μ/π separation using Convolutional Neural Networks for the MicroBooNE Charged Current Inclusive Cross Section Measurement

Jessica Nicole Esquivel

Bachelor of Science in Electrical Engineering and Applied Physics
St. Mary's University
San Antonio, TX, USA 2011

DISSERTATION

Submitted in partial fulfillment
of the requirements for the degree
Doctor of Philosophy in Physics

-*- DRAFT September 21, 2017 -*-

December, 2017
Syracuse University
Syracuse, New York

Abstract

The purpose of this thesis was to use Convolutional Neural Networks (CNN) to separate $\mu's$ and $\pi's$ for use in increasing the acceptance rate of $\mu's$ below the implemented 75cm track length cut in the Charged Current Inclusive (CC-Inclusive) event selection for the CC-Inclusive Cross-Section Measurement. In doing this, we increase acceptance rate for CC-Inclusive events below a specific momentum range.

Dedication

I dedicate this dissertation to the two important women in my life; My wife and my mom. Both have been there cheering me on giving me strength and love as I worked towards the hardest accomplishment I've ever done.

Jessica Nicole Esquivel

Acknowledgements

Of the many people who deserve thanks, some are particularly prominent, such as my supervisor…

Contents

List of figures

List of tables

"If the don't give you a seat at the table,
bring a folding chair."

— Shirley Chisholm

Chapter 1.

Introduction

This thesis will be a description of work done to further increase effiiciency and purity of the charged current inclusive cross section measurement using the MicroBooNE detector. It will also describe the MicroBooNE detector, what neutrinos are, the low energy excess and how MicroBooNE will contribute to this measurement, the charged current inclusive cross section measuremnt and it's importance as well as convolutional neural networks and how they can be used in μ/π separation.

More introduction

Chapter 2.

Neutrinos

2.1. What are Neutrinos

Neutrinos are one of the fundamental particles which make up the universe. They are also one of the least understood. Neutrinos are not affected by the electromagnetic forces because they do not have electric charge. Neutrinos are affected by a "weak" sub-atomic force of much shorter range than electromagnetism, and are therefore able to pass through great distances in matter without being affected by it. Until the late 90's, neutrinos were thought to have no mass. Due to their mass, neutrinos are also affected by gravity. Neutrinos are created by radioactive decay or nuclear reactions such as the ones that happen in the sun, in nuclear reactors or when cosmic rays hit atoms. There are three types of neutrinos, ν_e, ν_μ and ν_τ which correspond to their charged lepton pairs.

As previously stated, neutrinos are very weakly interacting; in fact, neutrinos can pass unscathed through a wall of lead several hundred light-years thick. Because neutrinos interact so rarely, studying neutrinos requires a massive detector and a powerful neutrino source. With that being said, we can only infer their existence when they interact in a detector. In a collision, distinct charged particles are produced with each type of neutrino. An electron neutrino will create an electron, a muon neutrino will create a muon, and a tau neutrino will create a tau. The track the particle leaves in the detector is how one figures out what type of neutrino interaction was âĂIJseen.âĂİ Liquid Argon Time Projection Chambers are the newest type of detectors being used to study neutrinos due to their excellent imaging and particle identification capabilities.

2.2. History of Neutrinos

The neutrino was first postulated by Wolfgang Pauli in 1931 to explain how beta decay could resolve the conservation of energy, momentum and angular momentum problem. Pauli suggested that this missing energy might be carried off, unseen, by a neutral particle (he called neutron) which was escaping detection. James Chadwick discovered a much heavier nuclear particle in 1932 that he also named neutron, leaving two particles with the same name. Enrico Fermi was the first person to coin the term neutrino (which means little neutral one in latin) in 1933 to fix this confusion. Fermi's paper, which was published in 1934, unified Pauli's neutrino with Paul Dirac's positron and Werner Heisenberg's neutron-proton model and his theory accurately explained many experimentally observed results. Wang Ganchang first proposed the use of beta capture to experimentally detect neutrinos and in 1959 Clyde Cowan and Frederick Reines published their work stating that they had detected the neutrino. The experiment called for antineutrinos created in a nuclear reactor by beta decay that reacted with protons producing neutrons and positrons: $\nu_e + p^+ \rightarrow n^0 + e^+$. Once this happens, the positron finds an electron and they annihilate each other and the resulting gamma rays are detectable. The neutron is detected by neutron capture and the releasing of another gamma ray. In 1962 Leon M. Lenderman, Melvin Schwartz and Jack Steinberger were the first to detect interactions of the muon neutrino. The first detection of the tau neutrino was announced in the summer of 2000 by the DONUT collaboration at Fermilab. In the late 1960s, many experiments found that the number of electron neutrinos arriving from the sun was around 1/3 to 1/2 the number predicted by the Standard Solar Model. This became known as the solar neutrino problem and remained unresolved for around thirty years. This problem was resolved by the discovery of neutrino oscillation and mass. [**?**]

2.3. Neutrino Oscillations

Neutrino oscillation was first predicted by Bruno Pontecorvo. It describes the phenomenon of a neutrino created with a specific lepton flavour (electron, muon or tau) that is later measured to have a different flavour. Neutrino oscillation is important theoretically and experimentally due to the fact that this observation implies that the neutrino has a non-zero mass, which is not part of the original Standard Model of particle physics. [**?**]

2.4. Solar Oscillations and the Solar Neutrino Problem

The solar neutrino flux derived from Bahcall's Standard Solar Model is shown in figure 2.1. Nuclear fusion and decay processes produce an abundant amount of neutrinos. The standard solar model predicts that these reactions produce several groups of neutrinos, each with differing fluxes and energy spectra. The figure also shows the ranges of detection of existing solar neutrino experiments in different shades of blue to illustrate that they sample different portions of the solar neutrino energy spectrum. Three of these experiments, plus a new one, are discussed below.

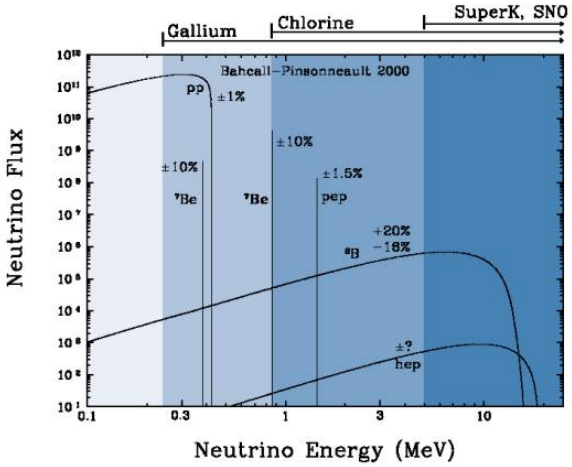


Figure 2.1.: The Standard Solar Model

Since neutrinos rarely interact with matter, they pass through the sun and the earth undetected. About 65-billion neutrinos from the sun stream through every square centimeter on the Earth every second, yet we are oblivious to their passage in our every-day lives. [**?**]

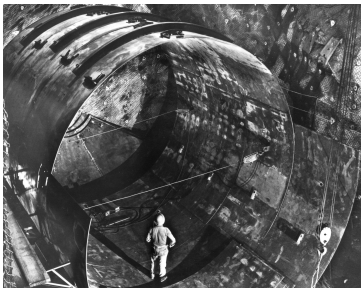
The first experiment to detect the effects of neutrino oscillation was the Ray Davis's Homestake Experiment. The detector was stationed in the Homestake Gold Mine in Lead, South Dakota. It was 1,478 meters underground and was 380 m^3. The detector was filled with perchloroethylene. Perchloroethylene was chosen because of it's high concentrations of chlorine. When an ν_e interaced with chlorine-37 atom, the atom would transform to argon-37 which was then extracted and counted. The neutrino capture reaction is shown in equation 2.1. Davis observed a deficit of about 1/3 the flux of solar neutrinos that was predicted by Bahcall's Standard Solar Model.

The unexplained difference between the measured solar neutrino flux and model predictions lead to the Solar Neutrino Problem. [**?**]

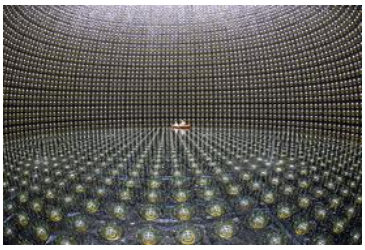
$$\nu_e + {}^{37}Cl \rightarrow {}^{37}Ar + e^- \tag{2.1}$$

While it is now known that the Homestake Experiment detected neutrinos, some physicist were weary of the results. Conclusive evidence of the Solar Neutrino Problem was provided by the Kamiokande-II experiment, a water Cherenkov detector with a low enough energy threshold to detect neutrinos through neutrino-electron elastic scattering. In the elastic scattering interaction the electrons coming out of the point of reaction strongly point in the direction that the neutrino was traveling, away from the sun. While the neutrinos observed in Kamiokande-II were clearly from the sun, there was still a discrepancy between Kamiokande-II and Homestake; The Kamiokande-II experiment measured about 1/2 the predicted flux, rather than the 1/3 that the Homestake Experiment saw.

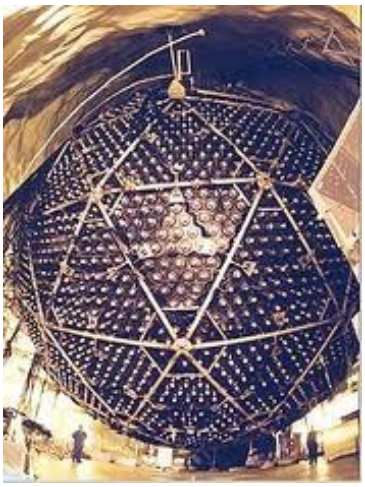
The solution to the solar neutrino problem was finally experimentally determined by the Sudbury Neutrino Observatory(SNO). The Ray Davis's Homestake Experiment was only sensitive to electron neutrinos, and the Kamiokande-II Experiment was dominated by the electron neutrino signal. The SNO experiment had the capability to see all three neutrino flavours. Because of this, it was possible to measure the electron neutrinos and total neutrino flux. The experiment demonstrated that the deficit was due to the MSW effect, the conversion of electron neutrinos from their pure flavour state into the second neutrino mass eigenstate as they passed through a resonance due to the changing density of the sun. The resonance is energy dependent, and is visible near 2MeV. The water cherenkov detectors only detect neutrinos above about 5MeV, while the radiochemical experiments were sensitive to lower energy (0.8MeV for chlorine, 0.2MeV for gallium), and this turned out to be the source of the difference in the observed neutrino rates at the two types of experiments. Figure 2.2 shows Homestake, Kamiokande-II and SNO experiments.



(a) Ray Davis's Homestake Experiment



(b) Kamiokande Experiment



(c) SNO Experiment

Figure 2.2.: Solar Neutrino Experiments

2.4.1. MSW Effect

The Mikheyev-Smirnov-WOlfenshein effect is a process which acts to modify neutrino oscillations in matter. The presence of electrons in matter changes the energy levels of the mass eigenstates of neutrinos due to charged current coherent forward scattering of the electron neutrinos. This coherent forward scattering is similar to the electromagnetic process with respect to the refractive index of light in a medium. Because of this MSW Effect, neutrinos in vacuum have a different effective mass than neutrinos in matter and because neutrino oscillations depend on the squared mass difference of the neutrinos, the neutrino oscillations are different in matter than in vacuum. This effect is important at the sun where electron neutrinos are produced. The neutrinos of high energy leaving the sun are in a vacuum propagation eigenstate ν_2 that has a very small overlap with the electron neutrino $\nu_e = \nu_1 cos(\theta) + \nu_2 sin(\theta)$ seen by the charged current reactions in Kamiokande-II and SNO. The discrepancy of the deficit between SNO, Kamiokande-II and Homestake is due to the energy of the solar neutrinos. The MSW effect "turns on" at about 2MeV and at lower energies, this MSW effect is negligible. [**?**]

2.5. Atmospheric Oscillations and the Atmospheric Neutrino Anomaly

Atmospheric neutrinos are neutrinos that stem from the decay hadrons coming from primary cosmic rays. The dominant part of the decay chain is shown in equations 2.2 and 2.3

$$\pi^+ \rightarrow \mu^+ \nu_\mu \mu^+ \rightarrow e^+ \nu_e \overline{\nu_\mu} \tag{2.2}$$

$$\pi^- \rightarrow \mu^- \overline{\nu_\mu} \mu^- \rightarrow e^- \overline{\nu_e} \nu_\mu \tag{2.3}$$

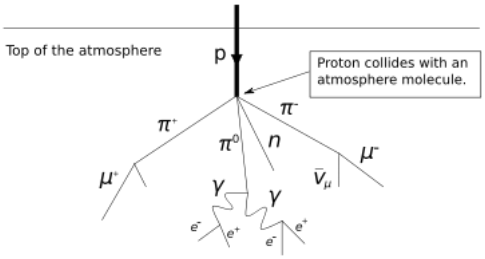


Figure 2.3.: Cosmic Ray Shower

Figure 2.3 shows the cosmic ray shower. In general, these neutrinos have energies from 1GeV to 100s of GeV and the ratio of ν_μs to ν_es equals to 2 (see equation 2.4)

$$R = \frac{(\nu_\mu + \overline{\nu_\mu})}{(\nu_e + \overline{\nu_e})} \tag{2.4}$$

There have been two types of detectors used to study atmospheric neutrinos: Water Cherenkov detectors and tracking calorimeters. Super-Kamiokande is the detector we will focus on. These atmospheric detector experiments measure the ratio of ν_μ to ν_e. They also measure the zenith angle distribution of the neutrinos. These experiments report a double ratio (shown in equation 2.5). This double ratio is the ratio measured in the detector to the ratio thats expected which is 2. If the double ratio equals to 1, the data agrees with the prediction. Various measurements from multiple experiments are shown in figure 2.4. Except for Frejus, all R measurements are less than 1. This discrepancy between the predicted R and the measured R became known as the

Atmospheric Neutrino Anomaly.

$$R = \frac{(N_\mu / N_e)_{DATA}}{(N_\mu / N_e)_{SIM}} \tag{2.5}$$

Experiment	Type of experiment	R
Super-Kamiokande	Water Cerenkov	0.675 ± 0.085
Soudan2	Iron Tracking Calorimeter	0.69 ± 0.13
IMB	Water Cerenkov	0.54 ± 0.12
Kamiokande	Water Cerenkov	0.60 ± 0.07
Frejus	Iron Tracking Calorimeter	1.0 ± 0.15

Figure 2.4.: Measurments of the double ratio for various atmospheric neutrino experiments

Kamiokande-II has the the capability of measuring the direction of the incoming neutrinos. The expectation of atmospheric neutrino detection is that the flux be isotropic due to the fact that atmospheric neutrinos can reach the detector from all directions. Kamiokande-II noticed that muon-like data did not agree well with this expectation. At low energies approximately half of the ν_μ are missing over the full range of zenith angles. At high energies the number of ν_μ coming down from above the detector seems to agree with expectation, but half of the same ν_μ coming up from below the detector are missing. This anomaly can be easily explained by neutrino flavour oscillations. Due to the fact that the neutrino travels less distance coming straight down into the detector (about 15km) than coming up from the bottom of the detector(13000km) changes the probability of oscillation. The probability of oscillation for the muon neutrinos coming down into the detector is roughly zero, whereas for neutrinos coming up, the oscillation probability is $sin^2(2\theta)$. Also, that fact that the electron-like events are not reduced, but the muon-like events are, suggests that the oscillation mode for atmospheric neutrinos is $\nu_\mu \rightarrow \nu_\tau$. Both the solar and atmospheric neutrino problems can be explained by neutrino oscillation so its fitting to derive this phenomenon mathematically. In the next two sections, two flavour and three flavour neutrino oscillation derivations will be explained.

2.6. Two Flavour Neutrino Oscillation Formulation

The flavour eigenstates can oscillate between eachother because they are composed of an add-mixture of mass eigenstates(ν_1,ν_2). Figure 2.5 shows the mass and flavour eigenstates rotated by an angle θ which is the mixing angle.

In matrix form the wavefunctions are:

$$\begin{pmatrix} \nu_\mu \\ \nu_e \end{pmatrix} = \begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix} * \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \tag{2.6}$$

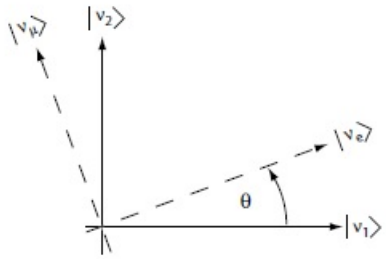


Figure 2.5.: The flavour eigenstates are rotated by an angle θ with respect to the mass eigenstates

Applying the time evolution operator to ν_μ:

$$|\nu_\mu(t)> = -sin\theta|\nu_1> e^{-i\frac{E_1 t}{\hbar}} + cos\theta|\nu_2> e^{-i\frac{E_2 t}{\hbar}} \tag{2.7}$$

where $E_1 = \sqrt{p^2c^2 + m_1^2c^4}$ and $E_2 = \sqrt{p^2c^2 + m_2^2c^4}$ and $p_1 = p_2$. For the time being, let us assume $\hbar = c = 1$. With this assumption: $E_1 = \sqrt{p^2 + m_1^2}$ and $E_2 = \sqrt{p^2 + m_2^2}$. The next modifications is to assume neutrinos are relativistic:

$$\gamma = \frac{E}{m_o^2c^2} = \frac{\sqrt{p^2c^2 + m_o^2c^4}}{m_oc^2} \gg 1 \tag{2.8}$$

because of this,

$$p \gg m_o \tag{2.9}$$

$$E = \sqrt{p^2 + m_o^2} = p\sqrt{1 + m_o^2/p^2} \simeq p + \frac{1}{2}\frac{m_o^2}{p} \tag{2.10}$$

where the binomial expansion is used. Now E_1 and E_2 can be written as:

$$E_1 \simeq p + \frac{1}{2}\frac{m_1^2}{p} \text{ and } E_2 \simeq p + \frac{1}{2}\frac{m_2^2}{p} \tag{2.11}$$

Now applying all these assumptions back into equation 2.7 gives us:

$$|\nu_\mu(t)>= -sin\theta|\nu_1> e^{-i\left(p+\frac{1}{2}\frac{m_1^2}{p}\right)t} + cos\theta|\nu_2> e^{-i\left(p+\frac{1}{2}\frac{m_2^2}{p}\right)t} \tag{2.12}$$

$$|\nu_\mu(t)>= e^{-i\left(p+\frac{1}{2}\frac{m_1^2-m_2^2}{p}\right)t}(-sin\theta|\nu_1> + cos\theta|\nu_2>) \tag{2.13}$$

Substituting $\Delta m^2 = m_1^2 - m^2 2$ and $t = \frac{x}{c} = x$ and $e^{-iz} = e^{-i\left(p+\frac{1}{2}\frac{m_1^2}{p}\right)t}$ gives us:

$$|\nu_\mu(t)>= e^{-iz}\left(-sin\theta|\nu_1> + cos\theta|\nu_2> e^{+ix\left(\frac{1}{2}\frac{\Delta m^2}{p}\right)}\right) \tag{2.14}$$

Finding the Probability for a $\nu_\mu \to \nu_e$:

$$P(\nu_\mu \to \nu_e) = |<\nu_e|\nu_\mu(t)>|^2 \tag{2.15}$$

Remembering that $<\nu_i|\nu_j>= \delta_{ij}$

$$<\nu_e|\nu_\mu(t)>= e^{-iz}\left(-sin\theta cos\theta + sin\theta cos\theta e^{\frac{i\Delta m^2 x}{p}}\right) \tag{2.16}$$

Taking the absolute value squared gives us:

$$P(\nu_\mu \to \nu_e) = |<\nu_e|\nu_\mu(t)>|^2 = e^{+iz}e^{-iz}sin^2\theta cos^2\theta\left(-1+e^{\frac{i\Delta m^2 x}{p}}\right)\left(-1+e^{\frac{i\Delta m^2 x}{p}}\right) \tag{2.17}$$

Since the neutrino is relativistic we can set $p = E_\nu$ and change $x = L$. Also recognizing the trigonometric relation $(1 - cos2\theta)/2 = sin^2\theta$ the above equation becomes:

$$P(\nu_\mu \to \nu_e) = sin^2 2\theta sin^2\left(\frac{\Delta m^2 L}{4E_\nu}\right) \tag{2.18}$$

All that's left to do now is re-introduce $\hbar$ and c doing this we get:

$$P_{\nu_\mu \to \nu_e}(L, E) = sin^2 2\theta sin^2\left(1.27\Delta m^2 \frac{L}{E_\nu}\right) \tag{2.19}$$

This equations has three important variables.

- The angle θ: This angle, as mentioned before, is called the mixing angle. It defines the difference between the flavour and the mass eigenstates. When $\theta = 0$ the mass and flavour eigenstates are identical and now oscillations occur.
- The mass squared difference, Δm^2: Again $\Delta m^2 = m_1^2 - m_2^2$. The reason this is an important variable is because it implies that for neutrinos to oscillate, neutrinos must have mass. Furthermore, the mass squared difference also tells us that the neutrino mass eigenstates must be different.
- L/E: This is the variable that is of most interest to experimental physicists due to the fact that it is the variable that we set. L is the distance between the source and detector and E is the energy of the neutrino. For a given Δm^2, the probability of oscillation changes with respect to L/E.

2.7. Three Flavour Neutrino Oscillation Formulation

Seeing the quantum mechanics involved in deriving the probability of a two flavour neutrino oscillation, it is now possible to formulate the three flavour neutrino oscillation. The three flavour neutrino oscillation formulation begins similarly to the two flavour, but there is the Pontecorvo-Maki-Nakagawa-Sakata matrix (PMNS) instead of the 2X2 matrix in the previous section. The PMNS matrix is show below:

$$\begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} * \begin{pmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.20}$$

where $c_{ij} = cos\theta_{ij}$ and $s_{ij} = sin\theta_{ij}$

Following the same steps as before we get:

$$P_{\alpha\to\beta} = \delta_{\alpha\beta} - 4\Sigma Re(U^*_{\alpha i}U_{\beta i}U\alpha jU^*_{\beta j})sin^2\left(\frac{\Delta m^2_{ij}L}{4E}\right) 2\Sigma Im(U^*_{\alpha i}U_{\beta i}U\alpha jU^*_{\beta j})sin\left(\frac{\Delta m^2_{ij}L}{2E}\right) \tag{2.21}$$

The main things to notice here are δ_{ij} which is the CP violating term and has not been measured yet, and θ_{13} which has just been measured. CP violation is a violation of the postulated CP-symmetry. CP-symmetry states that the laws of physics should be the same if a particle were to be exchanged with its antiparticle and then if the left hand side of a decay were switched with the right hand side.

Chapter 3.

The MicroBooNE Experiment

The purpose of this chapter is to discuss and understand the details of the MicroBooNE detector. A thorough understanding of MicroBooNE and the technology behind liquid argon time projection chambers is important for understanding results as well as understanding how images were made for use in deep learning efforts that will be outlined in later chapters.

3.1. Introduction

The experimental study of neutrinos is central to FermiLab's scientific program and offers the possibility for exciting discoveries that could reshape our understanding of the universe. Liquid Argon Time Projection Chambers (LArTPCs) are an exciting detector technology that provide excellent imaging and particle identification, and are now being used to study neutrinos. In these detectors, liquid argon (LAr) is the medium used to detect neutrinos. When a neutrino interacts with an argon atom, the charged particles that are produced ionize the LAr as they travel away from the interaction. By placing a uniform electric field throughout the LAr, the ionization is made to drift towards a set of anode planes, which consist of wires spaced very closely together collecting the ionized charge, which is subsequently read out by electronics connected to the anode wires. The collected ionization creates a spatial image of what happened in the detector on each anode plane. The $\nu - Ar$ interaction also produces scintillation light which is collected by photomultiplier tubes (PMTs) which allows the exact time of the neutrino interaction to be determined. The drift time of the ionization relative to the time of the original signal allows the signal to be projected

back along the drift coordinate, hence the name LArTPC. Having very small distances between each wire within an anode plane allows for very fine granularity and detail to be captured, and having multiple wire planes at different angles coupled with exact timing of neutrino interaction, provides independent two-dimensional views that can be combined into a three dimensional picture of the interaction. Once the charge signal is created on the anode planes, software analysis packages identify particles in the detector by using deposited energy on the wires along their track length.

MicroBooNE (Micro Booster Neutrino Experiment) is a 89 T active volume (180 T total mass) LArTPC which is then inserted into a cylindrical crysotat on axis of the Booster Neutrino Beam (BNB) stationed at Fermilab in Batavia, Illinois. The main components of MicroBooNE will be detailed in the upcoming sections. MicroBooNE is also an R&D detector that can be scaled up to a significanlty larger size, such as Deep Underground Neutrino Experiment (DUNE) which is roughly 40 kT compared to MicroBooNE at 180T [1]. Understanding LArTPC technology and detector physics is necessary to build a LArTPC the size of DUNE, and MicroBooNE has made many advances in developing this technology [**?**] [**?**].

3.2. Time Projection Chamber

MicroBooNE's Time Projection Chamber (TPC) is 10.3 m long (beamline direction), 2.3 m high and 2.5 m wide (which corresponds to the drift distance). The TPC is shown in figure **??**. MicroBooNE is the largest LArTPC currently running in the world [**?**]. This LArTPC has 3 wire planes: 1 plane that collects the ionization in the wires and is 0° to the virtical with 3456 wires spaced 3 mm apart, and 2 planes where the ionization drifts passed and induces a signal at $+/-60^{\}circ}$ to the vertical each with 2400 wires also spaced 3 mm apart. Each plane has a spacing also of 3 mm from eachother. The wires are then connected to detector specific circuit boards (ASICS) that are submered and operate inside liquid argon. The first two planes are the induction planes and the last is the collection. The electric field of the TPC is created using 64 stainless steel tubes shaped into rectangules around the TPC and held in place by G10 to form a field cage. The cathode is charged at a high voltage of -70 kV and this voltage is stepped down across the field cage tubes using a voltage divider chain with an equivalent resitance of 240 MΩbetween the tubes. The field cage tubes are separated by 4 cm from center to center.

3.3. Light Collection System

The light collection system is a crucial part for 3D reconstruction of particle interactions in the LArTPC. It is possible to reconstruct interactions using just the wire signals, but without the initial timing (t0) of an event, it is impossible to position the event along the drift direction. When a particle interaction occurs, the scintillation light created propogates within nanoseconds to the light collection system compared to the milliseconds it takes the ionized electrons from the interaction to reach the anode wire planes. Therefore we can precisely know where along the drift direction the particle interaction first took place. The scintillation light is also localized, so combining the PMT information with the wire plane information allows for cosmic background rejection happening outside the beam timing window.

The light collection system is made up of 32 Hammamatsu R5912-02mod cryogenic PMTs with a diameter of 8-inches. The PMTs are are located behind the 3 wire anode planes and provides 0.85% photocathode coverage. Each PMT has an acrylic plate mounted in front of it that is coated with a wave-length shifting material called TPB. The acrylic plates take in the scintillation light, at 128 nm, and re-emits it visible wavelengths visible to the PMTs, with a peak at 425 nm.

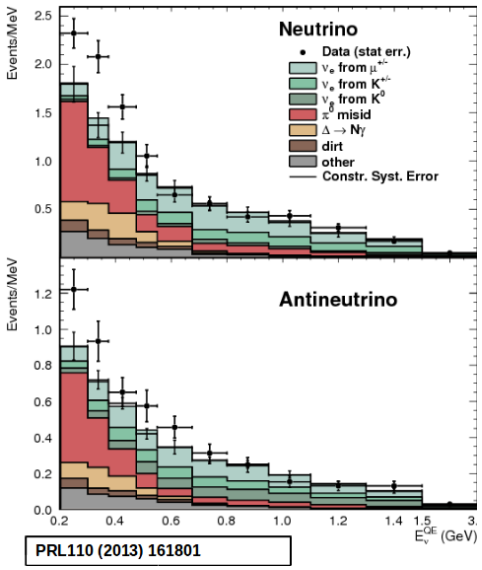
3.4. Electronics System

More MicroBooNE stuff. Possibly talk about rack protection system work I did i.e circuit board soudering? Talk about deconvolution paper Adam and I wrote a while back?

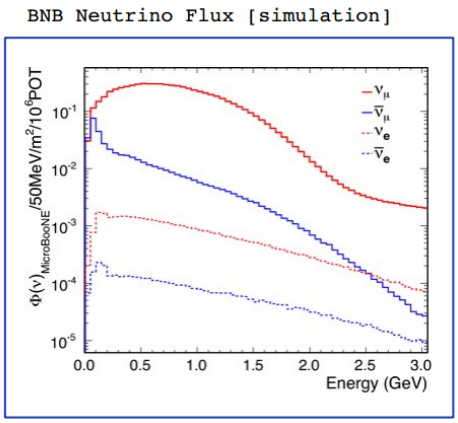
Chapter 4.

The Booster Neutrino Beam

The MicroBooNE detector is stationed at FNAL where it receives neutrinos from both the Booster Neutrino Beam (BNB) and Neutrinos from the Main Injector (NuMI) beams. MicroBooNE is on-axis for the BNB and off-axis by 135 mrad for NuMI. The BNB is a very pure ν_μ beam, with only 0.6% contamination from $\nu_e s$. The energy also peaks around 700 MeV which is desired based on the probability of oscillation equation which depends on the the value of L/E, where L is the distance of the detector from the neutrino beam and E is the energy of the neutrino beam. L/E was chosen to increase the probability of seeing neutrino oscillations in the MiniBooNE Low Energy Excess (LEE) range based on the probability of oscillation equation is $P_{\nu_\mu \to \nu_e}(L, E) = \sin^2 2\theta \sin^2\left(1.27\Delta m^2 \frac{L}{E_\nu}\right)$. The Short Baseline Neutrino (SBN) Program will further study non-standard neutrino oscillations and is composed of 3 LArTPCs including the MicroBooNE experiement [**?**]. MicroBooNE's second goal is to gain a cross section measurement on $\nu - Ar$. MicroBooNE's neutrino cross-section program will be the first $\nu - Ar$ cross-section in the 1 GeV energy range and one of only a few cross-section measurements of $\nu - Ar$ in the world. This has significance in part because of the large scale LArTPCs slated for construction over the next decade [1] [**?**].



(a) Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories



(b) Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories

Figure 4.1.: ?? Flux of BNB at FNAL.

...

Chapter 5.

Finding the first Neutrinos in MicroBooNE

Chapter about your neutrinoID here.

More neutrinoID stuff.

Chapter 6.

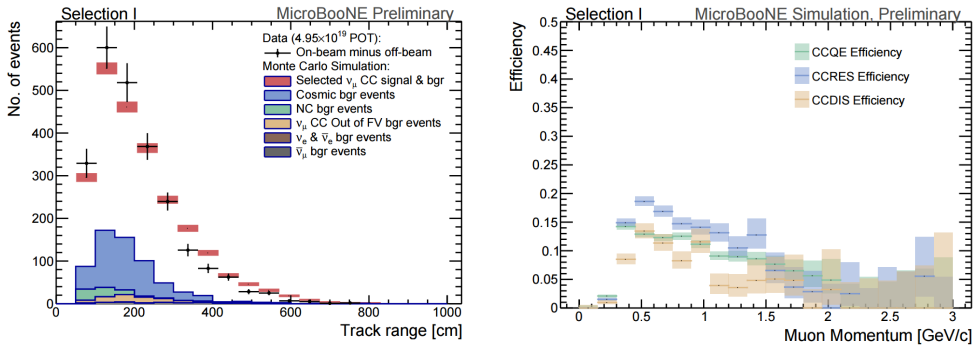
CC-Inclusive Cross Section Selection Filter

One of the cross-section measurements MicroBooNE can make is an inclusive charged-current cross-section measurement (referred to as CC-inclusive). CC-inclusive events consist of a neutrino exchanging a $W^{\pm}$ boson with an argon atom, producing a charged lepton and any number of other final state particles. In MicroBooNE's case, a CC-inclusive event will mostly have a defining muon track coming out of the vertex due to our neutrinos being predominately ν_{μ}s. A cross-section measurement is the energy dependent probability of $\nu - Ar$ interaction in the detector. Cross-sections however are independent of the intensity or focus of the particle beam so they can be compared among different experiments. A background for a CC-inclusive cross-section measurement are the neutral-current events that contain a pion. It is possible to have a neutral current interaction with a $\pi + p$ event signature that looks like a charged current $\mu + p$ event. Reconstruction tools implemented to date don't efficiently separate muons from pions. A common way to separate these two particles species is to implement a track length cut. On average, muons tend to have longer track lengths in LArTPCs so by requiring that the hypothesized lepton be above a threshold track length, it is possible to increase signal to background.

MicroBooNE requires fully automated event reconstruction and selection algorithms for use in the many physics measurements being worked on to date due to the large data rate MicroBooNE receives. Being able to automatically pluck out the neutrino interaction among a sea of cosmics proved to be challenging but was accomplished. MicroBooNE has developed two complementary and preliminary selection algorithms to select charged-current $\nu_{\mu} - Ar$ interactions. Both are fully automated

and cut based. The results below focus on the first selection and the "In-Progress" plots presented on the poster associated with this proceeding will focus on further improving this algorithm using Convolutional Neural Network (CNN) implementations. The full details can be found in MicroBooNE public note [**?**] and for more information of CNN implementation on MicroBooNE data refer to [**?**]. Selection I is based on cuts developed in a MC performance study described in [**?**]. It identifies the muon from a neutrino interaction without biasing towards track multiplicity. To combat cosmic and neutral current background, the analysis is strongly biased towards forward-going long tracks which are contained. This limits phase space and reduces acceptance.

The efficiency and purity are used as performance values of selection I. Efficiency is described as the number of selected true ν_μ CC events divided by the number of expected true ν_μ CC events. The purity is described as the number of selected true ν_μ CC events divided by the sum of itself and all the backgrounds. The efficiency of selection I is 12% and the purity is 39.7%. The poster related to this proceedings will focus on the last cut which requires the longest track to be longer than 75 cm. This cut has a passing rate of 30% w.r.t the previous cut and is implemented in part to separate charged-current events from neutral-current events that mimic our signal. Implementing a CNN for $\mu - \pi$ separation picks out differences in these two particles that are track range independent therefore eliminating the need for the 75 cm track length cut and increase efficiency and passing rate at low muon momentum. Figure **??** shows the track distribution of selection I and the lack of data below the 75 cm track length cut. Figure **??** shows the efficiency of selection I as a function of muon momentum.



(a) Track range distribution of selection I

(b) Selection efficiency as a function of the true muon momentum

Figure 6.1.: ?? Track range distribution for selection I. The track range is defined as the 3D distance between the start and end of the muon candidate track. No data is shown below 75 cm due to the track length cut described previously. ?? Efficiency of the selected events by process quasi-elastic (QE), resonant (RES), and deep-inelastic (DIS). Statistical uncertainty is shown in the bands and the distributions are a function of true muon momentum. The rise of the efficiency between 0 GeV and 0.5 GeV is due to the minimum track length cut and the decreasing efficiency for higher momentum tracks is caused by the containment requirement.

Chapter 7.

The importance of μ/π separation in MicroBooNE

μ/π separation chapter goes here.

More μ/π separation.

Chapter 8.

Convolutional Neural Networks

Image processing is processing of images by using any form of signal processing. Most image processing techniques treat an image like a 2-dimensional signal and apply standard signal processing techniques to it. Examples of image processing include Gaussian smoothing, edge detection, contouring and pattern recognition or high-level image processing like computer vision. By transforming each sample value for a single wire into a pixel value, it is possible to create an 2-D image from each LArTPC anode plane and hence apply image processing techniques. Gaussian smoothing is used to reduce the noise in an image. Edge detection algorithms look for discontinuities in an image and is the fundamental stage in pattern recognition or any higher level image processing, machine learning and computer vision due to the fact that it extracts the important features in an image. Once edge detection is applied to an image, a contouring algorithm can then be applied. A contour in image processing is an outline bounding an object of interest. A contour must be a closed line outlining in this case the edges found during edge detection. Once we have a bounded object, it is then possible to use these extracted features to train a computer vision algorithm like a Convolutional Neural Network (CNN) to learn that these features describe certain objects, and to then use these computer vision algorithms on random images.

When used for image recognition, convolutional neural networks consist of multiple layers of feature maps that extract different information on small portions of the input image. How many layers and feature maps is tunable to increase the matching process. The output of these collections are then tiled so that they overlap to gain a better representation of the original image and allow for translation. To truly understand CNNs, a breakdown of what convolution means with regards to imaging is necessary. Convolution of an image is using a map of some sort to extract features

Figure 8.1.: Applying a feature mask over a set of fashion items to extract necessary information for auto-encoding. Unnecessary information for example color or brand emblems are not saved. This feature map is an edge detection mask that leaves only shape information which helps to distinguish between different types of clothes.

from the input image by matrix multiplication. The map is multiplied to a small image patch and this is then saved. This is done until the whole image is processed and what is left is a feature map that has the important features extracted. An example of convolution for the process of object identification is shown in figure 8.1. In this figure you can see how an edge detection feature map is used to save only necessary information for recognizing different types of clothes. You can also see by having multiple feature maps you can get more detail or less detail from an image which can then simplify or complicate the object recognition task. Being able to distinguish between a shirt or a leg garment is as much information you want, having a feature map that extracts outline edge or shape information would be all that you need. But if instead you wanted to distinguish between a formal cocktail dress or a summer dress, more information would need to be saved equating to many more feature maps for one image. Rather than trying to come up with a scheme of feature maps to run over the image by hand, CNNs do this automatically. CNNs take input parameters, for example number of layers, number of units per layers, number of connections per unit, and uses these to create the feature maps. The layers build upon each-other, for example if we were creating a CNN for facial recognition the convolutional layers will start learning feature combinations off of the previous layers. The simple edges,

gradients, and corners of the first layers become things like eyes, noses, and hairs in later layers. This process is visualized in figure 8.2

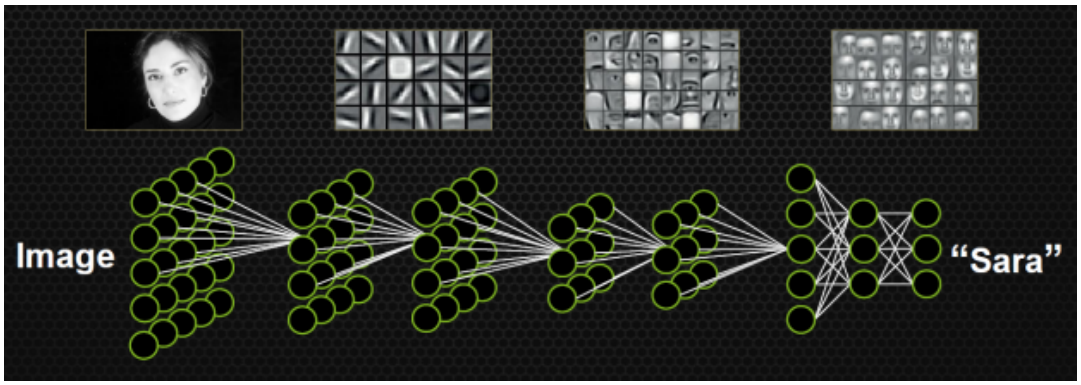


Figure 8.2.: Pictorial Representation of Convolutional Neural Networks as well as a visual representation on CNN's complexity of layer feature extraction

Chapter 9.

Hardware Frameworks

Chapter about different frameworks used/timing on each setup.

9.1. Syracuse CPU Machine setup

9.2. Syracuse University GPU Cluster Setup

...

Chapter 10.

Using Convolutional Neural Networks to separate $\mu's$ from $\pi's$

The work shown in these next sections are based on the previous work done described in [**?**]. That CNN (now referred to as CNN1075) was trained using single generated isotropic muons and pions from 0-2 GeV energy range. 1,075 muons and pions were used to train the network and 1,075 μ/π were used as a validation set. The accuracy is how well CNN1075 is doing by epoch and was 74.5%. The loss is gradient descent or minimization of the error of the weights and biases used in each neuron of each layer of CNN1075 and was 58% with a trend sloping downwards on the loss curve as well as a trend sloping upward in the accuracy curve. Due to the depth of the neural network framework, it was necessary to train with a larger dataset and for more epochs, however, the downward slope of the loss curve is an indication that once trained for longer with a higher training sample, neural networks can be used for μ/π separation. Updates in the image making and downsampling algorithm were made to fix issues that arose in CNN1075.

10.1. Image Making Scheme

The μ/π image dataset used to train and test the second CNN (now referred to as CNN10000) was created using single generated isotropic muons and pions from 0-2 GeV energ range. 10,000 muons and 10,000 pions were used for training and testing split 50%. The images were created based on wire number and time tick in the collection plane. Uboonecode v06_23_00 was used instead of v05_08_00 which was

used previously. The wire signal was the raw ADC value after noise filtering. Each collection plane grayscale image was 3456x1280x1 where 5 time ticks were pooled into 1 bin which is different than the previous dataset and was implemented due to the fact that the time ticks of an event went from 9400 to 6400 with the change of uboonecode version. The grayscale color standard is 8bit therefore the ADC value of wire and time tick was also downsampled due to the 12bit ADC value MicroBooNE has. To do this, the highest ADC pixel in the image was found and then this was divided by the rest placing all pixel values between 0-1. From there, all pixel values are then multiplied by 255. All images were made using a LArSoft module. Once the images were created, using and image manipulation framework called OpenCV images were read into a numpy array and cropped to the region of interest by only keeping rows and columns where all ADC values are higher than 0 and then resized it to 224x224 using OpenCV's resize function. This downsampling of ADC values creates a problem of information loss for example, a proton which is highly ionizing will have the same brightness as a minimum ionizing muon by virtue of how the images are created. Creating images that thoroughly depict the ADC scale for use in dE/dx particle identification has been implemented and retraining of CNN is underway. Issues that arose in CNN1075 that were fixed in CNN10000 include zero-padding images in X and Y that are smaller than 224X224 to eliminate over-zooming effect and fixing a bug that shifted pixels separated by a dead-wire region.

Images were also made from events that passed the cc-inclusive selection 1 filter right before the 75 cm track length cut and were classified using the CNN10000. The dataset used to create these images is the same one used in [**?**], prodgenie_bnb_nu_cosmic_uboone_mc These images were created using information from the track candidate that passed the filter. Only wire number and time ticks associated to the track candidate were drawn on the image to mimic a single particle generated image. These images were then classified using CNN10000. Two approaches were taken in making these images. The first was using the image normalization above where the maximum pixel in each image is used as a normalization constant to get all pixels between 0-1 then multiply all pixels by 255. As described above, this is the incorrect way to normalize; it should be normalized by dataset not by event, which is the second way the images were created. The results of CNN10000 performance are shown in section 10.2.

10.2. Convolutional Neural Network Training

The hyperparameters used for CNN10000 are shown. The batch size for the training and testing as well as the test iter were chosen to encompass the whole training/testing image set when doing accuracy/loss calculations. To do this, multiplying the test iter by the test batch size give you the amount of images used when calculating accuracy/loss curves. For reference, the accuracy and loss are defined as well.

- `train_batch_size: 100`
- `test_batch_size: 100`
- `test_iter: 100`
- `test_interval: 100`
- `base_lr: 0.001`
- `lr_policy: "step"`
- `gamma: 0.1`
- `stepsize: 1000`
- `display: 100`
- `max_iter: 10000`
- `momentum: 0.99`
- `weight_decay: 0.0005`
- `snapshot: 100`
- Accuracy: How often the CNN predicts the truth over total number of images
- Loss: Error between truth and prediction. Minimize loss by gradient descent to update weights and biases of CNN

The same architecure that was used to train CNN1075 was employed on CNN10000, Imagenet. Caffe [**?**] was the software package used for both CNNs. The differences include batch size and test_iter and momentum to account for the larger dataset. Both CNNs were trained on a CPU machine, Syracuse01. Further training will be done on a GPU cluster stationed at Syracuse University. Figure 10.1 shows the loss and

accuracy of CNN10000. There is around a 10% increase in accuracy from CNN1075 to CNN10000, 85%, and around a 20% decrease in loss, 36%.

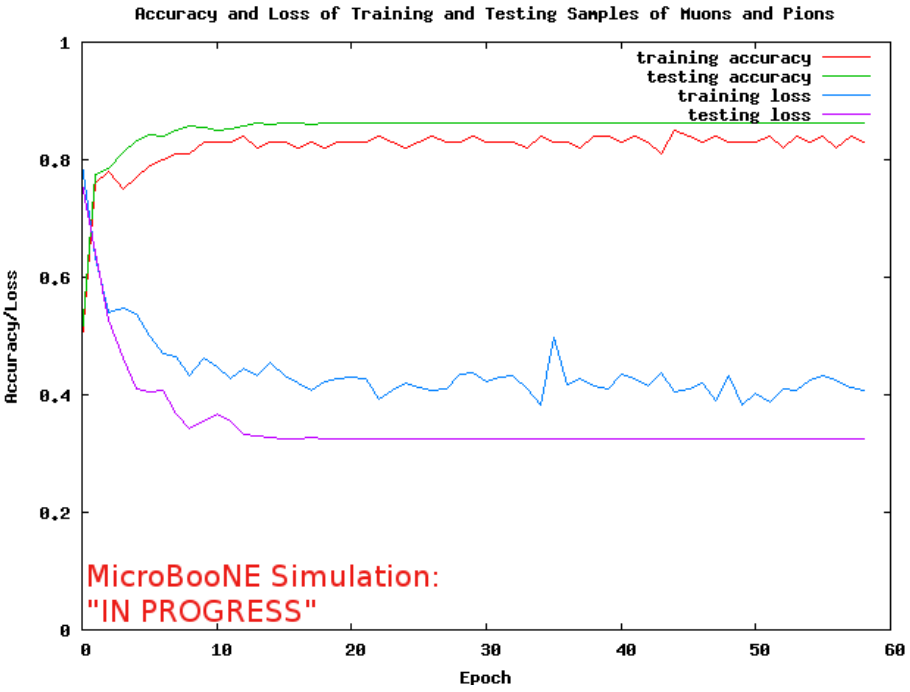
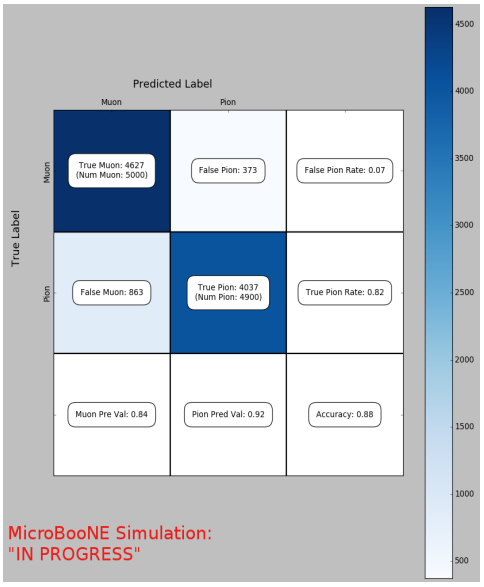


Figure 10.1.: Accuracy vs. Loss of ImageNet 2-output μ/π sample consisting of 10000 images each.

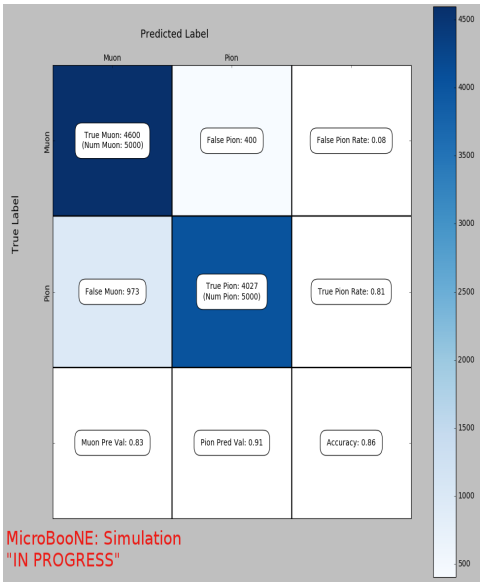
Figure 10.2 show a breakdown of μ/π separation for CNN10000. It also shows the network is not being overtrained due to the Accuracy of both the training and testing datasets being within .01% of eachother. The CNN is doing a very good job of classifying true muons as muons, and our loss increase from CNN1075 is due to the increase in accuratly classifing pions as pions.

10.3. Classification of MC data using Selection I Original CC-Inclusive Filter

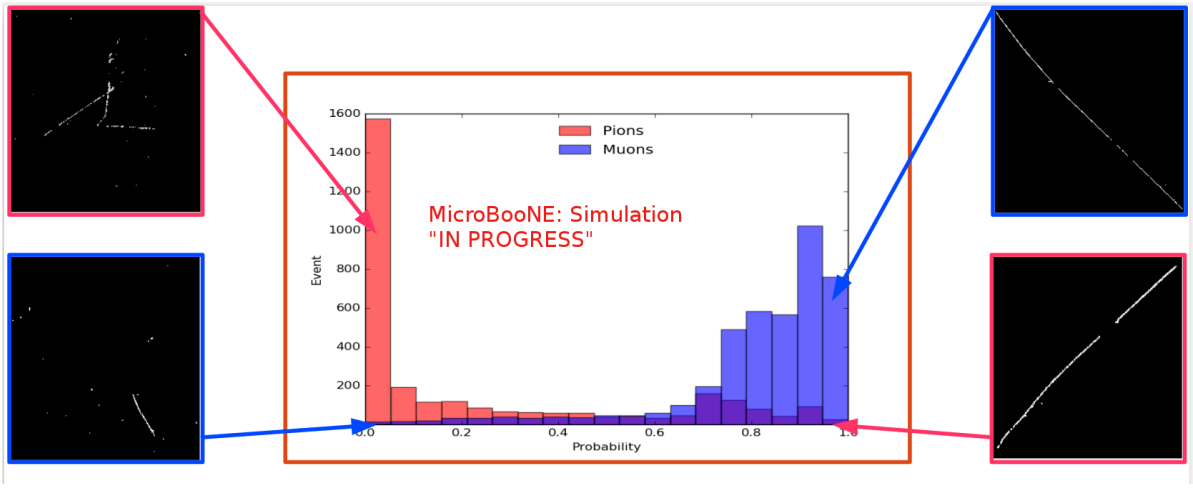
The next step that was taken was to use CNN10000 to classify track candidate images that were identified by the selection I original cc-inclusive filter described in [**?**]. Passing rates for each cut in cc-inclusive filter are show in figure 10.3. For the incorrect image making normalization dataset, out of 188,880 events, 7438 passed the cut right before 75 cm track length cut which is 3.9% of total data. Discrepancies in passing rates



(a) Confusion Matrix showing Accuracy of CNN using training data



(b) Confusion Matrix showing Accuracy of CNN using testing data



(c) Probability plot of muons and pions from testing set. Images surrounding histogram are a random event from lowest bin and highest bin for each particle.

Figure 10.2.: Description of confusion matrix varables: False pion rate = $false\pi/total\pi$ True pion rate = $true\pi/total\pi$ Accuracy = $(true\pi rate + true\mu rate)/2$ Pion prediction value = $true\pi/(true\pi + false\pi)$ Muon prediction value = $true\mu/(true\mu + false\mu)$ 10.2c The probability plot includes muons and pions that are classified as primary particles.

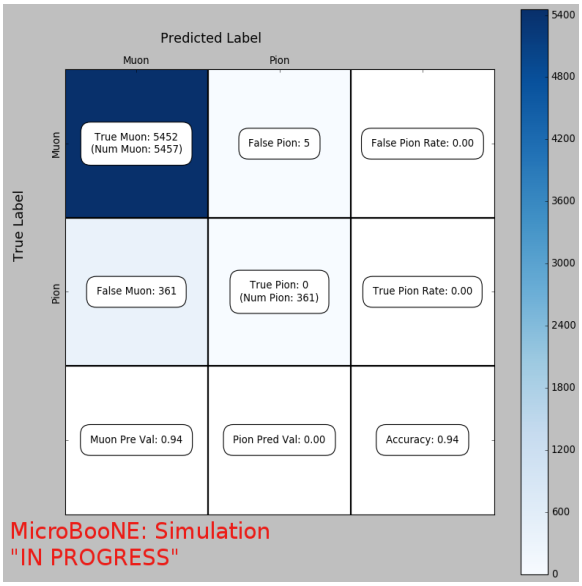
Table 3: **Selection I: Original** The table shows passing rates for the above described event selection. Numbers are absolute event counts and Cosmic background is not scaled appropriately. The BNB+Cosmic sample contains all events, not just ν_μ CC inclusive. The selected events are further broken up in the following subsection. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt the generated events (second percentage). In the BNB+Cosmic MC Truth column shows how many true ν_μ CC inclusive events (in FV) are left in the sample. This number includes possible mis-identifications where a cosmic track is picked by the selection instead of the neutrino interaction in the same event. The cosmic only sample has low statistics, but please note that it is not used in any plots, it is just for illustrating the cut efficiency. The last column Signal:Cosmic only gives an estimate of the ν_μ CC events wrt the cosmic only background at each step. For this number, the cosmic background has been scaled as described in chapter 2. Note that this numbers is not a purity, since other backgrounds can't be determined at this step.

	BNB + Cosmic				Cosmic only		Signal:
	Selection		MC-Truth				Cosmic only
Generated events	191362		45273		4804		1:22
$\geq$ 1 flash with $\geq$ 50 PE	136219	(71%/71%)	44002	(97%/97%)	2979	(62/%62%)	1:14
$\geq$ 1 vertex in FV	131170	(96%/69%)	43794	(99%/97%)	2805	(94%/58%)	1:13
$\geq$ 1 track within 5 cm of vertex	129784	(99%/68%)	43689	(99%/97%)	2756	(98%/58%)	1:13
flash matching of longest track	44775	(34%/23%)	23647	(54%/52%)	647	(23%/13%)	1:5.7
track containment	10114	(23%/5.3%)	6882	(29%/15%)	61	(9.4%/1.3%)	1:1.9
track $\geq$ 75 cm	7358	(73%/3.8%)	5801	(84%/13%)	31	(51%/0.6%)	1:1.1

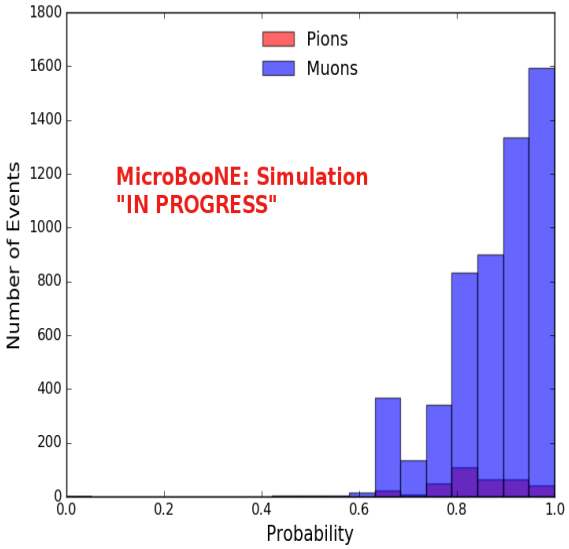
Figure 10.3.: Snapshot of passing rates of Selection I from CC-Inclusive Filter

are due to grid submission issues, however, this dataset is used to check if changes in image making normalization affects μ/π separation probability due to CNN10000 being trained with incorrectly image making normalized data. For the second dataset with correct image making normalization, out of 188,880 events, 9552 events passed the cut right before the 75 cm track length cut which is 5.1% passing rate and is comparable to figure 10.3. In time cosmics were also run over for efficiency and purity calculations. Out of 14395 in time cosmic events, 175 passed the cut right before the 75 cm track length cut which is a passing rate of 1.2% compared to 1.3% shown in figure 10.3.

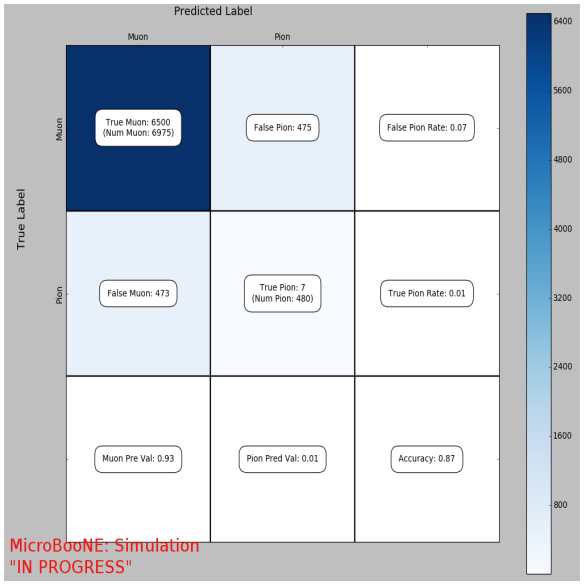
Figures 10.4a, 10.4b, 10.4c and 10.4d show the accuracy and μ/π separation of both the correct and incorrect normalized images. The confusion matrices are only composed of μ/π data. Other particles passed the cc-inclusive filter before the 75 cm track length cut and were all mis-id'ed as muons. Since CNN10000 has not seen any particles other than muons and pions, it makes sense that those get mis-id'ed. Figures 10.4b and 10.4d don't have μ/πseparation comparable to 10.2c, but 10.4b does skew to higher probabilities compared to 10.4d. This is to be expected and further work on quantifying the performance of CNN10000 should use the incorrect image making normalization. It is also expected that the separation isn't as defined as the testing dataset for CNN10000. CNN10000 was trained and tested using single particle muons and pions and the track candidate dataset come from BNB+Cosmic events, not to mentions all track candidates have passed the cc-inclusive filter that



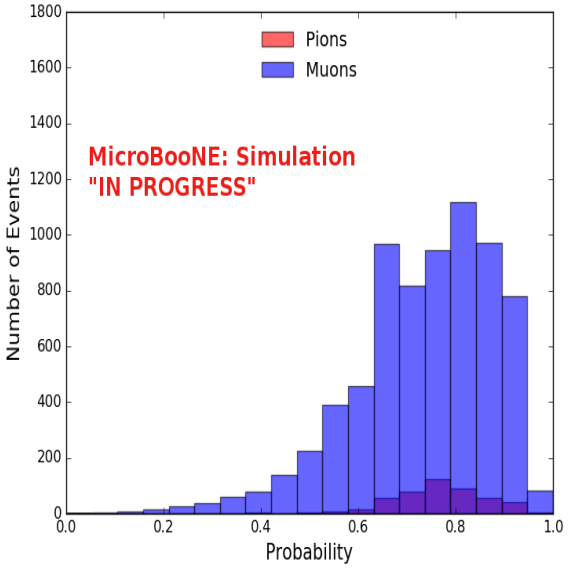
(a) Confusion Matrix showing Accuracy of CNN using data with wrong normilazion



(b) Probability plot showing μ/π separation of CNN using wrong normalization

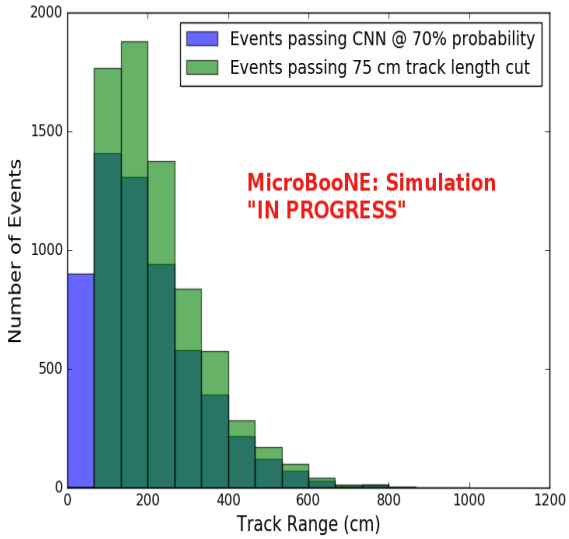


(c) Confusion Matrix showing Accuracy of CNN using data with correct normilazion

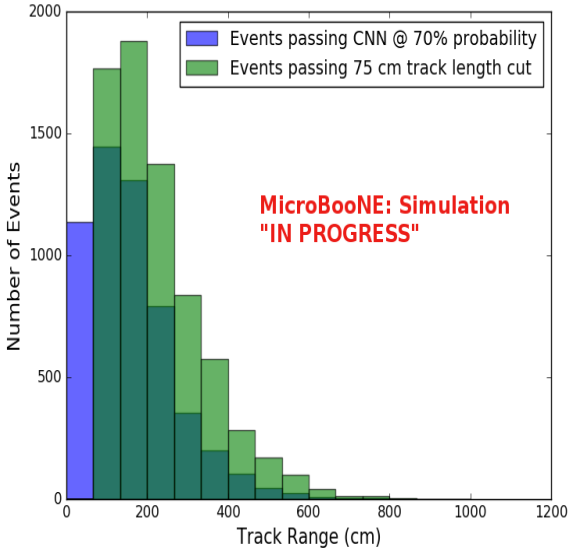


(d) Probability plot showing μ/π separation of CNN using correct normalization

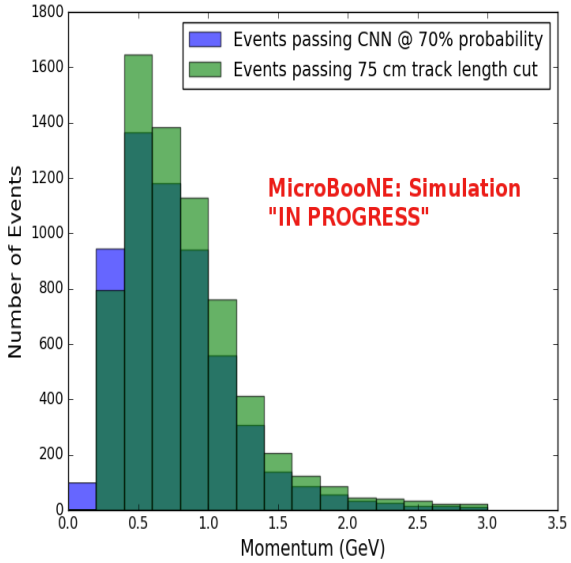
Figure 10.4.: Results of CNN10000 classification of track candidate images output from cc-inclusive filter.



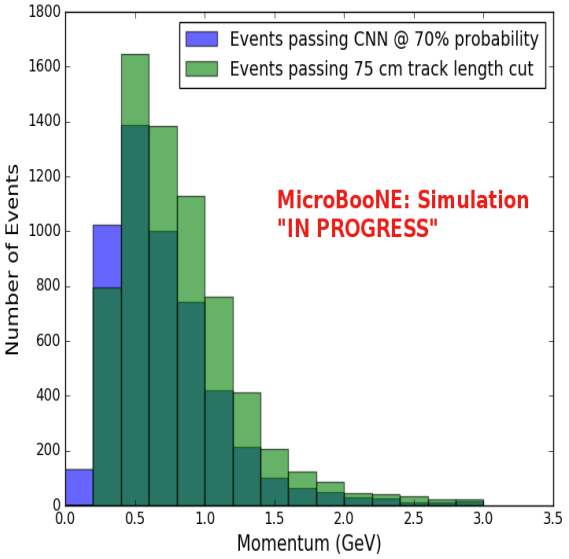
(a) Track range distribution of events from Selection I Original passing CNN with 70% accuracy using image data with wrong normilazion



(b) Track range distribution of events from Selection I Original passing CNN with 70% accuracy using image data with correct normilazion



(c) Momentum distribution of events from Selection I Original passing CNN with 70% accuracy using image data with wrong normilazion



(d) Momentum distribution of events from Selection I Original passing CNN with 70% accuracy using image data with correct normilazion

Figure 10.5.: CNN10000 distributions of track candidate images output from Selection I Original cc-inclusive filter with different image data normalizations

tags "muon-like" tracks therefore the pions in this sample look much closer in muon topology than the network has seen. Also, these images were made from wire and time ticks associated to hits from the track candidate that passed the cc-inclusive filter. This is different from the training images where a bounding box was drawn over the total μ or π interaction. Spurious energy deposition from a $\pi - Ar$ interaction is most likely not included in the BNB+Cosmic images due to the tracking algorithm. To remedy this, the neural network needs to see more "muon-like" pions and muons and pions from a neutrino interaction passing the cc-inclusive filter as well as a larger particle variety including protons, photons and electrons. Although μ/π separation is lacking, CNN10000 does an excellent job of classifying muons and using higher CNN probability can increase purity. Figures 10.5a, 10.5b, 10.5c and 10.5d show the track and momentum distributions for these two datasets. In both sets you have an increase in data in the bin below 75 cm and at bins below 0.5 GeV. These distributions were made with events classified with 70% probability of being a muon regardless of true particle type.

10.4. Classification of MC data using Selection I Modified CC-Inclusive Filter

Table 5: **Selection I: Modified** The table shows passing rates for the above described event selection. Numbers are absolute event counts and Cosmic background is not scaled appropriately. The BNB+Cosmic sample contains all events, not just ν_μ CC inclusive. The selected events are further broken up in the following subsection. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt the generated events (second percentage). In the BNB+Cosmic Mc Truth column shows how many true ν_μ CC inclusive events (in FV) are left in the sample. This number includes possible mis-identifications where a cosmic track is picked by the selection instead of the neutrino interaction in the same event. The cosmic only sample has low statistics, but please note that it is not used in any plots, it is just for illustrating the cut efficiency. The last column Signal:Cosmic only gives an estimate of the ν_μ CC events wrt the cosmic only background at each step. For this number, the cosmic background has been scaled as described in chapter 2. Note that this numbers is not a purity, since other backgrounds can't be determined at this step.

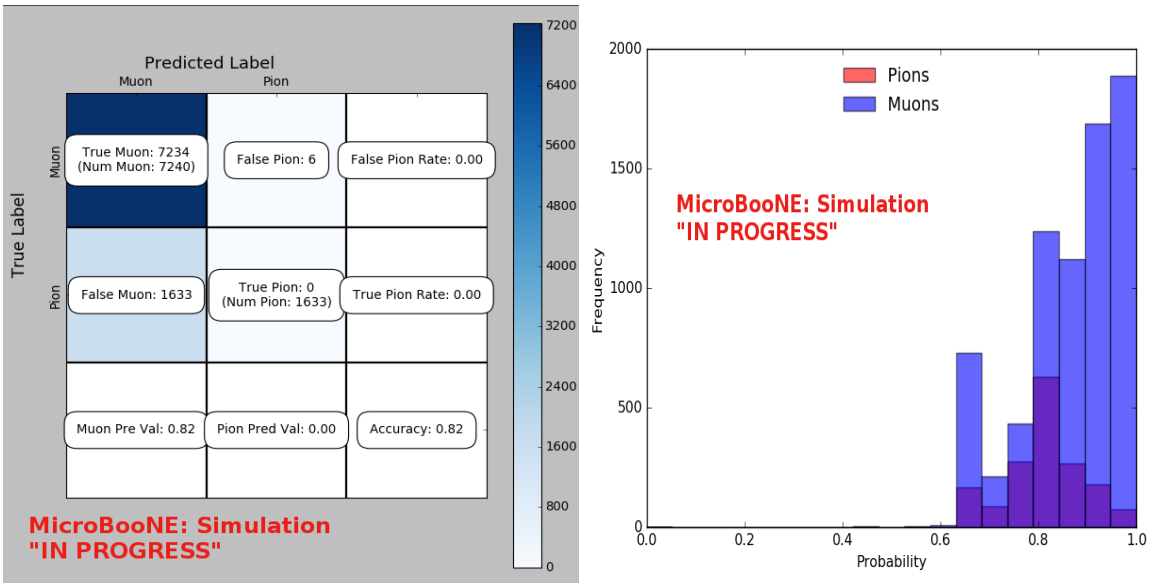
	BNB + Cosmic				Cosmic only		Signal:
	Selection		MC-Truth				Cosmic only
Generated events	191362		45273		4804		1:22
$\geq$ 1 flash with $\geq$ 50 PE	136219	(71%/71%)	44002	(97%/97%)	2979	(62%/62%)	1:14
$\geq$ 1 track within 5 cm of vertex	135830	(99%/71%)	43974	(99%/97%)	2975	(99%/62%)	1:14
vertex candidate in FV	79112	(58%/41%)	34891	(79%/77%)	1482	(50%/31%)	1:8.9
flash matching of longest track	40267	(51%/21%)	25891	(74%/57%)	340	(23%/7.1%)	1:2.8
track containment	19391	(48%/10%)	11693	(45%/26%)	129	(38%/2.7%)	1:2.3
track $\geq$ 75 cm	6920	(36%/3.6%)	5780	(49%/13%)	17	(13%/0.4%)	1:0.6

Figure 10.6.: Snapshot of passing rates of all cuts from Selection I Modified cc-inclusive filter

CNN10000 was also used to classify track candidate images that were identified by the selection I modified cc-inclusive filter described in [**?**]. Passing rates for each cut in this filter are shown in figure 10.6. As seen in section 10.3, wrong image normalization had a higher muon classification probability so all work done using selection I modified cc-inclusive filter was done using this normalization. Out of 188,880 events, 19,112 passed the cut right before the 75 cm track length cut which is a 10.1% passing rate and comparable to the 10% passing rate shown in figure 10.6. In time cosmics were also run over, out of 14,606 in time cosmics events, 302 passed the cut right before the 75 cm track length cut which is a 2.1% passing rate comparable to the 2.7% passing rate in the cc-inclusive tech-note. Figures 10.7a and 10.7b show the accuracy and μ/π separation. Both plots are only composed of muons and pions and like selection I original data, all other particles were id'ed as muons. Also like selection I original data, muons are being identified at a very high rate. Figure 10.8a shows the track range distributions of all events from selection I modified being classified by the CNN as a muon with a probability of 70% regardless of true particle type. We get entries for the CNN curve in the lowest bin and none for the 75 cm curve. To see how many true CC events were identified by CNN10000 breaking down figure 10.8a by event type was necessary. Figures 10.8b and 10.8c show track range distributions separated by signal and various backgrounds. Particle type was not taken into consideration in these plots so true CC event images can be any track candidate particle passing selection I modified cut right before track length cut including pions and protons.

To gain an even deeper understanding on how CNN10000 is performing, plotting these distributions with only muons and pions was done due to the fact that CNN10000 was trained with only those particles for μ/π separation. Figures 10.8d-10.9d show the stacked histograms of signal and background of the track range distributions with varying CNN probabilities starting from 70% and ending at 90% probability. With higher probabilities we get a purer sample in the lower bin but we end up losing events as well. Momentum distributions for all signal/background events are shown in figure 10.10.

Another check was to see if any true CC pions were passing through the cut right before the 75 cm track length cut. Figure 10.11 shows the comparison of the stacked track range distribution with only true CC muon signal versus the stacked distribution with true CC muons and pions signal. As you can see, we gain more events when plotting CC events with a particle type of either muons or pions due to the CNN classifying all pions in this dataset as muons. This is an interesting scenario and a

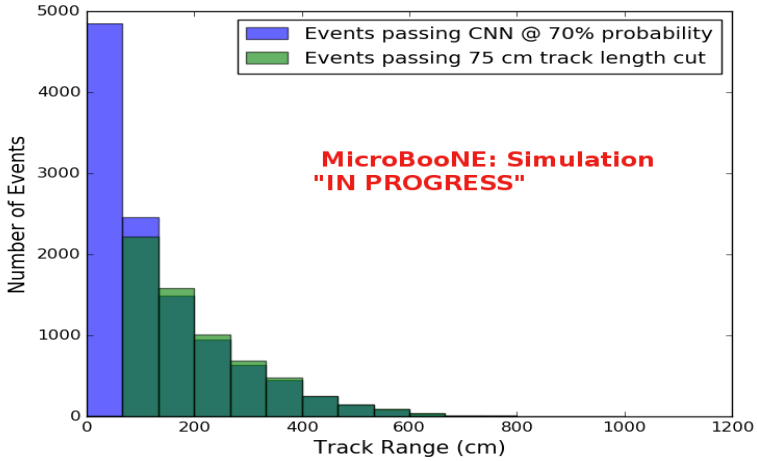


(a) Confusion Matrix for CNN10000 classified events from selection I modified

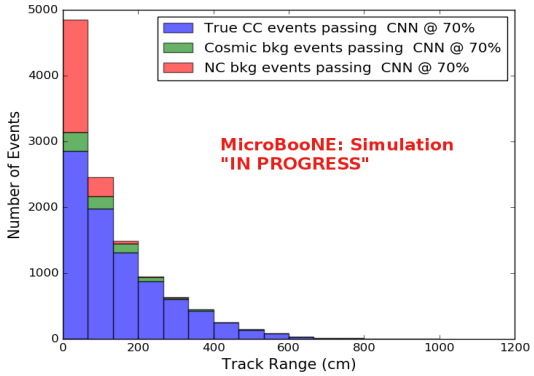
(b) Probability plot for CNN10000 classified events from selection I modified

Figure 10.7.: Confusion matrix and probability plot of events passing selection I modified cc-inclusive cuts right before 75cm track length cut

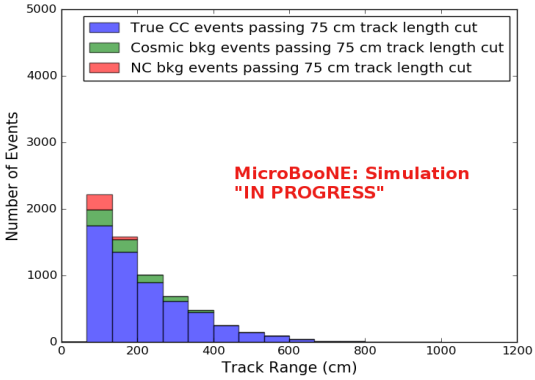
sample of topologies of these images are represented in figure 10.12, at least 3 tracks are coming out of the vertex for these types of events. With the 75 cm track length cut, the selection is cutting event topologies like this where the pion is the tagged track candidate. Figure 10.12a has a defined longer muon track, but because of dead wires through the track, the reconstructed range is 1. less than 75 cm and 2. shorter than the reconstructed pion whose length is also less than 75 cm. This is a very interesting event, but because of issues with the tracking algorithm, the 75 cm cut would get rid of this event. The CNN was able to recover this event only because it has classified all pions as muons. Figure 10.12b shows the second case to think about, the pion, while still less than 75 cm has a reconstructed track length longer than the muon. Again, the CNN recovered this event due to pions being classified as muons. Lastly, figure 10.12c shows a pion with a reconstructed track length greater than 75 cm and the muon. These three cases show that a broader question must be asked when training the network other than is it a muon or pion. There are different routes to recover interesting events like these. One route is to ask the network "Is it a CC event or is it an NC event?" and obtain an image dataset consisting of whole CC/NC events that will train the network to answer this question. The other route is to ask the network "Is this a $\mu/\pi/p/$ from a CC event or NC event and obtain an image dataset consisting



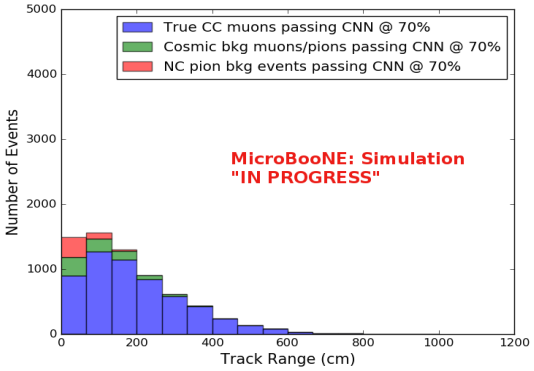
(a) Track range distribution of events from Selection I Modified passing CNN with 70% accuracy



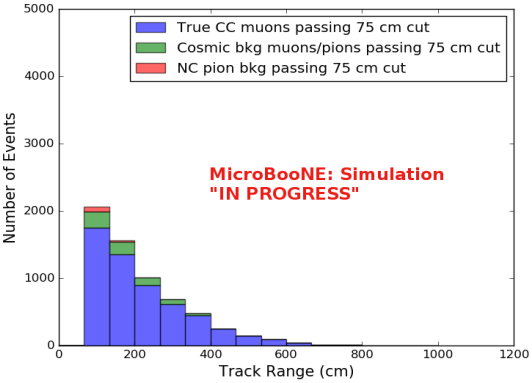
(b) Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy



(c) Stacked signal and background track range distributions from Selection I Modified passing 75 cm track length cut

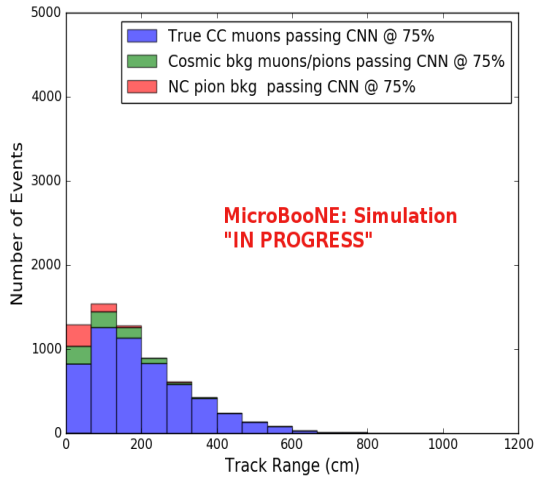


(d) Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 70% accuracy

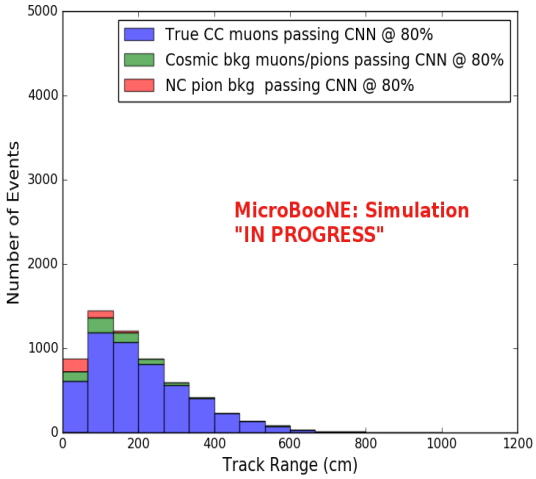


(e) Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing 75 cm track length cut

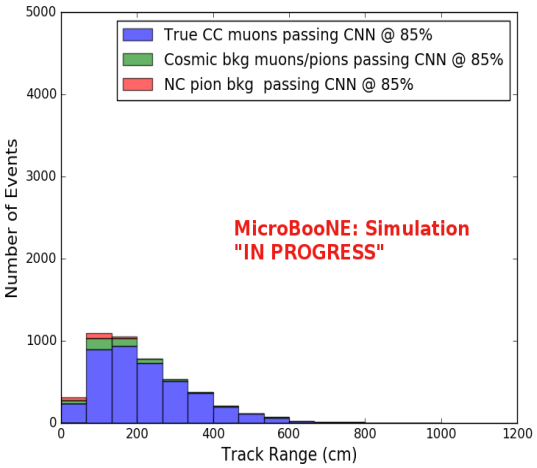
Figure 10.8.: CNN10000 distributions of track candidate images output from Selection I Modified cc-inclusive filter



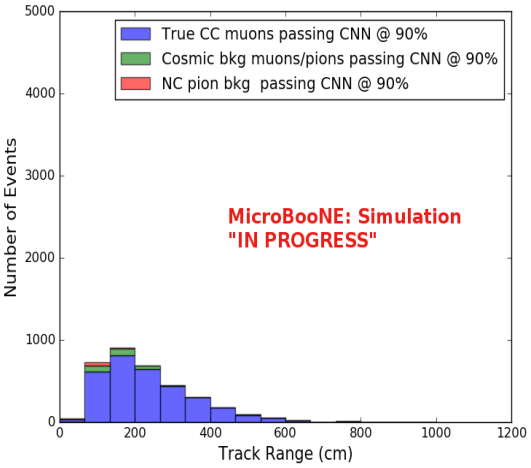
(a) Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 75% accuracy



(b) Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 80% accuracy

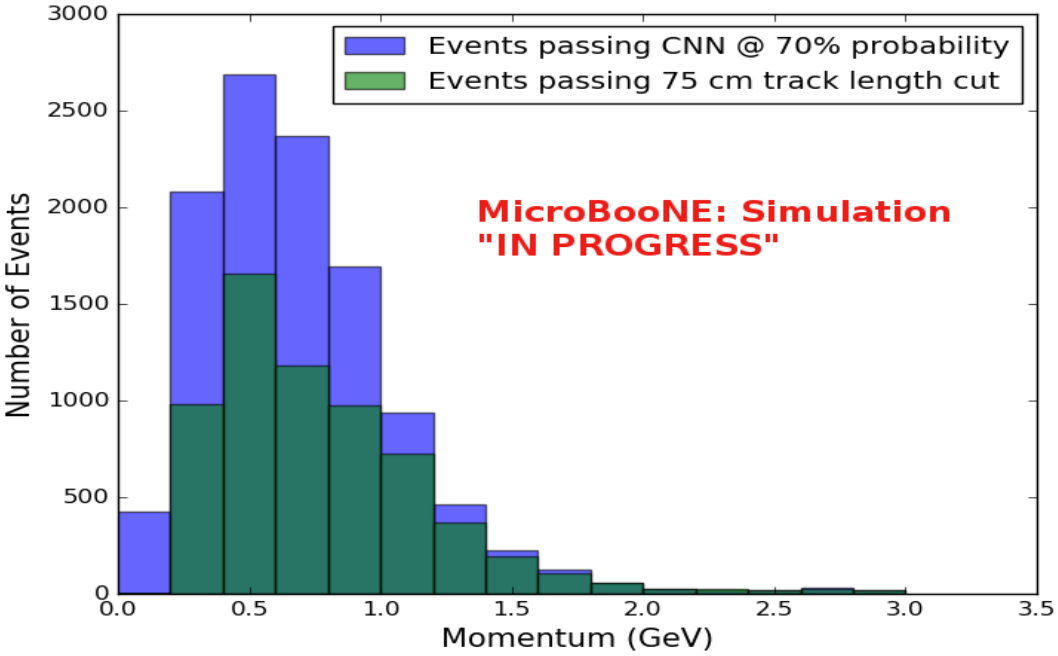


(c) Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 85% accuracy

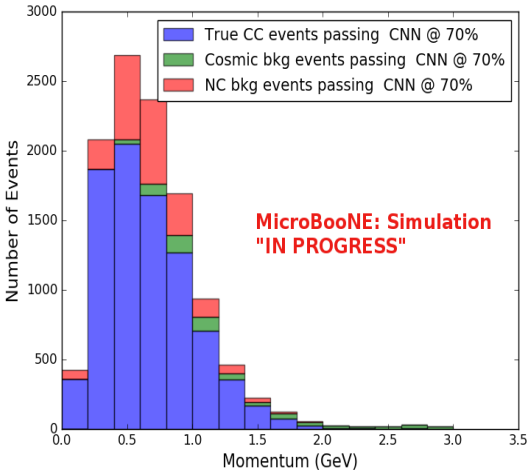


(d) Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 90% accuracy

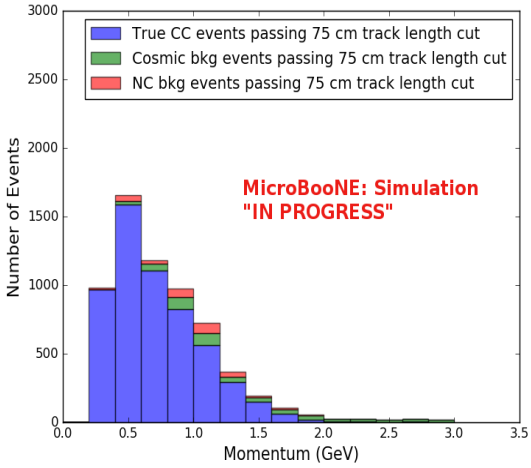
Figure 10.9.: CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter



(a) Momentum distribution of events from Selection I Modified passing CNN with 70% accuracy



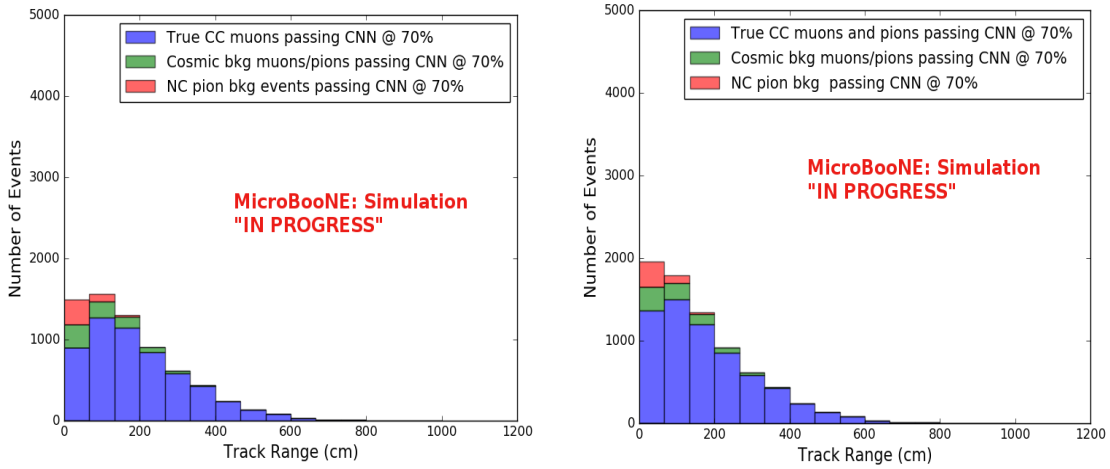
(b) Stacked signal and background momentum distributions from Selection I Modified passing CNN with 70% accuracy



(c) Stacked signal and background momentum distributions from Selection I Modified passing 75 cm track length cut

Figure 10.10.: CNN10000 momentum distributions of track candidate images output from Selection I Modified cc-inclusive filter

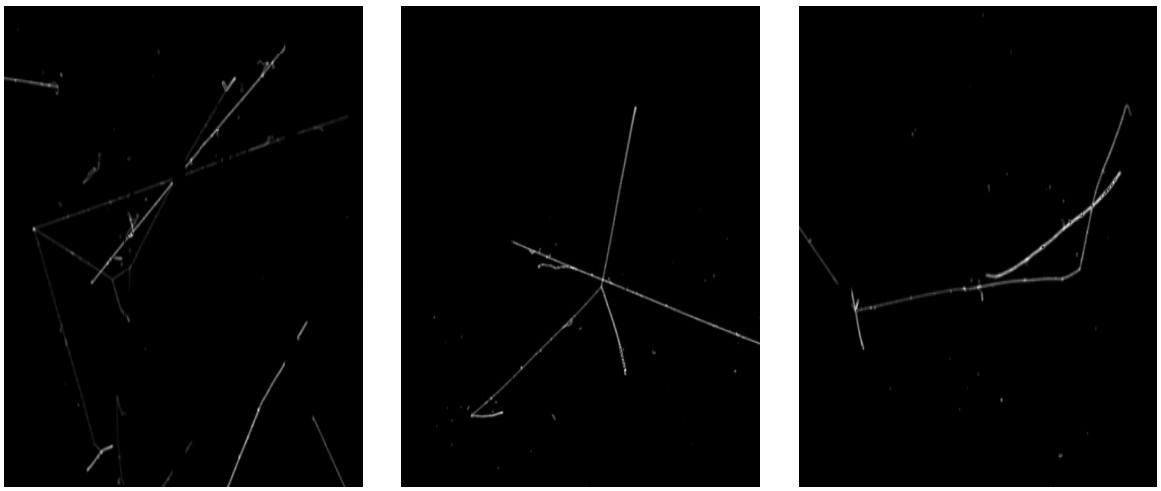
of primary particles from a CC/NC event. Both these paths will be explored in future work.



(a) Stacked signal μ/backround μ and π track range distribution of CNN @ 70%

(b) Stacked signal $\mu\&\pi$/backround $\mu\&\pi$ track range distribution of CNN @ 70%

Figure 10.11.: Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted



(a) Pion recontructed track range is less than 75 cm and longer than muon track due to dead wires

(b) Pion recontructed track range is less than 75 cm and larger than muon reconstructed track

(c) Pion reconstructed track range is greater than 75 cm and larger than muon reconstructed track

Figure 10.12.: Images of true CC events where the pion was the tagged track candidate

		BNB + Cosmics		Cosmic Only	Signal:
		Selection	MC Truth		Cosmic Only
75 cm Cut passing rates	Generated Events	191362	45723	4804	1:22
	Track Containment	19391 (48%/10%)	11693 (45%/26%)	129 (38%/2.7%)	1:2.3
	track $\geq$ 75 cm	6920 (36%/3.6%)	5780 (49%/13%)	17 (13%/0.4%)	1:0.6
CNN passing rates	Generated Events	188880	44689	14606	1:21
	Track Containment	19112 (/10%)	11554 (/26%)	302 (/2.1%)	1:1.73
	CNN cut @ 70% Probability	16502 (86%/8.7%)	10605 (92%/23%)	205 (68%/14%)	1:1.28
	CNN cut @ 83% Probability	7511 (46%/4.0%)	6142 (58%/14%)	32 (16%/0.2%)	1:0.4

Table 10.1.: Comparing passing rates of CNN at different probabilities versus 75 cm track length cut: Numbers are absolute event counts and Cosmic background is not scaled appropriately. The BNB+Cosmic sample contains all events. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt the generated events (second percentage). In the BNB+Cosmic MC Truth column shows how many true ν_μ CC-inclusive events (in FV) are left in the sample. This number includes possible mis-identifications where a cosmic track is picked by the selection instead of the neutrino interaction in the same event.The CNN MC True generated events were scaled wrt the MC True generated events for the 75 cm cut passing rates due to only running over 188,880 generated events versus the 191362 generated events. The last column Signal:Cosmic only gives an estimate of the ν_μ CC events wrt the cosmic only background at each step. For this number, the cosmic background has been scaled as described in [**?**]. Note that these numbers are not a purity, since other backgrounds canâĂŹt be determined at this step.

		#Events(Fraction)	#Events(Fraction)
		passing CNN @ 70% Probability	passing CNN @ 83% Probability
Signal	ν_μ CC events with true vertex in FV	10605(35%)	6142(61%)
Backgrounds	Cosmics Only Events	13573(45%)	2582(26%)
	Cosmics in BNB Events	2249(7.4%)	492(4.9%)
	NC Events	3412(11%)	778(7.7%)
	ν_e and $\bar{\nu}_e$ Events	139(0.5%)	32(0.3%)
	$\bar{\nu}_\mu$ Events	97(0.3%)	67(0.7%)

Table 10.2.: Signal and background event numbers at modified selection level with CNN cut estimated from a BNB+Cosmic sample and Cosmic only sample normalized to $5 * 10^{19}$ PoT. The last column gives the fraction of this signal or background type to the total selected events per CNN probability.

Table 10.1 shows the passing rates for the 75 cm track length cut and the CNN cut at 70% and 83%. The passing rates at the track containment level for the 75 cm track length cut compared to the CNN are comparable with only a 0.6% difference in the in time cosmic bin which may be due in part to the larger in time cosmic statistics used for the CNN dataset. These passing rates need to be comparable to then be able to compare the passing rates after the CNN cut to the 75 cm cut. Again, the same BNB+Cosmic sample was used for both selection I modified with 75 cm cut and

selection I modified with CNN cut. As it stands, a CNN cut at 83% probability has a MC true CC event passing rate of 14% compared to the 13% passing rate of the 75 cm track length cut. The Signal:Cosmic Only background is also reduced from 1:0.6 to 1:0.4 The total passing rate is also higher than the 75 cm cut, 3.6% vs 4.0%. Table 10.2 shows the breakdown of signal and backgrounds for the CNN at the different probabilities. We have a 61% signal passing rate with the CNN cut @ 83% versus the 53.8% signal passing rate of the 75 cm cut.

Based on these numbers, the following performance values of the modified selection with 75 cm cut versus modified selection with CNN @ 83% probability cut were calculated:

- Efficiency: Number of selected true ν_μ CC events divided by the number of expected true ν_μ CC events with interaction in the FV.
  - Selection I modified: 13%
  - Selection I modified with CNN cut @ 83% probability: 14%
- Purity: Number of selected true ν_μ CC events divided by sum of itself and the number of all backgrounds.
  - Selection I modified: 53.8%
  - Selection I modified with CNN cut @ 83% probability: 61%

Lastly, figure 10.14 shows a more representative performance of the CNN. Due to the fact that the CNN was trained on muons and pions, showing the performance of CC muon events versus NC pion events with respect to CNN probability gives a better picture of how the network is performing. Figure 10.14 shows that at 83% we are below the 75 cm cut NC pion threshold and still above the CC muon threshold. Using 83% probability not only reduced the NC pion background, it also dramatically reduced the in time cosmics and cosmics in the BNB.

10.5. Conclusions and Future Work

It was shown that even though CNN10000 was trained with single particle generated muons and pions, it performs fairly well at classifying track candidate images from BNB+Cosmic events. Events have been regained below the 75 cm track length cut and

Table 8: **Selection I: Modified** Signal and background event numbers at modified selection level estimated from a BNB+Cosmic sample and Cosmic only sample normalized to 5×10^{19} PoT. The last column gives the fraction of this signal or background type to the total 2189 selected events.

		#Events	
Signal	ν_μ CC events with true vertex in FV	1168	53.8%
Backgrounds	Cosmics only events	725	33.4%
	Cosmics in BNB events	144	6.6%
	NC events	75	3.5%
	ν_e and $\bar{\nu}_e$ events	4	0.2%
	$\bar{\nu}_\mu$ events	15	0.7%
	ν_μ CC events with true vertex outside FV	40	1.8%

Figure 10.13.: Snapshot of signal and background event numbers of Selection I modified from cc-inclusive note [**?**]

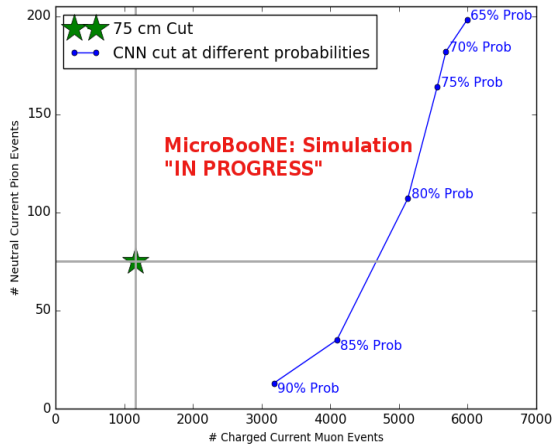


Figure 10.14.: CNN performance of classified muons and pions compared to the already implemented 75 cm track length cut

the momentum and track range distributions have similar shapes to the distributions of Selection I original and modified. Efficiencies and purities were calculated for selection I modified events before 75 cm track length cut with the CNN at 83% probability and are 14% and 62% respectively. Although the CNN doesn't have separation between muons and pions and although all particles passing CNN are classified as muon, increasing CNN probability allows us to increase the purity as well as maintain an efficiency comparable to the 75 cm track length cut all while recovering events below that 75 cm cut. Out of the 6142 events that passed the CNN @ 83% 1470 events were below the 75 cm cut, a recovery of 3.3% of data with an purity of 15%. Although these numbers are low, it is an improvement from the selecion I modified in both total efficiency and purity and an increase in phase space by recovering these events. The next steps are to train two additional neural networks. One with track candidate muon and pion images originating from BNB+Cosmic events and the second whole CC_{μ}/NC_{π} events to gauge CNN performance. These trainings are underway.

Chapter 11.

Using Convolutional Neural Networks on MicroBooNE Data

...

...

Chapter 12.

Comparing two CC-Inclusive Cross Section Selection Filters

...

...

Chapter 13.

Conclusion

Your Conclusions here.

Appendix A.

Pointless extras

"Le savant n'étudie pas la nature parce que cela est utile;
il l'étudie parce qu'il y prend plaisir,
et il y prend plaisir parce qu'elle est belle."
— Henri Poincaré, 1854–1912

Appendixes (or should that be "appendices"?) make you look really clever, 'cos it's like you had more clever stuff to say than could be fitted into the main bit of your thesis. Yeah. So everyone should have at least three of them…

A.1. Like, duh

Padding? What do you mean?

A.2. $y = \alpha x^2$

See, maths in titles automatically goes bold where it should (and check the table of contents: it *isn't* bold there!) Check the source: nothing needs to be specified to make this work. Thanks to Donald Arsenau for the teeny hack that makes this work.

Colophon

This thesis was made in LaTeX 2_ε using the "hepthesis" class [2].

Bibliography

[1] MicroBooNE, R. e. a. Acciari, (2015), arXiv:1512.06148.

[2] A. Buckley, The hepthesis LaTeX class.