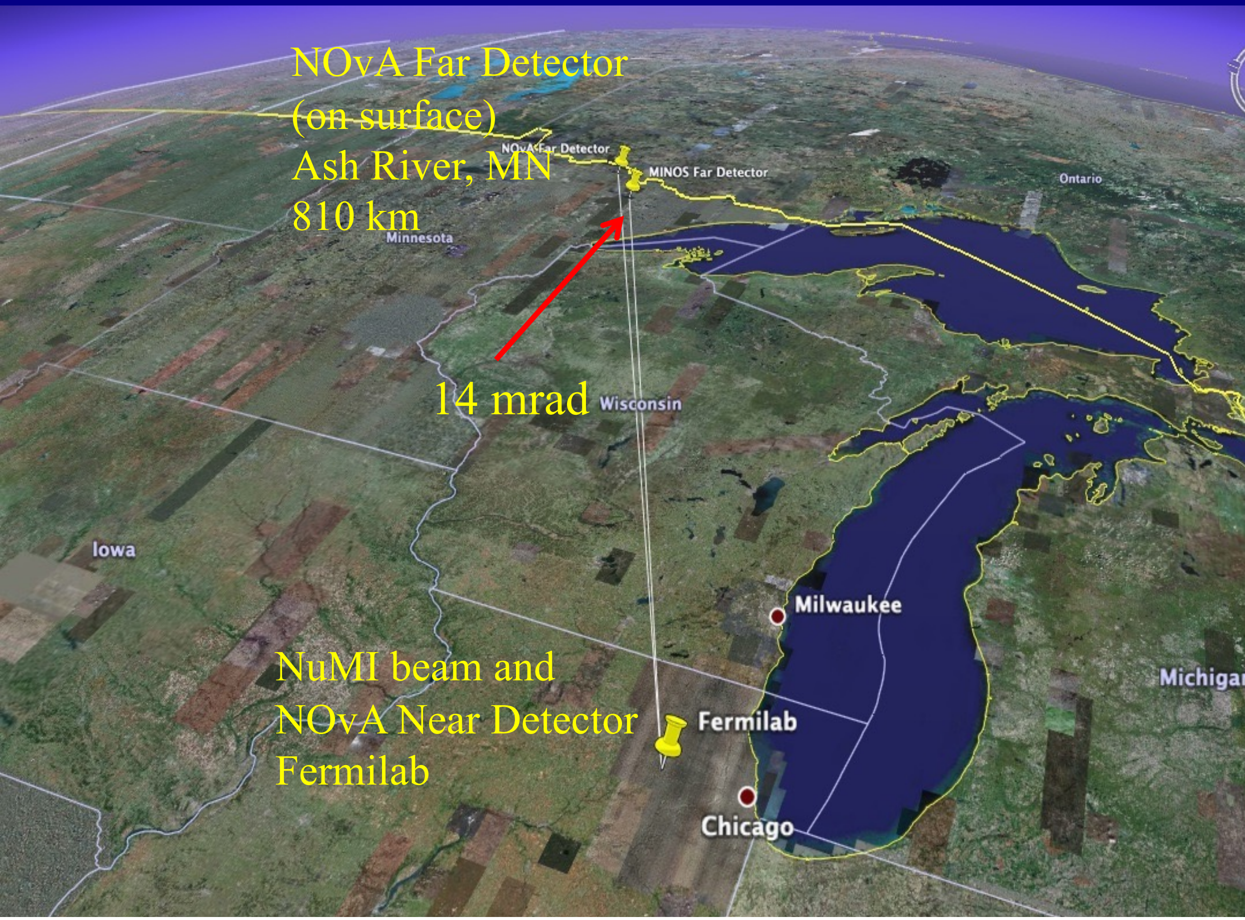
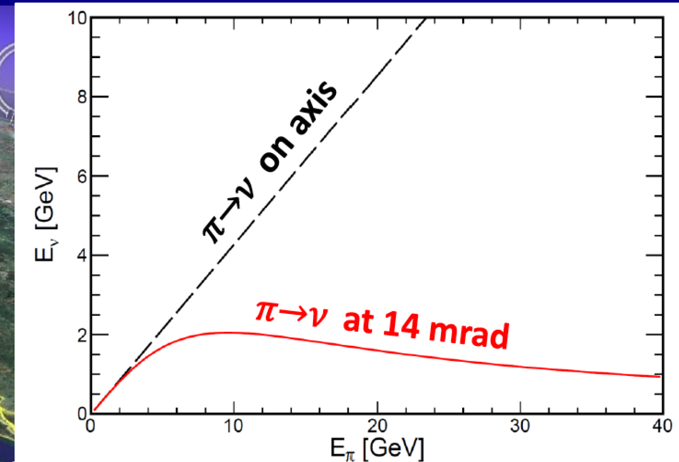
Deep Learning Reconstruction in the NOvA Experiment

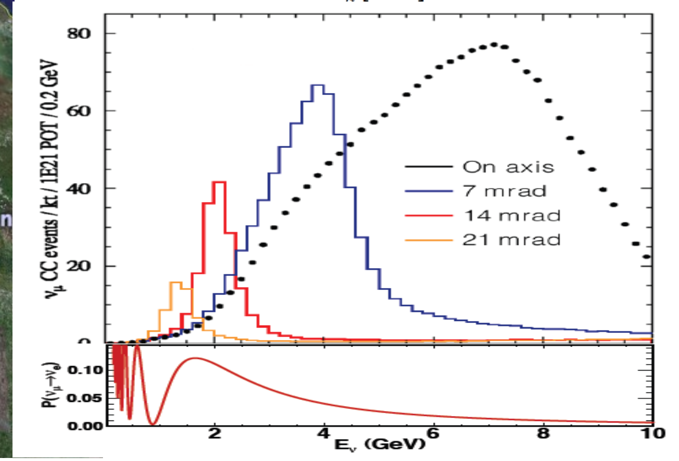
Jianming Bian
University of California, Irvine
May 12, 2022

CosSURF 2022, South Dakota Mines Campus

NuMI Off-Axis ν_e Appearance Experiment (NOvA)







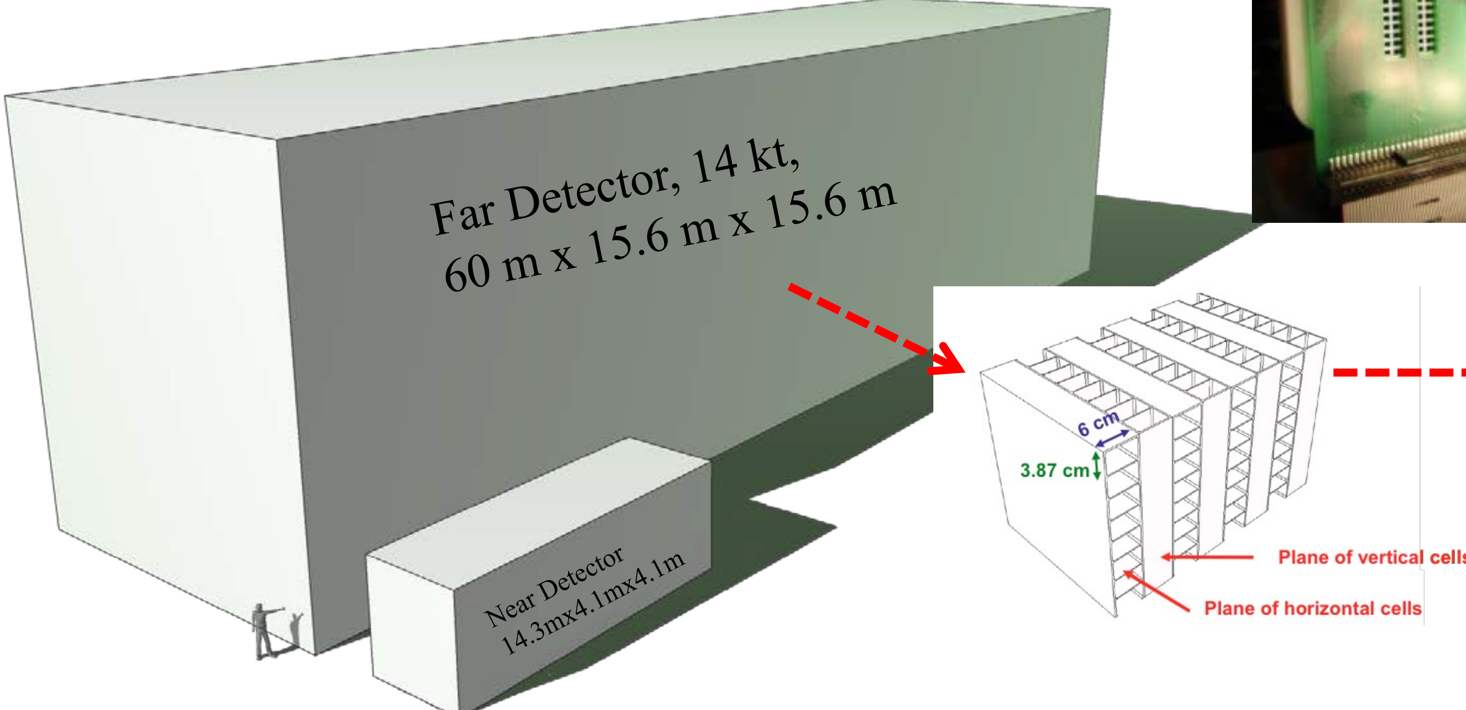
- Muon neutrino beam at Fermilab near Chicago
- Longest baseline in operation (810 km), large matter effect, sensitive to mass ordering
- Far/Near detector sited 14 mrad off-axis, narrow-band beam around oscillation maximum

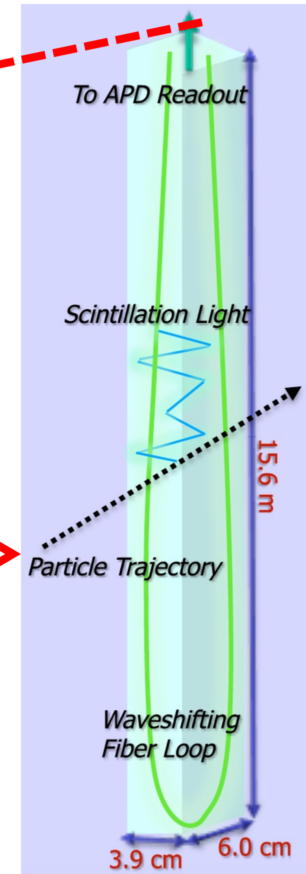
NOvA Detectors

Far Detector (FD):

- 14-kton, fine-grained
- 344k detector cells

0.3-kton functionally identical Near Detector (ND), ~20k cells

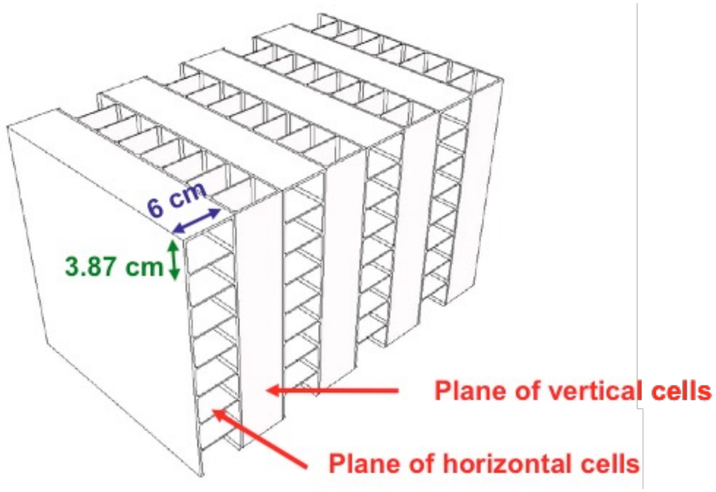


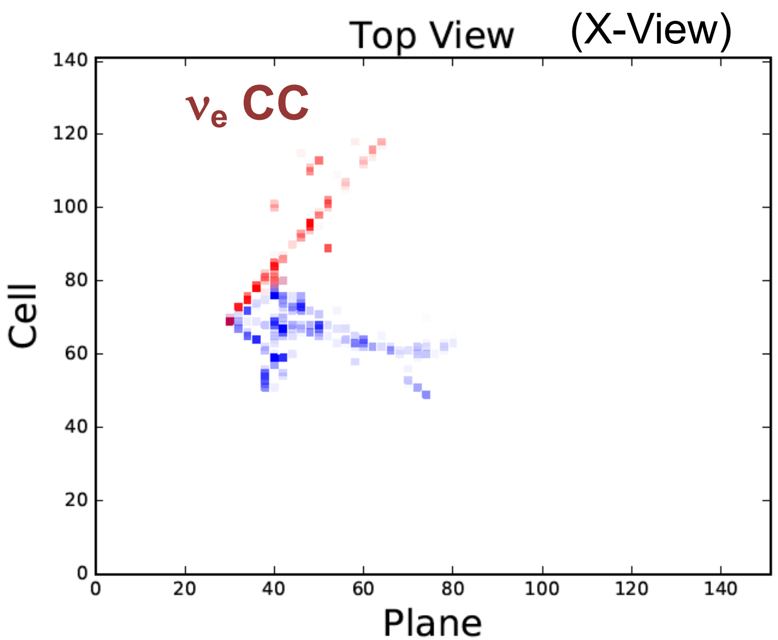


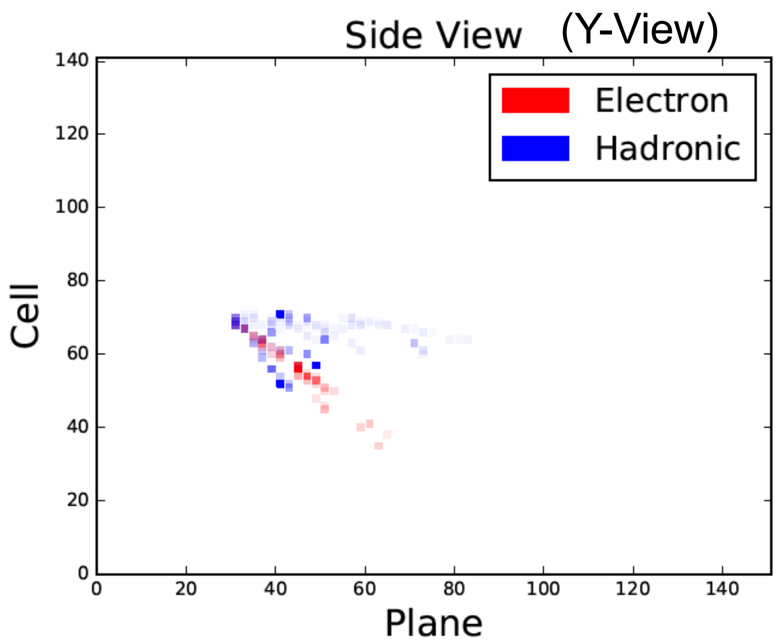
- Detectors are composed of PVC modules extruded to form long tube-like cells
- Each cell: filled with liquid scintillator, has wavelength-shifting fiber (WLS) routed to Avalanche Photodiode (APD)
- Cells arranged in planes, assembled in alternating vertical and horizontal directions → 3-D information of neutrino interactions

NOvA Event Images

- NOvA detector cells arranged in planes, assembled in alternating X and Y directions
- Produce a pair of pixel maps (Cell Number vs. Plane Number) for the X and Y view of each interaction

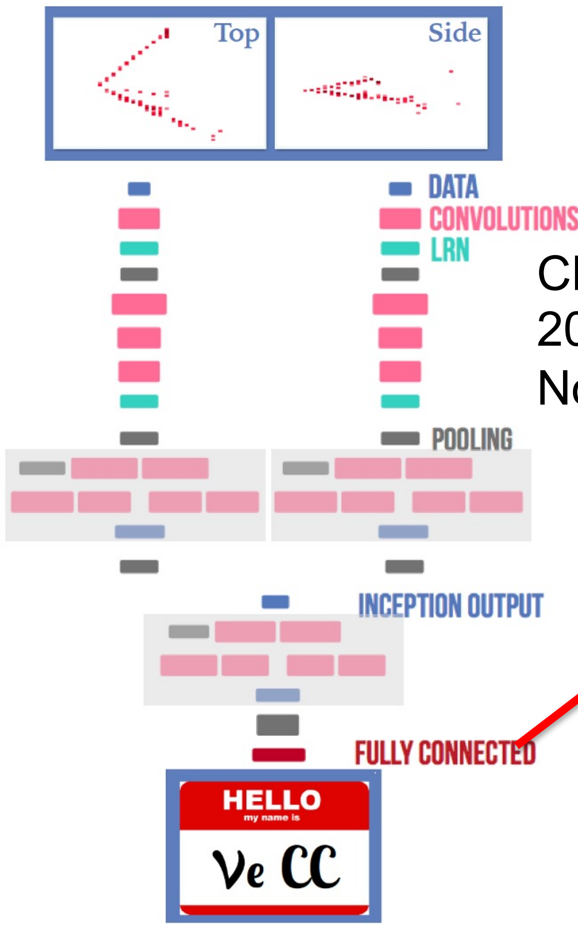






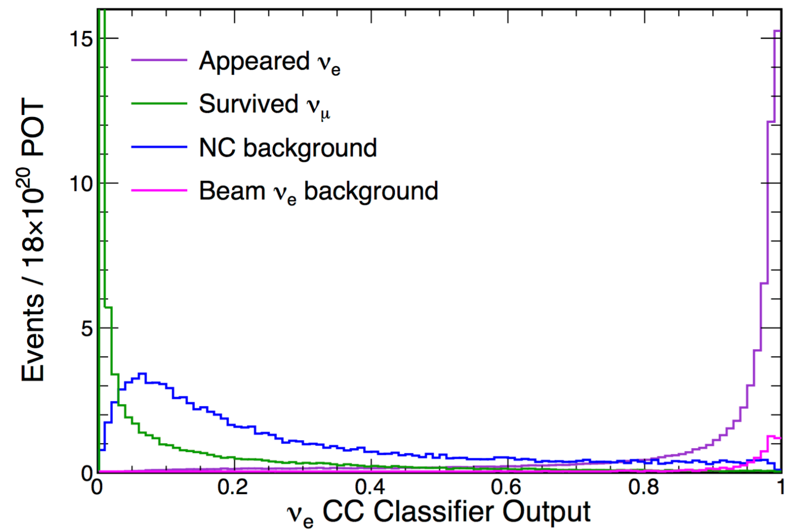
CNN based Event Classifier (CVN)

- CVN: a convolutional neural network (CNN), based on modern image recognition technology, extract features directly from pixel maps
- NOvA is the first HEP experiment to use CNNs to publish physics results: *Phys.Rev.Lett. 118 (2017)*
- Yielded an equivalent 30% increase in exposure than traditional methods



CNN architecture
2016: GoogLeNet
Now: MobileNet

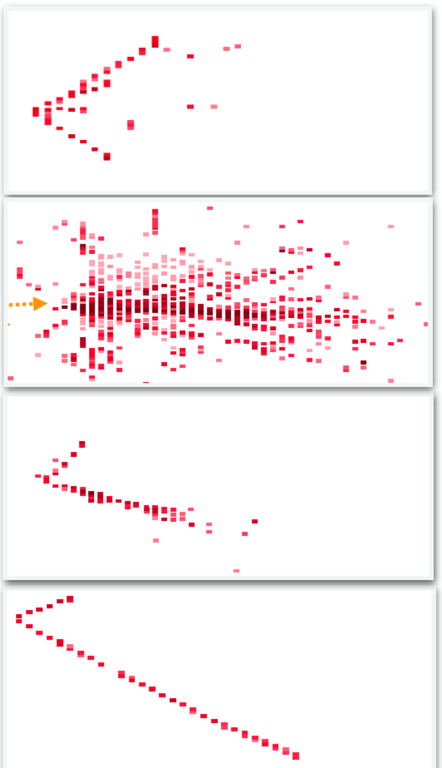
CVN output in the far detector MC

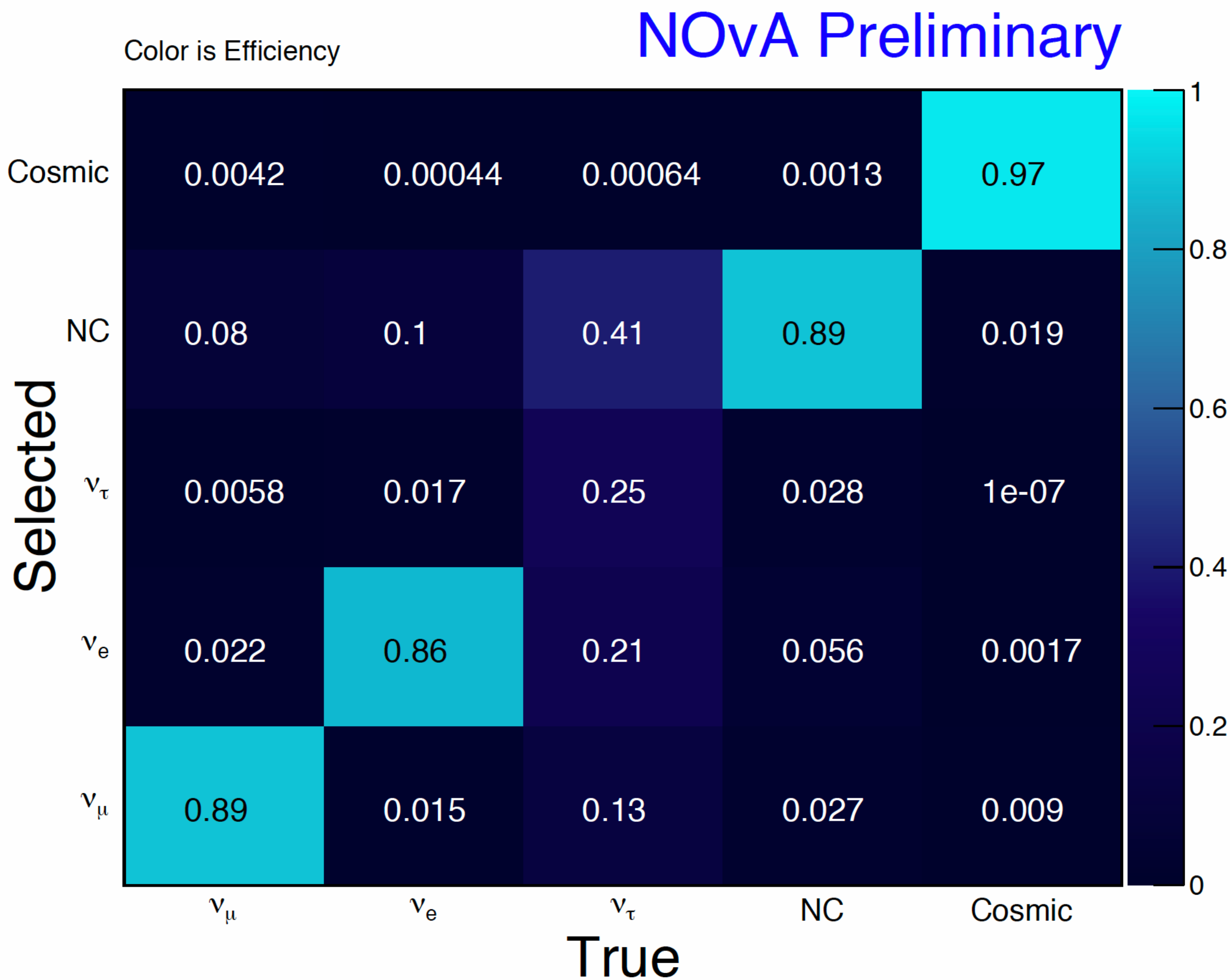


A. Aurisano et. al, JINST 11, P09001 (2016)

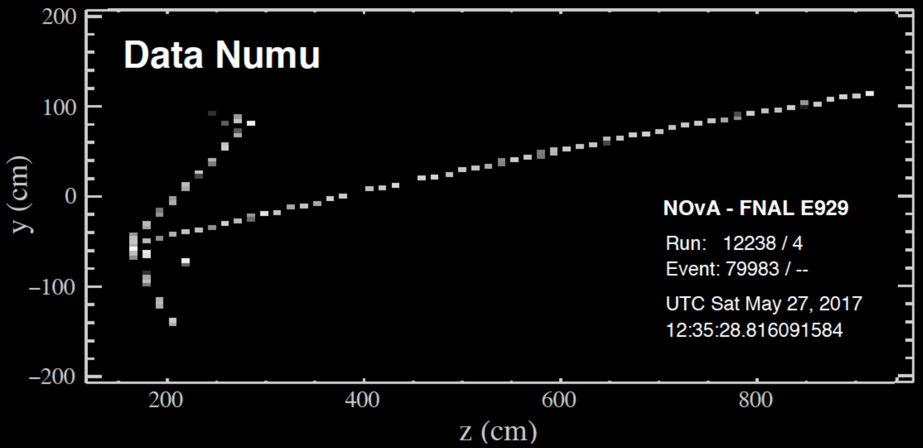
Select ν_μ ($\bar{\nu}_\mu$) CC and ν_e ($\bar{\nu}_e$) CC candidates from neutrino (antineutrino) beam with CVN in Near Detector (ND) and Far Detector (FD)

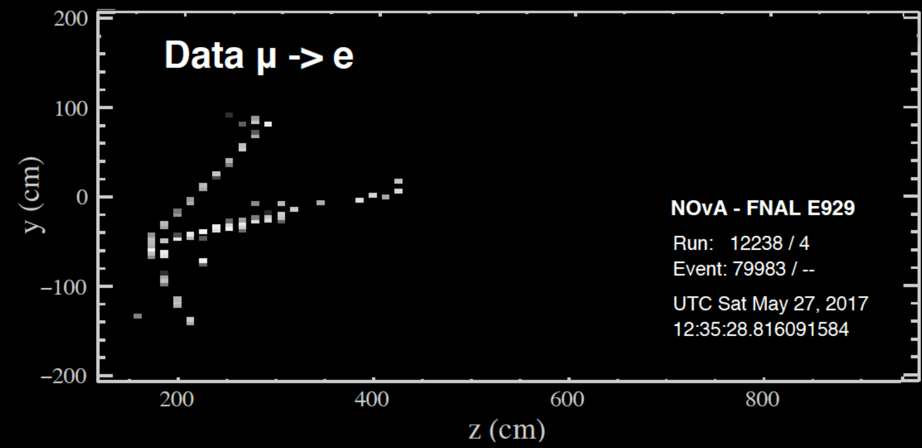
CNN based Event Classifier (CVN)

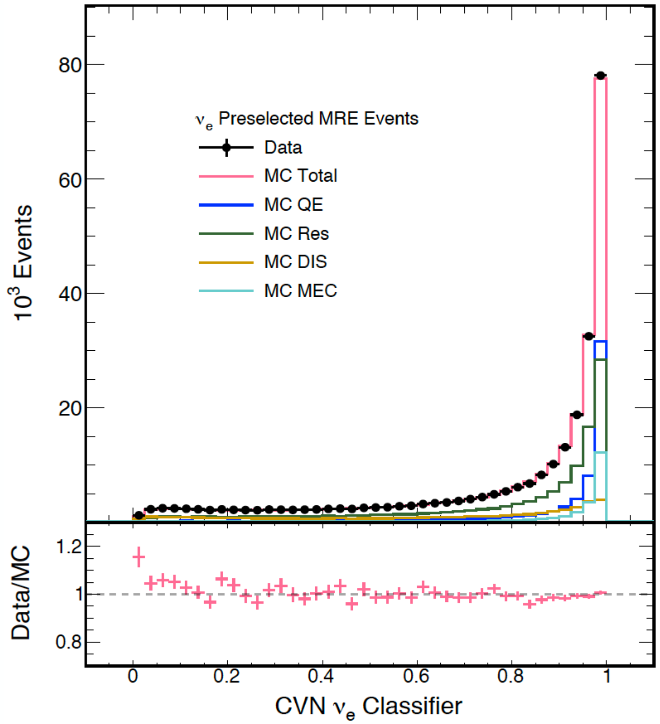




Example Data Check: MRE







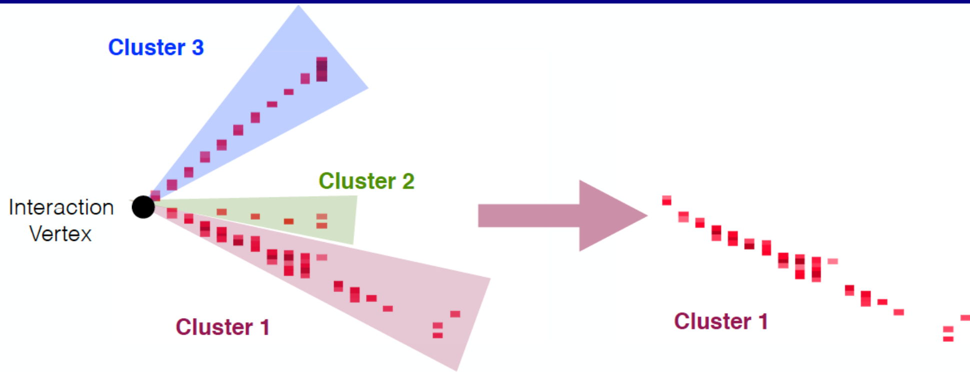
Muon Removed - Electron Added:

Select a muon neutrino interaction.

Remove the muon hits and replace with a simulated electron.

	Pre Selection	Full Selection	Efficiency
Data Events	486083	316009	0.6501
MC Events	511287	341119	0.6672

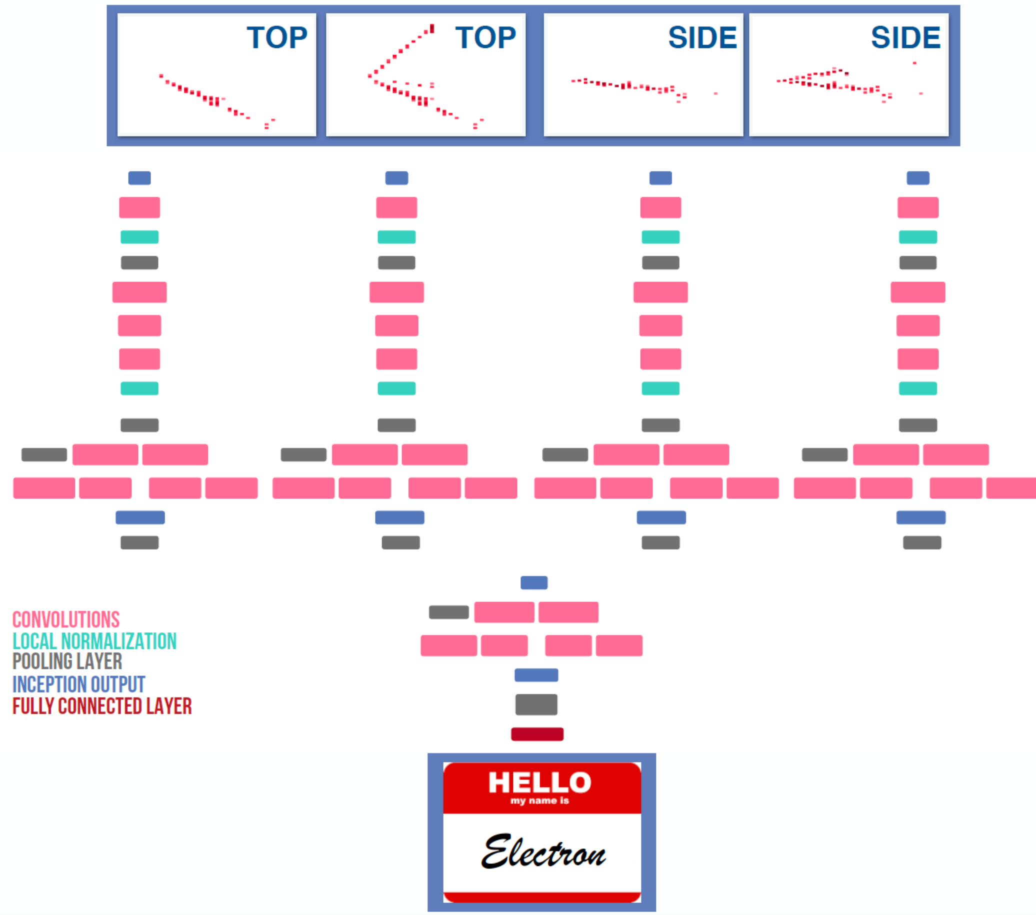
CNN based Particle Classifier (ProngCVN)



Single particles are currently separated using geometric reconstruction methods.

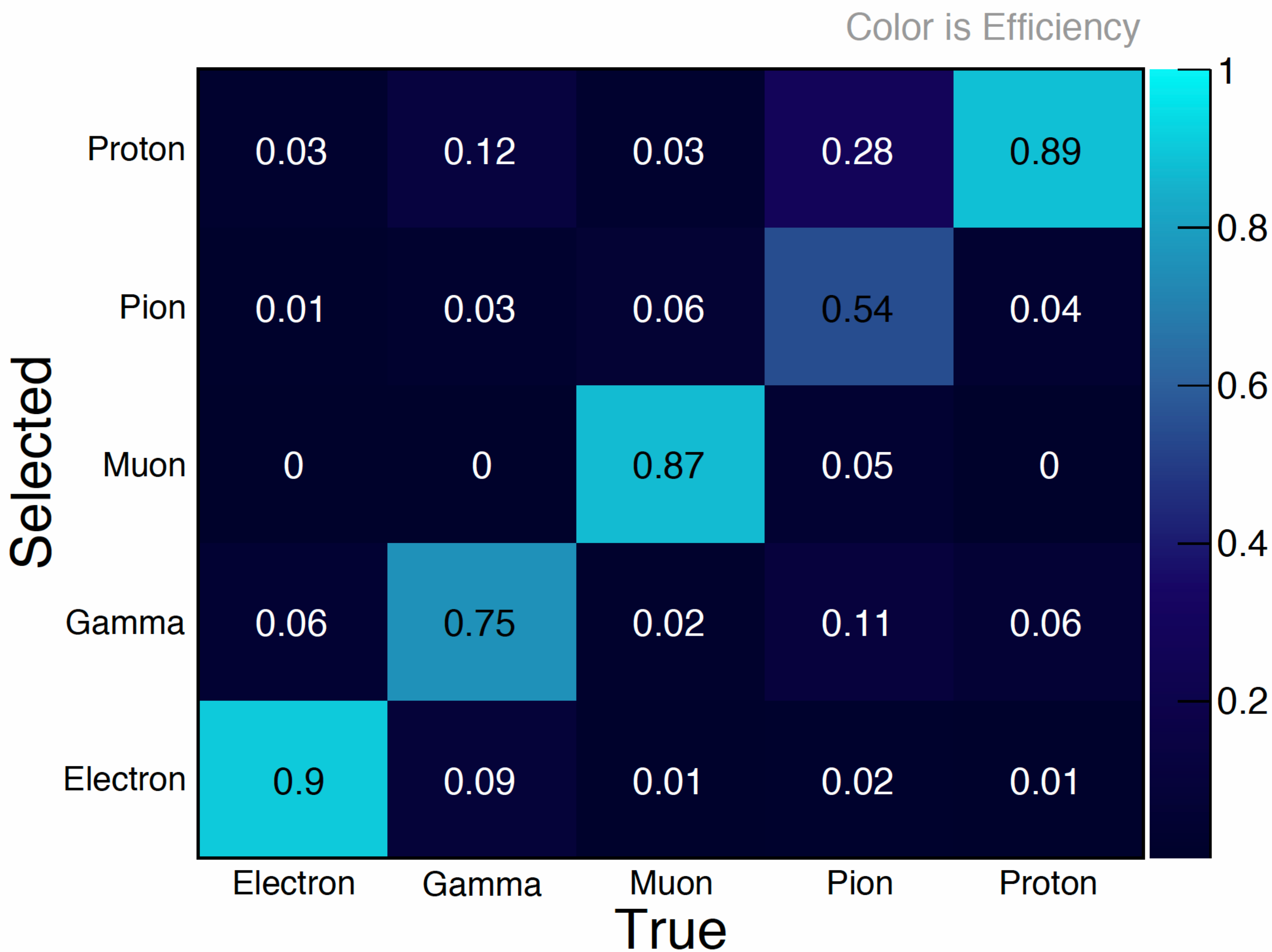
Classify particles using both views of the **particle** and both views of the entire **event**.

This shows the network **contextual information** about single particles.



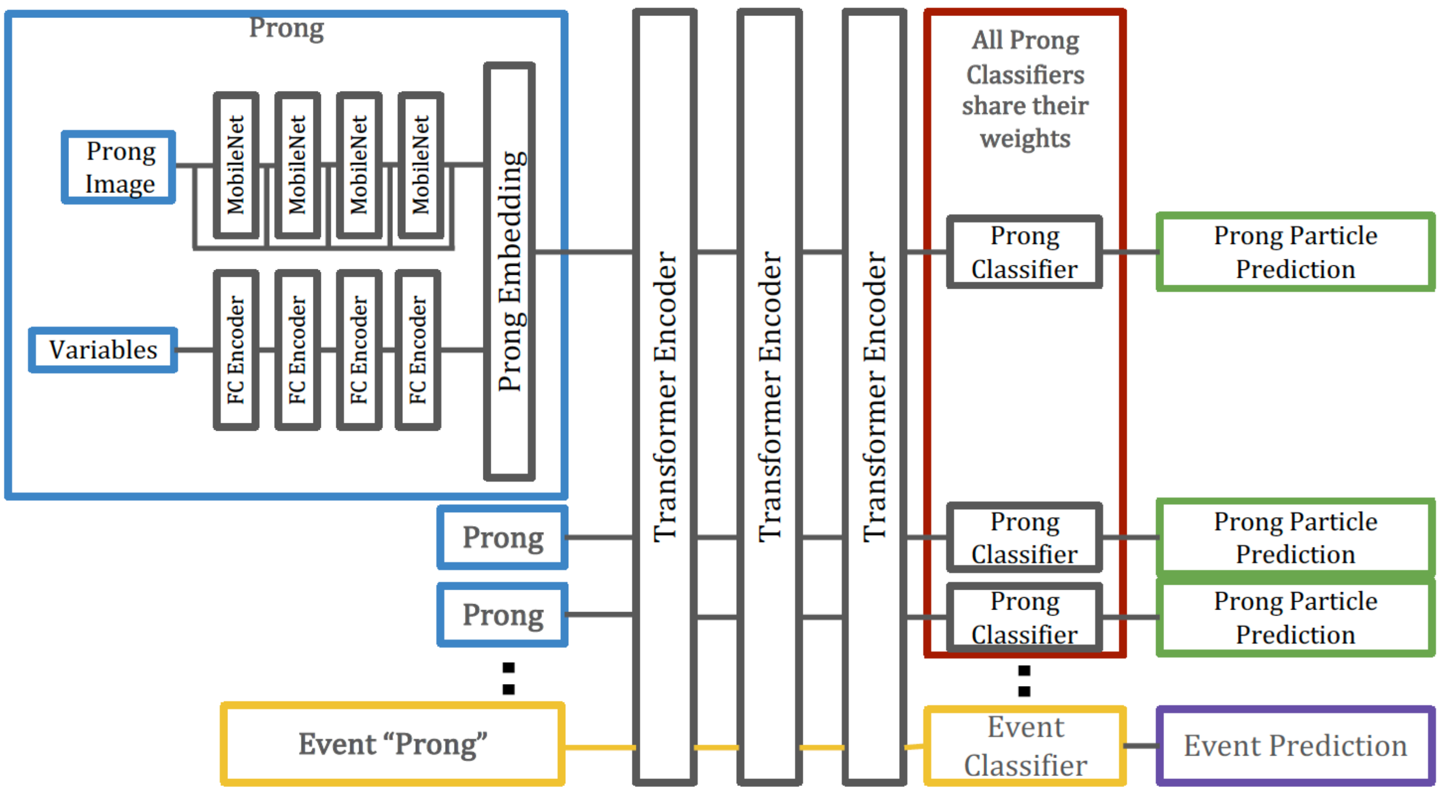
Phys.Rev.D 100 (2019) 7, 073005

CNN based Particle Classifier (ProngCVN)

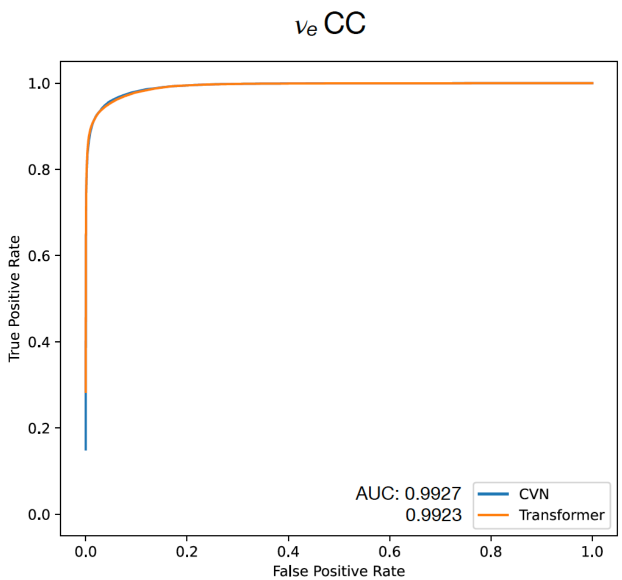


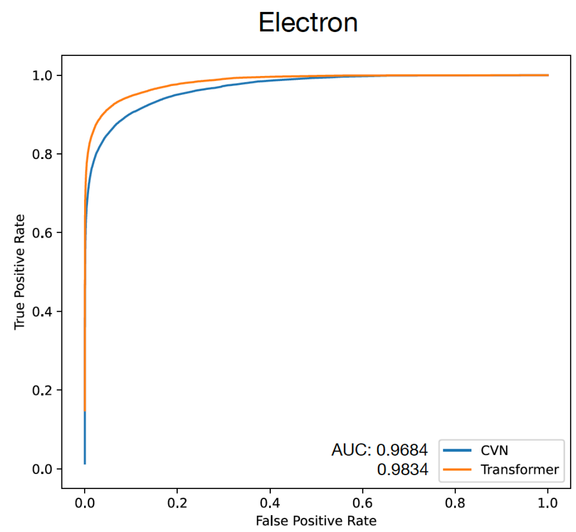
Transformer for both Event and Particle Classification (Transformer CVN)

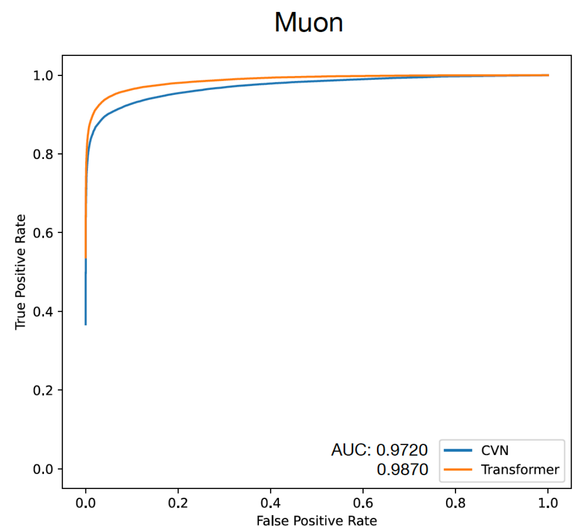
- Transformer is attention-based network trained on vector of objects, recently developed for Natural Language Processing in CS
- Deals with various types of inputs → combine pixel maps and particle level information to produce event and particle classification
- The attention mechanism in Transformers can be used to study correlations between inputs and outputs, makes each step in ML/AI based reconstruction checkable and explainable



Transformer for both Event and Particle Classification (Transformer CVN)



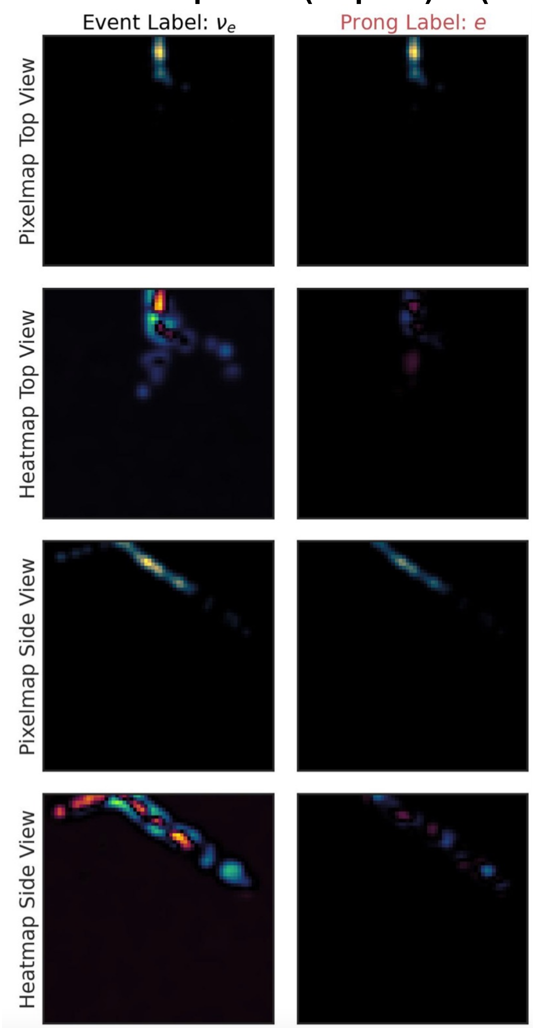




Transformer CVN's ROC curves for prong ID outperforms ProngCVN and event ID are nearly identical to event CVN.

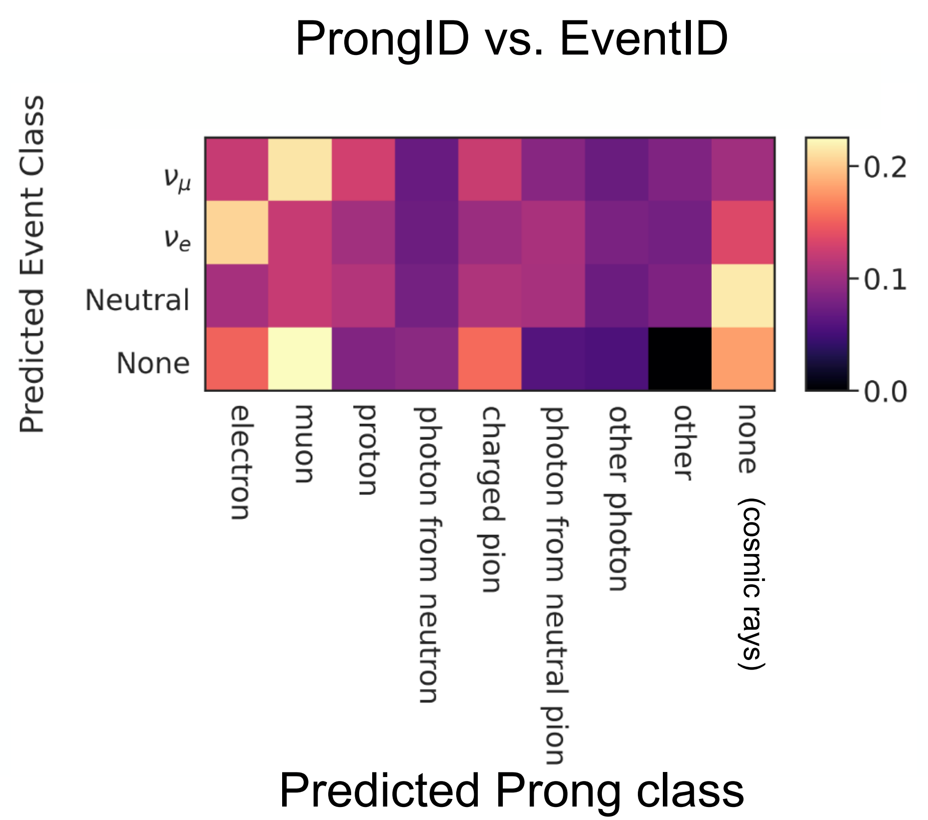
Example Impact Analysis with Transformer

Heat Map = d(input)/d(output)



Red: positively correlated

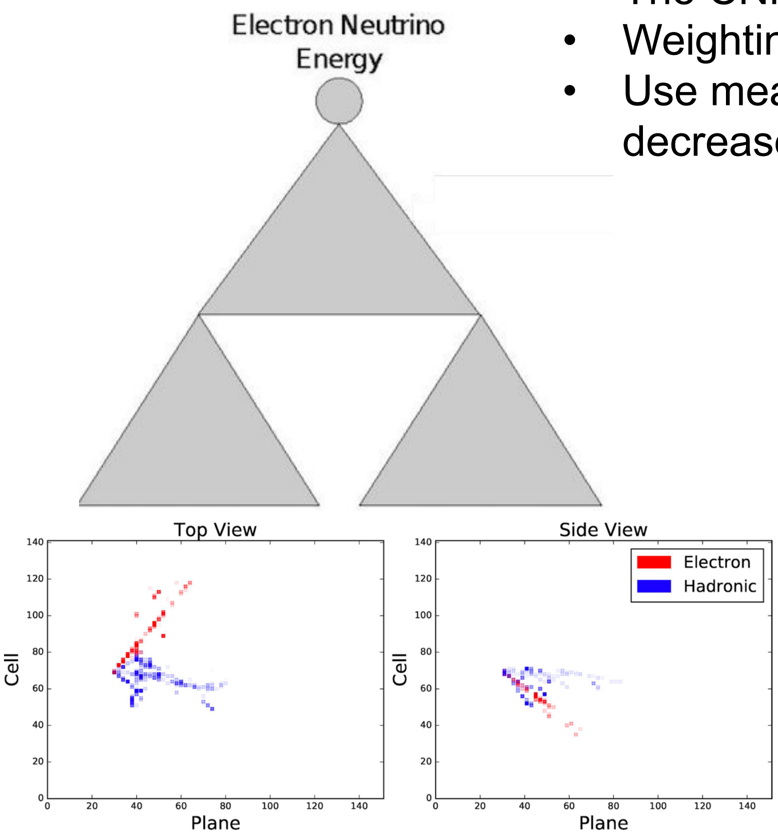
Blue: negatively correlated



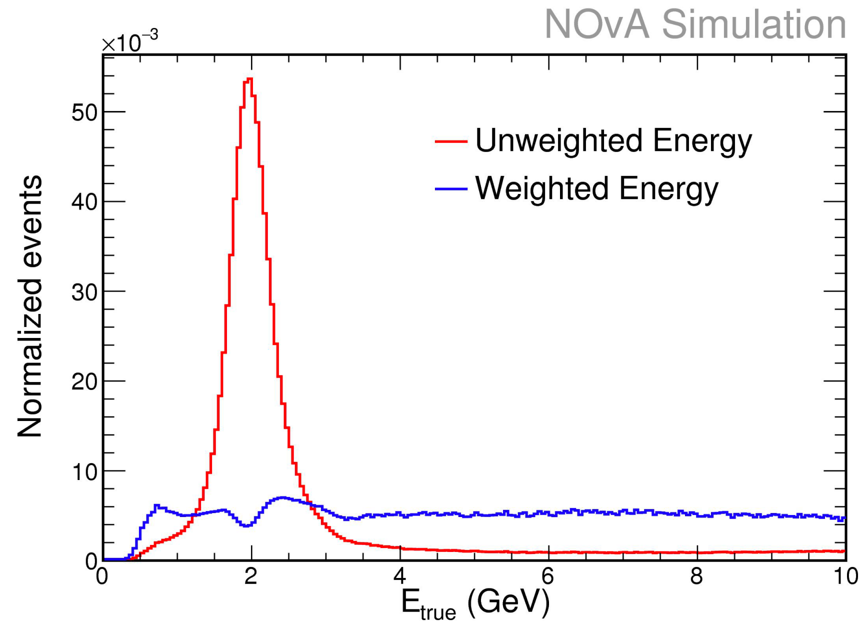
An ν_e CC event, with event and prongs identified correctly

Regression CNNs for Energy Estimation

- The CNN architecture used is an adapted ResNet
- Weighting scheme so the loss function sees a flat distribution
- Use mean absolute percentage error instead of square of errors to decrease the effects of outliers



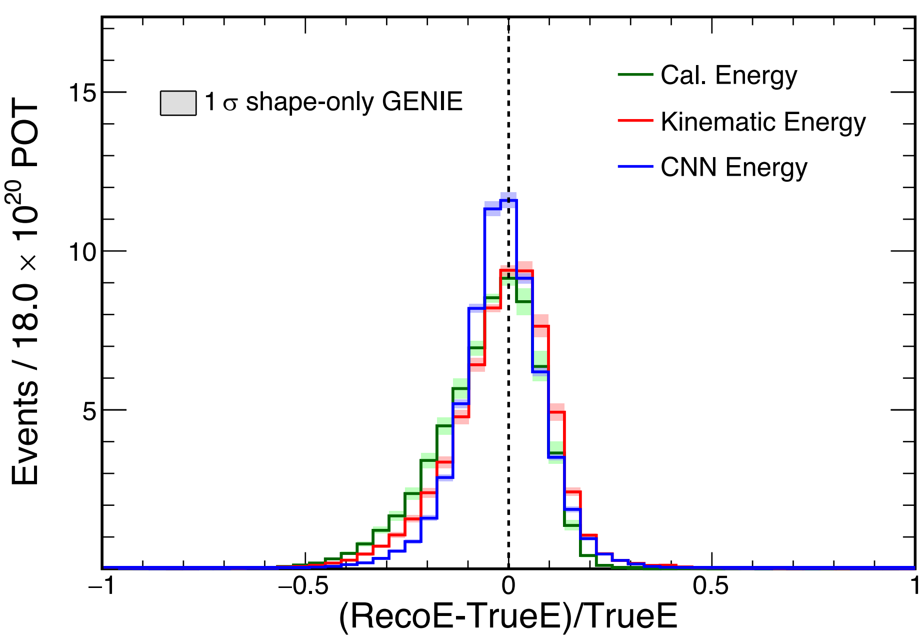
$$L(\mathbf{W}, \{\mathbf{x}_i, y_i\}_{i=1}^{n}) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{f_{\mathbf{W}}(\mathbf{x}_i) - y_i}{y_i}\right|$$

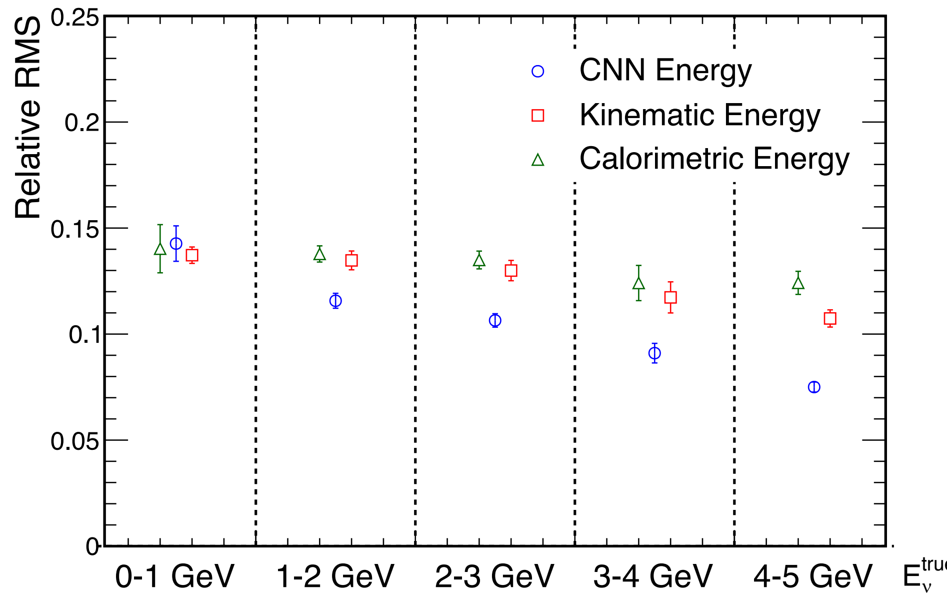


Phys. Rev. D **99**, no. 1, 012011 (2019)
doi:10.1103/PhysRevD.99.012011

Regression CNNs for Energy Estimation

- Compared with traditional kinematics-based energy reconstruction, regression CNN shows a better resolution
- Also shows smaller systematic uncertainties due to neutrino interaction simulation

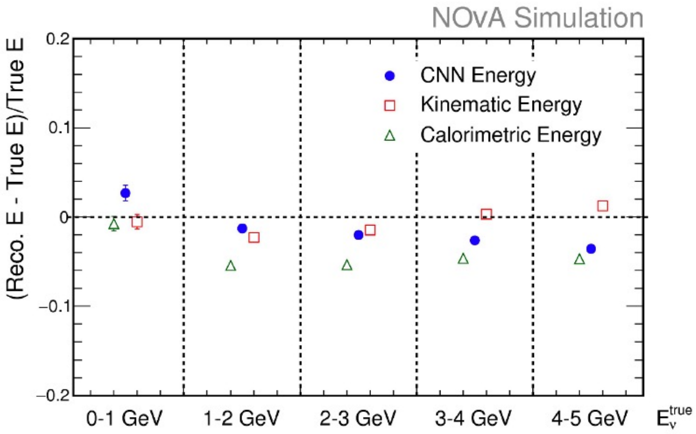


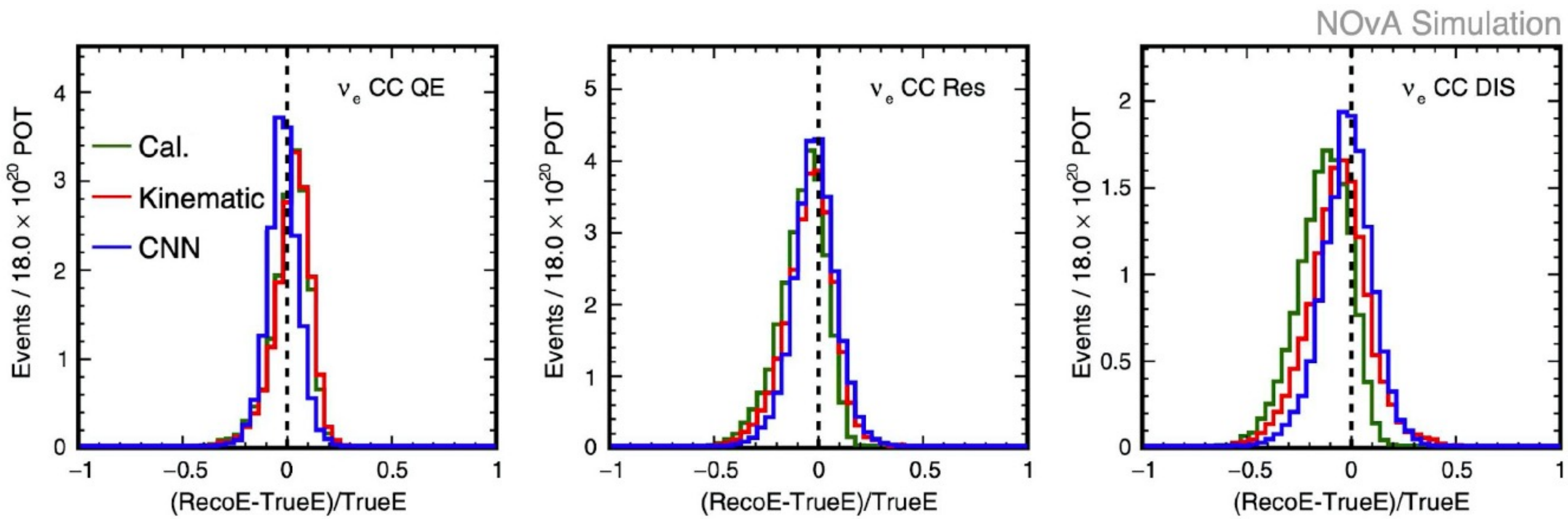


Also trained for electron energy, hadronic energy, ν_μ Energy, etc

Regression CNNs for Energy Estimation

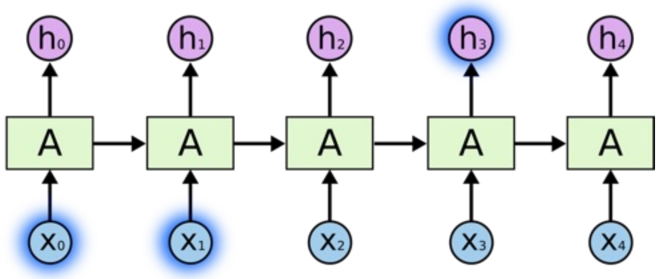
- Regression CNN energy shows good stability over interaction types
- Also shows comparable or less energy dependent bias
- The robustness of a CNN model can provide with a large degree of freedom

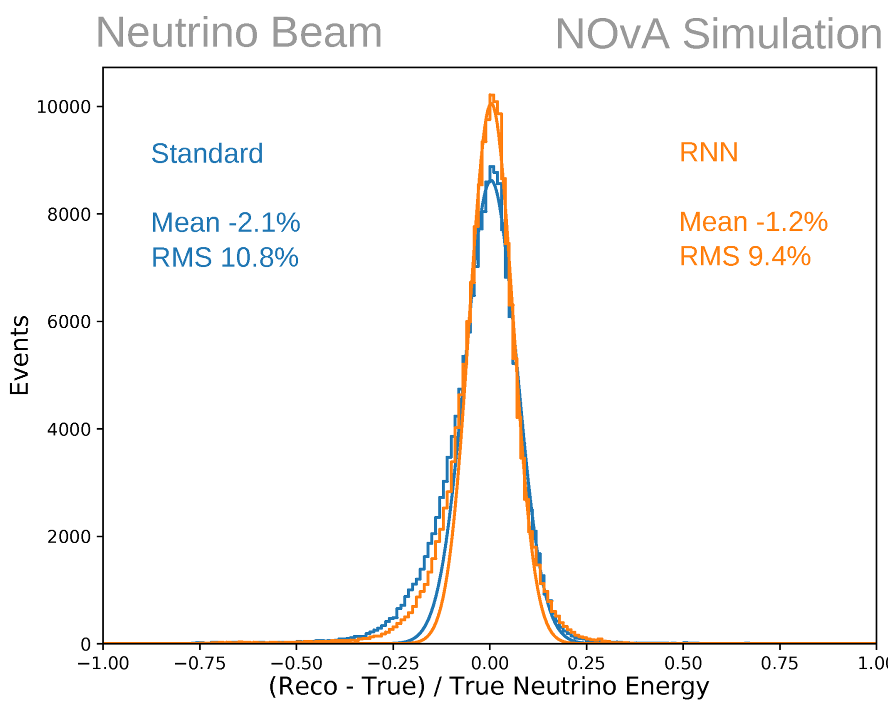




LSTM for Energy Estimation

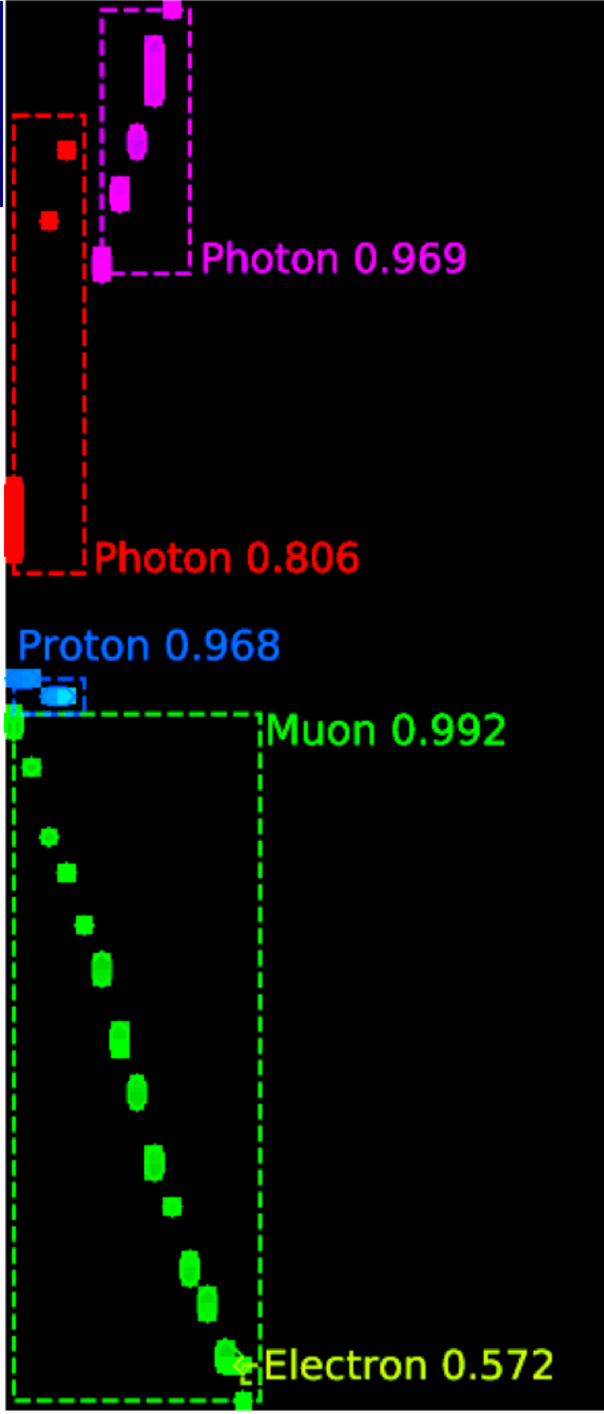
- Long Short-Term Memory (LSTM) is a type of recurrent neural network
- Takes a number of traditional reconstruction quantities as inputs
- Trained using calibration shifts to increase network resilience
- Resolution comparable with regression CNN





Full Event Reconstruction with Image Segmentation

- Full event reconstruct on a hit-by-hit basis using instance segmentation:
  - Bounds: Create a bounding box around each particle with a Region-based CNN (RCNN)
  - ID Score: Use a softmax function to classify the particle contained within each box
  - Clusters: Group together hits, identify hits, then individual hits are combined to form clusters
- Very powerful in PID and clustering efficiency, working on running at scale



Other Efforts Regarding Machine Learning in NOvA

- Sparse and Graphical Neural Networks
- Regression CNN for vertex reconstruction
- ResNet for cosmic ray rejection
- Understanding generator biases in deep learning models, by exploring other generators (NuWro, GIBUU, NEUT, etc)
- Improving traditional reconstruction with ML methods

Data Sample	Traditional Cosmic Rejection	Cosmic Rejection Neural Network
ν_e	93.21	99.71
$\overline{\nu_e}$	92.81	99.82
ν_μ	93.22	99.20
$\overline{\nu_\mu}$	92.82	99.20
ν NC	93.24	97.08
$\bar{\nu}$ NC	92.79	96.82
Cosmic ν	7.80	5.00

Summary

- NOvA is the first HEP experiment to use CNNs to publish physics results: *Phys.Rev.Lett. 118 (2017)*
- In NOvA, deep-learning has been developed to:
  - Identify events and final state particles from beam and cosmic ray backgrounds
  - Reconstruct neutrino energy, final state particle energy, and other kinematic variables
  - Perform full event reconstruction
- NOvA has been performing expansive data comparison, impact analysis, uncertainty studies and cross-checks to improve robustness and interpretability of ML tools

Thank you!