Particle Identification in a Liquid Argon Time Projection Chamber

Fermi National Accelerator Laboratory

Author:

Samuel Borer

Supervisor:

Dr. Jennifer Raaf

August 18, 2016

I. INTRODUCTION

In 1914, J. Chadwick used a magnetic spectrometer and photographic plates to measure the spectrum of electrons from β radioactivity. Contrary to earlier predictions, the data showed a continuous spectrum of β electrons[1]. This was further puzzling when C. D. Ellis and W. A. Wooster found that the mean energy in the β decay was only 1/3 of the energy expected[2]. In December of 1933, W. Pauli released a letter proposing a companion particle that would account for the rest of the energy missing from the decay. He called it the *neutron* because of its need to be electrically neutral and carry a spin 1/2. This name was changed to *neutrino* by Enrico Fermi in 1933 after J. Chadwick discovered a much larger neutral particle which he named the neutron (and stayed to become the neutron known today). There was strong speculation over whether or not these neutral particles could be experimentally confirmed due to their inability to exhibit electrostatic attraction.

It was not until 1945 that B. Pontecorvo noted that theoretically an electron neutrino can impact a Chlorine-37 atom and transform it into an Argon-37 atom[3]. This was not able to be confirmed experimentally at the time. Then in 1953, Cowan and Reines proposed and tested using liquid scintillation to detect neutrinos, and in 1956 announced they had discovered the electron neutrino[4]. A year later, Lee and Yang postulated that neutrinos, unlike their charged particle counterparts, were purely left-handed particles. This meant that weak interactions could violate parity[6].

In 1962, the muon neutrino was detected[6]. A few months later, Maki, Nakagawa and Sakata theorized the idea that these two neutrino types could mix and readily oscillate from one type to another. This could only be theoretically possible if both neutrino flavors had nonzero mass[7]. This was further developed when R. Davis set up his famous experiment in the Homestake mine, utilizing a 100,000 gallon tank of cleaning fluid to study the solar neutrino flux. This experiment measured a value that was 1/3rd the expected value. This later lead to the interpretation of neutrino oscillations. In July 2000, the DONUT Collaboration experimentally confirmed the existence of a third type of neutrino, the tau neutrino.

As the Standard Model developed, it predicted that the three neutrino flavors must be massless. Since there are no right-handed neutrinos or left-handed anti-neutrinos, then it could not pair with the Higgs field to give the particles mass. If neutrinos were to have mass, and no right handed neutrinos or left-handed anti-neutrinos were added to the Standard Model, then the model becomes non-renormalisable. When the Super-Kamiokande collaboration announced they had detected evidence for neutrino mass, it gave a palpable signal that there is physics beyond the Standard Model[8]. This new breaking of the Standard Model has fueled an increased interest in further research in neutrinos. Fermi National Accelerator Laboratory is continuing its long history of world-leading neutrino physics research[9].

II. LIQUID ARGON TIME PROJECTION CHAMBERS

In order to achieve transformative discoveries, Fermi National Accelerator Laboratory is preparing the Deep Underground Neutrino Experiment (DUNE). Neutrinos will be created at Fermilab's Long Baseline Neutrino Facility (LBNF) and sent 800 miles under the Earth's surface to a detector at the Sanford Underground Research Facility in Lead, South Dakota.

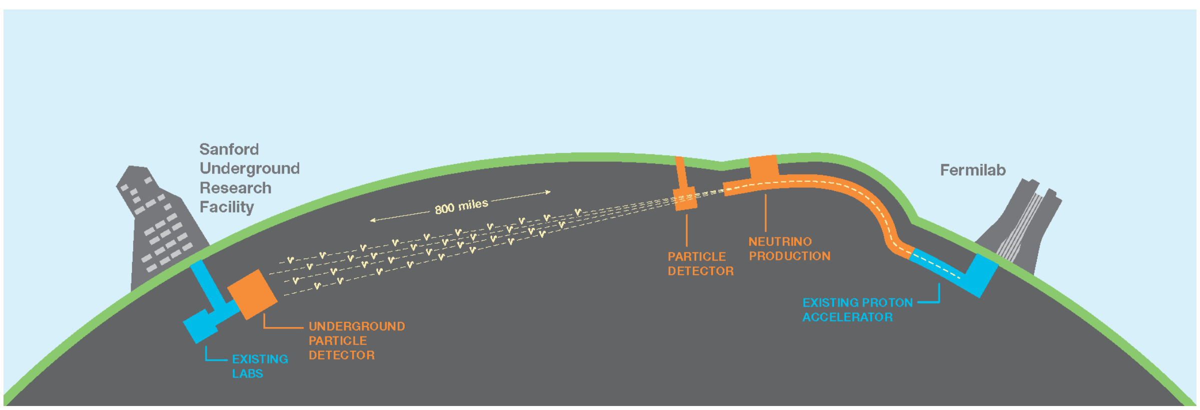

FIG. 1. Long Baseline Neutrino Facility Layout

This detector is a Liquid Argon Time Projection Chamber (LArTPC) detector. LArTPCs' ability for 3D imaging, particle identification capabilities, and calorimetric energy reconstruction makes it a great detector to study neutrinos. Since neutrinos are electrically neutral, they must be indirectly, from the byproducts of their interactions with the detector.

In order to prepare for DUNE's data analysis, set to begin in 2022, an experiment named Liquid Argon Time Projection Chamber In A Testbeam (LArIAT) is being used to characterize LArTPC's. By experimenting with a smaller LArTPC, we can utilize the lessons learned to the much larger DUNE. While LArIAT holds roughly 0.76 tons of argon, DUNE's LArTPC will hold 70,000 tons.

Housed at the Fermilab Test Beam Facility (FTBF), a tertiary beam of pions, protons, muons, kaons, electrons, and their respective antiparticles are sent through a set of wire chambers, bending magnets, cherenkov radiation detectors, and Time-Of-Flight (TOF) counters before entering the LArTPC.

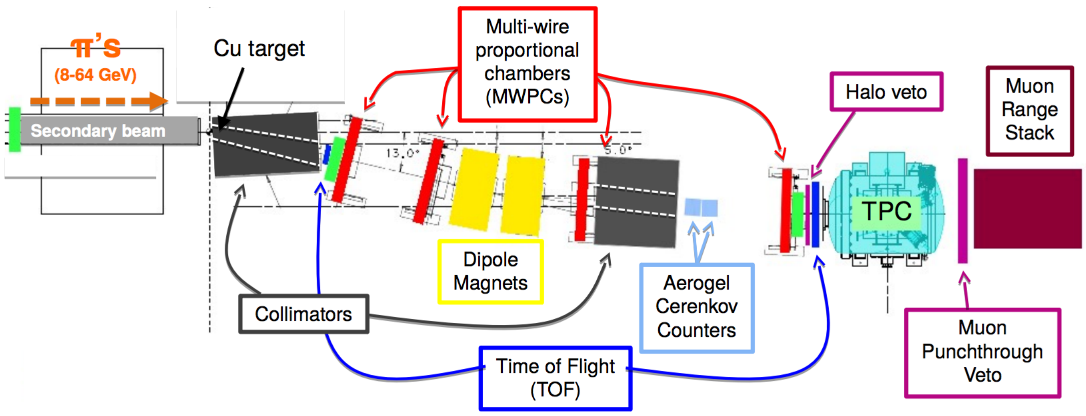

FIG. 2. LArIAT Beamline Diagram

As the particle beam enters the beamline, shown in Fig 2, it will pass through four multi-wire proportional chambers that track the trails of gaseous ionization to detect charged particles and photons, providing their location and trajectory. Using this information along with the difference in angles of the two dipole magnets allows for the reconstruction of particle momentum before entering the time projection chamber.

Additionally there are two time of flight (TOF) counters that allow a determination of the particle speed. Plotting the time of flight versus momentum allows for the separation of deuterons, protons, and kaons, while pions, electrons, and muons still present similar characteristics, shown in Fig 3. The aerogel Cherenkov counters, as well as the muon range stack, allow for the separation of muons and pions.

With these beamline detectors, it is possible to have presumptive particle classification before the particle beam reaches the TPC, however, there is great interest in the ability to classify particles without the beamline instruments. Neutrino experiments do not utilize beamline detection, and can only rely on TPC data. The goal of the project described in this paper is to test and quantify the effectiveness of a Convolutional Neural Network to classify particle events in a LArTPC. By developing machine learning software, a proof of concept can be achieved by training a neural network to identify different particle interactions and characterize each particle.

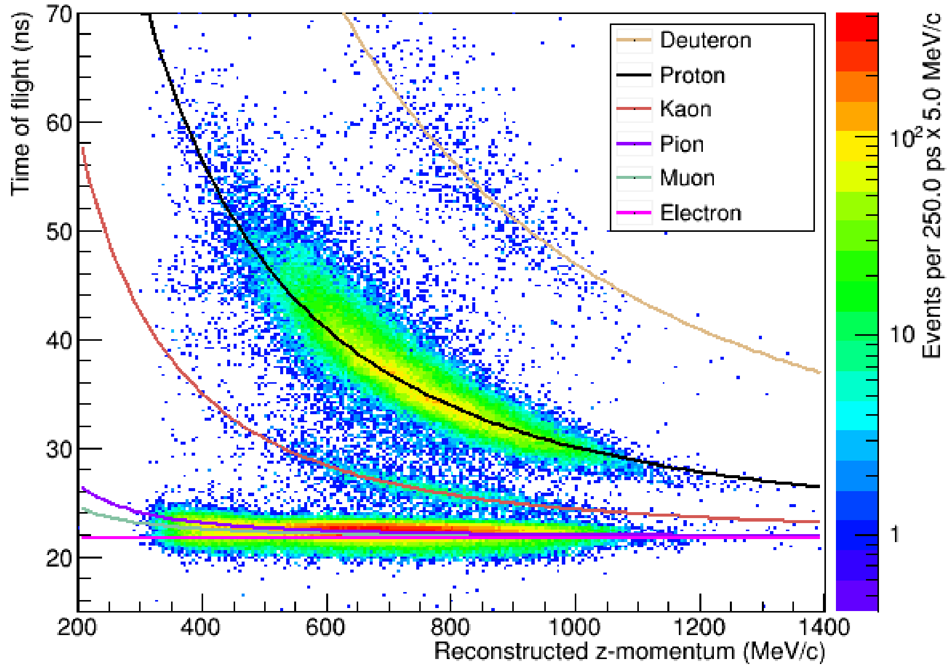

FIG. 3. Time of Flight vs. Momentum

Classifying particle interactions inside a detector is an extremely important need for high energy physics experiments. LArIAT is placed in a well understood beam of charged particles; by utilizing the beamline instruments, classification can be made for particles before they enter the LArTPC. By knowing the classification of the particles entering the TPC, training sets can be created for use by machine learning software. The neural networks trained on LArIAT with beamline information could aid in understanding how neural networks could be used for DUNE data where there is no beamline information.

III. DEEP LEARNING

The traditional neural network used in high energy physics analysis is the multilayer perceptron (MLP) [11]. The MLP consists of an input layer that takes in the available information, which is passed to one or more hidden layers, and then moved to an output layer. The purpose of the MLP is to approximate a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, where n is the dimensionality of the input $\overrightarrow{x}$ and m is the dimensionality of the output $\overrightarrow{f}$. With an MLP each layer is fully connected the the next, shown in Figure 4.

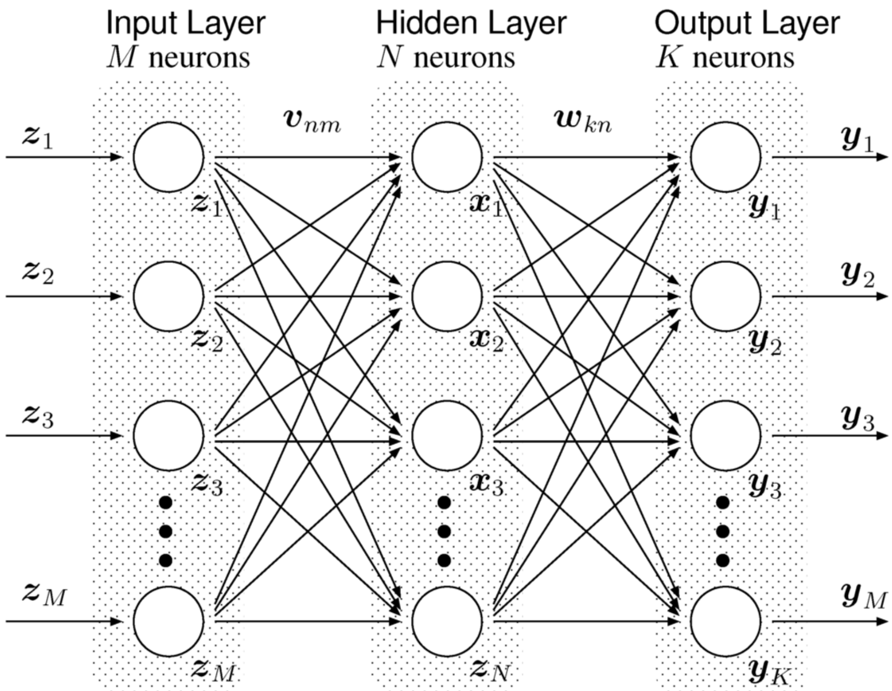

FIG. 4. Graphical representation of a MLP

Each node, with the exception of input nodes, is a neuron with a nonlinear activation function. A nonlinear activation function mimics the frequency of action potentials of biological neurons in the brain. These layers are connected with a certain weight w_{ij}, which describes the amount of influence that the input node will have on the layer. In supervised learning, the MLP is presented with examples where both the input and output are known. This known output is what is known as *ground truth*. A calculated *loss* can be made between the MLP output and ground truth. This difference in error can then be used to calculate a gradient as a function of the weights and bias. This is known as a *back-propagation algorithm*. Using the gradient, changes can be made to the weights and bias to minimize the loss. This, in a sense, is the MLP "learning" to correct for its errors[12]. While MLPs are a great introductory neural network model, there are problems with the MLP which make it ill-suited for the LArTPC experiments. First, when fed a large amount of raw inputs, it scales poorly. For image analysis, the input to a layer is a m x m x r image where m is the height and width of the image and r is the number of channels[13]. If you have a small image of only $32 \times 32 \times 3$ dimensions, a single fully-connected neuron in the first hidden layer of a MLP would have $32 \times 32 \times 3 = 3072$ weights. While this is manageable, if you move to a larger image, the fully connected layer does not scale. If you had a $200 \times 200 \times 3$ image for example, you would have 120,000 weights. The number of nodes needed in a layer for it to approximate complex functions increase beyond appropriate computing power. This can be mitigated with more hidden layers, each with fewer nodes, however it then becomes increasingly harder to train.

IV. CONVOLUTIONAL NEURAL NETWORKS

There have been many advances in the field of computer vision that has allowed for the optimization of architectures with multiple layers. These structures have had increasing success in image recognition and language processing. This has lead to advancements and the development of convolutional neural networks.

Convolutional neural networks were inspired by studies of the visual cortex of animals[14]. A study conducted by Harvard Medical School on macaque and spider monkeys found three classifications of cells in the visual cortex: simple, complex, or hypercomplex. The visual cortex creates sub-regions in the visual field called receptive fields. The simple cells look at a receptive field and are sensitive to straight, edge-like, features, while the complex and hypercomplex cells look at groupings of simple cells. These cells can then cover the entire visual field by tiling together the receptive fields. This is modeled in a convolutional neural network by a series of convolutional layers that are sensitive to a specific set of features from the input image and then pooling layers that will exploit dimensionality reduction and translational invariance. This is widely considered among the most promising architectures to mimick the mammalian visual cortex.

This is the advantage of a convolutional neural network over a traditional MLP. The MLP network's fully connected layers are not ideal for image classification because it does not factor in the spatial structure of the image. Regardless of where the input pixels originate from in the image, the MLP network will treat them with equal standing. A convolutional neural network utilizes spatial architecture to understand the spatial structure of the image, making it ideal for image analysis and classification. A successful CNN will take the local features of an input and convolve receptive fields using filters. The outputs are then sub-sampled and filtered repeatedly, creating a classification output. This particle classification output will serve to help differentiate tracks and clusters of particle interactions in liquid argon TPCs.

Each convolutional layer is characterized by its size and number of maps. The implementation of the CNN is being developed using the TensorFlow framework[15], using the Keras modular library[16]. TensorFlow is an open source software library that utilizes a data flow graphical representation to design multidimensional data arrays to communicate between neural network nodes.

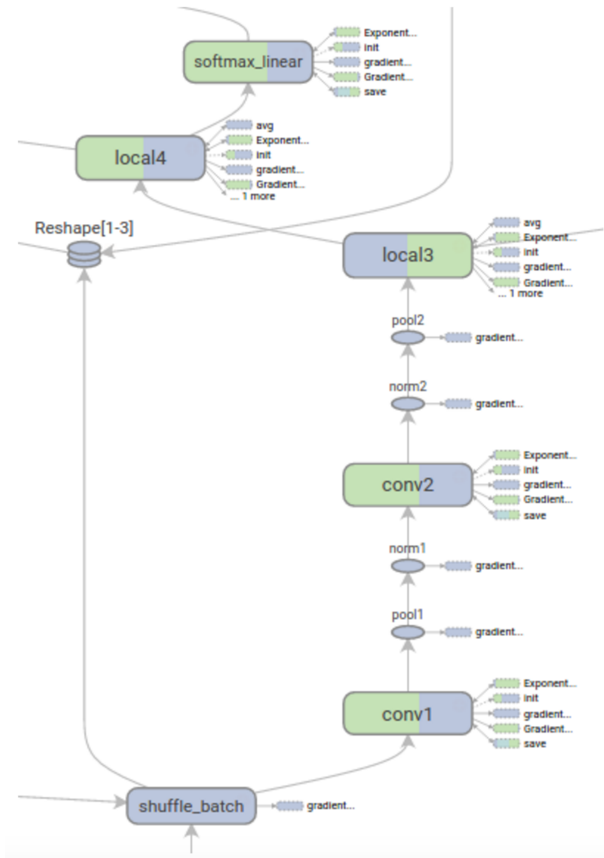

FIG. 5. An example of a TensorFlow graphical model

Keras is a deep learning library that can be run on top of the TensorFlow framework. The library is designed for fast prototyping on both CPU and GPU implementation.

V. CONCLUSION

Our future intent is to build a convolutional neural network that can be trained on known particle data in LArIAT to test the accuracy of particle classification. Using recent advancements in computer science, we will test the viability of using CNN's for data analysis. By creating training sets from LArIAT data, we can train our CNN to correctly identify particle interactions in liquid argon. This architecture can then be used for other liquid argon experiments, such as DUNE and MicroBooNE. Further documentation will follow with relevant findings.

[1] J. Chadwick. "Intensittsverteilung im magnetischen Spektrum von -Strahlen von Radium B+C" *Verhandlungen der Deutschen Physikalischen Gesellschaft*, **16**, 383-391(1914).

[2] C. D. Ellis and W. A. Wooster. *Proc. Royal Soc.*, **A117**, 109(1927).

[3] B. Pontecorvo *Chalk river Report* , **PD-205**, November 1946, unpublished.

[4] C. L. Cowan and F. Reines *Phys. Rev* , **90**, 492(1953).

[5] C. S. Wu; E. Ambler; R. W. Hayward; D. D. Hoppes; and R. P. Hudson "Experimental Test of Parity Conservation in Beta Decay" *Physical Review*, **105**(4): 1413-1415.

[6] G. Danby; J.-M. Gaillard; K. Goulianos; L. M. Lederman; N. B. Mistry; M. Schwartz; J. Steinberger "Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos" *Physical Review Letters*, **9**: 36.

[7] Z. Maki, M. Nakagawa and S. Sakata *Prog. Theo. Physics*, **28**, 247(1962).

[8] T Kajita, for the Super-Kamiokande collaboration "Atmospheric neutrino results from Super-Kamiokande and Kamiokande - Evidence for nu$_m uoscillations" XVIIIInternationalConferenceonNeutrinoPhysicsandAstrophysics, June$1998.

[9] Paolo Lipari. *Introduction to Neutrino Physics.*

[10] W. McCulloch and W. Pitts *A Logical Calculus of Ideas Immanent in Nervous Activity* , Bulletin of Mathematical Biophysics **5**(4): 115-133.

[11] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard et al. *Backpropagation applied to handwritten zip code recognition*, Neural Comput. 1(Dec., 1989) 541-551.

[12] Rosenblatt, Frank. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms.*, Spartan Books, Washington DC, 1961.

[13] Y. LeCun, Y. Bengio and G. Hinton *Deep Learning*, Nature 521 (2015) 436-444.

[14] D. Hubel and T. Wiesel *Receptive fields and functional architecture of monkey striate cortex*, Journal of Physiology 195 (1968) 215-243.

[15] *TensorFlow.* Computer Software. Vers. 0.10.0. Google Inc.

[16] F. Chollet. *Keras.* Computer Software. Vers. 1.0.7. GitHub.