Neutrinos & Neural Networks:

Reconstructing GeV Scale IceCube Neutrinos

Jessie Micallef
Michigan State University
micall12@msu.edu

ICECUBE

NSF

About Me - Portrait of a Scientist

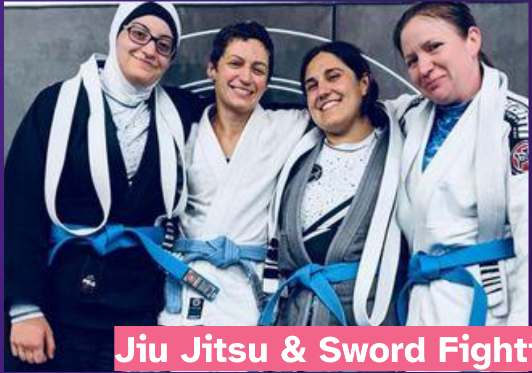
KOTO (Japan!)

IceCube

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

Comic Cons

My Cat Ahsoka



Jiu Jitsu & Sword Fighting

Table of contents

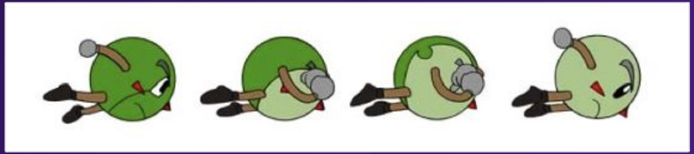
01

Neutrinos and IceCube

Why are neutrinos interesting?

How do we see them with IceCube?

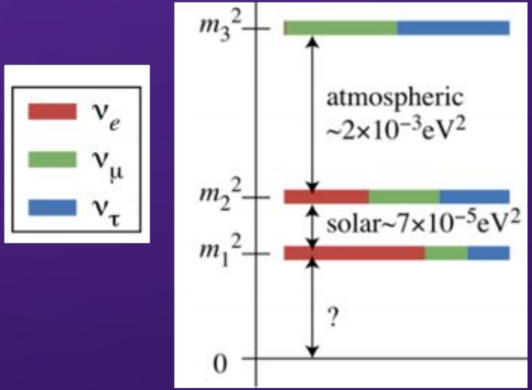
Neutrino Oscillation



https://dchooz.titech.jp.hep.net/nu_oscillation.html

http://docmadhattan.fieldofscience.com/2015/10/a-brief-history-of-neutrinos.html

- Neutral leptons with 3 flavors:
  - Electron
  - Muon
  - Tau
- Produced and interact in flavor states
- Propagate in mass states



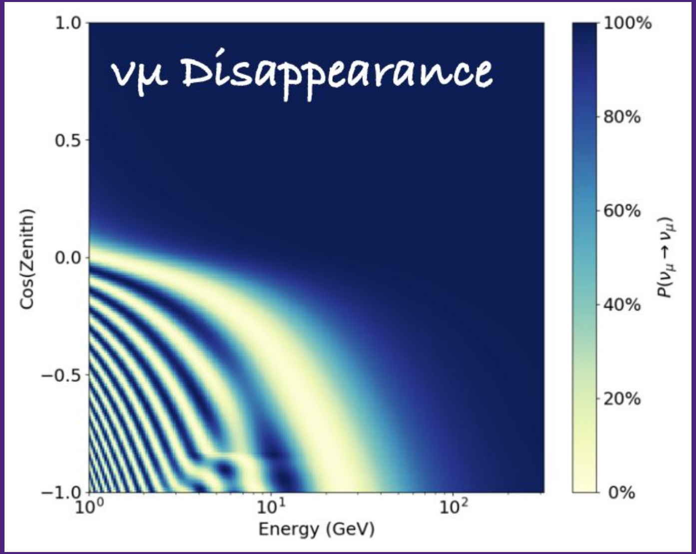
Where do we see Neutrino Oscillation?

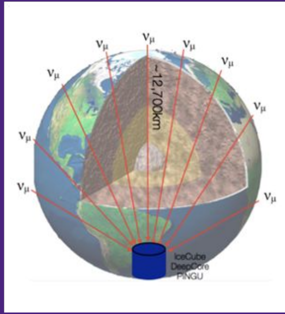
- Neutrino oscillation can be easily observed on Earth at GeV-scales
- To measure oscillation parameters…

$$P_{\alpha\to\beta}(L) \propto \sin^2\left(1.27\frac{\Delta m_{ij}^2 L}{E}\right)$$

… need to reconstruct the neutrino's

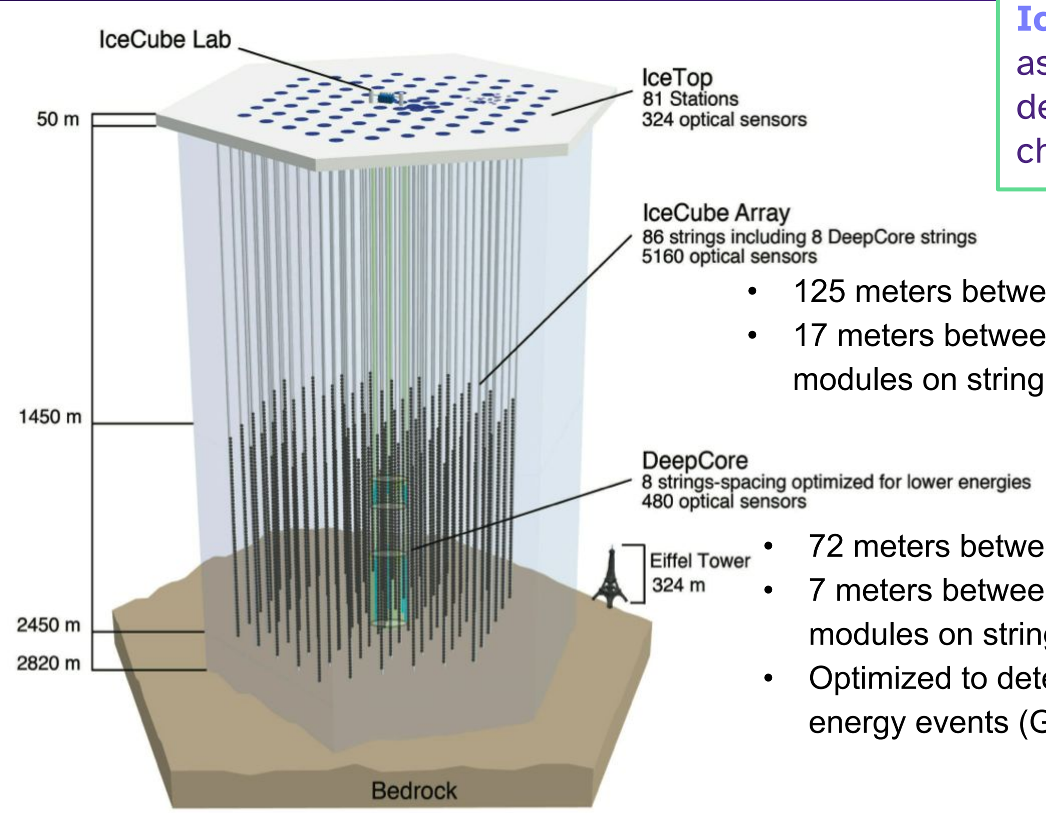
- Energy
- Distance (calculated from incident angle traveling through earth)
- Flavor





Plot Credit: PISA at https://arxiv.org/abs/1803.05390

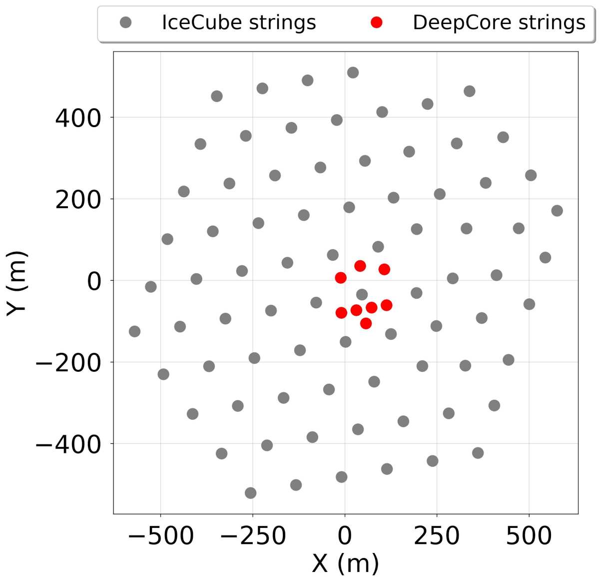
How Do We See Neutrinos?



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

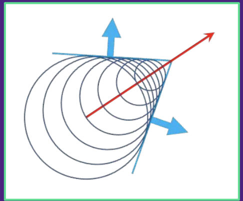
- 125 meters between strings
- 17 meters between optical modules on string

- 72 meters between strings
- 7 meters between optical modules on string
- Optimized to detect lower energy events (GeV-scale)



“Typical” Event Signatures in IceCube

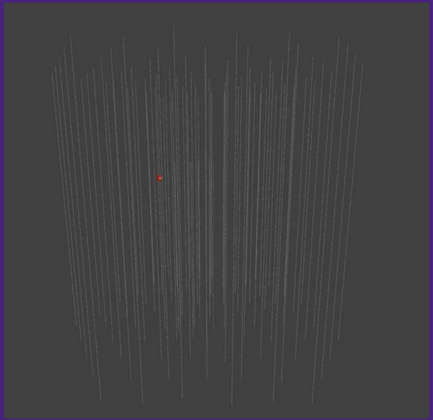
1. Neutrinos interact with nucleons in ice, emitting charged particles
2. Charged particles travel faster than the speed of light in ice, emitting blue light called Cherenkov radiation
3. Optical modules record pulse charges & times



quora_cherenkov_art

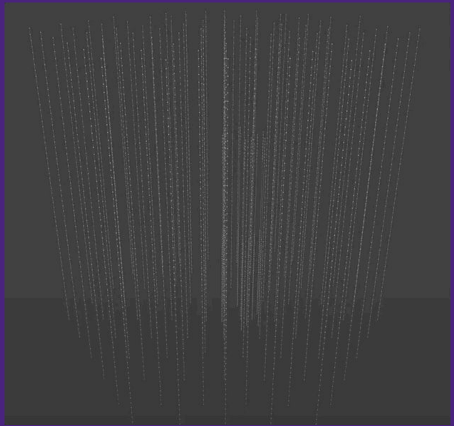
Track-like events:

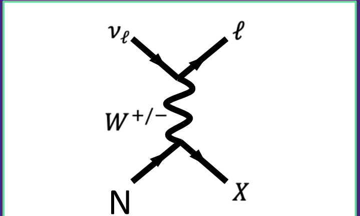
- Source: $\boldsymbol{\nu_\mu}$ **CC**
- Energy: 71 TeV



Cascade-like events:

- Source: $\boldsymbol{\nu_e}$ **CC**, ν_τ CC, all NC
- Energy: 2 PeV



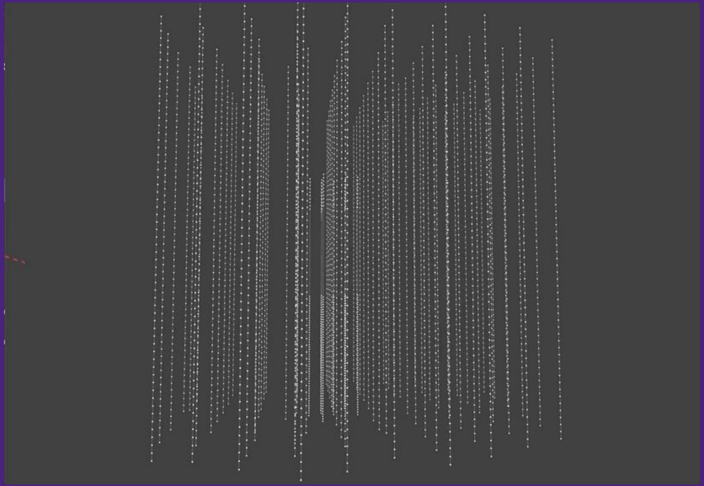


Charged Current (CC)

→ 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:



Challenging to determine
- Track or cascade
- Direction
- Energy

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

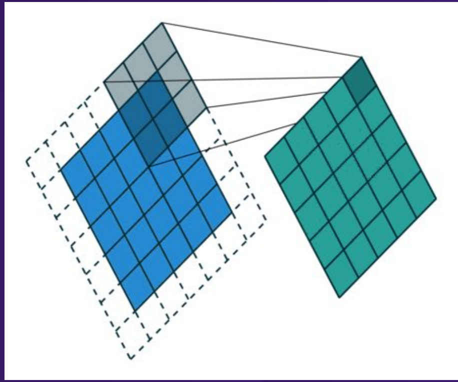
02

Neural Networks for IceCube

Goal of this work: optimize convolutional neural network to reconstruct neutrinos for 10 GeV-scale ν_μ CC and ν_e CC events

Convolutional Neural Networks (CNNs)

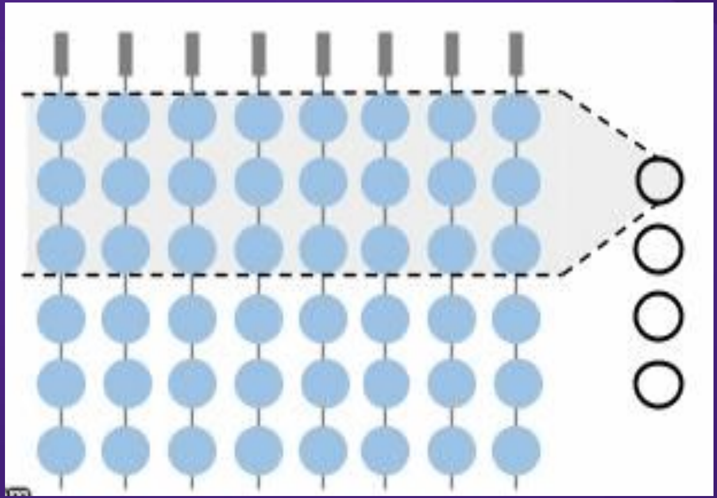
Convolutional Neural Network:



Training Layer 1 → Training Layer 2

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

CNN on IceCube String:

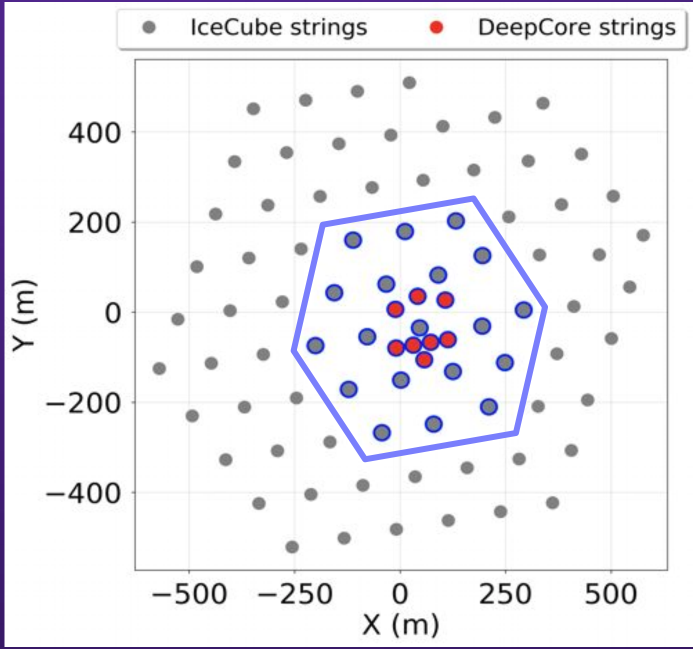


CNN kernel in depth going down optical modules

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

Preparing CNN for GeV-Scale Neutrinos

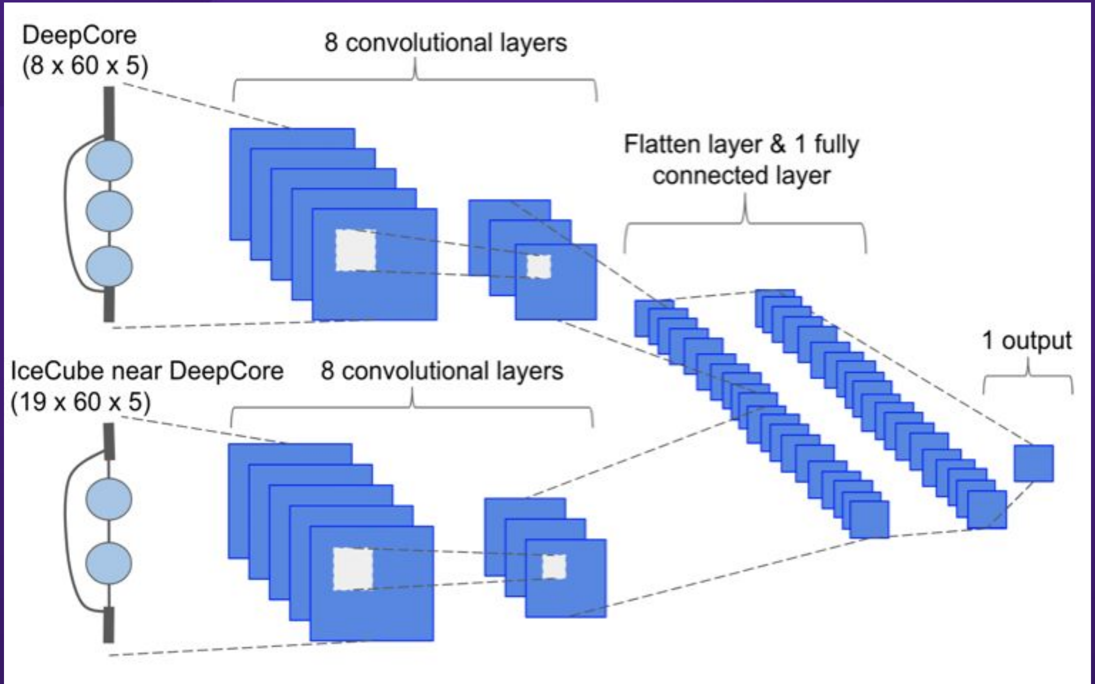
→ Only use DeepCore & nearby IceCube strings (kernel in depth only)
→ Noise cleaning applied & hit time within [-500, 4000] ns



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 σ 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?

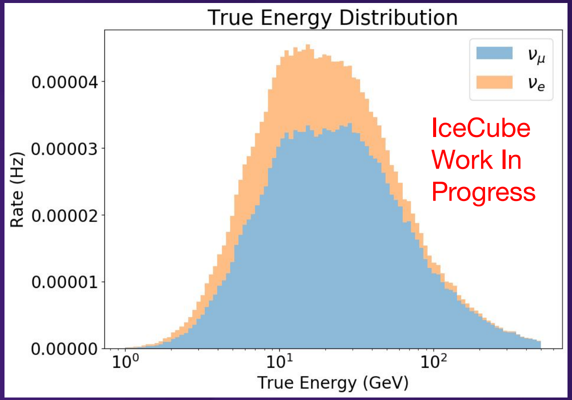
Testing Samples

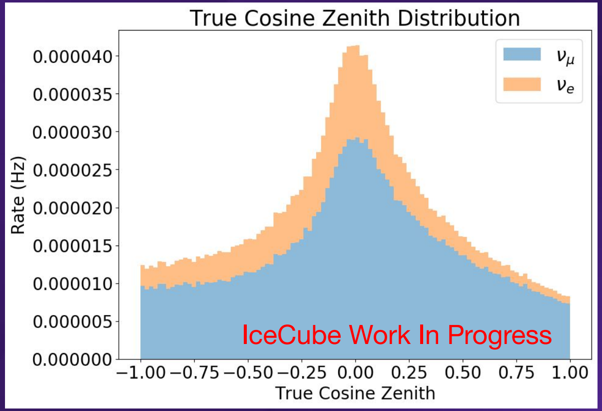
→ Testing sample with atmospheric flux & oscillation model weights applied

→ Distributions expected to be similar to data

→ Separate testing samples for ν_μ CC & ν_e CC

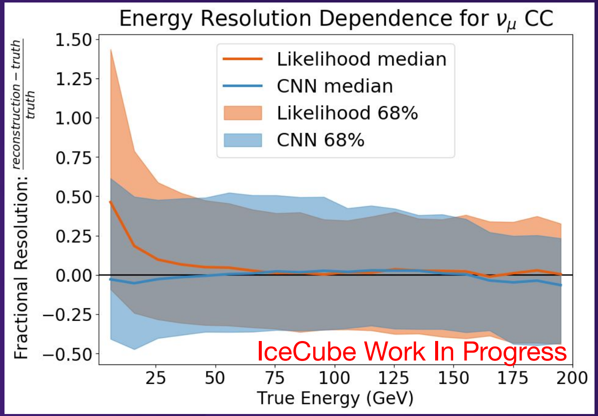
→ Vertex reconstruction resolution in backup

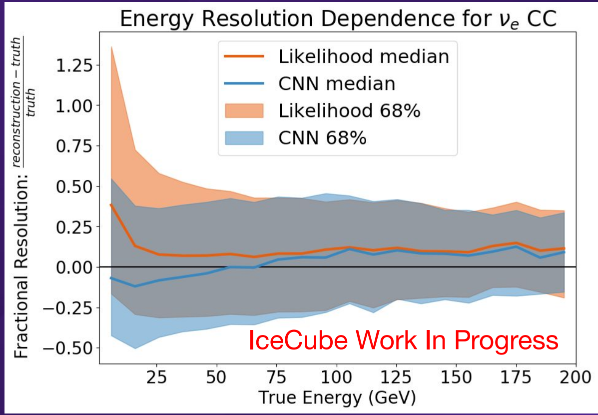




CNN Resolution: Energy

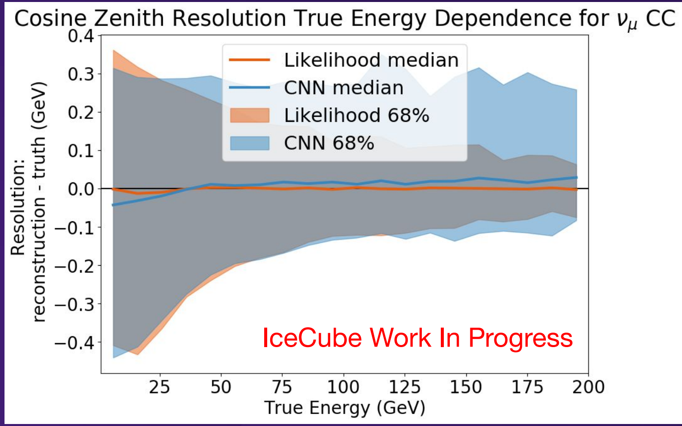
- CNN's energy prediction best at low energies!
  - Where majority of data expected
- CNN's energy resolution median near zero ✓
- Comparable to likelihood method ✓
- Similar resolution between ν_μ and ν_e

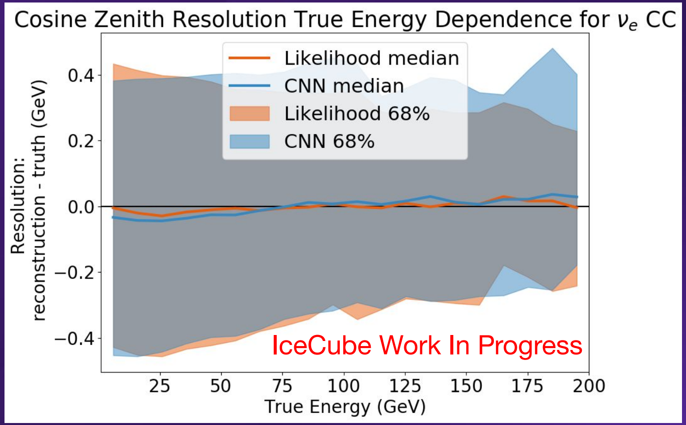




CNN Resolution: Cosine Zenith

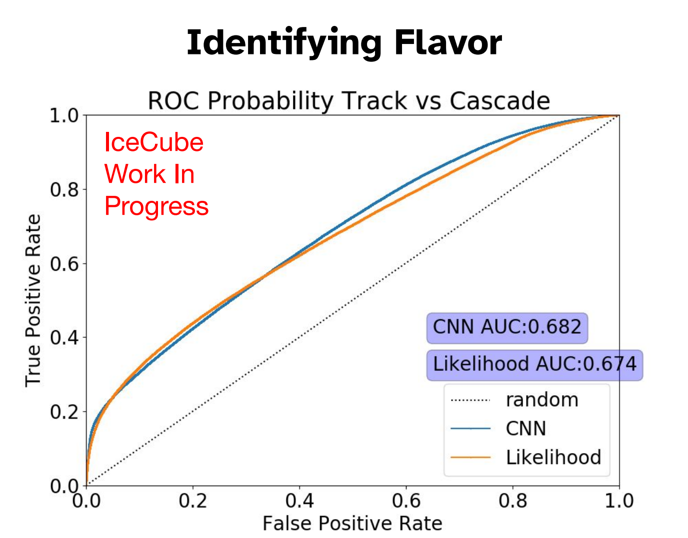
- CNN's zenith resolution median near zero ✓
- Comparable to likelihood method ✓
- Better resolution expected for ν_μ CC (tracks)
- CNN difference at high energies due to containment (leaving DeepCore array)
  - Cuts being explored using CNN vertex reconstruction

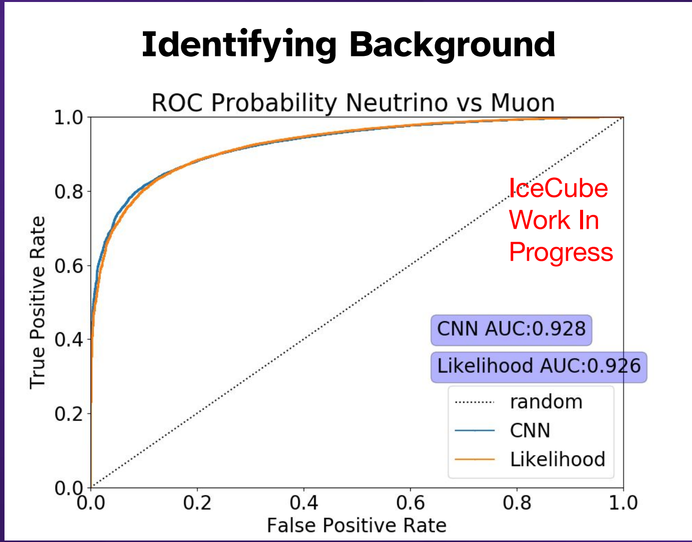




CNN Classifiers: Tracks and Muons

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





CNN Significantly Reduces Reconstruction Time

	Average time (s) per event	Events per day per single core
CNN on GPU	0.0077	11,000,000
CNN on CPU	0.27	320,000
Previous Likelihood-based method on CPU*	40	2,100

*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

04

The Future

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?

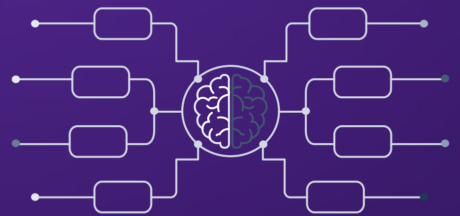
Future of IceCube Low Energy CNN

- 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)

  Measuring oscillations parameters

  - Vertex
  - Muon vs Neutrino

  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

DEEP UNDERGROUND NEUTRINO EXPERIMENT

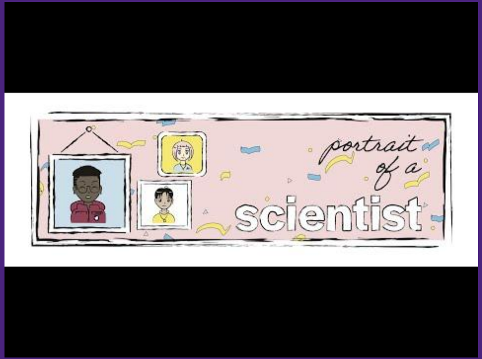
ICECUBE UPGRADE

⇨ 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

Video Montage Project

GOALS:
- Dismantle stereotypes
- Highlight diversity
- Demystify who is & can be a scientist

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 a video or picture at https://tiny.cc/POAS!

Questions? Email me at micall12@msu.edu or visit 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."

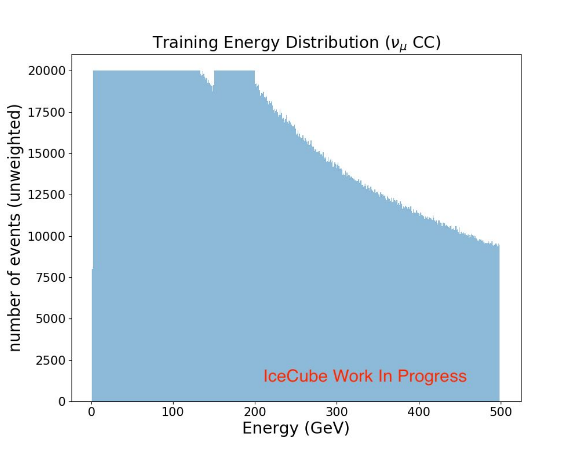
Backup

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

Training Samples Optimized Per Variable

Samples optimized for unbiased training. Examples:

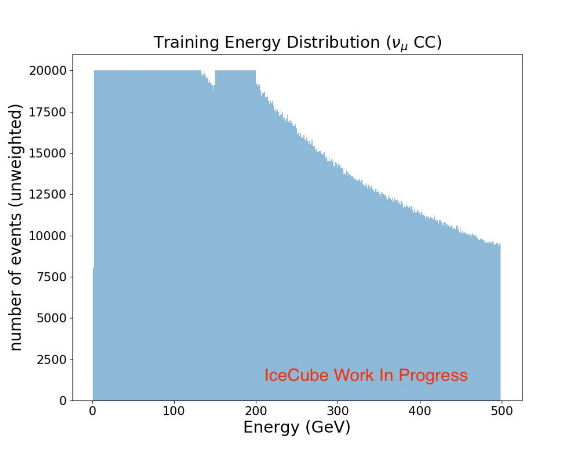
- ❖ CNN for Energy Reconstruction
  - ➢ Flat energy distribution in region of interest
    - ■ 1-200 GeV target
    - ■ Extended to 500 GeV
  - ➢ Use ν_μ CC events only
- ❖ Track vs Cascade (Flavor) Classifier
  - ➢ 50% Track and 50% Cascade per GeV
  - ➢ Includes all ν_μ , ν_e , CC, and NC



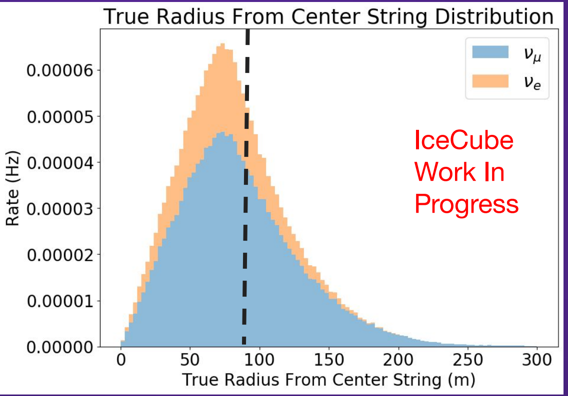
→ Details of other training samples in backup

Training Samples Optimized Per Variable

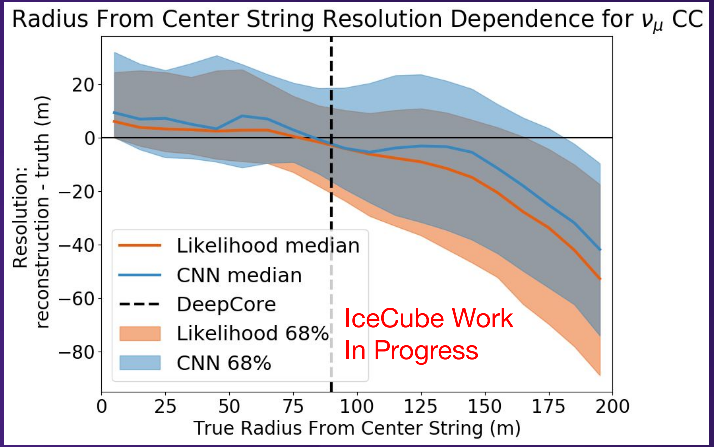
- Regressions trained with ν_μ CC events only
  - **Energy:** flat energy sample in region of interest
    - 1-200 GeV target
    - Extended to 500 GeV
  - **Zenith:** flat zenith sample across all angles
  - **Interaction Vertex:** uses flat energy sample
- Classifications trained with balanced samples
  - **Track vs Cascade**: 50/50 Track vs. Cascade
    - Includes ν_μ and ν_e ; CC and NC
  - **Muon vs Neutrino**: 40% μ and 40% ν_μ and 20% ν_e

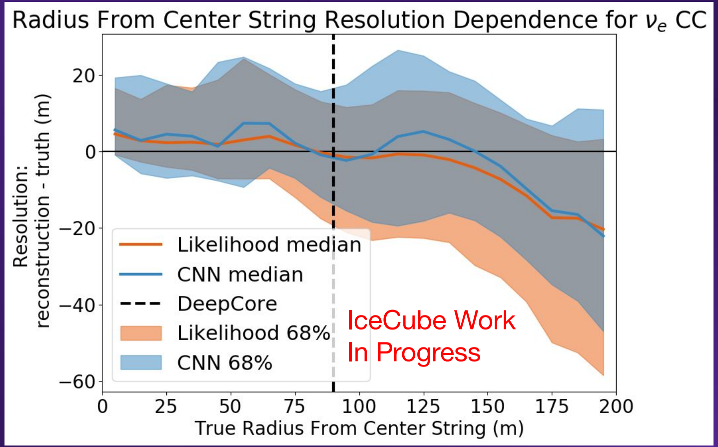


CNN Resolution: Interaction Radius

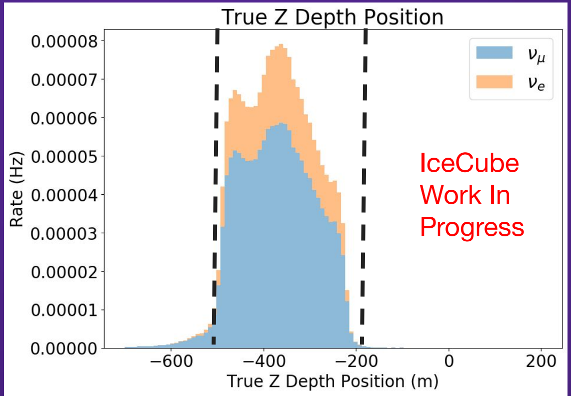


- Resolution best inside of DeepCore detector (roughly at dotted lines)
- Similar loss of resolution outside of region as likelihood method
- Using CNN Radius prediction as cut





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