Machine Learning for Particle Image Neutrino Detectors at the HEP Intensity Frontier

Kazuhiro Terao & Laura Domine
SLAC National Accelerator Lab.
ACAT @ Saas-Fee (11-15 Mar. 2019)
Outline:

- Neutrino detectors
- Machine learning applications
- Toward 3D ML-based data reconstruction
Me: Neutrino Physicist

- Neutrinos?
  - One of least understood elementary particles
Me: Neutrino Physicist

- **Neutrinos?**
  - One of least understood elementary particles
  - **They are everywhere**
    - 400 trillion neutrinos pass your body every second
    - Your body generates ~340 million neutrinos a day
  - They come from everywhere

![Image of neutrinos from various sources: Big Bang, SuperNova, AGN, Sun, Atmospheric, Earth, Accelerator, Reactor, Good Stuff.]

Inverse Beta Decay (IBD)

\[ \nu_e + p \rightarrow e^+ + n \]

by Reines & Cowan (Nobel Prize 1995)

First neutrino detection

Cd-doped water
0.4 ton, 100 PMTs
(1956)
Early days particle imaging

Bubble Chamber
Analog photographs to record trajectory of charged particles

The 'Neutrino Event'
Nov. 13, 1970 — World's first observation of a neutrino in a hydrogen bubble chamber
5ms of data at the NOvA Far Detector
Each pixel is one hit cell
Color shows charge digitized from the light

A 603MeV muon in Super-K.

Several hundred cosmic rays crossed the detector
Need for advanced algorithms for analyzing high resolution data with complex topologies. (goal: maximize physics output)
Hi-Res Particle Imaging

μBooNE

• High resolution photograph of charged particle trajectories
• Calorimetric measurement + scalability to a large mass

Liquid Argon Time Projection Chamber

~mm/pixel spatial resolution
~MeV level sensitivity

μ

Bubble Chamber

MicroBooNE
~87 ton (school bus size)
Topological shape difference is a major distinction for “shower” particles.
Trajectory ends are distinct, and useful for seeding particle clustering and trajectory fitting.
Many, local kinks caused by Multiple Coulomb Scattering process can be used for momentum estimation.
Small branches on muon-like trajectories are knocked-off electrons, useful key for the direction.
Energy deposition patterns ($dE/dX$) vary with particle mass & momentum, useful for analysis.
Energy deposition patterns (dE/dX) vary with particle mass & momentum, useful for analysis.
Hi-Res Particle Imaging

Cosmic Data: Run 6280  Event 6812  May 12th, 2016
Hi-Res Particle Imaging

Interaction vertex can be anywhere in LAr, varying in size (cm ~ meters)

Cosmic Data: Run 6280 Event 6812 May 12th, 2016
Hi-Res Particle Imaging

3D imaging LA (on-going R&D)
Machine Learning and Computer Vision
How to write an algorithm to identify a cat?

... very hard task ...
Development Workflow for non-ML reconstruction

1. Write an algorithm based on physics principles

A cat = collection of certain shapes
(or, a neutrino)
Development Workflow for non-ML reconstruction

1. Write an algorithm based on physics principles
2. Run on simulation and data samples
3. Observe failure cases, implement fixes/heuristics
4. Iterate over 2 & 3 till a satisfactory level is achieved
5. Chain multiple algorithms as one algorithm, repeat 2, 3, and 4.

A cat = collection of certain shapes
(or, a neutrino)

Partial cat (escaping the detector)

Stretching cat (Nuclear FSI)
Development Workflow for non-ML reconstruction
1. Write an algorithm based on physics principles
2. Run on simulation and data samples
3. Observe failure cases, implement fixes/heuristics
4. Iterate over 2 & 3 till a satisfactory level is achieved
5. Chain multiple algorithms as one algorithm, repeat 2, 3, and 4.

Machine Learning

• “Learn patterns from data”
  - automation of steps 2, 3, and 4

• “Chain algorithms & optimize”
  - step 5 addressed by design
Convolutional Neural Networks (CNNs)

CNNs are effective image feature extractors, and also data transformers.
CNN for “Image Classification”

Machine Learning in Neutrino Physics

NOvA Neutrino Event Topology

MicroBooNE Signal/Background

LArLIAT Signal vs. Background

NEXT Signal vs. Background
WHAT is WHERE in an image?
Machine Learning in Neutrino Physics

Image Context Detection

Mask R-CNN
arXiv:1703.06870
WHAT is WHERE and HOW in an image?
Machine Learning in Neutrino Physics

Interpretation of Contexts’ Correlation

"girl in pink dress is jumping in air."

NeuralTalk
github:karpathy/neuraltalk2
Object Detection for Neutrino Finding (MicroBooNE LArTPC)

Task: propose a rectangular box (location & size) that contains neutrino interaction.
Beyond Image Classification
Machine Learning in Neutrino Physics

Semantic Segmentation
- Recently published ... [arXiv:1808.07269](http://arxiv.org/abs/1808.07269)
- Pixel-level object classification
  - Separation of EM-particle from other types
  - Key input information for particle clustering
- First time deep neural network validated on LArTPC data
Pixel-level Feature Information
Machine Learning in Neutrino Physics

Region 1

Region 2

Region 3

MicroBooNE Data
Localized features at the pixel-level are useful to inspect correlation of data features & algorithm responses.
Localized features at the pixel-level are useful to inspect correlation of data features & algorithm responses.
ML-based 3D Data Reconstruction
Multi-Task Network Cascade

- **Chain of Segmentation + Detection**
  - Feature points: “shower start” and “track edges”
  - Classify each pixel into “shower” vs. “track”

- **Extension to 3D data**
  - Change in tensor dimensions, identical algorithms

**3mm/pixel resolution**

**Pixel distance** between the target truth point to the closest proposed point
Multi-task DNN for Physics Reconstruction
Introduce physical feature extraction tasks (auxiliary targets) to bias the data transformation path to support producing a logical conclusion. Optimize the whole reconstruction chain.
Data feature: generally sparse, locally dense image, and very large volume (1 E10-20 pixels)

Issues using standard CNNs

- **Inefficient** calculations ("zero" matrix elements)
- **Prohibitive resource** usage (memory, time)
- **Degraded performance**

... terrible scaling = garbage!
**Data feature**: generally sparse, locally dense image, and very large volume (1 E10-20 pixels)

**Got a solution :)**
(right: $768^3$ volume)

**Submanifold Sparse Conv. Net**
Talk by Laura Domine (Thursday)
Great for LArTPC and other domains!

<table>
<thead>
<tr>
<th>Type</th>
<th>HIP</th>
<th>MIP</th>
<th>Shower</th>
<th>Delta</th>
<th>Michel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc.</td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
<td>0.97</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Michel Electron Reconstruction

- Run spatial DBscan for MIP & Michel pixels (separately)
- Keep only Michel clusters which edge touches with an edge of a MIP cluster

Preliminary
- ID efficiency = 93.9%
- ID purity = 98.8%

Deep CNN for LARGE Detectors? (scalability)
Machine Learning for Particle Image Analysis

MicroBooNE Michel electron paper reports efficiency of 2% with purity of 80-90%, tuned for high purity (calibration)
Reproducible technique sharing is important...

- Submanifold Sparse Conv. Net for scalability
  - See Laura’s talk, and our benchmark ... arXiv: 1903.05663
  - Open data sample: DOI 10.17605/OSF.IO/VRUZP
  - Software stuck: Singularity or Docker container
  - Implementation: github repo

Toward HPC: contact them if you want help!

- SSCN + Horovod + custom MPI for production
- Corey Adams (ANL)
  - KNL/GPU nodes @ ALCF
- Eric Church, Jan F Strube, Alexander R. Hagen (PNNL)
  - SummitDev Intel Power8, now moving onto Power9
Experimental neutrino physics:

- **Detector trend: particle imaging**
  - LArTPC is the current frontier for imaging

- **Many applications from computer vision**
  - ML-based full data reconstruction being developed
  - Active but not mentioned: data/sim domain adaptation

- **Next few years**
  - Integration of ML-based reconstruction
  - Data/Simulation domain adaptations
  - Software stack development toward HPC
Thank you for listening and Thank YOU for organizing ACAT2019!
Back Up Slides
How image classification works
How image classification works

Intermediate Data Tensor
(low-resolution, high-level features)

down-sampling (encoding)

“Human Face”
How pixel segmentation works

- Combine “up-sampling” + convolutions
- Output: “learnable” interpolation filters

Intermediate Data Tensor
(low-resolution, high-level features)
How pixel segmentation works

- Combine “up-sampling” + convolutions
- Output: “learnable” interpolation filters

Intermediate Data Tensor
(low-resolution, high-level features)

Concatenation recovers spatial resolution information
Parasitic multi-task scheme for point prediction network (PPN) on U-ResNet ... 2D/3D agnostic

Concatenation recovers spatial resolution information
Relabeling study

*Relabeled dataset = only primary ionization is labelled as Michel electrons*

<table>
<thead>
<tr>
<th>Train data</th>
<th>Regular</th>
<th>Relabeled</th>
<th>Relabeled + Weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIP</td>
<td>98.0%</td>
<td>98.1%</td>
<td>98.1%</td>
</tr>
<tr>
<td>MIP</td>
<td>99.4%</td>
<td>99.2%</td>
<td>99.4%</td>
</tr>
<tr>
<td>EM shower</td>
<td>99.4%</td>
<td>97.9%</td>
<td>99.2%</td>
</tr>
<tr>
<td>Delta rays</td>
<td>85.7%</td>
<td>94.8%</td>
<td>96.0%</td>
</tr>
<tr>
<td>Michel electrons</td>
<td>56.6%</td>
<td>94.4%</td>
<td>94.7%</td>
</tr>
</tbody>
</table>
Number of pixels in candidate vs matched Michel cluster
Michel electrons energy spectrum reconstruction

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample size</strong></td>
<td>7105</td>
</tr>
<tr>
<td><strong>Identification purity</strong></td>
<td>98.8%</td>
</tr>
<tr>
<td><strong>Identification efficiency</strong></td>
<td>93.9%</td>
</tr>
<tr>
<td><strong>Cluster efficiency</strong></td>
<td>96.1%</td>
</tr>
<tr>
<td><strong>Cluster purity</strong></td>
<td>97.8%</td>
</tr>
</tbody>
</table>

First ML-based approach. Simulation only, next step is data!
Detectors
Detecting Neutrinos: BMB

We cannot observe neutrinos, but we can detect particles that come out of a neutrino interaction.
Evolution of Detectors

Inverse Beta Decay (IBD)

$\bar{\nu}_e + p \rightarrow e^+ + n$

by Reines & Cowan (Nobel Prize 1995)

Cd-doped water
0.4 ton, 100 PMTs

First neutrino detection
Evolution of Detectors

Cd-doped water
0.4 ton, 100 PMTs
(1956)

Birth of neutrino astrophysics!

KamiokaNDE Detector
3 kton ultra-pure water, 1000 20” PMTs
(shared Nobel Prize 2002)
Neutrino Oscillation Experiments (I)

Evolution of Detectors

Cd-doped water
0.4 ton, 100 PMTs (1956)

Ultra-pure water
3 kton, 1000 PMTs (1983)

Discovery of $\nu_{\text{atmo}}$ oscillation!

Super-KamiokaNDE
50 kton ultra-pure water, 11000 PMTs
(shared Nobel Prize 2015)
Evolution of Detectors

Cd-doped water
0.4 ton, 100 PMTs
(1956)

Ultra-pure water
3 kton, 1000 PMTs
(1983)

Ultra-pure water
50 kton, 11000 PMTs
(1996)

Discovery of $\nu_{\text{solar}}$ oscillation!

SNO
1 kton heavy-water Cherenkov,
9600 PMTs
(shared Nobel Prize 2015)
Evolution of Detectors

- **Cd-doped water**
  0.4 ton, 100 PMTs
  (1956)

- **Ultra-pure water**
  3 kton, 1000 PMTs
  (1983)

- **Ultra-pure water**
  50 kton, 11000 PMTs
  (1996)

- **Heavy water**
  1 kton, 9600 PMTs
  (1999)

**Reactor neutrino oscillation!**
(the solar model is right!)

KamLAND
1 kton liquid scintillator, 1900 PMTs
My first neutrino experiment
(undergraduate RA @ UC Berkeley)
Evolution of Detectors

- **Cd-doped water**
  - 0.4 ton, 100 PMTs
  - (1956)

- **Ultra-pure water**
  - 3 kton, 1000 PMTs
  - (1983)

- **Ultra-pure water**
  - 50 kton, 11000 PMTs
  - (1996)

- **Heavy water**
  - 1 kton, 9600 PMTs
  - (1999)

- **Liquid Scintillator**
  - 1 kton, 1900 PMTs
  - (2002)

- **Gd-doped liquid scintillator**
  - RENO, Daya Bay, Double Chooz

- **“Near” & “Far” design**
  - 2 x 16 ton detectors with 400 PMTs each (Double Chooz)

- **My Ph.D thesis!** (MIT)
  - “Last mixing angle” $\theta_{13}$ Experiments!
Neutrino Oscillation Experiments (I)

A 603 MeV muon in Super-K.
Neutrino Oscillation Experiments (I)

A 492MeV electron in Super-K.

It's a bit fuzzier.
Less topological information but excellent energy resolution
Neutrino Oscillation Experiments (II)

How can we do better?

Three important detector features for oscillation measurement

\[ P(\nu_\mu \rightarrow \nu_e) = \sin^2 2\theta \sin^2 \left( \frac{1.27 \Delta m^2 L}{E_\nu} \right) \]

**Good Energy Resolution**
- Precise \( E_\nu \) reduce oscillation uncertainty

**Large Mass (scalable)**
- “More” statistics to measure rare physics process

**Particle ID Capability**
- Better \( \nu \) identification background rejection
Challenge

S
Analysis
Challenges

There may be lots of backgrounds

 Cosmic Data: Run 6280 Event 6812 May 12th, 2016
Interaction vertex can be anywhere in LAr, varying in size (cm ~ meters)

Cosmic Data: Run 6280 Event 6812 May 12th, 2016
Analysis Challenges

Must identify event vertex + neutrino interaction topology (particle types)
Cluster energy depositions for an accurate calorimetry
Deal with optical illusions in 2D projections + 3D pattern recognitions