A Convolutional Neural Network for Multiple Particle Identification in the MicroBooNE Liquid Argon Time Projection Chamber

MicroBooNE Collaboration

June 2020

MICROBOONE-NOTE-1080-PUB

Abstract

We present the multiple particle identification (MPID) network, a convolutional neural network (CNN) for multiple object classification, developed by MicroBooNE. MPID provides the probabilities of e^-, γ, μ^-, $\pi^\pm$, and protons in a single liquid argon time projection chamber (LArTPC) readout. The network extends the single particle identification network previously developed by MicroBooNE [1]. MPID takes as input an image either cropped around a reconstructed interaction vertex or containing only activity connected to a reconstructed vertex, therefore relieving the tool from inefficiencies in vertex finding and particle clustering. The network serves as an important component in MicroBooNE's deep learning based ν_e search analysis. In this note, we present the network's design, training, and performance on simulation and data from the MicroBooNE detector.

Contents

1 Introduction

A series of liquid argon time projection chamber (LArTPC) detectors have been or are being deployed at Fermilab as part of the Short-Baseline Neutrino (SBN) program [2] along the Booster Neutrino Beamline (BNB [3]) and as part of the long-baseline program of the Deep Underground Neutrino Experiment (DUNE) [4]. The MicroBooNE experiment [5], part of the Fermilab SBN program, has been operating since 2015, collecting data accumulated during beam-on and beam-off time periods.

MicroBooNE operates a 170 ton (87 ton active) LArTPC placed 470 m from the BNB target at Fermilab. The LArTPC is 10.3 m long, 2.5 m wide and 2.4 m high. The detector has three readout wire planes with 2400 readout wires on the two induction planes and 3456 readout wires on the collection plane [6]. Wires are installed with two induction planes oriented at $\pm 60°$ with respect to the vertical collection plane at a wire pitch of 3 mm. An array of 32 PMTs are installed behind the collection plane to detect the scintillation light from argon ionization caused by charged final state particles from neutrino interactions [7]. The TPC readout time window is 4.8 ms and is digitized into 9600 readout time ticks. Charged particles in liquid argon produce ionization electrons, which drift to the readout wire planes in an electric field of 273 V/cm. It requires 2.3 ms for an ionization electron to drift across the full width of the detector.

The MicroBooNE LArTPC continuously records charge drifted and its arrival time on each wire. A software trigger, based on PMT signals, records an event triggered by the BNB beam spill if the interaction light detected by the PMT array is above a set threshold. Each event consists of data collected from 1.6 ms before the trigger and 3.2 ms after the trigger. Therefore, each event has three sets of TPC data for each wire on all three planes. The two induction planes have resolutions of 2400 wires $\times$ 6048 readout ticks, while the collection plane has a resolution of 3456 wires $\times$ 6048 readout ticks. Wire and time data can be converted into an image format (charge on each wire versus drift time) using the software toolkits LArSoft [8] and LArCV [9] while maintaining high resolution in wire, time and charge amplitude space. These information-rich LArTPC images are suitable for applying deep learning tools. In consideration of computing resources, images for deep learning tools are compressed along the time tick axis by a factor of six. Pixel values are merged by a simple sum. Images become 2400 wires $\times$ 1008 ticks and 3456 wires $\times$ 1008 ticks for the induction and collection planes, respectively. This corresponds to an effective position resolution of 3.3 mm [10] and 3 mm [6] along the time tick and wire number directions, respectively.

Convolutional neural networks (CNN), deep learning networks commonly applied to image processing applications, are currently used across neutrino and high energy physics experiments [11]. For accelerator neutrino experiments, NOvA has applied a CNN as a neutrino event classifier [12] in its $\nu_\mu \rightarrow \nu_e$ oscillation measurement [13, 14] and its neutral-current (NC) coherent π^0 production measurement [15]. NOvA has also demonstrated a context-enriched particle identification network [16]. MINERvA has developed CNN tools to determine neutrino interaction vertices and study possible biases due to models used in the large simulated training sample [17]. The NEXT experiment has also used a CNN classifier to perform particle content studies at candidate

neutrinoless double beta decay vertices [18].

A variety of deep learning techniques have been used in neutrino LArTPC experiments. In MicroBooNE, a CNN for assigning probabilities of particle identities for single particles in the MicroBooNE LArTPC has been demonstrated on simulated data in Ref. [1]. A semantic segmentation network for LArTPC data [19] has been used for π^0 event reconstruction [20], vertex finding, and track reconstruction [21]. The DUNE experiment has recently presented an updated long-baseline neutrino oscillation sensitivity study incorporating a CNN for neutrino event selection and background rejection [22].

In this article, we present our study in developing and applying a multiple particle identification (MPID) network with the task of multiple binary logistic regression problem solving in MicroBooNE. It is the first demonstration of the performance of a CNN on LArTPC data including systematic uncertainties, and the first particle identification network applied to LArTPC datasets. The MPID network extends the functionality of MicroBooNE's previously-described single PID CNN network [1]. It does not require pre-processing of image data to identify and filter selected pixels in an image assumed to be produced by a specific particle. The network provides simultaneous prediction scores for particle existence probabilities in the same image among five different particle species: electrons (e^-), photons (γ), muons (μ^-), charged pions ($\pi^\pm$) and protons (p). The network is a particularly useful tool for data analysis of particle interactions in LArTPC detectors, since the region of an interaction vertex often contains many particles.

The MPID algorithm can take as input a LArTPC image with a fixed 512×512 pixel scale. A detailed description of the network design and training for MPID is given in Section 2. When used in MicroBooNE's deep learning based low-energy ν_e (LEE 1e-1p) search analysis, the MPID network is primarily applied to images that contain candidate reconstructed neutrino interaction vertices as well as all reconstructed topologically connected activity. MPID predictions are derived based on the full information of all energy depositions topologically connected to the vertex, particularly the first few centimeters of final-state particles' trajectories, which are critical for particle identification. In the ν_e search, the network is also applied to more inclusive images roughly cropped around the interaction vertex. This is a new feature compared with the single PID network, which takes as input only images containing filtered, reconstructed hits. Cropping around the interaction vertex allows re-evaluation of charge missing from the former topologically-connected image, but is nonetheless present near the vertex, such as photon showers from final-state π^0s. This feature of the MPID network can help MicroBooNE suppress important photon backgrounds to a LEE search, as observed by MiniBooNE [23]. We demonstrate this feature's robustness against the presence of LArTPC activity such as cosmic ray tracks that are uncorrelated with signal features of interest.

In this note, we demonstrate the expected MPID performance on simulated test images containing one muon and one proton (1μ-1p) and one electron and one proton (1e-1p) in Section 3. Both are signal interactions of a MicroBooNE deep learning based LEE analysis. We then demonstrate the MPID's achieved level of μ^- and p identification in a filtered ν_μ charged current (CC) dataset and compare results between data and simulation in Section 4. We similarly demonstrate the MPID's e^- and γ separation capability using a filtered sample of π^0-containing events, and compare data and simulation results. In Section 5, we demonstrate the value of MPID for a LEE measurement in e^- and γ separation, e^- and μ^- separation and p and cosmic ray separation expected from a set of simulated ν_e interaction images.

2 Multiple particle convolutional neural network

2.1 Network design

The MPID network applies a typical CNN [24] structure for the task of multiple object classification, which is summarized in block diagram form in Fig. 1. Input images have a resolution of 512×512 (1.5 m×1.5 m) pixels, which generally matches the size of neutrino-induced activity in MicroBooNE. A series of ten convolutional layers are applied to the image for extracting high-level features.
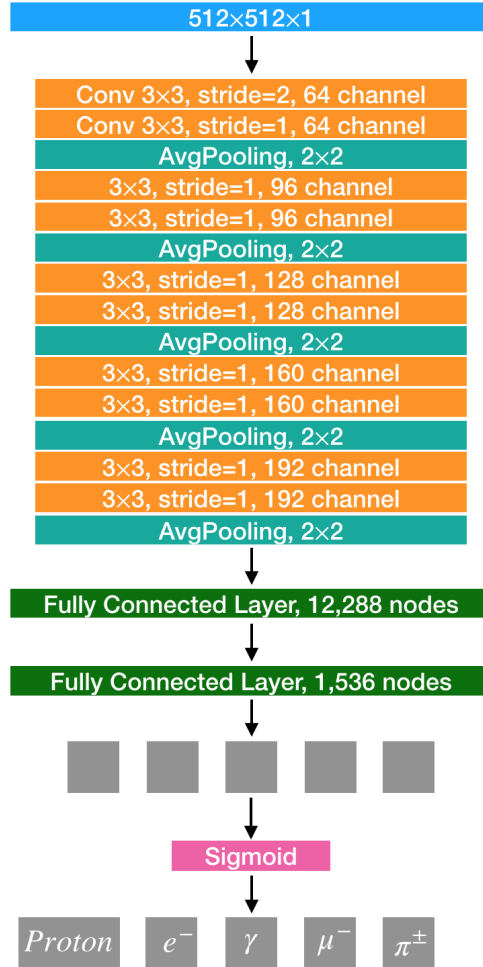

Figure 1: MPID network scheme. The output has five numbers. Each of the values is between 0 and 1, representing the probabilities of corresponding particles in the given LArTPC image.

The first convolutional layer has a stride (shift unit of the convolution calculation) of two with the goal of reducing the LArTPC images' sparsity and increasing feature abundance at the beginning of the algorithm. Following convolutional layers have a stride of one, a block of two convolution layers with a kernel size of three, followed by a pooling layer that is repeated five times. An average pooling layer is applied at every other convolutional layer to contract the spatial dimension. The following the pooling layer is a rectifier activation function (ReLU) [25] for adding non-linearities to the network, as well as a group normalization operator [26] to avoid early overfitting.

Two fully connected layers with 192×8×8 nodes and 192×8 nodes are applied to combine the features derived by convolutional layers. Output of the fully connected layers is a vector with five floating point numbers, each representing a confidence score for a target particle type to be present in an image. The score is interpreted as a normalized probability after applying a sigmoid function [27]. The algorithm is optimized by minimizing the sum of binary cross entropy loss [28] across target particle types. In this way the prediction categories are not exclusive between particles.

Figure 2 shows one example of the input and output of the MPID network during inference. In this case, the input image from has one e^- and one p concatenated at the same vertex, a typical signal interaction topology for an interaction-channel-exclusive 1e-1p search, as implemented in the MicroBooNE deep learning based LEE analysis. The MPID network calculates as output the five floating point numbers described in the previous paragraph, or "particle scores," that correspond to the inferred probability to have each type of particle present in the image. In this example, high scores of 0.99 and 0.98 are given for p and e^- in the image and low scores of 0.06, 0.01 and 0.02 are provided for γ, μ^- and $\pi^\pm$.
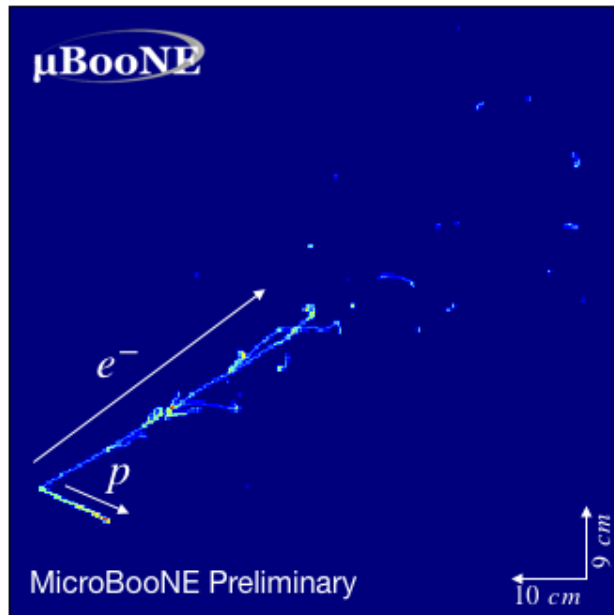

Figure 2: MPID example of an 1e-1p topology with a tabulated output of particle scores. This image is generated by concatenating a p and an e^- at the same vertex. Scores indicate high probabilities of having a p and e^- in the image. The image applied to MPID has 512×512 pixels. A zoom-in image of 250×250 pixels is shown for visualization.

Figure 3 shows another example of the input and output for the network during inference. The input image has one γ, one e^- and one p produced at same vertex, which would in principle be rejected in an exclusive 1e-1p search. Again, the MPID calculates scores that correspond to containing each particle in the image. High scores of 0.89, 0.95 and 0.85 are for p, e^- and γ in the image and low scores of 0.02 and 0.08 are found for μ^- and $\pi^\pm$ in the image. We also note for total clarity that the photon particle score is indicative not of the predicted *total number* of photons in the image, but rather the probability that *any* photons are present in the image. The former judgement, as well as the capability to identify the particle content of specific sub-features within an image, is not within the scope of the MPID algorithm.
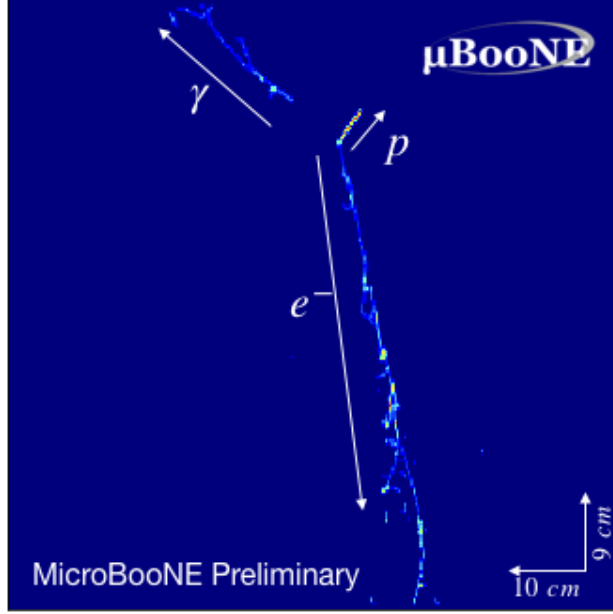
Figure 3: MPID example of an $1e$-1γ-$1p$ topology with a tabulated output of particle scores. This image is generated by concatenating three particles at the same vertex. Scores indicate higher probabilities of having p, e^- and γ in the image. The image applied to MPID has 512×512 pixels. A zoom-in image of 250×250 pixels is shown for visualization.

2.2   Training and Test Samples

Training and test samples for the MPID CNN are produced with a customized event generator that uses LArSoft [8] and LArCV [9]. Detector processes are simulated with the GEANT4 [29, 30, 31] simulation tool.

The first generator step produces a 3D vertex uniformly distributed in the MicroBooNE LArTPC. The second step generates a random number of particles from e^-, γ, μ^-, $\pi^\pm$, and p options. All particles are generated at the vertex from the first step with isotropic directions. The multiplicities for the total number of particles allowed in each image are randomly distributed between two and four. The multiplicity for each particle type is allowed to vary randomly between zero and two. Such a configuration will include as a subset final-state interaction vertex topologies that we are searching for or trying to reject in MicroBooNE analyses, such as $1e$-$1p$, 1μ-$1p$ and $1\gamma - 1p$, as well as non-signal ones, such as 2μ or $2e$. This generation strategy purposefully does not rely on any of the standard neutrino final-state generators [32] to avoid possible biasing the MPID network via inclusion of possibly-incorrect kinematic or multiplicity information provided by the generator. Moreover, this training model will produce a more robust particle identification tool capable of producing unbiased results for a much broader range of vertex-generating physics processes. Finally, high multiplicity topologies generated in this randomized training samples help the network to activate more nodes and learn more parameters for classification.

Each particle is generated with a single particle simulation package, where no neutrino interaction model kinematics are assumed. For 80% of the training and test samples, particles are simulated with kinetic energies between 60 MeV and 400 MeV for protons and between 30 MeV and 1 GeV for other particles. For the other 20% of the training and test samples, particles are simulated with kinetic energies between 30 MeV and 100 MeV for protons and between 30 MeV

and 100 MeV for other particles. Particles are generated with a flat energy distribution. Energy ranges are chosen based on the BNB neutrino energy distribution and the analysis priority towards the lower energy range. We generated 60,000 simulated events for training and 20,000 images for testing. The images are intentionally generated without overlaying cosmic rays on simulated images to retain separation capabilities for μ^-. Images used for training, testing and inference are from the collection plane only.

2.3 Network Training

The loss of the network is defined using the BCEWithLogitsLoss [28] function in PyTorch taking the output layer (five floating point number) as input. The BCEWithLogitsLoss function combines a sigmoid [27] operator with the binary cross entropy calculation. During training, we applied an initial learning rate of 0.001. Batch sizes of 32 and 16 are chosen for the training and test processing.Training is processed with one single NVIDIA 1080 Ti graphics card. Regularization methods of dropout [33] and group normalization [26] are applied to avoid early overfitting during training.
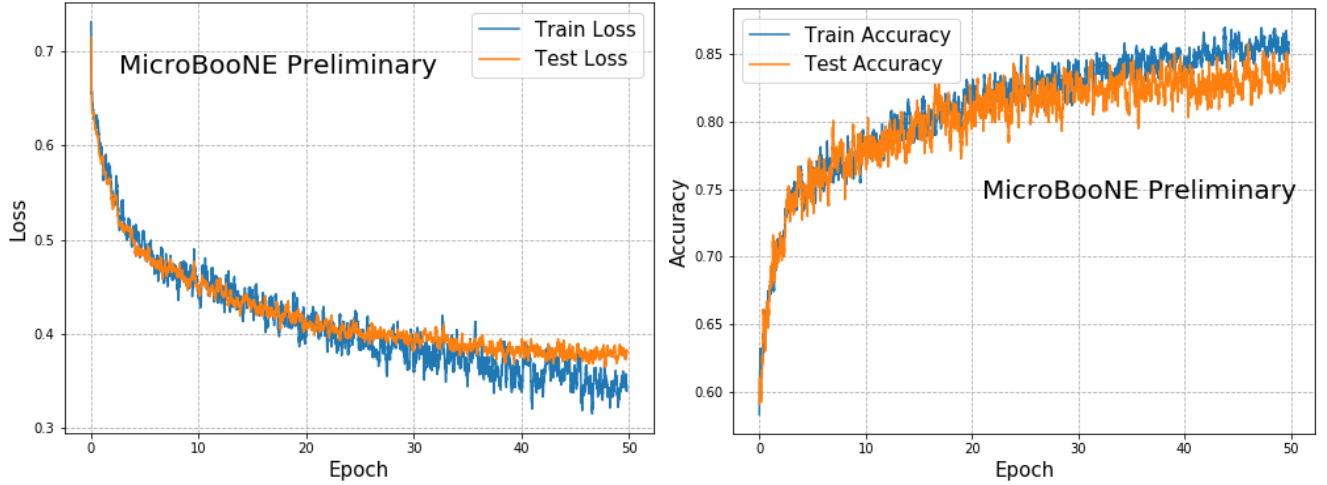

Figure 4: Losses of training and test events during training (top). Accuracies of training and test events during training (bottom).

An accuracy is calculated while the training is monitored for loss. Accuracy is defined as the fraction of predicted labels matching the truth labels with a threshold value of 0.5 per event. MPID training curves of accuracies and losses are shown in Fig. 4. After epoch 29, the test curve continues to improve but does not keep up with the training curve. With the consideration of not introducing bias from the training sample, we checked weights around epoch 29 and selected the one with best accuracy on the test sample.

2.4 MPID Occlusion Analysis

We applied an occlusion analysis [34] to determine whether the MPID network has calculated its predictions using image features associated with underlying physics for example, dE/dx at the first pixels of a particle (referred as the trunk region of a particle), as opposed to other extraneous features in the image. The strategy is to feed the network an image partially masked to check how the MPID responds to the masked image. The occlusion analysis places a 9×9 pixel box in the top
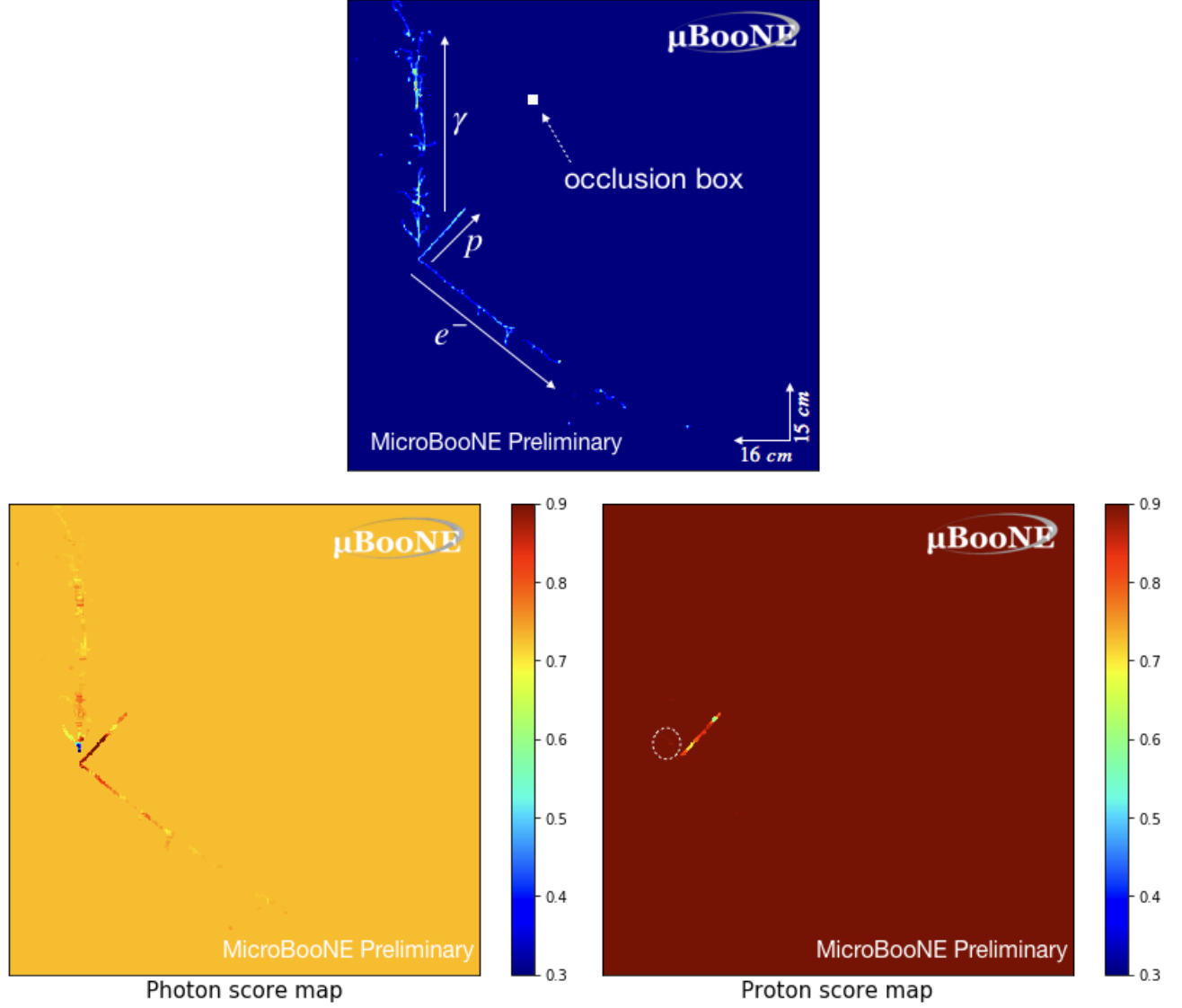
Figure 5: Simulated $1e$-1γ-$1p$ final state event example (top). p score map (bottom left), p scores decrease as occluded region crosses the p pixels. γ score map (bottom right), γ scores decrease as pixels in the broken shower are occluded and increase as the trunk region of the e- shower are occluded.

left corner of the image, which masks all pixels in the occlusion box with zero values. With this box placed, we then apply the MPID network to the masked image and plot at that center pixel the produced score value. This process is then repeated for each pixel as the occlusion box scans along the x and y axis of the image. Figure 5 shows an example of the occlusion box placed on the image. After scanning the whole image, we obtain score maps showing the MPID responses to each occlusion box placement location. A lowered score for a particular pixel in occlusion images indicates that the masked region contains topological information valuable for determining the identity of that particular particle.

A simulated interaction image with one e^-, one γ and one p at the same vertex, shown in Fig. 5 is chosen to demonstrate the occlusion analysis. The bottom left panel in Fig. 5 shows the p scores from the occlusion study on the input image. The p score drops significantly as the proton track's Bragg peak region, where strong p dQ/dx features exist, is masked. This indicates that the MPID network is properly identifying and leveraging features associated with the p's unique energy deposition density profile. It can also be seen that a few pixels in the circle with very high pixel values in the pictured shower are mildly misinterpreted as p-like features.

The bottom right panel in Fig. 5 shows the γ scores from occlusion analysis of the same input image. From the occlusion image, it is clear that a few key physics features of γ-containing images have been properly learned by the MPID network. There are two critical features in the particle trunk region for e^-/γ separation: the projected trunk region dE/dx difference and the presence or absence of a gap between the trunk and the interaction vertex. One can see the γ score drops to near 0.3 when the trunk region of the γ (rather than the gap region between γ and vertex) is masked. We also applied an occlusion analysis to images with single γ images to confirm that γ scores drop and e^- scores increase as the γ trunk region is masked. We observe in this example that the γ score also increases to near one when nearby pixels connecting the p and e^- are masked, since this produces more gaps between different particles. The e^- score does not change much as we move the occlusion box around since there are overwhelming e^--like features in the image from both the e^- and γ.

3   Performance on Simulation

To provide a first look at the capabilities of the trained MPID algorithm, we present particle score results returned from the test images generated using the same method applied in producing training images. This section is divided into discussions of individual final state vertex and particle topologies of interest to MicroBooNE physics analyses.

We generated two test samples with particles of 1μ-$1p$ and $1e$-$1p$ in the final state, which are not used in training. 10,000 events are generated in each sample. These samples are generated with the same customized event generator described in Section 2.2. Vertices are uniform in the detector with one proton and one corresponding lepton. Kinetic energies of the protons are between 50 MeV and 400 MeV, while kinetic energies of leptons are generated between 50 MeV and 1 GeV. The $1e$-$1p$ final state dataset has a similar final state as the target events of MicroBooNE's deep learning based LEE $1e$-$1p$ analysis. The 1μ-$1p$ dataset has a final state similar to a MicroBooNE ν_μ selection analysis, described in Section 4.1, that will be used to constrain the beam-intrinsic backgrounds (non $1e$-$1p$ ν_e interactions) in the LEE search. Studying the performance of MPID with these two simulated datasets shows the general level of expected performance of the network when signal (or background) pixels in input ν_e and ν_μ images are included (or removed).

3.1 1μ-1p Simulated Sample

Figure 6 shows stacked MPID scores of five particle hypothesis for the 1μ-$1p$ simulated test dataset. One can see that the MPID network provides good separation between track-like and shower-like particles with p and μ^- scores concentrated near one and e^- and γ piled up near zero. The plot also shows a good separation between μ^- and $\pi^\pm$ using MPID, with a low score distribution for $\pi^\pm$. Separation between μ^- and $\pi^\pm$ comes from $\pi^\pm$s have higher rates in nuclear scattering than μs and $\pi^\pm$s can have a kink point where they decay into μs as note in Ref. [1].
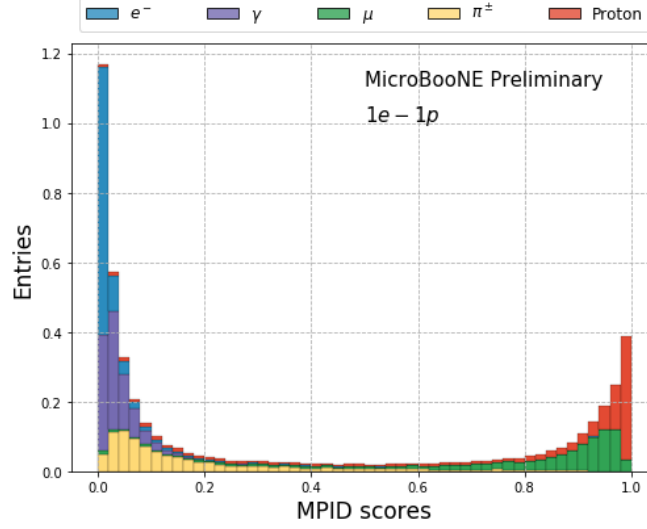

Figure 6: MPID score distributions for the probabilities of p, μ^-, e^-, γ, $\pi^\pm$ on the 1μ-$1p$ validation sample.
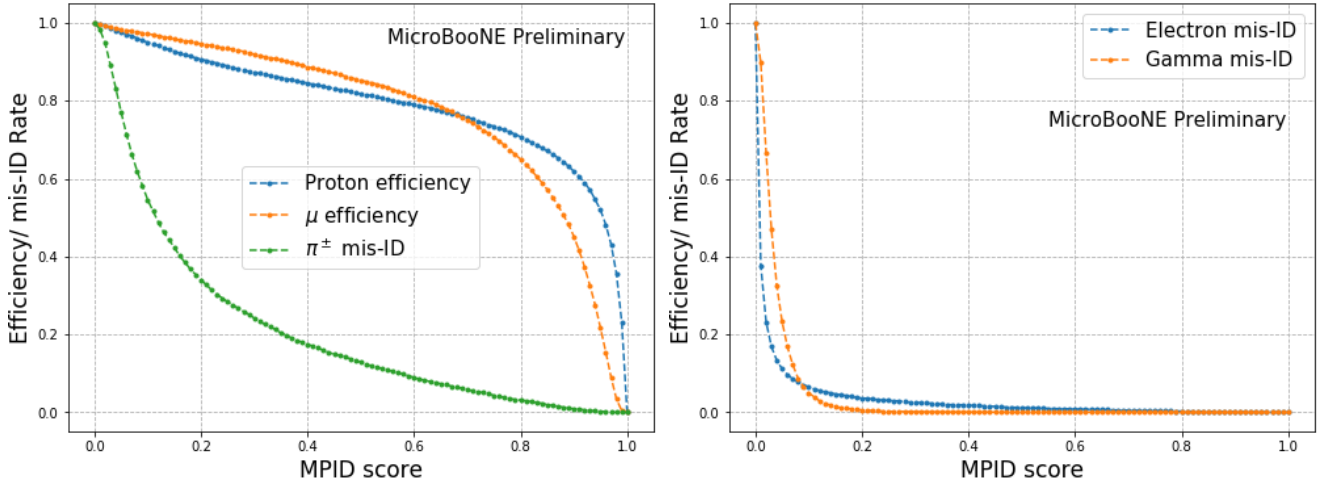

Figure 7: MPID efficiency and mis-ID curves for track-like particles (top) of p, μ^- and $\pi^\pm$ on the 1μ-$1p$ validation sample. MPID efficiency and mis-ID curves for shower-like particles (bottom) of e^- and γ on the 1μ-$1p$ validation sample.

To perform particle identification as part of a neutrino event selection analysis, a set of selections are usually applied to particle score variables; these cuts will have associated impacts on total signal selection efficiencies and on the probability of mis-identifying background events as

signal events. Figure 7 shows the correlation curve between efficiency/misidentification (mis-ID) rates and MPID scores for track-like particles in the 1μ-$1p$ dataset. The cut value for each particle score is varied between 0 and 1 with a step size of 0.01. For example, the green dotted line shows the efficiency of correctly predicting a p in the image at each p score cut value. Meanwhile, the orange dotted line shows the mis-ID rate of falsely predicting a $\pi^{\pm}$ in the image with different $\pi^{\pm}$ score values. Figure 7 also shows the correlation curve between mis-ID rates for shower-like particles in the 1μ-$1p$ dataset. The mis-ID rates are both extremely low for having either in the 1μ-$1p$ sample.

Table 1 provides another view of Fig. 7, showing the efficiencies of correctly labelling p and μ^{-}, as well as mis-ID rates of falsely labelling e^{-} and γ in the 1μ-$1p$ dataset at score cut values of 0.1, 0.3, 0.5, 0.7 and 0.9. At a cut score of 0.5, efficiencies of identifying a p and μ^{-} are at 82% and 85%, while mis-ID rates of shower-like particles are 1% or less. It is also worth noting that a cut score of 0.5 on $\pi^{\pm}$ produces a sufficiently low mis-ID rate at 13%.

Table 1: Efficiencies and mis-ID rates for 1μ-$1p$ sample at specific selection scores.

MPID Cut Value	Efficiency		mis-ID rate		
	p	μ^{-}	e^{-}	γ	$\pi^{\pm}$
0.1	0.95	0.97	0.07	0.05	0.54
0.3	0.87	0.92	0.03	0.00	0.24
0.5	0.82	0.85	0.01	0.00	0.13
0.7	0.76	0.75	0.00	0.00	0.06
0.9	0.62	0.45	0.00	0.00	0.01

Figure 8 shows the correlation between $\mu^{-}/\pi^{\pm}$ scores and the μ^{-} kinetic energy using the same 1μ-$1p$ simulation of 10,000 events. One can see that when the μ^{-} particles have low kinetic energy and produce fewer μ^{-}-like pixels in the image, the μ^{-} score is decreased. Meanwhile, $\pi^{\pm}$ scores for the same dataset appear to be comparatively stable across all tested muon energies.

3.2 $1e$-$1p$ Simulated Sample

Figure 9 shows stacked MPID score distributions for the simulated $1e$-$1p$ dataset. MPID correctly calculates high scores for signal particles of p and e^{-}. The network shows good separation between track particles in deriving low scores for μ^{-} and $\pi^{\pm}$. The MPID CNN also shows good separation between shower-like particles when e^{-}'s are present in the image: derived scores for γ are clustered close to zero, while e^{-}-like scores are clustered around unity.

The correlation curve between efficiency/ mis-ID rates and MPID scores for track-like particles in the $1e$-$1p$ dataset are given in Fig. 10. Efficiencies for p in the image are much higher than the mis-ID rates for containing a μ^{-} or $\pi^{\pm}$ in the image. The capability to discriminate between p and μ^{-} appears to be particularly high, while $p/\pi^{\pm}$ separation also remains high. This difference in performance between μ^{-} and $\pi^{\pm}$ should not be too surprising given the level of $\pi^{\pm}$-μ^{-} separation demonstrated in the previous section. Figure 10 also shows correlations between efficiency/ mis-ID rates and MPID scores for the shower-like particles in the $1e$-$1p$ dataset. This plot shows that MPID should produce good separation between e^{-} and γ containing events.

Table 2 shows the efficiencies for labelling p and e^{-}, and mis-ID rates for labelling μ^{-} and γ in the $1e$-$1p$ dataset with at each particle score cut values at 0.1, 0.3, 0.5, 0.7 and 0.9. Efficiencies of correctly finding the p and e^{-} are at 82% and 91% with particle score cut values at 0.5, while
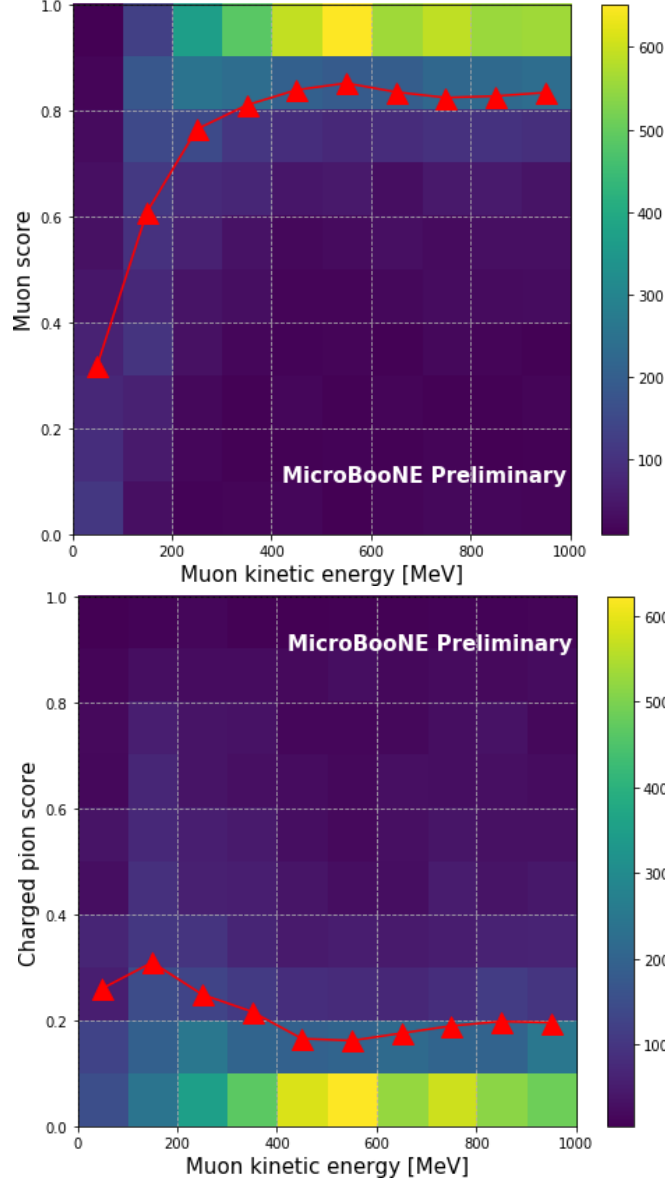
Figure 8: Muon score vs. muon kinetic energy (top) and charged pion score vs. muon kinetic energy (bottom) for the 1μ-1p simulation. Red dots indicate the average score in the vertical bin.

Table 2: Efficiencies and mis-ID rates for 1e-1p sample at specific selection scores.

MPID Cut Value	Efficiency		mis-ID rate		
	p	e^-	μ^-	γ	$\pi^{\pm}$
0.1	0.95	0.98	0.15	0.68	0.44
0.3	0.87	0.95	0.03	0.31	0.18
0.5	0.82	0.91	0.01	0.15	0.09
0.7	0.76	0.84	0.00	0.06	0.04
0.9	0.63	0.64	0.00	0.02	0.01
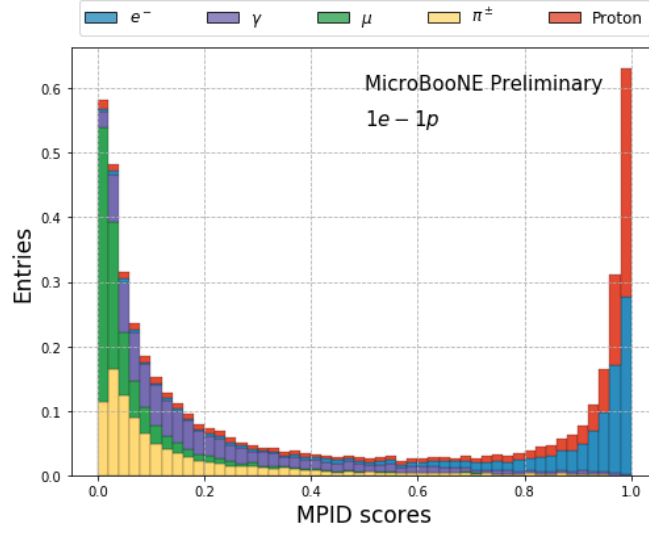
Figure 9: MPID score distributions for the probabilities of p, μ^-, e^-, γ, $\pi^\pm$ on the 1e-1p validation sample.
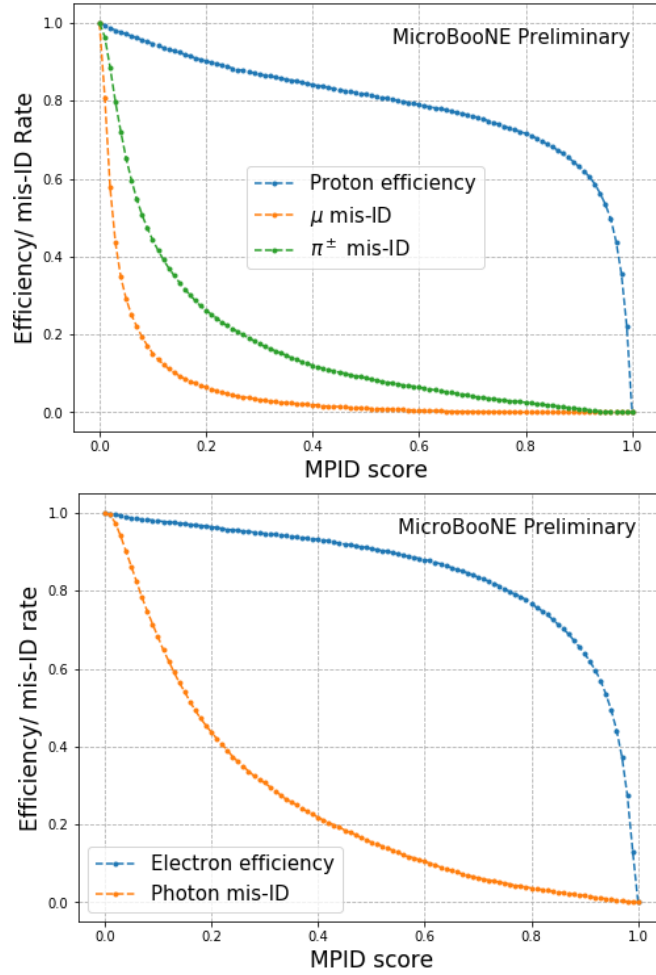

Figure 10: MPID efficiency and mis-ID curves for track-like particles (top) of p, μ^- and $\pi^\pm$ on the 1e-1p validation sample. MPID efficiency and mis-ID curves for shower-like particles (bottom) of e^- and γ on the 1e-1p validation sample.

impressively small mis-ID rates are found for μ^- (1%) and γ (15%). As in the previous section for 1μ-1p selection, $\pi^\pm$ mis-ID rates remain low, with a 9% Mis-ID at a cut score of 0.5.

Figure 11 shows the correlation between e^-/γ scores and e^- kinetic energy. One can see the MPID network has an overall high e^- score until the e^- kinetic energy approaches its critical energy in liquid argon and becomes less shower-like. In a related sense, μ^- scores for low energy 1e-1p interactions are found to be slightly higher than high energy ones. Meanwhile, the γ score for these events has a positive correlation with e^- kinetic energy, since high energy e^- are more likely to experience substantial amounts of radiative energy loss.
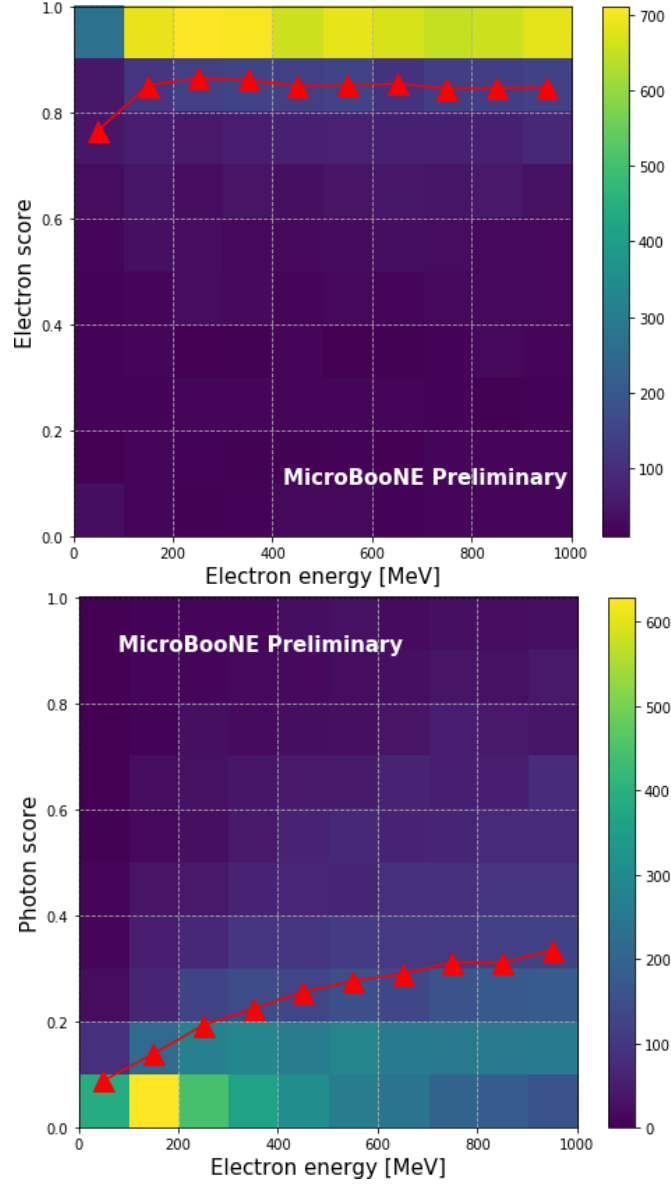

Figure 11: Electron score vs. electron kinetic energy (top) and photon score vs. electron kinetic energy (bottom) for the 1e-1p simulation. Red dots indicate the average score in the vertical bin.

4 Comparison of Data/Simulation Performance

We prepared two different MicroBooNE LArTPC data samples to validate the performance of the MPID network on data. The MPID network was not employed in the selection of these data samples. The first data sample is a 1μ-$1p$ enriched selection that uses a hybrid selection of a series of reconstruction algorithms [21] and MicroBooNE's semantic segmentation network [19]. This dataset is intended to be used in a MicroBooNE LEE $1e$-$1p$ analysis to provide a data-based constraint on the BNB neutrino beam's intrinsic ν_e contamination. The second sample contains ν_μ charged current interactions with a final-state π^0 ($\nu_\mu CC\pi^0$) as defined in Ref. [20]. This dataset was used to report the first measurement of neutral π^0 production in charged-current ν_μ interactions on argon. In this section we demonstrate that the MPID network works well on real LArTPC images. We show good agreement in MPID scores between data and simulation for the selected datasets.

To enable data/simulation comparisons for these two event classes, we simulate neutrino interactions using the GENIE v3.0.6 [32] neutrino Monte Carlo generator. To accurately include on-surface cosmogenic backgrounds present in all MicroBooNE LArTPC images, beam-off data containing only cosmic rays is overlayed on simulated neutrino interaction images. Beam-off data is taken with cosmic ray triggers. An overlay sample is a combination of GENIE simulated beam events and cosmic data events. This ensures that the reported particle score distributions for data and simulated images will be equally affected by the presence of cosmic rays.

In the study of the 1μ-$1p$ dataset, we apply the MPID network to processed images containing only wire signal activity associated with particles reconstructed at a candidate neutrino interaction vertex. In the study of $\nu_\mu CC\pi^0$ dataset, we instead apply the MPID network to images made with all pixels near the reconstructed vertex; in this case, particle scores are completely independent of any previous reconstruction. The data/simulation comparison therefore also demonstrates the robustness of the MPID network to both 'cleaned' (input images contain only the reconstructed interactions) and potentially 'polluted' (input images also contain cosmic rays) input image types.

4.1 1μ-$1p$ Enriched Data

The 1μ-$1p$ enriched dataset is selected from a set of MicroBooNE beam-on data corresponding to 4.4×10^{19} protons on target (POT) in the BNB beam. These events consist of exactly two reconstructed particles – ideally one p and one μ^- – at the candidate interaction vertex. The selection consists of two steps. The first step involves a set of preliminary cuts based on optical information and interaction topology cuts. Candidate 1μ-$1p$ interactions are required to have above a threshold number of photon-electrons in the beam trigger window to be signal. Interaction topology selections require candidates to be located inside the TPC with exact two fully-contained reconstructed tracks. Topology selections also require an opening angle greater than 0.5 radians. The second step involves two boosted decision trees (BDT) to make a final 1μ-$1p$ selection. The first BDT is trained to separate 1μ-$1p$ from the cosmic backgrounds using a simulated ν_μ sample and a beam-off cosmic ray only data. The second BDT is trained to separate 1μ-$1p$ from non-signal neutrino interactions (i.e non-CCQE ν_μ interactions, off-vertex ν_μ interactions and interactions missing more than 20% energy in reconstruction) using a simulated ν_μ sample. Details of preliminary selection and BDT selections will be documented in details in future publications. The selection of the dataset described above produces 478 data and 466 simulated input images for processing by the MPID network. In the simulated dataset, 94% of these images contain true neutrino interactions. Among these, 314 (67% of total images) events contain solely one reconstructable final-state p and

μ^-.

We produce the input images in three steps. First, the interaction vertex is located and any associated track-like particles are reconstructed using algorithms described in Ref. [21]; two and only two reconstructed tracks are required. Next, a 512×512 image is produced, centered at the pixel-weighted center of the reconstructed 1μ-1p event from a flat weight for non-zero pixels. Finally, to address noise-related features, a threshold is placed on the images with a minimum and maximum pixel value of 10 and 500, respectively. This procedure removes effects from pixels from unrelated interactions near the neutrino interaction vertex. Figure 12 shows an example of a 1μ-1p image fed into the MPID network. The image is from the collection plane.
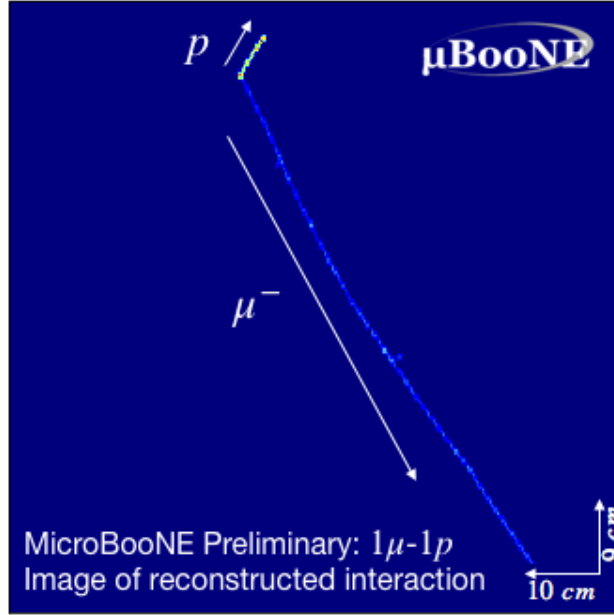

Figure 12: Example of the input data image from 1μ-1p selection. The image is centered at the non-zero pixel weight center. The image has 512×512 pixels. A zoom-in image of 250 × 250 pixels is shown for visualization.

The top image of Fig. 13 shows the p score for the selected candidate 1μ-1p interactions, broken down into the true physics origin of each imaged vertex. The simulation predicts that true 1μ-1p charged-current neutrino interactions should cluster at high p score, with background processes (particularly cosmic processes) more evenly distributed across the score axis. In the data, a distinct peak is present at high p score, providing a strong indication of proton(s) being present in most of the images.

The bottom sub-panel of this sub-figure shows the ratio of data and simulation versus the p score. We note that as we are primarily concerned with understanding the agreement in the distribution of scores from 0 to 1, discussion of the level of absolute agreement in normalization between data and simulation is beyond the scope of this study. For each point, the data's statistical uncertainty is shown, along with the systematic uncertainty associated with flux and cross-section uncertainties. Beam flux uncertainties are evaluated by re-weighting events according to the properties of the hadrons that decay to produce the neutrinos. Cross section uncertainties are evaluated by re-weighting events according to the properties of the neutrino's interaction with an argon nucleus. Detector uncertainties are in development and are expected to not have a dominant
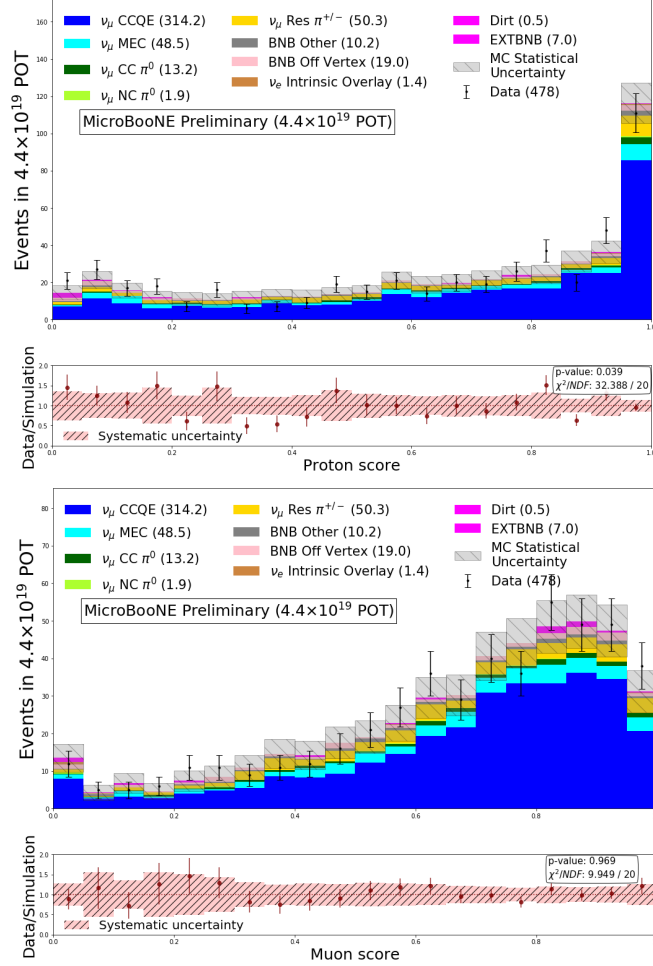
16

Figure 13: MPID proton score distribution (top) and muon score distribution (bottom) for selected 1μ-$1p$ interactions. Simulation-predicted score distributions show satisfactory agreement with those realized in the 1μ-$1p$ selection applied to MicroBooNE data. Plot error bars indicate data statistical errors, while hatched bands indicate statistical and/or systematic uncertainties in the simulated dataset. The χ^2 calculation incorporates contributions from systematic and statistical uncertainties. The breakdown interaction type is based on the predicted event classification for the initial neutrino interaction.

systematic effect on MPID scores for 1*e*-1*p* events. Good agreement is found between the data and simulation across the full range of *p* scores with flux and cross section uncertainties. This level of agreement was quantified by calculating the χ^2 between the data and simulation distributions in the top panel of Fig. 13. This χ^2 includes both statistical and systematic uncertainties in the data and simulation. A χ^2/NDF of 32.4/ 20 is found, indicating a comparable performance of the MPID on both data and simulation.

Figure 13 also shows the μ^- score distribution for the same selected 1*μ*-1*p* interactions. A majority of events are found in the higher score region, indicating that the MPID algorithm has correctly identified the presence of μ^- in these images. The bottom panel again shows the ratio of data to simulation in the μ^- score distribution; systematic error bars are similarly defined as for the simulated *p* score distribution. A χ^2/NDF of 9.9/ 20 is found between the two distributions, indicating a good data-simulation comparison for both *p* and μ^- scores.
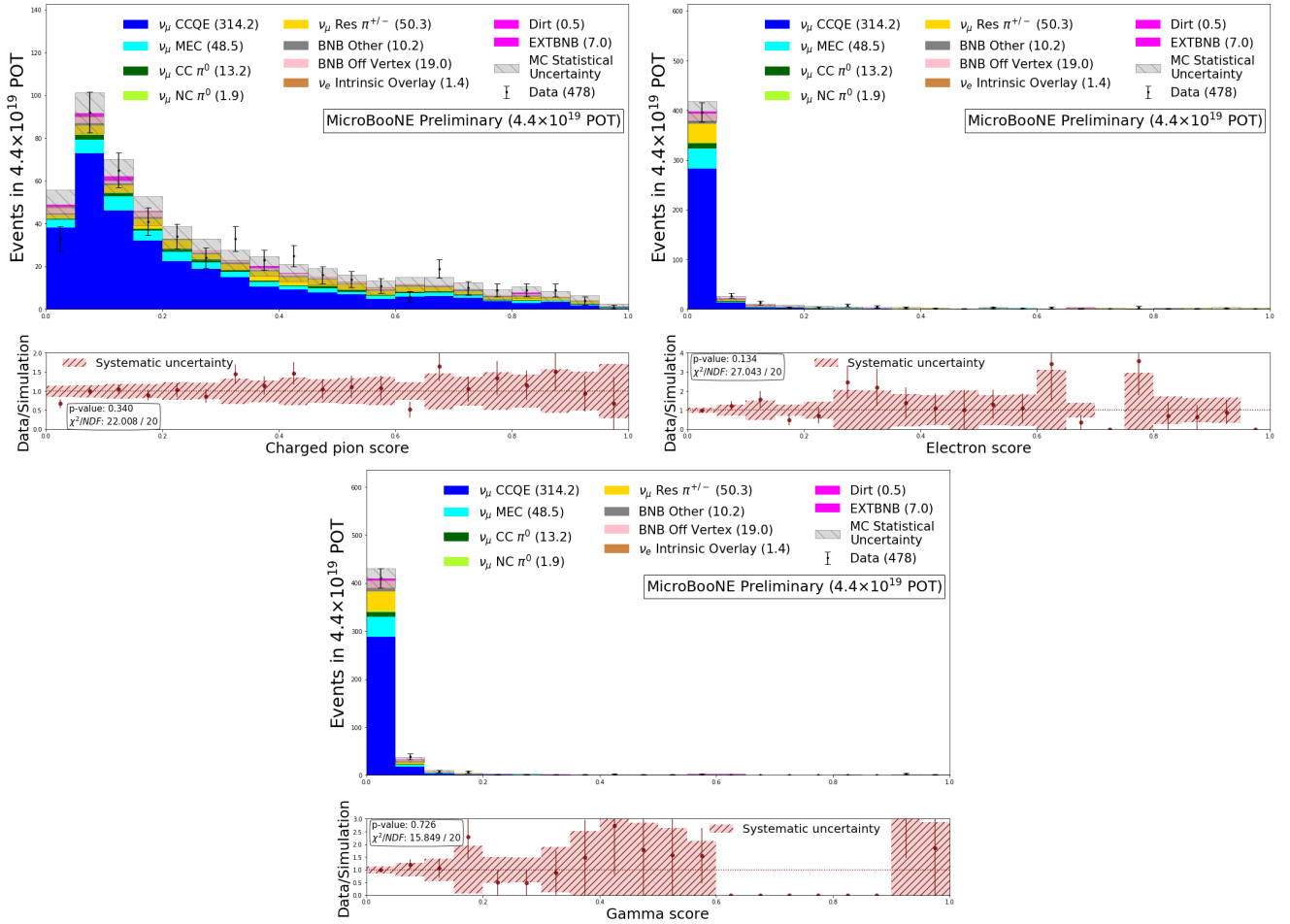

Figure 14: Charged pion score distribution (top), electron score distribution (mid), and photon score distributions (bottom) for selected 1*μ*-1*p* interactions. Score distributions agree with the 1*μ*-1*p* selection. Data and simulation agree well. χ^2 calculation include systematic and statistical uncertainties.

Figure 14 shows the score distributions for particle types expected to be absent from or contained in limited quantities in the selected 1*μ*-1*p* dataset: $\pi^{\pm}$, e^-, and γ. For γ and e^-, the score distributions are peaked very close to zero, since input images have only track-like particles, and because, as demonstrated in Section 3, discrimination between track-like μ and p particles and

shower-like γ and e^- particles is expected to be high. Scores for track-like $\pi^\pm$ particle scores are also clustered towards zero, but with a broader overall width; this result also matches the expectations of Section 3. The χ^2/NDF of 22.0/ 20, 27.0/ 20, and 15.8/ 20 for data/simulation comparisons for γ, $\pi^\pm$, and e^- indicate comparable performances of MPID on data and simulation.

The MPID network appears to provide similar performance on both data and simulated neutrino interaction images containing primarily track-like final-state particles. This similarity in performance is achieved despite the input image's reliance on other reconstruction algorithms to 'remove' pixel content not related to final-state particles connected to the candidate neutrino interaction vertex. This indicates that not only the MPID algorithm, but also the upstream reconstruction algorithms, treat data and simulated LArTPC images on an equal footing.

4.2 $\nu_\mu \mathrm{CC}\pi^0$ Enriched Data

A study of π^0-producing charged current ν_μ ($\nu_\mu \mathrm{CC}\pi^0$) interactions is useful in providing a similar data/simulation agreement validation for images that also contain shower-like objects, as is expected from charged-current ν_e interactions. For this study, we select events from the same dataset used in MicroBooNE's previous $\nu_\mu \mathrm{CC}\pi^0$ measurement [20]. The primary reconstruction toolkits used to develop selection metrics for these events are Pandora [35] and SSNet [19]. Selected events are primarily required to have two showers close to the interaction vertex. This requirement makes this dataset distinct from a $1e - 1p$ selection, where one and only one shower is allowed, which must be directly attached to the vertex. In this way, in studying MPID performance on the $\nu_\mu \mathrm{CC}\pi^0$ data sample, we demonstrate not only data/simulation performance, but also show how the network can help to reduce a major intrinsic background to the ν_e channel: π^0-producing interactions.

Input images from $\nu_\mu \mathrm{CC}\pi^0$ candidates are generated by cropping a 512×512 square image centered at the reconstructed interaction vertex, rather than at the image's pixel-weighted center as in the 1μ-$1p$ images. To ensure that π^0 decay γs are not scrubbed from the image, no additional pixel 'cleaning' is applied. This means that cosmic rays and other interactions unrelated to the vertex remain in input images, presenting an additional challenge to the MPID network's performance. The same noise filtering metric, as described for the 1μ-$1p$ dataset, is applied to the images. Figure 15 shows an example of a $\nu_\mu \mathrm{CC}\pi^0$-containing image fed into the MPID network. The image is from the collection plane.

The selection and dataset described above produces 2051 data and 2011 simulated input images for processing by the MPID network. According to the simulation, 41% of total events have $\nu_\mu \mathrm{CC}\pi^0$ interactions and 60% of events contain π^0-including interactions (including the $\nu_\mu \mathrm{CC}\pi^0$ interactions).

Figure 16 shows the score distribution for having any e^- in the images cropped from the $\nu_\mu \mathrm{CC}\pi^0$ sample. The score indicates a generally low probability of having e^--like features in the data and simulated images. As a comparison, Fig. 16 also shows the score distribution for having any γ-rays in the images. One can see a much higher score distribution for the γ existence case, as expected based on the event filtering criteria described above. Figure 17 shows the score distribution for having any μ^- in the $\nu_\mu \mathrm{CC}\pi^0$ images. The score generally indicates a high probability of having μ^--like features in data and simulation. In particular, it shows a difference between the CC-and NC-π^0 events in the low μ^- score region.

The bottom panels of Fig. 16 and Fig. 17 show the ratio of data to simulation versus e^-, γ and μ^- scores. As in the 1μ-$1p$ dataset, we scale the normalization of $\nu_\mu \mathrm{CC}\pi^0$ interactions based on preliminary measurements by the MicroBooNE collaboration. Systematic uncertainties are also
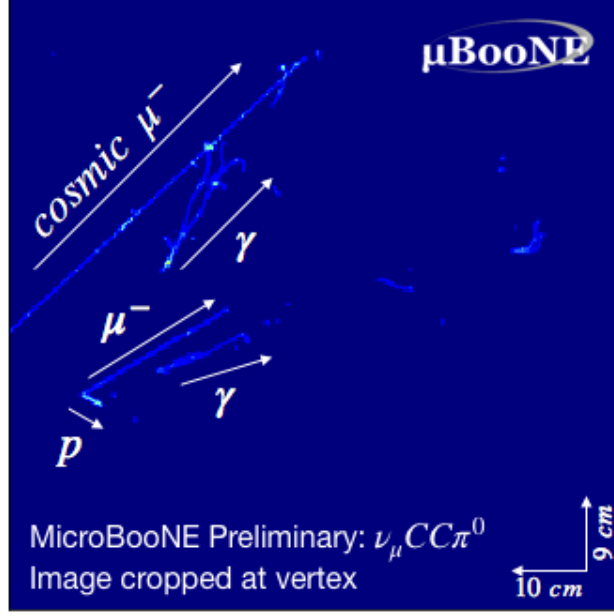
Figure 15: Example of the input data image from the $\nu_\mu CC\pi^0$ selection. The image is centered at the reconstructed vertex. The image has 512×512 pixels. A zoom-in image of 250×250 pixels is shown for visualization.

included in the same manner as described for the 1μ-$1p$ dataset. Good comparable performance can be seen between data and simulation with χ^2/NDFs of 43.6/ 39 for e^- score, 42.8/ 39 for γ score distributions and 24.0/ 39 for μ^- score. Thus, this study demonstrates that the MPID algorithm can reliably identify shower-related particle content in images without introducing biases between neutrino data and simulation predictions. This is achieved despite of the presence of incidental additional pixel activity being present in interaction candidate images.

5 Use of MPID In A Low Energy Excess Measurement

In the two previous sections, we have demonstrated the MPID network's utility in particle identification for both track and shower topologies in LArTPC images, as well as its equivalent performance on both data and simulated events. We will now apply the trained MPID network to simulated BNB ν_e and ν_μ interactions overlaid with beam-off cosmic event images to demonstrate the ability of the MPID network to aid in event selection for MicroBooNE's deep learning-based $1e$-$1p$ low-energy excess search.

5.1 Simulated Intrinsic ν_e vs. ν_μCCQE and $\nu_\mu\pi^0$

We generated simulated neutrino events to evaluate the performance of MPID in the $1e$-$1p$ selection in separating beam-intrinsic BNB backgrounds from ν_μCC quasi-elastic (CCQE) and neutrino interactions with one or more π^0s in the final state ($\nu_\mu\pi^0$). Samples for these three datasets are produced using the standard GENIE v3.0.6 [32] neutrino interaction generator and filtered using truth-level information. In these samples, we require the lepton kinetic energy be greater than 35 MeV and p kinetic energy greater than 60 MeV. The minimum kinetic energy thresholds were set in
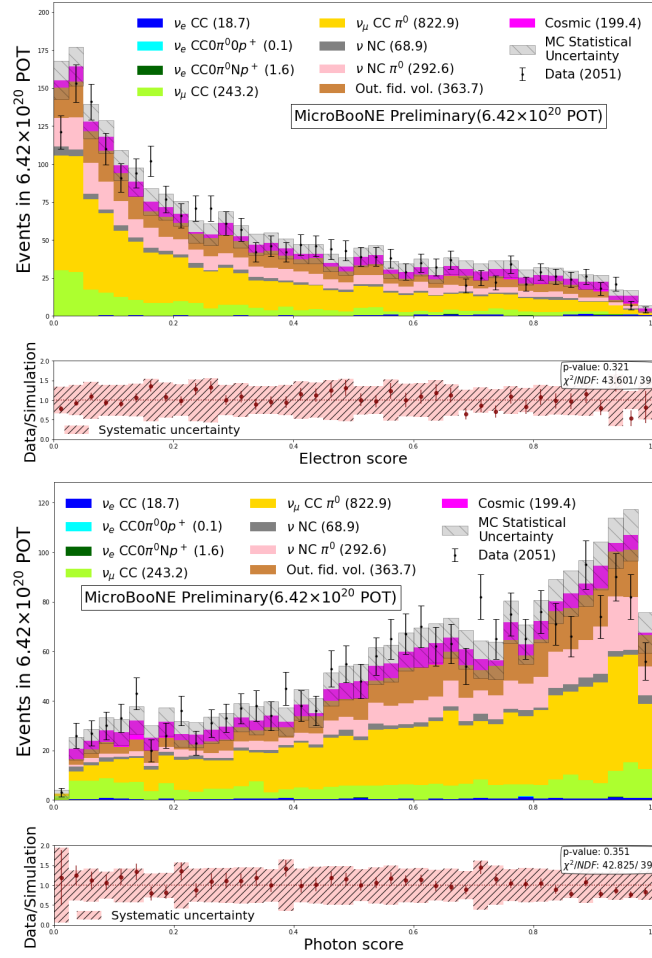
Figure 16: Electron score distribution (top) and photon score distribution (bottom) for selected $\nu_\mu CC\pi^0$ interactions. Score distributions agree with the $\nu_\mu CC\pi^0$ selection. Data and simulation agree well. χ^2 calculation include systematic and statistical uncertainties.
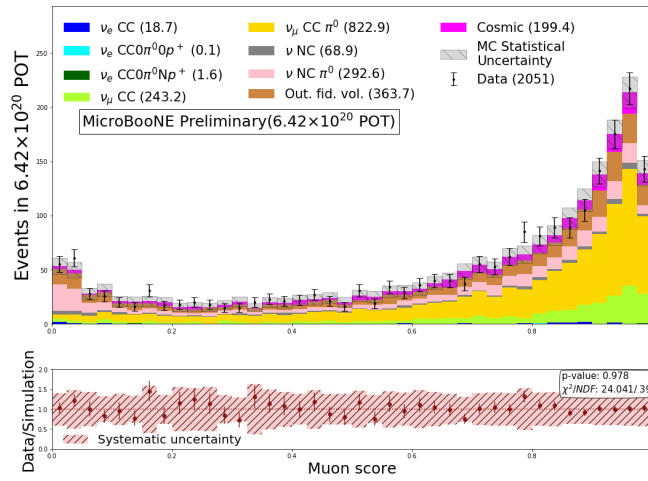

Figure 17: Muon score distribution for selected $\nu_\mu \pi^0$ interactions. Score distributions agree with the $\nu_\mu \pi^0$ selection. Data and simulation agree well. χ^2 calculation include systematic and statistical uncertainties.

order to choose events whose lepton and p trajectories are long enough to be reconstructed by our deep learning based vertex finding and particle reconstruction algorithms [21]. Samples are then processed using the reconstruction algorithms to identify candidate interaction vertices and nearby related particles. Finally, input images are generated with pixels from only the reconstructed interaction final-state particles; each interaction is required to have two particles at this stage. Images are centered at the pixel weighted center of reconstructed interactions. No other selection cuts beyond the truth-level filtration described above are applied to the samples.
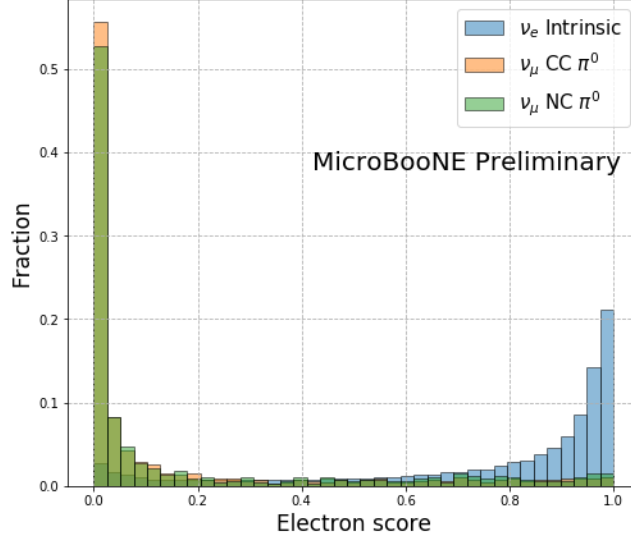

Figure 18: Electron score of ν_e intrinsic events and $\nu_\mu \pi^0$ events. Both datasets are generated with the GENIE neutrino generator and filtered using truth level information. Presented events have a reconstructed vertex.

Figure 18 shows the e^- score distribution of reconstructed events from ν_e and $\nu_\mu \pi^0$ datasets. A good separation is visible between these two event classes. For example, with only an e^- cut score of 0.5, 83% of $\nu_\mu NC\pi^0$ and 86% of $\nu_\mu CC\pi^0$ events are rejected, while 81% of true 1e-1p events are selected. It seems likely that further gains in background rejection could be achieved by also considering scores for other particles and by using differing input pixel image inclusion settings.

Previous discussion from the occlusion analysis presented in Section 3 provides some level of insight into the causes of the substantial discrimination shown in Fig. 18. In particular, ν_e interactions will contain a shower-like object with a trunk directly connected to another particle, a feature that was clearly noticed by the MPID network. This is not the case for most γ rays present in $\nu_\mu \pi^0$ interactions. Another critical parameter for separating e^-- and π^0-including events is the energy deposition density, dE/dx, along this vertex-connected shower trunk; the trunk region information is usually well-reconstructed, since it is almost always directly attached to neutrino candidate vertex. Some of the discrimination in Fig. 18 may thus also arise from the network's ability to discriminate a high trunk dE/dx for vertex-connected showers from quickly-converting π^0 γ rays.

The e^- score can also be applied to separate 1e-1p and 1μ-1p events. The separation is shown in Fig. 19. The ν_e and $\nu_\mu CCQE$ events are well separated using the e^- score calculated by the MPID network. For example, with only an e^- cut score of 0.2, 91% of true 1e-1p events are selected, while 95% of $\nu_\mu CCQE$ events are rejected. This discrimination ability almost certainly arises from the lack of shower-like topologies in the $\nu_\mu CCQE$ interaction images.
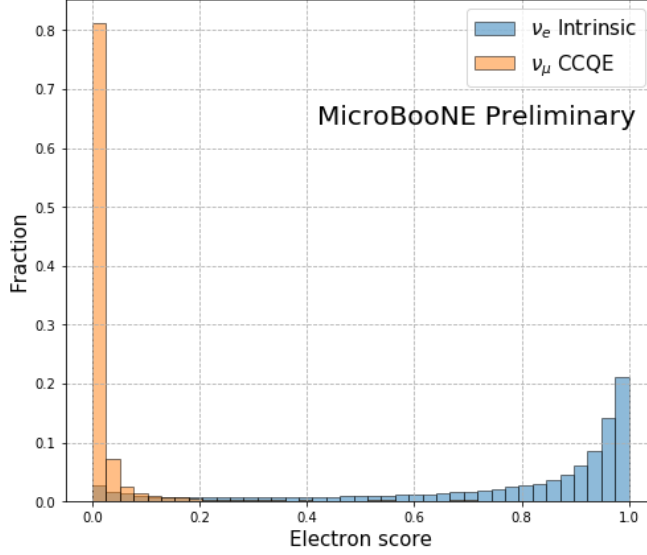
Figure 19: Electron score between ν_e intrinsic events and ν_μ CCQE events. Both datasets are generated with the GENIE neutrino generator and filtered using truth level information. Presented events have a reconstructed vertex.

5.2 Simulated Intrinsic ν_e vs. Cosmic Event

Due to the lack of substantial overburden and the long readout time, cosmic rays could provide a substantial background to a BNB-based 1e-1p ν_e measurement in MicroBooNE. As most of this cosmic ray activity is induced by μ^-, it is expected that the presence of a p in the signal's final state will aid in distinguishing the two categories. To test the MPID network's ability to discriminate the signal's p particle content, we generated a simulated intrinsic ν_e dataset with cosmic data overlay, in addition to another event set consisting purely of beam-off cosmic triggers. For both datasets, we applied the vertex finding and particle reconstruction algorithms developed for two-track events, as described in Ref. [21]; in particular, each image is required to have exactly two reconstructed particles connected to the candidate neutrino interaction vertex. As in the sub-section above, no selection cuts are applied beyond truth-level event filtration.

Figure 20 shows the p score distributions on images from the intrinsic ν_e dataset with cosmic overlay and the pure cosmics dataset. One can see that the majority of pure cosmic dataset events reconstructed as two-particle signals events have p scores below 0.2. Meanwhile, the majority of reconstructed ν_e intrinsic events have p scores near 1. For example, with only a p cut score of 0.5, 81% of true 1e-1p events are selected, while 79% of cosmic events are rejected.

Investigation of information from prior reconstruction stages and hand-scanning of event displays indicates that the small peak in p score close to zero in the ν_e intrinsic dataset is due to inefficiencies in p reconstruction as shown in Fig. 21(a) of Ref. [21]. Similar investigations show that the small peak of of p score close to one in the cosmic sample is introduced by cosmics with small incident angles relative to the collection plane; these non-p tracks are often topologically compressed by reconstruction algorithms, giving them the appearance of short tracks with a proton-like Bragg peak. Thus, future improvements in lower-level signal processing and particle reconstruction is likely to further improve the cosmic discrimination shown in Fig. 20.
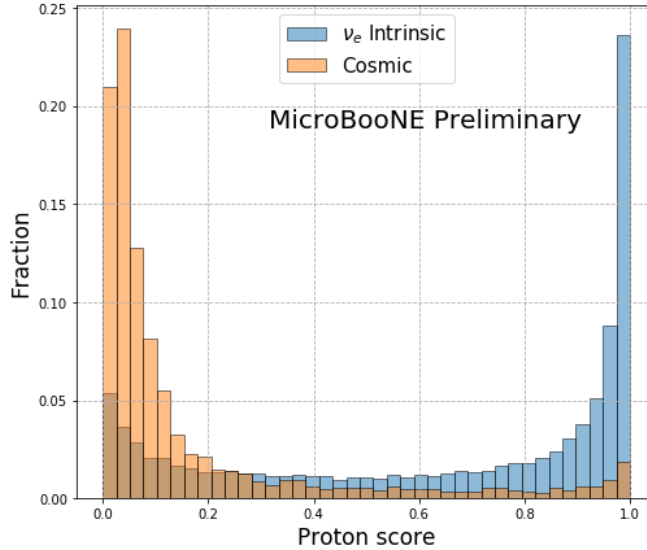
Figure 20: Proton score of ν_e intrinsic events and beam-off cosmic data. The ν_e dataset is generated with the GENIE neutrino generator and filtered using truth level information. Presented events have a reconstructed vertex.

6 Conclusion

We have developed a CNN-based multiple particle identification network, MPID, and applied it to images of event interactions in MicroBooNE data. This is the first demonstration of the performance of a CNN that incorporates systematic uncertainties in LArTPC data, and the first use of CNNs to perform particle identification on real LArTPC data. The network takes a 512×512 LArTPC image and calculates the probability scores for any particle in the image as p, e^-, γ, μ^-, and $\pi^\pm$. The training images are generated with a customized event generator that concatenates particles at the same vertex. The code for making the network and training sample are made available in MPID [36] and LArSoft [8].

10,000 1e-1p and 1μ-1p images are used to benchmark the network performance on simulated interactions. A score cut of 0.5 is applied on each particle. In the 1μ-1p sample, 8177 and 8522 events passed the selection for p and μ^-, or 82% and 85% in percentage. 117 and 4 events are mis-identified as e^- and γ, or 1% and 0% in percentage. In the 1e-1p sample, 8165 and 9079 events passed the selection for p and e^-, or 82% and 91% in percentage. 1546 and 106 events are mis-identified as γ and μ^-, or 15% and 1% in percentage.

Satisfactory agreement in all score distributions are found between data and simulation despite the many complexities of the MicroBooNE liquid argon TPC response, including inactive wire regions [37], electronics noise [37], signal processing [38, 39], and space charge effects [40].

We also demonstrated the metrics and performance of applying the MPID network on BNB beam data from MicroBooNE, which also illustrated the MPID network's clear capabilities in particle discrimination. When we take reconstructed vertex activity as input in filtered 1μ-1p candidate event images, MPID score distributions are indeed high for p and μ^-, and low for e^-, γ and $\pi^\pm$. When we instead take all pixel activity as input in filtered images containing π^0-produced γ rays, we see large differences between obtained e^- and γ scores. By applying these demonstrated particle identification capabilities to simulated BNB ν_e and ν_μ interactions, we have shown that this validated tool can play an important role in achieving a successful low-energy electron-like

excess measurement in MicroBooNE.

References

[1] R. Acciarri *et al.* (MicroBooNE Collaboration), Convolutional neural networks applied to neutrino events in a liquid argon time projection chamber, J. Instrum. 12, P03011 (2017).

[2] R. Acciarri *et al.* (ICARUS-WA104, LAr1-ND and MicroBooNE collaborations), A Proposal for a Three Detector Short-Baseline Neutrino Oscillation Program in the Fermilab Booster Neutrino Beam (2015), arXiv:1503.01520.

[3] A. A. Aguilar-Arevalo *et al.* (MiniBooNE Collaboration), The neutrino flux prediction at MiniBooNE, Phys. Rev. D79, 072002 (2009).

[4] R. Acciarri *et al.* (Dune Collaboration), Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report Volume 2: The Physics Program for DUNE at LBNF, arXiv:1512.06148.

[5] R. Acciarri *et al.* (MicroBooNE Collaboration), Design and Construction of the Micro-BooNE Detector, J. Instrum. 12, P02017 (2017).

[6] R. Acciarri *et al.*, Construction and assembly of the wire planes for the MicroBooNE Time Projection Chamber, J. Instrum. 10, T08006 (2015).

[7] J. Conrad, B.J.P. Jones, Z. Moss, T. Strauss and M. Toups, The photomultiplier tube calibration system of the MicroBooNE experiment, J. Instrum. 12, T06001 (2015).

[8] E. L. Snider and G. Petrillo, LArSoft: toolkit for simulation, reconstruction and analysis of liquid argon TPC neutrino detectors, J. Phys. Conf. Ser. 898, 042057 (2017).

[9] V. Genty, K. Terao, and T. Wongjirad, LArCV: LArTPC image processing/analysis framework for deep learning https://www.github.com/LArbys/LArCV.

[10] C. Adams *et al.* (MicroBooNE collaboration), A Method to Determine the Electric Field of Liquid Argon Time Projection Chambers Using a UV Laser System and its Application in MicroBooNE(2019), JINST 15 (2020) 07, P07010.

[11] A. Radovic *et al.*, Machine learning at the energy and intensity frontiers of particle physics, Nature 560, 41–48 (2018).

[12] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M.D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa and P. Vahle, A convolutional neural network neutrino event classifier, J. Instrum. 11 P09001–P09001 (2016).

[13] P. Adamson *et al.* (NOvA Collaboration), Constraints on Oscillation Parameters from ν_e Appearance and ν_μ Disappearance in NOvA, Phys. Rev. Lett. 118, 231801 (2017).

[14] P. Adamson *et al.* (NOvA Collaboration), New constraints on oscillation parameters from ν_e appearance and ν_μ disappearance in the NOvA experiment, Phys. Rev. D 98, 032012 (2018).

[15] M.A. Acero *et al.* (NOvA Collaboration), Measurement of Neutrino-Induced Neutral-Current Coherent π^0 Production in the NOvA Near Detector, Phys. Rev. D 102, 012004 (2020).

[16] F. Psihas, E. Niner, M. Groh, R. Murphy, A. Aurisano, A. Himmel, K. Lang, M. D. Messier, A. Radovic and A. Sousa, Context-Enriched Identification of Particles with a Convolutional Network for Neutrino Events, Phys. Rev. D 100, 073005 (2019).

[17] G. N. Perdue *et al.* (MINERvA Collaboration), Reducing model bias in a deep learning classifier using domain adversarial neural networks in the MINERvA experiment J. Instrum. 13, P11020–P11020 (2018).

[18] J. Renner *et al.* (NEXT Collaboration), Background rejection in NEXT using deep neural networks J. Instrum. 12, T01004–T01004 (2017).

[19] C. Adams *et al.* (MicroBooNE Collaboration), Deep neural network for pixel-level electromagnetic particle identification in the MicroBooNE liquid argon time projection chamber, Phys. Rev. D99, 092001 (2019).

[20] C. Adams *et al.* (MicroBooNE Collaboration), Reconstruction and measurement of $\mathcal{O}(100)$ MeV energy electromagnetic activity from $\pi^0 \to \gamma\gamma$ decays in the MicroBooNE LArTPC. J. Instrum. 15, P02007–P02007, (2020).

[21] P. Abratenko *et al.* (MicroBooNE Collaboration), Vertex-Finding and Reconstruction of Contained Two-track Neutrino Events in the MicroBooNE Detector, arXiv:2002.09375 (submitted to JINST).

[22] B Abi *et al.* (DUNE Collaboration), Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume II: DUNE Physics, arXiv:2002.03005 [hep-ex].

[23] A. A. Aguilar-Arevalo *et al.* (MiniBooNE Collaboration), Observation of a Significant Excess of Electron-Like Events in the MiniBooNE Short-Baseline Neutrino Experiment, Phys. Rev. Lett. 121, 221801 (2018).

[24] A. Krizhevsky *et al.*, Imagenet classification with deep convolutional neural networks, NIPS 25(2012) 1106.

[25] Rectifier neural networks, https://en.wikipedia.org/w iki/Rectifier_(neural_networks).

[26] Group Normalization, Yuxin Wu and Kaiming He, arXiv:1803.08494.

[27] PyTorch, Sigmoid, https://pytorch.org/docs/stable/nn .html.

[28] PyTorch, BCEWithLogitsLoss, https://pytorch.org/ docs/stable/nn.html.

[29] S.Agostinelli *et al.*, GEANT4—a simulation toolkit, Nucl. Instrum. Meth., A614,pp 87-104, 2010.

[30] J. Allison *et al.*, Geant4 developments and applications, IEEE Transactions on nuclear science, 53,1,pp 270-278, 2006.

[31] J. Allison *et al.*, Recent developments in GEANT4 Nucl. Instrum. Meth., A835, pp 186-225, 2016..

[32] C. Andreopoulos *et al.*, The GENIE Neutrino Monte Carlo Generator, Nucl.Instrum.Meth.A 614 (2010) 87-104.

[33] Nitish Srivastava *et al.*, Dropout: A Simple Way to Prevent Neural Networks from Overfitting, JMLR, 15(56):1929-1958, 2014.

[34] Matthew D Zeiler, Rob Fergus, Visualizing and Understanding Convolutional Networks, arXiv:1311.2901.

[35] J. S. Marshall and M. A. Thomson, The Pandora software development kit for pattern recognition, Eur. Phys. J. C 75, 439 (2015).

[36] R. An, Multiple Particle Identification Network, https: //github.com/ruian1/MPID.

[37] R. Acciarri *et al.* (MicroBooNE Collaboration), Noise Characterization and Filtering in the MicroBooNE Liquid Argon TPC, J. Instrum. 12, P08003 (2017).

[38] C. Adams *et al.* (MicroBooNE Collaboration), Ionization electron signal processing in single phase LArTPCs. Part I. Algorithm Description and quantitative evaluation with MicroBooNE simulation, J. Instrum. 13, P07006 (2018).

[39] C. Adams *et al.* (MicroBooNE Collaboration), Ionization electron signal processing in single phase LArTPCs. Part II. Data/simulation comparison and performance in MicroBooNE, J. Instrum. 13, P07007 (2018).

[40] P. Abratenko *et al.* (MicroBooNE Collaboration), Measurement of Space Charge Effects in the MicroBooNE LArTPC Using Cosmic Muons, arXiv:2008.09765.