Image Based Reconstruction and Deep Learning Tools in MicroBooNE

Rui An @ Illinois Institute of Technology
09-18-2019, DESY
On behalf of MicroBooNE Collaboration
Reconstruction and Machine Learning in Neutrino Experiments
Outline

• MicroBooNE developed an *imaged based event reconstruction chain*,
  ★ taking advantage of high resolution image data and computer vision software OpenCV,
  ★ targeting the exclusive $1 \text{lepton}-1 \text{Proton}$ topologies in the low energy excess region,

• Various deep learning networks applied by MicroBooNE,
  ★ Semantic segmentation network $\rightarrow$ pixel clustering $\rightarrow$ **SSNet**,
  ★ Classification network $\rightarrow$ particle identification $\rightarrow$ **MPID**,
  ★ Instance segmentation network $\rightarrow$ cosmic removal $\rightarrow$ **MaskRCNN**.
MicroBooNE LArTPC and Image

2,400 wire on U,V and 3,456 wires on Y plane.
4.8 ms readout.
~[wires]x[ticks /6 ] pixels after applying compression factor on wires and time.
Proton Event Topology

- Benefits: easier to reject cosmic and non-CCQE in LEE region using topology parameters.
- Visible requirement: proton > 60 MeV and lepton > 35 MeV.
Reconstruction Chain

1. Cosmic Tagging
2. Shower / Track Pixel Labeling
3. 2D & 3D Reconstruction
4. PID using CNN
5. 1μ1p
6. 1e1p

Cosmic Tagger | SSNet | Vertex Finding | Particle Clustering | MPID
The goal is to make pixel-level label decision between track and shower for LArTPC,

SSNet applies a combination of U-Net and ResNet. Feature maps concatenated between down and up sampling steps,

Applied the SparseNet, training speed boosted by 5,

Training sample size: 100,000 events. Test sample size: 20,000 events.
SSNet applied a weighted loss for the LArTPC image sparsity.

Achieved very low error rates across different samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>ICPF mean</th>
<th>ICPF 90%</th>
<th>Shower</th>
<th>Track</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>1.9</td>
<td>4.6</td>
<td>4.1</td>
<td>2.6</td>
</tr>
<tr>
<td>$\nu_e$</td>
<td>6.0</td>
<td>13.8</td>
<td>7.6</td>
<td>13.8</td>
</tr>
<tr>
<td>$\nu_\mu$</td>
<td>3.9</td>
<td>4.5</td>
<td>14.2</td>
<td>4.3</td>
</tr>
<tr>
<td>1e1p</td>
<td>2.2</td>
<td>5.7</td>
<td>2.8</td>
<td>4.0</td>
</tr>
<tr>
<td>1$\mu$1p-LE</td>
<td>2.3</td>
<td>2.2</td>
<td>6.2</td>
<td>2.4</td>
</tr>
<tr>
<td>1e1p-LE</td>
<td>3.9</td>
<td>11.5</td>
<td>3.8</td>
<td>8.0</td>
</tr>
</tbody>
</table>

*ICPF: Incorrectly Classified Pixel Fraction*
SSNet Output

- SSNet is used to provide vertex seed.
We developed two reconstruction chains respectively for,

- $1\mu 1$Proton (this talk) $\rightarrow$ using track pixel only images.
- $1e 1$Proton (backup slides) $\rightarrow$ using track and shower pixel images.

Two libraries developed:

- Liquid Argon Open Computer Vision (LArOpenCV),
  - library for image manipulation, reconstruction and analysis,
  - interface between LArTPC data and OpenCV,
- Liquid Argon Computer Vision (LArCV),
  - bridge LArTPC data and TensorFlow, PyTorch and caffe etc,
  - Image data format and processing tool.
• $1\mu1P$ Proton vertex finding uses track-only images,
• Calculated with OpenCV.
1μ1P Vertex Candidates

- Vertex Candidate #1: contour defect points,
- Vertex Candidate #2: PCA (principal component analysis) crossing points.
1μ1P Angular Metric

**Angle Definition**

- $\theta$: angle of seed and crossing points.
- $\phi$: angle of PCA crossing points.

**Find local minimum of $|\theta - \phi|$**

**Scan d$\theta$ Over 3 Planes**

$\Delta\theta = |\theta - \phi|$
1μ1p Vertex Finding Performance

- Average efficiency of finding vertex at ~52%.
- Vertex finding resolution for $1\mu 1p \rightarrow$ less than 0.73cm for 68% events.
- Shower vertex finding procedure in backup slides.
Track Particle Clustering

- Pixels are clustered based on their distance and angle under the polar coordinate.

- Clusters across three planes are further evaluated by matching pixels over time tick and wire crossing.
Reconstruction Example

MicroBooNE Simulation Preliminary
1\(\mu\)1Proton Example

MicroBooNE Simulation Preliminary
1\(e\)1Proton Example
Multi-particle PID Network

- Multi-particle PID network is a CNN network application, extended work of the single-particle PID network in arXiv: 1611.05531.
- Final layer in a sigmoid function and predicts probabilities of particles in the input image.
- Training image, simulated multiple particles coming from one vertex,
  - avoid bias from neutrino mode,
  - learns more from richer topology information.
- Training sample size: 50,000 events. Test sample size: 40,000 events.
- Better as,
  - does not require a vertex resolution,
  - does better on Pi0 present events (hard to reconstruct detached shower).

P  e\(^-\)  γ  μ\(^-\)  π\(^\pm\)  each ∝ (0,1)
Multi-particle PID Network

<table>
<thead>
<tr>
<th></th>
<th>$e^-$</th>
<th>$\gamma$</th>
<th>$\mu^-$</th>
<th>$\pi^\pm$</th>
<th>$P^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>0.16</td>
<td>0.86</td>
<td>0.38</td>
<td>0.77</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$e^-$</th>
<th>$\gamma$</th>
<th>$\mu^-$</th>
<th>$\pi^\pm$</th>
<th>$P^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.92</td>
<td>0.23</td>
<td>0.35</td>
<td>0.07</td>
<td>0.74</td>
</tr>
</tbody>
</table>
Multi-particle PID Network

MicroBooNE Simulation Preliminary
One Shower Example

<table>
<thead>
<tr>
<th>$e^-$</th>
<th>$\gamma$</th>
<th>$\mu^-$</th>
<th>$\pi^\pm$</th>
<th>$P^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.82</td>
<td>0.4</td>
<td>0.15</td>
<td>0.02</td>
<td>0.72</td>
</tr>
</tbody>
</table>

MicroBooNE Simulation Preliminary
Two EM Showers Example

<table>
<thead>
<tr>
<th>$e^-$</th>
<th>$\gamma$</th>
<th>$\mu^-$</th>
<th>$\pi^\pm$</th>
<th>$P^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.64</td>
<td>0.71</td>
<td>0.12</td>
<td>0.23</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Cosmic Tagger | SSNet | Vertex Finding | Particle Clustering | MPID
• Good separation between shower-like and track-like particles,
• Good separation between track-like particles:
  ★ proton, muon and charged pion.
• Good separation between track-like particles:
  ★ electron and gamma,
• Using 1μ1P as a sideband study for MPID data MC comparison.

• Good comparison between MicroBooNE’s open data and beam simulation + cosmic.

• Syst. error includes applying MPID to same events with various E-field, SCE, channel noise etc.

• In proton score distribution, there is a bump at 1 introduced by neutrino events.
New Reconstruction Chain

1. Cosmic Tagging
2. Shower / Track Pixel Labeling
3. 2D & 3D Reconstruction
   - PID using CNN
   - $1\mu1p$
   - $1e1p$
• MicroBooNE has large number of dead wires.
• Infill network tries to fill these region with realistic values and improve tracking performance for cosmic muon.
• Training on off-beam data plus dead wire pattern crop.
• U-net + encoder and decoder layers.
Infill Network

- MicroBooNE simulation example. Colored boxes are dead wire regions.
- Pixels are predicted in reasonable region.
• Predicated pixel ADCs are in a reasonable region and have good efficiencies.
Mask RCNN

- **Mask Regional-CNN**, a model well known for its good performance on instance segmentation, can provide:
  - ★ ROIs for objects,
  - ★ Label for the object,
  - ★ Pixels for the object.

- Researched with MicroBooNE data in two ways:
  - ★ Mask RCNN-Ancestor: Neutrino vs. Cosmic.
  - ★ Mask RCNN-Particle:
    Four particle types of $e^-$-like, $\mu^-$, $\pi^\pm$ and proton.
• Training image size of 512x832 cropped from full event display.
• Training sample has 93% cosmic, 3% neutrino and 4% other.
• Training with ResNet-50 pre-trained weights.
• Applied Mask RCNN-Ancestor to MicroBooNE data.
• ~90% particle covered by the proposed clusters of network output.
• Proposed cluster maps strongly to one particle
• Great tool for cosmic pixel tagging.
• DBScan applied to EM shower particles to break large sparse ROIs introduced by EM showers,

• Four particle types in training sample, $P$, $\mu^-$, $\pi^\pm$ and electron-like shower piece,

• Applied similar pixel weighted as SSNet and a class weighted loss (to unbias the dominating electron labels from DBScan),

• Trained with ResNet-101 weights from COCO, improving the ROI proposal on tracks.
Mask RCNN-Particle

- Mask RCNN-particle shows potential in improving particle clustering, especially for event with detached shower.
- Challenge: the center-based ROI proposal step does not seem ideal for particles coming from one vertex (e.g. compared to corner-based ROI proposal).
Conclusion

• MicroBooNE has developed an imaged-based reconstruction tool for $1\mu 1$Proton and $1e1$Proton topologies,
  ★ on the way to a LEE result,
  ★ being applied to higher multiplicity topologies.

• A various of deep learning networks and how to apply the output have been researched by MicroBooNE.

• Good demonstrations of applying deep learning to LArTPC.
Back UP
Precuts applied to reconstructed flash in a short beam spill window (93.75 ns) to reject:
1. Random single PE noise.
2. Single PMT noise.
3. Flash from Michel electron prior to beam spill window.

Achieved:
Neutrino efficiency: > 96%
Backgrounds rejection: > 75%
Dedicated cosmic tagging tool developed to tag cosmic objects and find ROI.
1. Cosmics down beam direction cross the front and end wires
2. Cosmics across top-bottom have unique triplet of wires of three planes.
3. Cosmics traversing TPC have the max $\Delta t$ between the first and last charges on the track.
3D Track Reconstruction

- This reconstruction uses computer vision and clustering tools to find 3D-consistent vertices, and a 3D stochastic best neighbor search.
Occlusion Analysis on MPID

- Electron Score Map
- Proton Score Map
- Gamma Score Map
- Electron Score Map

μBooNE Simulation
3D Track Reconstruction

V Plane

time [0.5 μs]

Y Plane

time [0.5 μs]
1e1P Vertex Finding

- Step 1: Find edge points on track,
- Step 2: Keep edge point crossing shower pixels,
- Step 3: Match edge points across planes

2. Merge showers based on pixel location (length, angle) in a polar coordinate.

3. Reconstructed showers in fit cones.
MPID Training Sample

- Training sample generated with customized event generator,
  - Create “3D interaction vertex”,
  - One vertex per event, random, uniformly distributed in TPC,
  - Random particle multiplicity [1, 4] particles per event,
  - Random, isotropic particle momentum directions,
  - Random particle types from, $P, e^-, \gamma, \mu^-$ and $\pi^\pm$,
  - Two particle types mixtures,
    - 80% events with kinetic energy in [100,1000]MeV and proton in [100,400]MeV,
    - 20% events with kinetic energy in [30,100]MeV and proton in [40,100]MeV.

- ~45,000 events for training and ~40,000 for validation.

- Why not train with overlay image for training, (1) Cosmic muon would make the network fail for detecting neutrino induced muon (2) Michel & deltas would make the MPID fail for detecting neutrino-induced electron.
MPID Data MC Comparison

MicroBooNE Preliminary

\[ \chi^2 / \text{NDF} = 2.93 / 19, \ P\text{-value} = 1.0000 \]

\[ \chi^2 / \text{NDF} = 4.04 / 19, \ P\text{-value} = 0.9999 \]
MPID Data MC Comparison

MicroBooNE Preliminary

χ² / NDF = 1.71 / 19, P-value = 0.9655

χ² / NDF = 7.03 / 19, P-value = 1.0000