Measurement of Atmospheric Muon Neutrino Disappearance using CNN Reconstructions with IceCube

Shiqi Yu
Michigan State University

NuFact 2022
Neutrino Oscillation

- Produced and detected in 3 flavor states;
- Propagate in mass states;
- Described by PMNS matrix.

\[
\begin{pmatrix}
\nu_e \\
\nu_\mu \\
\nu_\tau
\end{pmatrix} = 
\begin{bmatrix}
U_{e1} & U_{e2} & U_{e3} \\
U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\
U_{\tau 1} & U_{\tau 2} & U_{\tau 3}
\end{bmatrix}
\begin{pmatrix}
\nu_1 \\
\nu_2 \\
\nu_3
\end{pmatrix}
\]

Some parameters need to be better measured: $\theta_{23}$, $\Delta m^2_{32}$

Atmospheric & LBL  reactor & LBL  Solar
Neutrino Oscillation

- Atmospheric muon neutrino beam from cosmic ray interactions;
- Neutrino distance of travel (L) calculated using arrival direction (zenith)

Reconstruction is critical for studying atmospheric neutrino oscillation at 10-GeV scales

\[
P(\nu_\mu \rightarrow \nu_\mu) \approx 1 - \sin^2(2\theta_{23}) \sin^2\left(\frac{1.27 \Delta m^2_{32} L}{E}\right)
\]
IceCube Neutrino Observatory

- 1 km$^3$ neutrino detector deep under South Pole ice;
- 5160 digital optical modules (DOMs) detect Cherenkov photons emitted during neutrino interactions;
- DOMs record pulse charges & times
- **DeepCore**: denser configured sub-detector, can observe **GeV-scale neutrinos**;
List of Reconstructed Variables

Reconstructions:

- Energy
- Direction (L)
- PID
- Interaction vertex
- Muon classifier

Analysis binnings

Selections

Track-like events:
\( \nu_\mu \) CC, 17% \( \nu_\tau \) CC

Cascade-like events:
\( \nu_e \) CC, NC, \( \nu_\tau \) CC

\( \nu_\mu(65.4\text{GeV}) \rightarrow \mu_\nu(62.7\text{GeV}) + \text{hadrons} \)

\( \nu_e(67\text{GeV}) \rightarrow e_\nu(57.5\text{GeV}) + \text{hadrons} \)
**GeV-Scale CNN Architecture**

- Only use DeepCore & nearby IceCube strings
- Five CNNs trained & optimized separately

**Inputs:** 5 summarized variables
- sum of charges
- time of first (last) pulse
- charge weighted mean (std.) of times of pulses

**Regression:**
- Energy, direction, interaction vertex

**Classification:**
- PID, muon classifier
Training Samples

- Balanced MC samples;
- Energy, direction, interaction vertex are trained on $\nu_\mu$ CC events (signal);
- PID and muon classifiers are trained on balanced samples.

7 million

Muon Classifier

4.2 million in total

$\text{NuE}$
20.0%

$\nu_\mu$
40.0%

$\text{NuMu}$
40.0%

PID: $\nu_\mu$ CC

6 million in total

Cascade
50.0%

$\nu_\mu$ CC
50.0%

$\nu_\mu$ NC

Testing Samples

- Nominal MC sample with flux, xsec, and oscillation weights applied;
- Testing on signal ($\nu_\mu$ CC) and major background ($\nu_e$ CC);
- Baseline: current reconstruction method (likelihood-based)

K. Leonard IceCube plenary talk
Reconstruction Performance

- Flat median against true neutrino energy and zenith;
- CNN has comparable resolution to current method, and better at low energy (majority of sample)
**Performance: Vertex**

- Selecting events starting near DeepCore;
- Comparable purities in selected $\nu_\mu$ CC samples.

![CNN Prediction](chart1)

![Likelihood-based Prediction](chart2)
Performance: Muon and PID Classifiers

- Comparable performance to the current methods:
  - Similar AUC values.
  - Hard to identify track from cascades at low energy → less DOMs see photons.
## Performance: Speed

<table>
<thead>
<tr>
<th></th>
<th>Second per file (~3k events)</th>
<th>Time for full sample assuming 1000 cores</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN on GPU</td>
<td>21</td>
<td>~ 13 minutes</td>
</tr>
<tr>
<td>CNN on CPU</td>
<td>45</td>
<td>~ 7.5 hours</td>
</tr>
<tr>
<td>Current Likelihood-based method (CPU only)</td>
<td>120,000</td>
<td>~ 46 days</td>
</tr>
</tbody>
</table>

- CNN runtime improvement: ~6,000 times faster;
  - CNNs are able to process in parallelize with clusters → can be even faster!
- Big advantage: large production of full Monte Carlo simulations ~$O(10^8)$.
Preliminary Sample

- Event processings up to final level shared with the current analysis: K. Leonard IceCube plenary talk
- Final sample: high signal ($\nu_{\mu}$ CC) and low background (noise and cosmic-ray muon) rates (~0.6%).
Measuring Oscillation

Measure atmospheric muon neutrino disappearance in 3D binning: reconstructed [energy, cos(zenith), PID]

- PID discriminates $\nu_\mu$ CC vs. all other interactions
- Robust against systematic uncertainties

IceCube Work in Progress

Cascade-like

Mixed

Track-like
Oscillation Sensitivity

Oscillation analysis using CNN reconstruction has similar sensitivity (black) as IceCube’s current likelihood-based analysis (red)

- Sensitivities projected from DeepCore 2021 (golden) [K. Leonard IceCube plenary talk]
- 6000 times faster!
- Apply to future detector — the Upgrade
- Analysis is unblinding, new results soon!
Future

The Upgrade detector:

- More densely instrumented strings in the center
- DOM: multiple PMT designs
- **Target deploying 2024/25**

New reconstruction methods needed:

CNN is one solution
Conclusion

- CNNs are used for multipurpose reconstructions for IceCube oscillation analysis:
  - Energy, direction, interaction vertex;
  - PID (numu CC vs. background neutrinos), muon classifier.
- Approximately 6000 times faster in runtime than the current method;
  - Big advantage for IceCube full production → large atmospheric neutrino sample.
- CNNs have better or comparable performances to the current reconstruction method;

- Ongoing and future work:
  - numu disappearance analysis using CNN reconstructions;
  - Optimizations on CNN itself;
    - Train for “ending point”, etc.
  - Implement it for future experiment → Upgrade.
Thank you!
Training Samples

Energy: nDOM $\geq 7$
Muon: nDOM $\geq 4$; 5–200 GeV
Muon, PID, Vertex: nhits $\geq 8$ hit 5-200 GeV
Zenith: full containment cut on true vertexes, 5-300 GeV
Performance: Direction

- Direction bias flat against true energy;
- Comparable to current method;
- Better resolution for $\nu_\mu$ CC (signal);
- High energy (>100 GeV) neutrinos leaving DeepCore
  - Need containment cut: interaction vertex reconstruction.
Performance: Energy

- Flat median against true neutrino energy;
  - CNN has better resolution at low energy (majority of sample)
- Comparable performance to current method at higher energy and in background;

\[
\nu_\mu \text{ CC} \quad \text{and} \quad \nu_e \text{ CC}
\]
Performance: Zenith

- Flat median against true direction;
- Comparable to current method in both signal and background.

Performance: Zenith
(Contained, 5-300 GeV Sample)
Performance: Zenith (Analysis Samples)
Performance: Vertex

Efficiency matrixes

![Efficiency matrixes](image)

- CNN Prediction
  - **Inside**: 94.48% of truth, 5.52% of truth
  - **Outside**: 42.75% of truth, 57.25% of truth

- Likelihood Prediction
  - **Inside**: 97.53% of truth, 2.47% of truth
  - **Outside**: 51.94% of truth, 48.06% of truth
Systematic Effect: Neutrino Flux Model

Neutrino flux spectral index variation has different signature to expected oscillation signal

- **Cascade-like**
- **Mixed**
- **Track-like**

Fit for spectral index among other model systematics

\[ N_{\sigma} = \frac{N_{\text{pulled}} - N_{\text{nominal}}}{\sqrt{N_{\text{nominal}}}} \]

IceCube Work in Progress

Flux model systematic: Neutrino flux spectral index changed by $+1\sigma$
Physics Motivations: Neutrino Oscillations

\[
\begin{bmatrix}
\nu_e \\
\nu_\mu \\
\nu_\tau
\end{bmatrix} = U_{PMNS} \times \begin{bmatrix}
\nu_1 \\
\nu_2 \\
\nu_3
\end{bmatrix}
\]

- Neutrino flavor eigenstates are superpositions of mass eigenstates.
- Relations described by PMNS matrix.

\[P(\nu_\mu \rightarrow \nu_\mu) \approx 1 - \sin^2(2\theta_{23}) \sin^2\left(\frac{1.27 \Delta m_{32}^2 L}{E}\right)\]

- Most parameters are well measured.
- Some parameters need to be better measured: $\theta_{23}$ and $\Delta m_{32}^2$
IceCube Oscillation Results

Main results + current projection on sensitivity
Kayla’s plenary on Monday

We’ll show an alternative way of doing numu disappearance using convolutional neural networks: 6000 times faster in runtime; similar sensitivity; portable for the future experiment, the Upgrade.