PAPER

Direction reconstruction using a CNN for GeV-scale neutrinos in IceCube

Published by IOP Publishing for Sissa Medialab

Received: *July 23, 2021*
Accepted: *September 13, 2021*
Published: *November 8, 2021*

Very Large Volume Neutrino Telescope 2021
May 18–21, 2021
Valencia, Spain

Direction reconstruction using a CNN for GeV-scale neutrinos in IceCube

S. Yu on behalf of the IceCube collaboration[1]

Department of Physics and Astronomy, Michigan State University, 426 Auditorium Road, East Lansing, U.S.A.

E-mail: yushiqi2@msu.edu

Abstract: The IceCube Neutrino Observatory is designed to observe neutrinos interacting deep within the South Pole ice sheet. It consists of 5160 digital optical modules, which are arrayed over a cubic kilometer from 1450 m to 2450 m depth. At the lower center of the array is the DeepCore subdetector. It has a denser configuration which lowers the observable energy threshold to about 10 GeV and creates the opportunity to study neutrino oscillations with low energy atmospheric neutrinos. A precise reconstruction of neutrino direction is critical in the measurements of oscillation parameters. In this contribution, I will discuss a method to reconstruct the zenith angle of 10-GeV scale events in IceCube using a convolutional neural network and compare the result to that of the current likelihood-based reconstruction algorithm.

Keywords: Analysis and statistical methods; Data processing methods; Pattern recognition, cluster finding, calibration and fitting methods

[1]Full author list and acknowledgments are available at icecube.wisc.edu.

https://doi.org/10.1088/1748-0221/16/11/C11001

Contents

1 Introduction

The IceCube Neutrino Observatory is a Cherenkov detector located deep under the Antarctic ice. There are 5160 digital optical modules (DOMs) which make up 78 IceCube strings and 8 DeepCore strings (figure 1). The DeepCore strings are more densely instrumented and located at the lower center of the IceCube string array. The DeepCore subdetector lowers the observable energy threshold to approximately 5 GeV, providing an opportunity to study neutrino oscillations in IceCube. It is sensitive to the neutrino mixing angle θ_{23} and mass splitting Δm^2_{32}, which can be measured by studying ν_μ disappearance using atmospheric neutrinos created in cosmic ray air showers. Neutrino oscillation probabilities depend on the ratio of neutrino energy to neutrino travel distance, which can be inferred using the incident neutrino zenith angle (θ_{zenith}). Precisely measuring θ_{zenith} is critical in measuring oscillation parameters.

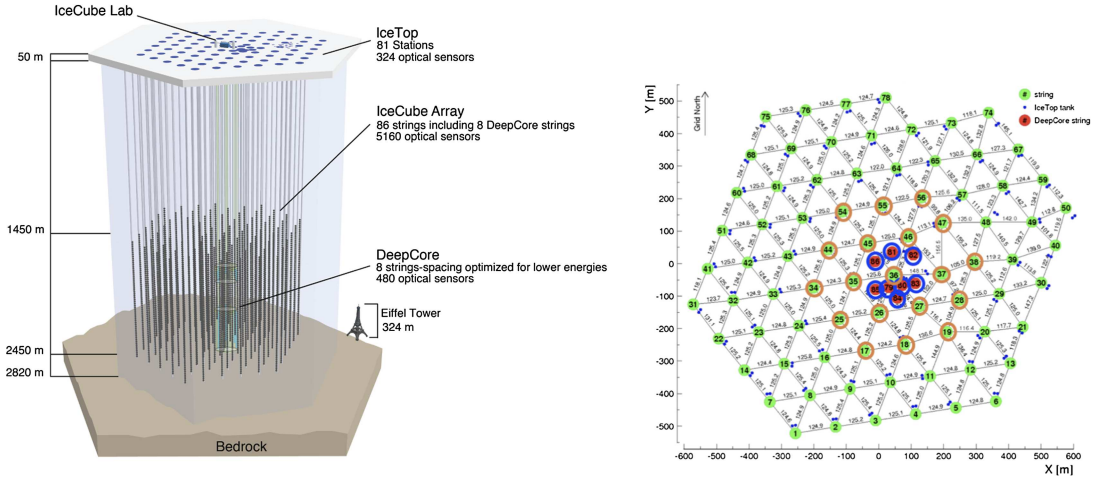

Figure 1. IceCube Neutrino Observatory at the South Pole (left) and top view of detector strings (right) with 8 DeepCore strings (red filled) and 19 IceCube strings (orange circled) used as input to CNN.

When neutrino interactions take place within the detector, relativistic charged particles are produced and propagate in the ice, emitting Cherenkov photons which are detected by the DOMs

2021 JINST 16 C11001

and converted into a series of electrical pulses, or hits. A convolutional network (CNN) is employed to reconstruct θ_{zenith} by using these hits.

2 Convolutional neural network

CNNs are broadly used in modern physics experiments for particle identification [1] and reconstruction [2, 3]. The CNN employed for θ_{zenith} reconstruction has the structure as shown in figure 2.

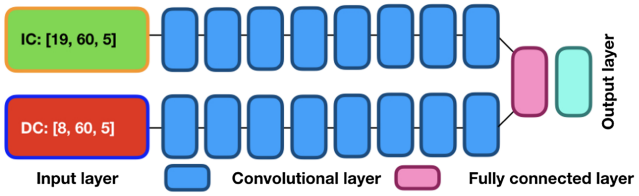

Figure 2. Structure of CNN with input shape of (number of strings, 60 DOMs, 5 variables).

Most neutrino interactions deposit light signals in both DeepCore and the surrounding IceCube strings. Since the DeepCore and the IceCube strings have different configurations, two separate input layers are used to separate the hits in the neutrino event from the eight DeepCore strings and the 19 surrounding IceCube strings, as shown in the left panel of figure 2. For each of the 60 DOMs on a string, 5 variables are calculated from the pulse series: sum of charges, time of the first hit, time of the last hit, charge weighted mean of pulse time, and charge weighted standard deviation of pulse time. The two sets of input variables pass via 8 convolutional layers separately and the outputs are concatenated at the end of the convolutional layers. The output layer delivers the value of θ_{zenith} for the neutrino event. The training sample is a simulated ν_μ charged-current (CC) Monte-Carlo dataset with true neutrino energy between 5–150 GeV and true θ_{zenith} distribution as shown in figure 3. A total of 3479000 events were used to train the CNN at the high performance computing center at Michigan State University, requiring approximately 6 days and over 600 epochs to converge. The loss function of the CNN is the mean absolute error. The training and validation loss curves are shown in the right plot in figure 3.

3 Performance

The performance of the CNN method is discussed by comparing to the standard likelihood-based reconstruction method using a ν_μ CC sample. Selections applied on reconstructed variables include: neutrino energy in the range 5 to 300 GeV, z-coordinate of neutrino event vertex in the range -500 to -200 m, and $\rho_{36} < 300$ m, where ρ_{36} is the horizontal distance in the xy plane between the reconstructed neutrino interaction vertex and string 36.

The plots in figure 4 show the 1D distributions of $\cos(\theta_{\text{zenith}})$. The CNN method has more events concentrated at the equator ($\cos\theta = 0$) while the likelihood-based method over-distributes events near the poles ($\cos\theta = \pm 1$). This could be due to the zenith distribution of the training sample and might be resolved in future by using a training sample with flat true θ_{zenith} distribution.

2021 JINST 16 C11001

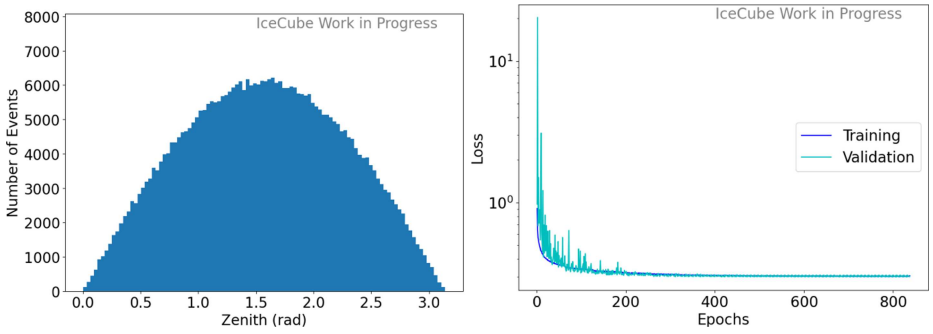

Figure 3. True θ_{zenith} distribution of training sample (left); training (blue) and validation (teal) loss curves (right).

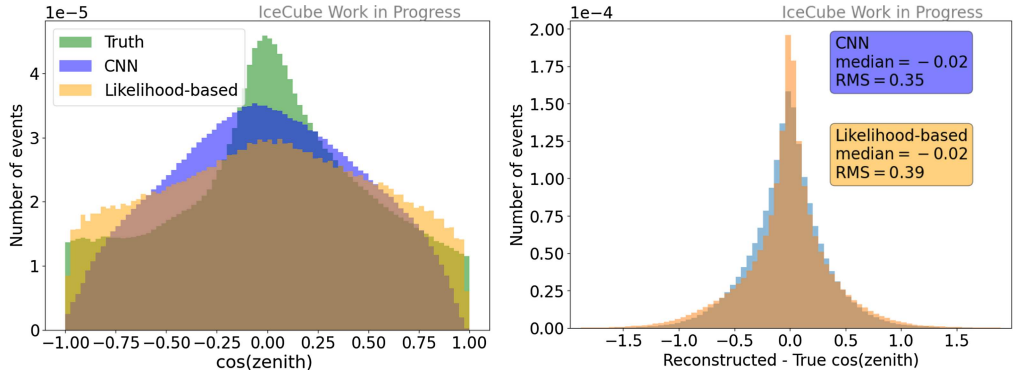

Figure 4. 1D distributions of $\cos(\theta_{\text{zenith}})$ (left) and $\cos(\theta_{\text{zenith}})$ reconstruction error (right) with blue (orange) representing CNN (likelihood-based) reconstructed $\cos(\theta_{\text{zenith}})$ and green representing true $\cos(\theta_{\text{zenith}})$ of true ν_μ CC events.

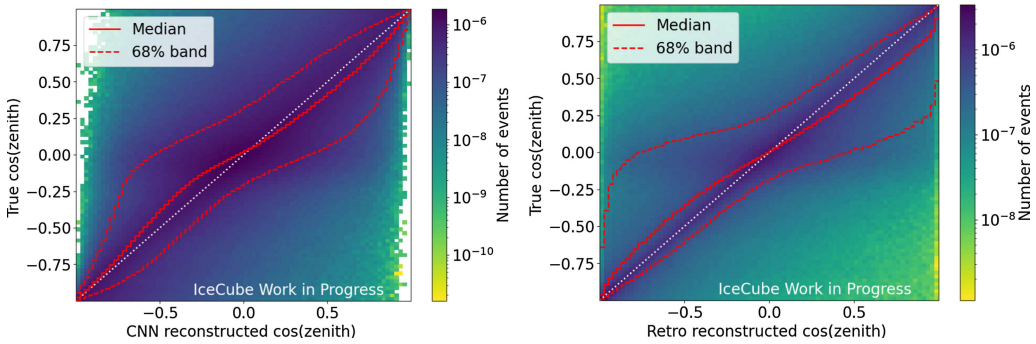

Figure 5. 2D distributions of true vs. CNN (left) or likelihood-based (right) reconstructed $\cos(\theta_{\text{zenith}})$ with median (solid) and contours (dashed) of 68% of events in vertical slices.

The overall RMS of CNN method is smaller than that of the likelihood-based method by 10.3%. In the 2D distributions of true versus reconstructed $\cos(\theta_{\text{zenith}})$ (see figure 5) the CNN method has a similar median contour and narrower 68%-contours than those of the likelihood-based method.

2021 JINST 16 C11001

The CNN and likelihood-based methods have a similar bias against true or reconstructed $\cos(\theta_{\text{zenith}})$ (see figure 6) and true neutrino energy (see figure 7).

Figure 6. 1D slices of reconstructed — true vs. true (left) or reconstructed (right) $\cos(\theta_{\text{zenith}})$ with blue (orange) representing CNN (likelihood-based) result, solid curve representing median, and shaded area containing 68% of events.

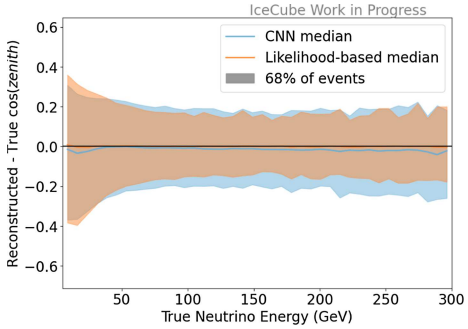

Figure 7. 1D slices of reconstructed — true vs. true neutrino energy with blue (orange) representing CNN (likelihood-based) result, solid curve representing median, and shaded area containing 68% of events.

As listed in table 1, the CNN method can run on both CPU and GPU and on a GPU can achieve a speed-up of 10000 compared to the likelihood-based method. Rapid processing is crucial for the oscillation analysis of the very large neutrino dataset provided by IceCube-DeepCore.

Table 1. Processing speed of CNN and likelihood-based methods.

Second/Event	GPU	CPU
CNN	0.0044	0.108
Likelihood-based	-	44.97

2021 JINST 16 C11001

4 Conclusion

The CNN method provides a comparable performance to the current likelihood-based method on O(100)-GeV neutrino direction reconstruction, improving the overall RMS in the direction reconstruction by 10% on the ν_μ sample. The bias against either true or reconstructed $\cos(\theta_{\text{zenith}})$ slices are comparable between the two methods. With the help of GPU cluster, the CNN method is up to 10000 times faster than the current method in processing, easing the computational burden required for IceCube oscillation analyses. Future studies will use a training sample with a flat true θ_{zenith} distribution.

References

[1] A. Aurisano et al., *A convolutional neural network neutrino event classifier,* 2016 *JINST* **11** P09001 [arXiv:1604.01444].

[2] IceCube collaboration, *Reconstruction techniques in IceCube using convolutional and generative neural networks, EPJ Web Conf.* **207** (2019) 05005.

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

2021 JINST 16 C11001