Neutrino Energy Reconstruction with a Regression CNN in DUNE

Neutrino Physics and Machine Learning (NPML)

Wenjie Wu (For the DUNE collaboration)

University of California, Irvine

Jul. 22, 2020
Introduction

➤ DUNE is a long baseline neutrino experiment

➤ Primary physics goals
  ➤ Measuring the oscillation patterns of $\nu_\mu$ and $\bar{\nu}_\mu$
    - CP-violating phase
    - Neutrino mass ordering, the octant of $\theta_{23}$, etc.

➤ Supernova burst neutrinos

➤ BSM processes (baryon number violation, NSI, etc.)

Neutrino Energy Reconstruction
  * Kinematics-based method
  * Regression Convolutional Neural Network (CNN)
Kinematics-based method

\[ E_\nu = E_{\text{cor lep}} + E_{\text{cor had}} \]

- \( \nu_e \) CC energy: divide event into reconstructed shower with highest charge and hadronic energy
- \( \nu_\mu \) CC energy: divide event into longest reconstructed track and hadronic energy
- Hadronic/Electron energy: electron lifetime (wire-by-wire) and recombination (constant) corrected calorimetric energy
Kinematics-based method

Electron shower energy
Calorimetric energy calibrated with MC

Muon momentum (Longest track contained)
By track range, calibrated by MC

Muon momentum (Longest track exiting)
By multi-Coulomb scattering, calibrated by MC

Hadronic energy
By reconstructed hits not in the muon track or electron shower, calibrated by MC

Nick Grant, Tingjun Yang, DPF2017
Energy reconstruction has many challenges due to missing energy caused by argon impurities, nonlinear detector energy responses, invisible energy, hadron identities (mass), and overlaps between lepton and hadron interactions.

Regression CNN:

- Neural networks have shown state of the art performance in HEP classification and regression tasks in problems with high dimensionality.
- Use a convolutional neural network directly on the pixel maps.
Data Generation

- In the DUNE FD module, three wire readout planes collect the ionization charge that is generated when charged particles traverse the liquid argon volume.
- Raw detector waveforms are deconvolved to obtain the charge information.
- The position of the charge observed in each of the three planes is combined with the drift time to create three views of each interaction.
Regression CNN Architecture

- Architecture modified from UCI’s NOvA Regression CNN energy estimator (Phys. Rev. D. 99.012011)

- Mean absolute percentage error

\[
L \left( W, \{ x_i, y_i \}_{i=1}^n \right) = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{f_W(x_i) - y_i}{y_i} \right|
\]

- Images are concatenated as channels to allow for extra convolutional layers afterwards

- All convolutional layers use ReLU

- Hyper-parameters are not fully optimized due to computational constraints

Inception: https://arxiv.org/abs/1409.4842
Input Images for $\nu_e$

- Use ADC counts and TDC units from Wire instead of using the reconstructed hits
- Three pixel maps: 280x400
- Merged 6 TDC ticks: real covered space is 1680 ticks and 400 wires
  - Make the same physical dimensions of the x- and y-axis
- The pixel map size is chosen to contain 90% of hits on average
ν _e_ CC Energy Reconstruction

Energy resolution

- Applied the trained model to the official ν _e_ MC samples
- Fiducial volume cut is applied with true vertex information
- Fit with Gaussian in the range (-1, 1)
  - Kinematic-based method: σ = 13.1 %
  - RegCNN: σ = 7.2 %
$\nu_e$ CC Energy Reconstruction

Energy resolution vs. True energy

- Mean and RMS of each True energy bin
- RegCNN has smaller RMS over the energy range (0, 6) GeV
- RegCNN over-estimates for low energies due to low statistics
- Best solution: use flat flux energy spectrum to enrich low energy neutrinos in the training

![Graphs showing energy resolution vs. true energy]
Since flat flux sample not available, re-weight individual events to give low energy events larger impacts in the training

Re-weight samples in the loss function

\[
L \left( W, \{x_i, y_i\}_{i=1}^n \right) = \frac{1}{\sum_j^N \omega_j} \sum_i^n \omega_i L \left( W, x_i, y_i \right)
\]

If the weights are highly imbalanced, this can impact the efficiency of the stochastic gradient descent

Instead, sample \((x_i, y_i)\) with probability:

\[
p_i = \frac{\sqrt{\omega_i}}{\sum_j^N \sqrt{\omega_i}}
\]

Use a loss function with weights:

\[
L \left( W, \{x_i, y_i\}_{i=1}^n \right) = \frac{1}{\sum_j^N \sqrt{\omega_j}} \sum_i^n \sqrt{\omega_i} L \left( W, x_i, y_i \right)
\]
$\nu_e$ CC Energy Reconstruction

Weighted training and result

- Similar energy resolution: 7.2% → 7.3%
- Reduced bias in low energy region
$\nu_e$ CC Energy Reconstruction

Energy resolution with different interaction modes

- RegCNN shows good performance for different interaction modes
- Fit with Gaussian in the range (-1, 1)

<table>
<thead>
<tr>
<th></th>
<th>RegCNN</th>
<th>Kinematic</th>
</tr>
</thead>
<tbody>
<tr>
<td>QE</td>
<td>5.3%</td>
<td>9.5%</td>
</tr>
<tr>
<td>Res</td>
<td>8.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>DIS</td>
<td>9.4%</td>
<td>15.2%</td>
</tr>
</tbody>
</table>
Input Images for $\nu_\mu$

- Three pixel maps: 280x400
- For $\nu_\mu$ interactions, leptonic portion is characterized by very long $\mu$-tracks
- In order to contain most of the hits, further lower the image resolution
  - Merged 24 TDC ticks and 7 wires: real covered space is 6720 ticks and 2800 wires
As a first step, performed the reconstruction for events with contained tracks

RMS of kinematic-based method: 19.0% and RegCNN: 12.5%

Moving to study events with exiting muon track
Summary

- Developed regression CNN models to reconstruct $\nu_e$ CC and $\nu_\mu$ CC events
- RegCNN shows promising results with better energy resolution for both neutrinos
- For $\nu_e$ CC: 13.1%→7.3%, for contained $\nu_\mu$ CC:19.0%→12.5%
- In the near future:
  - Train a model on un-contained $\nu_\mu$ CC events
  - Eliminate energy dependence of bias by re-weighting during training