Incorporating Advances in Machine Learning for Reconstruction in T2K and Super-Kamiokande Experiments

Félix Cormier
28/05/2024
CAP Congress - PPD Session
Super-Kamiokande

SuperK is a Water Cherenkov detector

- Large 50 kton water tank lined with PMTs optimal for detecting neutrino interactions
- Acts as far detector for T2K long-baseline experiment
- Can also detect solar, atmospheric or supernova neutrinos
- Cherenkov light from neutrino interaction products can lead to particle identification, reconstruction
The WatChMaL Collaboration

- The Water Cherenkov Machine Learning Collaboration is a cross-experiment group
- Common data generation, pre-processing and training frameworks are shared by members
- Currently have members in experiments
  - T2K/SuperK
  - Water Cherenkov Test Experiment (WCTE)
  - Intermediate Water Cherenkov Detector (IWCD)
  - Hyper-Kamiokande

See talk Friday by B. Jamieson

Talk Monday by S. Yousefnejad
Water Cherenkov data

- As products of neutrino interactions travel in water, they produce Cherenkov rings
  - These are imaged by the PMTs
- Products are often electrons (muons) from $\nu_e$ ($\nu_\mu$) interactions
  - Electrons will produce larger rings due to multiple interactions and showering
  - Muons (and $\pi^+$) will often have thinner rings
- Data is in the form of integrated charge, time of individual PMTs
Data generation

- Data generation is important for training deep learning networks in order to avoid biases
- We generate data using official SuperK simulation software, SKDETSIM
  - Generate samples of electrons, muons and $\pi^+$
    - **Energy is uniformly sampled** from 0 to 1 GeV above Cherenkov threshold
    - Uniform in vertex position & direction
• Currently high energy event reconstruction in SuperK is done using the \textit{fitQun} algorithm
  • This algorithm depends on likelihood minimization, using PMT charges and time to construct the likelihood function
  • Makes assumptions about what the data will look like
  • Calculation complexities make this algorithm difficult to extend to e.g. different particle hypotheses
  • Is very slow to compute - hard to scale
• By contrast, machine learning algorithms can learn this complexity by training over simulated data
  • Can e.g. speed up calculations by over order of magnitude
ResNet Architecture

- **ResNet**
  - Convolutional layers over 2D image-like input PMT data & subsequent layers build features
  - Residual connections between layers help with vanishing gradients

![Diagram Illustrating ResNet Architecture](image)

- PMT Charge
- PMT Time

\[ F(x) + x \]

\[ F(x) \]

weight layer

relu

identity

Figure 2. Residual learning: a building block.
Processing Data for ResNet

- Tests over many Water Cherenkov detectors have shown **ResNet to have better performance** at both classification and regression.

- One of the challenges of ResNet is to project 3D cylindrical data from SuperK into a 2D image.

- Each PMT is linked to its 3D position in space using a dictionary.
  - Each 3D position is then unrolled into a 2D image, with each PMT becoming a pixel.
Classification Networks

• Electrons and muons can be easily distinguished in most cases
  • Classical algorithms **already have very high accuracy** in distinguishing these events

• Muons and $\pi^+$ are more difficult to distinguish
  • Cherenkov rings from the initial particles virtually indistinguishable
  • Algorithm must use subtle pion hadronic interactions to distinguish between the events
  • Increased classification performance could improve T2K muon disappearance results
    • Potential to impact mass hierarchy measurements
Classification Results

- Run a 3-class classification network using ResNet to classify between \( e, \mu, \pi^+ \)
- Show results for \( \mu \) (signal) vs. \( \pi^+ \) (background)
- Can scan across all the class outputs to see how background rejection and signal efficiency vary
  - 100x **better background rejection** at same signal efficiency for \( \mu \) vs. \( \pi^+ \)
  - Currently studies ongoing to understand what the network is learning for \( \mu \) vs. \( \pi^+ \) that fITQun did not learn - hadronic interactions of the charged pion?

![Graph showing the comparison between ResNet (AUC=0.9702) and fITQun in terms of muon tagging efficiency and background rejection.](image)
Regression Networks

• An important part of event reconstruction is finding the particle’s initial **position, direction** and **momentum**
  • Position and direction reconstruction performance can increase efficiency of cuts based on detector location
  • Momentum reconstruction can help reconstruction of initial neutrino energy
• Better resolution in these kinematic variables can have large effects on efficiency during analysis
• Furthermore, better resolution and smaller bias of the reconstruction algorithms can reduce systematic uncertainties
Regression Analysis

- We construct & train 6 individual networks to reconstruct
  - Electron: position, direction, momentum
  - Muon: position, direction, momentum
- To gauge performance of each network, calculate the residuals of the reconstructed against true value
  - For position: look at 3D distance between true and reconstructed
  - For direction: look at angle between true and reconstructed unit vectors
  - For momentum: look at residual percentage: \((p_{true} - p_{reco})/p_{true}\)
- Calculate
  - **Resolution**: 68.3% quantile of the residual
  - Bias: median of the residual
Regression Results

- Can look at position along longitudinal, transverse direction with respect to true particle direction
  - See that getting better transverse resolution very hard
- Can also analyze regression results as function of underlying variables
  - Visible energy - energy above Cherenkov threshold
  - Towall - distance from initial vertex to detector wall, along true particle direction
- Overall: **10-100% improvement** in resolution depending on particle/variable
Other WatChMaL Results

IWCD

Above 200 MeV
~99.9% electron efficiency with 0.1% muon mis-ID

~ 6 cm electron position resolution

~ 2.6° electron direction resolution

~ 5 cm muon position resolution

~ 1.4° muon direction resolution

~ 70% separation of electrons and gammas converting to e⁺/e⁻ pair

~ 6% electron momentum resolution

~ 2% muon momentum resolution
Conclusions & Next Steps

• As part of the WatChMaL Collaboration, applied Deep Learning techniques for Super-Kamiokande event reconstruction

• Use the patterns of PMT charge and time to learn underlying phase space for both classification and regression
  • Classification between electrons, muons and pions shows great improvement over classical reconstruction methods
  • Regression for vertex position, direction and momentum show 10-100% improvement in resolution, and large reductions in bias for most cases

• WatChMaL analyses on other Water Cherenkov detectors ongoing, also showing good performance

• Next steps will include studying the role of adversarial training on systematic uncertainty reduction
Thank you!
Today will show current results for classification and regression:
- Classification: 3-class, e/μ/π+
- Regression: position, direction and momentum for e/μ

Data sets are simulated such that:
- Isotropic position, direction in SuperK
- Uniform visible energy distribution 0-1 GeV

Training uses the ResNet34 architecture:
- Inputs are all PMT charges, times flattened onto 2D ‘image’
Classification training

- Classification returns a score between 0 and 1 for each class (Softmax)
  - Electron/muon/π+
  - 3 scores per event which add up to 1
  - Here can see muon/π+ separation

![Score vs Events Diagram](Image)
Training - ResNet

- Based on prior HyperK experience & resSIM, the ResNet ML architecture provided the best performance.
- ResNet is a CNN which requires input data to be projected to a 2D map.
  - See examples of this projection on the plots here.
  - Decision on how to unroll cylinder arbitrary - see later for how we deal with this.
Double cover & Transforms

• One way which was found to work well to remove effects due to arbitrary choice of cylinder unrolling is double cover padding

```
01            CBALKJIHGFED
23            01  32
ABCDEF       -->  DEFGHIJKLABC
GHIJ         MNOPQRSTUVWX
KLM          PQRSTUWXYZMN
01            23  10
32            45  76
54            67  54
ONMXWVUSTRQP
```

• Another augmentation to the data is transforms, where we flip horizontally, vertically and do a front/back reflection

• Allows us to artificially inflate data, leading to higher stats
Position & Direction Regression

- Loss per iteration plots for position and direction
  - Each iteration is a batch which is ~100 events
  - Plots corresponds to ~10 epochs
- Shown is training loss and validation loss
  - Best is when validation loss is best yet during training
The graph shows the muon rejection as a function of electron tagging efficiency. The blue line represents the ResNet model with an AUC of 0.9968, and the red dot indicates the fitQun data point.
• Muon/Pi+ Classification performance as function of truth visible energy
  • Truth visible energy defined as initial particle energy over Cherenkov threshold
Electron - Momentum

Event Vertex for Global Axis

- \( \mu = 0.00626 \pm 0.0005 \) [MeV]
- \( \text{Quat.} = 0.04724 \pm 0.0004 \) [MeV]

Momentum Bias [%]

- ML
- \( \text{STQun} \)

Momentum Resolution [%]

- ML
- \( \text{STQun} \)
Electron - Position

Event Vertex for Longitudinal Axis

- ST2K: Electrons
  - $\mu = 0.27 \pm 0.1$ [cm]
  - Quant. = 17.50 $\pm$ 0.1 [cm]

Event Vertex for Transverse Axis

- ST2K: Electrons
  - $\mu = 2.46 \pm 0.03$ [cm]
  - Quant. = 7.06 $\pm$ 0.04 [cm]

Event Vertex for Longitudinal Axis

- ML: Electrons
  - $\mu = 0.03 \pm 0.06$ [cm]
  - Quant. = 12.45 $\pm$ 0.08 [cm]

Event Vertex for Transverse Axis

- ML: Electrons
  - $\mu = -0.14 \pm 0.03$ [cm]
  - Quant. = 6.84 $\pm$ 0.04 [cm]
Electron - Direction

Event Vertex for Angle Axis

$\text{fTQun}_{\text{Electrons}}$
$\mu = 1.93 \pm 0.03$
Quant. = $2.96 \pm 0.03$

Event Vertex for Angle Axis

$\text{ML}_{\text{Electrons}}$
$\mu = 1.54 \pm 0.02$
Quant. = $2.31 \pm 0.02$
Muon - momentum

Event Vertex for Global Axis

Muon - momentum

Event Vertex for Global Axis

Event Vertex for Global Axis
Muon - position

Event Vertex for Global Axis

Event Vertex for Longitudinal Axis

Event Vertex for Transverse Axis

Event Vertex for Global Axis

Event Vertex for Longitudinal Axis

Event Vertex for Transverse Axis

Muon

fitQun_Muons

μ = 14.51 ± 0.2 [cm]

Quant. = 19.55 ± 0.1 [cm]

fitQun_Muons

μ = -0.86 ± 0.2 [cm]

Quant. = 13.55 ± 0.1 [cm]

fitQun_Muons

μ = -2.30 ± 0.1 [cm]

Quant. = 8.94 ± 0.07 [cm]

ML_Muons

μ = 12.87 ± 0.08 [cm]

Quant. = 17.27 ± 0.1 [cm]

ML_Muons

μ = -0.29 ± 0.07 [cm]

Quant. = 9.01 ± 0.07 [cm]

ML_Muons

μ = 0.07 ± 0.07 [cm]

Quant. = 9.97 ± 0.09 [cm]

May 28th 2024

Félix Cormier - CAP Congress (PPD)
Muon - position - visible energy

May 28th 2024
Muon - direction

Event Vertex for Angle Axis

- fitQun_Muons
  - $\mu = 1.38 \pm 0.07$
  - Quant = 1.97 $\pm 0.01$

- ML_Muons
  - $\mu = 1.19 \pm 0.06$
  - Quant = 1.64 $\pm 0.01$

May 28th 2024