A multi-purposed reconstruction method based on machine learning for atmospheric neutrino at JUNO

Teng Li[1], **Hongyue Duyang**[1], **Jiaxi Liu**[2], **Zhen Liu**[2], **Xiaohan Tan**[1], **Zekun Yang**[1] **and Fanrui Zeng**[1]

[1]Institute of Frontier and Interdisciplinary Science, Shandong University, Qingdao, Shandong, China

[2]Institute of High Energy Physics, CAS, 19B Yuquan Road, Shijingshan District, Beijing, China

E-mail: tengli@sdu.edu.cn

Abstract. The Jiangmen Underground Neutrino Observation (JUNO) experiment is designed to measure the neutrino mass order (NMO) using a 20-kton liquid scintillator detector to solve one of the biggest remaining puzzles in neutrino physics. Regarding the sensitivity of JUNO's NMO measurement, besides the precise measurement of reactor neutrinos, the independent measurement of the atmospheric neutrino oscillation has great potential to enhance the sensitivity in the combined analysis. This heavily relies on the event reconstruction performance at high energy (GeV) level, including the angular resolution of the incident neutrino, the energy resolution, as well as the accuracy of the flavor identification etc. In this contribution, we present a multi-purposed reconstruction algorithm for high energy particles in JUNO based on machine learning method. This includes extracting effective features from tens of thousands of PMT waveforms, as well as the development of two types of machine learning models (spherical GNN and planar CNN/Transformer). Novel techniques, such as improving the model convergence speed and eliminating reconstruction bias by maintaining the rotation-invariance are also discussed. Preliminary results based on JUNO simulation present reconstruction precision at an unprecedented level, showing great application potential for other large liquid scintillator detectors as well.

1. Introduction

The Jiangmen Underground Neutrino Observation (JUNO) [1, 2] is a multipurpose experiment designed to determine the neutrino mass order (NMO) and to measure several neutrino oscillation parameters precisely. Figure 1 shows the central detector of JUNO, which is a 20 kton liquid scintillator (LS) detector equipped with about 18,000 20 inch PMTs and 25,000 3 inch PMTs distributed on the sphere, placed about 700 meters underground.

The sensitivity of JUNO's NMO measurement mainly comes from the detection of reactor neutrinos from the Yangjiang and Taishan Nuclear Power Plants, located at 53 km away from JUNO. However, the measurement of the atmospheric neutrino oscillation could greatly boost the NMO sensitivity in a joint analysis. This requires precise reconstruction of the GeV level atmospheric neutrino events in the JUNO detector, including,

- The incident angle of the neutrino used to calculate the oscillation baseline.
- The flavor of the neutrino (particle identification).

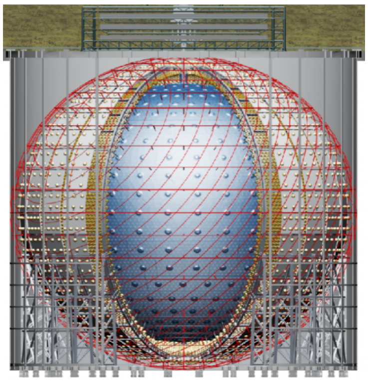

Figure 1. Schematic view of the JUNO detector

- The visible deposit energy of the neutrino in the LS.

Reconstructing the atmospheric neutrino are non-trivial tasks for LS detectors due to their complex interactions in the LS. In particular, since the angular information of particles in the LS is mainly carried by the Cherenkov light annihilated in the overwhelming number of scintillator photons, reconstructing the precise incident angle of neutrinos is extremely challenging based on traditional methods.

In the context of the rapid development of big data processing as well as the deep learning techniques within the high energy physics experiment domain, a feasible way it to take advantage of machine learning models, which are good at extracting hidden information from enormous amount of data. Base on this idea, in this work we propose a general multipurpose reconstruction algorithm for GeV-level particles in LS detectors based on machine learning method. The rest of this paper is organized as follows. In section 2 the basic idea of the method, as well as the machine learning models will be briefly introduced. In section 3 we present the data sample and the performance of our models. The status and outlook of this work is summarized in section 4.

2. Methodology

2.1. Feature engineering

In the LS detector, the light received by a PMT is the superposition of the scintillation light emitted by all tracks in one readout. How the amount of light received by a PMT evolves as a function of time (waveform) depends upon the properties of these tracks, such as the PMT's angle with respect to the track direction, the distance between the PMT and the track, the position of track start point and end point, as well as the dE/dx of the track (decided by the particle ID). In order to reconstruct the properties of these tracks (direction, PID, energy etc.), the basic idea is to use the waveform collected from all PMTs to train the machine learning models, to predict the properties of these tracks.

However, given the large number of PMTs in the JUNO central detector, its quite difficult to feed the machine learning models with the entire waveform from all PMTs. A feasible solution is to extract a few characteristic features (also known as the feature engineering) that reflect the event topology in the detector from the waveform. This method essentially converts the 3-dimention input into several 2-dimention input channels.

The specific features will reflect the properties of the tracks to be reconstructed. Based on what's needed by the neutrino oscillation measurement, the following features are extracted from the waveform of PMTs surrounding the JUNO central detector.

- The time (relative to the triggering time) when the first photon hits the PMT, which contains information of the distance between the track and the PMT, as well as the track angle with regard to the PMT.

- The time (relative to the triggering time) when the ADC value of the waveform reaches the peak, which reflects the information of the track length.
- The ratio of the peak ADC value and the time when ADC reaches the peak (slope of the waveform), reflecting the angle between the track and the PMT.
- The total number of p. e. collected by the PMT, reflecting the energy deposition topology in the LS.

2.2. Machine learning models

Taking the features extracted from the PMT waveform as input, a few state-of-art machine learning models are built to reconstruct the properties of atmospheric neutrinos (direction, energy and PID), including a spherical graph-convolution neural network (GCNN), Deepsphere [3], and two planar convolution neural network (CNN)/Transformer models, EfficientNetV2-S [4] and CoAtNet [5].

Since the PMTs of JUNO central detector are distributed on the sphere, a natual choice is to construct graphs on the sphere directly, with PMTs seen as graph vertices. Then convolution operations can then be performed on the graphs to extract useful signals. Based on this idea, a GCNN model based on Deepsphere is developed. Figure 2 shows the structure of the Deepsphere model. The input data is graphs constructed on the detector sphere. The graph vertices are defined as the input features, while the edges are defined based on the averaged distance between neighbour vertices. Subsequent to the input features, a set of Chebyshev convolution layer and pooling layer are followed. Then a fully connected layer is followed before the prediction block.

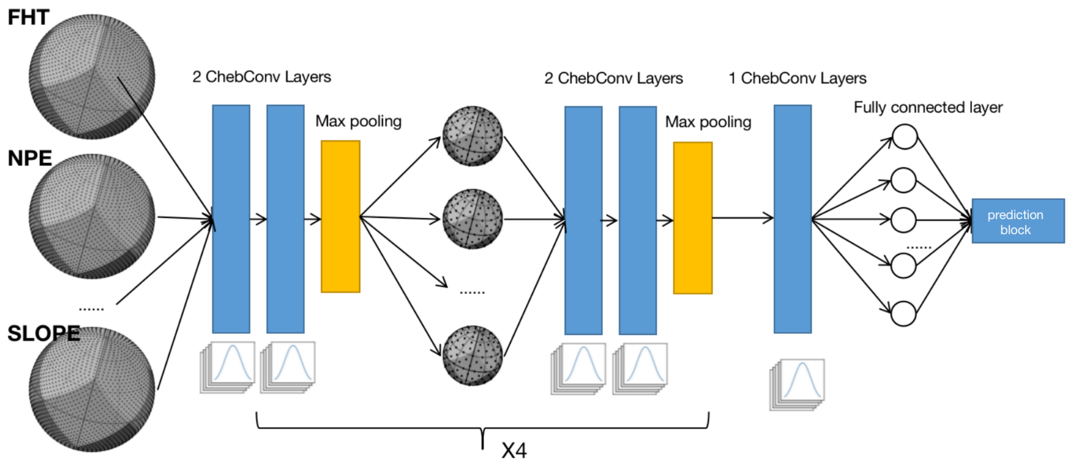

Figure 2. Structure of the Deepsphere model

Besides the GCNN model, an alternative method is to perform a projection from the sphere to a planar surface, which converts the PMT features into a set of 2D figures, then convolution neutral network (CNN) or transformer models are trained to predict the neutrino properties. In this study, a typical CNN structure, EfficientNetV2-S as well as a transformer based model, CoAtNet are studied.

The structure of the prediction block of the machine learning models is decided by the variable to be predicted. For example, for the direction reconstruction, the prediction block implements the regression of the zenith angle. While for the particle flavor identification, the prediction block is a softmax layer with standard cross-entropy loss function. One thing to note is that, to suppress the bias of the angle regression, the models predict the (x,y,z) coordinates of the

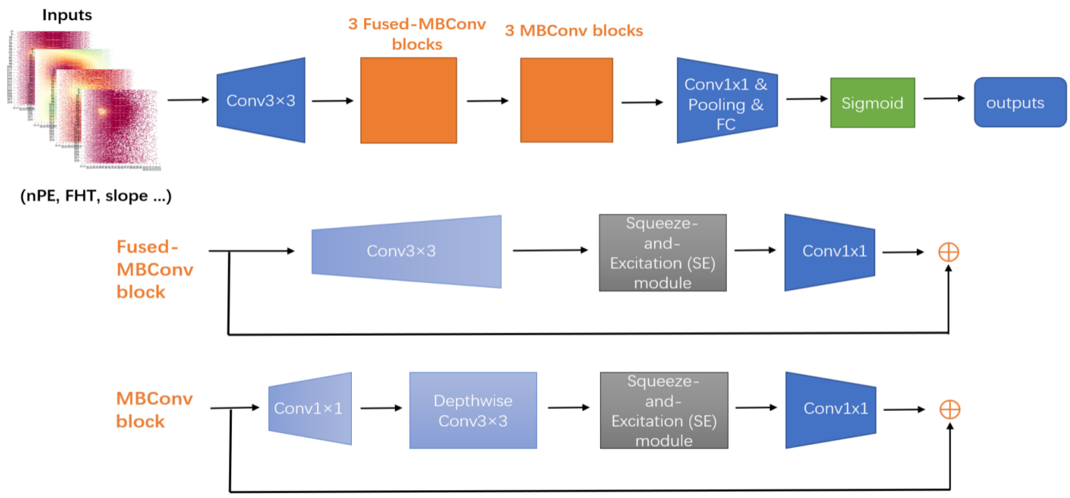

Figure 3. Structure of the EfficientNetV2-S model

endpoint of the angle vector instead of predicting zenith and azimuth angles directly. The loss function is defined as the distance between true endpoints and predicted ones.

3. Performance

The training and test dataset used in this study is simulated based on the Honda atmospheric neutrino flux [6]. The detector and PMT response is simulated via the JUNO offline software (JUNOSW) [7]. For the directionality, energy and interaction vertex reconstruction tasks, a total number of 135,000 nu_{mu} events with energy ranging from 1 GeV to 20 GeV are selected, 95,000 of which is used as training set, while the rest 40,000 as testing set. For the PID task, sub-samples of ν_μ, ν_e events from the charged current (CC) and events from the neutral current (NC) are divided. The models are trained to predict the correct label from ν_μ CC, ν_e CC and NC.

Figure 4 shows the performance of direction reconstruction on the test dataset. Both spherical and planar models provide similar zenith angular resolution, which gets better as the neutrino energy increases. The overall Gaussian fit of the zenith angular resolution is 9.96 and 11.71 degrees for the Deepsphere and EfficientNetV2-S models, respectively.

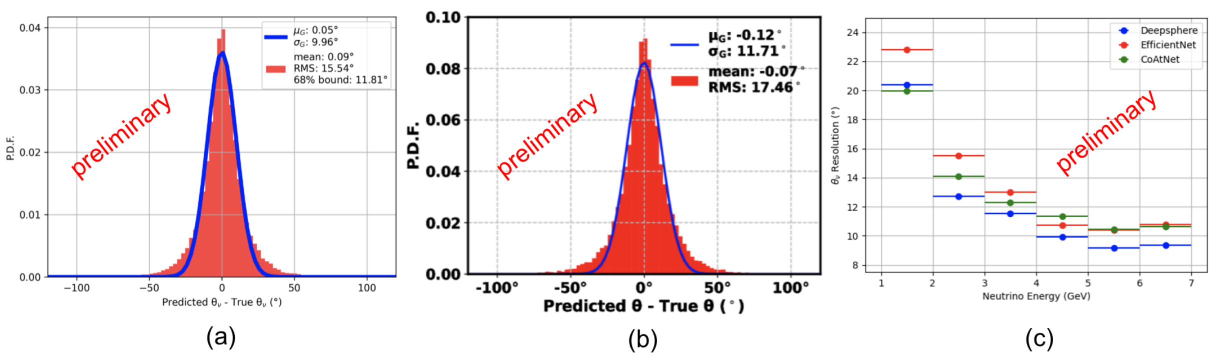

Figure 4. (a) Overall zenith angle resolution of the Deepsphere model. (b) Overall zenith angle resolution of the EfficientNetV2-S model. (c) Zenith angle resolution as a function of neutrino visible energy.

While reconstructing the visible energy, the training set is re-weighted for a flat energy spectrum to suppress the regression bias. The performance of the Deepsphere model is summarized in Figure 5. The energy resolution (up to 1%) also gets better while the visible energy increases, as more kinetic energy is carried by the leptons from neutrinos interactions with larger energy.

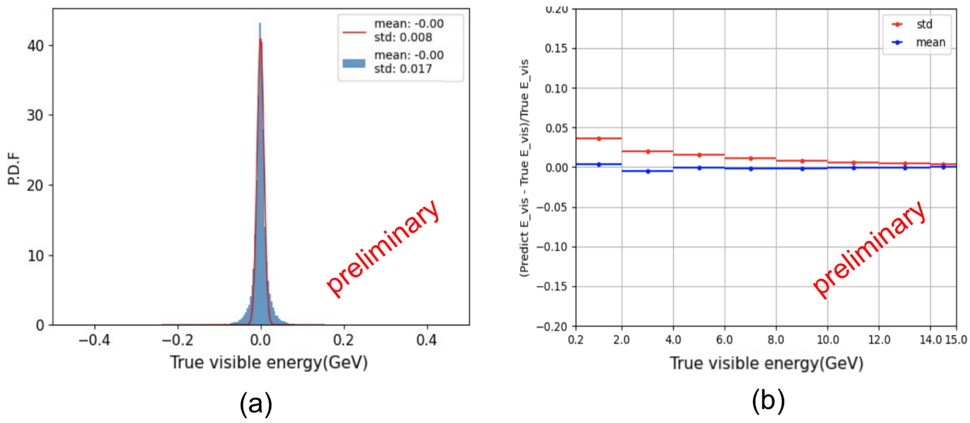

Figure 5. (a) Overall visible energy resolution of ν_μ CC events from Deepsphere. (b) Visible energy resolution as a function of neutrino visible energy.

Figure 6 shows the confusion matrix of the 3-label PID task of the Deepsphere model. The overall efficiency and purity of ν_μ CC events are 82% and 85%, while the efficiency and purity of ν_e CC events are 91% and 67%. Since the default score cut is used, it can be further optimized for the efficiency and purity trade-off.

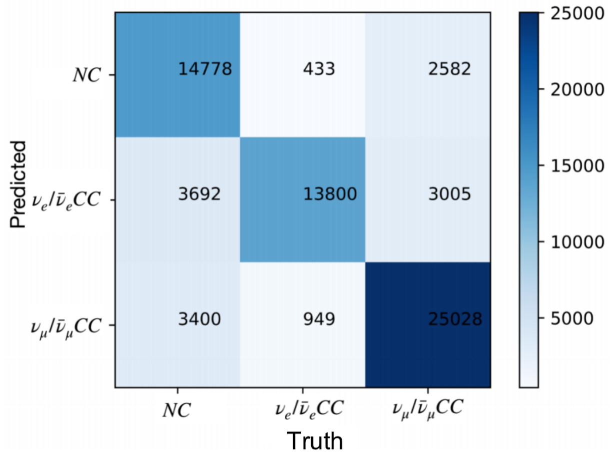

Figure 6. Confution matrix of the PID task by Deepsphere.

4. Summary

In this work, a general reconstruction approach for GeV level particles in the LS detector is introduced. The method is based on extracting features from PMTs distributed on the detector and training machine learning models to predict desired properties of the particles. Multiple machine learning models (including spherical GNN, CNN and Transformer) are developed and cross-validated. Based on the MC sample produced by JUNO, multiple properties (including directionality, energy and PID) of atmospheric neutrinos are reconstructed precisely. For non-trivial reconstruction tasks, such as the neutrino directionality, the performance is unprecedent.

5. Acknowledgments

This work is in supported by National Key R&D Program of China under Contract No. 2020YFA0406304 and National Natural Science Foundation of China (NSFC) under Contracts Nos. 12025502, 12105158, 12175124, 12188102.

References

[1] JUNO Collaboration, JUNO physics and detector, Prog.Part.Nucl.Phys. 123 (2022), 103927
[2] JUNO Collaboration, Neutrino Physics with JUNO, J.Phys.G 43 (2016) 3, 030401
[3] N. Perraudin et al., DeepSphere: Efficient spherical convolutional neural network with HEALPix sampling for cosmological applications, Astronomy and Computing 27 (2019) 130-146
[4] Mingxing Tan and Quoc V. Le, EfficientNetV2: Smaller Models and Faster Training, arXiv:2104.00298
[5] Z. H. Dai et al., CoAtNet: Marrying Convolution and Attention for All Data Sizes, arXiv:2106.04803
[6] M. Honda et al., Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data, Phys. Rev. D 75 (2007), 043006
[7] X. T. Huang et al., Offline data processing software for the JUNO experiment, PoS ICHEP2016 (2017) 1051