mvprado@icecube.wisc.edu

Deep Learning Classifier for Low-Energy Events in IceCube

Maria Prado Rodriguez for the IceCube Collaboration
Wisconsin IceCube Particle Astrophysics Center

Abstract

Tau neutrino appearance from neutrino oscillations of atmospheric muon neutrinos is studied by the DeepCore subarray, the densely-instrumented region of IceCube, an ice-Cherenkov neutrino detector 1.5 kilometers below the surface in the South Pole. These studies probe the unitarity of the PMNS matrix. Distinguishable event signatures in this region include track-like and shower-like events. Because the contribution of tau neutrinos manifests as a statistically significant excess of shower-like events, accurate event classification is crucial. However, at the low energies relevant to the oscillation maximum, separation of tracks and showers is challenging. This poster shows an ongoing study of a deep learning event classifier that currently achieves an accuracy comparable to that of previously used methods, with still large room for improvement. We show that DNNs can learn complex features in DeepCore data at hit level (i.e., not relying on reconstructed quantities) that differentiate the signal types.

IceCube Experiment

IceCube is an ice-Cherenkov neutrino detector located at the South Pole. It consists of a lattice of light sensors embedded in Antarctic ice ~1.5 kilometers below the surface. Neutrinos interact with the ice via charged-current and neutral-current interactions and produce detectable charged particles that emit Cherenkov radiation. Our digital optical module (DOM) sensors convert this radiation into a current with photomultiplier tubes and transfer the digitized current up the cable to the lab.

The DeepCore subarray is the densely-instrumented region of IceCube. With a 7-meter vertical DOM separation for the eight strings, DeepCore studies include neutrino events with energies as low as 5 GeV. At these energies, we look at tau neutrino appearance from oscillated atmospheric muon neutrinos, measuring θ_{23} and Δm^2_{32} in order to assess whether the mixing angle is maximal. Figure 1 shows the latest IceCube tau neutrino appearance results in blue. As part of this study, a tau neutrino normalization value is obtained as shown in figure 2. This normalization value probes the unitarity of the PMNS matrix via $|U_{\tau_3}|^2$. It is defined as "the ratio of the measured ν_τ flux to that expected assuming best-fit oscillation and other nuisance parameters for that ν_τ normalization (arXiv:1901.05366v1)."

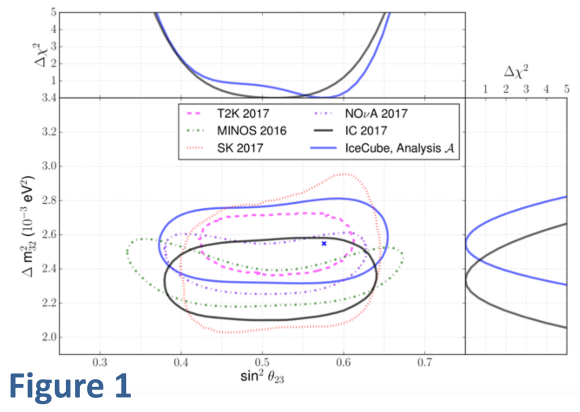


Figure 1

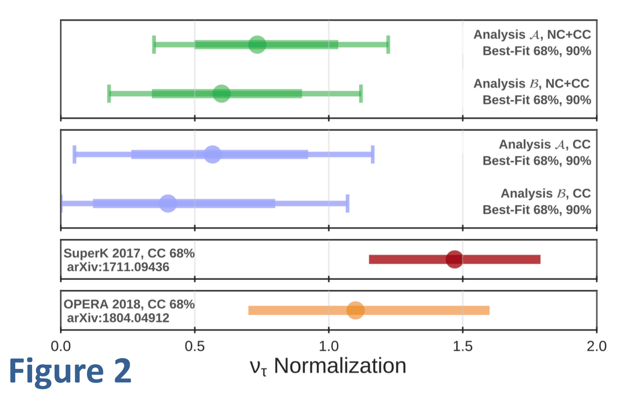
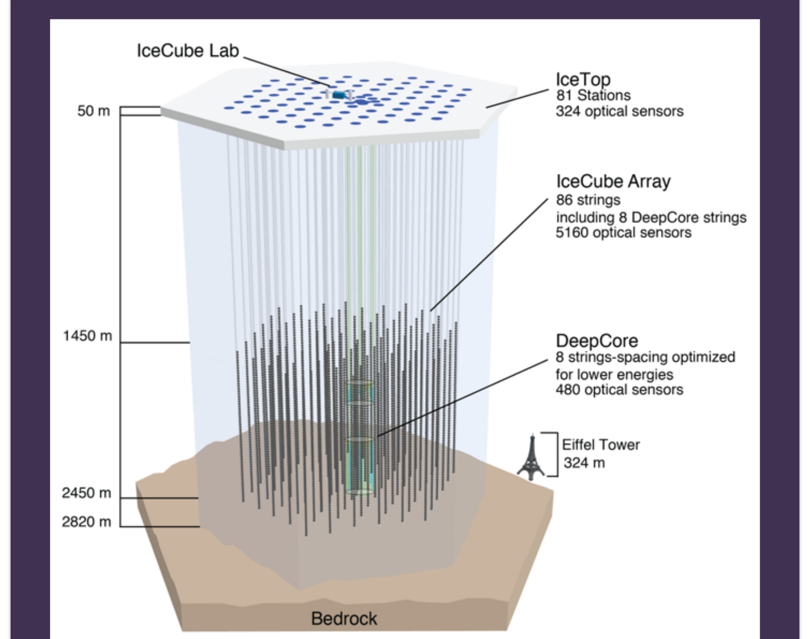


Figure 2



Our Research

Distinguishable neutrino event signatures in DeepCore include *track-like* and *shower-like* events. Because the contribution of tau neutrinos manifests as a statistically significant excess of shower-like events, accurate event classification is crucial. However, at low energies, the small number of hit channels makes particle ID difficult using conventional techniques. This research explores deep learning techniques to develop a DeepCore event classifier.

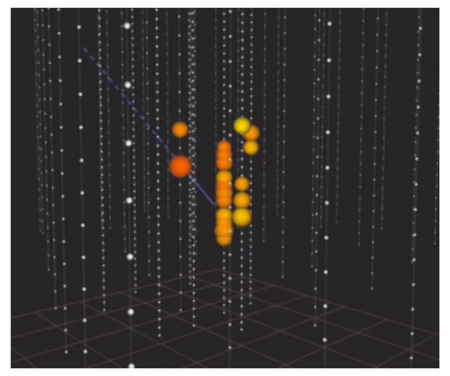
Current Event Classification Method

BDT method

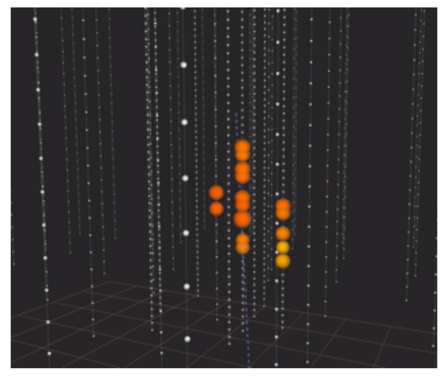
Feeds reconstructed quantities into an XGBoost Boosted Decision Tree (BDT) for learning. The table below shows the reconstructed quantities used as input variables for the BDT and their separation power.

Feature name	Feature importance
reconstructed shower energy	0.081
reconstructed track length	0.657
reconstructed $\cos(\theta_{zenith})$	0.068
reconstructed $\cos(\theta_{zenith})$ uncertainty	0.099
LLH ratio of shower+track reco to shower only reco	0.096

Track in DeepCore with energy of 26 GeV



Shower in DeepCore with energy of 30 GeV

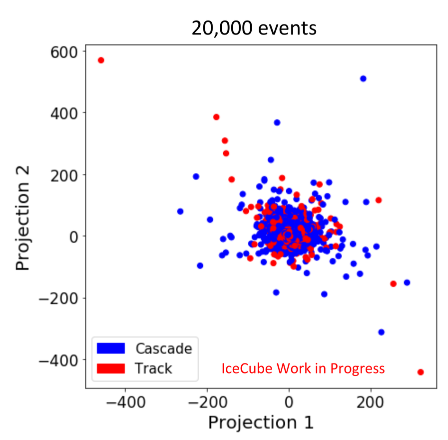


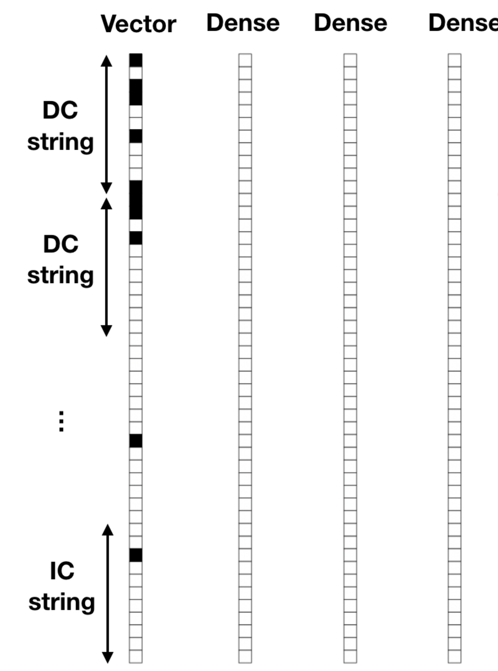
Tracks: muon neutrinos and muon anti-neutrinos from charged-current interactions as well as 17% of tau neutrino charged-current interactions which decay into muons
Showers: all other neutrino interactions

Neural Network Methods

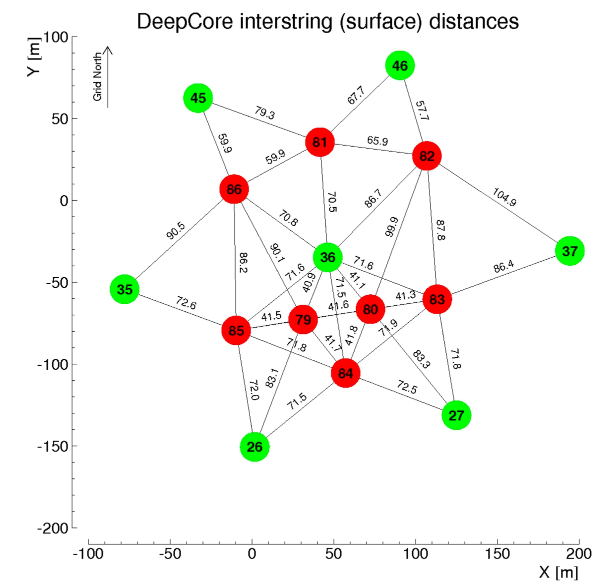
1D Classifier:

- 1D vector input concatenating strings
- **532 input vector**
- ~84,000 free trainable parameters
- 2D projections of our data show that data is not inherently indistinguishable





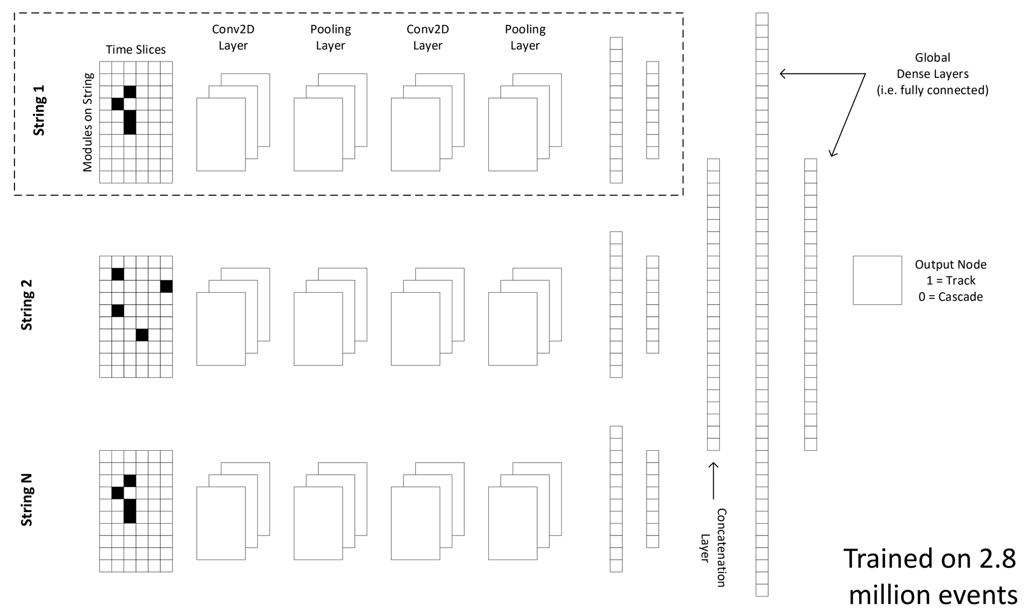
Trained on 2.6 million events



As seen in the figure above, DeepCore is anything but spatially symmetric. This poses problems when trying to use any kind of conventional CNN.

2D Classifier:

- 15 separate inputs each representing one string
- All DeepCore strings share weights, all IceCube strings share weights
- String images are:
  DC: (z, t) = (50, 25)
  IC: (z, t) = (21, 25)
- ~200,000 free trainable parameters



Trained on 2.8 million events

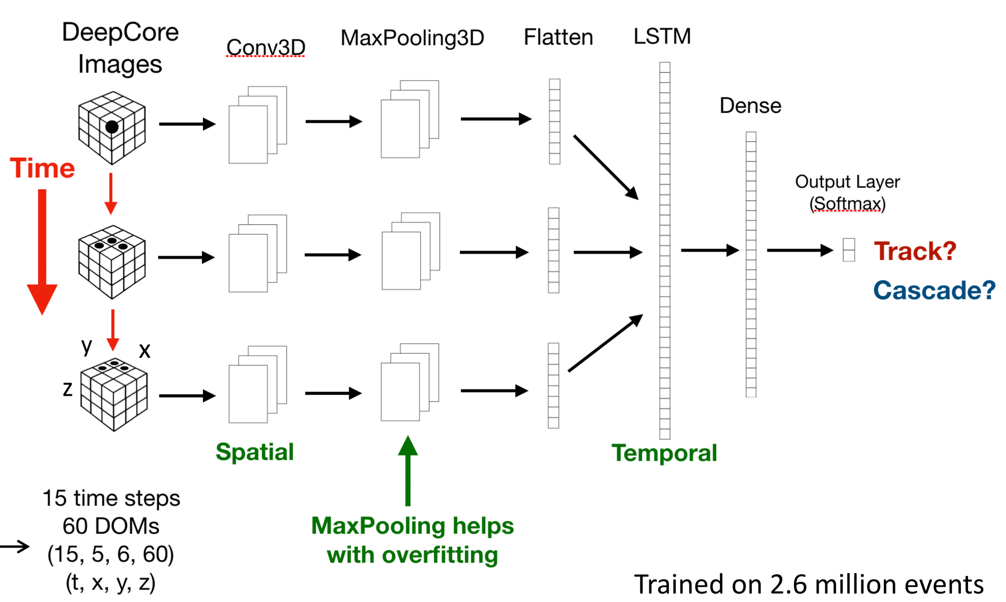
4D Classifier:

- Grid padded with zeros in locations without strings for x-y symmetry as well as for mimicking spatial distances between strings
- Preserves symmetry features, but with inaccurate x-y-z spacing
- ~12 million free trainable parameters

DeepCore strings
IceCube strings

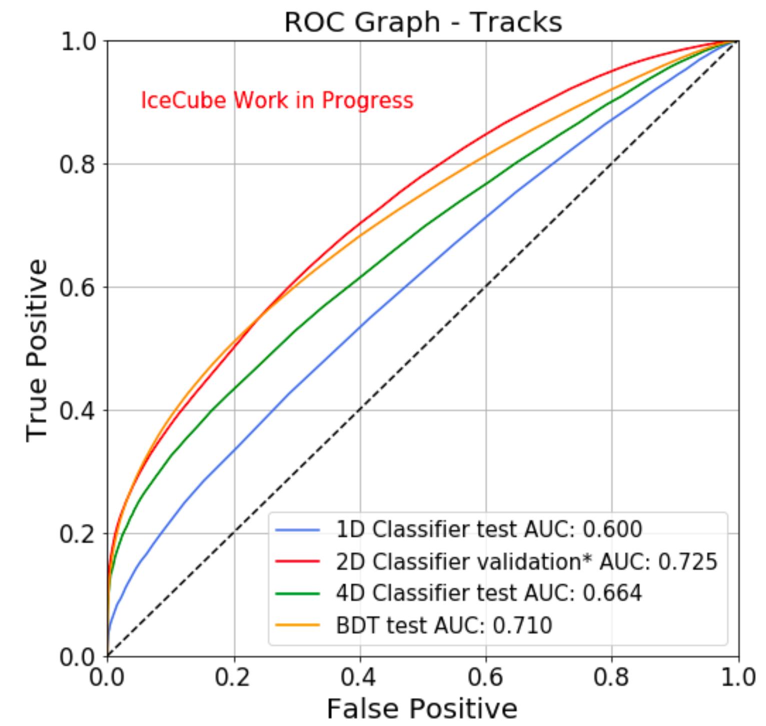
0	45	0	46	0
0	86	81	82	0
35	0	36	0	37
85	79	80	83	0
0	0	84	0	0
0	26	0	27	0

Images TopView



Trained on 2.6 million events

Performance



*The 2D classifier curve was formed using a different simulation of DeepCore neutrino events to the other two classifiers. It also does not include any atmospheric neutrino event weights while the other three curves do. However, its performance is not expected to change very much once these changes are made.

Observations

The Receiver Operating Characteristic (ROC) plot above shows how many tracks were classified correctly (true positive) along with how many showers were classified incorrectly (false positive). Perfect classification would lie on the top left-hand corner. Performance of the four classification methods is measured using the area under the curve (AUC): the greater the area under the curve, the better the classification.

The 2D classifier already shows comparable performance to the current BDT method. Seemingly, convolutions over a time and z dimension appear to extract some separation power in the data. Given the irregularities in symmetry of DeepCore, perhaps this is due to the comparatively low number of approximations in both the time and z dimensions. Moving forward, the 2D classifier could potentially improve its performance by stacking all the string images as a "channels" dimension. In a "channels" dimension, ordering of strings does not matter and, still using a 2D convolution, any cross-channel patterns could be recognized.

Towards the IceCube Upgrade

- Seven more strings will be added with even shorter DOM spacing
- Will help us reach lower energy events (1-10 GeV) as well as improve the current analysis

Figure 3

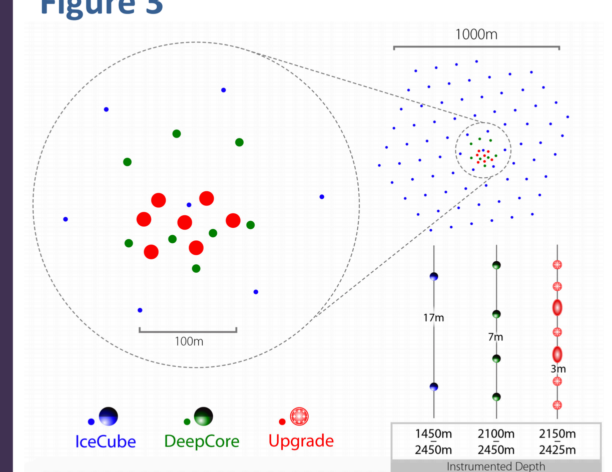
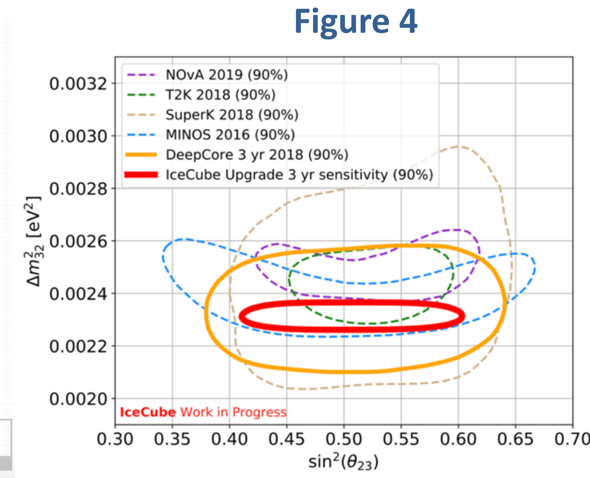


Figure 4



References

- [1] & [2] IceCube Collaboration, Phys. Rev. D 99, 032007 (2019), arXiv:1901.05366
- [3] & [4] IceCube Collaboration, (2019), arXiv:1908.09441v1