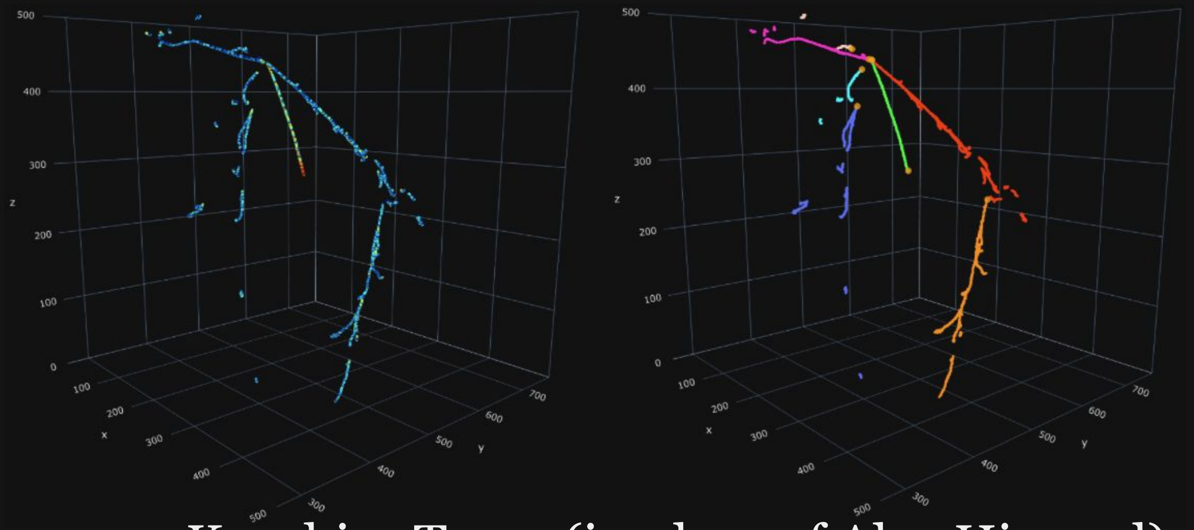
Machine Learning in Particle Physics

Original image credit: xkcd

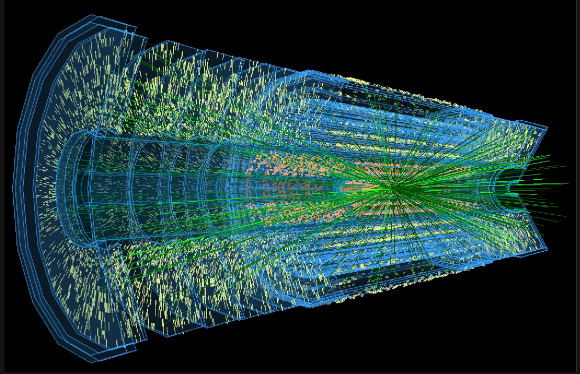


Kazuhiro Terao (in place of Alex Himmel)
SLAC National Accelerator Laboratory

COMHEP 2021

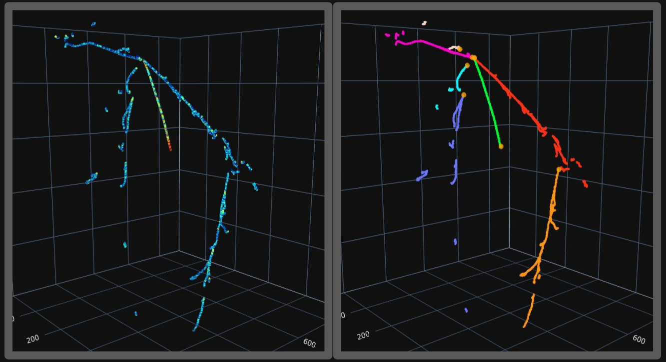
Energy Frontier

- Multi-modal detector
- Good reconstruction, big-data analysis



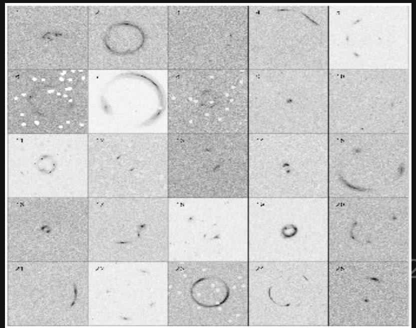
Intensity Frontier

- Big monolithic detector for imaging
- Challenging reconstruction



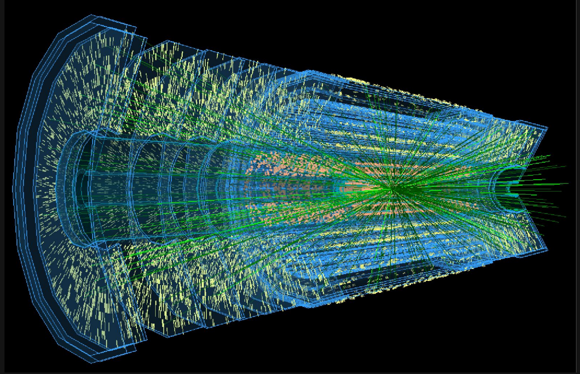
Cosmic Frontier

- Many low-resolution images
- Need for complex statistical inference



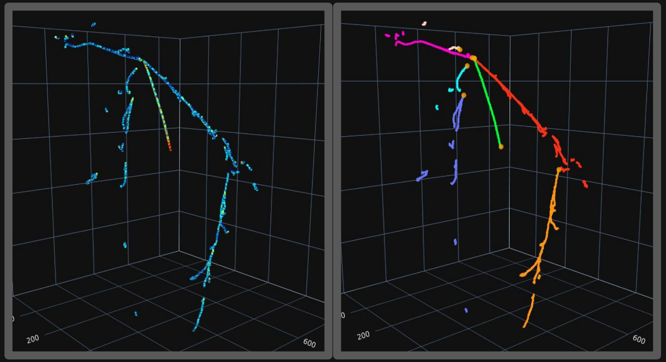
Energy Frontier

- Multi-modal detector
- Good reconstruction, big-data analysis



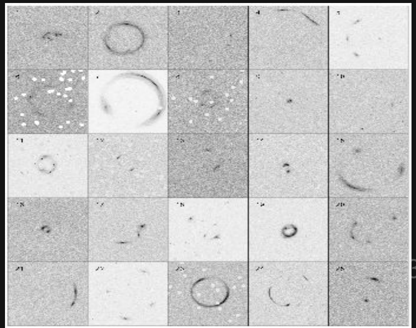
Intensity Frontier

- Big monolithic detector for imaging
- Challenging reconstruction

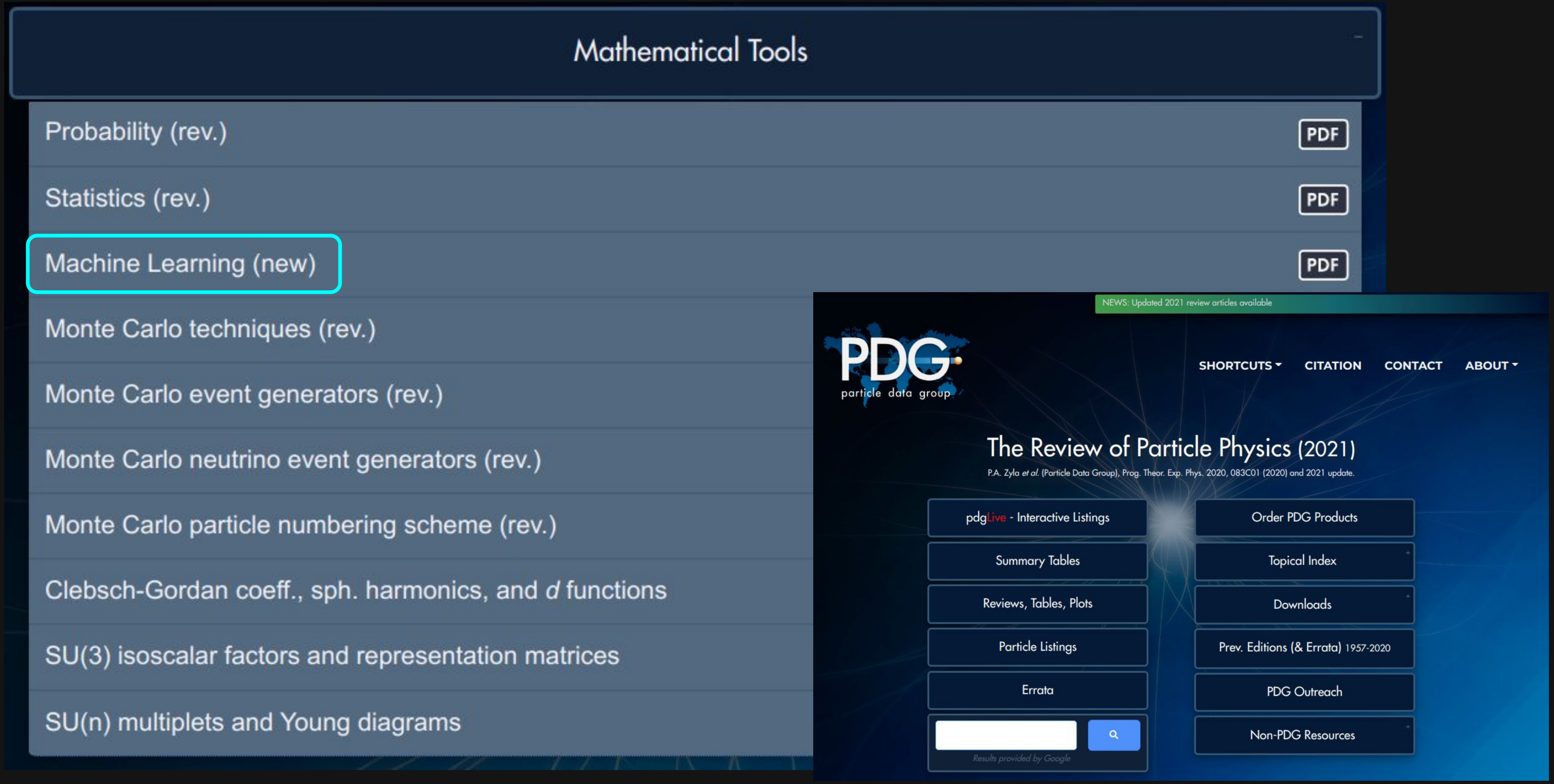


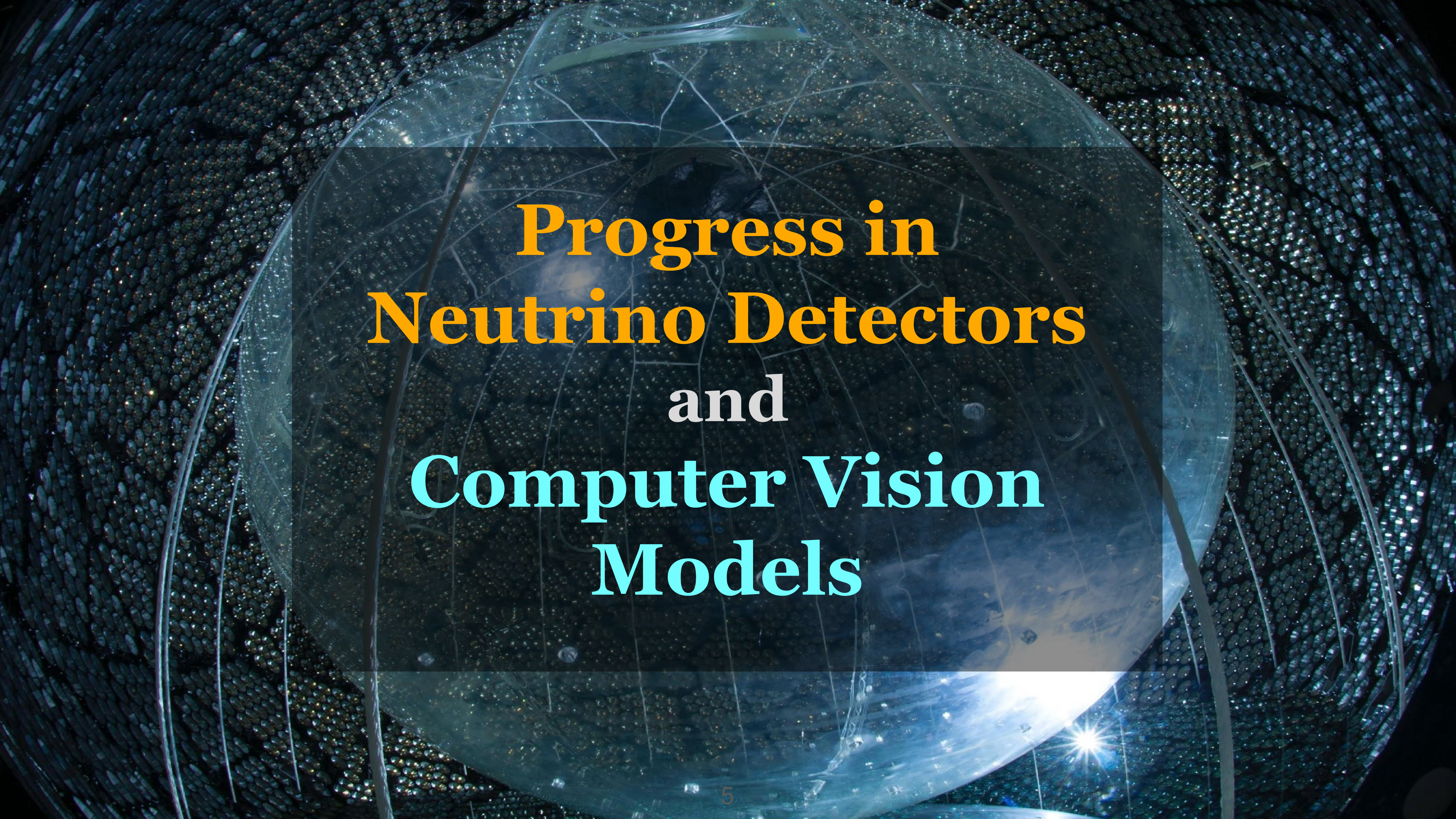
Cosmic Frontier

- Many low-resolution images
- Need for complex statistical inference



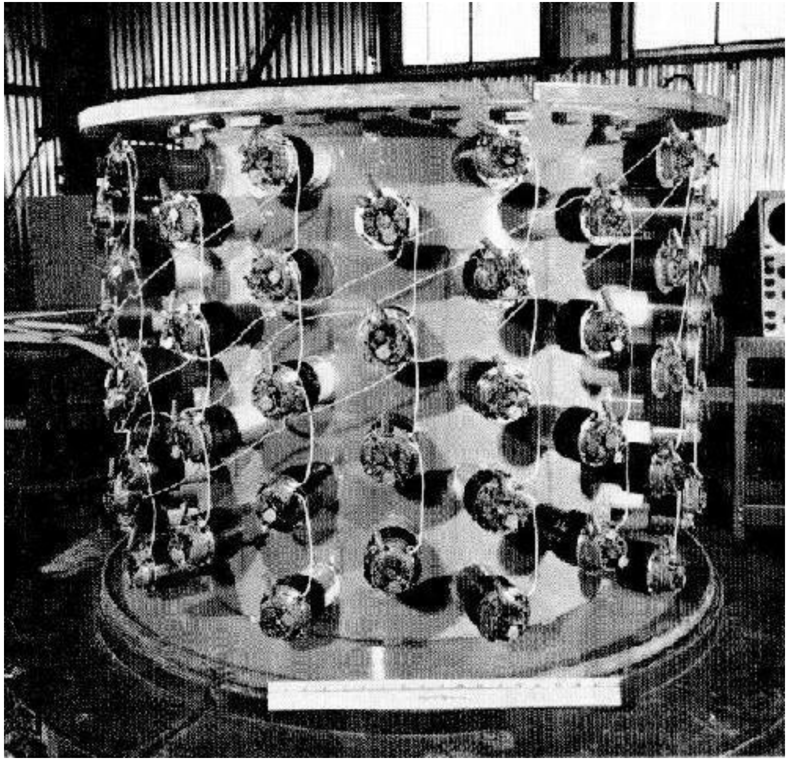
Ads: “Machine Laerning” review for particle physics now available in Particle Data Group review (new in 2021)!

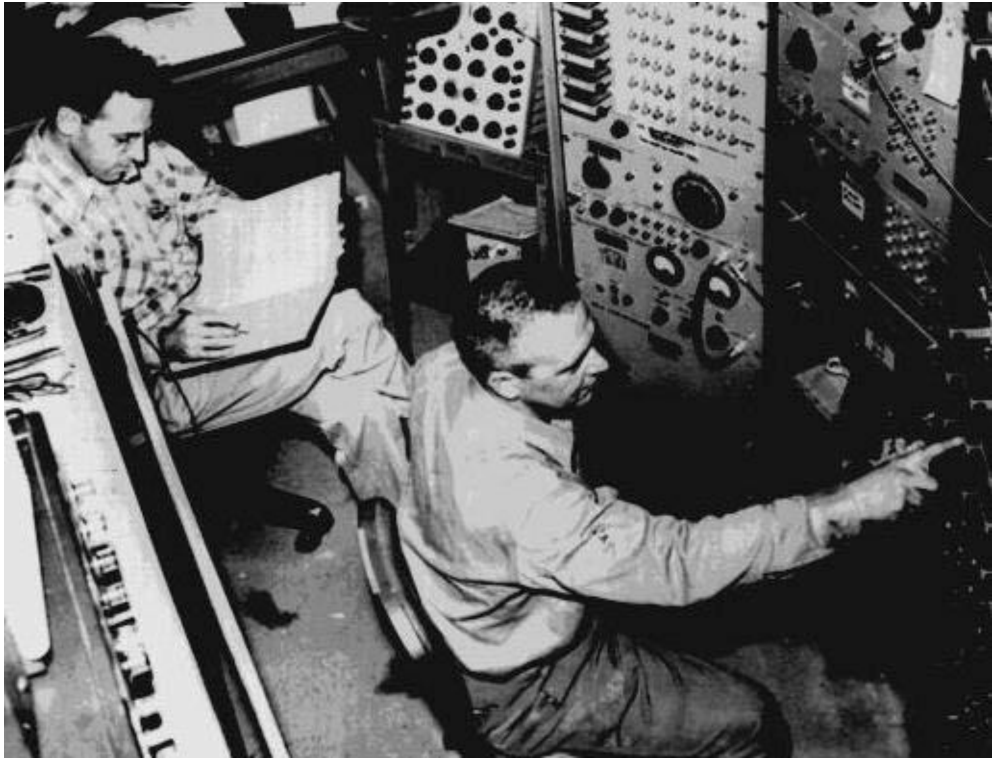




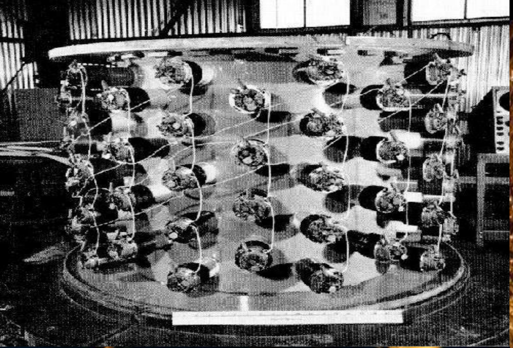
Progress in Neutrino Detectors and Computer Vision Models

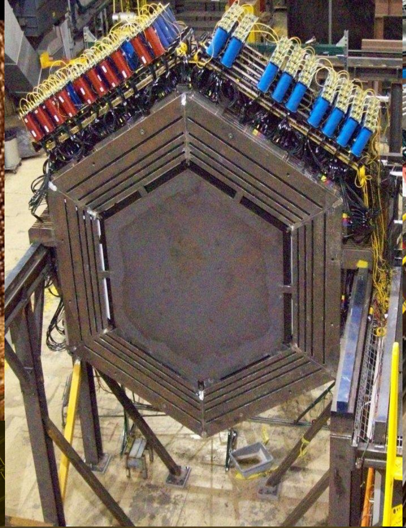
“Discovery”

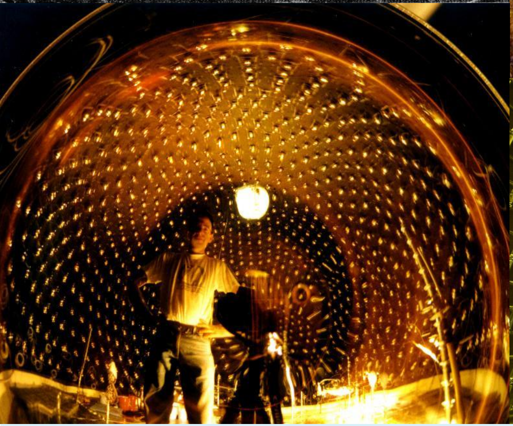


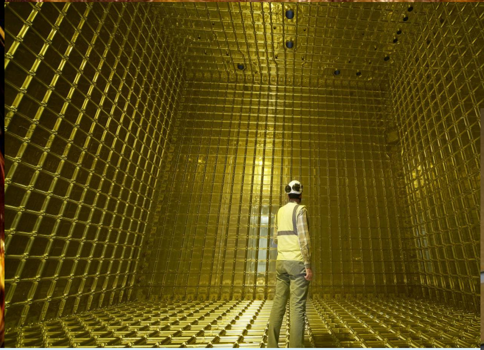


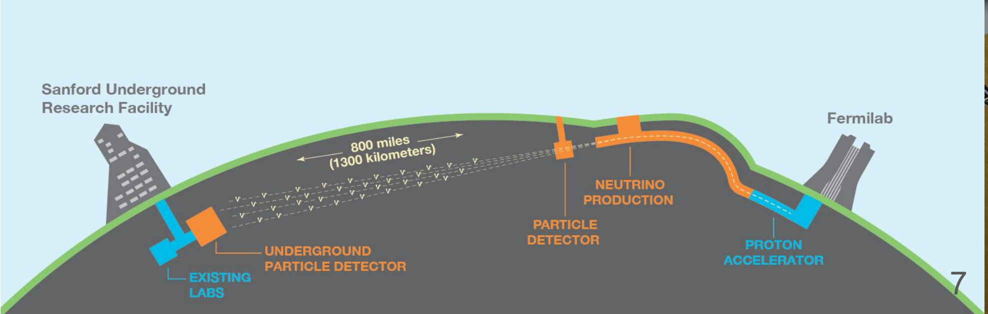
State-of-the-Art >60 years ago

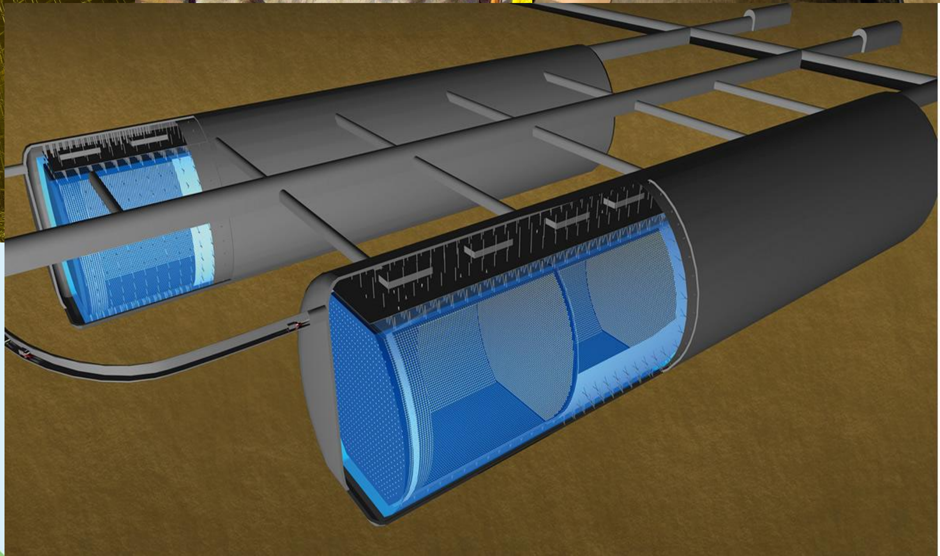


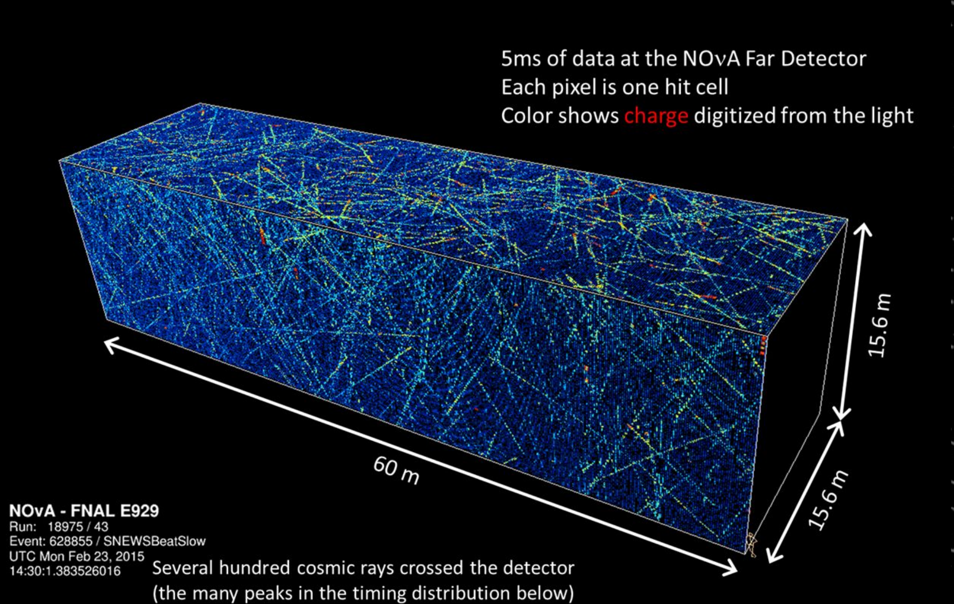
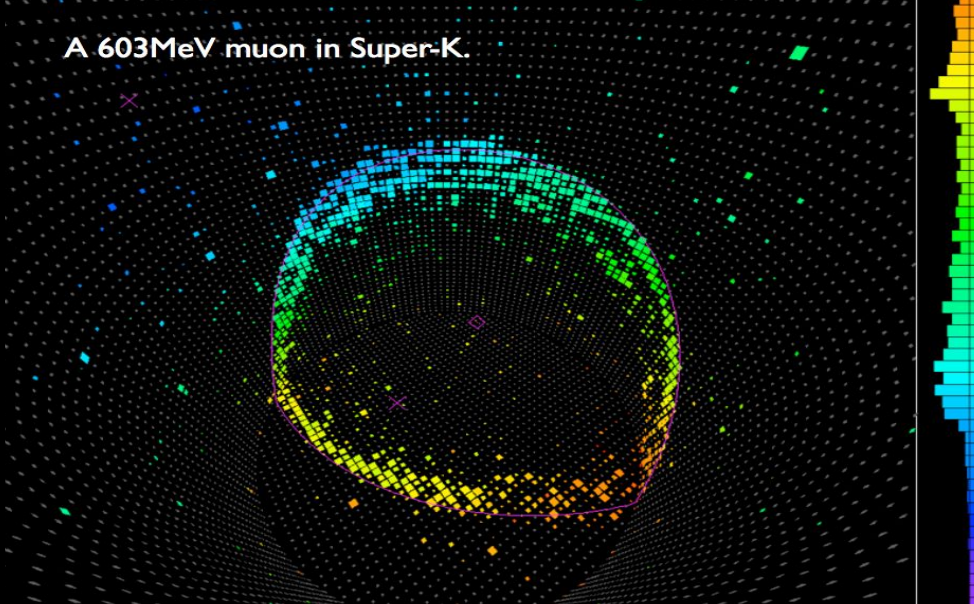
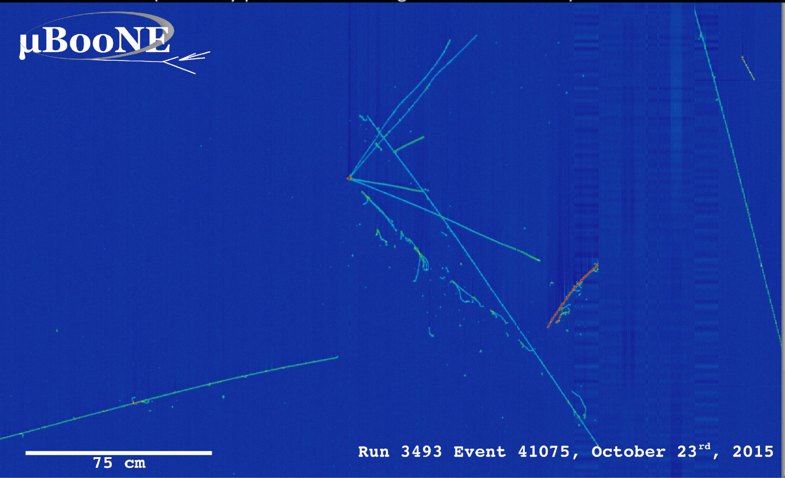
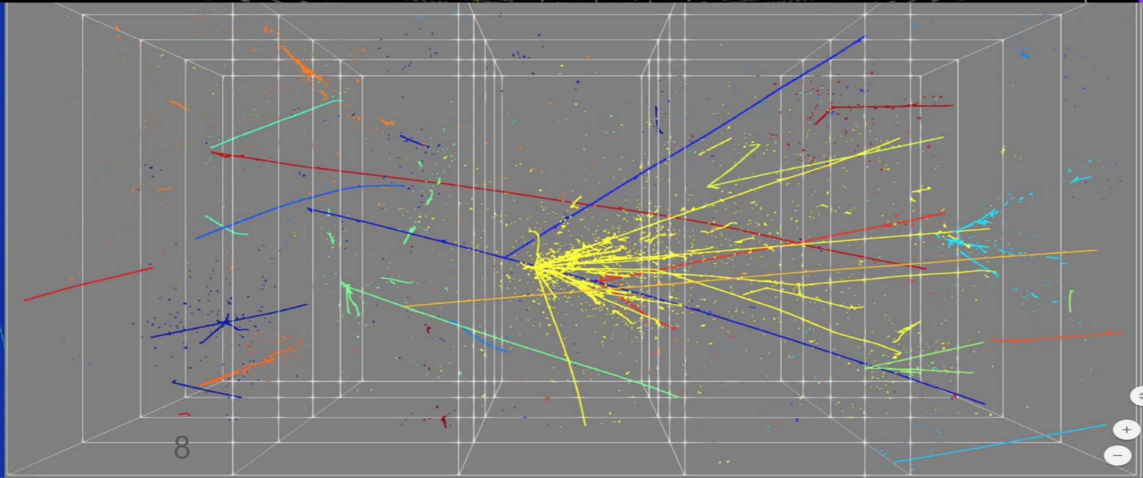


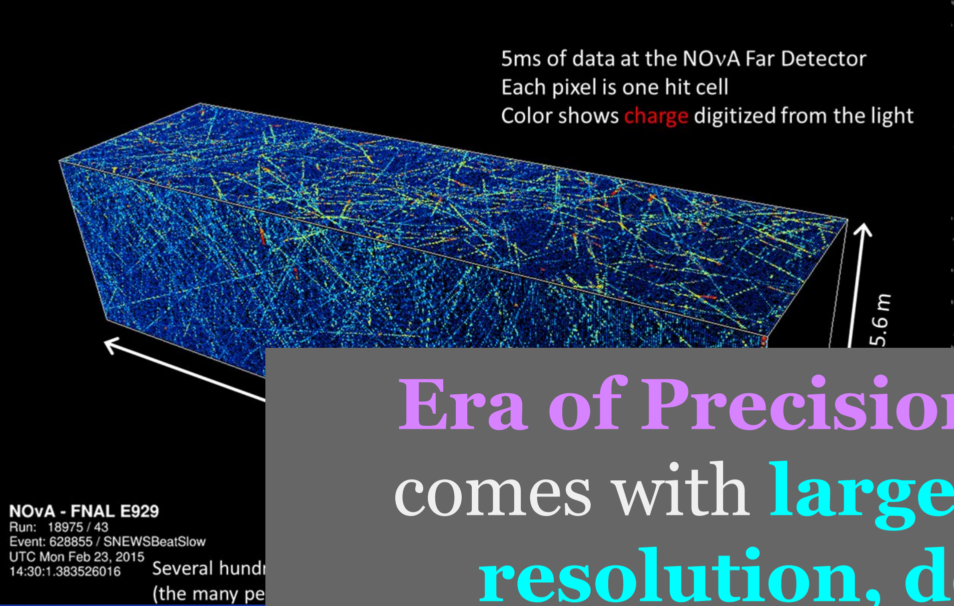


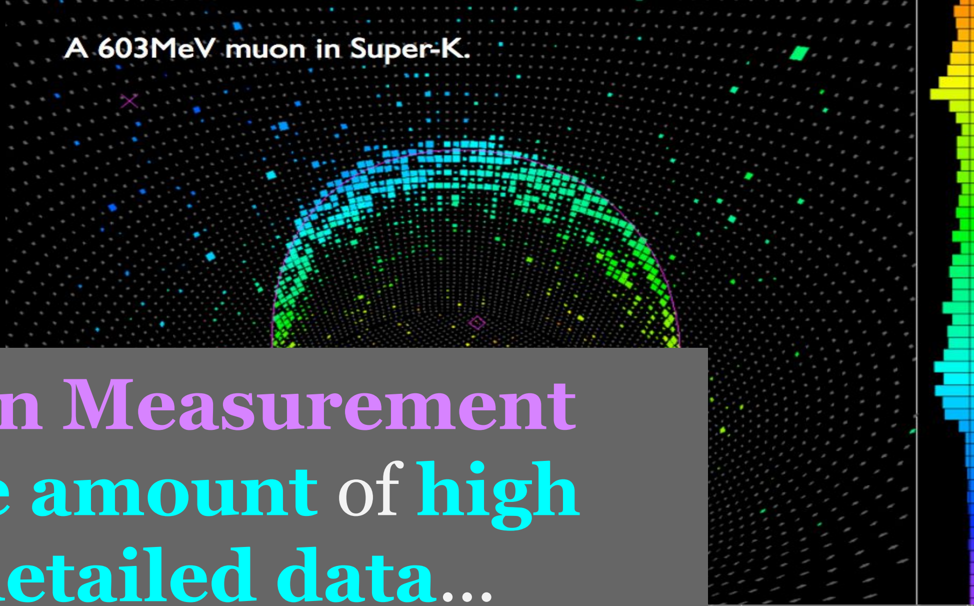



Sanford Underground
Research Facility
800 miles
(1300 kilometers)
Fermilab
NEUTRINO
PRODUCTION
PARTICLE
DETECTOR
PROTON
ACCELERATOR
UNDERGROUND
PARTICLE DETECTOR
EXISTING
LABS

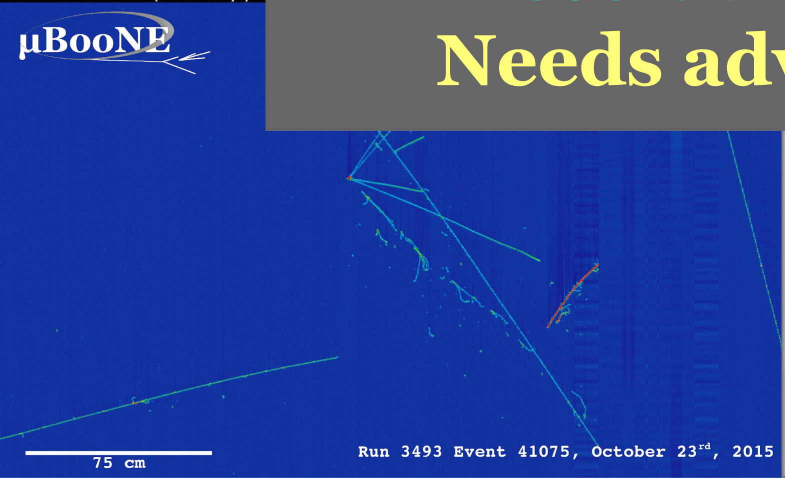


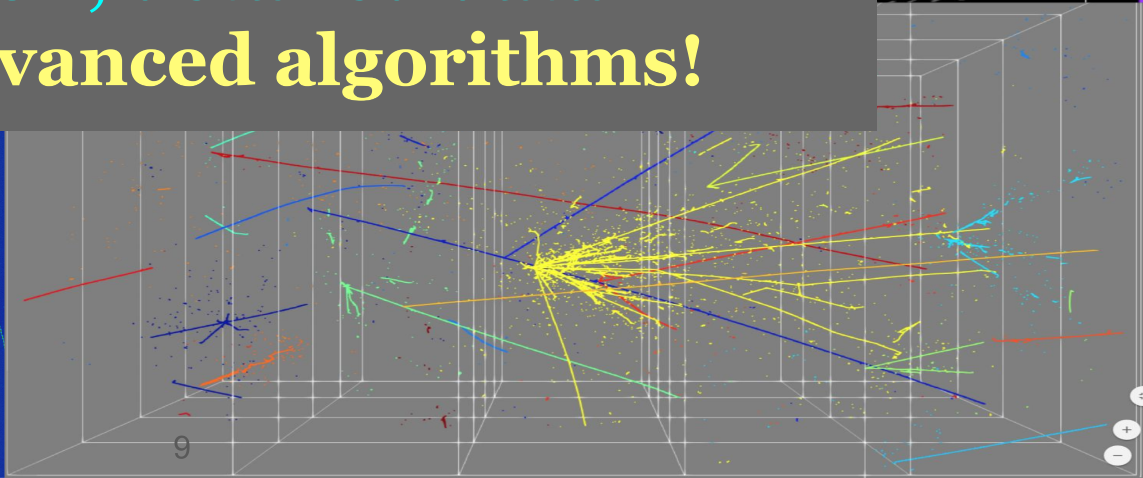

5ms of data at the NOνA Far Detector
Each pixel is one hit cell
Color shows charge digitized from the light
15.6 m
15.6 m
60 m
NOvA - FNAL E929
Run: 18975 / 43
Event: 628855 / SNEWSBeatSlow
UTC Mon Feb 23, 2015
14:30:1.383526016
Several hundred cosmic rays crossed the detector
(the many peaks in the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

A 603MeV muon in Super-K.

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






Era of Precision Measurement
comes with **large amount** of **high resolution, detailed data**…
Needs advanced algorithms!





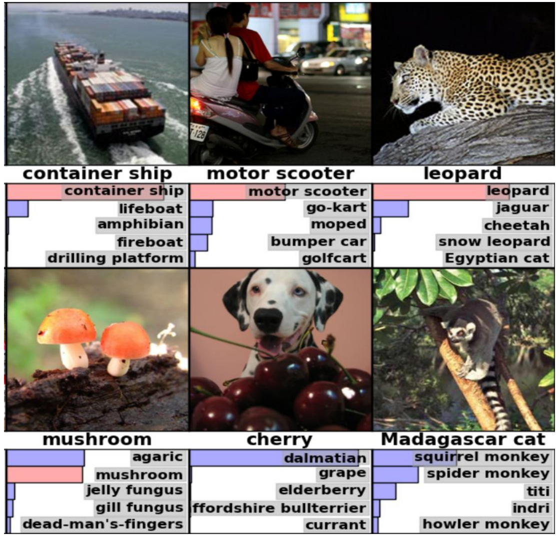
Machine Learning and Computer Vision

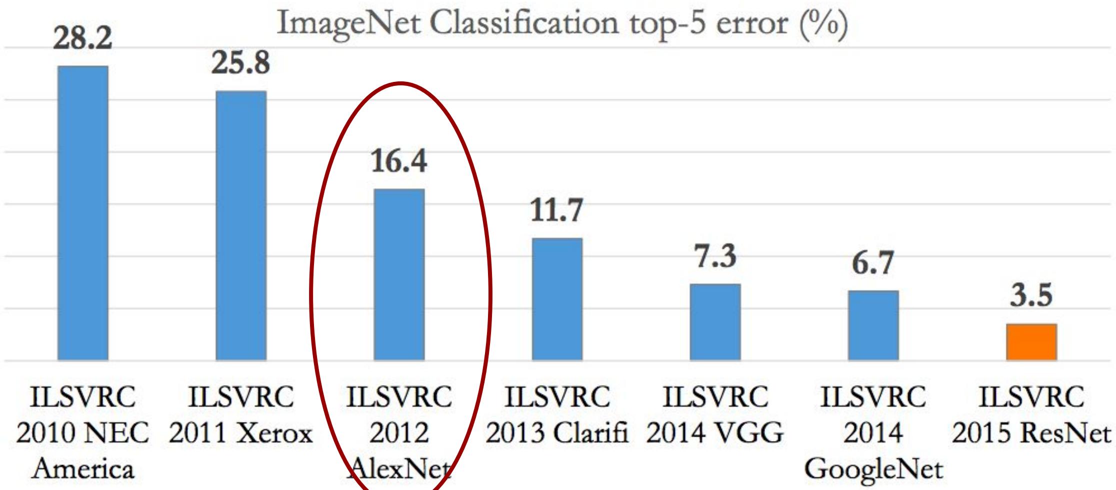
Image analysis can be difficult...

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

Machine Learning and Computer Vision

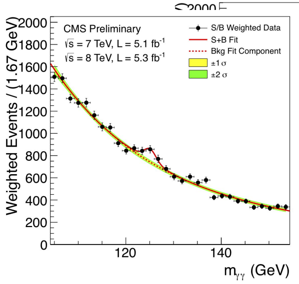
“Deep Learning” sparked 2012
sparked the wave of modern machine learning (ML) and AI

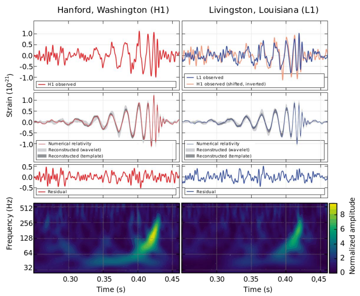


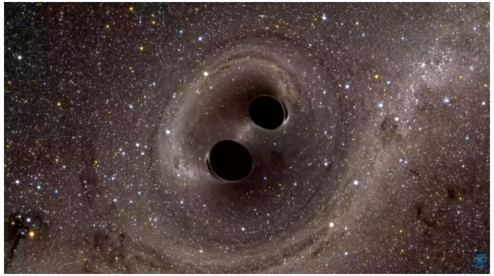


Machine Learning in Neutrino Physics & HEP

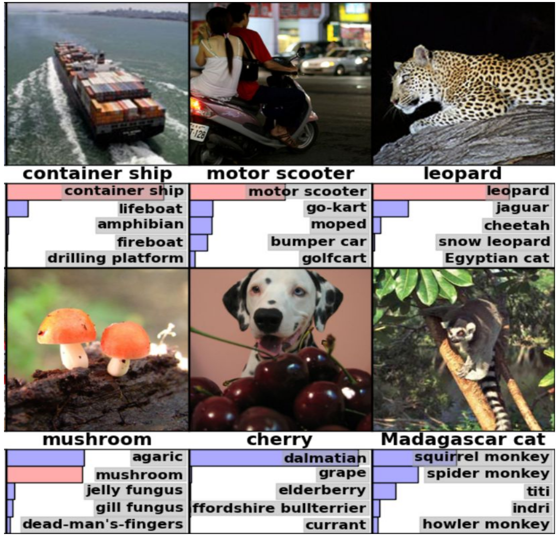
Machine Learning and Computer Vision

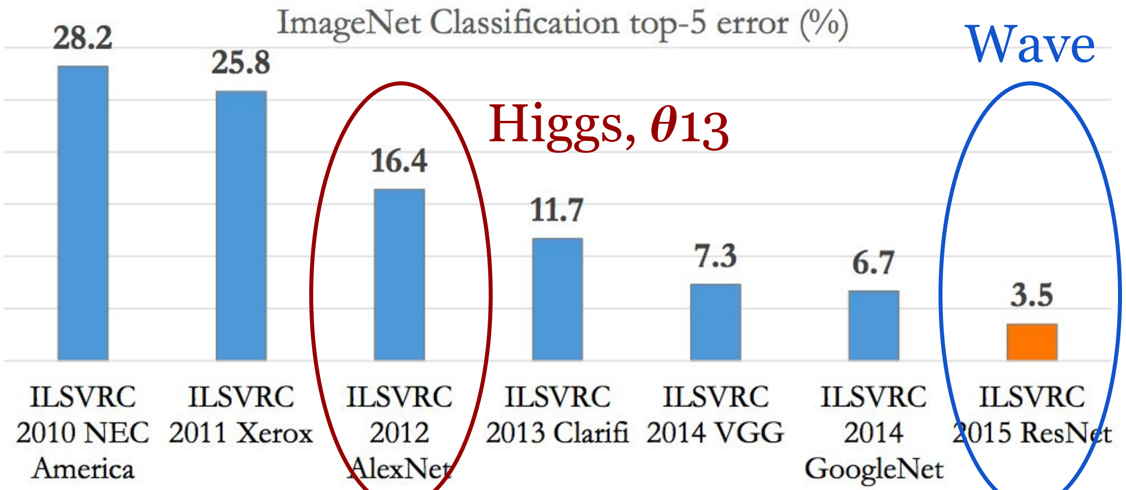






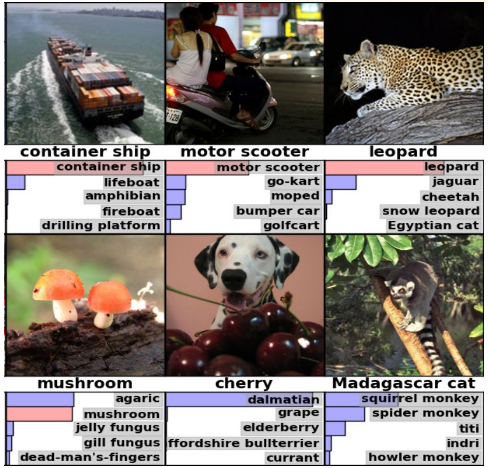
Gravitational Wave

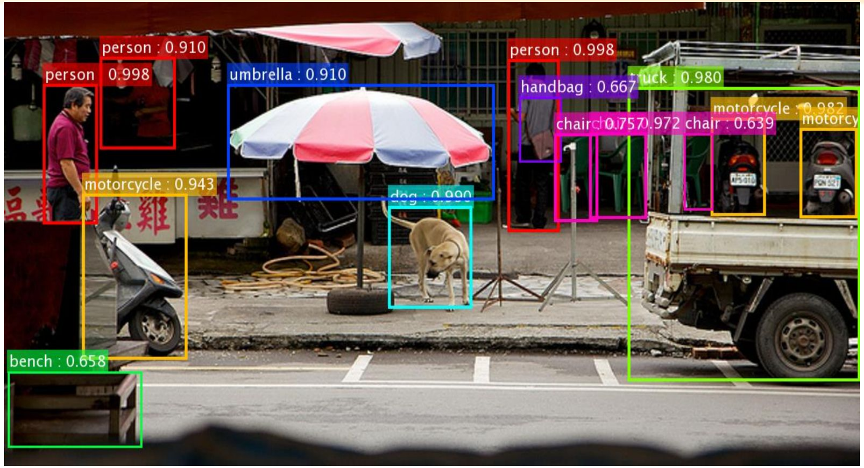


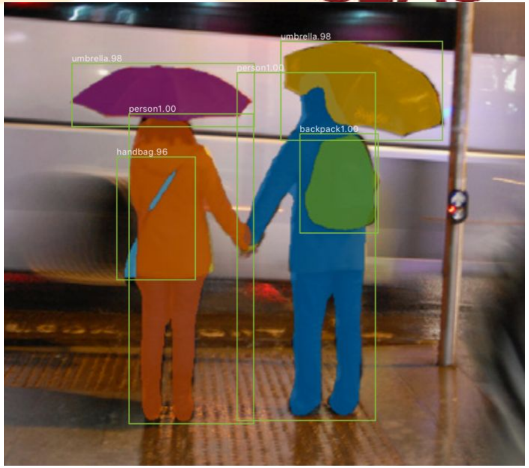


Higgs, θ13

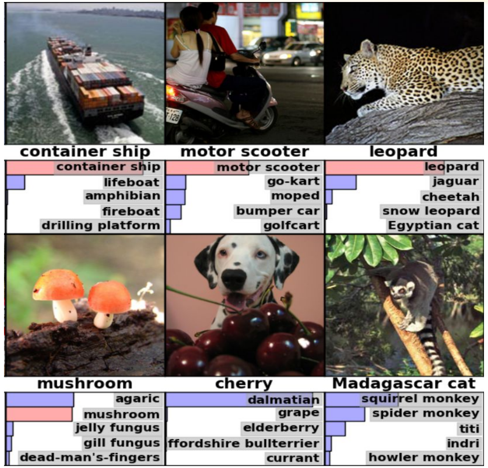
Machine Learning and Computer Vision

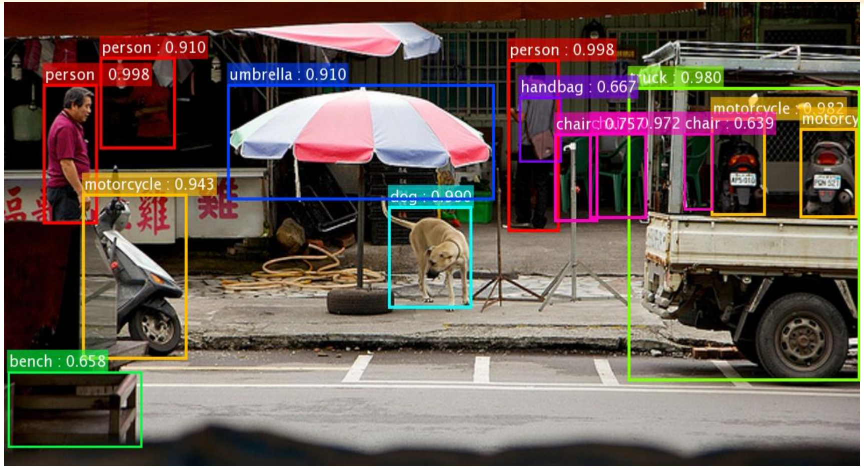


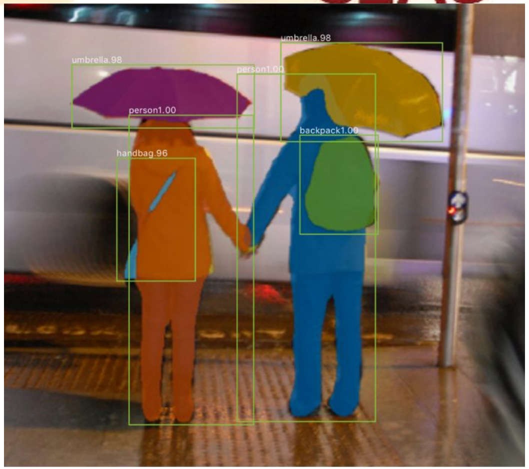




Machine Learning and Computer Vision

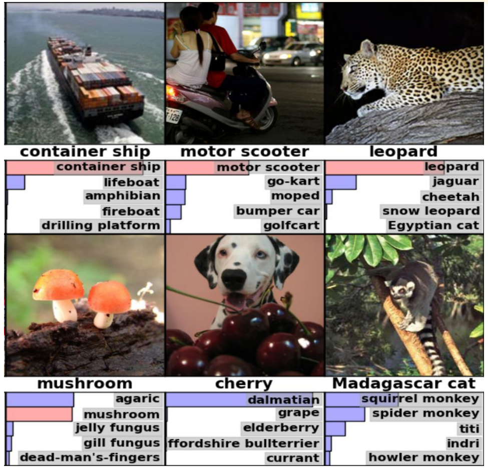


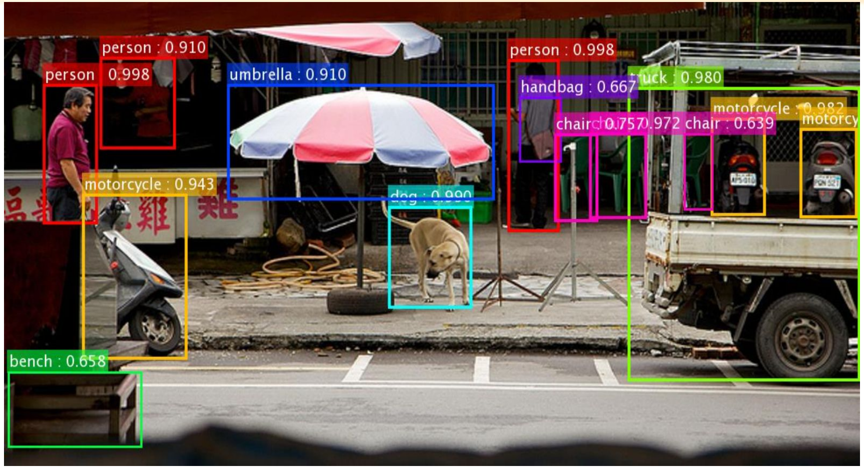


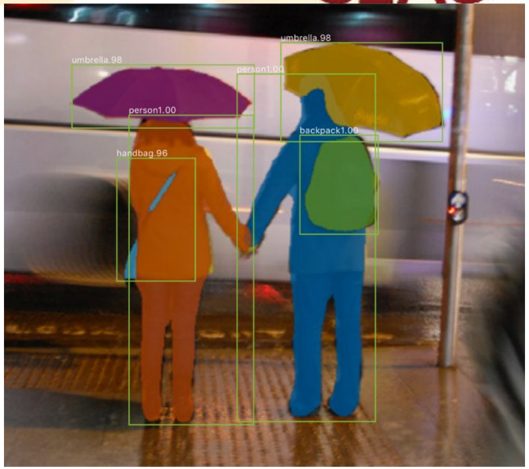


Machine Learning in Neutrino Physics & HEP

Machine Learning and Computer Vision

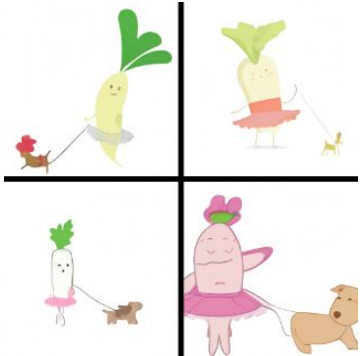






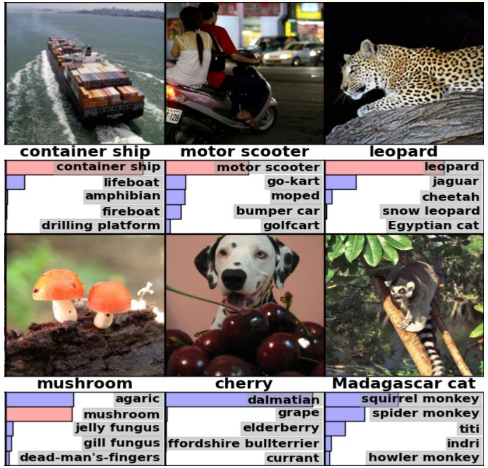
"girl in pink dress is jumping in air."

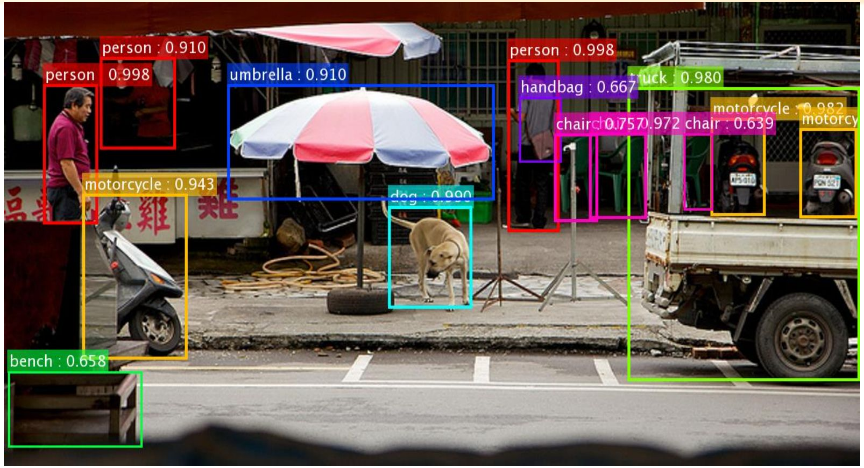
A baby daikon radish in a tutu walking a dog

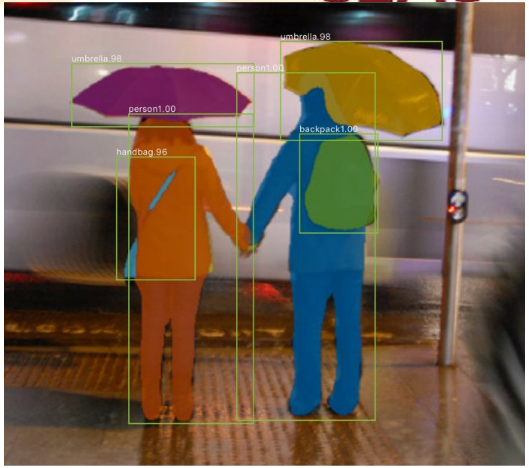


Machine Learning in Neutrino Physics & HEP

Machine Learning and Computer Vision

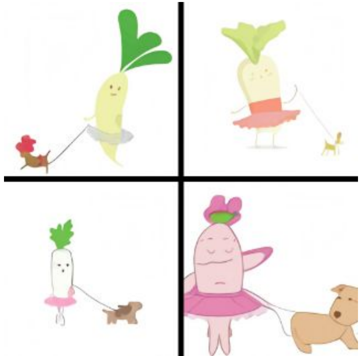


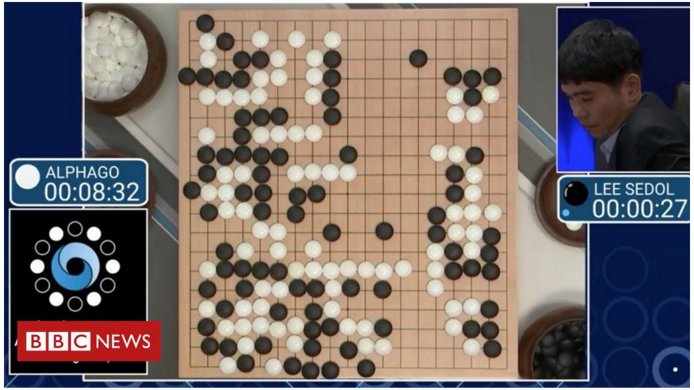




"girl in pink dress is jumping in air."

A baby daikon radish in a tutu walking a dog



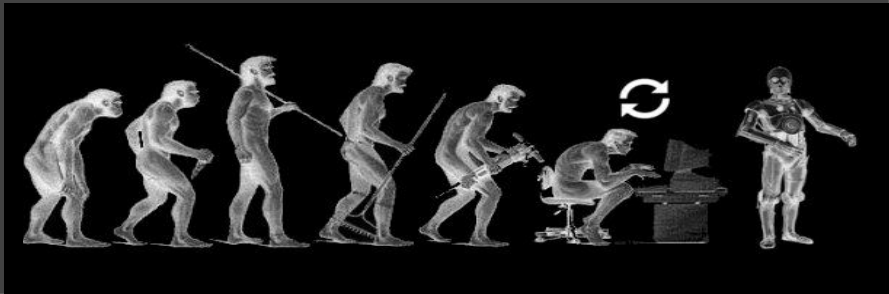


Outline for the rest

Can we apply ML for image analysis?

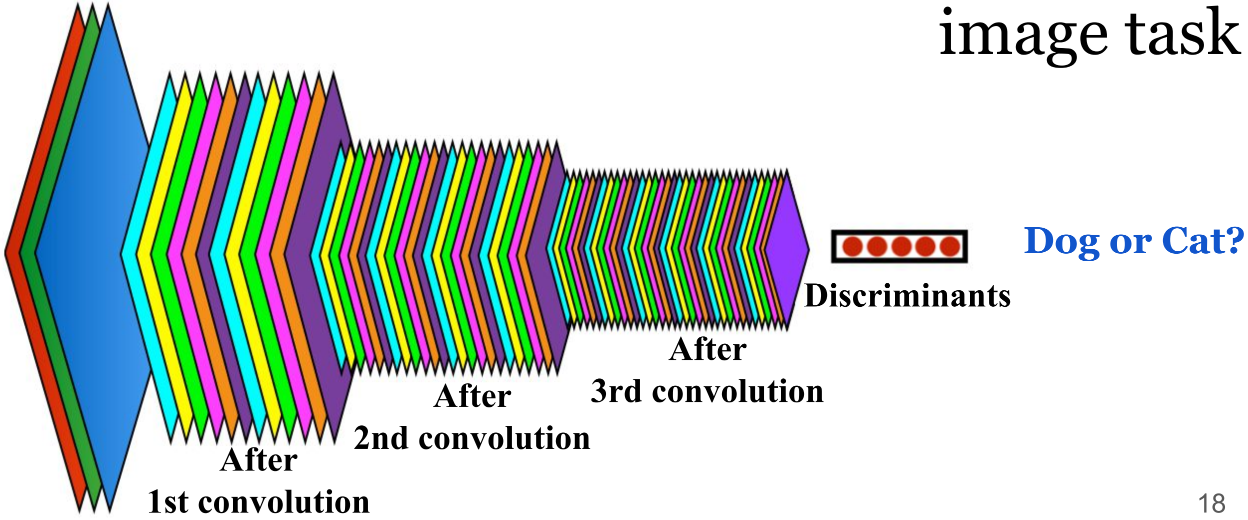
Can we reconstruct physics features with ML?

Next: ML & simulation



Deep Neural Network for Image Analysis

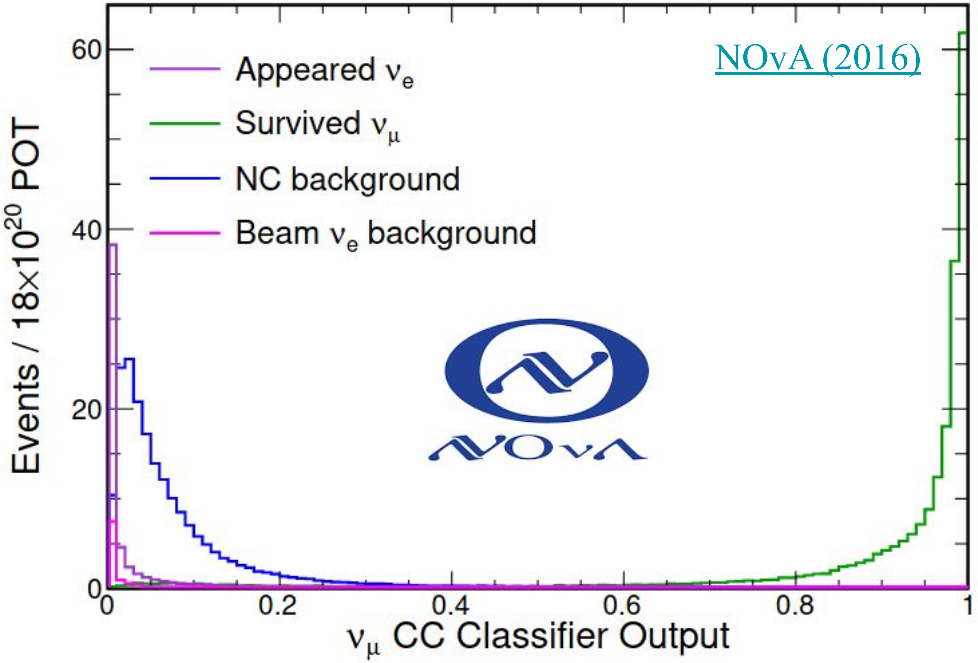
Convolutional Neural Network (CNN)
Learns to transform 2D image data into an array of useful features to address an image task

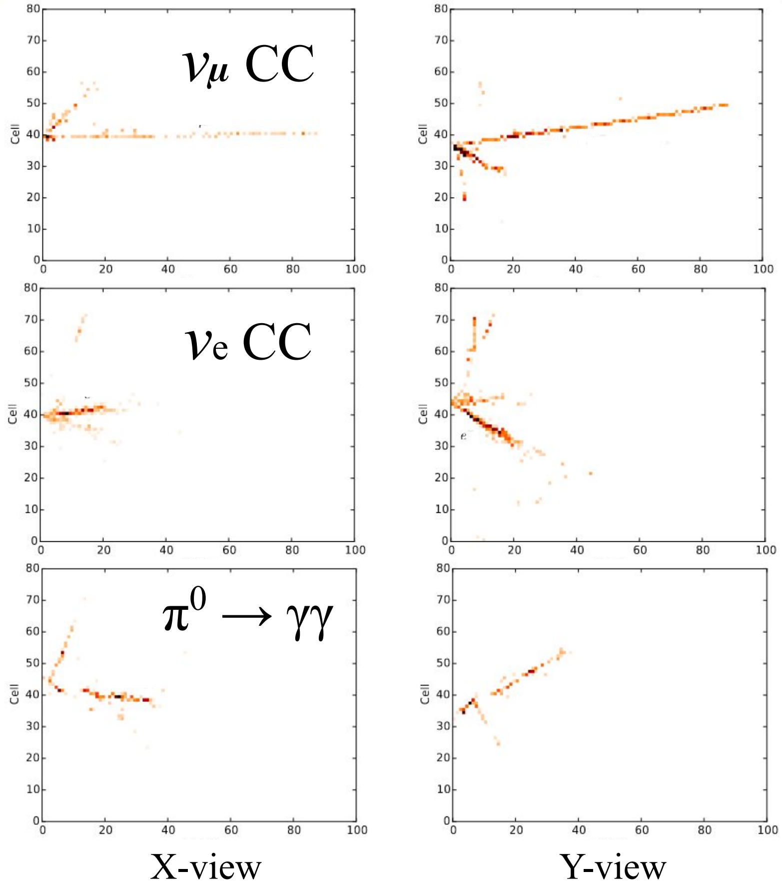


Machine Learning in Neutrino Physics & HEP
Deep Neural Network for Image Analysis

SLAC

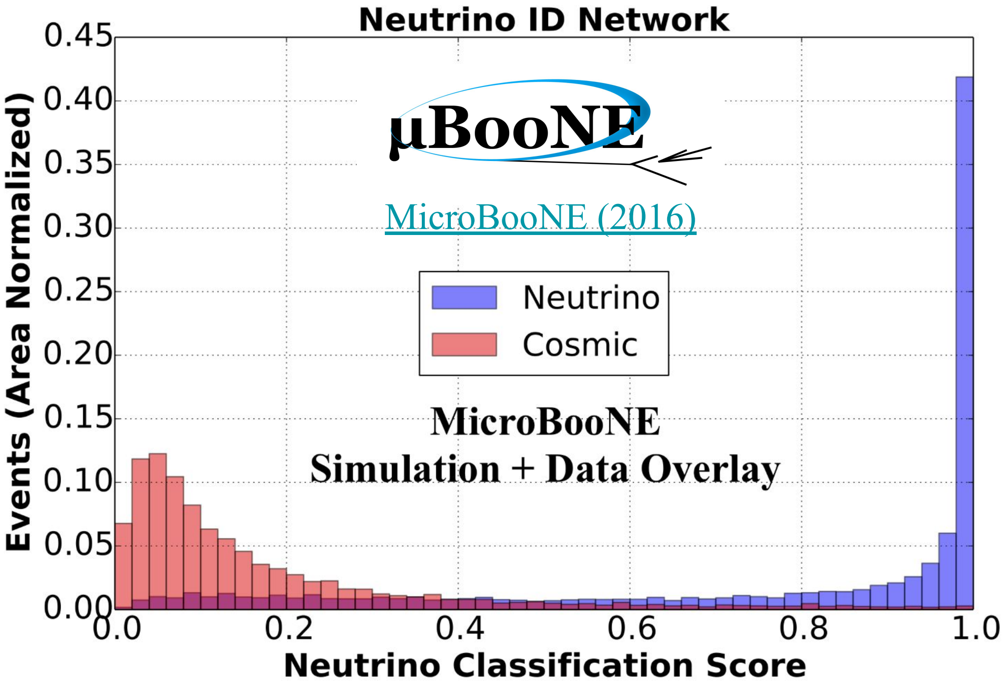
First attempt: CNN image classifier for neutrino interaction classification

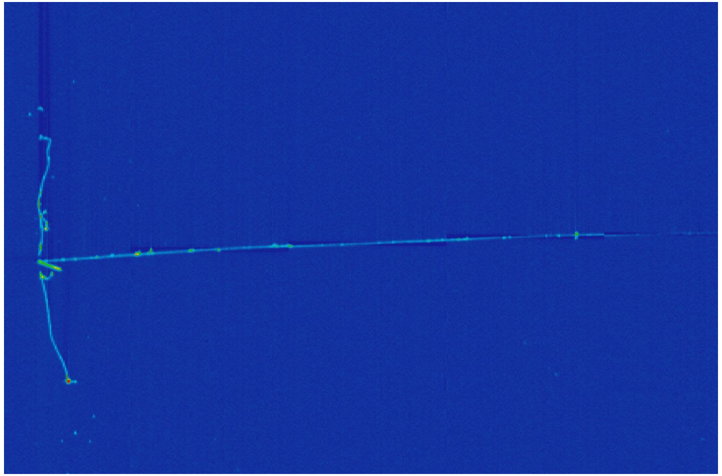


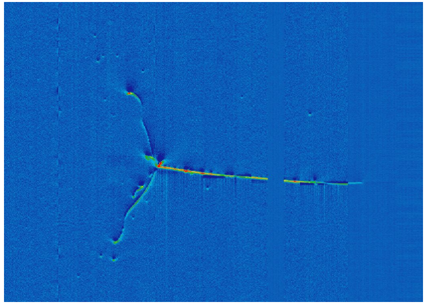


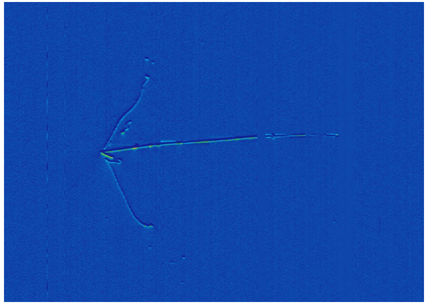
Machine Learning in Neutrino Physics & HEP
Deep Neural Network for Image Analysis

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





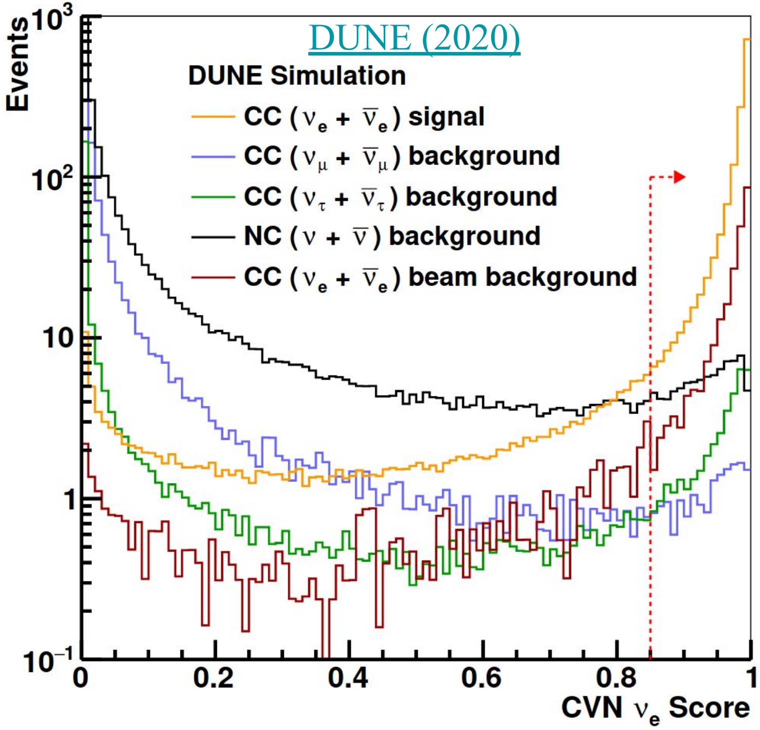


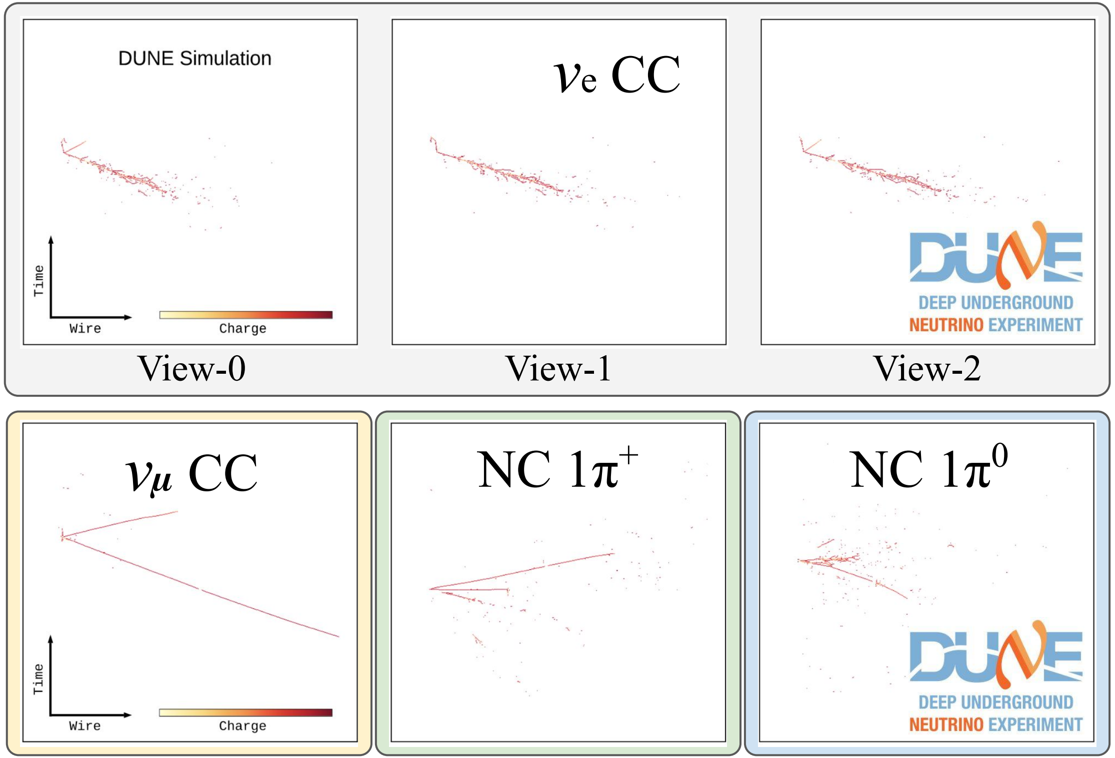


Machine Learning in Neutrino Physics & HEP

Deep Neural Network for Image Analysis

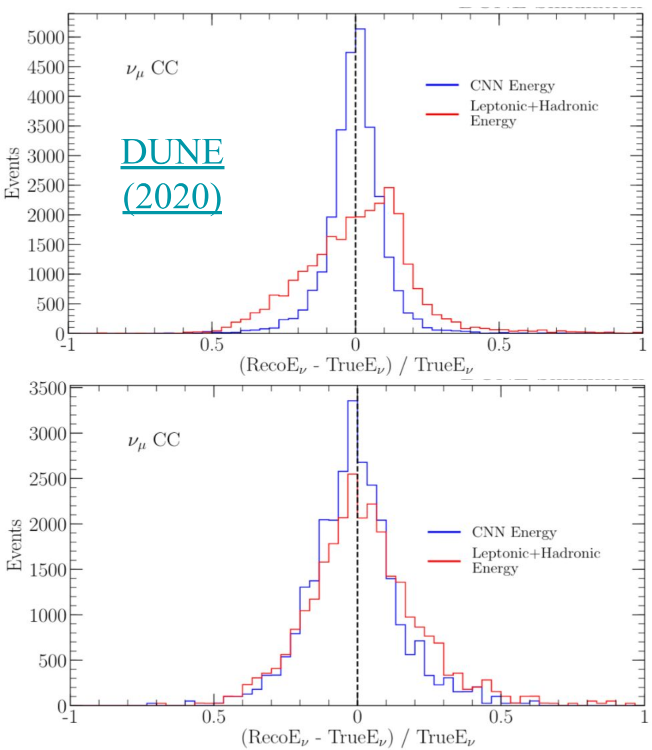
CNN image classification remains to date as a strong approach

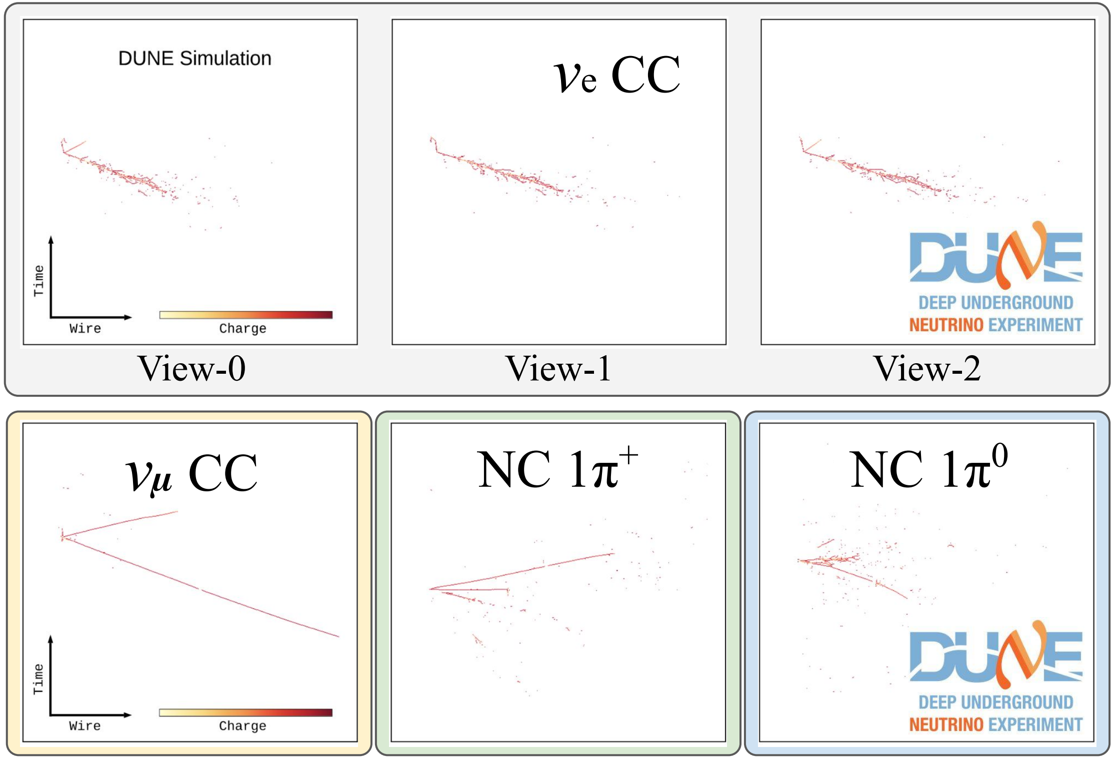




Machine Learning in Neutrino Physics & HEP
Deep Neural Network for Image Analysis

… and CNN also used for an image-level regression (e.g. neutrino energy)





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

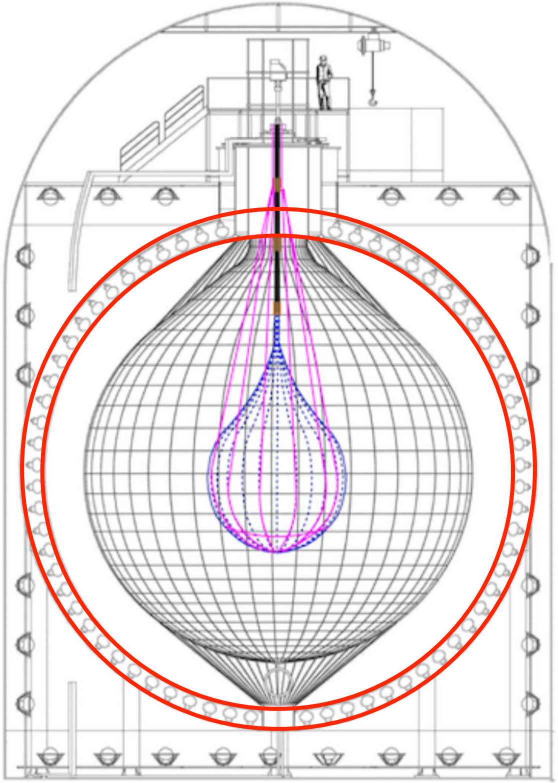
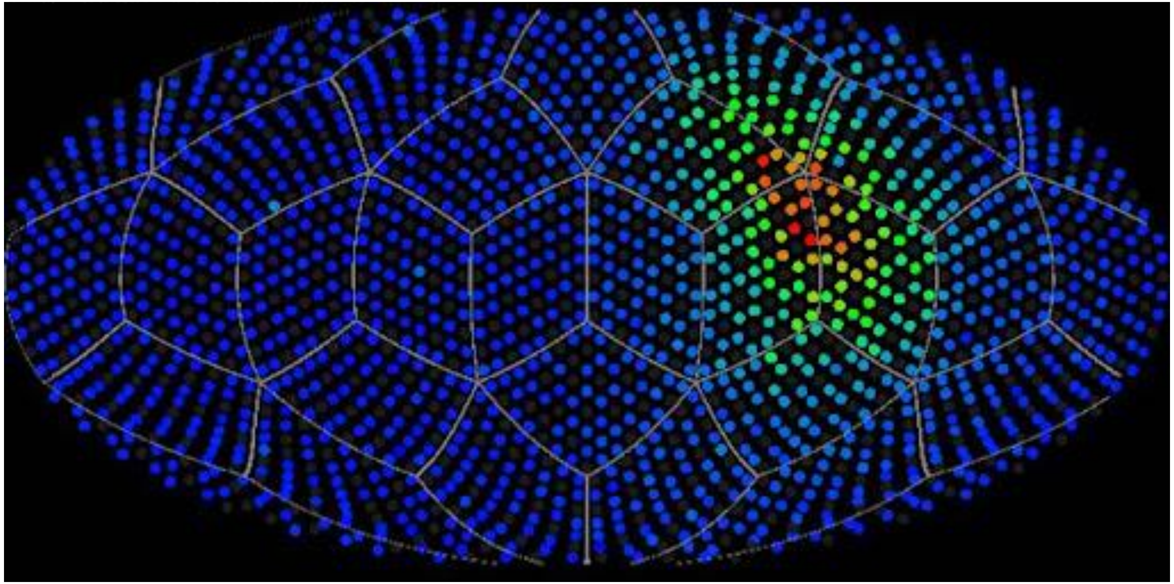
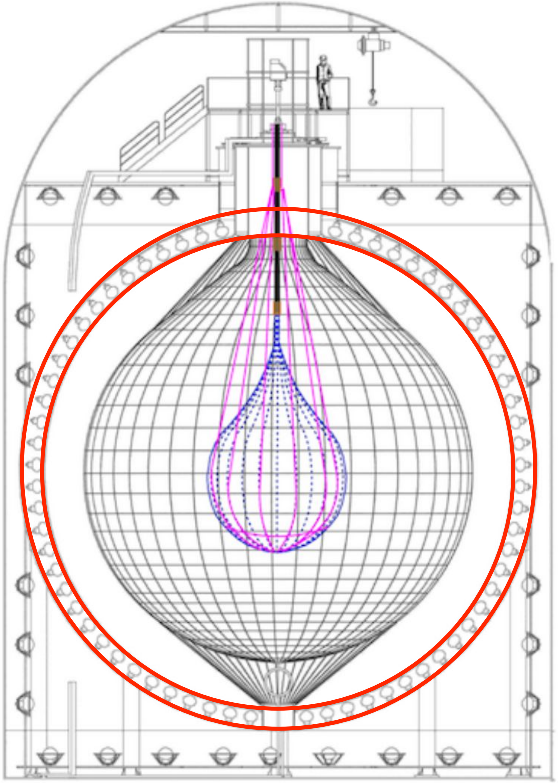


Image data recorded in non-cartesian grid



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

SLAC



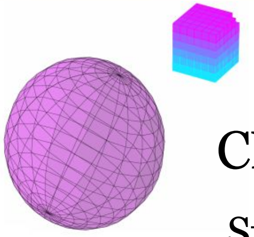
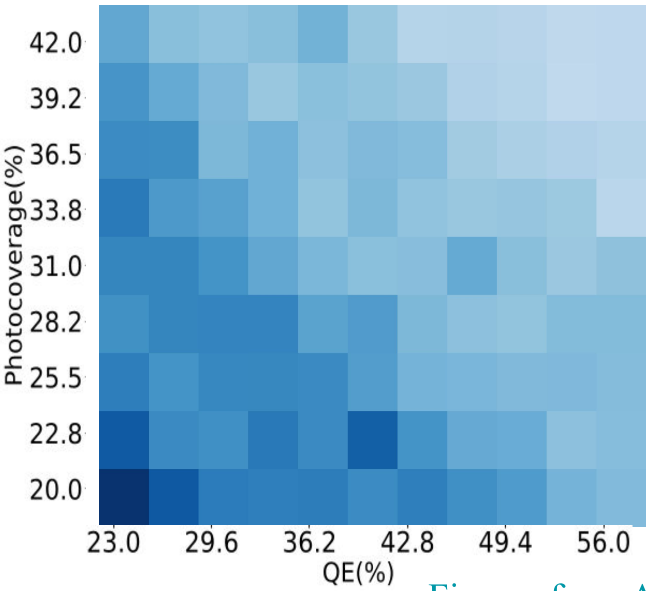
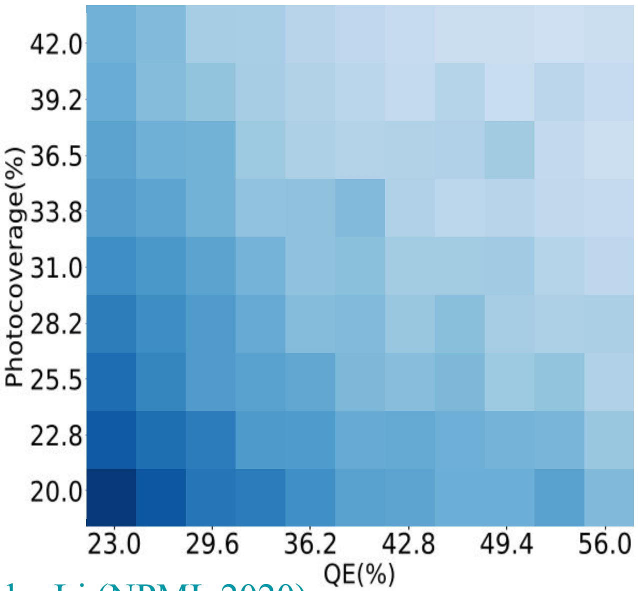


Image data recorded in non-cartesian grid
CNN with appropriate equivalance (rotation)

Standard CNN



Spherical CNN

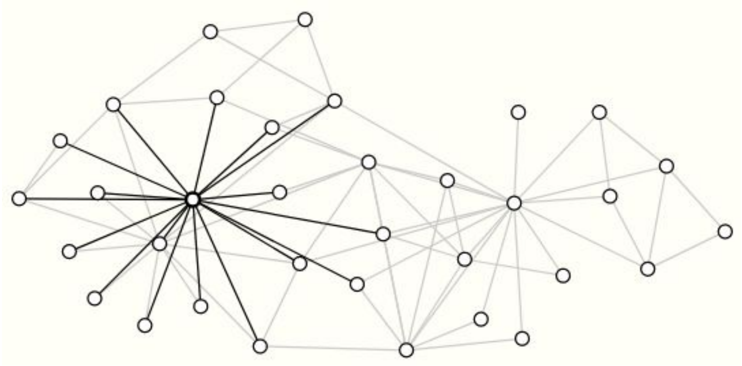


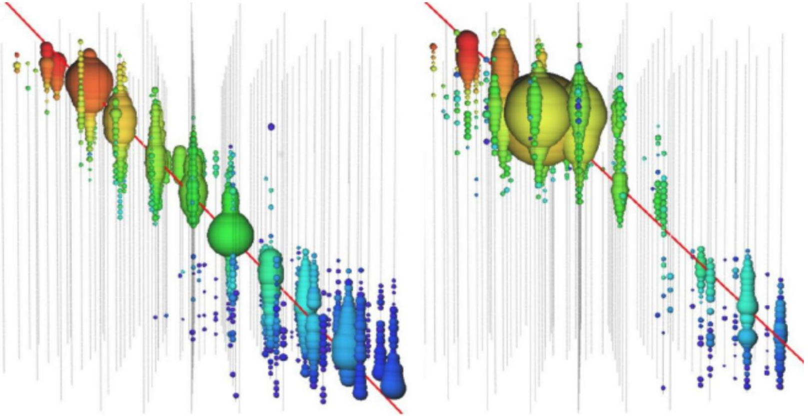
Figures from Aobo Li (NPML 2020)

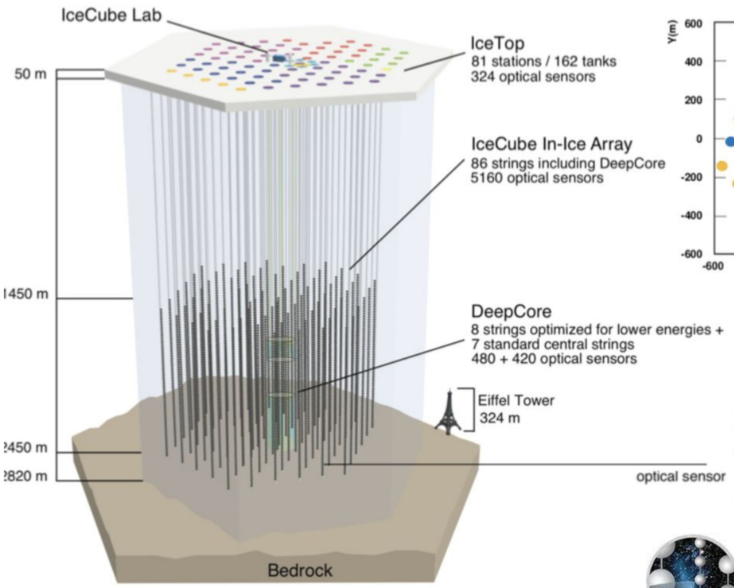
Machine Learning in Neutrino Physics & HEP

Deep Neural Network for Image Analysis

Graph as a flexible representation suited for non-square-grid / multi-modal data.

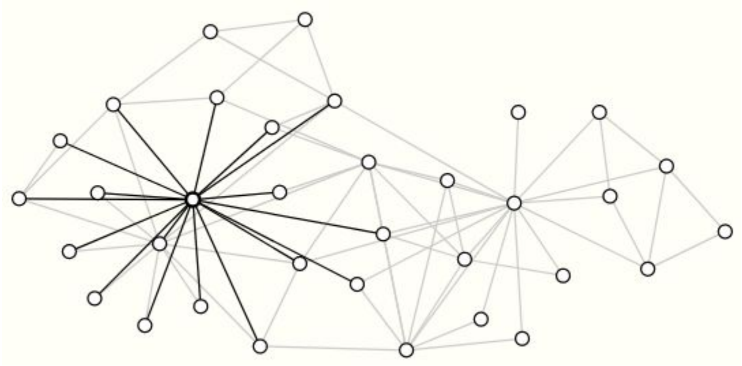


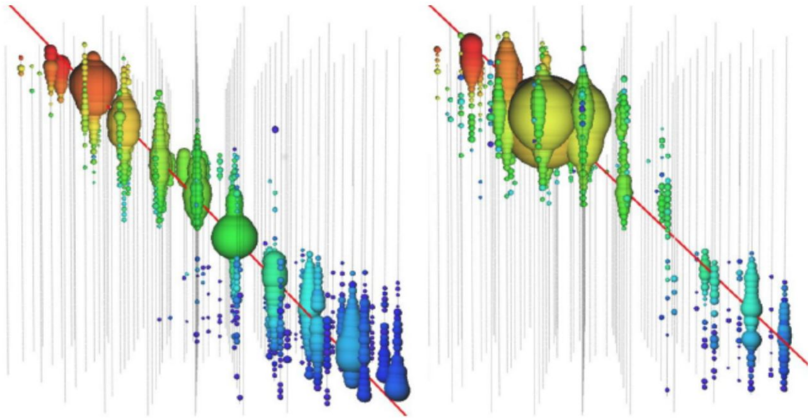




ICECUBE
SOUTH POLE NEUTRINO OBSERVATORY

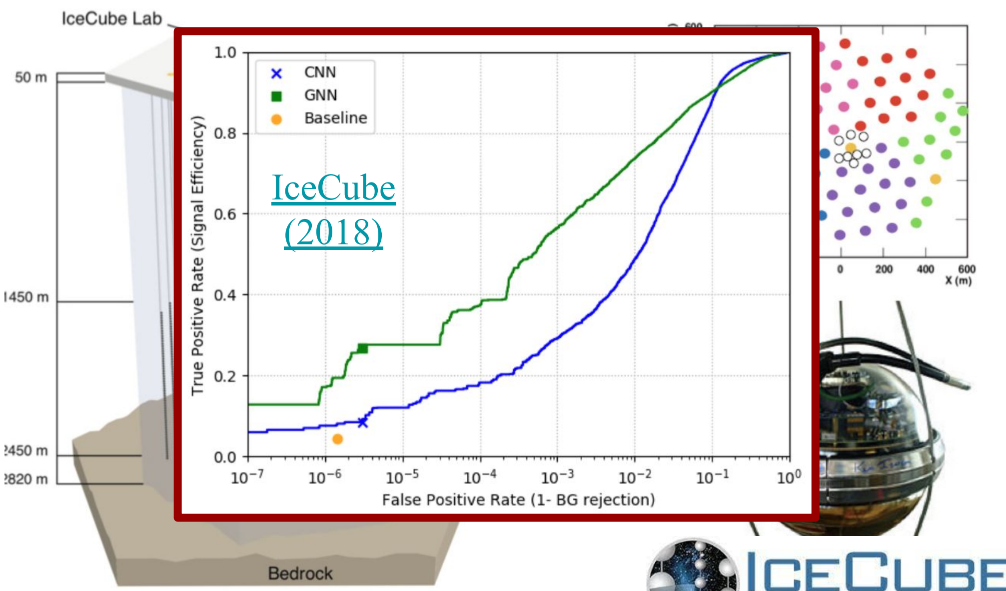
Machine Learning in Neutrino Physics & HEP
Deep Neural Network for Image Analysis





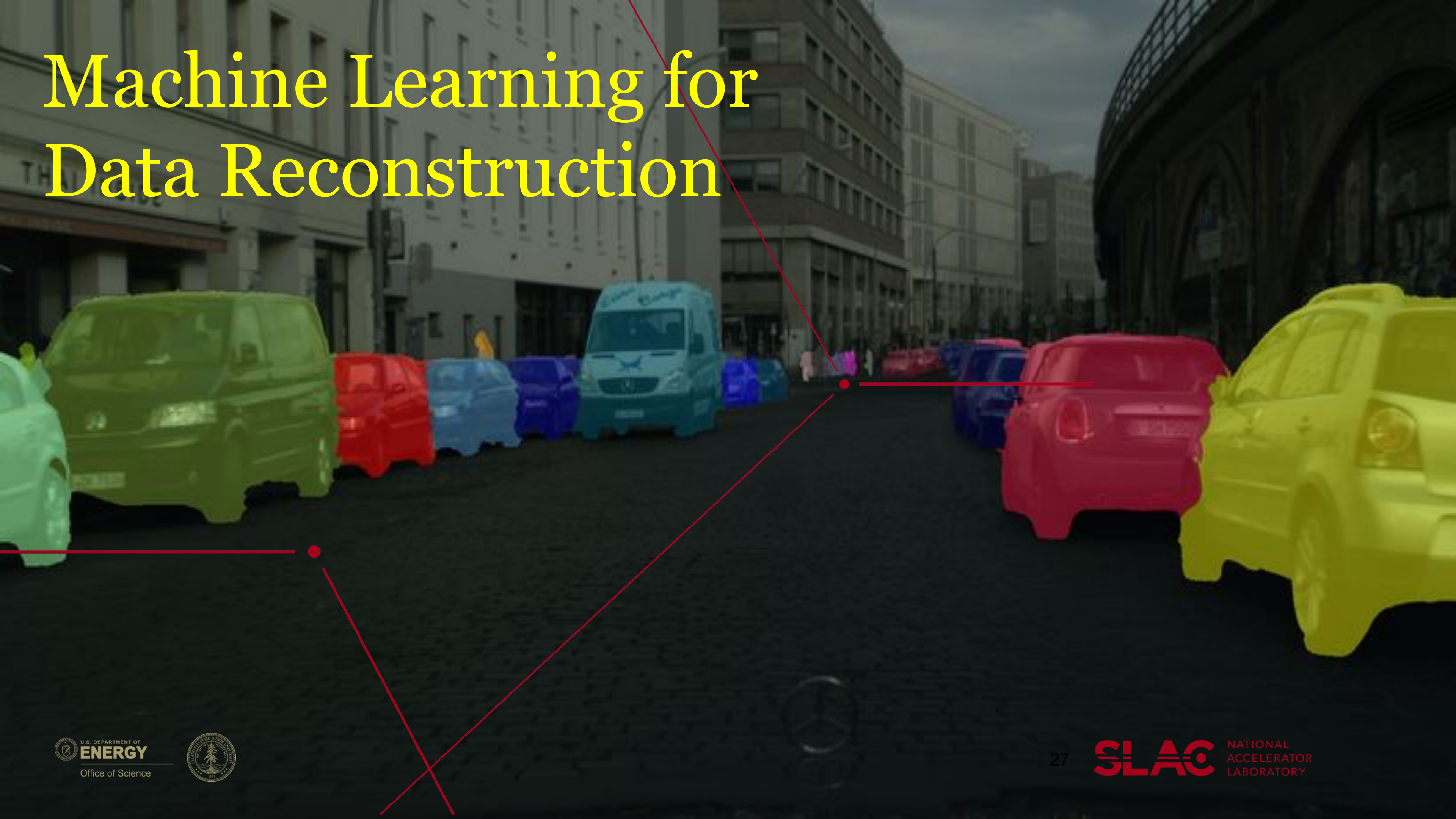
Graph as a flexible representation suited for non-square-grid / multi-modal data.

Graph Neural Network for neutrino ID



[IceCube (2018)]

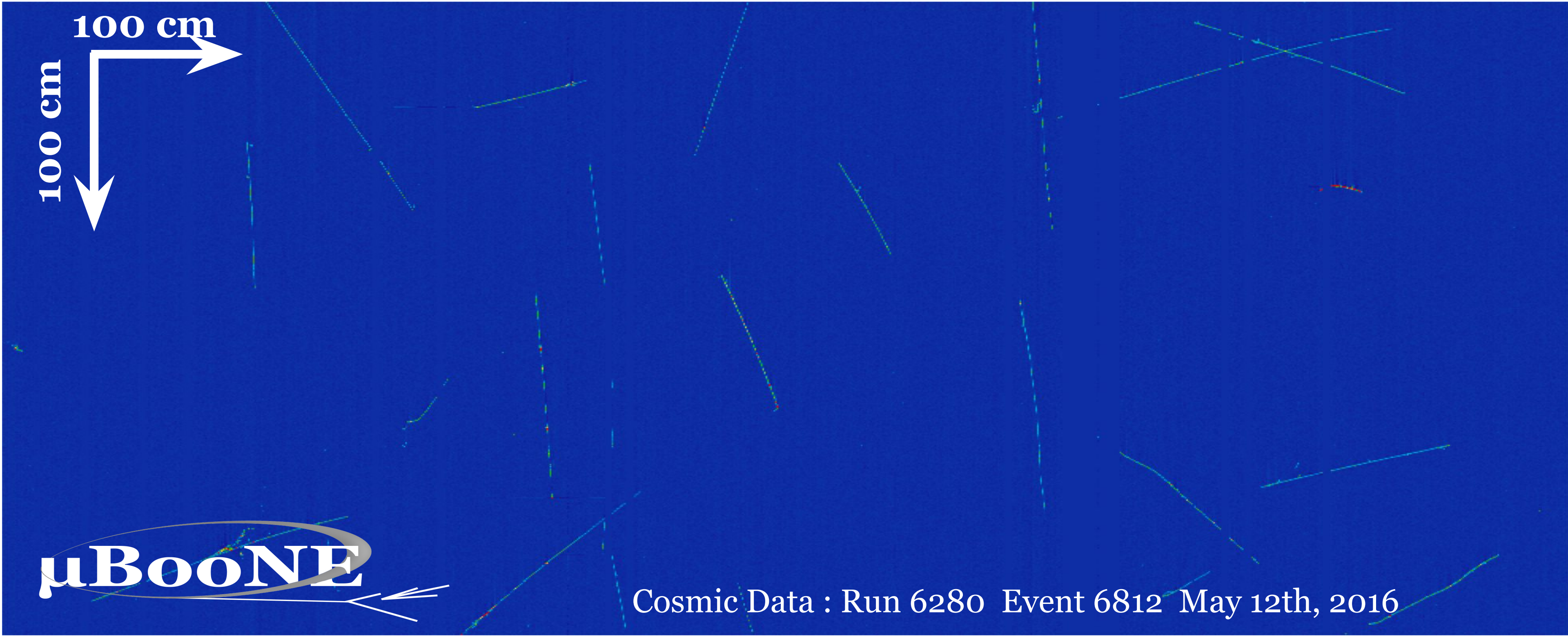
Machine Learning for Data Reconstruction



Machine Learning in Neutrino Physics & HEP
Deep Neural Network for Data Reconstruction

SLAC

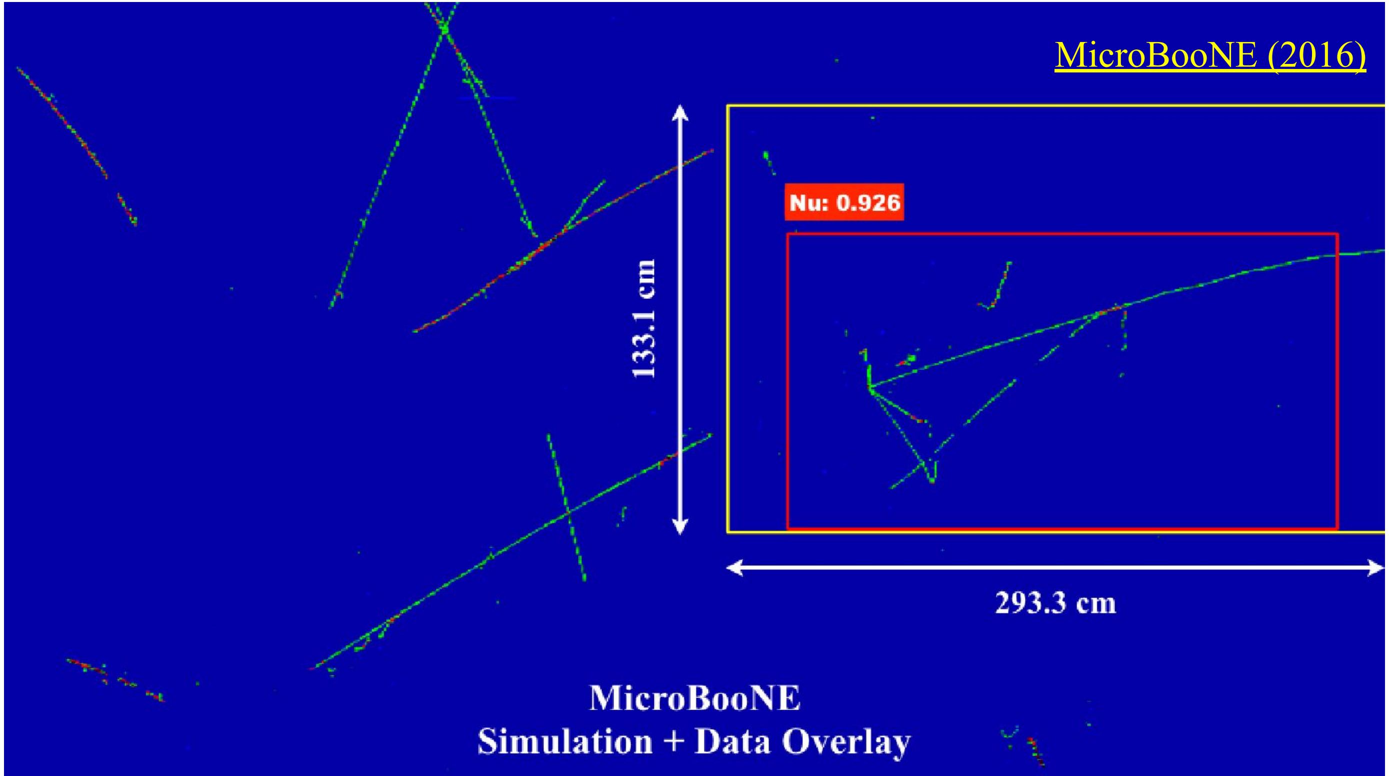
Signal v.s. Background? …. more like "Signal in Background"



Deep Neural Network for Data Reconstruction

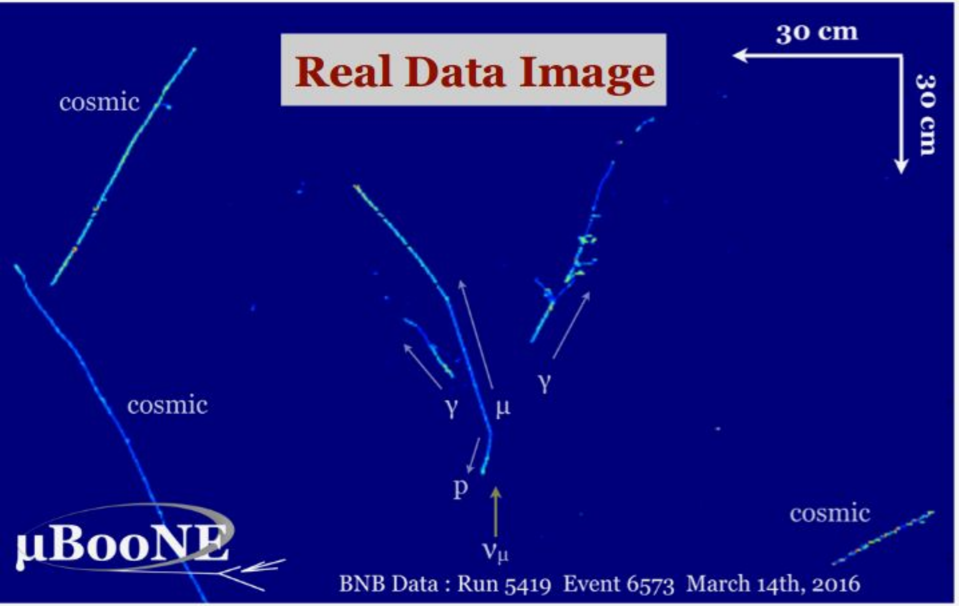
SLAC

Tell "where" the signal is

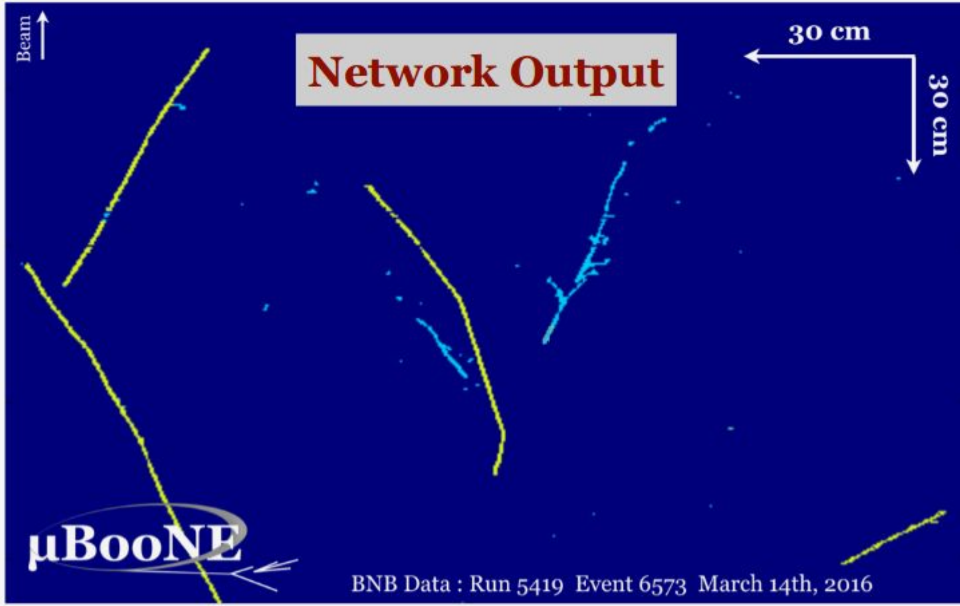


Deep Neural Network for Data Reconstruction

Even at the pixel-level! … then onto a "full reconstruction chain"



Network Input

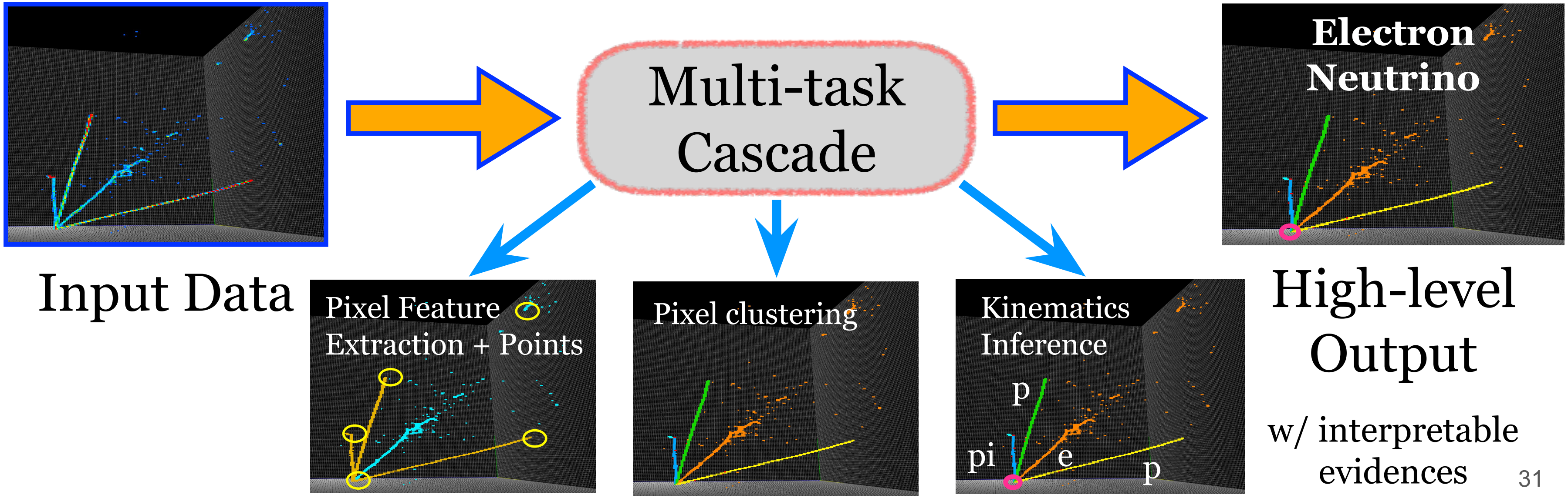


Network Output

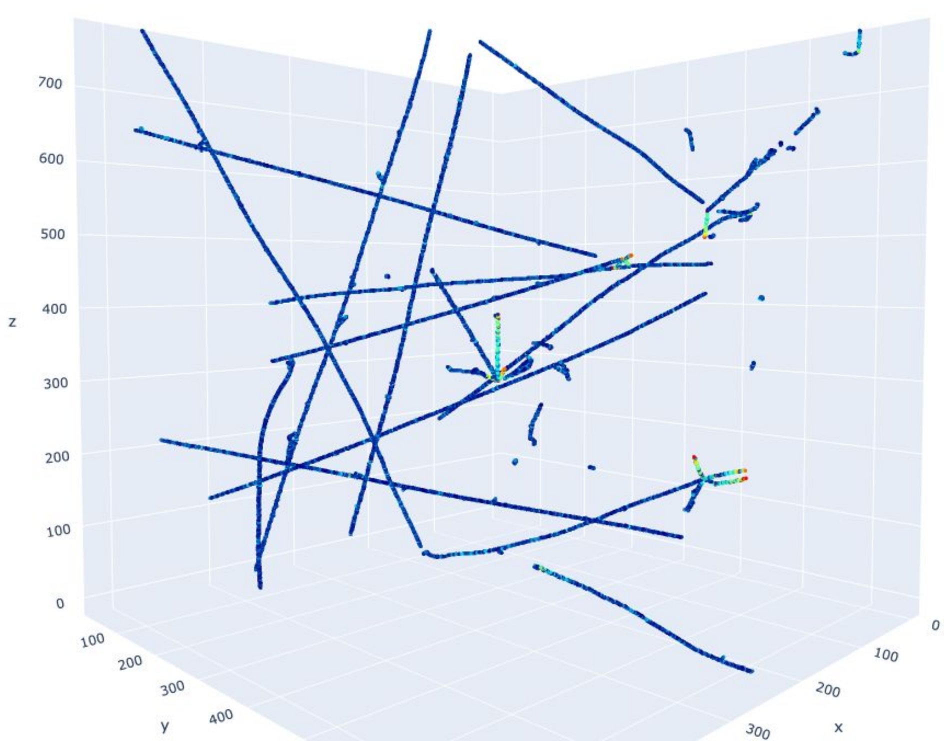
PRD 99 092001
arXiv:1808.07269

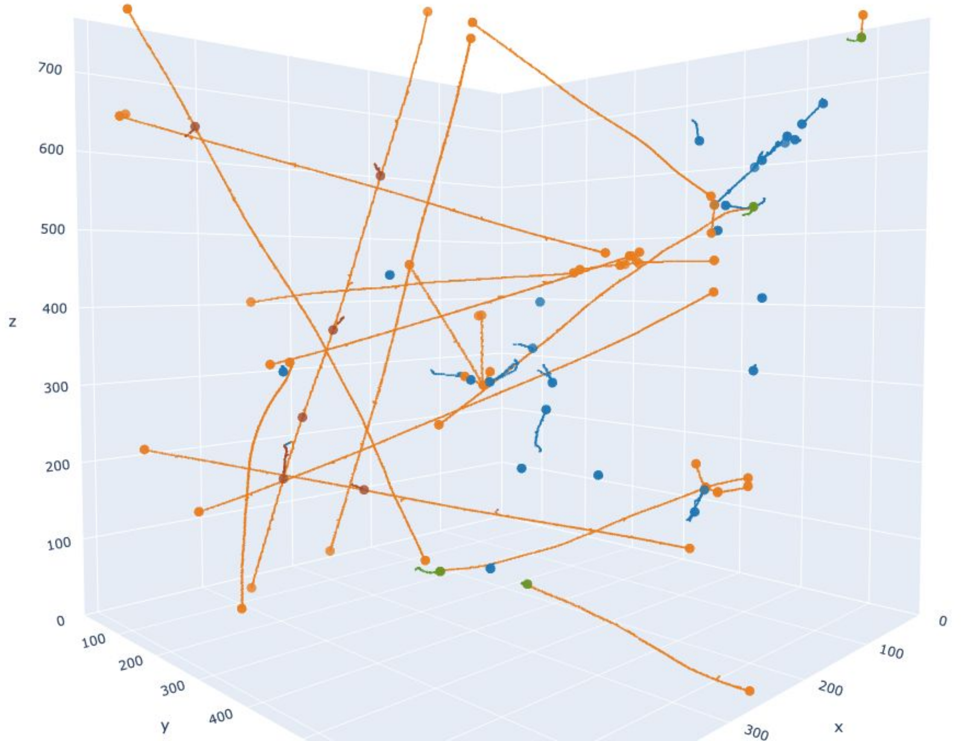
Machine Learning for Data Reconstruction

- **Goal**: high level abstract information (like image classification)
- **How**: design the algorithm = data transformation architecture that extracts a hierarchy of physically meaningful features (evidences)

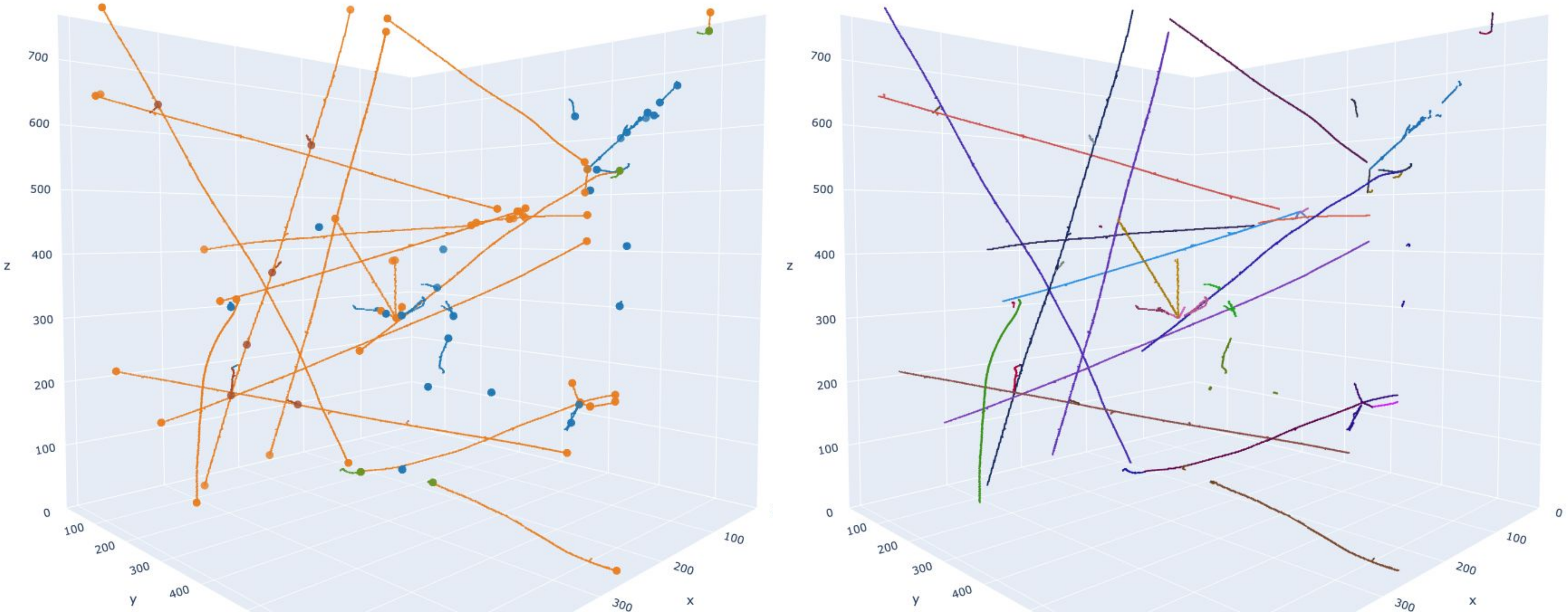


End-to-End Reconstruction for 3D Imaging Detector

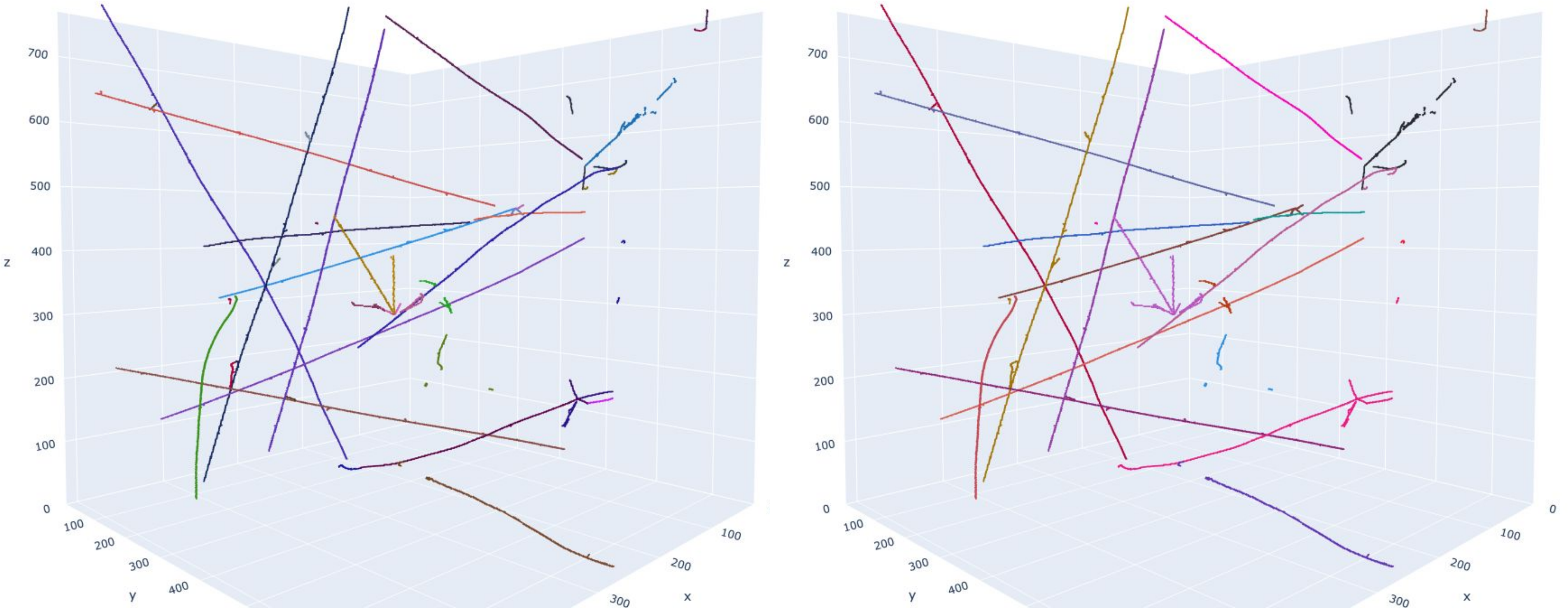




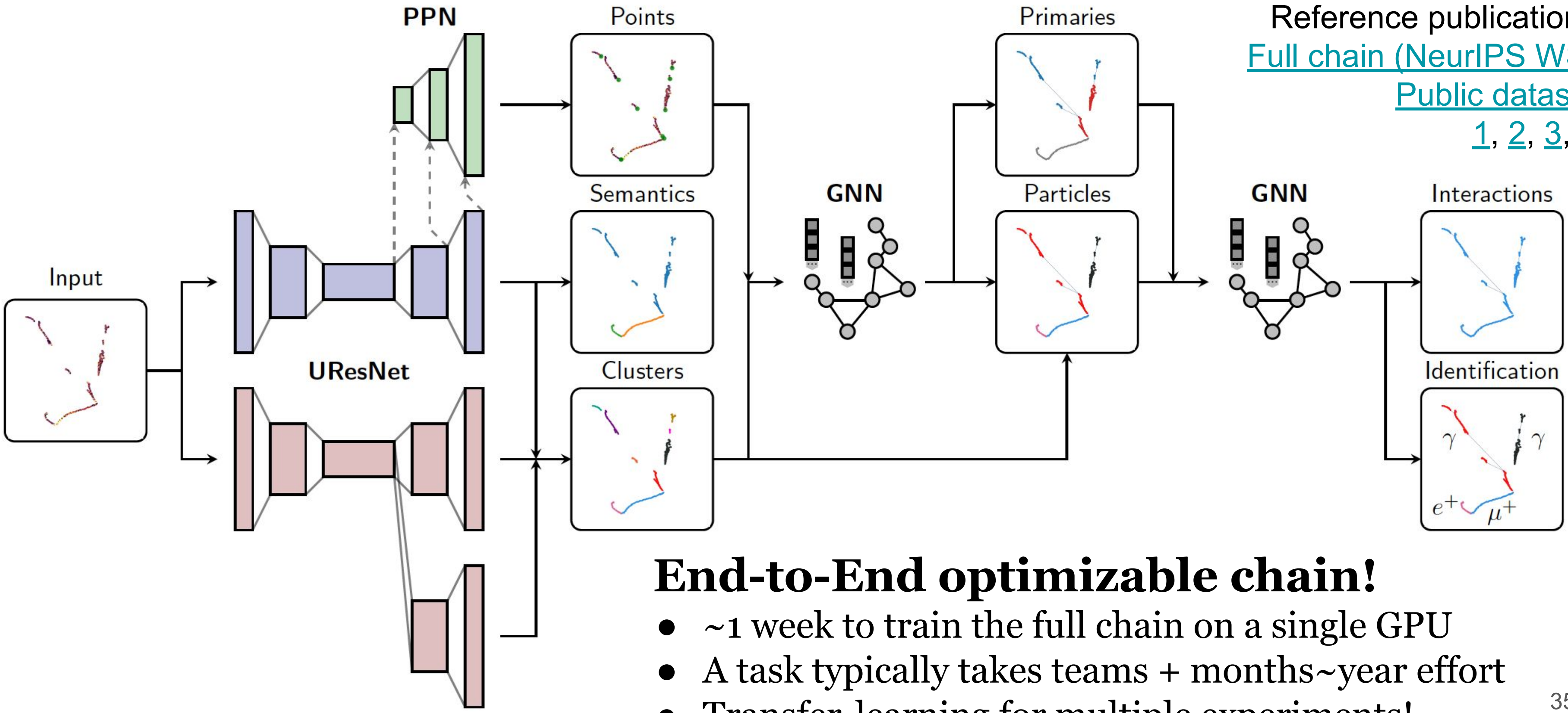
End-to-End Reconstruction for 3D Imaging Detector



End-to-End Reconstruction for 3D Imaging Detector



Deep Neural Network for Data Reconstruction

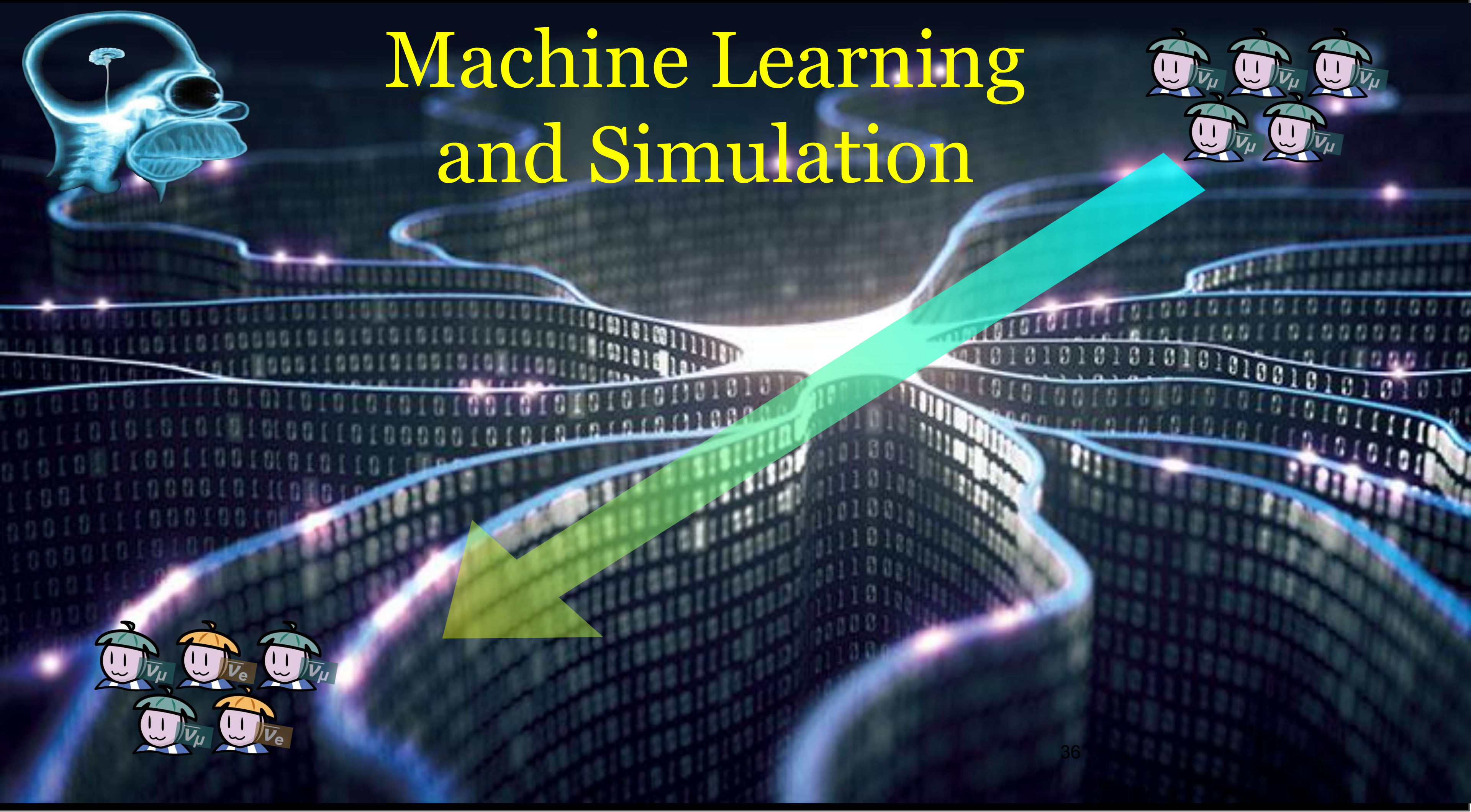


Reference publications
Full chain (NeurIPS WS)
Public dataset
1, 2, 3, 4

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!

Machine Learning and Simulation



Next Step: Innovative Simulator

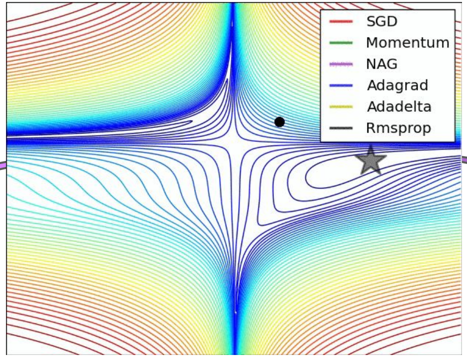
Recent success in machine learning … much are backed by **deep learning**

Machine Learning in Neutrino Physics & HEP

Next Step: Innovative Simulator

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

Next Step: Innovative Simulator

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

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

… which is enabled by computing hardwares & differentiable programming

Yann LeCun

January 5, 2018

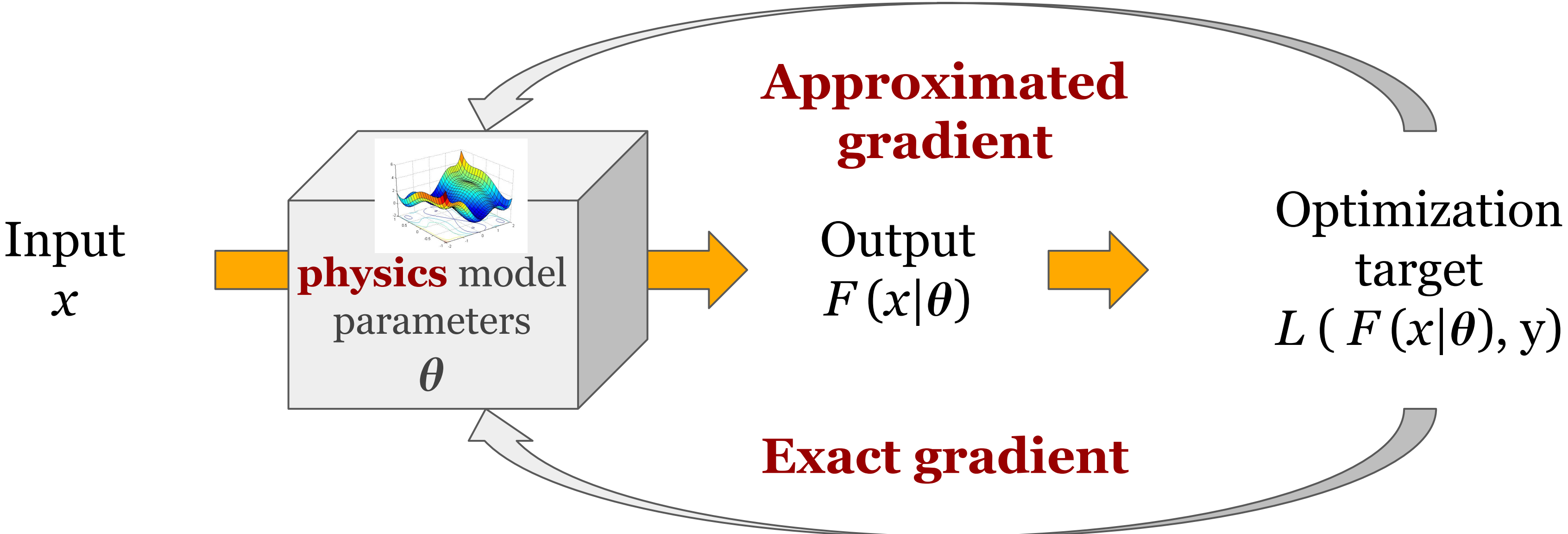
OK, Deep Learning has outlived its usefulness as a buzz-phrase.
Deep Learning est mort. Vive Differentiable Programming!

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

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

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

… which is enabled by computing hardwares & differentiable programming



Next Step: Innovative Simulator

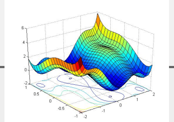
Recent success in machine learning … much are backed by **deep learning**

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

… which is enabled by computing hardwares & differentiable programming

Approximated gradient

Input
x



physics model parameters
θ

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

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

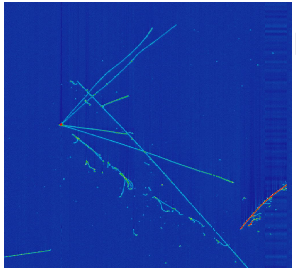
Exact gradient

Next Step: Innovative Simulator

Traditionally: experiments collect data, we infer physics from data

But: in this process, we draw likelihood using simulation (physics models)

"**Reconstruction**" as a process of inferring
a high(er) level physics quantities from raw data.

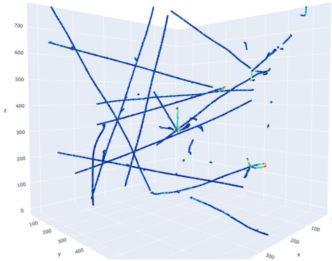


Detector Output
(ADC)

Reconstruction

Detector
physics model
parameters

Calibration



Reconstructed
(Calibrated dE/dX)

Physics knowledge
(models)

$$\mathscr{R}_{\text{ICARUS}} = \frac{A_B}{1+k_B\cdot(dE/dx)/\mathscr{E}}$$

$$Q = Q_0\exp(-v_{\text{drift}}t/\tau)$$

$$\sigma_t^2(t) \simeq \sigma_t^2(0) + \left(\frac{2D_L}{v_d^2}\right)t$$

...

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

Detector Simulation

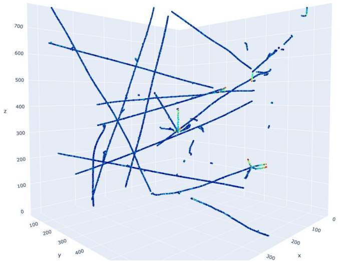
$$\mathscr{R}_{\text{ICARUS}} = \frac{A_B}{1 + k_B \cdot (dE/dx)/\mathscr{E}}$$

$$Q = Q_0 \exp(-v_{\text{drift}} t/\tau)$$

$$\sigma_t^2(t) \simeq \sigma_t^2(0) + \left(\frac{2D_L}{v_d^2}\right) t$$

...

Simulation Input
(true dE/dX)



"**Simulation**" takes in target physics information and convolve effects associated with the measurements (i.e. nuisance)

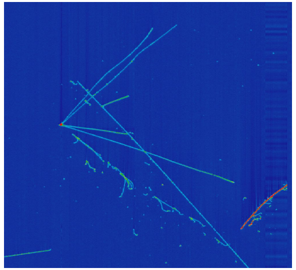
Physics knowledge
(models)

$$\mathscr{R}_{\text{ICARUS}} = \frac{A_B}{1 + k_B \cdot (dE/dx)/\mathscr{E}}$$

$$Q = Q_0 \exp(-v_{\text{drift}} t/\tau)$$

$$\sigma_t^2(t) \simeq \sigma_t^2(0) + \left(\frac{2D_L}{v_d^2}\right) t$$

...



Detector Output
(ADC)

Reconstruction

Detector physics model parameters

Calibration

Reconstructed
(dE/dX)

Machine Learning in Neutrino Physics & HEP

Next Step: Innovative Simulator

Detector Simulation

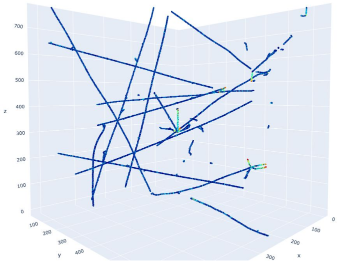
$$\mathscr{R}_{\text{ICARUS}} = \frac{A_B}{1 + k_B \cdot (dE/dx)/\mathscr{E}}$$

$$Q = Q_0 \exp(-v_{\text{drift}} t/\tau)$$

$$\sigma_t^2(t) \simeq \sigma_t^2(0) + \left(\frac{2D_L}{v_d^2}\right) t$$

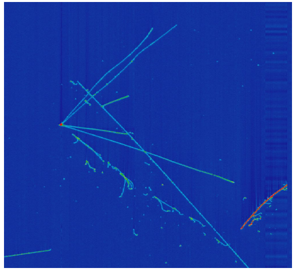
...

Simulation Input (true dE/dX)



Innovative Simulator

Develop techniques that can infer the physics from data (Simulation-based Inference)



Detector Output (ADC)

Reconstruction

Detector physics model parameters

Calibration

Reconstructed (dE/dX)

Physics knowledge (models)

$$\mathscr{R}_{\text{ICARUS}} = \frac{A_B}{1 + k_B \cdot (dE/dx)/\mathscr{E}}$$

$$Q = Q_0 \exp(-v_{\text{drift}} t/\tau)$$

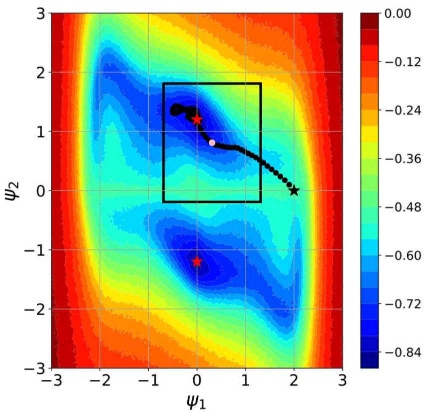
$$\sigma_t^2(t) \simeq \sigma_t^2(0) + \left(\frac{2D_L}{v_d^2}\right) t$$

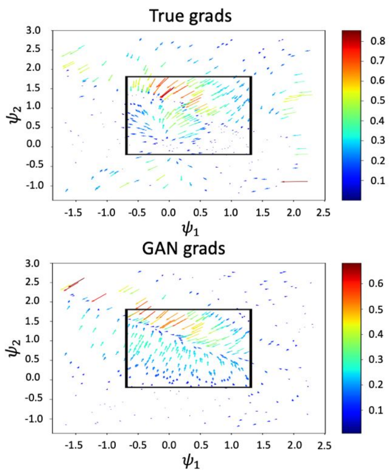
...

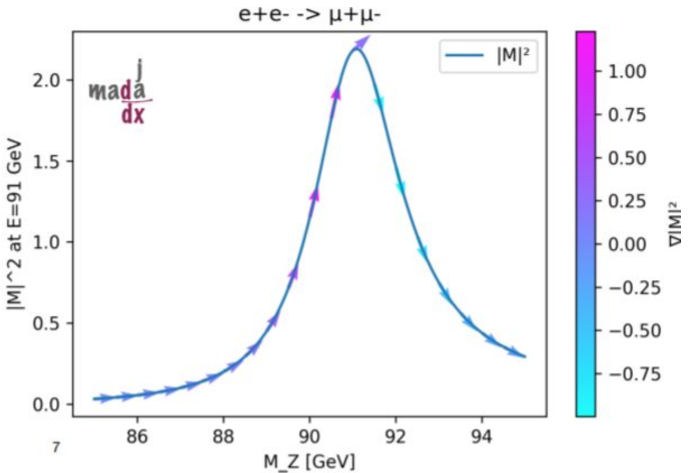
Next Step: Innovative Simulator

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.







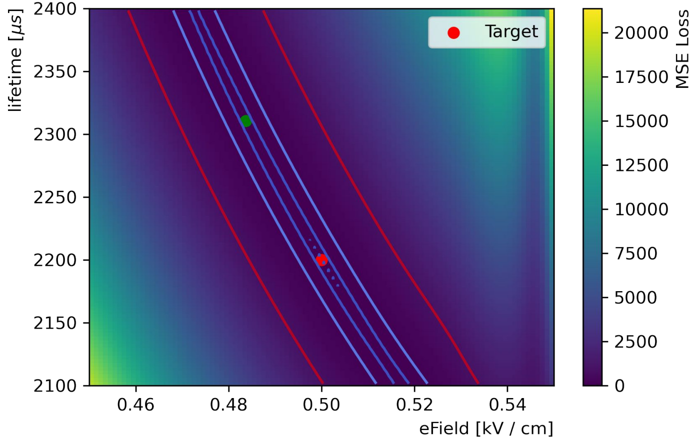
Left: surrogate model for magnet optimization

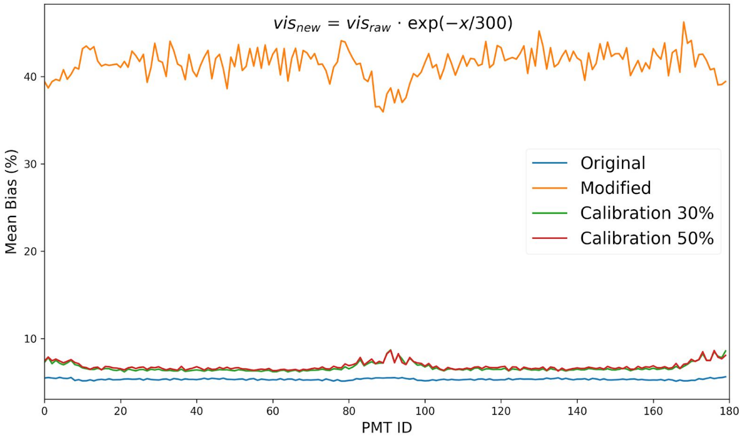
Right: differentiable matrix element calculation (MadJax)

Next Step: Innovative Simulator

Simulator-based Inference being explored across HEP frontiers

- Automation of detector model parameter inference (“calibration”)
- Detector design optimization for future experiments (DUNE/Hyper-K)
- Surrogate for event generator studies (100M - 1B event re-weighting)
- Exploitation of fully-differentiable analysis chain (sim+reco)





… wrapping up …

Final remarks

- Machine learning is spreading across applications
- Applications in neutrino physics used Computer Vision models for analysis and reconstruction
- Active research in physics inference methods and innovative simulation tools, HEP cross-frontier effort
- ML is interdisciplinary: please see the new PDG review for a wider scope of applications for other frontiers!

Closing

SLAC

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

Satellite targeted young folks for short technical talks, the main workshop invited collaborations for a summary. 20+40 = 60 talks all recorded :)

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

Interest? go check out slides and recorded talks!

... + start conversation to collaborate on ML application development with neutrino physicists :)

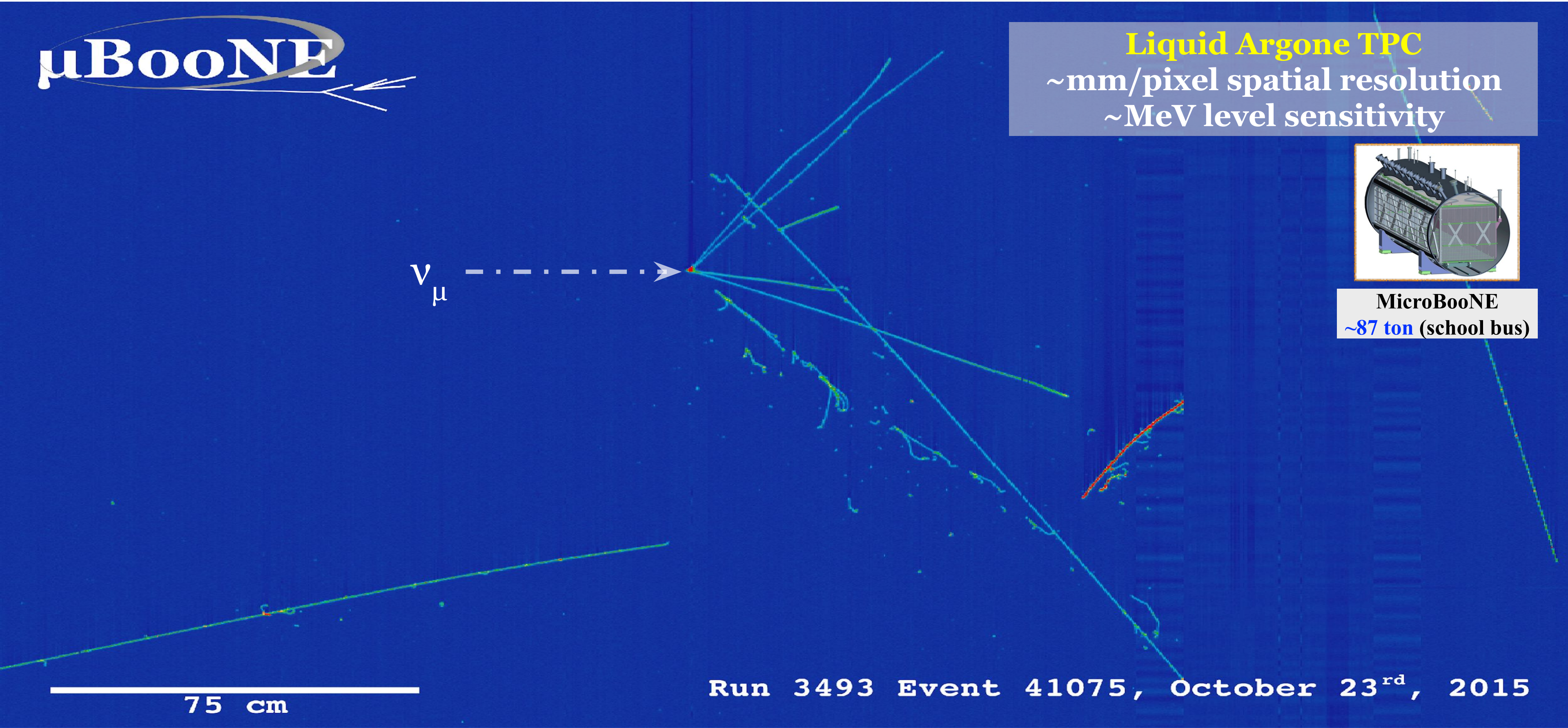
Planning to hold another in 2022!

Questions?

Back-ups

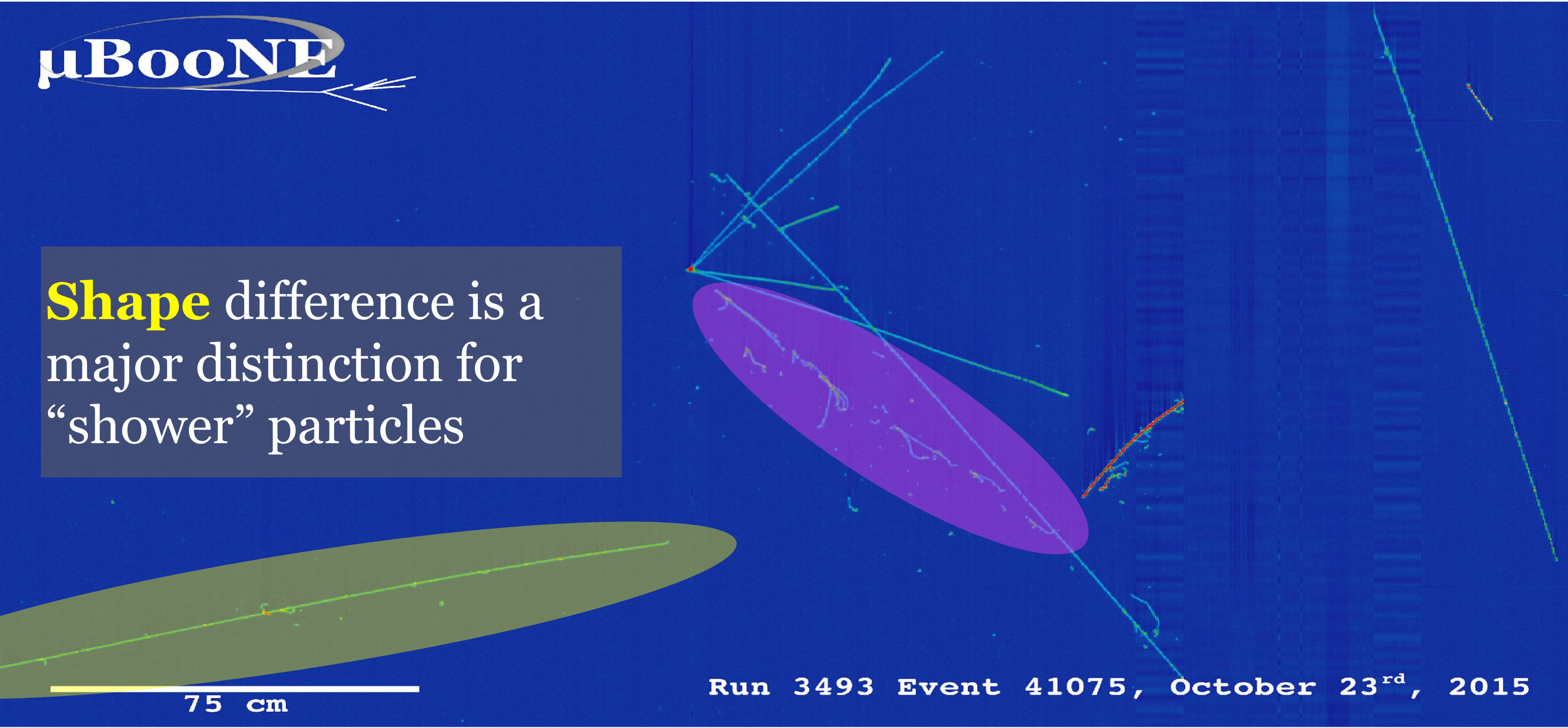
Machine Learning & Computer Vision in Neutrino Physics

Time Projection Chambers



Time Projection Chambers

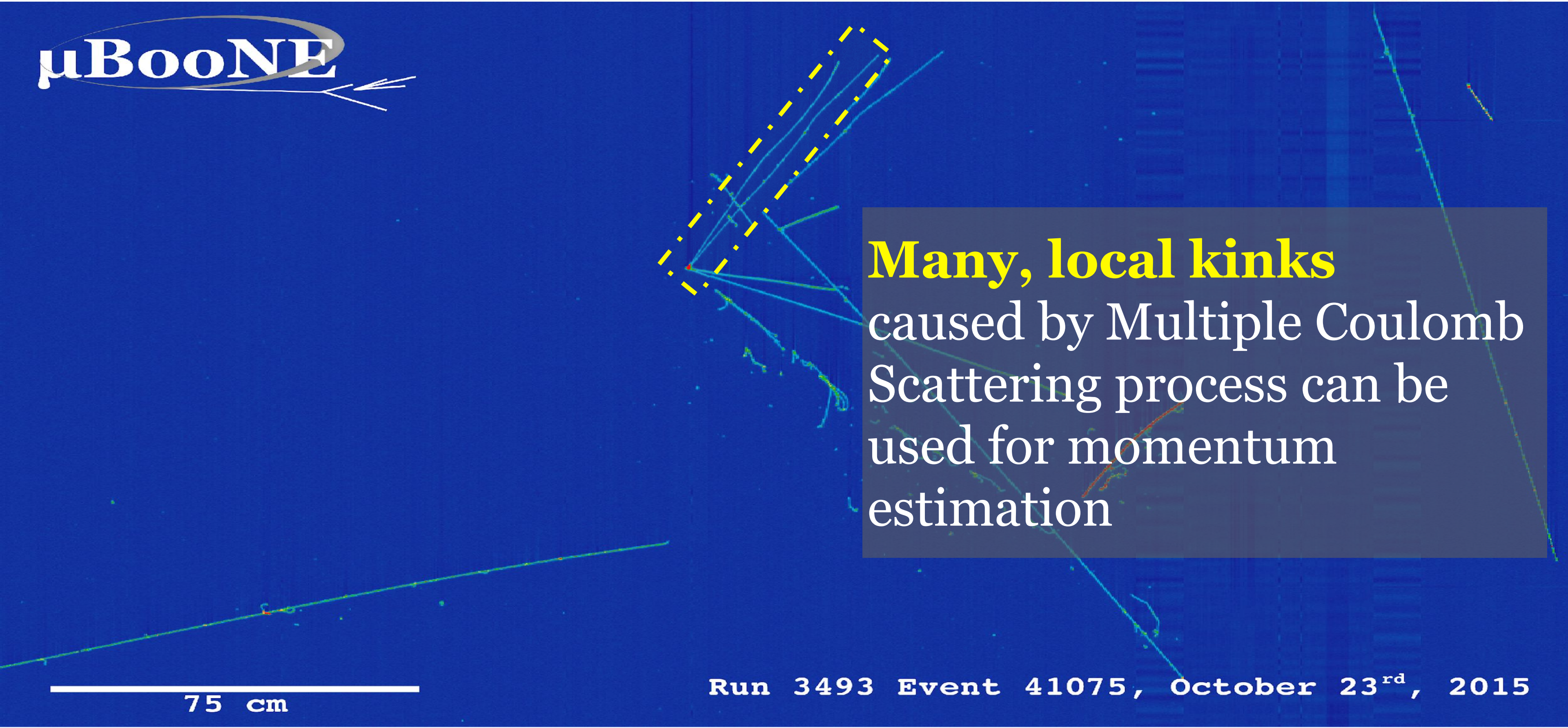
SLAC



Shape difference is a major distinction for "shower" particles

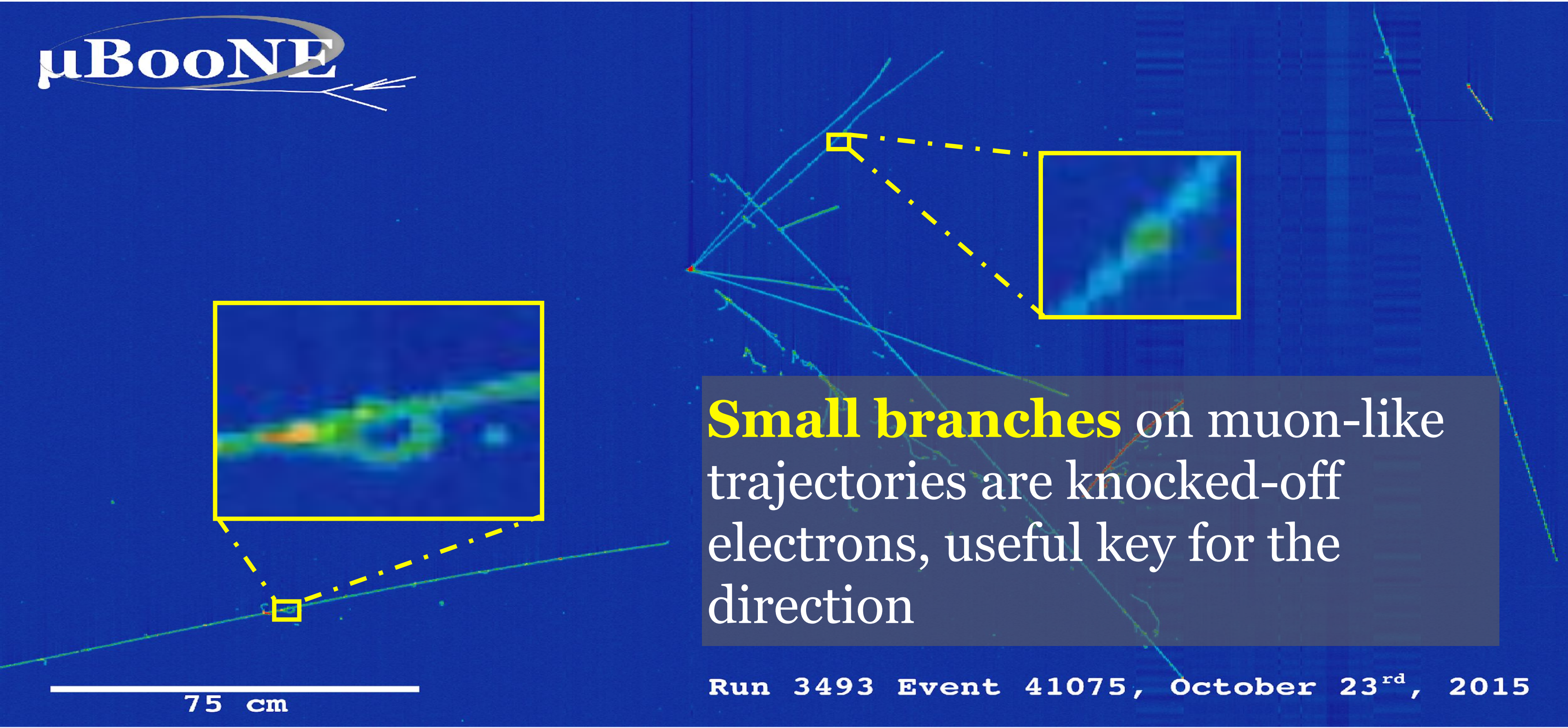
Machine Learning & Computer Vision in Neutrino Physics

Time Projection Chambers



Many, local kinks caused by Multiple Coulomb Scattering process can be used for momentum estimation

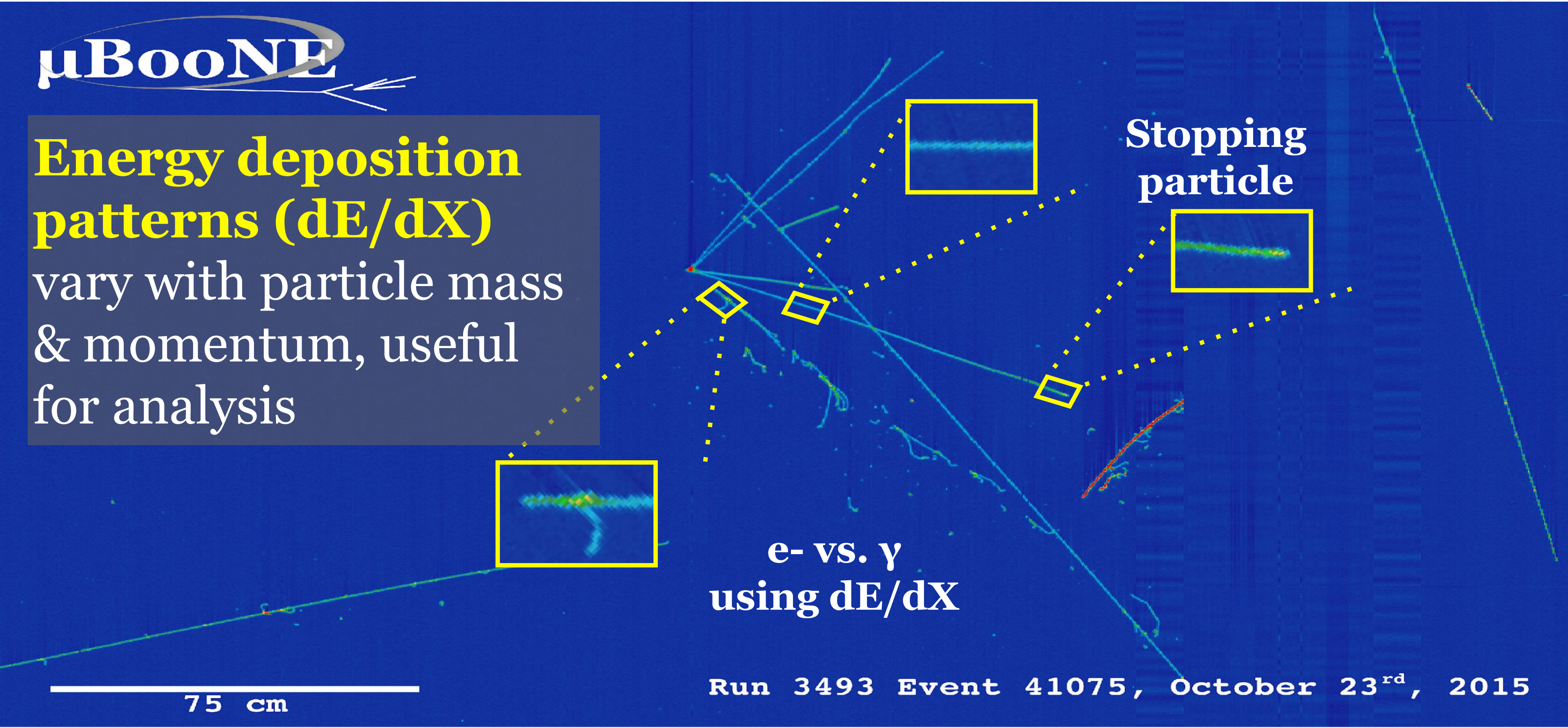
Machine Learning & Computer Vision in Neutrino Physics
Time Projection Chambers



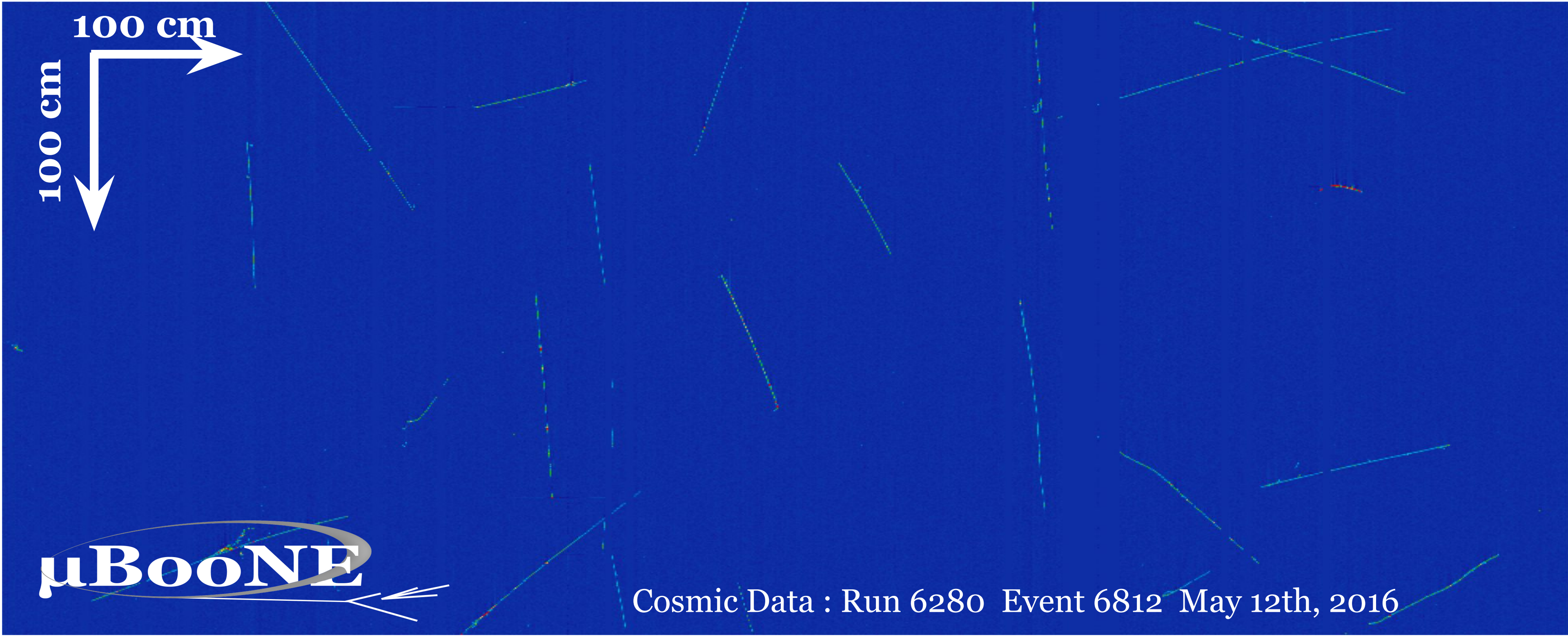
Small branches on muon-like trajectories are knocked-off electrons, useful key for the direction

Machine Learning & Computer Vision in Neutrino Physics
Time Projection Chambers

SLAC

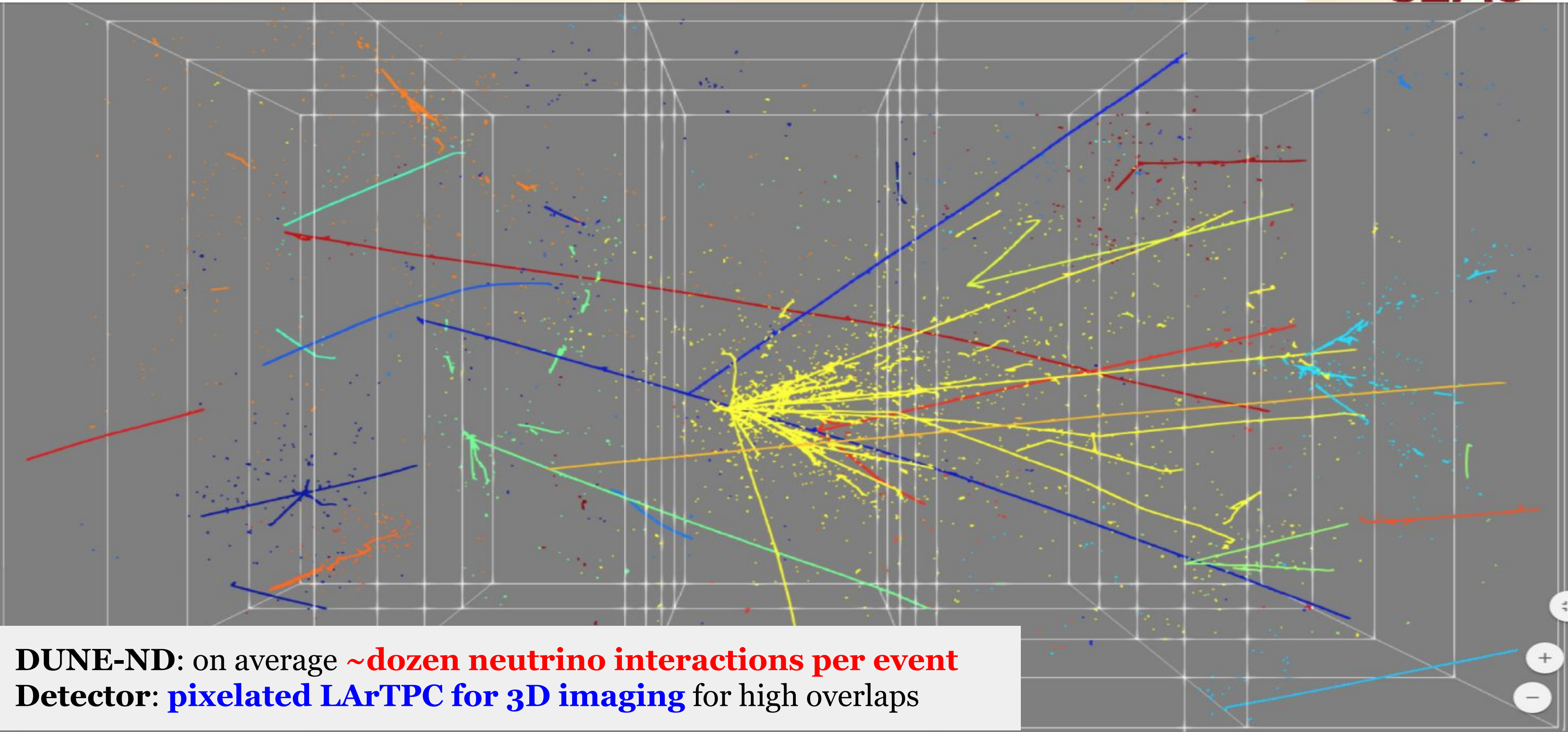


Nope :) In this detector, <1% beam neutrino interacts



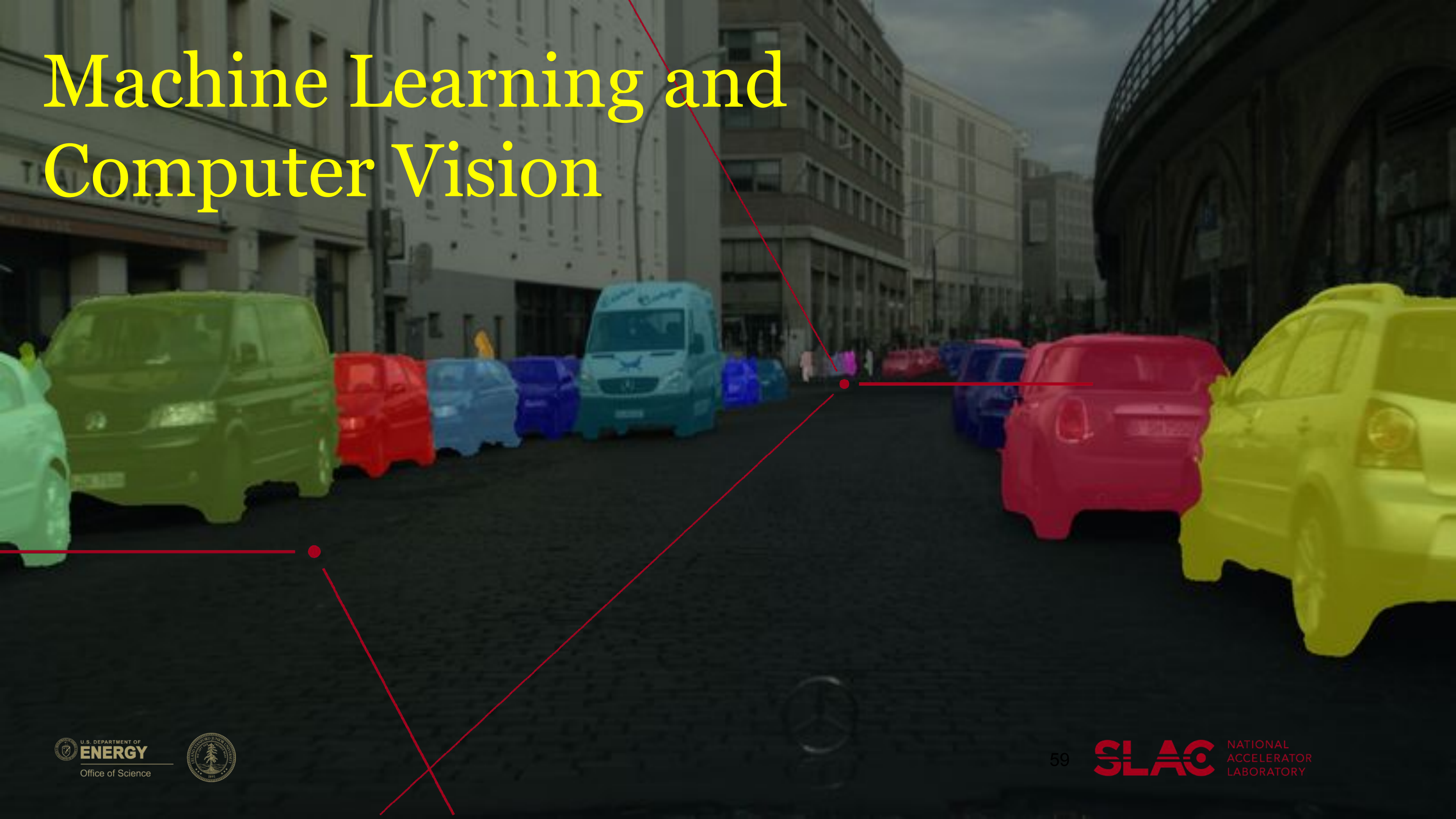
Machine Learning & Computer Vision in Neutrino Physics
Time Projection Chambers (3D ones)

SLAC



DUNE-ND: on average **~dozen neutrino interactions per event**
Detector: **pixelated LArTPC for 3D imaging** for high overlaps

Machine Learning and Computer Vision



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

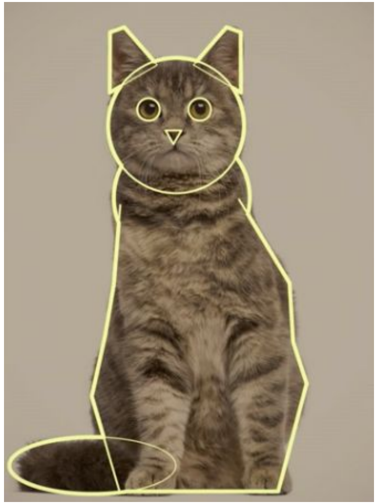
How to write an algorithm to identify a cat?

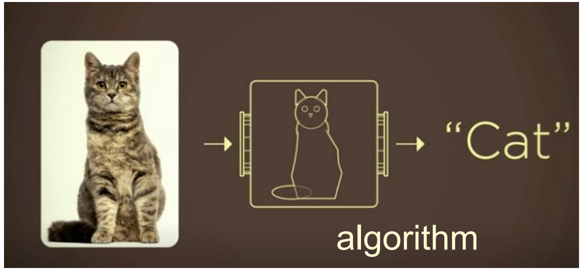
... very hard task ...

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 (or, a neutrino) = collection of certain shapes

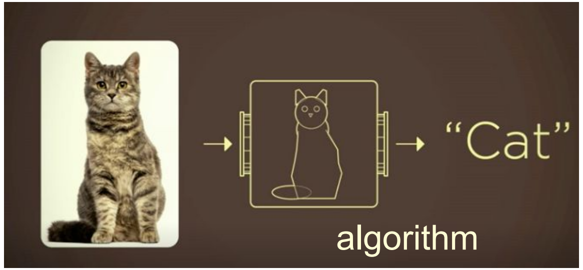
Images courtesy of Fei Fei Li's TED talk

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)

Stretching cat (Nuclear Physics)



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.

"Machine learning"

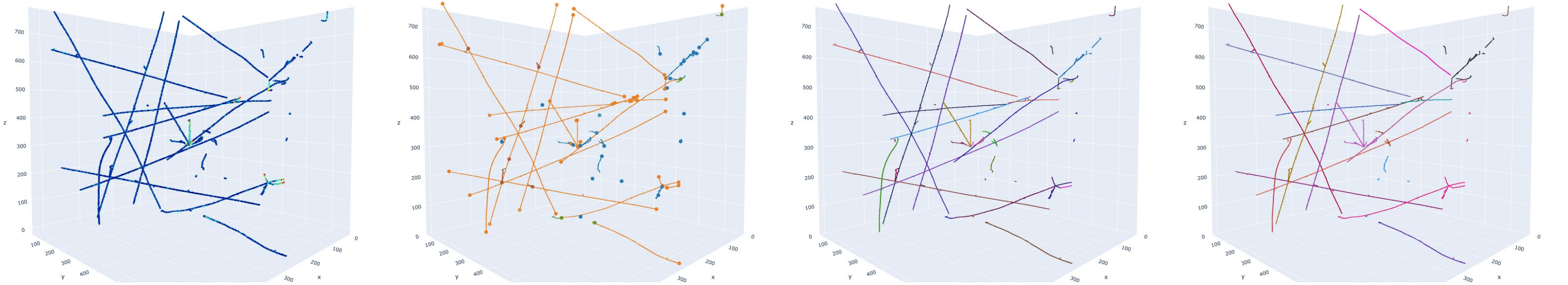
- Model instead of explicit programming
- Automatization of steps 2-4
- Multi-task optimization possible (step 5)

Next: what kind of ML algorithms?

Machine Learning for Data Reconstruction

- **Goal**: high level abstract information (like image classification)
- **How**: design the algorithm = data transformation architecture that extracts a hierarchy of physically meaningful features

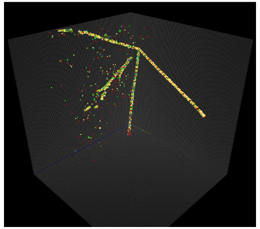
Input | Pixel Feature | Pixel Clustering | Particle Clustering

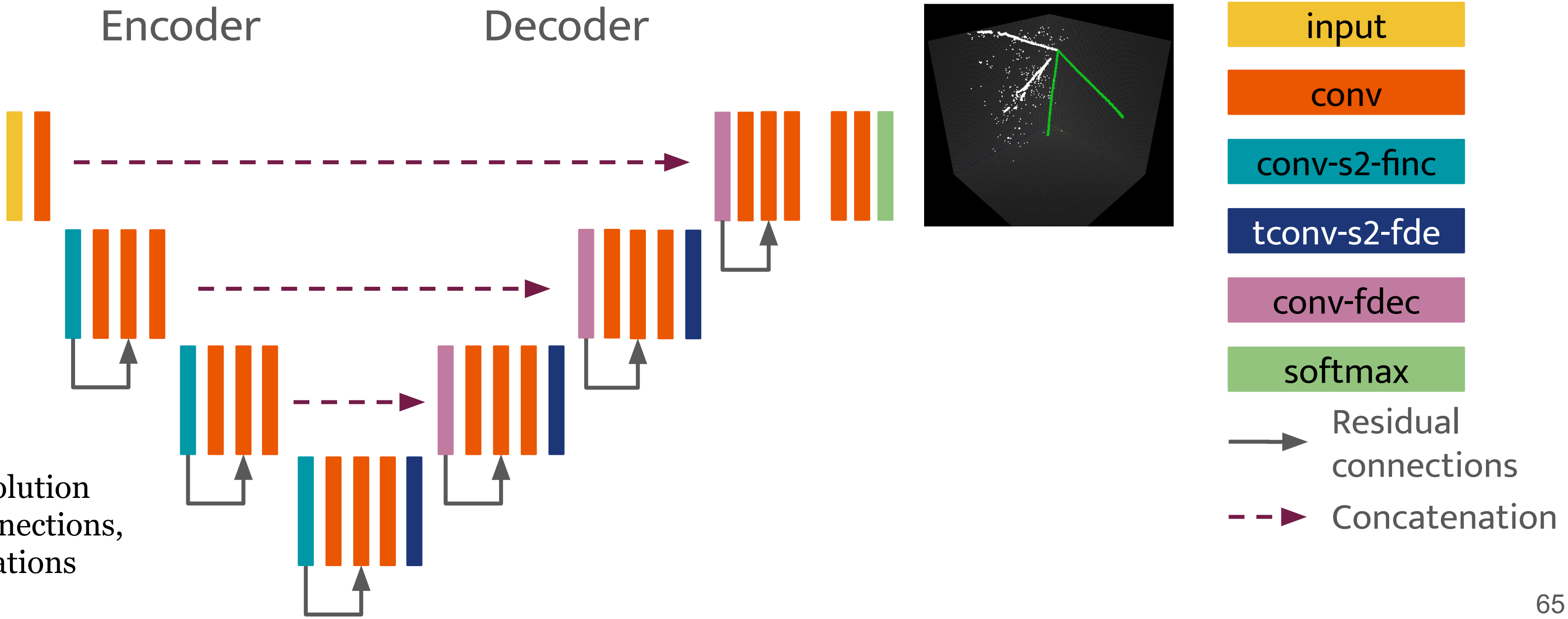


The Rest: describe the chain for 3D

ML-based Neutrino Data Reconstruction Chain
Stage 1-a: Pixel Feature Extraction + Scalablility

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





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

Image credit: Laura Domine @ Stanford

"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

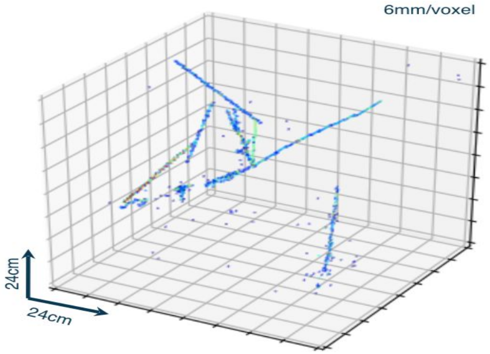
"Applying CNN" is simple, but is it scalable for us?

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

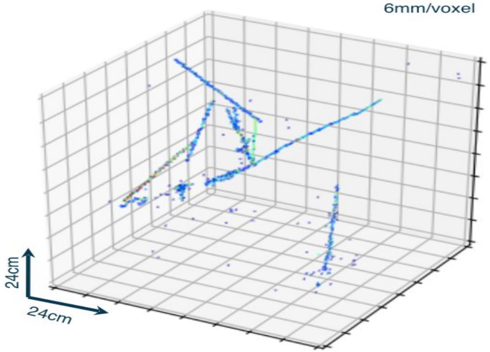
SLAC

"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

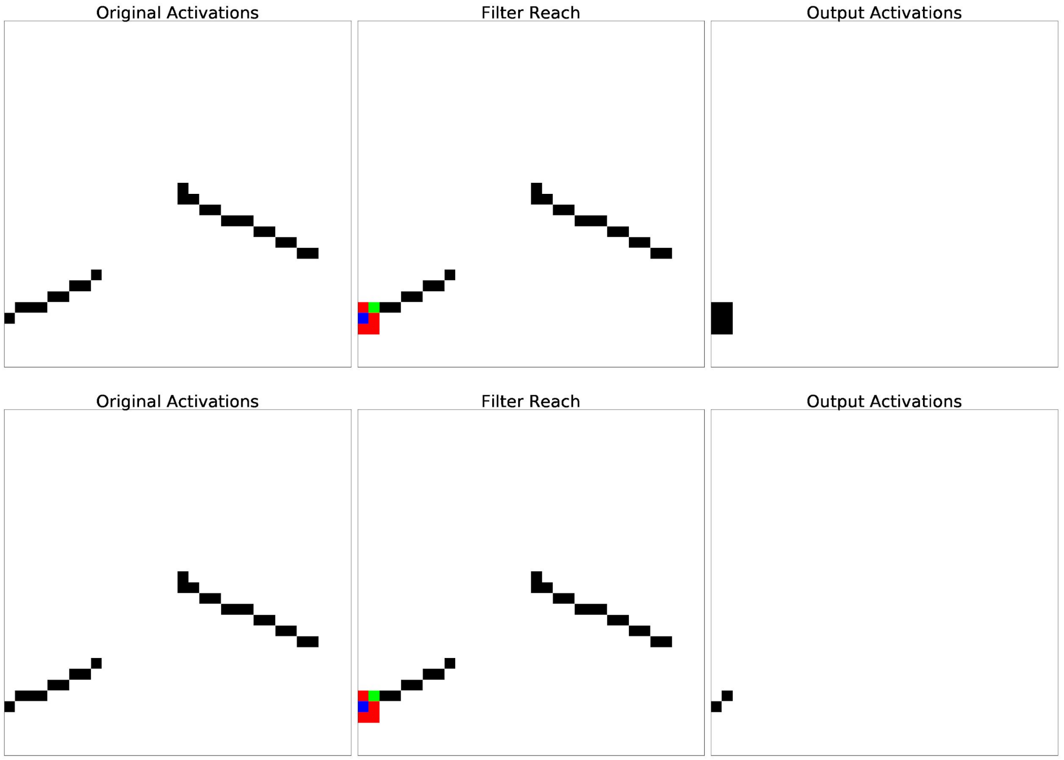
- **Scalability for larger detectors**
  - Computation cost increases linearly with the volume
  - But the number of non-zero pixels does not

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



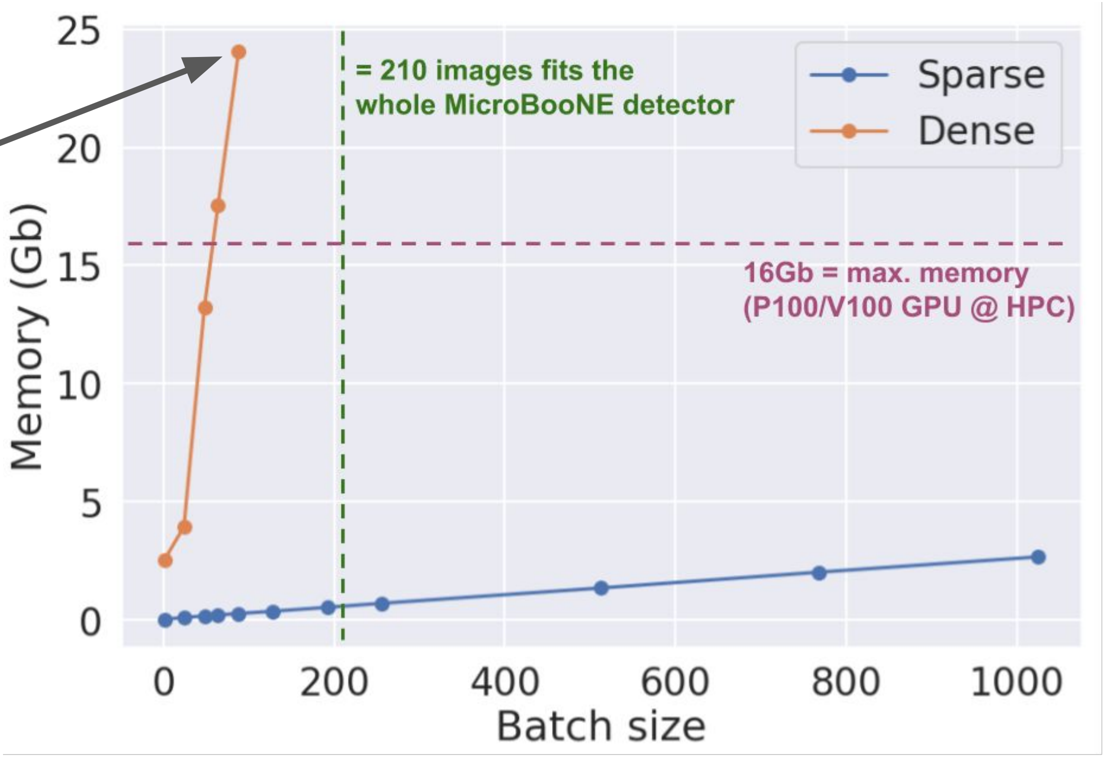
ML-based Neutrino Data Reconstruction Chain

Stage 1-a: Pixel Feature Extraction + Scalablility

SLAC

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-based Neutrino Data Reconstruction Chain
Stage 1-a: Pixel Feature Extraction + Scalablility

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

Sparse Sub-manifold Convolutional NN

- **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

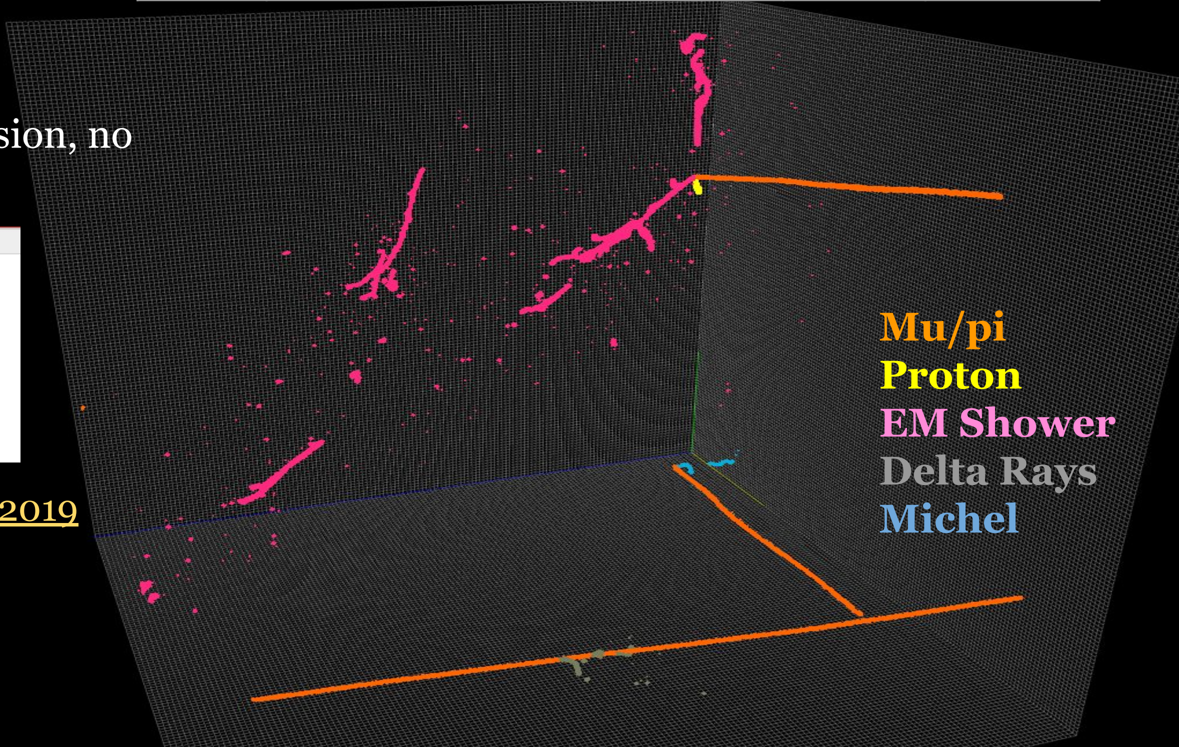
Laura Dominé, Kazuhiro Terao

(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

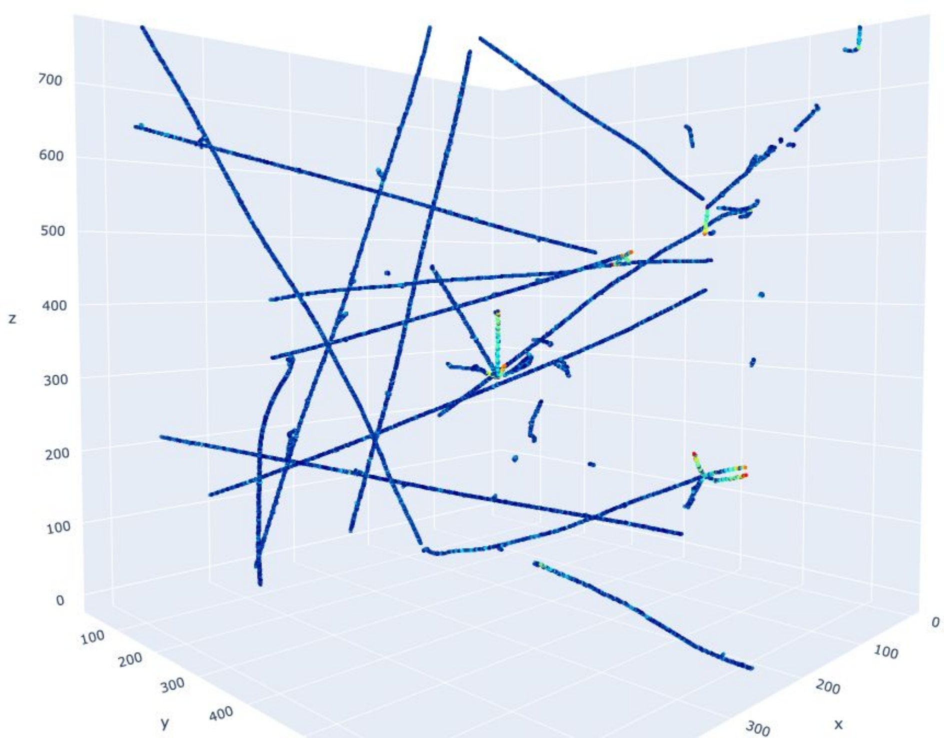
PhysRevD.102.012005 presented @ ACAT 2019

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

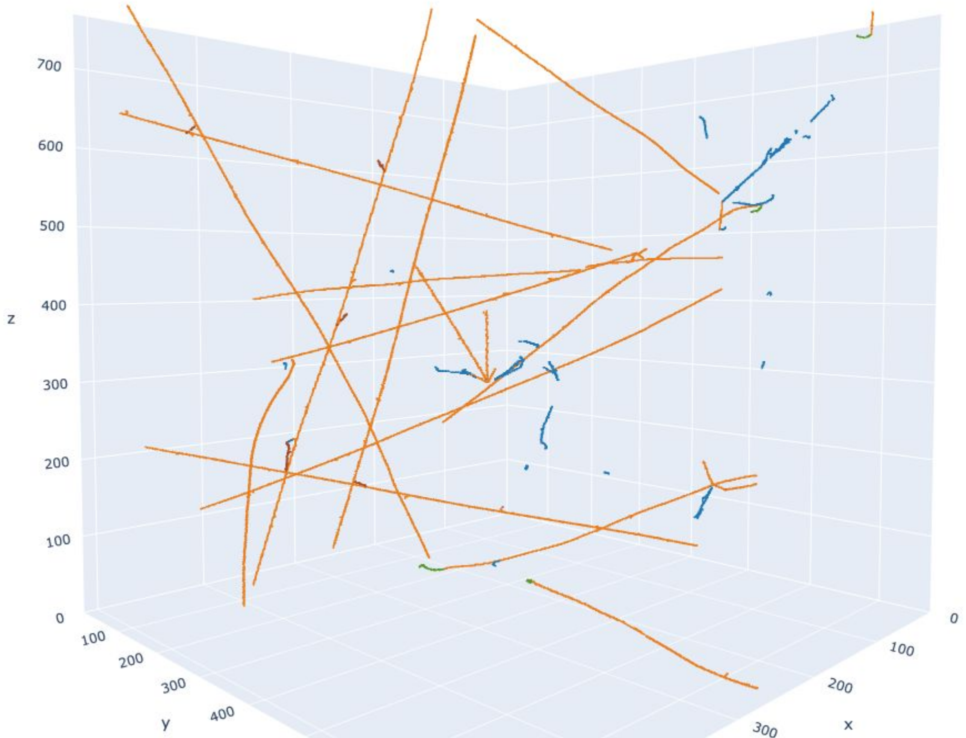


ML-based Neutrino Data Reconstruction Chain
Stage 1-a: input & output

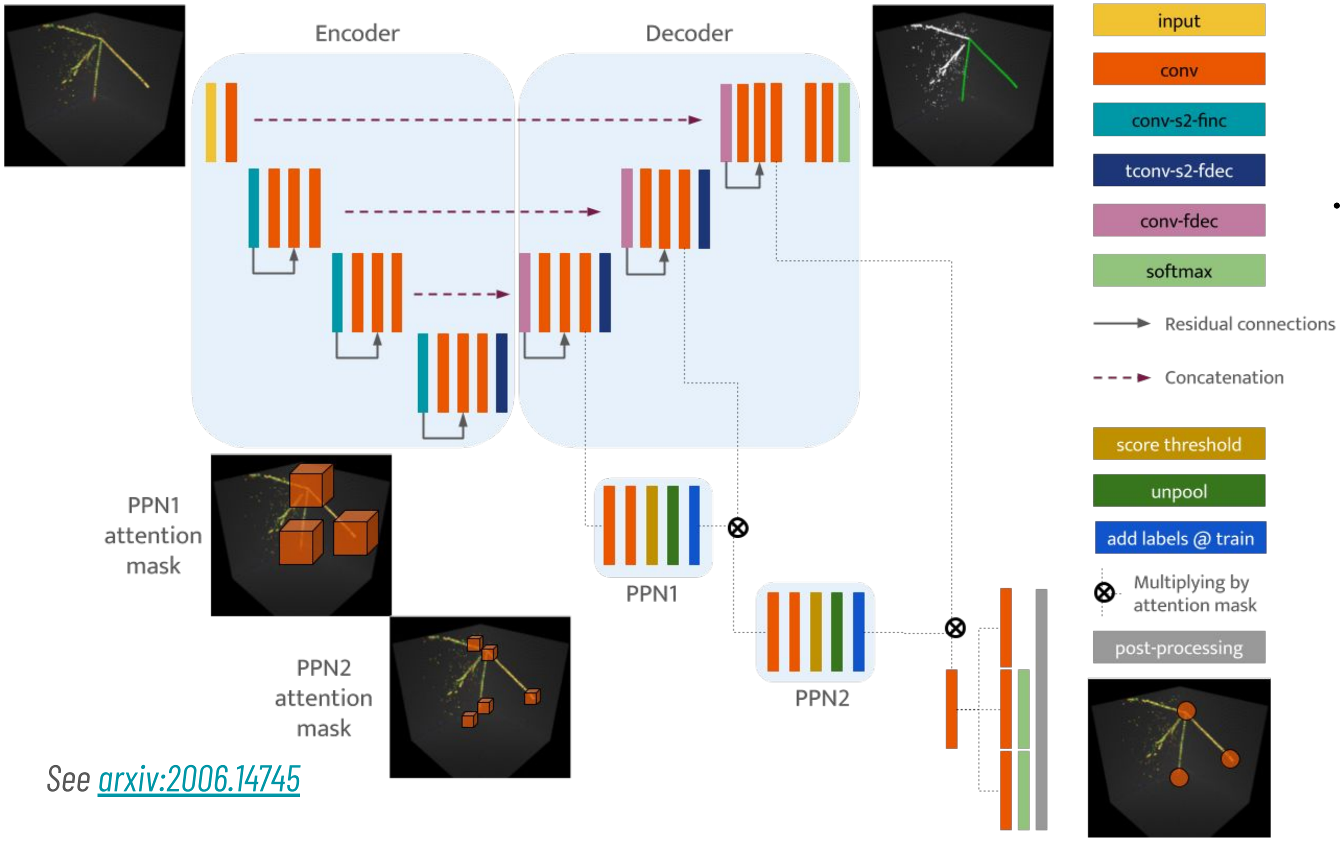
Stage 1-a Input



Stage 1-a Output



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



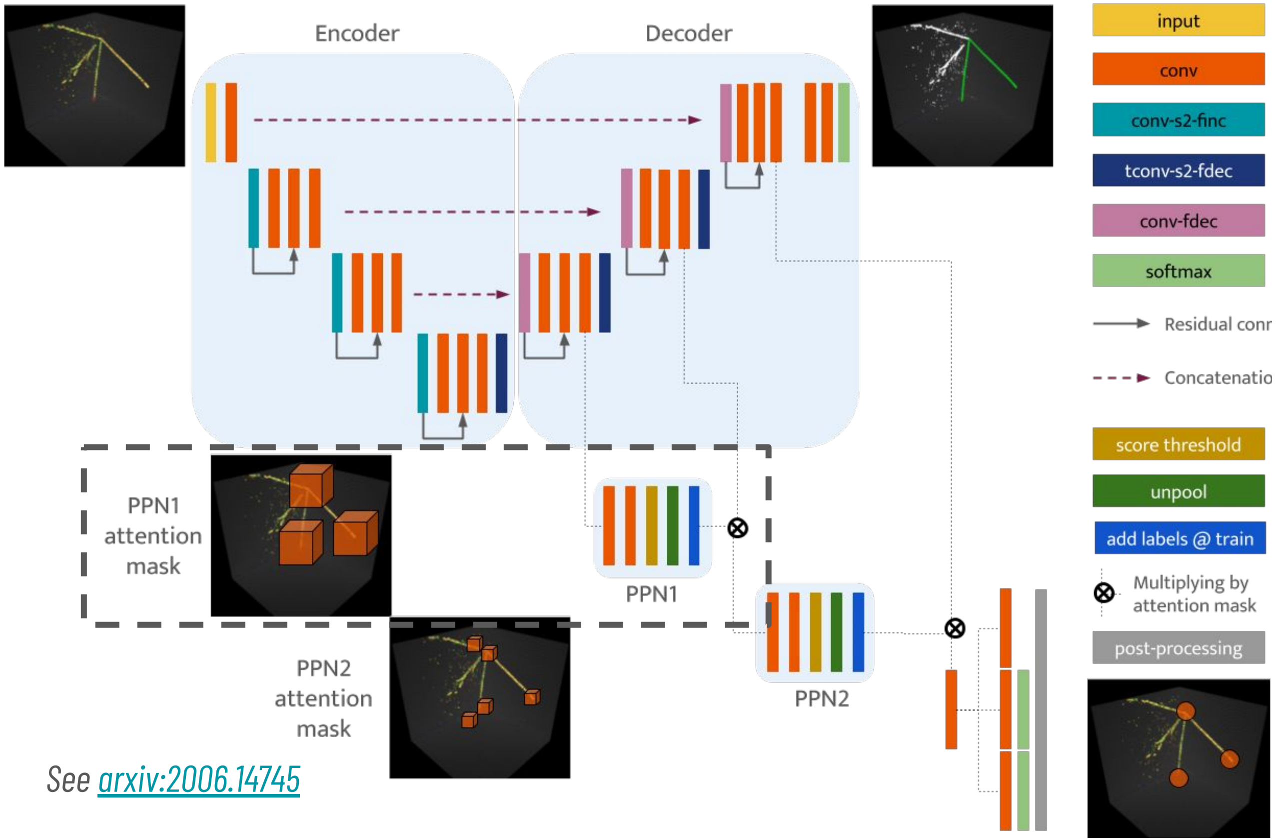
Point Proposal Network (PPN)

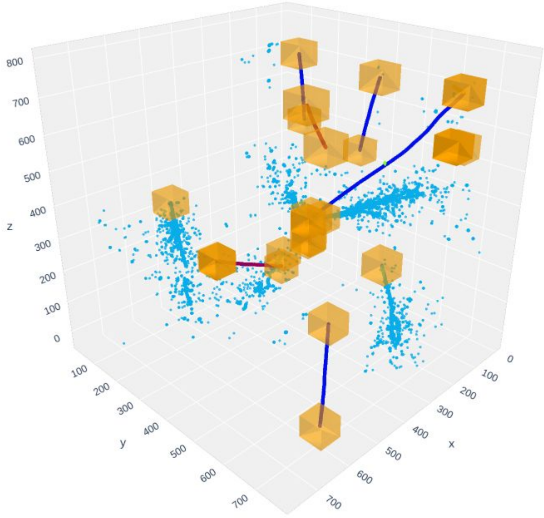
... extension of U-ResNet with 3 CNN blocks

See arxiv:2006.14745

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

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



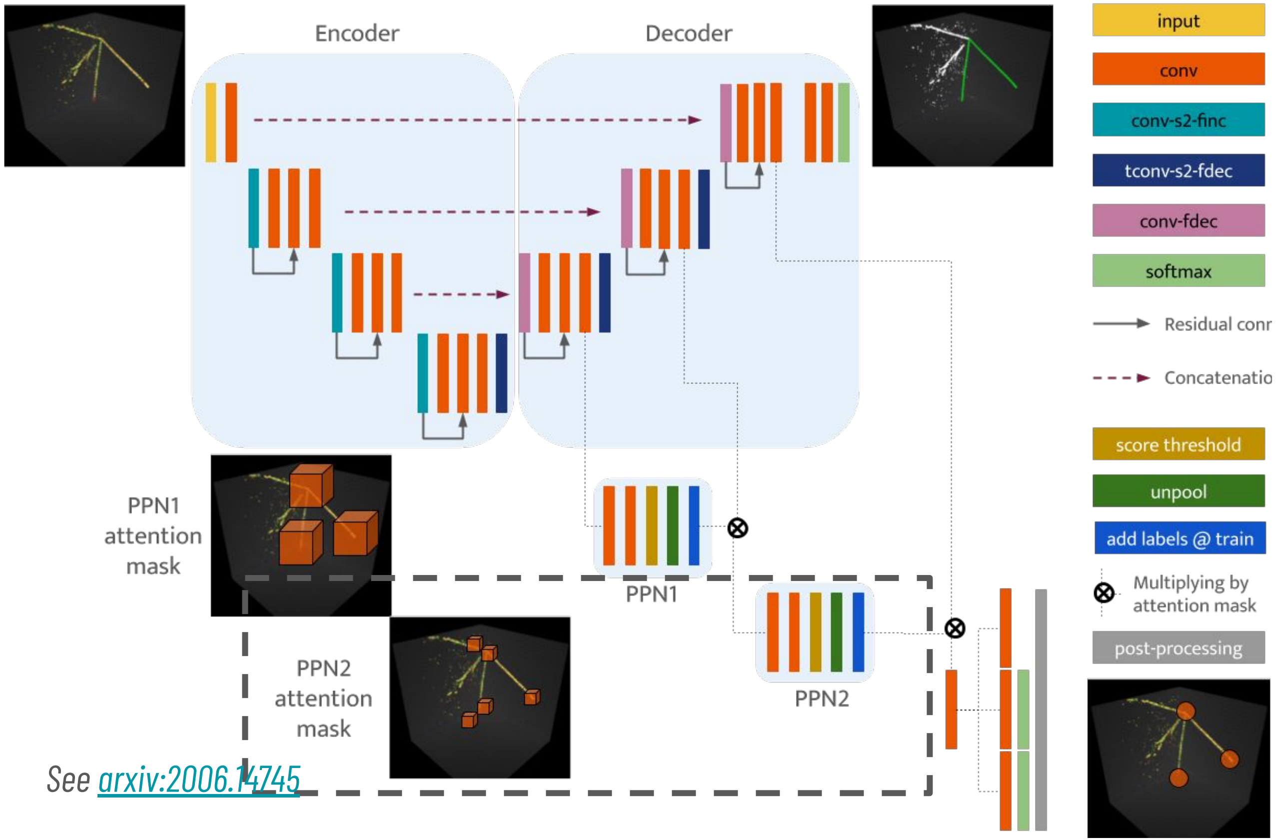


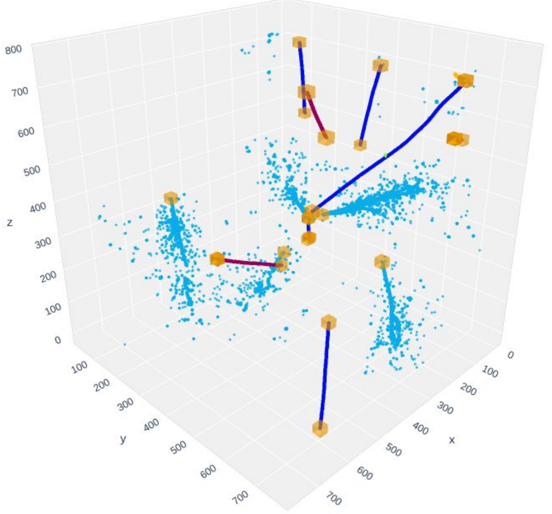
PPN1 generates an attention mask at the lowest resolution

See arxiv:2006.14745

ML-based Neutrino Data Reconstruction Chain

Stage 1-b: Particle Endpoint Prediction





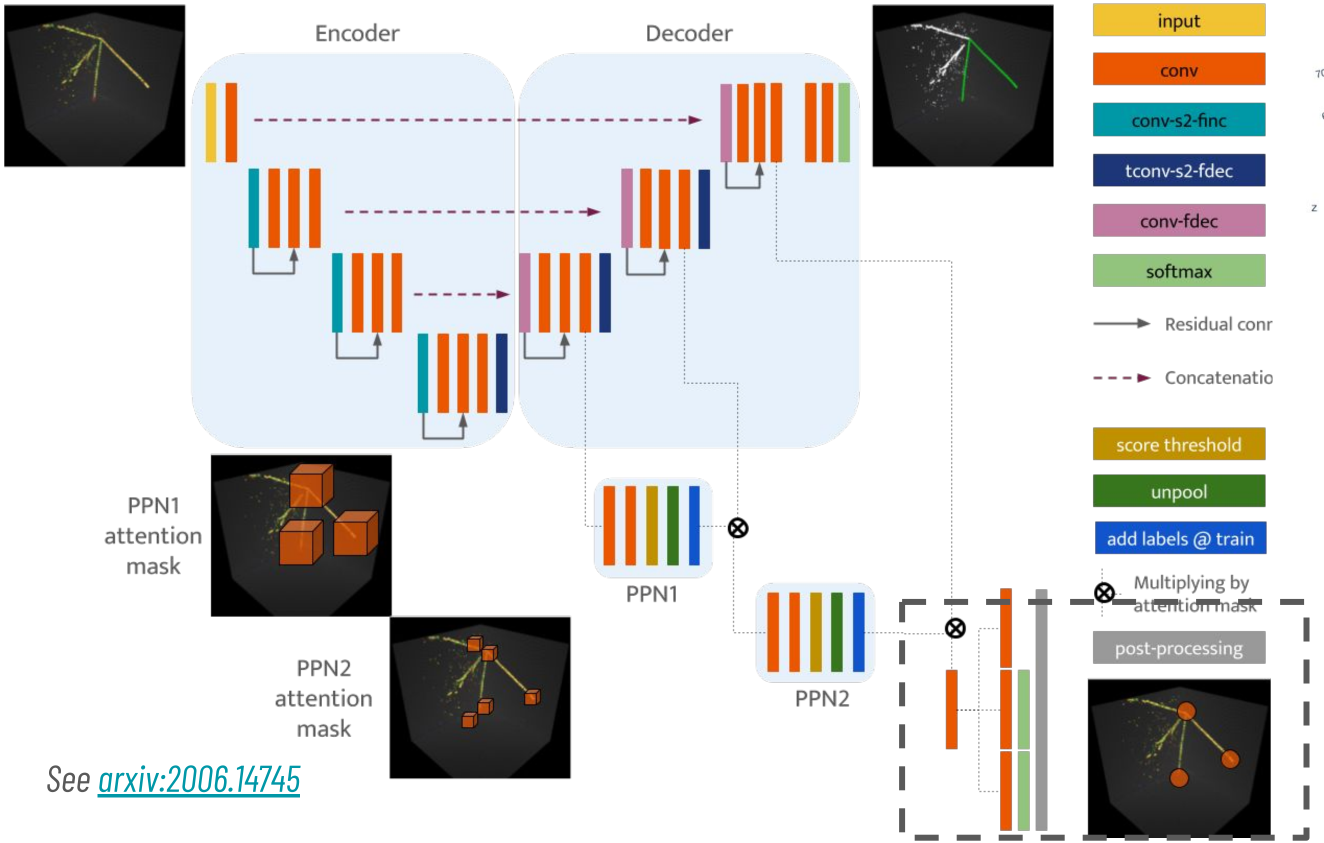
PPN2 generates an attention mask at the intermediate resolution

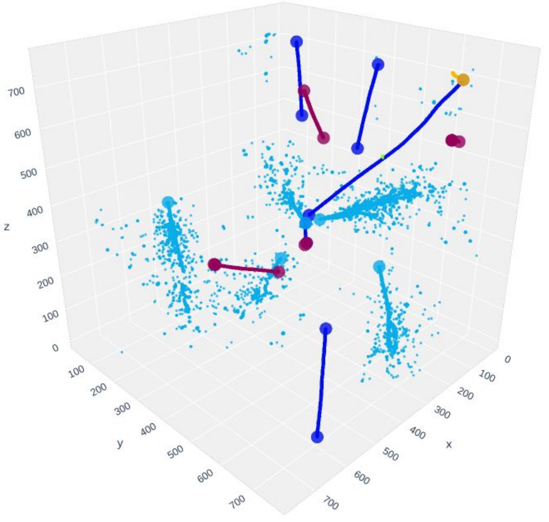
See arxiv:2006.14745

ML-based Neutrino Data Reconstruction Chain

Stage 1-b: Particle Endpoint Prediction

SLAC





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

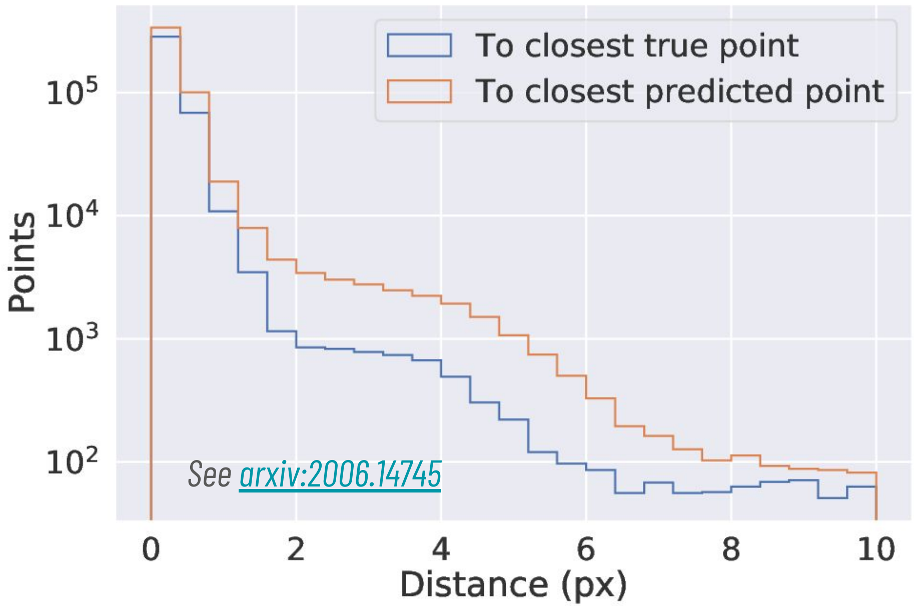
See arxiv:2006.14745

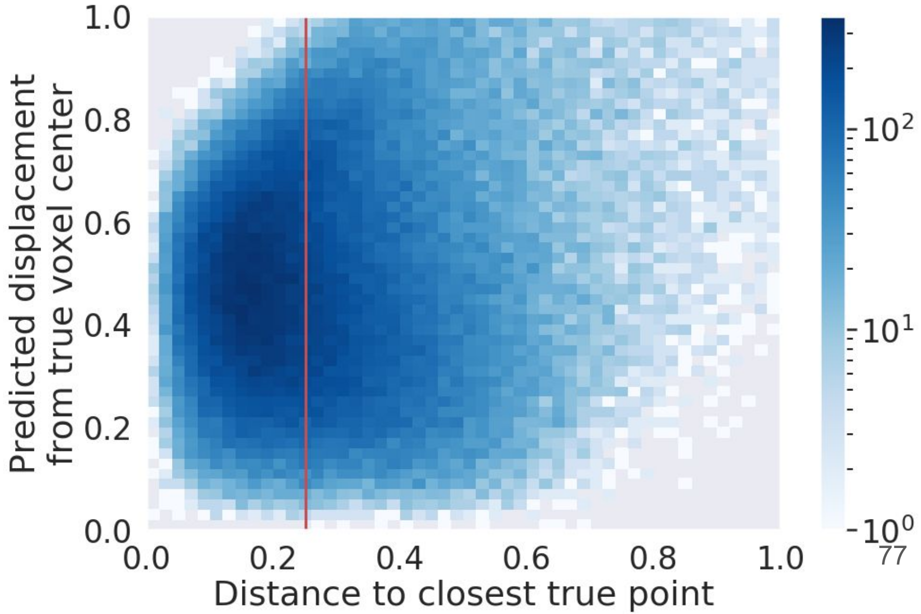
ML-based Neutrino Data Reconstruction Chain
Stage 1-b: Particle Endpoint Prediction

SLAC

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

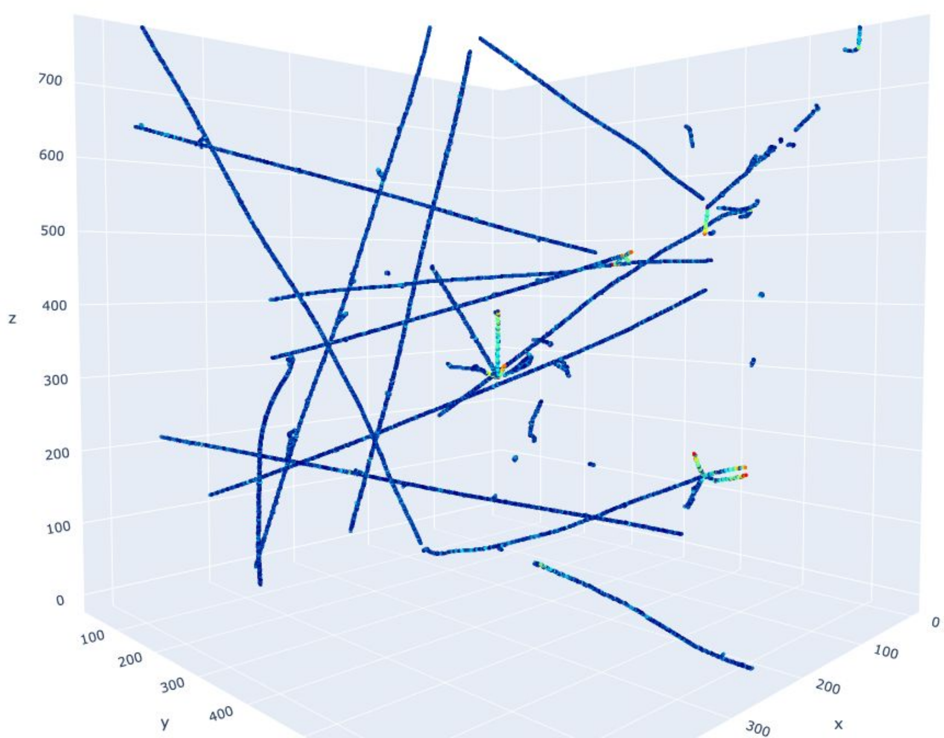




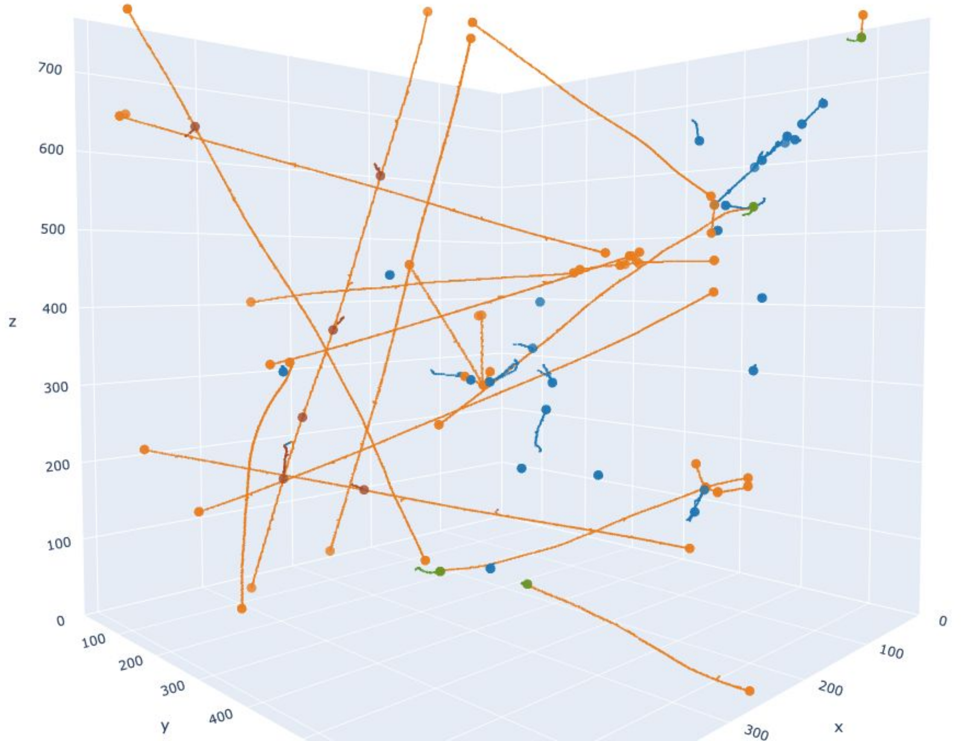
ML-based Neutrino Data Reconstruction Chain
Stage 1: input & output

SLAC

Stage 1 Input



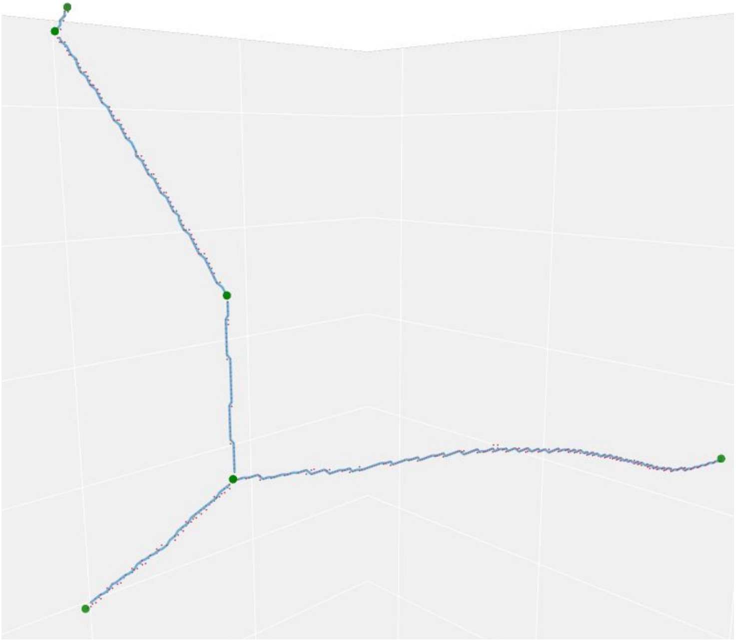
Stage 1 Output

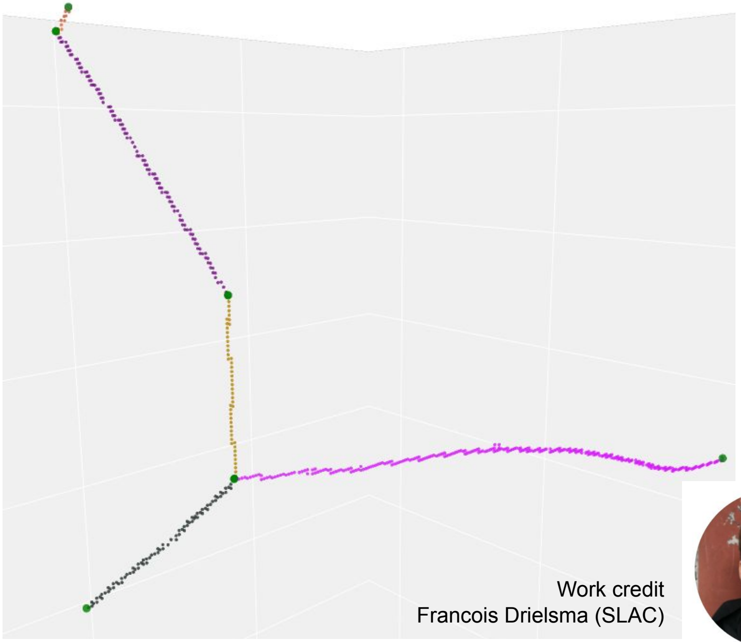


ML-based Neutrino Data Reconstruction Chain
Stage 2-a: Dense Pixel Clustering

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)

Learnable approach: 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

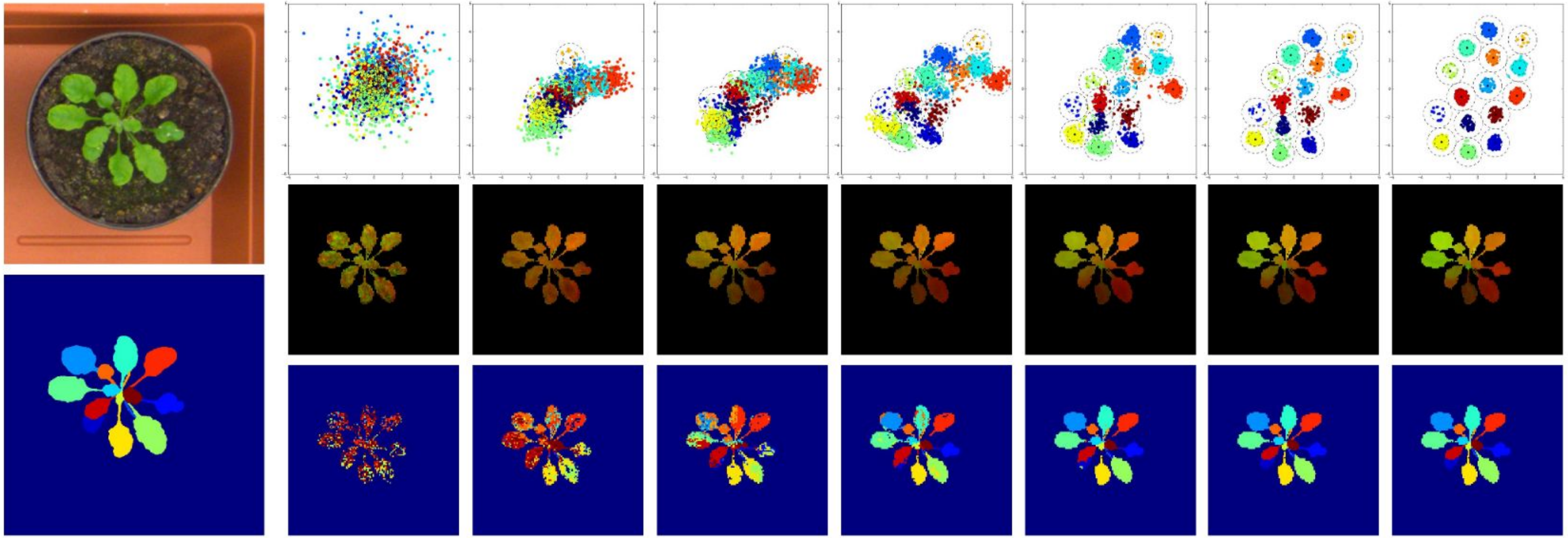
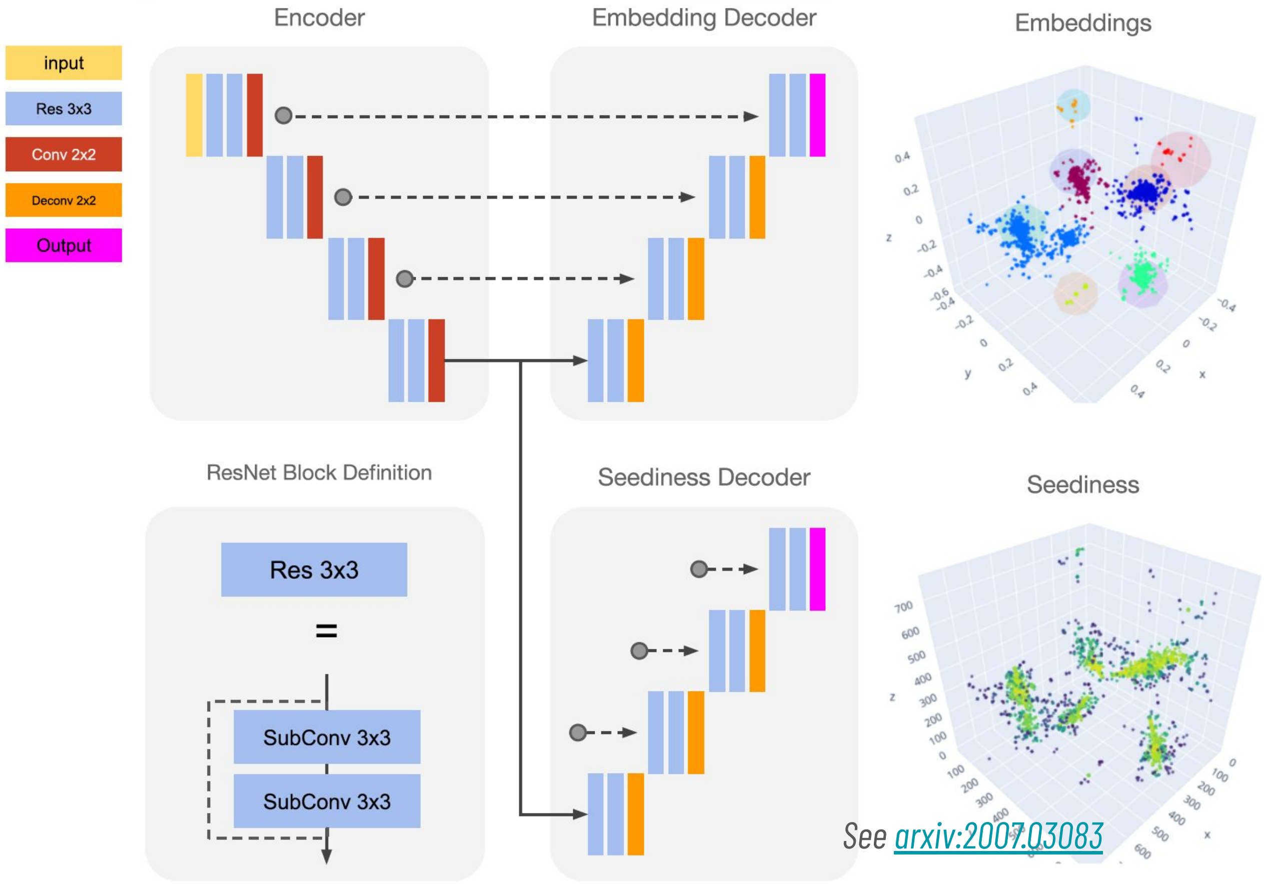


Image credit: arXiv 1708.02551

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

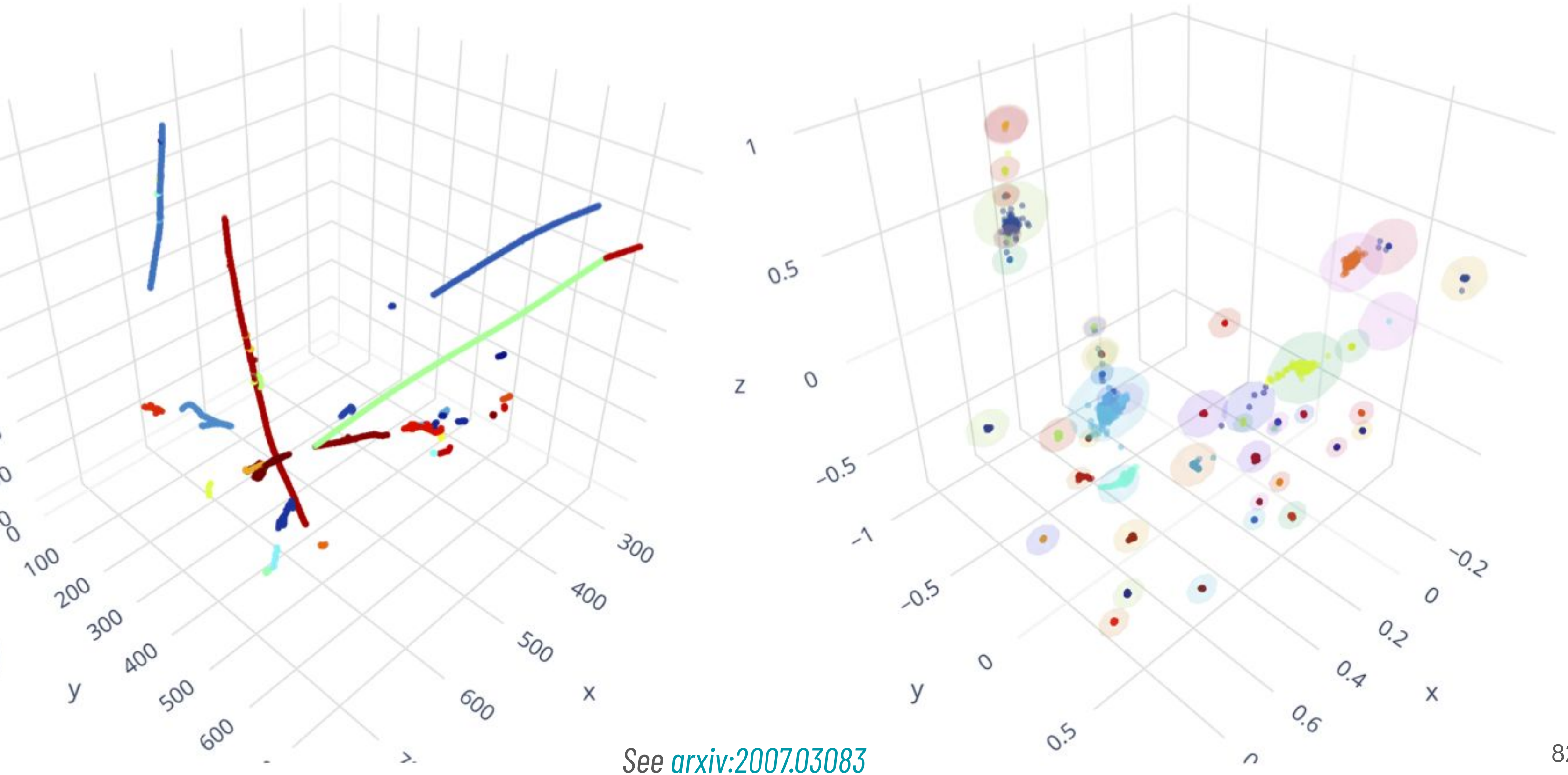


See [arxiv:2007.03083](https://arxiv.org/abs/2007.03083)

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

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

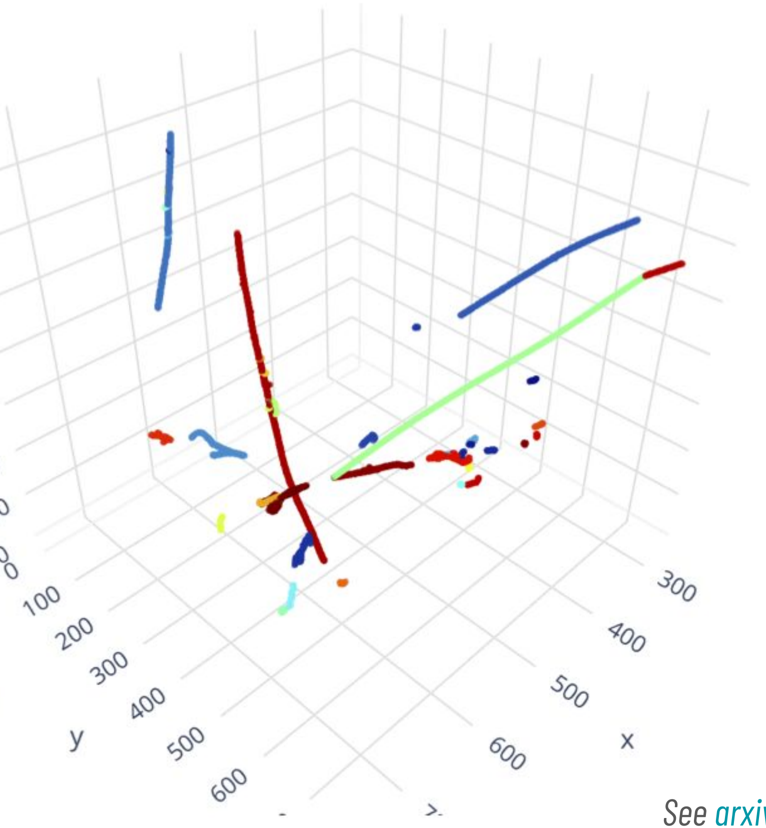


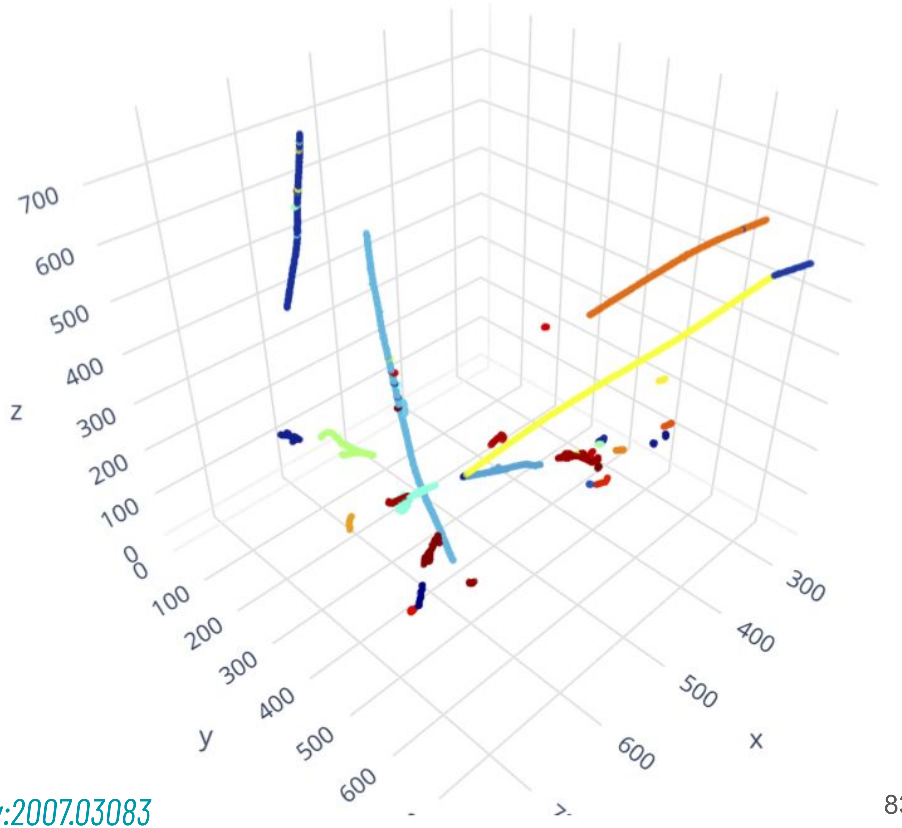
Work credit:
Dae Heun Koh (Stanford)

See arxiv:2007.03083

ML-based Neutrino Data Reconstruction Chain

Stage 2-a: Dense Pixel Clustering





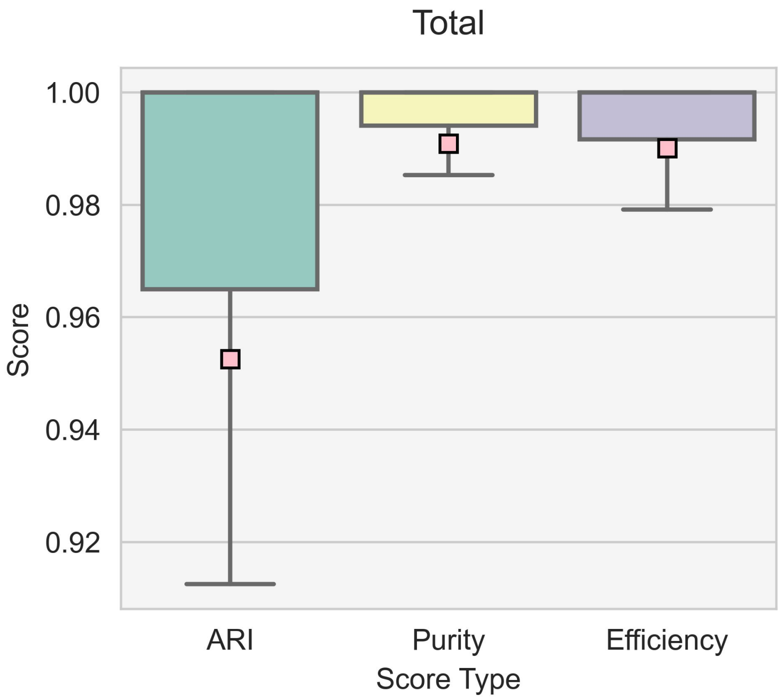
Work credit:
Dae Heun Koh (Stanford)

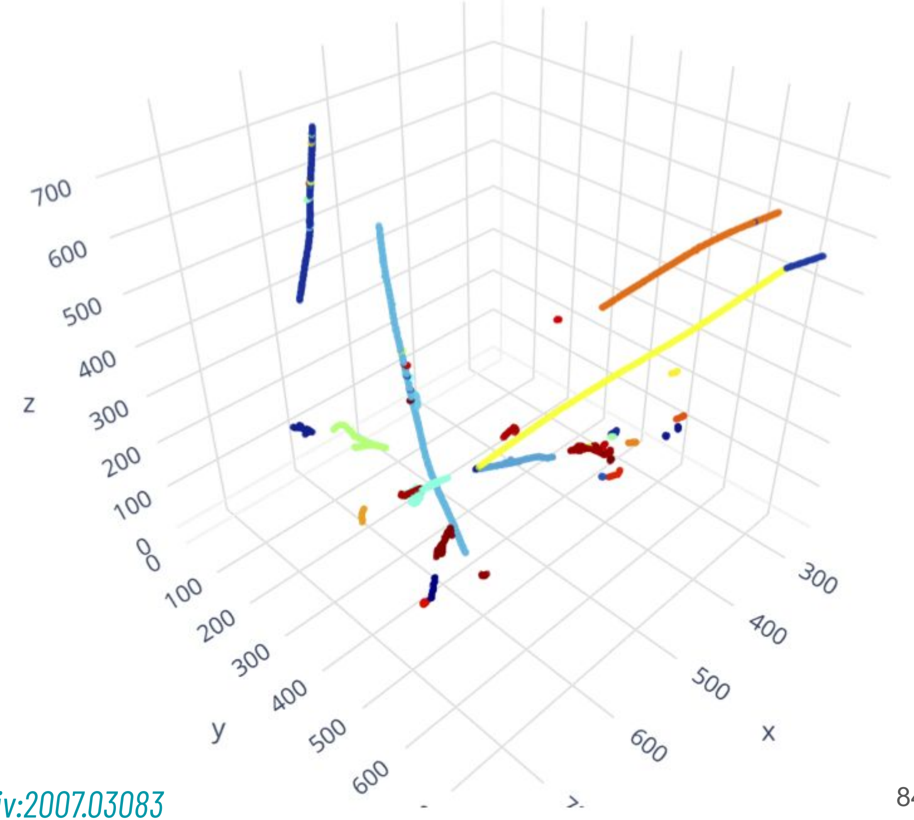
See arxiv:2007.03083

Stage 2-a: Dense Pixel Clustering

SLAC

Pixels clustered into trajectory fragments using SPICE

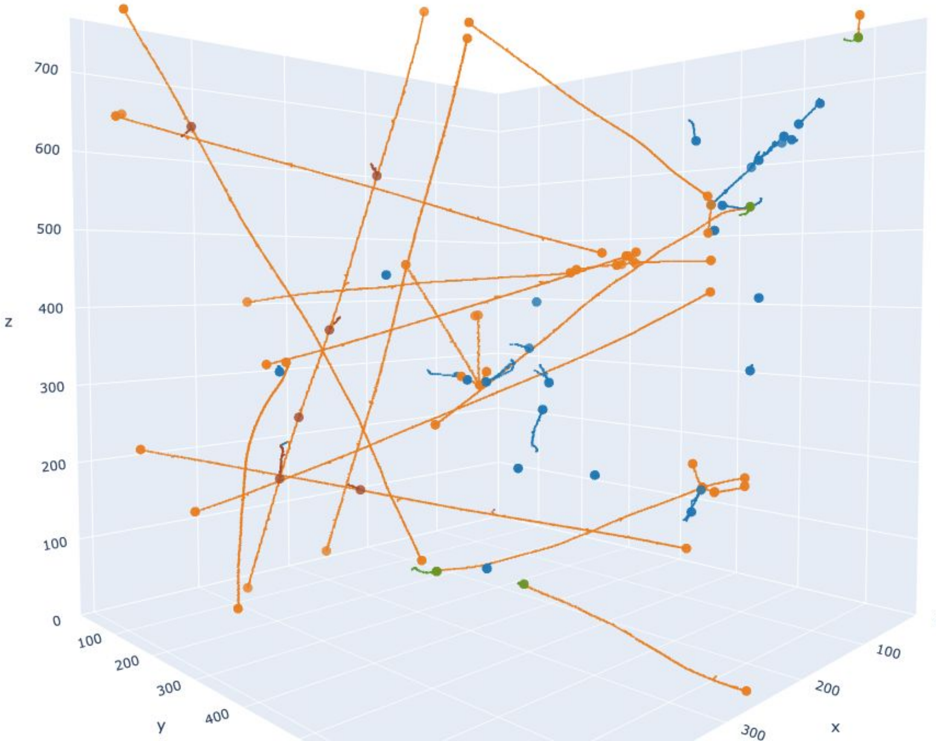




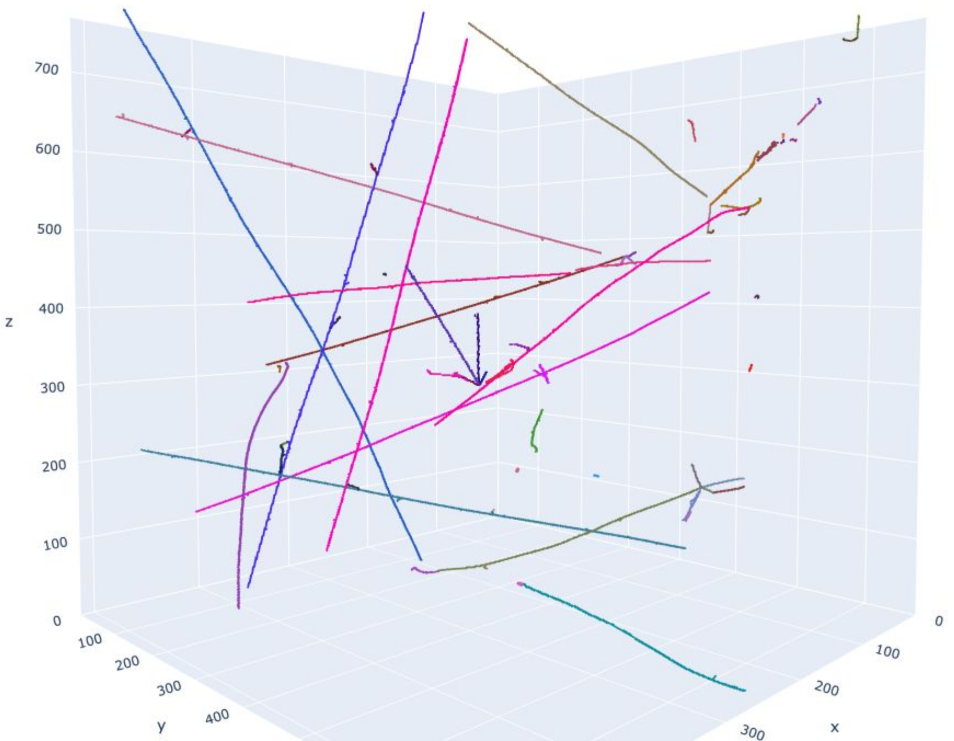
See arxiv:2007.03083

ML-based Neutrino Data Reconstruction Chain
Stage 2-a: input & output

Stage 2-a Input

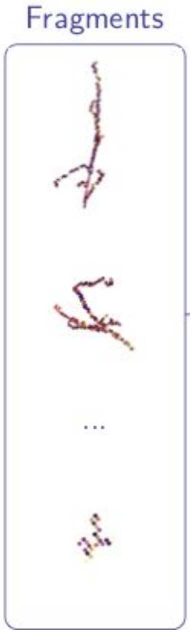


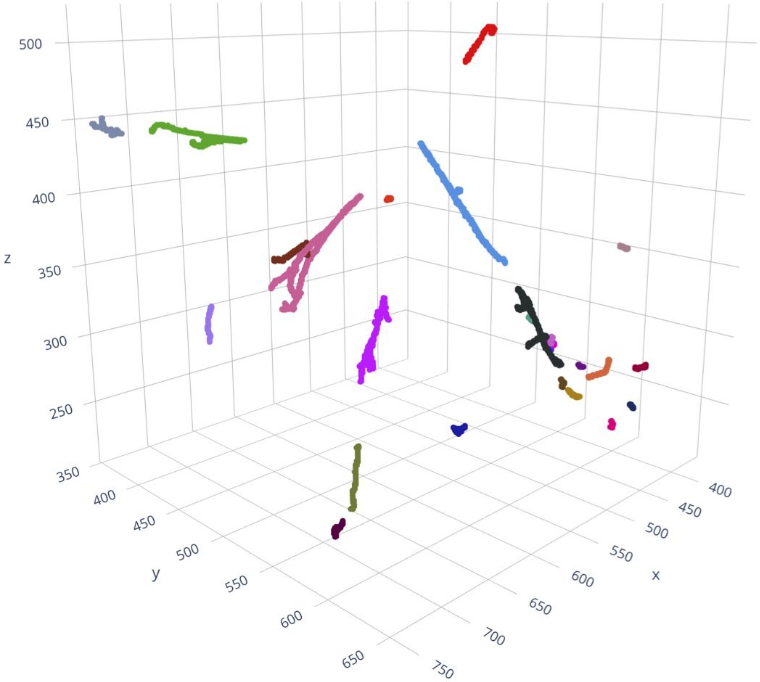
Stage 2-a Output



Stage 2-b: Sparse Fragment Clustering

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



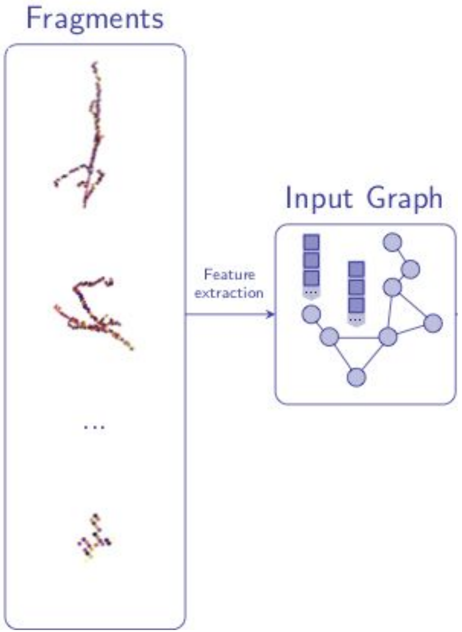


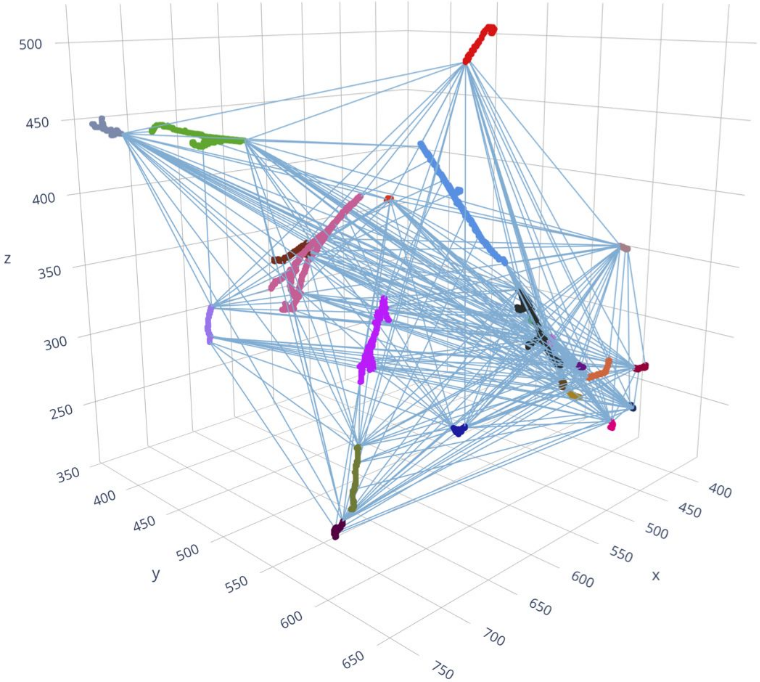
Stage 2-b: Sparse Fragment Clustering

SLAC

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

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

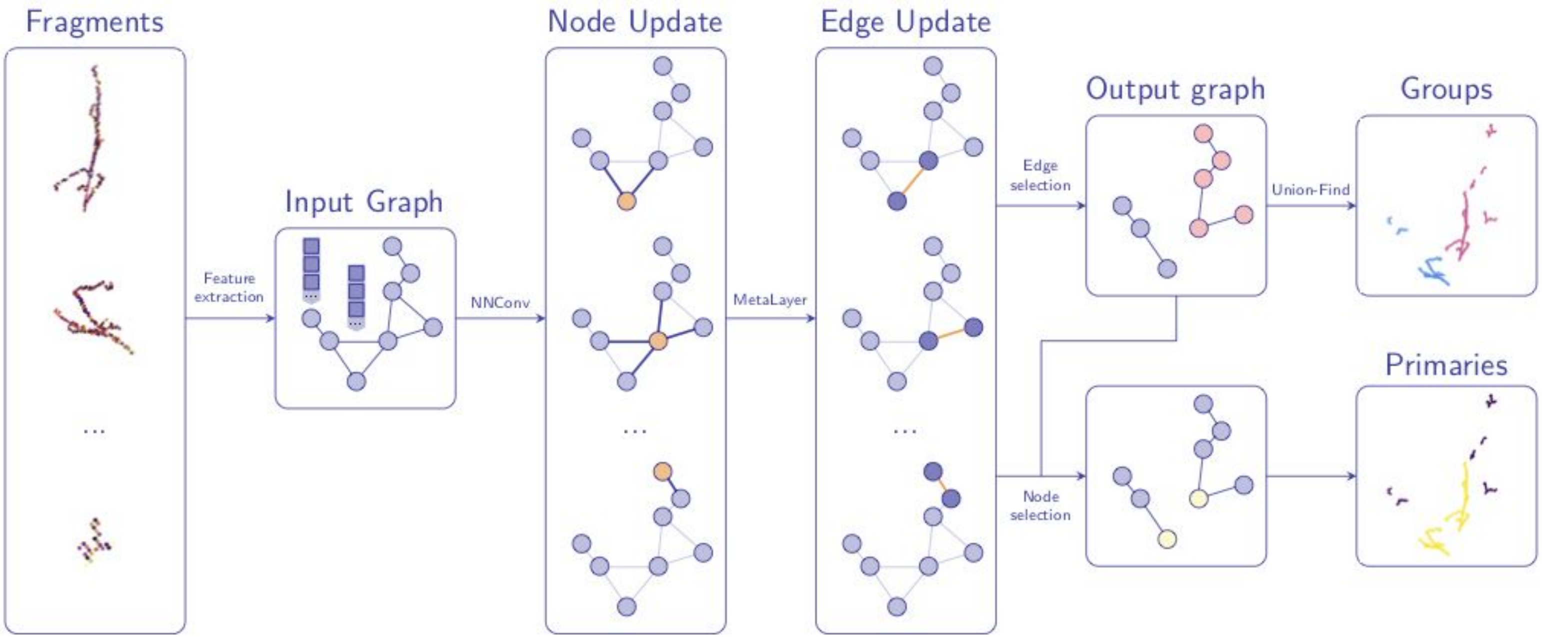




ML-based Neutrino Data Reconstruction Chain
Stage 2-b: Sparse Fragment Clustering

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)



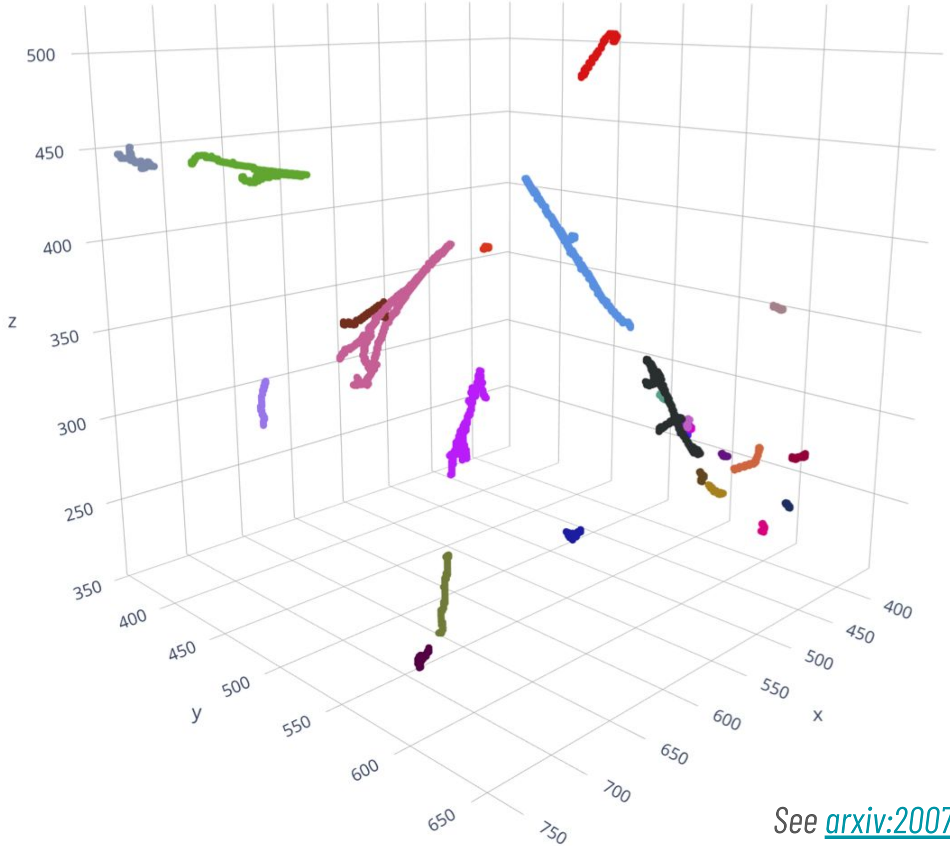
ML-based Neutrino Data Reconstruction Chain
Stage 2-b: Sparse Fragment Clustering

SLAC

Graph-NN for Particle Aggregation (GrapPA)

Input:

- Fragmented EM showers



See arxiv:2007.01335

Stage 2-b: Sparse Fragment Clustering

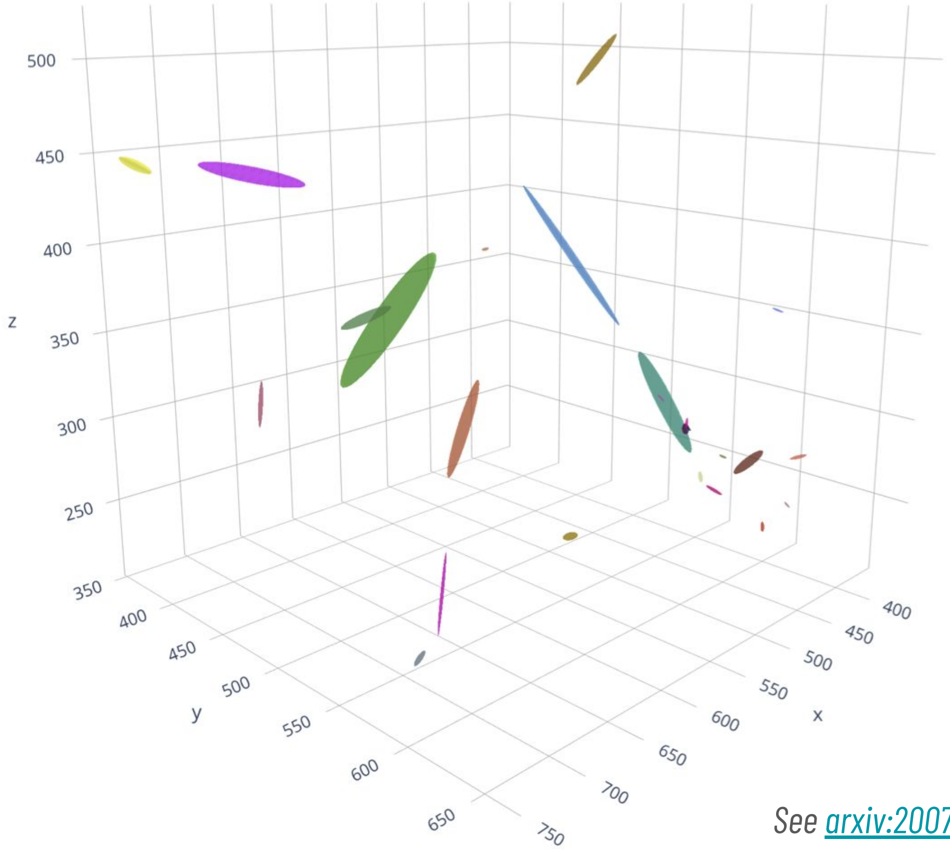
Graph-NN for Particle Aggregation (GrapPA)

Input:

- Fragmented EM showers

Node features:

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



See arxiv:2007.01335

Stage 2-b: Sparse Fragment Clustering

Graph-NN for Particle Aggregation (GrapPA)

Input:

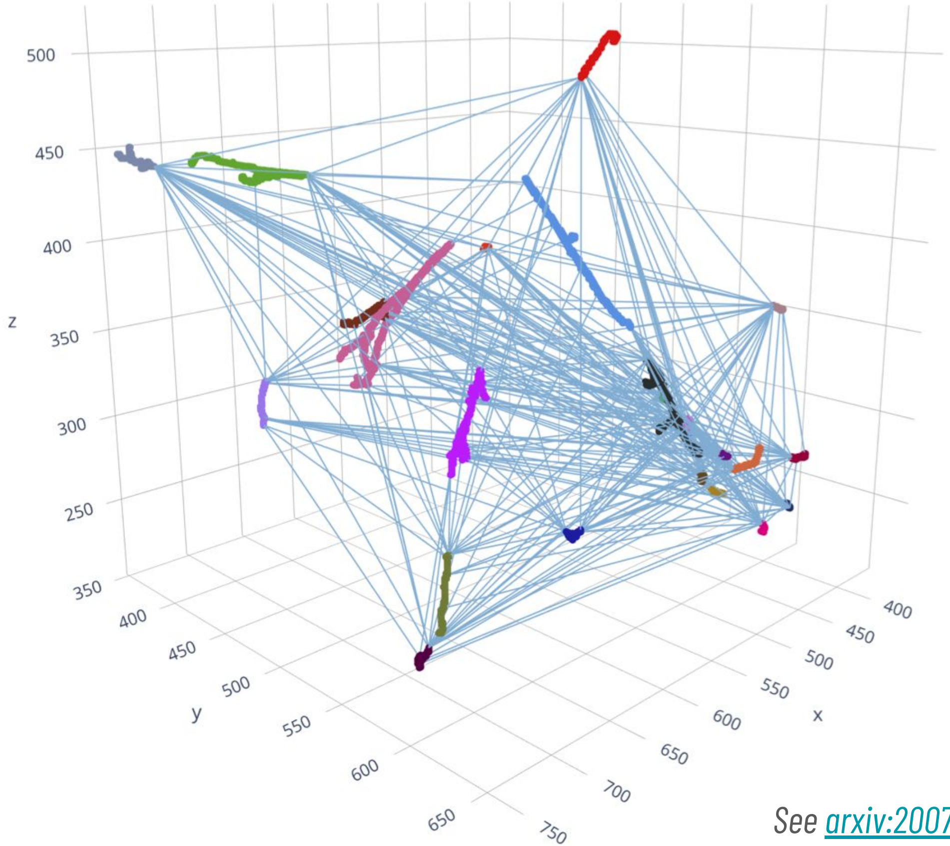
- Fragmented EM showers

Node features:

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

Input graph:

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



See arxiv:2007.01335

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

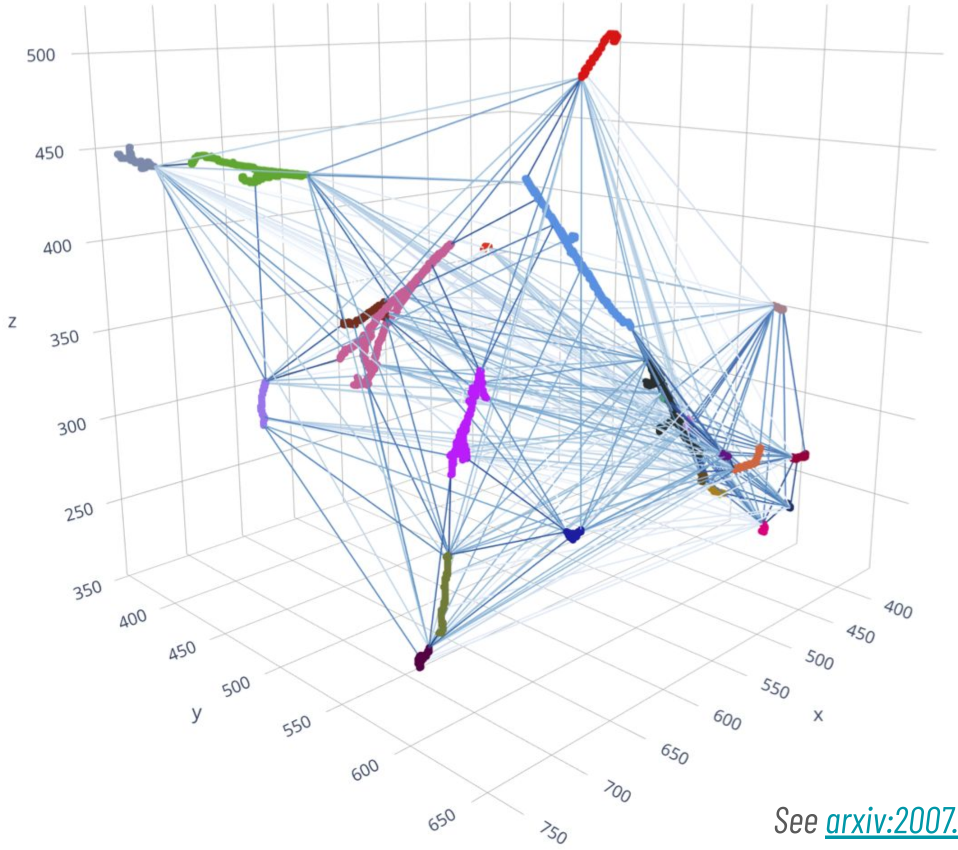
Graph-NN for Particle Aggregation (GrapPA)

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 arxiv:2007.01335

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

SLAC

Graph-NN for Particle Aggregation (GrapPA)

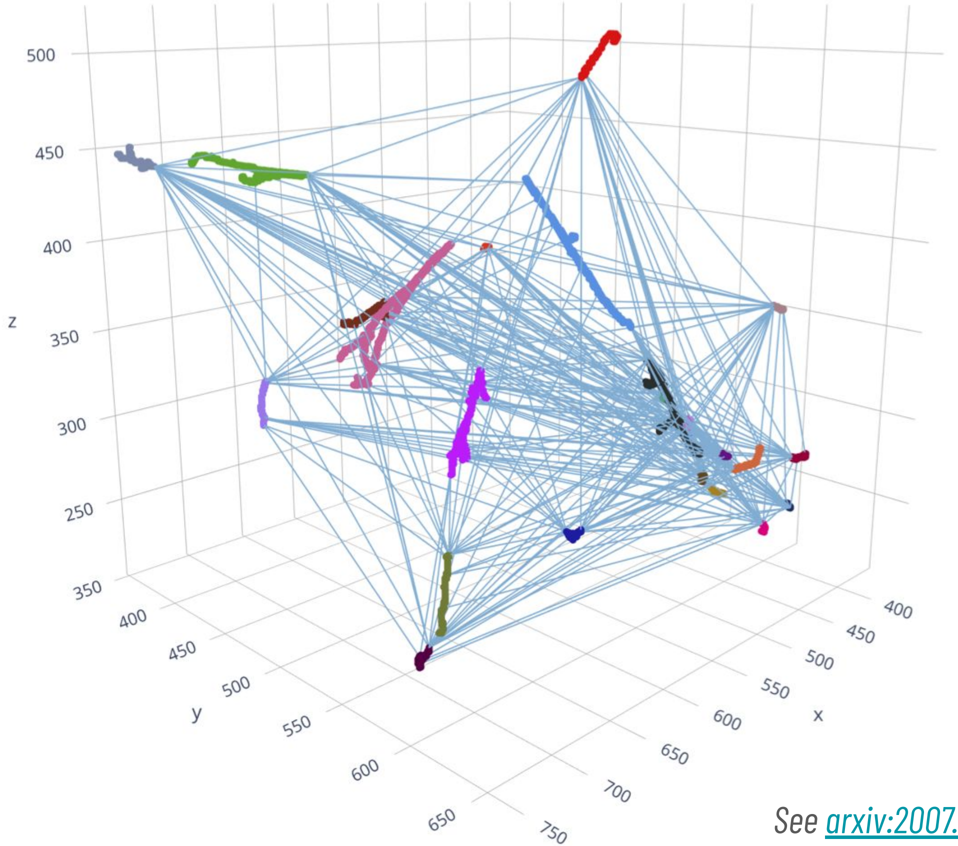
Message passing (MP):

- 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

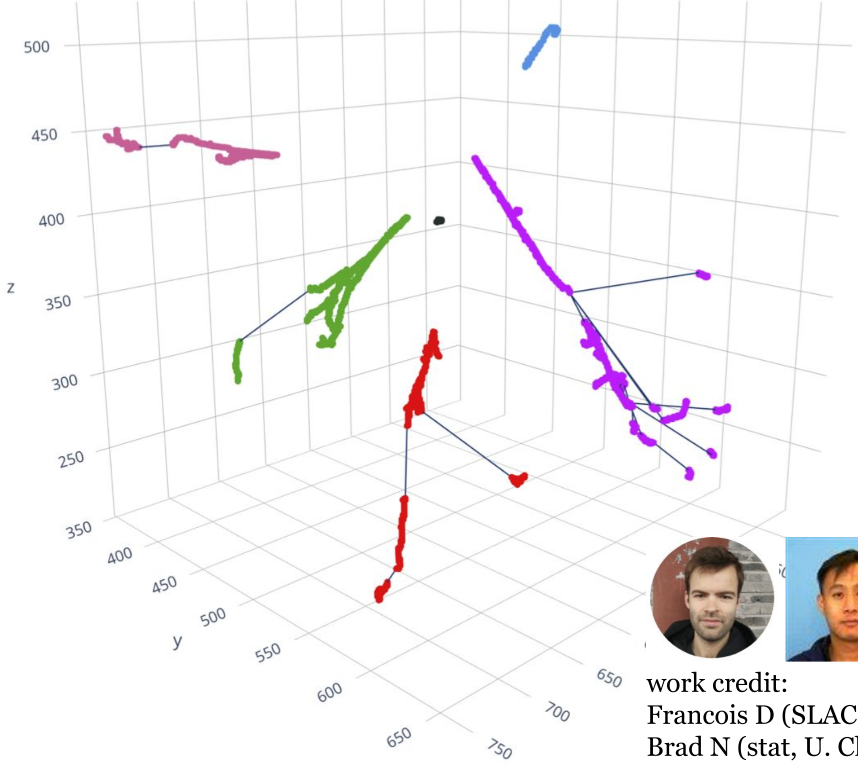
For more studies, see our paper



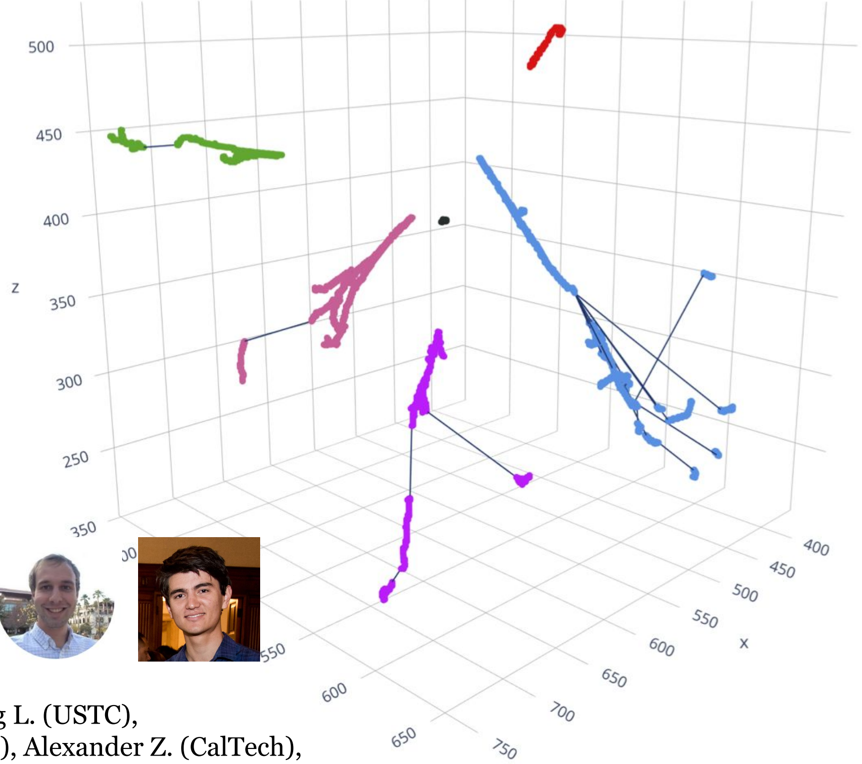
See arxiv:2007.01335

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

Target Label



Edge Prediction

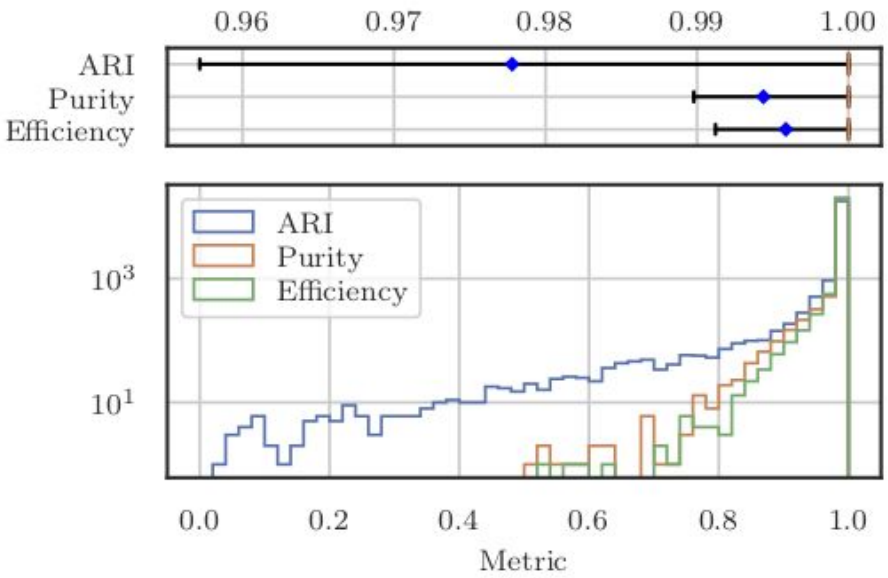


work credit:
Francois D (SLAC), Qing L. (USTC),
Brad N (stat, U. Chicago), Alexander Z. (CalTech),

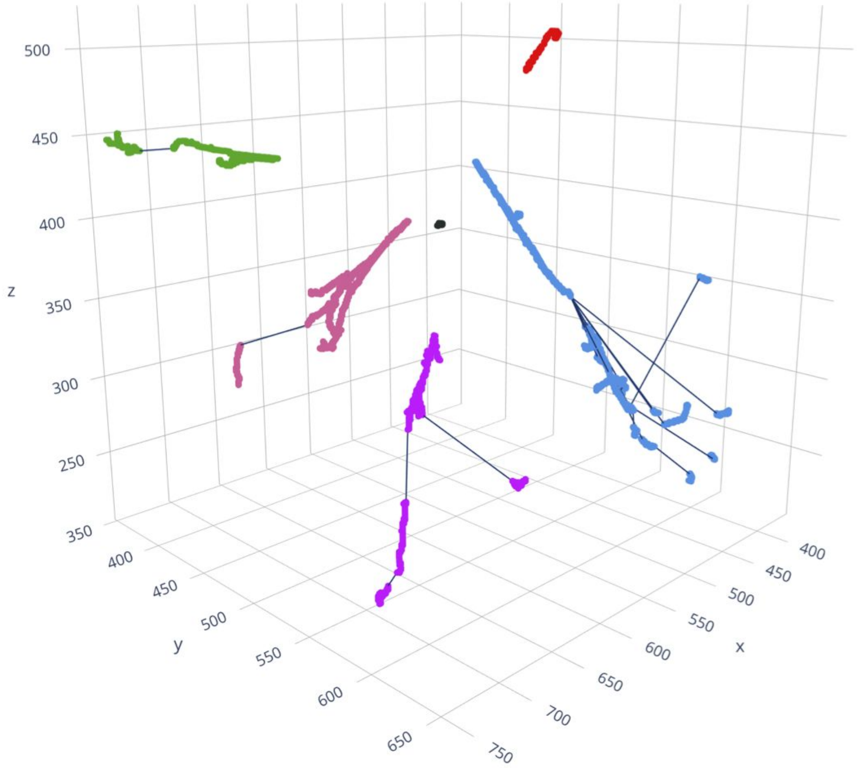
Stage 2-b: Sparse Fragment Clustering

Clustering using GrapPA

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



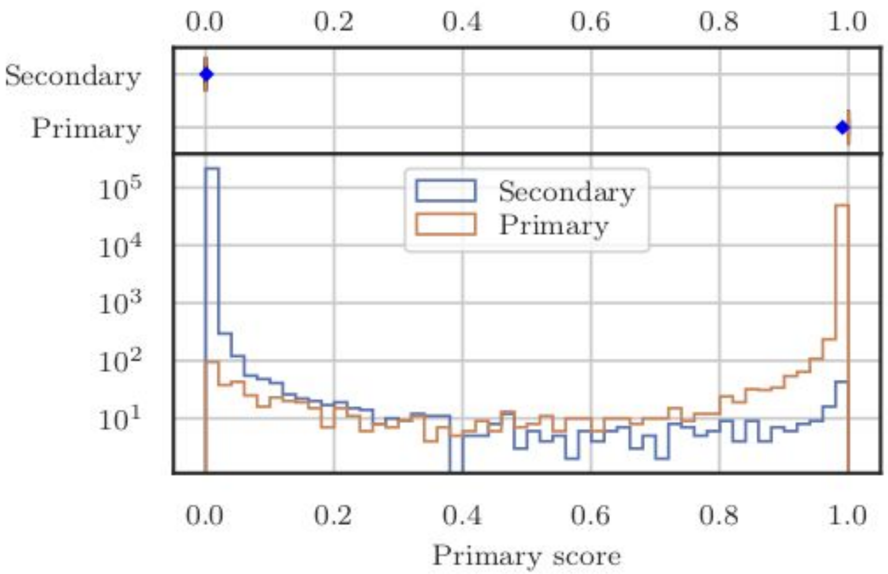
Edge Prediction



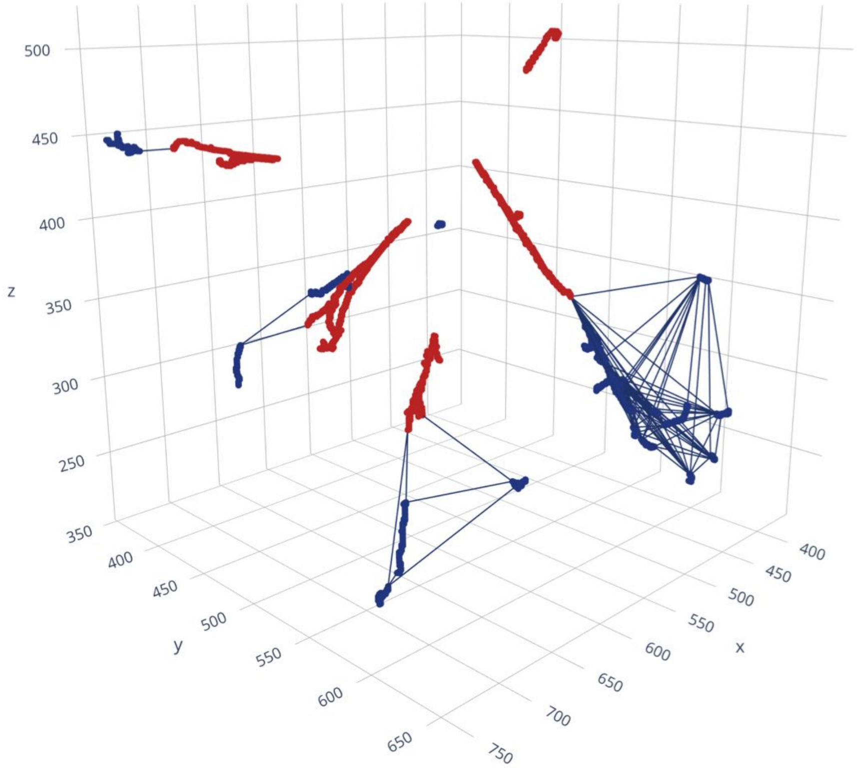
Stage 2-b: Sparse Fragment Clustering

Start ID using GrapPA

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

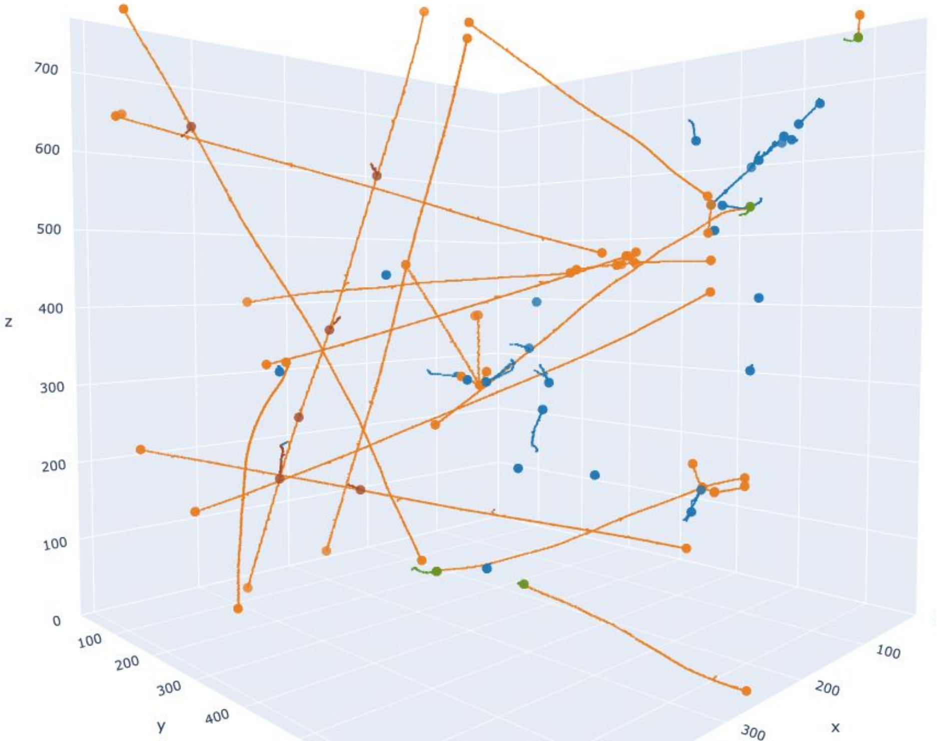


Node prediction

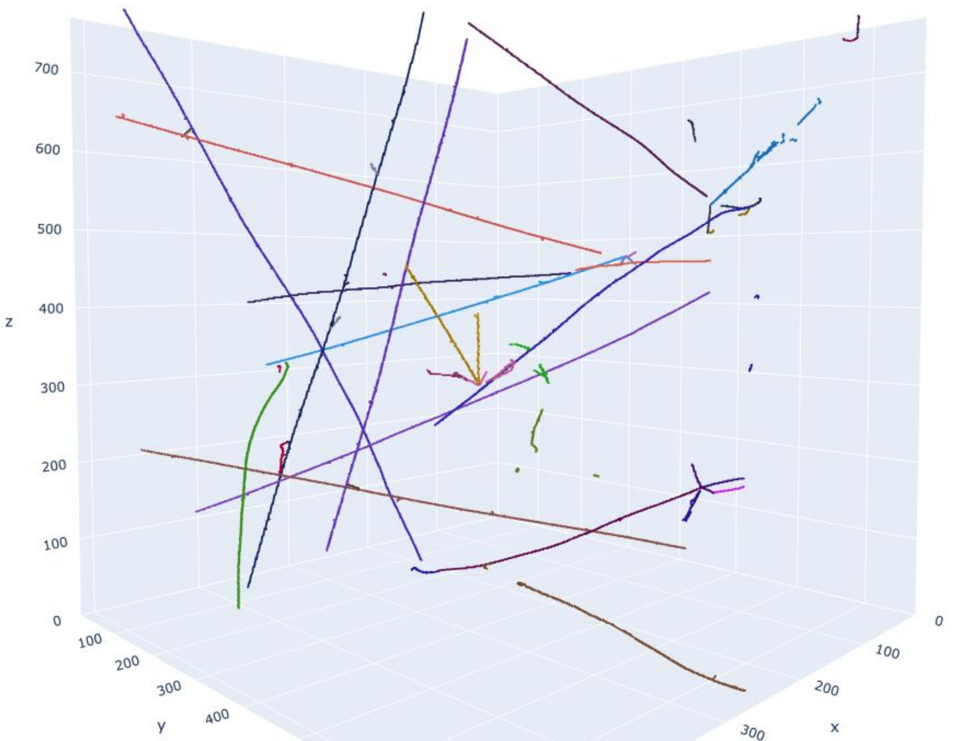


ML-based Neutrino Data Reconstruction Chain
Stage 2: input & output

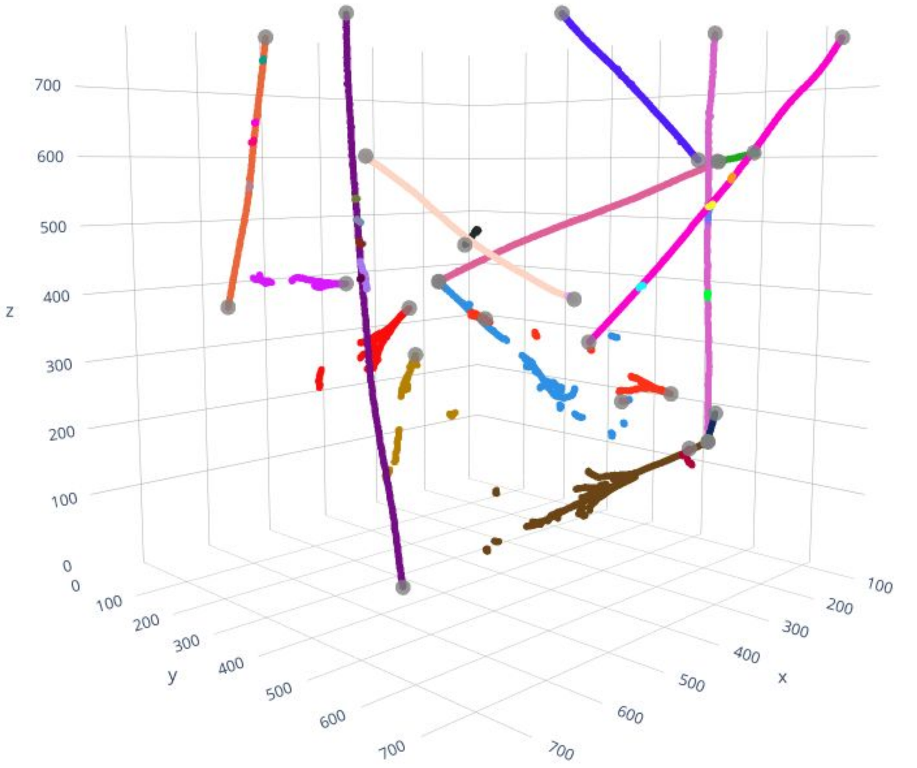
Stage 2 Input



Stage 2 Output



ML-based Neutrino Data Reconstruction Chain
Stage 3: Interaction Clustering



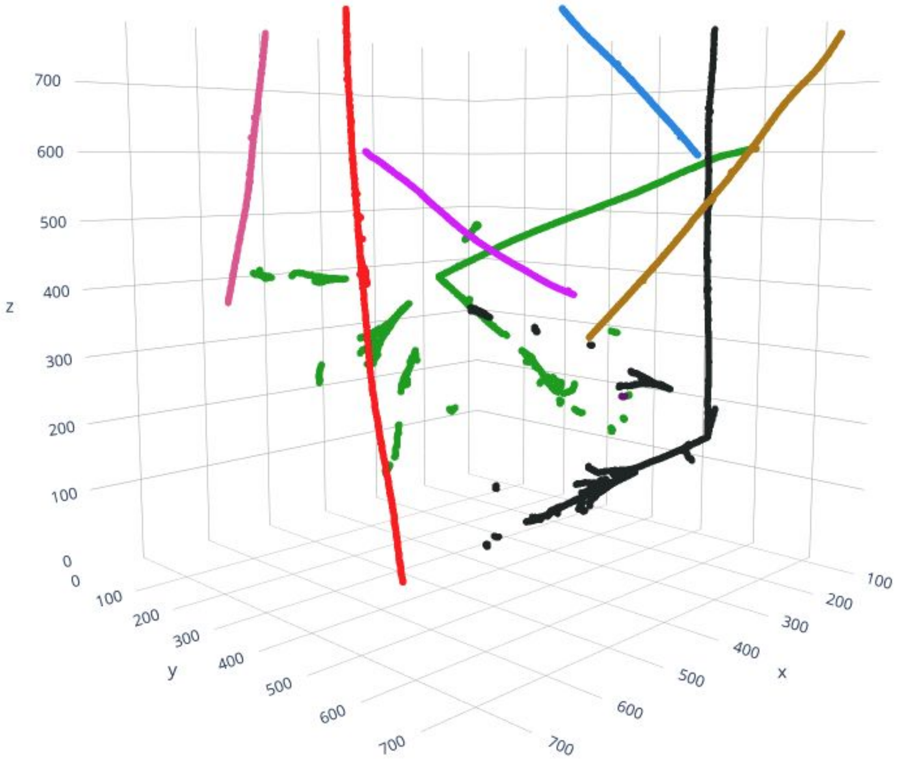
Identifying Each Interaction?

This task can be casted to the same task already solved using GrapPA!

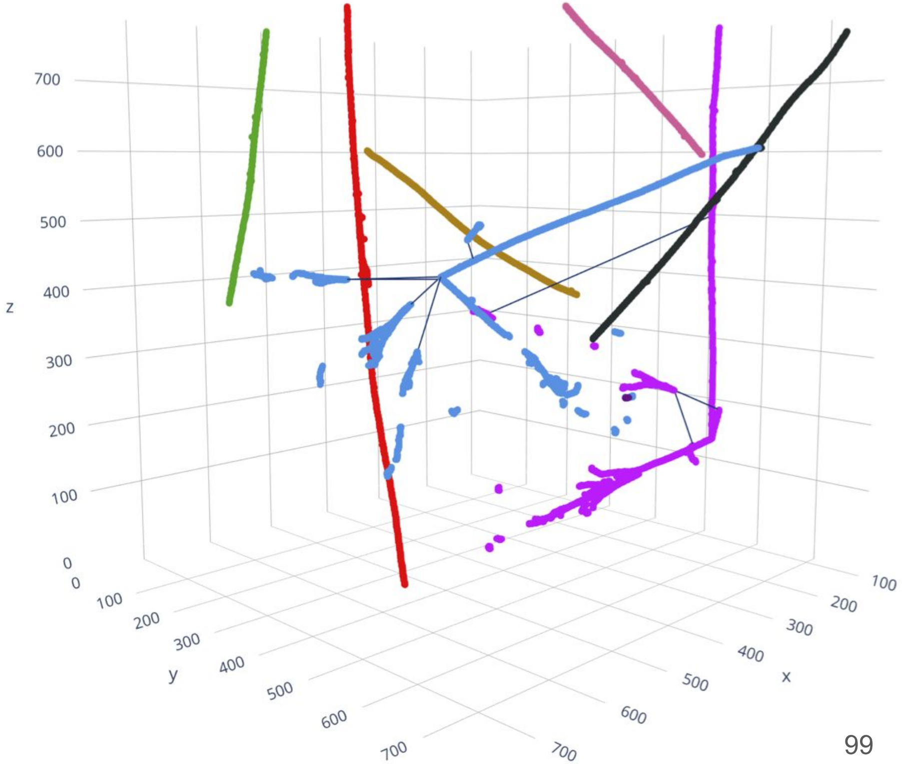
- Interaction = a group of particles that shared the same origin (i.e. neutrino interaction)
- Edge classification to identify an interaction
- Node classification for particle type ID

ML-based Neutrino Data Reconstruction Chain
Stage 3: Interaction Clustering

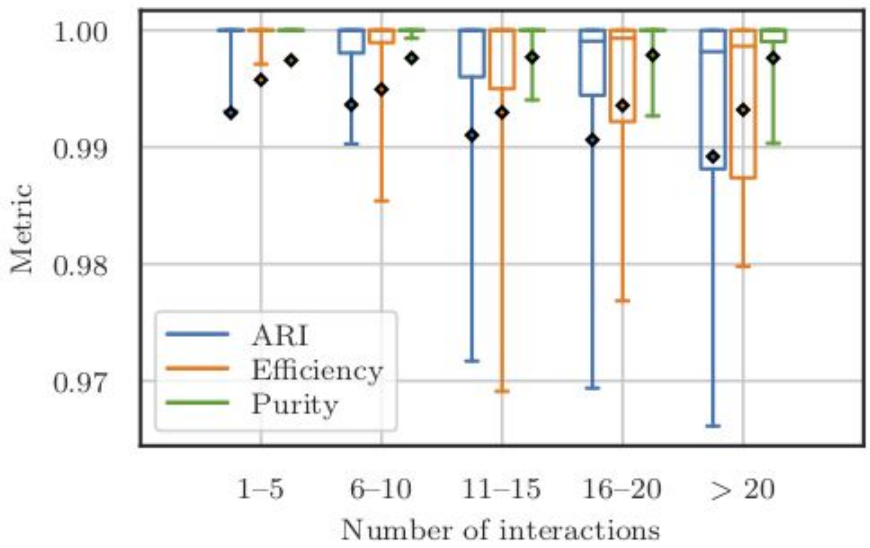
Target Group



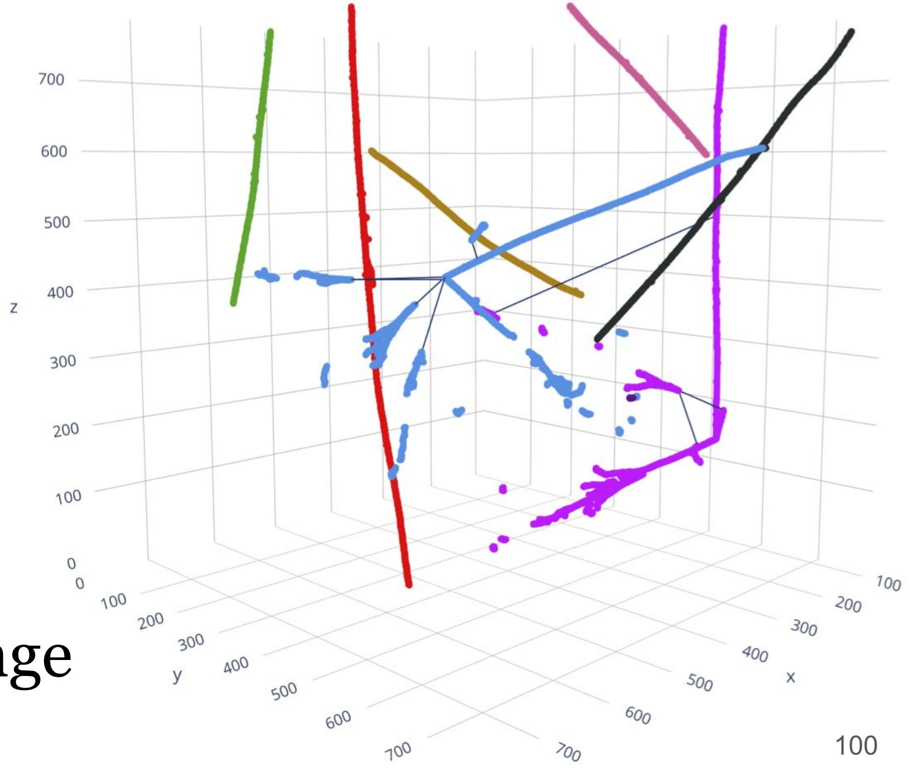
Predicted Interaction



ML-based Neutrino Data Reconstruction Chain
Stage 3: Interaction Clustering



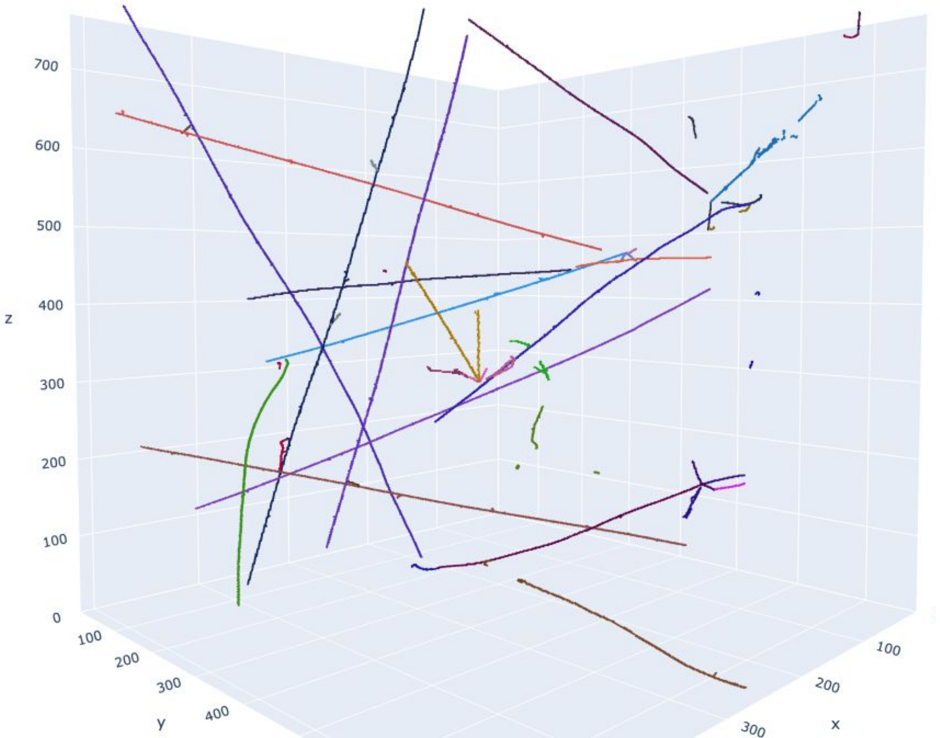
Predicted Interaction



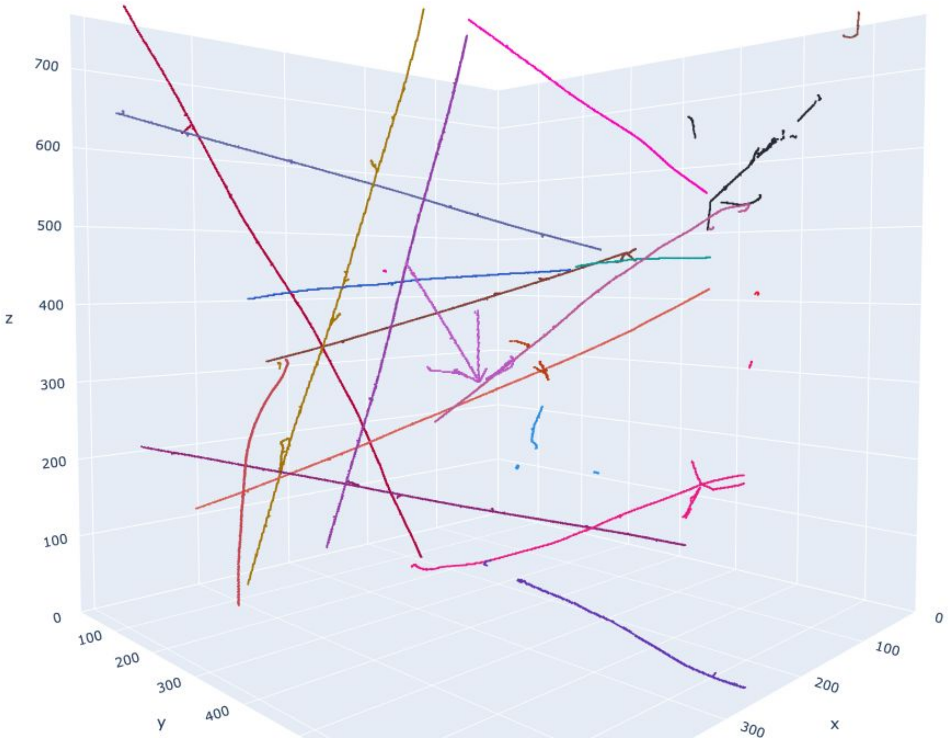
Promising result to address DUNE-ND reconstruction challenge (~20 neutrino pile-up)

ML-based Neutrino Data Reconstruction Chain
Stage 3: input & output

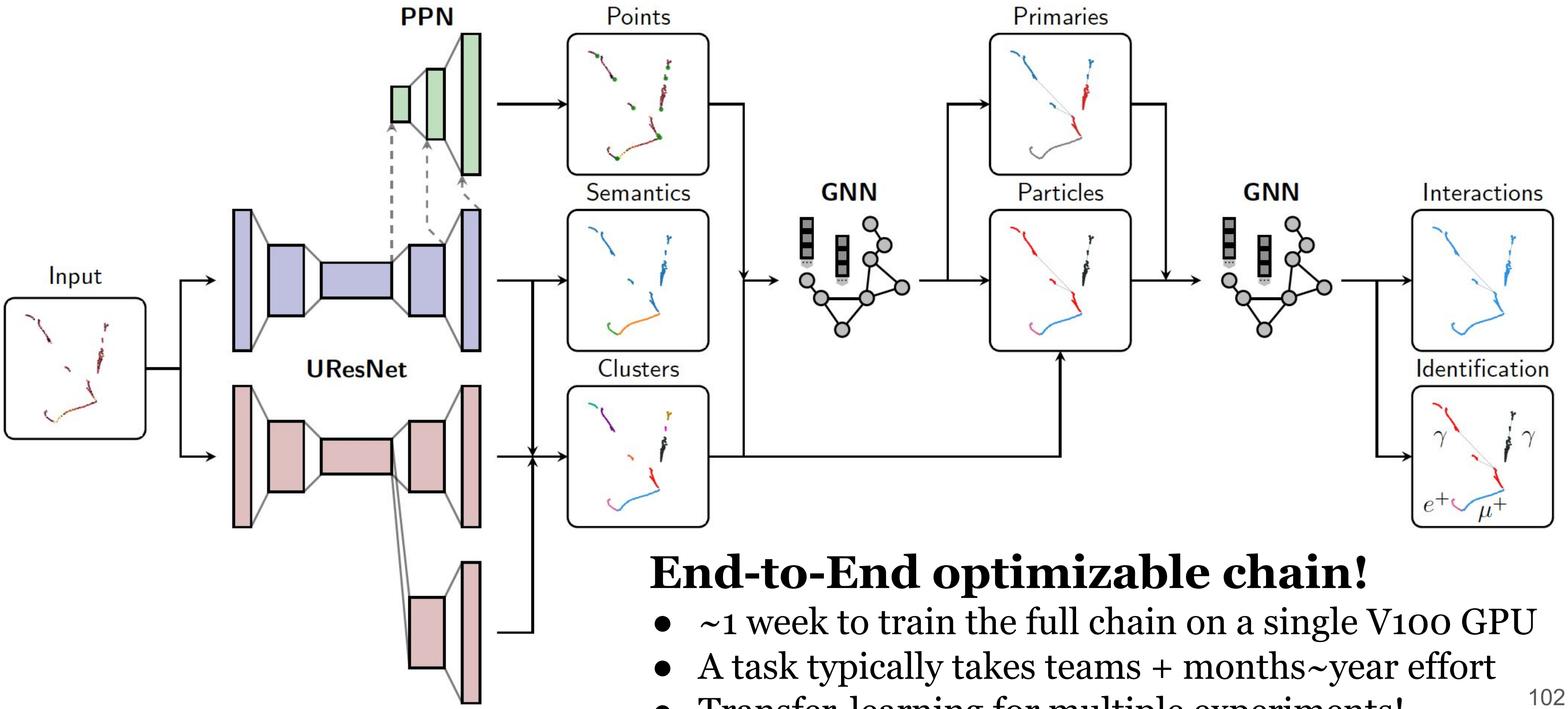
Stage 3 Input



Stage 3 Output



ML-based Neutrino Data Reconstruction Chain
Wrapping up...



End-to-End optimizable chain!

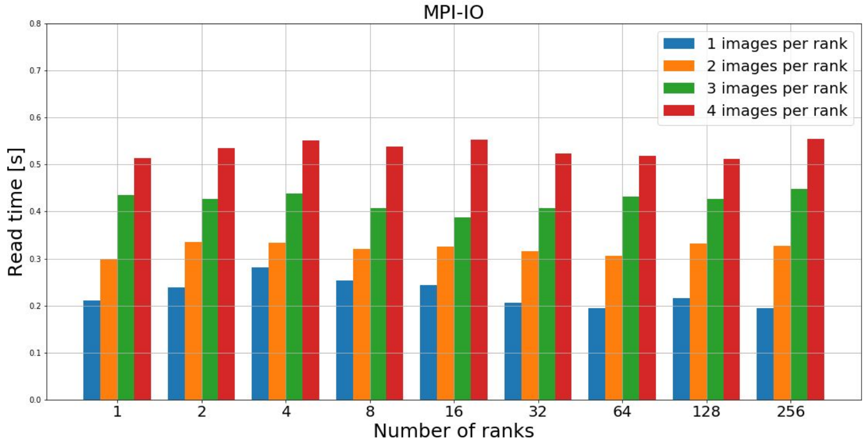
- ~1 week to train the full chain on a single V100 GPU
- A task typically takes teams + months~year effort
- Transfer-learning for multiple experiments!

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

Wrapping up...

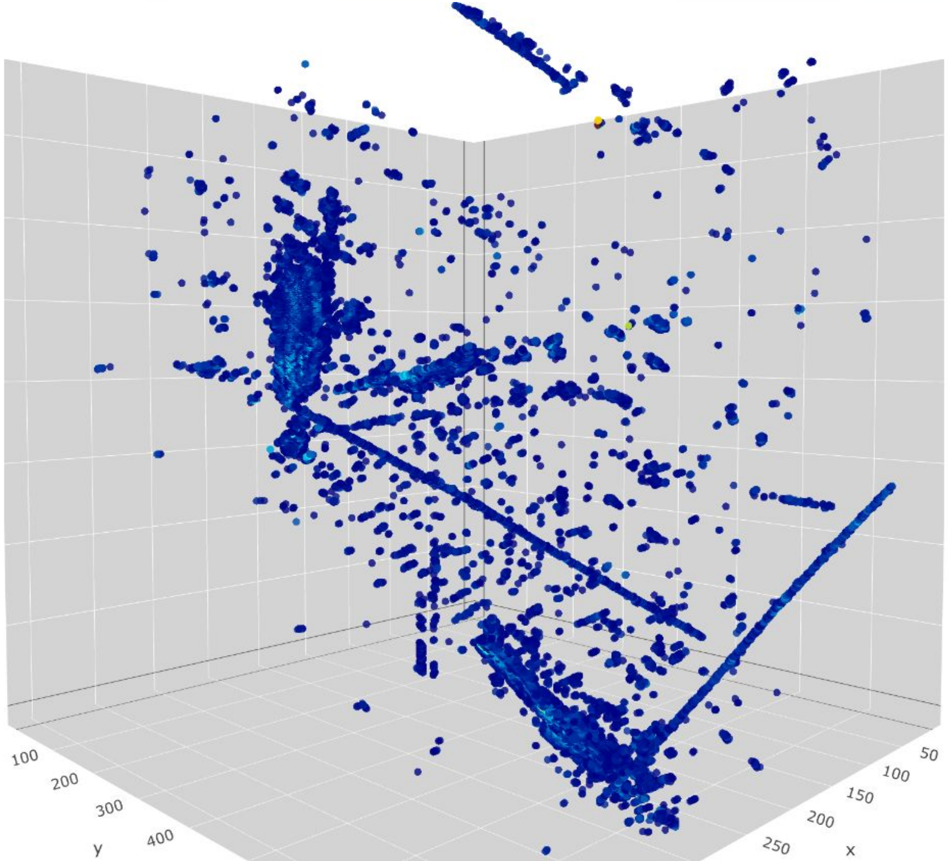
Summary

- **Neutrino detector trend: hi-res. particle imaging**
- **Analysis trend: computer vision algorithms**
  - Benefit the hi-resolution image = lots of heuristics (in non-ML)
  - ML-based approach has shown strong promise
- **ML-based data reconstruction approach**
  - especially for “busy” detectors … my research :)
  - Working on implementing inductive-bias/causality (“physics”)
- **Other active areas**: data/sim domain discrepancy adaptation
  - minimize the discrepancy, identify the source, quantify uncertainty

2D=>3D

Machine Learning & Computer Vision in Neutrino Physics

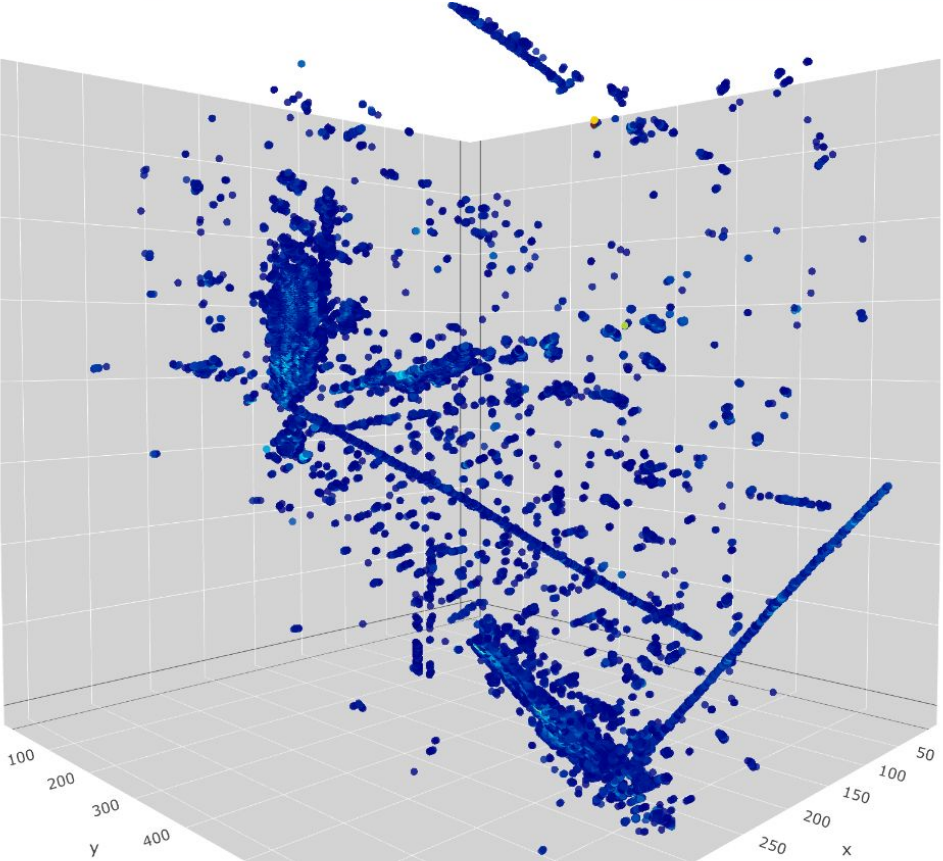
Bonus: isochronous ghost point removal

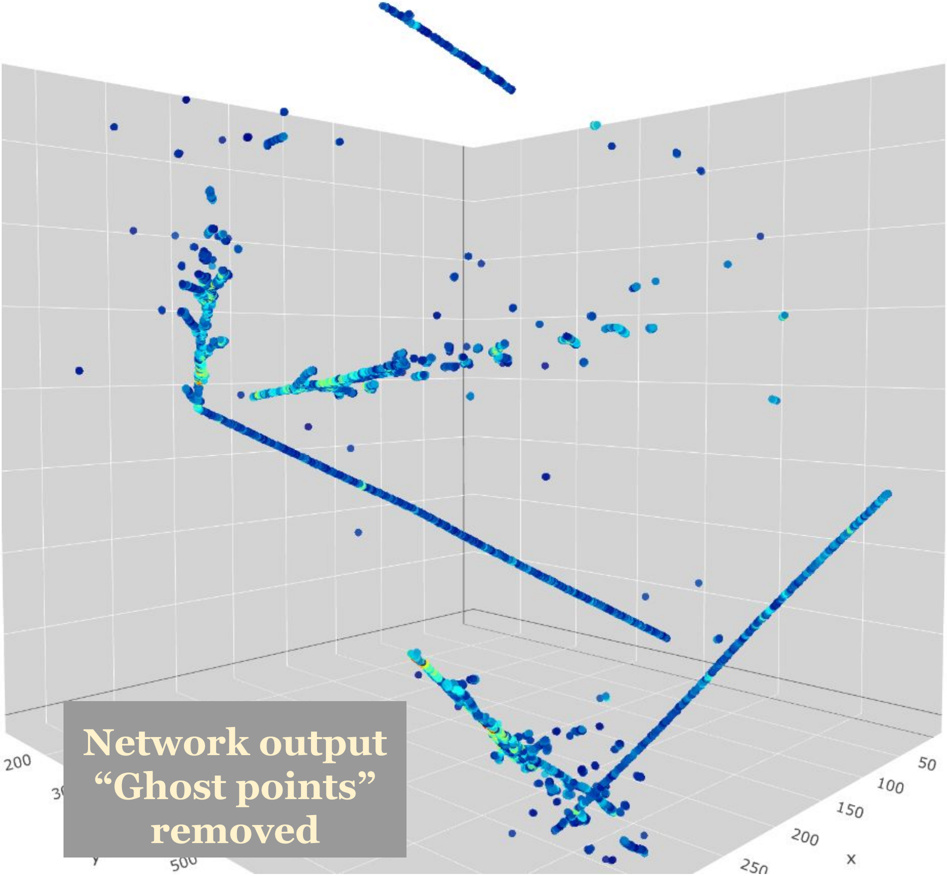


ICARUS Detector
Reconstructed 3D points

work credit:
Laura Domine
Patrick Tsang

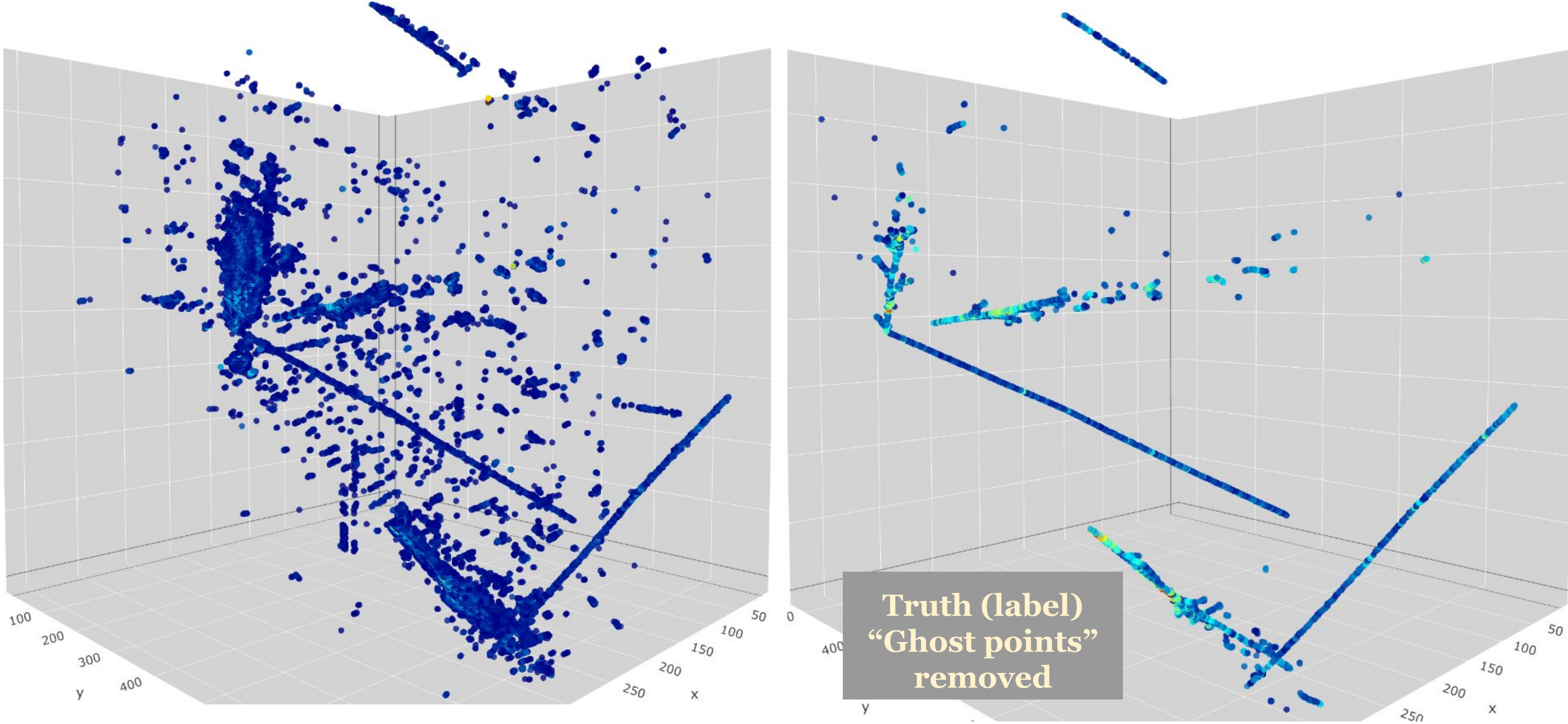
Bonus: isochronous ghost point removal





Machine Learning & Computer Vision in Neutrino Physics

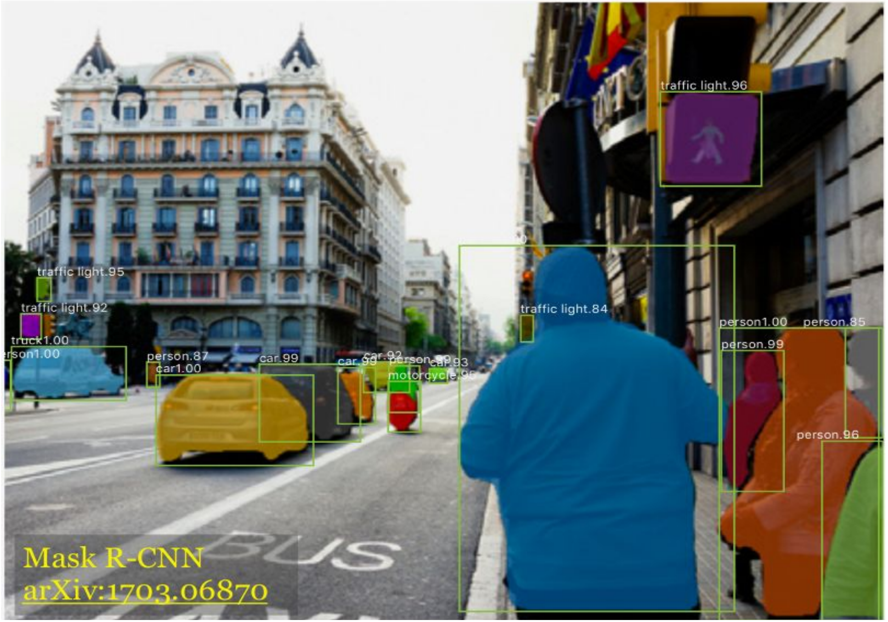
Bonus: isochronous ghost point removal



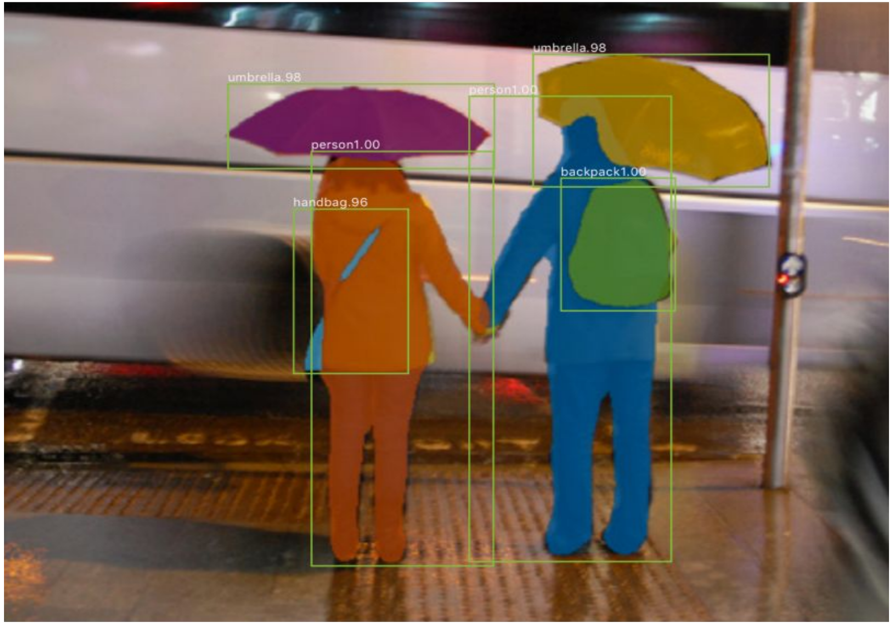
SPICE

Instance+Semantic Segmentation

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

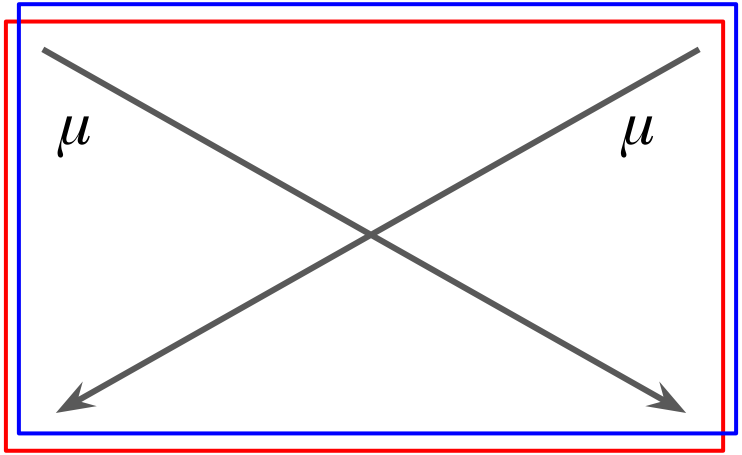


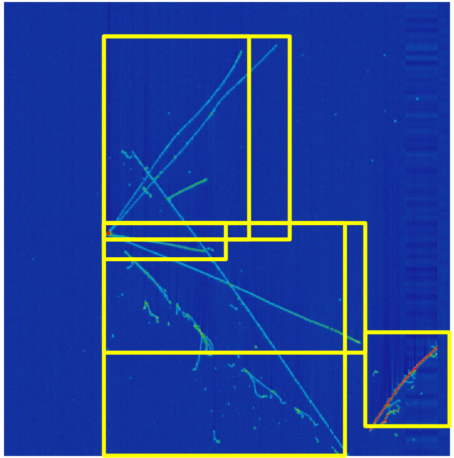
Mask R-CNN
arXiv:1703.06870



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)





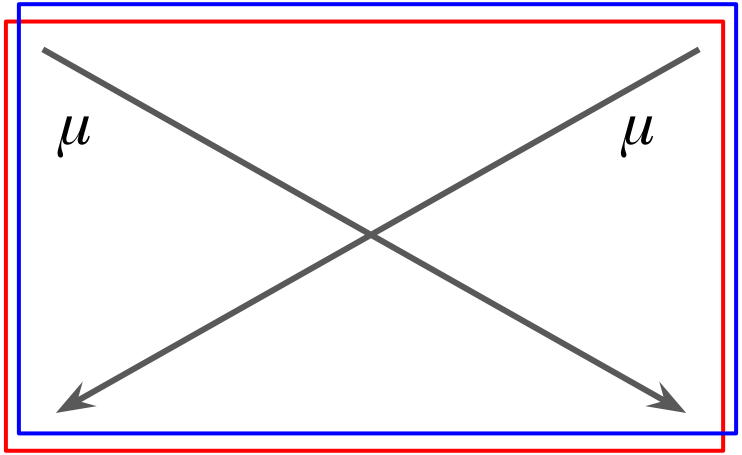
Occlusion issue

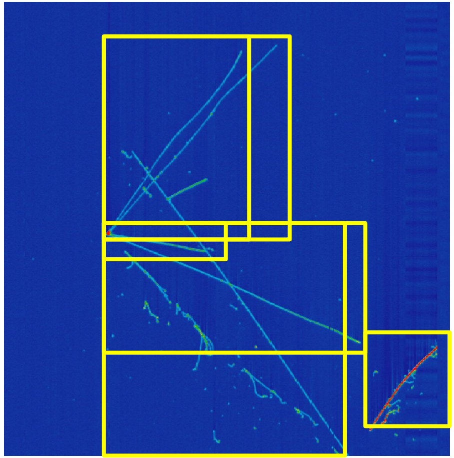
The overlap rate of particles is very high especially for our signal (neutrinos) with an event vertex.

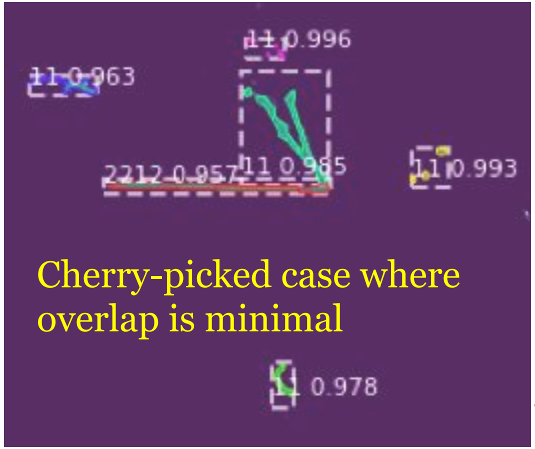
ML-based Neutrino Data Reconstruction Chain
Stage 2: Particle & Interaction Clustering

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

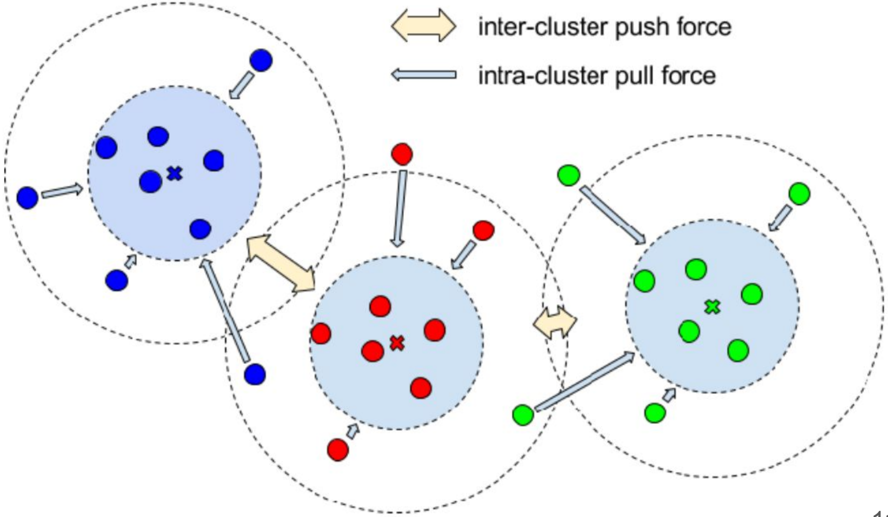
- Three component loss: pull together points that belong to the same cluster, keep distance between clusters, and regularization

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

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

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

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

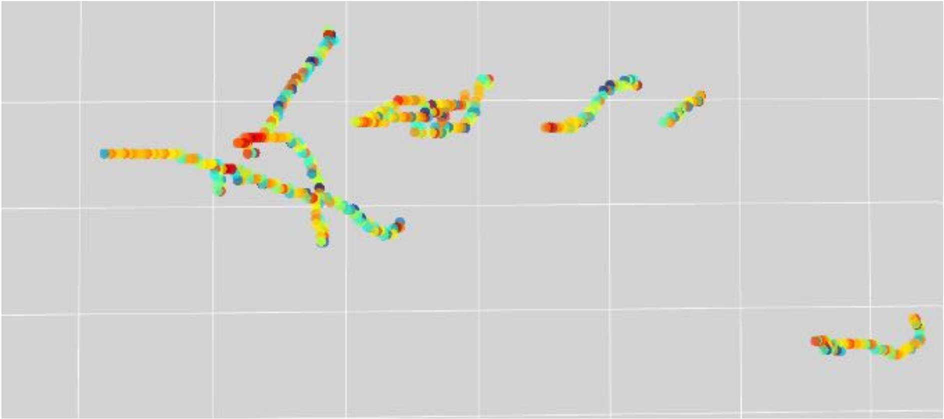


Equation credit: Dae Hyun K. @ Stanford

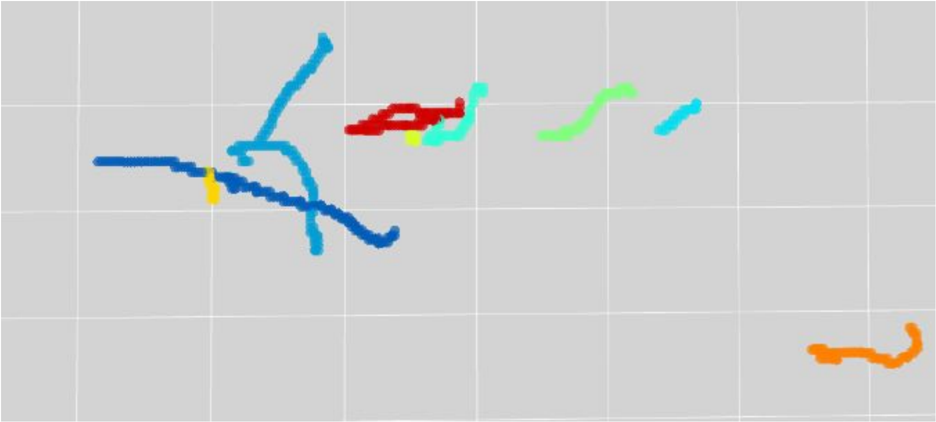
Image credit: arXiv 1708.02551

Instance+Semantic Segmentation

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



Input: 3D pixel energy depositions

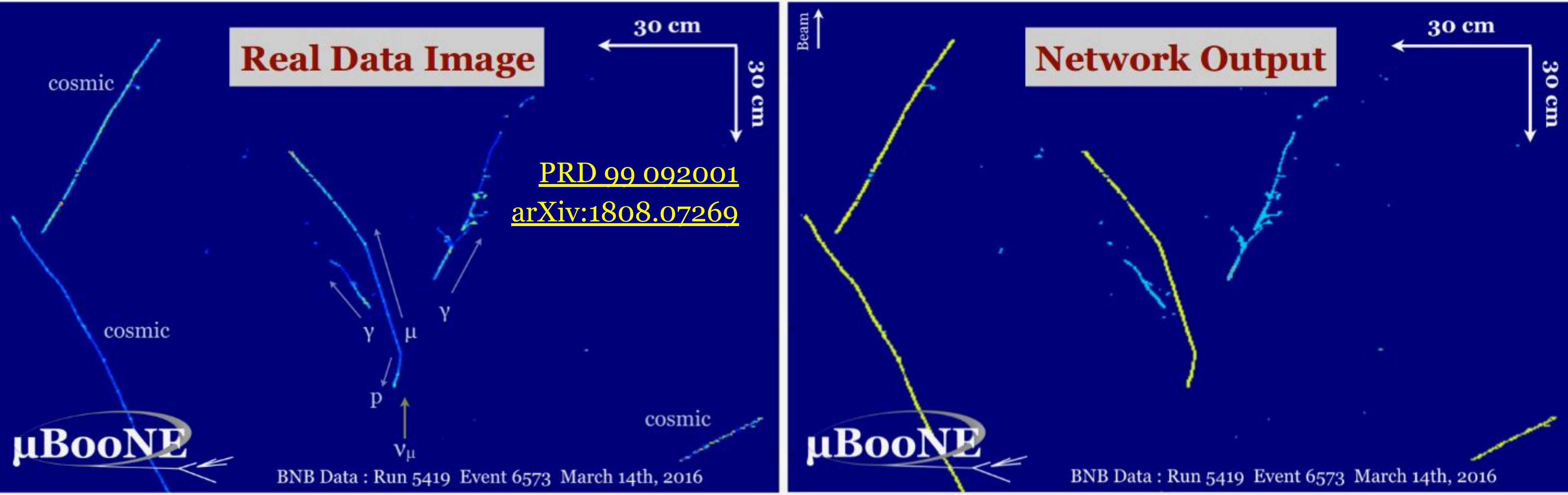


Output: 3D pixel clusters (DBScan in hyperspace)

Machine Learning & Computer Vision in Neutrino Physics

Semantic Segmentation for Pixel-level Particle ID

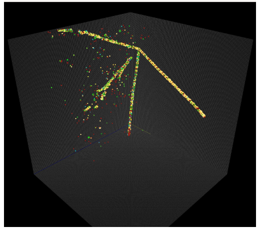
Separate electron/positron energy depositions from other types at raw waveform level.
Helps the downstream clustering algorithms (**data/sim comp. @ arxiv:1808.07269**)

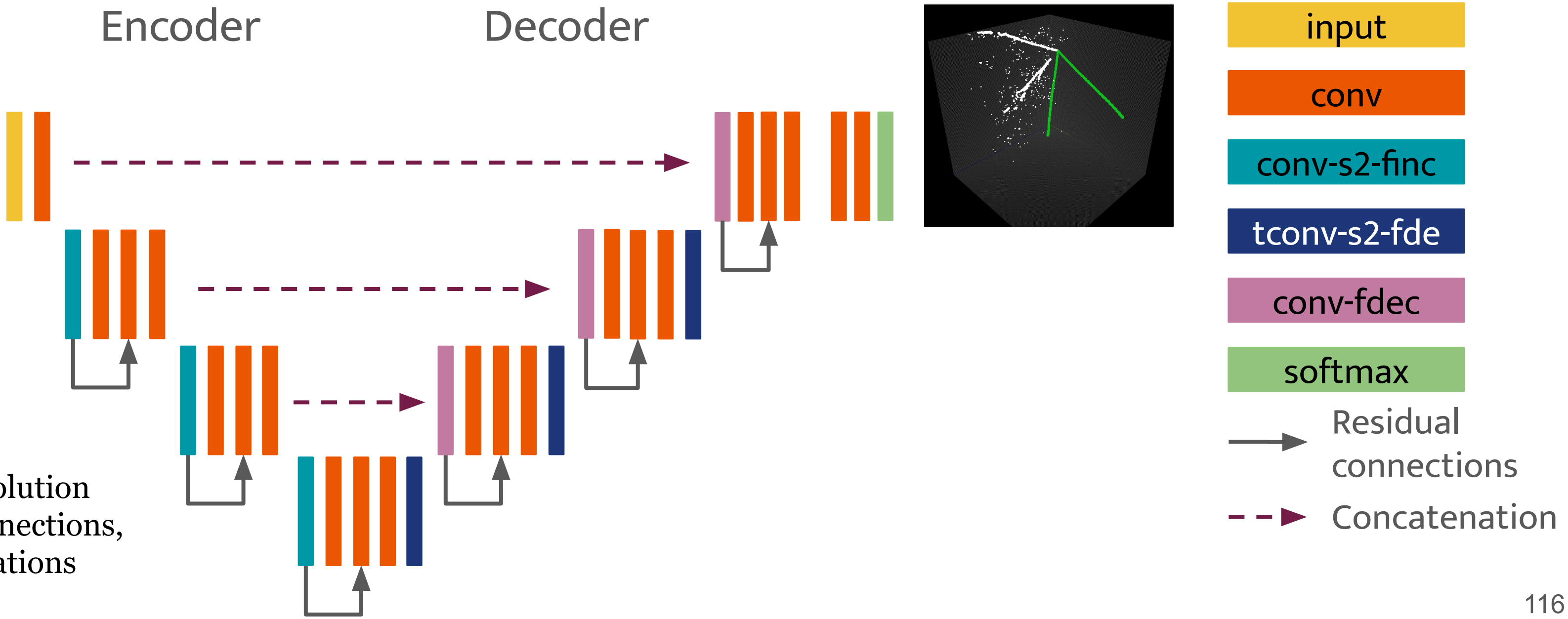


Network Input **Network Output**

Semantic Segmentation for Pixel-level Particle ID

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

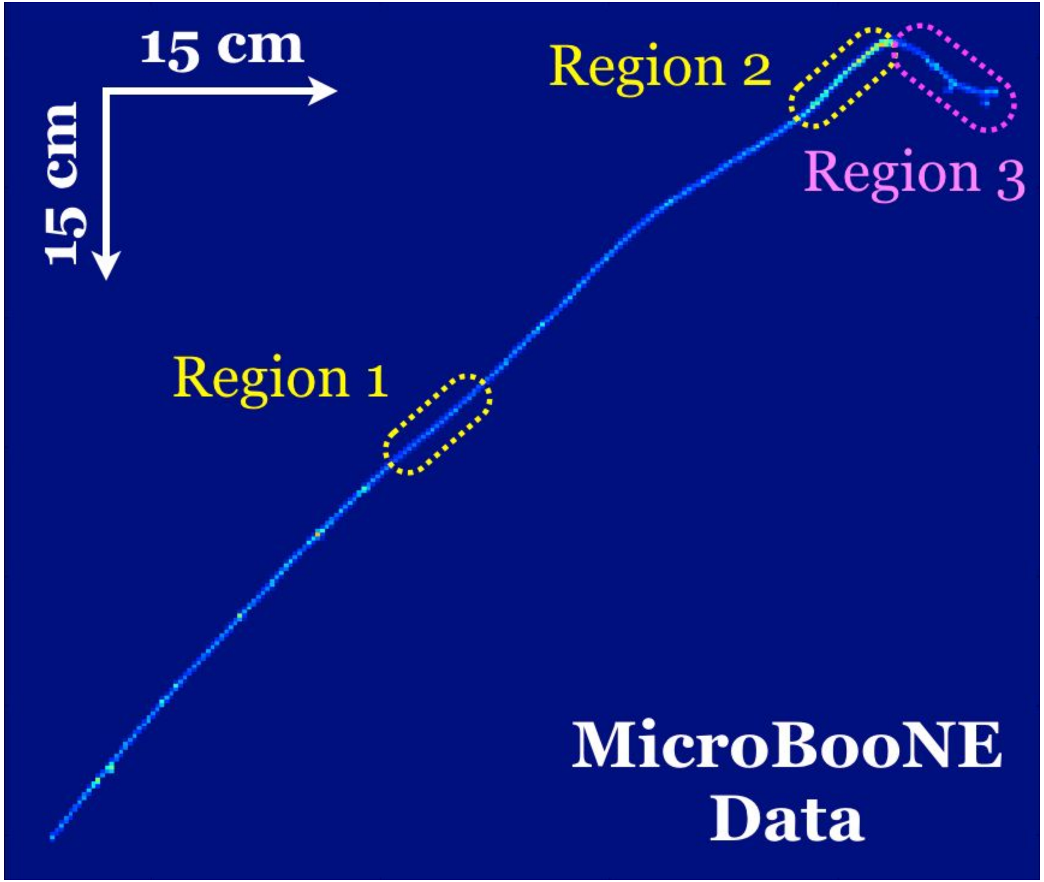




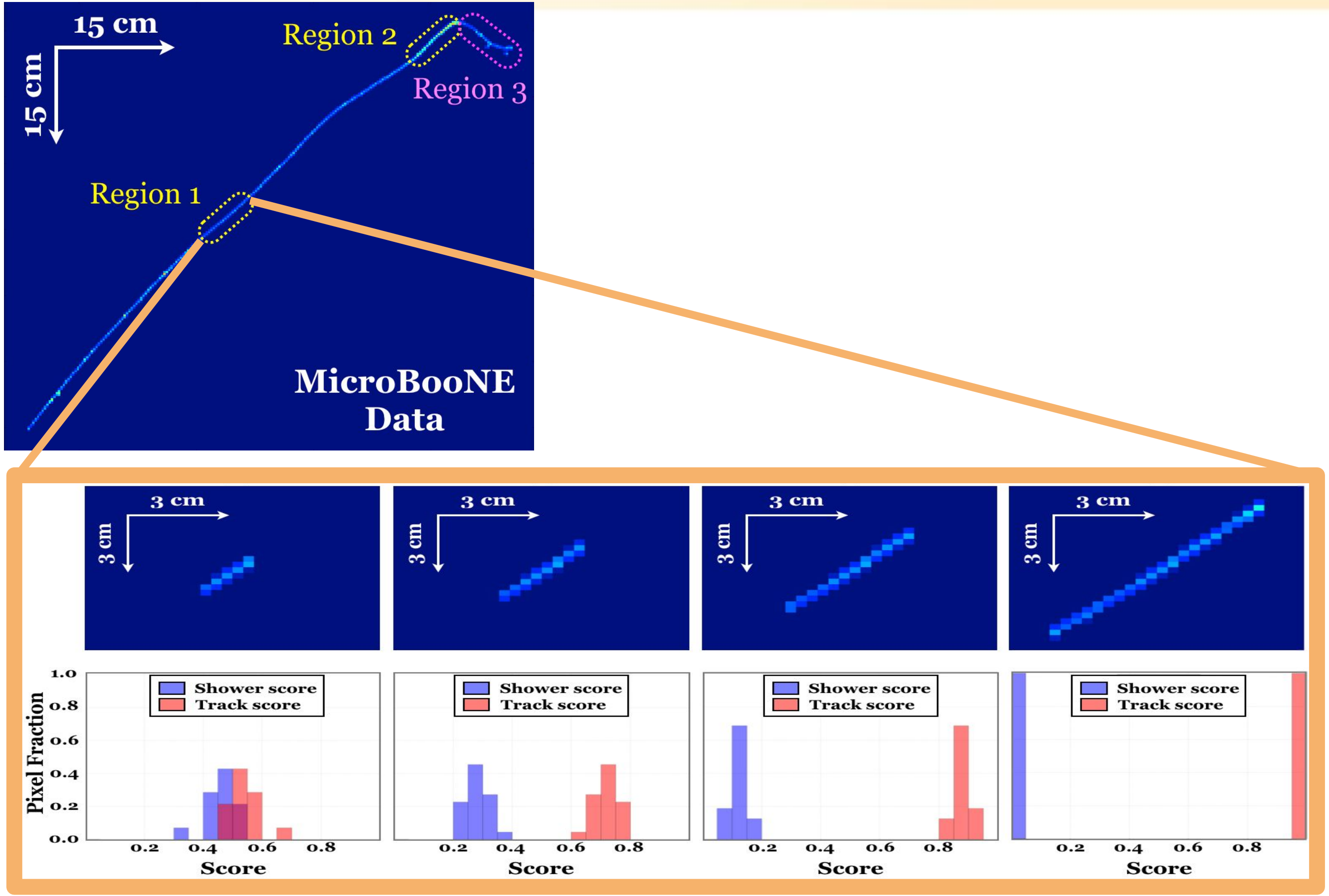
Number of strided convolutions, convolution layers, residual connections, differ in impementations

Image credit: Laura Domine @ Stanford

Fun Playing with Semantic Segmentation



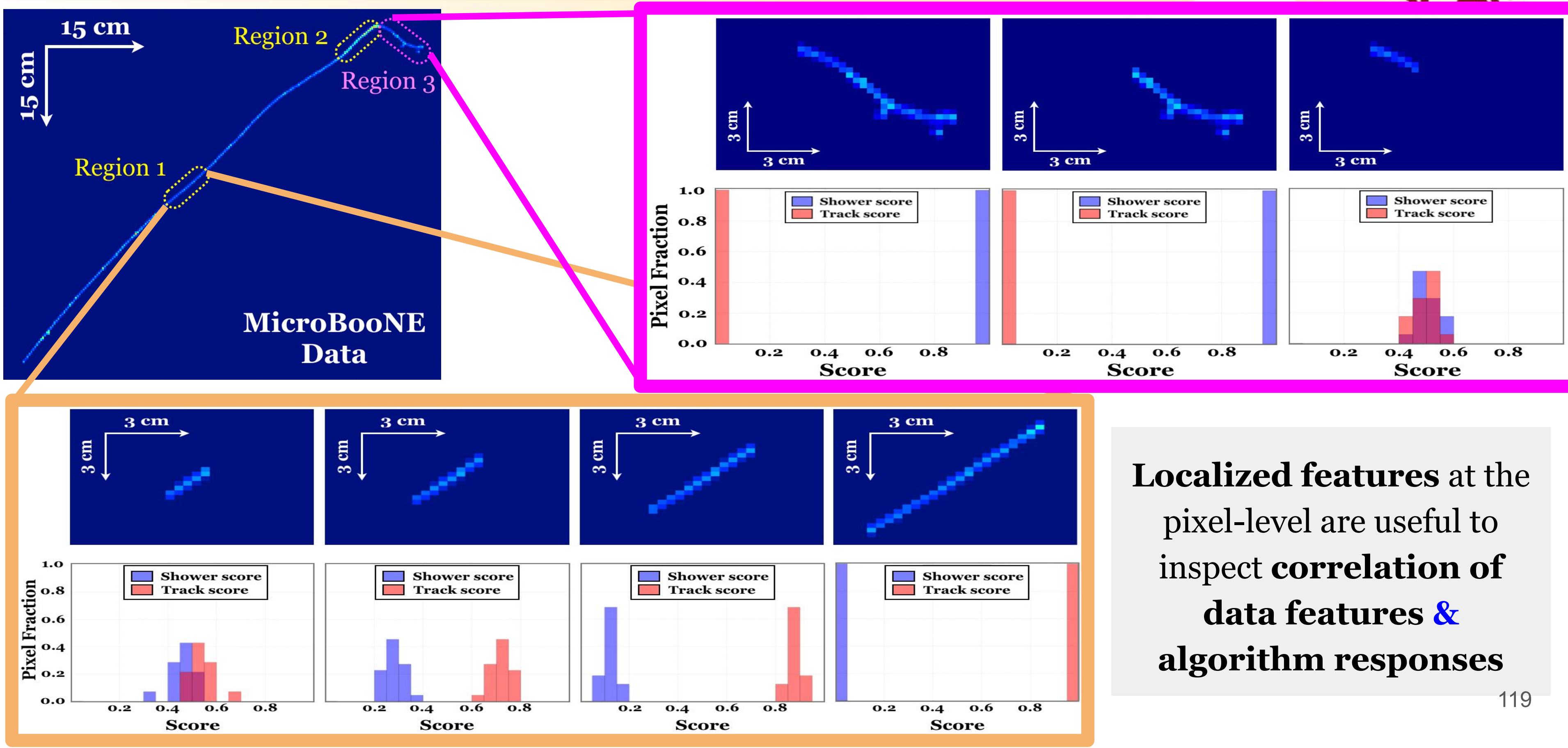
Fun Playing with Semantic Segmentation



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**