Neutrinos & Neural Networks: Reconstructing GeV Scale IceCube Neutrinos

Jessie Micallef
Michigan State University
micall12@msu.edu
About Me - Portrait of a Scientist

KOTO (Japan!)

IceCube

Co-Chair of Michigan State Conference for Undergraduate Women in Physics 2019

Comic Cons

Concerts (Croatia!)

Jiu Jitsu & Sword Fighting
Neutrinos and IceCube

Why are neutrinos interesting?
How do we see them with IceCube?
Neutral leptons with 3 flavors:
- Electron
- Muon
- Tau

Produced and interact in flavor states
Propagate in mass states
Where do we see Neutrino Oscillation?

- Neutrino oscillation can be easily observed on Earth at GeV-scales
- To measure oscillation parameters...

\[ P_{\alpha \to \beta}(L) \propto \sin \left( 1.27 \frac{\Delta m^2_{ij} L}{E} \right) \]

... need to reconstruct the neutrino’s

- Energy
- Distance (calculated from incident angle traveling through earth)
- Flavor

Plot Credit: PISA at https://arxiv.org/abs/1803.05390
How Do We See Neutrinos?

IceCube Neutrino Observatory: Detects astrophysical and atmospheric neutrinos to determine their sources and measure neutrino characteristics, such as oscillation parameters.

- 125 meters between strings
- 17 meters between optical modules on string
- 72 meters between strings
- 7 meters between optical modules on string
- Optimized to detect lower energy events (GeV-scale)
“Typical” Event Signatures in IceCube

1. Neutrinos interact with nucleons in ice, emitting charged particles
2. Charged particles travel faster than the speed of light in ice, emitting blue light called Cherenkov radiation
3. Optical modules record pulse charges & times

Track-like events:
- Source: $\nu_\mu$ CC
- Energy: 71 TeV

Cascade-like events:
- Source: $\nu_e$ CC, $\nu_\tau$ CC, all NC
- Energy: 2 PeV

→ Can make a “picture” or video, so can we use image recognition?
- Yes! Successful convolutional neural network for reconstructing high energy cascade events in IceCube: arXiv:2101.11589v1
Tackling 10 GeV-Scale Neutrino Events in IceCube

Typical 10 GeV scale event:

- Less light produced per event means fewer optical modules record pulses
- Must leverage DeepCore array
- Need to optimize neural network specifically for these events

Challenging to determine:
- Track or cascade
- Direction
- Energy
Neural Networks for IceCube

Goal of this work: optimize convolutional neural network to reconstruct neutrinos for 10 GeV-scale $\nu_\mu$ CC and $\nu_e$ CC events
Convolutional Neural Networks (CNNs)

Shadow is the kernel moving across the nearby inputs as it searches the entire layer 1 and outputs a weighted layer 2.

CNN kernel in depth going down optical modules

CNN Gif Credit: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
Preparing CNN for GeV-Scale Neutrinos

→ Only use DeepCore & nearby IceCube strings (kernel in depth only)
→ Noise cleaning applied & hit time within [-500, 4000] ns

Inputs: 5 variables that summarize all pulses hitting optical module
- Sum of charge
- Time of first hit
- Time of last hit
- Charge weighted mean of times
- Charge weighted $\sigma$ of times
GeV-Scale CNN Architecture

Five separate CNNs trained & optimized for “single” output.

**Regressions:**
1. Energy
2. Zenith
3. Interaction Vertex → (x, y, and z)

**Classifications:**
4. Track vs Cascade (flavor)
5. Muon vs Neutrino

→ Everything we need for oscillations analysis (+ more!)
03 Results of CNN

How well does the CNN do?
How does the CNN compare to current likelihood-based methods?
Testing Samples

→ Testing sample with atmospheric flux & oscillation model weights applied
→ Distributions expected to be similar to data
→ Separate testing samples for $\nu_\mu$ CC & $\nu_e$ CC

→ Vertex reconstruction resolution in backup
CNN Resolution: Energy

- CNN’s energy prediction best at low energies!
  - Where majority of data expected
- CNN’s energy resolution median near zero ✓
- Comparable to likelihood method ✓
- Similar resolution between $\nu_\mu$ and $\nu_\text{e}$
CNN Resolution: Cosine Zenith

- CNN’s zenith resolution median near zero ✓
- Comparable to likelihood method ✓
- Better resolution expected for $\nu_\mu$ CC (tracks)
- CNN difference at high energies due to containment (leaving DeepCore array)
  - Cuts being explored using CNN vertex reconstruction
CNN Classifiers: Tracks and Muons

- CNN comparable to likelihood-based method’s classification
  - Larger AUC is better!
- Expect difficulty distinguishing tracks and cascades at these low energies

**Identifying Flavor**

ROC Probability Track vs Cascade

- IceCube Work In Progress

CNN AUC: 0.682
Likelihood AUC: 0.674

**Identifying Background**

ROC Probability Neutrino vs Muon

- IceCube Work In Progress

CNN AUC: 0.928
Likelihood AUC: 0.926
CNN Significantly Reduces Reconstruction Time

<table>
<thead>
<tr>
<th></th>
<th>Average time (s) per event</th>
<th>Events per day per single core</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN on GPU</td>
<td>0.0077</td>
<td>11,000,000</td>
</tr>
<tr>
<td>CNN on CPU</td>
<td>0.27</td>
<td>320,000</td>
</tr>
<tr>
<td>Previous Likelihood-based method on CPU*</td>
<td>40</td>
<td>2,100</td>
</tr>
</tbody>
</table>

*Likelihood-based method outputs 8 reconstructed variables

- $10^4$ runtime improvement possible in serial!
- Having access to computer clusters, can parallelize the process
- Turns processing into a day instead of weeks to months
What do these CNN results mean for the future of particle physics?

What other things do we need to be mindful about for our future?
Future of IceCube Low Energy CNN

- CNN provides competitive resolution to likelihood based method
- Faster runtimes important for large atmospheric neutrino data sample!
- CNN handles multipurpose reconstructions & classifications
  - Energy
  - Cosine Zenith
  - Track vs Cascade (flavor)
  - Vertex
  - Muon vs Neutrino

Measuring oscillations parameters
Removing background

Ongoing and future work:
- More variables! Ending vertex, interaction time, etc.
- More optimizations -- improved and bigger training samples
- Training for uncertainty -- does the CNN know when it does badly?
Future of Machine Learning & Particle Physics

➢ Machine learning is advancing quickly! Take advantage of new methods
  ○ IceCube exploring RNNs, GNNs, hybrid methods, and more
  ○ Use for...
    ■ Reconstruction
    ■ Generating simulation
    ■ Online filters and triggers

➢ Can be fast and accurate, but requires time to optimize & test
  ○ Lots of work for undergrads, grads, and postdocs!

➢ Relevant for other particle physics experiments too! Especially next generations

⇒ Training next generation of scientists is important!
  ○ Develop skills of new researchers in physics and machine learning
  ○ Support diverse researchers to promote thriving research environment
**PROJECT:** Collect & edit video montages finishing the phrase “I’m a scientist and I also...” with the following
- Race/Ethnicity/Culture
- LGBTQ+/Gender Identity
- Hobbies/Interests/Passions
- And more!

**CONTRIBUTE:**

Contribute a video or picture at [https://tiny.cc/POAS](https://tiny.cc/POAS)

Questions? Email me at micall12@msu.edu or visit [https://www.facebook.com/PortraitOfAScientist](https://www.facebook.com/PortraitOfAScientist)
Thanks!

Do you have any questions?

micall12@msu.edu

https://www.facebook.com/PortraitOfAScientist

https://tiny.cc/POAS

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon and infographics & images by Freepik

Please keep this slide for attribution

Special Thanks to CNN Contributors:
- Shiqi Yu (CNN Zenith)
- Julia Willison (CNN Vertex)

Acknowledgements:
“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.”
Backup
Training Samples Optimized Per Variable

Samples optimized for unbiased training. Examples:

- **CNN for Energy Reconstruction**
  - Flat energy distribution in region of interest
    - 1-200 GeV target
    - Extended to 500 GeV
  - Use $\nu_\mu$ CC events only

- **Track vs Cascade (Flavor) Classifier**
  - 50% Track and 50% Cascade per GeV
  - Includes all $\nu_\mu$, $\nu_e$, CC, and NC

→ Details of other training samples in backup
Training Samples Optimized Per Variable

- Regressions trained with $\nu_\mu$ CC events only
  - **Energy:** flat energy sample in region of interest
    - 1-200 GeV target
    - Extended to 500 GeV
  - **Zenith:** flat zenith sample across all angles
  - **Interaction Vertex:** uses flat energy sample

- Classifications trained with balanced samples
  - **Track vs Cascade:** 50/50 Track vs. Cascade
    - Includes $\nu_\mu$ and $\nu_e$; CC and NC
  - **Muon vs Neutrino:** 40% $\mu$ and 40% $\nu_\mu$ and 20% $\nu_e$
CNN Resolution: Interaction Radius

- Resolution best inside of DeepCore detector (roughly at dotted lines)
- Similar loss of resolution outside of region as likelihood method
- Using CNN Radius prediction as cut
CNN Resolution: Interaction Z Position

- Resolution best inside of DeepCore detector (roughly at dotted lines)
  - Low statistics outside
- Similar loss of resolution outside of region as likelihood method
- Using CNN Z prediction as cut