SISSA

PAPER

Neutral pion reconstruction using machine learning in the experiment at $\langle E_\text{v} \rangle$ 6 GeV

To cite this article: A. Ghosh *et al* 2021 *JINST* **16** P07060

View the article online for updates and enhancements.

Neutral pion reconstruction using machine learning in the MINERvA experiment at $\langle E_\nu \rangle \sim 6\,\text{GeV}$

A. Ghosh,[a,b,1,*] **B. Yaeggy,**[a] **R. Galindo,**[a] **Z. Ahmad Dar,**[c] **F. Akbar,**[c] **M.V. Ascencio,**[d]
A. Bashyal,[e] **A. Bercellie,**[f] **J.L. Bonilla,**[g] **G. Caceres,**[b] **T. Cai,**[f] **M.F. Carneiro,**[e,b]
H. da Motta,[b] **G.A. Díaz,**[f,d] **J. Felix,**[g] **A. Filkins,**[h] **R. Fine,**[f] **A.M. Gago,**[d] **T. Golan,**[i,f]
R. Gran,[j] **D.A. Harris,**[i,k] **S. Henry,**[f] **S. Jena,**[l] **D. Jena,**[i] **J. Kleykamp,**[f] **M. Kordosky,**[h]
D. Last,[m] **T. Le,**[n,o] **A. Lozano,**[b] **X.-G. Lu,**[p] **E. Maher,**[q] **S. Manly,**[f] **W.A. Mann,**[n] **C. Mauger,**[m]
K.S. McFarland,[f] **B. Messerly,**[r] **J. Miller,**[a] **L.M. Montano,**[s] **D. Naples,**[r] **J.K. Nelson,**[h]
C. Nguyen,[t] **A. Olivier,**[f] **V. Paolone,**[r] **G.N. Perdue,**[i,f] **M.A. Ramírez,**[m,g] **H. Ray,**[t]
D. Ruterbories,[f] **C.J. Solano Salinas,**[u] **H. Su,**[r] **M. Sultana,**[f] **V.S. Syrotenko,**[n] **E. Valencia,**[h,g]
M. Wospakrik,[t] **C. Wret,**[f] **K. Yang**[p] **and L. Zazueta**[h]

[a] *Departamento de Física, Universidad Técnica Federico Santa María,*
  Avenida España 1680 Casilla 110-V, Valparaíso, Chile

[b] *Centro Brasileiro de Pesquisas Físicas,*
  Rua Dr. Xavier Sigaud 150, Urca, Rio de Janeiro, Rio de Janeiro, 22290-180, Brazil

[c] *AMU Campus, Aligarh, Uttar Pradesh 202001, India*

[d] *Sección Física, Departamento de Ciencias, Pontificia Universidad Católica del Perú,*
  Apartado 1761, Lima, Perú

[e] *Department of Physics, Oregon State University, Corvallis, Oregon 97331, U.S.A.*

[f] *University of Rochester, Rochester, New York 14627 U.S.A.*

[g] *Campus León y Campus Guanajuato, Universidad de Guanajuato,*
  Lascurain de Retana No. 5, Colonia Centro, Guanajuato 36000, Guanajuato México.

[h] *Department of Physics, College of William & Mary, Williamsburg, Virginia 23187, U.S.A.*

[i] *Fermi National Accelerator Laboratory, Batavia, Illinois 60510, U.S.A.*

[j] *Department of Physics, University of Minnesota — Duluth, Duluth, Minnesota 55812, U.S.A.*

[k] *York University, Department of Physics and Astronomy, Toronto, Ontario, M3J 1P3 Canada*

[l] *Department of Physical Sciences, IISER Mohali,*
  Knowledge City, SAS Nagar, Mohali - 140306, Punjab, India

[m] *Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104*

[n] *Physics Department, Tufts University, Medford, Massachusetts 02155, U.S.A.*

[o] *Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854, U.S.A.*

[1] Now at University of Padova.

[*] Corresponding author.

^{p} *Oxford University, Department of Physics, Oxford, OX1 3PJ United Kingdom*

^{q} *Massachusetts College of Liberal Arts, 375 Church Street, North Adams, MA 01247*

^{r} *Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, U.S.A.*

^{s} *Departamento de Fisica, CINVESTAV,*
  Av. Centro de Investigación y de Estudios Avanzados del Instituto Politecnico Nacional,
  Col. San Pedro Zacatenco, 07360 Mexico City, DF, Mexico

^{t} *University of Florida, Department of Physics, Gainesville, FL 32611*

^{u} *Facultad de Ciencias, Universidad Nacional de Ingeniería, Apartado 31139, Lima, Perú*

  E-mail: anushree.ghosh@unipd.it

ABSTRACT: This paper presents a novel neutral-pion reconstruction that takes advantage of the machine learning technique of semantic segmentation using MINERvA data collected between 2013–2017, with an average neutrino energy of 6 GeV. Semantic segmentation improves the purity of neutral pion reconstruction from two γs from $70.7 \pm 0.9\%$ to $89.3 \pm 0.7\%$ and improves the efficiency of the reconstruction by approximately 40%. We demonstrate our method in a charged current neutral pion production analysis where a single neutral pion is reconstructed. This technique is applicable to modern tracking calorimeters, such as the new generation of liquid-argon time projection chambers, exposed to neutrino beams with $\langle E_\nu \rangle$ between 1–10 GeV. In such experiments it can facilitate the identification of ionization hits which are associated with electromagnetic showers, thereby enabling improved reconstruction of charged-current ν_e events arising from $\nu_\mu \rightarrow \nu_e$ appearance.

Contents

1 Introduction

MINERvA's extensive data set is used to investigate a wide variety of neutrino interaction channels, including both neutral and charged pion production. Neutral pion production is challenging to measure because the neutral pion reconstruction relies on identifying two photon-like objects in an event, and until now the primary approach to reduce backgrounds has been to simply cut on the neutral pion invariant mass. This is particularly true for the MINERvA dataset taken between 2013-17, called NuMI [1] Medium Energy Beam data, where most neutral pions are produced with a kinetic energy between 0.015–2.0 GeV at $\langle E_\nu \rangle \sim 6$ GeV. This dataset can be contrasted with the dataset for 2009-2012, called Low Energy Beam data, where most neutral pions are produced with a kinetic energy between 0.015–1.0 GeV at $\langle E_\nu \rangle \sim 3.5$ GeV. The increased number and energy of

particles produced by the neutrino interaction at $\langle E_\nu \rangle \sim 6\,\text{GeV}$ creates a complicated environment for neutral pion reconstruction. For the results presented here, we use a subset of the Medium Energy Beam data corresponding to 1.65×10^{20} protons on target.

Machine Learning (ML) reconstruction applications traditionally have taken one of two approaches [2–6]. In the first approach, the entire record from an event is provided to the machine learning algorithm (MLA) [7, 8]. This may prove problematic for exclusive or semi-exclusive measurements. In such measurements, many or most particles produced by the interaction are reconstructed. A challenge for this technique is the possibility that the different algorithms may claim conflicting hypotheses for the same information. In the second approach, which grew out of the log-likelihood based reconstructions [9, 10], only a subset of the data is provided to the MLA. Here the problem is no longer that information may be used to make incompatible hypotheses, but rather that valuable information may not be provided to the algorithms.

Deep Convolutional Neural Networks [11, 12] (DCNN) using techniques like semantic segmentation [13] make it possible to identify individual objects in an input image where pixels in the image are related by some regular metric (in [4] a distance weighted graph network is used for irregular geometry). These developments could revolutionize the reconstruction of events in some high energy particle physic experiments, in cases where the data can be represented by images with regular geometry, by allowing DCNN methods to model the activity in the detector as a step in the characterization of the interaction. This improves the measurement of energy that is deposited in the detector and motivates a third approach, where MLA provides context to the reconstruction [14].

In semantic segmentation, the MLA classifies individual bits of information (ionization hits in MINERvA) as one of several interesting classes. The classification then can be used as a filter for the hits used in the reconstruction algorithms (in [15] a CNN was used to construct a hit filter for track pixel seeds used as input to the CMS track reconstruction algorithm). The hits and information can be provided to further MLAs, however the clearest way to test the impact of the MLA responsible for semantic segmentation is to use it as a filter for the current state-of-the-art algorithm.

The ability to identify electromagnetic-like energy depositions within complicated event topologies is also important for the identification of electrons in charged-current electron-neutrino interactions. Improving the identification of such showers would improve both the efficiency and purity of measurements of electron neutrino appearance in current and future experiments, which are a crucial component of the worldwide effort to measure neutrino oscillations.

1.1 Overview of the paper

The paper is organized as follows. First, section 2 provides a short review of details of the MINERvA detector, details of the NuMI Medium Energy Beam data set and neutral pion production analyses in MINERvA. Section 3 describes the development of a machine learning algorithm for the semantic segmentation of energy deposition within the detector based on the class of particle which deposited that energy. Section 4 discusses the neutral pion reconstruction algorithm and an extension using semantic segmentation as a filter of electromagnetic-like hits. Section 5 presents conclusions and discusses implications for other experiments.

2 Background

The main goal of the MINERvA experiment is to measure neutrino-nucleus interaction cross-sections with a high degree of precision. Resonance production is the most abundant interaction mechanism at energies relevant to DUNE [16], a flagship neutrino experiment using a neutrino beam whose energy is comparable to that of MINERvA. Thus, measurements of baryon resonance production yielding final state pions are crucial.

In this work we study π^0 reconstruction algorithms applied to neutrino interactions on hydrocarbon (CH) target using the NuMI Medium Energy Beam dataset. To study the performance of these algorithms, we use a sample of candidates in the semi-exclusive channel $\nu_\mu + CH \longrightarrow \mu^- + \pi^0 +$ *nucleons*. Performance in an alternative semi-inclusive channel, where instead of just nucleons anything is allowed other than a neutral pion, is considered in appendix B.

2.1 The MINERvA Experiment

Neutrinos are produced by decays of charged pions produced by 120 GeV protons incident on a fixed carbon target and focused with two magnetic horns to create the NuMI beam. The muons associated with those same decays are stopped by approximately 200 meters of rock, leaving a beam of neutrinos. The MINERvA detector is located directly in front of the MINOS near detector [17, 18]. The core of the detector is a volume with a hexagonal cross section with 120 modules along the beam direction (z), 5 m long, and 1.7 m of apothem, as shown in figure 1. Starting from upstream and proceeding downstream, the full detector consists of multiple nuclear targets (He, Pb, Fe, C, H_2O), a central scintillator tracker inner detector, an electromagnetic calorimeter (ECAL), a hadronic calorimeter (HCAL), and the MINOS near detector, which serves as a magnetized iron spectrometer to identify the charge and momentum of muons.

The inner detector is made up of 8.3 tons of polystyrene (CH) strips divided into planes in three orientations: X, U and V (U and V are offset by ±60 degrees from the orientation of the X-view). These orientations allow the planes to be divided into three classes defining "views", the X-view, U-view and V-view. Each view can also be represented as a two dimensional image, i.e. the X-view can be represented as a x-z image. The planes are ordered in the pattern UXVXUXVX such that the sampling frequency of the X-view is double that of the U-view and V-view along the z-axis. Further details are available in [17, 19].

2.2 Neutrino-nucleus interactions

In pursuit of a suite of cross-section measurements covering $\nu_e/\bar{\nu}_e$ and $\nu_\mu/\bar{\nu}_\mu$ interactions [20–26], MINERvA has made a series of resonance production measurements including of neutral pion production in the NuMI Low Energy Beam dataset, $\langle E_\nu \rangle \sim 3.5$ GeV. Extending these measurements to the NuMI Medium Energy Beam dataset, $\langle E_\nu \rangle \sim 6$ GeV, provides an opportunity to expand the phase space of the earlier measurement, but also introduces challenges given the higher intensity and higher mean energy of the Medium Energy beam.

A neutrino interaction can produce a pion through several different processes. One process proceeds through baryon resonance production, in which the neutrino excites the target nucleon to a baryon resonance state that then decays to a nucleon and a pion in $\sim 10^{-23}$ seconds. Another process is deep inelastic scattering (DIS), in which the neutrino interacts with one of the quarks inside the
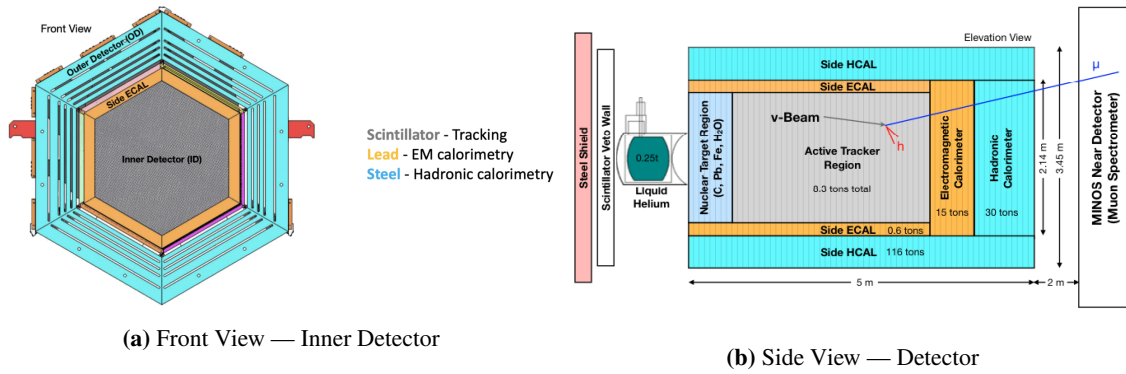
(a) Front View — Inner Detector

(b) Side View — Detector

Figure 1. a) The inner detector has four regions, which are the nuclear target region, active tracking region (tracker), downstream electromagnetic calorimeter (ECAL) and downstream hadronic calorimeter (HCAL) region. b) Side view of the MINERvA detector with highlighted regions: nuclear targets, active tracker, calorimeters and MINOS near detector. The nuclear target region consists of passive material (Pb, Fe and C) interspersed with scintillator. From [19].

nucleon and the quark then hadronizes to form one or multiple pions. The main characteristic of DIS events is that the struck quark is ejected from the nucleon as opposed to being merely transformed within the nucleon, unlike other pion production processes. Pion production from a nucleon not involving a resonance is commonly called a non-resonant process and it is treated as a sub-sample of DIS processes. A third process is coherent pion production from nuclei, a process in which a pion is created from neutrino-nucleus interaction where the target nucleus remains unchanged in its ground state after the scattering.

2.3 Neutral pion production in MINERvA

In the semi-exclusive neutral pion production analyses, the signal is $\pi^0 + X + \mu$ where X can not be a meson. An example of this analysis is given by [27]. This type of analysis provides the central context for our application and the discussion in section 4. Also examined in this work is a less selective analysis, here referenced as semi-inclusive with figures and details in appendix B, which is defined by a signal $\pi^0 + X + \mu$ where X is not a neutral pion. An example of a neutral pion production event relevant for semi-exclusive analyses is provided in figure 2.

Other charged current neutral pion production analyses, not examined in this work, are variations of the measurements described above. A more selective analysis is defined by a signal $\mu^- + \pi^0 + P + X$ where X can not be a meson or any energetic proton. An example of such an analysis is given by [28]. In other analyses X is minimal, such as in anti-neutrino analyses [29] and in analyses of neutral current neutral pion production [30].

2.4 GENIE and GEANT4 simulation in MINERvA

Detailed event simulation is essential for the development of reconstruction algorithms and analysis of MINERvA data. We use an event generator to simulate neutrino-nucleus interactions and a detector simulation to simulate the response of the MINERvA detector to the products of those interactions. Our event generator is GENIE v2.12.6 [32], and our detector simulation package is
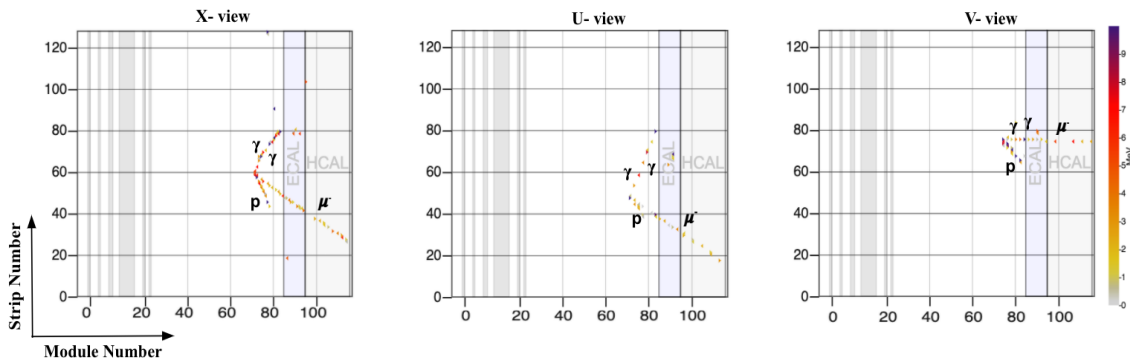
Figure 2. Arachne display of a simulated DIS event with a single neutral pion in the final state. Arachne is the MINERvA event viewer and is useful for developing and analyzing reconstructions, evaluating data quality during the MINERvA's data run, and providing a natural image representation of the MINERvA data, [31].

based on GEANT4 v9.4.2 [33]. GENIE models neutrino interactions, while the final state particles generated by GENIE are propagated through the MINERvA detector by GEANT4.

The charge current (CC) pion production processes relevant to this work proceed via baryon resonance production or by DIS. For the baryon resonance channels, GENIE uses the Rein and Sehgal model without interference between neighboring resonances [34]. It models DIS using the Bodek and Yang model [35]. The nuclear medium is simulated by GENIE based on a modified version of the relativistic Fermi gas model that includes short-range nucleon-nucleon correlations [36]. GENIE models final-state interactions of pions and nucleons (FSI) within the primary nucleus using an effective cascade model [37].

GENIE output for the final state particle kinematics after FSI is used as the input for GEANT4, which propagates particles in discrete steps. In each step, the energy deposited in the detector is calculated due to ionization and radiation based on the particle type. After each step, GEANT4 uses interaction models to determine whether any of the particles interacted in the detector. If an interaction occurs, GEANT4 produces outgoing particles according to the interaction model.

3 Semantic segmentation of energy deposition within MINERvA

Our machine learning algorithm was developed to provide context for further reconstructions in MINERvA, in particular the neutral pion reconstruction. We built a fully connected network (FCN) [13] to realize a semantic segmentation of the hits within the views of MINERvA. This provides additional information for each hit to physics reconstruction algorithms in MINERvA, such as the neutral pion reconstruction algorithm. While this machine learning algorithm was developed with the goal of providing context to the neutral pion reconstruction, it was developed in an analysis independent way, only assuming that the neutrino interaction was a charged current interaction where a muon was measured in the MINOS spectrometer. Alternative machine learning models are described appendix A. These includes ones where the requirement of a regular geometry is bent by including as additional layers the other MINERvA views, which are spatially shifted and which require extrapolation relative to the X-view.

3.1 Data description

The dataset used to train this MLA was collected as described in [38] but with an additional label. In [38] the data was simulated with flux from the Medium Energy configuration of the NuMI beam. The simulation has been tuned based on previous MINERvA analyses (for a recent description, [39]) but each analysis using sides bands for a final correction as a final step which has not been done here during the development of the reconstruction. The data shown represents approximately 10% of the total data available to the semi-exclusive neutral pion production analysis, and the MC statistics shown is about twice that of the data. The MC used is semi-inclusive and includes the target region.

The additional label vector was the full image but with each pixel labeled as "electromagnetic(EM)", "neutron", "non-electromagnetic(non-EM)", or "null". This label looks up the history of the highest energy deposited in that pixel and selects the source that produced that energy deposit, which is not necessarily the particle that was produced directly in the neutrino interaction. As such the label defines the source energy which was deposited in the detector and is not necessarily directly related to particles that were produced in the neutrino interaction. If the immediate history was a γ, e or π^0, the pixel was defined as being from the "EM" class. If the immediate history was a neutron, the pixel was defined as being from the "neutron" class. The remaining pixels were defined as "non-EM" unless there was no immediate history in which case they were classed as "null". This "null" class includes activity in the detector unrelated to the neutrino interaction. The majority of the "null" class pixels have no energy in the pixel, but some pixels assigned the "null" class may have energy due to cross-talk or other backgrounds (including activity in the detector from cosmic rays or residual activity from earlier neutrino interactions). Each pixel is assigned one and only one class.

3.2 Network description

In order to reconstruct neutral pions via the semantic segmentation approach, we labeled each pixel of the image as "EM", "Neutron" or "non-EM". The pixel was left without a label (NULL) if it was not given one of these three labels. This ML task can be described as multi-class classification with 3 classes. The energy originating from neutrons was explicitly labeled due to the possibility that hits from the neutron could be mistaken for a low energy γ shower and to provide the opportunity to reconstruct neutrons at a later stage. We differentiate the pixel labels from the simulation given in 3.1 from those that the classifier returns (EM, Neutron, non-EM and NULL) to emphasize that the classifier has 3 classes and not 4.

The MINERvA events are sparse, i.e. the majority of pixels in any given event, ($\sim 99\%$), are not assigned and are thus left unlabeled. The package LarCV [40] was adopted to save the views as an image in the ROOT [41] file format and to facilitate the use of the U-ResNet network [42, 43]. The implementation of the U-ResNet model used is described by figure 3. The U-ResNet model was initially developed within the MicroBooNE collaboration [44], and combines U-Net [45] with ResNet [46] to create a fully convolutional network for image or particle physics event segmentation. The skip connections of ResNet allow deeper networks with more expressive power, or the breadth of the domain of functions that the network can compute, to be trained without running into the problem of vanishing gradients during back-propagation based machine learning [47].

Each ResNet module consists of two convolution layers and each of these layers is followed by a batch normalization operation [48] and a rectified-linear unit (ReLU) activation function. In figure 3, black arrows indicate the direction of tensor data flow and the brown dashed lines indicate concatenation operations to combine the output of convolution layers from the encoding path to the decoding path. The final output has the same spatial dimension as the input with a depth defined by the number of labels.

As described in section 2, the data can be represented in three views: X, U and V, which may be interpreted as images. Each image has two channels: hit time and energy. Three separate semantic segmentation models were trained using the same algorithm with prediction applied to each view separately.

The detailed description of the production of images containing hit and energy information at the detector can be found in [38]. For the training of the whole detector configuration, both nuclear target and tracker regions were used. But for the semantic prediction and the neutral pion reconstruction, only the tracker region was used.

Figure 4 shows the confusion matrix obtained by training over one million simulated Monte Carlo (MC) events and validating over fifty thousand simulated events. The validation set is independent of the training sample but comes from the same MINERvA MC dataset. A confusion matrix represents the correlation between reconstructed and true values. The labels classified correctly according to the simulation are represented by the diagonal value, whereas miss-assignments are represented by the off-diagonal cells. Figure 4 represents the row-normalized matrix, which can be interpreted as the fraction of events reconstructed under a label (row value) that matches the one that truly originated the event. The values along the row-normalized matrix diagonal are above 70, which indicates the goodness of the network's ability to predict the data.

We applied semantic segmentation to the X-view only. We made this decision based on numerous factors, including the dominance of the X-view in the baseline reconstruction, the transparency of the impact of the application of the semantic segmentation to the neutral pion reconstruction and the requirement of a regular geometry for the use of a FCN for semantic segmentation. In the reconstruction of the neutral pion the X-view is reconstructed first and the sampling of the X-view is twice that of the other views, increasing the importance of the X-view for the reconstruction. Details about an alternative where we apply semantic segmentation in all three views is provided in appendix A. The neutral pion reconstruction would need to be modified to include application to the U and V views, complicating the comparison and evaluation.

Figure 5 shows the comparison between true MC events and their corresponding predictions from the MLA. The left panel of the figure 5 is a visualization of MC events using a web-based event viewer called Arachne [31]. The right panel of the figure 5 is the corresponding predicted hits of the event. The colors of the pixel signify the different labels where 0, 1, 2 and 3 represent label null, EM-like, neutron and non-EM-like, respectively.

4 Neutral pion reconstruction

Here we present the baseline neutral pion reconstruction algorithm in the context of the semi-exclusive analysis and compare it to the machine learning-based algorithm in a step-by-step manner. The baseline neutral pion reconstruction algorithm in MINERvA is based on the identification of
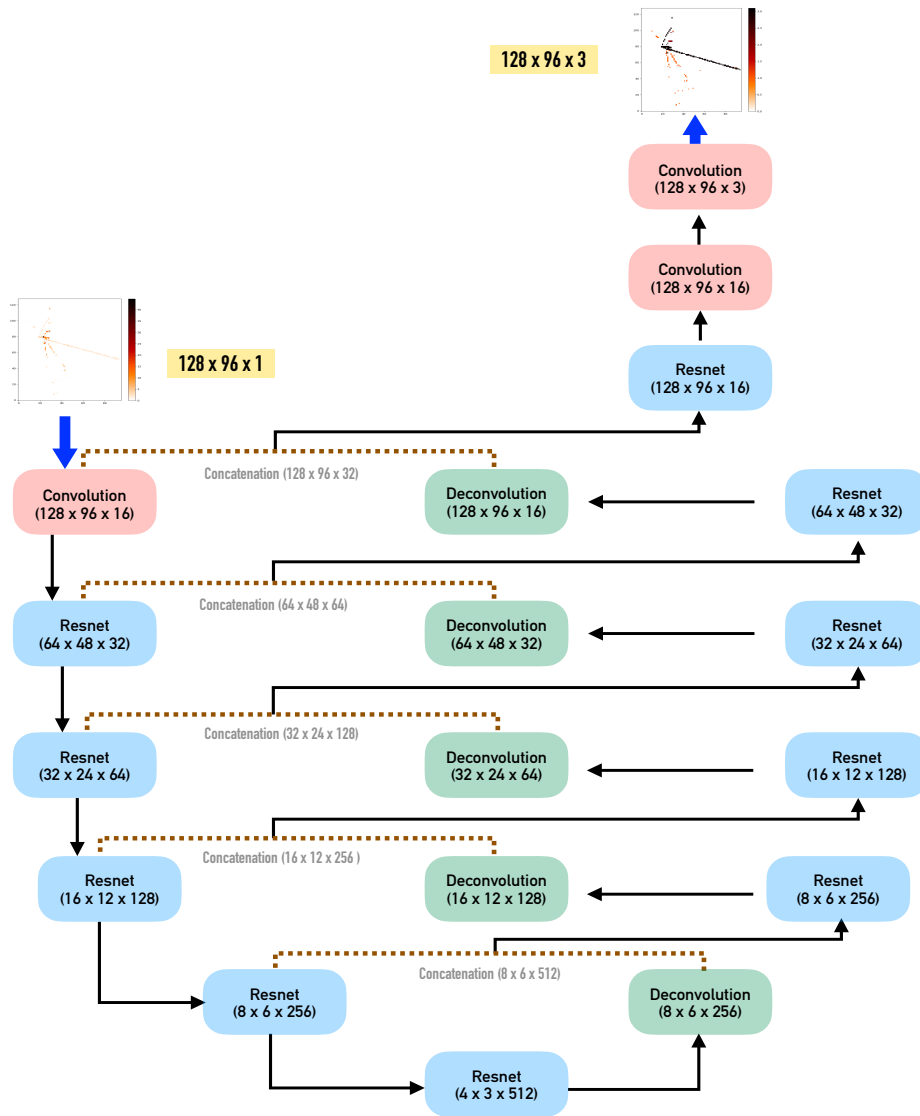
Figure 3. UResNet model architecture. The input image (left) is encoded using ResNet modules along the left before decoding on the right resulting in the semantic segmented image (top right). During decoding, the output of the encoding ResNet modules is combined at each decoding step. For each step, the width X height X (semantic) depth is provided. This diagram is for the energy only x-view with 4 class output.

two electromagnetic shower candidates as γ candidates produced via the decay $\pi^0 \to \gamma\gamma$. The neutral pion reconstruction algorithm is designed to identify and measure the energy and direction of the two electromagnetic candidate showers; the algorithm proceeds in four stages.

The first stage is a hit filter, which removes hits which are used by the muon or proton reconstructions, considered likely to be from optical or electronic cross-talk, or are in the HCAL. The remaining hits are available to the second stage, the electromagnetic shower candidate reconstruction. The only change for the machine learning-based algorithm is to modify this hit filter to remove hits which are determined by the machine learning algorithm to be not arising from electromagnetic activity.
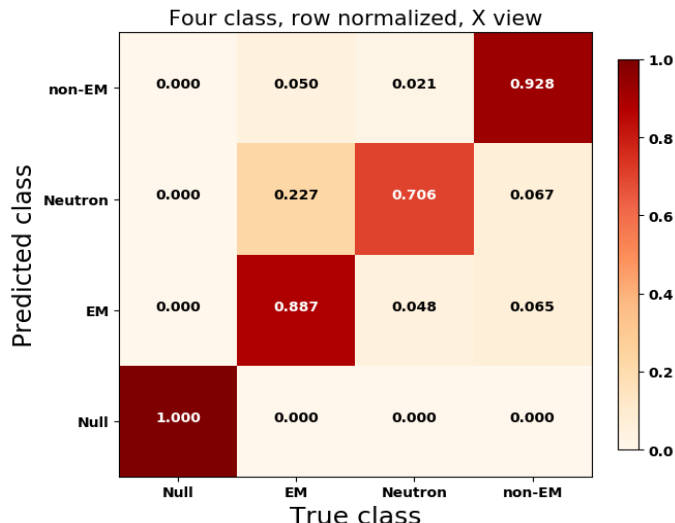
Figure 4. Row normalized confusion matrix with four labels: "null", EM, "neutron", non-EM. Only X-views are used. This confusion matrix is for a sample representative of the full training sample, including the nuclear target region.

In the second stage, the hits are grouped into candidate showers under the hypothesis that they originated from the electromagnetic cascade of an energetic γ originating in the neutrino interaction vertex.

While the neutral pion analyses primarily focus on the reconstruction of neutral pions in neutrino interactions with only a single neutral pion (and therefore two electromagnetic showers), reconstructions of one or three electromagnetic shower candidates are also kept. Three electromagnetic candidate reconstructions are kept due to the likelihood that one of these cannot be reconstructed into a γ candidate. Single electromagnetic candidate reconstructions are kept because the single electromagnetic candidate can possibly be divided into two overlapping showers.

The electromagnetic shower candidates are then available for the third stage, the γ candidate reconstruction. As part of the requirement to be able to be reconstructed as a γ, only shower candidates comprised of hits from two or three views are considered.

In the final stage, the reconstruction of the γ candidates from the electromagnetic candidates is realized. Here, properties of energetic γs are used, such as the conversion length in the scintillator and the average energy per unit length deposited by the shower candidate, $\frac{dE}{dX}$. Finally, the neutral pion candidate is reconstructed from the γ candidates using the invariant mass, $m_{\gamma\gamma} = \sqrt{2E_{\gamma 1}E_{\gamma 2}(1 - \cos\theta)}$ where $E_{\gamma 1}$ and $E_{\gamma 2}$ are the energy for γ candidate 1 (most energetic) and for γ candidate 2 (less energetic) respectively, θ is the angle between the γ candidates.

Each of the four stages is described in more detail below. Figure 5 shows an event that illustrates the challenge to γ and neutral pion production in multi-GeV CC inclusive interactions.

The procedure that we use for the application of semantic segmentation to the reconstruction of a neutral pion is to add an additional step to the hit selection, as described below in section 4.1. Having such a simple and obvious application of ML to an existing reconstruction has distinct advantages, including the advantage of transparency when assigning and propagating systematic uncertainties.

The simplicity and clearness of the application can be contrasted with the expressive complexity and opacity of the semantic segmentation model or even the relative algorithmic complexity of the Γ score used in the reconstruction of γ candidates. A benefit of applying semantic segmentation in such a way, instead of, for example, an a more complicated likelihood algorithm during the reconstruction, is that the reconstruction and associated uncertainties can be compared at every point in the reconstruction.
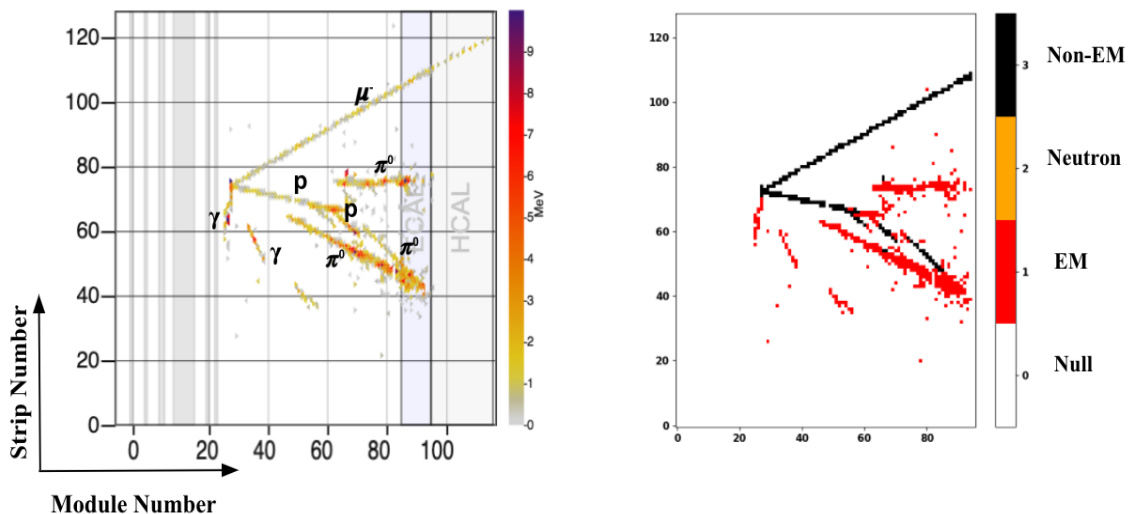

Figure 5. Left: a visualization of a simulated DIS neutral pion production event using the web-based tool Arachne [31] tool. Right: a semantically-segmented image of the same event. For the medium energy dataset, many events have γs with trajectories very close to each other, thereby increasing the difficulty of distinguishing γ pairs. The higher the π^0 energy is, the smaller the opening angle is, raising the complexity of distinguishing the γs' pair. A common approach when faced with this type of issue is by rejecting these mis-reconstructed events, which reduces the efficiency of pi^0 reconstruction.

4.1 Hit selection

Clusters are a set of nearby hits in the same plane; those not used to create tracks (generally charged pions, protons, and muons) are made available for the electromagnetic shower reconstruction. The hits in the cluster must have an associated time within 25 ns of the identified muon track. Additionally, low-activity clusters that have less than three photo-electrons per hit are removed, as they are most likely to originate from cross-talk. Moreover, clusters in the HCAL are not included in the available hits.

For the semantic segmentation based reconstruction, we place a requirement that the considered hits must have probability to be electromagnetic-like, as determined by the MLA, greater than a threshold of 0.5 which is the expected optimum due to the interpretation of the MLA output as a probability of the given class. This improves the purity of the sample during the formation of available clusters. This additional filter in the hit selection is the only difference between the baseline reconstruction and the reconstruction that uses ML.

4.2 Electromagnetic shower (EM) candidate reconstruction

Clusters are then assigned to electromagnetic showers via a Bayesian algorithm that groups the available clusters of the X-view into conical regions emanating from the neutrino interaction vertex. The assumption is that these clusters arise from an electromagnetic shower originating at the neutrino interaction vertex. In figure 6a we show the results of the Bayesian algorithm applied to the semi-exclusive analysis. This allows one to associate a group of clusters with a shower that is characterized by the angle between the shower and the longitudinal axis of the detector. These groupings of clusters are then considered in the next stage of the electromagnetic shower reconstruction.

Clusters of the U-view and V-view are then included to form electromagnetic shower candidates. The position of the centroid for the electromagnetic shower candidate is given by $\vec{R} = \frac{\sum_c E_c \vec{r}_c}{\sum_c E_c}$ where E_c is the energy of the cluster in photo-electrons and $\vec{r}_c$ is the position of the vector with respect to the interaction vertex.

The first topological condition for the electromagnetic shower candidate is that each candidate must have hits in at least two views. This allows three-dimensional reconstruction of the direction of the shower. In figure 6b these are referred to as "Selected Showers". The corresponding figures for the electromagnetic shower candidate reconstruction with ML are in figure 7 demonstrating improved purity.
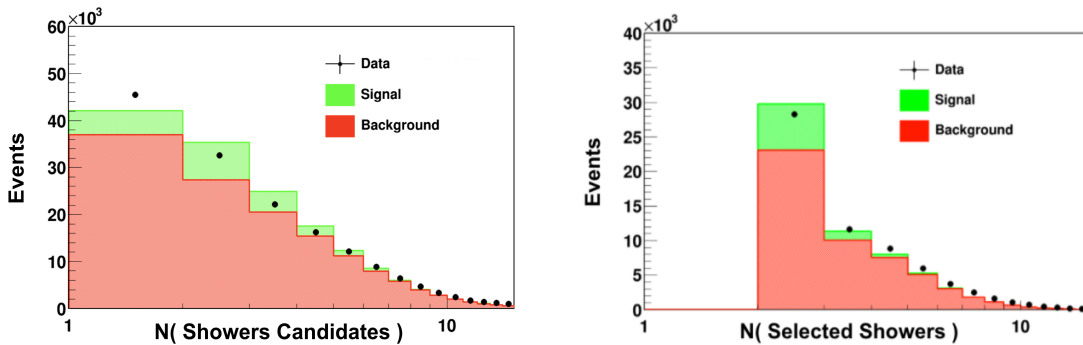

Figure 6. Baseline reconstruction in the semi-exclusive analysis: a) Number of electromagnetic shower candidates selected by a Bayesian algorithm. b) Showers from a) for which a three-dimension reconstruction is possible. The number of events in the simulation is normalized to the data. Systematic uncertainties are not shown.

4.3 γ reconstruction

The electromagnetic shower candidates are then reconstructed into γ candidates. As part of this reconstruction, a second topological condition is applied. This requirement, that the distance from the interaction vertex to the centroid of the shower candidate be at least 14 cm, is motivated by the conversion length of a γ in scintillator being 40 cm.

In analyses of the NuMI Low Energy Beam dataset [27–29] the second requirement was only applied to the highest energy γ candidate. In the NuMI Medium Energy Beam dataset, neutral pions are produced at higher energy, as are the other particles produced in the interaction. This results in complications in the reconstruction of γs due to the production of more particles in the
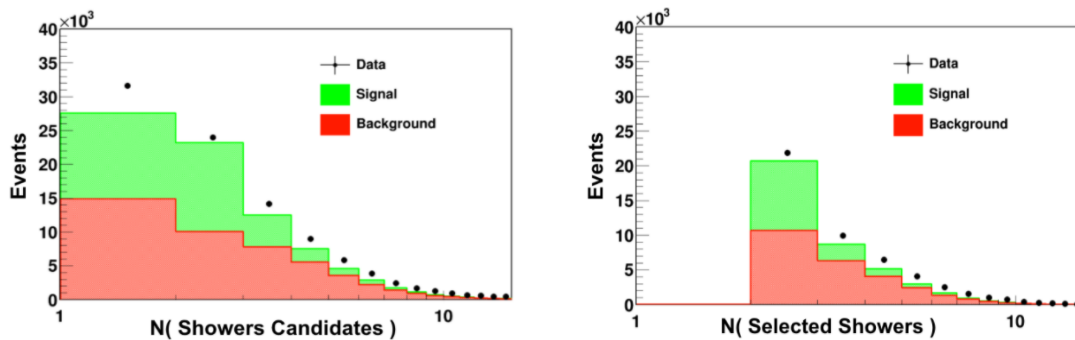
Figure 7. Machine learning reconstruction in the semi-exclusive analysis: a) Number of electromagnetic shower candidates selected by a Bayesian algorithm. b) Showers from a) for which a three-dimension reconstruction is possible. The number of events in the simulation is normalized to the data. Systematic uncertainties are not shown. Shown are the distributions before tuning of the Monte-Carlo simulation and determination of the systematic uncertainty.

neutrino interaction, increased interactions of energetic particles with the detector material, and the production of high energy γs through stochastic processes in the electromagnetic shower.

These complications, exacerbated by passive material in the detector, can result in the production of a high energy γ early in the electromagnetic shower, the production of more neutrons and the production of energetic particles not originating from the neutrino vertex. This produces more γ candidates during the electromagnetic shower reconstruction and increases the energy of the candidate γs, both due to γs being at higher energy and due to overlap between the electromagnetic shower from the γ and other energy from non-electromagnetic sources. Additionally, the more energetic neutral pions often have a narrower opening angle between the γs. The strategy that we have followed to address such challenges for neutral pion reconstruction in the NuMI Medium Energy Beam dataset is to improve the selection of γ candidates during the γ reconstruction as described below.

A third topological feature was developed for the γ candidates. This feature is based on the total deposited visible energy in the γ candidate per unit length, given by:

$$\frac{dE}{dx} = \frac{E_{\text{vis}}}{l} \tag{4.1}$$

where E_{vis} is the visible energy and l is the length of the γ candidate. Figures 8 and 9 show reconstructed energy versus $\frac{dE}{dx}$ for γ candidates using the baseline and ML-based reconstructions, respectively.

In both cases, γ candidates created by true photons have a clear correlation between reconstructed energy and $\frac{dE}{dx}$. The mean and standard deviation of the $\frac{dE}{dx}$ distributions for γ candidates produced by true photons are used to create a Γ score given by:

$$\Gamma_{\text{score}} = \frac{\frac{dE}{dx}_{\text{obs}} - \langle \frac{dE}{dx} \rangle_i}{\sigma_{\frac{dE}{dx}, i}}, \tag{4.2}$$

where $\frac{dE}{dx}_{\text{obs}}$ is the $\frac{dE}{dX}$ of a γ candidate and $\langle \frac{dE}{dx} \rangle_i$ and $\sigma_{\frac{dE}{dx}, i}$ are the mean and standard deviation of the simulated dE/dx distribution for true photons, evaluated in the reconstructed energy bin
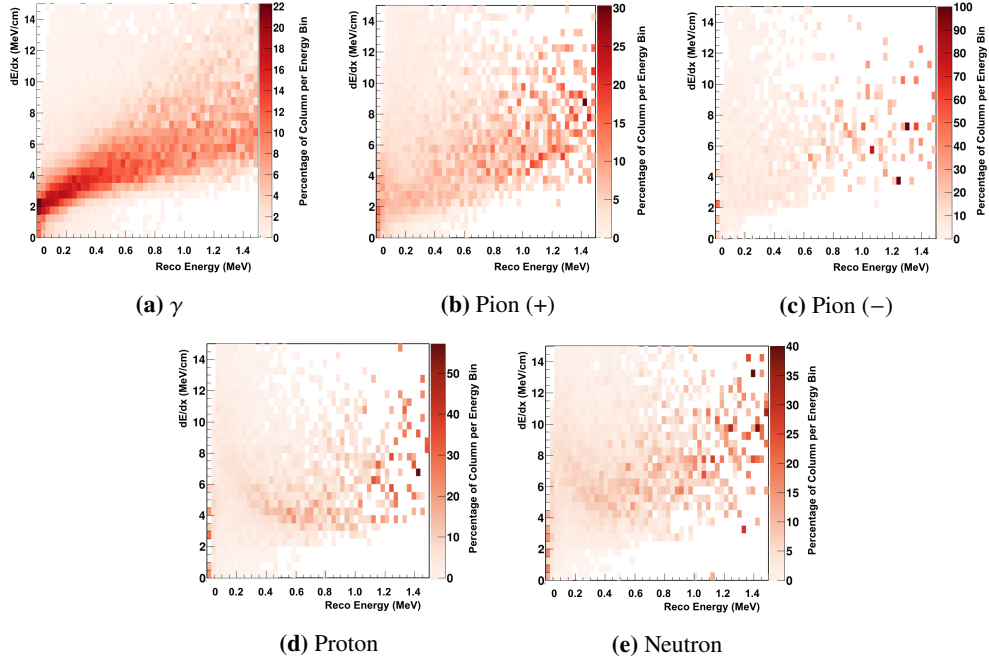
(a) γ **(b)** Pion $(+)$ **(c)** Pion $(-)$

(d) Proton   **(e)** Neutron

Figure 8. Reconstructed energy versus $\frac{dE}{dx}$ for the primary γ candidate in simulated events in the semi-exclusive analysis with **the baseline reconstruction**. The candidates are classified according to the particle which produced the electromagnetic shower, (a) to (e). There is a clear correlation between energy and $\frac{dE}{dx}$ for candidates produced by true photons (a).
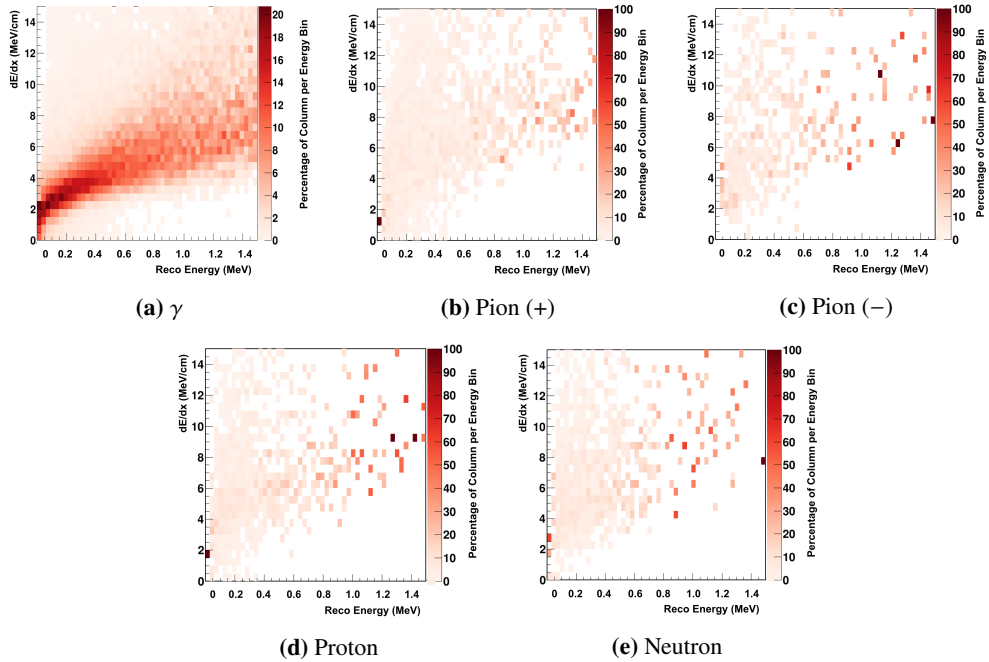
(a) γ **(b)** Pion $(+)$ **(c)** Pion $(-)$

(d) Proton   **(e)** Neutron

Figure 9. Reconstructed energy versus $\frac{dE}{dX}$ for the primary γ candidate in simulated events in the semi-exclusive analysis with **the ML reconstruction**. The candidates are classified according to the particle which produced the electromagnetic shower, (a) to (e). There is a clear correlation between energy and $\frac{dE}{dX}$ for candidates produced by true photons (a).

corresponding to the γ candidate. The distribution of γ candidate scores in the semi-exclusive analysis is shown in figure 10 for both the baseline and ML reconstruction.

For γ candidates that are not associated with photons, the Γs score is observed to be very asymmetric, while the score for true γs behaves approximately as a Gaussian. In the baseline reconstruction, this provides the third topological requirement that the γ candidate has a score of $\Gamma_{\mathrm{score}} < 2$. Having a score of $\Gamma_{\mathrm{score}} > 2$ indicates that the γ candidate is not likely to be a photon. For the ML-based reconstruction, Γ score is not a useful discriminator, presumably because $\frac{dE}{dx}$ has already been taken into account by the ML algorithm.

In figure 10 we see that the match between simulation and data is similar for both techniques for γ candidates with Γ score ≤ 2.0, but that data is more often under simulation for the baseline reconstruction and over simulation for the semantic segmentation based reconstruction in the peak area of -2 to 2. The neutral pion reconstruction includes the cut on Γ score of less than 2 and this is reflected in the final plots and numbers. In the high tail, with high uncertainty we see that, the semantic segmentation based reconstruction over predicts the data with a linear dependence on the Γ score.

Finally, the energy of the γ candidates is corrected by a scale factor to correct the simulation for the detector response due to electromagnetic activity. To do this we used energetic γs from the semi-inclusive neutral pion analysis; as described in [49]. Analysis of this scale factor is an important contribution to the systematic uncertainty of analyses which depend on measurements of electromagnetic activity.

4.4   Neutral pion reconstruction

The neutral pion analyses select events with one and only one neutral pion, so the first part of the neutral pion reconstruction is to require that two and only two reconstructed γs remain after γ reconstruction. The second component of the neutral pion reconstruction is the calculation of the neutral pion candidate invariant mass, $m_{\gamma\gamma}$, and selection based on that invariant mass. The neutral pion candidate invariant mass is given by

$$m_{\gamma\gamma} = \sqrt{2E_{\gamma 1}E_{\gamma 2}\left(1 - \cos\Theta_{\gamma\gamma}\right)} \tag{4.3}$$

where $E_{\gamma 1(2)}$ is the reconstructed energy of the most (least) energetic γ and $\Theta_{\gamma\gamma}$ is the opening angle between the reconstructed γs. The neutral pion invariant mass is shown in figure 11 for the semi-exclusive analysis, where we can see that the two γs are most often reconstructed into a neutral pion candidate with $60 < m_{\gamma\gamma} < 200\,\mathrm{MeV}$. There is also an evident shift between data and simulation whose source is currently unknown, but it is present with both the baseline and ML reconstructed samples, and is also present in the semi-inclusive analysis. That analysis has a different background than the semi-inclusive, due to both the change of signal definition and the change in event selections and is discussed in appendix B.

The simulated invariant mass distribution in figure 11 is divided according to the true identity of the particles creating the photon candidates. The percentage of events in each category is given in table 1 for the entire 0–700 MeV mass range and for the 60–200 MeV range where most true π^0 candidates reside. Within the 60–200 MeV range, $70.7 \pm 0.9\%$ of reconstructed π^0 candidates are reconstructed from two true photons in the simulation, while this number rises to $89.3 \pm 0.7\%$ with the ML-based algorithm.
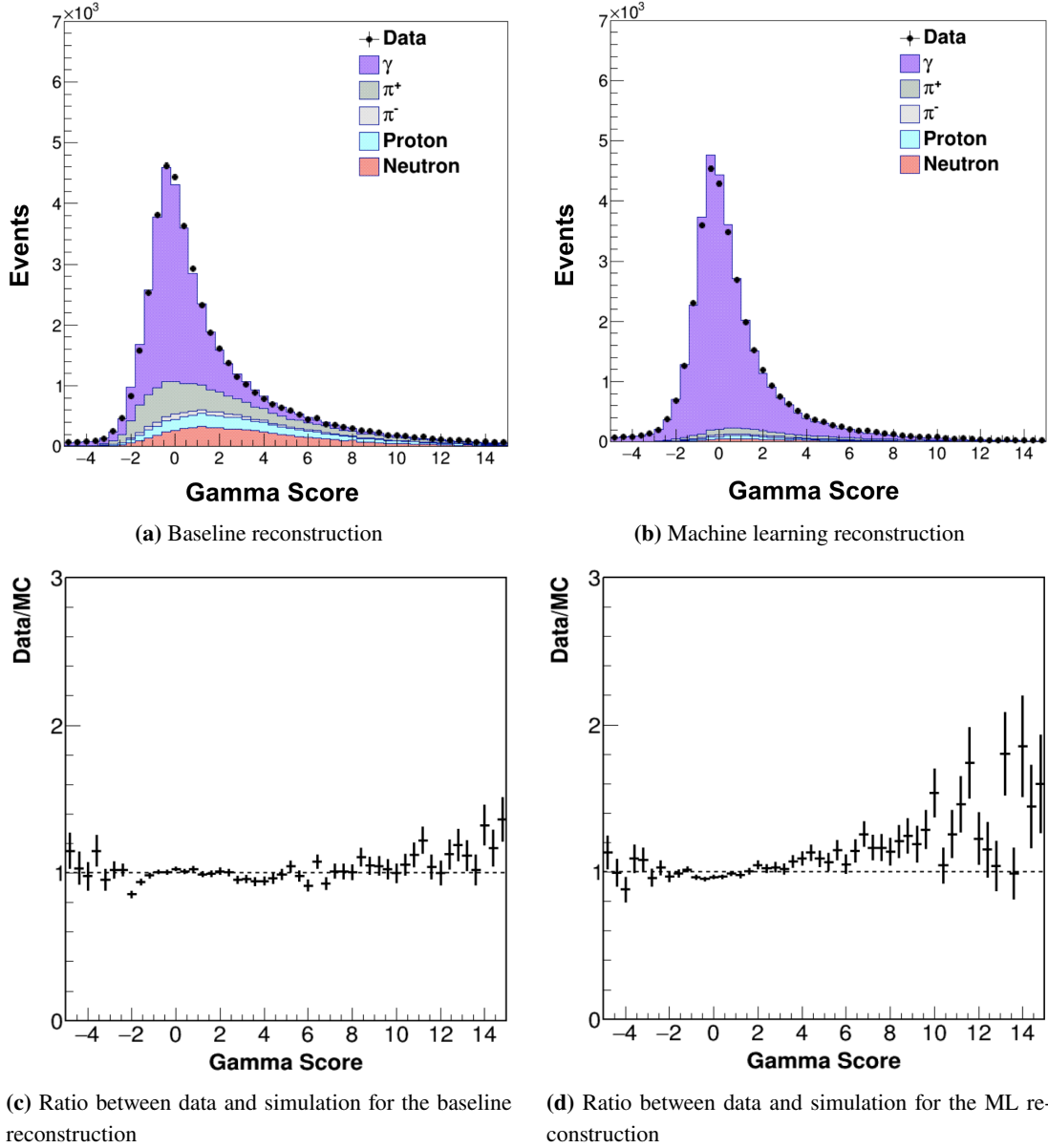
(a) Baseline reconstruction

(b) Machine learning reconstruction

(c) Ratio between data and simulation for the baseline reconstruction

(d) Ratio between data and simulation for the ML reconstruction

Figure 10. Γ_{score} for the semi-exclusive analysis in data and simulation. The simulation is classified based on the true particle that contributed the greatest fraction of energy to the γ candidate. In the baseline reconstruction a Γ score of less than 2 is a signature of a photon. When using the ML prediction, the Γ score is not a useful metric to improve the purity of the γ reconstruction. The number of events in the simulation is normalized to the data.

While the neutral pion reconstruction obviously also provides the neutral pion kinetic energy, momentum and angle with respect to the neutrino beam axis, this paper will not include a detailed discussion of such quantities other than to note the evident relationship between improving the reconstruction of the γs and improving the measurement of these kinematic quantities, which will be the subject of future MINERvA publications.

(a) Baseline reconstruction

(b) Machine learning reconstruction

(c) Ratio between data and simulation for the baseline reconstruction

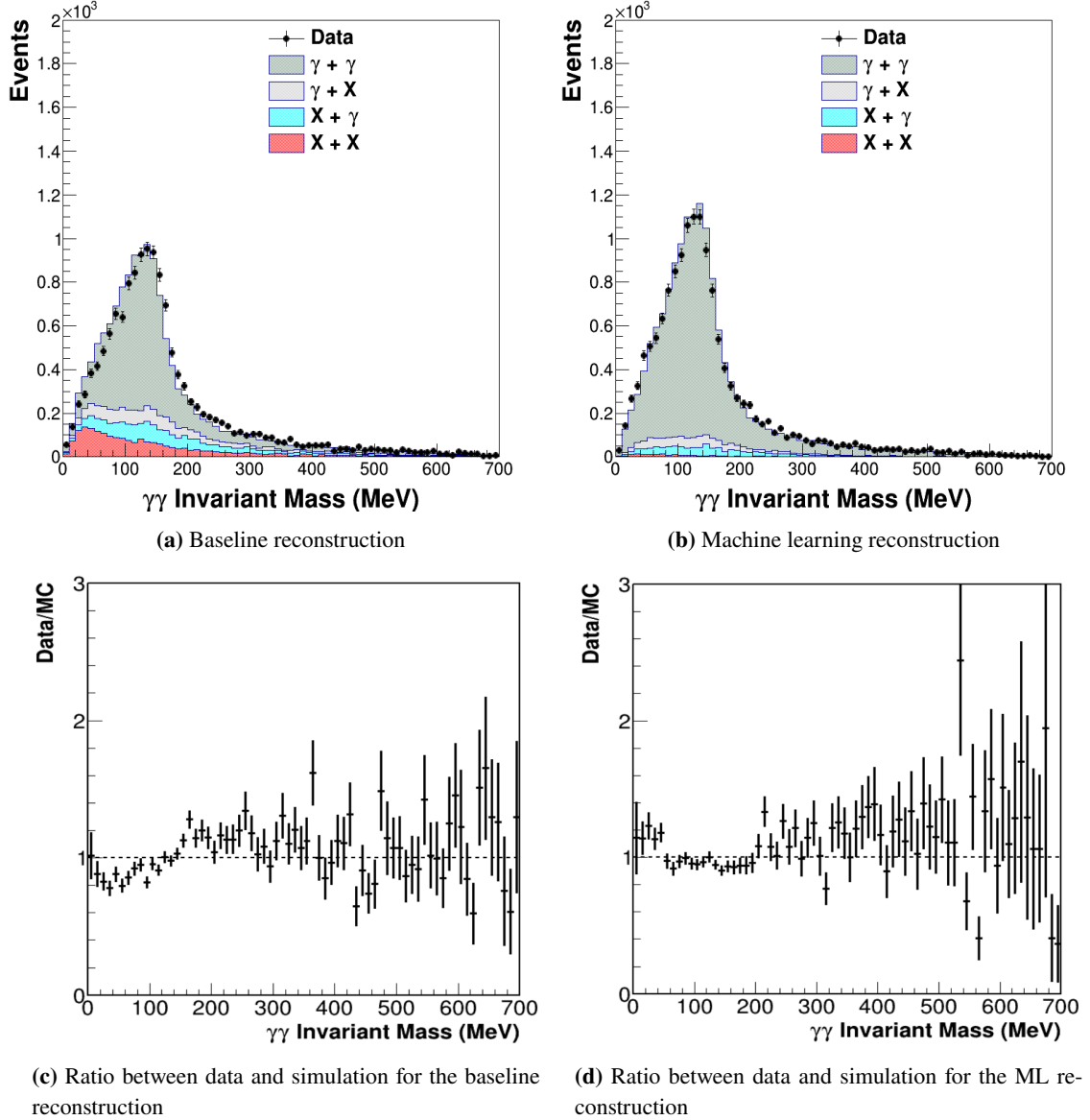**(d)** Ratio between data and simulation for the ML reconstruction

Figure 11. Neutral pion invariant mass distribution between [0 − 700] MeV for the semi-exclusive analysis in data and simulation. The simulation is classified based on the true particle that contributed the greatest amount of energy to the primary and secondary γ candidates. X corresponds to the absence of a γ. Note the significant improvement in the size of the $\gamma + \gamma$ category and decrease in all other categories when semantic segmentation provides context. From this result our final event selection are events between [60 − 200] MeV as shown on table 1. The number of events in the simulation is normalized to equal that of the data.

Event	Full Distribution	[0–700] MeV	Selection Distribution	[60–200] MeV
Category	Baseline	Machine Learning	Baseline	Machine Learning
$\gamma + \gamma$	62.5 ± 0.8	87.1 ± 0.6	70.7 ± 0.9	89.3 ± 0.7
$\gamma + X$	11 ± 2	6 ± 2	9 ± 2	5 ± 3
$X + \gamma$	12 ± 2	6 ± 3	10 ± 2	5 ± 3
$X + X$	14 ± 2	1^{+6}_{-1}	10 ± 2	1^{+7}_{-1}
Total	100	100	100	100

Table 1. Fraction (in per cent) of simulated events by category in the neutral pion invariant mass distributions between [0–700] MeV and between [60–200] MeV which is the region where the efficiency of the selection is better, from figure 11. uncertainties are statistical.

5 Discussion & conclusions

The new approach keeps more real neutral pion signal and rejects more background than the baseline approach. This is clearly observed in figures 10 and 11 (with associated table 1). The reconstruction purity is observed to increase from $70.7 \pm 0.9\%$ to $89.3 \pm 0.7\%$ in the signal region [60–200] MeV of figure 11 as shown in table 1, while the efficiency of the reconstruction increases by approximately 40%. The only change in the new approach is the application of semantic segmentation to provide context in the hit filter stage of the reconstruction. This improvement results in improvements to both the efficiency and purity in the charged current neutral pion production analysis, which is the subject of a forthcoming publication by MINERvA.

5.1 Discussion

A similar technique to the one described in section 3 was used by MicroBooNE, both for semantic segmentation of electromagnetic showers [43] and for the application of that semantic segmentation as a hit filter for a neutral pion reconstruction [50]. While the techniques are similar in that they are both based on semantic segmentation, the experimental environments are quite different. The mean energy of NuMI Medium Energy Beam dataset for MINERvA is $\langle E_\nu \rangle \sim 6$ GeV while for MicroBooNE $\langle E_\nu \rangle \sim 800$ MeV, causing MINERvA to have a much richer sample in the resonant and DIS region. The detector technologies, MINERvA's segmented scintillator versus MicroBooNE's liquid Argon TPC, are also very different. The success of the technique in both cases indicates that ML-based semantic segmentation can benefit π^0 reconstruction in a wide variety of neutrino detectors and beams.

The behavior with Γ score > 2 in figure 10b is consistent with an imperfect process, energy and spatial simulation being used to train the ML algorithm and/or used as the input to the score. While this distribution is primarily γ candidates, for the semi-exclusive pion production analysis it is primarily background and the discrepancy will be accounted for in the systematic uncertainty analysis. The significant reduction in the background contamination shown in table 1 should provide an improvement in the uncertainty analysis that should make up for the systematic associated with the ML algorithm which may be the source of this behavior. The reconstruction includes a cut removing candidates with Γ score > 2 and this is reflected in figure 11b.

While figure 10b shows a mismatch between data and simulation that increases linearly for a Γ score greater than 2, the hypothesis that this arises due to bias coming from the semantic

segmentation due to an unexpected mis-identification of hits caused by γs can be disfavored because such hypothesis would suggest a correction that is more proportional to the number of events or linear in Γ score. While linearity is seen above a Γ score of 2, it is not linear in the peak region between -2 and 2 where the vast majority of the γ candidates reside.

The sample in the semi-inclusive analysis is significantly different than that of the semi-exclusive analysis due to the exclusion of a significant fraction of events with at least one charged pion. That the same machine learning is able to provide a similarly strong filter for reconstruction of neutral pions from two γs for both the semi-inclusive and semi-exclusive analyses with a similar match between data and simulation provides additional evidence that the machine learning algorithm is successfully filtering out energy from charged pions.

5.2 Conclusions

We have described various methods for neutral pion reconstruction: one is the baseline approach, one is a ML-based approach. The paper describes those approaches for a semi-exclusive analysis, with a different approach presented in the appendix. These measurements were done in a beam with E_ν around 6 GeV. We have also described the development of a semantic segmentation-based filter on the hits used in the neutral pion reconstruction and demonstrated that it allows a significant reduction of mis-reconstructed γs and thus mis-reconstructed neutral pions. We expect significant increases in the final purity and efficiency of the analyses due to the improvement of the reconstruction purity from approximately $70.7 \pm 0.9\%$ to $89.3 \pm 0.7\%$ in the peak region (60 to 200 MeV) of the invariant mass distributions (uncertainty is statistical).

Additionally, we have seen that the semantic segmentation-based filter reduces instances where energy from other particles is included in the reconstructed pions. This allows the energy to be fully available to other reconstructions which enables an improved estimate of the neutrino energy. This approach thus provides a advantage over using multiple full event machine learning algorithms which may count features towards multiple quantities. It also is an advantage over the baseline reconstruction which lacked the sophistication necessary to separate energy from a γ from that of other particles within the reconstruction cone (see section 4.2).

This semantic segmentation based approach could be extended to include identifying the energy originating from protons, charged pions and more exotic hadrons. Such an extension would be particularly interesting for semi-exclusive measurements of neutrino-induced meson production. Semantic segmentation should be used to provide context to reconstructions in next generation experiments in High Energy Physics and should enable improved identification of ν_e appearance events.

Acknowledgments

A   Alternative Machine Learning models

We also studied an alternative model where the neutron was included in the "other" class during the training of the ML algorithm. In this study, the data was labeled with three classes and the segmentation was done by labeling each pixel as being one of two classes: "electromagnetic" (EM-like) or "non-electromagnetic" (non-EM-like). The "null" labeled pixels were left unlabeled. In this study the ML task is described as multi-class classification with 2 classes. This model more closely follows the approach described in [43].

In another alternative model, we also combined those three images into one by layering them [51]. We trained one network using those combined images and obtained the prediction for three views separately. In figure 13 we show result of the layered 4-class semantic segmentation model for one charged current neutral pion production event.

Figure 12a shows the confusion matrix obtained by training over images with three labels such "null", EM-like and non-EM-like. On the other hand, figure 4 shows the confusion matrix with four labels such as "null", EM-like, "neutron" and non-EM-like. The comparison between these two figures brings the conclusion that with four labels, the purity for EM-like hits improves by 5% where as for non-EM-like hits the purity decreases by 3%. Figure 12b shows the confusion matrix where the images are formed by layered three views together ƒand are labeled into four classes such as "null", EM-like, "neutron" and non-EM-like. Comparing figure 12b with figure 4 we see that the purity for EM-like hits and non-EM-like hits are decreased by 5% and 3% respectively for the combined image case whereas for "neutron" the purity is increased by 10%. An example of an event with a neutron is provided in figure 14.

From our studies of alternative semantic segmentation models, including the "neutron" class improves the performance but layering did not, at least not to an X-view only application. Additionally, the paucity of hits in U and V may complicate the application of the semantic segmentation due to the baseline reconstruction reconstructing candidate showers in the U and V-views before combining them into a candidate 3D shower. A better application of semantic segmentation for U and V may be as a correction to the energy of the candidate γ and not as a hit filter.

B   Semi-inclusive analysis

In figure 15 we see the invariant mass for both the baseline reconstruction and the reconstruction using the ML prediction for the semi-inclusive analysis. The semi-inclusive analysis is similar to the semi-exclusive analysis, one exception being that the semi-exclusive analysis has additional filters
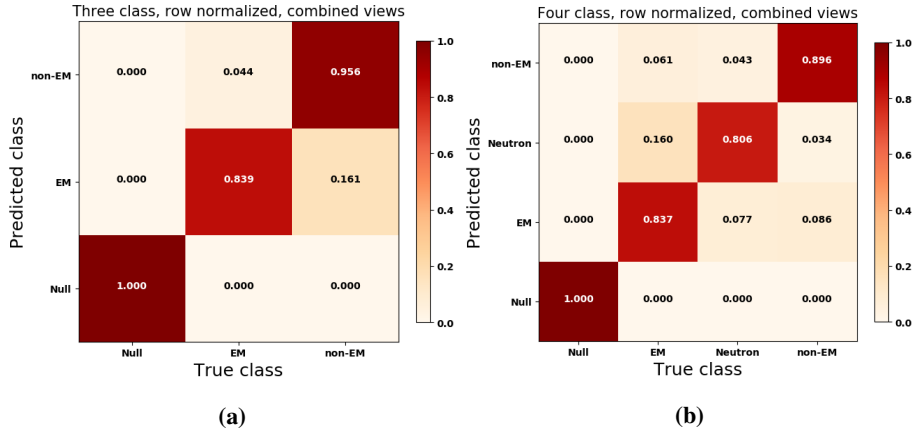
Figure 12. (a) Row-normalized confusion matrix with three labels: "Null", EM-like, non-EM-like. Only X-views are used. (b) Row Normalized confusion matrix with four labels: "null", EM-like, "neutron", non-EM-like shower where three views (X,U and V) are combined. These confusion matrices are for a sample representative of the full training sample, including the nuclear target region.
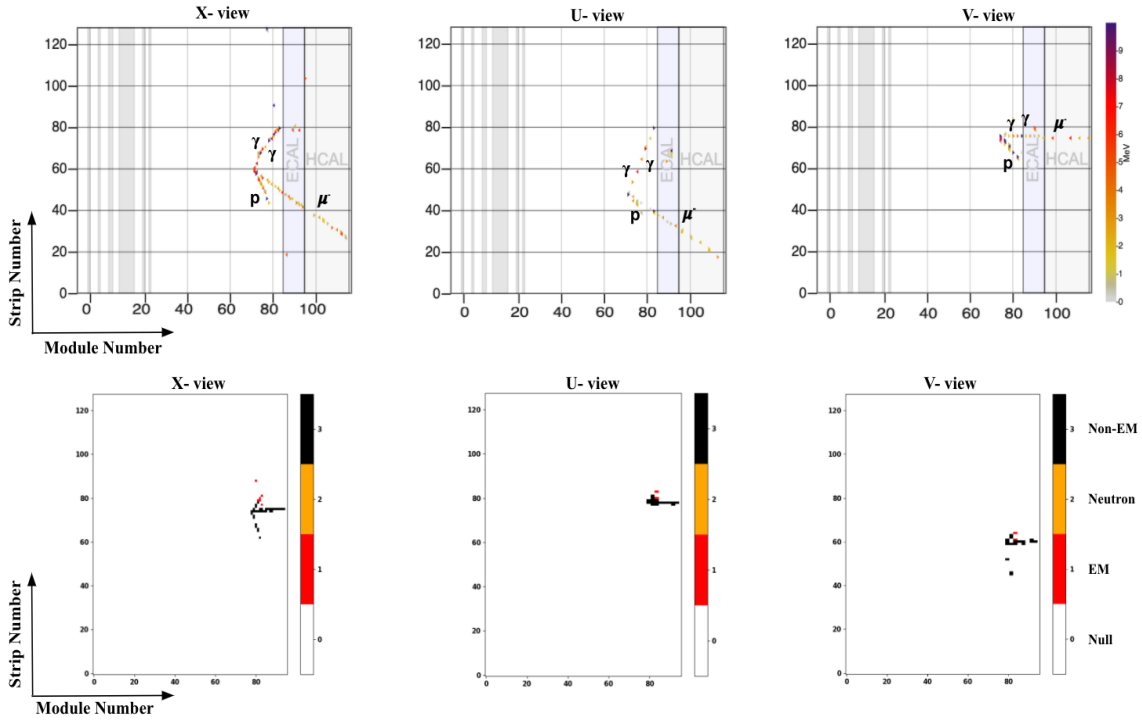

Figure 13. The lower panel represents the corresponding predicted images obtained by ML-based approach where the input images are formed by layered the three views together. The color bar represents the integer values of each class.

designed to remove events with charged mesons, especially the positively charged pion (π^+). By comparing figure 15 with figure 11, we see that the additional filters in the semi-exclusive analysis remove a significant number of the mis-reconstructions from non-γs. However, the reconstructions using semantic segmentation have a similar purity for both analyses.
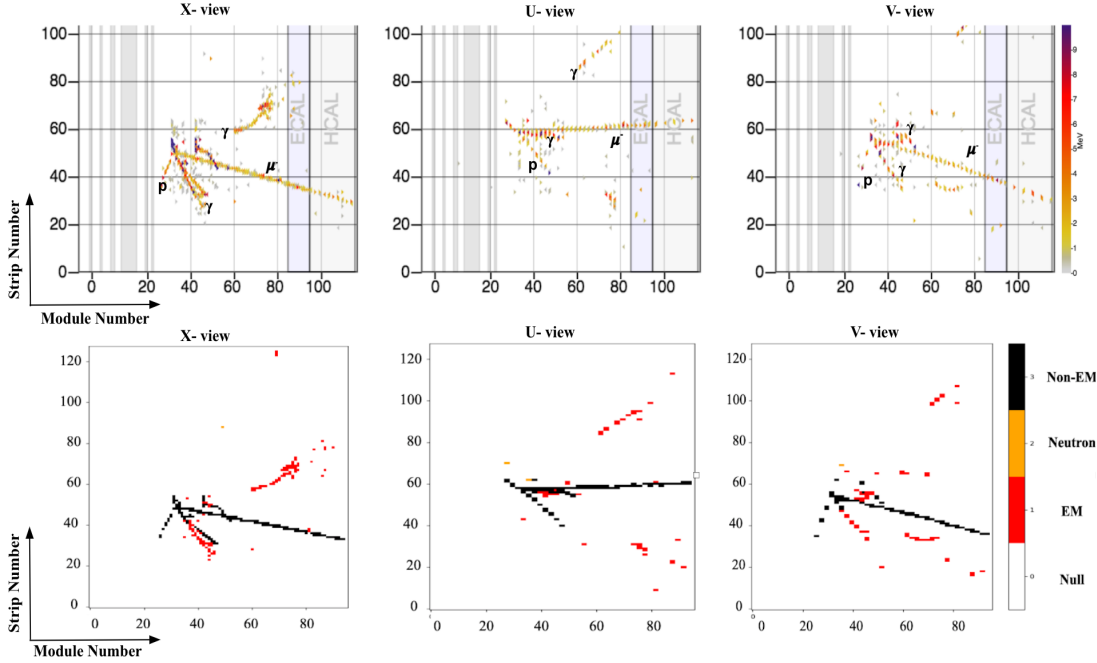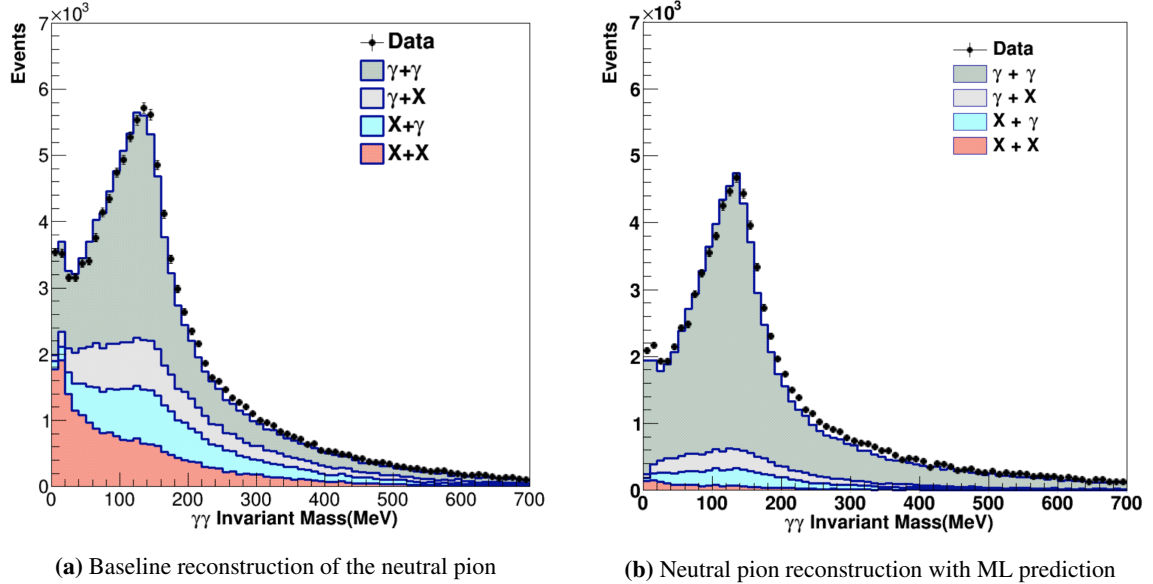
Figure 14. Simulation of neutral pion production in MINERvA for DIS event with a neutron in the FS. Here the three columns show the X, U and V-view of an event respectively. The upper panel is the visualization of a MC event on web-based tool Arachne. The lower panel represents the corresponding predicted images obtained by ML-based approach where the input images are formed by layered the three views together. The color bar represents the integer values of each class.

(a) Baseline reconstruction of the neutral pion

(b) Neutral pion reconstruction with ML prediction

Figure 15. Neutral pion invariant mass distribution between [0–700] MeV for the semi-inclusive analysis in data and simulation. The simulation is classified based on the true particle that contributed the greatest amount of energy to the primary and secondary γ candidates. X corresponds to the absence of a γ. Note the significant improvement in the size of the $\gamma + \gamma$ category and decrease in all other categories when semantic segmentation provides context. The number of events in the simulation is normalized to equal that of the data.

Additionally, despite the very disparate samples in the two analyses, the application of semantic segmentation to the reconstruction does not significantly impact the agreement between data and simulation. This provides of evidence for the success of our thesis, that semantic segmentation provides a strong hit filter which allows for the reconstruction of neutral pions independent of the existence of other energy within the detector which would normally have impacted the reconstruction.

References

[1] P. Adamson et al., *The NuMI Neutrino Beam*, *Nucl. Instrum. Meth. A* **806** (2016) 279 [`arXiv:1507.06690`].

[2] F. Psihas, M. Groh, C. Tunnell and K. Warburton, *A Review on Machine Learning for Neutrino Experiments*, *Int. J. Mod. Phys. A* **35** (2020) 2043005 [`arXiv:2008.01242`].

[3] D. Guest, K. Cranmer and D. Whiteson, *Deep Learning and its Application to LHC Physics*, *Ann. Rev. Nucl. Part. Sci.* **68** (2018) 161 [`arXiv:1806.11484`].

[4] S.R. Qasim, J. Kieseler, Y. Iiyama and M. Pierini, *Learning representations of irregular particle-detector geometry with distance-weighted graph networks*, *Eur. Phys. J. C* **79** (2019) 608 [`arXiv:1902.07987`].

[5] M. Andrews, M. Paulini, S. Gleyzer and B. Poczos, *End-to-End Physics Event Classification with CMS Open Data: Applying Image-Based Deep Learning to Detector Data for the Direct Classification of Collision Events at the LHC*, *Comput. Softw. Big Sci.* **4** (2020) 6 [`arXiv:1807.11916`].

[6] D. Belayneh et al., *Calorimetry with deep learning: particle simulation and reconstruction for collider physics*, *Eur. Phys. J. C* **80** (2020) 688 [`arXiv:1912.06794`].

[7] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M.D. Messier, E. Niner et al., *A Convolutional Neural Network Neutrino Event Classifier*, 2016 *JINST* **11** P09001 [`arXiv:1604.01444`].

[8] MINOS collaboration, *Measurement of single π^0 production by coherent neutral-current ν Fe interactions in the MINOS Near Detector*, *Phys. Rev. D* **94** (2016) 072006 [`arXiv:1608.05702`].

[9] NOvA collaboration, *Measurement of neutrino-induced neutral-current coherent π^0 production in the NOvA near detector*, *Phys. Rev. D* **102** (2020) 012004 [`arXiv:1902.00558`].

[10] MINERvA collaboration, *Charged Pion Production in ν_μ Interactions on Hydrocarbon at $\langle E_\nu \rangle =$ 4.0 GeV*, *Phys. Rev. D* **92** (2015) 092008 [`arXiv:1406.6415`].

[11] Y. LeCun, L. Jackel, L. Bottou, A. Brunot, C. Cortes, J. Denker et al., *Comparison of learning algorithms for handwritten digit recognition*, in proceedings of *International conference on artificial neural networks*, Paris, France, 9–13 October 1995, pp. 53–60.

[12] A.S. Razavian, H. Azizpour, J. Sullivan and S. Carlsson, *CNN features off-the-shelf: An astounding baseline for recognition*, in proceedings of *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Columbus, OH, U.S.A., 23–28 June 2014, pp. 512–519 [`arXiv:1403.6382`].

[13] J. Long, E. Shelhamer and T. Darrell, *Fully convolutional networks for semantic segmentation*, in proceedings of *IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, U.S.A., 7–12 June 2015, pp. 3431–3440 [`arXiv:1411.4038`].

[14] F. Psihas, E. Niner, M. Groh, R. Murphy, A. Aurisano, A. Himmel et al., *Context-Enriched Identification of Particles with a Convolutional Network for Neutrino Events*, *Phys. Rev. D* **100** (2019) 073005 [arXiv:1906.00713].

[15] CMS collaboration, *Convolutional Neural Network for Track Seed Filtering at the CMS High-Level Trigger*, *J. Phys. Conf. Ser.* **1085** (2018) 042040.

[16] DUNE collaboration, *The DUNE Far Detector Interim Design Report Volume 1: Physics, Technology and Strategies*, arXiv:1807.10334.

[17] MINERvA collaboration, *The MINERvA Data Acquisition System and Infrastructure*, *Nucl. Instrum. Meth. A* **694** (2012) 179 [arXiv:1209.1120].

[18] MINOS collaboration, *The Magnetized steel and scintillator calorimeters of the MINOS experiment*, *Nucl. Instrum. Meth. A* **596** (2008) 190 [arXiv:0805.3170].

[19] MINERvA collaboration, *Design, Calibration, and Performance of the MINERvA Detector*, *Nucl. Instrum. Meth. A* **743** (2014) 130 [arXiv:1305.5199].

[20] MINERvA collaboration, *Measurement of Muon Antineutrino Quasielastic Scattering on a Hydrocarbon Target at Ev~3.5 GeV*, *Phys. Rev. Lett.* **111** (2013) 022501 [arXiv:1305.2234].

[21] MINERvA collaboration, *Measurement of Muon Neutrino Quasielastic Scattering on a Hydrocarbon Target at $E_\nu \sim 3.5\,GeV$*, *Phys. Rev. Lett.* **111** (2013) 022502 [arXiv:1305.2243].

[22] MINERvA collaboration, *Measurement of Ratios of ν_μ Charged-Current Cross Sections on C, Fe, and Pb to CH at Neutrino Energies 2-20 GeV*, *Phys. Rev. Lett.* **112** (2014) 231801 [arXiv:1403.2103].

[23] MINERvA collaboration, *Measurement of electron neutrino quasielastic and quasielasticlike scattering on hydrocarbon at $\langle E_\nu \rangle = 3.6\,GeV$*, *Phys. Rev. Lett.* **116** (2016) 081802 [arXiv:1509.05729].

[24] MINERvA collaboration, *Measurement of Neutrino Flux from Neutrino-Electron Elastic Scattering*, *Phys. Rev. D* **93** (2016) 112007 [arXiv:1512.07699].

[25] MINERvA collaboration, *Measurement of Quasielastic-Like Neutrino Scattering at $\langle E_\nu \rangle \sim 3.5\,GeV$ on a Hydrocarbon Target*, *Phys. Rev. D* **99** (2019) 012004 [arXiv:1811.02774].

[26] MINERvA collaboration, *High-Statistics Measurement of Neutrino Quasielasticlike Scattering at 6 GeV on a Hydrocarbon Target*, *Phys. Rev. Lett.* **124** (2020) 121801 [arXiv:1912.09890].

[27] MINERvA collaboration, *Measurement of ν_μ charged-current single π^0 production on hydrocarbon in the few-GeV region using MINERvA*, *Phys. Rev. D* **96** (2017) 072003 [arXiv:1708.03723].

[28] MINERvA collaboration, *Probing nuclear effects with neutrino-induced charged-current neutral pion production*, *Phys. Rev. D* **102** (2020) 072007 [arXiv:2002.05812].

[29] MINERvA collaboration, *Single Neutral Pion Production by Charged-Current $\bar{\nu}_\mu$ Interactions on Hydrocarbon at $\langle E_\nu \rangle = 3.6\,GeV$*, *Phys. Lett. B* **749** (2015) 130 [arXiv:1503.02107].

[30] MINERvA collaboration, *Evidence for Neutral-Current Diffractive π^0 Production from Hydrogen in Neutrino Interactions on Hydrocarbon*, *Phys. Rev. Lett.* **117** (2016) 111801 [arXiv:1604.01728].

[31] MINERvA collaboration, *Arachne - A web-based event viewer for MINERvA*, *Nucl. Instrum. Meth. A* **676** (2012) 44 [arXiv:1111.5315].

[32] C. Andreopoulos et al., *The GENIE Neutrino Monte Carlo Generator*, *Nucl. Instrum. Meth. A* **614** (2010) 87 [arXiv:0905.2517].

[33] GEANT4 collaboration, *GEANT4–a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.

[34] D. Rein and L.M. Sehgal, *Neutrino Excitation of Baryon Resonances and Single Pion Production*, *Annals Phys.* **133** (1981) 79.

[35] A. Bodek and U.K. Yang, *Higher twist, xi(omega) scaling, and effective LO PDFs for lepton scattering in the few GeV region*, *J. Phys. G* **29** (2003) 1899 [`hep-ex/0210024`].

[36] A. Bodek and J.L. Ritchie, *Further Studies of Fermi Motion Effects in Lepton Scattering from Nuclear Targets*, *Phys. Rev. D* **24** (1981) 1400.

[37] S.A. Dytman and A.S. Meyer, *Final state interactions in GENIE*, *AIP Conf. Proc.* **1405** (2011) 213.

[38] MINERvA collaboration, *Reducing model bias in a deep learning classifier using domain adversarial neural networks in the MINERvA experiment*, 2018 *JINST* **13** P11020 [`arXiv:1808.08332`].

[39] MINERvA collaboration, *Tuning the GENIE Pion Production Model with MINERvA Data*, *Phys. Rev. D* **100** (2019) 072005 [`arXiv:1903.01558`].

[40] LarCV, https://github.com/DeepLearnPhysics/larcv2, 2017

[41] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth. A* **389** (1997) 81.

[42] U-ResNet, https://github.com/DeepLearnPhysics/u-resnet, 2018

[43] MicroBooNE collaboration, *Deep neural network for pixel-level electromagnetic particle identification in the MicroBooNE liquid argon time projection chamber*, *Phys. Rev. D* **99** (2019) 092001 [`arXiv:1808.07269`].

[44] MicroBooNE Collaboration, https://microboone-exp.fnal.gov/.

[45] O. Ronneberger, P. Fischer and T. Brox, *U-net: Convolutional networks for biomedical image segmentation*, in proceedings of *Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany, 5–9 October 2015, pp. 234–241 [`arXiv:1505.04597`].

[46] K. He, X. Zhang, S. Ren and J. Sun, *Deep Residual Learning for Image Recognition*, in proceedings of *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, U.S.A., 27–30 June 2016, pp. 770–778 [`arXiv:1512.03385`].

[47] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, *Gradient-based learning applied to document recognition*, *Proc. IEEE* **86** (1998) 2278.

[48] S. Ioffe and C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, in proceedings of *32nd International Conference on International Conference on Machine Learning*, Lille, France, 6–11 July 2015, pp. 448–456 [`arXiv:1502.03167`].

[49] MINERvA collaboration, *Constraint of the MINERvA medium energy neutrino flux using neutrino-electron elastic scattering*, *Phys. Rev. D* **100** (2019) 092001 [`arXiv:1906.00111`].

[50] MicroBooNE collaboration, *Reconstruction and Measurement of O(100) MeV Energy Electromagnetic Activity from $\pi^0 \to \gamma\gamma$ Decays in the MicroBooNE LArTPC*, 2020 *JINST* **15** P02007 [`arXiv:1910.02166`].

[51] L. Song, F. Chen, S.R. Young, C.D. Schuman, G. Perdue and T.E. Potok, *Deep Learning for Vertex Reconstruction of Neutrino-Nucleus Interaction Events with Combined Energy and Time Data*, in proceedings of *IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, U.K., 12–17 May 2019, pp. 3882–3886 [`arXiv:1902.00743`].