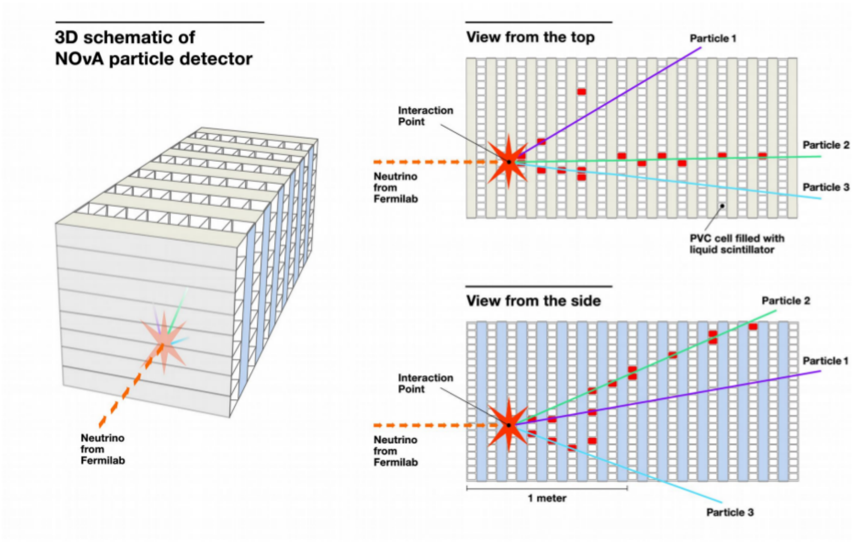
NOvA’s Convolutional Neural Network

Sam Kohn
290E Fall 17
6 December 2017

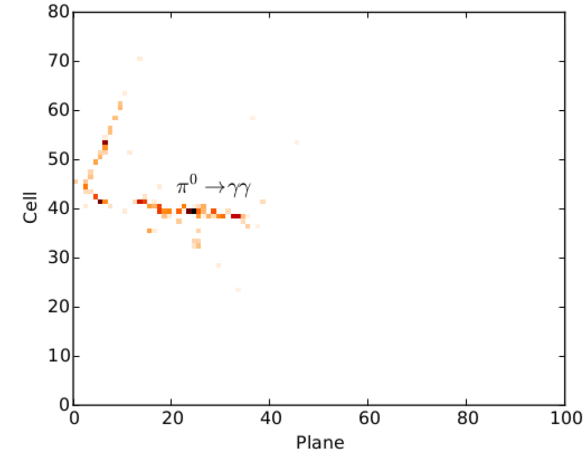
The NOvA Experiment

- Long-baseline accelerator neutrinos
- Near-far comparison between Fermilab and Ash River, MN (810 km away)
- ν_e appearance → θ_{13} and δ_{CP}
- ν_μ disappearance → θ_{23}
- Scintillating cells arranged in 2 views

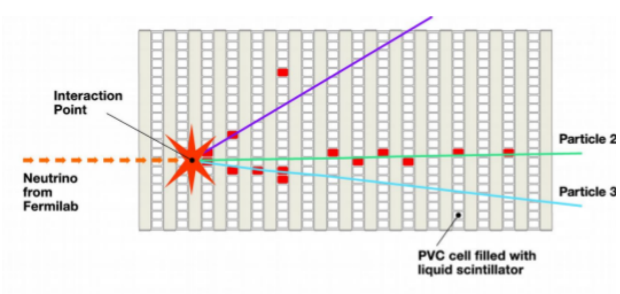


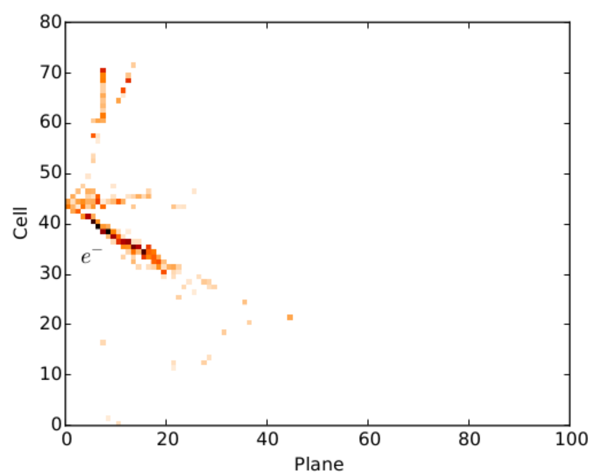
Different neutrino interactions

- Charged current e and μ
  - Resonant
  - Deep inelastic
  - Quasi-elastic
- Neutral current (any flavor)

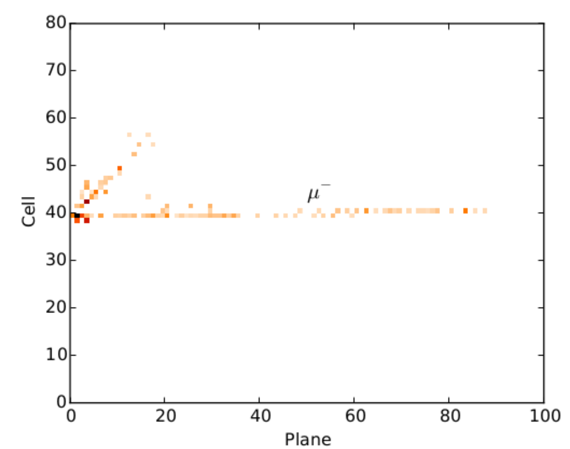


NC (simulation)





ν_e CC (simulation)



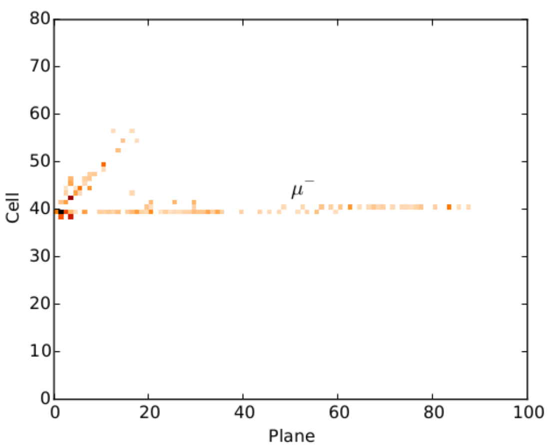
ν_μ CC (simulation)

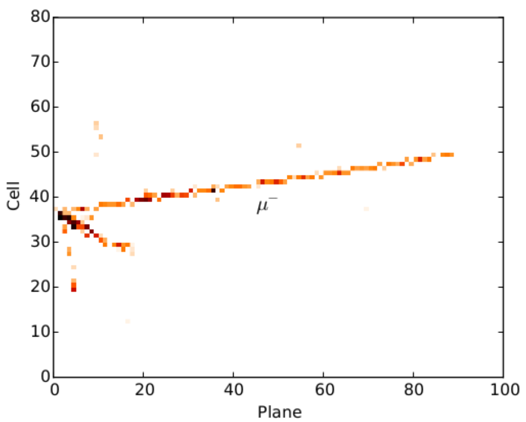
Traditional methods

- LID (Likelihood-based selector)
  - Reconstruct tracks and showers
  - Reconstruct energy
  - Into a standard MLP neural network (multi-layer perceptron aka “plain old” neural network)
- LEM (Library event matching)
  - Generate millions of simulated events
  - Use figure of merit to compare real events to library
- These work OK but there are problems

Obstacles

- LID relies on track and shower reconstruction
  - Uses dE/dx along shower, for example
  - These are 2 views of the same event…how many tracks are there?

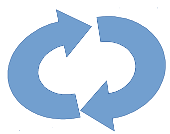




- LEM is expensive
  - 77×10^6 events!
- Is there a better way?

Machine learning!

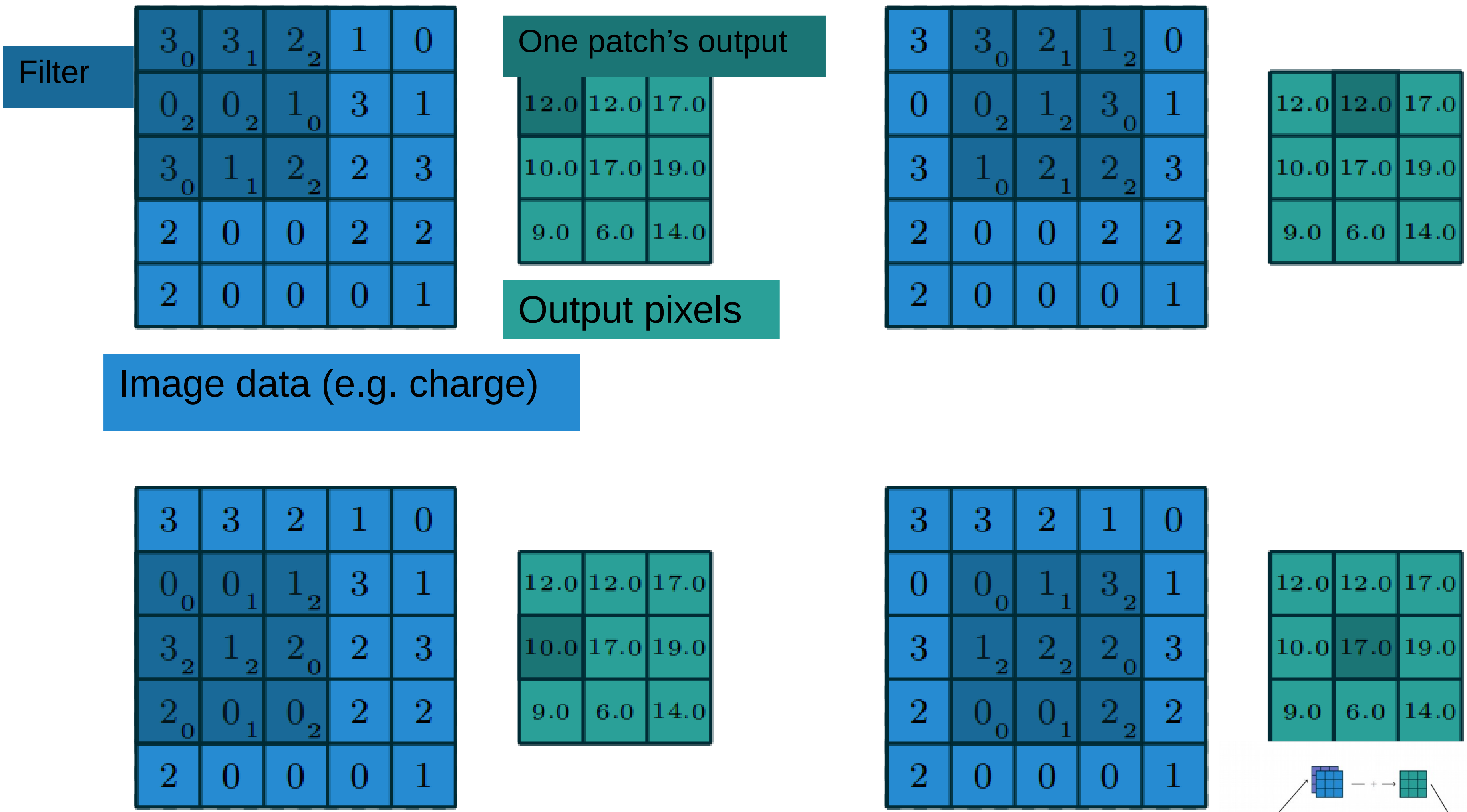
- Yes it's a fad…doesn't hurt to try
- Basic idea
  - Construct a function of many parameters

  - Input: 1 physics event; output: event classification
  - Fit that function to all of your data (aka train it)
- Implementation
  - (Matrix multiplication + non-linearity) x 
  - "Fully connected"
  - There are other architectures too

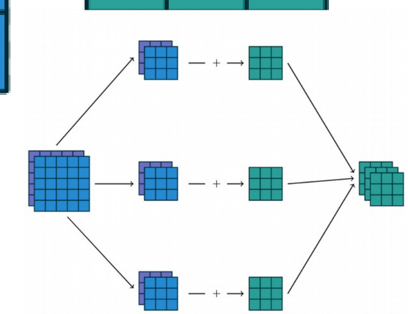
Convolutional networks (CNNs)

- Useful for images or image-like data
- Stack of images = “channels” (e.g. RGB)
- Convolution operation
  - One large matrix → many small “filters”
  - Apply each filter to each patch of input image
  - Each filter generates one channel of output
- Benefits:
  - translational invariance
  - semantically similar to how humans recognize images (and we’re pretty good at it)

Convolution example



This is a simple example because the image has only 1 channel.
A filter has access to all of the channels in an image.



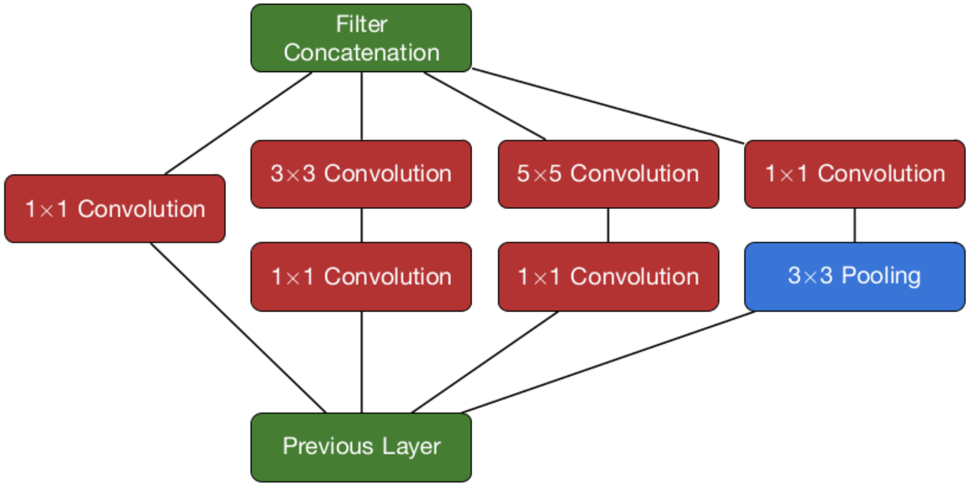
What.

- There’s such a thing as 1x1 convolutions
- Better to think of it as 1x1xN
- Purpose is to compress an image with N channels down to a single channel
- So with $M < N$ different 1x1 convolutional filters you can compress an N-channel image into an M-channel one → saves computational power

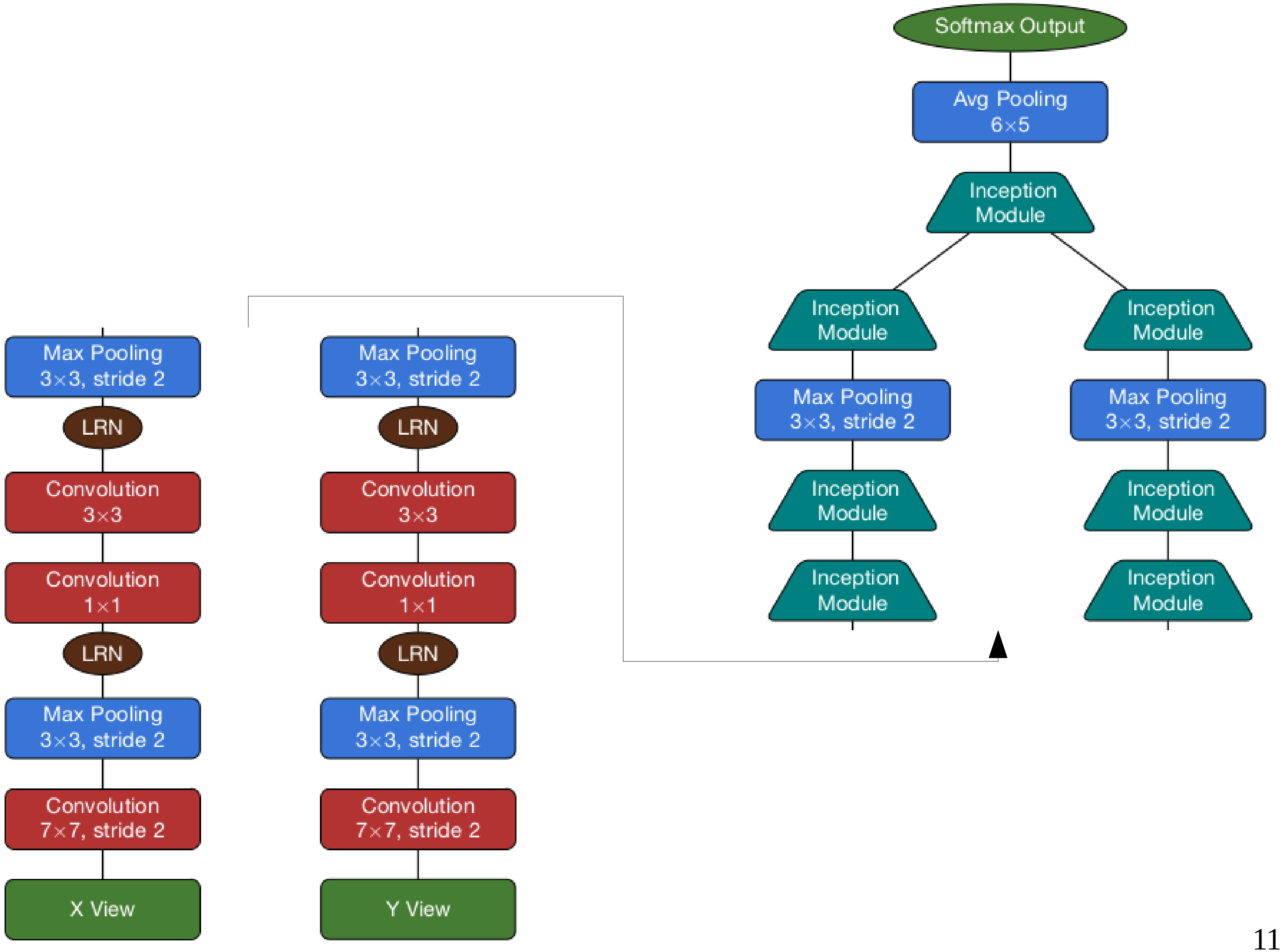
Let’s go deeper…

- Google researchers invented the “inception module”
- A mini-CNN that can be embedded inside another CNN. Without adding much computing time

The look on my face when my family asks “what are neutrinos made out of”



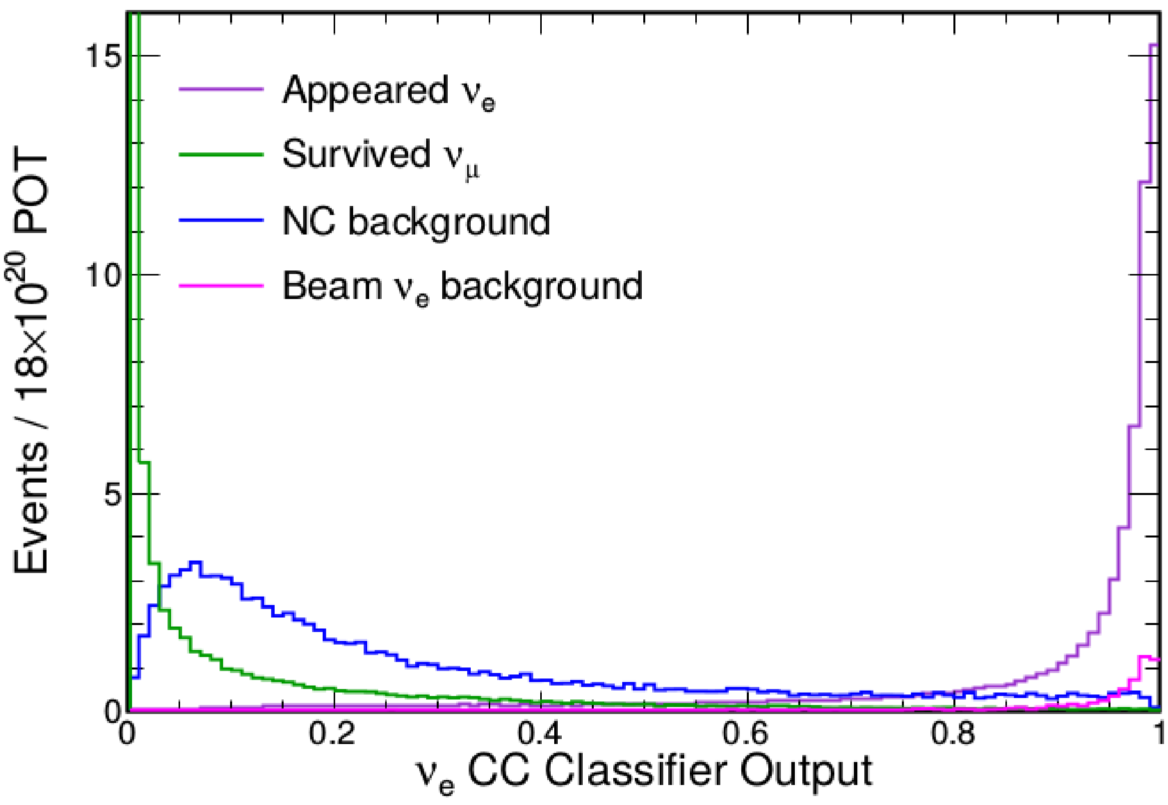
Putting it all together



So does it work?

- First some asides
- Training this CNN took 4.7M simulated events — compare to 77M for the LEM technique
- Extra training events were created by adding noise to and/or mirror-reflecting existing events
- No track or shower reconstruction (like in LID)
- Energy reconstruction is performed on each cell but not on tracks or showers

Yes, it works



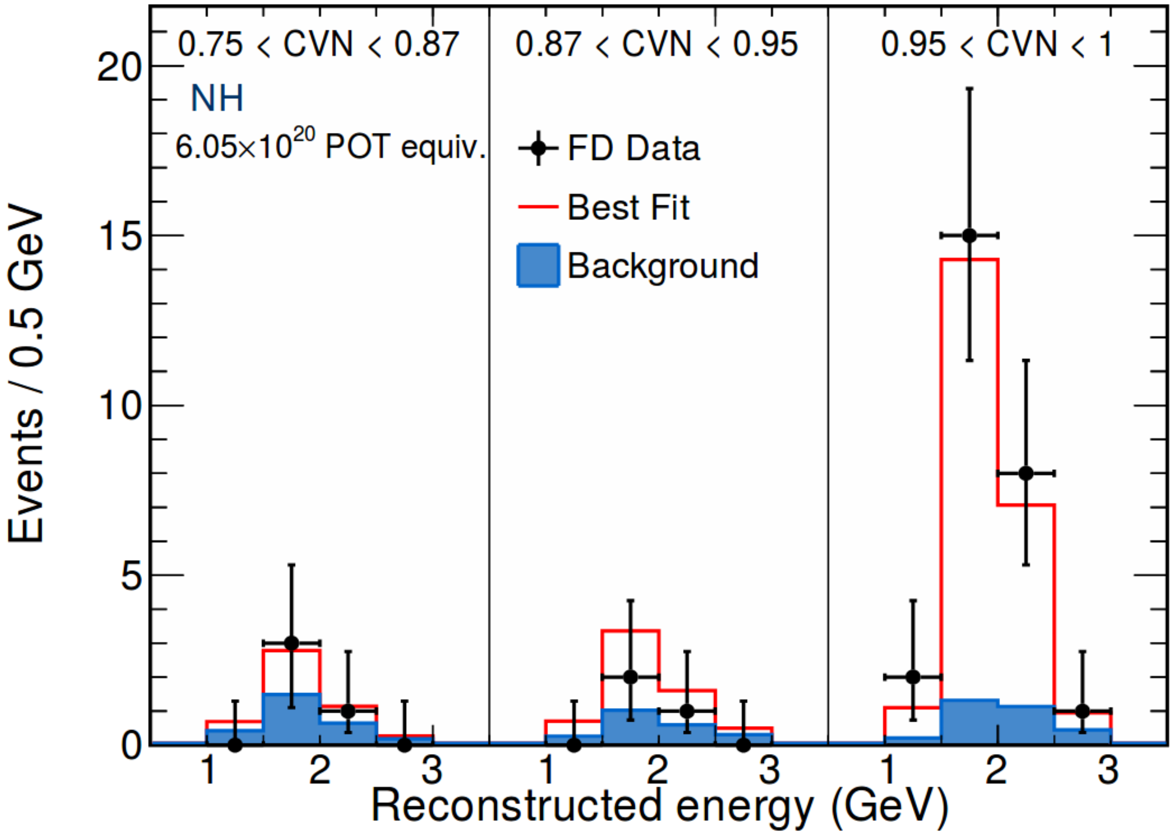
Appeared ν_e
Survived ν_μ
NC background
Beam ν_e background
Events / 18×10^20 POT
ν_e CC Classifier Output
0
5
10
15
0.2
0.4
0.6
0.8
1

Event better than LEM and LID

- Matches the efficiency of these methods for ν_μ at 57% and beats the efficiency for ν_e 49% to 35%
- Now in use for standard NOvA physics results

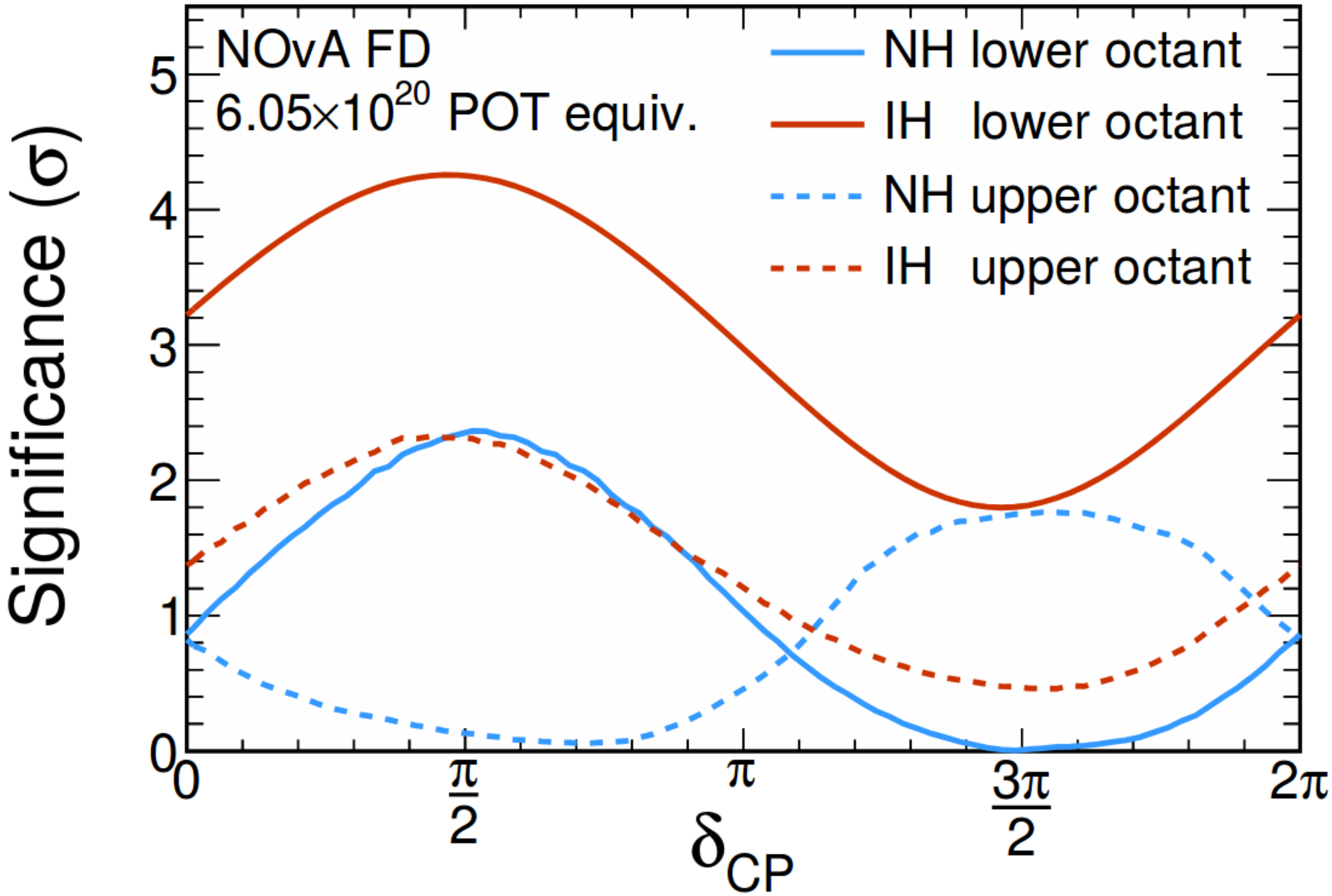
For this analysis a new ν_e CC classifier was developed to select a signal sample with improved purity and efficiency. The Convolutional Visual Network (CVN) [14] is a convolutional neural network and was designed using deep learning techniques from the field of computer vision [15, 16]. Recorded hits in the detectors are formed

ν_e appearance results



The latest parameter plot

- NOvA, May 2017



- Looks like Normal Hierarchy and maximal CP violation are winning…

Any questions?

- Thanks for listening
- Even during reading week
- When you could be studying or sleeping
- Or compiling code

ALL PHYSICS CODE

References

1. A. Aurisano and A. Radovic and D. Rocco et al., JINST 11 (2016) (arXiv: 1604.01444v1)
2. (Convolution images) https://arxiv.org/pdf/1603.07285v1.pdf
3. Phys. Rev. Lett. 118, 231801 (arXiv: 1703.03328)