SISSA

PAPER

Using convolutional neural networks to reconstruct energy of GeV scale IceCube neutrinos

View the article online for updates and enhancements.

You may also like

Using convolutional neural networks to reconstruct energy of GeV scale IceCube neutrinos

J. Micallef[*] **on behalf of the IceCube collaboration**[1]

Michigan State University,
567 Wilson Rd, Lansing, MI 48824, U.S.A.

E-mail: micall12@msu.edu

ABSTRACT: The IceCube Neutrino Observatory, located under 1.4 km of Antarctic ice, instruments a cubic kilometer of ice with 5,160 optical modules that detect Cherenkov radiation originating from neutrino interactions. The more densely instrumented center, DeepCore, aims to detect atmospheric neutrinos at 10 GeV scales to improve important measurements of fundamental neutrino properties such as the oscillation parameters and to search for non-standard interactions. Sensitivity to oscillation parameters, dependent on the distance traveled over the neutrino energy (L/E), is limited in IceCube by the resolution of the arrival angle (which determines L) and energy (E). Event reconstruction improvements can therefore directly lead to advancements in oscillation results. This work uses a Convolutional Neural Network (CNN) to reconstruct the energy of 10 GeV scale neutrino events in IceCube, providing results with competitive resolutions and faster runtimes than previous likelihood-based methods.

KEYWORDS: Analysis and statistical methods; Cherenkov detectors; Neutrino detectors

---

[*]Corresponding author.

Contents

1 Introduction

The IceCube Neutrino Observatory, located at the South Pole, detects atmospheric neutrinos that have interacted in the ice and bedrock. IceCube uses 5160 Digital Optical Modules (DOMs), attached to cables called strings that are arranged in a hexagonal array, to detect the Cherenkov light from neutrino interactions [1]. The center-most DOMs are more densely arranged, in an array called DeepCore, to extend the detection energy down to GeV scales [2]. These 10-GeV scale neutrinos are at energies higher than accelerator and reactor experiments can reach, giving a unique look at the neutrino oscillation region at these energies. Likelihood-based reconstruction methods are currently used to reconstruct the energy (E) and arrival angle (determines L), which are the necessary variables to constrain the oscillations parameters. An alternate solution is to apply machine learning methods to improve the resolution and runtime of the energy reconstruction.

2 Convolutional Neural Network architecture and training

Neutrinos can be detected anywhere within the IceCube detector volume, so that events are spatially invariant, which is a symmetry that convolution neural networks (CNNs) excel at identifying [3, 4]. IceCube has successfully implemented a CNN for high-energy events (> 100 GeV) [5, 6], but a CNN must be specifically optimized to handle low-energy neutrinos because the sparse DOM layout and the limited scattered light at these energies provide less input information for reconstruction.

The low-energy CNN only uses the 8 DeepCore strings and nearby 19 IceCube strings, focusing the CNN on the low-energy optimized region of the detector (figure 1). All hits per DOM are summarized into 5 charge and timing variables, which are separately given to the CNN for the DeepCore and IceCube strings because they have different z-spacing between DOMs, creating a two-branched network architecture. The CNN then uses its convolutional kernel to span the vertical, or z-depth, of the string; no xy-plane convolution is applied since the DeepCore strings are in an irregular geometry. After 8 convolutional layers, each followed by a dropout and batch normalization layer (see ref. [7] for more detail), the two CNN branches are concatenated to train on one fully connected layer, giving a single output: the reconstructed energy.
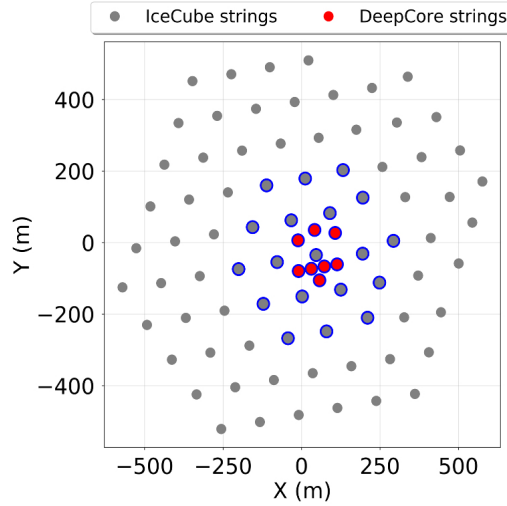
Figure 1. Top view of the IceCube detector, with DeepCore strings in red, IceCube strings in gray, and a blue highlight on all strings used for the low-energy CNN.

To train the CNN, a sample of more than 6 million unweighted, charge-current (CC) ν_μ Monte Carlo (MC) events was generated in GENIE [8], a neutrino MC generator with an emphasis on few-GeV neutrino interaction physics. Since ν_e events look like ν_μ events with high inelasticity, the generation for the training sample focused only on ν_μ events. The training sample has a flat energy distribution from 1-200 GeV, in the target region for an oscillation analysis, and a nearly flat energy distribution from 200–500 GeV. CNNs are known for their difficulty extrapolating beyond their training samples [9], so the energy is extended to 500 GeV to avoid issues at the sample boundary. To test the CNN, an independent GENIE sample was generated with ∼ 4 million unweighted ν_μ CC events and ∼ 1.6 million unweighted ν_e CC events. These samples are weighted by both the atmospheric flux and standard oscillation parameters, so that the CNN is tested on data-like distributions.

3 CNN performance and runtime

Figure 2 shows the CNN's energy prediction compared to the truth for the ν_μ CC (left) and ν_e CC (right) testing samples. For the ν_μ CC events, the CNN does well at low energies, with it's red, solid median line staying near the 1:1 perfect diagonal until about 150 GeV. At energies greater than 150 GeV, the CNN begins to slightly underestimate the true energy on average. At high energies, the performance is degraded by muon tracks that leave the volume of DOMs used for the CNN. Training the CNN with more high energy events generated above the target region (200–500 GeV) has improved resolution in past studies. Other options that are being explored for future improvements are expanding the CNN to include more IceCube DOMs or removing events that leave the detector volume from the final data set.

For the ν_e CC events, the CNN has a slight overestimation of the energy after 100 GeV. Since the CNN is trained on ν_μ CC events only, it only sees the hadronic component of the cascade, which has less light per unit energy than electromagnetic cascades. This could account for the CNN's underestimation of the high energy ν_e events.

Figure 3 shows a direct comparison of the energy-dependent fractional resolution for the CNN and the current likelihood-based reconstruction used for low-energy IceCube events. The CNN shows the most improvement at low energies ($< 25\,\text{GeV}$) for both the ν_μ CC (left) and ν_e CC (right) testing samples, where the majority of the atmospheric neutrino data is expected. The CNN also has comparable resolution with the likelihood-based method up to $200\,\text{GeV}$ for both testing samples.
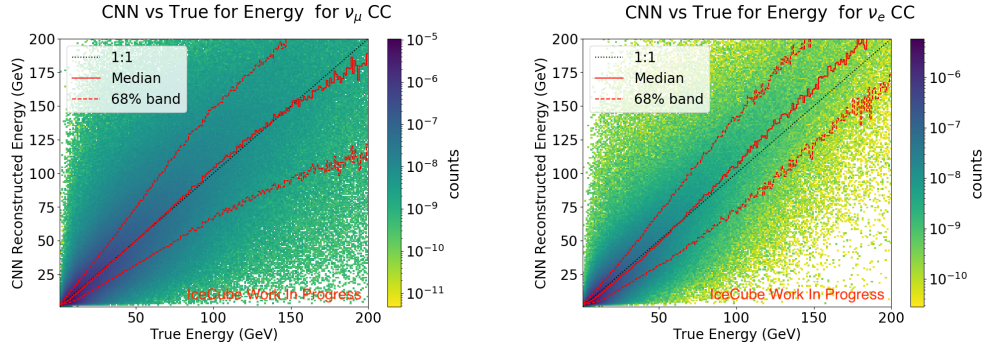

Figure 2. CNN predicted energy vs. true energy for ν_μ CC events (left) and ν_e CC events (right).
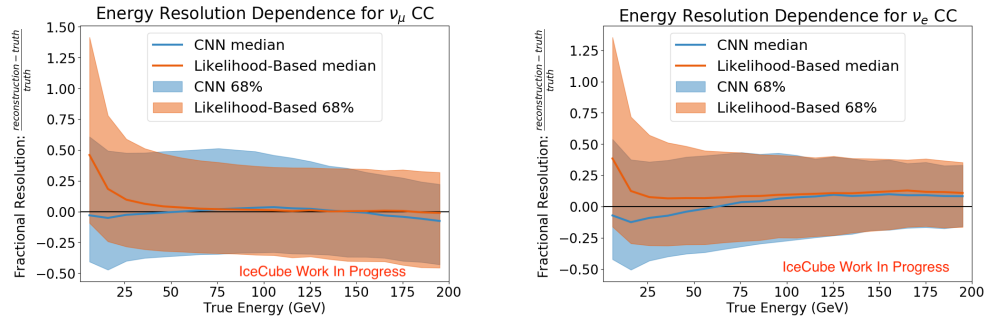

Figure 3. Fractional energy resolution as a function of true neutrino energy, for ν_μ (left) and ν_e (right).

The CNN and likelihood-based reconstructions' runtimes are in table 1. Both the CNN's GPU and CPU runtimes are faster per event than the likelihood-based reconstruction (CPU capability only). It should be noted that the likelihood-based method reconstructs 8 variables simultaneously, but even if the CNN is run in serial to reconstruct 8 variables, its runtime would increase by less than a factor of 10. Speed is particularly important for IceCube's high statistics atmospheric data sample, which the CNN can reconstruct on a computer cluster in minutes.

Table 1. Runtimes for the CNN (using GPU and CPU) and likelihood-based method (using CPU only).

Method	Average Time per Event (seconds)	Events per Day per Single Core
CNN on GPU	0.0077	11,000,000
CNN on CPU	0.27	320,000
Likelihood-Based	40	2,100

4 Conclusion

The low-energy CNN has comparable energy resolution to current likelihood-based reconstruction methods and a speedup of 10^4 in runtime when applied to 10 GeV scale IceCube events. The success of the low-energy CNN architecture extends beyond energy reconstruction, with ongoing explorations of applications spanning other variable reconstructions and classifications, such as neutrino direction [10], which are necessary for neutrino oscillation analyses.

Acknowledgments

References

[1] IceCube collaboration, *The IceCube neutrino observatory: instrumentation and online systems*, 2017 *JINST* **12** P03012 [arXiv:1612.05093].

[2] IceCube collaboration, *The design and performance of IceCube DeepCore*, *Astropart. Phys.* **35** (2012) 615 [arXiv:1109.6096].

[3] A. Krizhevsky, I. Sutskever and G.E. Hinton, *ImageNet classification with deep convolutional neural networks*, *Commun. ACM* **60** (2017) 84.

[4] E. Kauderer-Abrams, *Quantifying translation-invariance in convolutional neural networks*, arXiv:1801.01450.

[5] R. Abbasi et al., *A Convolutional Neural Network based cascade reconstruction for the IceCube neutrino observatory*, 2021 *JINST* **16** P07041.

[6] IceCube collaboration, *Deep learning in physics exemplified by the reconstruction of muon-neutrino events in IceCube*, *PoS* **ICRC2017** (2018) 1057.

[7] IceCube collaboration, *Reconstructing neutrino energy using CNNs for GeV scale IceCube events*, *PoS* **ICRC2021** (2021) 1053 [arXiv:2107.11446].

[8] C. Andreopoulos et al., *The GENIE neutrino Monte Carlo generator*, *Nucl. Instrum. Meth. A* **614** (2010) 87 [arXiv:0905.2517].

[9] D.C. Psichogios and L.H. Ungar, *A hybrid neural network-first principles approach to process modeling*, *AIChE J.* **38** (1992) 1499.

[10] IceCube collaboration, *Direction reconstruction using a CNN for GeV-scale neutrinos in IceCube*, *PoS* **ICRC2021** (2021) 1054 [arXiv:2107.02122].