Application of Convolutional Neural Networks to Reconstruct GeV-Scale IceCube Neutrino Events

Jessie Micallef for the IceCube Collaboration
Michigan State University

NEUTRINO 2020

IceCube Neutrino Observatory

The IceCube Neutrino Observatory, instrumenting a cubic kilometer of Antarctic ice, searches for astrophysical and atmospheric neutrinos. It comprises of 5160 digital optical modules (DOMs) that each contain a photomultiplier tube, arranged on strings in a hexagonal array (Figure 1).

The more densely instrumented center, DeepCore, is optimized to measure the lower energy neutrinos in the 10-GeV scale [1]. These energies are important for

- Oscillation parameter measurements
- Non-standard interactions
- Sterile neutrino searches.

Improving reconstruction speed and resolution for these events are important for future IceCube studies.

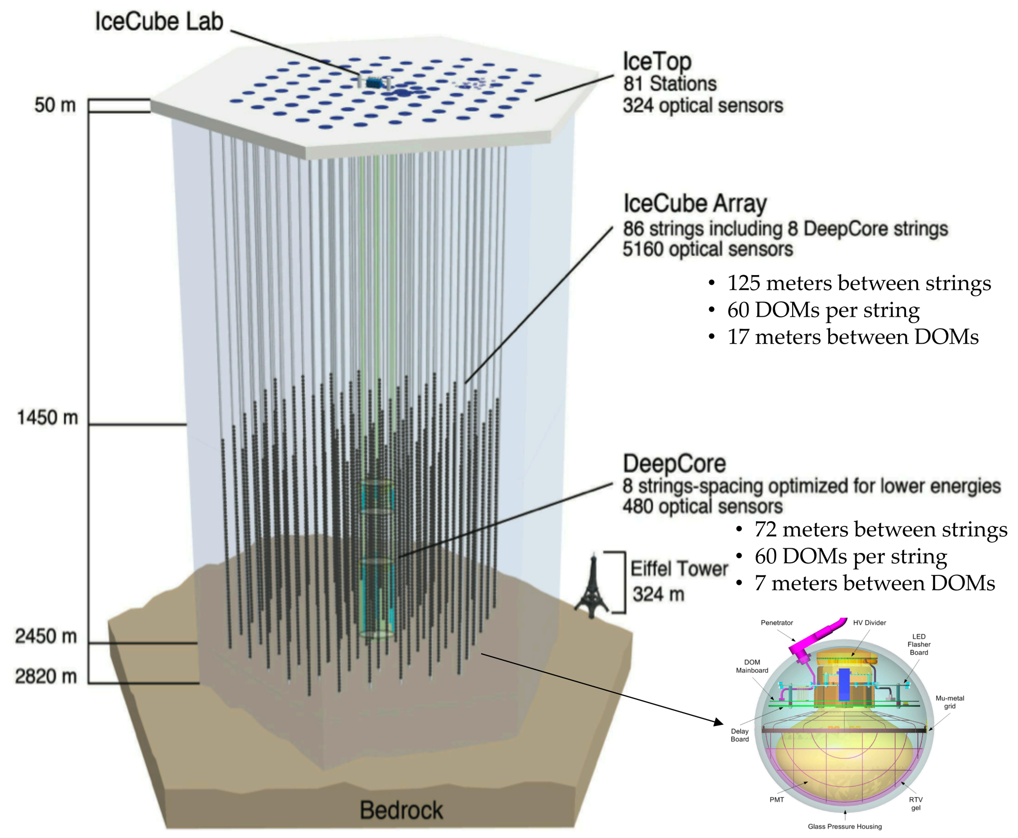

Figure 1: Side view of IceCube detector with a close-up of a digital optical module (DOM)

Goal: Apply a CNN to improve the speed and resolution of reconstructing energy and direction for 10s of GeV-scale muon neutrino events in IceCube.

Execution Speed

The CNN predicts the reconstructed variables in 1.5 - 0.5 ms on average per event. The previously used, likelihood-based method takes seconds to minutes per event. Thus, the CNN could improve the runtime by 6 orders of magnitude!

Energy Resolution Comparison

The CNN's energy resolution is slightly more precise at low energies with a narrower spread, but the bias saturates at energies > 70 GeV compared to the previous likelihood-based reconstruction [2].

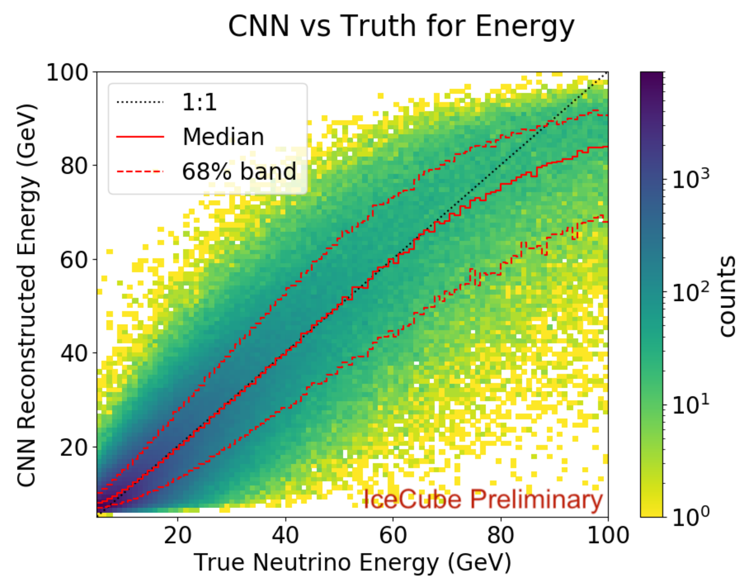

Figure 5. CNN reconstructed energy performance

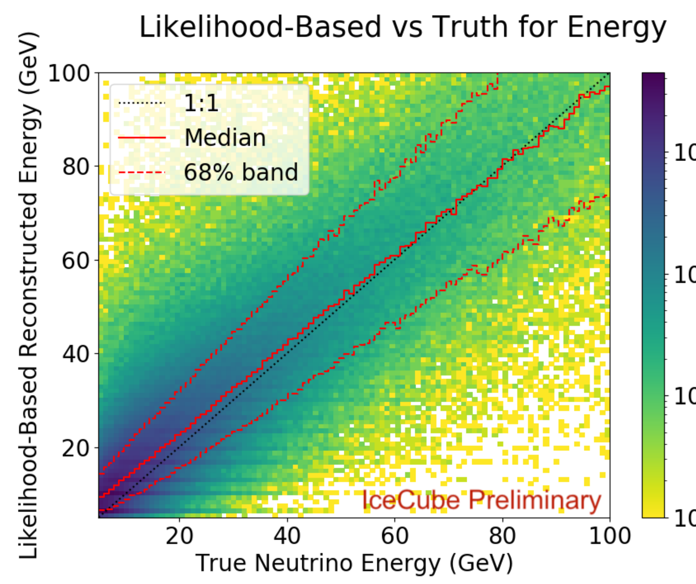

Figure 6. Likelihood-based reconstructed energy performance

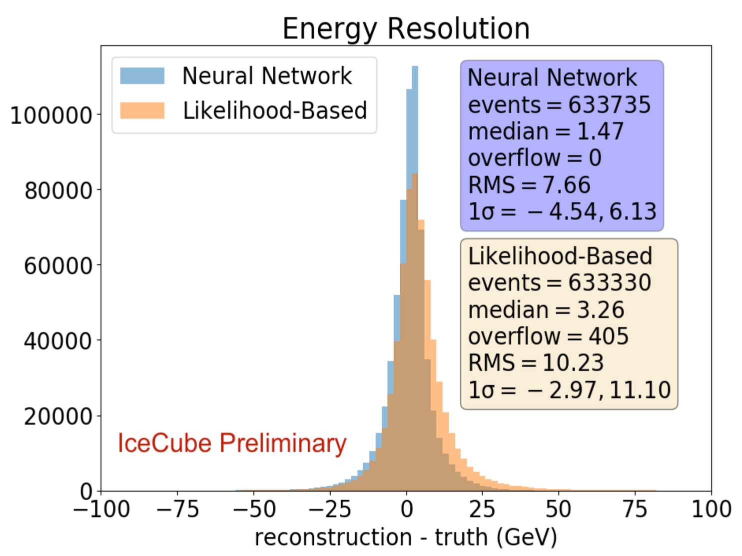

Figure 7. Absolute resolution taking the difference between the CNN and previous method to truth

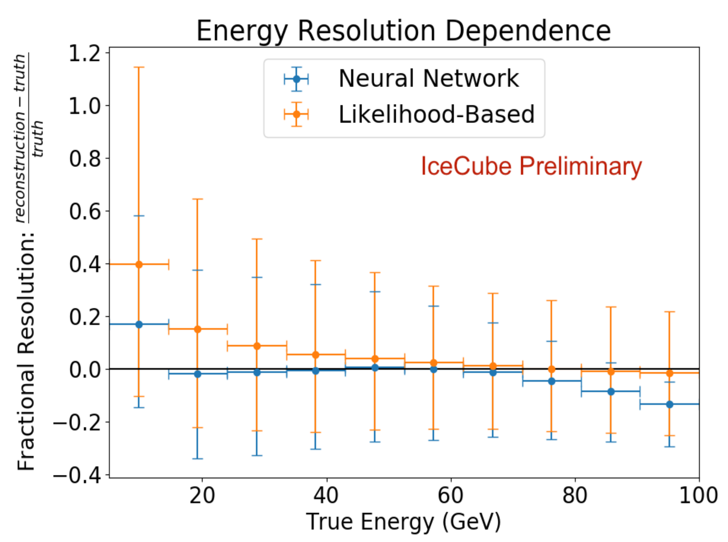

Figure 8. Fractional resolution evaluated for sections of energy, with the datapoint showing the median resolution and error bar showing 68% containment

Applying CNN to Focus on IceCube's DeepCore

The Convolutional Neural Network (CNN) relies on nearby neighbors to affect the training weight of the next layer. Often used in image recognition, nearby pixels are important to recognize edges and shapes.

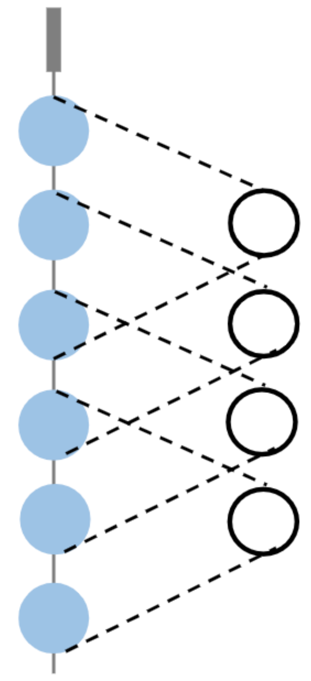

Figure 2 CNN used in z-direction over DOMs

The CNN is applied only in z-depth where nearby DOMs on all strings are part of a convolutional kernel, or window, that affects to weights of the next training layer (Figure 2). Since DeepCore is optimized for low energy events [1], only the centermost strings on IceCube are used for the low energy CNN, circled in blue in Figure 3.

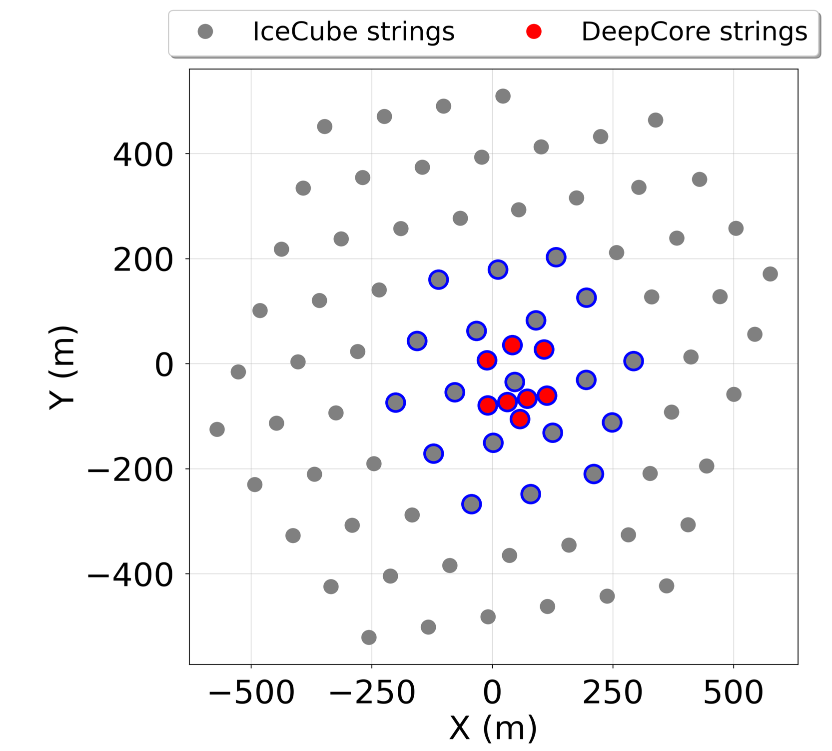

Figure 3: Top view of the IceCube detector. Blue circles indicate strings that are used for CNN

Low Energy CNN Architecture

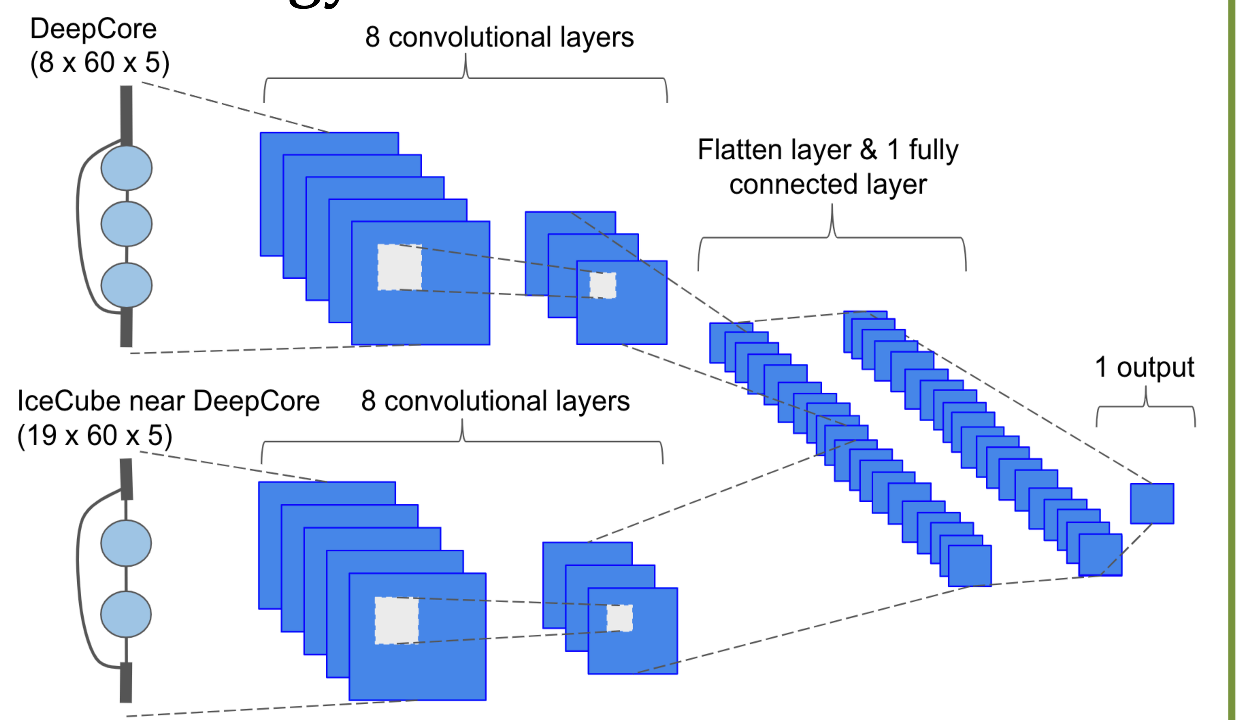

Figure 4: CNN architecture used for low energy IceCube events

Notable CNN training features:

- Split into two parallel branches
  - Separates the DeepCore and IceCube strings which have different z-depth DOM spacing
- Trains for one variable only
  - The loss is optimized for energy and cosine zenith (direction) individually
- Training sample is energy and direction independent
  - Current results shown here are for a charge current muon neutrino sample

Direction Resolution Comparison

The CNN's direction resolution is very comparable to the previous likelihood-based reconstruction [2], with slightly narrower spread at the lowest energies and broader spread at energies > 70 GeV.

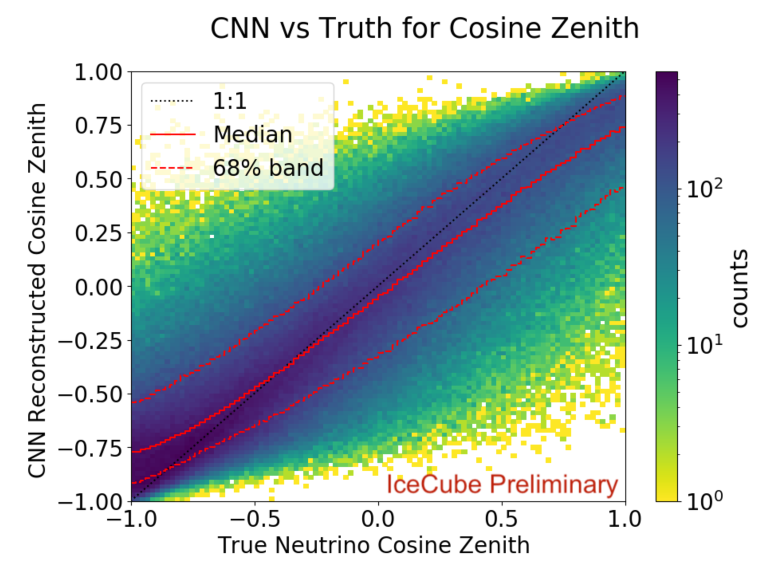

Figure 9. CNN reconstructed cosine zenith performance

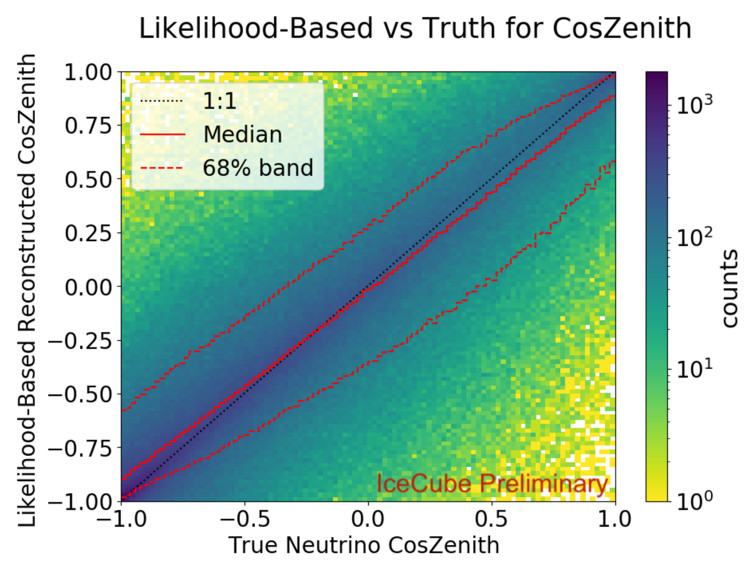

Figure 10. Previous likelihood-based reconstructed cosine zenith performance

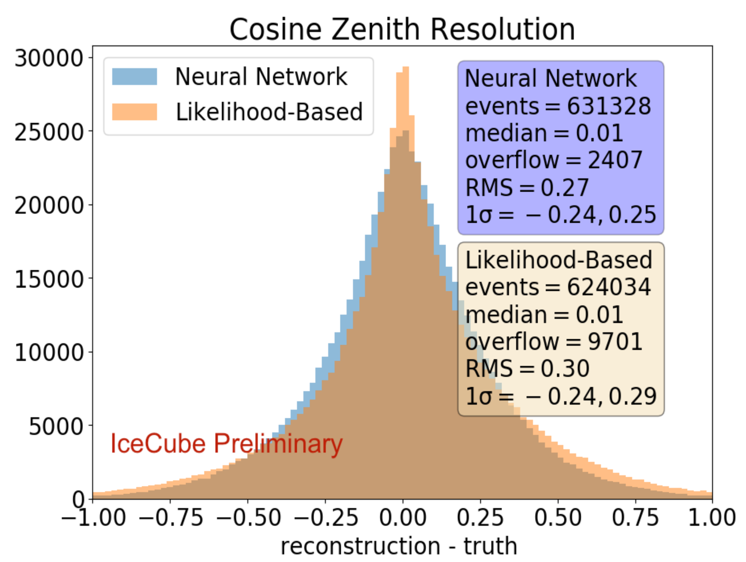

Figure 11. Absolute resolution taking the difference between the CNN and previous method to truth

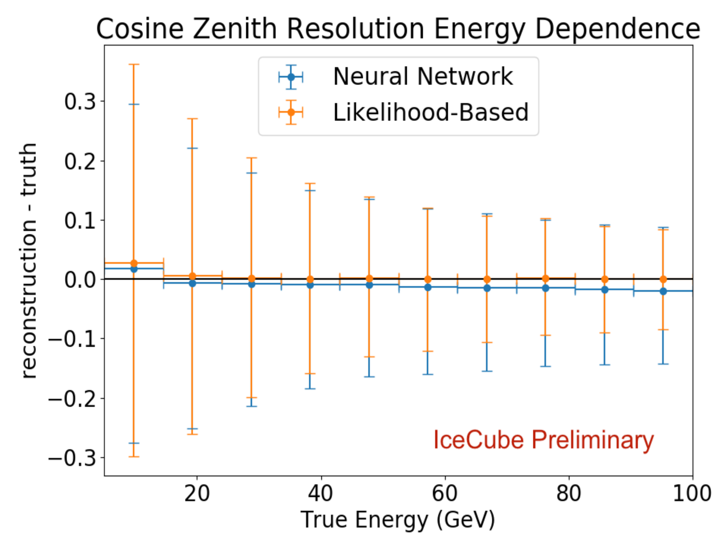

Figure 12. Fractional resolution evaluated for sections of energy, with the datapoint showing the median resolution and error bar showing 68% containment

Advantages of Low Energy CNN

- ✓ CNN resolution for energy and direction are comparable to likelihood-based method
- ✓ CNN can reconstruct 1,000,000 times faster
- ✓ Training is done at earlier stage of data processing pipeline than where other reconstructions are applied

Future Work:

- Improve resolution of CNN further
- Extend sample to include other neutrino flavors or energies
- Explore robustness against systematics

References

[1] M. Aartsen *et al.* *JINST* Vol. 12. 2017. DOI 10.1088/1748-0221/12/03/P03012
[2] M. Leuermann. PhD RWTH Aachen University. 2018. DOI 10.18154/RWTH-2018-231554

Acknowledgments

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE1848739. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.