LArCV Implementation at ICEBERG

Shekhar Mishra
On Behalf of ICEBERG
Neutrino Division, Fermilab

AI-Neutrino @ICEBERG Meeting
04/05/2023

Fermilab DUNE

Physics goals

- LArTPC is designed to study neutrino (particle) interaction by reconstructing the interaction as fine-grained images.

- In AI-Neutrino R&D for DUNE we are working to identify:

  - Michel electrons (coming from muon decay)

  - Track-like particles (e.g. muons, protons, pions)

  - ^{39}Ar Beta Decay

  - Electromagnetic showers (e.g. electrons, photons = gammas)

  - Delta rays (electrons knocked off from hard scattering)

  - Low energy depositions

- AI-Neutrino R&D at ICEBERG is working to identify "these" at the 100 Gbits/sec data input rate at the pipeline timeframe speed.

What is LArCV

- MicroBooNE originally started the development of Liquid Argone Open Source Computer Vision (LArOpenCV).

  - An OpenCV based pattern recognition software for LArTPC.

- LArCV has been derived for utilization in other LArTPC experiments.

  - 2D image and 3D volumetric data representations + file I/O

  - 2D/3D data processing algorithms + development environments for new algorithms

  - data processing framework to read/write/analyze/modify a large chunk of data with algorithms

  - extensive Python interfaces that allow bindings to Python code (jupyter notebook, caffe, TensorFlow, PyTorch)

- LArCV uses ROOT/HDF5 file as an input and created 3D images

  - Its I/O can be modified to take data from memory

🐝 Fermilab DUNE

LArCV Steps

- Input Data at present comes from MC in ROOT/HDF5 format.

    - In DUNE-DAQ there are packages to provide the raw data buffer

- Input: 3D images are constructed using the wire data (2D images, three views)

    - In MC path a ROOT file is an input to LArCV

    - A I/O parser creates a 3D input for the downstream networks.

- Pixel-wise classification is used to identify and classify images into physics processes

- Point of interest (essentially start and end points) is identified by using PPN.

- Identification of Interaction and particle types are done Graph Neural Network

🔷 Fermilab DU(VE)

Dataflow in LArCV

ROOT & LArCV

LArCV & MLReco3D
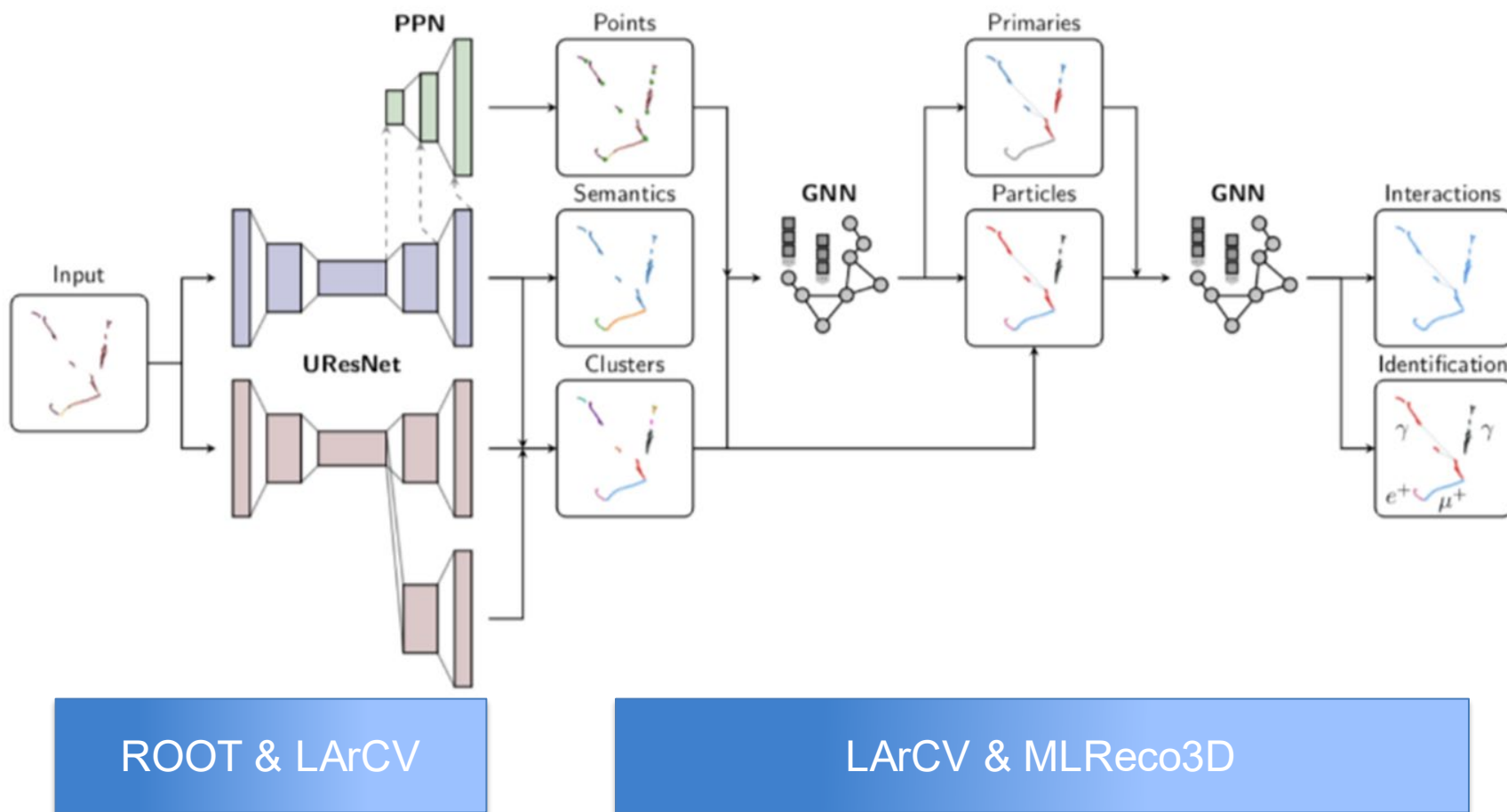
Why LArCV at ICEBERG

- During 2019-2021 at ICEBERG we took more than 1000 cosmic runs

  - We analyzed enough data to address hardware questions.

  - Rest are not analyzed due to man power limitation.

  - DUNE will generate huge volume of data.

- An ICEBERG colleague, who was very much involved with ICEBERG DAQ presented a poster on ML for ProtoDUNE-SP data.

  - https://nusoft.fnal.gov/nova/nu2020postersession/pdf/posterPDF-512.pdf

- This is a tool we were looking for:

  - Patrick Volunteered to help.

- LArCV is in use for Offline analysis development at ProtoDUNE, ICARUS, SBND and DUNE (FD and NDLAr).

  - Pool of physicists and computer professional

I. ProtoDUNE Single Phase

- one of the two prototypes for the Deep Underground Neutrino Experiment (DUNE) at CERN Neutrino Platform
- the detector is a 7.2 x 6.0 x 6.9 m liquid argon time projection chamber (LArTPC) w/ ~740 tons of liquid argon
- beam data was collected Sep - Nov 2018

Fig 1: ProtoDUNE-SP

II. Machine Learning Based Reconstruction

- UResNet[1]: a convolution neutral network using U-Net[2] & ResNet[3]
  - allow voxel-level features extraction
  - remove fake 3D space points and to separate track / shower voxels from the prediction of network decoder output
- Point Proposal Network (PPN)
  - an extension of UResNet along the decoder path
  - detect the starting/end point(s) of a shower/track

[1] L. Domine & K. Terao, arXiv:1903.05663
[2] O. Ronneberger et al., arXiv:1505.04597
[3] K. He et al., arXiv:1512.03385

Fig 2: UResNet + PPN Architectures

III. Building 3D Space Points in LArTPC

- 3D points are reconstructed by matching three 2-D views of wire projections
- fake (ghost) points are introduced because of the combinatorics of matching 2D hits to 3D space
- ghost points are identified by UResNet based on pattern recognition

(a) True Energy Deposition (b) Reconstructed 3D Points (c) Network Output

ProtoDUNE-SP Simulation

Fig 3: Ghost Points Removal

Track
Shower 1
Shower 2

ProtoDUNE-SP Simulation

Fig 4: An Example of π^0

IV. Neutral Pion Reconstruction in ProtoDUNE-SP

- aims to reconstruct neutral pion π^0 mass produced from charged pion of the ProtoDUNE beam data
- provides a standard candle (mass peak at 134.98 GeV /c^2) to gauge the performance of the reconstruction algorithm

Procedures

- remove ghost points from 3D image (UResNet)
- identify shower/track voxels (UResNet)
- require at least 2 shower starting points (PPN)
- estimate shower direction using principal component analysis (PCA)
- match the two showers backward to a track
- group shower voxels from the estimated direction and opening angle (θ)
- estimate shower energies (E_1, E_2) using voxel occupancy

$$m_{\pi^0} = \sqrt{2 E_1 E_2 (1 - \cos\theta)}$$

Reco. m_{π_0}
132.1 ± 24.6
[MeV/c^2]

ProtoDUNE-SP Simulation

π_0 invariant mass [MeV/c^2]

Fig 5: Reconstructed π^0 Mass

Run 5387 - Event 39479

Fig 6: An candidate for $\pi^0 \to \gamma\gamma$ from 1 GeV charged pion beam

On April 11, 2022 Requested Patrick (SLAC)

ICEBERG: 2 Paths of Development

- LArTPC Data from MC Data

    - LArTPC ROOT MC data to establish the LArCV framework.

    - ICEBERG Single-Muon ROOT MC data

- LArTPC data from WIB

    - ICEBERG is integral part of DUNE CE, DUNE-DAQ and DUNE-Offline analysis development.

    - We are working to understand how LArCV can access the data stored by DUNE-DAQ in the server memory.

    - Data format, layout and Pipeline

    - AI/ML steps in the pipeline time frame.

    - Data output….

🔷 **Fermilab** DUNE

LArCV on wc.fnal.gov

- LArCV and lartpc_mlreco3d on wc.fnal.gov

  - We are using a Singularity file to understand AI/ML

    - /work1/dune/users/mishra/larcv/larcv2_ub20.04-cuda11.3-cudnn8-pytorch1.10.0-extra.sif

  - The framework is working: Michel Electron in LArTPC MC Data
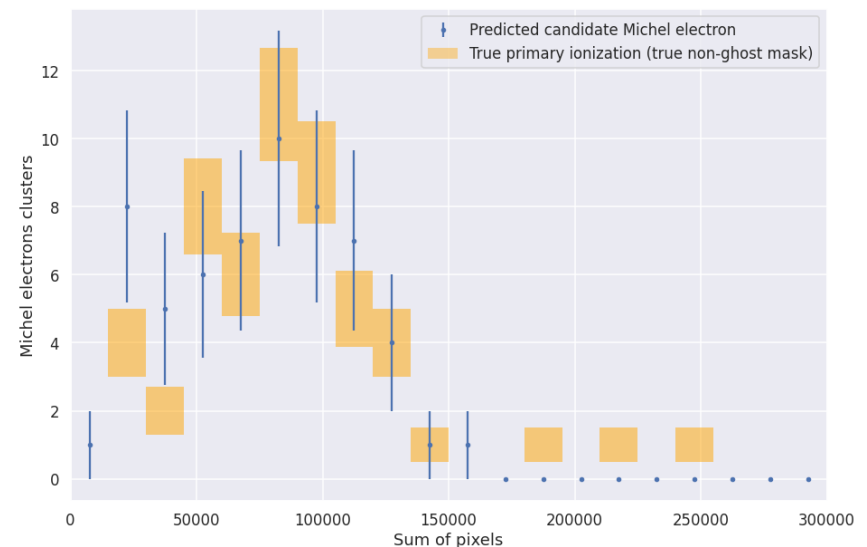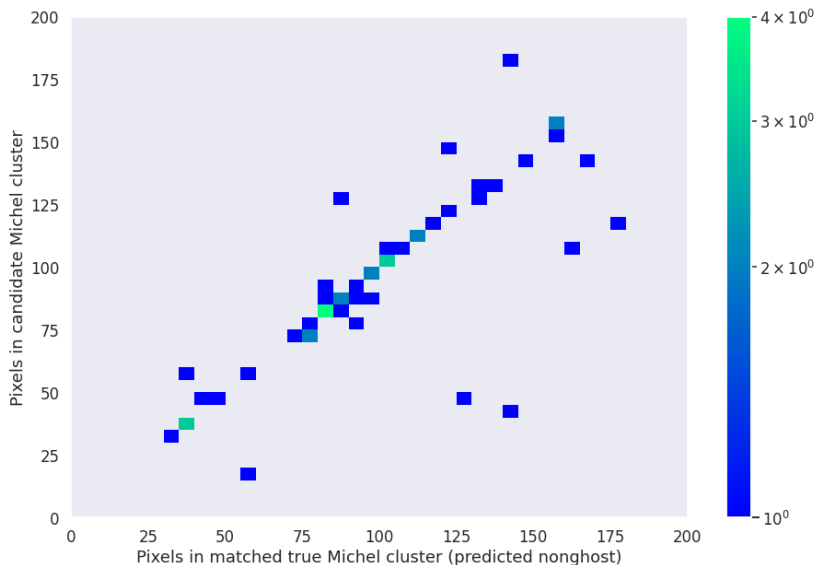
- Working on generating ICEBERG Root file as an input.

ICEBERG TPC

- ½ DUNE Anode Plane Assembly (APA) 1280 channels.

  - 10 Front End Mother Board

- 1/6 DUNE APA Wire length

- Two Drift Volume of 30 cm each

- Two cathode Planes and 1" spacing Field Cage

  - 500 v/cm field gradient

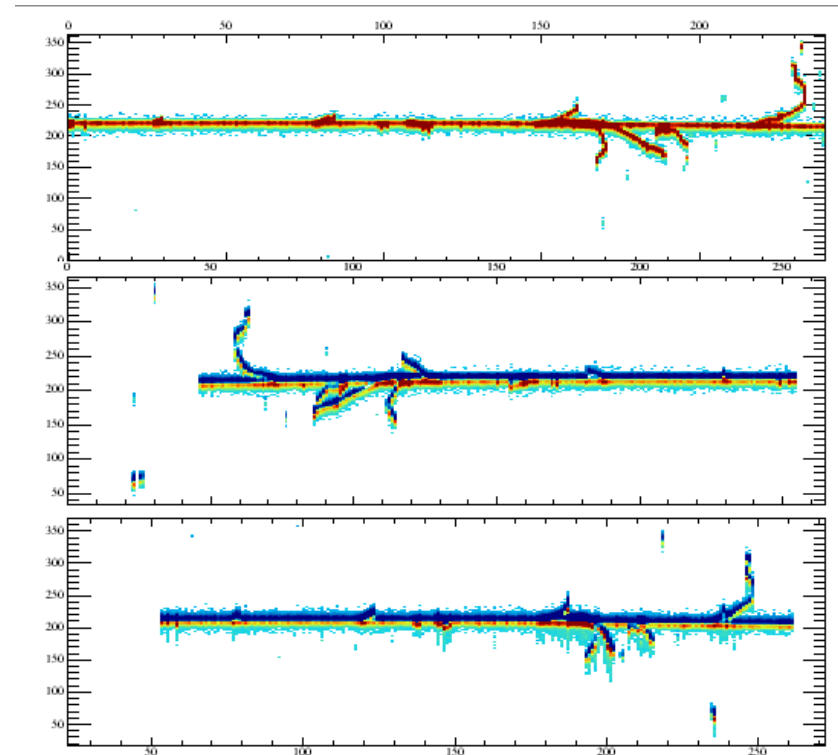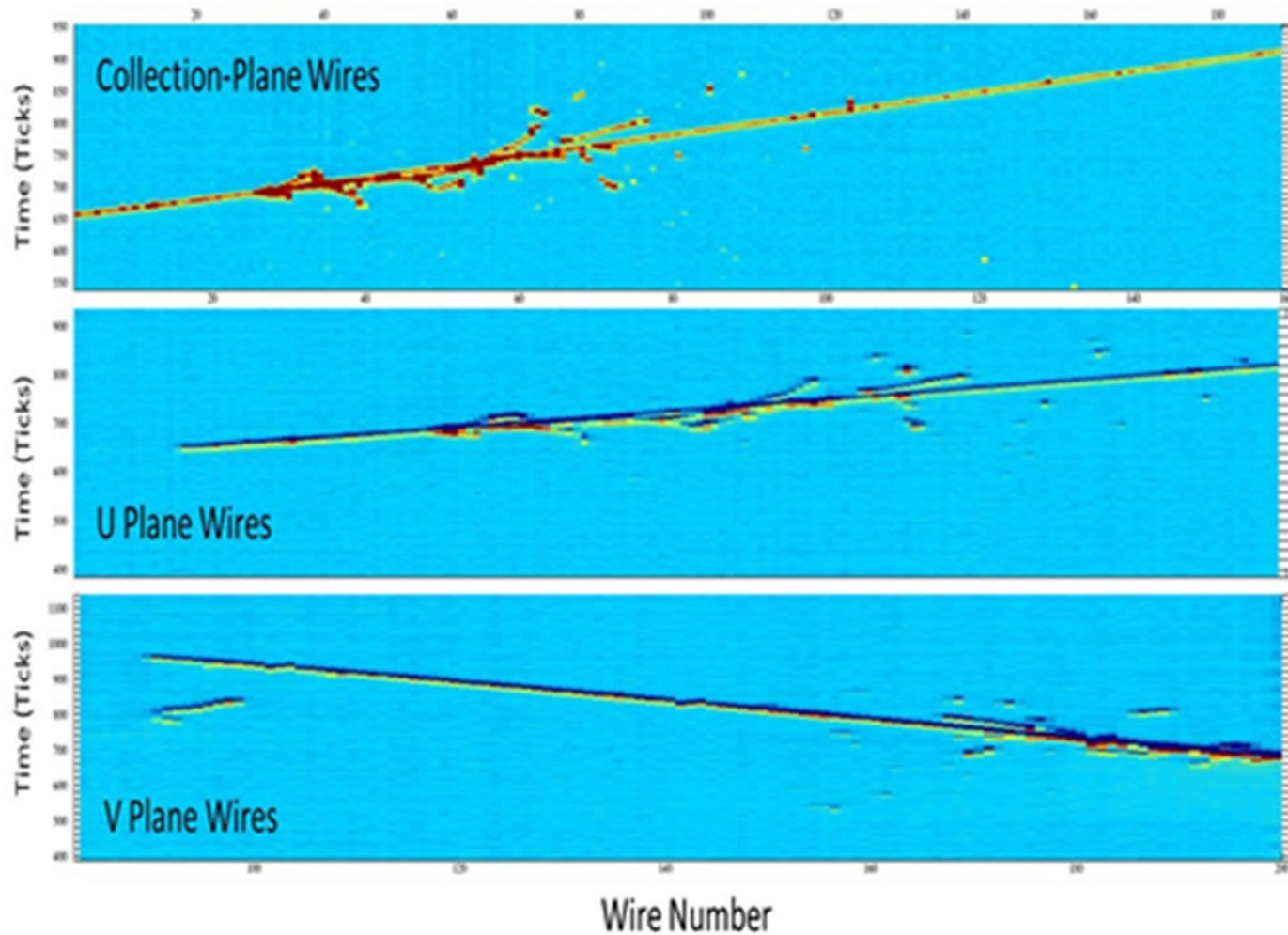- 2 Slots for Photon Detectors

ICEBERG ROOT MC Data

- We have established a procedure to generate MC data:

  - for ICEBERG FEMB-WIB layout

  - DUNE-WIB (Frame Format 14)

  - Read it through DUNEsw and analyze.

- DUNE Wiki

https://wiki.dunescience.org/wiki/AI-Neutrino_Project_at_ICEBERG

Tom's Talk

🔷 Fermilab DU(VE

An ICEBERG Cosmic Track

OS and Hardware limitations

- All the LArCV developments have been done on Ubuntu Linux and latest versions of many packages.

  - Fermilab uses SL7.9

  - Fermilab is moving to Alama Linux

- LArCV on Alama Linux VM running of Window 11
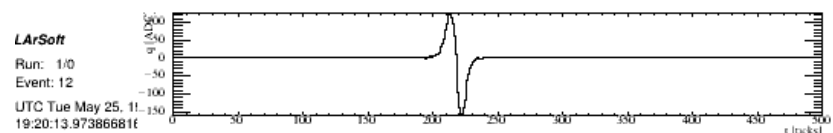
Status of LArCV

- LArCV compilation errors on AL9.1 have been fixed.
  - Essentially due to OpenCV3 → OpenCV4
- LArCV compilation warning are being fixed.
  - Essentially due to different versions of C++ compilers
- DUNEsw/LarSoft is build on SL7.9 and is not supported on AL9.1
  - Compilation on AL is being worked on
- Authors have mlreco3D have reported that it has stopped working on CPU since April 2022.
  - It only works on GPU hence the use of wc.fnal.gov
- LArCV2 → LArCV3
  - We have installed LArCV3 on AL9.1 VM.
    - Have not tried to run a physics analysis
  - LArCV3 is moving from ROOT format to HDF5.

Thoughts of AI/ML at ICEBERG

- I am only working on LArCV implementation at ICEBERG in the DUNE-DAQ hardware and DUNEsw (DUNEtpc and Larsoft).

- Colleagues interested in other implementation of AI-Neutrino @ ICEBERG

  - Assume that ICEBERG would provide a 100 Gbits/sec fiber with WIB data.

    - WIB Firmware

  - We could Plug that fiber into any server with NIC and

    - FPGA

    - FELIX

    - …..

- ICEBERG with CE and PD Consortium is working to setup a vertical slice test of DUNE-FD-1 with DUNE-WIB later in Summer 2023.

  - ICEBERG will make that data available as a HDF5 file for offline analysis.

🔶 **Fermilab** DU(N)E