

PAPER • OPEN ACCESS

The Convolutional Visual Network for Identification and Reconstruction of NOvA Events

To cite this article: Fernanda Psihas and NOvA Collaboration 2017 *J. Phys.: Conf. Ser.* **898** 072053

View the [article online](#) for updates and enhancements.

The Convolutional Visual Network for Identification and Reconstruction of NOvA Events

Fernanda Psihas, For the NOvA Collaboration.

Address: 727 E. Third St., Swain Hall West, Room 117, Bloomington, IN 47405-7105

E-mail: psih@fnal.gov

Abstract. In 2016 the NOvA experiment released results for the observation of oscillations in the ν_μ and ν_e channels as well as ν_e cross section measurements using neutrinos from Fermilab's NuMI beam. These and other measurements in progress rely on the accurate identification and reconstruction of the neutrino flavor and energy recorded by our detectors. This presentation describes the first application of convolutional neural network technology for event identification and reconstruction in particle detectors like NOvA. The Convolutional Visual Network (CVN) Algorithm was developed for identification, categorization, and reconstruction of NOvA events. It increased the selection efficiency of the ν_e appearance signal by 40% and studies show potential impact to the ν_μ disappearance analysis.

1. Introduction to NOvA's Measurement

NOvA is a long baseline neutrino oscillations experiment whose main purpose is to study the appearance of electron neutrinos in a muon neutrino source. It does so by measuring the neutrino signal close to its source, at Fermilab, as well as 810 km away, at Ash River, MN. Oscillations measurements rely on the identification of the flavor of neutrino interactions in typically low statistics samples. Thus, high efficiency event identification techniques are a crucial component of these analyses.

The NOvA detectors are two low-Z material sampling calorimeters, designed to optimize electron identification. They are composed of alternating vertical and horizontal planes of liquid scintillator. This array allows us to produce two views of the detector activity, which in detector coordinates are a view on the XZ plane (from the top) and one on the YZ plane (from the side).

1.1. NOvA Neutrino Events

NOvA events are typically 550 microsecond readouts of the detector electronics, centered around the ~ 10 microsecond neutrino beam spills from Fermilab's Main Injector. The interactions are isolated by correlating groups of hits in time and space in order to separate them from the rest of the activity in the detector. Images from the NOvA event display such as Figure 1 typically display hits in color, representing either the time or deposited charge of each hit.

Neutrino interactions in these events can be classified by analyzing the topology and energy deposition profile of the hits from them. The main approach at signal identification employed



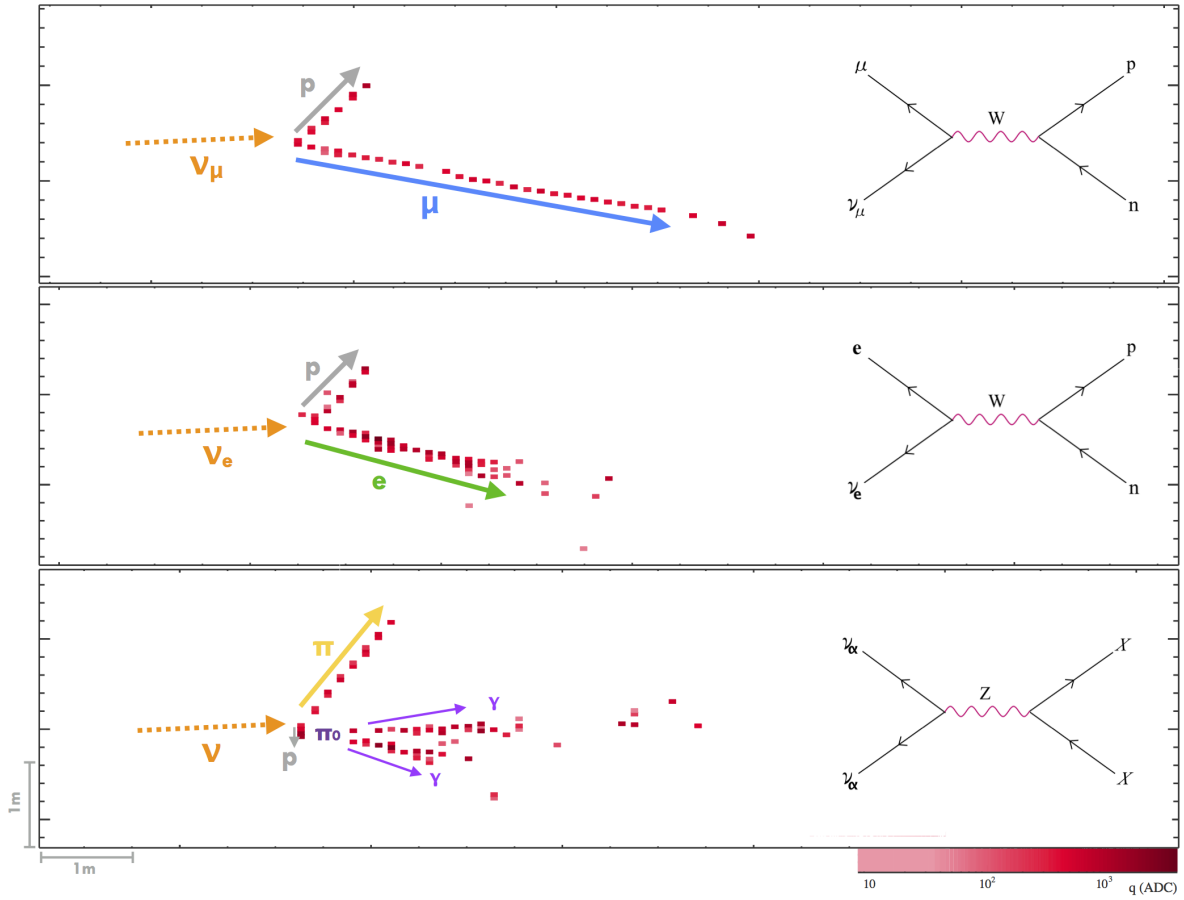


Figure 1. NOvA characteristic data events. Side views of 3x11 meter sections of the detector. The color of the hits indicates deposited charge (measured in ADC counts). The neutrino neutral current interactions (bottom), as well as the charged current interactions for electron (middle) and muon (top) flavor are each the main signal on NOvA's neutral current, ν_e appearance and ν_μ disappearance analyses, respectively. This makes the classification of these events the crucial first step for these analyses.

for our first analyses[1, 2] was done in two main steps. First, reconstruction algorithms make a geometrical separation of each particle's contribution to the event. Then, identification algorithms extract physics information, i.e. dE/dx and projected trajectory, from each particle's contribution (given as a cluster of hits) and attempt to identify the leptonic component of the interaction¹ by using neural networks trained on these features.

2. The CVN Convolutional Neural Network

2.1. Advantages of Convolutional Neural Networks

Deep learning algorithms[7] have been successful in tasks like image recognition[6, 9]. These networks—and in particular convolutional neural networks (CNNs)—present several advantages with respect to the traditional identification methods described in Section 1. Not only do traditional algorithms rely heavily on the efficiency of the geometric separation of the components, they are also limited in that the features they employ for identification are only those

¹ As seen in Figure 1 the outgoing lepton carries the same flavor as the original neutrino by lepton conservation.

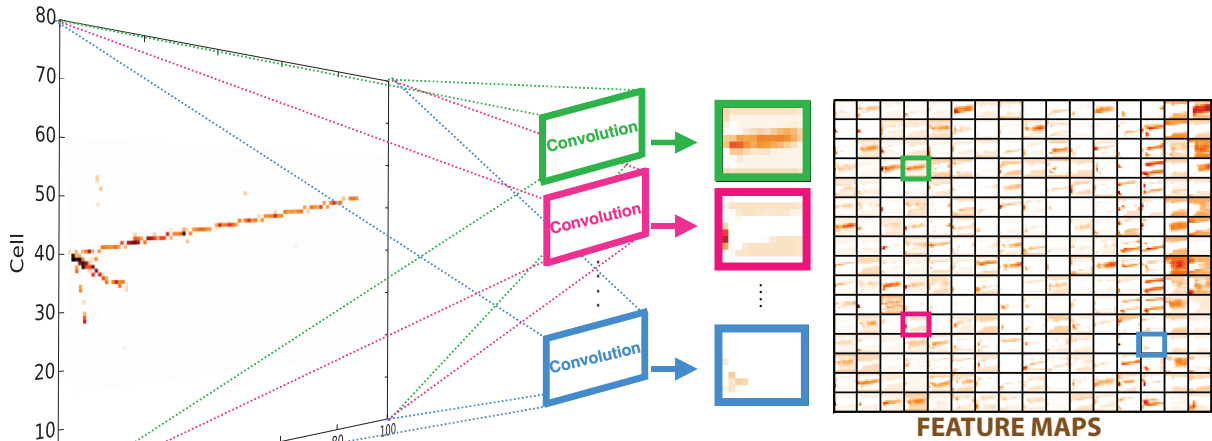


Figure 2. Pixel map for simulated ν_μ CC event. This diagram exemplifies how the convolutional operations extract features from the images. Some track activity is visible in one of the feature maps.

which have been previously selected by the designer, equally folded in with the efficiency of their extraction. In turn, convolutional neural networks—in their implementation for image classification[10]—eliminate the need for both previous separation of the event into components and extraction of predefined features. The following is a brief explanation of the main features of CNN architectures:

- **Convolutional Layers:** Employ the use of feature extraction kernels of various types to extract features from the image. Kernels operate on images to extract different features (averages, edge effects, etc.)². The network then learns and trains on correlations between these feature maps for events of each type.
- **Inception Modules:** Multiple kernels of different dimensions are used to extract features at multiple scales simultaneously. Figure 2 shows a set of feature maps produced by an inception module.
- **Pooling Layers:** The feature maps are down-sampled by replacing regions with the output of a function (average, maximum value, etc.).
- **Dropout Layers:** Select a random sub-sample of existing connections to be used at every iteration by randomly resetting weights.

2.2. CVN

We have developed a Convolutional Visual Network (CVN) based on existing implementations of machine learning[3] which classifies neutrino interactions by categories. The architecture of the CVN is depicted in Figure 3. We use a Siamese-style architecture optimized for categorization of thirteen types of events, labeled by neutrino flavor and interaction type. These categories can

² The kernels also evolve through local response normalization as training progresses.

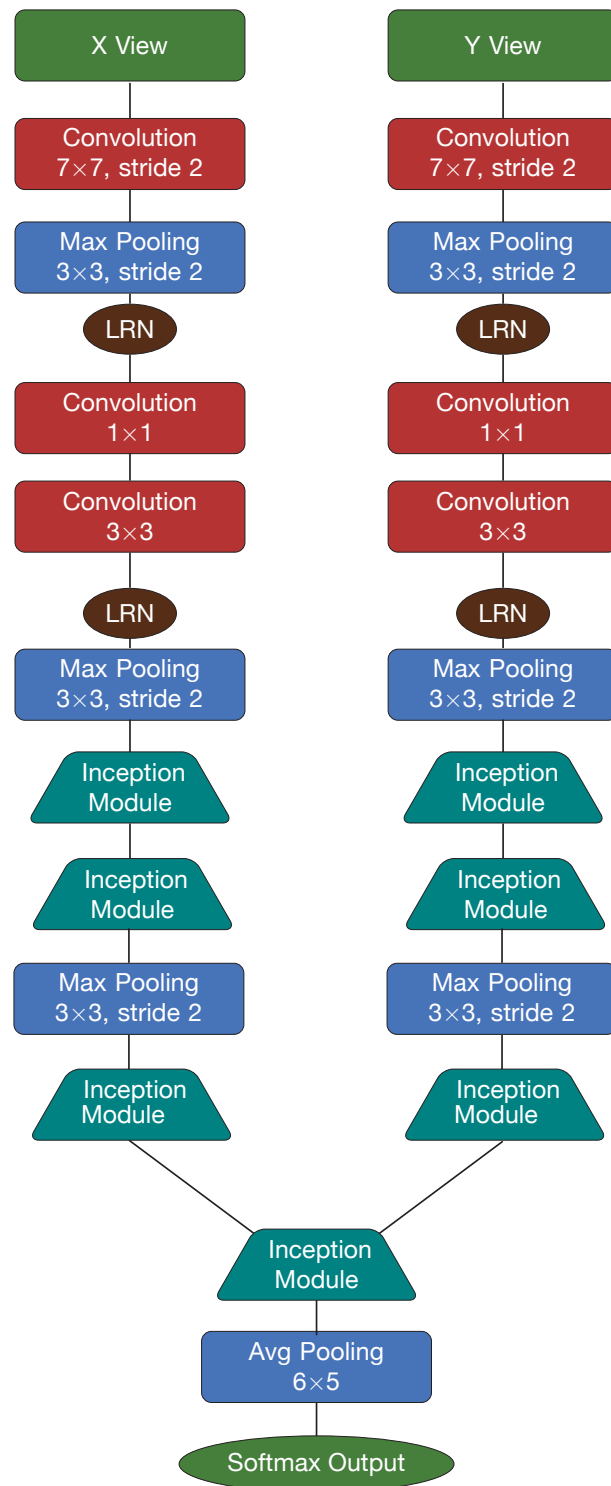


Figure 3. Diagram of the CNN architecture used for event classification. Starting with the input at the top, the network has separate branches for the XZ and YZ views of the event. Each branch undergoes successive convolution, pooling, and local response normalization (LRN). Inception modules are used in downstream layers. The two views are merged and passed through a final inception module, and pooled. The output of the network comes from softmax units.

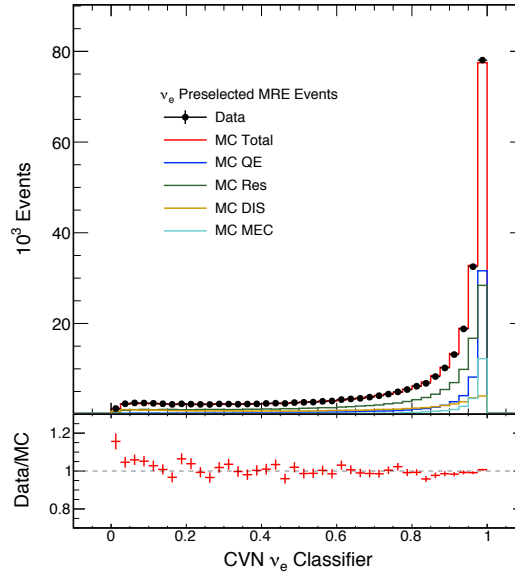


Figure 4. The CVN ν_e PID distribution for muon-removed electron (MRE) events in the near detector after basic preselection has been applied.

then be labeled as signal or background for each analysis independently.

The CVN network is trained on pixel maps of simulated events and labels corresponding to the type of simulated interaction. The maps are analogous to images but constructed from the output of the detector. One example of a side-view pixel map is shown in Figure 2. Both the side and top views of each event are trained on to provide further context of the topology. This implementation was trained using the caffe framework[5].

3. Performance of CVN

3.1. Performance Tests on Data

One test of the performance of the classifier on data uses the Muon Removed - Electron (MRE) technique. It begins by selecting a muon neutrino interaction with traditional identification methods. The hits from the muon are then removed from the event and they are replaced by hits from a single simulated electron of matching momentum. Data/MC comparisons using this technique (Figure 4) show less than 1% difference in efficiency.

3.2. Results Using CVN

The CVN event classifier is implemented as a selector for the ν_e appearance analysis presented in the summer of 2016[1]. It improved the efficiency of ν_e CC events by 40% with respect to the traditional algorithm[4]. This is the first implementation of CNNs in a high energy physics result, followed by NOvA's neutral current neutrino disappearance analysis. A CVN-based ν_μ disappearance analysis of NOvA's initial dataset reveals potential improvements on our sensitivities as well as signal efficiency[8] and this technique is currently being incorporated into the main branch of this analysis as well.

4. CVN for Reconstruction

While the current implementation of CVN focuses on event classification, it is possible to train a network for single particle identification through the clusters of hits it contributes to the event.

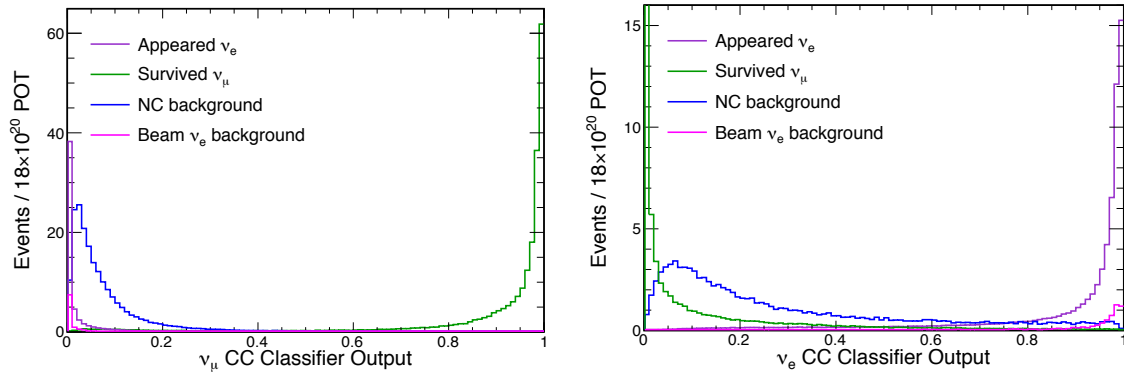


Figure 5. Signal and background CVN PID distributions for the ν_μ disappearance (left) and the ν_e appearance (right) analyses.

In this case a similar architecture to Figure 3 is used, with four rather than two main branches. This network trains on two views of the full event as well as two views of the isolated cluster. These clusters need to be previously defined as a collection of hits by some other algorithm. Adding the single cluster views to the network input allows it to combine features extracted from the particle signature itself as well as in context with the rest of the interaction. Currently the input clusters to our particle CVN classifier come from traditional cluster reconstruction[4] but there is ongoing work to disentangle the network from traditional reconstruction completely.

The CVN particle classifier is currently in development, but initial studies already show promise, especially for e^-/π^0 and μ/p discrimination, as seen by their respective efficiencies in Figure 6. Optimization of this classifier is ongoing and preliminary versions of it are already being implemented in studies of cross sections and energy reconstruction. Work is also ongoing to design a network for hit-by-hit identification within an event. This type of identification could influence the existing approaches to reconstruction.

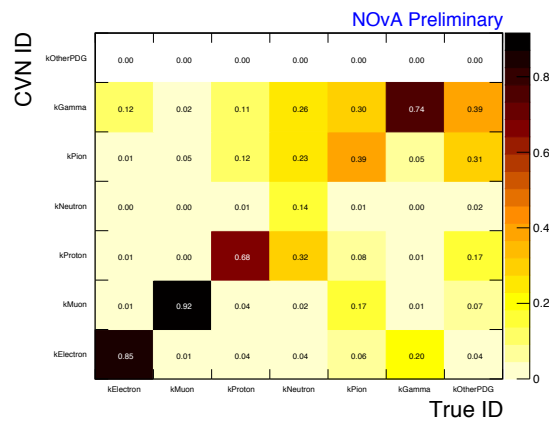


Figure 6. CVN for Prong ID tests on 50% purity clusters from simulated NuMI events. The entries along the diagonal represent efficiencies of categorizing by the best PID value across categories rather than cutting on a PID value.

References

- [1] P. Adamson et al. First measurement of electron neutrino appearance in NOvA. *Phys. Rev. Lett.*, 116(15):151806, 2016.
- [2] P. Adamson et al. First measurement of muon-neutrino disappearance in NOvA. *Phys. Rev.*, D93(5):051104, 2016.
- [3] 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. *JINST*, 11(09):P09001, 2016.
- [4] M. Baird, J. Bian, M. Messier, E. Niner, D. Rocco, and K. Sachdev. Event Reconstruction Techniques in NOvA. *J. Phys. Conf. Ser.*, 664(7):072035, 2015.
- [5] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*, 2014.
- [6] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25*, pages 1097–1105, 2012.
- [7] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521:436–444, 2015.
- [8] Dominick Rosario Rocco. *Muon Neutrino Disappearance in NOvA with a Deep Convolutional Neural Network Classifier*. PhD thesis, Minnesota U., 2016.
- [9] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. 2015.
- [10] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott E. Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. *CoRR*, abs/1409.4842, 2014.