Machine Learning for Neutrino Identification

Saúl Alonso-Monsalve
CERN

Symposium on Artificial Intelligence for Science, Industry and Society (AISIS)

UNAM, Mexico City

21 October 2019

Overview

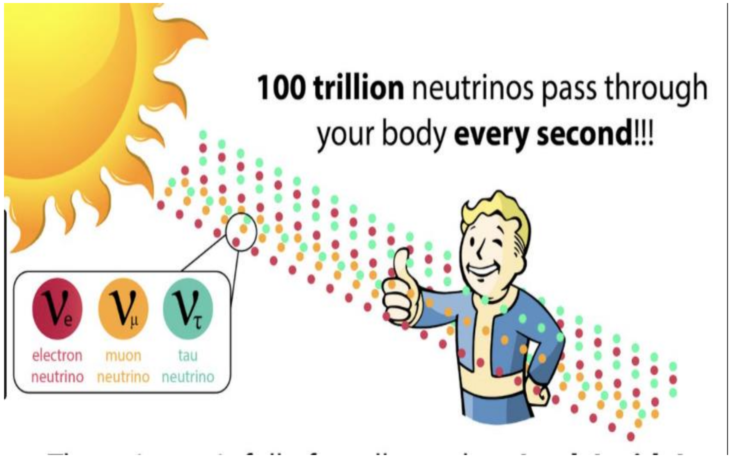
- Introduction to neutrinos.
- Machine Learning in neutrino experiments:
  - NuMI Off-axis νe Appearance (NOvA).
  - Micro Booster Neutrino Experiment (MicroBooNE).
  - Deep Underground Neutrino Experiment (DUNE).
  - Tokai to Kamioka (T2K).
- Moving to the edge:
  - Inference on TPU.
  - Inference on FPGA.
- Summary.

Overview

- **Introduction to neutrinos.**
- Machine Learning in neutrino experiments:
  - NuMI Off-axis νe Appearance (NOvA).
  - Micro Booster Neutrino Experiment (MicroBooNE).
  - Deep Underground Neutrino Experiment (DUNE).
  - Tokai to Kamioka (T2K).
- Moving to the edge:
  - Inference on TPU.
  - Inference on FPGA.
- Summary.

Neutrinos

- Neutrinos are **fundamental particles**.
  - They belong to the Standard Model of Physics.
  - Neutrinos were first detected in 1956.
- Neutrinos are ghostly particles.
  - ~100 trillion (10^{14}) neutrinos pass through your body for every second of your life!
  - There are a billion neutrinos for each atom in the Universe. There are ~3×10^8 neutrinos per cubic meter.
- Neutrinos are still **mysterious particles**.
- Neutrinos come from "everywhere".
  - Solar neutrinos.
  - Atmospheric neutrinos.
  - Relic/supernova neutrinos.
  - Nuclear reactor created neutrinos.
  - Accelerator created neutrinos.
  - Geoneutrinos, radioactive decay (even from your body).



Neutrino oscillations

- Important discovery in 1998: **neutrino oscillations**.
  - *"Neutrino oscillation is a quantum mechanical phenomenon whereby a neutrino created with a specific lepton flavour (electron, muon, or tau) can later be measured to have a different flavour. The probability of measuring a particular flavour for a neutrino varies between 3 known states as it propagates through space."*
  - Neutrino oscillations only possible if neutrinos have a **non-zero** mass.
  - 2015 Nobel Prize:
    - Takaaki Kajita, Art McDonald:
    - *"For the discovery of **neutrino oscillations**, which shows that neutrinos have mass."*

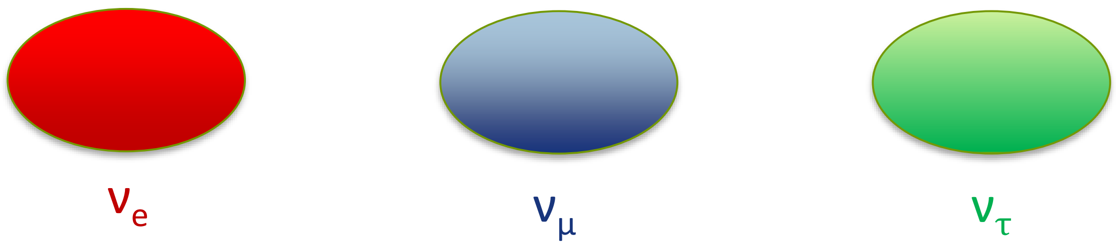
Takaaki Kajita
SuperKamiokaNDE

Art McDonald
SNO

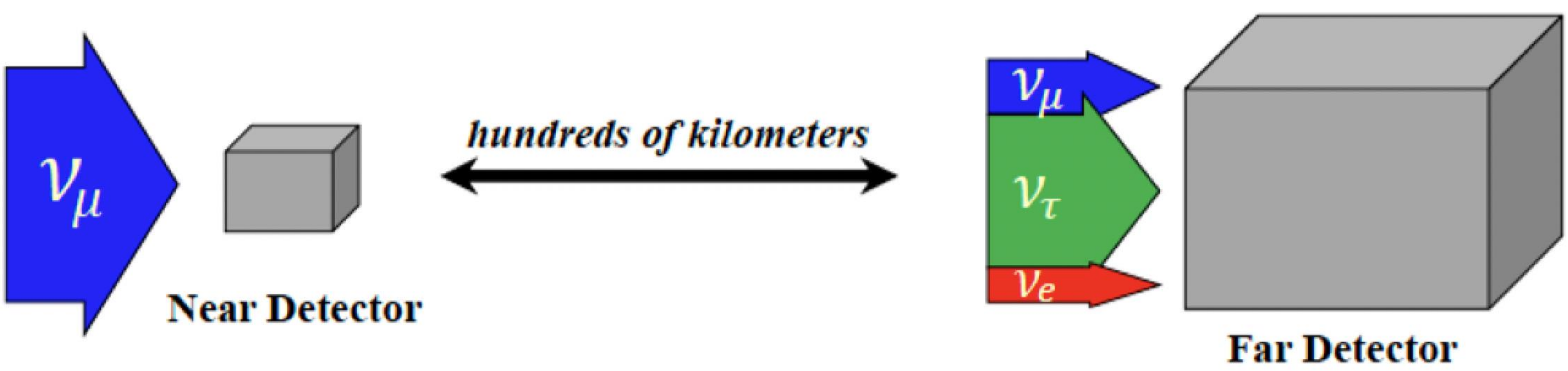
3 More Nobel prizes for Neutrinos since 1988:
- *2002: R. Davis and M. Koshiba.*
- *1995: F. Reines.*
- *1988: L. Lederman, M. Schwartz and J. Steinberger.*

Neutrino flavour

- Neutrino flavour or interaction states:
  - Electron neutrino ν_e (“nue”), muon neutrinos ν_μ (“numu”), tau neutrinos ν_τ (“nutau”).



- Measuring *CP*-violation allow us to understand how neutrinos (and anti-neutrinos) oscillate.

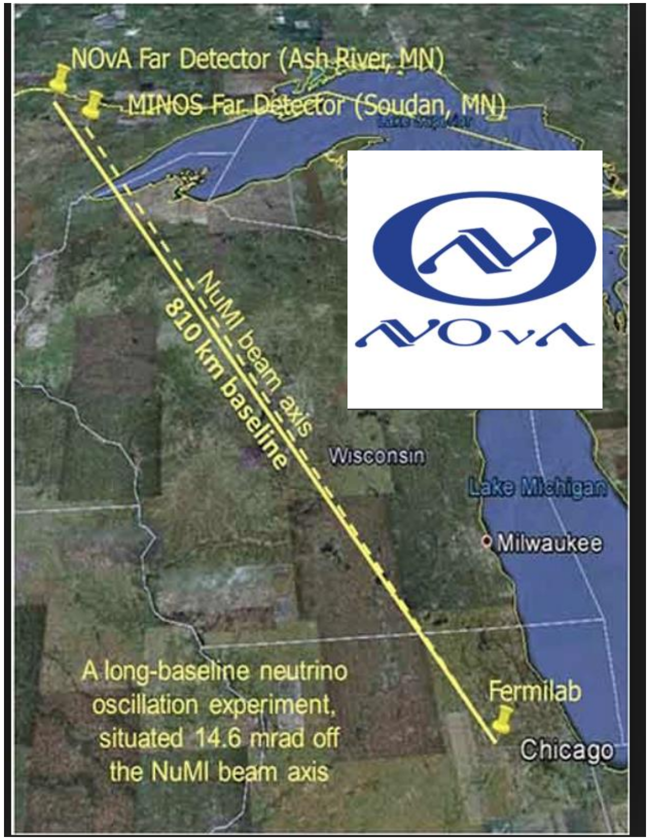


Overview

- Introduction to neutrinos.
- **Machine Learning in neutrino experiments:**
  - **NuMI Off-axis ve Appearance (NOvA).**
  - Micro Booster Neutrino Experiment (MicroBooNE).
  - Deep Underground Neutrino Experiment (DUNE).
  - Tokai to Kamioka (T2K).
- Moving to the edge:
  - Inference on TPU.
  - Inference on FPGA.
- Summary.

NOvA

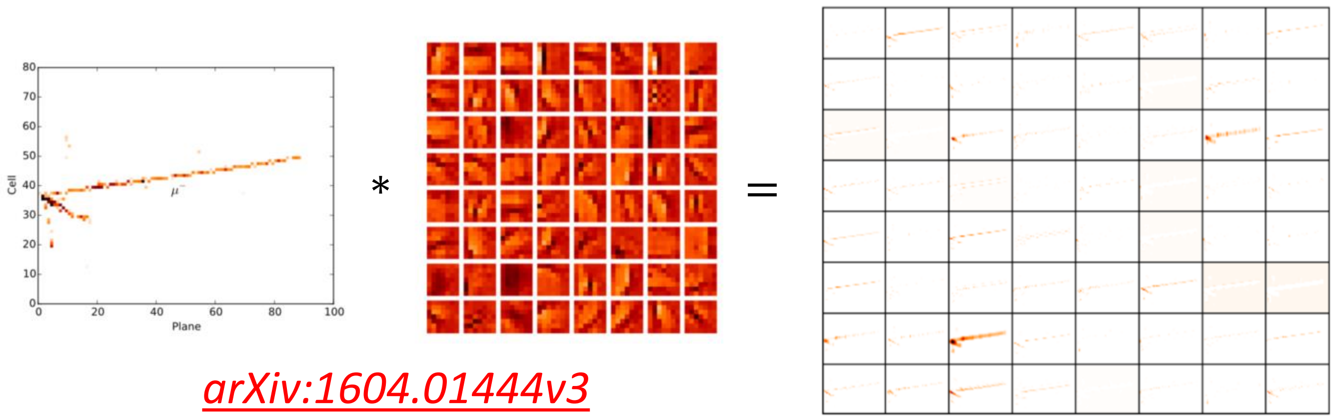
- **NuMI Off-axis ν_e Appearance (NOvA)** is a long baseline neutrino oscillation experiment.
- Measures the neutrino signal close to its source, at Fermilab, as well as 810 km away, at Ash River, MN.
  - Aims to make precision measurements of neutrino oscillation parameters via the disappearance of ν_μ and appearance of ν_e from neutrino oscillation.
- Pioneering on using convolutional neural networks (CNN) for neutrino identification:
  - *A. Aurisano et al., "A Convolutional Neural Network Neutrino Event Classifier", arXiv:1604.01444v3.*



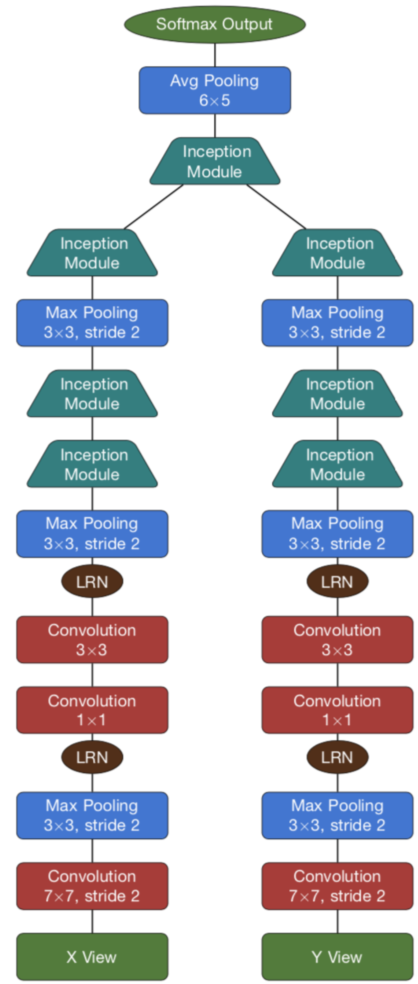
NB: I will only write neutrino from now on, but the same is applicable for antineutrinos

Deep Learning in NOvA

- Initial CNN approach inspired by the GoogleNet architecture (InceptionV1).
  - The goal was to identify particle interactions in sampling calorimeters.
  - Called CVN (Convolutional "Visual" Network).



arXiv:1604.01444v3



- The NOvA CVN classifier outperformed other algorithms in use by the NOvA collaboration!

Overview

- Introduction to neutrinos.
- **Machine Learning in neutrino experiments:**
  - NuMI Off-axis ve Appearance (NOvA).
  - **Micro Booster Neutrino Experiment (MicroBooNE).**
  - Deep Underground Neutrino Experiment (DUNE).
  - Tokai to Kamioka (T2K).
- Moving to the edge:
  - Inference on TPU.
  - Inference on FPGA.
- Summary.

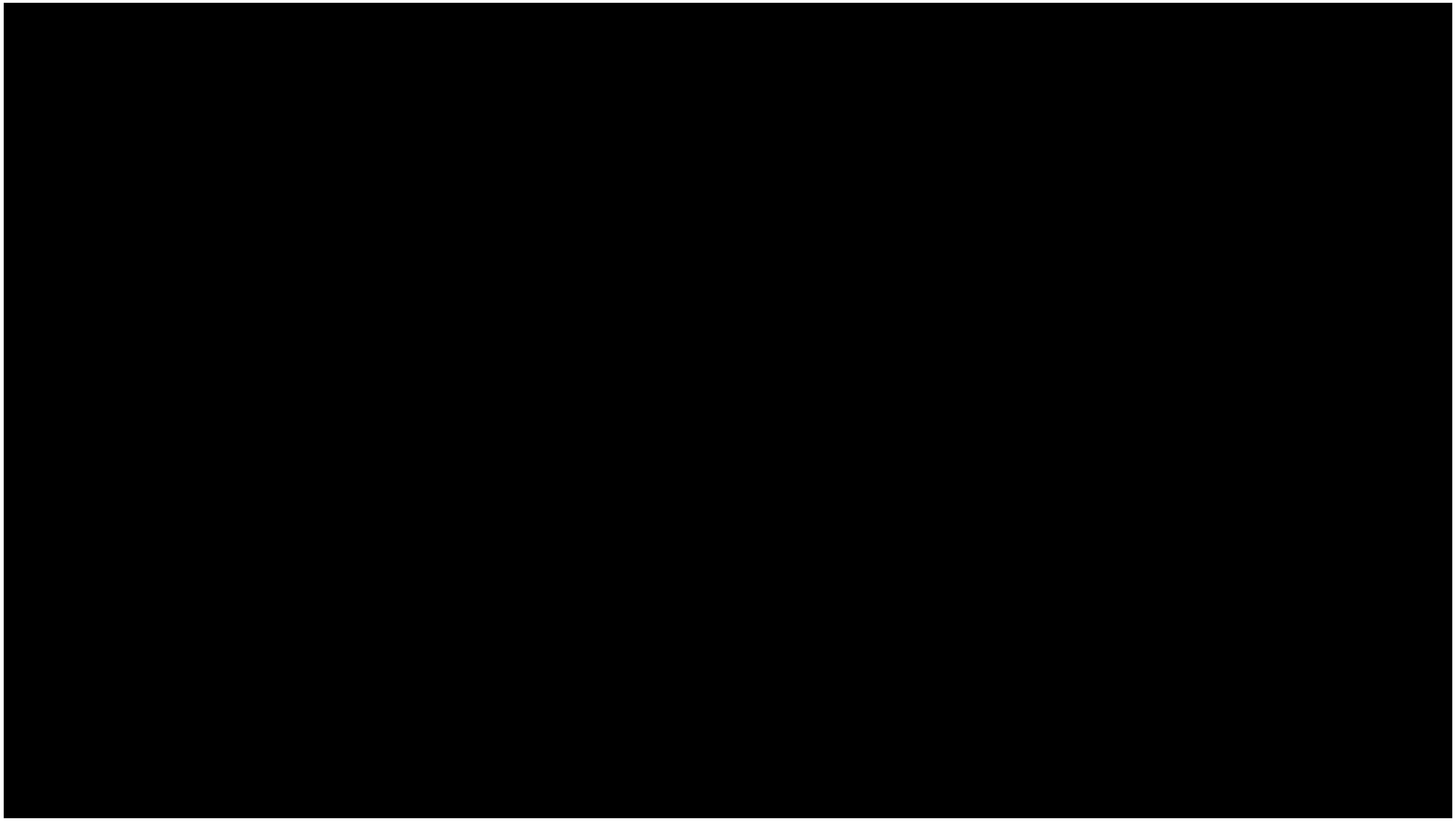
MicroBooNE

- The **Micro Booster Neutrino Experiment (MicroBooNE)** will investigate the low energy excess events observed by the MiniBooNE experiment, measure a suite of low energy neutrino cross sections, and investigate astro-particle physics.

- MicroBooNE is a large 170-ton Liquid-Argon Time Projection Chamber (LArTPC) neutrino experiment located on the Booster neutrino beamline at Fermilab.

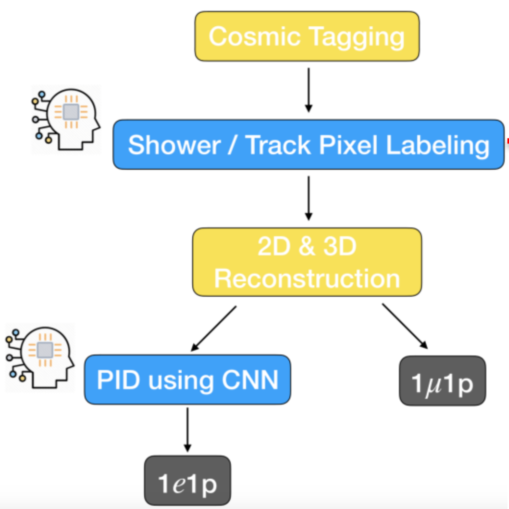
LArTPC

- Liquid-Argon Time Projection Chamber (LArTPC).
  - This provides “images” of each neutrino interaction.

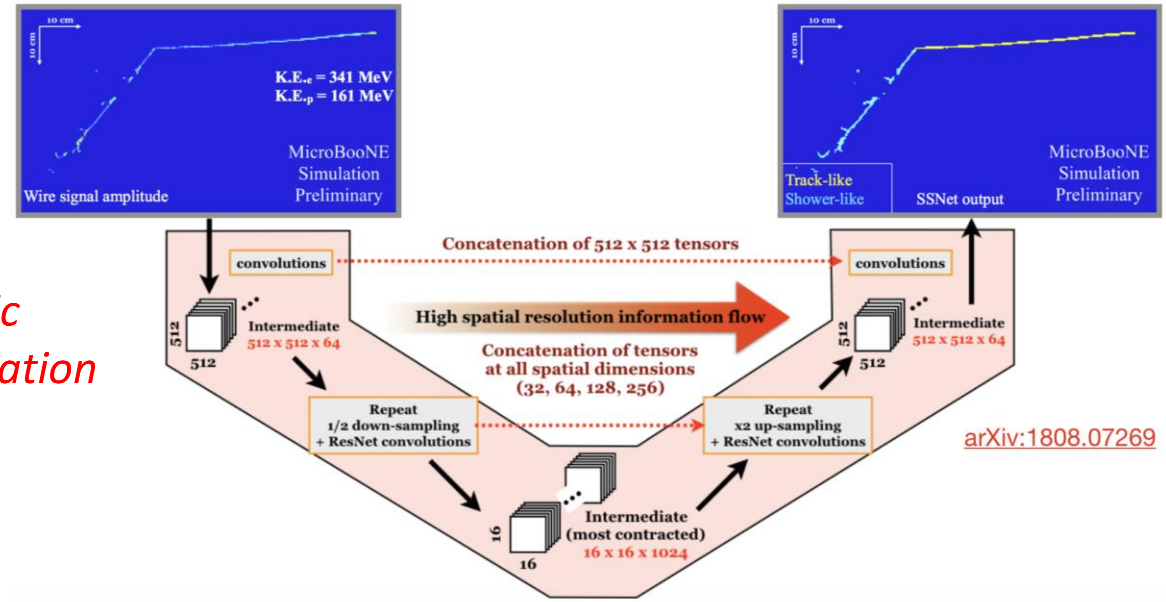


Deep Learning MicroBooNE

- MicroBooNE developed an imaged based event reconstruction chain.



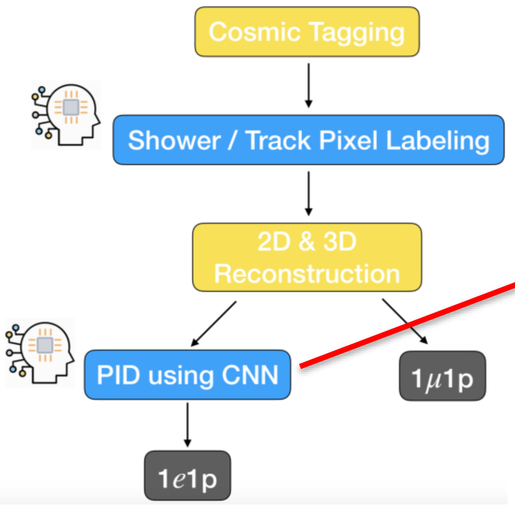
SSNet (Semantic Segmentation network)



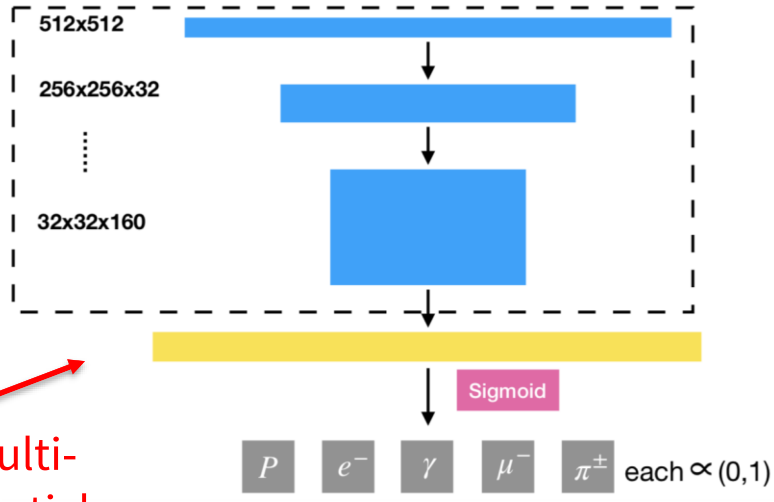
- See Rui An's talk at the reconstruction and Machine learning in Neutrino Experiments workshop (Hamburg, September 2019):
  - https://indico.desy.de/indico/event/21853/session/2/contribution/46

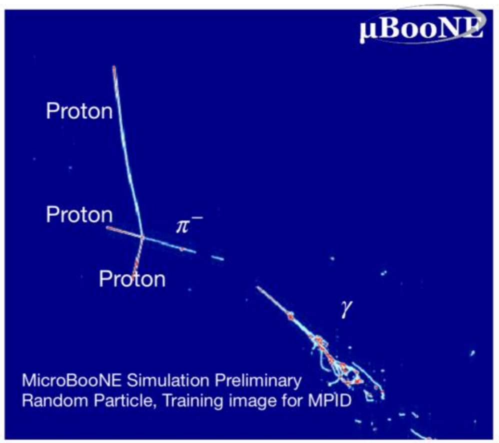
Deep Learning MicroBooNE

- MicroBooNE developed an imaged based event reconstruction chain.



Multi-particle PID Network





arXiv: 1611.05531

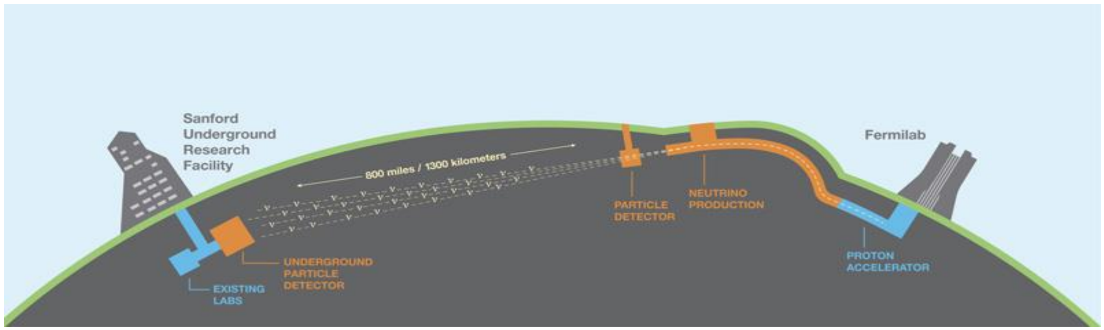
- See Rui An's talk at the reconstruction and Machine learning in Neutrino Experiments workshop (Hamburg, September 2019):
  - https://indico.desy.de/indico/event/21853/session/2/contribution/46

Overview

- Introduction to neutrinos.
- **Machine Learning in neutrino experiments:**
  - NuMI Off-axis νe Appearance (NOvA).
  - Micro Booster Neutrino Experiment (MicroBooNE).
  - **Deep Underground Neutrino Experiment (DUNE).**
  - Tokai to Kamioka (T2K).
- Moving to the edge:
  - Inference on TPU.
  - Inference on FPGA.
- Summary.

DUNE

- The **Deep Underground Neutrino experiment (DUNE)** is a next-generation neutrino oscillation experiment.



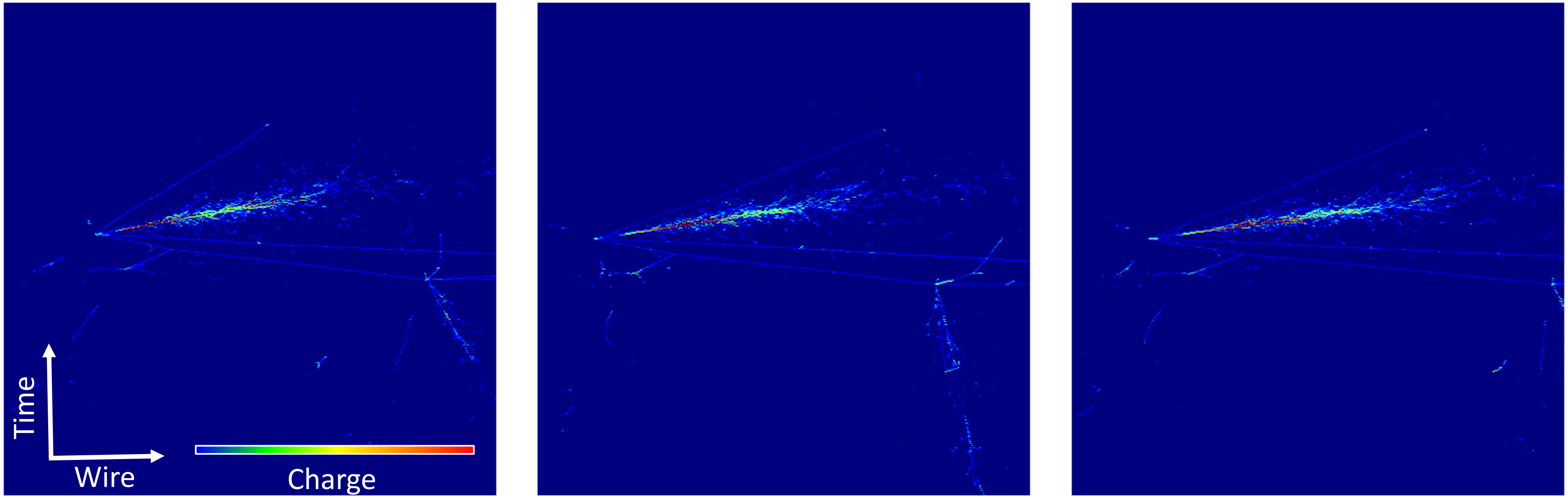
DEEP UNDERGROUND NEUTRINO EXPERIMENT

- Far Detectors (FD) are 800 miles from the neutrino beam source.
  - Four modules, each with 10,000 ton of liquid argon.
- High power neutrino beam produced at Fermilab.
- Measure *CP*-violation.
  - Primary goal: classify the neutrino flavour as ν_e ,ν_μ , ν_τ or NC.
  - We use a convolutional neural network (CNN) to perform the classification.
  - See my talk at the Reconstruction and Machine learning in Neutrino Experiments workshop (Hamburg, September 2019):
    - https://indico.desy.de/indico/event/21853/session/2/contribution/35

Far Detector Data

- The Far Detectors contain three wire readout planes.
  - This provides three “images” of each neutrino interaction.

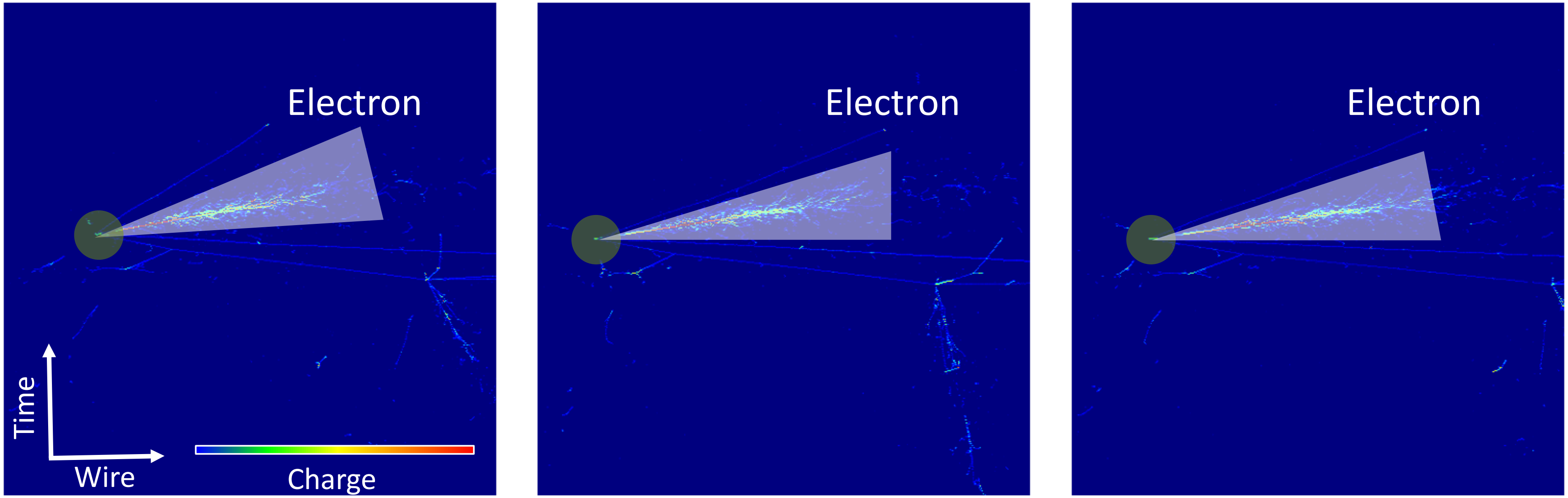
- Official simulated electron neutrino interaction (signal).



Far Detector Data

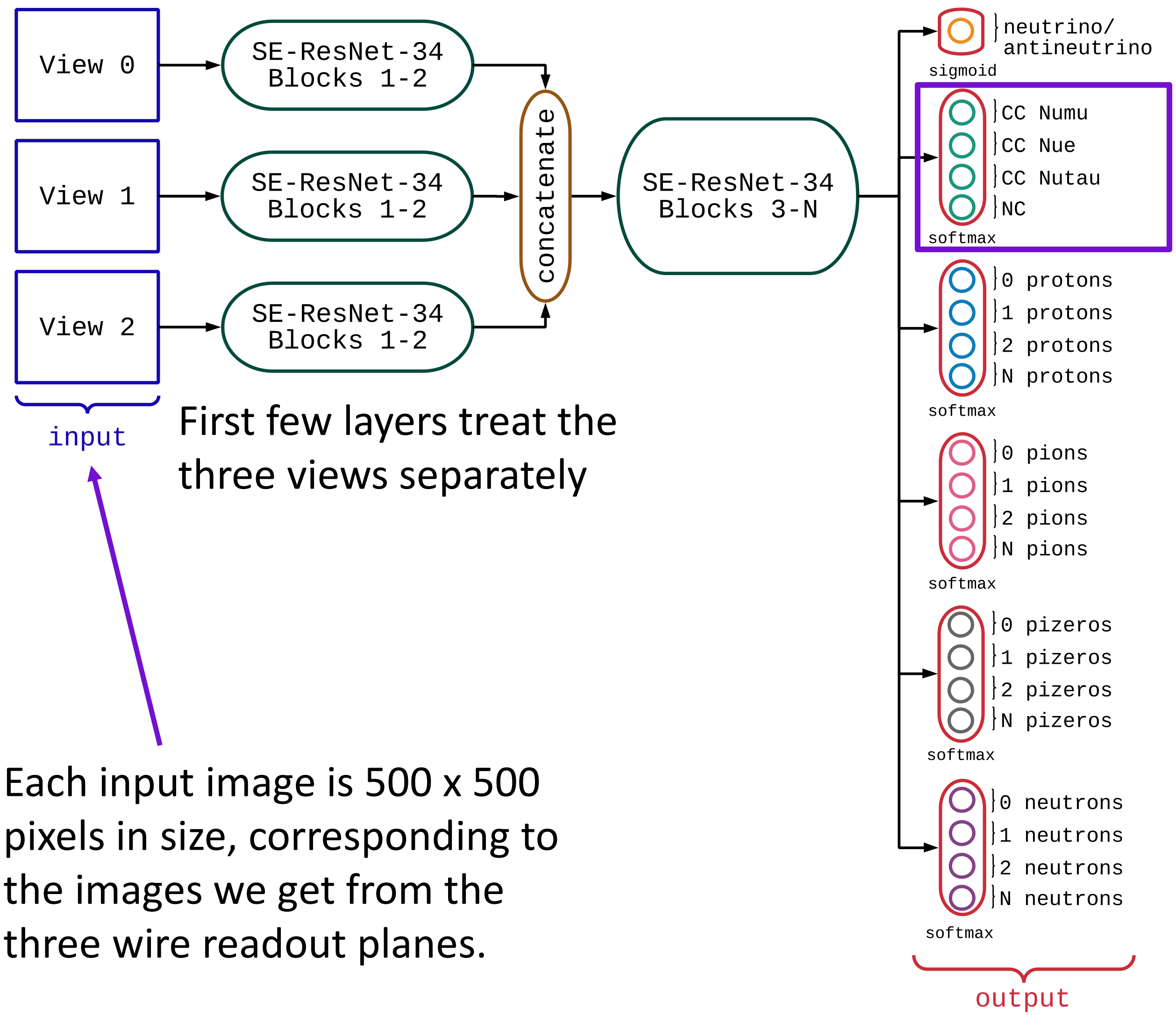
- The Far Detectors contain three wire readout planes.
  - This provides three “images” of each neutrino interaction.

- Official simulated electron neutrino interaction (signal).



- Electron produces the highlighted shower, beginning at the vertex.

DUNE CVN Architecture Overview

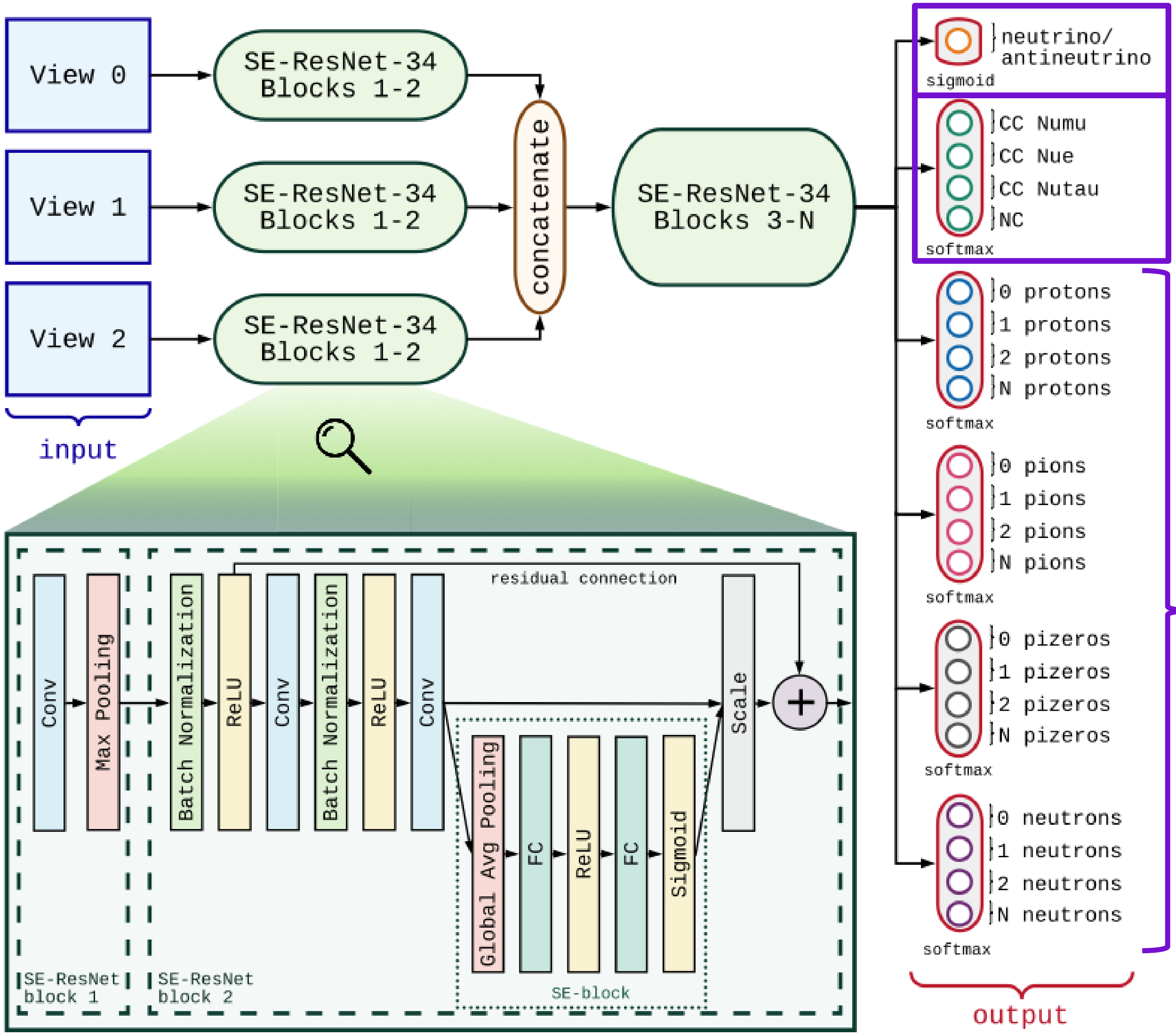


First few layers treat the three views separately

Primary output: Flavour identification

Each input image is 500 x 500 pixels in size, corresponding to the images we get from the three wire readout planes.

DUNE CVN Architecture Overview



Neutrino / antineutrino

Primary output: Flavour identification

Secondary outputs: Particle counting for exclusive final states

Training and Using the CVN

- Use millions of images (~10M images) of simulated neutrino interactions with the true neutrino flavour known.
  - Allows the CNN to learn the features of each type of neutrino interaction.
  - Tested on a fully independent sample.

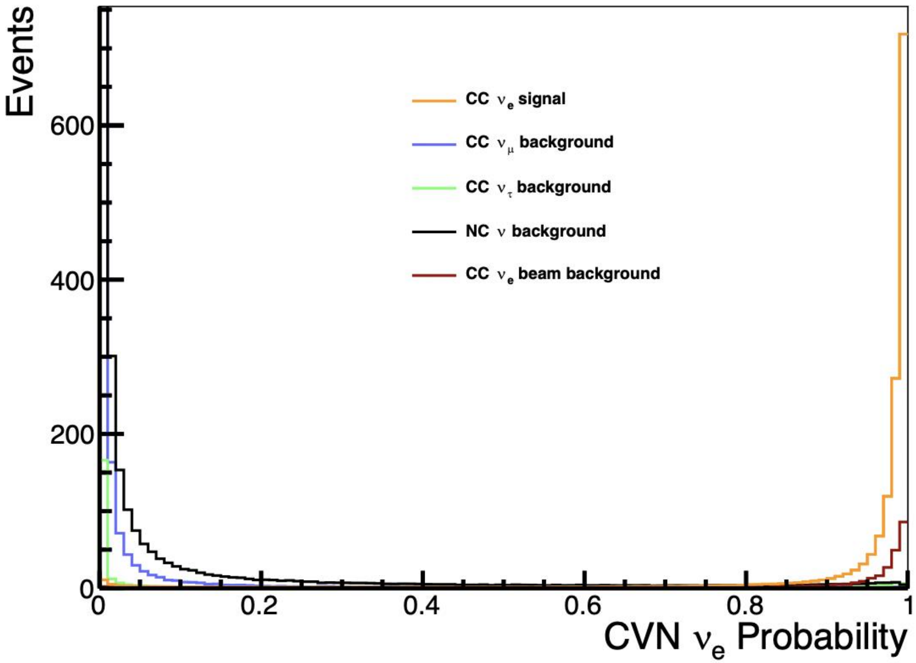
- Once the CVN is trained it is applied to images with no truth information attached – eventually the experimental data.

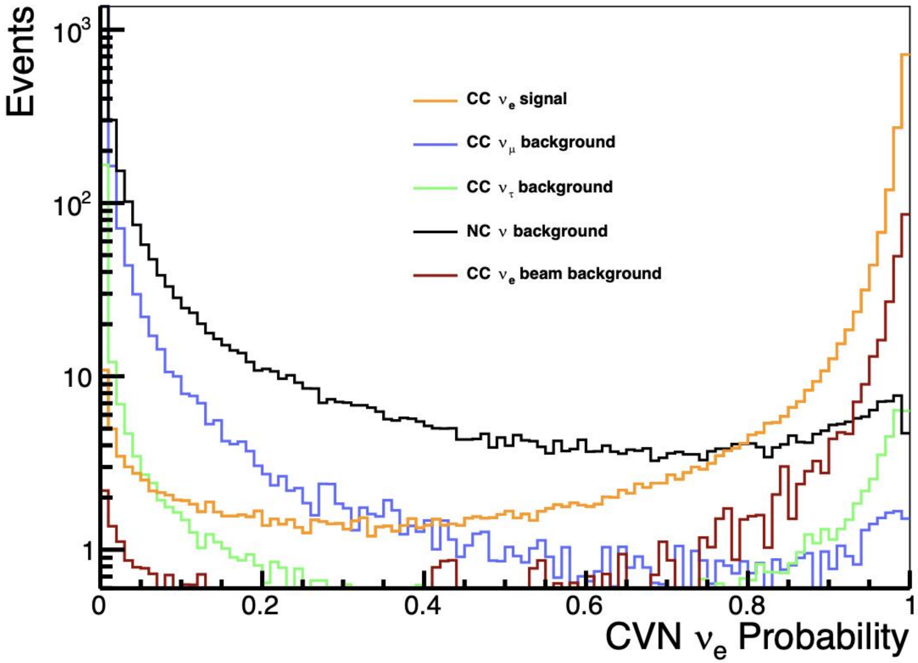
- The CVN gives probabilities for each event to be the following:
  - Charged-current ν_e, ν_μ, ν_τ and neutral-current (all flavours).
  - Outputs sum to one.
  - Use these probabilities for the event selection.

- The ArgoNeuT experiment is using the DUNE CVN for its analysis.
  - See backup slides.

Selecting Electron Neutrinos

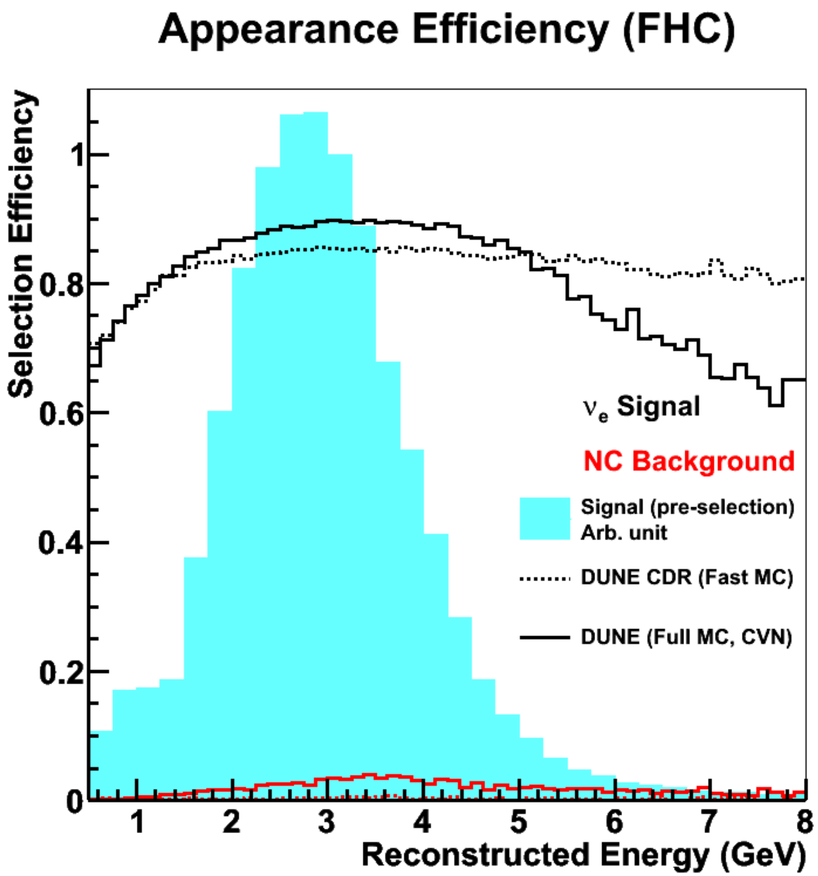
- Electron neutrino probability spectra from the DUNE CVN.
  - Curves combine neutrinos and antineutrinos.





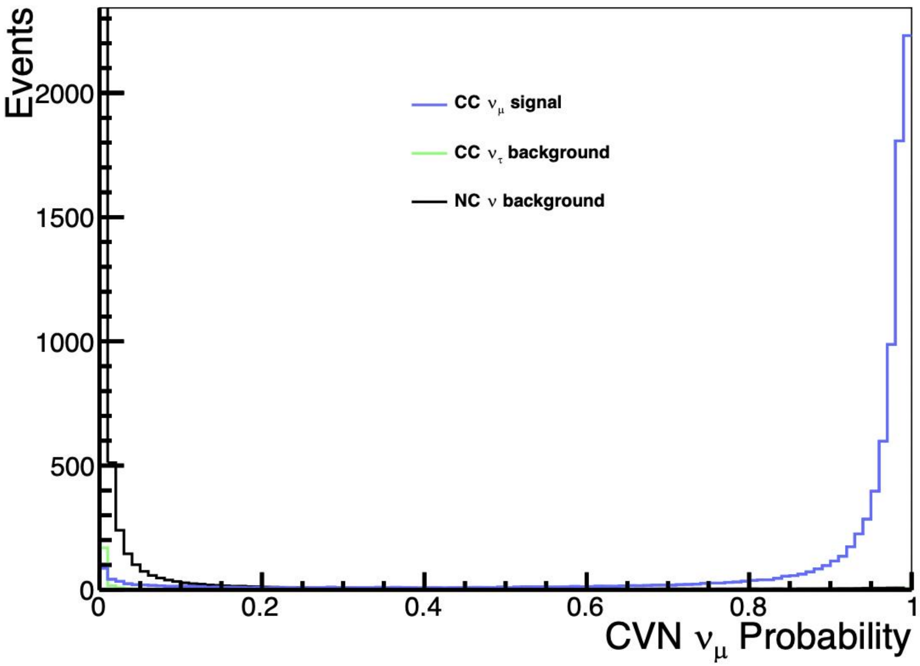
Electron Neutrino Efficiency

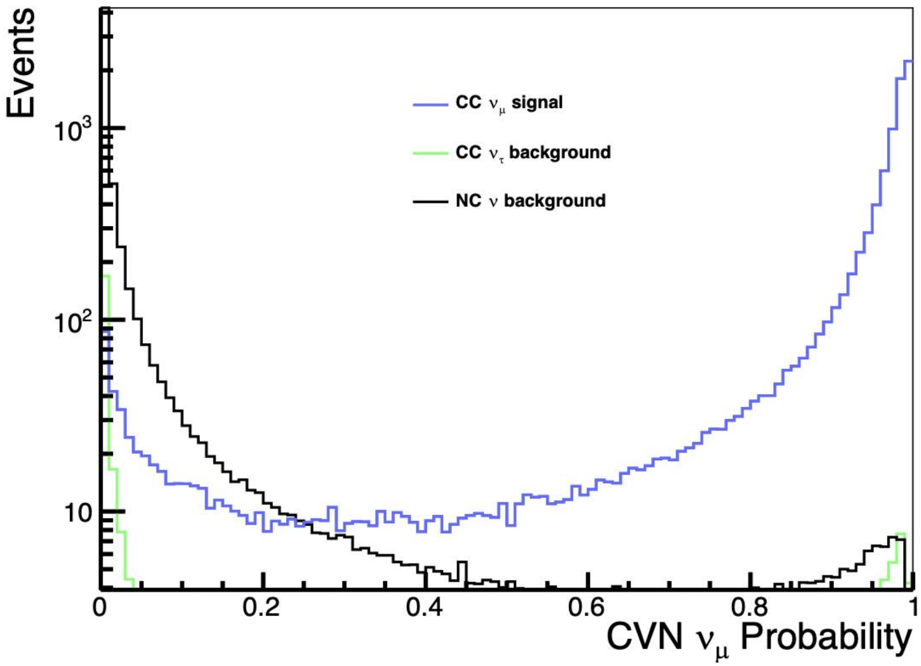
- Select all events that are more than 85% likely to be electron neutrinos.
- Over 90% selection efficiency in the flux peak.
- Efficiency better for antineutrinos due to typically cleaner final state (neutron instead of proton).



Selecting Muon Neutrinos

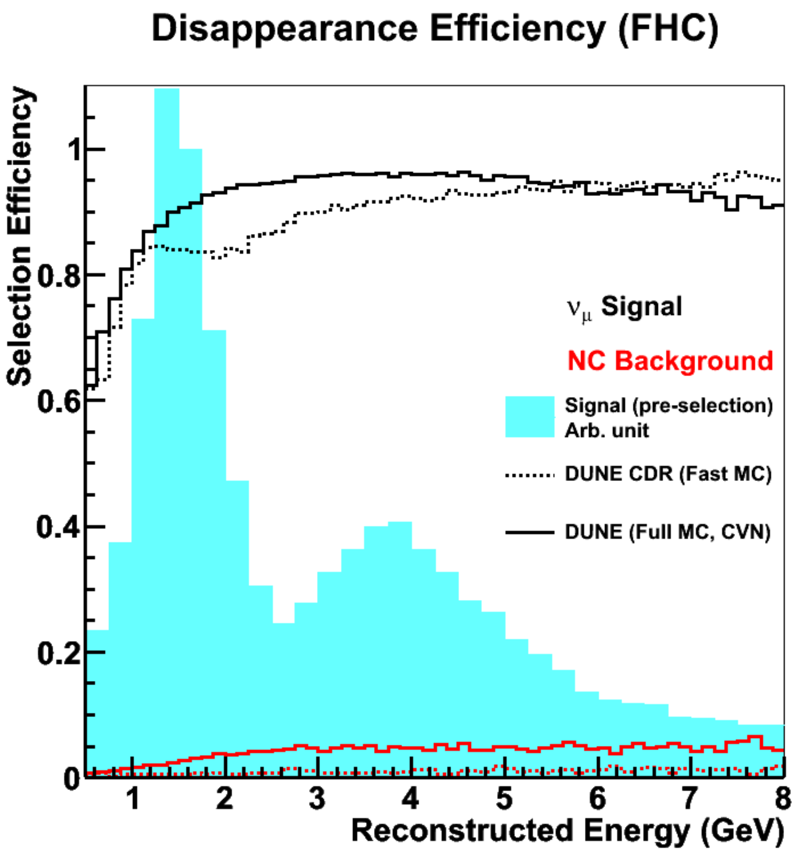
- Muon neutrino probability spectra from the DUNE CVN.
  - Curves combine neutrinos and antineutrinos.





Muon Neutrino Efficiency

- Select all events that are more than 50% likely to be muon neutrinos.
- Over 90% selection efficiency in the flux peak.
- Efficiency better for antineutrinos due to typically cleaner final state (neutron instead of proton).



Generative Models in the DUNE Photon Simulation

- Current DUNE photon simulation:
  - Input parameters: x, y, z.
  - Output: photon detector system as a 6x20 pixel image, where each pixel gives the visibility of one photon detector.

- In the current simulation, the entire geometry is stored in memory.
  - The current library is too big to store in memory.
  - The idea is to have higher resolution and cover a larger volume, both of which will make it impossibly large.

- The approach is to try the fast-simulation segment from our Model-Assisted GAN (arXiv:1812.00879) to speed things up.
  - Modification of a Generative Adversarial Network (GAN).

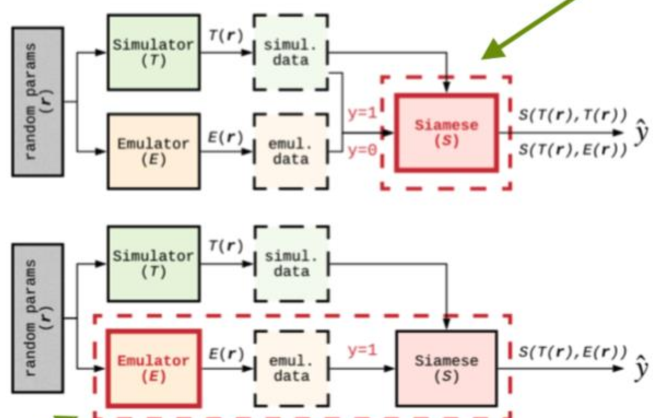
- Trained on 3M images.

Other Machine Learning Applications: DUNE Photon Simulation

- The procedure follows the pre-training stage of the MAGAN:

This network tells us how similar the images are

These are the (x,y,z) positions



(a) Adversarial pre-training: the Siamese network learns the similarity between the simulator and the emulator images; the emulator learns to make emulated data to mimic simulated data.

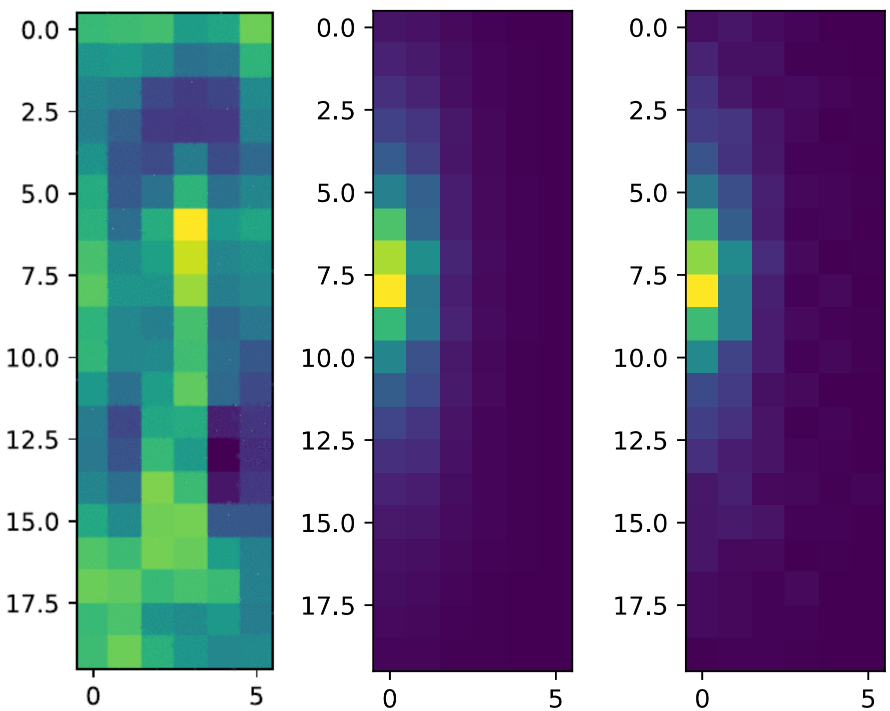
The emulator is the network that mimics the simulator

The Simulator (T) in this case is the photon library

Other Machine Learning Applications: DUNE Photon Simulation

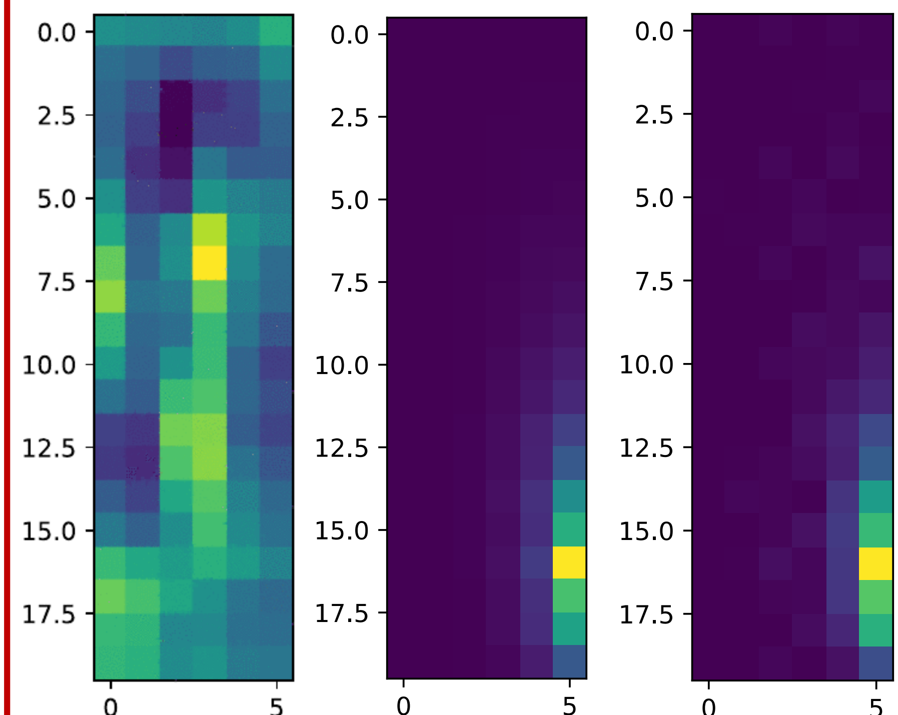
Example 1:

Training process | GAN output (8K iter.) | Simulated



Example 2:

Training process | GAN output (8K iter.) | Simulated

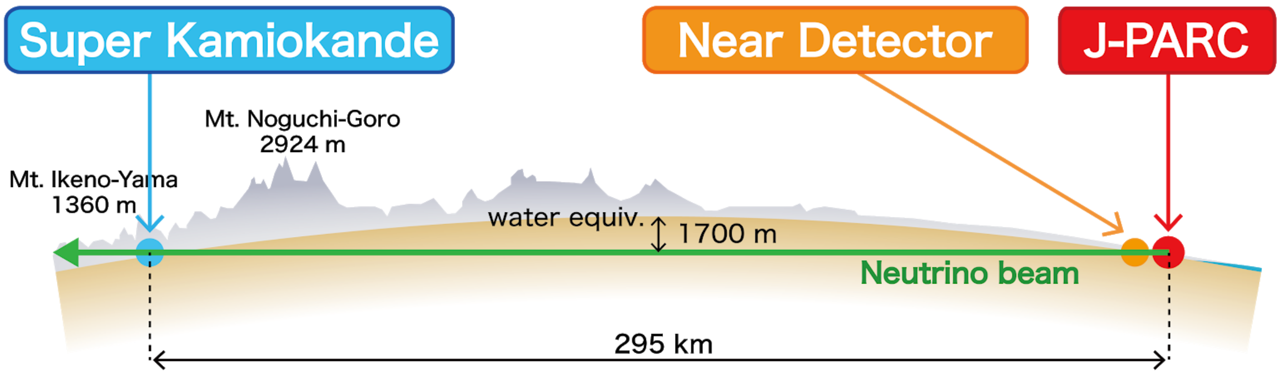


Overview

- Introduction to neutrinos.
- **Machine Learning in neutrino experiments:**
  - NuMI Off-axis νe Appearance (NOvA).
  - Micro Booster Neutrino Experiment (MicroBooNE).
  - Deep Underground Neutrino Experiment (DUNE).
  - **Tokai to Kamioka (T2K).**
- Moving to the edge:
  - Inference on TPU.
  - Inference on FPGA.
- Summary.

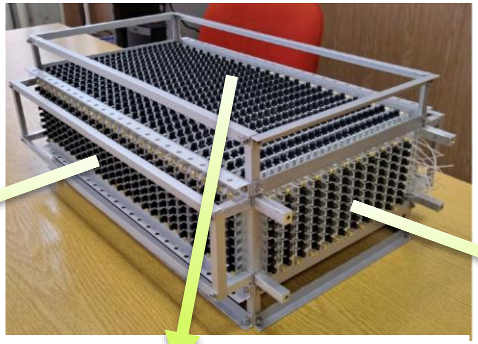
Deep Learning in T2K

- **Tokai to Kamioka (T2K)** is a long-baseline neutrino experiment in Japan, and is studying neutrino oscillations.
- Two separate detectors:
  - Near Detector ND280: measures the number of muon neutrinos in the beam before any oscillations occur and characterizes the physical properties of the beam.
  - Super Kamiokande: very large cylinder of ultra-pure water, detects muon neutrino after oscillating.

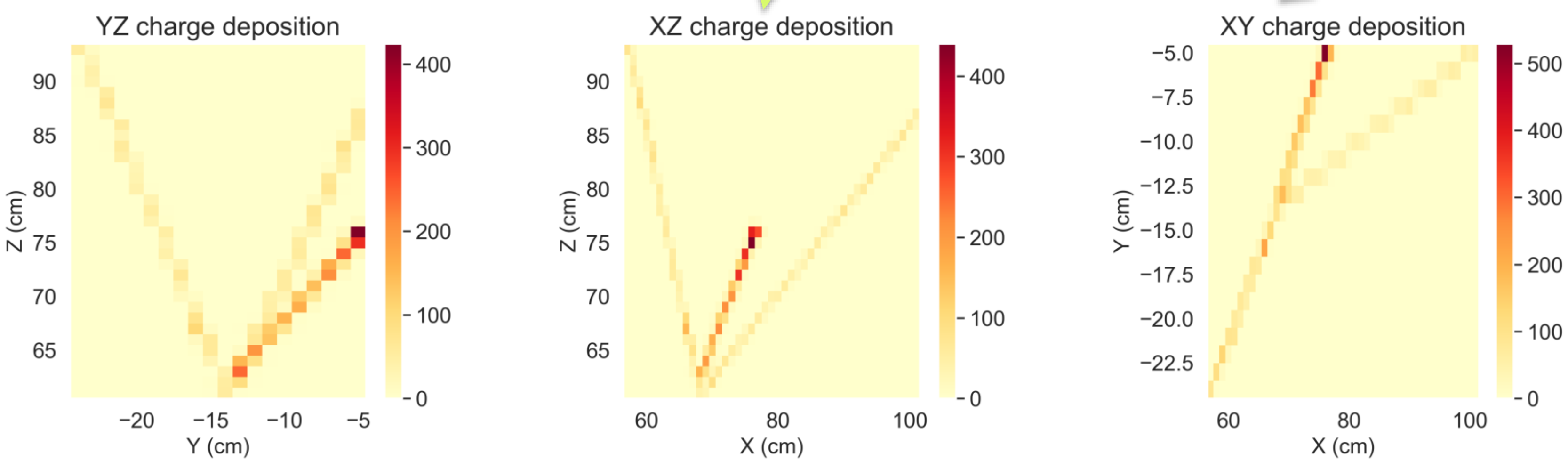


SFGD

- The ND280 upgrade will introduce a new active target, the SuperFGD (SFGD).
- Three 2D charge deposition views (XY, XZ, YZ):

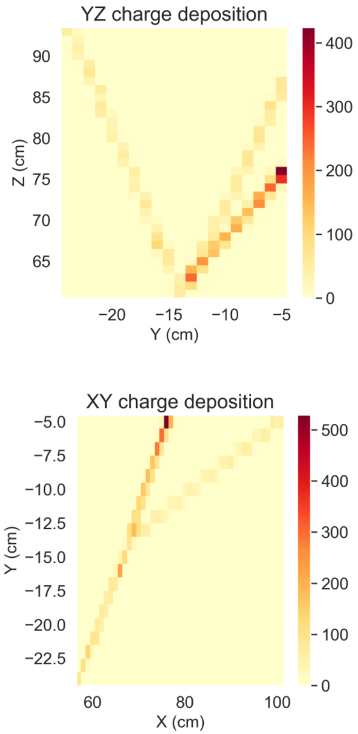


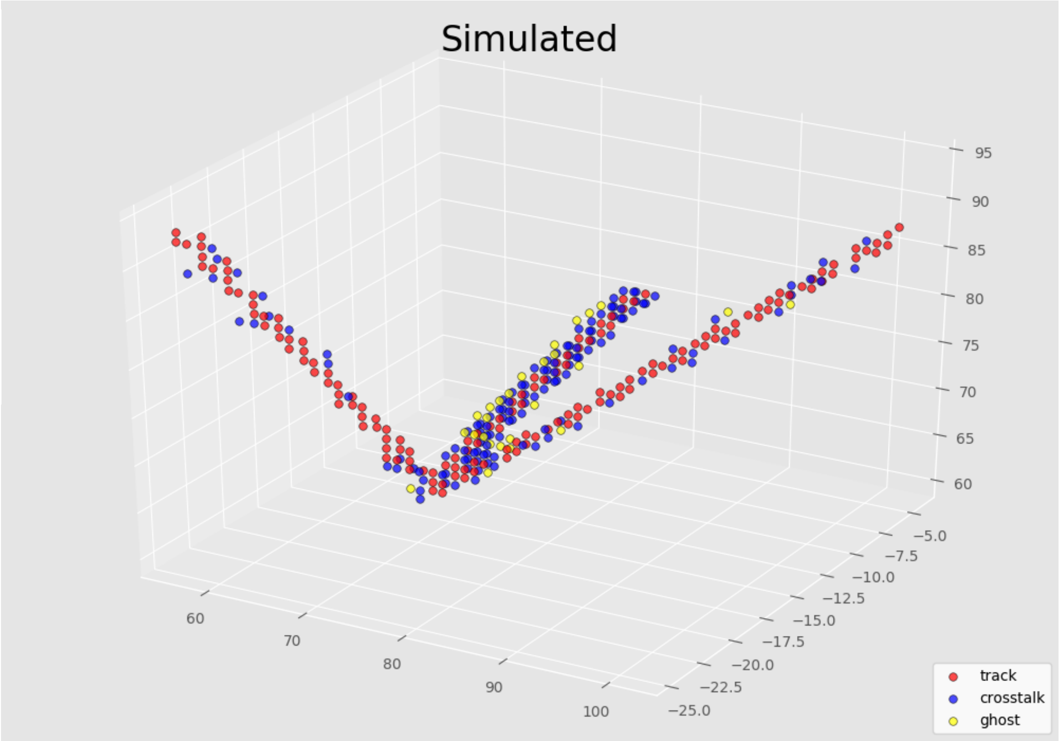
*Simulated event



SFGD 2D to 3D

- Matching the common axis 2-to-2 in the three views XY, XZ, YZ we obtain the 3D information.





- Drawback: non-physical voxels appear due to lack of information during the 2D to 3D reconstruction algorithm, called ghost voxels.

Problem Description

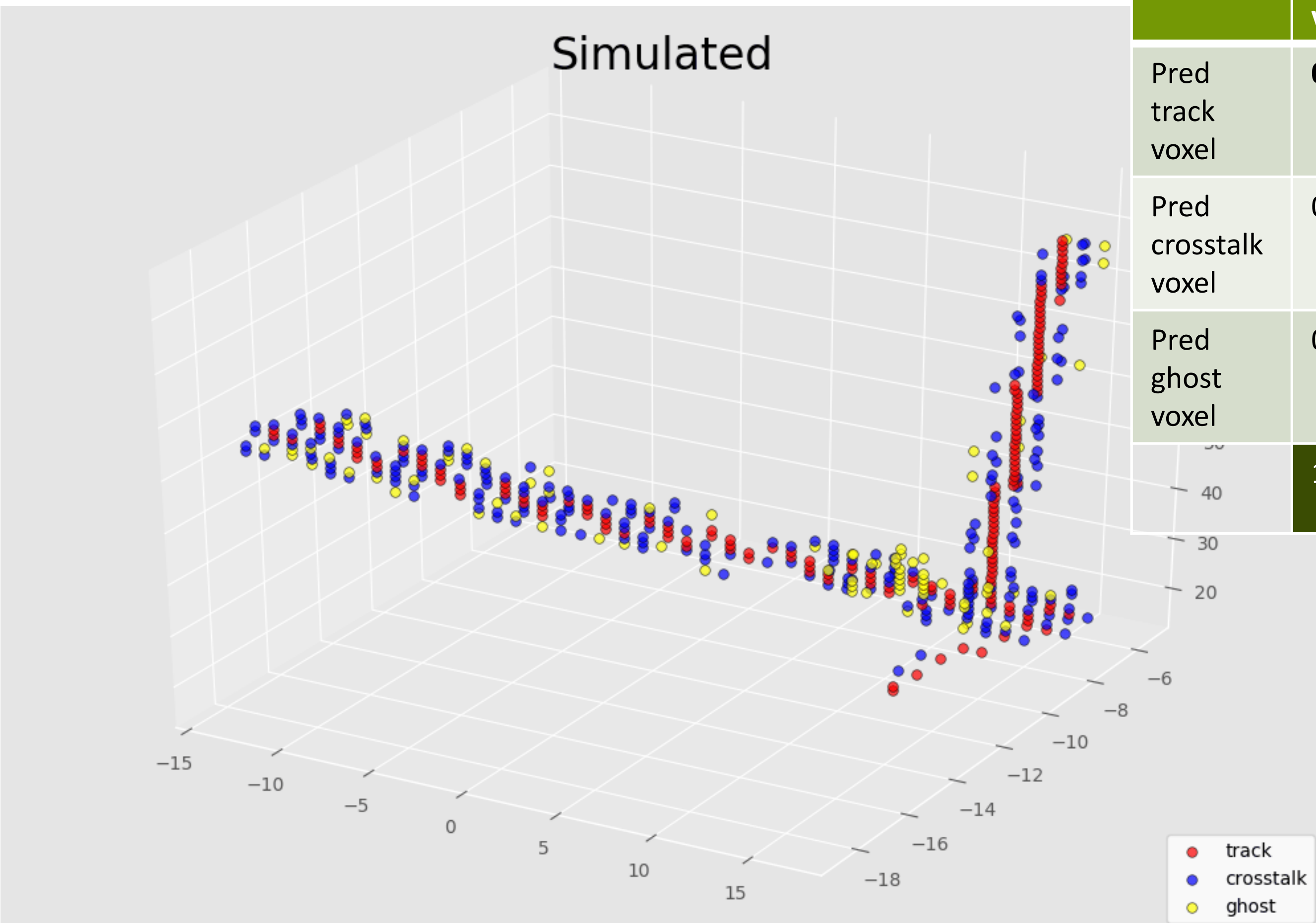
- After 3D-matching, we classify each individual 3D voxel into one of the following:
  - **Track voxel**: a cube where the track has passed by.
  - **Crosstalk voxel**: a cube with a real deposition but where any track has passed through it (all light comes from cube-to-cube optical crosstalk).
  - **Ghost voxel**: a cube that does not have any real deposition and is formed from the 2D ambiguity when reconstructing the 3D event.

- Approach: use a supervised deep learning algorithm (GraphSAGE*) to perform the classification task.
  - The approach based around a graph neural network (GNN) handles each individual voxel as a list of variables (physics information) associated to it.
  - See S. Pina-Otey's talk at the ND 280 Upgrade Meeting: https://indico.cern.ch/event/842568/contributions/3578802/

**W. L. Hamilton, R. Ying, J. Leskovec, Inductive Representation Learning on Large Graphs, arXiv:1706.02216.*

Results (GraphSAGE)

Event : simulated vs pred. (GIF image*):

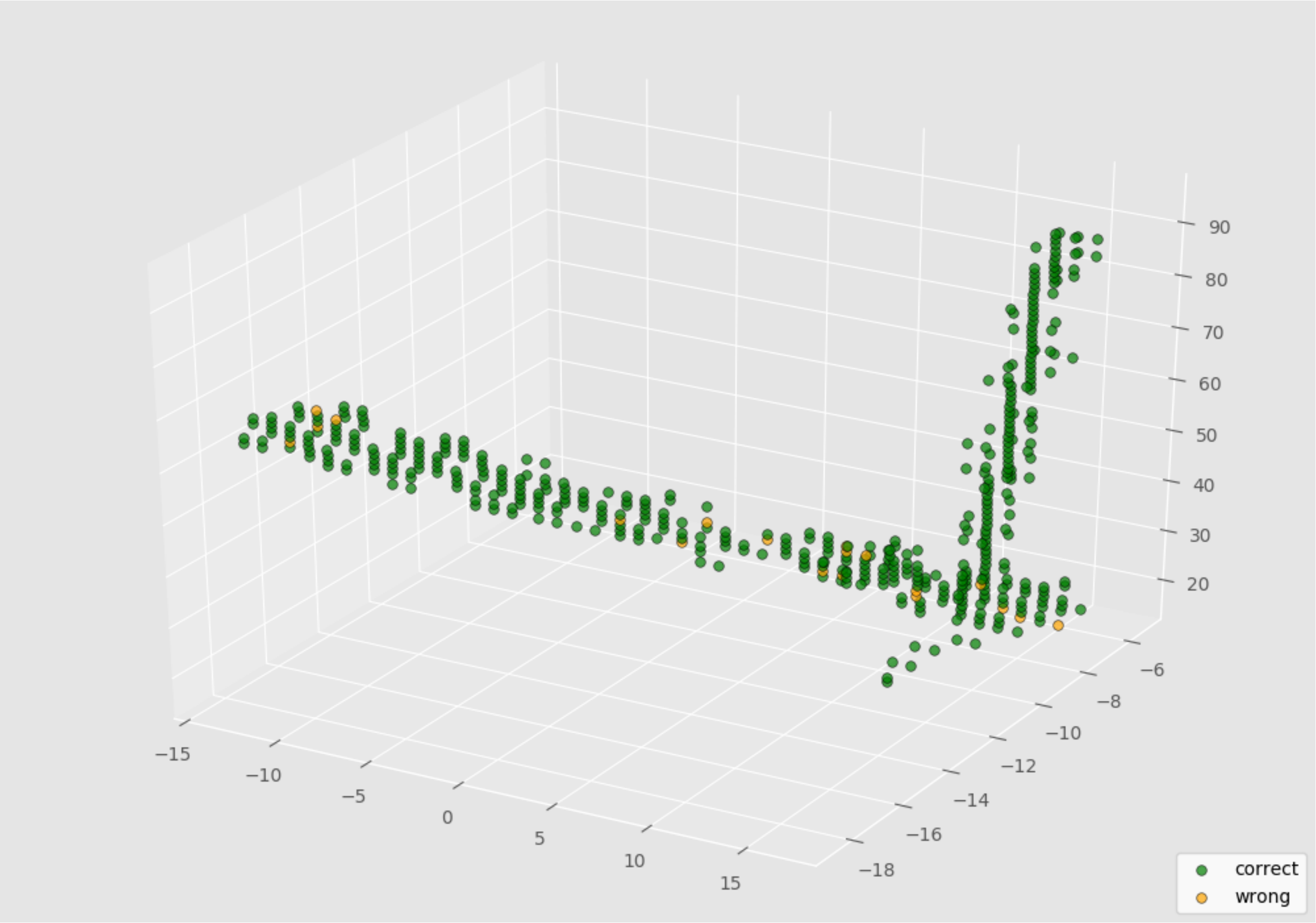


	True track voxel	True crosstalk voxel	True ghost voxel
Pred track voxel	**0.970**	0.022	0.000
Pred crosstalk voxel	0.024	**0.955**	0.050
Pred ghost voxel	0.006	0.023	**0.950**
	1.000	1.000	1.000

**slide show* to see the animated GIF

Results (GraphSAGE)

Event: correct classified voxels.



Overview

- Introduction to neutrinos.
- Machine Learning in neutrino experiments:
  - NuMI Off-axis ve Appearance (NOvA).
  - Micro Booster Neutrino Experiment (MicroBooNE).
  - Deep Underground Neutrino Experiment (DUNE).
  - Tokai to Kamioka (T2K).
- **Moving to the edge:**
  - Inference on TPU.
  - Inference on FPGA.
- Summary.

Moving to the Edge

- **Edge computing** is the practice of processing data near the edge of your network, where the data is being generated, instead of in a centralized data-processing warehouse.

- Fermilab-Google collaboration.
  - Goal: study the performance of deep learning tasks using different hardware: CPU, GPU, and Edge TPU (Tensor Processing Unit from Google).

- CERN Openlab-Micron collaboration.
  - Goal: develop new deep learning models and use Micron's FPGAs for fast inference online reconstruction for DUNE and CMS.

Overview

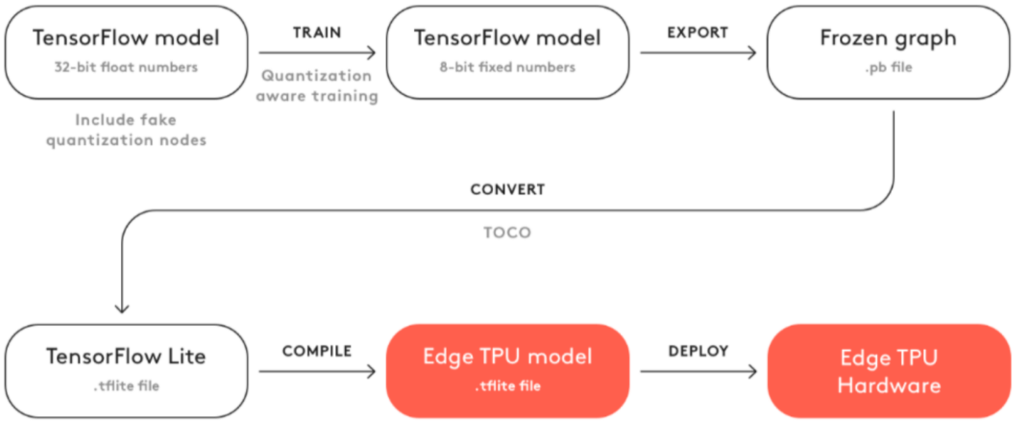
- Introduction to neutrinos.
- Machine Learning in neutrino experiments:
  - NuMI Off-axis νe Appearance (NOvA).
  - Micro Booster Neutrino Experiment (MicroBooNE).
  - Deep Underground Neutrino Experiment (DUNE).
  - Tokai to Kamioka (T2K).
- **Moving to the edge:**
  - **Inference on TPU.**
  - Inference on FPGA.
- Summary.

Fermilab - Google Collaboration

- Specifications:

	CPU	GPU	Edge TPU
Model	Intel(R) Core(TM) i5-6500 CPU @ 3.20GHz	NVIDIA Tesla K80 (from Google Colab)	Coral Edge TPU
TDP*	65 w (16 w per core)	300 w	2 w
Price (USD)	200	5,000	80

- Generating the right model:



- See S. Vergani's talk at the September 2019 DUNE Collaboration meeting:
  - https://indico.fnal.gov/event/21445/session/1/contribution/95/material/slides/0.pdf

*Thermal Design Power (TDP) represents the average power, in watts, the processor dissipates when operating at Base Frequency with all cores active under an Intel-defined, high-complexity workload.

Results

- Tested using ResNet-50 on MNIST dataset:

	CPU (Intel(R) Core(TM) i5-6500 CPU @ 3.20GHz)	GPU (NVIDIA Tesla K80)	Coral Edge TPU
Categorical accuracy	97%	97%	95%
Total inference time (10k images)	142 s	14.7 s	356 s
Inference per image	14 ms	1.5 ms	35 ms

- Tested using the DUNE CVN for neutrino identification (50 test images):

	CPU (Intel(R) Core(TM) i5-6500 CPU @ 3.20GHz)	GPU (NVIDIA Tesla K80)	Coral Edge TPU
Categorical accuracy	88%	86%	88%
Total inference time (10k images)	22 s	1 s	5 s
Inference per image	431 ms	20 ms	100 ms

- Costs: $cost/inference = time/inference \times TDP \times cost\ of\ energy = K \times cost\ of\ energy$

	CPU (Intel(R) Core(TM) i5-6500 CPU @ 3.20GHz)	GPU (NVIDIA Tesla K80)	Coral Edge TPU
K factor (ResNet-50 on MNIST 56x56 images)	0.21	0.45	0.07
K factor (DUNE 500x500 images)	6.9	6	0.2

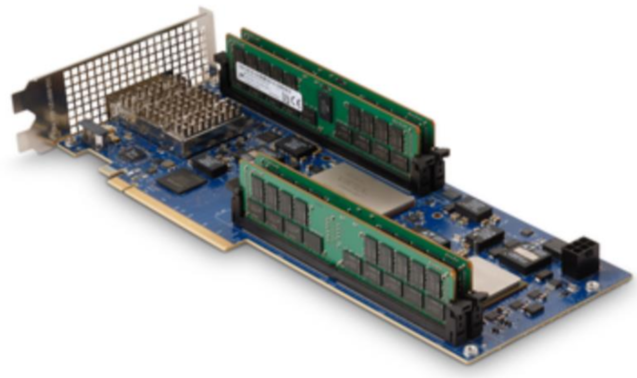
- GPU appears to be by far the fastest piece of hardware.
- Edge TPU performs better with bigger images
- Edge TPU showed the smallest cost per inference and CPU showed the biggest cost per inference.

Overview

- Introduction to neutrinos.
- Machine Learning in neutrino experiments:
  - NuMI Off-axis νe Appearance (NOvA).
  - Micro Booster Neutrino Experiment (MicroBooNE).
  - Deep Underground Neutrino Experiment (DUNE).
  - Tokai to Kamioka (T2K).
- **Moving to the edge:**
  - Inference on TPU.
  - **Inference on FPGA.**
- Summary.

CERN Openlab - Micron Collaboration

- SB-852 (FPGA ready for machine learning!):
  - Xilinx Virtex Ultrascale+ UV9P.
  - 64GB DDR4 SODIMM.
  - High-bandwidth.
  - Low-latency.



- FWDNXT:
  - No need to VHDL programming.
  - Any framework*.
  - Any network*.
- Already ran the DUNE CVN on the FPGA.
  - Same results in GPU and FPGA.
- Future plans:
  - Measure time and energy.
  - Integrate the FPGA in the protoDUNE-SP DAQ.
  - Test how far we can go in the data selection or even in fast online reconstruction.
- See M.J. Rodríguez's talk at the DUNE Data Selection Working Group Meeting:
  - https://indico.fnal.gov/event/21955/contribution/0/material/slides/0.pdf

Overview

- Introduction to neutrinos.
- Machine Learning in neutrino experiments:
  - NuMI Off-axis νe Appearance (NOvA).
  - Micro Booster Neutrino Experiment (MicroBooNE).
  - Deep Underground Neutrino Experiment (DUNE).
  - Tokai to Kamioka (T2K).
- Moving to the edge:
  - Inference on TPU.
  - Inference on FPGA.
- **Summary.**

Summary

- Machine Learning, and Deep Learning in particular, provide many powerful mechanisms for classifying input data from many different fields, including high-energy physics and neutrino experiments in particular.

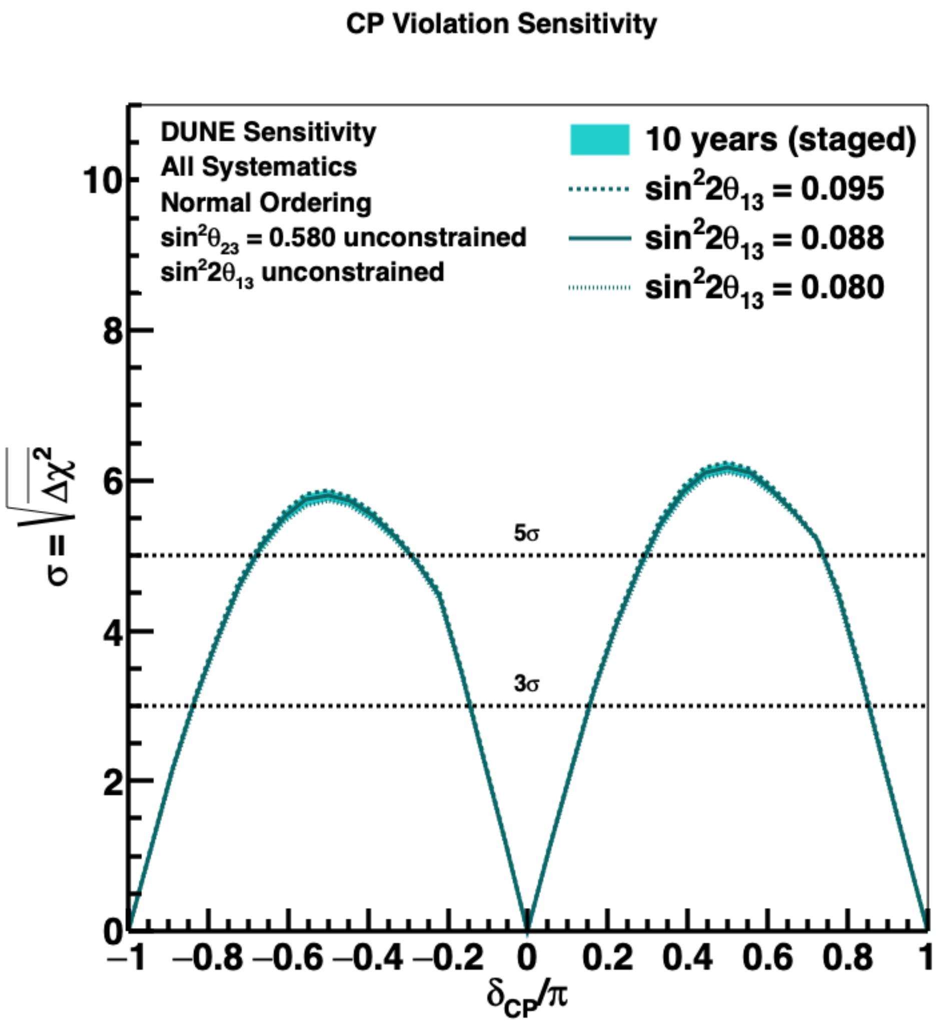
- Showed some machine learning applications in a number of worldwide neutrino experiments: NOvA, MicroBooNE, DUNE, [ArgoNeuT] and T2K.
  - Many other neutrino experiments: Icarus, MINOS, MINERvA, IceCube, K2K, MiniBooNE…
  - Inference via edge computing : two current projects.

- Current (and future) work focused on using more sophisticated techniques (e.g., sparse CNNs, Graph Neural Networks) for semantic segmentation of neutrino interactions.

Backup Slides

DUNE CP-Violation Sensitivity – Dec 2018

- Same selection criteria:
  - ν_e selection: $P(\nu_e) > 85\%$.
  - ν_μ selection: $P(\nu_\mu) > 50\%$.
- Exceeded the DUNE conceptual design report sensitivity.
  - Very big milestone for DUNE!



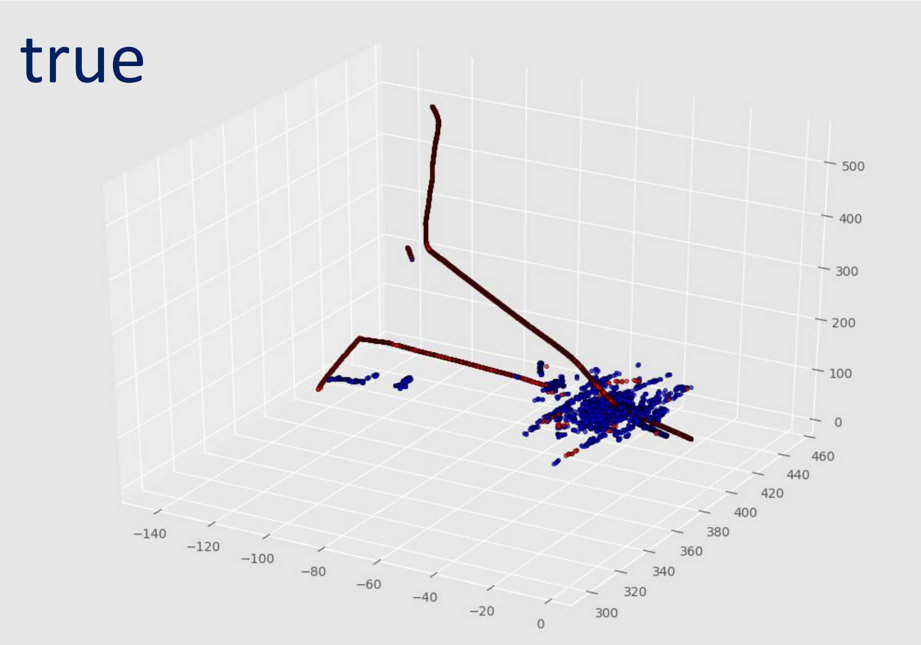
Other Machine Learning Applications: ProtoDUNE Track vs Shower Identification

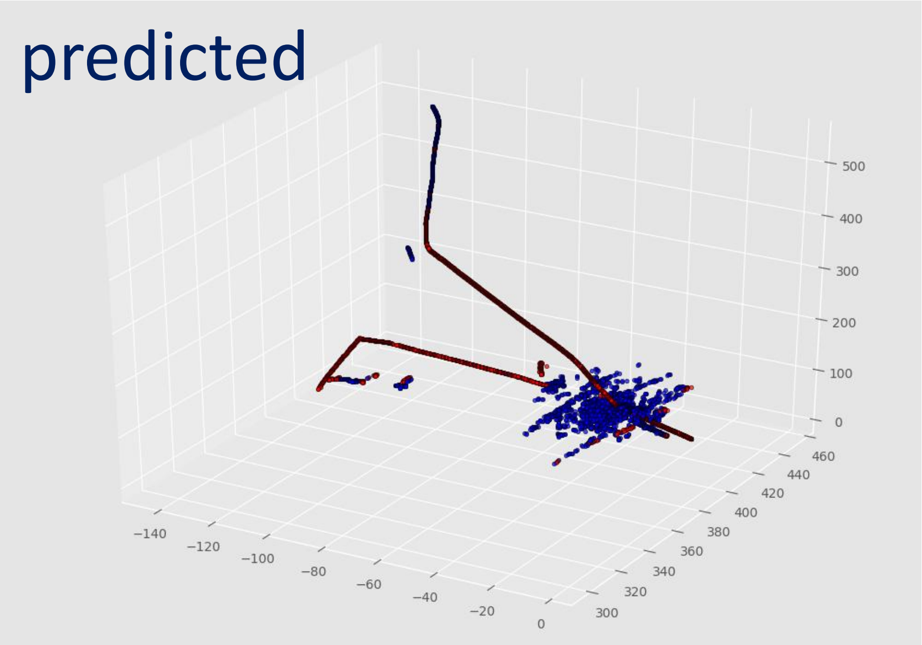
- The approach is to use the GraphSAGE algorithm (*arXiv:1706.02216*) to label track vs shower hits in ProtoDUNE 3D images.
  - Each image is internally stored as a graph.
    - Graphs are much smaller than images (less size on disk).
    - Neighbourhood (adjacency) can be used during the learning process.
  - Model has a constant number of parameters.
    - Training on graphs of different sizes.
    - Inference on unseen graphs (graphs that were not seen during the training).
  - Can be applied to any node of the graph.

Other Machine Learning Applications: ProtoDUNE Track vs Shower Identification

- Example:

Red: track hits. **Blue**: shower hits.





T2K SFGD Results (GraphSAGE)

- **Training on 6k events.**
- **Confusion matrix (from 60k events):**

	True track voxels	True crosstalk voxels	True ghost voxels	
Pred track voxels	**5,140,704**	167,236	13,089	5,321,029
Pred crosstalk voxels	286,890	**5,001,600**	124,886	5,413,376
Pred ghost voxels	16,561	140,140	**1,401,551**	1,558,252
	5,444,155	5,308,976	1,539,526	**12,292,657**

Observation: a 0.96% of voxels cannot be calssified by GraphSAGE due to not having any edge in the graph.

T2K SFGD Results (GraphSAGE)

- **Purity (left) vs Efficiency (right)**

	True track voxels	True crosstalk voxels	True ghost voxels	
Pred track voxels	**0.9661**	0.0314	0.0025	1.0000
Pred crosstalk voxels	0.0530	**0.9239**	0.0231	1.0000
Pred ghost voxels	0.0106	0.0899	**0.8995**	1.0000

	True track voxels	True crosstalk voxels	True ghost voxels
Pred track hits	**0.9443**	0.0315	0.0085
Pred crosstalk hits	0.0527	**0.9421**	0.0811
Pred ghost hits	0.0030	0.0264	**0.9104**
	1.0000	1.0000	1.0000

ArgoNeuT

- **Argon Neutrino Teststand (ArgoNeuT)** is a joint NSF/DOE R&D project at Fermilab to expose a small-scale liquid argon time projection chamber (LArTPC) to the NuMI neutrino beam.
- Just starting using the same DUNE CVN architecture for distinguishing between ν_e and ν_μ interactions.
- Some differences with the DUNE version:
  - Two input views instead of three.
  - Images of size 240x1800 instead of 500x500.
- Very promising preliminary results.
  - ~80% ν_μ accuracy.
  - ~93% ν_e accuracy.