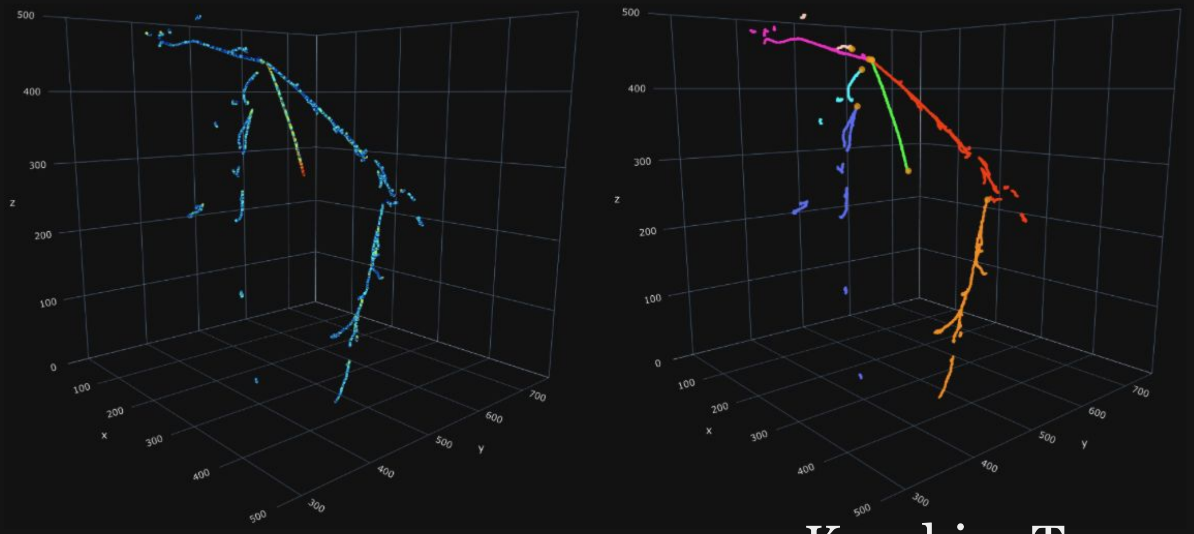
Machine learning to find ghost particles in big data

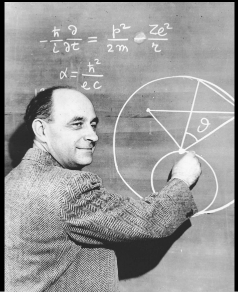
Original image credit: xkcd



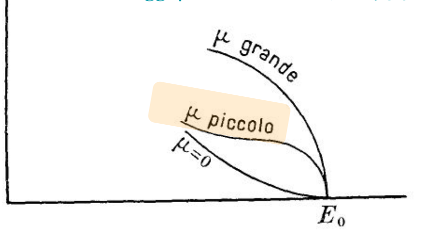
Kazuhiro Terao
SLAC National Accelerator Laboratory
Inst. for AI & Fundamental Interactions (colloquium)

Neutrinos

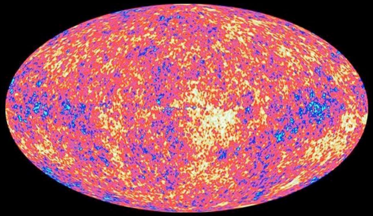
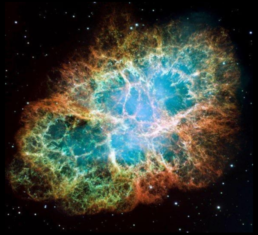
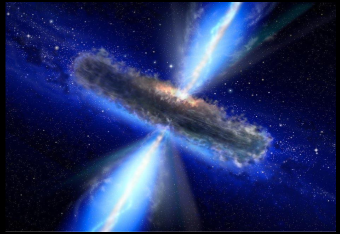
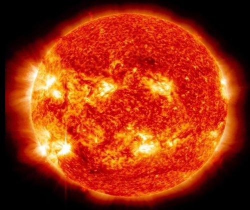
(weakly interacting slim ghosts)



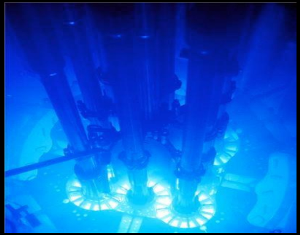
[Fermi, E. Tentativo di una Teoria Dei Raggi β. Nuovo Cim 11, 1 (1934)]



Neutrinos

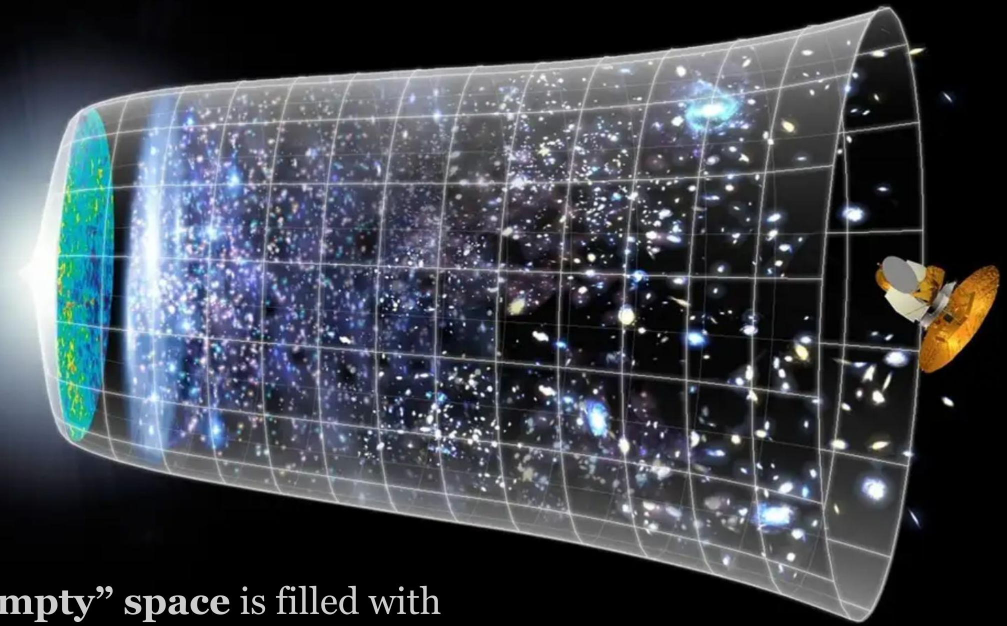










Neutrinos
are produced
everywhere = natural
physics messengers

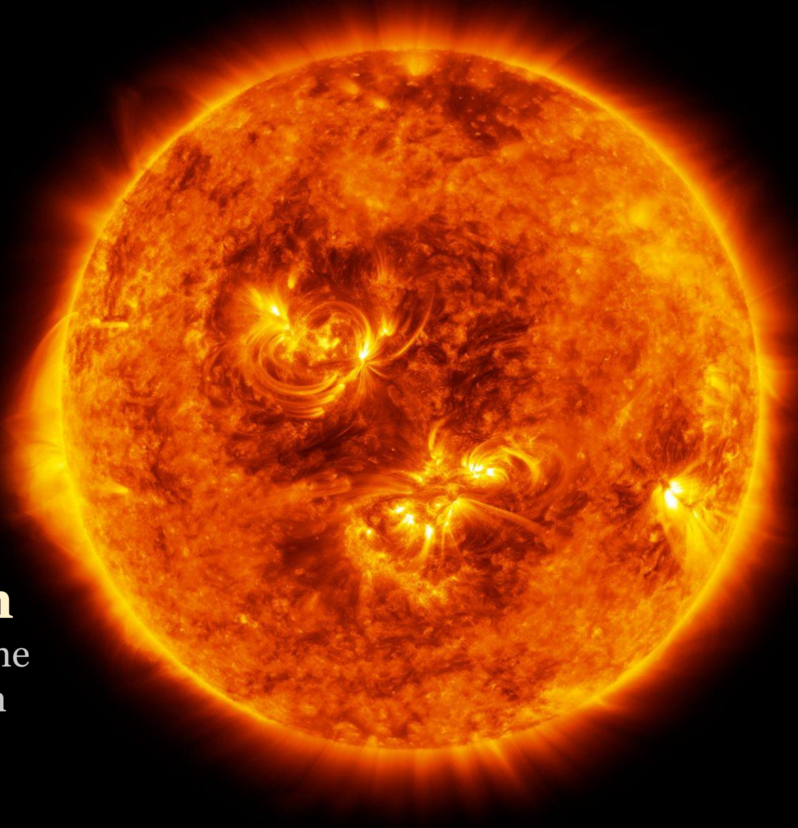
Neutrinos



The "empty" space is filled with
100/cm^3 relic neutrinos
produced 0.2 second after Big Bang

Neutrinos
are the most
abandoned matter
particles we know in
the universe

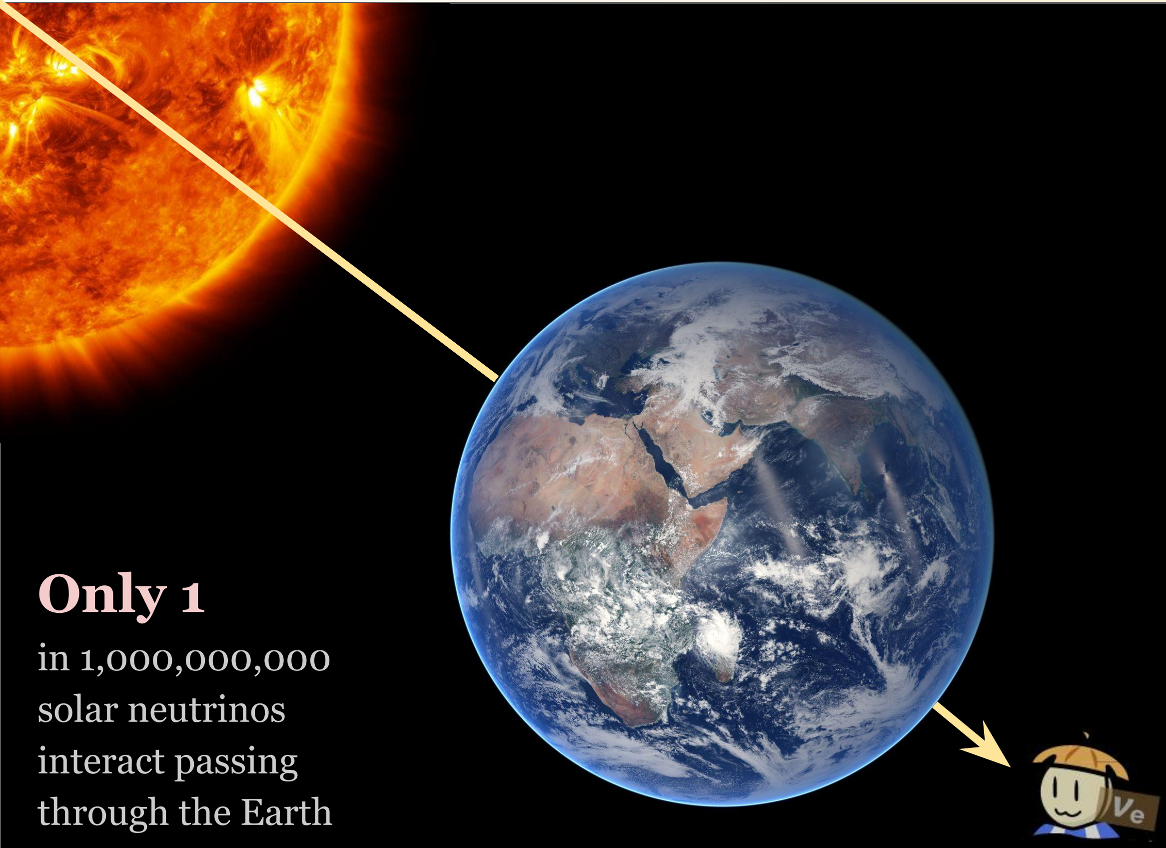
Neutrinos



100 billion
neutrinos from the Sun pass through your thumbnail every second

Neutrinos
are the most
abandoned matter
particles we know in
the universe

Neutrinos



Only 1

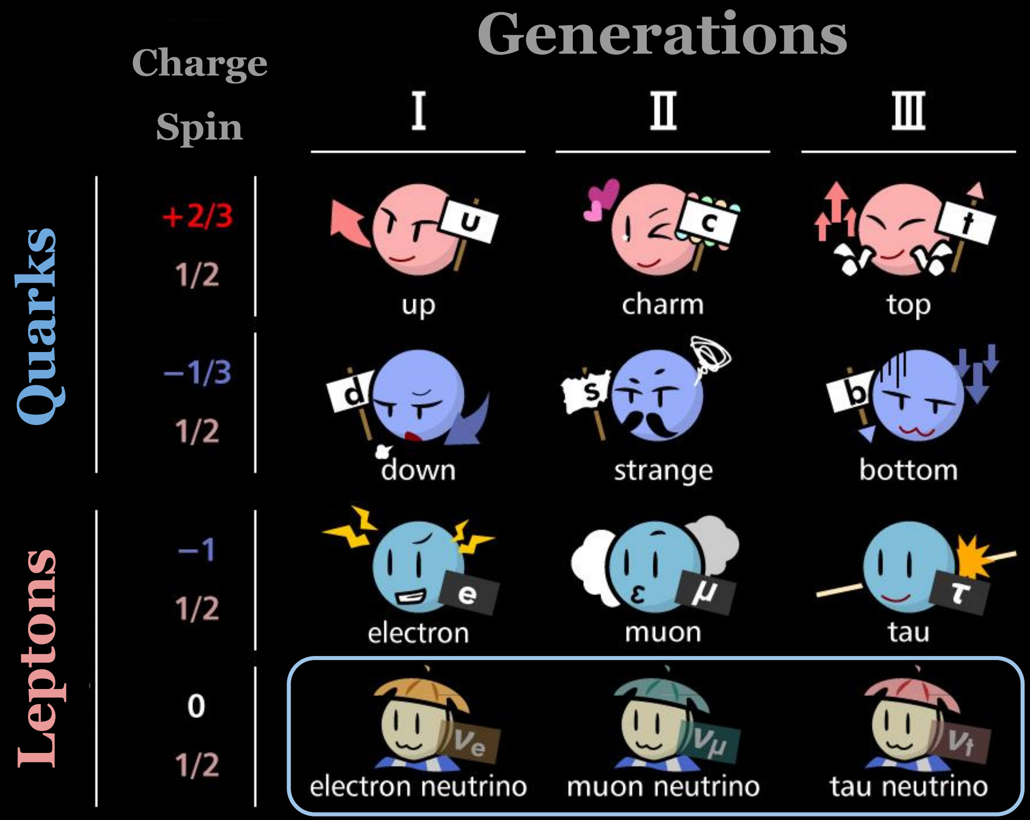
in 1,000,000,000 solar neutrinos interact passing through the Earth

Neutrinos

are **ghostly**

ML for Analyzing Big Image Data in Neutrino Experiments

Neutrinos



Neutrinos

the most ghostly and lightest matter particle in the "**Standard Model**"

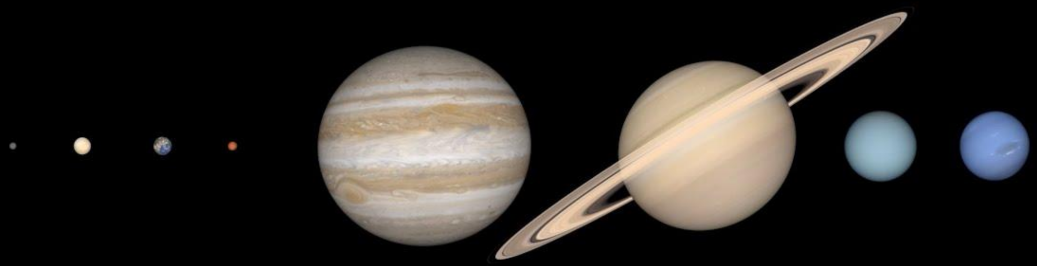
Image credit: higgstan

ML for Analyzing Big Image Data in Neutrino Experiments

Neutrinos

Proton ~ 1,000,000,000 eV

Up quark ~2,000,000 eV



Electron ~ 500,000 eV

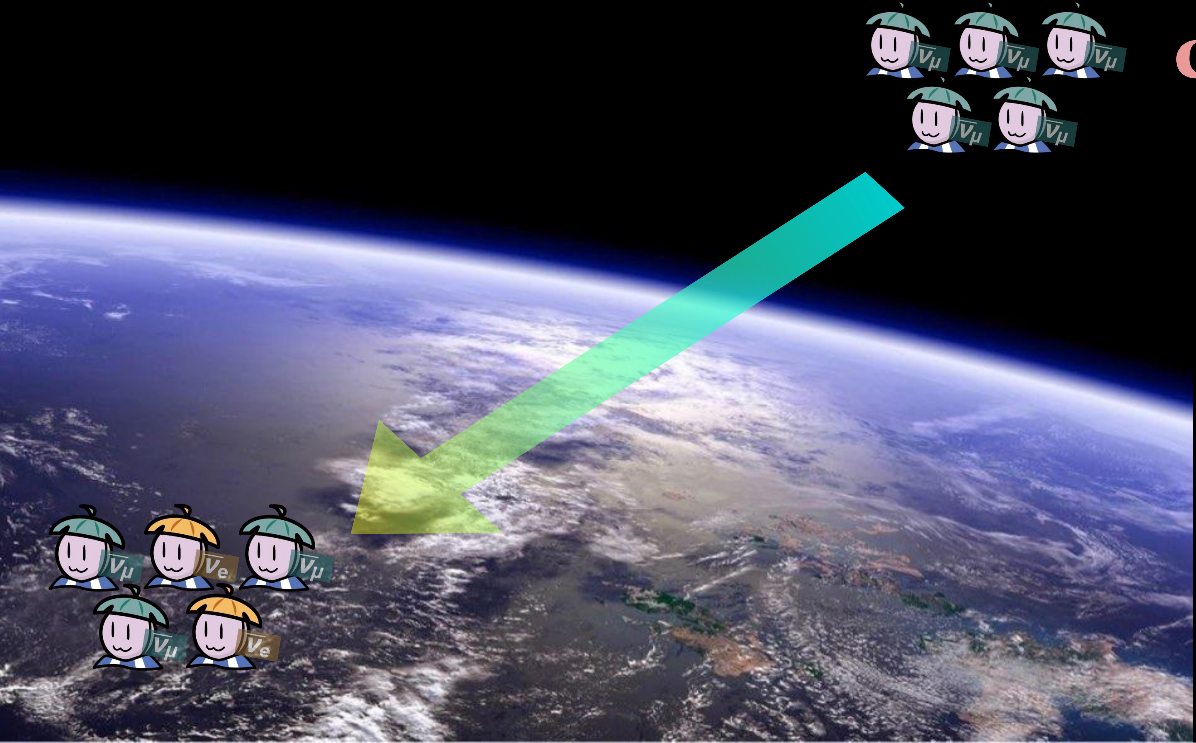
Neutrino < 0.2 eV

Neutrinos
amusingly too small mass
compared to siblings

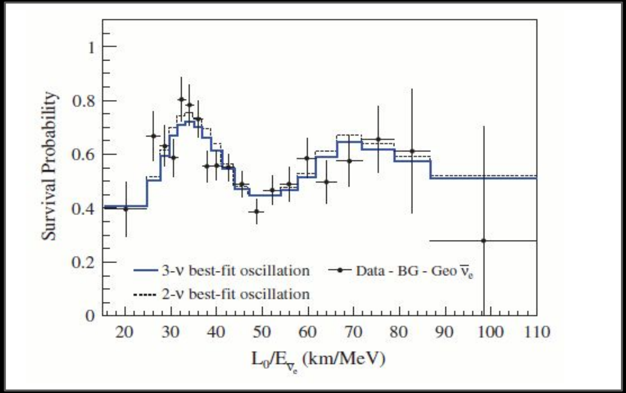
Image credit: higgstan

ML for Analyzing Big Image Data in Neutrino Experiments

Neutrinos



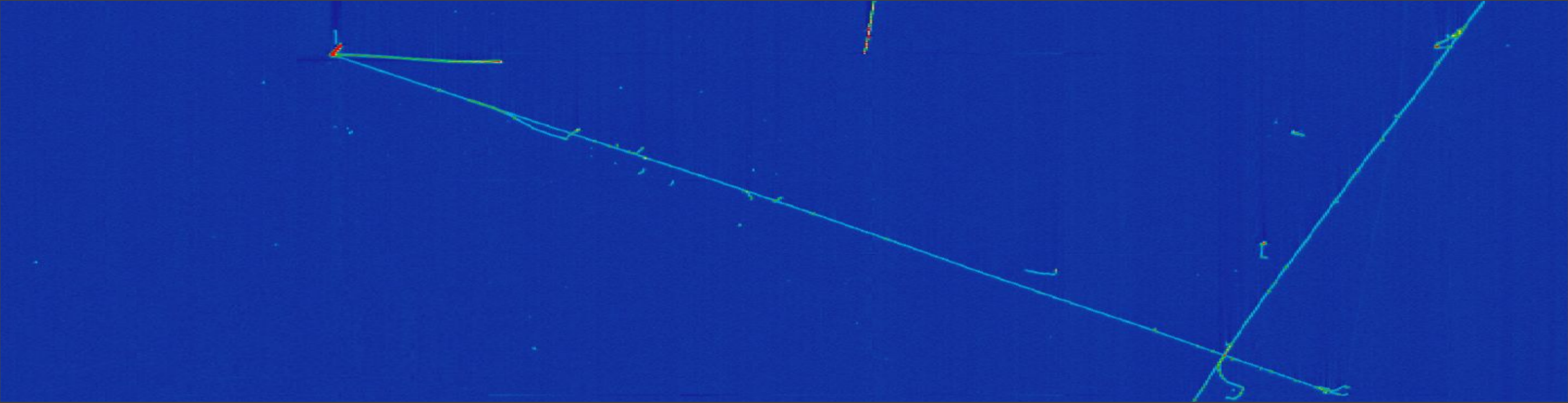
Neutrinos
change their flavor
as they travel
(oscillation)



Neutrinos

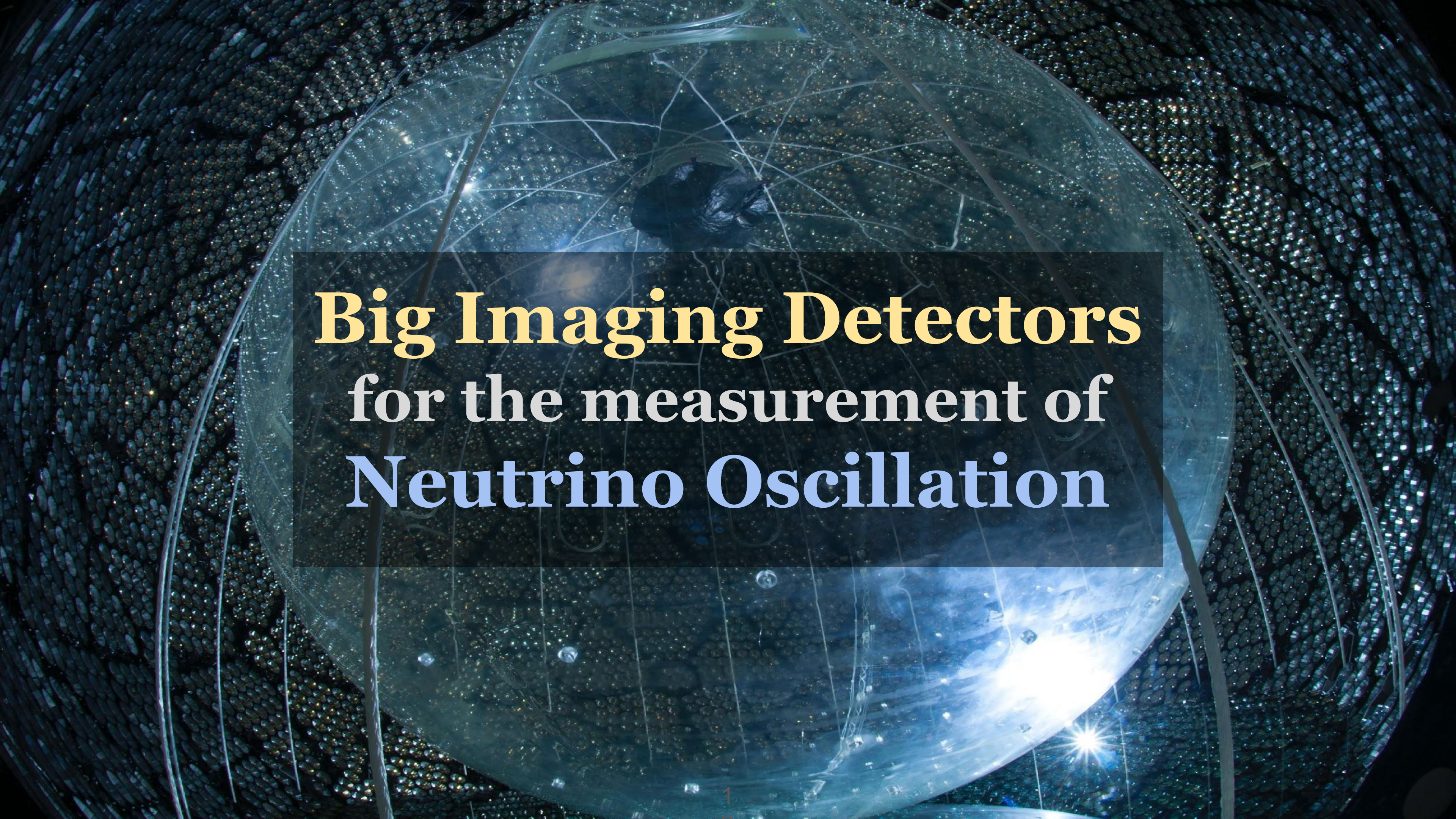
Flavor Oscillation

measurements might shed light to a question, how the universe has evolved to the present?



Outline

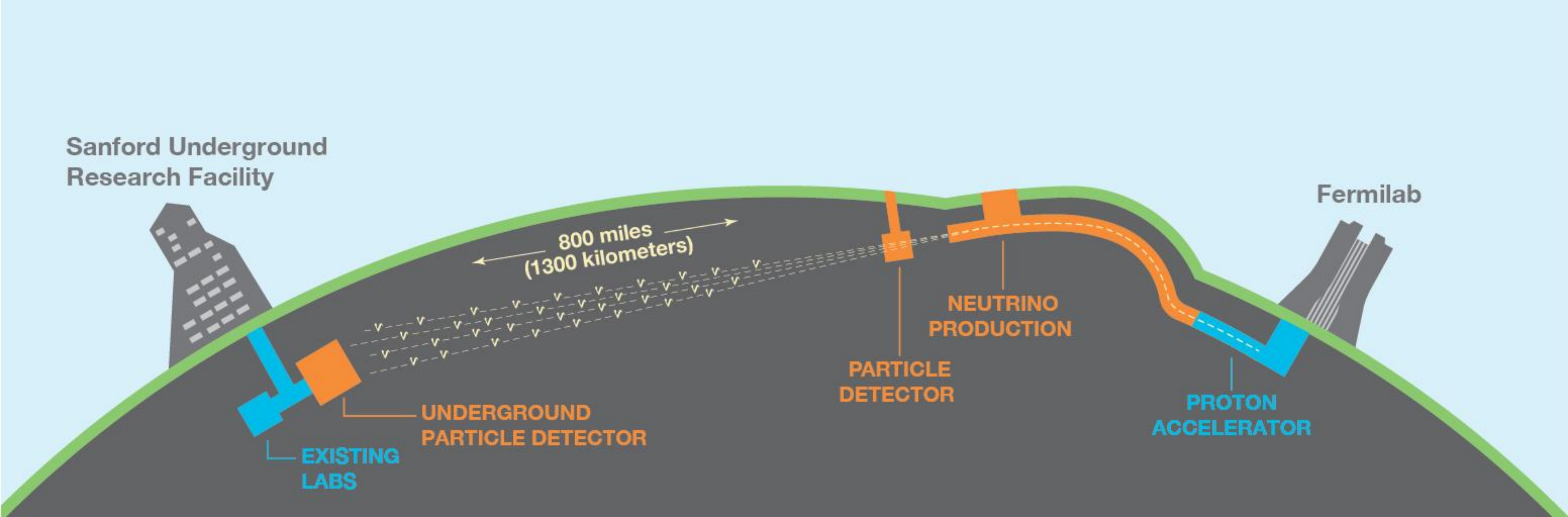
1. Neutrinos oscillation experiments
2. Machine learning for big image data from neutrino detectors
3. Machine learning for physics model optimization
4. Summary



Big Imaging Detectors
for the measurement of
Neutrino Oscillation

Neutrino Oscillation Experiments

two detectors to measure oscillated & unoscillated flux



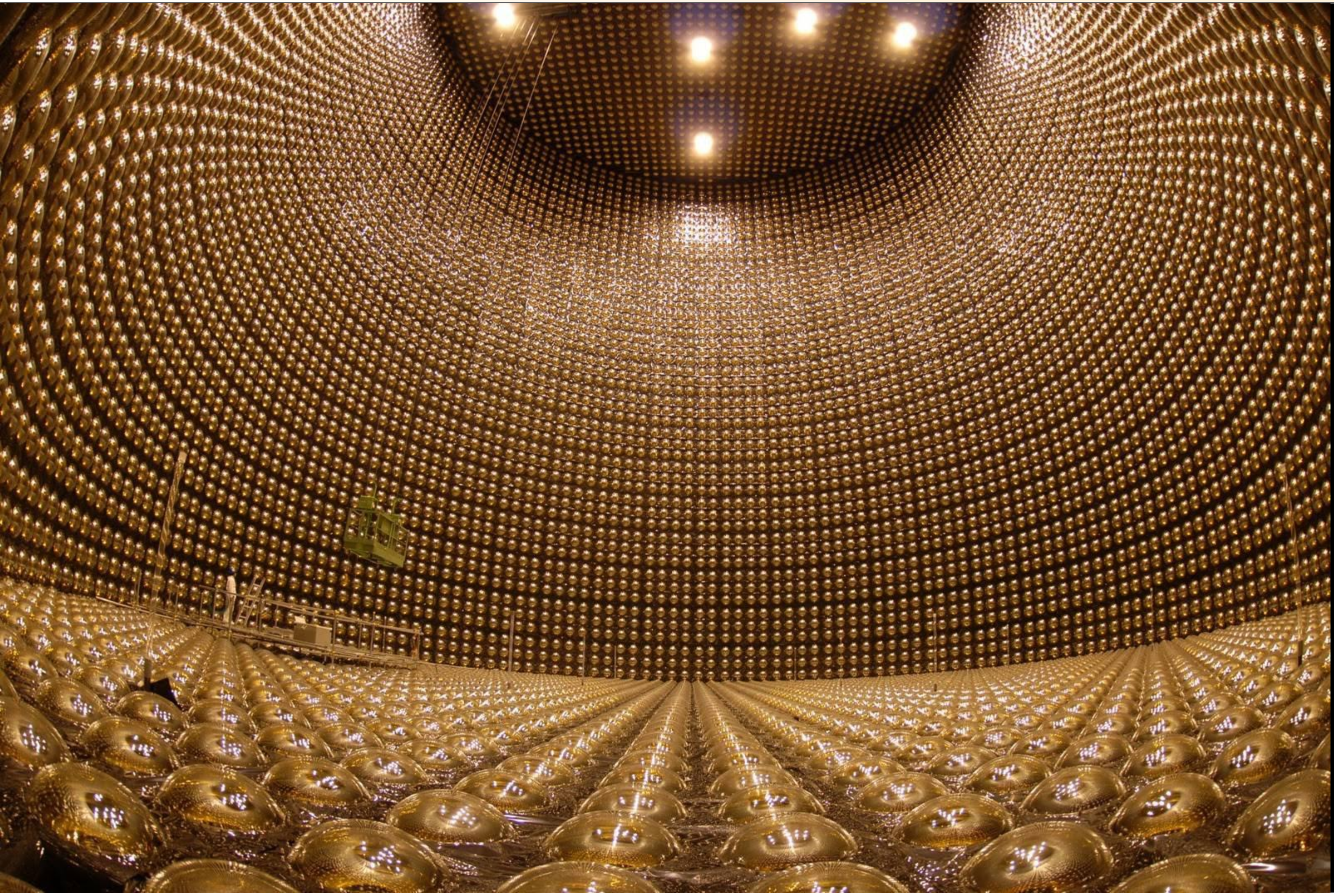
ML for Analyzing Big Image Data in Neutrino Experiments

Neutrino Oscillation Experiments

SLAC

Accelerator
well understood
neutrino source
for precision
measurement

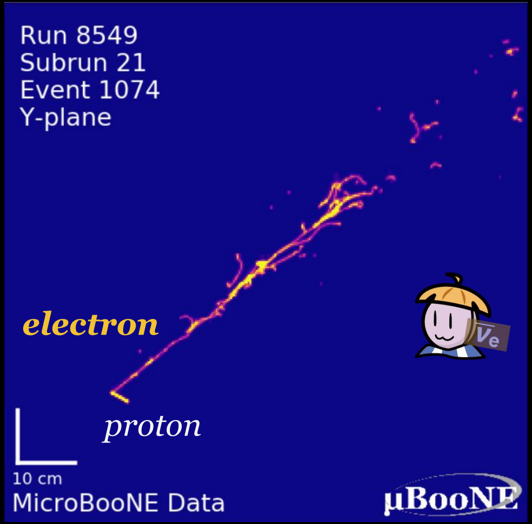
ML for Analyzing Big Image Data in Neutrino Experiments
Neutrino Oscillation Experiments

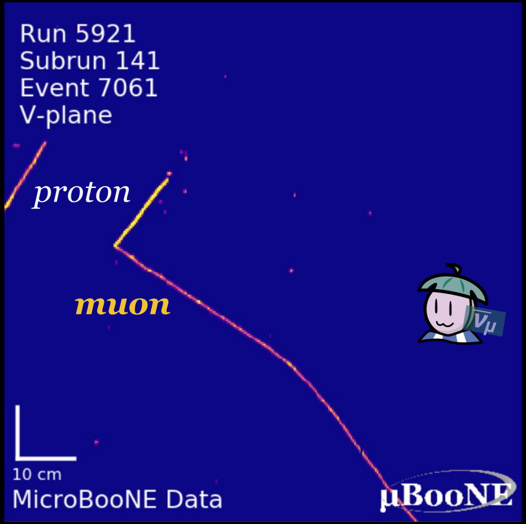


Detectors must be **BIG**

50,000 ton
ultra-pure water watched by 11,000 PMTs in Super-Kamiokande (1996)

Neutrino Oscillation Experiments

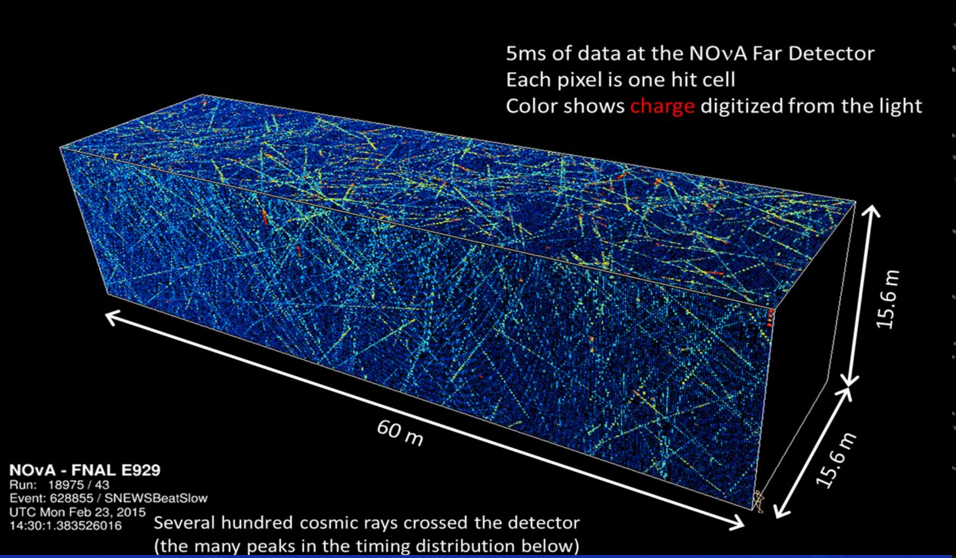


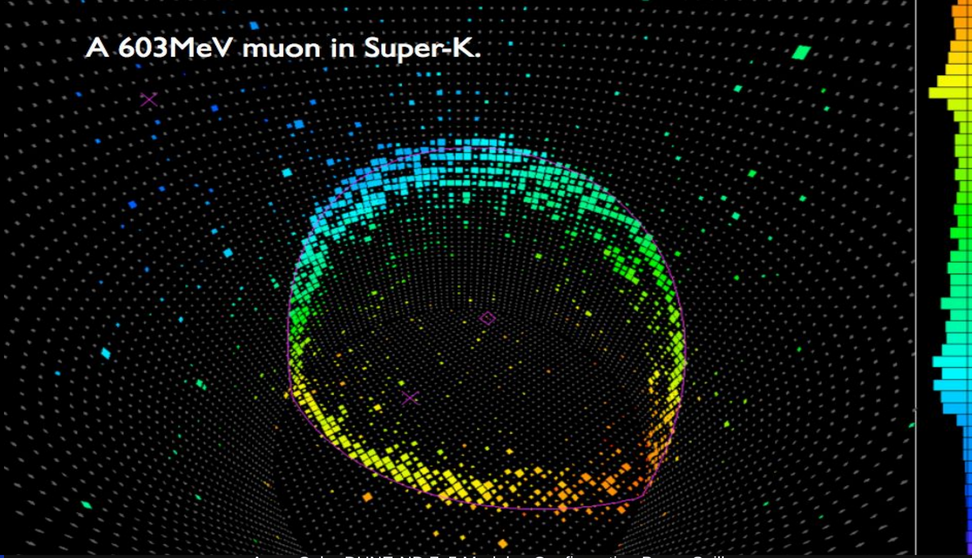


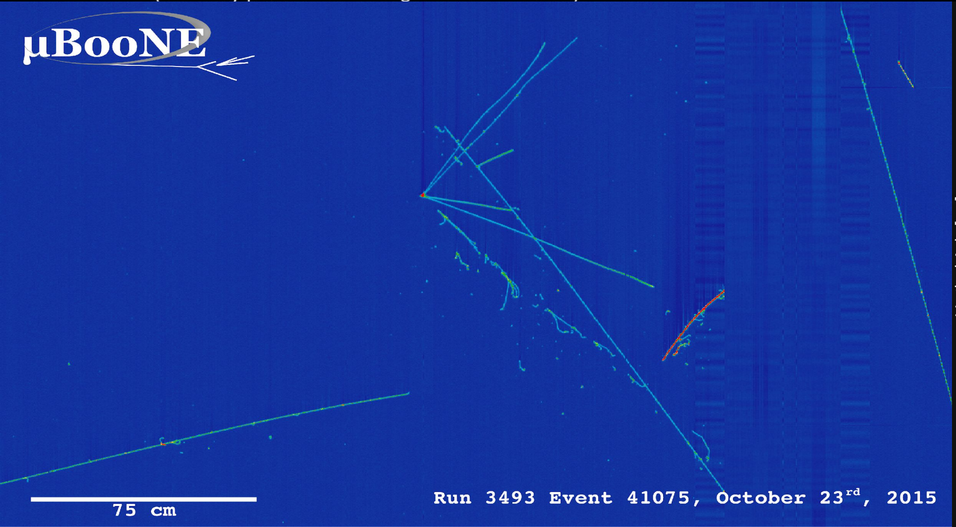
ν_e creates electron (e)

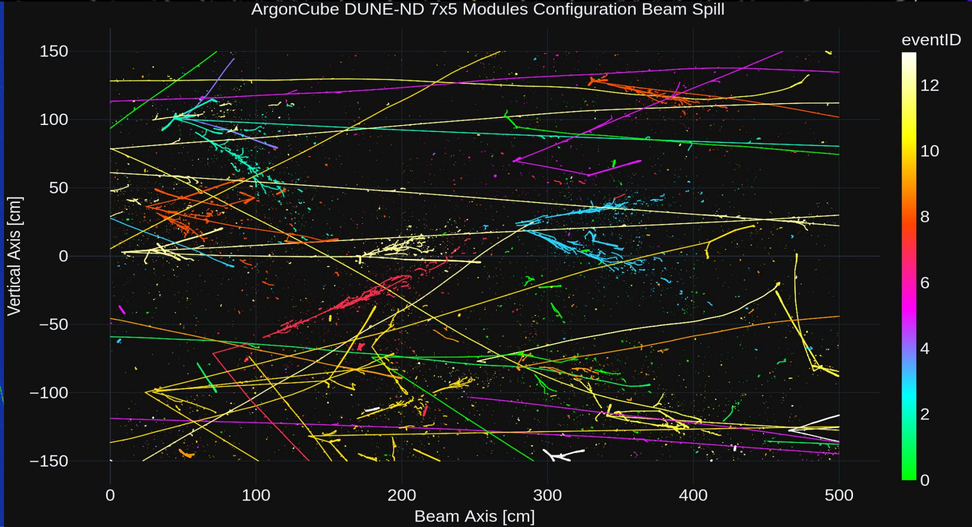
ν_μ creates muon (μ)

Detectors must be capable of measuring **type** & **energy**


5ms of data at the NOvA Far Detector
Each pixel is one hit cell
Color shows charge digitized from the light
15.6 m
60 m
15.6 m
NOvA - FNAL E929
Run: 18975 / 43
Event: 628855 / SNEWSBeatSlow
UTC Mon Feb 23, 2015
14:30:1.383526016
Several hundred cosmic rays crossed the detector
(the many peaks in the timing distribution below)


A 603MeV muon in Super-K.


μBooNE
75 cm
Run 3493 Event 41075, October 23rd, 2015


ArgonCube DUNE-ND 7x5 Modules Configuration Beam Spill
Vertical Axis [cm]
Beam Axis [cm]
eventID
150
100
50
0
−50
−100
−150
0
100
200
300
400
500
12
10
8
6
4
2
0

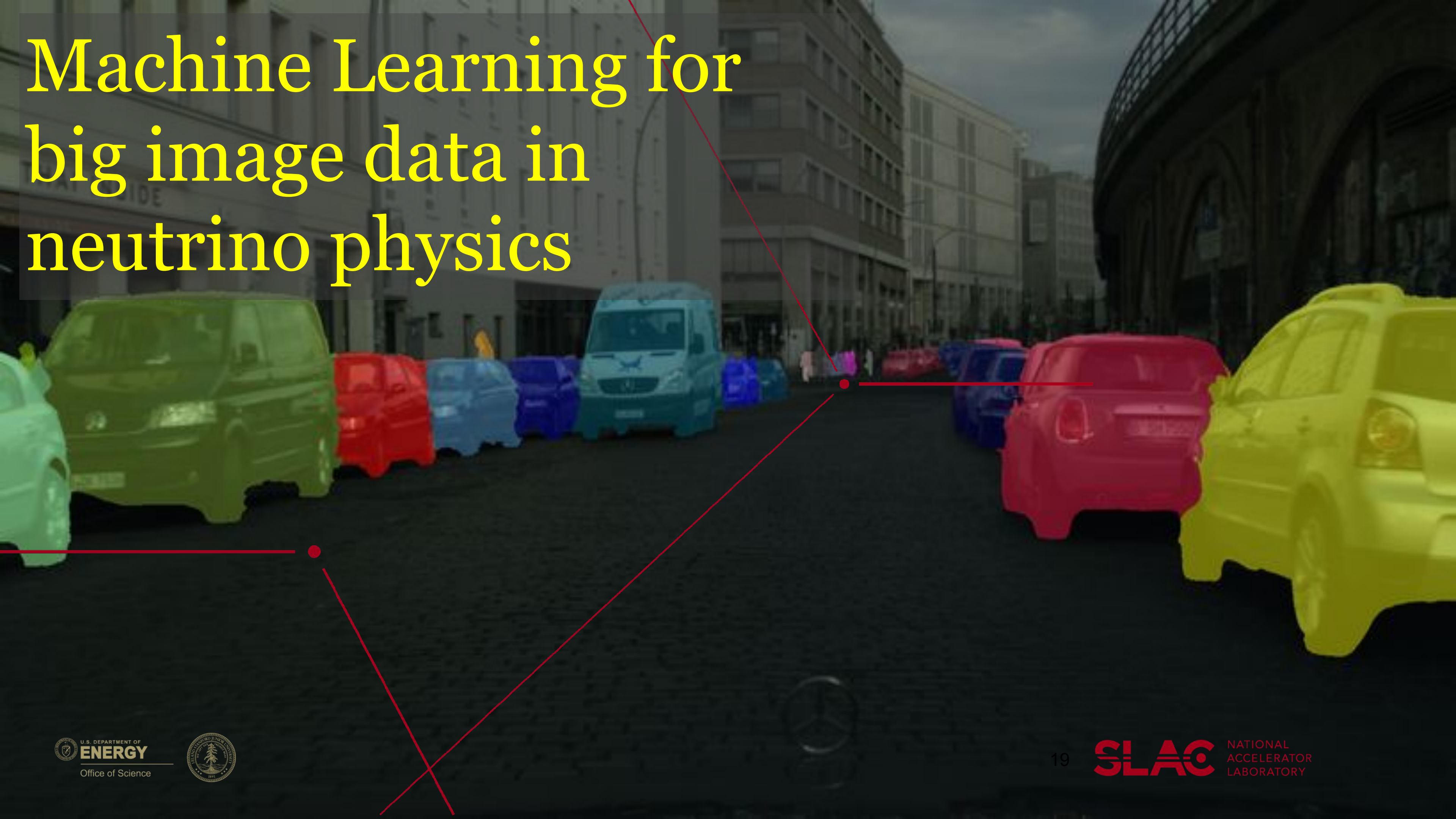
Present/Future Challenges

Lack of quality physics reconstruction for big image data

Slow, manual ("by-hand") workflow for development & tuning

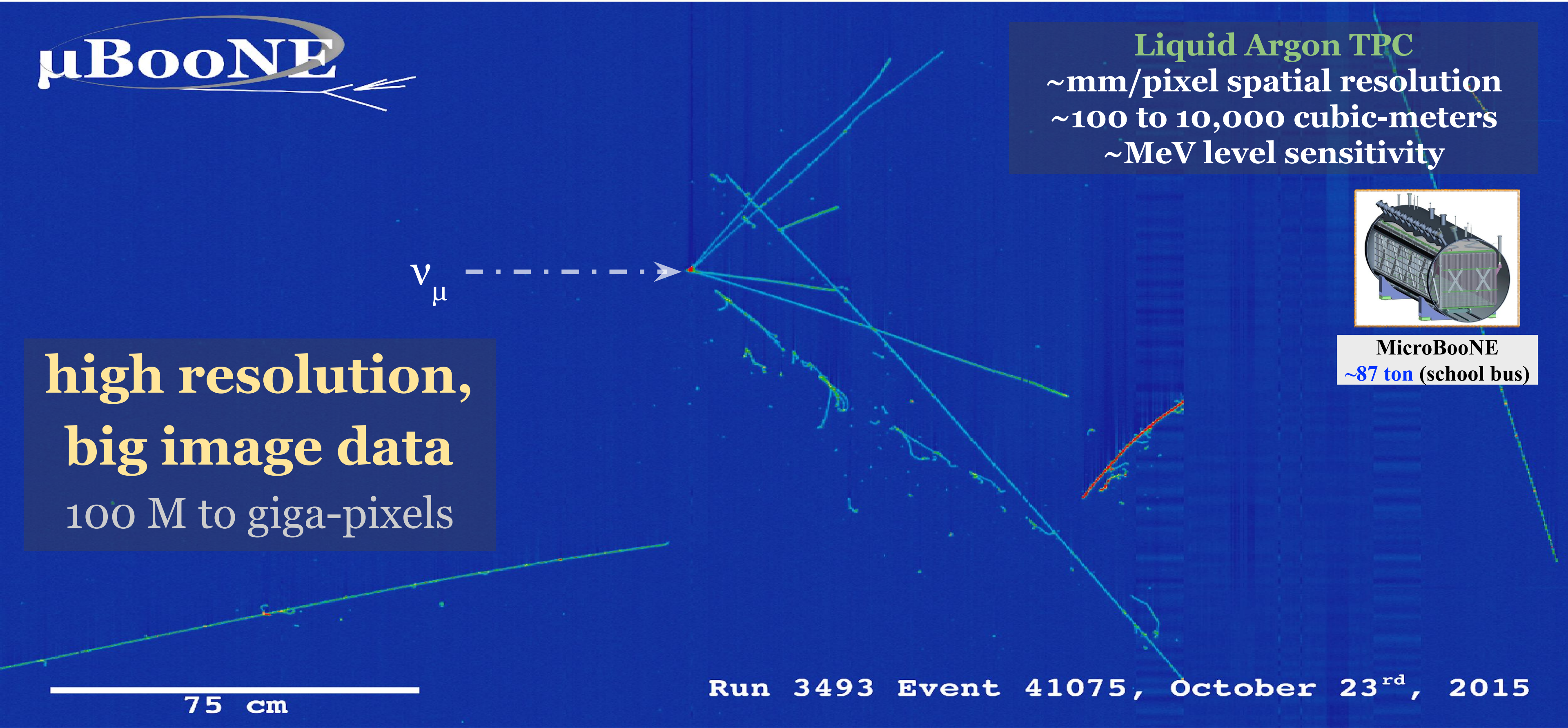
Imperfect physics modeling

Machine Learning for big image data in neutrino physics



ML for Analyzing Big Image Data in Neutrino Experiments

Challenges in particle imaging neutrino detectors



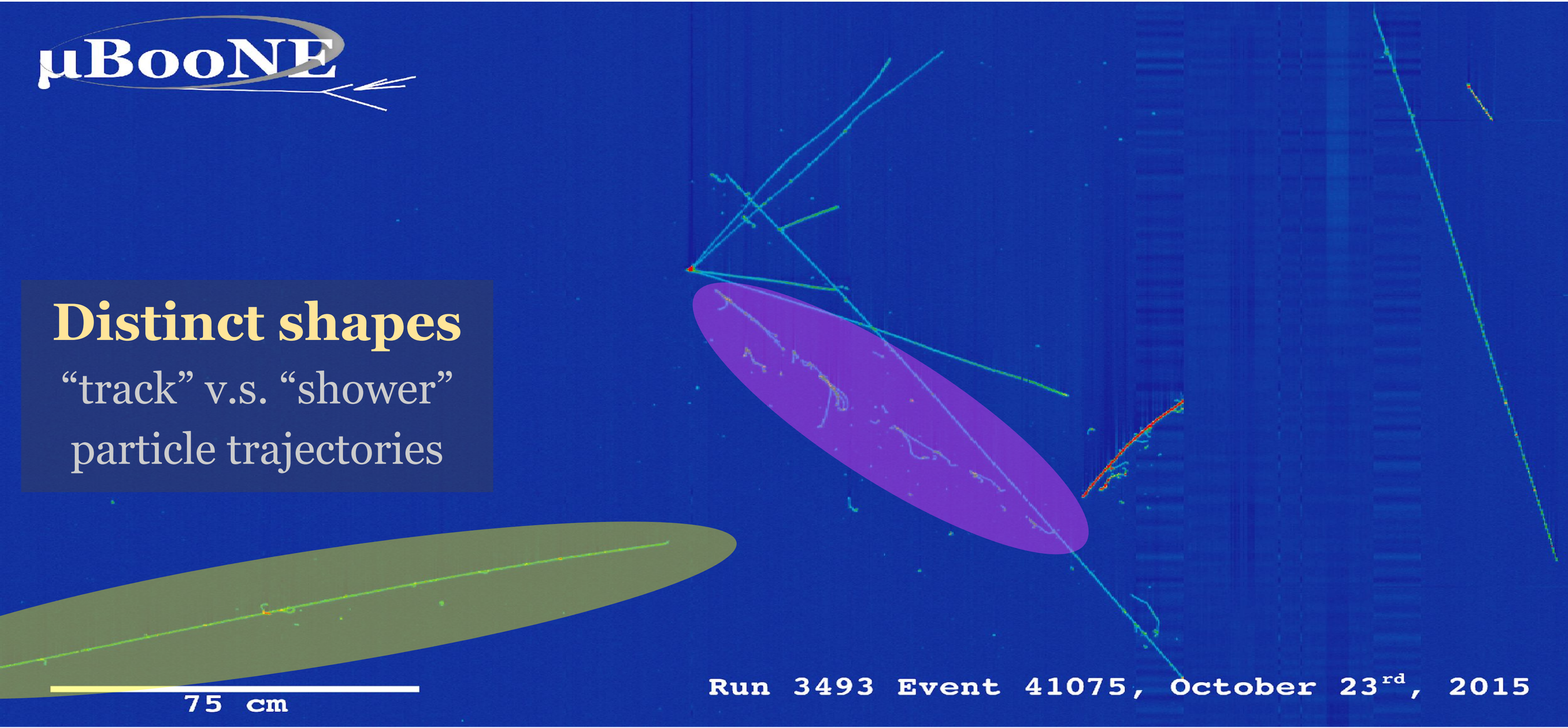
Liquid Argon TPC
~mm/pixel spatial resolution
~100 to 10,000 cubic-meters
~MeV level sensitivity

MicroBooNE
~87 ton (school bus)

high resolution,
big image data
100 M to giga-pixels

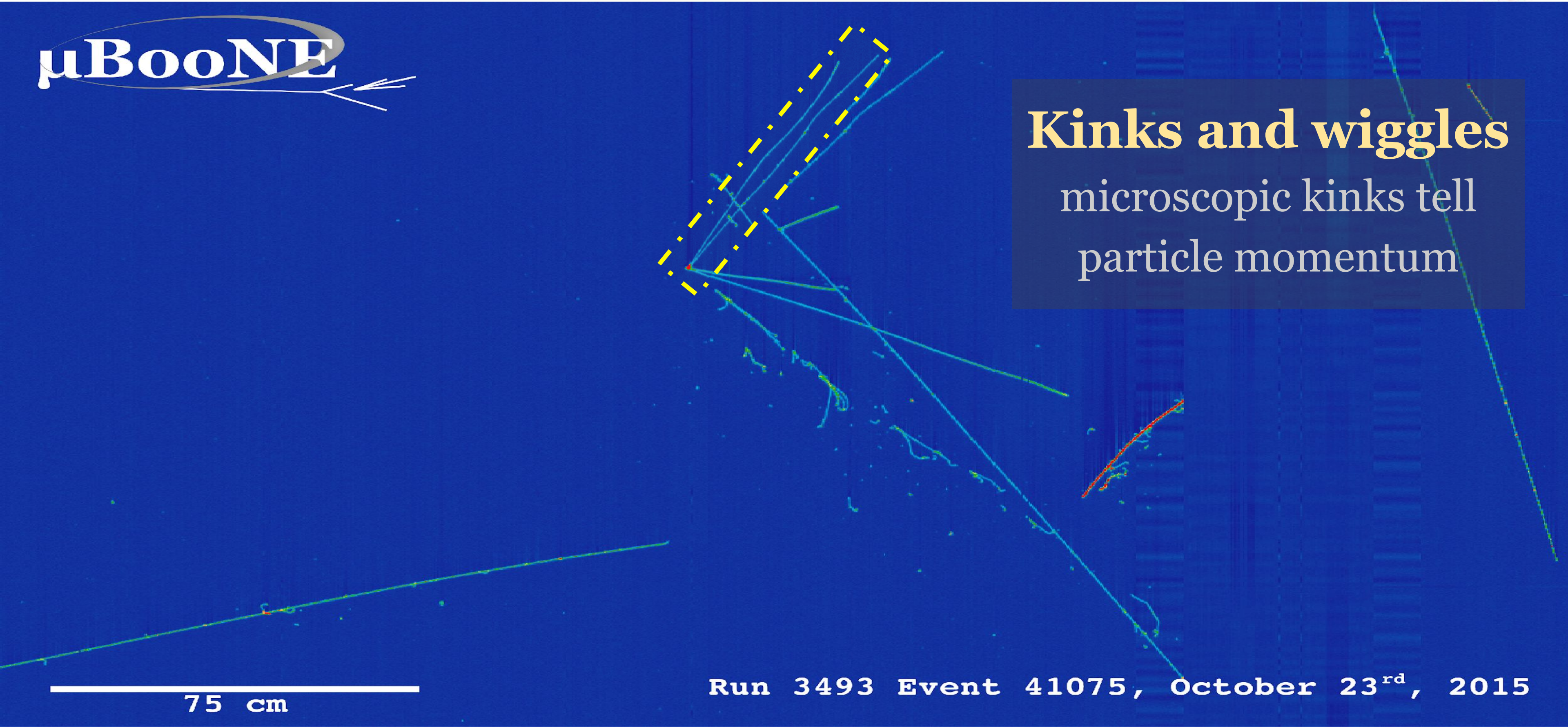
ML for Analyzing Big Image Data in Neutrino Experiments

Challenges in particle imaging neutrino detectors



ML for Analyzing Big Image Data in Neutrino Experiments

Challenges in particle imaging neutrino detectors

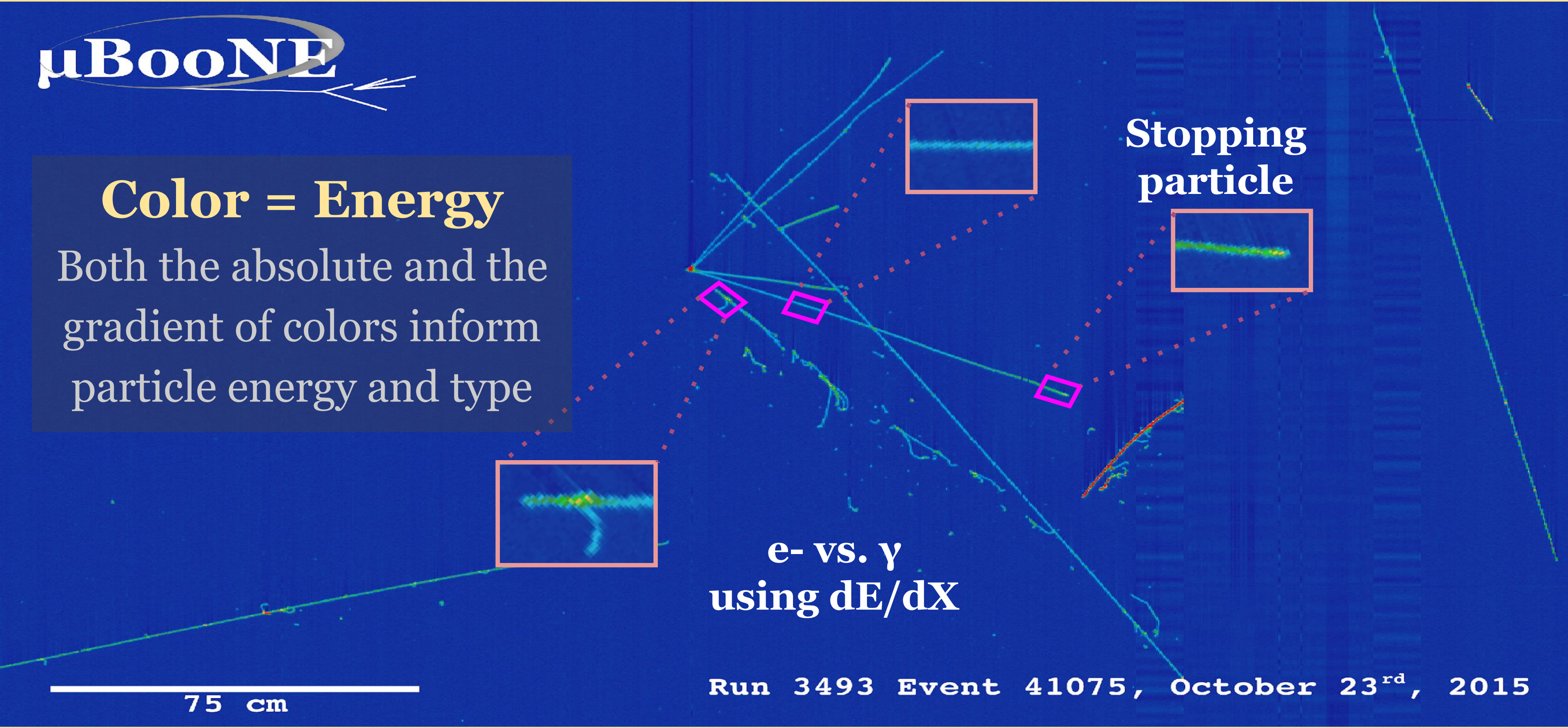


ML for Analyzing Big Image Data in Neutrino Experiments
Challenges in particle imaging neutrino detectors

SLAC

ML for Analyzing Big Image Data in Neutrino Experiments

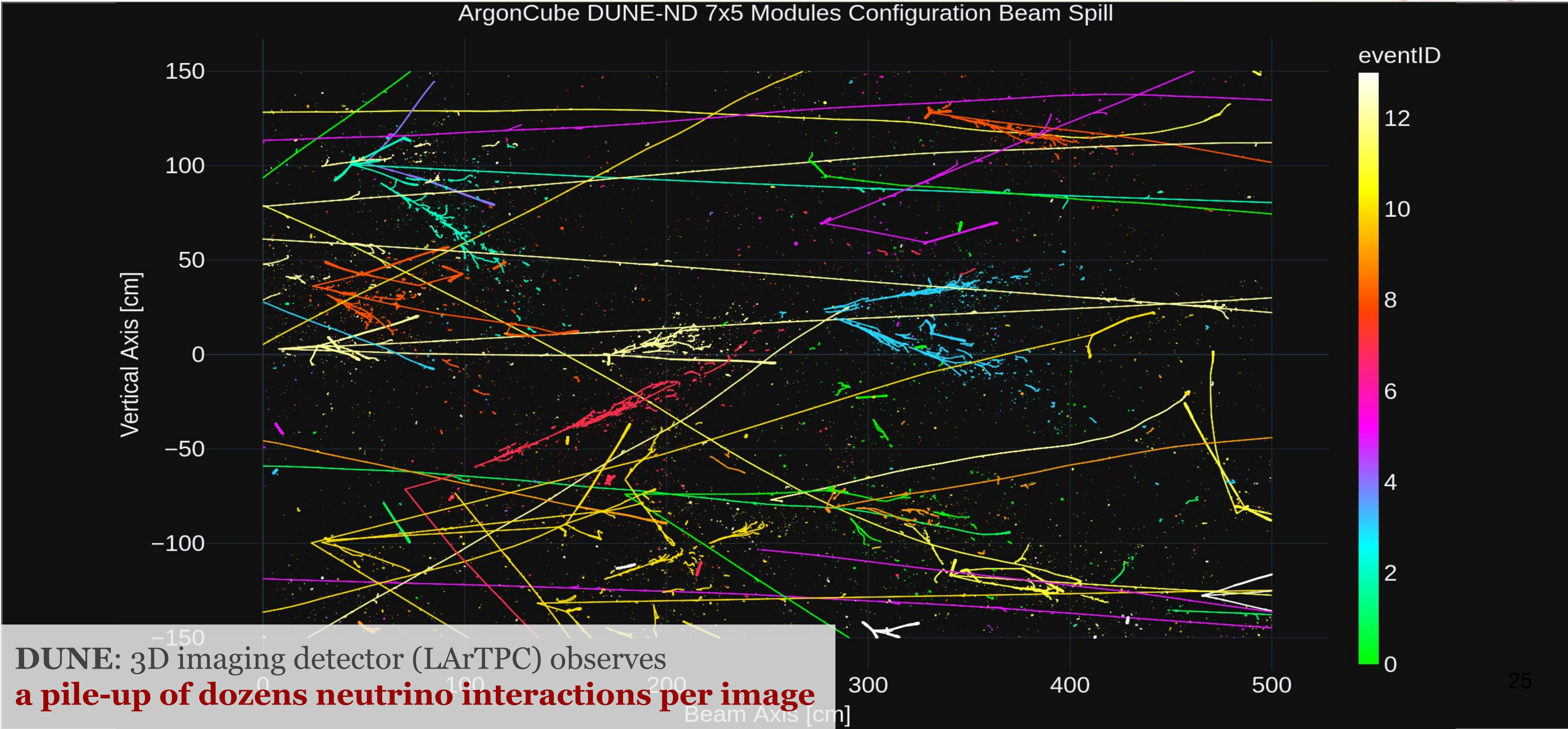
Challenges in particle imaging neutrino detectors



ML for Analyzing Big Image Data in Neutrino Experiments

Challenges in particle imaging neutrino detectors

SLAC



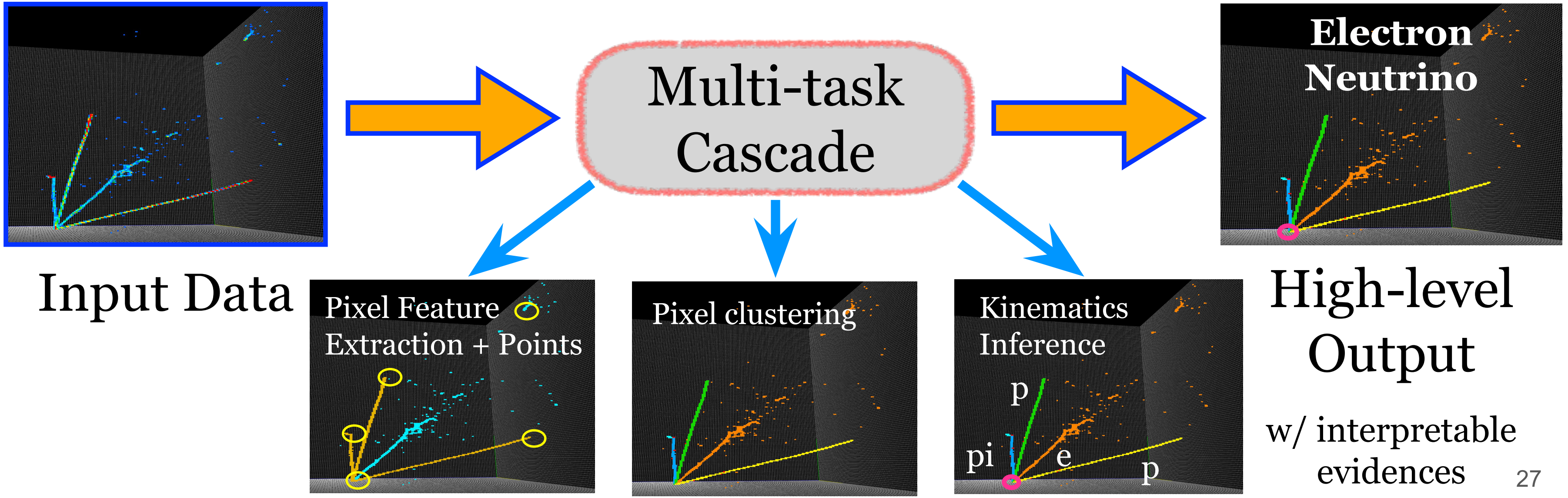
DUNE: 3D imaging detector (LArTPC) observes
a pile-up of dozens neutrino interactions per image

ML-Based LArTPC Data Reconstruction

End-to-end data reconstruction using ML

Machine Learning for Neutrino Image Data Analysis

- **Goal**: particle-level type and energy reconstruction
- **How**: extract physically meaningful, hierarchical features (evidences) by chaining multiple ML models designed for each task

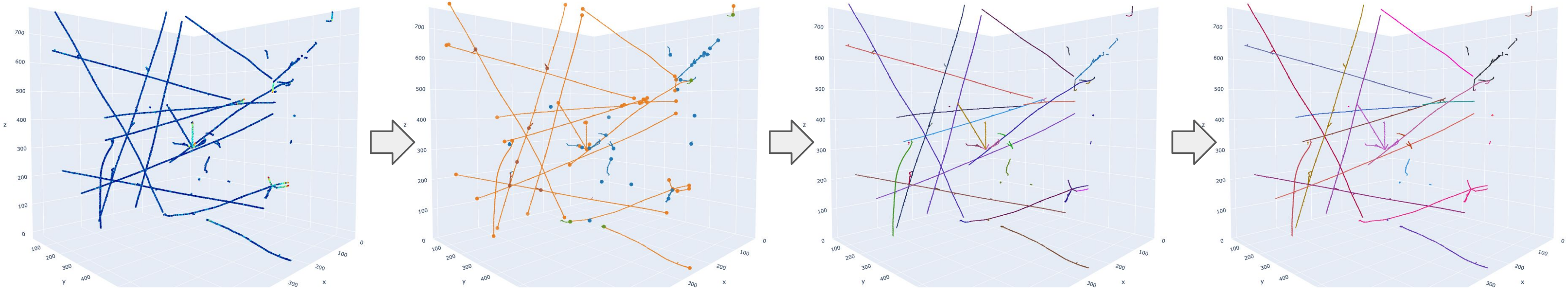


End-to-end data reconstruction using ML

Machine Learning for Neutrino Image Data Analysis

- **Goal**: particle-level type and energy reconstruction
- **How**: extract physically meaningful, hierarchical features (evidences) by chaining multiple ML models designed for each task

Input → Pixel Features → Pixel Clustering → Particle Clustering



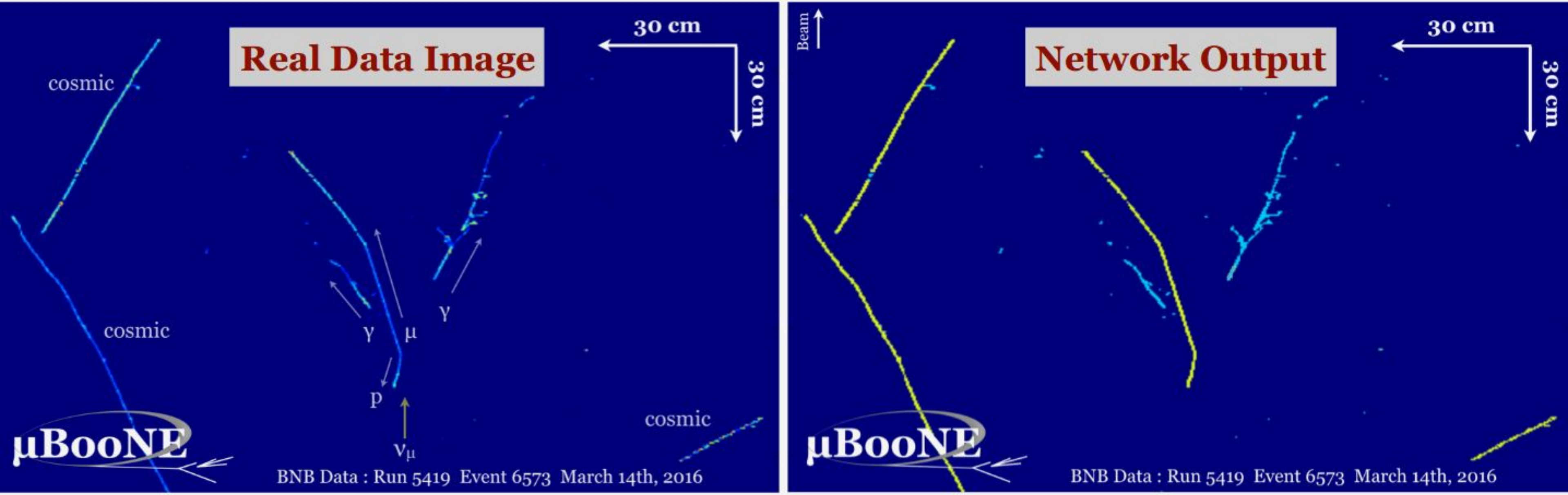
Three major stages of reconstruction

ML for Analyzing Big Image Data in Neutrino Experiments

Stage 1: Pixel-level Feature Extraction + Scalablility

Distinguish 2 distinct topologies: **showers v.s. tracks** (for the next stage = clustering)

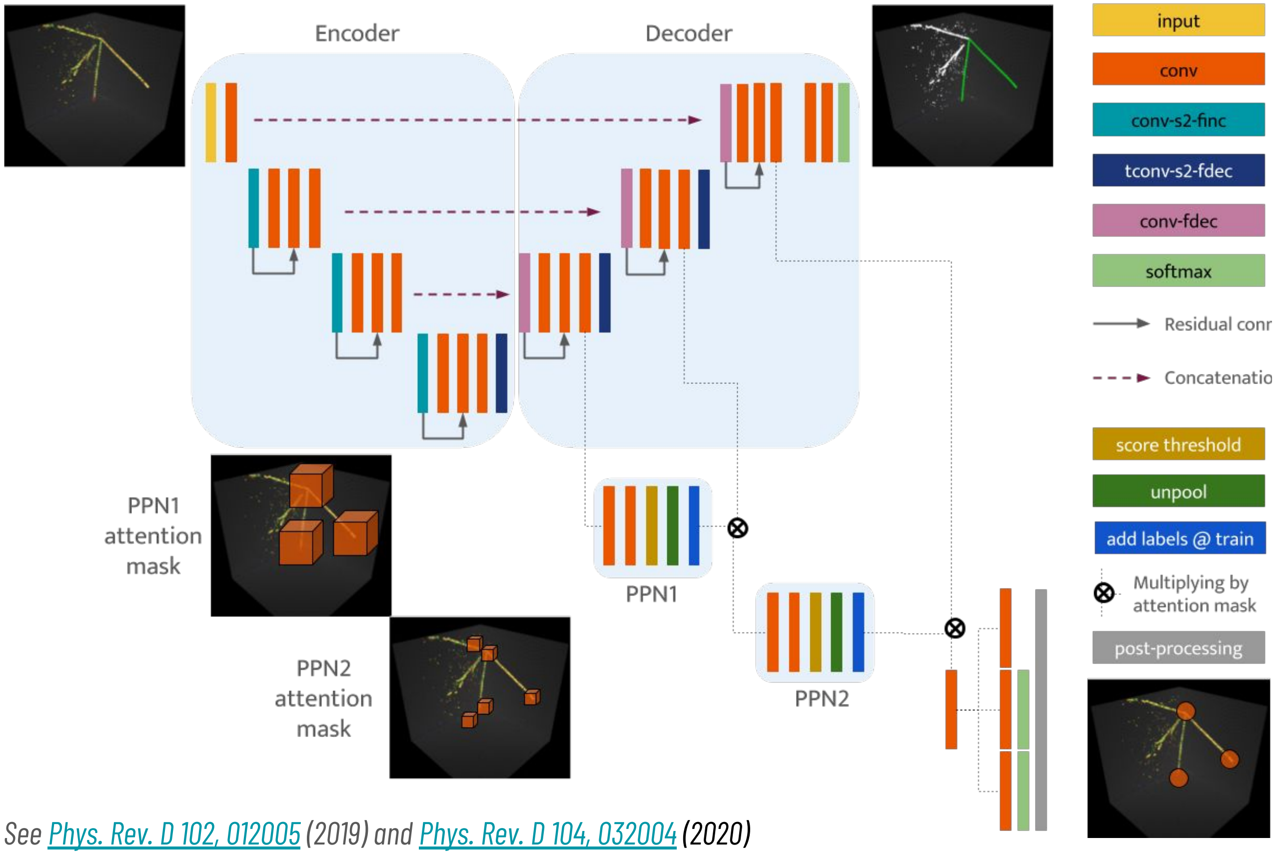
Identify trajectory **edge points** (track start/end, shower start)

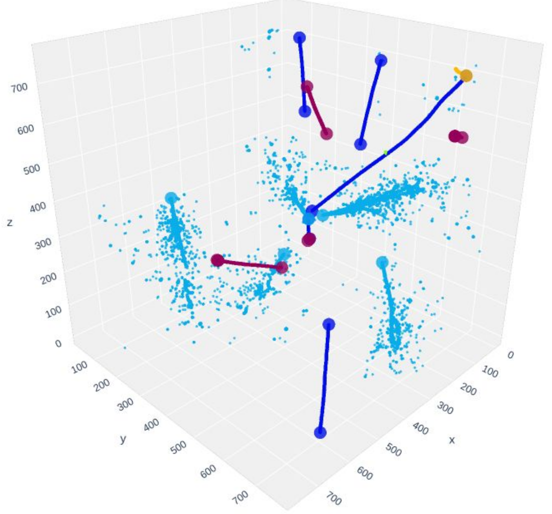


Network Input PRD 99 092001 (2018) **Network Output**

ML for Analyzing Big Image Data in Neutrino Experiments

Stage 1-b: Particle Edge-point Prediction





Semantic segmentation
(U-Net + residual conn.)

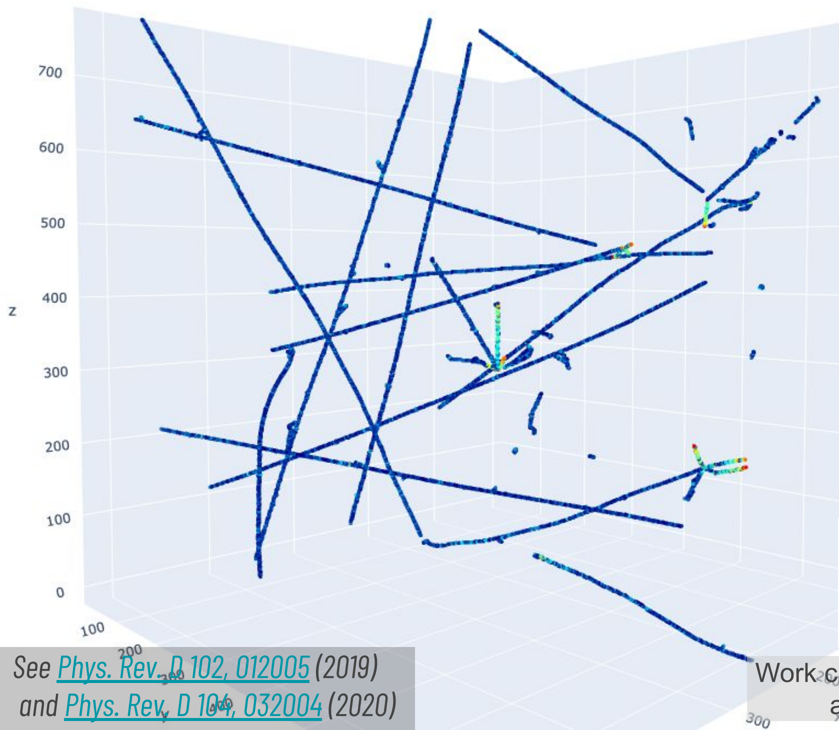
Edge point detection
(Faster R-CNN)

Sparse tensor operation
(Minkowski Engine)

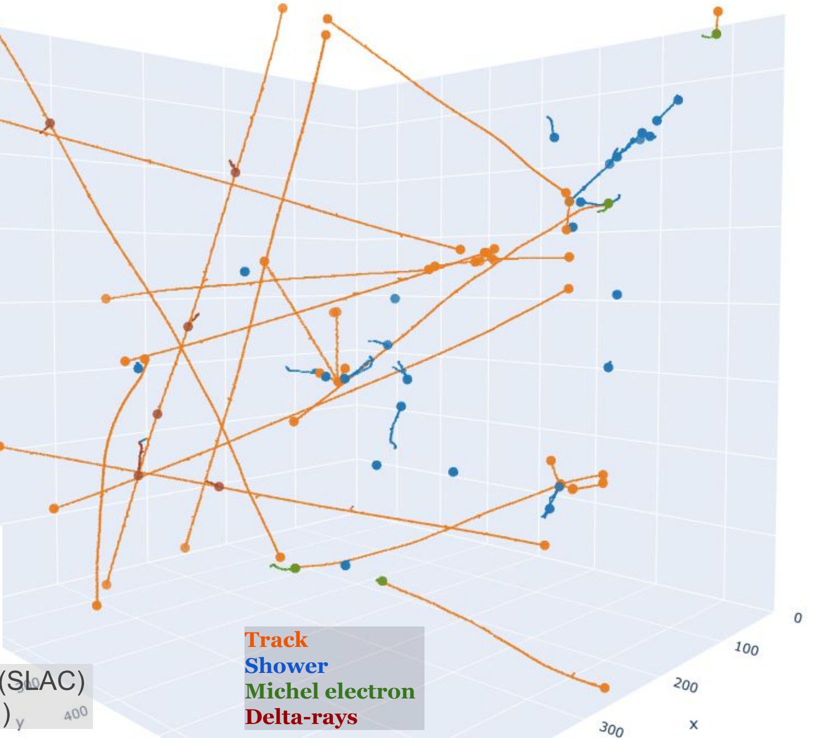
See Phys. Rev. D 102, 012005 (2019) and Phys. Rev. D 104, 032004 (2020)

ML for Analyzing Big Image Data in Neutrino Experiments
Stage 1: input & output

Stage 1 Input



Stage 1 Output



Work credit (from left) Ran I. (SLAC) and Laura D. (Stanford)

See Phys. Rev. D 102, 012005 (2019) and Phys. Rev. D 104, 032004 (2020)

Clustering in the embedding space

- Use CNN to learn a transformation function from the 3D voxels to the embedding space where clustering can be performed in a simple manner

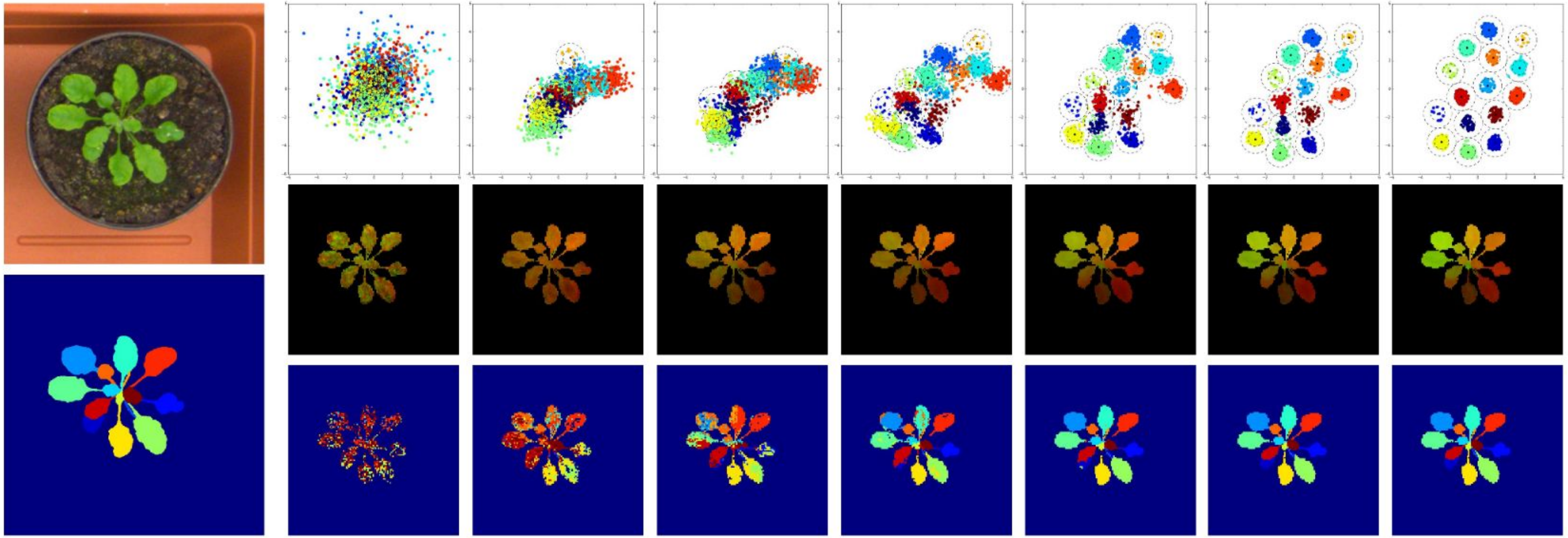
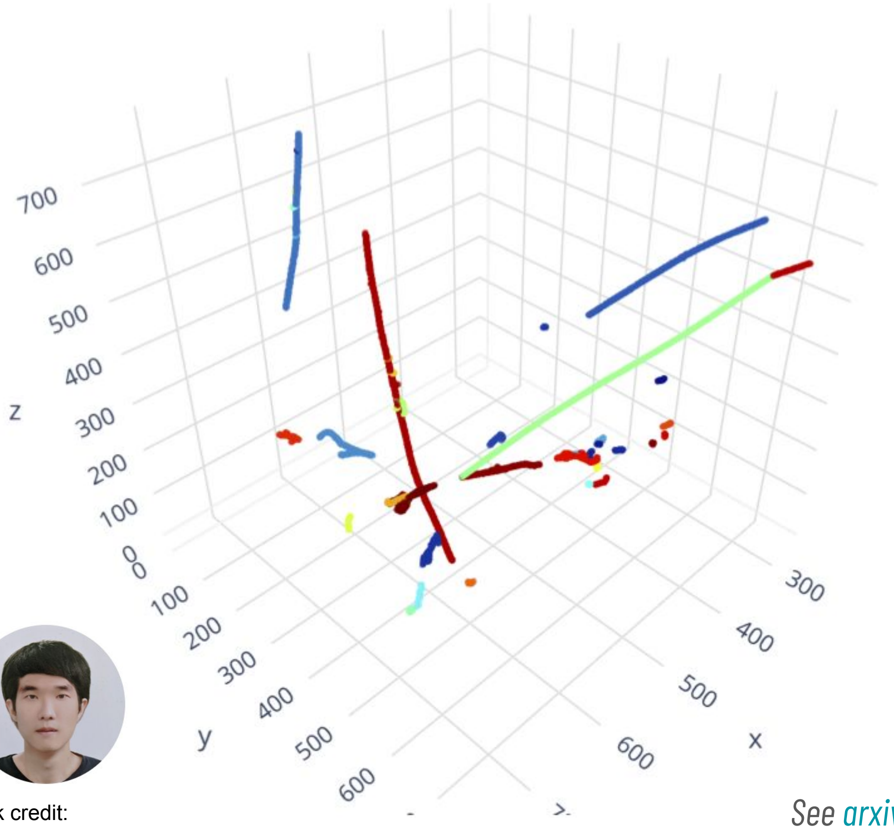
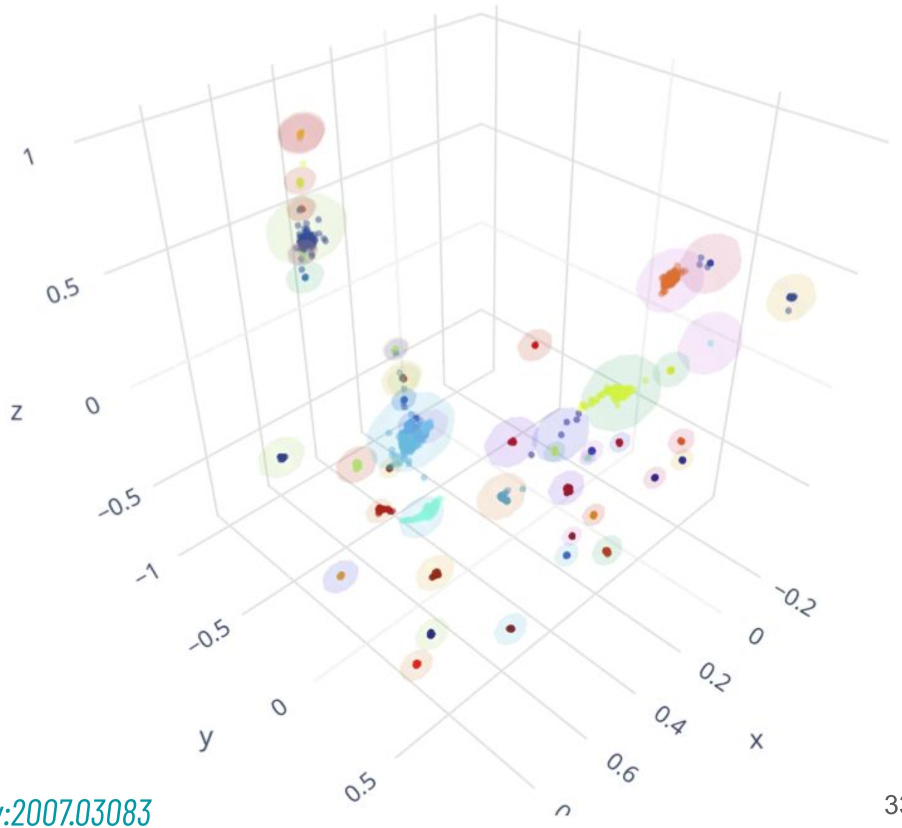


Image credit: arXiv 1708.02551

Stage 2-a: Dense Pixel Clustering



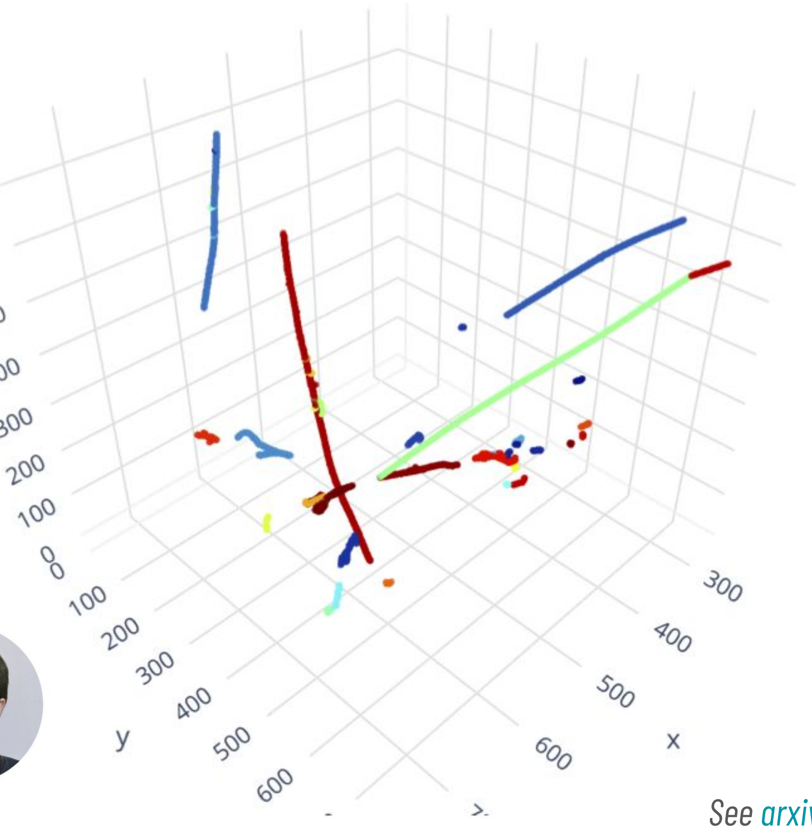


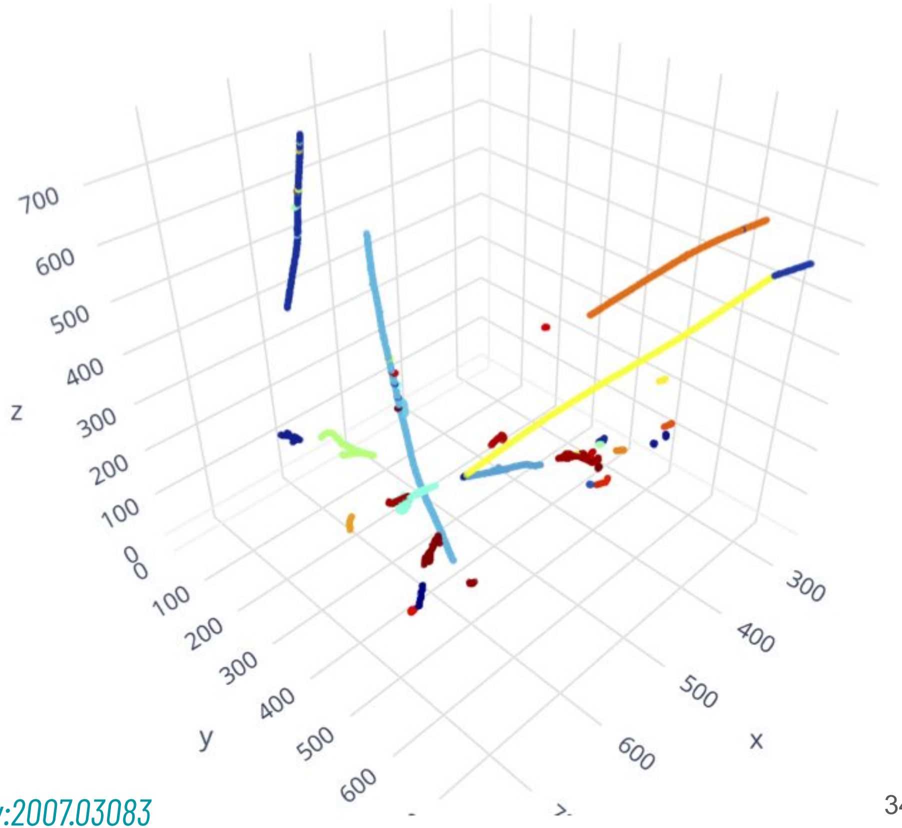
Work credit:
Dae Heun Koh (Stanford)

See arxiv:2007.03083

ML for Analyzing Big Image Data in Neutrino Experiments

Stage 2-a: Dense Pixel Clustering





Work credit:
Dae Heun Koh (Stanford)

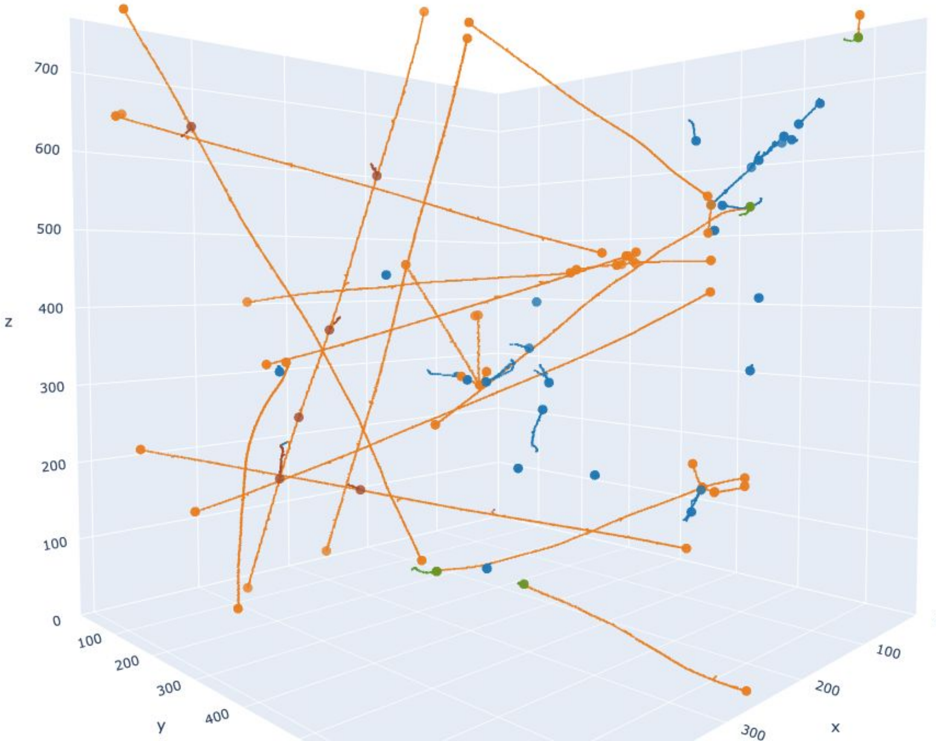
See arxiv:2007.03083

ML for Analyzing Big Image Data in Neutrino Experiments

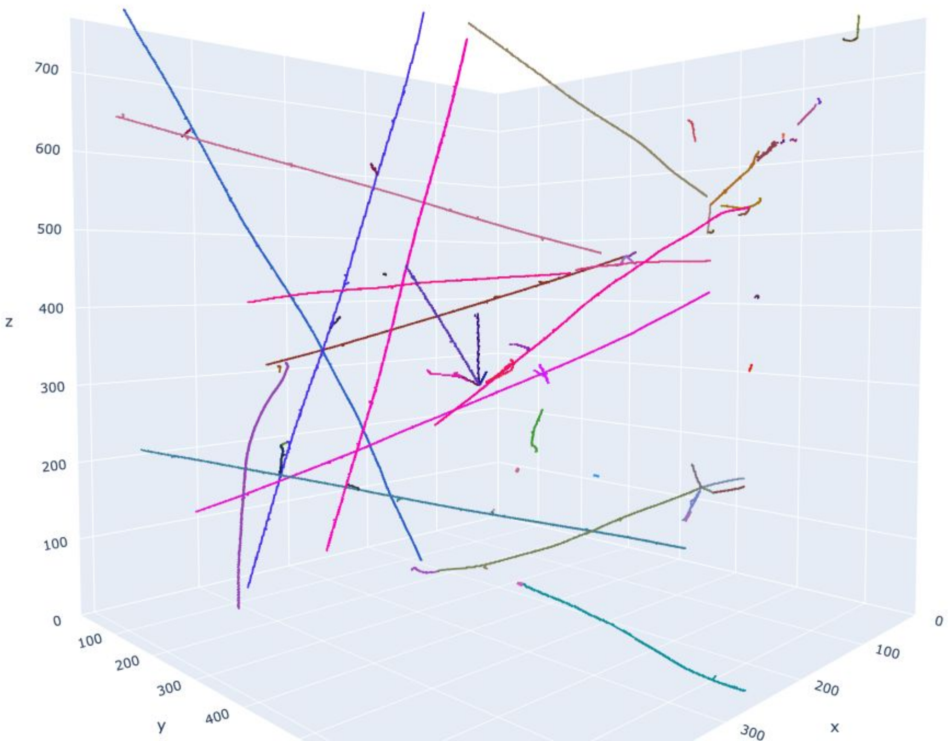
Stage 2-a: input & output

SLAC

Stage 2-a Input



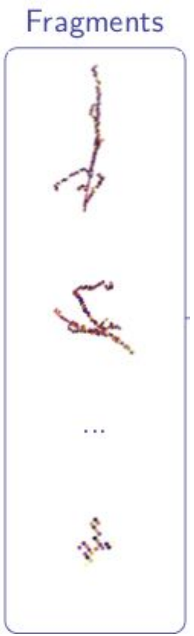
Stage 2-a Output

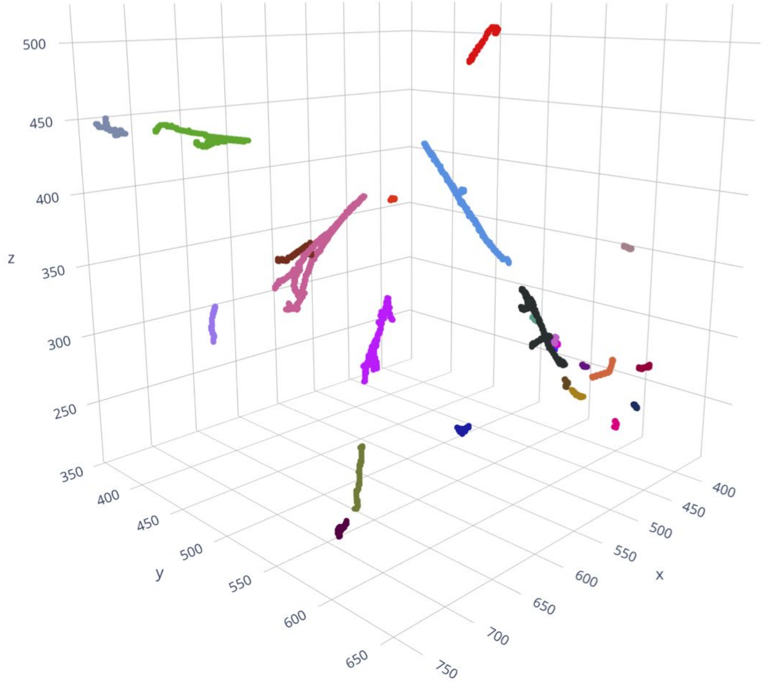


ML for Analyzing Big Image Data in Neutrino Experiments

Stage 2-b: Sparse Fragment Clustering

Identifying 1 shower … which consists of **many fragments**

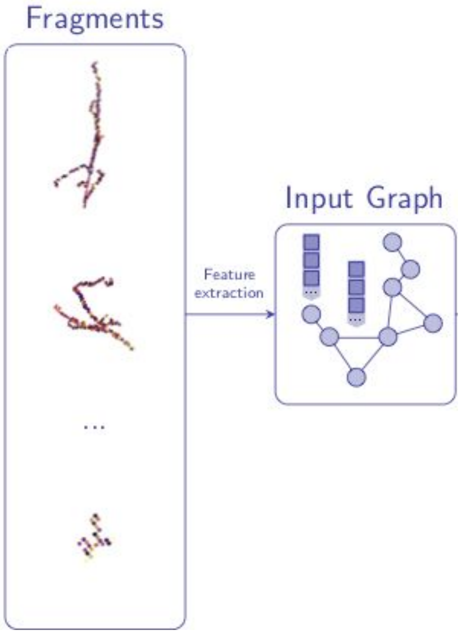


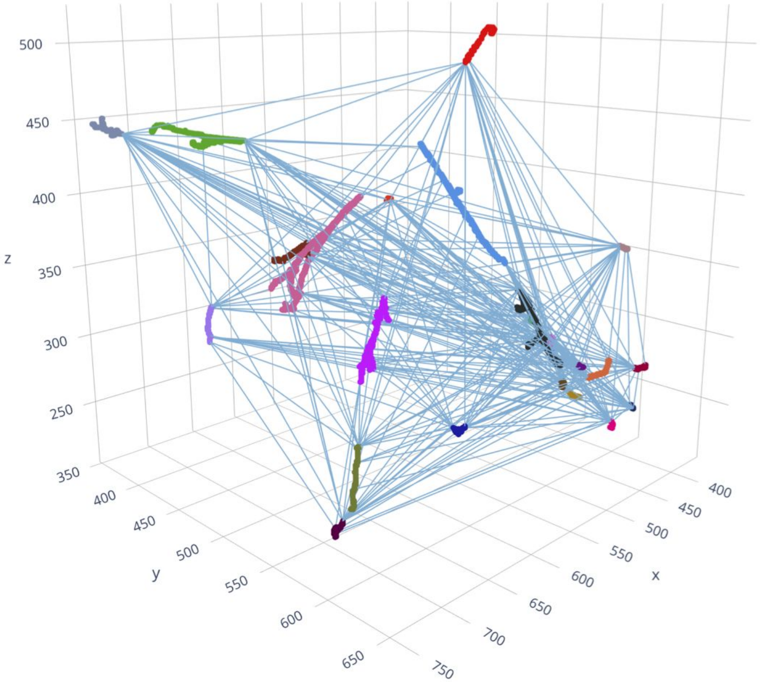


Stage 2-b: Sparse Fragment Clustering

Identifying 1 shower ... which consists of **many fragments**

- Interpret each fragment as a graph node + edges connect nodes in the same cluster

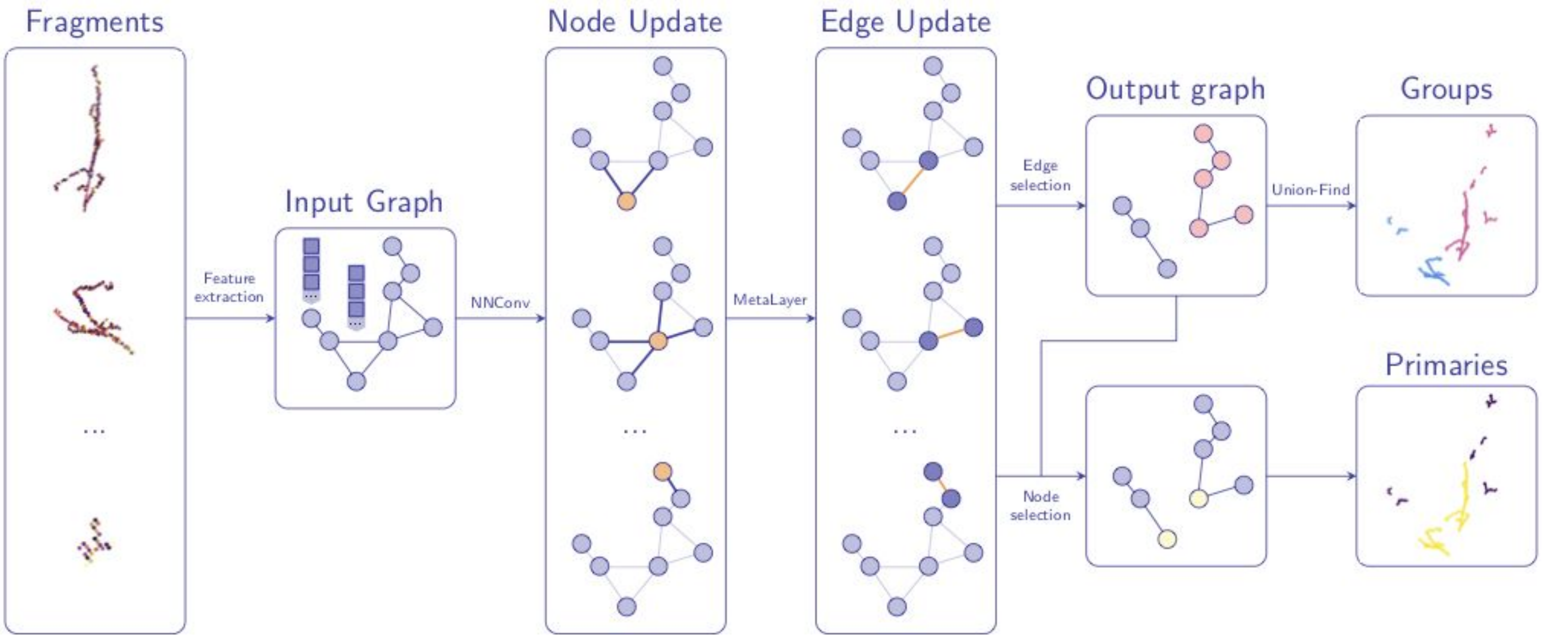




Stage 2-b: Sparse Fragment Clustering

Identifying 1 shower … which consists of **many fragments**

- Interpret each fragment as a graph node + edges connect nodes in the same cluster
- Cast the problem to a classification of node (e.g. particle type) and edge (clustering)

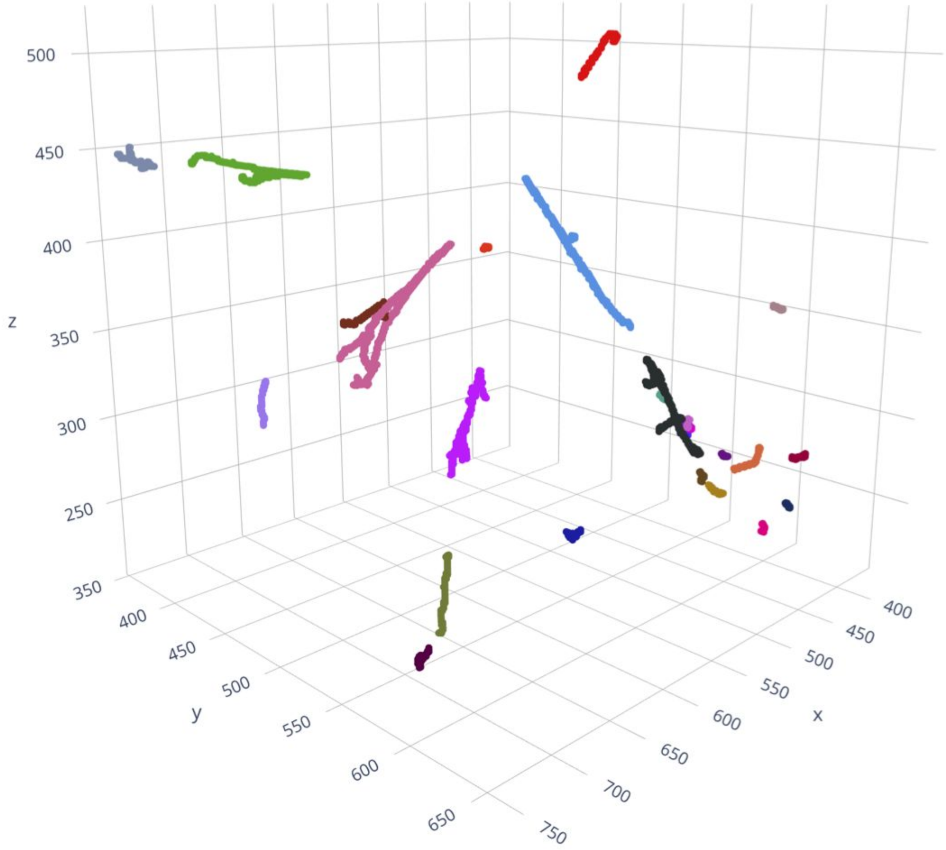


Stage 2-b: Sparse Fragment Clustering

Graph-NN for Particle Aggregation (GrapPA)

Input:

- Fragmented EM showers



See Phys. Rev. D 104, 072004

Stage 2-b: Sparse Fragment Clustering

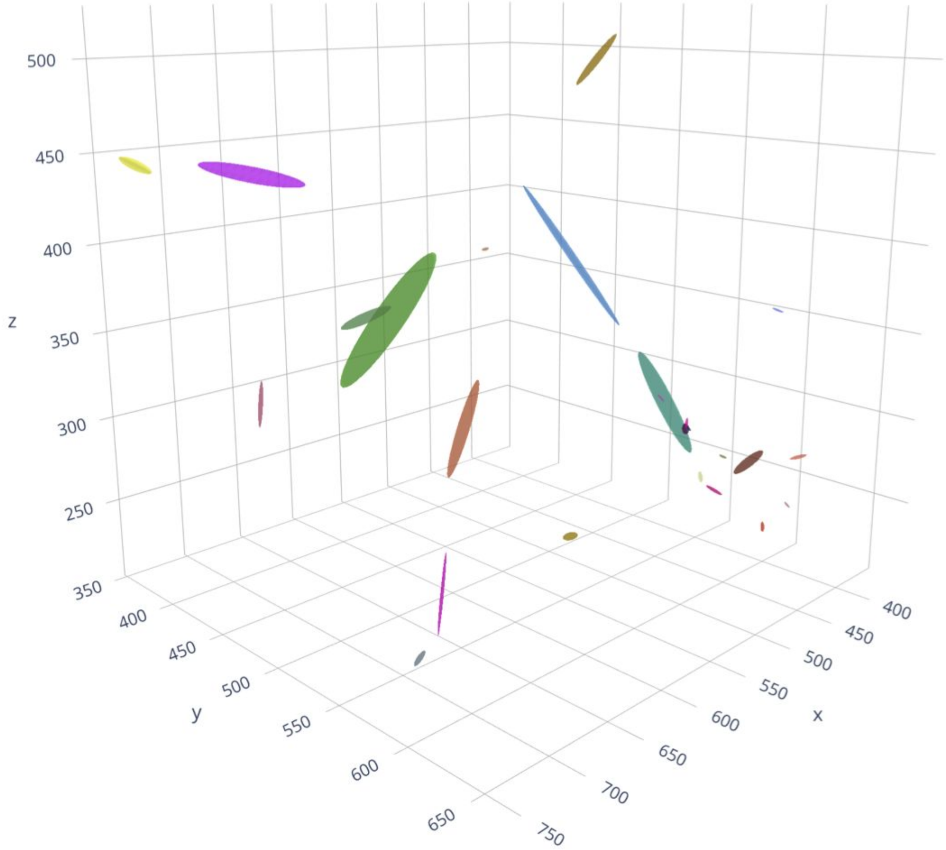
Graph-NN for Particle Aggregation (GrapPA)

Input:

- Fragmented EM showers

Node features:

- Centroid, Covariance matrix, PCA
- Start point, direction (PPN)



See Phys. Rev. D 104, 072004

Stage 2-b: Sparse Fragment Clustering

Graph-NN for Particle Aggregation (GrapPA)

Input:

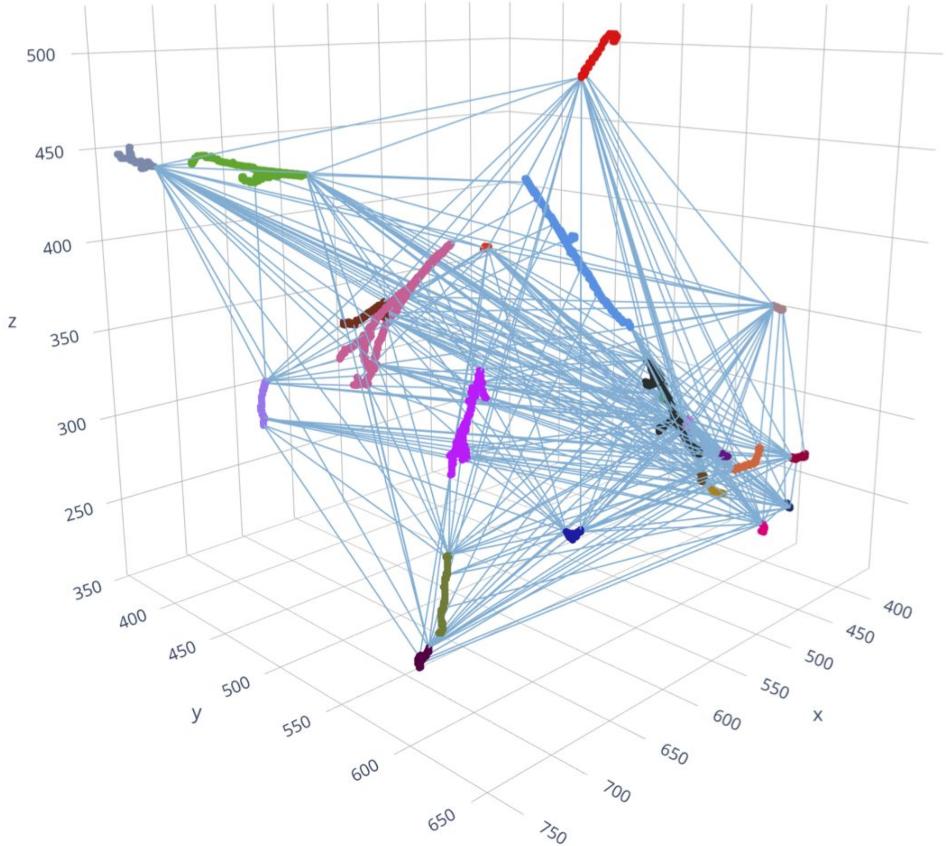
- Fragmented EM showers

Node features:

- Centroid, Covariance matrix, PCA
- Start point, direction (PPN)

Input graph:

- Connect every node with every other node (complete graph)



See Phys. Rev. D 104, 072004

Stage 2-b: Sparse Fragment Clustering

Graph-NN for Particle Aggregation (GrapPA)

Input:

- Fragmented EM showers

Node features:

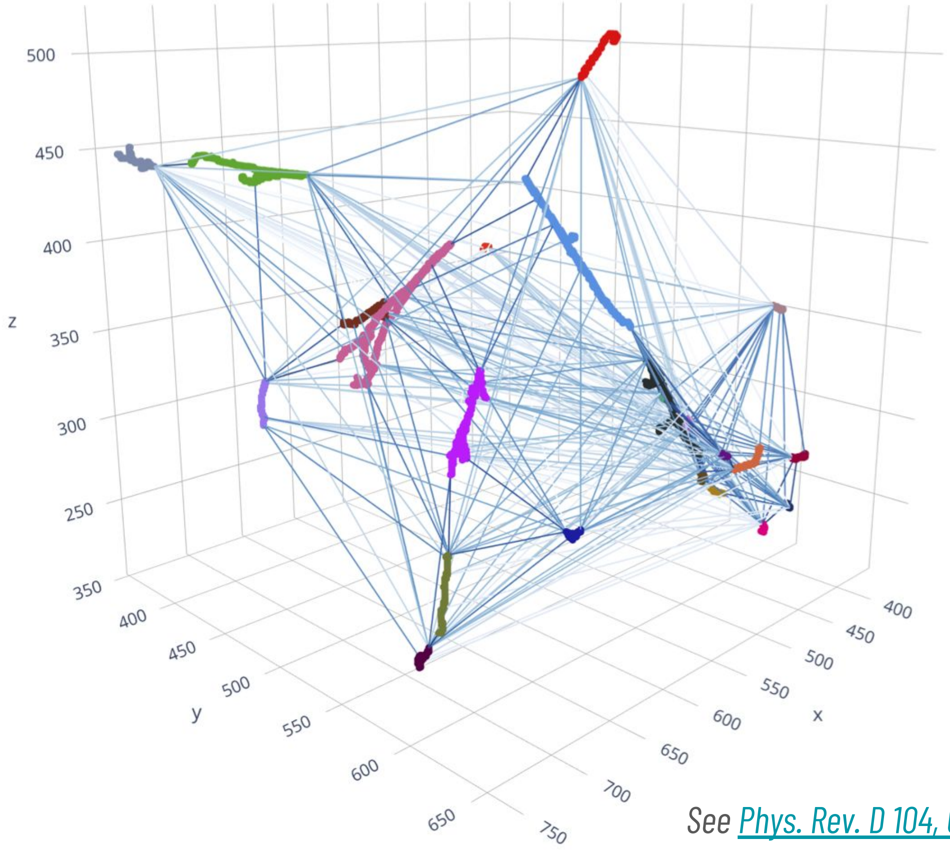
- Centroid, Covariance matrix, PCA
- Start point, direction (PPN)

Input graph:

- Connect every node with every other node (complete graph)

Edge features:

- Displacement vector
- Closest points of approach

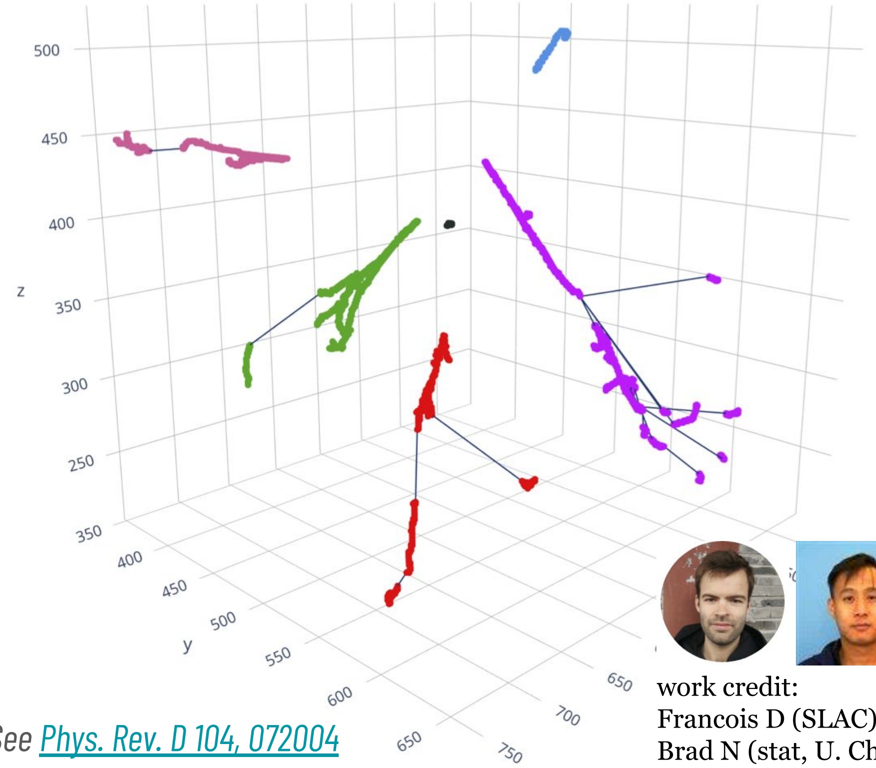


See Phys. Rev. D 104, 072004

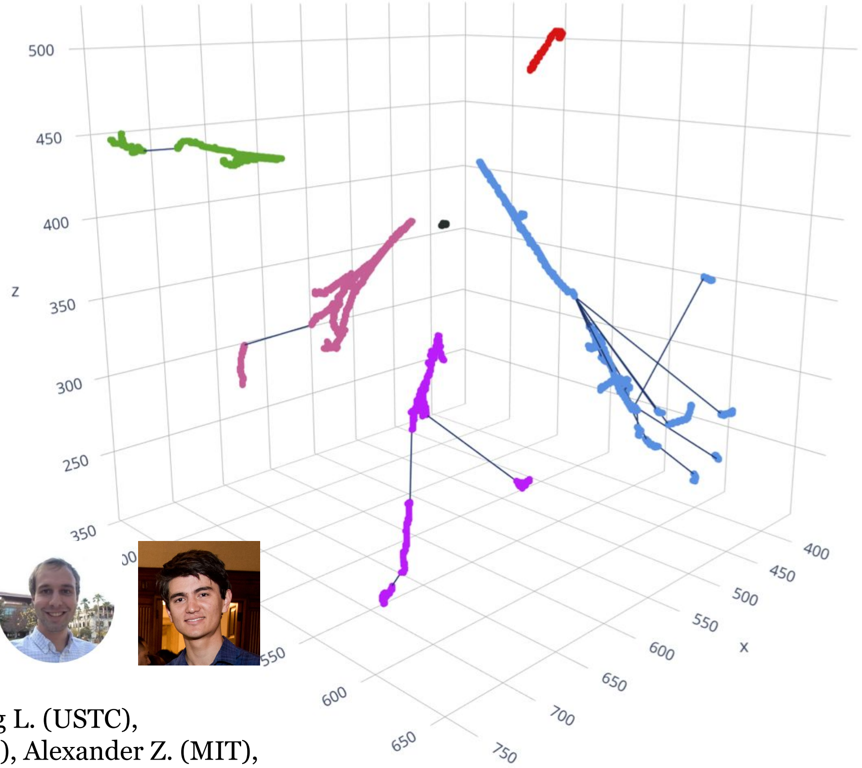
ML for Analyzing Big Image Data in Neutrino Experiments

Stage 2-b: Sparse Fragment Clustering

Target



Prediction



work credit:
Francois D (SLAC), Qing L. (USTC),
Brad N (stat, U. Chicago), Alexander Z. (MIT),

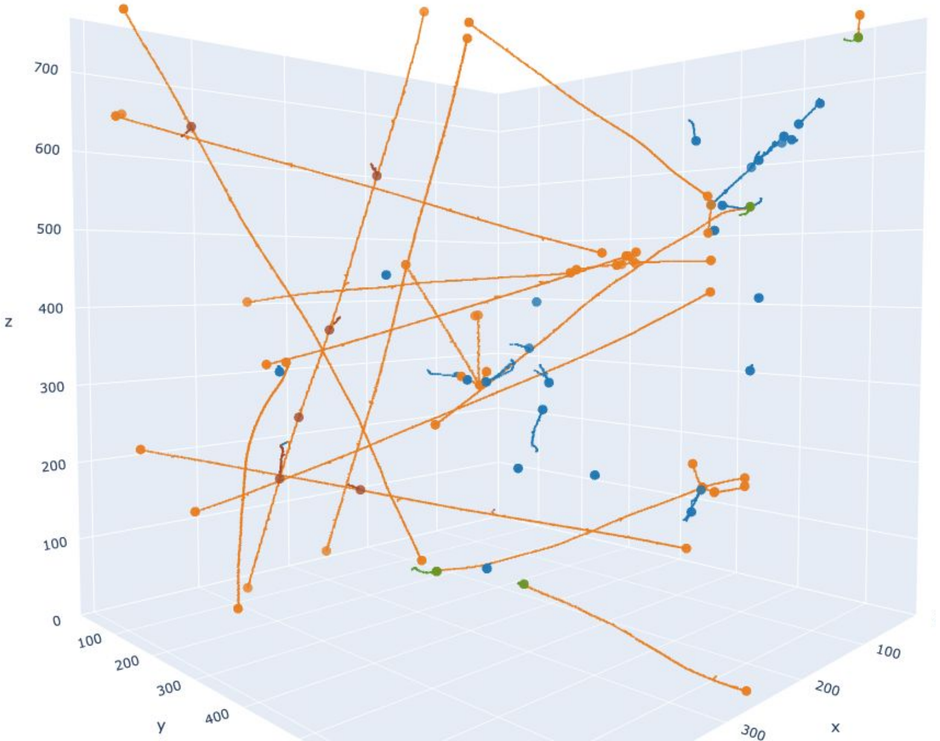
See [Phys. Rev. D 104, 072004]

ML for Analyzing Big Image Data in Neutrino Experiments

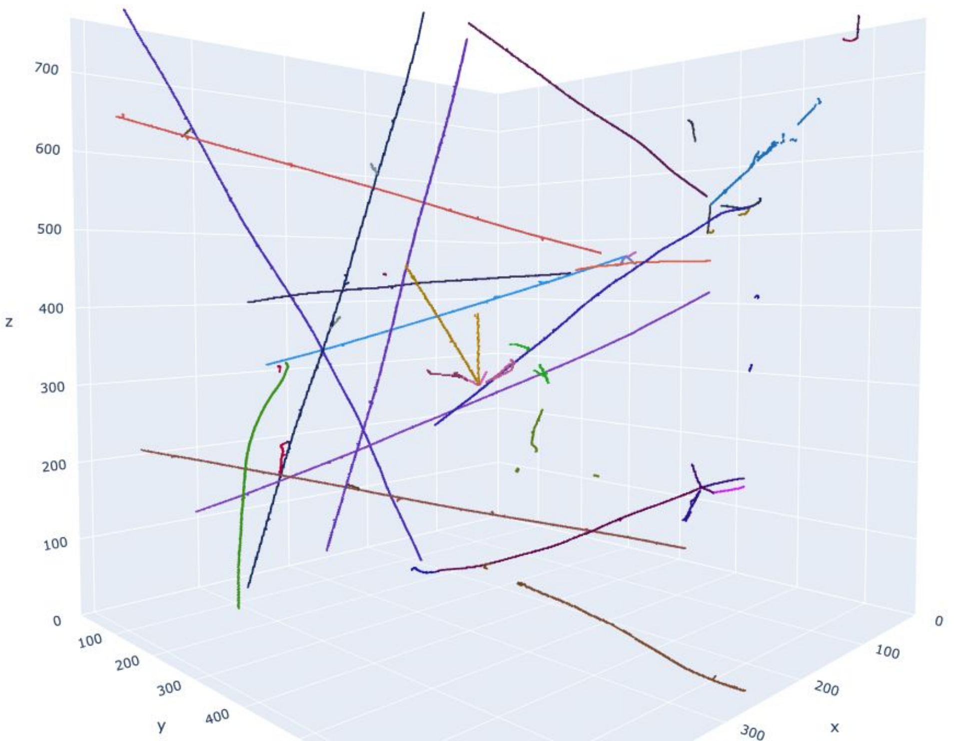
Stage 2: input & output

SLAC

Stage 2 Input



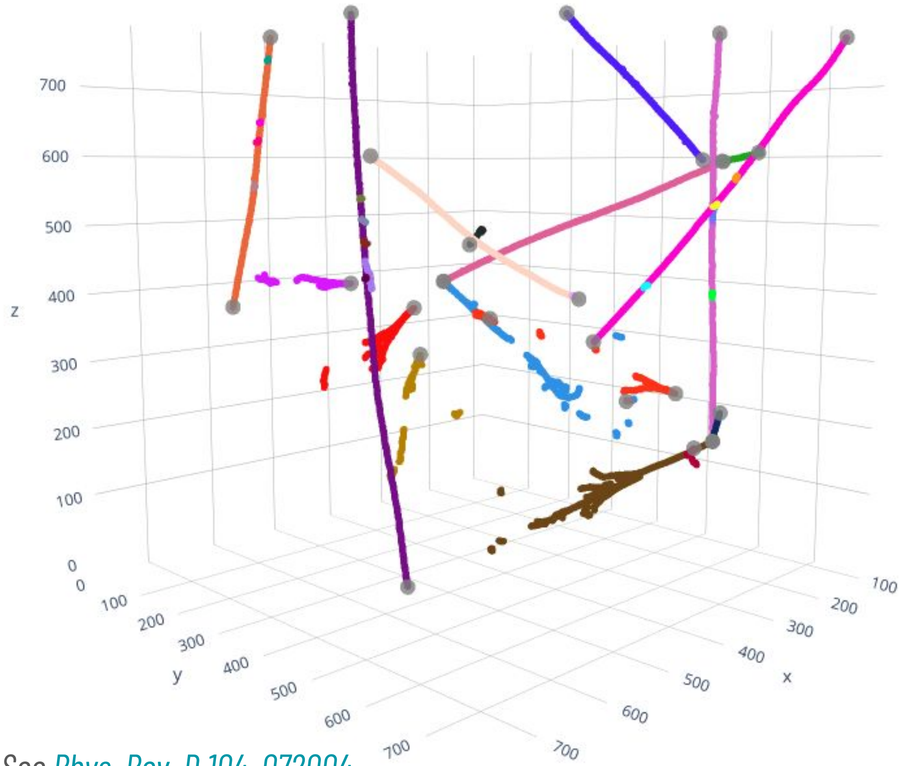
Stage 2 Output



ML for Analyzing Big Image Data in Neutrino Experiments

Stage 3: Interaction Clustering

SLAC



Identifying Each Interaction?

Grouping task = re-use GrapPA!

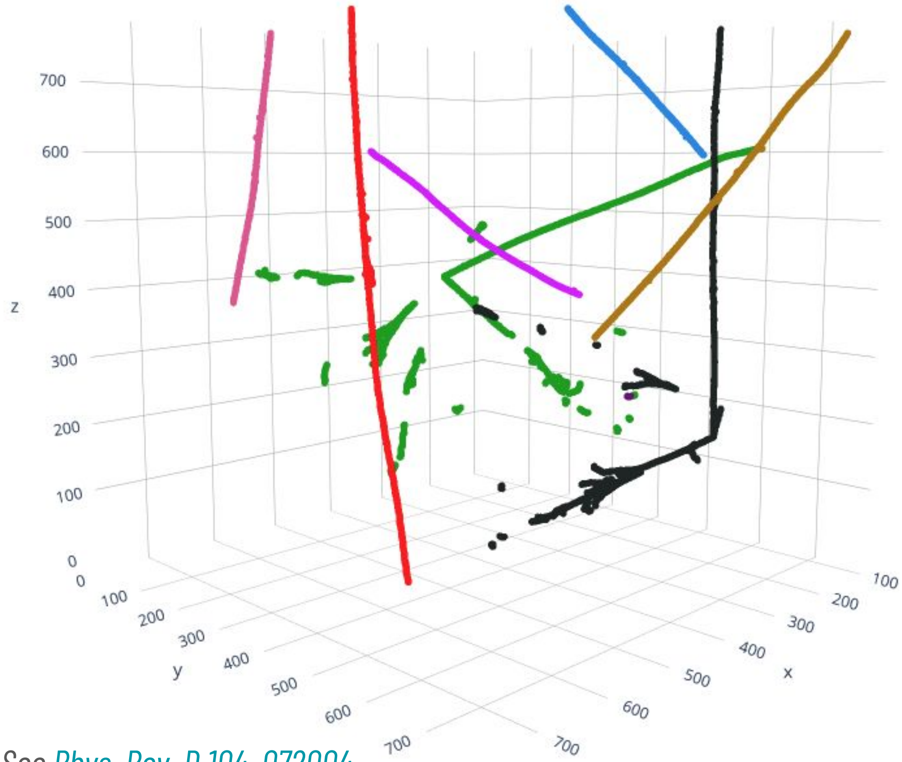
- Interaction = a group of particles that shared the same origin (i.e. neutrino interaction)
- Edge classification to identify an interaction
- Node classification for particle type ID

See [*Phys. Rev. D 104, 072004*]

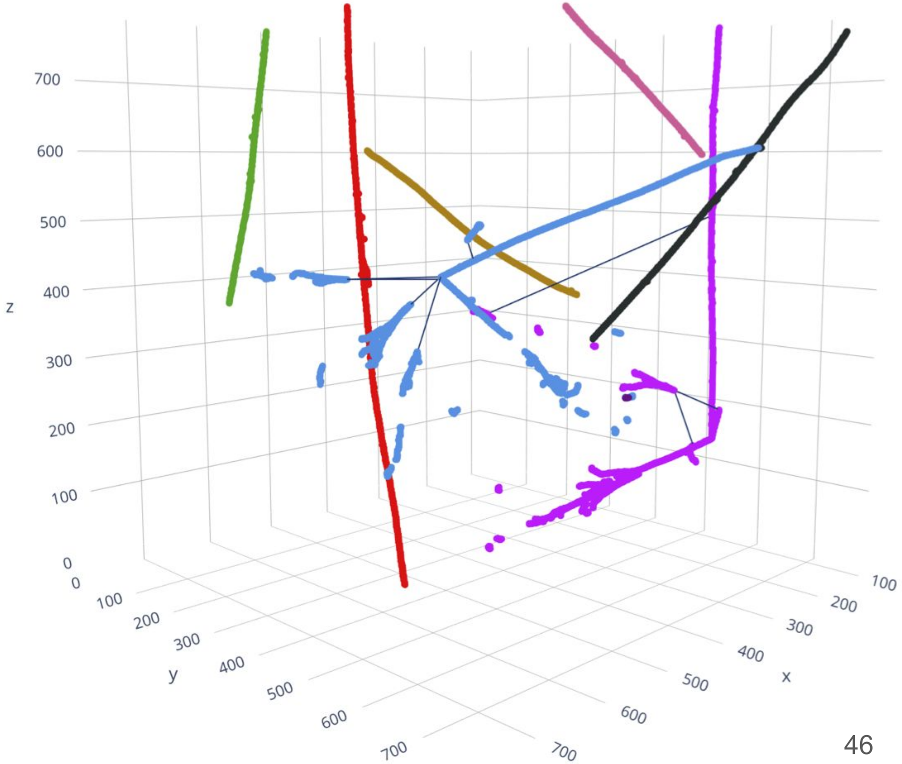
ML for Analyzing Big Image Data in Neutrino Experiments
Stage 3: Interaction Clustering

SLAC

Target Group

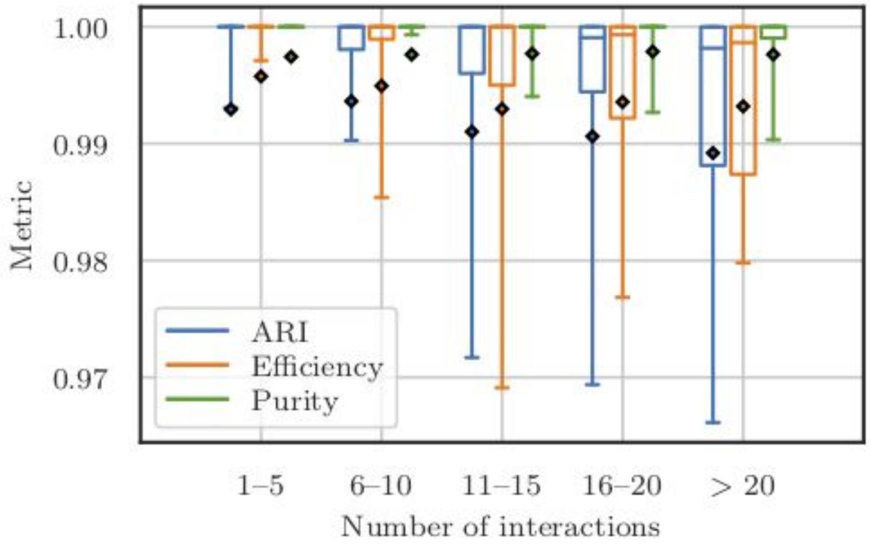


Predicted Interaction

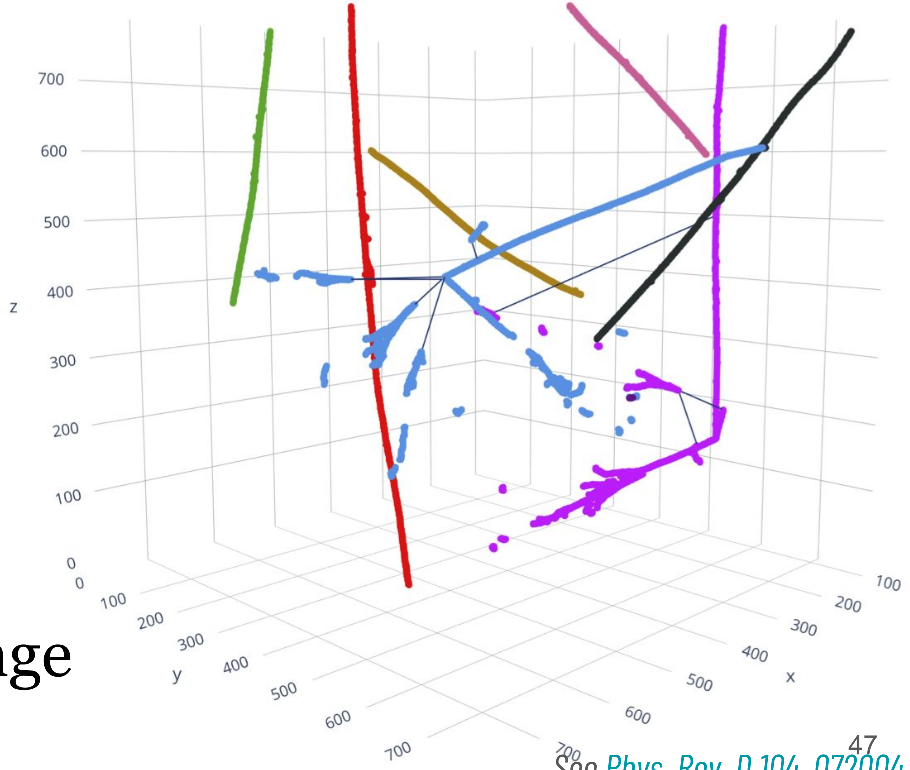


See [Phys. Rev. D 104, 072004]

ML for Analyzing Big Image Data in Neutrino Experiments
Stage 3: Interaction Clustering



Predicted Interaction



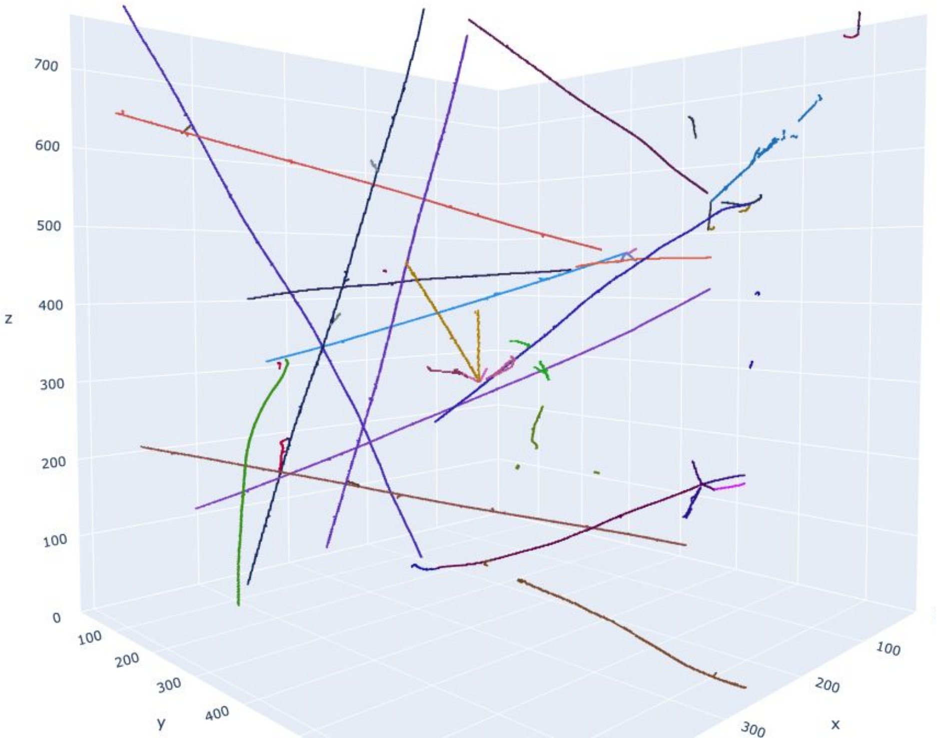
Promising result to address DUNE-ND reconstruction challenge (~20 neutrino pile-up)

See Phys. Rev. D 104, 072004

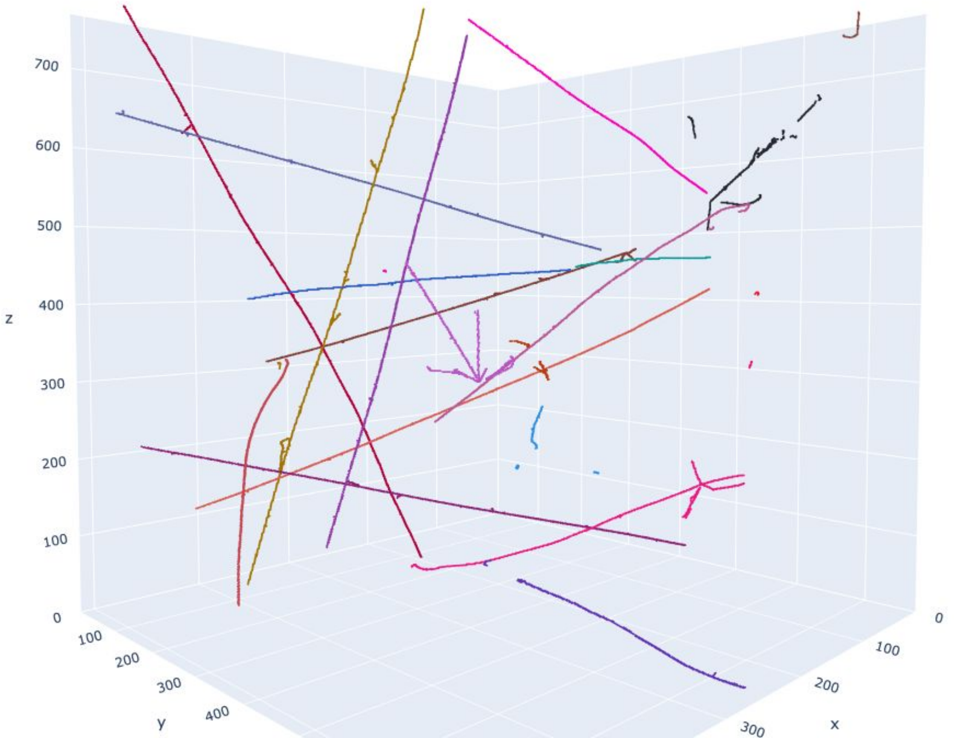
ML for Analyzing Big Image Data in Neutrino Experiments

Stage 3: input & output

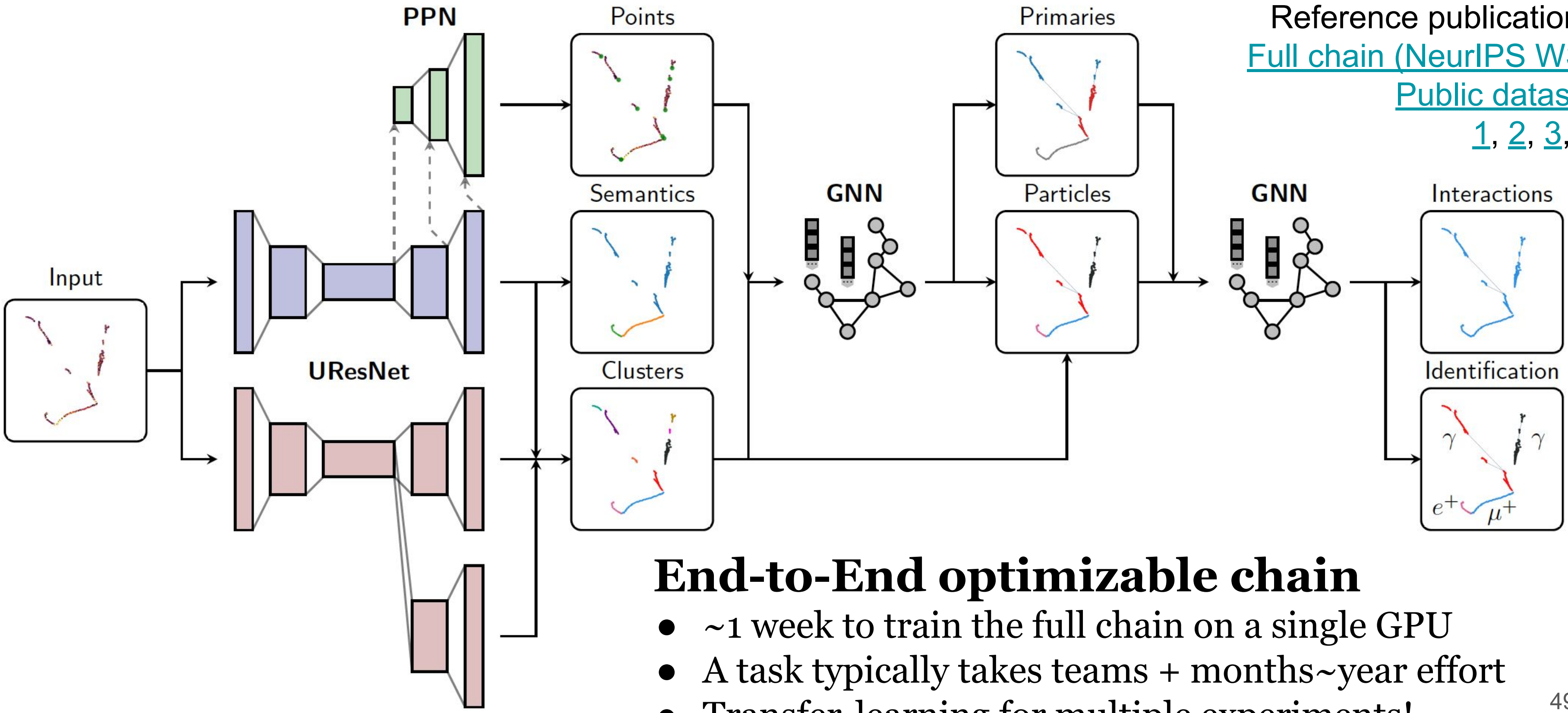
Stage 3 Input



Stage 3 Output



ML for Analyzing Big Image Data in Neutrino Experiments
Deep Neural Network for Data Reconstruction



Reference publications
[Full chain (NeurIPS WS)]
[Public dataset]
[1], [2], [3], [4]

End-to-End optimizable chain

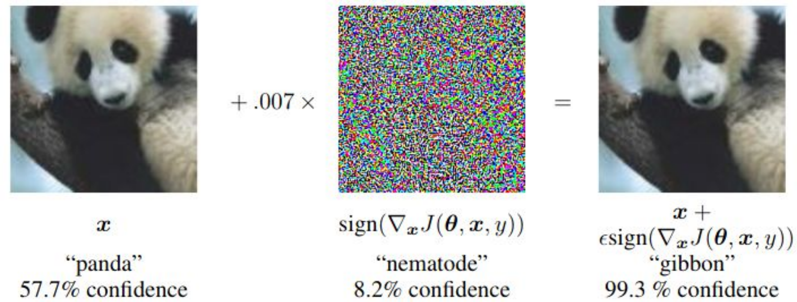
- ~1 week to train the full chain on a single GPU
- A task typically takes teams + months~year effort
- Transfer-learning for multiple experiments!

Physics model tuning

ML for Tuning Physics Models

Physics model tuning

Explaining and harnessing adversarial examples



The Catch

Used supervised optimization with simulated particle interactions, which may be imperfect (i.e. domain shift)

Present/Future Challenges

Lack of quality physics reconstruction for a big image data

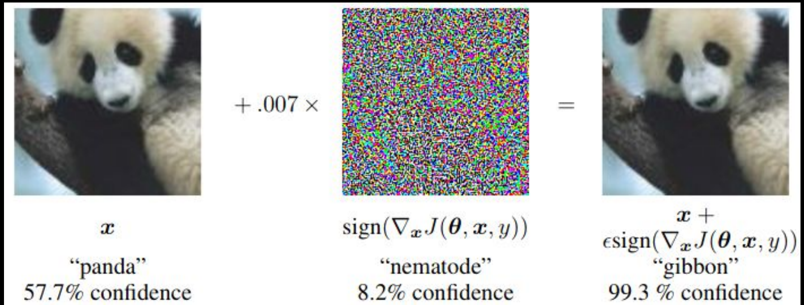
Slow, manual ("by-hand") workflow for development & tuning

Imperfect physics modeling

From an earlier slide

Physics model tuning

Explaining and harnessing adversarial examples

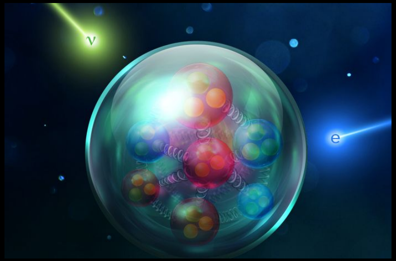


The Catch

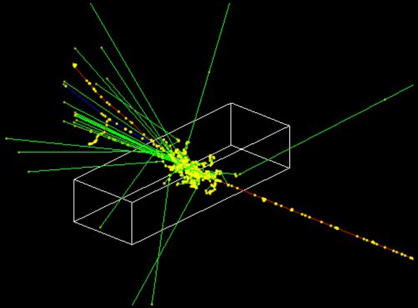
Used supervised optimization with simulated particle interactions, which may be imperfect (i.e. domain shift)

= multiple iterations of manual tuning

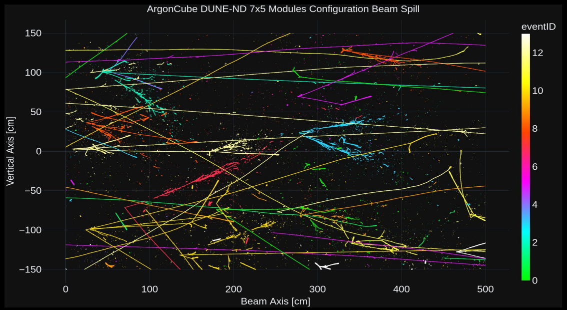
Most typical: detector mis-modeling



Fundamental particle interactions



Interaction with the detector volume



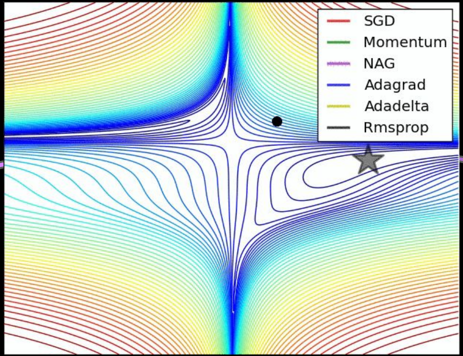
Detector response simulation

ML for Analyzing Big Image Data in Neutrino Experiments

Physics model tuning

Recent success in machine learning … much are backed by **deep learning**

… for which, one key success is **gradient-based optimization**



Analysis & reconstruction using neural networks

Input
x

ML (NN) parameters
$\boldsymbol{\theta}$

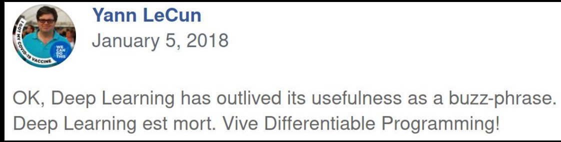
Output
$F\,(x|\theta)$

Optimization target
$L\,(\,F\,(x|\theta),\,y)$

Physics model tuning

Recent success in machine learning … much are backed by **deep learning**

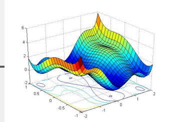
… for which, one key success is **gradient-based optimization**


Yann LeCun
January 5, 2018

OK, Deep Learning has outlived its usefulness as a buzz-phrase.
Deep Learning est mort. Vive Differentiable Programming!

Approximated gradient

Input
x



physics model
parameters
θ

Output
$F(x|\theta)$

Optimization
target
$L\,(\,F\,(x|\theta),\,y)$

Exact gradient

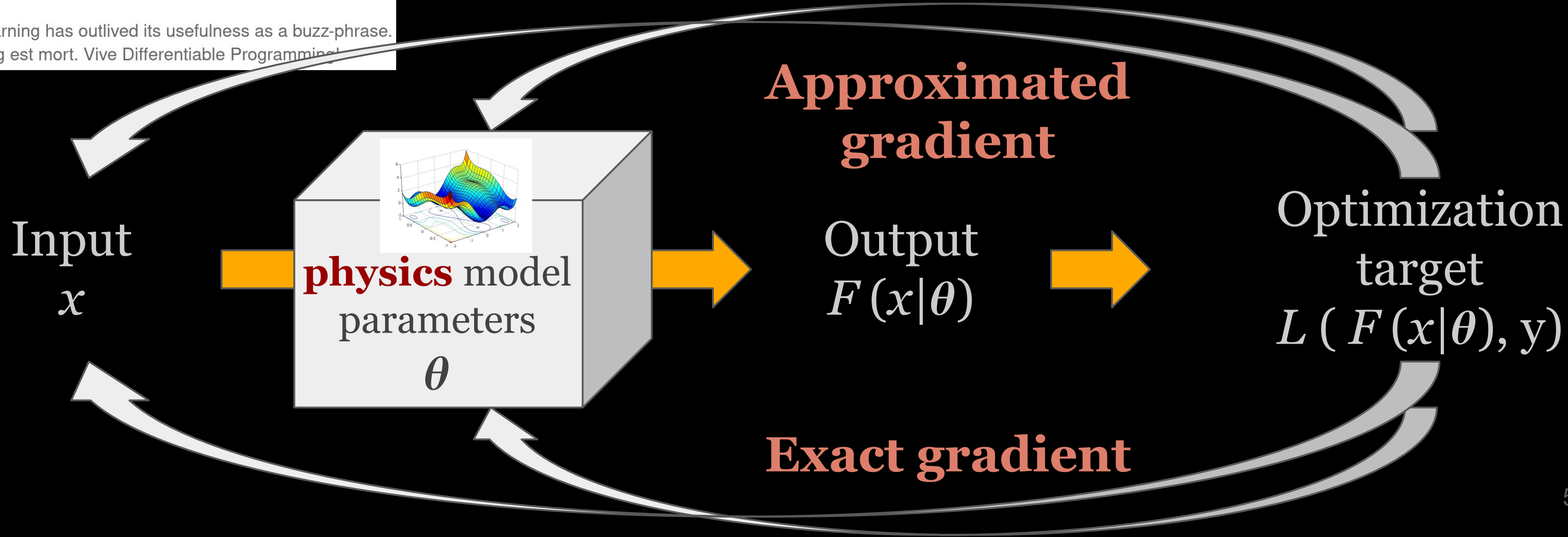
Physics model tuning

Recent success in machine learning … much are backed by **deep learning**

… for which, one key success is **gradient-based optimization**


Yann LeCun
January 5, 2018

OK, Deep Learning has outlived its usefulness as a buzz-phrase.
Deep Learning est mort. Vive Differentiable Programming!



Physics model tuning

Example Application
for
Modeling Detector Physics

ML for Analyzing Big Image Data in Neutrino Experiments

Differentiable detector simulator

Photo-multiplier tubes (PMTs) detect scintillation photons

Optical Photon Transport

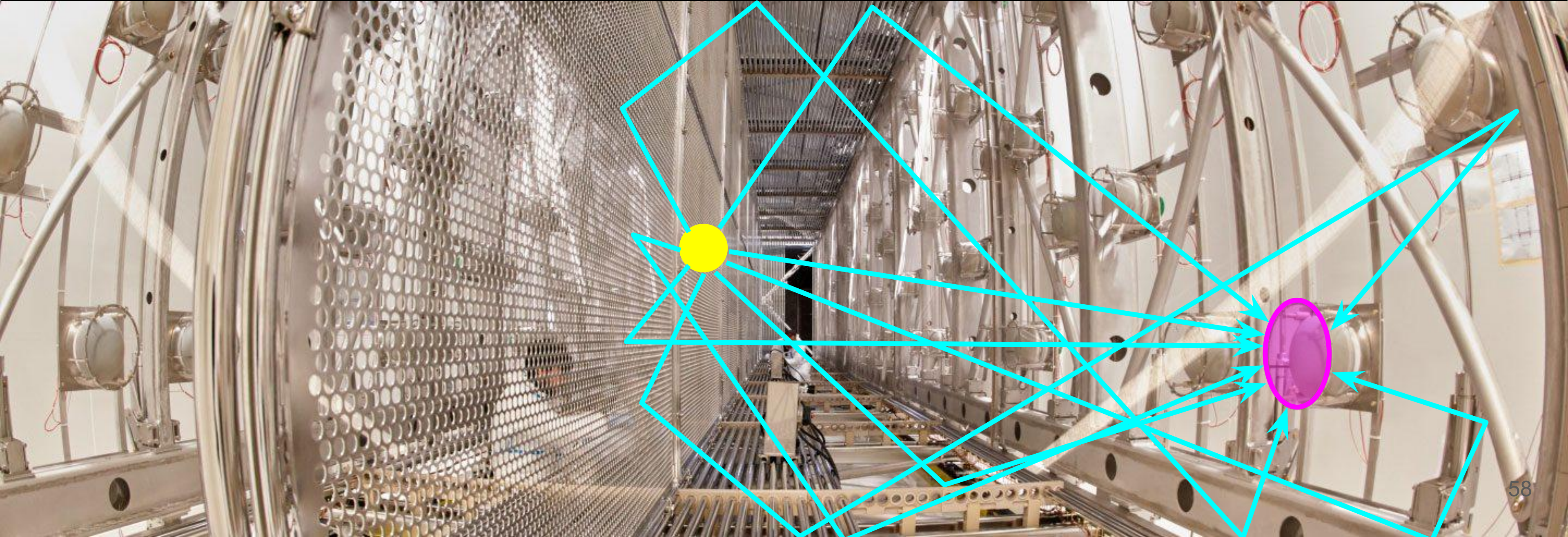
ML for Analyzing Big Image Data in Neutrino Experiments

Differentiable detector simulator

Photo-multiplier tubes (PMTs) detect scintillation photons produced isotropically from an Argon atom

1 meter muon produces **> 4M photons**

Optical Photon Transport



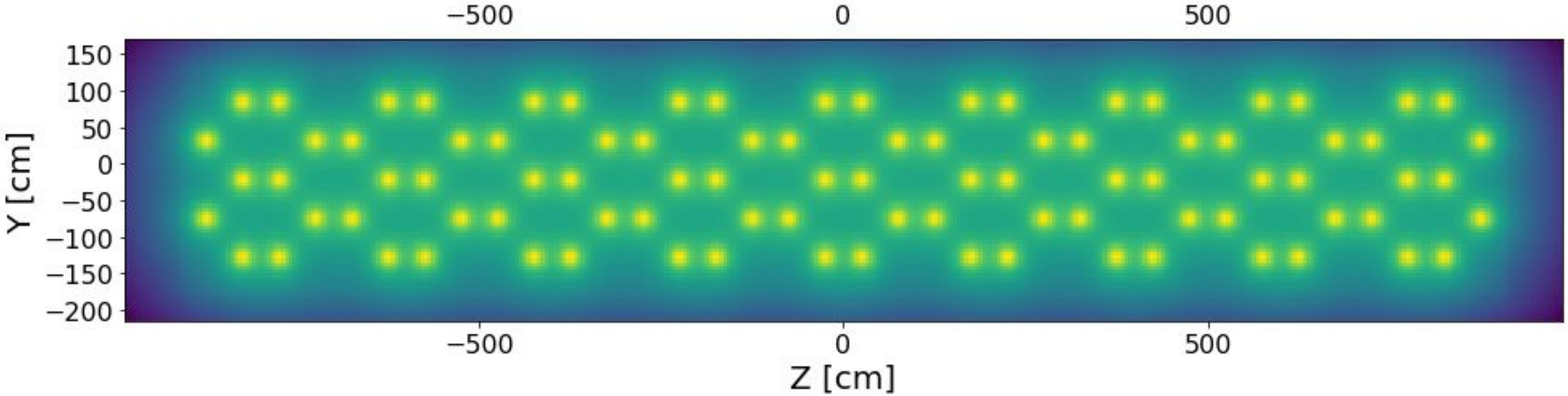
ML for Analyzing Big Image Data in Neutrino Experiments

Differentiable detector simulator

A marginalized "**Visibility Map**" for 3D voxelized volume used to estimate photon count at each PMT

Issue: static, not scalable

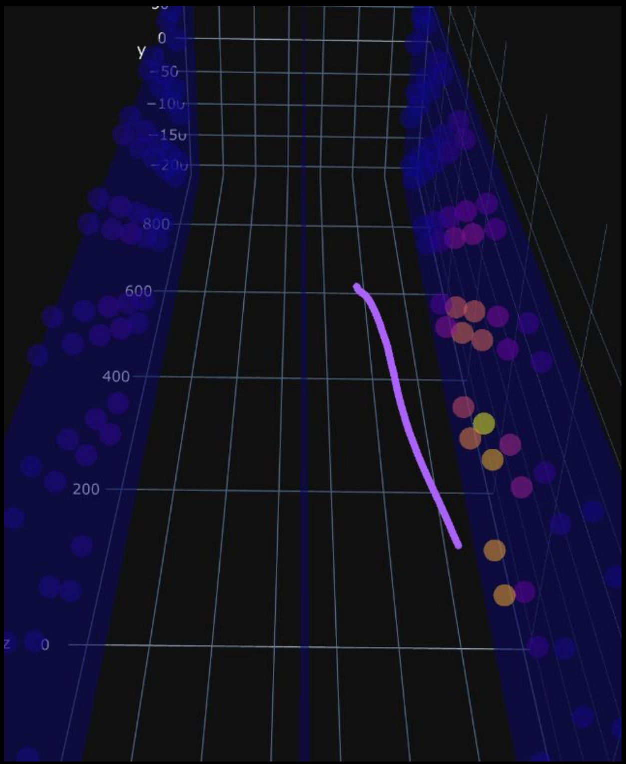
Optical Photon Transport



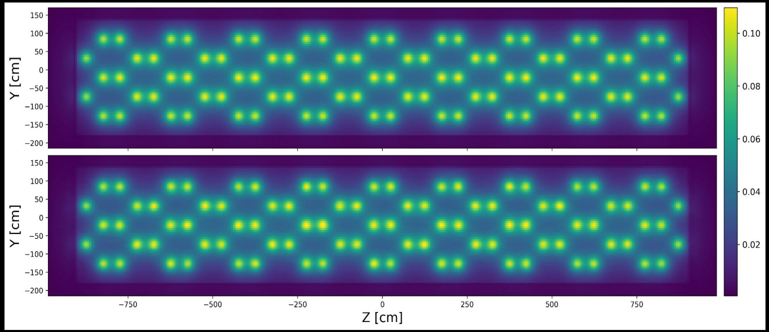
Example: ICARUS detector, 2D slice of a 3D map

ML for Analyzing Big Image Data in Neutrino Experiments

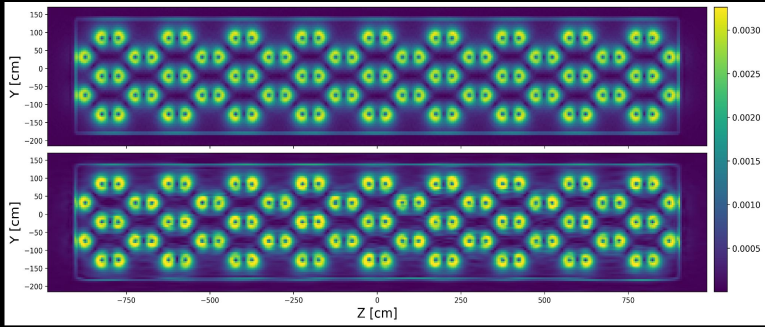
Differentiable detector simulator



Static map (top) v.s. SIREN



Gradient map (top, sobel filter) v.s. SIREN

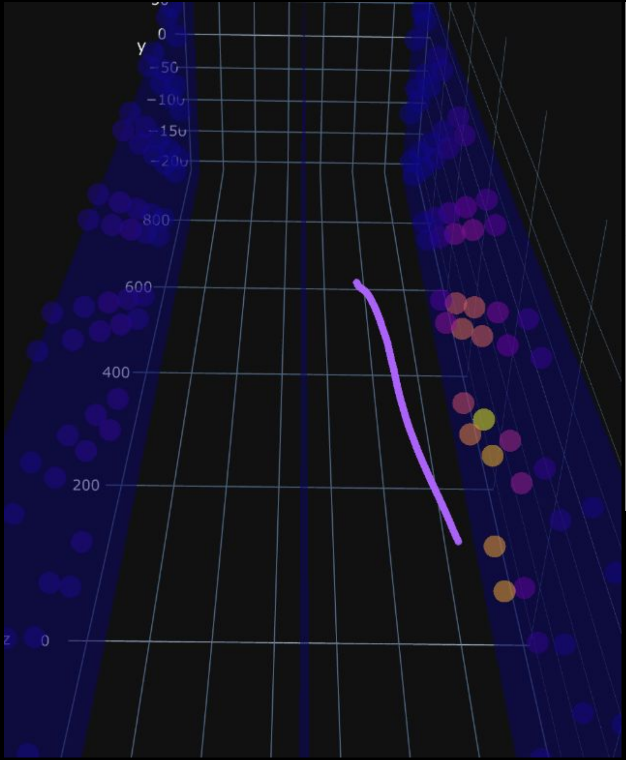


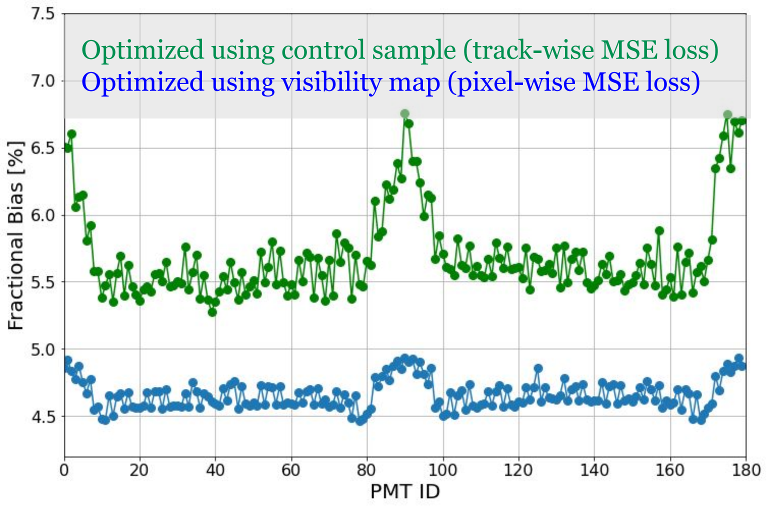
Optical Photon Transport using **Differentiable Surrogate (SIREN)**

Neural scene representation (alternative: NeRF inc. differentiable rendering)

ML for Analyzing Big Image Data in Neutrino Experiments

Differentiable detector simulator





Work credit (from left): Olivia P. (UC Berkeley), Minjie L. (SLAC), Patrick T. (SLAC), , Gordon W. (Stanford CS), Chuan L. (Lambda Labs)

Optical Photon Transport using **Differentiable Surrogate (SIREN)**

Neural scene representation (alternative: NeRF inc. differentiable rendering)

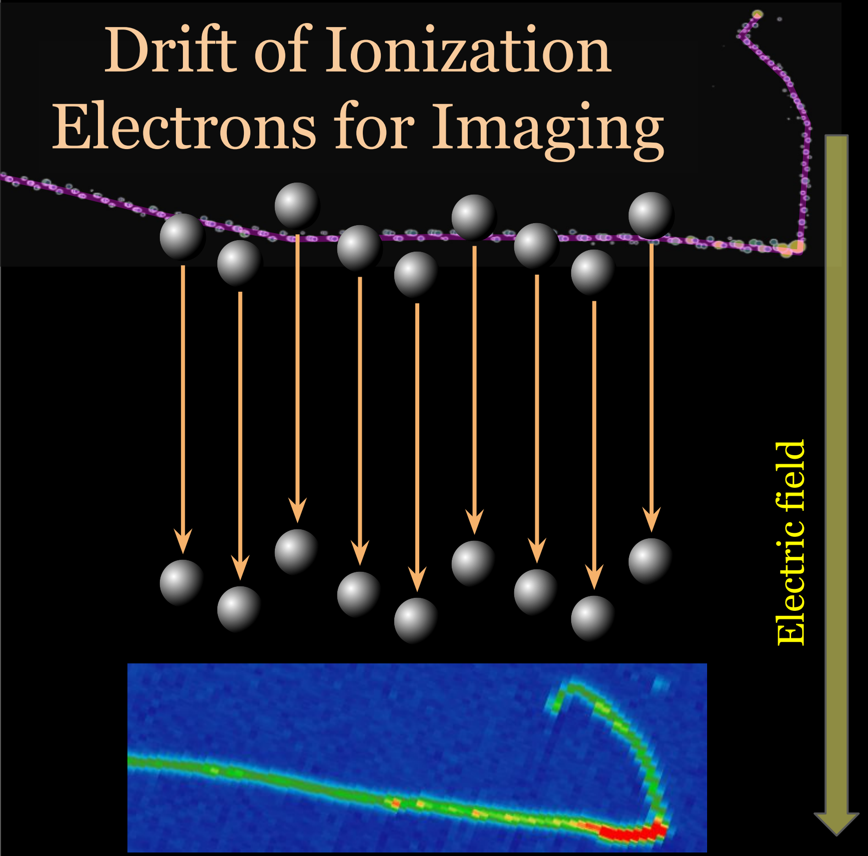
Drift of Ionization Electrons for Imaging

Differentiable detector simulator

Drift of Ionization Electrons for Imaging

1. Particle ionize Argon

Differentiable detector simulator



1. Particle ionize Argon
2. Ionization electron drift in E-field at a constant velocity, some charge lost due to capture
3. Imaging by charge-sensitive plane (detectors) at the anode

Tuning simulation = extract physics model parameter values from data

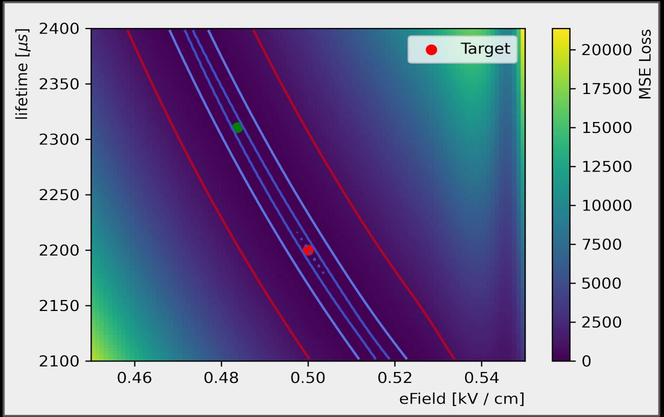
ML for Analyzing Big Image Data in Neutrino Experiments

Differentiable detector simulator

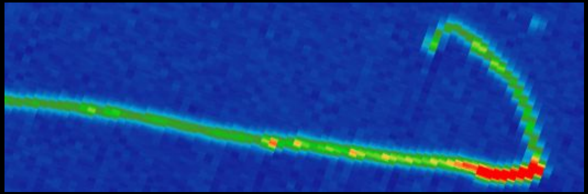
Drift of Ionization Electrons for Imaging

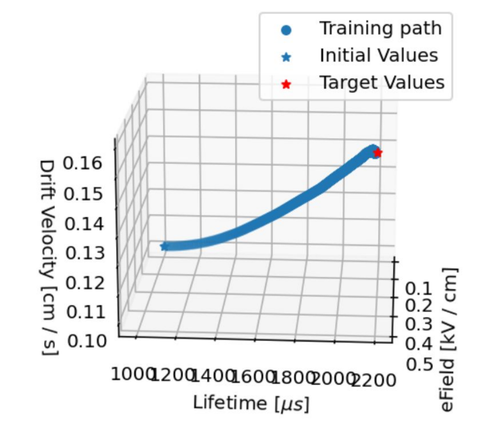
Differentiable Simulator

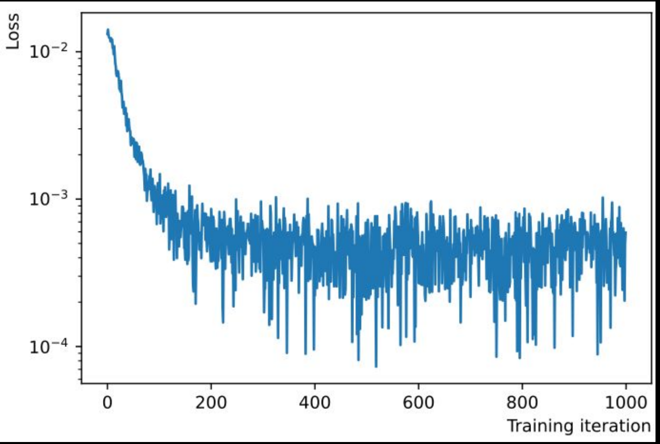
using explicit gradient calculation using AD-enabled tools (JAX/Pytorch)



Work credit due (from left):
SLAC-ML: Youssef N., Sean G., Daniel R.
SLAC-neutrino: Yifan C.
LBNL-neutrino: Roberto S.







Differentiable detector simulator

Beyond detector physics modeling

- Neutrino-nucleus event generator
  - Diff. simulator for neutrino interaction, hadronization, etc.
  - Modeling many-body particle interactions inside a nucleus
- Modeling of particle passage through medium (e.g. stochastic "shower")
- Fast surrogate to enable testing of new models with *very high* statistics

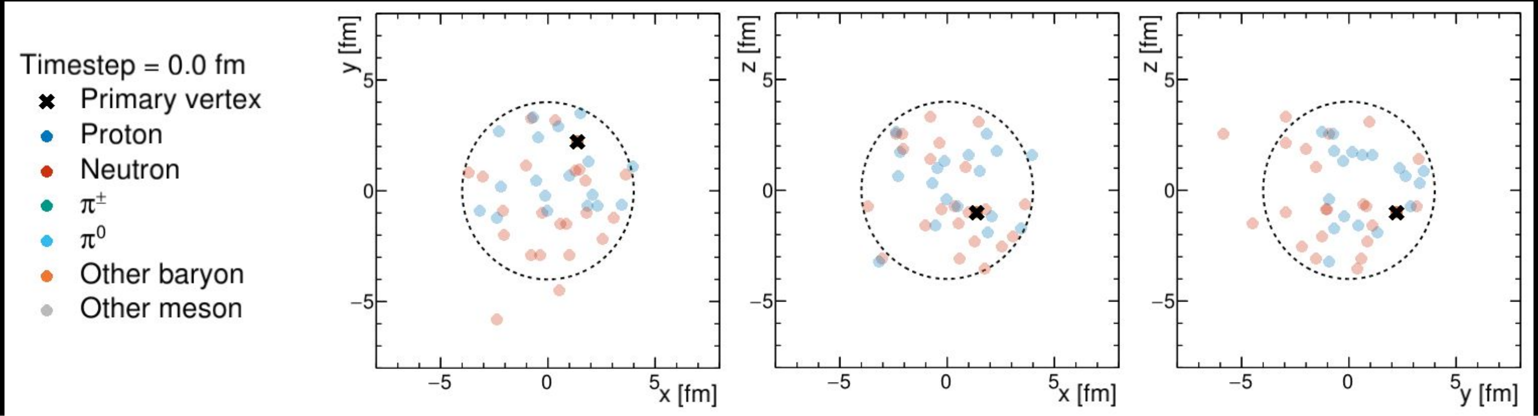


Image credit/collaborator Callum Wilkinson (LBNL)

… wrapping up …

Wrapping up...

Summary

- **Neutrino detector trend: hi-res. particle imaging**
- **ML, in particular computer vision, + reconstruction**
  - ML-based approach has shown strong promise + tuning automation
  - Extension skipped in this talk: calibrated uncertainty quantification
- **Emerging area: differentiable physics modeling**
  - part of a larger trend, simulation-based inference
  - detector physics modeling a primary target to automate tuning
  - event generator will be a new frontier of active research (my view)

Thank you for your attention!

Back-up slides

Reconstruction Details

Stage 1-a: Pixel Feature Extraction + Scalablility

“Applying CNN” is simple, but **is it scalable for us?**

CNN applies **dense matrix operations**

In photographs, **all pixels are meaningful**

grey pixels = dolphins,
blue pixels = water, etc…

Stage 1-a: Pixel Feature Extraction + Scalablility

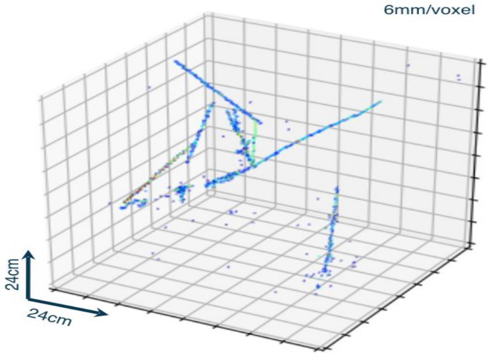
"Applying CNN" is simple, but **is it scalable for us?**

LArTPC data is generally sparse, but locally dense

CNN applies **dense matrix operations**

In photographs, **all pixels are meaningful**

grey pixels = dolphins,
blue pixels = water, etc…



Empty pixels = no energy

<1% of pixels are non-zero in LArTPC data

Zero pixels are meaningless!

Figures/Texts: courtesy of Laura Domine @ Stanford

Figure credit: Laura Domine @ Stanford

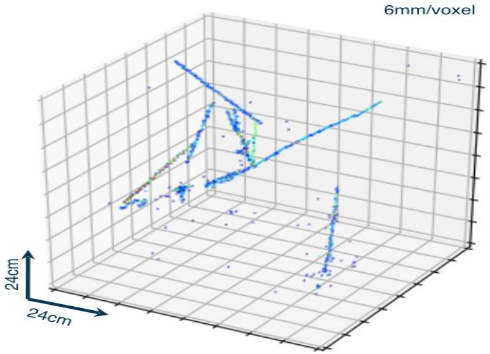
Stage 1-a: Pixel Feature Extraction + Scalablility

"Applying CNN" is simple, but is it scalable for us?

LArTPC data is generally sparse, but locally dense

CNN applies **dense matrix operations**

In photographs, **all pixels are meaningful**



<1% of pixels are non-zero in LArTPC data

Zero pixels are meaningless!

Figures/Texts: courtesy of Laura Domine @ Stanford

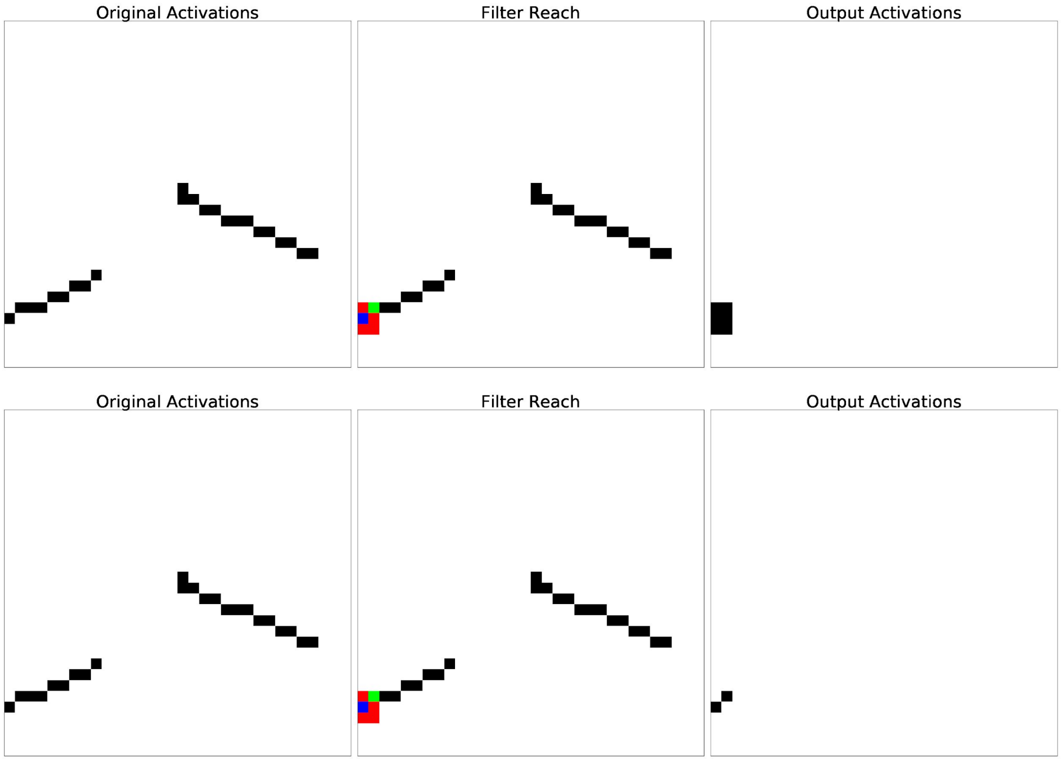
- **Scalability for larger detectors**
  - Computation cost increases linearly with the volume
  - But the number of non-zero pixels does not

Figure credit: Laura Domine @ Stanford

Sparse Submanifold Convolutions

Only acts on an active input pixels + can limit output activations for only the same pixels.

- 1st implementation by FAIR
- 2nd implementation by Stanford VL
  - ... also supported in NVIDIA now



ML-based Neutrino Data Reconstruction Chain

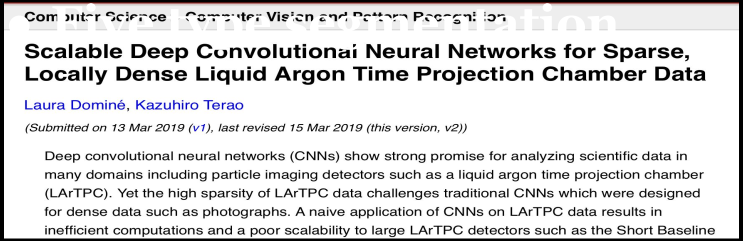
Stage 1-a: Pixel Feature Extraction + Scalablility

SLAC

Type	Proton	Mu/Pi	Shower	Delta	Michel
Acc.	0.99	0.98	0.99	0.97	0.96

CNN on sparse tensors (MinkowskiEngine)

- **Public LArTPC simulation**
  - Particle tracking (Geant4) + diffusion, no noise, true energy


Computer Science > Computer Vision and Pattern Recognition

Scalable Deep Convolutional Neural Networks for Sparse, Locally Dense Liquid Argon Time Projection Chamber Data

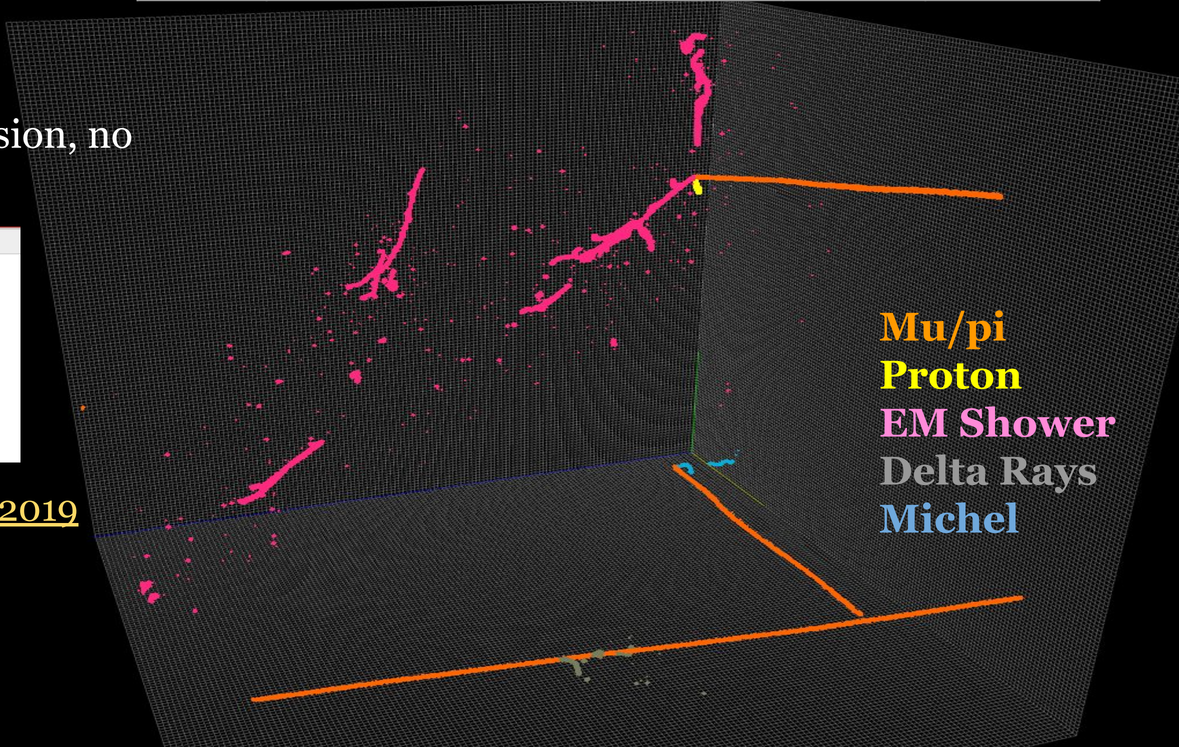
Laura Dominé, Kazuhiro Terao

(Submitted on 13 Mar 2019 (v1), last revised 15 Mar 2019 (this version, v2))

Deep convolutional neural networks (CNNs) show strong promise for analyzing scientific data in many domains including particle imaging detectors such as a liquid argon time projection chamber (LArTPC). Yet the high sparsity of LArTPC data challenges traditional CNNs which were designed for dense data such as photographs. A naive application of CNNs on LArTPC data results in inefficient computations and a poor scalability to large LArTPC detectors such as the Short Baseline

PhysRevD.102.012005 presented @ ACAT 2019

- Memory reduction ~ 1/360
- Compute time ~ 1/30
- Handles large future detectors

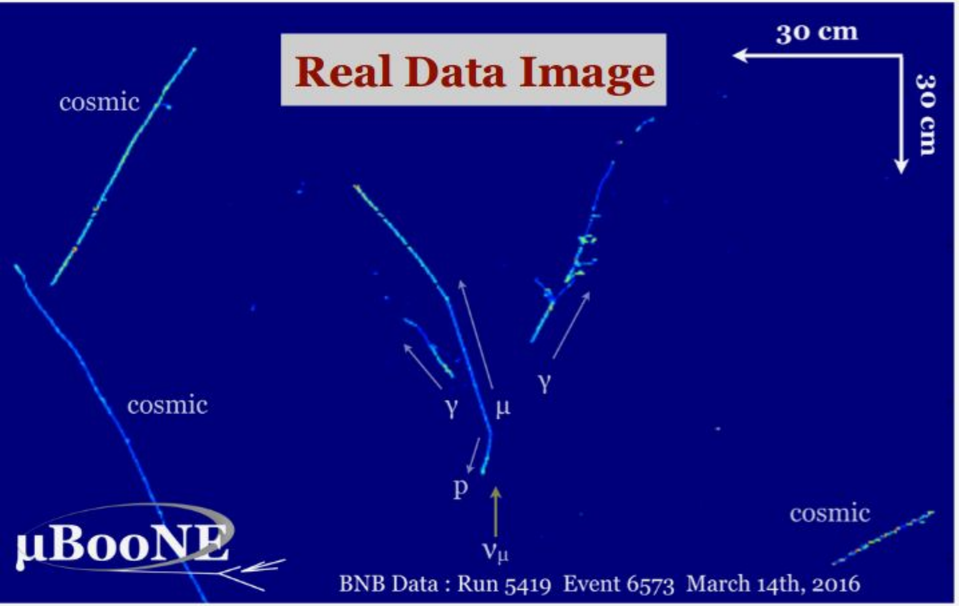


ML for Analyzing Big Image Data in Neutrino Experiments

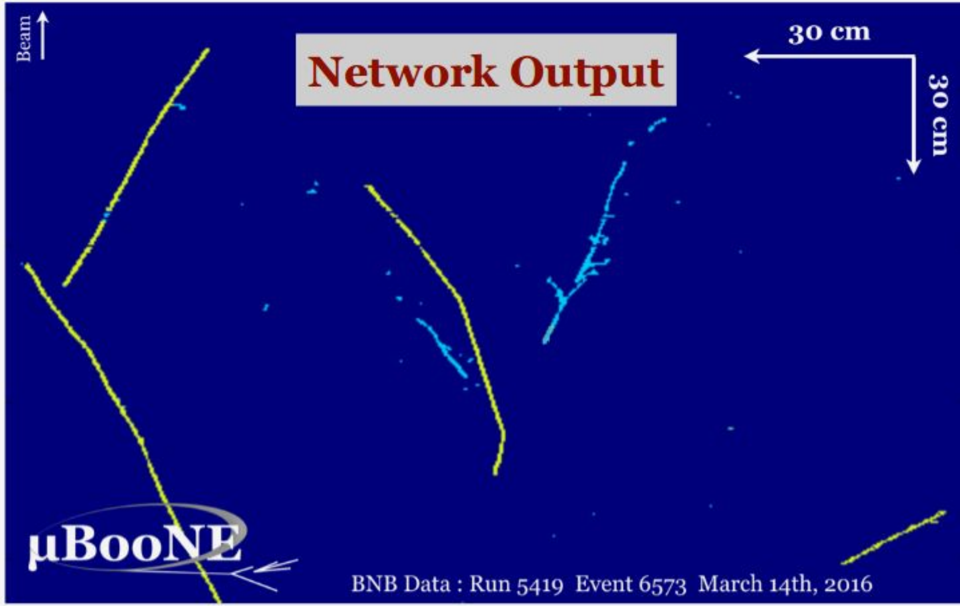
Stage 1-a: Pixel Feature Extraction + Scalablility

Distinguish 2 distinct particle topologies: **showers v.s. tracks**

Critical to deploy different algorithms for clustering pixels in the next stage.



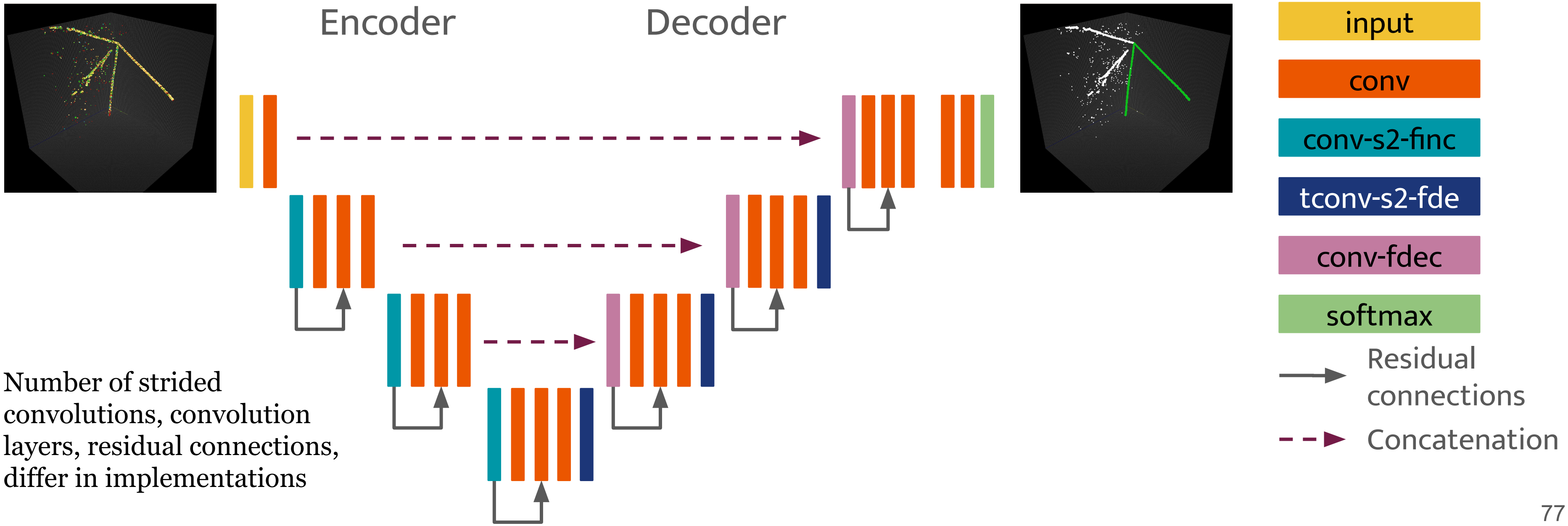
Network Input



Network Output

PRD 99 092001
arXiv:1808.07269

Architecture: **U-Net + Residual Connections**



Number of strided convolutions, convolution layers, residual connections, differ in implementations

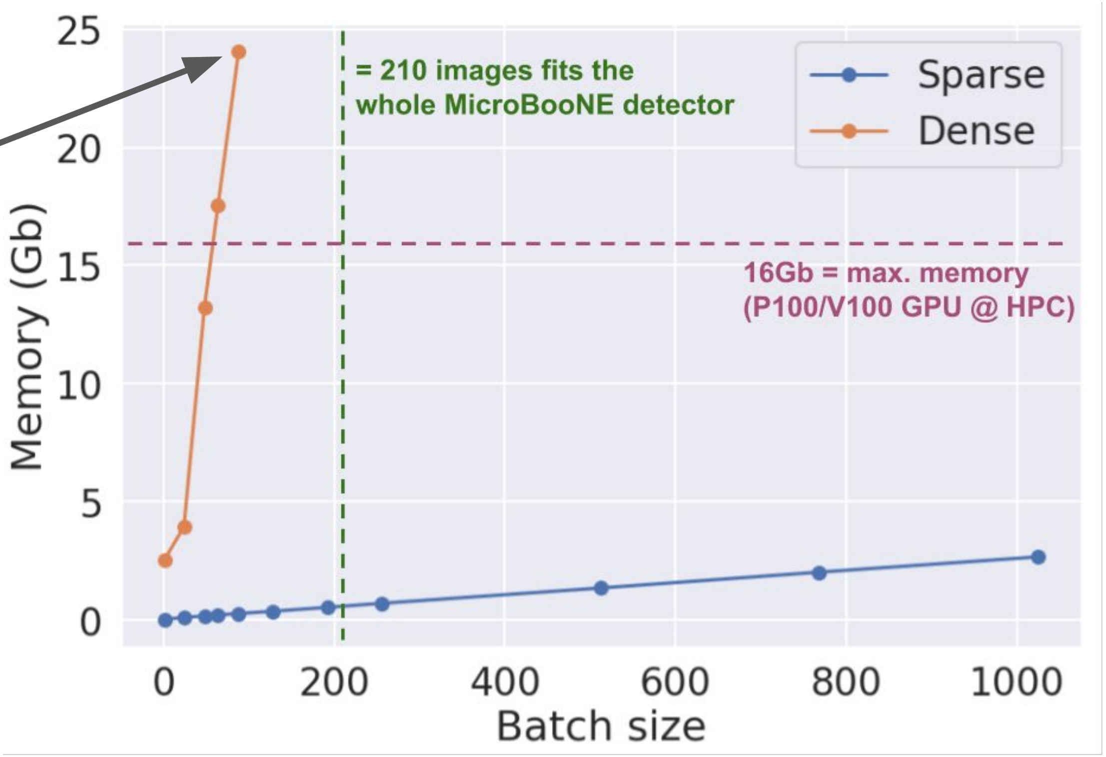
Image credit: Laura Domine @ Stanford

Stage 1-a: Pixel Feature Extraction + Scalablility

SLAC

Sparse U-ResNet fits more data in GPU + good scalability

@batch size 88 sparse uses **93x less memory** than dense and computation is **3x faster**



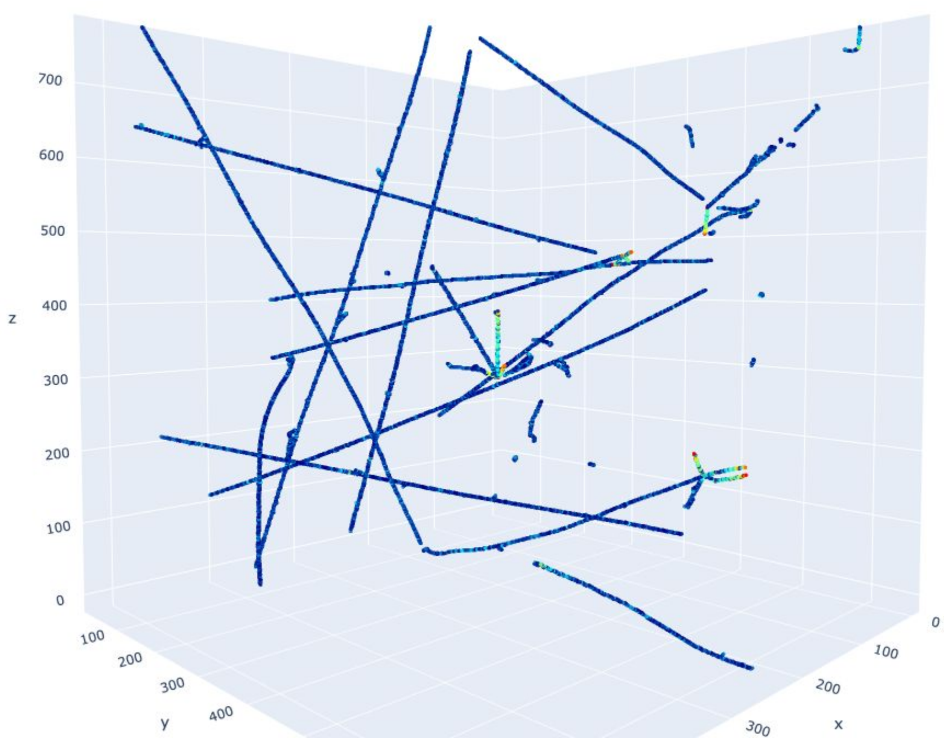
Can handle easily the whole ICARUS detector which is x6 larger than MicroBooNE.

DUNE-FD is piece of cake (larger volume but less non-zero pixels)

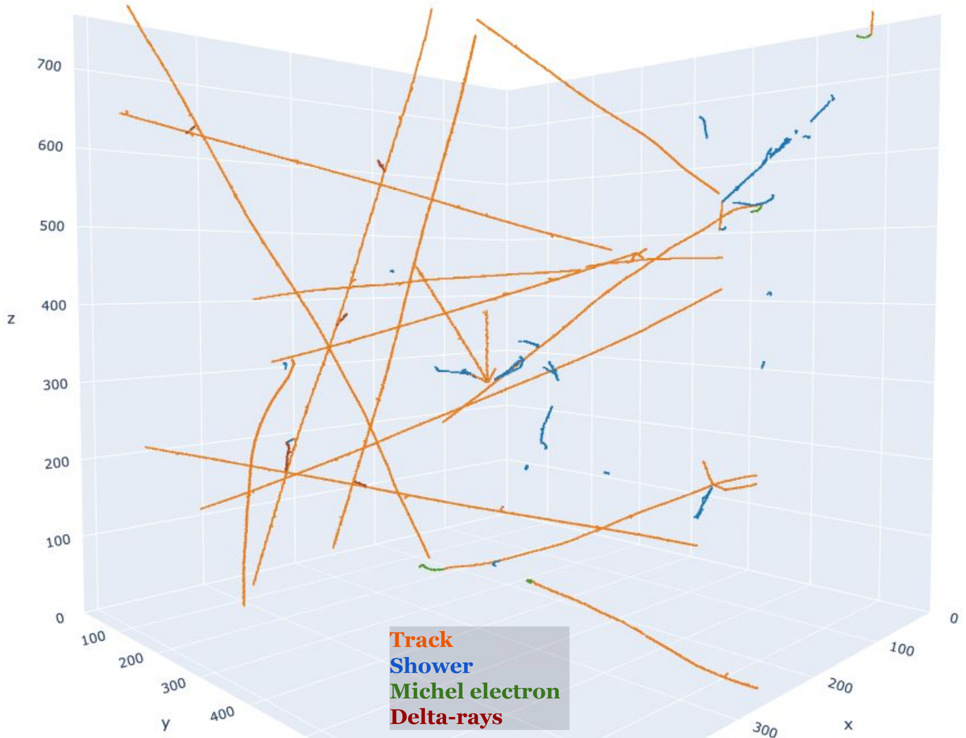
Work credit: Laura Domine (Stanford) and Ran Itay (SLAC)

ML for Analyzing Big Image Data in Neutrino Experiments
Stage 1-a: input & output

Stage 1-a Input

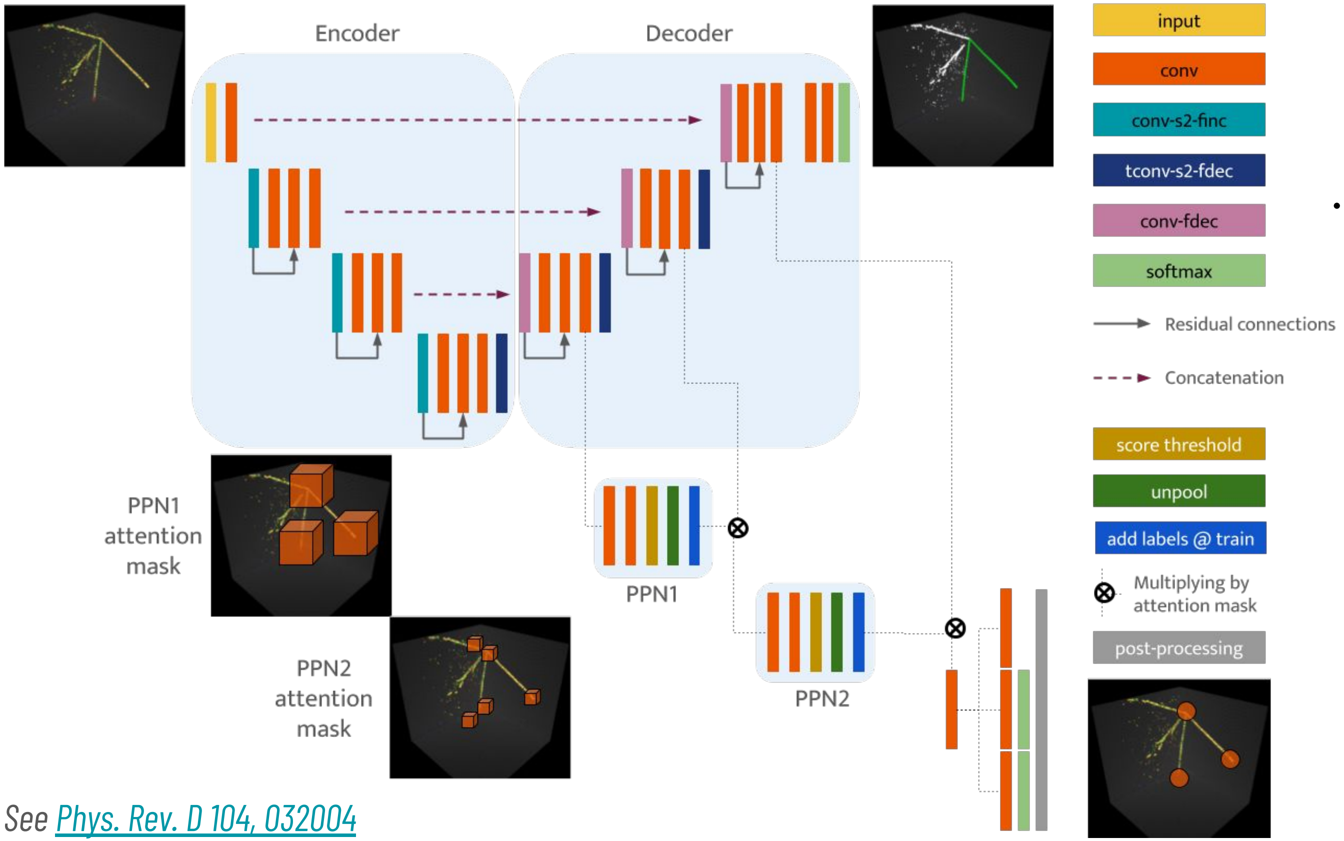


Stage 1-a Output



ML for Analyzing Big Image Data in Neutrino Experiments

Stage 1-b: Particle Edge-point Prediction



Point Proposal Network (PPN)

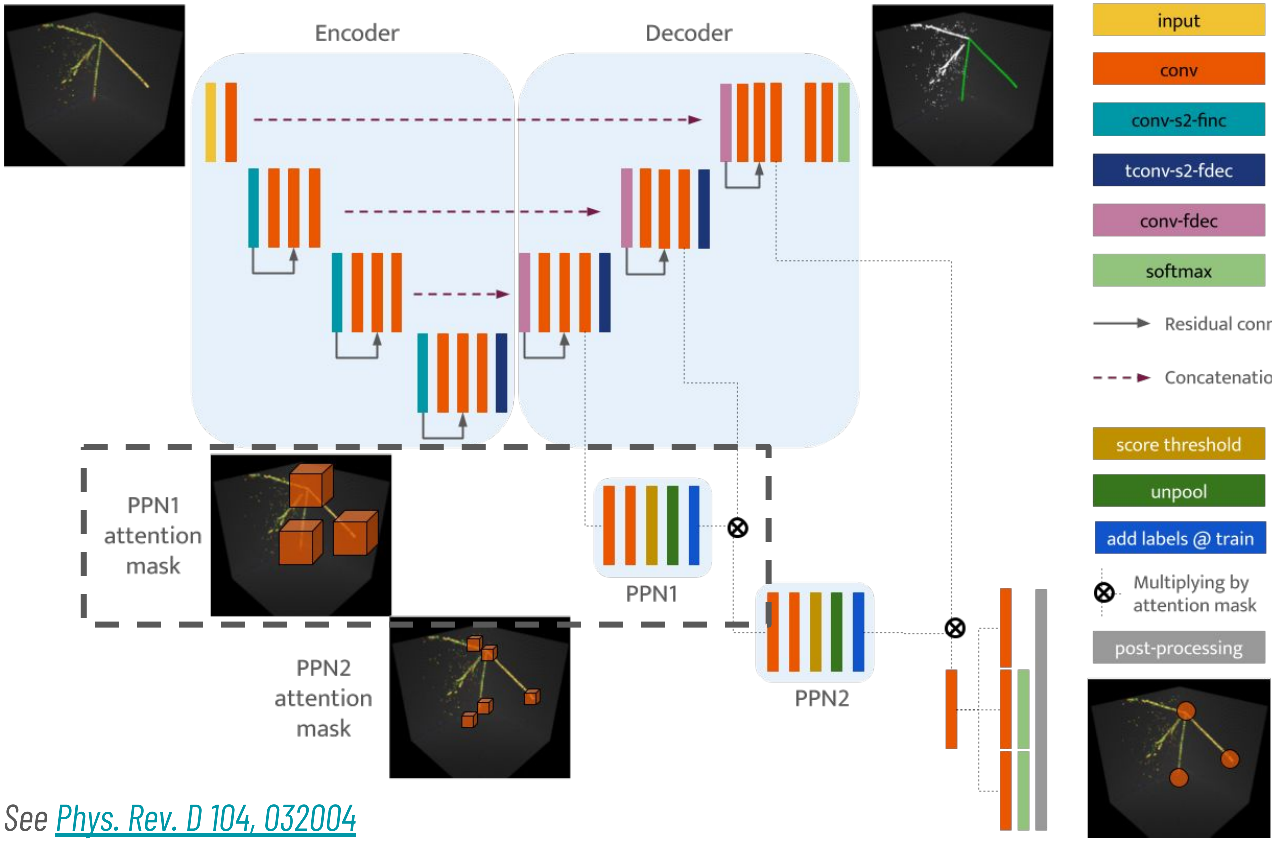
… extension of U-ResNet with 3 CNN blocks

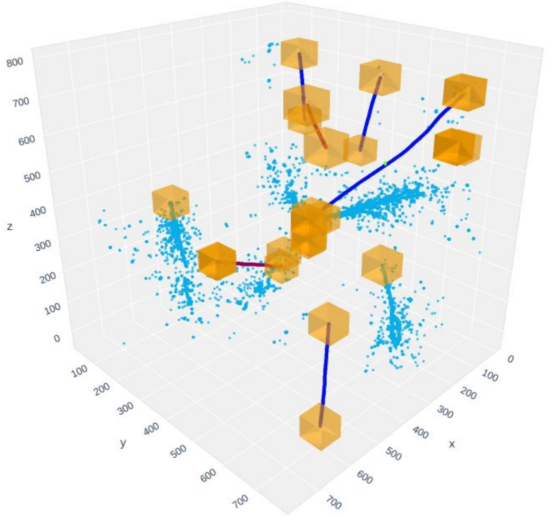
See Phys. Rev. D 104, 032004

Work credit: Laura Domine (Stanford) and Patrick Tsang (SLAC)

ML for Analyzing Big Image Data in Neutrino Experiments

Stage 1-b: Particle Edge-point Prediction



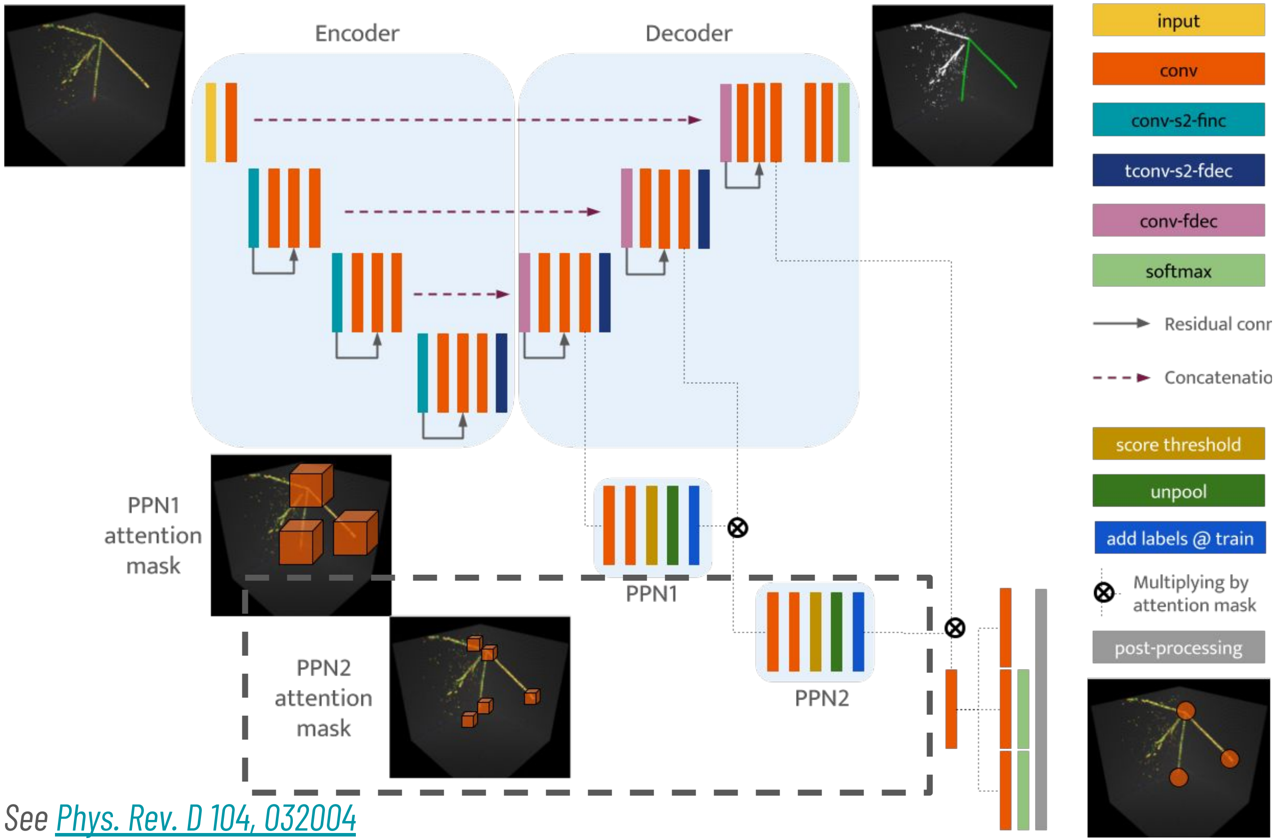


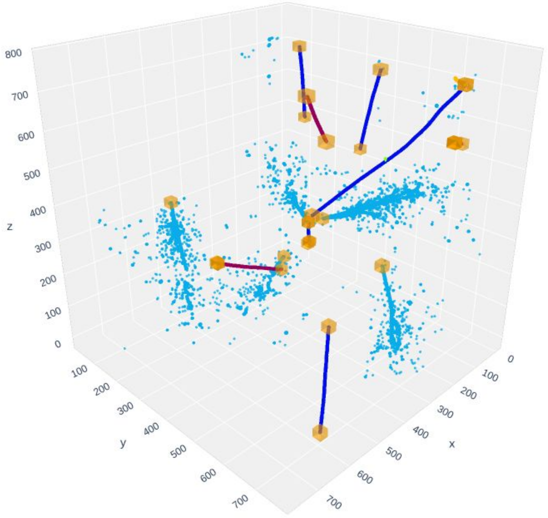
PPN1 generates an attention mask at the lowest resolution

See [Phys. Rev. D 104, 032004](#)

ML for Analyzing Big Image Data in Neutrino Experiments

Stage 1-b: Particle Edge-point Prediction

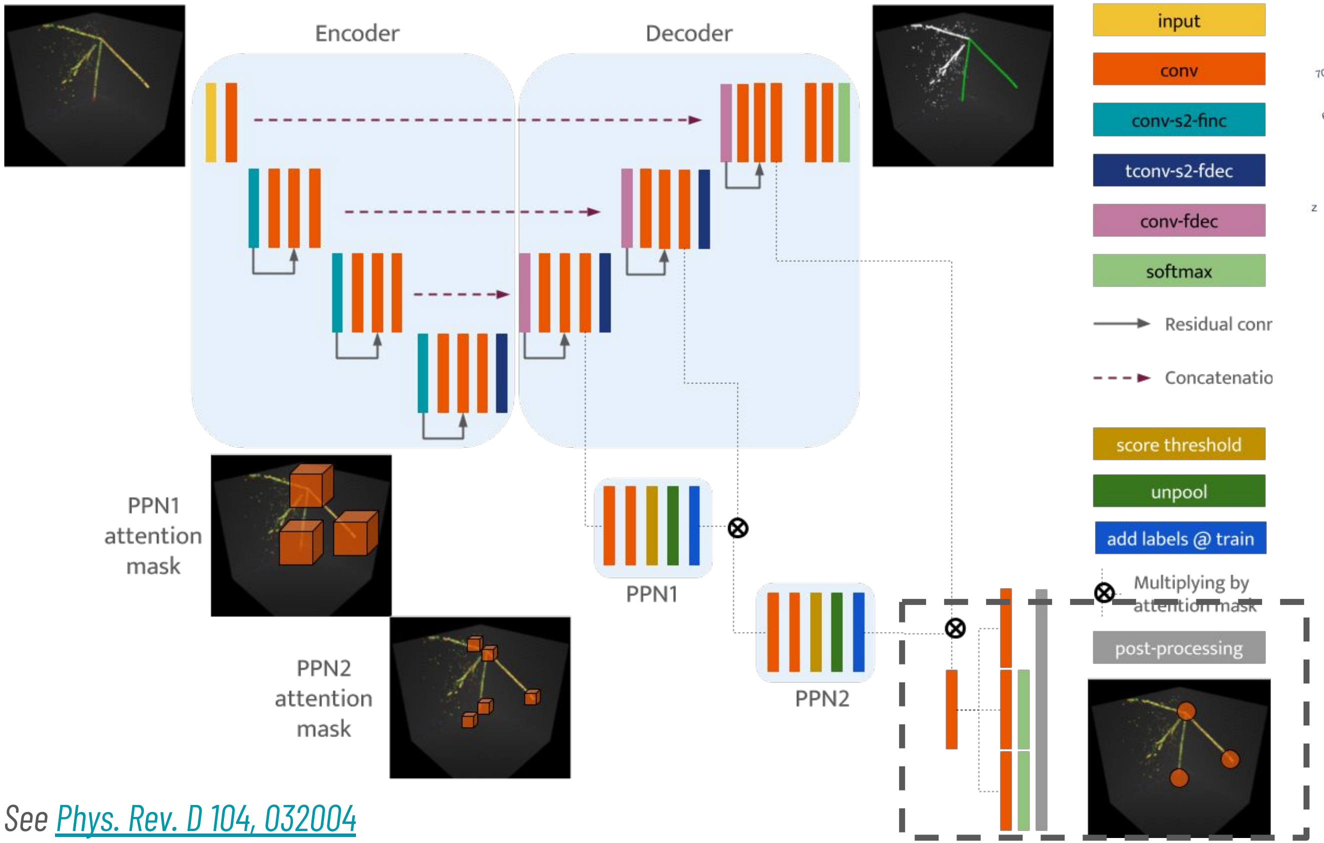


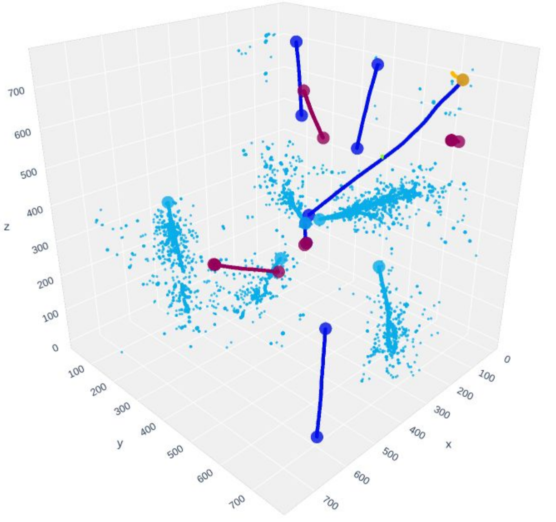


PPN2 generates an attention mask at the intermediate resolution

See Phys. Rev. D 104, 032004

ML for Analyzing Big Image Data in Neutrino Experiments
Stage 1-b: Particle Edge-point Prediction





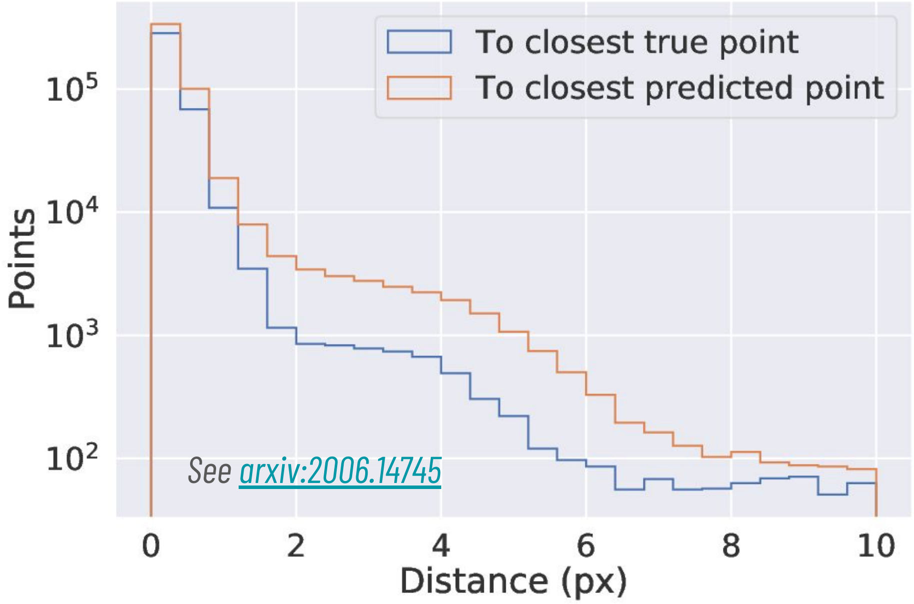
PPN makes the final prediction (point type + coordinate regression)

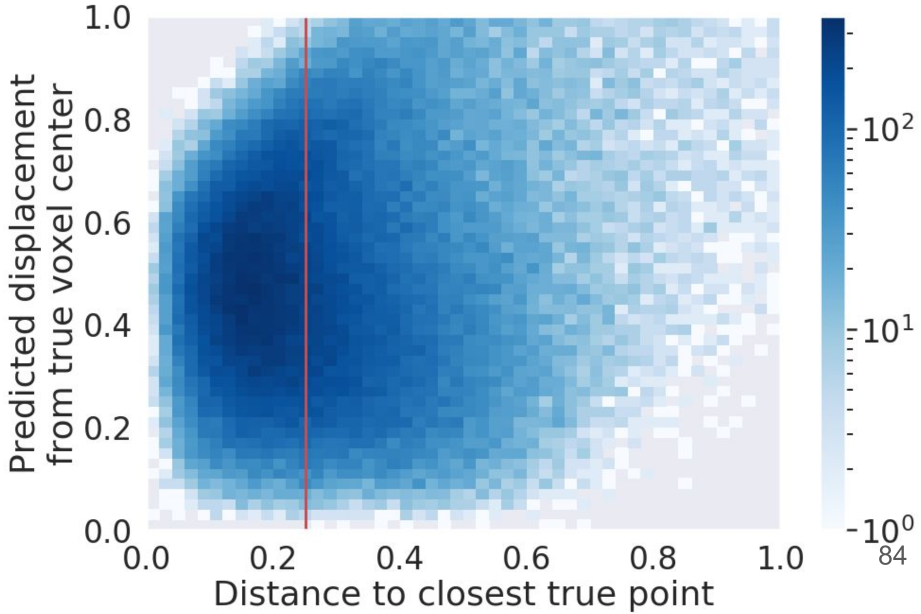
See [Phys. Rev. D 104, 032004]

ML-based Neutrino Data Reconstruction Chain
Stage 1-b: Particle Endpoint Prediction

96.8% of predicted points within 3 voxels of a true point

- 68% of true points found within the radius of 0.12 cm
- Traditional (nominal) reconstruction method finds 90% of predicted points within 17 voxels, and 68% of true points found within the radius of 0.74cm

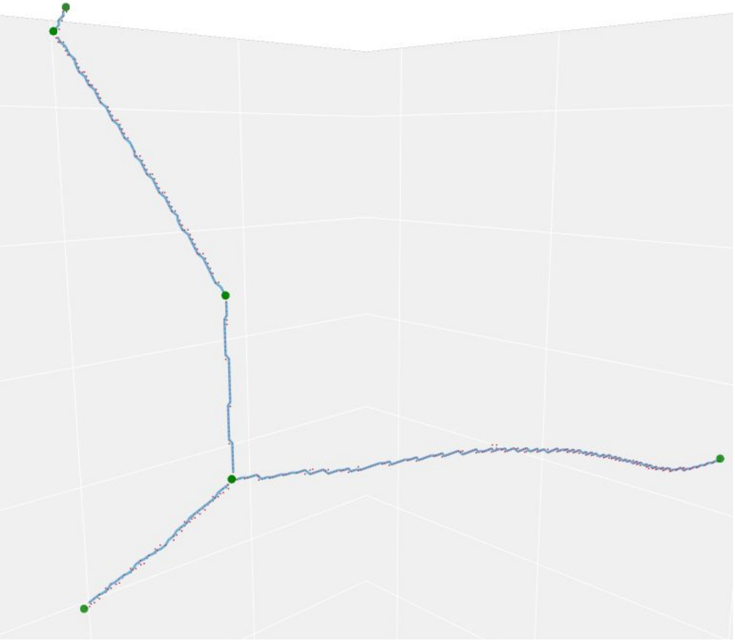


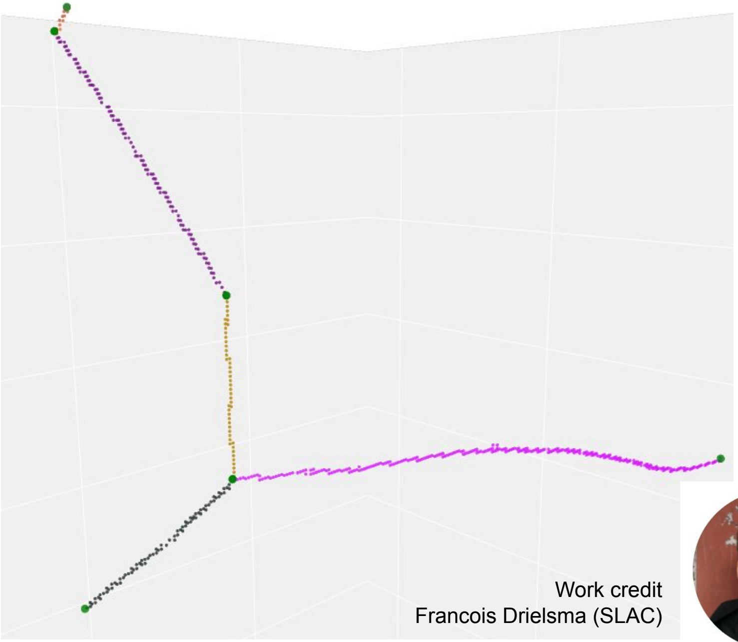


ML-based Neutrino Data Reconstruction Chain
Stage 2-a: Dense Pixel Clustering

Simple approach: path-finding between PPN points

- MST to find the “shortest” path between PPN points to cluster pixels
- **Works well! BUT** it depends on PPN performance directly + not learnable

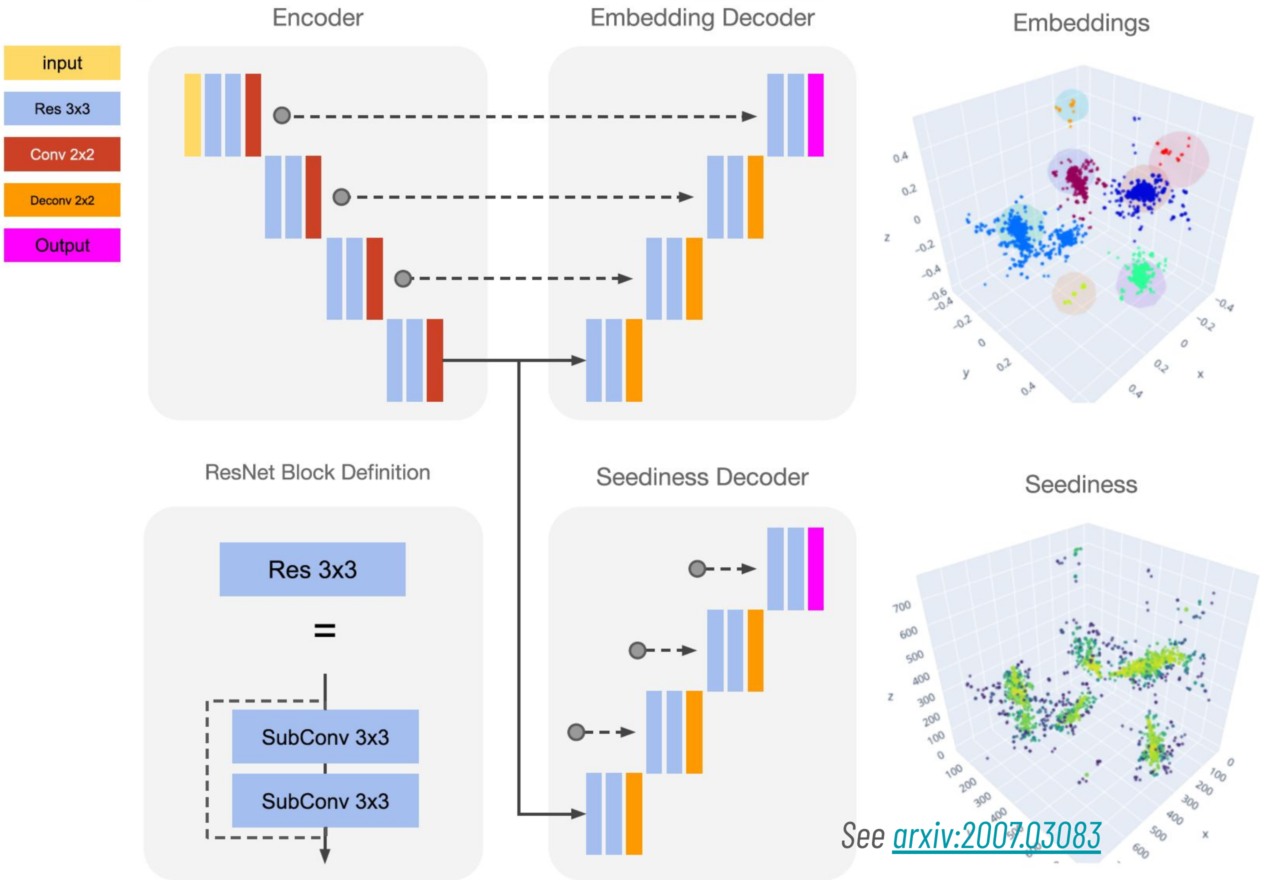




Work credit
Francois Drielsma (SLAC)

ML for Analyzing Big Image Data in Neutrino Experiments

Stage 2-a: Dense Pixel Clustering



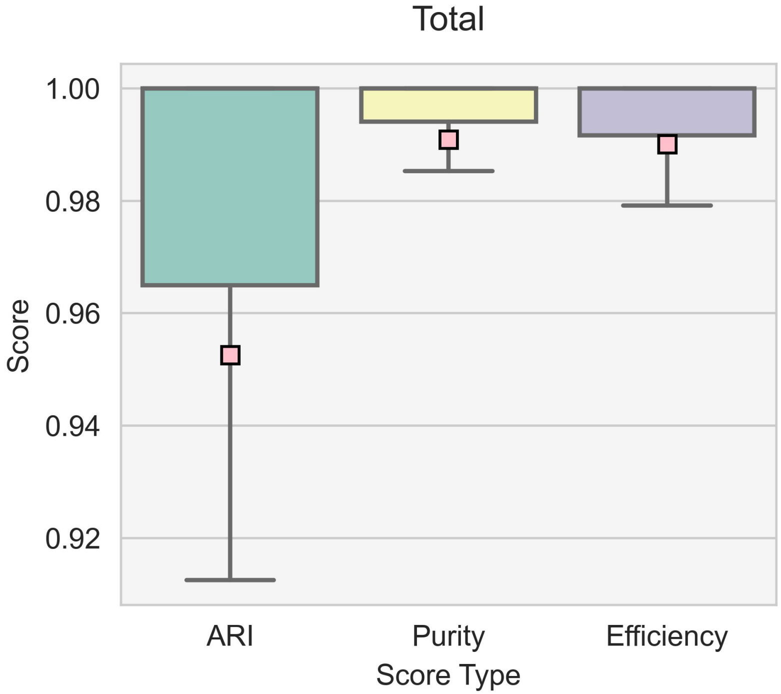
See arxiv:2007.03083

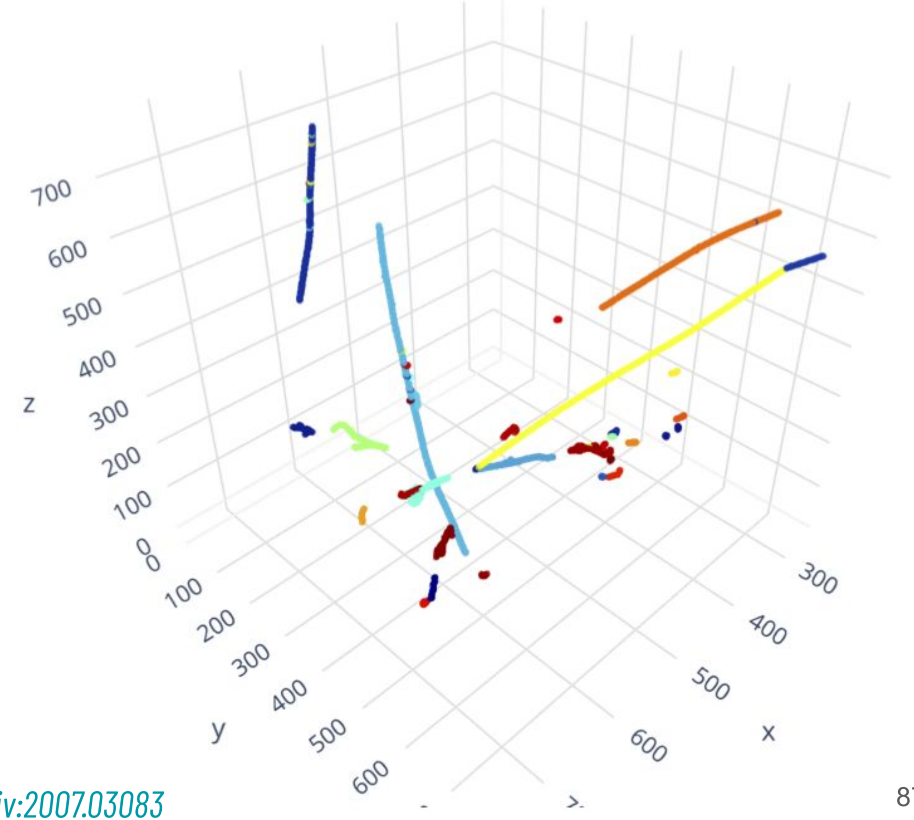
Scalable Particle Instance Clustering using Embedding (SPICE)

- Embedding decoder learns transformation
- Seediness decoder identifies the centroids
- During training, loss is conditioned so that the points that belong to the same cluster follow a normal distribution

Stage 2-a: Dense Pixel Clustering

Pixels clustered into trajectory fragments using SPICE





See arxiv:2007.03083

ML-based Neutrino Data Reconstruction Chain
Stage 2-b: Sparse Fragment Clustering

Graph-NN for Particle Aggregation (GrapPA)

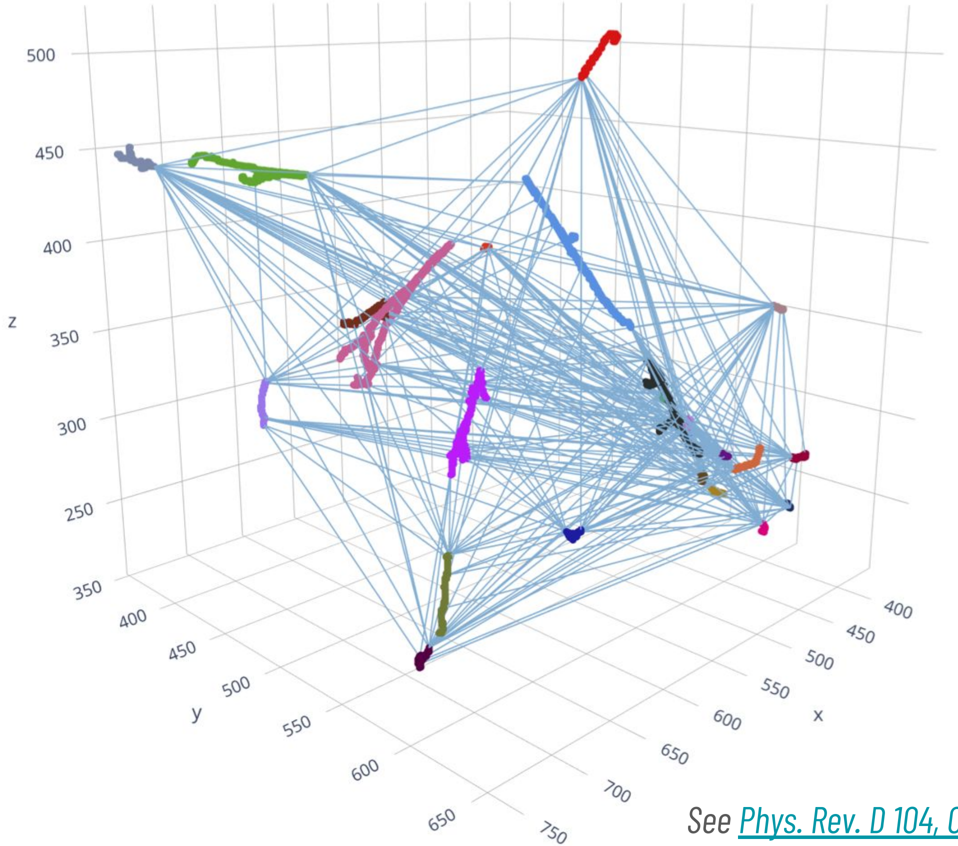
Message passing (MP):

- Meta layer (arxiv:1806.01261)
- Essentially two 3-layer MLPs (BatchNorm + LeakyReLU) for edge feature update and node feature update
- 3 times MP (=Edge+Node feature update)

Target:

- Prediction of adjacency matrix representing valid edges (=true partition)
- Apply cross-entropy loss

For more studies, see our paper

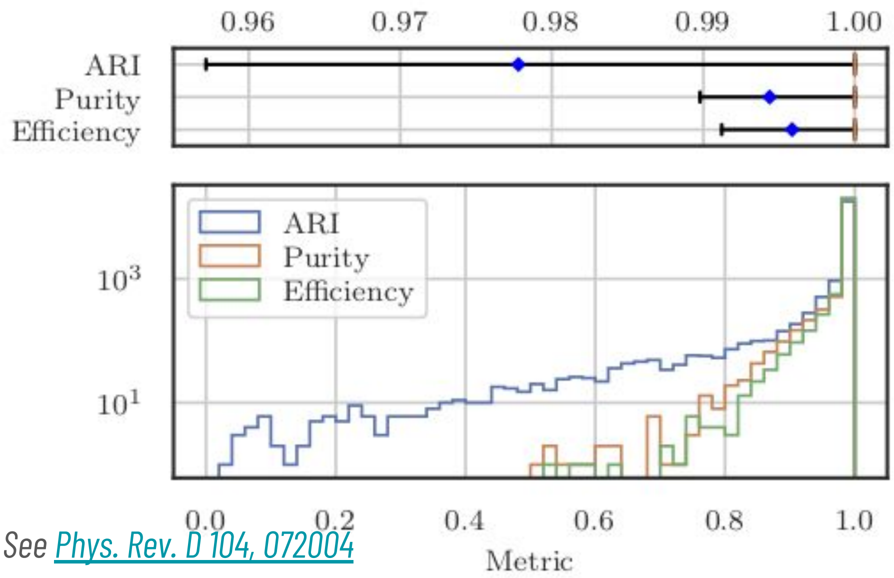


See Phys. Rev. D 104, 072004

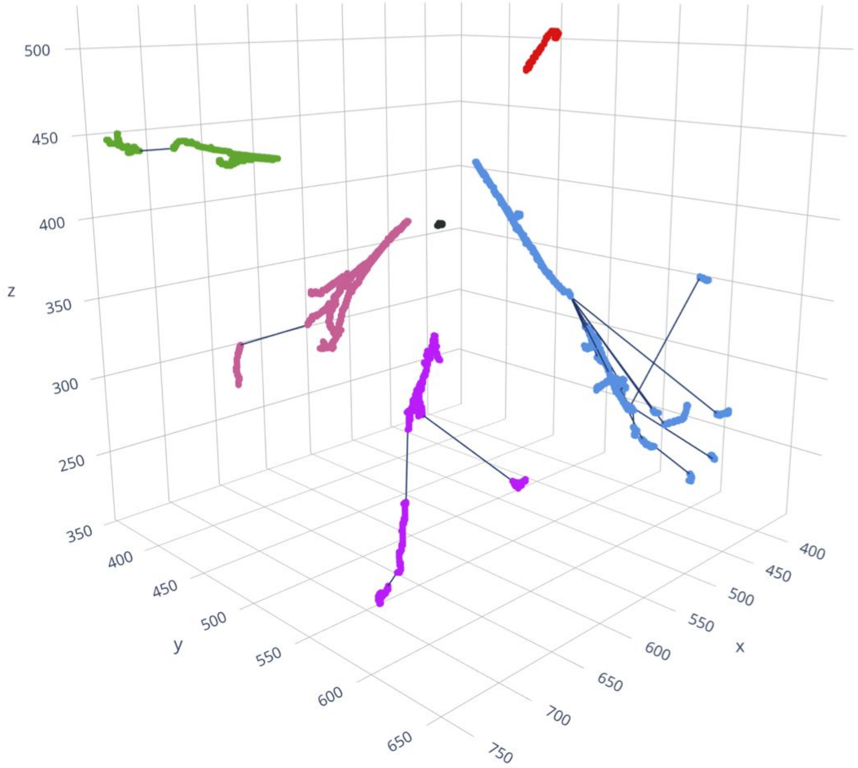
Stage 2-b: Sparse Fragment Clustering

Clustering using GrapPA

- Mean purity and efficiency > 99%
- Sufficient for moving to the next stage (particle analysis)



Edge Prediction

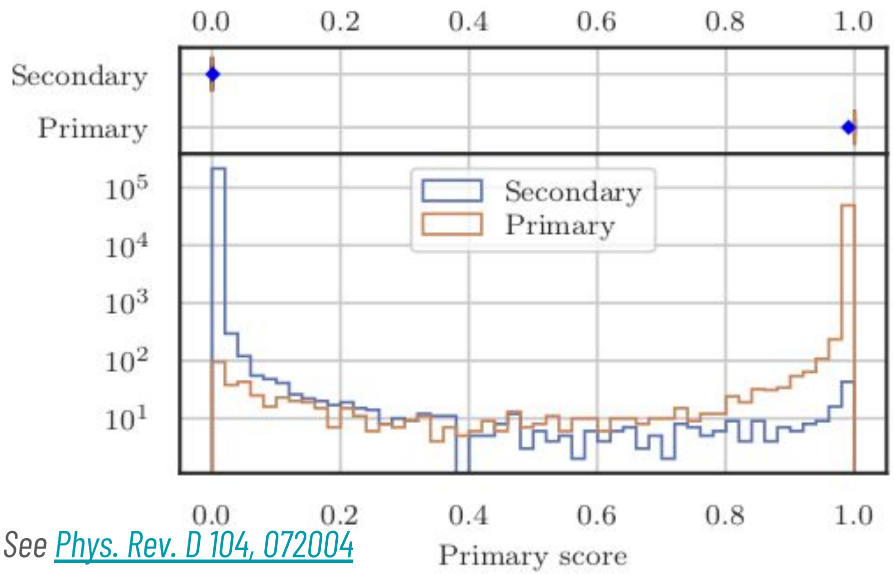


See Phys. Rev. D 104, 072004

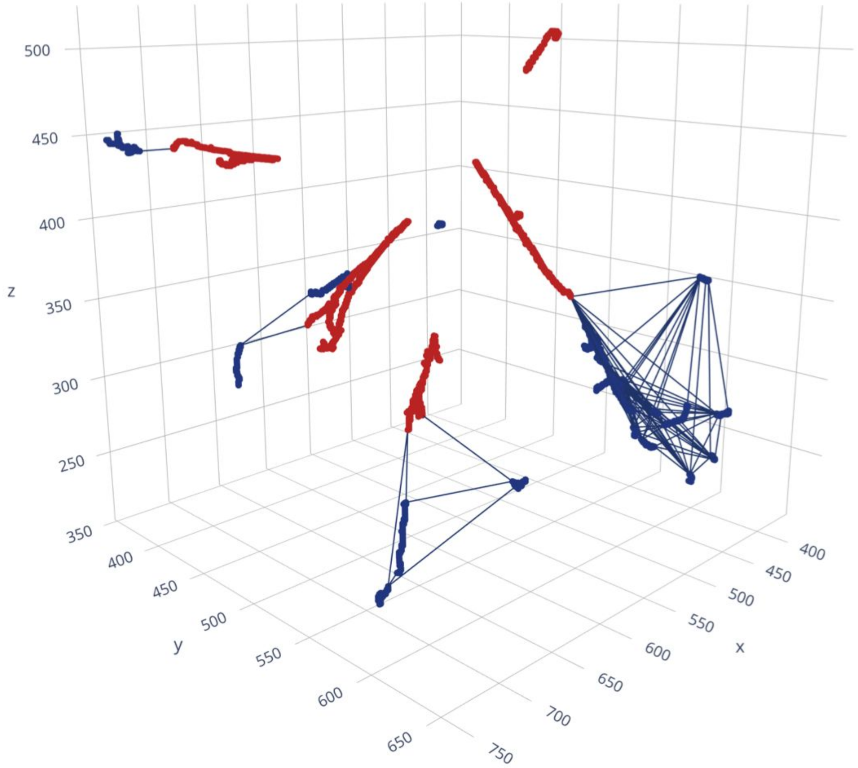
Stage 2-b: Sparse Fragment Clustering

Start ID using GrapPA

- Important to identify the "primary fragment" (=shower start)
- >99% classification accuracy



Node prediction



See Phys. Rev. D 104, 072004

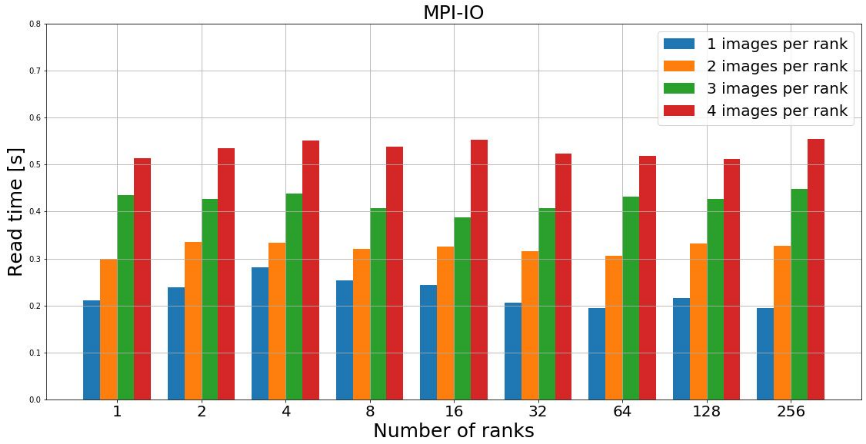
HPC Application

Wrapping up...

Inter-experimental collaborative work

- Open simulation sample
  - **Open real data?** Soon! (3D proto-type R&D @ SLAC)
- **Open software development**
  - Fast, distributed IO, optimized for sparse data

Work credit:
Corey Adams (ANL)
Marco del Tutto (FNAL)



- Custom HDF5 format for sparse data for fast IO
- Custom API for data distribution using MPI
  - Using Horovod, good scaling @ ~100 GPUs test setup (with InfiniBand interconnect)

Custom development among hobby-coders from SLAC/ANL/FNAL, lead by Corey Adams @ ANL

Collaboration

Neutrino Physics and Machine Learning Workshop

Reminder… :)

SLAC

Nu2020 Satellite (indico link) + Main Workshop (indico link)

PROJECT 8

ICECUBE
SOUTH POLE NEUTRINO OBSERVATORY

EXA TrkX

μBooNE

NOvA

T2K

WatChMaL

KamLAND

Enriched Xenon Observatory
EXO
for double beta decay

Wire-Cell

Hyper-K

DIDACTS

ν SONIC

DeepLearnPhysics
Research Collaboration

COHERENT
SNS

PROSPECT

next

ANNIE
Accelerator Neutrino Neutron Interaction Experiment

DUNE
DEEP UNDERGROUND
NEUTRINO EXPERIMENT

Image Analysis in Neutrino Physics

You can find a cat? You can find a neutrino!

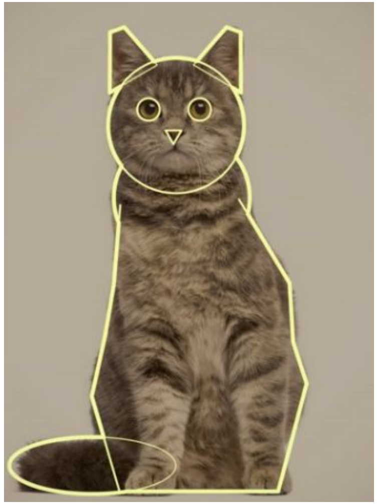
How to write an algorithm to identify a cat?

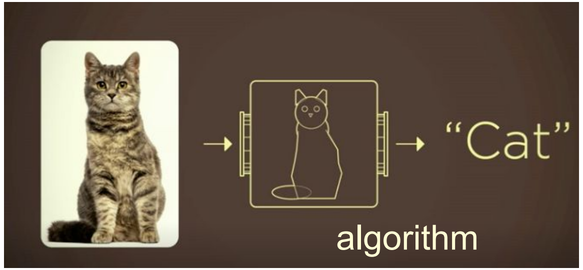
... very hard task ...

16	08	67	15	83	09
37	52	77	23	22	74
35	42	48	72	85	27
68	36	43	54	21	33
79	60	10	25	54	71
18	55	38	73	50	47

Development Workflow for non-ML reconstruction

1. Write an algorithm based on physics principles





A cat = collection of certain shapes
(or, a neutrino)

Images courtesy of Fei Fei Li's TED talk

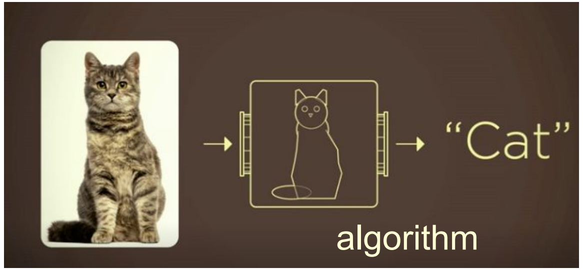
Development Workflow for non-ML reconstruction

1. Write an algorithm based on physics principles
2. Run on simulation and data samples
3. Observe failure cases, implement fixes/heuristics
4. Iterate over 2 & 3 till a satisfactory level is achieved
5. Chain multiple algorithms as one algorithm, repeat 2, 3, and 4.

Partial cat
(escaping the detector)

Images courtesy of Fei Fei Li's TED talk

Stretching cat (Nuclear Physics)



A cat = collection of certain shapes
(or, a neutrino)

Development Workflow for non-ML reconstruction

1. Write an algorithm based on physics principles
2. Run on simulation and data samples
3. Observe failure cases, implement fixes/heuristics
4. Iterate over 2 & 3 till a satisfactory level is achieved
5. Chain multiple algorithms as one algorithm, repeat 2, 3, and 4.

"Machine learning"

- Model instead of explicit programming
- Automatization of steps 2-4
- Multi-task optimization possible (step 5)

Next: what kind of ML algorithms?

Image Classifications: a lot of applications

Especially great for: **“a rare event in a quiet detector”**

- **Quiet** = can assume “almost always neutrino”
  - e.g.) no cosmic-ray background
- **Rare** = “only 1 neutrino”

Image Classifications: a lot of applications

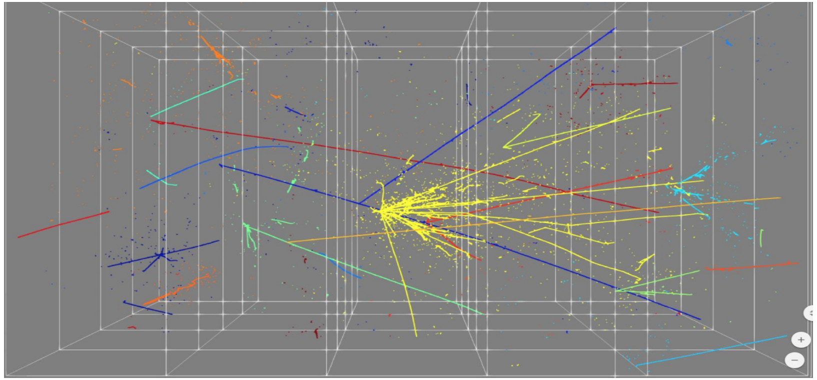
SLAC

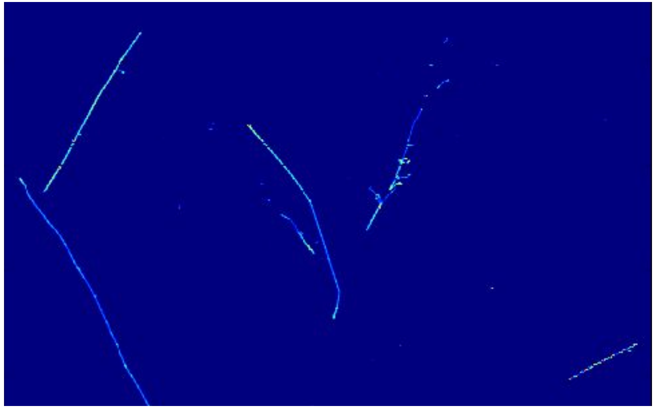
Especially great for: **"a rare event in a quiet detector"**

- **Quiet** = can assume "almost always neutrino"
  - e.g.) no cosmic-ray background
- **Rare** = "only 1 neutrino"
  - the same "image classification architecture" can be applied for…
    - neutrino flavor (topology) classification
    - energy regression (image to one FP32 value)
    - vertex regression (image to three FP32 value)
    - etc. …

Image Classifications: a lot of applications

Especially great for: **“a rare event in a quiet detector”**



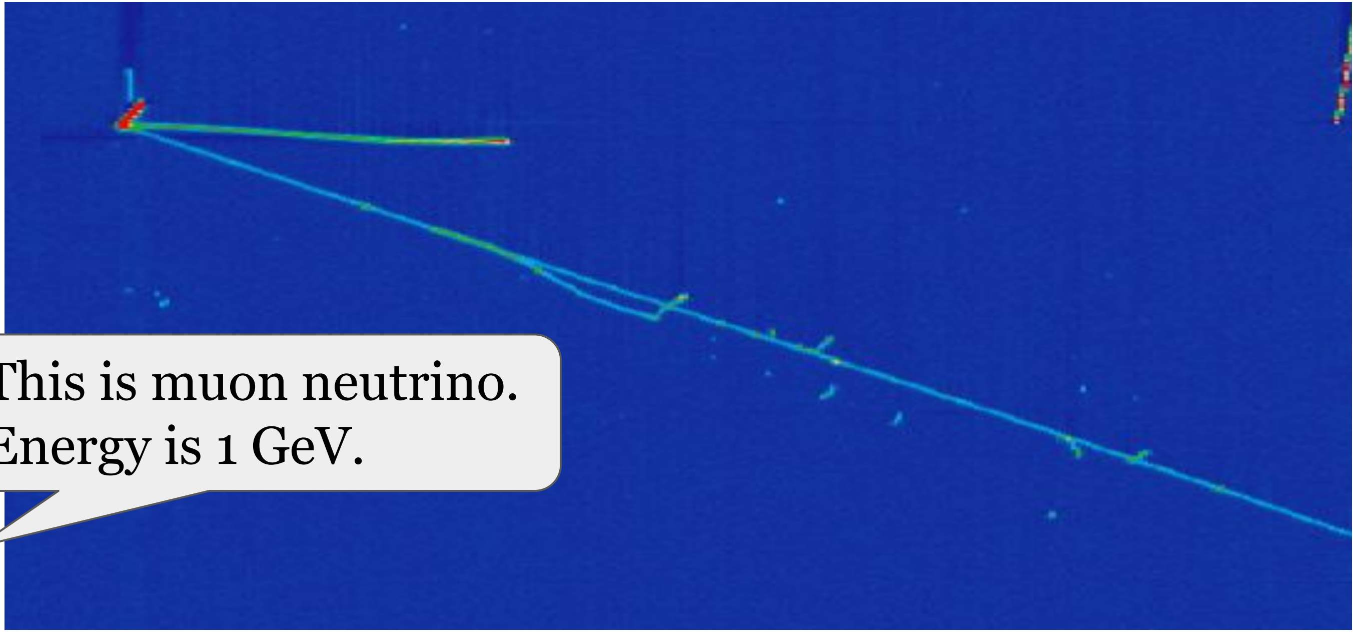


... **but most of LArTPC detectors are not** ...

- MicroBooNE, ICARUS, SBND, ProtoDUNE ... physics in next 5 years
  - Busy: typically dozens of cosmic rays in each event
- DUNE-ND
  - Not rare (busy): a dozen of neutrino interaction pile-up in each event

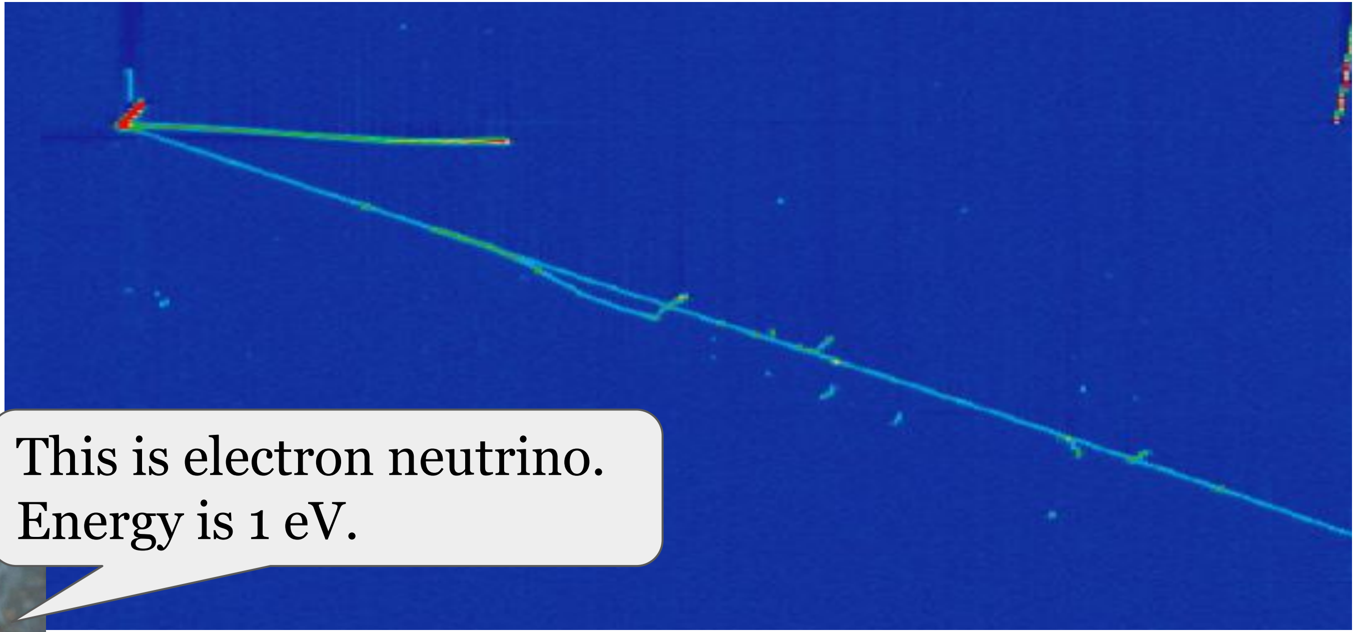
Why Data Reconstruction

Image classification/regression: straight to "**flavour & energy**"



Why Data Reconstruction

… but also challenging: a huge single-step of information reduction

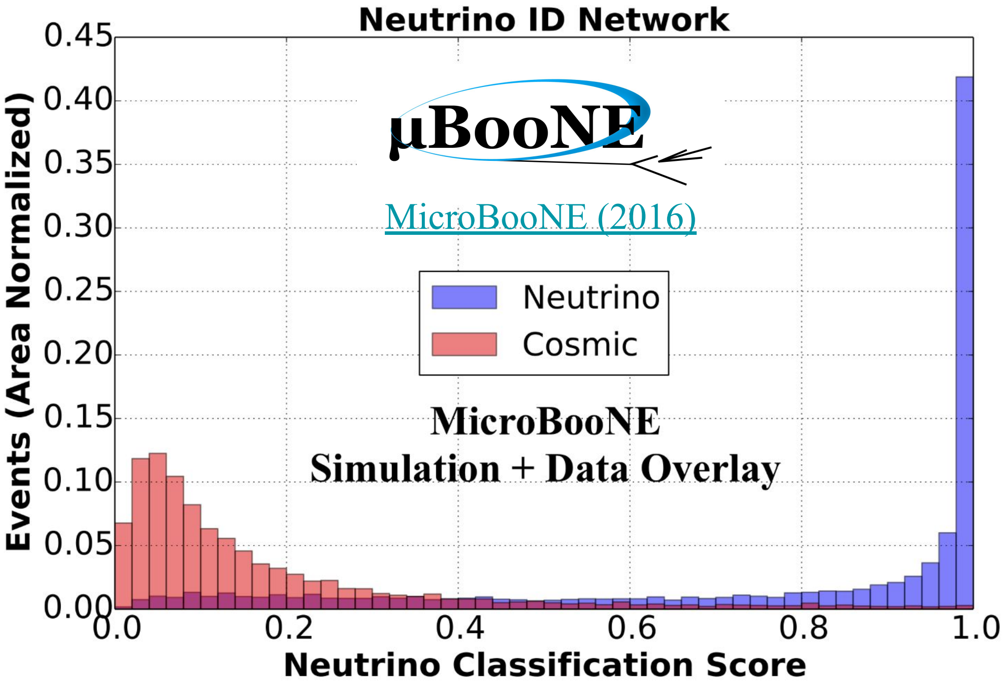


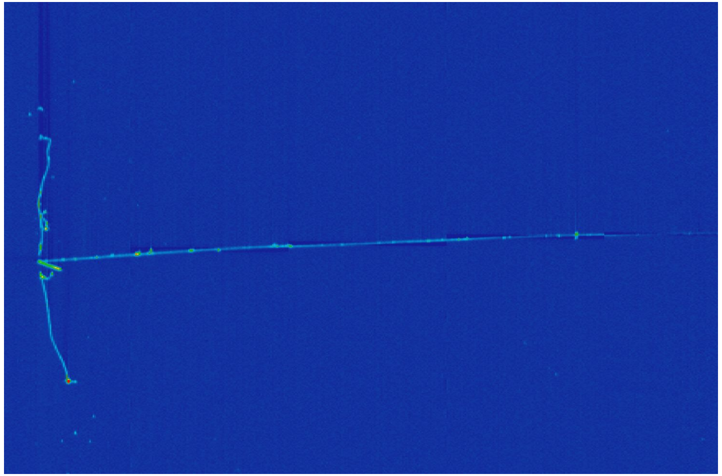
… **would be nice to know why you thought so …**

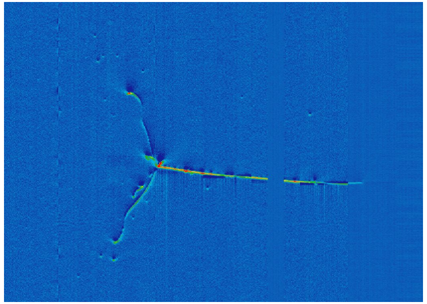
Deep Neural Network for Image Analysis

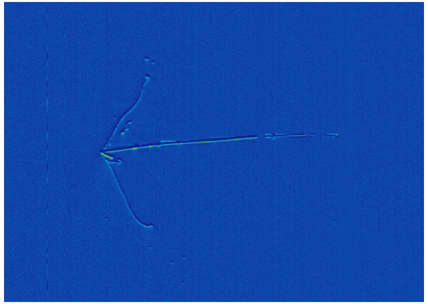
SLAC

First attempt: CNN image classifier
for signal v.s. background classification



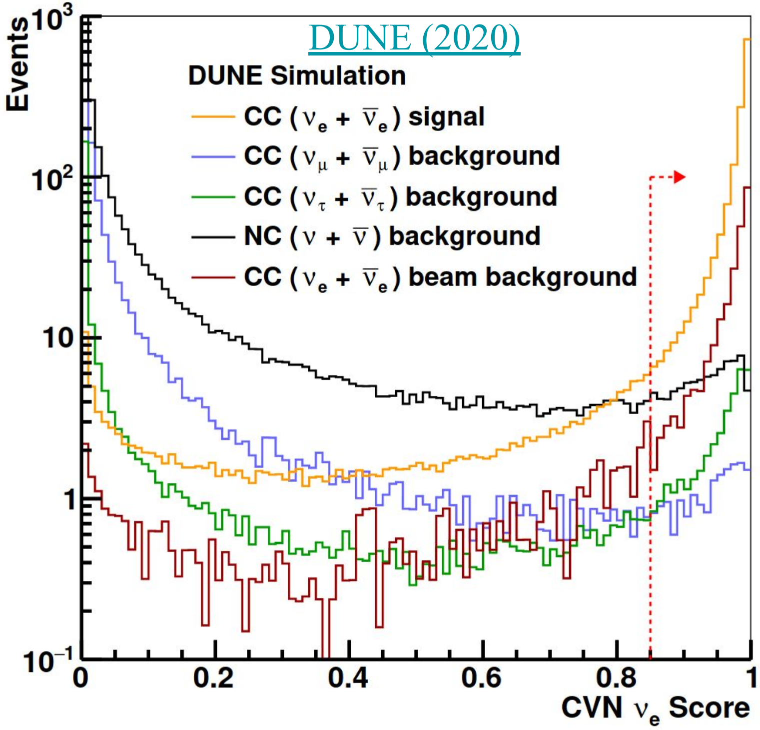


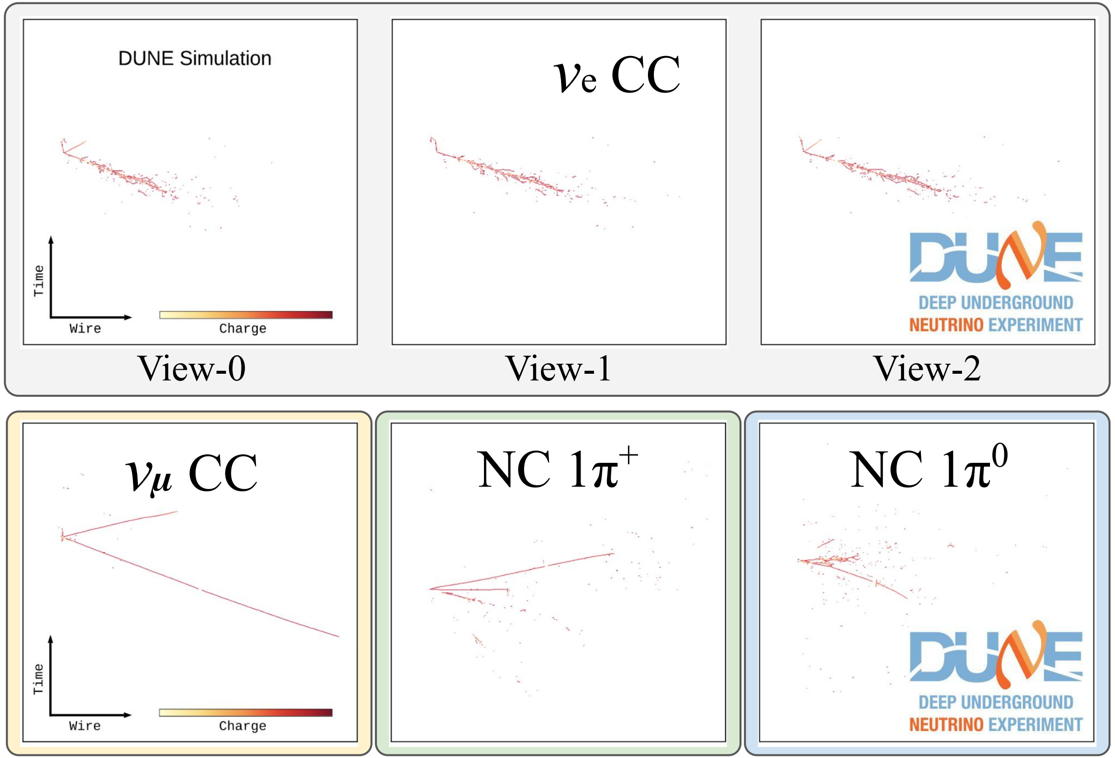




Machine Learning in Neutrino Physics & HEP
Deep Neural Network for Image Analysis

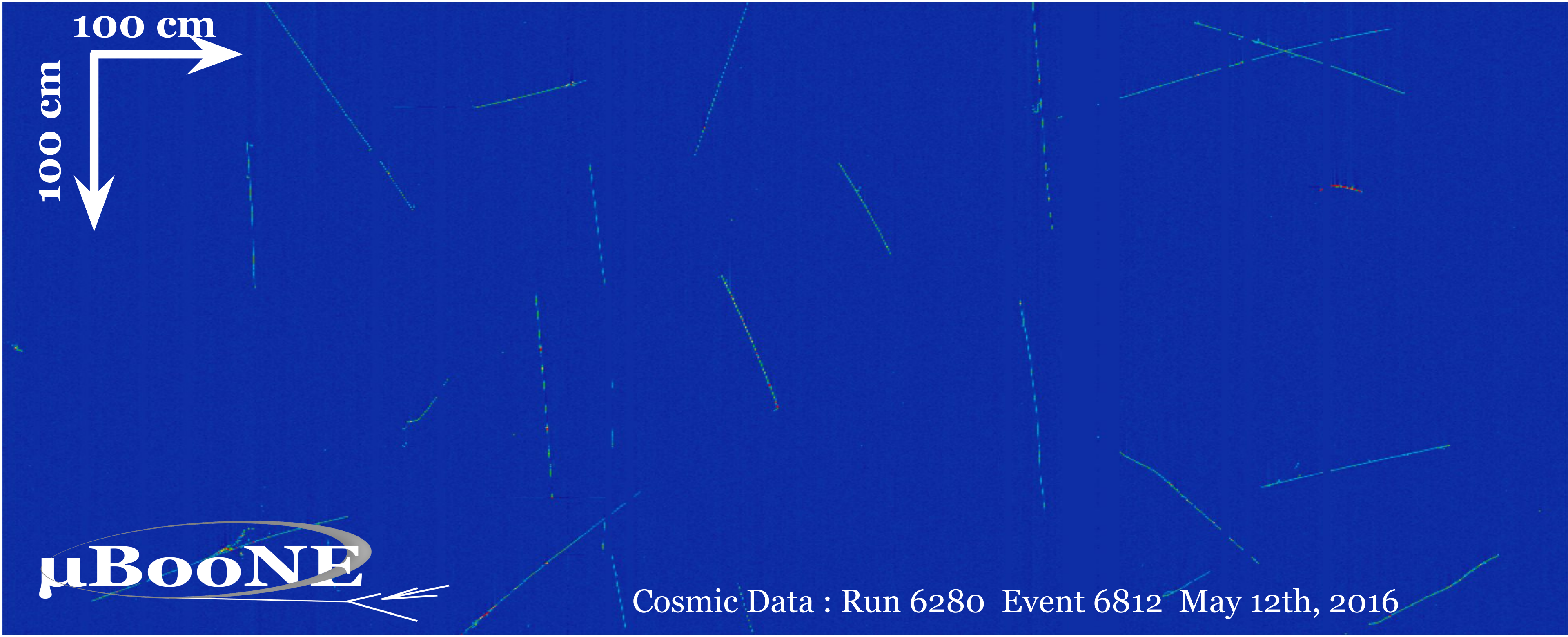
CNN image classification remains to date as a strong approach





ML for Analyzing Big Image Data in Neutrino Experiments
Challenges in particle imaging neutrino detectors

Rare Signals

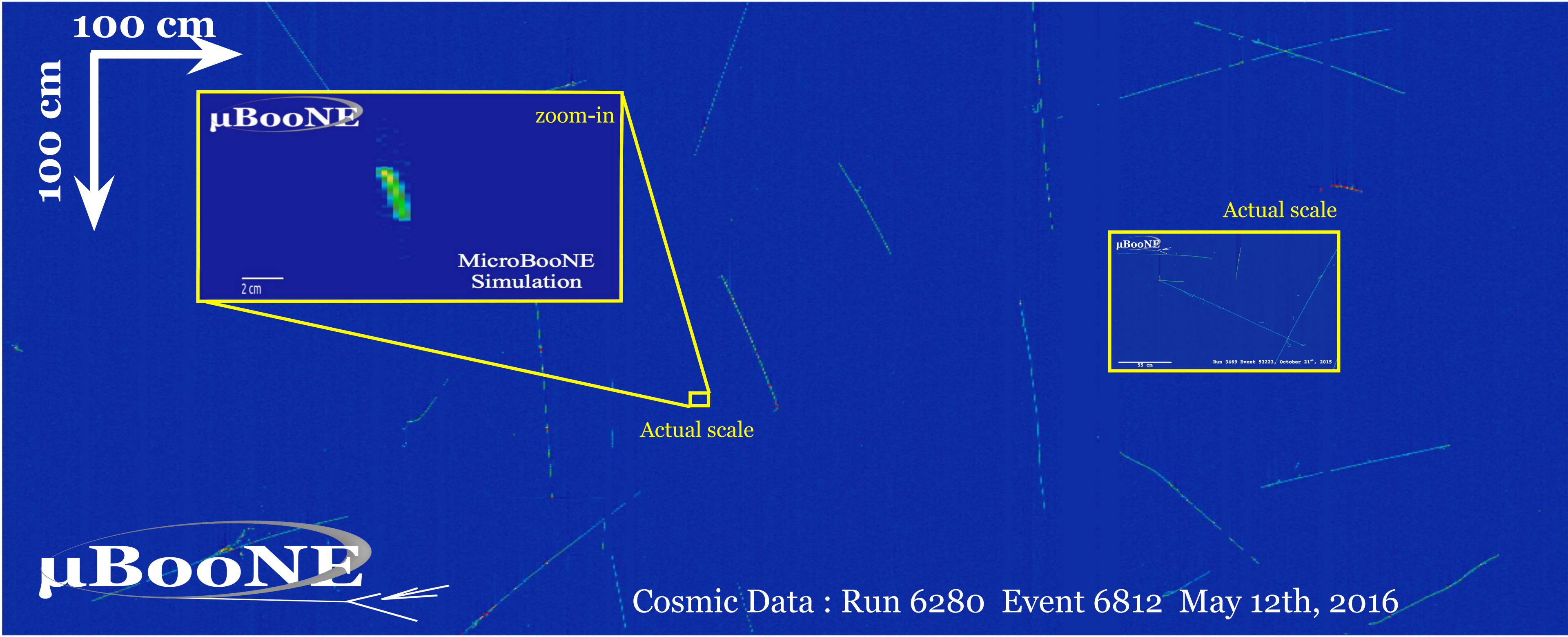


ML for Analyzing Big Image Data in Neutrino Experiments

Challenges in particle imaging neutrino detectors

SLAC

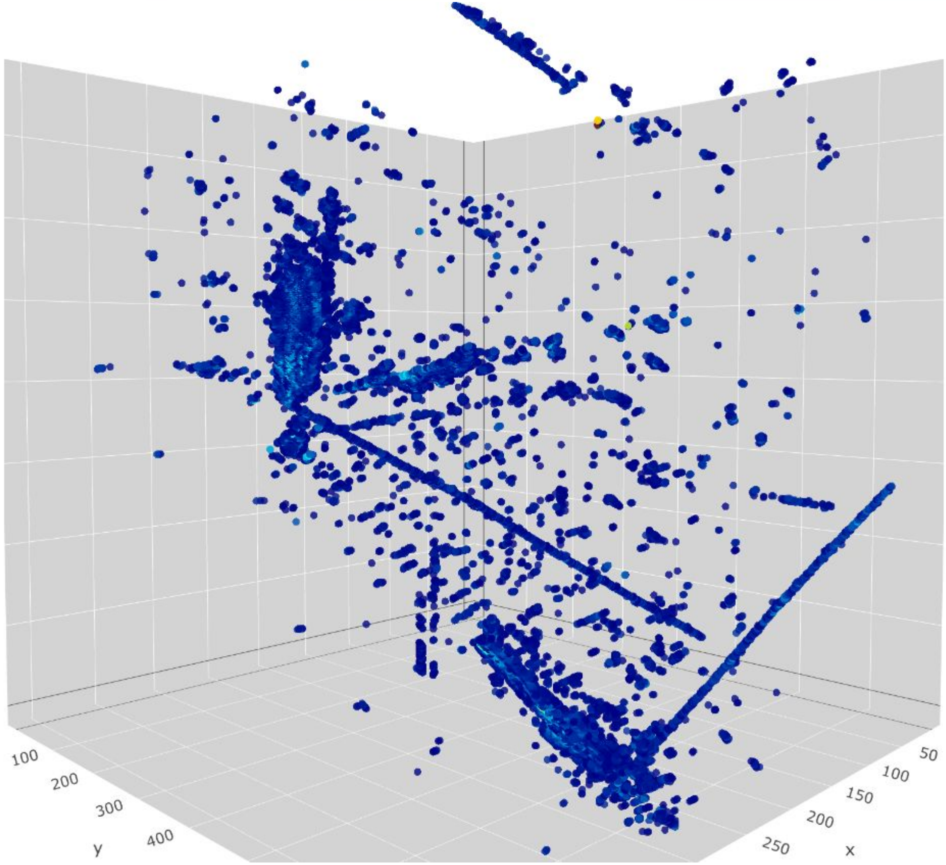
Many Backgrounds



Backup Slides

2D=>3D

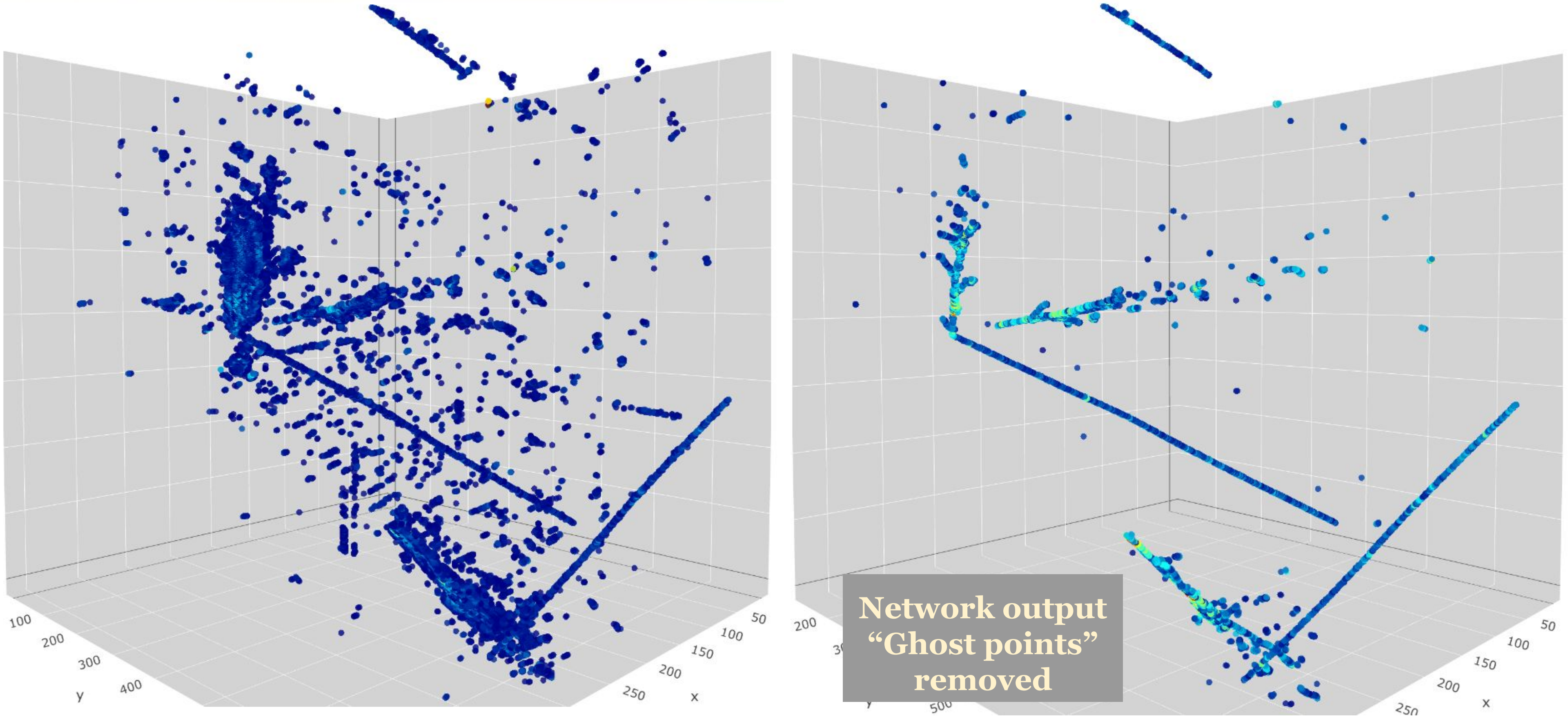
Bonus: isochronous ghost point removal



ICARUS Detector
Reconstructed 3D points

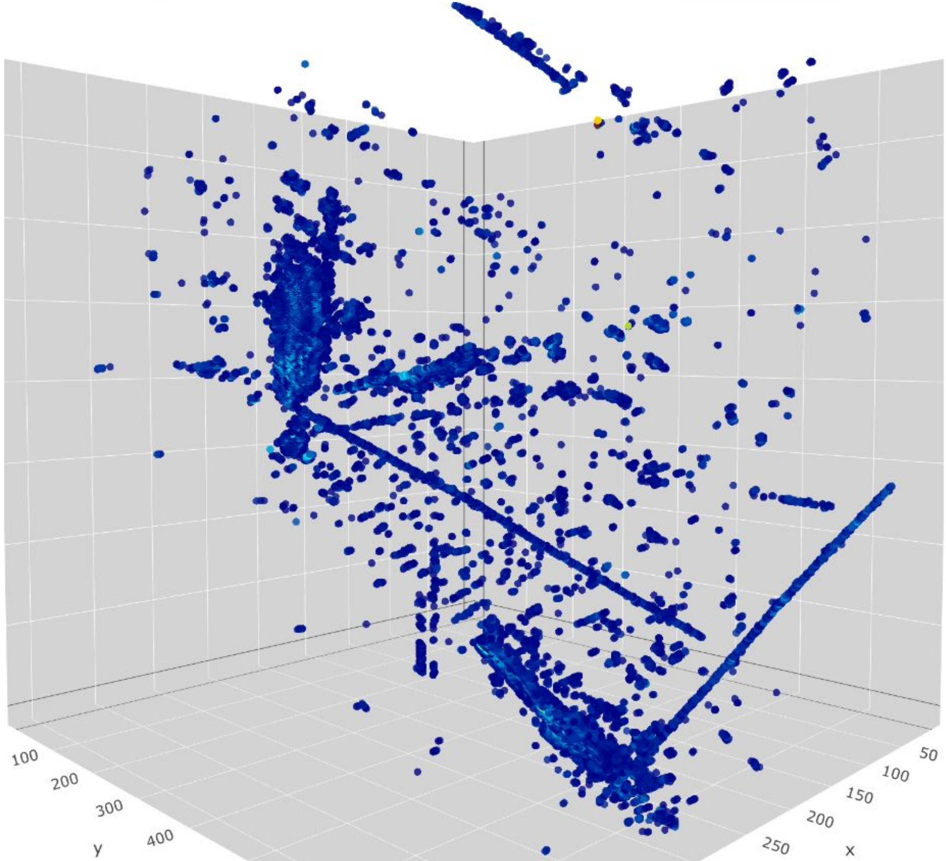
work credit:
Laura Domine
Patrick Tsang

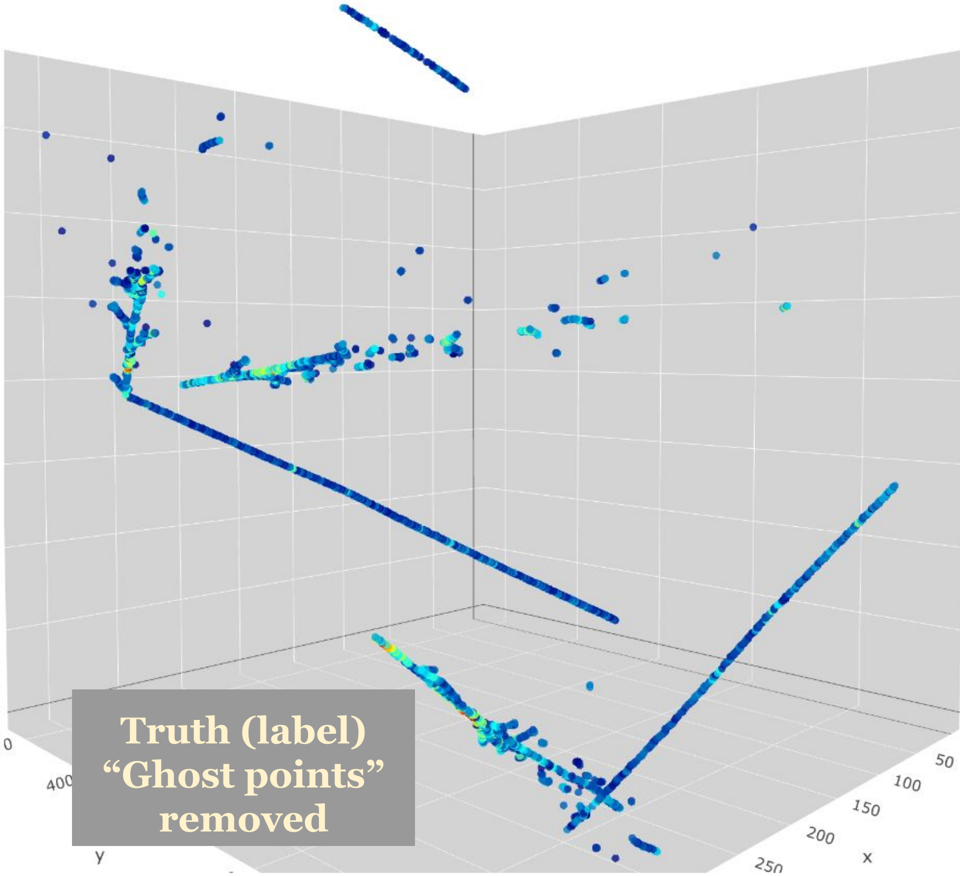
Bonus: isochronous ghost point removal



Bonus: isochronous ghost point removal

SLAC



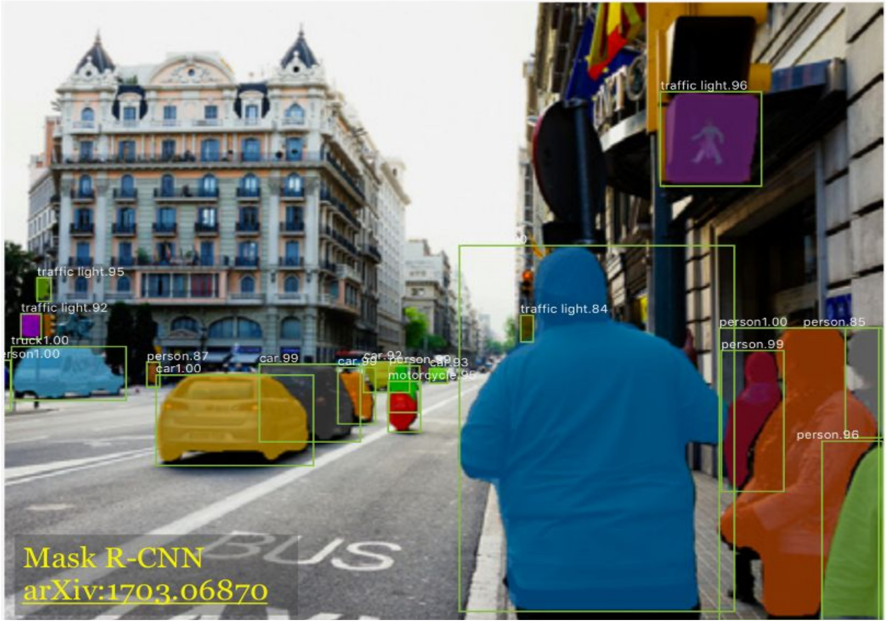


SPICE

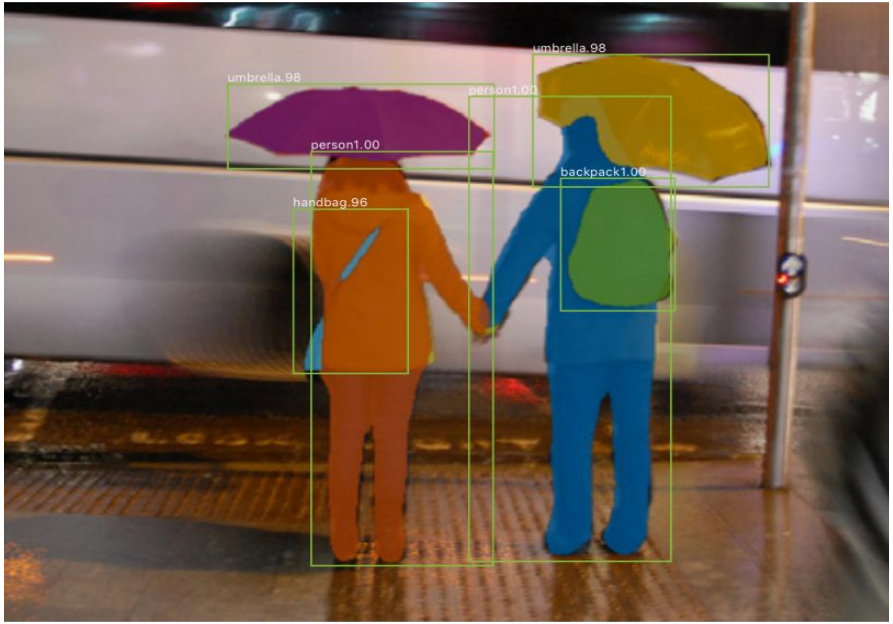
SLAC

Instance+Semantic Segmentation

- **Mask R-CNN** … a popular solution, many applications in science/industries
  - Object (=instance) detection + instance pixel masking in a bounding box

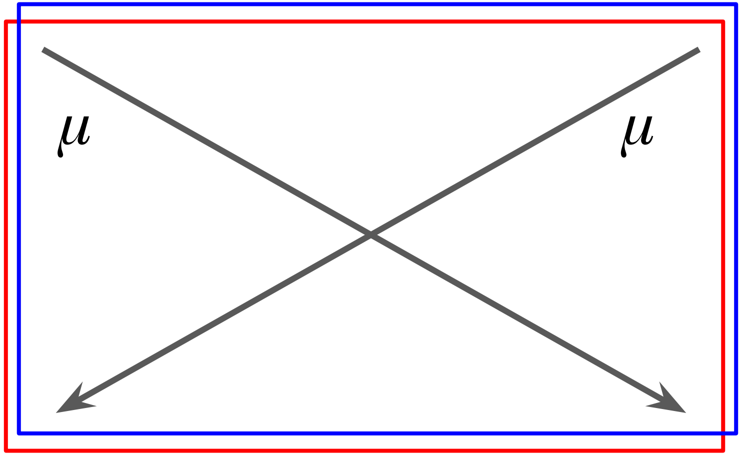


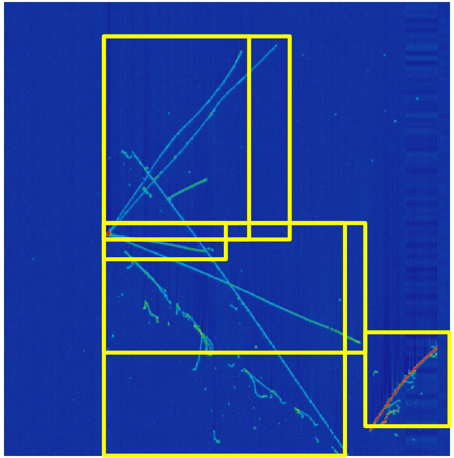
Mask R-CNN
arXiv:1703.06870



Instance+Semantic Segmentation

- **Mask R-CNN** … a popular solution, many applications in science/industries
  - Object (=instance) detection + instance pixel masking in a bounding box
  - **Issue**: instance distinction is affected by BB position/size
  - Another family: Single-Shot-Detection (SSD) based (not covered here)





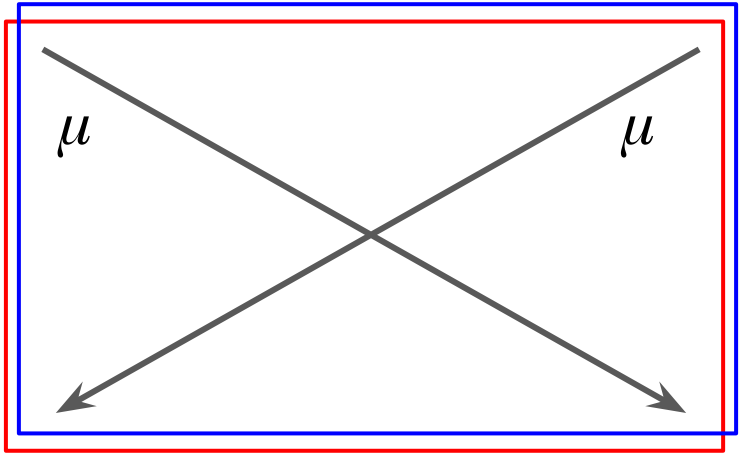
Occlusion issue

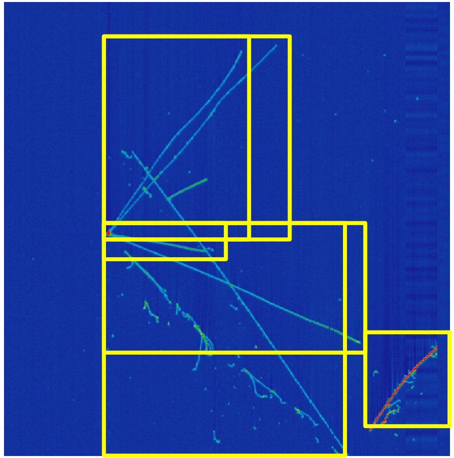
The overlap rate of particles is very high especially for our signal (neutrinos) with an event vertex.

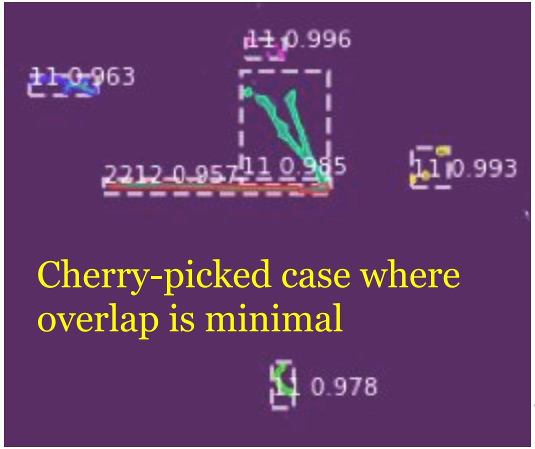
ML-based Neutrino Data Reconstruction Chain
Stage 2: Particle & Interaction Clustering

Instance+Semantic Segmentation

- **Mask R-CNN** … a popular solution, many applications in science/industries
  - Object (=instance) detection + instance pixel masking in a bounding box
  - **Issue**: instance distinction is affected by BB position/size
  - Another family: Single-Shot-Detection (SSD) based (not covered here)







Instance+Semantic Segmentation

- Three component loss: pull together points that belong to the same cluster, keep distance between clusters, and regularization

$$L = \alpha L_{var} + \beta L_{dist} + \gamma L_{reg},$$

$$L_{var} = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{N_c}\sum_{i=1}^{N_c}\left[\max\left(0, \|\mu_c - x_i\| - \delta_v\right)\right]^2$$

$$L_{dist} = \frac{1}{C(C-1)}\sum_{\substack{c_A, c_B = 1\\ c_A \neq c_B}}^{C}\left[\max\left(0, 2\delta_d - \|\mu_{c_A} - \mu_{c_B}\|\right)\right]^2$$

$$L_{reg} = \frac{1}{C}\sum_{c=1}^{C}\|\mu_c\|$$

Equation credit: Dae Hyun K. @ Stanford

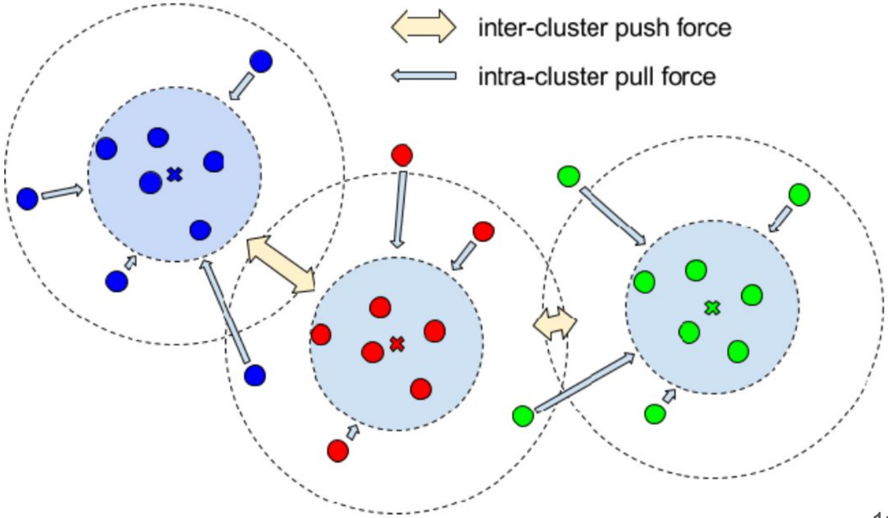


Image credit: arXiv 1708.02551

Why Data Reconstruction

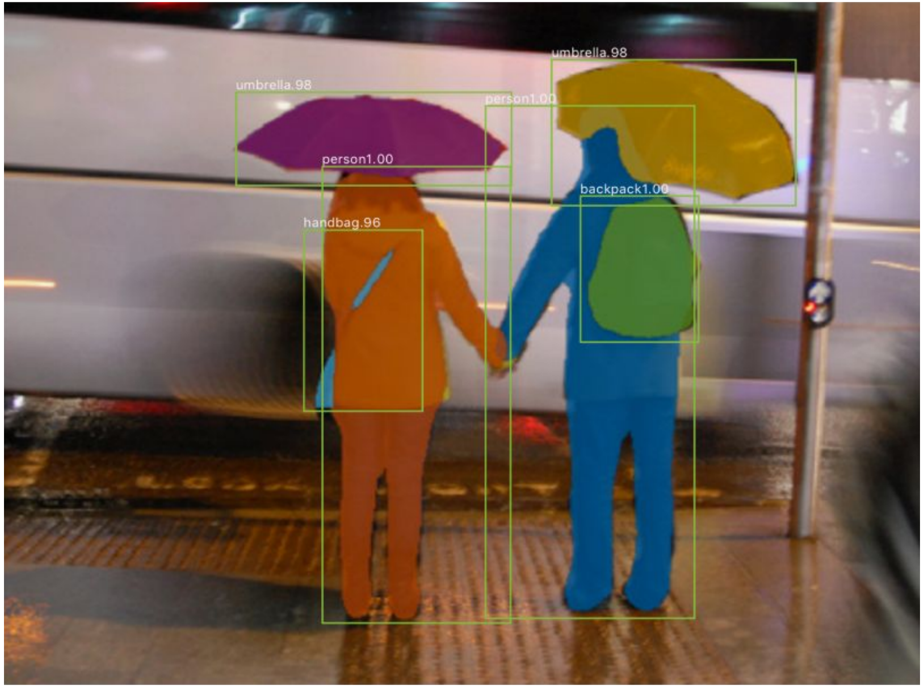


Image Context Identification

Why Data Reconstruction

"girl in pink dress is jumping in air."

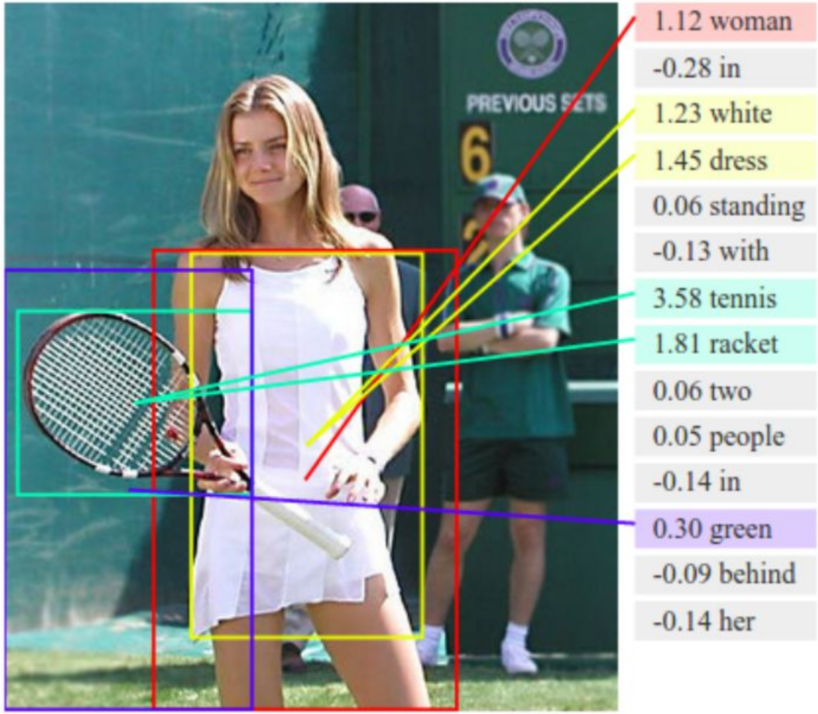
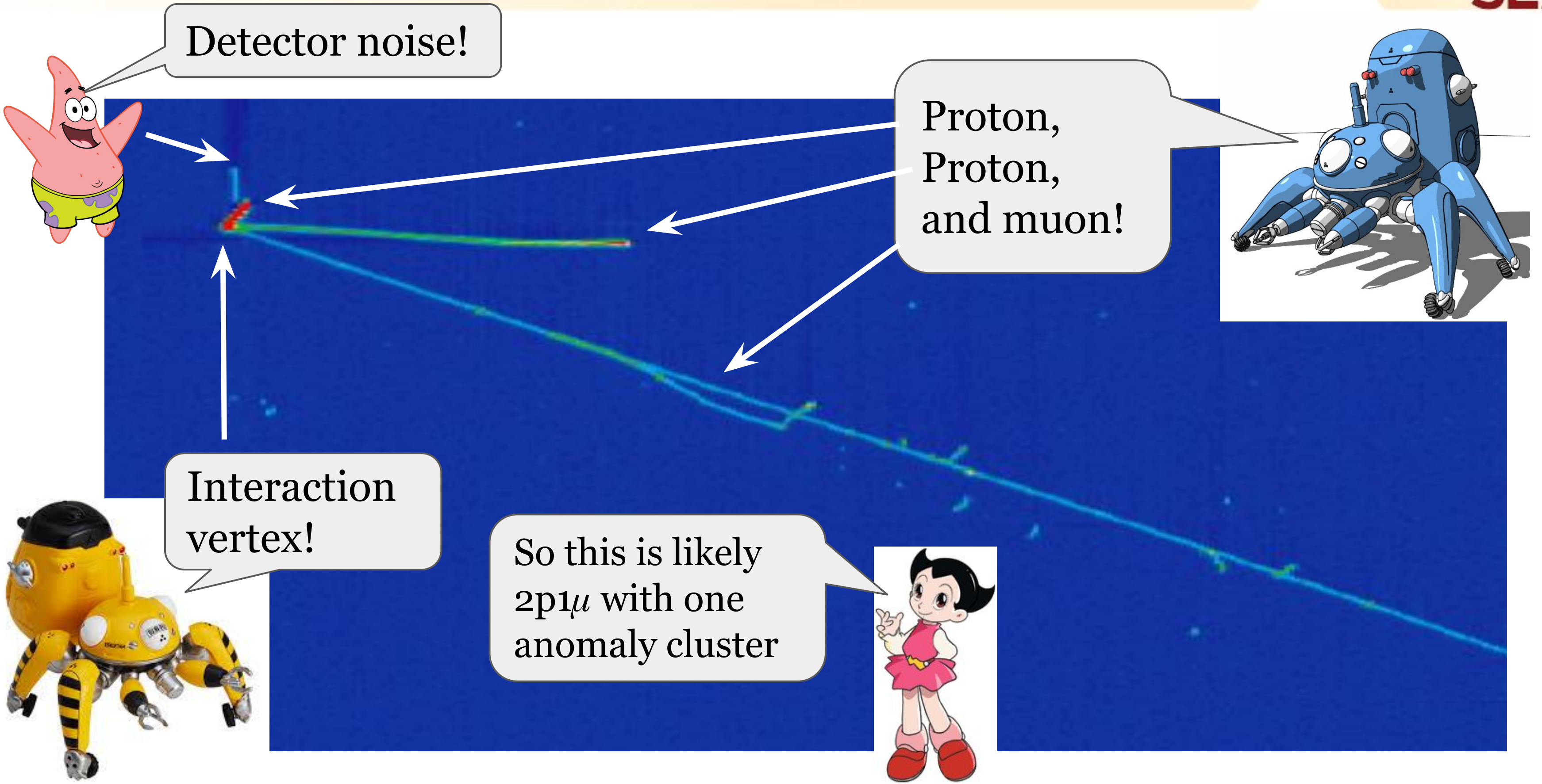


Image Context Correlation/Hierarchy Analysis

Machine Learning & Computer Vision in Neutrino Physics

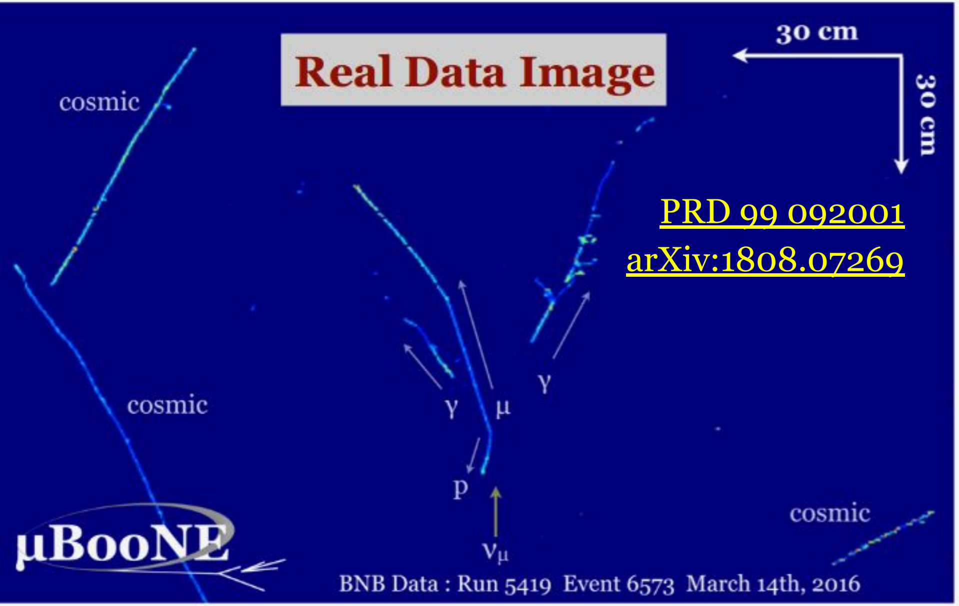
Why Data Reconstruction



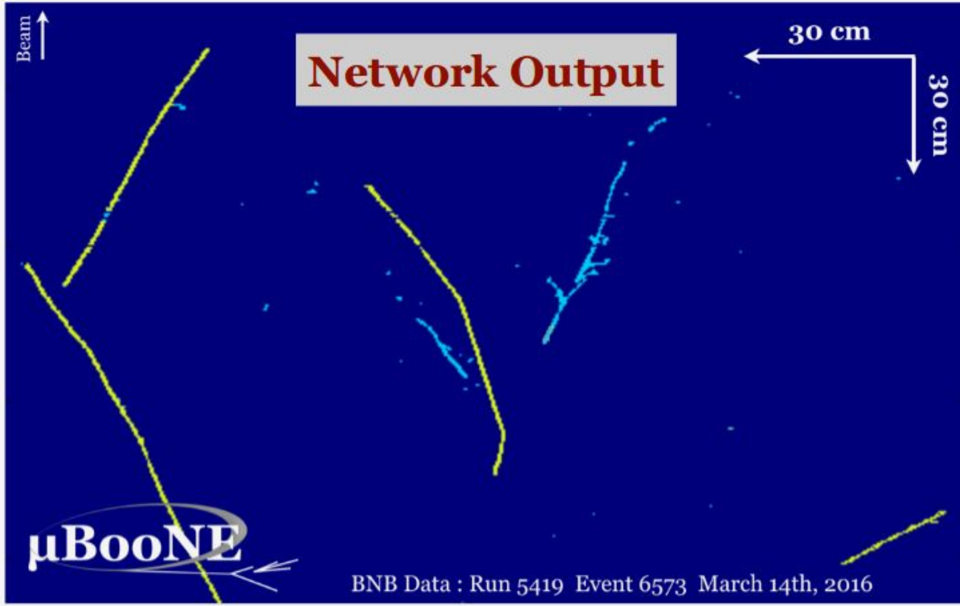
Segmentation Data

Machine Learning & Computer Vision in Neutrino Physics
Semantic Segmentation for Pixel-level Particle ID

Separate electron/positron energy depositions from other types at raw waveform level.
Helps the downstream clustering algorithms (**data/sim comp. @ arxiv:1808.07269**)



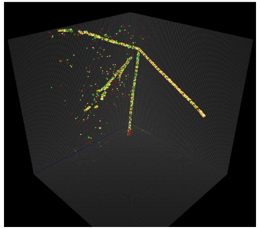
PRD 99 092001
arXiv:1808.07269

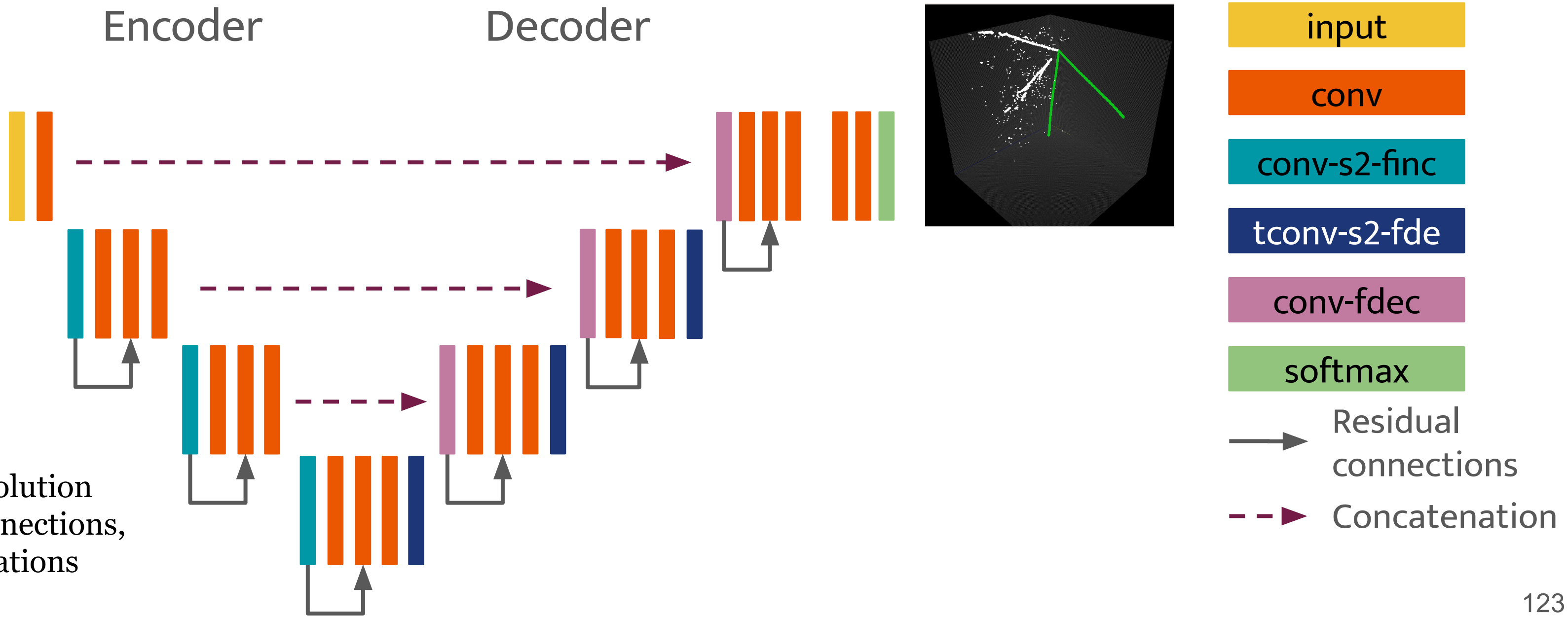


Network Input **Network Output**

Semantic Segmentation for Pixel-level Particle ID

Architecture: **U-Net + Residual Connections**

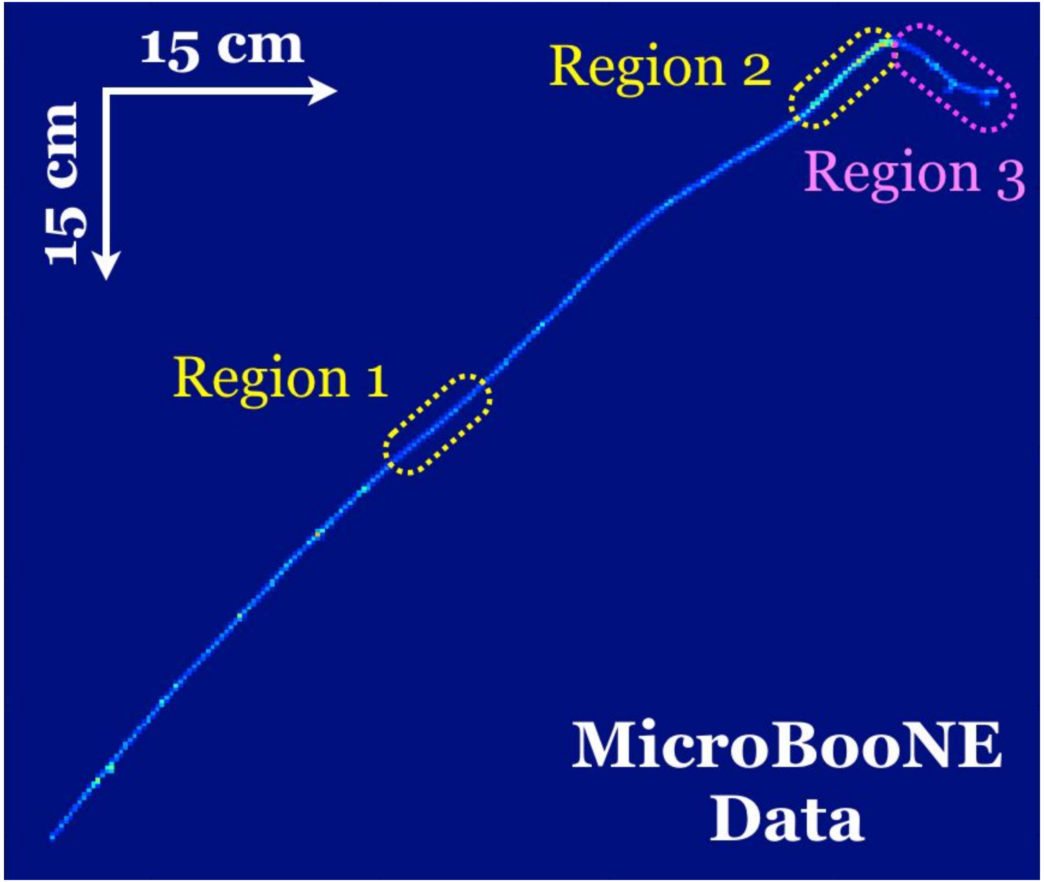




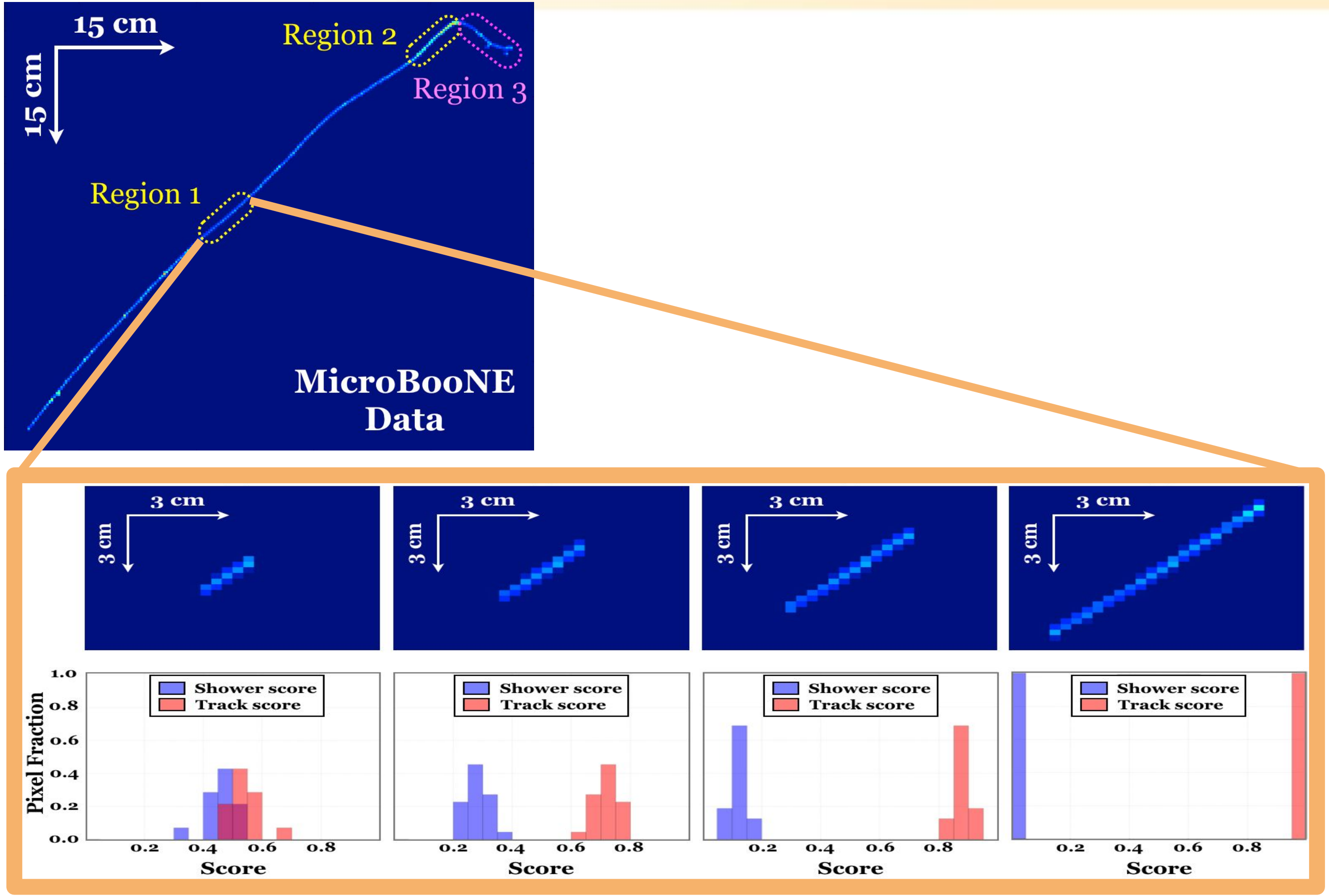
Number of strided convolutions, convolution layers, residual connections, differ in impementations

Image credit: Laura Domine @ Stanford

Machine Learning & Computer Vision in Neutrino Physics
Fun Playing with Semantic Segmentation



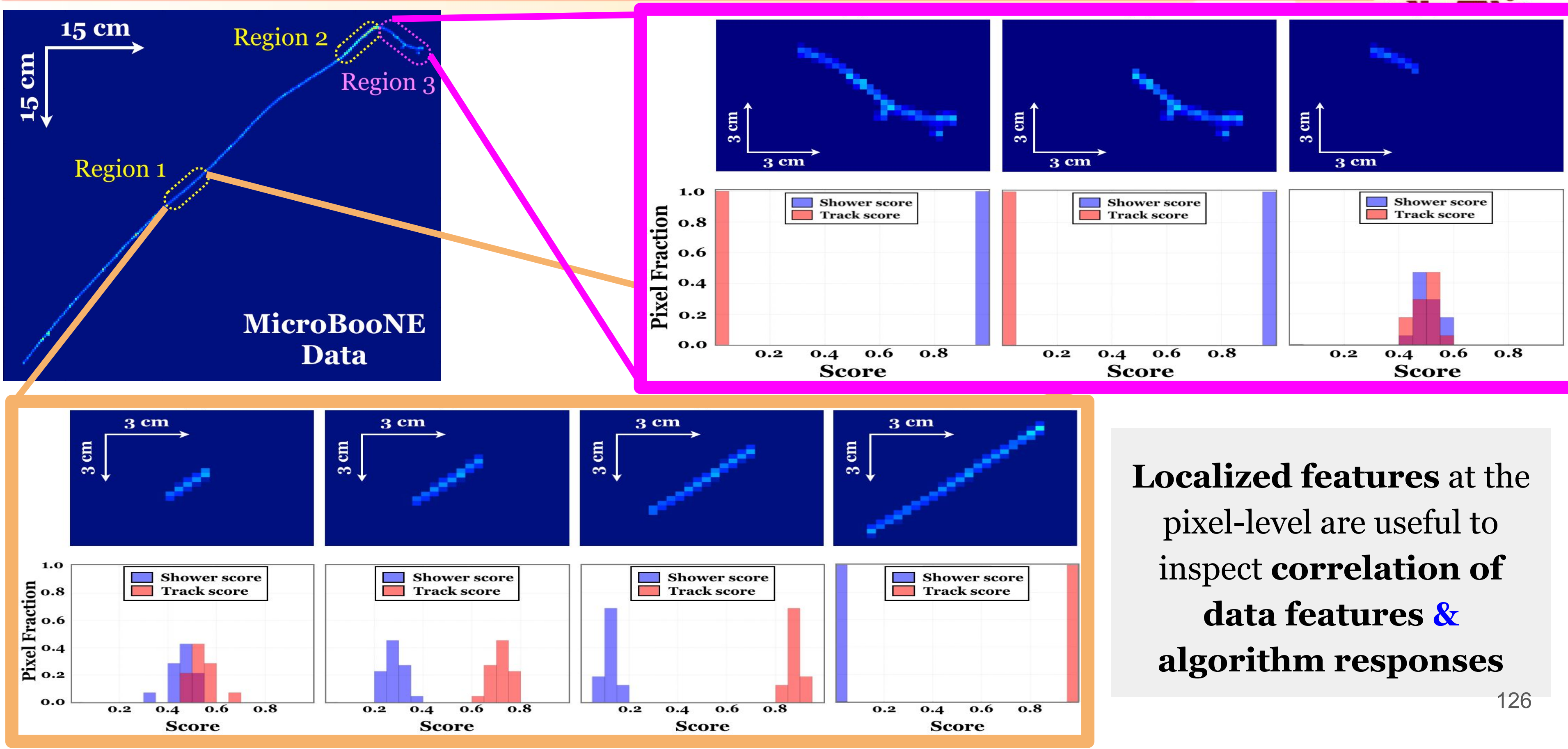
Fun Playing with Semantic Segmentation



Localized features at the pixel-level are useful to inspect **correlation of data features & algorithm responses**

Machine Learning & Computer Vision in Neutrino Physics

Fun Playing with Semantic Segmentation



Localized features at the pixel-level are useful to inspect **correlation of data features & algorithm responses**

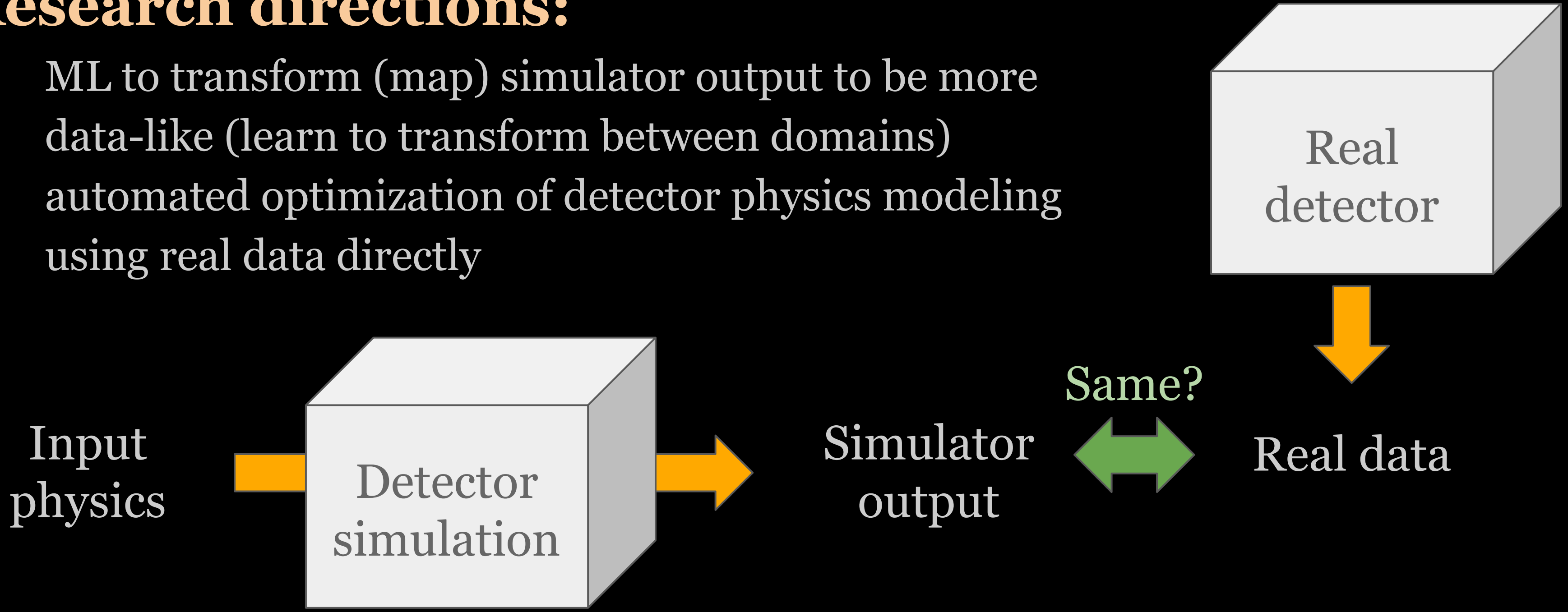
Misc. Slides

Physics model tuning

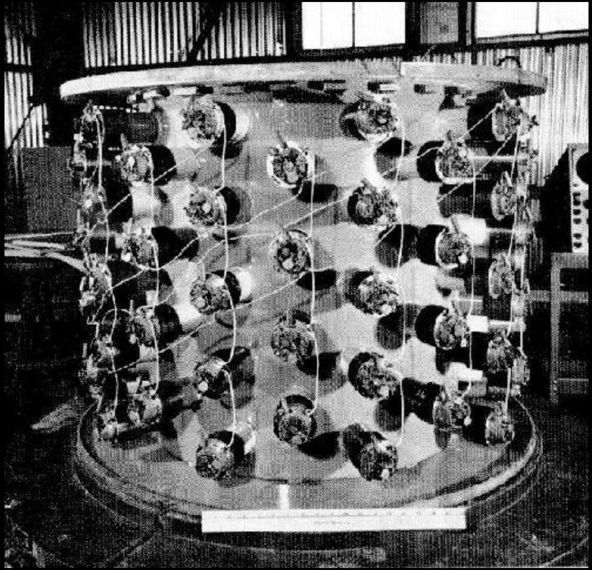
SLAC

Research directions:

- ML to transform (map) simulator output to be more data-like (learn to transform between domains)
- automated optimization of detector physics modeling using real data directly



Why neutrinos?



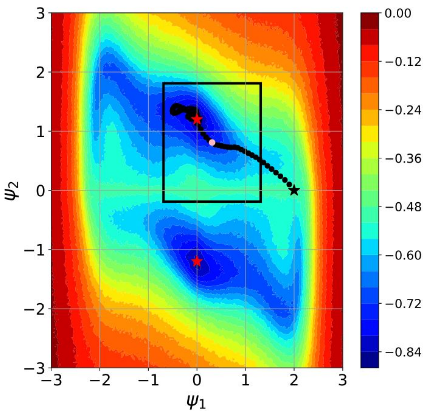
Cd-doped water
0.4 ton, 100 PMTs
(1953)

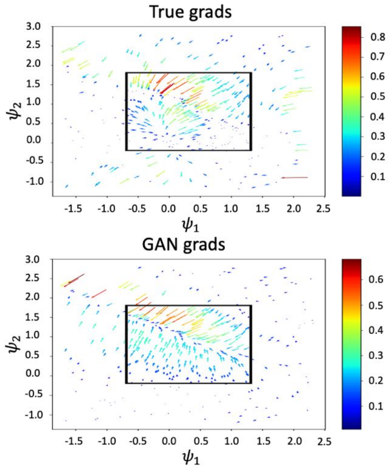
Inverse Beta Decay (IBD)
$\bar{\nu}_e + p \rightarrow e^+ + n$
from a nuclear reactor
(Reines & Cowan)

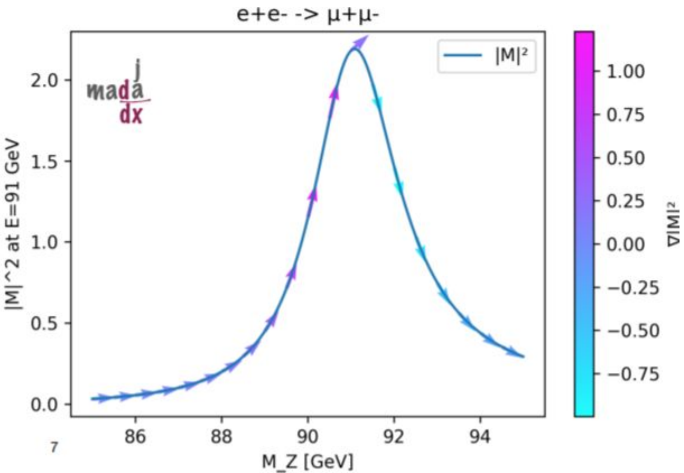
Machine Learning in Neutrino Physics & HEP
Next Step: Innovative Simulator

E.g. Differentiable Simulator

- Exploit model derivatives to enable new inference techniques
  - Surrogate (neural network) model to approximate gradients
  - Exact gradient using differentiable programming (ML) frameworks
- Applications: physics inference, design optimization, decision control, etc.







Left: surrogate model for magnet optimization

Right: differentiable matrix element calculation (MadJax)