Regression CNN based Neutrino Energy Reconstruction

Lars Hertel, Ilsoo Seong, Jianming Bian
University of California, Irvine

2019/01/29

DUNE collaboration meeting, CERN
Introduction

Neutrino Energy reconstruction at DUNE:

• Crucial because oscillations and differential cross-sections depend on it
• Challenging due to complexity in detector response and final state particle kinematics.
• Tradition method: Kinematics-based Energy Reconstruction:
  \[ E(\nu) = E_{\text{lep}}^{\text{cor}} + E_{\text{had}}^{\text{cor}} \]

• This talk: Regression Convolutional Neural Network (CNN) based Energy Reconstruction: deep learning with direct waveform inputs from U/V/Y planes
Kinematics-based $\nu_\mu$ CC and $\nu_e$ CC Total Energy in DUNE LArSoft

Nick Grant, Tingjun Yang, DPF2017

True $\nu_\mu$ CC events with contained track

True $\nu_\mu$ CC events with exiting track

Estimate muon momentum from multi-Coulomb scattering (MCS).

Fiducial cuts based on true vertex, at least one track/shower required

<table>
<thead>
<tr>
<th></th>
<th>$\nu_\mu$ CC</th>
<th>$\nu_e$ CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Longest reco track (contained)</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>Longest reco track (exiting)</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>Reco shower with highest charge</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>Hadronic energy</td>
<td>39</td>
<td>49</td>
</tr>
</tbody>
</table>
Regression Convolutional Neural Network for Neutrino Energy

- Kinematics-based method complicated by event topology, invisible energy and identities (mass) of hadrons, shower/track reconstruction quality etc.
- Convolutional Neural Networks (CNNs) with raw pixel inputs have demonstrated success in Classification problems such as event identification (CVN identifier in NOvA and DUNE, image segmentation/prong identifier at MicroBooNE, vertices plane identifier at MINERvA)
- **Regression** CNN solves continuous variables, haven’t developed in HEP
- Developed **Regression** CNN based method for energy reconstruction at DUNE
- Can also be extended to solve other reconstruction in DUNE

*CNNs take raw pixel inputs, using all detector information with acceptable computing cost.*

**Convolutional neural network**

Traditional artificial neural network
Charge Information: Direct waveform inputs

We use ADC counts and TDC units from Wire instead of using the reconstructed hits

The hits from “gaus hit” sometimes fail to correctly reconstruct hits

Goodness of Fit shows that minimum 4% of hits per event correspond to such hits
Pixel Map inputs

- Three input pixel maps: U-T, V-T, and Z-T
- Pixel map size has been chosen to contain 90% of hits on average
- Coarsed TDC ticks to make same physical dimensions of the x- and y-axis of the pixel map
- Pixel map size: 280x400 (actual covered space: 1680 ticks x 400 wires) → 6 ticks are merged
CNN Architecture

- Architecture modified from UCI’s NOvA Regression CNN energy estimator (Pierre Baldi, Jianming Bian, Lars Hertel, Lingge Li, PhysRevD.99.012011)
- Loss: \( L(W, \{x_i, y_i\}_{i=1}^n) = \frac{1}{n} \sum_{i=1}^n \left| \frac{f_W(x_i) - y_i}{y_i} \right| \) Optimize energy resolution and reduce impacts from outliers.
- One linear output unit
- No regularization applied
- Hyperparameter optimization software SHERPA used

Plane U
- 3xConv2d
- Max Pool
- Inception
- Max Pool
- Inception
- Max Pool
- Inception
- Max Pool
- Inception
- Max Pool
- Average Pool
- Energy

Plane V
- 3xConv2d
- Max Pool
- Inception
- Max Pool
- Inception
- Max Pool
- Inception
- Max Pool
- Inception
- Max Pool
- Average Pool
- Energy

Plane Z
- 3xConv2d
- Max Pool
- Inception
- Max Pool
- Inception
- Max Pool
- Inception
- Max Pool
- Inception
- Max Pool
- Average Pool
- Energy
NueCC Energy Resolution

- Training with 1M events, without energy dependent weight
- Applied the model into the Nue MCC10
- Fiducial cuts based on true vertex, at least one track/shower required
- Fit with Gaussian in (-1,1)
- Sigma of RegCVN: 7.1%, Std.: 13.1%

![Graph showing energy resolution](image)

Fiducial cuts based on true vertex, at least one track/shower required
Fiducial cuts: true vertex $\rightarrow$ reco hits location

Regression CNN could solve uncontained energy

Fit with Gaussian close to peaks

Sigma of RegCVN: 5.8%, Std.: 8.0%

Fiducial cuts based on reco hits location,
At least one track/shower required
Energy Resolution vs True Energy

- Mean and RMS of energy resolution
- Regression CNN shows less bias than and smaller RMS
- Still need improvement in the low energy region → weighted training
Weighted Training

- Network over-estimates for low energies.

- Bias due to low statistics in low energy neutrino in training data.

- Best solution: Use flat flux energy spectrum to enrich low energy neutrinos in the training.

- Since flat flux sample currently not available, re-weight individual events to give low energy events larger impacts in the training.
Weighted Training

Re-weigh samples in the loss function:

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

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:

\[
L(W, \{(x_i, y_i)\}_{i=1}^n) = \frac{1}{\sum_j^n \sqrt{\omega_j}} \sum_i^n \sqrt{\omega_i} L(W, x_i, y_i)
\]
Result of Weighted Training

- Similar energy resolution: 7.6% (Std.: 13.1%)
- With weighted training

Fiducial cuts based on true vertex,
At least one track/shower required
Energy resolution for different interaction modes

Regression CNN: Reconstructed to true energy ratio not sensitive to interaction modes:

**CNN Resolution:**  5.2% (QE), 8.3% (RES), 9.4% (DIS)

**Kinematic method:** 9.5% (QE), 13.1% (RES), 15.2% (DIS)

Weighted training
Fiducial cuts based on true vertex,
At least one track/shower required
Systematic uncertainties in Regression CNN $\nu_e$ CC Energy Reconstruction at NOvA

- Regression CNN systematic uncertainties from neutrino interactions evaluated by GENIE parameter reweighting at NOvA
- The regression CNN shows smallest systematic uncertainties from the simulation of neutrino interactions
- Varying calibration by 5% changes output $\nu_e$ CC Energy by 4.5%
- At DUNE, will use similar method and ProtoDUNE data to estimate systematic errors
\( \nu_\mu \) Energy

- At DUNE, the overall Numu resolution for from the traditional method is 20% in current DUNE LArSoft.
- Extending the regression CNNs to Numu muon energy reconstruction in NOvA and DUNE is highly needed.
- DUNE has a broad energy spectrum, and traditional (kinematic) methods have difficulty to reconstruct low-energy and uncontained muons in high-energy.
- Did a preliminary training for contained \( \nu_\mu \) Energy
- Working on uncontained \( \nu_\mu \) - The uncontained muon energy can only be estimated from scattering angles in multi-Coulomb scattering (MCS). Traditional fitting methods to determine the path lengths and random scattering angles in MCS are not accurate. These problems can be solved by the high-resolution direct pixel map inputs in regression CNNs.
Pixel Map inputs for $\nu_\mu$ Energy (contained)

- Three input pixel maps: U-T, V-T, and Z-T
- Pixel map size has been chosen to contain 90% of hits on average
- Coarse TDC ticks and wires
- Pixel map size: 280x400, (actual covered space: 6720 ticks x 2800 wires) → Merged 7 wires and 24 ticks
$\nu_\mu$ energy training (contained, preliminary)

- Training samples are generated with the true fiducial volume cut on x and z axes (contained)
- Small training samples: 240k
- Trained up to 25 epochs:
  - the val_loss does not improve due to the small sample we used
- The plot is after the true vertex cut
- RMS for contained $\nu_\mu$ energy: 19% $\rightarrow$ 12%
Summary

- $\nu_e$ CC/ $\nu_\mu$ CC regression CNN energy reconstruction developed at DUNE
- Promising $\nu_e$ CC energy resolution of 7% (13% from kinematics-based method)
- Contained $\nu_\mu$ CC energy training improve resolution from 19% → 12%
- Energy scale shows small dependence on true energy (with weighted training) and interaction modes.
- Developing regression CNN based uncontained $\nu_\mu$ and shower/track energy reconstruction, optimizing training for high resolution $\nu_\mu$ input pixel maps.
- Systematic uncertainty studies underway, will use protoDUNE data
Backup
Kinematics-based Energy Reconstruction

\[ E(\nu) = E_{\text{lep}}^{\text{cor}} + E_{\text{had}}^{\text{cor}} \]

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

Longest track contained within detector, calculate momentum by range

Using Monte Carlo, estimate reco track momentum as (range - intercept) / gradient.

Longest track exits detector, estimate its momentum from multi-Coulomb scattering (MCS).
Divide track into segments of equal length and fit a straight line to each segment. Scattering angle is angle between successive segments.
$\nu_e$ CC energy: Electron Shower Energy

Electron shower energy: Calorimetric energy calibrated with Monte Carlo

\[ \text{Gradient} = 0.99 \]
\[ \text{Intercept} = -0.02 \]

\[ \text{Mean} = 0.00 \]
\[ \sigma = 0.08 \]

Nick Grant, Tingjun Yang, DPF2017
Hadronic Energy

- Estimate the hadronic energy from reconstructed hits that are not in the muon track or electron shower. Make calibration using Monte Carlo.

- Reco hadronic energy tends to be too low since neutral particles are not reconstructed in the DUNE far detector. On average 40% of the energy is invisible due to neutron scattering etc, and there are fluctuations in this from event to event, and this limits energy resolution.

Nick Grant, Tingjun Yang, DPF2017
Fiducial cuts with reco hits and Global Wire

- Select Nue CC MC events: No hits near the edge of the FD
- Global Wire
  - modified the GlobalWire algorithm in the “BlurredClusteringAlg” module
  - It shows a clean and continuous event in the one image.

Local Wire using “gaus hit”

Global Wire

Hits on two TPCs
Pass through APA