 146

Abstract

Radiation physics is typically troubled by the accompanying of an interesting radiation field with an uninteresting one, presented as the signal and the background respectively. In some instances this background is a consequence of the signal of interest and its interaction with the world around it. These issues can be alleviated by clever experimental design that take advantage of the highly predictable way in which radiation interacts with the world. This predictability can lead to structured data which lends itself to data driven techniques which enable the identification of signal and background. One such technique that can exploit the structure of radiation detection data is machine learning, which is the focus of this thesis.

The first application was event selection onboard a satellite, called Advanced Energetic Pair Telescope (AdEPT), being developed for photon polarimetry. This application utilizes an adaptation of Google's `GoogLeNet` which was trained off of simulated data produced in a simulation developed in `GEANT4`. The performance of this adaptation, named `GammaNet`, was investigated and found to achieve the desired 99.99 % background rejection while maintaining signal sensitivity ranging from 0.1±0.1 % to 69±2 %. The signal sensitivity range depends on the down sampling rate implemented and the energy of the incident photon.

The other application explored was neutron and photon separation in EJ-301 and EJ-309 neutron sensitive scintillators. This application, being experimental and not simulation based, required the implementation of an already existing classification routine to generate the labelled data necessary for machine learning.

This was accomplished by utilizing Tail-To-Total (TTT) Pulse Shape Discrimination (PSD) and algorithmically fitting the resultant photon and neutron populations. With this technique it was possible to correctly identify 75.3±0.5 % of the neutrons while removing 99.9999 % of the photons. The neutron identification rate was found to depend on the neutron source used, but was mostly accounted for when considering the differences in neutron energies for those sources.

Both machine learning applications had the feature extraction technique, Gradient-weighted Class Activation Mapping (Grad-CAM), applied to them. This technique produces a spatially correlated activation mapping for each class inside the neural network, identifying regions of the input that produced the greatest confidence to an output classification. The features observed from both applications aligned with the intuition of what comprises the dominant features of each signal and background.

Acknowledgements

Firstly I would like to express my gratitude to my supervisor Dr. Soo Hyun Byun for providing me with intellectual freedom, opportunities, and continued support since my first having you as a professor for MedPhys 4R06. Being able to freely pursue a project I thought was interesting and engaging made my time spent working on it feel effortless. I would also like to thank Dr. Andrei Hanu and Dr. Stanley Hunter for initiating my interests in machine learning by welcoming me at NASA for the summer before starting my graduate experience. Having been given this opportunity introduced me to the seemingly endless possibilities afforded by machine learning. Lastly I would like to thank Dr. Alan Chen and Dr. David Chettle for sitting on my committee and providing me with feedback throughout this process.

I would also like to express thanks to the McMaster medical physics community at large. It will take more than a pandemic to make me forget the camaraderie shared by us all. Camaraderie such as the countless late night studying sessions trying to complete quantum mechanics assignments or finishing off a lab report the night before it's due. I'm also grateful for the stress releases that we all got to enjoy together like going to the gym with Matt Bernacci and Matt Wong, or taking a few swings at softball with Dr. David Chettle and all the other students that joined the team. Having a support network that both helped me strive for the best in my education, and maintain my sanity throughout, is what enabled my success.

Lastly I would like to thank Zohra Tarar for sticking it through with me during

the arduous and seemingly endless process of writing! Always having someone to turn to during tough times is the essence of companionship.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Richard Garnett, declare that this thesis titled, "The Application of Machine Learning to Event Classification in Radiation Detectors" and the work presented in it are my own.

Chapter 1

Introduction

1.1 Motivation

In radiation physics detecting only a subset of a radiation field can be of key importance. Most often this task is complicated by the interaction of the field of interest with matter present around it, such as the generation of Bremsstrahlung radiation by beta emitting isotopes. This is an issue because simply having a radiation detector immersed in the radiation field of interest will impact the field and create secondary radiation that may or may not be what is the primary signal of interest. This could be, as previously mentioned, Bremsstrahlung radiation being generated in the detector while intending to measure beta decays. In some situations these issues can be mitigated with various techniques that allow you to differentiate between the signal and the background on a physical basis such as charge, which would be the case in measuring beta decays, with photons being the background signal. In a situation as such, measuring just the neutrally charged photon component or the positively or negatively charged beta emissions can be

accomplished with coincidence measurements. Two concentric radiation detectors can be run in coincidence and if a signal is detected in both detectors simultaneously, this is likely a charged particle. If, conversely, only one interaction is observed in either of the two detectors, this radiation interaction likely was caused by a neutrally charged particle. These inferences can be made because charged particles deposit energy in their surrounding media effectively continuously, given the extent of their electromagnetic fields, leading to a signal being generated in any sensitive volume of the radiation detector that is traversed. Neutrally charged particles on the other hand have negligible electromagnetic fields, leading to their interactions being collisional and hence probabilistic. The probabilistic nature of a neutral particle's interaction inside the sensitive volume leads to discontinuous energy deposition and a possibility of that radiation traversing the radiation detector without producing a signal. This is demonstrated in Fig. 1.1, where a concentric plastic scintillator is used as a coincidence detector with the main instrument, a tissue equivalent proportional counter being used for measuring charged and neutral particle doses in Low Earth Orbit (LEO) [1]. This setup provides a simple means of discriminating between charged and neutral particles, though the rate of discrimination between charged and neutral is not perfect. In one case the charged particle could impinge on only one of the two radiation detectors, leading to an anti-coincident signal which would be interpreted as a neutral particle. Another failure case would be a neutral particle interacting with both detectors producing a coincident signal which then would be interpreted as a charged particle. Clever design of the coincidence system can be employed to minimize these effects, but in practice they cannot be avoided completely.

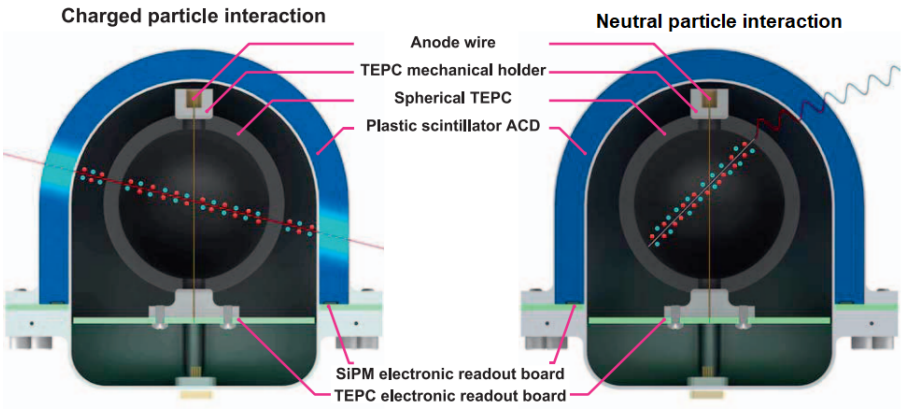


Figure 1.1. Depiction of the difference in how energy deposition occurs for neutral and charged radiation quanta [1]. In the case of a charged particle, seen on the left, energy deposition occurs along the entirety of the trajectory. For the neutral particle, seen on the right, energy deposition occurs once and generates a secondary particle which is then able to be detected in the sensitive volume.

In some situations physical solutions to separation of radiation fields are not feasible or even possible, one such instance would be the separation of neutrons from photons. The differentiation of neutron and photon radiation is impossible given the previous solution because both particles possess the same neutral charge, leading to similarly discontinuous energy deposition. The solution to these more complicated situations is to look at the result of an interaction, namely electron recoils for photon interactions and proton recoils for neutron interactions. In some radiation detection media, the generation of excited states inside the detector are dictated by the density of the generation of those states. This difference results in differing signal shapes for different radiation interactions and this can be exploited for Particle Identification (PId) with methods called Pulse Shape Discrimination (PSD) [2]. The situation of separating neutron signals from photon signals is of key importance whenever measuring neutron fields because of the

presence of photons from inelastic neutron scattering in the surrounding material as well as photons present from the neutron source itself [3, 4, 5, 6].

One other situation where a physical solution is not feasible is in high energy radiation fields where PId is desired. In radiation physics it is always of great importance to consider the distance a particle may travel on average within a medium before depositing all of its energy known as the range. If the particle generated by an interaction inside a detector has a range greater than the physical extent of the detector it will on average exit the sensitive volume of the detector before exhausting all of its energy. This would eliminate the possibility of using coincidence detection for PId in the case of neutral and charged radiation because the recoil particle from neutral interactions would behave in the same way as the charged radiation, traversing both detection volumes and producing a coincidence signal resulting in misidentifying the neutral particle as a charged particle. This could be the case in high energy particle accelerators like the Large Hadron Collider experiment, or in spaceborne instruments for measuring extraterrestrial radiation, the solution in the former is to increase the mass and density of the interaction medium or to increase the size of the detector to greater than the range of the radiation of interest. This solution of increases to the detector size may not be possible in spaceborne instruments because of their imposed limits on mass and energy available in space. The solutions in these limiting situations would be complex signal processing routines to achieve PId, such as PSD. It is this requirement for complex signal processing, as well as the abundance of simulation and the flexibility of radiation detection experiments, that lends the problem of PId so well to machine learning. A great deal of interest has re-emerged in recent years around

machine learning for its applications in complex image classification tasks such as computer vision, facial recognition, and autonomous piloting of vehicles. In the scope of classification of natural images machine learning should be more than capable of handling the relatively small solution space of PId.

1.2 Machine Learning and Neural Networks

The beginning of machine learning occurred in 1959 when Arthur Samuel of IBM coined the term for an algorithm used to play checkers [7]. This algorithm would evaluate the board state and attempt to maximize the likelihood of winning given its available legal moves. The learning aspect of this program came from the ability of the program to alter its evaluation when playing against itself. One version of the algorithm would be kept fixed while another was altered. If the altered version of the algorithm won more frequently, it would be allowed to remain and the adversary would then be altered. Further research building on this and the concept of how neural pathways in the brain behave lead to the advent of the perceptron [8, 9, 10]. This was the first machine learning algorithm built up to model neural behaviour in an attempt to learn, giving birth to the first neural network. The perceptron was used for computer vision and achieved little success, leading to a decrease in research within this area for decades. Somewhat more recently this area was reinspired by results from the Imagenet Large Scale Visual Recognition Challenge (ILSVRC) [11] where computer vision algorithms were put to the task of classifying natural images, such as photographs, and localizing objects within them. These fields have since grown so profoundly that implementations of neural networks can be found in self driving cars, noise cancellation, natural language

processing, and advertising.

Machine Learning is classified as either supervised or unsupervised learning. Supervised learning is when ground truth labels for data are used to train the machine learning algorithm, and unsupervised learning is when unlabelled data is used in training. A ground truth label is an identifier for the content of the data that is taken in the machine learning algorithm as the truth for which class each datum belongs to. Both methods of training a machine learning algorithm require large amounts of data for significant results to be obtained, but supervised training is so far the more prominent method. The introduction of ILSVRC as an open source labelled database of natural images was a significant impetus for the reemergence of machine learning and began the popularization of Convolutional Neural Networks (CNNs) for computer vision applications, with CNNs being the dominant architecture in the competition since 2012 being capable of superhuman performance in image recognition [11]. A CNN is a subset of machine learning wherein the neural network is composed of primarily convolution operations. The ILSVRC database and challenge contains roughly 1.2 million images belonging to 1,000 different categories, with performance in the challenge being measured on successfully classifying those images as belonging to their respective category.

Alongside the increased performance and popularity of CNNs afforded by the availability of open sourced ground-truth-labelled image databases was the increasing power of Graphics Processing Units (GPUs) and the parallelizability made possible with GPU architecture. The increased performance capabilities of GPUs over Central Processing Units (CPUs) enabled the increases in size of CNNs which is credited for the increases in performance for image classification. The 2012

winners of ILSVRC, AlexNet, were able to nearly double the performance of the previous state of the art in the ILSVRC when classifying images to 1,000 categories. AlexNet achieved this with a network that contained 60 million trained parameters [12], which can be compared to the 17,000 parameters of LeNet which was designed to read hand written digits from images [13]. This comparison, though not exactly transferable, shows that a larger CNN is more capable, being able to classify the significantly more information dense natural images to 1,000 categories compared to hand written digits which belong to only 10 categories.

The successes of machine learning, and more specifically CNNs, in the complex task of classifying natural images leads to perhaps an obvious implementation of CNNs to PId in radiation detection. The focus of the research herein discussed is to develop and further PId capabilities in two main projects: pair production identification in a spaceborne time projection chamber (TPC) and neutron and photon separation in mixed fields. Both projects preclude the use of more traditional identification techniques and thus require more intricate signal analysis to be able to perform the desired PId.

1.3 Pair Production Classification

The first of the two projects included in this thesis was the development of an event selection routine for use in background rejection onboard a proposed photon polarimeter called the Advanced Energetic Pair Telescope (AdEPT) [14]. This mission's primary radiation detector is a TPC which is designed to generate 2 perpendicular projection images of the radiation track, along the x-axis and the y-axis shown in Fig. 1.2.

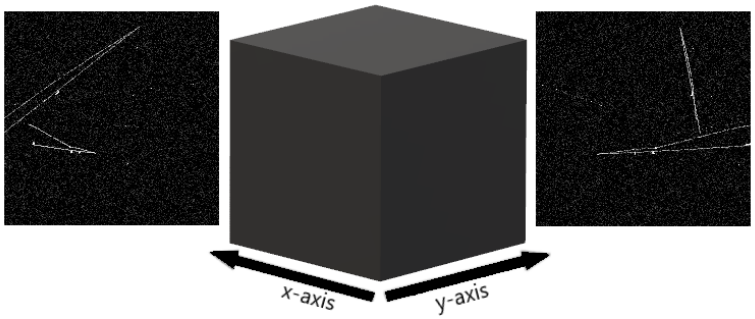


Figure 1.2. Diagram demonstrating the cubic sensitive volume of the simulations for AdEPT and the subsequent projection images output by the simulation.

The method for finding the polarization of incoming photons in this detector is to measure the angle of the plane connecting the electron and positron produced by the pair production interaction [15, 16], shown in Fig. 1.3 as the angle Φ. The TPC designed for use in this satellite is filled with Ar and CS_2 with a combined pressure of 1140 *torr* at 25 $^{\circ}C$. The energy range for the photons intended to be measured by this experiment are between 5 and 200 MeV, which would produce electron and positron pairs with energies just below half that, less the rest mass of the electron and positron pair. The range of these resultant particles in the TPC would be well over the length of the proposed sensitive volume, which precludes the use of a coincidence and anti-coincidence setup as discussed in Section 1.1, before considering the weight and size constraints imposed by spaceflight. Further requirements are imposed on the event selection requirements when considering the radiation environment of LEO, where the intended photon signal will be outnumbered 10,000:1 by cosmic rays, solar radiation, and trapped radiation in the Van Allen belts. These requirements setup the unique situation where a potential solution is the use of a CNN for real time event selection.

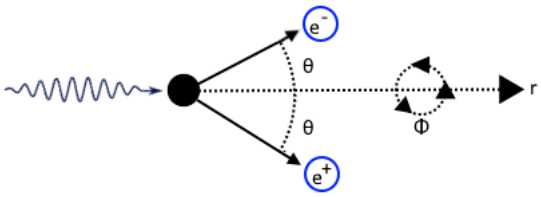


Figure 1.3. Graphic demonstrating the generation of a positron and electron pair through pair production. Angle Θ represents the angle of separation between the two pairs. Φ is the angle of rotation about the r-axis, which would dictate the polarization of the incident photon.

Machine learning applications for imaging detectors such as TPCs have seen success in several experiments [17, 18, 19, 20, 21, 22, 23, 24, 25]. At the time of the research for this project, the only known example was an implementation of Google's GoogLeNet [26] for use in classifying neutrino interactions for the NOvA experiment [27]. The application of GoogLeNet in this instance outperformed what was currently being used by the NOvA experiment for event classification without the need for event reconstruction. This success provided the inspiration for the CNN used in event selection for the AdEPT mission.

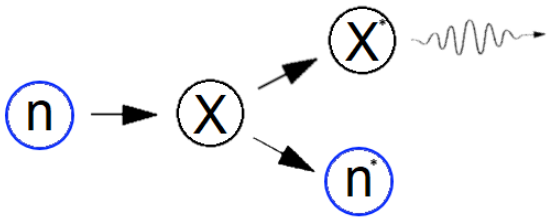
During the time of this research the satellite was still in the developmental stages and not yet functional so all results were found utilizing simulations developed in the GEANT4 Monte Carlo toolkit [28, 29]. The requirement to carry out simulations to develop the training and testing data lends itself very well to supervised learning for the development of a CNN based event selection algorithm. The use of the GEANT4 toolkit allows for control over every process that occurs within the simulation making it trivial to generate a training and testing data set with perfectly accurate ground truths assigned to each event. Given though that these are only simulations this work serves as a proof of concept that given

a sufficient training and testing data set a CNN could meet the requirements set out by the AdEPT mission. With that in mind, the simulations were designed with a simple geometry to represent the gas volume in AdEPT that would be sensitive to radiation, leaving out the other structures of the satellite that may still impact the radiation field present inside the TPC. Aside from this simplification, all relevant physics was included in the simulation and the radiation field was chosen to mimic what would be encountered in the LEO proposed for the AdEPT mission. There were two classes chosen for the CNN: positive or negative. The positive case would be any image containing a pair production interaction, and the negative case being any other interaction type possible with the incident radiation. Given that the CNN was to identify pair production or lack there of, the network was coined `GammaNet`. To generate the data for training and testing sets the simulation was run in two batches, the first batch with at least one pair production interaction being forced in the sensitive volume from incident photons, and the second set not having any interaction types forced from incident cosmic rays . A more thorough discussion of the AdEPT mission, the goals that were proposed for `GammaNet`, the simulation and the simulated radiation environment, and the results of this research are found in Chapter 2.

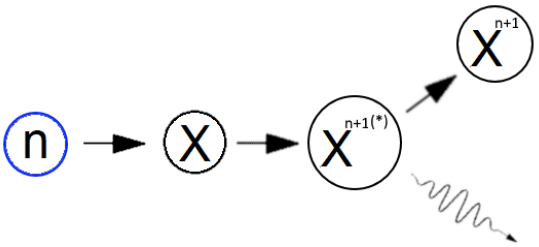
1.4 Neutron and Photon Separation

The second project was the development of a CNN based solution for PSD of neutrons and photons. This is a significant research question because the presence of a secondary photon field is inescapable whenever a neutron field is present. This

secondary photon field is generated primarily through two types of neutron interaction in matter, inelastic scattering, and neutron capture. In the case of inelastic neutron scattering a neutron will collide with an atom, excite it to an intermediate excited state, which then results in either a prompt or delayed emission of a photon as shown in Fig. 1.4a. The other interaction, neutron capture, is when a neutron collides with an atom and is absorbed by the atom creating a new nucleus. This new nucleus then can emit photons throughout its nuclear decay chain or through prompt or delayed emission again resulting from the nucleus being brought into an excited state which then decays to the ground state, as shown in Fig. 1.4b.



(a)



(b)

Figure 1.4. Diagram depicting the two neutron interactions of interest resulting in photon emission. Fig. 1.4a represents the inelastic scattering of a neutron off of a target nucleus resulting in prompt photon emission. Fig. 1.4b shows the absorption of a neutron resulting in a new isotope of the original target nucleus, which is left in a potentially excited state resulting in prompt photon emission. These are not exhaustive of the set of possible neutron interactions, which could include fission, proton emission, alpha emission, or other combinations of hadron emissions.

There are effectively innumerable pathways for a neutron field to create a subsequent photon field which results in intense mixed fields in the presence of a neutron source. Even if there were effectively no matter between the neutron detector and the neutron source, a secondary photon field would still be generated via interaction with the neutron detector itself. One solution may be to then have a neutron detector that is entirely insensitive to photons, but this too is impractical in practice as photons interact with all matter. Perhaps the most popular solution through insensitivity to photons is gaseous ^{3}He neutron detectors, which are often used as a gold standard for neutron and photon separation. A ^{3}He neutron detector is able to achieve an intrinsic neutron efficiency of 1.2E-1 % while having an intrinsic photon efficiency of 1.7E-7 % [30] for source and detector geometries as dictated in **??**, with intrinsic efficiency is defined by Eq. (1.1). The interaction mechanism for neutron detection in a ^{3}He neutron detector is the thermal neutron (n,p) reaction, which has a cross section roughly 1E3 times greater than the cross section for photon interactions. Additionally, the reaction products from a ^{3}He neutron capture are ^{1}H and ^{3}H with 0.764 MeV of energy. This results in a characteristic energy signal from each neutron interaction of 0.764 MeV, which can be reduced by wall effects where some energy is deposited in insensitive detector media. This discrepancy in sensitivities, combined with energy gating of signals, is what leads to the 1E6 factor of difference in intrinsic efficiency for neutrons and photons.

$$\epsilon_{int} = \frac{C}{\phi A} \tag{1.1}$$

Where C is the count rate in $\frac{\#}{s}$ found by the radiation detection system, ϕ is the fluence rate of the radiation of interest, $\frac{particle}{s \cdot cm^2}$, that are impinging on the radiation detection system's sensitive volume, and A is the area of the detector that is being impinged upon by the radiation in cm^2.

The relative inaccessibility of ^{3}He and its price have contributed to the increased popularity of PSD in recent years as a means of neutron and photon separation. PSD doesn't depend on the physical insensitivity of the detector to neutrons, but instead uses the differentiating pulse shape features to identify the neutron and photon components of a mixed radiation field. In this research the specific PSD capable detectors utilized are EJ-301 and EJ-309 which possess very similar characteristics, with EJ-309 having slightly poorer performance in PSD [31]. EJ-301's use as a neutron sensitive detector initially came about as Ne-213, both of which are a Xylene based liquid scintillator. It was found that some organic scintillators, namely Xylene, displayed delayed scintillation when being interacted with by highly ionizing radiation [32, 33]. This difference in scintillation time was exploited to enable neutron and photon separation with early work finding it possible to achieve a 9.5 % intrinsic efficiency for 2 MeV neutrons while minimizing photon sensitivity to 0.007 % intrinsic efficiency [32]. This capability made Xylene based liquid scintillators very popular for use in experiments where neutron and photon separation were desired.

The PSD enabling characteristic in Xylene based liquid scintillators is the generation of excited singlet and triplet states, with the singlet states decaying with a characteristic time of 3 ns and the triplet states decaying with characteristic times of 32 and 270 ns. The generation of the triplet state occurs by the transfer

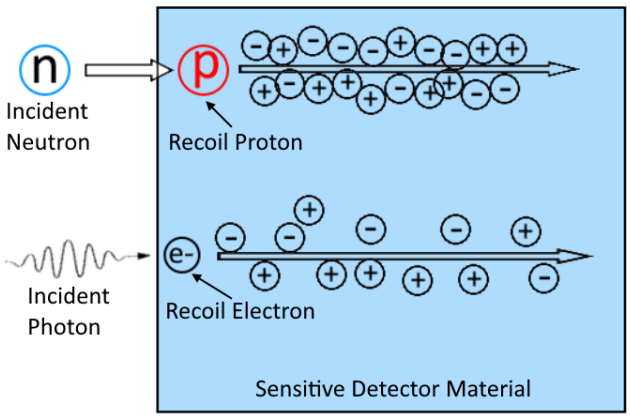


Figure 1.5. Figure demonstrating the differences in LET based on the type of ionizing particle inside the scintillation medium.

of energy from singlet states, which has an increased occurrence with increases in density of singlet states. This dependence on density of singlet states for triplet state formation expresses itself as an increase in proportion of triplet states when the scintillator is interacted with by particles with higher Linear Energy Transfer (LET). The LET is the amount of energy deposited per unit distance in the scintillator, demonstrated in Fig. 1.5.

When looking to neutron and photon discrimination in EJ-301 and EJ-309, the primary differences are in the recoil particles with protons for the former and electrons for the latter. The primary interactions in these scintillators are Compton scatter for the incident photons and elastic scattering off of the hydrogen nuclei of the scintillator for incident neutrons. In the case of Compton scatter, the recoil electron can have energy ranging from 0 to the maximum allowed by a 180° Compton scatter. Elastic neutron scattering from a hydrogen nucleus results in a recoil proton with an energy ranging from 0 up to the initial neutron energy,

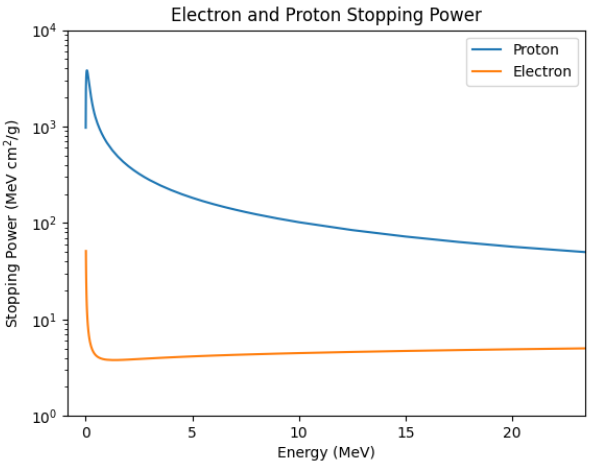


Figure 1.6. Plot of the stopping power, which is a measure of LET, for electrons and protons over a range of energies using data provided by the National Institute of Standards and Technology [34].

similarly to Fig. 1.4a without the excitation of the proton. These both result in a continuum of energies detected in the scintillator which, for non-relativistic energies of protons, results in a much higher LET for recoil protons as shown in Fig. 1.6. This causes the signals produced by neutrons in EJ-301 and EJ-309 to be characteristically longer than those of photon interactions.

There exist many techniques to leverage these differences in time characteristics to perform PSD with the most common being the tail-to-total method [35]. This PSD method is the main point of comparison for traditional PSD methods in this work, with various machine learning based solutions also utilized for comparison. Several other attempts at using machine learning for PSD in neutron and photon separation have been made, with a lot of success [36, 37, 38, 39, 40, 41], but no applications of CNNs had been found at the time of writing. This lack of application of CNNs to pulse shape discrimination likely is due to the signal for

PSD being 1-D and the main use case for CNNs is computer vision, a 2-D task. The adaptation of 2-D CNNs to 1-D tasks is trivial in most machine learning toolkits with each function being adjustable to 1, 2, or 3-D inputs, this work was carried out in `Caffe` [42] and `TensorFlow` [43] where such functionality exists.

To enable CNN based PSD for these liquid scintillators it is necessary to produce training and testing sets with photon and neutron classification labels defined. In this work there exists three possible solutions to this problem, namely the use of already existing PSD methods, experimental design to exploit other physical characteristics of photons and neutrons such as velocity, and lastly would be simulation. Each of these potential solutions are further explored in the chapter covering the development of the CNN solution Chapter 3.

1.5 Outline

This thesis contains 6 chapters, the first of which is this introduction describing the motivations and the basics of the principals relevant to this research. Subsequent to the introduction is Chapter 2 which gives a much more detailed look at the research that was done in developing, testing, and interpreting the results that were achieved with the implementation of the adaptation of `GoogLeNet`. This was done with simulated data, which is also further discussed in Chapter 2. Following this Chapter 3 outlines the development and application of a similarly adapted version of `GoogLeNet` to experimental data measuring neutron and photon sources. The goal of this work was to quantify the PSD performance achieved with the adapted CNN, to outperform classical tail-to-total PSD, compare performance between various machine learning approaches, and identify the importance of sampling rate

and bit depth. This work is then followed up by Chapter 4 which investigates different techniques for generating the labels used in training and testing the CNN for PSD. Further to this Chapter 4 also investigates the failure cases of the CNN, the application or generalization of the CNN to neutron sources not used in training, and further investigates the performance. Lastly Chapter 5 contains a brief discussion and conclusion of all results as well as areas for future improvements.

Chapter 2

Simulation Based Application: GammaNet

2.1 Preface

The core content of this chapter has been published in Nuclear Instruments and Methods Section A. All simulations, Machine Learning (ML) computations, and data analysis were performed by the author under the supervision of coauthors. The manuscript was prepared by me and revised by coauthors. In this thesis, Sections 2.2 to 2.5 were altered. Section 2.2 was greatly expanded upon to provide more thorough description of the limitations on event classification imposed by the physical requirements of Advanced Energetic Pair Telescope (AdEPT). Sections 2.3 to 2.5 were all changed to improve readability, to improve context for subsequent chapters, and to add in additional description perhaps not suitable for journal publication.

2.2 AdEPT Mission and Objectives

The measurement of astrophysical photons is greatly important to the understanding of our universe because they represent a tool to investigate the most extreme non-thermal processes in the Universe. As it stands currently the greatest amount of attention has been paid to astrophysical photons ranging from roughly 20 MeV to 300 GeV with the largest satellites being the AGILE [44] and Fermi [45] space telescopes. The AGILE mission's high energy sensitivity extends from 30 MeV to 50 GeV and Fermi's energy sensitivity is from 20 MeV to 300 GeV. Imaging for AGILE and Fermi are both enabled by the 3-D track reconstruction of pair production interactions in the sensitive volume of the satellite's detectors. The primary element of these detectors are silicon-strip detectorss (SSDs) [44, 45] which provide 2-D readout of charge generated in the detector and with the position in the 3rd dimension being generated by stacking of multiple layers of SSDs. In addition to these SSDs is the use of a thin high Z element used to increase pair production probability in the detector volume, such as Tungsten.

Neither of these missions were designed for measurement of polarization of photons or for measurements within the medium energy range, where many astrophysical objects exhibit unique behaviour. For example, of interest would be the presence or absence of polarization in one of the sources of these high energy photons, inverse Compton scattering. When there are intense sources of photons and relativistic particles, inverse Compton scattering can occur where energy is transferred from the relativistic particle to the photon, as shown in Fig. 2.1. The presence or absence of polarization in these photons can provide information about the processes that generated the relativistic particles [46, 47]. An additional source

of non-polarized high energy photons would be inverse Compton scattering from the decay of dark matter candidates [48].

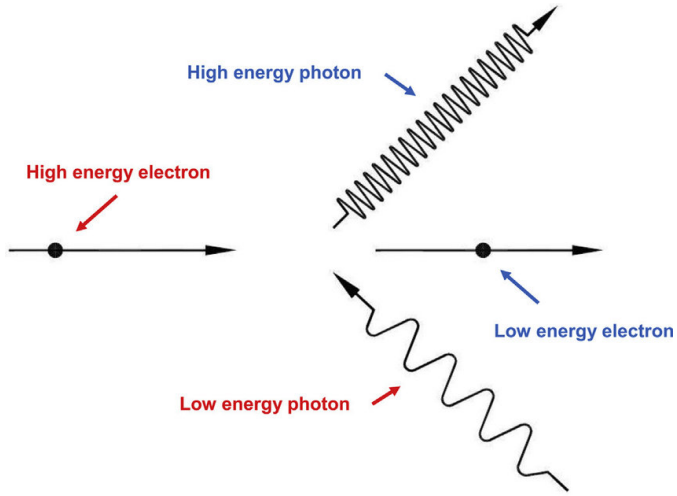


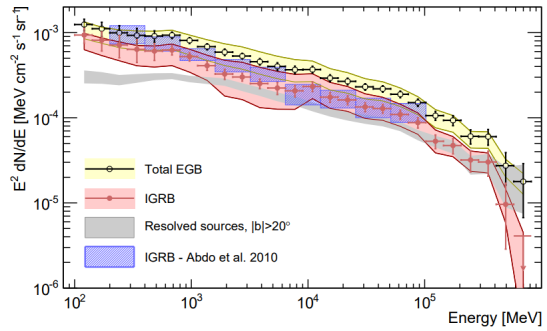
Figure 2.1. Diagram depicting the process of inverse Compton scattering wherein a relativistic electron imparts energy on an incident photon increasing the energy of the photon drastically. Image taken from [49].

The difficulty behind measuring medium energy photons is the significant contribution to photon interactions from both pair production and Compton scattering, each interaction possessing largely different signatures. These competing processes would require greatly different instrumentation and algorithms to observe the polarization of the incident photon [14, 50, 51]. Optimization of an instrument to be able to effectively measure both pair production and Compton scatter on-board a satellite is prohibitive at this time. The challenge is further exacerbated by the Galactic Cosmic Ray (GCR) background [52], which is an extragalactic source of charged atomic nuclei at extremely high kinetic energy. The GCR background cannot be effectively shielded for on satellites given their high kinetic energy, which can extend to several TeV per nucleon. Even at a modest energy of 10 GeV the Continuous Slowing Down Approximation (CSDA) range

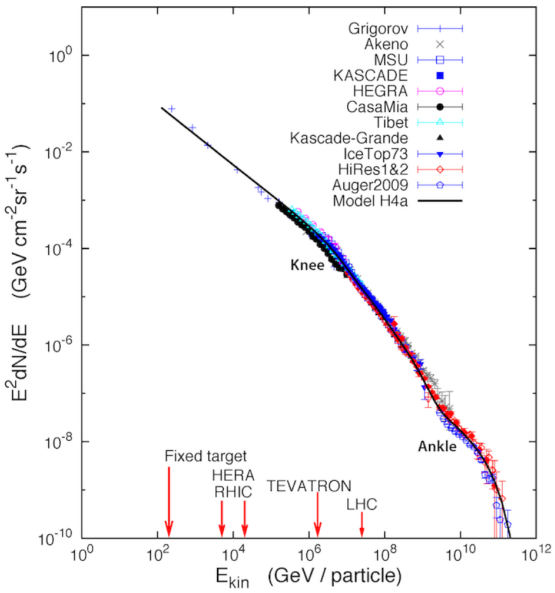
would be 7 m in lead. In addition, the fluence of GCR particles exceeds the astrophysical gamma-ray flux by approximately four orders of magnitude, as can be seen in Fig. 2.2.

Next-generation telescopes are being developed with the goal of characterizing the complete signature of extraterrestrial photons including their direction, energy, arrival time, and polarization. The most promising space missions (AdEPT [14], HARPO [55], and SMILE-I/II [56, 57]) proposed to explore the photon sky in the medium energy range are based on low-density gaseous time projection chamber (TPC) technologies that enable precise, three-dimensional tracking of particle interactions.

The AdEPT mission is one such medium energy gamma-ray polarimeter, which is being designed to measure photons with energies from 5 MeV to 200 MeV. The science data for AdEPT will consist of pair production interactions, with a background composed primarily of GCR and Compton scatter interactions. Compton scatter, though a photon interaction of interest for characterizing the medium energy gamma-ray spectra, is considered background for the AdEPT mission. Compton scatter is considered background because the AdEPT instrument is not designed to measure polarization for this interaction, so these interactions would form incomplete data. The AdEPT TPCs takes advantage of the Micro-Well Detector (MWD) technology augmented with the negative ion drift technique [58] to construct an instrument with the largest volume that can be accommodated in the rocket fairings currently available to MIDEX missions, 8 m^3. The active gas volume of the TPC is bounded on the top and bottom faces by an array of MWDs defining the 400 μm X- and Y-coordinate spatial resolution of the TPC [14]. The



(a)



(b)

Figure 2.2. a) shows the fluence rate of extraterrestrial photons [53] which can be compared to the fluence rate of GCRs in b) [54]. The relative scales are off by a factor of 1000 given that b) is normalized to GeV instead of MeV in a).

uniform electric field in the active volume provides a constant ionization charge drift velocity. Measurement of a relative arrival time of the signals on the detector strips provides the third, Z-coordinate. The use of the negative ion drift technique in the AdEPT TPC design [14] effectively reduces electron drift diffusion in the gas, making possible drift distances up to 1 m. With the applied electric field, ionization charge can traverse the Z dimension of the detector within a maximum of 50 ms. A schematic of the TPC designed for use in AdEPT is shown in Fig. 2.3.

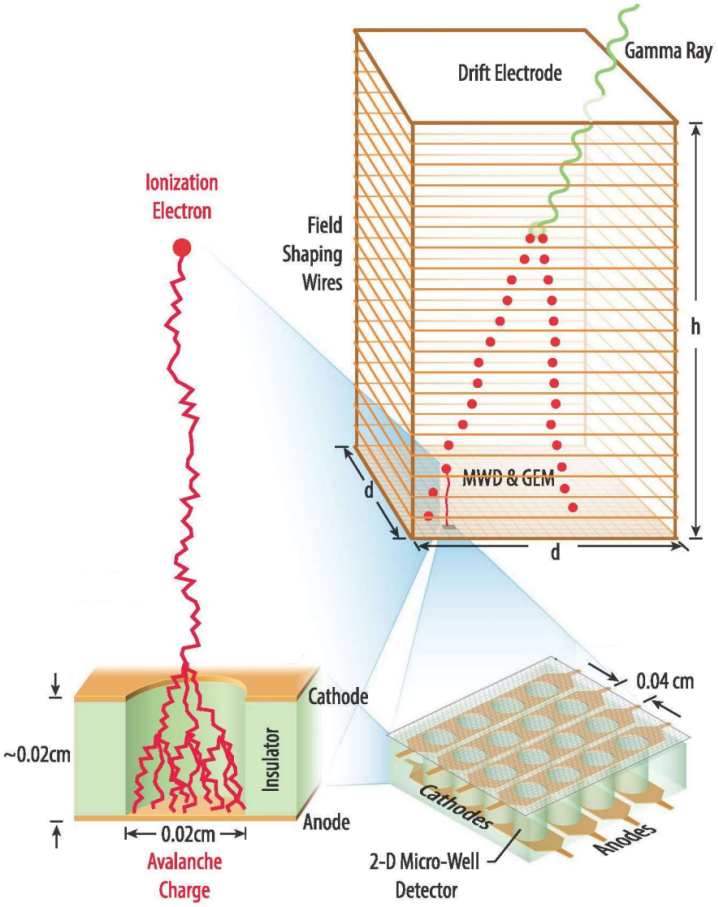


Figure 2.3. Diagram showing the main design features of the TPC to be used in AdEPT, the concept of how the interaction volume is reconstructed for a pair production event, and the MWD readout plane.

The use of the negative ion drift technique precludes the use of an anti-coincidence system, as used in HARPO [55]. The reason that this is the case is because a single event occurring any significant distance away from the MWD readout plane will take an appreciable amount of time to drift to the readout plane. This results in a coincidence window for any interaction of up to 50 ms, which with the rate of background GCRs expected in Low Earth Orbit (LEO) would result in effectively all events being seen as charged particles. In addition to this, with the energy range of interest and the density of gas filling the TPC, the positron and electron pair produced in pair production would have a range greater than the length of the TPC. This would result in pair production events activating the anti-coincidence detector for every interaction. The inability to use anti-coincidence to eliminate the background events results in a large and practically unavoidable data rate. This requires an alternative on-board processing approach for discrimination of GCR tracks and photon interactions.

The 8 m^3 version of AdEPT is estimated to produce an uncompressed data rate of $\sim$16 $Gbps$, which is much greater than what is capable for current satellite communication. Currently the Fermi Large Area Telescope mission [45, 59, 60] achieves an average science data downlink of 1.5 $Mbps$, while planned communications methods aim to achieve an average 50 $Mbps$ downlink [61]. The range of possible average downlinks leaves two to four orders of magnitude difference between the raw data rate and communications data rate for the AdEPT mission. The proposed solution is to use a Convolutional Neural Network (CNN) running on-board the spacecraft in real time to discriminate gamma-ray interactions of interest from the abundant GCR background.

The design and science goals of the AdEPT mission setup an easily stated objective for this event classification algorithm: obtain the highest signal sensitivity while achieving a background classification of 99.69-99.99 %. This classification should also be able to be conducted within 50 ms onboard a satellite, though the hardware to be available onboard is not yet known so a comparison to commercially available flight hardware is necessary. One such available flight computer is Innoflight's Compact Flight Computer 500, which is radiation hardened up to 30 krad and is space rated. Additionally, National Aeronautics and Space Administration (NASA) is investigating the suitability of System on a Chip (SoC) solutions available from NVIDIA [62]. This background rejection rate and classification inference time would result in a science data rate that would be manageable by satellite downlink.

2.3 Monte Carlo Simulations of AdEPT

The AdEPT mission, at the time of this work, was in the developmental stages and as such the development and testing of the CNN event classifier was conducted with simulation data. The simulation of the response and readout of the AdEPT instrument was carried out using the `GEANT4` Monte Carlo toolkit [28, 29]. Simulations in `GEANT4` require four primary components, the first is a description of the geometries involved in the simulation such as any components that will measure the radiation known as the sensitive volume. Then it is necessary to describe what physics will be involved in the simulation, which tells the simulation what radiation interactions to account for such as pair production. Next is the description of what radiation sources will be used in the simulation such as gamma rays from an

isotopic source. Lastly is the description of what information will be reported by the simulation, such as secondary ionizations inside the sensitive volume caused by a specific particle type.

2.3.1 Detector Geometry

Due to the availability of only simulated data at the time of this research, this work serves as a proof of concept that the proposed CNN could perform adequately for event selection. Given this is a proof of concept the simulated geometry of the AdEPT satellite only consisted of the appropriate gas mixture making up the sensitive volume representative of the TPC. This volume contained 1.5 atmospheres of Ar and CS_2 at a temperature of 293 degrees K with a sub-scale size of 25x25x25 cm^3, and full-scale size of 2x2x2 m^3. The simulations were carried out with the two varying sensitive volume sizes in order to identify an appropriate level of downsampling for the output of `GammaNet`. The full-scale volume was used to determine the performance of `GammaNet`, given that this would be more representative of the output from AdEPT. In this work, downsampling is defined as the process of taking an N x N region of the output image, averaging it, and applying it to a single pixel in the output, as shown in Fig. 2.4. This scales down the image by a factor of N^2, which is necessary for this work because the time to train and run classification for any CNN is strongly correlated to the image size passed to it. The full size AdEPT TPC will produce images of 5000 x 5000 pixels, which would be prohibitively slow in terms of both training and time to classification during operation.

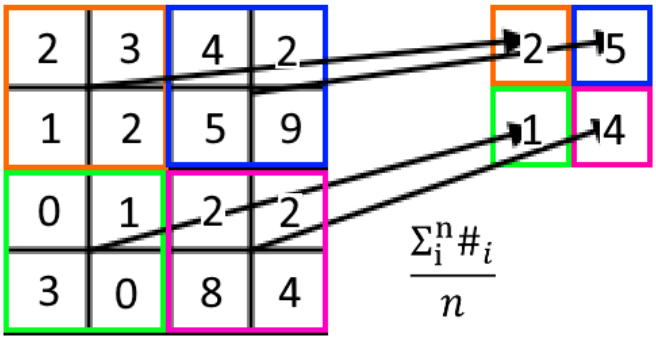


Figure 2.4. Figure depicting the method for downsampling the output from the simulation. In the event of non-integer dimensionality for the output, 0 valued padding was added to enable integer dimensionality.

2.3.2 Physics

The physics included in the simulation account for the different types of interactions between source particles and the Ar gas. These include hadronic physics for the interaction of GCR protons, electromagnetic physics for the interaction of gamma rays, electrons, protons, and other secondary particles, and the photo-absorption ionization model to accurately model the primary ionization and energy loss of relativistic charged particles in low density media. When the data set for the signal positive case was being run, any incident photon that impinged on the detector's sensitive volume was forced to undergo an interaction along it's initial trajectory with the distance selected randomly. `G4AdEPTSim` produces the ideal response of the AdEPT TPC, reporting the number of ionization electrons, their X-, Y-, and Z-coordinates, and the energy deposited in the active volume by a single incident particle. This situation only provides the error in the signals that are the result of variations in ionization, which can be characterized by Poisson statistics. This level of simulation wouldn't account for the noise introduced by the readout electronics or losses of signal from recombination of ions. The latter

of these sources of error were omitted but the electronic noise was added into the simulation results. Lastly would be the error present in the models of high energy interactions used to simulate such interactions in `Geant4`, which are built off of sometimes incomplete or unavailable data for high energy interactions. This is an unavoidable source of error for the background signals simulated herein, but will be alleviated in the future when this work is adapted for experimental data collected at high-energy particle accelerators.

2.3.3 Simulated Radiation Environment

To properly model the source radiation that AdEPT will encounter in flight, it is important to know what elevation and inclination it will fly at. AdEPT is proposed for launch into LEO with a 550 km altitude and a 28 degree inclination. The background environment in such an orbit is well-known and consists predominantly of GCRs, cosmic diffuse radiation, atmospheric gamma-ray emissions, reactions induced by albedo neutrons, and background produced by satellite materials activated by fast protons, alpha particles, and albedo neutrons [63, 64, 65, 66, 67]. In the 0.1 to 200 MeV energy range, the instrument background is dominated by charged particles in the Van Allen belt impinging on the spacecraft, cosmic diffuse radiation, and atmospheric gamma-ray emissions. The signal for the AdEPT TPC will be the medium energy photons, which exists either as an isotropic diffuse photon background or from astrophysical objects such as active galactic nuclei.

Given that this work is is a proof of concept, some simplifications to the background and signal sources considered were made. The background events for the

simulation consisted of only GCR protons with the energy spectrum from the Space Environment Information System for the expected AdEPT orbital conditions. GCR protons were selected as the background because they comprise the majority of the GCR fluence, and the energy range present will sufficiently incorporate the relevant background signals. Further to this, the aim of this work is to develop and characterize an event selection routine and not to determine characteristics of performance for event reconstruction or imaging, so the signal portion of the simulation was also simplified. The instrument aims to detect 5–200 MeV photons through pair production, so signal simulations were evenly sampled between these 5 and 250 MeV with steps of 25 MeV, starting at 5 MeV and resuming at 25 MeV, with an addition of 225 and 250 MeV being sampled for additional data.

Outside of simulating the energy distribution of the radiation environment for signal and background is the angular and spatial distribution of the incident radiation. In this work the background is more significantly comprised of GCR protons, so an isotropic distribution of events was chosen. The signal was also simulated with an isotropic distribution, in spite of there being structure to extraterrestrial photons. The signal was left isotropic to avoid imparting directional bias in the event selection routine during training, where bias in a training data set can greatly impact ML results. To obtain an isotropic source geometry `G4AdEPTSim` models the simulated events using a spherical volume source with a radius of 22 cm for the sub-scale version, and 1.73 m for the full-scale version. This spherical source is concentric with the active volume. The origin of the simulated particles are evenly distributed on the surface of this sphere. Each particle is initiated with

a momentum vector facing inward of the sphere with the azimuthal angle evenly sampled between 0 and 2π, and the polar angle is sampled according to the cosine law. This results in a uniform distribution of the radiation within the simulation volume volume, as shown in Fig. 2.5

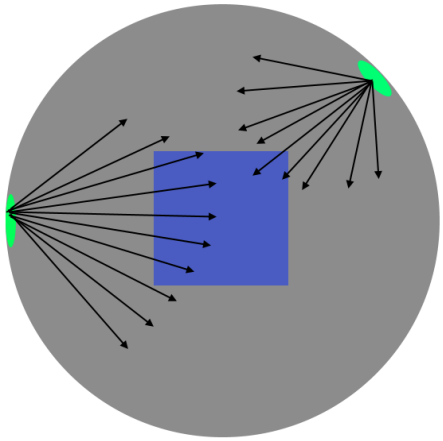


Figure 2.5. Diagram depicting the source geometry with the sensitive volume concentric with the spherical surface source. The source location is randomly chosen on the surface of the sphere, azimuthal angle is chosen evenly between 0 and 2 π, which is omitted in the drawing, and then the polar angle is sampled from the cosine law.

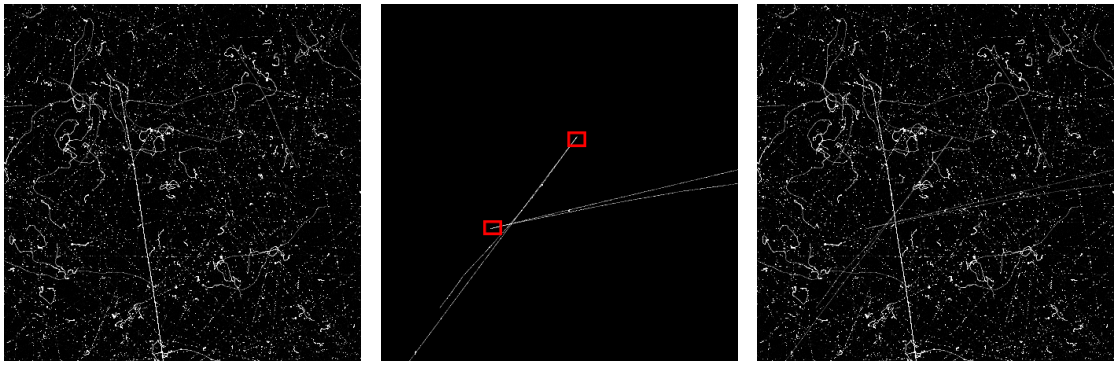
2.3.4 Event Recording

When running a simulation, there is one output per run of the simulation. In the case of `G4AdEPTSim` a run represents one of the 50 ms collection windows for AdEPT where any radiation interacting within the TPC within that collection window will be read out. Since the simulation emulates a 50 ms collection window, each run of `G4AdEPTSim` contains 375 incident GCR protons for signal and background, with an additional two incident photons for signal. The signal and background data sets both contain the 375 GCR protons because even when the signal event of pair production occurs within the TPC the background events will

still be present with the same detection rate. The sub-scale simulation runs consisted of five incident GCR protons or two incident gamma rays to account for the reduced surface area relative to the full-scale instrument. The number of incident particles were chosen in each case to fit the expected number of primary tracks within the 50 ms collection window given the AdEPT instrument parameters [14]. There are two incident photons for both simulations because the anticipated pair production rate in the full size simulation is less than one, although there is still the probability of two pair production events occurring within one collection window. For the photon events an interaction is forced if the path intersects the sensitive volume of the detector. This interaction type is forced as pair production for the signal events, and Compton scatter for the investigation of `GammaNet`'s sensitivity to Compton scatter. The source geometry allowed for the possibility of particles to miss the active volume, but results were only recorded if at least one particle interacted with the active volume. The source geometry used allows for a varying number of tracks to be recorded from each simulation run, although the number of simulated particles was constant between runs.

Per simulation run the number of ionization electrons in 400 x 400 x 400 μm^3 voxels was recorded, where a voxel is the discretized unit of volume used in 3-D simulations. This voxel size was chosen to correspond to the nominal resolution of the AdEPT instrument. The number of ionization electrons in each voxel is then projected onto the XZ and YZ planes to generate images. In order to more accurately simulate the images that would be produced by the AdEPT TPC electronic noise was added to the signal output for each set of images. The addition

of electronic noise was performed by adding a randomly generated number of electrons to each pixel of an image. The number of electrons added to each pixel was generated from a normal distribution with standard deviation of two and a mean of zero. In addition to electronic noise, background events were added to every photon image in the form of GCR protons. To do the background event addition, GCR proton images were generated with electronic noise and photon images without. Each photon image then had a unique GCR image added to it. Photon images were generated without the addition of electronic noise to ensure GCR images and the composite photon images would have a constant amount of electronic noise. Fig. 2.6 shows an example of the process used for generating the pair production data set, where an image containing two pair production events is added to a background GCR image with two tracks.



(a) **(b)** **(c)**

Figure 2.6. XZ projection of the sensitive volume of the AdEPT simulation. a) GCR background image containing several proton tracks with added electronic noise. b) gamma-ray image, containing two pair production events with the vertices outlined in red for illustrative purposes. c) Combination image that would be used for training and testing `GammaNet`. These simulation images have had their contrast adjusted for better viewing in this paper.

Correctly labeled image sets were generated from these simulations for both

training and testing of `GammaNet`. The training image sets contained 1.1x10^6 pair production images and 10^6 background GCR proton images. The testing image sets contained 1.5x10^3 pair production images, 1.5x10^3 Compton scatter images, and 10^6 background GCR images. The Compton scatter images were included in testing, but not training, as an additional source of background. `GammaNet` was found to be less sensitive to the Compton scatter images than pair production. The intuition when applying a CNN to this classification problem was that the CNN would be able to pick up on the discerning characteristic of pair production events compared to GCR proton tracks. These pair production signatures are further discussed in the results, Sections 2.5 and 2.6.

2.4 GammaNet

The application of ML to event classification in radiation detection is a natural progression of the field given that radiation detectors produce highly structured signals. These signals are often dependent on the nature of interacting radiation, and the type of interaction undergone. High energy physics projects, such as the Large Hadron Collider, are already implementing ML applications for event classification [68, 69]. There has also been implementations of machine vision for image classification in radiation imaging detectors using CNNs to classify neutrino interactions at Fermilab and the Ash River Laboratory [70]. `GammaNet` was inspired by the successes of this CNN application, with the CNN showing an increased performance compared to the state-of-the-art algorithms currently deployed for classification of neutrino events at Fermilab. Specifically, there was a relative increase of 40% sensitivity for electron neutrino signals, going from 35% to 49%

[70]. However, the AdEPT instrument does not require as much information about the background radiation as the NOvA detector, allowing `GammaNet` to classify to two classes as opposed to NOvA's 13. It is believed that this reduction in the number of classes helped achieve greater background rejection, whereas the inclusion of additional classes such as Compton scatter or hadroninc processes would complicate both the data set generation and the classification problem.

2.4.1 GammaNet Architecture

Two neural network architectures were used in this work, `GoogLeNet` and `VGG16`. Both architectures represented what was currently the state of the art performance for image classification in the Imagenet Large Scale Visual Recognition Challenge (ILSVRC) [71, 72] and so posed a good point of comparison. The `VGG16` architecture is a traditional end to end network architecture where each step is performed consecutively from input to output. The `GoogLeNet` architecture represents one of the first instances of non-linear operations, where `GoogLeNet` introduces a Network in Network (NiN) structure that was named the inception module. The idea of the inception module is that it performs multiple convolutions at the same depth in the layer simultaneously with each having a different kernel size, where the kernel is the NxN matrix with which the input is convolved. Fig. 2.7 demonstrates the implementation of an inception module. The NiN architecture was thought to be beneficial for event classification because of the ability of it to observe features that exist at different spatial extents simultaneously at different depths in the network. It was found that the `GoogLeNet` architecture outperformed `VGG16`, with performance comparisons between the two discusses in the results section of this chapter Section 2.5.

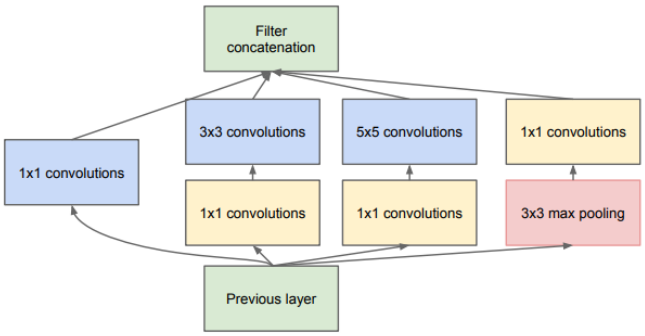
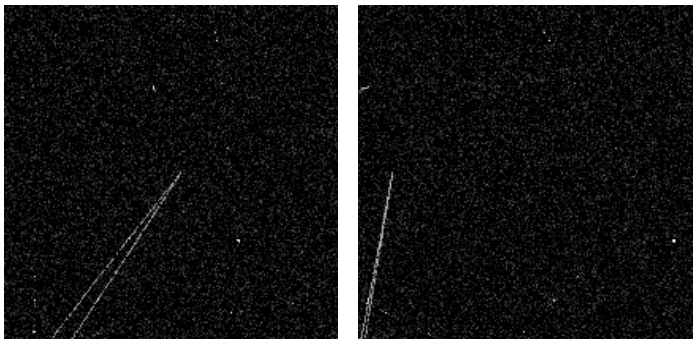


Figure 2.7. Diagram of the inception module from `GoogLeNet` [26] as was implemented in `GammaNet` for this work.

The typical input to a CNN is a NxMx3 image, with the last dimension being the three typical colour channels. The simulation data, as well as the eventual experimental data, are the two XZ and YZ projection images of an event in the AdEPT TPC. To perform classification on an event these two images are input into two instances of the same `GammaNet` instance. The output of the classification on these individual projection images is then compared as a Boolean or operation, with the classification of a pair production being present being a 1 and the absence of the pair production being a 0 i.e. signal or no signal. Using an or comparison of these two projection images is important for the performance of `GammaNet` because of the geometry of the signal in the detector that pair production creates. The key signature of a pair production interaction is the presence of the symmetrical positron and electron ionization track, as shown in Fig. 1.3. If the two tracks overlap on the axis over which the projection is taken, the resultant projection image will appear only as a single track, though the alternate projection may still show the characteristic pair production signal. An example of this situation is presented in Fig. 2.8, where in Fig. 2.8a there is clear separation of the positron and electron track, but in Fig. 2.8b the tracks overlap giving the appearance of a

singular ionization track.



(a) **(b)**

Figure 2.8. XZ and YZ projections of the same event generated in the sub-scale simulation, with a downsampling rate of 3. In a), the XZ projection, a well separated pair production track is shown in the lower half of the image. In b), the YZ projection, an overlapping pair production track is shown in the lower left of the image.

With the unique constraints for the AdEPT mission in mind, mainly the 50 ms timing constraint for inference and the power and mass budgets of space flight, modifications to `GoogLeNet` were required for use as `GammaNet`. These modifications were aimed at reducing the computational burden that is imposed by deep learning CNN architectures such as `GoogLeNet` where depth directly correlates to intensity. To reduce the computational intensity the number of inception modules used was reduced from the original 9. Three variations of the reduction were tested, namely 1, 2, and 3 inception modules with the same architecture as `GoogLeNet` for the remainder of the network [26]. The performance obtained by the three variants is presented in Table 2.1, with the classifications being determined by which ever classification output was the highest. Notably the last layer of a CNN produces a probability that the input belongs to each class, which together are normalized to unity. In the case of `GammaNet` this is just the presence or absence of pair production, which is considered signal or background respectively. From

Table 2.1 it can be seen that the greatest background rejection rate was observed for a network with only 1 inception module, and the greatest signal sensitivity was found for 2 inception modules.

These results are counterintuitive when considering the general trend over time of increased performance with increased depth in neural networks, where depth is the number of consecutive layers. This trend, when sufficient depth has been reached, is found to reverse and performance begins to greatly diminish [73]. This is observed for networks with depths surpassing 30, which is well outside the depth used here. These results are all applicable to natural images, the extent to which this applies to radiation detection images is unknown. It is expected that the observed performance in Table 2.1 can be explained by considering the complexity of the radiation images. Natural images contain important information in nearly every pixel with an additional 3 colour channels, this can be compared to the sparsity of information in the projection images of the AdEPT TPC in Fig. 2.8 which are also only gray scale. This very limited information density, and complexity, is believed to allow the CNN to learn all necessary features with a much more limited depth of network. This would enable the observed comparable performances with only one inception module.

Lastly, it was necessary to implement mixed precision in this work to achieve the classification accuracies presented. The simulation images were stored in the Portable Network Graphics (PNG) format with a bit depth of 8. The default data format for computation and storage of parameters in the ML library used, `NVCaffe`, is single precision or 32 bit. It was found that single precision wasn't sufficient to achieve the background classification desired, so the inner product and

Table 2.1. Tabulated results of `GammaNet` pair production sensitivity and background rejection rate for differing numbers of inception modules. Pair Production sensitivity reported as highest of the 5–250 MeV energy sets whereas background rejection rate was calculated from only one set.

Number of Inception Modules	Pair Production Sensitivity (%)	Background GCR Rejection Rate (%)
1	93.17	**96.30**
2	**94.28**	93.91
3	93.47	96.10

the softmax layers, shown in Fig. A1.1 m) and l), were stored and computed in double precision.

2.4.2 Training

In supervised learning training occurs by passing a number of inputs through the network, called the batch number, and having the network produce a classification for each input. After this the difference between the input ground truth and the classification is compared with a loss function. The gradient for the loss function is then calculated with respect to the parameter of each layer and these parameters are then updated based off of that gradient. This process is repeated a fixed number of times, typically with the aim of reaching a steady state in performance when classifying a set of testing data reserved for validation. The overall training is governed by a set of parameters that are independent of the CNN architecture called hyperparameters. These hyperparameters are used by `NVCaffe` during training to determine aspects such as the learning rate, the number of iterations, the batch size etc [42]. The hyperparameters used during training in this work can be found in Appendix A0.2. The outcome of training is shown in Fig. 2.9 and it shows that

training proceeded well for the instance of `GammaNet` with 1 inception module, with convergence on greater than 98 % accuracy on the testing data set. Training was carried out for 5 million iterations for each network described in Table 2.4, with the results for those not included in Fig. 2.9 being omitted because each plot was effectively the same.

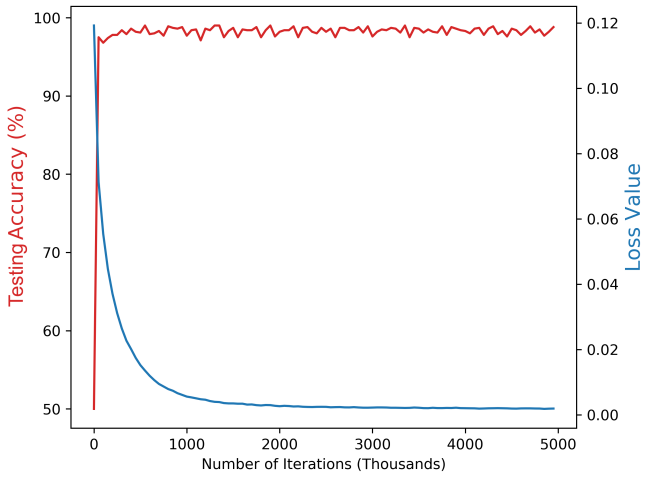


Figure 2.9. Graph of the training results for `GammaNet` with 1 inception module. This training data was generated with the sub-scale simulation, with a downsampling rate of 3x. The left axis contains the accuracy of `GammaNet` on the validation data set, and the right axis contains the loss value averaged over every 50k training iterations.

2.5 Results

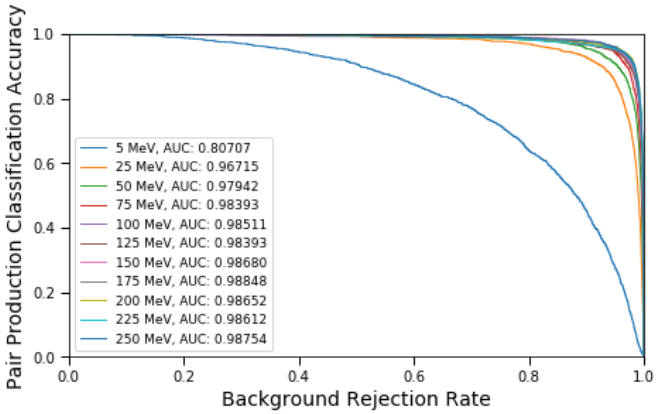
2.5.1 ROC Analysis

To quantify the performance of `GammaNet` as an event classifier Receiver Operating Characteristic (ROC) analysis was applied to the testing data set [74]. ROC analysis is performed by having the classifier provide an output for every event in a testing set, with each output containing the probability of the input belonging

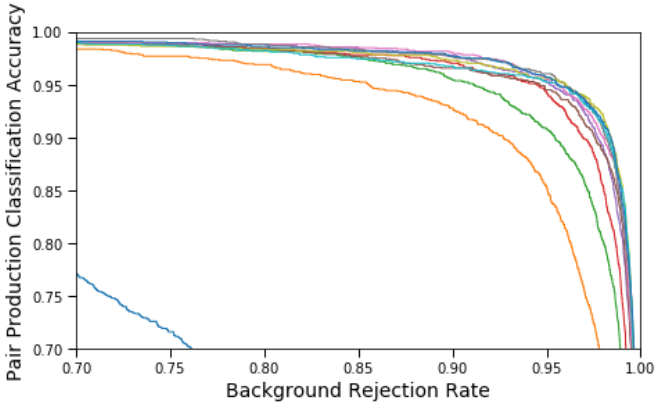
to the signal or background classes. These are then sorted in descending order of the output probability for the signal class, which in the case of this work is pair production present. This results in a sorted list of the probability that each input belongs to the signal positive class, which can be cross checked against the associated ground truth for the input. The threshold, which is the probability at which a positive or negative classification is made, is then iterated through the sorted list. At each iteration the threshold becomes the next lowest positive probability and that event is then considered positive. This allows for the number of true positives and a false positives to be identified for every threshold value. True positive is the case where an event with a ground truth of positive is classified as positive. False positive is the alternate situation where an event with a ground truth of negative is classified as positive. These values are then normalized to the size of the input data set giving the rates of false positive and true positive classification.

Plotting the True Positive Rate (TPR) and False Positive Rate (FPR) for each threshold value encountered gives a useful visual representation of the behaviour of the classifier which is shown in Fig. 2.10. The area under the ROC curve has a theoretical maximum of 1, and a minimum of 0, both cases would represent a perfect classifier with the latter situation occurring if labels are accidentally inverted. The worst case scenario would be randomized guessing for classification and with a balanced set should result in a straight line between 1 and 0, with an area under the curve of 0.5. The values obtained for each energy range are demonstrated in Fig. 2.10a. The ROC and its respective plot are important for identifying the threshold level to operate `GammaNet` at in order to achieve the

99.69–99.99 % background rejection rate desired for AdEPT.



(a)



(b)

Figure 2.10. Fig. 2.10a The ROC curves generated for each pair production data set as classified by `GammaNet`, using 11x downsampled images. The AUC is provided in the legend for each incident gamma-ray energy. Fig. 2.10b A subsection of Fig. 2.10a is presented to display the nuanced features of the plot.

2.5.2 Downsampling Investigation

With the use of the ROC analysis, the impact of downsampling the simulated images on classification performance can be investigated. The results of this are displayed in Table 2.2, with downsampling rates between 1–11 are investigated

Table 2.2. Pair production sensitivity of `GammaNet`, for sub-scale simulation images, given the desired background rejection rate with differing factors of downsampling. The data rate limits are sampled between the proposed minimum and maximum as described in Section 2.2. The background rejection rates listed are calculated by using the ratio of the raw data rate and the data rate limit, assuming the signal is approximately entirely background. Each data set was generated from the sub-scale simulation, using the given downsampling rate. `GammaNet` was then trained and tested on those data sets. The reported pair production signal sensitivities are the average sensitivity for the energies simulated. Error was calculated using binomial statistics with a 95% confidence interval.

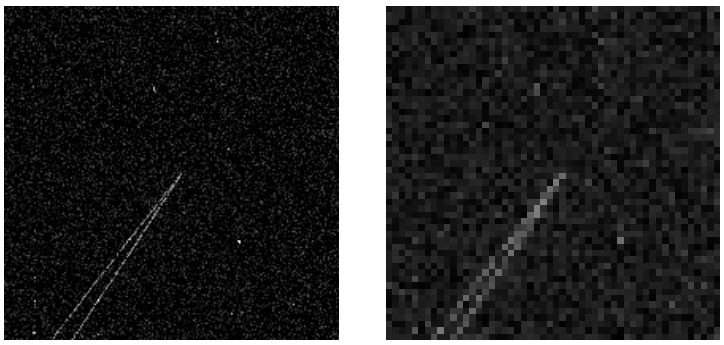
		Downsampling Rate					
		1	3	5	7	9	11
Data Rate Limit (Mbps avg.)	Background Rejection Rate (%)	Signal Sensitivity (%)					
1.5	99.99±0.002	43±2	65±2	50±2	41±2	40±2	28±2
5	99.97±0.003	47±3	73±2	65±2	54±2	49±2	45±2
10	99.94±0.005	54±3	78±2	73±2	62±2	56±2	56±2
20	99.87±0.007	64±2	84±2	81±2	74±2	67±2	66±2
30	99.81±0.009	68±2	87±2	84±2	78±2	72±2	71±2
40	99.75±0.01	71±2	89±2	86±2	81±2	75±2	74±2
50	99.69±0.01	74±2	90±1	87±2	83±2	78±2	77±2

using the sub-scale simulation. From these results it is shown that downsampling rates between 3 and 5 provide improved signal sensitivity at every background rejection rate, and with only marginal losses past this for the 99.99 % background rejection rate. When considering the 99.94 % or lower background rejection rates, every level of downsampling provides improved performance over the unaltered images. This result may seem counterintuitive given the downsampling rate effectively reduces the resolution of the images, but it is believed that the downsampling increases performance by reducing the sparsity of the images. One such consequence of this sparsity can occur at the original 400 μm resolution where

the ionization tracks from the electron or positron can have discontinuities. This can especially be a problem in the low energy range for pair production where the electron and positron are Minimum Ionizing Particles (MIPs). With a 5 MeV incident photon generating a 2 MeV electron and positron, the energy loss per unit distance using the CSDA is 3.9 $\frac{keV}{cm}$ which at the 400 µm resolution is 1.56E-1 keV. With a mean energy per ion pair in Argon gas of 26.4 eV [75] this gives on average 5 ion pairs generated per 400 µm distance, which could result in many discontinuities in the track as can be seen in Fig. 2.8a which was downsampled by a factor of 3 already.

While intermediate levels of downsampling provides higher performance for every background rejection rate, there are losses in pair production sensitivity for very high downsampling rates. This occurrence can be explained by the decreases in apparent separation angle caused by high degrees of downsampling shown in Fig. 2.11. In addition to this reduction in apparent separation of the two tracks, the averaging that occurs begins to reduce the contrast between the ionization track and the background noise.

These responses to downsampling are of benefit to `GammaNet` for both the increases in sensitivity for most situations and the faster inference times afforded by higher rates of downsampling. Given the large image size generated by the full-scale simulation, a downsampling rate of 11 was used for the remainder of the work when utilizing the full-scale simulation. Training of `GammaNet` on the full-scale simulation data at a downsampling rate of 11 took 30 days of compute time, proving investigating `GammaNet's` performance on lower downsampling rates with the full-scale simulation data to be prohibitively time consuming. This work was



(a) **(b)**

Figure 2.11. Demonstration of the decrease in apparent separation angle with increases in downsampling. Fig. 2.11a demonstrates the pair production at a 3 times downsampling rate, and Fig. 2.11b shows the same image at 11 times downsampling. Adjustments to brightness in the images are made for visibility.

completed with the use of a NVIDIA GTX 1080 Graphics Processing Unit (GPU) which possess roughly 8.8 TFLOPS of processing power.

Table 2.3. Pair Production sensitivity for `GammaNet` and `VGG16` at varying background rejection rates corresponding to anticipated downlink speeds. Performance comparison results were generated using the sub-scale simulation data, with a downsampling rate of 3.

Background Rejection Rate (%)	`GammaNet` Pair Production Sensitivity (%)	`VGG16` Pair Production Sensitivity (%)
99.990±0.002	65±2	28±2
99.97±0.003	73±2	38±2
99.94±0.005	78±2	46±2
99.87±0.007	84±2	57±2
99.81±0.009	87±2	61±2
99.75±0.01	89±2	64±2
99.69±0.01	90±1	66±2

2.5.3 Architecture Comparison

A cursory investigation between the performance of `GammaNet` relative to other neural network architectures was performed. In this investigation another neural network architecture was chosen, `VGG16` [71], given it outperformed `GoogLeNet` in the ILSVRC. `VGG16` was trained in the exact same manner as `GammaNet` and the performance of the two networks were compared. The performance comparison between `GammaNet` and `VGG16` was carried out with a downsampling rate of 3, and with data produced from the sub-scale simulation. Section 2.5.2 provides the results from each network when classifying the sub-scale simulation data, with `GammaNet` shown to largely outperform `VGG16` over the entire range of background rejection rates investigated. This result is not what would be anticipated given that `VGG16` was found to be the better performing neural network in the ILSVRC competition, but the task of event classification for AdEPT utilizes more sparse images. This increased image sparsity perhaps loans itself better to the NiN architecture employed by `GoogLeNet` allowing for improved performance on this task. These results show that `GammaNet` is more well suited for classifying background images than is `VGG16`, which ultimately is the primary task of `GammaNet` for AdEPT.

2.5.4 Compton Scatter Sensitivity

The performance of `GammaNet` when classifying Compton scatter events was of interest as well given that it is the main photon interaction contributing to background in the AdEPT instrument, and the similarity in track structure compared to pair production. The rate of misclassification for Compton scatter events as pair production events provides information about the features that `GammaNet` uses for

classifying the input. The main differentiation between the pair production and Compton scatter tracks is the presence of only a singular track for Compton scatter and the absence of the vertex from pair production. Support for the importance of these features for classification is shown in Fig. 2.12. As the incident gamma-ray energy increases, so too does the signal sensitivity for pair production. The increase in signal sensitivity is due to the increased energy of the positron-electron pair producing more linear tracks, closer in proximity, and with more distinct vertices. This is supported by the negligible increase in signal sensitivity for Compton scatter events.

The minor increase in sensitivity for Compton scatter is likely due to the increased incident photon energy producing higher energy recoil electrons on average, which would produce increased linearity in the ionization tracks. This produces a much lesser response than does pair production though because it represents only one of the characteristics of pair production signals. Additionally the energy of the recoil electron in Compton scatter doesn't take the entire incident photon's energy, instead having a continuum of possible energies unlike in pair production where all energy is transferred to the creation of the pair and their kinetic energy. These characteristics result in a much less significant increase in signal sensitivity for Compton scattering as the incident energy increases.

Due to the large raw data rate and the limits of satellite communications, it is required to achieve a background rejection rate of 99.99% to 99.69% in order for the data to be transmitted. To achieve this background rejection rate, the threshold for a pair production event classification has to be set quite high, 0.9999986, which results in a number of pair production events being misclassified as background

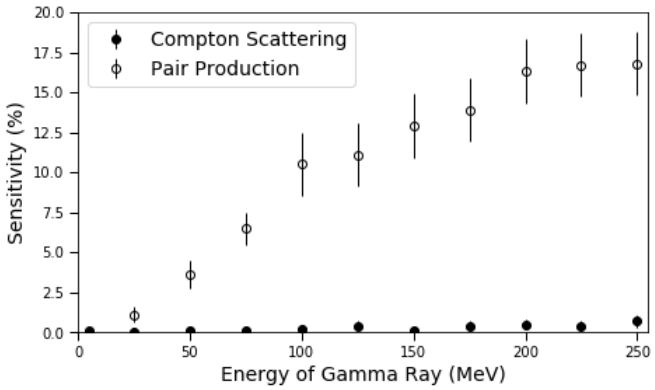


Figure 2.12. Plot of the sensitivity for Compton scatter and pair production image sets as the energy of the incident gamma ray varies, using a downsampling rate of 11 on the full-scale simulation. These sensitivities were calculated using a threshold value that generated a 99.990±0.002% background rejection rate. Errors were calculated using binomial statistics with a 95% confidence interval.

events. Table 2.4 shows the average rate at which `GammaNet` classifies pair production and Compton scatter events as a positive event, given different background rejection rates. These results were generated using the full-scale simulation with a downsampling rate of 11. The classification accuracies were averaged over the energies simulated for pair production and Compton scatter. It is shown in Table 2.4 and Fig. 2.12 that at the proposed 99.99% background rejection rate, we obtain a pair production sensitivity between 0.1±0.1% and 17±2%, depending on incident photon energy, with an average of 10±1%. For the best case scenario of 99.69% background rejection, the signal sensitivity increases to a range of 1.1±0.5% to 69±2%, again depending on incident photon energy, with an average of 52±2%. In both cases, the sensitivity to Compton scatter is quite small, which is beneficial for the mission due to Compton scatter representing background for the AdEPT mission. The relatively low sensitivity to pair production events at low energy will

reduce the effectiveness of the instrument, but this impact can be mitigated during mission design by implementing image compression, where these calculations were done assuming no compression.

Table 2.4. Pair production and Compton scatter sensitivity at varying background rejection rates corresponding to anticipated downlink speeds. The GCR proton background rejection rate was calculated for one set of background images. Each data point for Compton scatter and pair production sensitivity were generated by averaging the sensitivity over all simulated gamma-ray energies. All data here were generated using the full-scale simulation with a downsampling rate of 11.

Background Rejection Rate (%)	Pair Production Sensitivity (%)	Compton Scatter Sensitivity (%)
99.990±0.002	10±1	0.3±0.3
99.97±0.003	16±2	0.4±0.3
99.94±0.005	26±2	0.7±0.4
99.87±0.007	37±2	1.3±0.6
99.81±0.009	44±2	1.7±0.6
99.75±0.01	47±2	1.9±0.7
99.69±0.01	52±2	2.2±0.7

2.5.5 False Positive and False Negative Events

In this study, the test set of GCR protons contained 10^6 events, with twice as many images. Operating at 99.99% background rejection results in 100 GCR proton events being classified as a pair production event, considered false positives events. Given that `GammaNet` classifies both projections of an image and compares the results with a Boolean or, a false positive occurs when at least one of the two projections is classified as positive. Fig. 2.13 shows 4 of the GCR proton events that resulted in false positive classifications. The opposite situation where a pair production event is classified as background is a false negative event. Fig. 2.14 shows 4 pair production events that resulted in `GammaNet` producing the lowest response for pair production classification out of the testing set. The projection images shown in Fig. 2.13 are the projections that resulted in the false positive

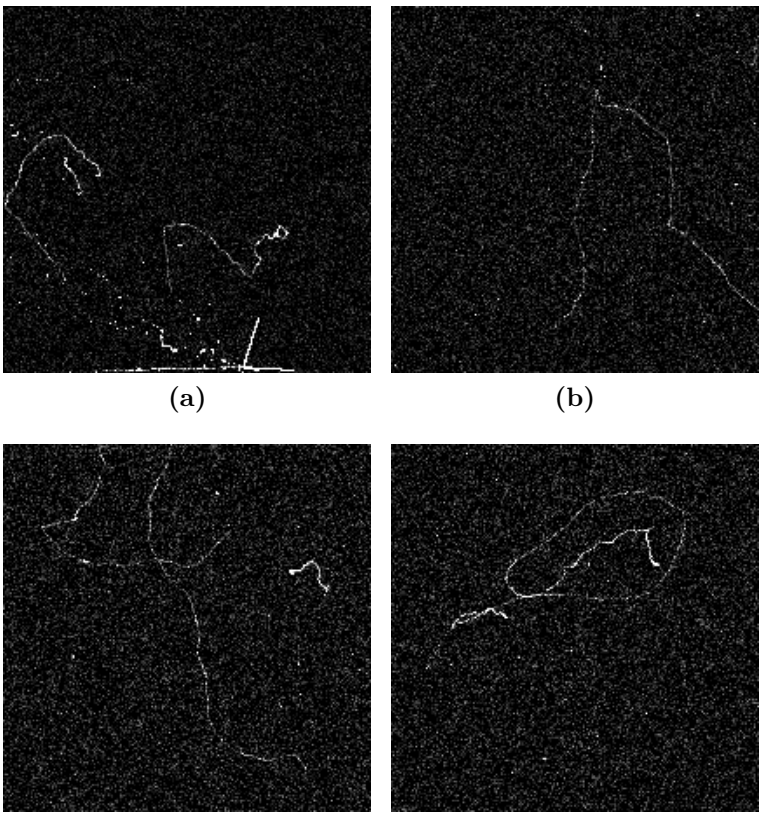


Figure 2.13. Projection images of simulated GCR proton events that resulted in false positive classifications. Only the projection image resulting in the false positive is shown, the alternate projection is not included because no event produced a false positive in both projections.

classification, with no single GCR proton event producing a false positive in both projection images. The projection images shown in Fig. 2.14 are the projections that were found to be more representative of the features that resulted in false negative classifications. In the false positive images, Figs. 2.13a to 2.13d, extended delta-ray tracks are observed with at least one point of track crossing. Where a delta ray is the occurrence of a secondary particle with enough energy to

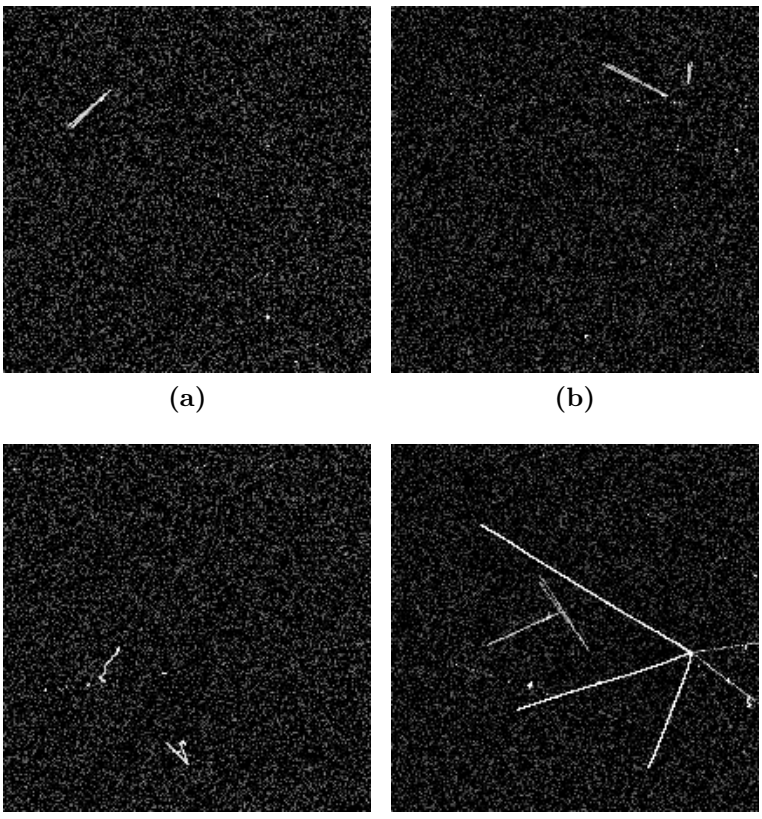


Figure 2.14. Projection images of pair production events that produced the lowest response in `GammaNet` for the pair production event class. The projection shown is most representative of the cause for false negative classification.

create a significant ionization trail which diverges from the original ionizing particle's vector. This observation demonstrates that `GammaNet` responds to extended contiguous tracks, and track crossings, as signals of pair production events. In addition, Fig. 2.13a contains a pair production event occurring from a GCR proton track which results in a false positive classification, showing `GammaNet` responds significantly to the vertex of a pair production event. In the false negative images, Figs. 2.14a to 2.14d, three features can be observed in the pair production images: short apparent track length in Figs. 2.14a to 2.14c, overlapping of the two tracks

making it appear as a singular track in Figs. 2.14a and 2.14b, and deep inelastic scattering events in Fig. 2.14d. Each of these characteristics are indicative of situations where the ionization track would be made to look more like a background event than pair production.

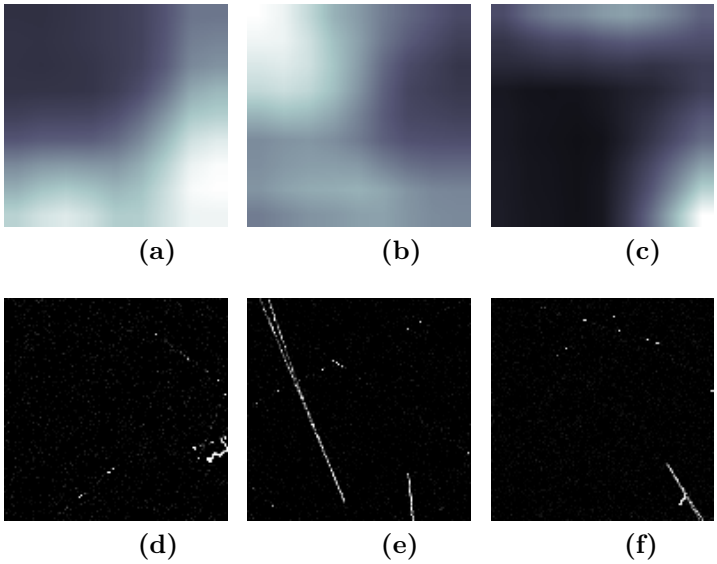


Figure 2.15. a-c) Images generated by the Gradient-weighted Class Activation Mapping (Grad-CAM) algorithm that demonstrate the features that GammaNet utilizes for classifying images as background or signal. d-f) The simulation images used to generate the respective Grad-CAM images, with d) being a background event and e) and f) being signal events.

2.6 GammaNet Visualization

As the use of CNNs becomes more prevalent in research, it is of increasing interest how the CNN performs the classification and what features of the input it uses to do so. Grad-CAM [76] is a recent algorithm developed to answer these questions by providing an activation map for input images that shows the regions the CNN used most within the image during classification. Fig. 2.15 shows the Grad-CAM images

generated for GammaNet with one background image, Fig. 2.15d, and two signal images, Figs. 2.15e and 2.15f. These images were generated using the sub-scale simulation of AdEPT because the lower track density provides interpretable results. Fig. 2.15a shows that for the background class, GammaNet utilized sparsely ionizing tracks and delta rays present in Fig. 2.15d, resulting correctly in a background classification. Fig. 2.15b demonstrates that for the signal class, GammaNet utilizes the separate, nearly parallel, tracks of the pair production event preferentially over the overlapped pair production event at the bottom of Fig. 2.15e, resulting in an accurate positive classification. Lastly, Fig. 2.15c results in a false background classification of the pair production image, Fig. 2.15f, with GammaNet using the sparsely ionized GCR tracks and the delta generated from the pair production track.

2.7 Conclusion

The event classification requirements of the AdEPT mission dictate a background rejection rate between 99.99% and 99.69% which must be achieved within a 50 ms time window determined by the instrument collection rate. `GammaNet`, using mixed precision enabled by `NVCaffe`, was able to achieve a background rejection rate of 99.990±0.004%. These results were achieved using the full-scale simulation, classifying on images downsampled at 11x. The time for inference was found to be on average 6.8 ms utilizing a NVIDIA GTX 1080 GPU, which has 8.2 TFLOPS of single precision compute performance. This implies that, as is, `GammaNet` would require 1.1 TFLOPS of single precision compute available to it from the on-board flight computer. The AdEPT mission is still in the development stages, and thus

the flight computer has not been chosen. Commercially available flight computers are capable of meeting this demand. Additionally NASA is investigating the use of commercial SoC solutions that possess greater than 1 TFLOPS performance [62]. In its current iteration, `GammaNet` is not prohibitively compute intensive for use as an on-board event classifier.

It was found that, in the best case situation of 99.69% background rejection, signal sensitivity for pair production will range from 1.1±0.5% to 69±2% for 5 to 250 MeV incident gamma rays. This range becomes 0.1±0.1% to 17±2% for the worst case scenario requiring 99.99% background rejection. The low sensitivity lowers the effectiveness of the AdEPT instrument, however these values were generated using conservative estimates. These results show that `GammaNet` achieves the desired background rejections of AdEPT, making it a serious consideration for use on-board the satellite for event classification.

These performance estimates include no image compression, and downlink bandwidth afforded by current and near future satellite communication [45, 59, 60, 61]. No image compression was used as a conservative assumption due to the data handling system for the AdEPT satellite not yet being decided. Simple lossless compression afforded by the PNG format produces compression ratios nearing 2 for the simulation images used in this study. As more systems aboard the AdEPT satellite are designed and implemented, more precise determination of the operational parameters of `GammaNet` can be achieved. Reductions in the raw data rate will allow `GammaNet` to operate at a lower background rejection rate, affording increased pair production sensitivity.

Grad-CAM was implemented for `GammaNet` in order to discern the features that `GammaNet` uses during classification of the simulation images. The results from this application support the supposition that `GammaNet` utilizes features that are characteristic of background for the respective classification, such as lower ionization density relative to the pair production tracks and the presence of delta rays. For the positive class of pair production the network responds strongly to semi-parallel tracks that are close in proximity, indicative of energetic pair production events.

Chapter 3

Experiment Based Application: NeutronNet

3.1 Introduction

The identification and quantification of neutron fields is inherently a difficult task. As was outlined in Section 1.4 this difficulty is due to two things: the lack of charge of photons and neutrons, and the generation of an intense secondary photon field by the principal neutron field. The lack of charge for both particles precludes the use of coincidence or anti-coincidence detection systems for event separation. Furthermore the generation of a secondary photon field by neutrons interacting with all matter makes it impossible to simply remove the photon field by shielding the detector from it. These issues left only two possible solutions: having a detector that is sensitive to neutrons and relatively insensitive to photons, or Pulse Shape Discrimination (PSD). Given the increasing difficulty of obtaining ^{3}He as well as its limitations, PSD has recently become the defacto solution for neutron and

photon event separation.

The remainder of the research contained in this thesis looks at the application of Machine Learning (ML) techniques to the task of PSD for neutron and photon event classification. This work is intended to be an extension of the research outlined in Chapter 2 in the sense that it is ML as applied to Particle Identification (PId). The limitations imposed by the instrument for Advanced Energetic Pair Telescope (AdEPT) not being completed is the inability to conduct experiments to validate and further develop `GammaNet`, which was a simulation based study. Simulation based studies, especially in the context of ML, are quite enticing. The capabilities afforded by simulation include having complete control over the contents of the physics involved, confidence in the labelling of results, and they are easily extended to arbitrary size for consideration of statistics. These benefits though are opposed by the unavoidable differences between reality and simulation, limiting simulation based studies to the realm of proof of concept. To fully develop any application requires experimental data, which is the focus of this research going forward. A review of the state of ML as applied to neutron and photon separation found there was as yet no application of Convolutional Neural Network (CNN) based ML algorithms as discussed in Section 1.4.

3.1.1 Data Generation

With each application of supervised ML there needs to be a commensurate labelled data set to train the algorithm off of. In the case of neutron photon separation, this includes a neutron data set and a photon data set, with the former presenting complications outlined in Section 1.4. The photon data set is relatively trivial

to produce because of the availability of isotopic photon sources and the limited neutron background present at sea level [77, 78]. There are three practical solutions to generating labeled neutron data: simulation of neutron detection events, Time of Flight (ToF) tagging, and PSD derived neutron and photon separation.

During this work it was found that the simulation capabilities of GEANT4 were limiting in terms of fully simulating the time dependent characteristics of signals for the proposed EJ-301 scintillator used in this work. The toolkit is limited to simulating scintillation with two principal time constants for scintillation light, but EJ-301 possesses three scintillation components. In addition to this incompatibility, which could be corrected for with additional programming, there was also limited or no literature found for the ratios of excitation for these states from different interacting particles. These issues combined made a better argument for developing the training and testing data sets for the ML application by means of ToF tagging and use of traditional PSD methods.

The enabling physical characteristic for ToF event tagging is the difference in velocity for photons and neutrons. Photons obviously travel at the speed of light and neutrons, with our energy range of interest being well below relativistic, travel at a fraction the speed of light. This difference in velocity can be carefully exploited to provide the identity of the particle that gets detected at a known distance from the source, if the emission of the neutron and photon occur at a well known time. Two such neutron sources would be those where neutron emission are accompanied by characteristic prompt photon emissions, or where an intermittent neutron emission is possible such as with pulsed accelerator based sources. The former was attempted with the ^{9}Be(α, n) neutron sources available at McMaster University,

namely ^{238}Pu^{9}Be and ^{241}Am^{9}Be. The latter would have been the objective of a lab visit to neutron metrology labs such as Physikalisch-Technische Bundesanstalt in Germany or the National Physics Laboratory in the United Kingdom, but these were made inaccessible because of the COVID-19 pandemic.

The ^{9}Be(α, n) reaction produces a neutron and a ^{12}C, with approximately 60 and 15 % chance of leaving the ^{12}C in it's first and second excited states [79], shown in the level diagram Fig. 3.1. The coincident detection of the neutron and one of these prompt photons would have an associated time delay governed by the velocity of each particle. This time difference is given by Eq. (3.1).

$$\Delta t = \frac{d}{c - v_n} or \frac{d_\gamma}{c} - \frac{d_n}{v_n} \tag{3.1}$$

With the time between coincidence signals being Δt, c being the speed of light, v_n being the velocity of the neutron, and d_x being the source to detector distances. The first solution holds if the distance between the source and each coincidence detector are kept equal, and the second being the solution for non-equal distances between detector and source.

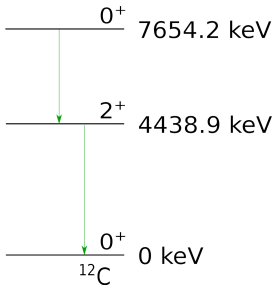


Figure 3.1. Energy level diagram for the ^{12}C nucleus. De-excitation of the second state occurs as a cascade, with each level decay promptly.

The neutron and photon separation results achieved with the ToF method for event tagging, which are discussed in Section 3.3, were not considered to be adequate. To remedy this data set separation for neutron and photon events was performed by applying traditional PSD techniques. A multitude of PSD techniques exist to try and perform PId, the most common and traditional being Tail-To-Total (TTT) [35]. Newer techniques look at more complex analysis of the signal including wavelet transforms, pulse gradients etc. to perform PSD with greater success but higher computational overhead [80, 81, 82, 83]. Given the objective with this application was to generate a pure enough training data set for neutron and photon classification TTT was found to be sufficient.

TTT produces event separation by taking the charge collected in the total signal, and then compares it to the charge collected in just the tail. The tail of the event is defined as everything following some fixed time after event start. With proper parameterization of the delay time from event start, identification of neutron and photon interactions can be achieved. This separation is done by plotting the energy of the event against the ratio of the tail and total charges. In EJ-301 the signal duration for a neutron interaction is characteristically longer in duration than that of a photon event, as discussed in Section 1.4. The resultant signal shape from the two interaction types is shown in Fig. 3.2. It is evident, that for uniform signal maxima, more of the total charge resides in the tail of a neutron signal than does a photon signal.

From the plot of TTT data there are two evident populations, with the top most population representing photons and the bottom representing neutrons. This determination is made because the closer to 1 on the y-axis a signal resides, the

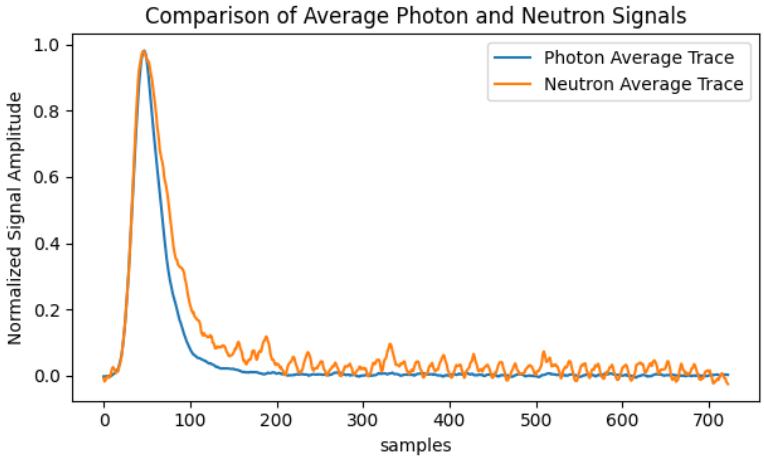


Figure 3.2. Average signals over the photon and neutron data sets. Signal event times were calculated to align pulses, and signal height was normalized to one before averaging. Noise present in the neutron signals is more pronounced due to low signal amplitudes relative to noise.

more of it's charge is present in the beginning of the signal. Conversely to this, being lower on the y-axis represents more of the signal being collected in the tail of the event. A representative TTT chart is shown in Fig. 3.3, which was generated by collecting events from an $^{241}Am^{9}Be$ neutron source, which is further discussed in Section 3.2.1. Given that, at low energy, the two populations begin to overlap, a line separating neutron and photon events needs to be made. Several attempts were made at separating these two populations, with good success found for manually drawing a conservative line. A conservative separation being one that more readily identifies events as photons, given that was thought to be the less detrimental contamination for labelling events.

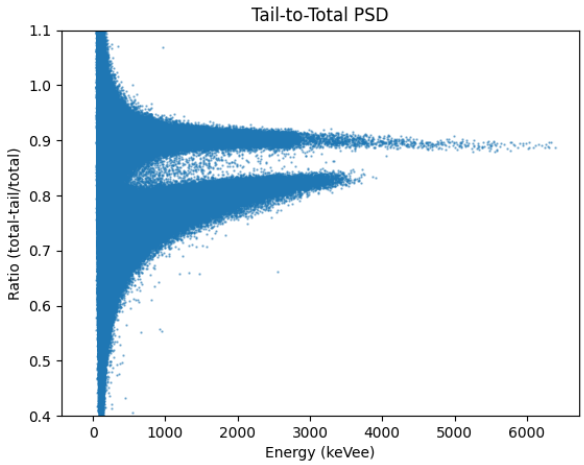


Figure 3.3. TTT plot of events recorded with the EJ-301 liquid scintillator and the ^{241}Am^{9}Be neutron source. Note that events at low energy with non-physical ratios of greater than 1 are caused by the tail of the signal having an overall negative value. Energy, as displayed here in $keVee$, is the measure of energy deposited in the scintillating material in terms of total charge measured for an event and calibrated to a linear fit of the Compton edges of ^{24}Na, ^{60}Co, and ^{137}Cs.

3.1.2 Performance Characterizations

Once a suitable data set of labelled neutron and photon events was generated, it was necessary to identify a performant ML architecture. To do so, 4 different ML architectures were implemented and tested. These architectures included more recent CNNs as well as older fully connected architectures, to investigate whether higher complexity was necessary. In addition to this the impact of sampling rate and bit depth on ML performance were investigated to complement research in the literature on these impacts on traditional PSD [84].

3.2 Methods

3.2.1 Experimental Setup

The scintillators used in these experiments were an EJ-301 liquid scintillator coupled to a Hamamatsu R7724 Photomultiplier Tube (PMT) and a $LaBr_3(Ce)$ scintillator coupled to a Hamamatsu R6233 PMT. The EJ-301 liquid scintillator was chosen for its neutron sensitivity and PSD performance, while the $LaBr_3(Ce)$ was chosen for its energy resolution as well as scintillation time. These scintillators would be used for ToF measurements and therefore both require very quick signal characteristics for effective sufficient time resolution. The advantage for the energy resolution offered by $LaBr_3(Ce)$ over the faster timing of another EJ-301 liquid scintillator is the ability to not only time gate events for ToF but also energy gate for the characteristic energies of the ^{12}C de-excitation photons shown in Fig. 3.1.

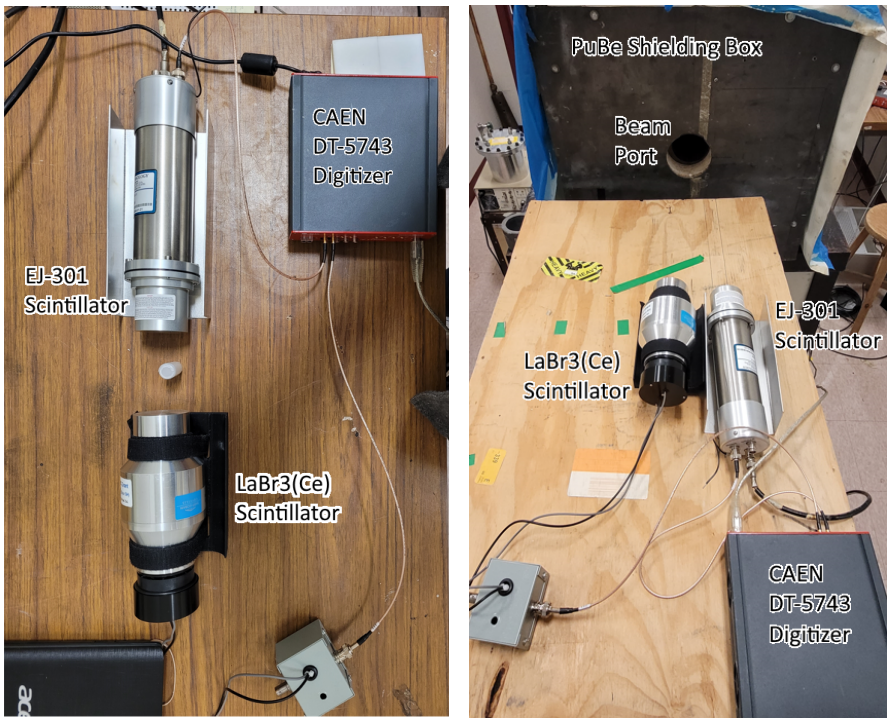
The signals from the scintillators were recorded using a CAEN DT-5743 digitizer. The CAEN DT-5743 digitizer was chosen for both its 12 bit depth and 2.5 V peak-to-peak input range, and its sampling frequency of 3.2 GHz. A rising edge threshold of 10 mV was in each measurement with the CAEN DT-5743 digitizer. The research found in the literature explored the impact on PSD of varying frequencies and bit depths, but the frequencies sampled were limited to below 2 GHz, with only 10 bits of resolution available at that rate [84, 85]. This digitizer would enable investigating a much wider range of sampling rates with a bit depth covering the useful range found in [84, 85]. During the time of the literature review and purchasing of components, more capable digitizers did exist such as the

ADQ7, which samples at 5 or 10 GHz with 14 bits of depth, though the difference in costs were significant.

Energy calibration and photon data set signals were collected from ^{24}Na, ^{60}Co, and ^{137}Cs sources. Neutron data, including ToF measurements, were collected from ^{238}Pu^{9}Be and ^{241}Am^{9}Be isotopic neutron sources. Both of these sources emit neutrons through the ^{9}Be(α,n) reaction and have very similar neutron spectra with accompanying ^{12}C de-excitation photons. The discrepancies between these neutron spectra can be attributed to the energy of the α decay for each isotope having a difference of roughly 100 keV, as well as construction of the source with grain size impacting the energy loss of the α before reacting with the target ^{9}Be. The impact of these differences on the work conducted here is believed to be inconsequential.

The experiments were conducted in three configurations with two variations each. Firstly the energy calibrations were conducted for each detector simply with the photon sources placed on the detector face. The next configuration was the coincidence analysis conducted as shown in Fig. 3.4a. Lastly was the coincidence measurements of the ^{238}Pu^{9}Be source, which were setup as shown in Fig. 3.4b. Of note in Fig. 3.4b is the the borehole in the shielding box, which is where the ^{238}Pu^{9}Be neutron source is placed during measurement, which is located at the end of the table in the top of the picture. The variations on these two coincidence experiments was the replacement of the $LaBr_3$(Ce) scintillator with a second EJ-301 later on. This replacement was done because the timing resolution was significantly worse when using the $LaBr_3$(Ce) scintillator due to its PMT which is slower and has a larger Transit Time Spread (TTS) when compared the the PMT used in the EJ-301 scintillator [86, 87]. For each experimental setup the electronics chain

includes a desktop high voltage power supply, the CAEN DT-5743 digitizer, the PMTs, and the readout laptop.

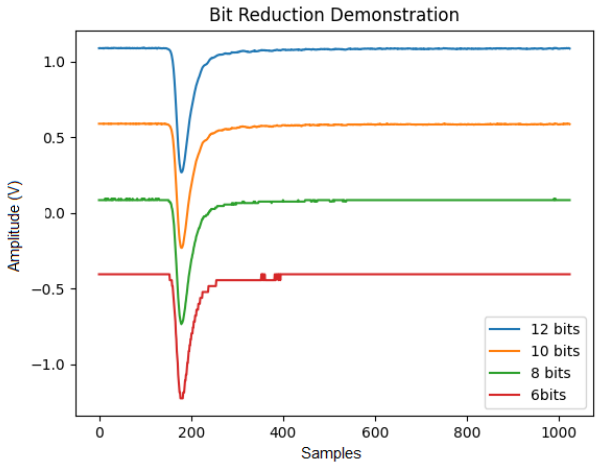


(a) **(b)**

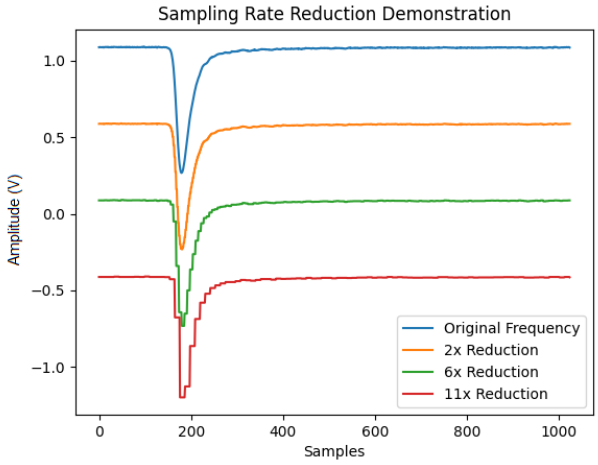
Figure 3.4. Images of the measurement setups used to generate neutron and photon data sets, perform coincident timing measurements, and to perform energy calibrations.

3.2.2 Event Timing and Energy Calibration

The event timing in this work was performed by finding the intersection point of a linear fit to the baseline before signal start and a quadratic fit to the rising edge of the signal. The linear fit was performed on a baseline of 25 samples and the quadratic fit was between 10 and 90 % of the signal max. The full length of the signal is 1024 samples or 320 ns total, which allows for some pulse pileup to occur. To avoid contamination in the training set from pulse pileup an algorithm was



(a)



(b)

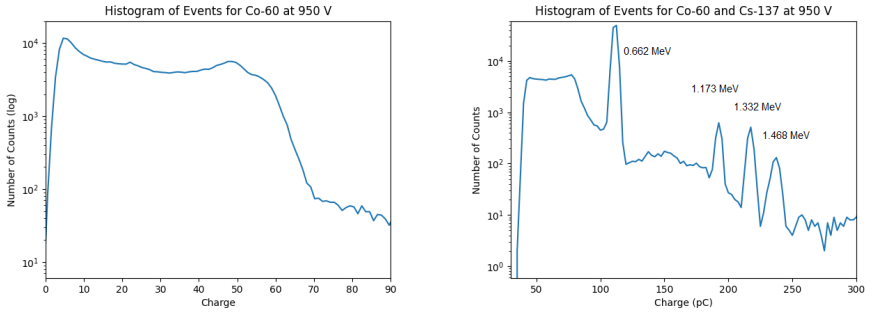
Figure 3.5. Plots of a signal demonstrating the reduction of bit depth from the original 12 bits Fig. 3.8a to 6 bits in steps of 2, and the same event with reductions in sampling frequency from 3.2 GHz Fig. 3.8b with 2, 6, and 11x reductions. A horizontal offset is provided for each representation of the event of 0.5 V for the sake of ease of viewing.

implemented to detect these events and remove the respective signals from the data set. The determination of a pulse pileup event was made following finding the event time. A signal was said to contain a pileup if the signal level rose above 15 times the noise level following a delay of 150 samples after the determined event time.

The energy of an event was determined by integrating all values in the sample following the determined event time. This is then converted to charge by considering the sampling time and the 50 Ω resistance of the PMT output. Plotting a histogram of all of these charges for a data set can then be used to calibrate for energy. The complication in this matter arises from the dominant photon interaction within a liquid scintillator being Compton scatter. This interaction creates a continuum of energy depositions within a scintillator ranging from 0 to a maximum dictated by the energy of the incident photon following Eq. (3.2).

$$E_{dep} = E_{\gamma} * (1 - \frac{1}{1 + (\frac{E_{\gamma}}{m_e c^2})(1 - cos\theta)}) \tag{3.2}$$

This continuous energy deposition prohibits the easy identification of characteristic photons that is possible with inorganic scintillators where full energy deposition is typical. A plot comparing the energy deposition histograms for the EJ-301 and $LaBr_3$(Ce) scintillators is shown in Fig. 3.6. In the $LaBr_3$(Ce) histogram, Fig. 3.6b, the relation to energy and charge is clearly discernible and calibration is trivial. For the EJ-301 histogram, Fig. 3.6a, characterizing the relationship between energy deposition and charge requires the identification of the Compton edge where the maximum energy deposition occurs. In this work the fit for the



(a) **(b)**

Figure 3.6. Charge histograms for the EJ-301 scintillator, Fig. 3.6a, and the $LaBr_3(Ce)$ scintillator Fig. 3.6b. The source involved in Fig. 3.6a is just the ^{60}Co source, for Fig. 3.6b a ^{60}Co and ^{137}Cs source were measured with an additional peak from internal decay of ^{138}La in the scintillator.

Compton edge was achieved by fitting to a convolution of a quadratic Heaviside function and a Gaussian function as outlined in [88]. This fit accounts for the Compton continuum as well as the Gaussian broadening effect of the scintillator and electronics.

Prior to determining event time and applying the TTT algorithm, each event was filtered with a finite impulse response filter. This filter simply took a running average of N signals by iterating through a pulse adding the newest element and removing the oldest element and averaging those N elements for the signal output. This filter was applied to improve the PSD results by achieving better timing results, an example filtered pulse is shown in Fig. 3.7, with the filtered signal having a 50 mV vertical offset for viewing.

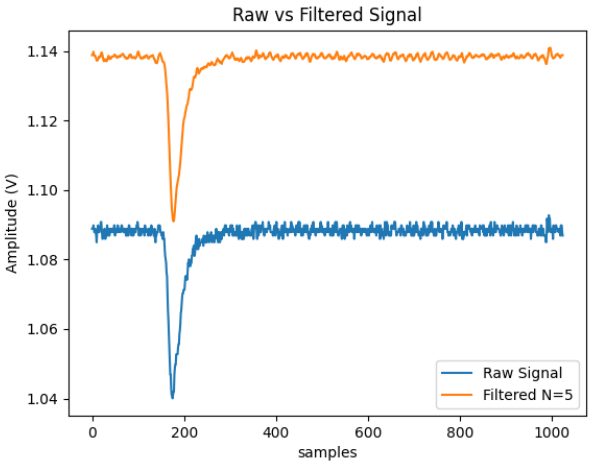


Figure 3.7. Plot of a signal acquired with the EJ-301 liquid scintillator and the resultant filtered signal with a 50 mV offset.

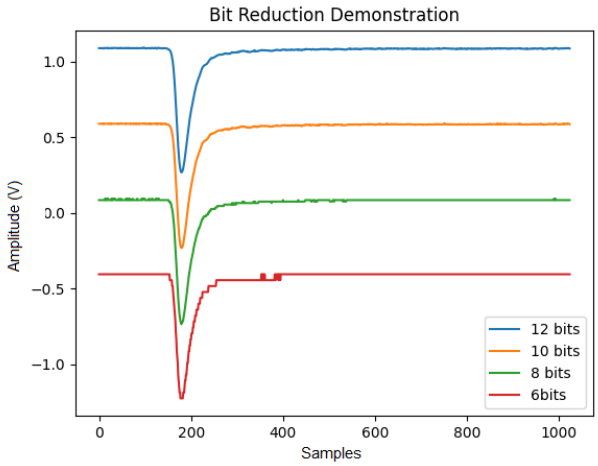
3.2.3 Bit Depth and Sampling Rate Reduction

Of interest in this work was the impact of sampling rate and bit depth on the performance of the ML solution. To enable this investigation events were recorded with full bit depth and sampling rate afforded by the DT-5743 digitizer, being 12 bits and 3.2 GHz. New versions of a data set were then generated with artificially reduced sampling rate and bit depth. This was opted for instead of reproducing each measurement with multiple digitizers at varying bit depths and sampling rates because it controls for other undesired variability such as electronic noise, dynamic range, and differential nonlinearity that change between digitizers. The generation of reduced sampling rate data sets was performed by taking every Nth value, with n being the ratio between the original sampling rate and the desired reduced sampling rate. For the reduction in bit depth the dynamic range of 2.5 Volts was divided into the number of steps possible for the desired bit depth and then values obtained in the measurement at 12 bits are rounded to the nearest value.

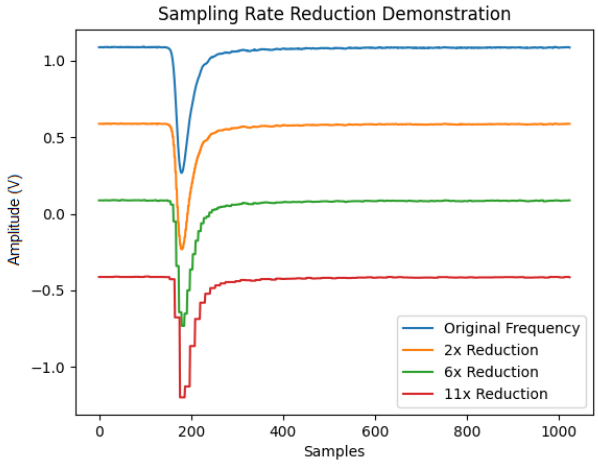
A demonstration of these reductions is shown in Fig. 3.11, with Fig. 3.8a showing the bit depth reduction and Fig. 3.8b showing the reduction in sampling rate. Bit depth reductions were done performed with integers numbers of resultant bits between 11 and 3. Sampling rate reductions were performed with integer reduction ratios of between 2 and 64 to achieve effective sampling rates of 1600 MHz to 50 MHz. Of note, the data reductions mentioned here were performed following event separation into neutron and photon data sets. This is important because the reductions in sampling rate and bit depth would severely impact the event selection routine used for separating neutrons and photons into their respective training and validation data.

3.2.4 Machine Learning Architecture Comparisons

To find an effective ML solution for event separation four algorithms were tested in this research: MultiLayer Perceptron (MLP), Recurrent Neural Network (RNN), `AlexNet`, and `GoogLeNet` [89, 90, 91, 92]. MLPs represent some of the older ML architectures and are designed around fully connected layers which are akin to inner products, though with much higher dimensionality. RNNs, `AlexNet`, and `GoogLeNet` all represent the more recent trend of machine vision architectures being built up from convolution operations called CNNs. Out of the three CNNs, `GoogLeNet` and `AlexNet` were modified to include only 1-Dimensional operations, and the input for the RNN was modified to be a 32x32 input. Additional modifications were made to `GoogLeNet` based off of the performance shown in Chapter 2, with a depth of three inception modules used in this work.



(a)



(b)

Figure 3.8. Plots of a signal demonstrating the reduction of bit depth from the original 12 bits Fig. 3.8a to 6 bits in steps of 2, and the same event with reductions in sampling frequency from 3.2 GHz Fig. 3.8b with 2, 6, and 11x reductions. A horizontal offset is provided for each representation of the event of 0.5 V for the sake of ease of viewing.

3.2.5 Machine Learning Training

In order for a ML algorithm to be able to perform classification it first is necessary to train it with data representative of the classification problem it will be applied to. The training performed in this work represents supervised ML, where the training data is already labelled for which class it belongs to which in this case is neutron or photon. Again, the generation of a photon data set is mostly trivial given the limited neutron background that is experienced near sea level [77, 78]. In the limiting case of low activity photon sources this can become an issue when trying to determine PId performances with only 1:1,000,000 error rates, as will be discussed in Chapter 4. With high enough activity sources, the neutron rate is low enough that contamination in the photon data set is negligible as shown in Fig. 3.9, which was collected with ^{60}Co and ^{137}Cs sources. In this work the TTT PSD results presented are found using a delay time of 24 ns to define the tail portion of the signal.

The issue of generating training and validation data for neutron classification is more difficult. As is the problem that this work aims to address, whenever neutron sources are present a significant secondary photon field is generated. In order to be able to successfully train a ML algorithm it is necessary to have the purest possible training data in terms of correct labelling. To achieve this in this work PSD was performed on measurements of neutron sources and the resultant neutron population and photon population were identified by manually drawing regions around them as can be seen in Fig. 3.10.

When drawing the bounds around the neutron and photon region attention was paid to broadening the region as energy decreased, but to avoid the region

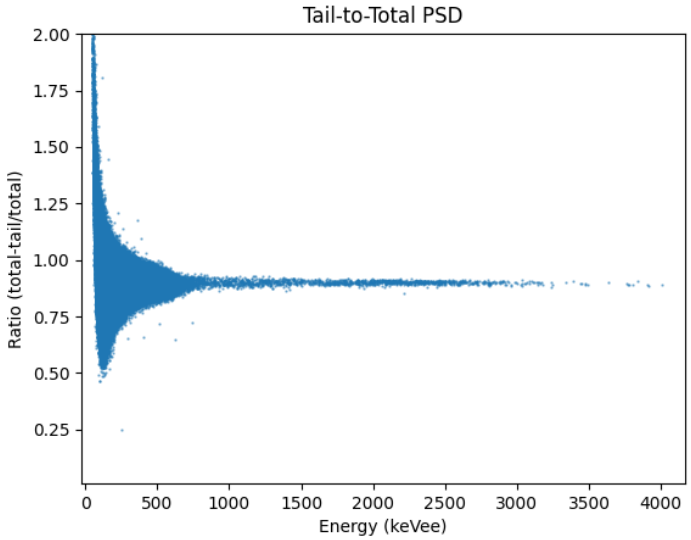


Figure 3.9. TTT PSD generated from photon data collected with a ^{60}Co and ^{137}Cs photon source.

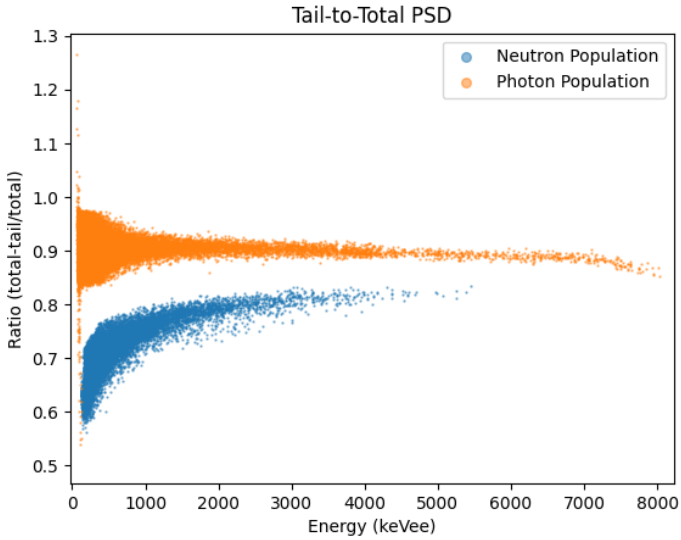


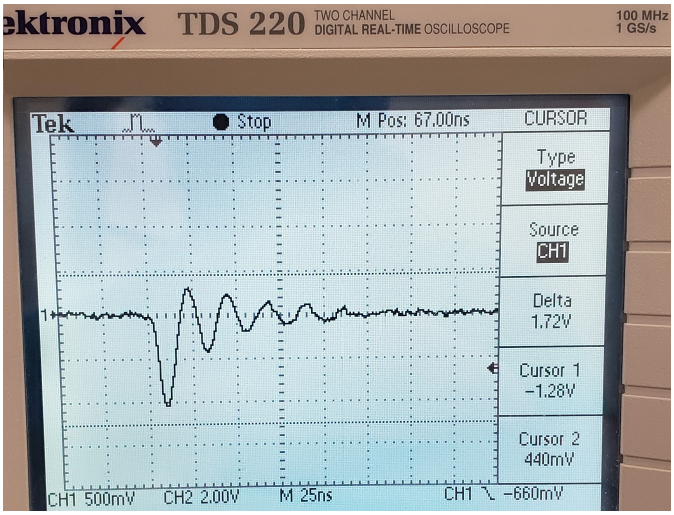
Figure 3.10. Plot of the TTT PSD results when data collected from an Am-Be neutron source were ingested. Manual regions were drawn around the neutron and photon populations to separate for training and validation of ML algorithms.

of overlap between the populations as much as possible. This avoidance was important to reduce the amount of improperly labelled events during training that would negatively impact performance of the ML algorithms. Further to avoiding this region of overlap for training data set generation, balancing the number of neutron and photon events improved performance, likely by avoiding generating a bias for one event over the other.

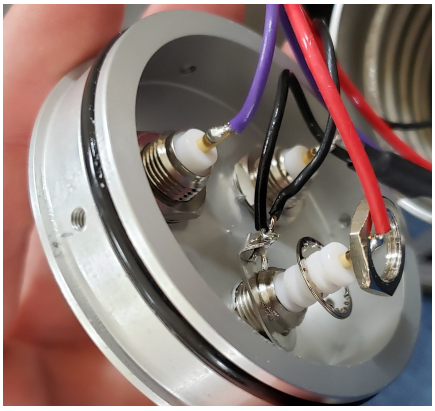
3.3 Results

3.3.1 Equipment Verification

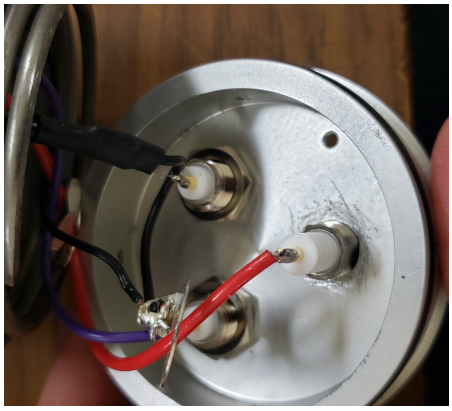
Upon receipt of the EJ-301 scintillator, tests were performed to observe the signal quality. These aimed at characterizing the noise, identifying if any power supply issues occurred, and validating there were no issues with signal. When investigating the signal significant signal reflections were observed indicating an issue with grounding in the PMT which are shown in Fig. 3.11a. The grounding issue was being caused by the anodized aluminum layer on the PMT end cap, which was too thick for the grounding connection to be sufficient. This issue was alleviated by roughing the surface of the PMT end cap exposing the aluminum for a better pathway to ground, shown in Fig. 3.11c. This issue is typically avoided by installation of a washer with teeth between the end cap and the grounding lug, which is intended to penetrate the anodized aluminum layer. It can be seen in Fig. 3.11b that this was incorrectly installed between the fastening nut and the grounding lug.



(a)



(b)



(c)

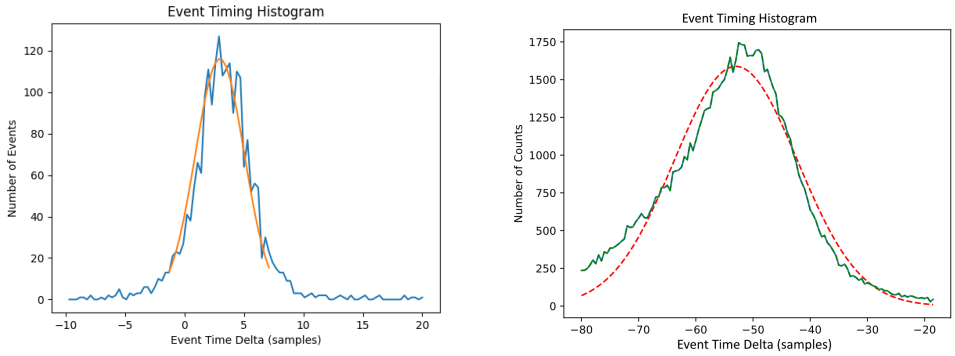
Figure 3.11. Image of the signal reflection, Fig. 3.11a, which was caused by improper installation of the washer and grounding lug, Fig. 3.11b, and the implemented fix to improve the ground contact Fig. 3.11c.

3.3.2 Timing Resolution

The initial approach taken to separate neutron and photon events was through implementation of ToF. This process requires accurate timing information to be able to correctly gate for neutron or photon events. To identify the timing resolution of the system and algorithms, coincident measurements of a ^{24}Na source were made with a pair of EJ-301 scintillators as well as with an EJ-301 scintillator and the $LaBr_3$(Ce) scintillator. The primary difference between these sets are the PMTs involved, with both being the Hamamatsu R7724 for the EJ-301 scintillators and R6623 for the $LaBr_3$(Ce) scintillator. The Hamamatsu R7724 has a transit time of 29 ns and a TTS of 1.2 ns and the Hamamatsu R6233 has a transit time of 52 ns and a TTS of 8.5 ns [87]. This was observed in experiment when finding the timing resolution of the setup including the $LaBr_3$(Ce) scintillator, shown in Fig. 3.12b, which results in a σ of 10.78 samples or 3.37 ns. When using the two EJ-301 scintillators this improves drastically, as shown in Fig. 3.12a, which results in a σ of 2.06, corresponding to 0.644 ns. When referring to the TTS, where Hamamatsu reports the Full Width at Half Maximum (FWHM), these values are 7.93 and 1.52 ns for the measurements including the $LaBr_3$(Ce) scintillator and without it respectively. The event timing resolution is within what should be expected given the published values from Hamamatsu [87].

3.3.3 Event Selection

ToF event tagging was attempted with coincidence measurements of the ^{238}Pu^{9}Be neutron source. The ToF measurements were made with the EJ-301 scintillator as the neutron signal and the $LaBr_3$(Ce) scintillator acting as the coincidence



(a) **(b)**

Figure 3.12. Event timing histograms for measurement of coincident photons from ^{24}Na by two EJ-301 scintillators Fig. 3.12a, and an EJ-301 scintillator combined with a $LaBr_3$(Ce) scintillator Fig. 3.12b. Note the negative values in time differences in Fig. 3.12b are caused by the longer transit time of the Hamamatsu R6233 PMT used with the $LaBr_3$(Ce) scintillator.

trigger for the ^{12}C de-excitation photons. The gating of events were made on both time of arrival for the EJ-301 signal as well as energy deposition in the $LaBr_3$(Ce) scintillator. The $LaBr_3$(Ce) scintillator was used, in spite of its worse timing resolution, for its superior energy resolution allowing for identification of full energy deposition from the ^{12}C de-excitation photons. The resultant ToF spectrum achieved with this setup is shown in Fig. 3.13, showing two distinct peaks.

The larger peak is photon coincidences, likely from detection of the two prompt photons from ^{12}C de-exctitation or cross talk from Compton scattering in either of the two scintillators due to inadequate shielding of each detector. The following peak is neutron coincidences, with relative arrival times of roughly 50-80 samples or 15-25 ns. The energy spectrum from the two detectors, before and after the energy gating, is shown in Fig. 3.14.

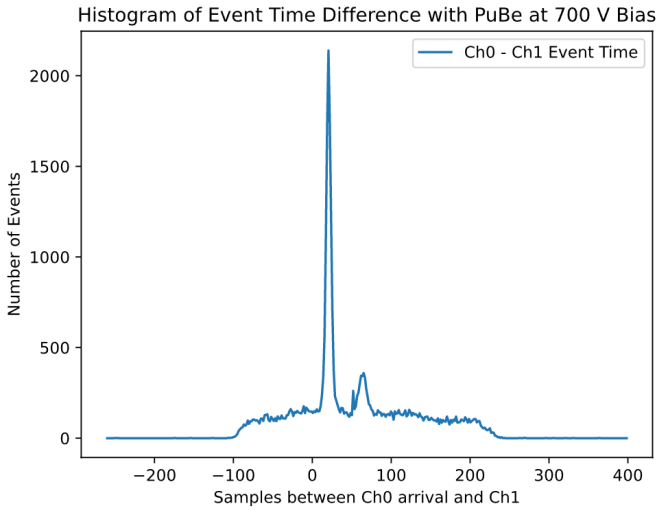
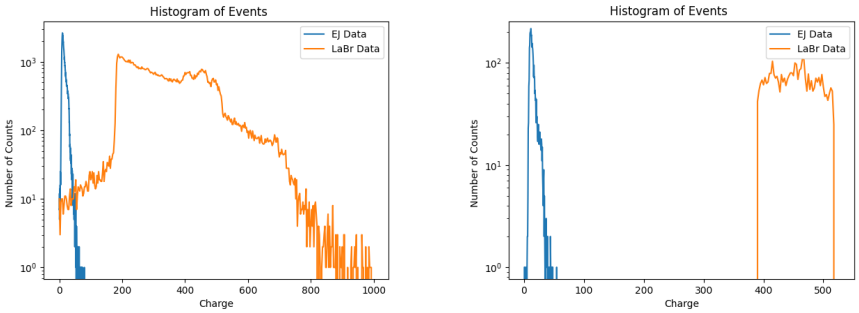


Figure 3.13. Histogram of the time difference for coincident events measured in a $LaBr_3(Ce)$ scintillator and an EJ-301 neutron sensitive liquid scintillator. The event timing window on the CAEN DT5743 for coincidence was set to 50 ns.



(a) **(b)**

Figure 3.14. Histograms of collected charge for events detected within a 50 ns coincidence window with EJ-301 and $LaBr_3(Ce)$ scintillators exposed to a $^{238}Pu^{9}Be$ neutron source.

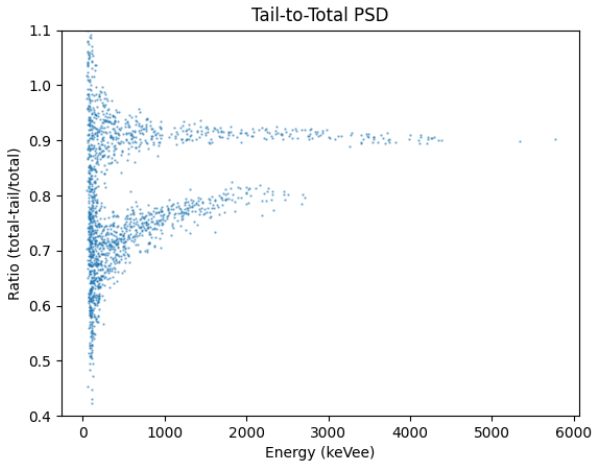


Figure 3.15. TTT PSD results for the ToF tagged "neutron" events from the ^{238}Pu^{9}Be neutron source.

The success of the ToF method was investigated by applying TTT PSD to the events determined to be neutrons by the ToF setup. These results, shown in Fig. 3.15, indicate that there was a large degree of photon contamination in the supposed neutron data set. This is obvious from the observation of two very well separated populations, as was observed in Fig. 3.10. This degree of contamination was found to be too great for machine learning and so this method of data set generation was abandoned. Label generation was instead performed by manually separating the neutron and photon populations in TTT PSD, as was indicated in Section 3.2.5.

3.3.4 Performance Characterization

Performance of PSD methods are typically quantified by using the Figure of Merit (FoM) [93], which plots the histogram of the pulse shape metric and fits a Gaussian to each population. From this, the FoM is calculated as the distance

between the two populations divided by the sum of their full width at half max, shown in Fig. 3.16. This performance metric does not work well for measuring the performance of ML classifiers because the output is the probability that the input belongs to one of the predetermined classes, in this case neutron or photon. The outputs observed in this work are extremely bi-modal, as seen in the histogram of outputs for the `GoogLeNet` adaptation used in this research Fig. 3.17. To compare the performance of standard PSD methods and the ML algorithms, the sensitivity to neutrons at varying photon rejection rates are investigated through use of the Receiver Operating Characteristic (ROC) curve [94].

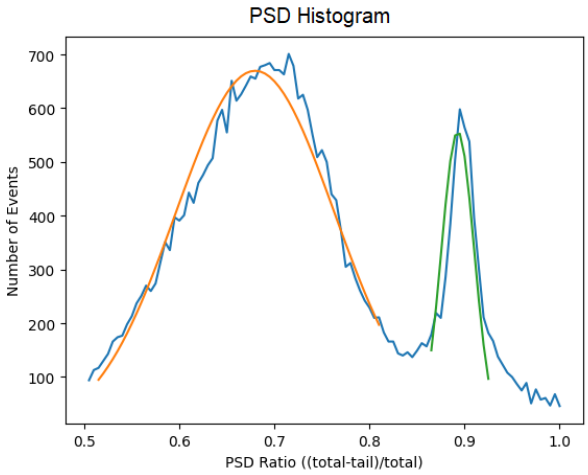


Figure 3.16. Histogram of the PSD metric for a measurement of the ^{238}Pu^{9}Be neutron source, achieving a FoM of 1.1.

ROC curves are applicable for quantifying binary classifier performance, such as the ML algorithms applied in this work. ROC curves are applied here in the same manner as for Section 2.5, but the classifications are for neutrons and photons. In the case of this application neutrons represent the positive classification and photons represent the negative classification. A resultant ROC curve is shown in Fig. 3.18 for the `GoogLeNet` adaptation. For this application the True Positive

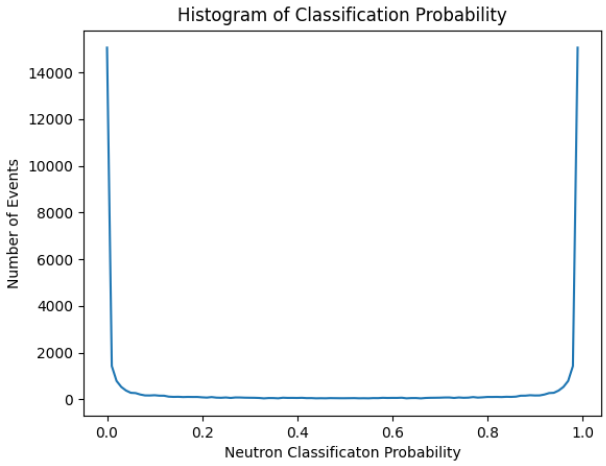


Figure 3.17. Histogram of neutron classification probability output by the neural network when classifying events recorded from the un-shielded ^{238}Pu^{9}Be neutron source.

Rate (TPR) and False Positive Rate (FPR) can then be identified as the number of neutrons and photons correctly identified over their respective data sets. The TPR can then be thought of interchangeably as signal sensitivity and the FPR as background rejection rate in the case of neutron detection.

When characterizing the performance of the ML algorithms care was taken in not providing a false sense of performance by only providing the network with well separated events in regards to PSD. This is significant because these events are already well managed by traditional PSD methods, consequently they are not of great interest when looking to out perform those methods. To avoid this issue, the photon validation data set was taken from the unaltered ^{60}Co and ^{137}Cs data sets. This photon set, below roughly 340 $keVee$, begins to impinge on the neutron population that was selected and then split for training and validation of the neutron class. This overlap is apparent when comparing Figs. 3.9 and 3.10.

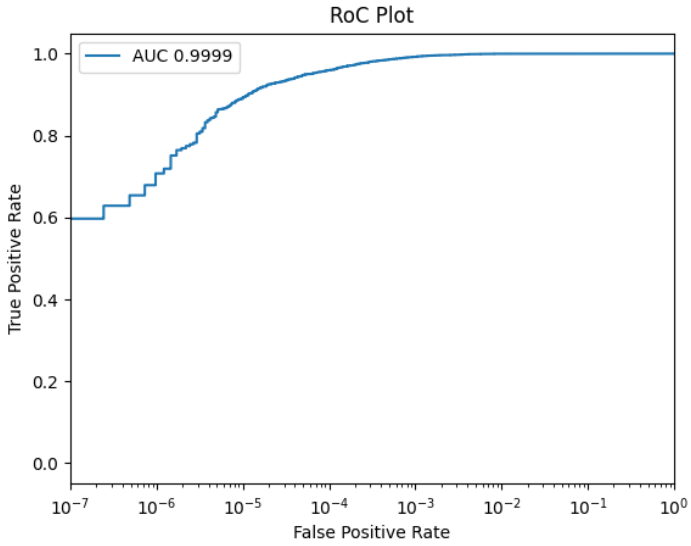


Figure 3.18. ROC curve demonstrating the classification performance of the `GoogLeNet` implementation.

In spite of this large degree of overlap in signal shape at low energy deposition the algorithm still achieves a TPR of 69.17 % while correctly rejecting 99.9999 % of the photons. Specifically, below 200 $keVee$ where both populations are entirely overlapping, the algorithm identifies 21.48 % of the events from the neutron population to be neutrons. This is significant because even with only the neutron source present, this region of the neutron population will be contaminated with secondary photons generated from the neutron field. This can be observed by comparing the extent of the photon population in Fig. 3.9 and the neutron population in Fig. 3.10. The photon population extends, at low energy, down to a ratio of 0.5 which entirely overlaps with the neutron population. The neutron population was selected to extend, at low energy, between a ratio of 0.7 and 0.5, making it impossible to encompass only neutrons in the validation data set. Avoiding this inseparability was the objective of implementing ToF, but as discussed the results

of this were insufficient.

3.3.5 Algorithm Comparisons

Many studies have looked at the comparison of differing methods in PSD to one another [95, 96, 97], with results for both traditional PSD methods as well as applications of ML, with the exception of any significant applications of CNNs. In this work the classification performance was determined utilizing four ML algorithms as well as TTT PSD. The performance when using region bounding with TTT PSD is seen in Fig. 3.19, where a region was manually drawn to get the best performance possible. This required a lower energy limit of 200 $keVee$ and achieved a TPR 0.78 of at a FPR of 6.4E-4. This is a very high degree of performance, but is limited to the already well separated region, and doesn't extend to low energy deposition. As soon as energies below 200 $keVee$ are considered, the FPR increase greatly.

The comparison of performance for each ML algorithm, shown in Table 3.1, demonstrates that when considering a low FPR, the `GoogLeNet` adaptation achieves the greatest TPR at 69.17 % for a FPR of 1E-06. Of note though is that each algorithm that implements a CNN architecture outperforms the more basic MLP. Also of note is that the RNN implementation outperforms `GoogLeNet` when not requiring FPRs below 1E-5.

Of interest in the performance of the `GoogLeNet` adaptation is the energy range of the neutrons that were identified as neutrons. To out perform the PSD provided by TTT, the identification of neutrons needs to extend below 200 $keVee$. To

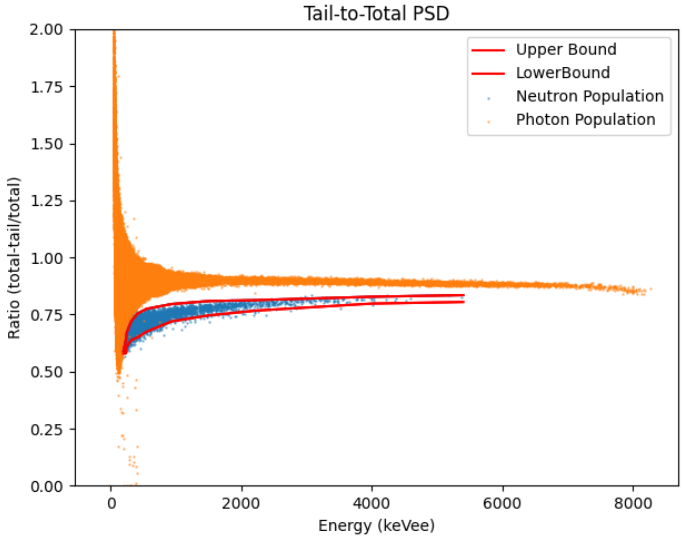


Figure 3.19. Manually bounded regions for determining classification of neutron or photon based off of TTT PSD.

observe this the neutrons correctly classified at the FPR of 1E-6 were saved and then put through the TTT algorithm for observation, shown in Fig. 3.20.

The lowest energy observed in the validation neutron set is 127 $keVee$ and for the correctly classified neutrons is 164 $keVee$. This shows that the ML approach to PSD out performs TTT in the low energy regime. At higher energies it is obvious that the ML solution fails to correctly identify some of the neutrons. A breakdown of these losses over the validation set are shown in Table 3.2.

These results show that the implementation of `GoogLeNet` used here outperforms the more traditional TTT PSD when energy deposition is below 1000 $keVee$. The loss of performance at greater than 1000 $keVee$ could be corrected for by applying both techniques gating on energy deposition for a combined approach. This

Table 3.1. TPR of each ML architecture relative to a given FPR. Data collected when trained and validated on the same data set.

FPR	TPR			
	RNN	MLP	`AlexNet`	`GoogLeNet`
1E-06	0.1346	0.0005	0.0969	0.6917
1E-05	0.9184	0.0220	0.2440	0.8882
1E-04	0.9961	0.4549	0.4181	0.9593
1E-03	0.9997	0.9820	0.5918	0.9919
1E-02	1	0.9996	0.8655	0.9998
1E-01	1	1	0.9673	1
2E-01	1	1	0.9766	1
5E-01	1	1	0.9864	1

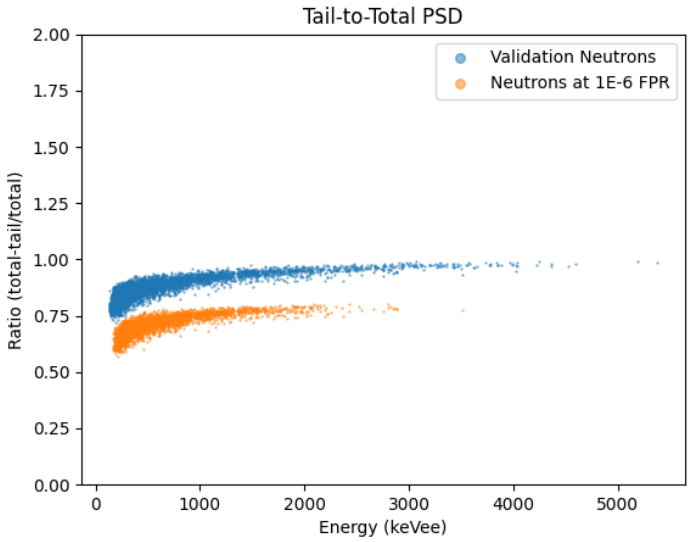


Figure 3.20. TTT plot of the validation neutron set in blue, and the neutrons correctly classified at a FPR of 1E-6 by the `GoogLeNet` implementation in orange. A vertical offset for the validation neutron data set was provided to avoid overlap to enable comparison.

would represent a minor improvement though given the marginal performance difference between 1000 and 3000 $keVee$ and the sparsity of events recorded at these higher energy depositions. Further to this, the shortcomings of `GoogLeNet` in this

Table 3.2. Table demonstrating the number of neutrons correctly identified by the competing algorithms for given energy ranges.

Energy Range (keVee)	Number of Neutrons		
	Total	TTT	`GoogLeNet`
0-200	419	0	90
200-340	955	596	741
340-1000	1723	1442	1550
1000-3000	920	757	515
3000-max	77	73	1

energy range could be mitigated by increasing the number of training examples within this energy range, which represents a small subset of the total number of neutron events.

3.3.6 Bit Depth and Sampling Rate Dependence

Looking at the performance of `GoogLeNet` when reductions in the bit depth and sampling rate are made displays a similar trend in both reductions, as seen in Table 3.3. For bit reduction, as would be expected, when reducing the bit depth of the input signals there is a reduction in the TPR. This trend is also observed for the reductions in sampling rate. In both cases, there is a marked increase in the TPR at a given reduction level which then continues the trend of lower TPRs with further reductions.

In the case of the bit depth reduction this increase in TPR occurs at 7 bits and can be explained by considering the dynamic range of the system of 2.5 V. Data collection on the CAEN DT-5743 used in this work was triggered on a pulse height of 10 mV, therefore when discretizing to 7 bits there is a possibility of lower energy deposition events to be entirely removed when converting from 12 bits. This issue

Table 3.3. Table of the TPRs achieved for varying degrees of reduction in sampling rate and bit depth. Only the TPR achived at a FPR of 1E-6 is displayed for the sake of brevity, but the trend stays the same when varying from a FPR of 1E-6 up to 1.

FPR	Bit Depth Reduction TPRs								
	11	10	9	8	7	6	5	4	3
	0.5760	0.5595	0.1903	0.0801	0.4215	0.2993	0.1332	0.0744	0.0084
1E-6	Sampling Rate Reduction TPRs								
	2x	4x	6x	8x	11x	16x	21x	32x	64x
	0.4265	0.4139	0.3182	0.4942	0.4863	0.2397	0.3861	0.2845	0.0754

is further accentuated by the method used to convert to lower bit depths, which looked at the discretization of 2.5 V for each target bit depth and then rounded up or down to the nearest value for each data point in a signal. If the baseline of the event resides immediately next to the round up point but never crosses it, a negative polarity signal less than the resolution will be effectively removed.

This possibility impacts the validation photons more than the validation neutrons because of radiation sources used as well as the interaction mechanism for each radiation. The photon sources used to generate the validation set have an effective maximum energy deposition of 1117 keV through 180° Compton scatter for ^{60}Co, whereas the neutron energy can range up to several 1000 $keVee$. Secondly the main mechanism for neutron energy deposition for EJ-301 is elastic scattering with protons, which results in uniform energy deposition. The main mechanism for photons to interact with EJ-301 is Compton scattering, which at the energies used here is preferentially forward scattering. These two considerations result in significantly more low energy events for photons than neutrons, resulting in more photon events being removed from the reduction in bit depth

than neutrons. This is supported by looking at the number of lost photon events at 7 bits, which amounted to 10.53 % of the events, whereas no neutron events were lost. The preferential loss of photons over neutron events artificially inflates the TPR by making it significantly easier for the neural network to achieve the 1E-6 FPR utilized here.

For the case of the increase in FPR when reducing sampling rate the increase is observed at a reduction of 8x in the sampling rate. The original sampling rate of 3.2 GHz corresponds to a time resolution of 0.3125 ns, which when reducing by a factor of 8 results in a time resolution of 2.5 ns. This results in inadequate sampling of the primary component of light emitted from EJ-301, which is reported with a decay constant of 3.16 ns [31]. The delayed light component has a decay constant of 270 ns [31] which is the primary differentiation between neutron events and photon events, which would still be fully represented at the reduced time resolution of 2.5 ns at the 8x reduction in sampling rate. It is expected that the inability for the primary light yield component to be adequately sampled, while the delayed light yield is fully sampled, that enables `GoogLeNet` to improve in performance in separating the two populations. Overall the trend over the range explored remains consistent with the reductions in bit depth and sampling rate.

Of further interest is the decrease in performance of the TTT PSD observed with reductions in bit depth and sampling rate. This reduction in performance is shown in Fig. 3.21, where it is quite visible that the overlap in the two populations now occurs at significantly higher energies of up to 700 $keVee$ when observing the 8x reduction in sampling data set. Using the same region bounding as in Fig. 3.19 gives a FPR of 0.0475 with a TPR of 0.5474 compared to 0.8246 for `GoogLeNet`.

This reduced performance at lower sampling rates and bit depths for traditional PSD methods is supported by previous research exploring the two factors [84].

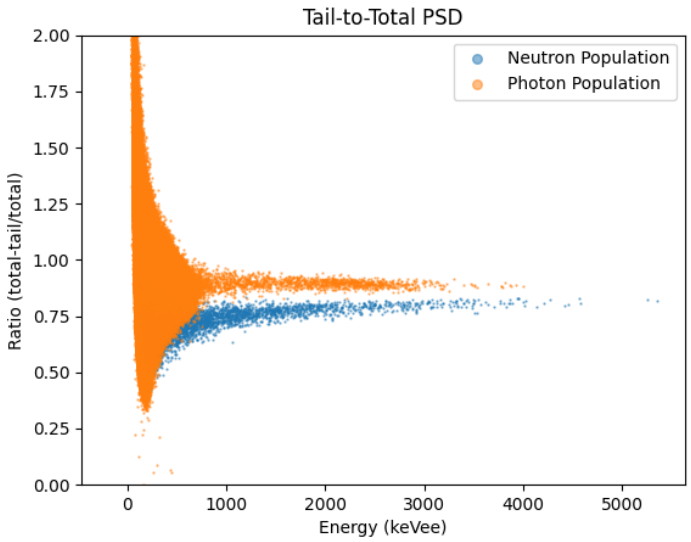


Figure 3.21. TTT plot of the 8x reduced sampling rate data set. Blue is the neutron population and orange is the photon population. The photon data set consists of 1 million events recorded with a ^{60}Co and ^{137}Cs photon source.

3.4 Conclusion

Comparisons were made between TTT PSD and four alternative ML algorithms as applied to PSD. It was found that of these four ML algorithms, the CNN based off of an adaptation of `GoogLeNet` provided the greatest performance for PSD. Comparison of `GoogLeNet`'s performance directly with that of the TTT PSD method shows that where the latter cannot effectively differentiate events below 200 $keVee$, the former obtains a TPR of 0.2148. When comparing the performance at various FPRs it was not feasible to achieve a 1E-6 FPR for TTT, but this was achievable with `GoogLeNet`. Comparison of the performance was instead done at

a FPR 6.4E-6 for TTT and 1E-6 for `GoogLeNet`, where the latter was found to outperform between 0-1000 $keVee$ and the former outperformed from 1000 $keVee$ onward.

The dependence of performance on the bit depth and the sampling rate was investigated by artificially down-sampling both aspects. The anticipated trend of reduced performance with reductions in bit depth and sampling rate were observed with an unanticipated increase in performance observed at 7 bits and 8x reduction in sampling rate. These are both anticipated to be artifacts from the implementation in reductions that impact the photon data set more than the neutron data set. The reduction to 7 bits resulted in a significant portion of the low energy photon signals to be reduced to zero, effectively removing the events that propose the most significant challenge for PSD. The reduction of sampling rate by a factor of 8 represents the point where the 3.16 ns component of the EJ-301 signal is no longer adequately sampled according to Nyquist's theory. Where Nyquist's theorem states that to adequately sample a signal you need to sample at a frequency twice greater than the highest frequency component of the signal [98].This is suspected to aid `GoogLeNet` by increasing the relative amount of information contained in the neutron signal's primary time component of 270 ns.

Chapter 4

NeutronNet: Generalization and Characterization

4.1 Introduction

Of large concern in this research was the applicability of a Machine Learning (ML) based solution which was trained on one neutron source but applied to classification for another neutron source. This would help in alleviating concerns of overfitting, which is a major issue for ML. In situations of overfitting an algorithm performs exceedingly well on test data but severely fails on new data. This is likened to the situation of fitting an n-dimensional polynomial to n data points; a perfect fit can be achieved but further extrapolation or interpolation fails entirely. To assess the applicability of the ML solution to varying neutron sources, measurements of multiple neutron sources were made. From these data sets training was performed for each source. For each trained instance testing was performed with each neutron data set.

4.1.1 Experimental Setup

In this research data from two experimental setups were utilized with five different neutron sources. These experiments were performed at McMaster University, Lawrence Berkeley National Laboratory (LBNL), and Bubble Technologies Industries (BTI).

McMaster and BTI Experimental Setup

The data collected at McMaster University and BTI were collected with an EJ-301 liquid scintillator and digitized with a CAEN DT-5743 digitizer. Three isotopic photon sources were also measured, ^{60}Co, ^{24}Na, and ^{137}Cs. These photon sources were used for energy calibration as well as for generation of the photon class for training and characterizing the Convolutional Neural Network (CNN)'s performance. At BTI a ^{252}Cf spontaneous fission neutron source was measured. At McMaster University three neutron sources were measured: a ^{238}Pu^{9}Be and ^{241}Am^{9}Be (α,n) neutron source, and a DD neutron generator. Measurements conducted at McMaster University follow the same setup as described in Chapter 3. Measurements at BTI involved a similar irradiation geometry though in a free-field room. Multiple sets of measurements of the ^{252}Cf source were made with different shielding configurations of High-Density PolyEthylene (HDPE) and lead, as shown in Fig. 4.1.

The ^{252}Cf source has an exposed and shielded position, where in the shielded position it is recessed several feet into a concrete well. When the source is exposed it is at the top of the pipe where the lead is coiled around the rope suspending the source. The use of the lead is to reduce the amount of photons impinging on

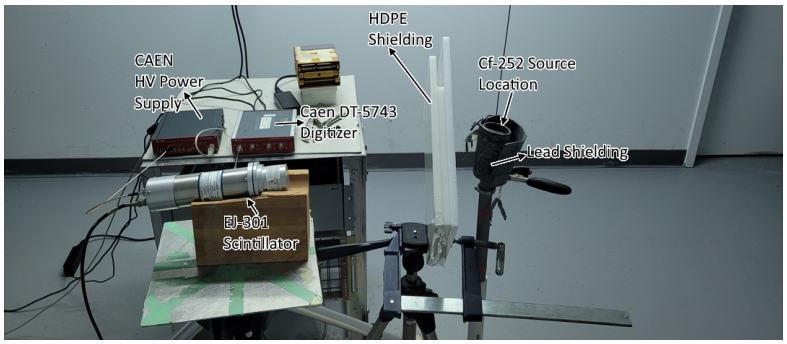


Figure 4.1. Image of the experimental setup used to measure the ^{252}Cf neutron source, with lead shielding wrapped around the source's position when exposed to the detector and with HDPE shielding in place.

the scintillator, and the HDPE is used to change the energy spectrum of neutrons reaching the scintillator. The wooden block the scintillator is resting on was used to bring the scintillator to level with the exposed position of the neutron source, allowing the lead and HDPE to shield the solid angle subtended by the scintillator.

Lawrence Berkeley National Laboratory

The data provided by LBNL was generated using an 5-inch EJ-309 liquid scintillator and digitized by a PIXIE-16 digitizer from XIA, LLC. The experimental setup was created for Associated Particle Imaging (API) and is described in details in [99]. To quickly summarize, neutrons and alpha particles are generated in a Deuterium Tritium (DT) reaction. Coincidence measurements are made between the alpha particle and a gamma-ray resulting from inelastic scattering of the neutron allowing one to calculate the 3D location of the scattering center with a resolution of several centimeters. The coincidence measurement enables drastic background reduction by filtering the data by location. The neutron data set used in this work utilizes the EJ-309 liquid scintillator as the scattering target and

data was taking in triple coincidence (alpha particle, EJ-309 inelastic scatter, and gamma ray) resulting in a very low random background event rate and therefore an ideally pure neutron data set. In practice, the application of Tail-To-Total (TTT) Pulse Shape Discrimination (PSD) shows a significant amount of photons in the neutron data sets. Subsequent measurements were made with ^{60}Co and ^{137}Cs for energy calibration and as additional sources of photons.

4.1.2 CNN Architecture

From the work outlined in Chapter 3, it was found that an adaptation of `GoogLeNet` performed the best out of a variety of Neural Networks (NNs) for PSD. The modifications used inChapter 3 are the same as were used for this work, with an adaptation for 1-D data and a reduction to three inception modules.

4.2 Methods

In order to characterize the performance achieved by *GoogLeNet* Receiver Operating Characteristic (ROC) curves [100] were generated. Similarly to Chapter 3, neutron classification is considered the positive case and photon classification is the negative case. In addition to ROC curves the TTT PSD method is implemented for visualizing neutron and photon data sets and for separation of neutron and photon data. As discussed in Chapter 3 the TTT method takes the entire charge deposition of a signal and compares that value to the charge present in the portion of the tail some time after the event start. For both EJ-301 and EJ-309 neutron signals contain a larger portion of the signal's charge in the tail than that of photons, as was shown in Fig. 3.2.

4.2.1 Photon Data Set

The *GoogLeNet* adaptation used here is trained using supervised training, where the data provided to the algorithm for training must already be labeled. In Chapter 3 it was outlined that this presents an issue for this work because of the difficulty in obtaining correct event labelling for both photon and neutron detection events. This difficulty presents itself as contamination in each data set with photons being labelled as neutrons and vice versa. The stage where this poses the largest problem is during training but mislabelling during performance validation also presents a problem, especially at the levels investigated here of 1E-6 False Positive Rate (FPR). At this low of an FPR data sets of millions of events are required to validate the True Positive Rate (TPR) performance.

For the photon data set the source of contamination is background neutrons which ideally present an insignificant contribution to the data set [101]. Given that the background neutron rate is fixed, a higher strength photon source and accompanying count rate will make this background neutron rate comprise a smaller fraction of the events recorded. When using the CAEN DT-5743 digitizer in list mode, which is necessary for this work, the counting rate is limited to 700 s^{-1}. This limitation is further worsened by the low activity of available ^{60}Co sources resulting in a significant degree of neutron events being recorded when measuring the ^{60}Co source, as can be seen in Fig. 4.2. In this TTT PSD plot you can also see events in excess of 10 $MeVee$ which represents cosmic ray muons interacting with the detector, further showing the impact of the low activity ^{60}Co source.

This measurement of 2 million events with the ^{60}Co source took a total of 13 hours. There are two populations present in Fig. 4.2, the more prevalent photon

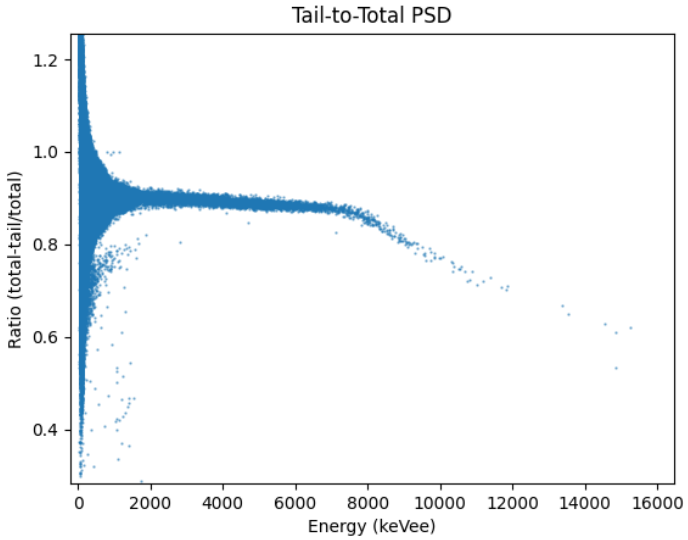


Figure 4.2. TTT PSD of data collected from a ^{60}Co source over 13 hours. High energy signals can be observed which are indicative of cosmic ray detections.

population above, and the neutron population below. In the neutron population, using the region drawing presented in Section 3.2.5, 411 of the 2 million events are identified as neutrons. With this level of contamination in the photon data set for testing, it would be impossible to achieve any TPR while maintaining the desired 1E-6 FPR. To correct for this a region was drawn around the photon population to exclude the neutron population. In addition to these ^{60}Co events, photons were collected from a ^{137}Cs and ^{24}Na source. The source activities used here were significantly higher, resulting in no observable neutron population. The TTT PSD of the full photon data set for testing is shown in Fig. 4.3, with a total data set size of 4.2 million events.

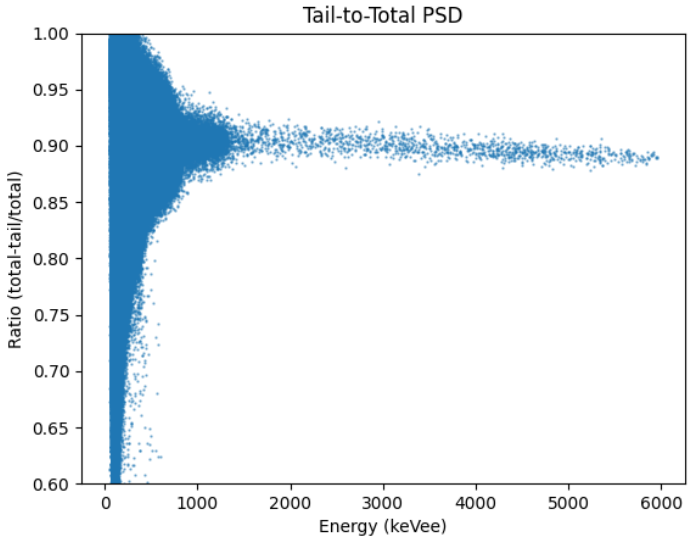


Figure 4.3. TTT PSD of the 4.2 million photon events collected for use in identifying the performance of `GoogLeNet` at PSD. Events beyond 6000 $keVee$ were removed on the basis of the maximum neutron energy of roughly 4000 $keVee$ in order to minimize the presence of signal clipping.

Attempts to fit the neutron population with the method described in Section 4.2.3 were made, but the discrepancy in population size makes the fit ineffective and results in the majority of photon events being removed.

4.2.2 Neutron Data Set

The second presentation of data set contamination is photons being present in the neutron data sets. This issue can be corrected for in a similar fashion to the photon data set contamination, but with less success. The method used is discussed more thoroughly in Section 4.2.3. The difficulty in removing the photon component from the neutron population comes from trying to retain as many low energy neutron events as possible for training. Low energy neutrons are of importance for

training and validation of *GoogLeNet*'s performance because that is the region of interest for improving upon PSD. This low energy region for the neutrons though can be somewhat dominated by low energy photons, which is apparent in Fig. 4.2 where the TTT ratio extends from beyond 1.2 to below 0.4. This contamination is unavoidable using classical methods and can lead to reduced performance in identifying neutrons by falsely labelling some of the training data as neutrons.

4.2.3 Training and Validation

To generate the training data for each neutron source, data were collected individually from each source. TTT PSD data were then generated and events were binned through energy ranges. From this, histograms of TTT ratio, or the shaping parameter, were generated and two Gaussian distributions were fit to the resultant populations as seen in Fig. 4.4. Using these fits a cut off was chosen for the shaping parameter of each population. This cut off was chosen at the point where the overlap from the opposing population was equal to 0.1 %, resulting in each data set having 99.9 % purity. The resultant two populations from this method are shown in Fig. 4.5. The data were then separated into an 80:20 split of training and validation data for use during training. Training was then performed on each data set separately to make an instance of *GoogLeNet* trained on each neutron source. In addition to the neutron and photon data sets from PSD separation of the neutron measurements, photons from the region of overlap were also provided as training data for the photon classification. This region is between a shaping parameter of 0.6 and 0.8 with energy less than 200 $keVee$. The addition of a pure photon data set within the region of overlap between the two populations greatly improved the ability of *GoogLeNet* to differentiate between the two event types.

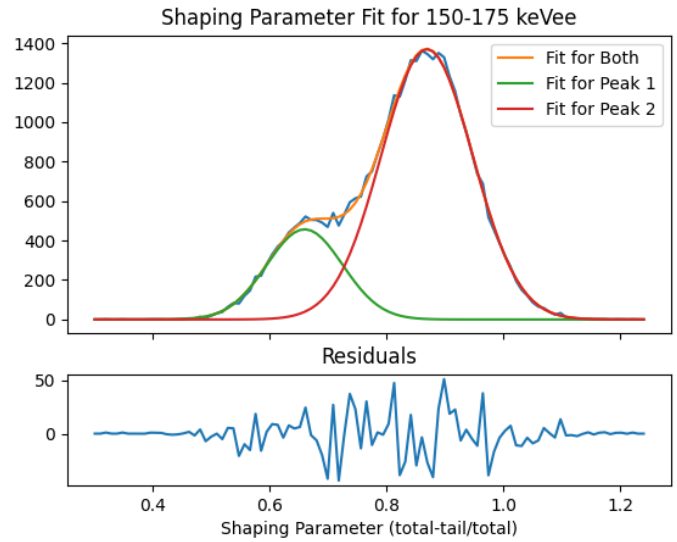


Figure 4.4. Plot of the histogram of shaping parameter for all events with energy deposition ranging from 150-175 $keVee$ with fits and resultant residuals from the fit.

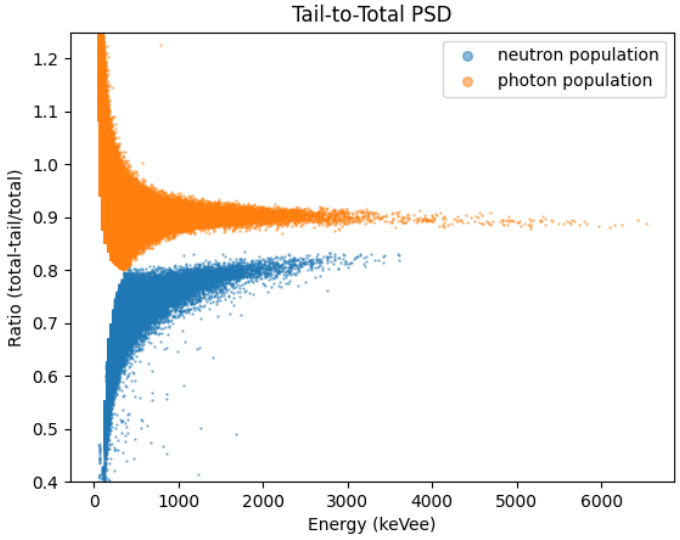


Figure 4.5. TTT PSD plot demonstrating the separation of neutron and photon data for use in training and validation of *GoogLeNet*. This data set was generated from the ^{252}Cf neutron source.

Training the implementations of *GoogLeNet* performed well across each data set, reaching a steady state performance within 5 epochs, with an epoch being one pass over the entirety of the data set. A plot of the training and validation accuracy over each epoch is shown in Fig. 4.6 and is characteristic of the performance for each data set. The most stable results were achieved when applying an initial learning rate of 0.001 with an exponential decay occurring every epoch with a decay rate of 0.96.

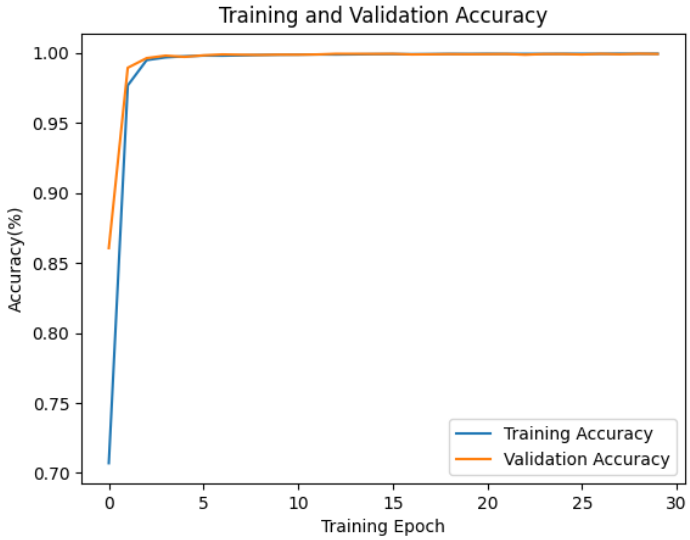


Figure 4.6. Plot of the training and validation accuracy achieved during training, with training accuracy being accumulated over an epoch and validation accuracy being evaluated following each epoch.

Following training off of one neutron set, classification was performed on the data collected from the various neutron sources. This investigation, further discussed in the results, demonstrated dependence in performance on which neutron source was used for training. To reduce this bias, and improve classification performance when applied to other sources, transfer learning was applied [102, 103].

Transfer learning takes an already trained NN and fixes the learned parameters for some subset of the NN and then retrains the remainder of the parameters on new data. This technique allows for training a NN on a limited data set with greater success, provided the classification task is similar enough. In this work what was done was an instance of *GoogLeNet* trained on one neutron source had all but the fully connected layer at the end fixed, and training was carried out with a different neutron source.

4.3 Results

Utilizing the ROC data enables quantifying the TPR for every FPR, where the former is the rate neutrons are correctly identified as neutrons and the latter is the rate photons are falsely identified as neutrons. The most significant and easily verifiable of the two performance metrics is the FPR. This is because a photon data set can be made significantly more pure than a neutron data set by measuring only a photon source. The only contamination present is background neutron events, as was discussed in Section 4.2.1. Given this availability, each ROC curve was generated with 4.2 million photon events for the McMaster and BTI data, and 1.2 million photon events for the LBNL.

4.3.1 PMT Bias

The performance when classifying events generated from the DT neutron source with varying PMT bias are shown in Table 4.1. Two significant characteristics can be seen, the highest performance is achieved at 1500 V bias, and a marked decrease in performance occurs at 1550 V. The peak in performance at 1500

Table 4.1. TPRs values achieved at various FPRs with data collected from the DT neutron source with varying Photomultiplier Tube (PMT) biases.

PMT Bias (V)	TPR			
	FPR: 1E-6±2E-6	1E-5±5E-6	1E-4±2E-5	1E-3±6E-5
1300	0.3481±0.0008	0.7596±0.0008	0.9645±0.0005	0.9927±0.0005
1350	0.3394±0.0008	0.8654±0.0007	0.9862±0.0005	0.9954±0.0005
1400	0.4428±0.0009	0.9694±0.0005	0.9920±0.0005	0.9960±0.0005
1450	0.3036±0.0008	0.8406±0.0007	0.9860±0.0005	0.9987±0.0005
1500	0.8099±0.0007	0.8917±0.0005	0.9862±0.0005	0.9965±0.0005
1550	0.0517±0.0004	0.0961±0.0005	0.5269±0.0009	0.9916±0.0005
1600	0.1173±0.0005	0.2464±0.0007	0.9465±0.0005	0.9896±0.0005

V is believed to occur because this is where maximum energy events begin to clip, which allows for the DT events to fully utilize the entire dynamic range of the digitizer. Where signal clipping is when the voltage level to be digitized exceeds the capabilities of the op-amp in the circuit leading to non-linearities in the signal. The full utilization of the dynamic range gives the greatest representation of the events, and maximizes separability of events based off of TTT PSD. The subsequent decrease in performance at a bias of 1550 and 1600 V are believed to be caused by the increased presence of clipping in these signals. Further to this, the relative performance gain between 1550 and 1600 V is due to the 1600 V data set having a lower proportion of low energy signals, which *GoogLeNet* struggles with in classifying. The higher proportion of low energy signals in the 1550 V data set then causes the performance to diminish relative to the 1600 V data set.

These observations are further supported by looking at the change in the Figure of Merit (FoM) as the PMT bias is changed. The FoM, as used here, gives a measure of the performance of the TTT PSD method at separating events into

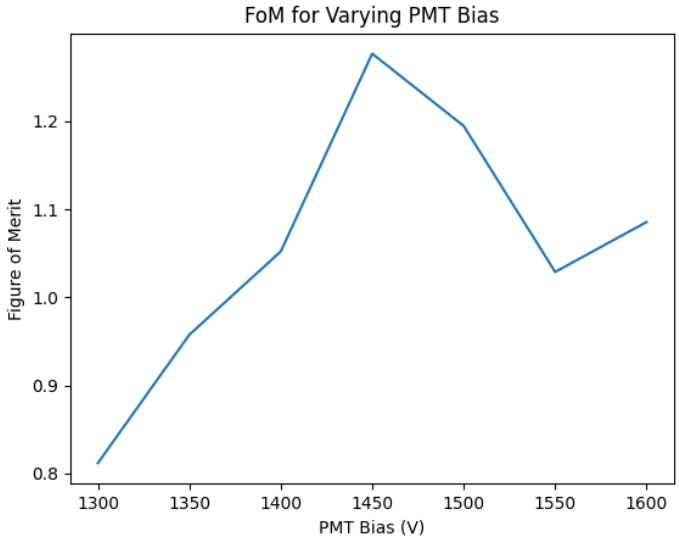


Figure 4.7. FoM values achieved with TTT PSD at varying PMT biases. FoM was calculated by averaging the FoM calculated for two energy bands per data set.

neutrons and photons. The FoM can be calculated according to Eq. (4.1).

$$FoM = \frac{S}{\delta_1 + \delta_2} \tag{4.1}$$

With S being the separation of the two Gaussians fit to the histogram of the TTT shaping parameter, and δ being the full width at half maximum of each Gaussian. The change in FoM with PMT bias is shown in Fig. 4.7, with a trend of increasing FoM with increasing bias. A subsequent decrease in FoM is observed when clipping begins to degrade separation.

Table 4.2. TPR values achieved by *GoogLeNet* following training on the various data sets and being applied to different neutron data. The TPRs for *GoogLeNet* when validating on neutrons from the same source used to train are in bold. These TPRs values are reported at a FPR of 1E-6.

Validation Source	Training Source				
	^{241}Am^{9}Be	^{238}Pu^{9}Be	^{252}Cf	DD	DT
^{241}Am^{9}Be	**0.5621**	0.6669	0.6570	0.5300	0.7545
^{238}Pu^{9}Be	0.5887	**0.6903**	0.6670	0.5302	0.7652
^{252}Cf	0.6259	0.6951	**0.9061**	0.8291	0.6612
DD	0.4968	0.5902	0.9033	**0.9090**	0.6098
DT	0.0000	0.0000	0.0000	0.0000	**0.7154**

4.3.2 Neutron Source Dependency

The performance of *GoogLeNet* when trained off of the various neutron sources can be seen in Table 4.2. From these data a few significant differences are apparent when applying the *GoogLeNet* implementation that was trained off of one neutron source to another. Firstly, the data collected at LBNL from the DT neutron source utilized a different scintillator with a different geometry. The DT neutron data utilized a 5"x5" right cylindrical EJ-309 scintillator compared to a 3"x3" right cylindrical EJ-301 scintillator used for the other neutron source measurements. The pulse shape characteristics for these two scintillators are for the most part the same, but degraded PSD performance for EJ-309 has been found [104], as well as decreases in PSD for larger detector geometries [105]. Lastly, the electronics used at LBNL vary from those used for the McMaster and BTI measurements. These variations lead to an understandable difficulty in *GoogLeNet* being trained on the data collected at McMaster or BTI and being applied to the LBNL data. Conversely, when *GoogLeNet* is trained on the LBNL data it performs well with

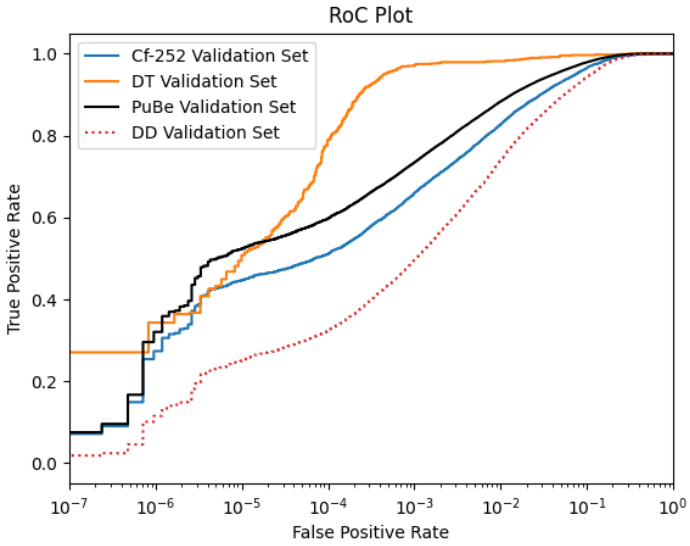


Figure 4.8. ROC curves generated for various validation data sets following retraining of a *GoogLeNet* instance. Initial training was performed with ^{238}Pu^{9}Be data and subsequent retraining was done with DT data. ^{241}Am^{9}Be validation data is omitted because of the similarity in performance to ^{238}Pu^{9}Be.

classifying neutron events from McMaster or BTI. No explanation for the lack of symmetry is available at this time. One other discrepancy in this data exists in the differences in performance when trained off of the (α,n) neutron sources or the DT and ^{252}Cf neutron sources. similarly to the case with the DT data, there is an asymmetry apparent, though to less of an extreme. These discrepancies are somewhat resolved when transfer learning is applied.

4.3.3 Transfer Learning

Transfer learning in this work was performed by freezing all but the final layer of *GoogLeNet*, the fully connected layer, and retraining the network on the alternate data set. When doing this for the worst case of *GoogLeNet* trained on

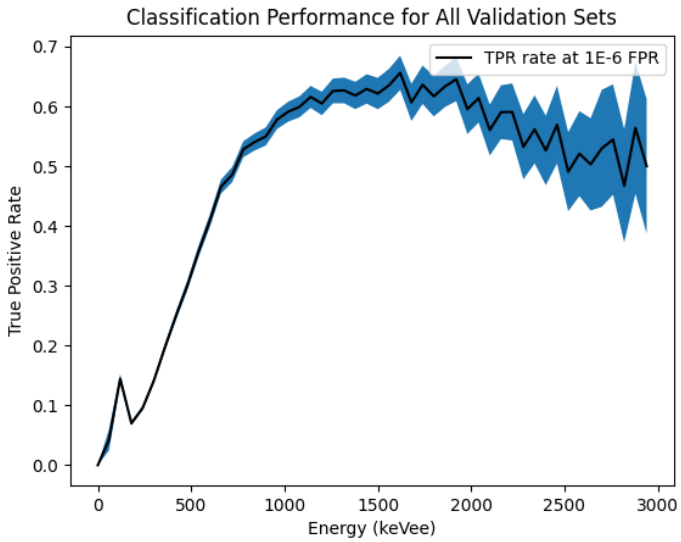


Figure 4.9. Plot of the TPR achieved at a 1E-6 FPR for varying energies. This data was generated by classifying ^{238}Pu^{9}Be, ^{252}Cf, DD, and ^{241}Am^{9}Be validation data with an instance of *GoogLeNet* that was initially trained on ^{238}Pu^{9}Be data and then retrained with DT data.

data collected with the EJ-301 scintillator and applied to data collected with the EJ-309 scintillator the performance of the network on all sources is relatively normalized. Specifically the performance when originally trained on ^{238}Pu^{9}Be data and retrained on DT data is shown in Fig. 4.8. The number of photons with which performance was validated varied between the DT data and the remaining neutron source data, with 1.2 million photons for the former and 4.2 million photons for the latter. When considering the performance at a FPR of 1E-6 it can be seen that for ^{238}Pu^{9}Be, ^{252}Cf, DT, and DD the TPRs are 32.1 $\pm$ 0.3, 27.4 $\pm$ 0.9, 34 $\pm$ 3, and 11.6 $\pm$ 0.6 % respectively. The discrepancies between these performance levels can be explained to an extent by considering the energy distribution of events for each source in combination with the classification performance's dependence on energy.

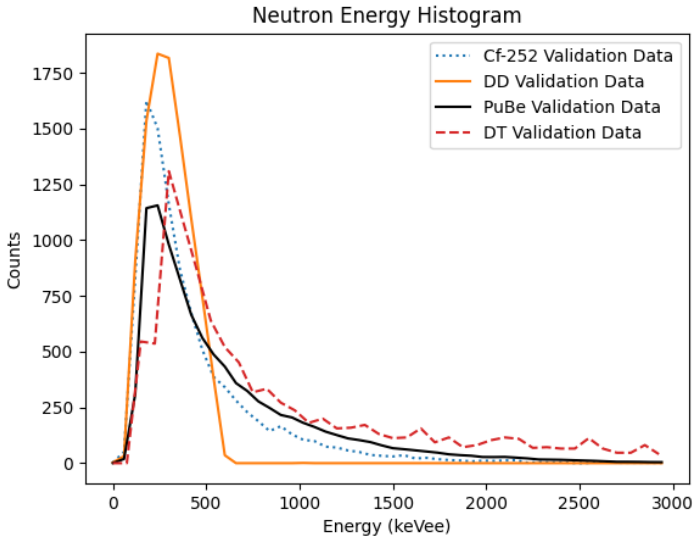


Figure 4.10. Energy deposition histogram for each neutron validation data set. ^{241}Am^{9}Be is omitted for visibility due to the similarity to the ^{238}Pu^{9}Be data. Each histogram is normalized in area to the ^{252}Cf data set. The number of bins is reduced for DT to account for its relatively small data set.

4.3.4 Energy Dependence and Correction

The classification performance exhibited by this instance of *GoogLeNet* when considering energy is shown in Fig. 4.9. This data demonstrates that there is significantly more difficulty in classifying low energy data for this instance of *GoogLeNet* where initial training data was from the ^{238}Pu^{9}Be neutron source and transfer learning was performed with data from the DT neutron source. Considering this then with the observed neutron energy distribution of the validation data sets, shown in Fig. 4.10, explains why there is diminishing classification performance when classifying the ^{252}Cf and DD data compared to ^{238}Pu^{9}Be and DT, with the former having a higher proportion of low energy events. When accounting for these discrepancies by multiplying the energy dependant TPRs in Fig. 4.9 by the

respective histograms the cumulative theoretical TPR is 31.1 ± 0.2, 24.0 ± 0.3, 37 ± 2, and 16.4 ± 0.3 % for ^{238}Pu^{9}Be, ^{252}Cf, DT, and DD validation sets. This gives a relative discrepancy between the observed TPR and that expected of 2.99 – 41.52 %. This degree of match between observed and expected TPR suggests that the majority of the variation in observed TPRs can be accounted for by variation in energy distribution, though there is still a large unexplained discrepancy for the DD validation data set.

Following this same process with the ^{252}Cf data, retraining the *GoogLeNet* instance first trained off of the ^{238}Pu^{9}Be data, shows a large degree of agreement in TPRs for each data set collected with the EJ-301 liquid scintillator. The ROC curves for each source are shown in Fig. 4.11. The TPRs observed at a 1E-6 FPR were 75.3 ± 0.9, 76.0 ± 0.3, and 65.8 ± 0.9 % for ^{252}Cf, ^{238}Pu^{9}Be, and DD respectively. This largely resolves the asymmetry observed in Table 4.2 for *GoogLeNet* instances trained off of DD or ^{252}Cf data where significantly higher TPRs were achieved for DD and ^{252}Cf than ^{238}Pu^{9}Be or ^{241}Am^{9}Be. The inability to effectively classify the DT validation data set still remains, with no response until a FPR of 3E-3.

4.3.5 Energy Dependence in Training

The energy dependence of TPR for the instance of *GoogLeNet* retrained with ^{252}Cf data and initially trained on ^{238}Pu^{9}Be is shown in Fig. 4.12. These results are vastly different compared to Fig. 4.9, which suggests that the energy response is dependent on the energy distribution of events provided during training. The relative abundance of high energy events from the DT data imparts increased

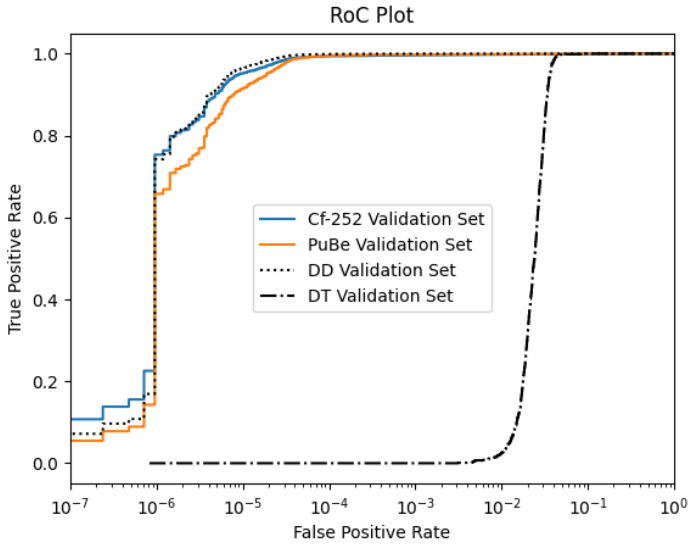


Figure 4.11. ROC curves generated for each validation data set by the *GoogLeNet* instance which was initially trained off of the ^{238}Pu^{9}Be data set and then retrained with the ^{252}Cf data set. Results from the ^{241}Am^{9}Be validation data set are omitted for visibility given their similarity to ^{238}Pu^{9}Be.

sensitivity for high energy events when classifying, with the peak of roughly 60 % occurring above 1200 $keVee$. Conversely the abundance of low energy events observed in the ^{252}Cf data produces improved sensitivity for low energy events, with up to 80 % TPR below 1000 $keVee$ seen in Fig. 4.12. Combining again the energy histogram of each source with the energy dependant TPR observed in Fig. 4.12 for the instance of *GoogLeNet* that was trained on ^{238}Pu^{9}Be then retrained of 252Cf data we can compare the observed and anticipated cumulative TPRs. The anticipated TPRs are 73.5 ± 0.8, 75.1 ± 0.3, 72.8 ± 0.8 % for ^{252}Cf, ^{238}Pu^{9}Be, and DD respectively. These results are within 1.2 – 9.6 % of the observed values stated previously. Of note is the discrepancy between the expected and observed TPR for DT where an anticipated TPR of 72 ± 3 % is found, but the observed value is 0 % at the stated 1E-6 FPR.

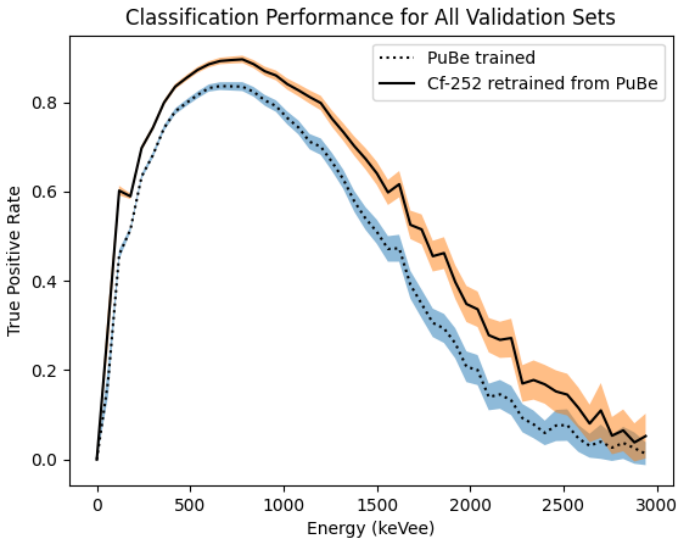


Figure 4.12. Plot of the TPR achieved at a 1E-6 FPR for varying energies. This data was generated by classifying ^{238}Pu^{9}Be, ^{252}Cf, DD, and ^{241}Am^{9}Be validation data with two instances of *GoogLeNet*. The first was initially trained on ^{238}Pu^{9}Be data and then retrained with ^{252}Cf data, the second is the initial instance trained off of ^{238}Pu^{9}Be data.

4.3.6 Source Shielding Variations

Multiple sets of measurements were made with the ^{252}Cf neutron source, where varying thicknesses of moderator in the form of HDPE as well as a thickness of lead were used to change the incident neutron and gamma ray field. One thickness of lead was utilized, being 0.25 inches, and 0 – 5 plates of 0.5 inch HDPE were used. To observe the response of an instance of *GoogLeNet* that was trained on ^{252}Cf data the newly measured data sets were left unaltered and classification was performed at 1E-6, 1E-5, and 1E-4 FPRs. The resultant TPRs are shown in Fig. 4.13 with observed values appearing much lower than those shown in Table 4.2. This apparent diminishing performance is due to the classification being performed on all data, not the algorithmically separated neutron data as described in the

methods.

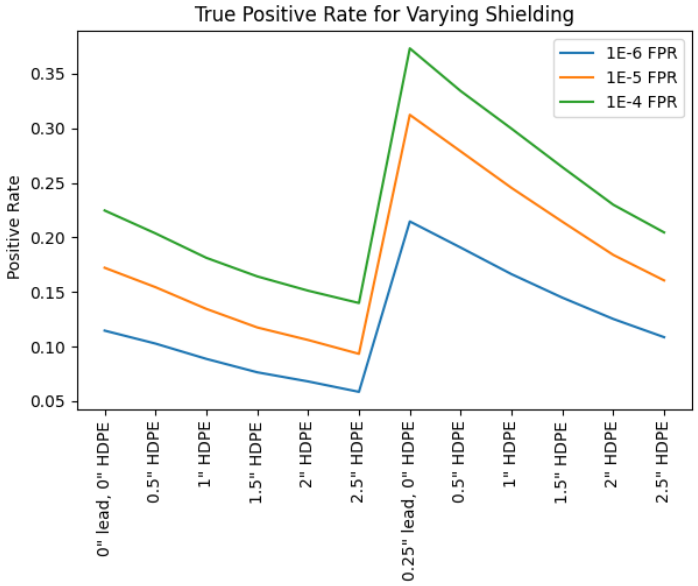


Figure 4.13. TPRs achieved at varying FPRs for ^{252}Cf data collected with varying shielding setups. Each data set was measured with 500 thousand events and with the same digitizer and PMT settings.

Two readily apparent trends are displayed by Fig. 4.13, increases in HDPE thickness reduces the ability to correctly classify events, and the presence of the lead greatly increases the number of neutron classifications. The first trend of decreasing positive rate with increasing thickness of HDPE is explained by the increase in moderation of the neutron fluence from ^{252}Cf. As more HDPE is included the average neutron energy at the detector is reduced and subsequently a greater proportion of the neutron events recorded have lower energies. These lower event energies present a lower neutron classification rate as shown in Figs. 4.9 and 4.12, resulting in a reduction of overall neutron classification rate. Similarly to this, the introduction of lead reduces the photon fluence at the detector face, which

consequently increases the relative abundance of neutron events. This results in an increase in the neutron classification rate of nearly double.

4.3.7 Neutron Intrinsic Efficiency

It is important to recognize when looking at these results that instead of the previously coined TPR this is the positive rate, which is more akin to neutron sensitivity. This is because the classification is being performed on the entire data set, which contains both photons and neutrons. A positive rate of 100 % in this situation would be a poor result considering that all photons in the input were classified as neutrons. A qualitative way of inspecting whether or not photons from the ^{252}Cf data set are being classified as neutrons is to perform TTT PSD on the events classified as neutrons and photons independently and inspect for misclassification that would be caught by TTT PSD. Figure 4.14 shows the resultant PSD for events classified as neutron and photon from the ^{252}Cf data set that was collected with 1 inch of HDPE and 0.25 inches of lead.

The neutron classifications shown in Fig. 4.14 extend to as low as 30 $keVee$ and reaches 3300 $keVee$, with no events overlapping with the region that would be readily identifiable as a photon. Looking at the events classified as photons shows a easily identified neutron population that could be properly identified as neutrons by a traditional PSD method. These false negative events number roughly 1200, which would generate an increase of 1.4 % neutron classifications. Of significance is the ability of the instance of *GoogLeNet* to separate neutron and photon events down into the region where a traditional PSD method would fail, where both populations overlap each other. The support that remains that these are correctly

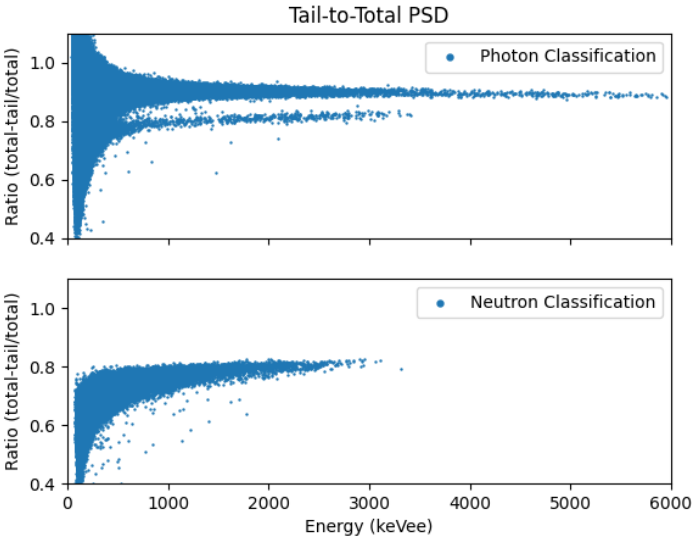


Figure 4.14. TTT PSD of the resultant photon and neutron classified events from 500 thousand events recorded from the ^{252}Cf source. This set was recorded with 0.25 inches of lead and 1 inch of HDPE between the source and detector. Classification was performed at a 1E-6 FPR.

identified neutrons is that the data was collected at a FPR of 1E-6 which was generated by classifying 4.2 million photon events, and that the trend observed in Fig. 4.13 follows what would be expected with impacts of shielding on the neutron and photon fields.

The intrinsic neutron efficiency at the 1E-6 FPR can be obtained in this configuration by using the neutron classification rate in Fig. 4.13 for the 0 lead and 0 HDPE setup. This is then compared to the neutron rate from the ^{252}Cf source, which is 1.4E6 n/s, resulting in this setup with 1411 n/s impinging on the detector. Taking into account the number of events recorded, the duration of the measurement, and the rate limiting of the digitizer the rate of neutron classifications is 176.3 n/s giving an effective intrinsic efficiency of 12.49 % at a 1E-6 FPR.

4.3.8 Visualization

Using the algorithm introduced in Section 2.6 for `GammaNet`, adapted for 1-D inputs, allows for the visualization of which regions of the pulse are most significant for each classification output. These mappings of significance are indicative of where `GoogLeNet` has learned to search for information from the input for producing classifications. For this visualization the instance of `GoogLeNet` trained off of ^{238}Pu^{9}Be data was utilized. An average neutron event was created from the events classified as neutrons from the ^{252}Cf data set at the 1E-6 FPR similarly to Fig. 4.14. This resulted in an average over roughly 130,000 neutron events, with an accompanying 130,000 photon events from the validation data set used to generate an average photon event. The Gradient-weighted Class Activation Mapping (Grad-CAM) results for the average photon and neutron events are shown in Fig. 4.15.

A few significant features in Fig. 4.15 are readily identified. The most significant is the relative width and timing of the main feature between the average photon and average neutron events. The response of `GoogLeNet` for neutron classification shows a peak slightly delayed from the input signal peak, shown in Figs. 4.15a and 4.15c. Similar to this peak is the trough observed for photon classification in response to the average neutron and average photon event in Figs. 4.15b and 4.15d. Both the trough and peak responses possess a larger width for the average neutron event than for the average photon event. This difference in duration aligns perfectly with the intuition behind the mechanism enabling PSD in EJ-301 scintillators, namely that neutron interactions generate a longer signal duration. This increased signal duration produces an increased duration of activation for `GoogLeNet` in

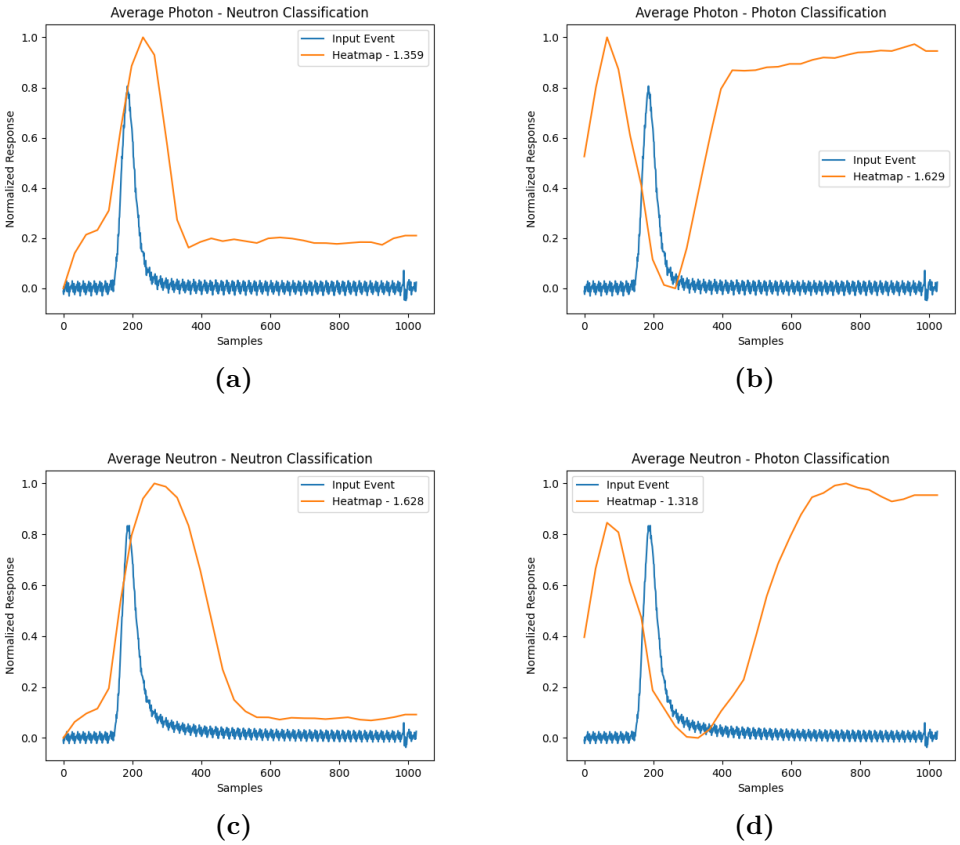


Figure 4.15. Activation mappings from Grad-CAM for the average neutron and photon events. Figs. 4.15a and 4.15b shows the average photon event for neutron and photon classification respectively. Figs. 4.15c and 4.15d shows the average neutron event for neutron and photon classification respectively.

response to a neutron signal. Conversely the average photon event is shorter in duration and therefore the extent of activation in `GoogLeNet` is reduced. In addition to this variation in duration of activation is the delay in the peak of activation observed. The greatest point of differentiation between a photon and neutron signal in EJ-301 occurs when the majority of the scintillation from the fast decay component has already occurred. For TTT PSD this is presented as the time delay used for defining the tail of the signal, and in the activation of

`GoogLeNet` this is presenting as a delay in peak activation.

The second most notable feature is the value of peak activation for the correct classifications and incorrect classifications. Each plot in Fig. 4.15 has the activation normalized for plotting, with the value normalized to presented in the legend of each figure. In the cases where classification is correct the activation maximum 1.629 for Fig. 4.15b and 1.628 for Fig. 4.15c. For the incorrect classification cases, this maximum is reduced to 1.318 for Fig. 4.15d and 1.359 for Fig. 4.15a. This variation in activation maxima is indicative of the response of `GoogLeNet` to the input events, with the average photon producing a lesser response for neutron classification and vice versa. This demonstrates qualitatively that `GoogLeNet` has learned the features indicative of each class of event.

The last feature of interest is the difference in shape for neutron and photon classification activation mapping, with the former being fairly close to an inverse of the latter. From this difference it appears that the peak region, shown in Figs. 4.15a and 4.15c, is the most significant region used for producing a neutron classification. This region providing the most significance for neutron classification is logical given that this is where the transition from the fast decay component to the intermediate decay component for EJ-301 occurs. The greatest activation for photon classification however occurs at the beginning and end of the signal, shown in Figs. 4.15b and 4.15d. This is believed to occur because of a tendency for the photon events to trigger the CAEN-DT5743 digitizer sooner than neutron events, creating a bias for event times as seen in Fig. 4.16. The dependence on the tail portion of the signal for photon classification is less intuitive, but is likely indicative of the importance of a lack of signal being present in the tail, where

for a photon there should be minimal presence of the long decay component in EJ-301.

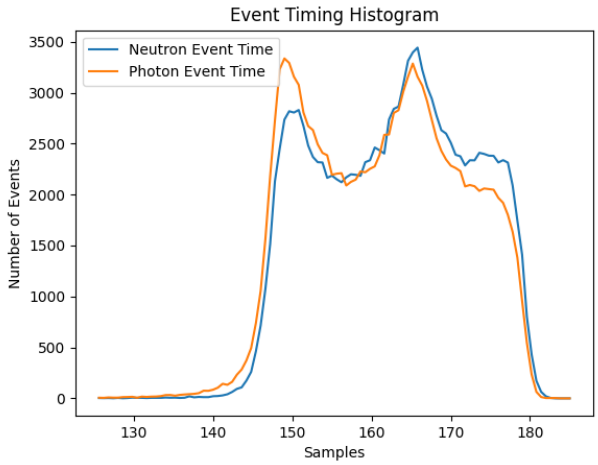


Figure 4.16. Histogram of the calculated event time for the neutron and photon events used to create the average events used in Fig. 4.15.

4.3.9 Signal Aberrations

Two features of note in the average neutron and photon signals from Fig. 4.15 are the presence of very periodic noise and an errant feature at 987 samples. The noise was occurring with a period of 200 MHz or 16 samples, which can be explained by the method of timing in the digitizer. The CAEN DT-5743 digitizer uses a 200 MHz clock signal fed through 16 delay lines [106] to generate an effective 3.2 GHz sampling rate. It is believed that this periodic signal is generated by non-linearities in the delay line circuits producing a very periodic offset in the recorded voltage. When combined with electronic noise for an individual signal this is inconsequential, but when averaging over 130,000 signals the electronic noise is averaged out and the non-linearities become apparent. This is backed up

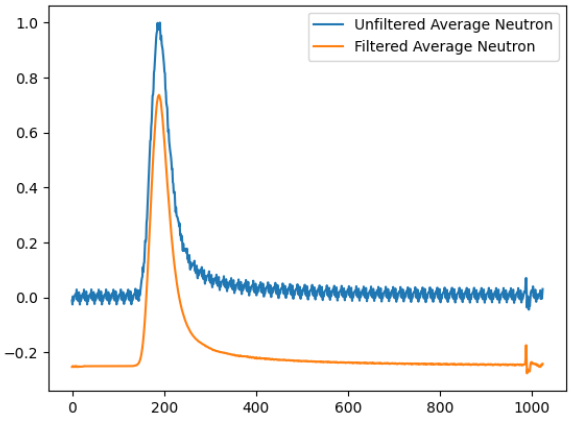


Figure 4.17. The average neutron signal from Fig. 4.15 with the repeated non-linearity from the delay lines removed. An offset between the signals is provided for viewing.

by taking the first 16 signals from the signal, which exists in the baseline of the signal, and subtracting those values from every N/16 set of samples. The result of this subtraction is shown in Fig. 4.17.

The feature observed at 987 samples is thought to be a result of a logic pulse produced in the digitizer for use in external triggering, which occurs with a set delay following a trigger event in the digitizer. Both features, though apparent in these averages, are not easily discernible within normal signals so are more of a curiosity at this point than an issue to be remediated. Additionally, the average signals were generated following normalization of each signal so the absolute scale of these features aren't well represented in Fig. 4.17. Observing the non-linearity without rescaling of the signals shows a peak-to-peak value of 3.9 mV, and the logic pulse artifact has a peak-to-peak value 5.8 mV for an average signal peak-to-peak value of 272.5 mV. At a lower signal amplitude average of 137.2 mV these values are reduced significantly to 0.98 mV and 3.9 mV respectively. The

observed dependence on signal amplitude for the non-linearity is expected, but the dependence displayed by the logic pulse artifact is as of yet not understood.

4.4 Conclusion

An adapted version of *GoogLeNet* was investigated for suitability for application in Particle Identification (PId) for use with multiple neutron sources and at varying levels of specificity. The impact of PMT bias on the ability of *GoogLeNet* to train on and identify data collected from a DT neutron source was investigated. It was found that the best performance was found at 1500 V with marked depreciation in performance when signal clipping began. This lines up with expectations as the FoM is greatest in the range of 1450-1500 V producing the most robust training data in this range.

The ability of an instance of *GoogLeNet* that was trained on one neutron data set to be able to classify data from other neutron sources was investigated. It was found that within one experimental setup there was good generalization to the various neutron sources investigated, but without symmetry in the deficits observed. An extreme asymmetry was observed when training and classifying off of the data collected in different experiments. The application of transfer learning greatly made up for these asymmetries. Additional attention to the differences in neutron spectra of each source accounted for the majority of the remaining inconsistencies. Notable differences in the energy dependence of the classification capabilities were observed for instances of *GoogLeNet* trained on the various neutron sources.

It was found that when performing classification with an instance of *GoogLeNet*

trained on one neutron source that the changes in classification performance followed changes in the neutron spectrum. Variations in neutron moderation and photon shielding were represented in exactly the way anticipated by the changes in radiation field. Further qualitative inspection of the shape of TTT PSD for the events classified as neutrons and photons by the *GoogLeNet* instance showed the capability of *GoogLeNet* at PId. It was found that, at the 1E-6 FPR specificity level, that some neutrons were missed that simple PSD methods would have recovered, though this only represents a 1.4 % relative increase. It was also demonstrated that events within the region of overlap for neutron and photon populations in TTT PSD could be separated, down to the minimum energy detected of 30 $keVee$, while maintaining the 1E-6 FPR.

Investigations into the regions of a pulse that produce the largest response in the neural network were conducted by using the Grad-CAM algorithm on average neutron and photon events. The average neutron event was generated from the ^{252}Cf source and the average photon event was generated from the validation photon set, both being averaged over 130,000 signals. Features observed from Grad-CAM aligned with intuition for what comprises the primary features of a photon and neutron detection event in EJ-301 liquid scintillators. During this investigation two anomalies were detected in the average photon and neutron events which are believed to be caused by the circuit design of the CAEN DT-5743 digitizer. These anomalies relate to the delay line design which enables the 3.2 GHz sampling rate and a logic pulse generated on event trigger. These anomalies are not anticipated to impact individual events as they represent a relatively small amplitude and are only apparent when electronic noise is averaged out over multiple samples.

Chapter 5

Discussion and Conclusions

The work presented in this thesis encompasses the development and quantification of performance for machine learning based solutions to two event classification problems. The first problem, outlined in Section 2.2, was the classification of pair production events recorded by the Advanced Energetic Pair Telescope (AdEPT) mission. The AdEPT instrument is a large volume time projection chamber (TPC) designed to perform gamma-ray polarimetry in the medium energy range of 5-200 MeV in a Low Earth Orbit (LEO). During the time of this work the AdEPT instrument was not yet functional so this research was conducted using simulated data and stands as a proof of concept that the technique, with a sufficiently large and labelled training data set, could achieve the requisite background rejection rate. These Monte Carlo simulations were carried out using GEANT4. The simulation parameters, described in Section 2.3, comprised a simplified geometry and radiation spectrum, with realistic outputs consisting of projection images of ionization inside the TPC with additional electronic noise. The simulation data was broken down into a signal set and a background set, with the former consisting of

pair production images overlaid with a Galactic Cosmic Ray (GCR) proton image. The background events were either Compton scatter images, also with an overlay of a GCR proton image, or a GCR proton image.

The control of the content of a simulation afforded by `GEANT4` was crucial in the development of the machine learning solution for event selection onboard AdEPT. The ability to create data sets of signal and background with complete certainty over the content of each enables generation of perfectly labeled training data. This correctly labelled training data is crucial supervised learning, which enabled the application of Convolutional Neural Networks (CNNs) to this problem. The correctly labeled data is also invaluable when determining performance levels of classifiers, where incorrectly labelled data artificially inflates or deflates the accuracy of a classifier. It was found in this work that the greatest performance, when accounting for limitations imposed by satellites, was achieved with a reduced depth implementation of `GoogLeNet` dubbed `GammaNet`. The performance requirements set by the AdEPT mission was to achieve a background rejection rate of 99.99 % while being able to classify the input images within 50 ms. `GammaNet` achieved this background rejection rate while maintaining a range of signal sensitivity dependent on the incident photon energy, with 0.1±0.1% to 17±2% signal sensitivity for energies ranging from 5-250 MeV. It was found that with updated satellite communications rates the onboard background rejection rate could be eased to 99.69 % without loss of science data, which would improve signal sensitivity to a range of 1.1±0.5% to 69±2% over the same energy range. The compute requirements of `GammaNet` to achieve a 50 ms inference time were found to be 1.1 TFLOPS, which is within the compute power of commercially available System on a Chips (SoCs)

that were being investigated for space readiness at the time of this research.

The other major focus of this thesis was the application of CNNs to event selection for advancing Pulse Shape Discrimination (PSD). This work was developed with the EJ-301 and EJ-309 liquid scintillators which enable PSD for neutrons and photons through different pulse time characteristics for each particle. In both scintillators there are three primary decay times for scintillation, 3.16, 32.3, and 270 ns. Neutron interactions are characterized by producing more scintillation light with the longer decay component, whereas photon interactions contain mostly the short decay component. As with the AdEPT implementation, the most significant task in generating a machine learning based solution for event classification was generation of the labelled training and validation data. Initially labelling the data was investigated with Time of Flight (ToF) analysis, which was found to be insufficient when performing Tail-To-Total (TTT) PSD on the separated data. Labelling was then followed up with manually drawing regions around the TTT PSD data to select the photon and neutron populations produced by TTT. This method produced good results but with a degree of variability introduced by manually drawing the regions. This was later improved upon by binning events by energy and then fitting two Gaussian distributions to the TTT results to identify the neutron and photon populations. This method was found to provide the same quality results as manually drawing the regions, but with reduced variability.

To identify a viable machine learning solution multiple architectures were investigated. Of the tested architectures three were CNNs and one was a fully connected design, with each being adapted for inference of 1-D signals. The architecture found to be most successful was the adaptation of `GoogLeNet` explored for

use with AdEPT, with a depth of three inception modules. The performance of `GoogLeNet` was compared to the traditional TTT PSD method and it was found to be limited to a practical minimum False Positive Rate (FPR) of 6.4E-4. At this 6.4E-4 FPR TTT achieves a True Positive Rate (TPR) of 0.78 with a minimum energy of 200 $keVee$ where the photon and neutron populations begin to overlap. These same limits were not found with the `GoogLeNet` implementation, where achieving a 1E-6 FPR was seemingly trivial. At the 1E-6 FPR level `GoogLeNet` achieved a TPR of 0.6917 with a minimum energy of 127 $keVee$.

In addition to comparing `GoogLeNet` to traditional PSD it was important to investigate the impacts of variations in Photomultiplier Tube (PMT) bias, sampling depth, and sampling frequency. It was found that there is generally an increase in classification performance with increased PMT bias up to the point where signal clipping or distortion occurs, where performance then decreases drastically. Additionally, a similar trend is observed for reduced bit depth and sampling frequency, where lowering either of these factors decreases performance. There exists some discrepancies in this trend at bit depths of 7 and at 8x sampling rate reductions. The presence of this deviation in the bit depth trend is believed to be caused by an artifact in the reduction technique, which predominantly impacts photon signals rather than neutrons. The deviation in the trend for sampling rate reductions is believed to be caused by the reduced sampling rate's inadequate representation of the fast component of the signal, which disproportionately impacts photon signals.

Machine learning is often referred to as a black box because of the disconnect between the result and the input. In this research Gradient-weighted Class Activation Mapping (Grad-CAM) was implemented to produce an interpretation of

the key features employed by `GoogLeNet` when performing event classification. For `GammaNet` it was found that the most significant identifier of background events was sparse or discontinuous ionization indicative of GCR protons and small high ionization density tracks indicative of delta rays. For deep inelastic scattering events from GCR protons the secondary particles generated created vertices which were correctly identified as background, likely because of the much higher ionization density than electron positron pair production. For the identification of pair production the apparent features were vertices, long parallel ionization tracks, and when a continuous track exited frame. The application of Grad-CAM to the `GoogLeNet` implementation for neutron and photon classification found similar successes. For neutron classification the apparent focus for `GoogLeNet` was the duration of the signal as indicated by a wider activation region for the average neutron signal compared to the average photon signal. Also noted in the neutron classification case is the activation at the rising edge of the signal being slightly higher for the average photon than the average neutron, likely caused by a slower rise time for the average neutron signal. For photon classification the activation in `GoogLeNet` was focused on the return to baseline and the beginning of the leading edge of the signal.

5.1 Future Work

The largest deficiencies with the research done for `GammaNet` relate to the incomplete status of the detector and satellite. Without a fixed computational budget, data rate availability, and real science data, the possible space of solutions for this problem are too large. The work stands as a proof of concept, but with these

shortcomings resolved a real solution could be developed and potentially one with greater performance than was found in this work. Further to these problems is the possibility to use newer neural network architectures which perform region bounding. With the use of these networks it would be possible to identify the sub-region of the event images that pertain to the pair production event and save only that for telemetry, affording an additional level of data rate reduction. Lastly, once the satellite mission is solidified the data compression will be known, allowing a higher FPR which would greatly increase the signal sensitivity. The use of PNG file format provides a near 16x reduction in file size for the full scale simulation images, and is a lossless compression format.

The research done implementing `GoogLeNet` as a neutron and photon classifier could be improved upon primarily in four ways. The first improvement would be to improve the ToF experimental results or gather neutron data at a pulsed neutron source with the aim of gathering more pure neutron training data. These two efforts would both depend heavily on improved ToF, which could be improved with use of a PMT with a smaller time spread or by operating the PMT at a higher bias to minimize the time spread. In addition, ToF results would likely benefit from a better free field environment, where less scattering and secondary field generation would occur, reducing the rate of false coincidences. Lastly, the use of a pulsed neutron source would minimize the amount of secondary field generation where, if the detector is sufficiently far and the pulse short enough, the secondary field would essentially be the photon background. This improved data set could then be used for training and compared to the results currently presented to discern if more pure data is necessary or beneficial.

Another potential improvement would be the use of an over-voltage protection circuit for use with the McMaster and Bubble Technologies Industries (BTI) experimental setup. This was not implemented in this work so the bias was left at a conservative level for that setup to avoid damage to the electronics. With the use of an over-voltage protection circuit more data could have been gathered and compared to the Lawrence Berkeley National Laboratory (LBNL) data, perhaps providing even better results for the EJ-301 setup. Additional work could have been done to optimize the electronics chain, where improvements to the PMT and digitizer could have reduced noise and improved PSD performance for TTT these could also improve the performance for machine learning. The improvements from better electronics could have lead to a more pure training data set, which relied on TTT PSD, or simply reduced false positive rates during classification by `GoogLeNet`. The last obvious improvement would be to recreate the experiment in reduced bit depth and sampling rates with actual measurements carried out at the desired fidelity. Most apparent of the discrepancies would be the programmatic reduction in bit depth and how it wouldn't correctly represent the non-linearities introduced per bit when discretizing voltage measurements.

A final improvement for both applications discussed in this research would be the implementation of newer neural network architectures. The most performant architectures, at the time of this research, were implemented but progress in the field of machine learning is continuous. A point of comparison is a newly published article for an image classifier that performs roughly 12 % better than the full scale `GoogLeNet` [107]. The draw back of these newer architectures is that there is ever increasing size and complexity to balance improved performance, where this can be

limiting for their usefulness, especially in the case of AdEPT. Further to this is the advent of faster accelerators for machine learning, where the rate of improvement between generations of Graphics Processing Units (GPUs) has been roughly 50 % for Machine Learning (ML) workloads.

Appendix A

Chapter 2 Supplement

A0.1 `GammaNet` Architecture

The architecture used for `GammaNet` is shown schematically in Fig. A1.1 and is comprised of convolution, Rectified Linear Unit (ReLU) [108], maximum or average pooling, Local Response Normalization (LRN) [12, 109], dropout [110], concatenation, inner product, and softmax [111, 12] operations. All of these operations come preprogrammed in `NVCaffe`, a platform for developing and programming Convolutional Neural Networks (CNNs) [42], which was used for the development of `GammaNet`.

A0.2 `GammaNet` Hyperparameters

The hyperparameters used in training `GammaNet` are as follows:

test_iter: 1000

test_interval: 50000

base_lr: 0.0001

display: 1000

max_iter: 10000000

lr_policy: "step"

gamma: 0.96

momentum: 0.9

weight_decay: 0.0002

stepsize: 320000

snapshot: 49000

snapshot_prefix: "/path/to/your_prefered_directory"

solver_mode: GPU

net: "/path/to/your_network.prototxt"

test_initialization: false

average_loss: 40

iter_size: 1

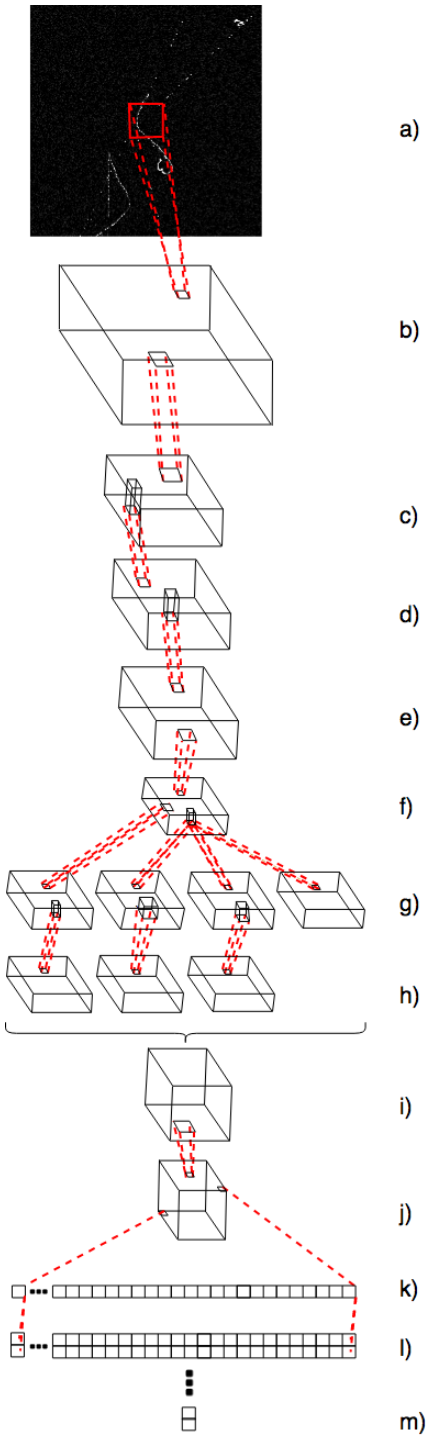


Figure A1.1. Diagram depicting the architecture used for `GammaNet`. All necessary functions are included in the `NVCaffe` library.
a) input to the network of an AdEPT simulation image.
b) first layer made of a 7x7 convolution with a stride of 2. The convolution is followed by a ReLU operation, where all negative values are made to be 0.
c) 3x3 max pooling layer with a stride of 2. The 3x3 max pooling is followed by a LRN operation, where the output is normalized along the depth.
d) 1x1 convolution with stride of 1 followed by a ReLU.
e) 3x3 convolution with a stride of 1 followed by a ReLU and LRN operation.
f) 3x3 max pooling layer with a stride of 2.
g) inception module used by `GoogLeNet` [26], part 1, from left to right: 3 1x1 convolutions of stride 1 and a 3x3 max pooling with a stride of 1.
h) part 2 of inception module, from left to right: 3x3 convolution of stride 1, a 5x5 convolution of stride 1, and a 1x1 convolution of stride 1.
i) concatenation along the depth of the previous 3 operations in h).
j) 7x7 average pooling with a stride of 1. The average pooling is followed by a dropout operation, randomly setting values in the output to 0 with a defined probability.
k) the flattening of j) into a vector.
l) inner product between vector k) and the parameters of l) where there is a set of parameters for each class contained in the output, with two classes in the case of `GammaNet`. This step was carried out in double precision.
m) 2 values output by the softmax operation. The softmax operation was carried out in double precision.

Bibliography

[1] A. Hanu et al. NEUDOSE: A CubeSat Mission for Dosimetry of Charged Particles and Neutrons in Low-Earth Orbit. *Radiation research* 187(1) (2017), 42–49.

[2] M. L. Roush, M. Wilson, and W. F. Hornyak. Pulse shape discrimination. *Nuclear Instruments and Methods* 31(1) (1964), 112–124.

[3] T. Gozani et al. Principles and applications of neutron-based inspection techniques. *Nuclear Instruments & Methods in Physics Research* 261(1-2) (2007), 311–315.

[4] J. Verbus et al. Proposed low-energy absolute calibration of nuclear recoils in a dual-phase noble element TPC using DD neutron scattering kinematics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 851 (2017), 68–81.

[5] D. Lea. Secondary gamma rays excited by the passage of neutrons through matter. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences* 150(871) (1935), 637–668.

[6] M. Heil et al. A 4πBaF2 detector for (n, γ) cross-section measurements at a spallation neutron source. *Nuclear Instruments and Methods in Physics*

Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 459(1-2) (2001), 229–246.

[7] A. L. Samuel. Some studies in machine learning using the game of checkers. *IBM Journal of research and development* 3(3) (1959), 210–229.

[8] F. Rosenblatt. Perceptron simulation experiments. *Proceedings of the IRE* 48(3) (1960), 301–309.

[9] F. Rosenblatt. *Principles of neurodynamics. perceptrons and the theory of brain mechanisms*. Tech. rep. Cornell Aeronautical Lab Inc Buffalo NY, 1961.

[10] D. O. Hebb. *The organization of behavior: A neuropsychological theory*. Psychology Press, 2005.

[11] O. Russakovsky et al. Imagenet large scale visual recognition challenge. *International journal of computer vision* 115(3) (2015), 211–252.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. 2012, 1097–1105.

[13] Y. LeCun et al. Comparison of learning algorithms for handwritten digit recognition. In: *International conference on artificial neural networks*. Vol. 60. Perth, Australia. 1995, 53–60.

[14] S. D. Hunter et al. A pair production telescope for medium-energy gamma-ray polarimetry. *Astroparticle Physics* 59 (2014), 18–28.

[15] G. Wick. Detection of gamma-ray polarization by pair production. *Physical Review* 81(3) (1951), 467.

[16] C. Yang. Possible experimental determination of whether the neutral meson is scalar or pseudoscalar. *Physical Review* 77(5) (1950), 722.

[17] M. Collaboration et al. Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber. *Journal of Instrumentation* 12 (2017).

[18] G. Fortino et al. Digital signal analysis based on convolutional neural networks for active target time projection chambers. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* (2022), 166497.

[19] M. P. Kuchera et al. Machine learning methods for track classification in the AT-TPC. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 940 (2019), 156–167.

[20] P. Abratenko et al. Cosmic ray muon clustering for the MicroBooNE liquid argon time projection chamber using sMask-RCNN. *arXiv preprint arXiv:2201.05705* (2022).

[21] T. Jyoti and V. Bhatnagar. MACHINE LEARNING FOR PARTICLE CLASSFICATION. *Durg V. Rai, Shiva Sharma and Manisha Rastogi* 69 (2019), 79.

[22] M. Collaboration et al. Deep neural network for pixel-level electromagnetic particle identification in the MicroBooNE liquid argon time projection chamber. *Physical Review D* 99(9) (2019), 092001.

[23] R. Acciarri et al. Convolutional neural networks applied to neutrino events in a liquid argon time projection chamber. *Journal of instrumentation* 12(03) (2017), P03011.

[24] P. Ai, D. Wang, G. Huang, and X. Sun. Three-dimensional convolutional neural networks for neutrinoless double-beta decay signal/background discrimination in high-pressure gaseous Time Projection Chamber. *Journal of Instrumentation* 13(08) (2018), P08015.

[25] L. Dominé, K. Terao, D. Collaboration, et al. Scalable deep convolutional neural networks for sparse, locally dense liquid argon time projection chamber data. *Physical Review D* 102(1) (2020), 012005.

[26] C. Szegedy et al. Going deeper with convolutions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015, 1–9.

[27] A. Aurisano et al. A convolutional neural network neutrino event classifier. *Journal of Instrumentation* 11(09) (2016), P09001.

[28] S. Agostinelli et al. GEANT4 a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506(3) (2003), 250–303.

[29] J. Allison et al. Geant4 developments and applications. *IEEE Transactions on nuclear science* 53(1) (2006), 270–278.

[30] R. T. Kouzes et al. Neutron detection gamma ray sensitivity criteria. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 654(1) (2011), 412–416.

[31] *EJ-301/309 Product description*. https://eljentechnology.com/products/liquid-scintillators/ej-301-ej-309. Accessed: 2022-01-01.

[32] F. Brooks. A scintillation counter with neutron and gamma-ray discriminators. *Nuclear Instruments and Methods* 4(3) (1959), 151–163.

[33] R. Owen. The decay times of organic scintillators and their application to the discrimination between particles of differing specific ionization. *IRE Transactions on Nuclear Science* 5(3) (1958), 198–201.

[34] M. Berger, J. Coursey, and M. Zucker. *ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.21)*. en. 1999-01-01 1999.

[35] S. Pozzi, M. Bourne, and S. Clarke. Pulse shape discrimination in the plastic scintillator EJ-299-33. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 723 (2013), 19–23.

[36] A. J. Jinia et al. An Artificial Neural Network System for Photon-Based Active Interrogation Applications. *IEEE Access* 9 (2021), 119871–119880.

[37] A. D. Kaplan et al. A neutron-gamma pulse shape discrimination method based on pure and mixed sources. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 919 (2019), 36–41.

[38] T. Sanderson et al. Machine learning for digital pulse shape discrimination. In: *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)*. IEEE. 2012, 199–202.

[39] R. Wurtz et al. Methodology and performance comparison of statistical learning pulse shape classifiers as demonstrated with organic liquid scintillator. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 901 (2018), 46–55.

[40] T. Tambouratzis, D. Chernikova, and I. Pázsit. A comparison of artificial neural network performance: the case of neutron/gamma pulse shape discrimination. In: *2013 IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA)*. IEEE. 2013, 88–95.

[41] G. Liu, M. Aspinall, X. Ma, and M. Joyce. An investigation of the digital discrimination of neutrons and γ rays with organic scintillation detectors using an artificial neural network. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 607(3) (2009), 620–628.

[42] Y. Jia et al. Caffe: Convolutional architecture for fast feature embedding. In: *Proceedings of the 22nd ACM international conference on Multimedia*. 2014, 675–678.

[43] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015.

[44] M. Tavani et al. The AGILE mission. *Astronomy & Astrophysics* 502(3) (2009), 995–1013.

[45] W. Atwood et al. The Large Area Telescope on the Fermi gamma-ray space telescope mission. *The Astrophysical Journal* 697(2) (2009), 1071.

[46] C. Dermer, R. Schlickeiser, and A. Mastichiadis. High-energy gamma radiation from extragalactic radio sources. *Astronomy and Astrophysics* 256 (1992), L27–L30.

[47] H. Krawczynski. The polarization properties of inverse compton emission and implications for blazar observations with the gems x-ray polarimeter. *The Astrophysical Journal* 744(1) (2011), 30.

[48] A. V. Belikov and D. Hooper. Contribution of inverse Compton scattering to the diffuse extragalactic gamma-ray background from annihilating dark matter. *Physical Review D* 81(4) (2010), 043505.

[49] M. D. Wright. Microbeam radiosurgery: an industrial perspective. *Physica Medica* 31(6) (2015), 601–606.

[50] F. Lei, A. Dean, and G. Hills. Compton polarimetry in gamma-ray astronomy. *Space Science Reviews* 82(3-4) (1997), 309–388.

[51] M. Forot et al. Polarization of the Crab pulsar and nebula as observed by the Integral/IBIS telescope. *The Astrophysical Journal Letters* 688(1) (2008), L29.

[52] J. Simpson. Elemental and isotopic composition of the galactic cosmic rays. *Annual Review of Nuclear and Particle Science* 33(1) (1983), 323–382.

[53] M. Fornasa and M. A. Sanchez-Conde. The nature of the diffuse gamma-ray background. *Physics Reports* 598 (2015), 1–58.

[54] *Askaryan Radio Array*. https://ara.wipac.wisc.edu/science. Accessed: 2022-01-01.

[55] D. Bernard et al. HARPO: a TPC as a gamma-ray telescope and polarimeter. In: *SPIE Astronomical Telescopes+ Instrumentation*. International Society for Optics and Photonics. 2014, 91441M–91441M.

[56] A. Takada et al. Observation of diffuse cosmic and atmospheric gamma rays at balloon altitudes with an electron-tracking Compton camera. *The Astrophysical Journal* 733(1) (2011), 13.

[57] K. Ueno et al. Development of the balloon-borne sub-MeV gamma-ray Compton camera using an electron-tracking gaseous TPC and a scintillation camera. *Journal of Instrumentation* 7(01) (2012), C01088.

[58] C. Martoff et al. Negative ion drift and diffusion in a TPC near 1 bar. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 555(1-2) (2005), 55–58.

[59] C. Meegan et al. The Fermi gamma-ray burst monitor. *The Astrophysical Journal* 702(1) (2009), 791.

[60] R. A. Cameron. Fermi Large Area Telescope operations: progress over 4 years. In: *SPIE Astronomical Telescopes+ Instrumentation*. International Society for Optics and Photonics. 2012, 84481J–84481J.

[61] B. Robinson et al. TeraByte InfraRed Delivery (TBIRD): a demonstration of large-volume direct-to-Earth data transfer from low-Earth orbit. In: *Free-Space Laser Communication and Atmospheric Propagation XXX*. Vol. 10524. International Society for Optics and Photonics. 2018, 105240V.

[62] W. Powell et al. Commercial Off-The-Shelf GPU Qualification for Space Applications (2018).

[63] G. D. Badhwar. The radiation environment in low-Earth orbit. *Radiation research* 148(5s) (1997), S3–S10.

[64] E. R. Benton and E. Benton. Space radiation dosimetry in low-Earth orbit and beyond. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 184(1-2) (2001), 255–294.

[65] D. Zhou, D. O'Sullivan, E. Semones, and W. Heinrich. Radiation field of cosmic rays measured in low Earth orbit by CR-39 detectors. *Advances in Space Research* 37(9) (2006), 1764–1769.

[66] G. Weidenspointner et al. The cosmic diffuse gamma-ray background measured with COMPTEL. In: *AIP Conference Proceedings*. Vol. 510. 1. American Institute of Physics. 2000, 467–470.

[67] R. C. Henry. Diffuse background radiation. *The Astrophysical Journal Letters* 516(2) (1999), L49.

[68] P. Baldi, P. Sadowski, and D. Whiteson. Searching for exotic particles in high-energy physics with deep learning. *Nature communications* 5 (2014), 4308.

[69] P. Baldi et al. Jet substructure classification in high-energy physics with deep neural networks. *Physical Review D* 93(9) (2016), 094034.

[70] A. Aurisano et al. A convolutional neural network neutrino event classifier. *Journal of Instrumentation* 11(09) (2016), P09001.

[71] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

[72] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Thirty-first AAAI conference on artificial intelligence*. 2017.

[73] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, 770–778.

[74] T. Fawcett. An introduction to ROC analysis. *Pattern recognition letters* 27(8) (2006), 861–874.

[75] E. Shibamura et al. Drift velocities of electrons, saturation characteristics of ionization and W-values for conversion electrons in liquid argon, liquid argon-gas mixtures and liquid xenon. *Nuclear Instruments and Methods* 131(2) (1975), 249–258.

[76] R. R. Selvaraju et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017, 618–626.

[77] M. Florek et al. Natural neutron fluence rate and the equivalent dose in localities with different elevation and latitude. *Radiation protection dosimetry* 67(3) (1996), 187–192.

[78] B. Wiegel et al. Spectrometry using the PTB neutron multisphere spectrometer (NEMUS) at flight altitudes and at ground level. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 476(1-2) (2002), 52–57.

[79] J. Scherzinger et al. Tagging fast neutrons from an 241Am/9Be source. *Applied Radiation and Isotopes* 98 (2015), 74–79.

[80] W. Bo et al. Digital pulse shape discrimination methods for n-γ separation in an EJ-301 liquid scintillation detector. *Chinese Physics C* 39(11) (2015), 116201.

[81] R. Lang, D. Masson, J. Pienaar, and S. Röttger. Improved pulse shape discrimination in EJ-301 liquid scintillators. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 856 (2017), 26–31.

[82] M. Nakhostin and P. Walker. Application of digital zero-crossing technique for neutron–gamma discrimination in liquid organic scintillation detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 621(1-3) (2010), 498–501.

[83] D. I. Shippen, M. J. Joyce, and M. D. Aspinall. A wavelet packet transform inspired method of neutron-gamma discrimination. *IEEE Transactions on Nuclear Science* 57(5) (2010), 2617–2624.

[84] D. Cester et al. Pulse shape discrimination with fast digitizers. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 748 (2014), 33–38.

[85] M. Flaska, M. Faisal, D. D. Wentzloff, and S. A. Pozzi. Influence of sampling properties of fast-waveform digitizers on neutron- gamma-ray, pulse-shape discrimination for organic scintillation detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 729 (2013), 456–462.

[86] O. Kalekin, J. Steijger, H. Peek, K. Consortium, et al. Photomultipliers for the KM3NeT optical modules. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 695 (2012), 313–316.

[87] *Photomultiplier Tubes and Assemblies*. https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/etd/High_energy_PMT_TPMZ0003E.pdf. Accessed: 2022-01-01.

[88] M. Safari, F. A. Davani, and H. Afarideh. Differentiation method for localization of Compton edge in organic scintillation detectors. *arXiv preprint arXiv:1610.09185* (2016).

[89] C. Szegedy et al. Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, 2818–2826.

[90] A. Krizhevsky, I. Sutskever, and G. Hinton. *2012 AlexNet*. 2012.

[91] W. Zaremba, I. Sutskever, and O. Vinyals. Recurrent neural network regularization. *arXiv preprint arXiv:1409.2329* (2014).

[92] M. W. Gardner and S. Dorling. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment* 32(14-15) (1998), 2627–2636.

[93] R. Winyard, J. Lutkin, and G. McBeth. Pulse shape discrimination in inorganic and organic scintillators. I. *Nuclear Instruments and Methods* 95(1) (1971), 141–153.

[94] J. N. Mandrekar. Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology* 5(9) (2010), 1315–1316.

[95] T. Tambouratzis, D. Chernikova, and I. Pázsit. A comparison of artificial neural network performance: the case of neutron/gamma pulse shape discrimination. In: *2013 IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA)*. IEEE. 2013, 88–95.

[96] R. Wurtz et al. Methodology and performance comparison of statistical learning pulse shape classifiers as demonstrated with organic liquid scintillator. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 901 (2018), 46–55.

[97] M. J. Balmer, K. A. Gamage, and G. C. Taylor. Comparative analysis of pulse shape discrimination methods in a 6Li loaded plastic scintillator. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 788 (2015), 146–153.

[98] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE* 37(1) (1949), 10–21.

[99] A. Unzueta Mauricio et al. An all-digital associated particle imaging system for the 3D determination of isotopic distributions. *Rev. Sci. Instrum.* 92(6) (June 2021), 063305. ISSN: 0034-6748.

[100] J. N. Mandrekar. Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology* 5(9) (2010), 1315–1316.

[101] M. Kowatari et al. Evaluation of the altitude variation of the cosmic-ray induced environmental neutrons in the Mt. Fuji area. *Journal of nuclear science and technology* 42(6) (2005), 495–502.

[102] H.-C. Shin et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging* 35(5) (2016), 1285–1298.

[103] M. Hussain, J. J. Bird, and D. R. Faria. A study on cnn transfer learning for image classification. In: *UK Workshop on computational Intelligence*. Springer. 2018, 191–202.

[104] L. Stevanato, D. Cester, G. Nebbia, and G. Viesti. Neutron detection in a high gamma-ray background with EJ-301 and EJ-309 liquid scintillators. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 690 (2012), 96–101.

[105] P. Kendall et al. Comparative study of the pulse shape discrimination (PSD) performance of fast neutron detectors. In: *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. IEEE. 2014, 1–6.

[106] *WaveCatcher Family User's Manual*. https://www.hep.ucl.ac.uk/pbt/wikiData/manuals/WaveCatcher/WaveCatcherFamily_V1.2.pdf. Accessed: 2022-01-01.

[107] M. Ding et al. DaViT: Dual Attention Vision Transformers. *arXiv preprint arXiv:2204.03645* (2022).

[108] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010, 807–814.

[109] G. E. Hinton et al. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580* (2012).

[110] N. Srivastava et al. Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research* 15(1) (2014), 1929–1958.

[111] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks* 8(1) (1997), 98–113.