Deep Neural Network Applications for Particle Imaging Detectors

Kazuhiro Terao
SLAC National Laboratory

Outline for 20 minutes

1. About me & neutrinos
2. Image of particles
3. Application of DNNs
4. Summary

Neutrino physics since college :)

Me: Neutrino Physicist

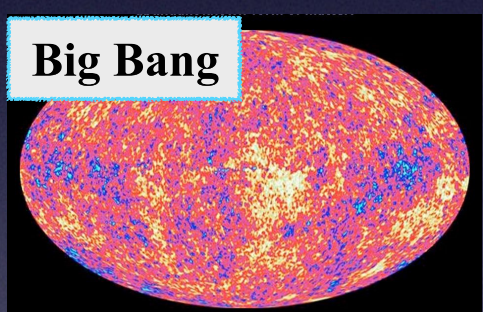
- **Neutrinos?**
  - Least Understood elementary particles
  - **They are everywhere**
    - 400 trillion neutrinos pass your body every second
    - Your body generates ~340 million neutrinos a day

Our Sun emits
10^{38} neutrinos per second

Me: Neutrino Physicist

- **Neutrinos?**
  - Least Understood elementary particles
  - **They are everywhere**
    - 400 trillion neutrinos pass your body every second
    - Your body generates ~340 million neutrinos a day
  - **They come from everywhere**

Our Sun emits 10^{38} neutrinos per second


Big Bang


SuperNova

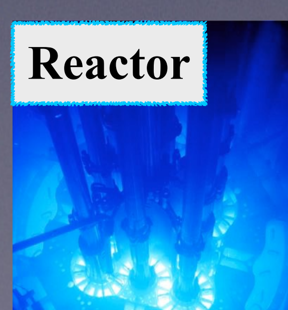

AGN


Sun


Atmospheric

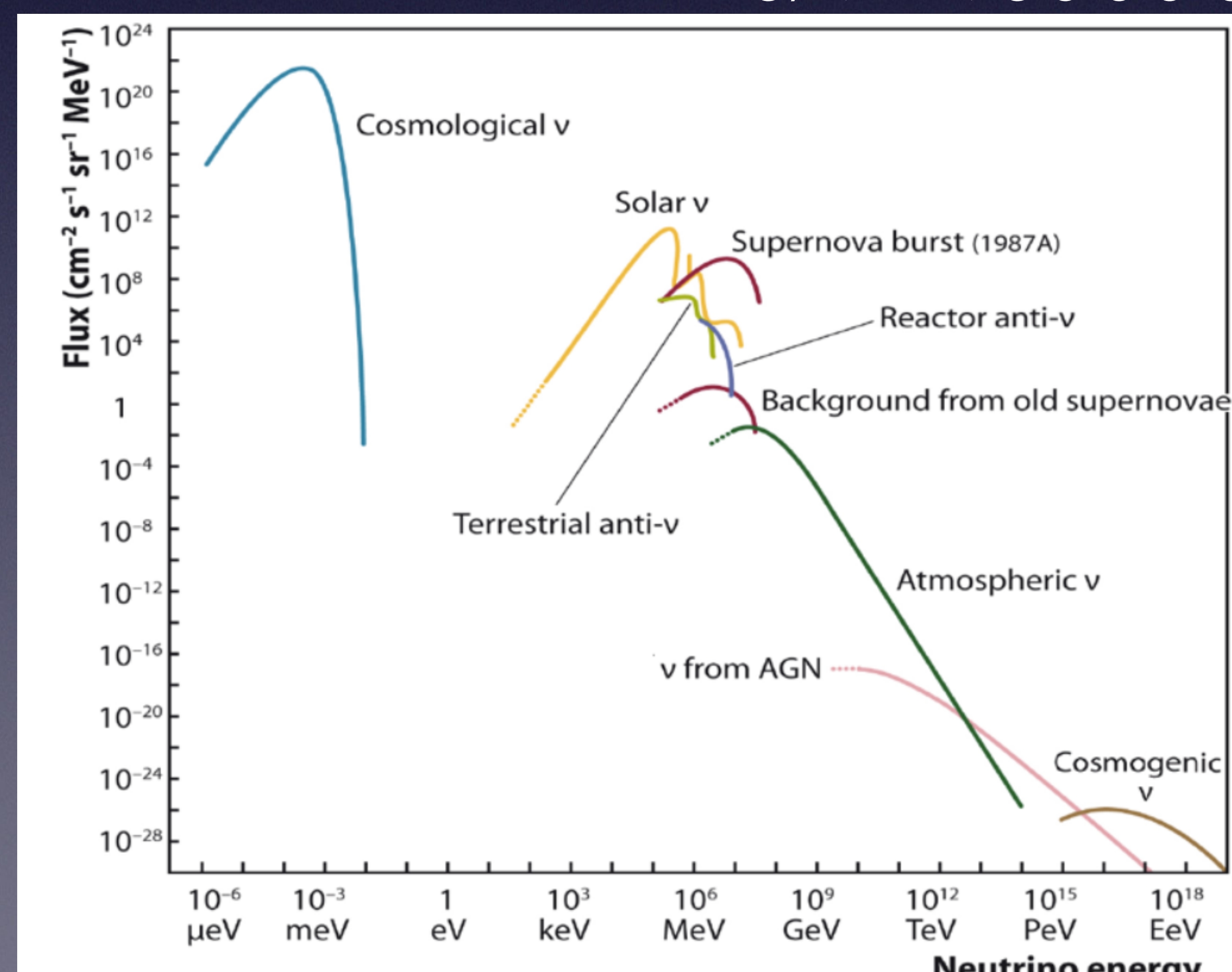

Earth


Accelerator

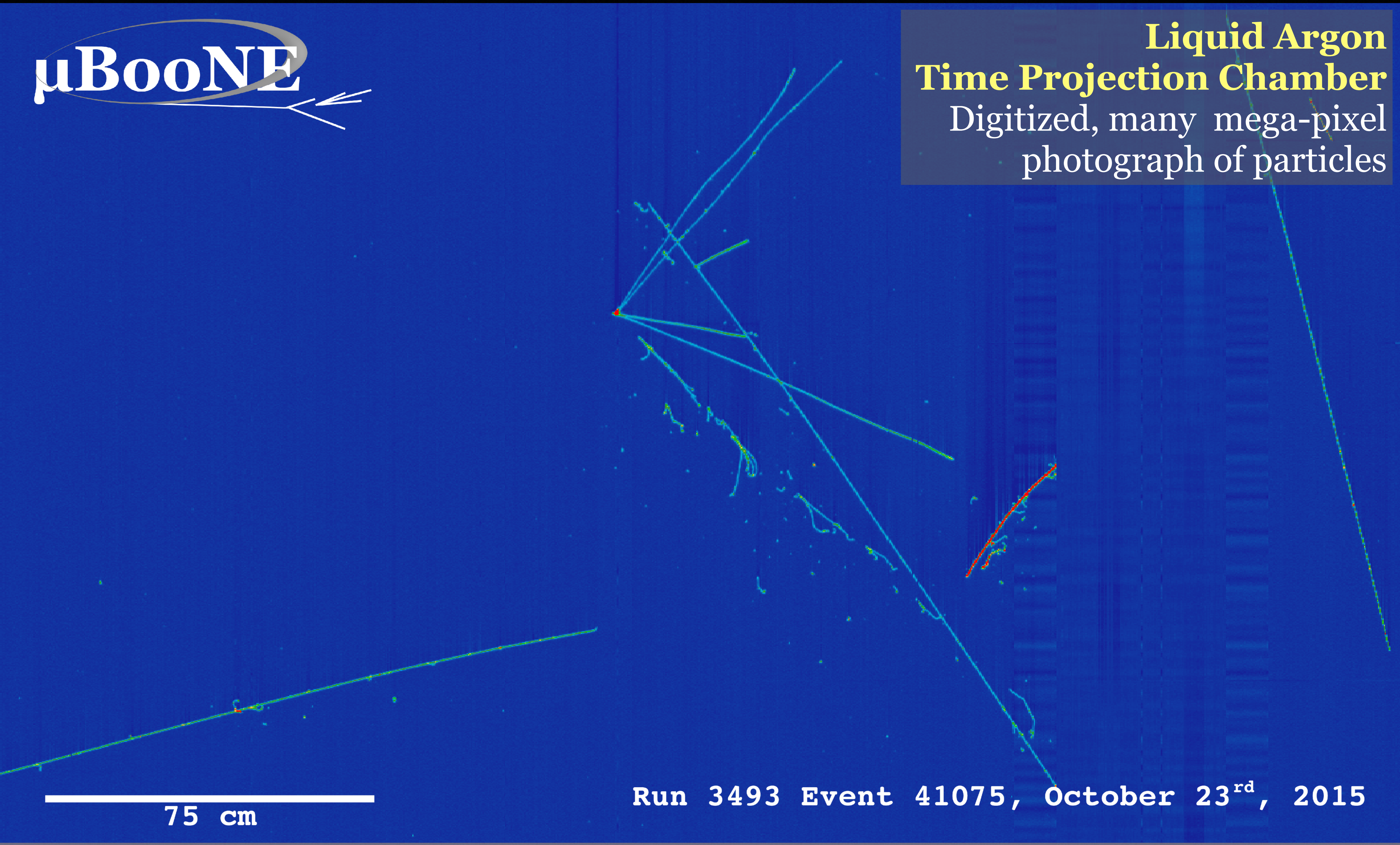

Reactor


Good Stuff

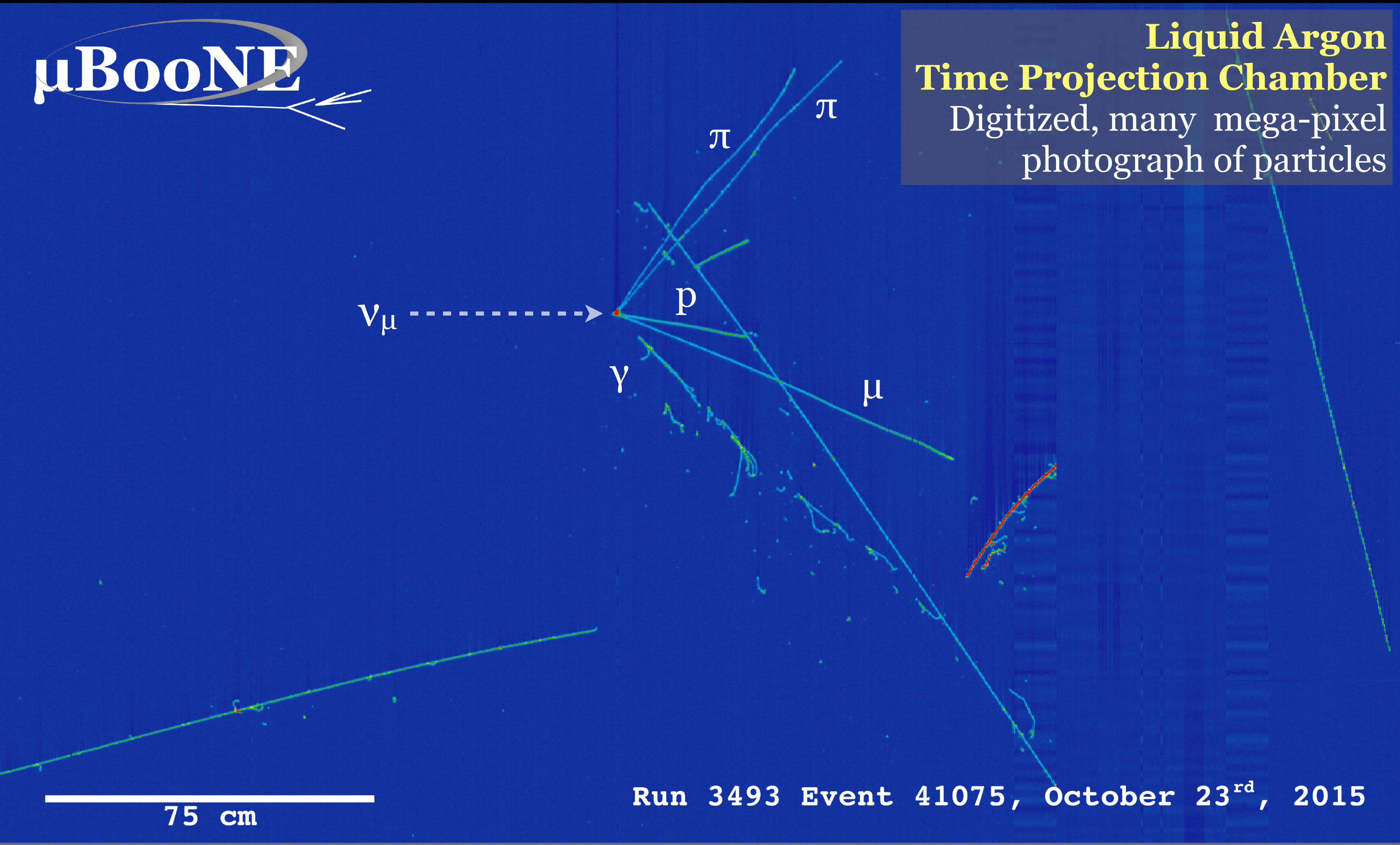
EPJ H37 (2012) 3:515-565



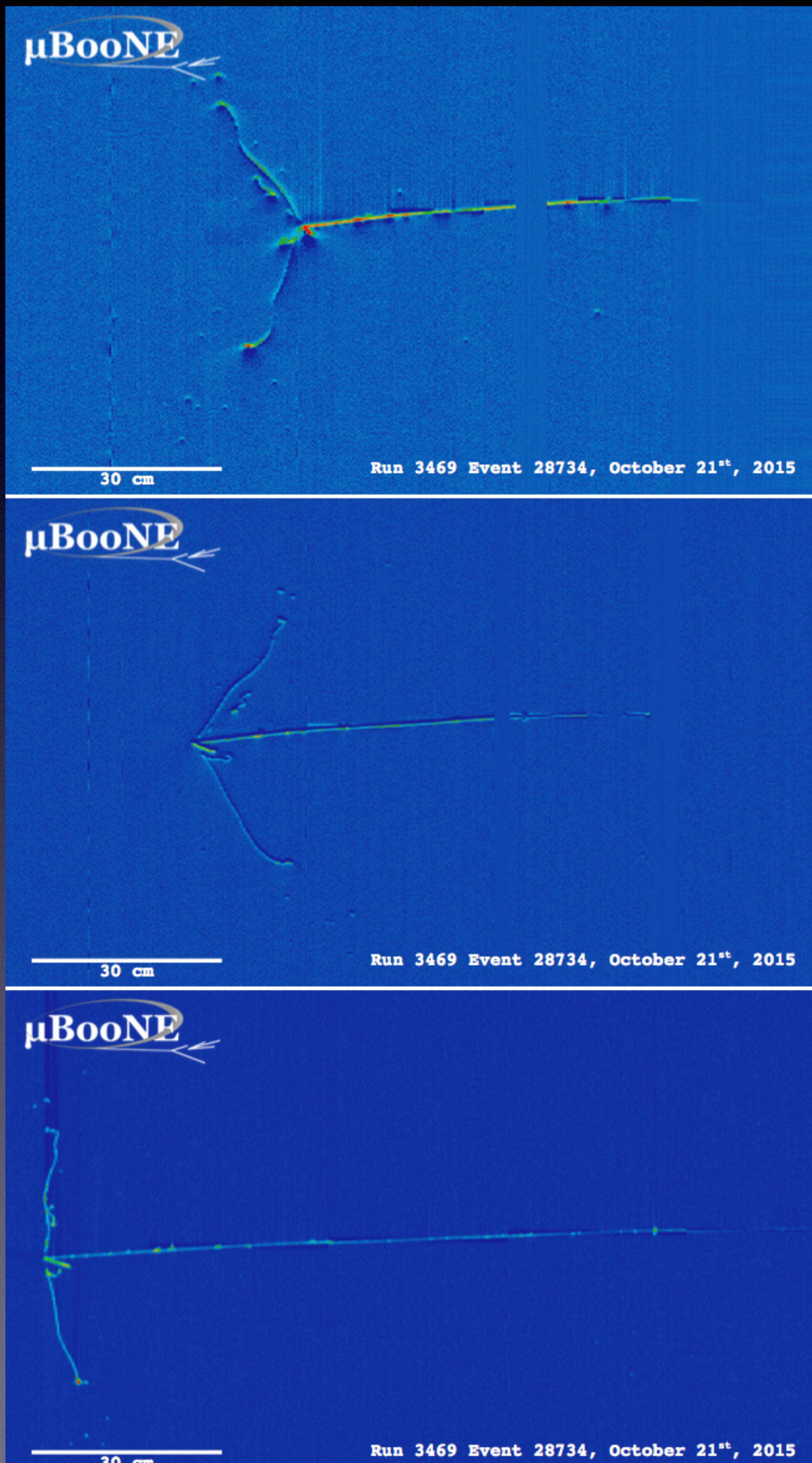
Detecting Neutrinos: Today



Detecting Neutrinos: Today

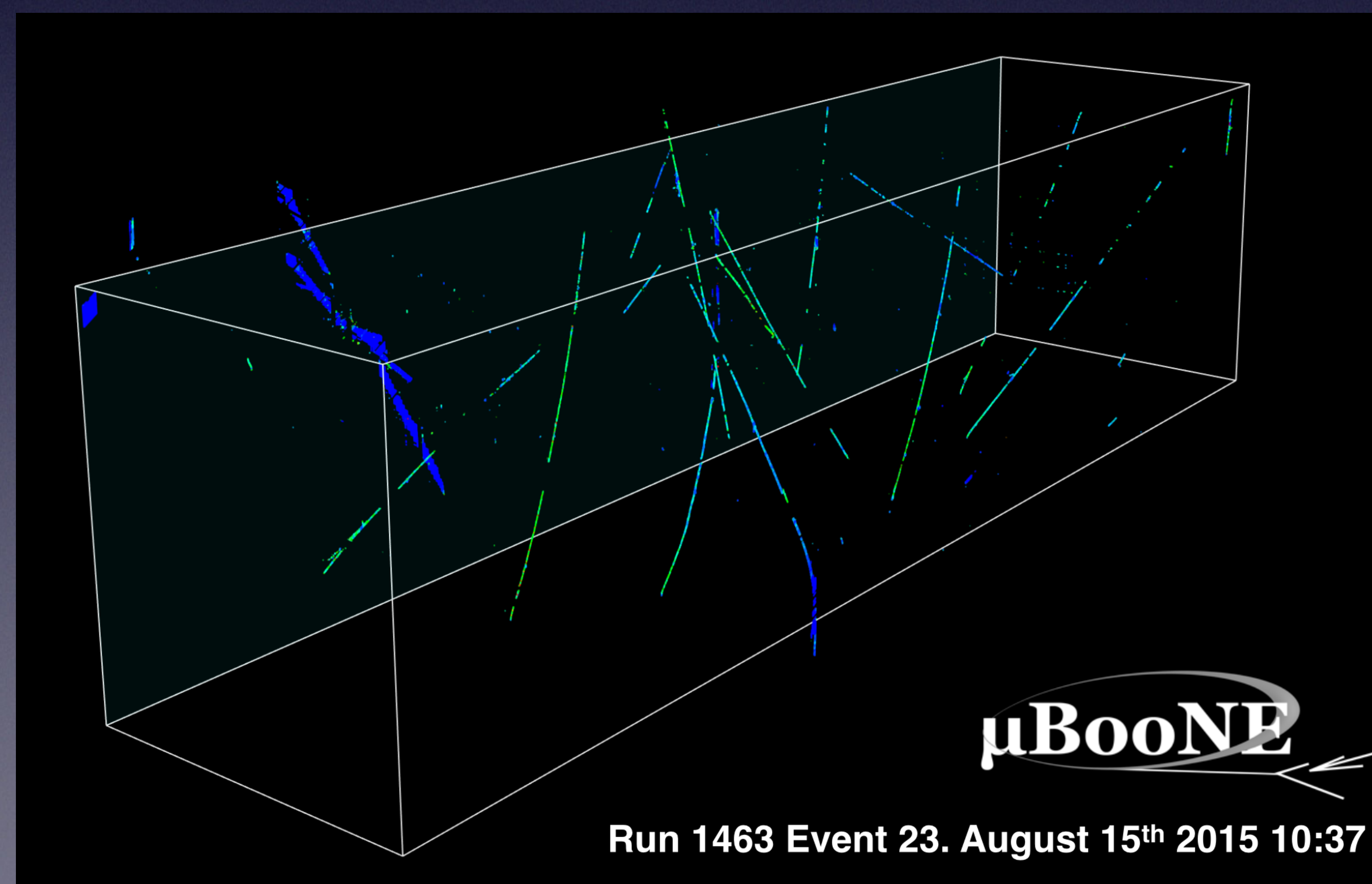


Detecting Neutrinos: Today



Type-A
Multiple 2D Projections

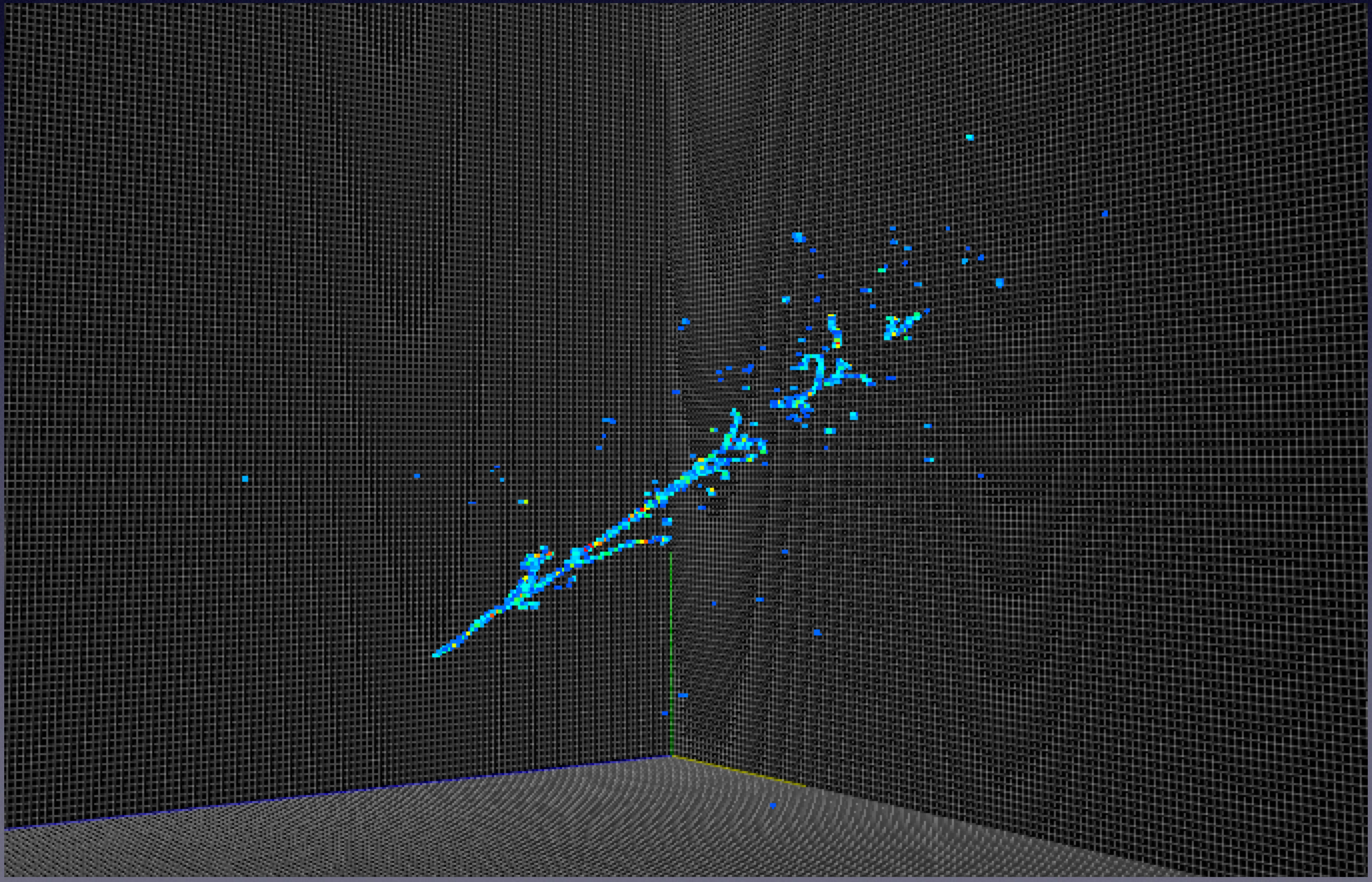
Main stream in neutrino LArTPC detectors, require data reconstruction for 3D imaging



Detecting Neutrinos: Today

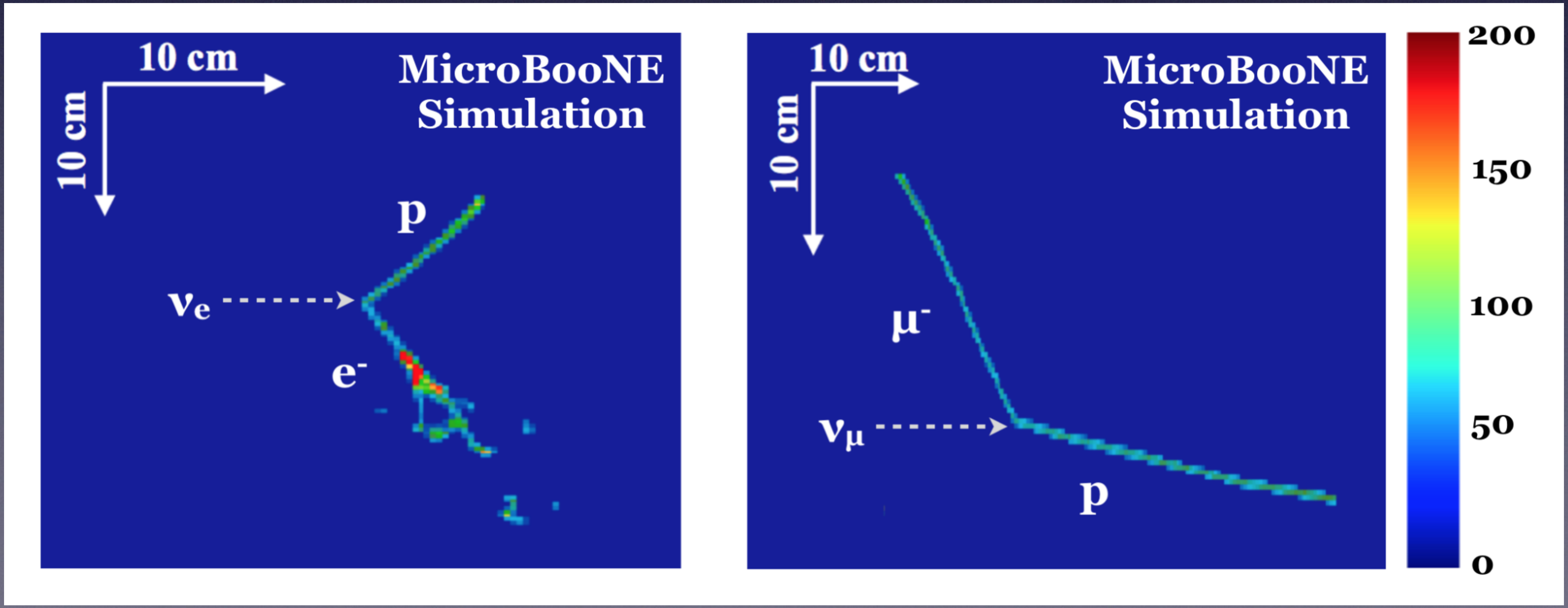
Type-B
3D Imaging Detector

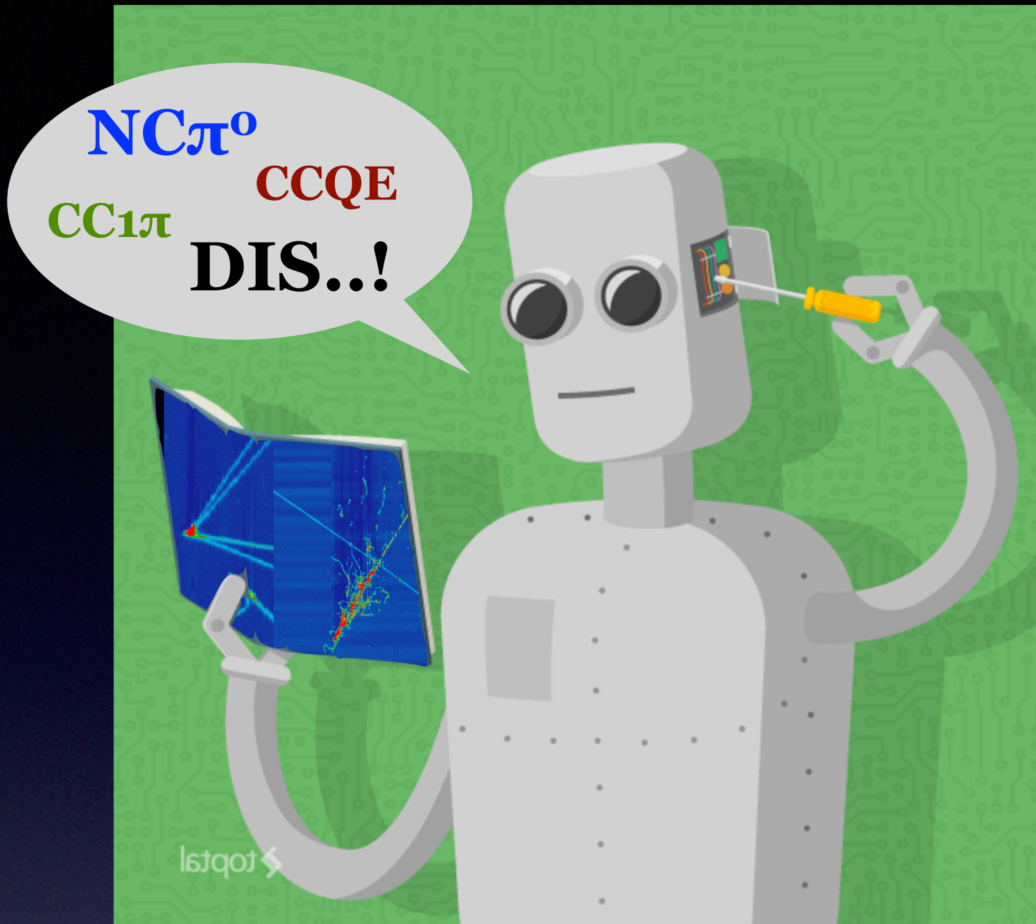
Next generation LArTPC, currently
R&D on-going

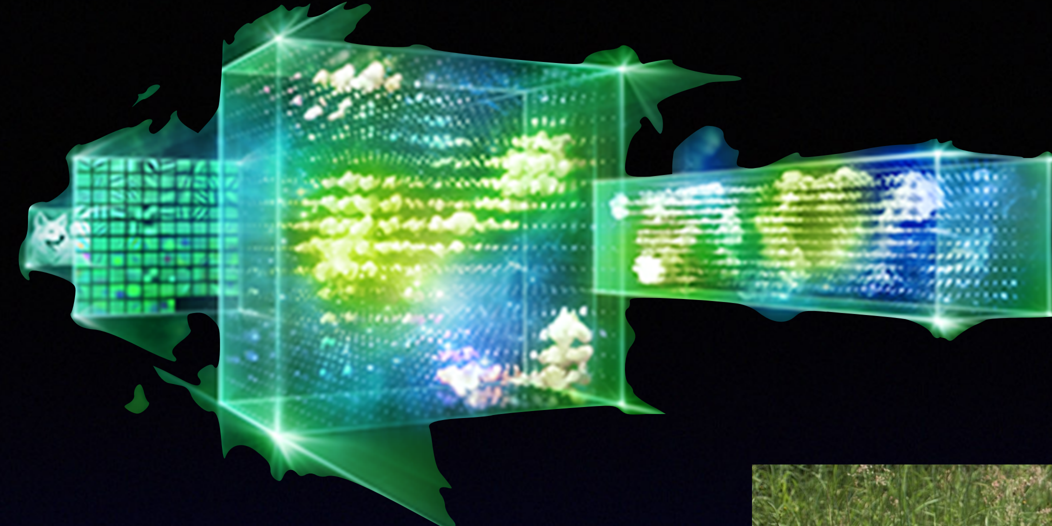


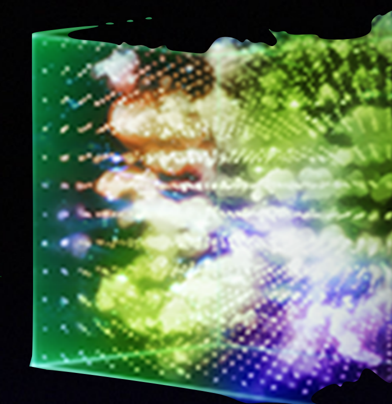
Analysis Goals

- **Find neutrino** interaction vertex
- Identify neutrino ***type***
- Reconstruct neutrino ***energy***









Outline for 20 minutes

1. About me & neutrinos
2. Image of particles
3. **Application of DNNs**
4. Summary

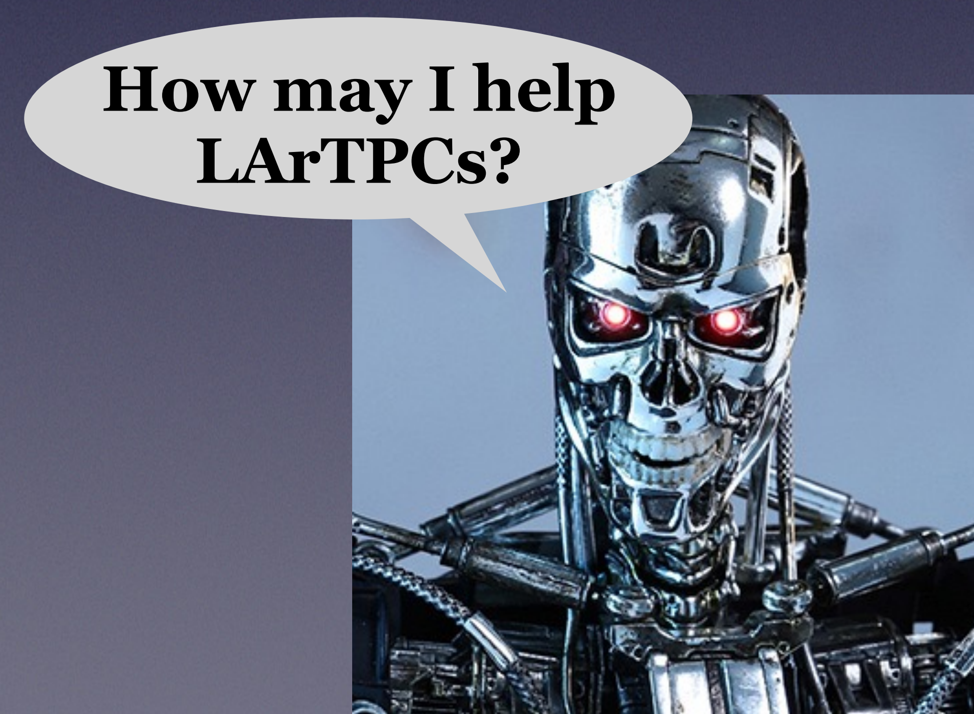
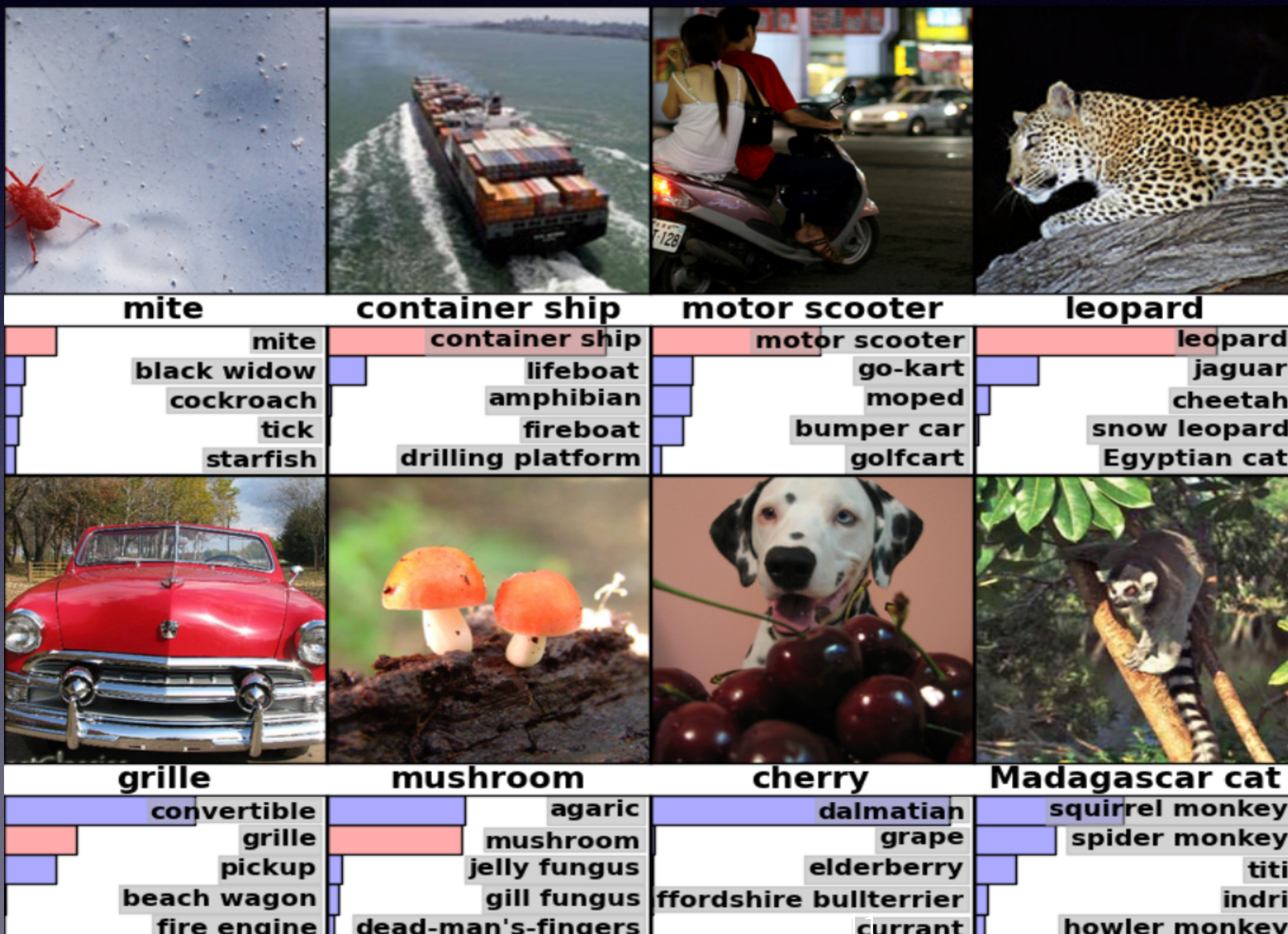
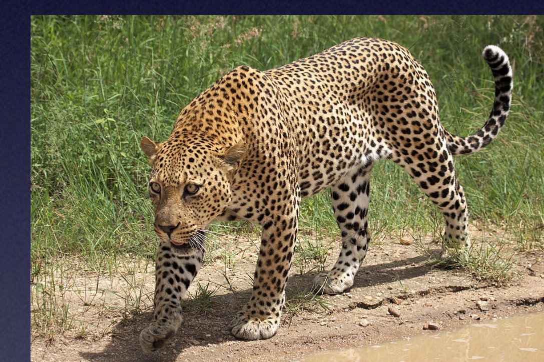


Image Classification by DNNs

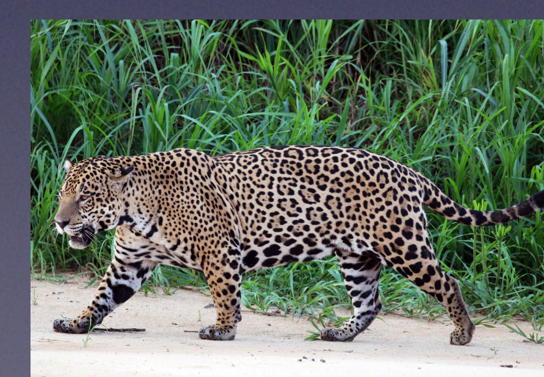
DNN has been the driver for the recent advancement in computer vision, the first breakthrough in image classification tasks (clearly better than me).



For my reference



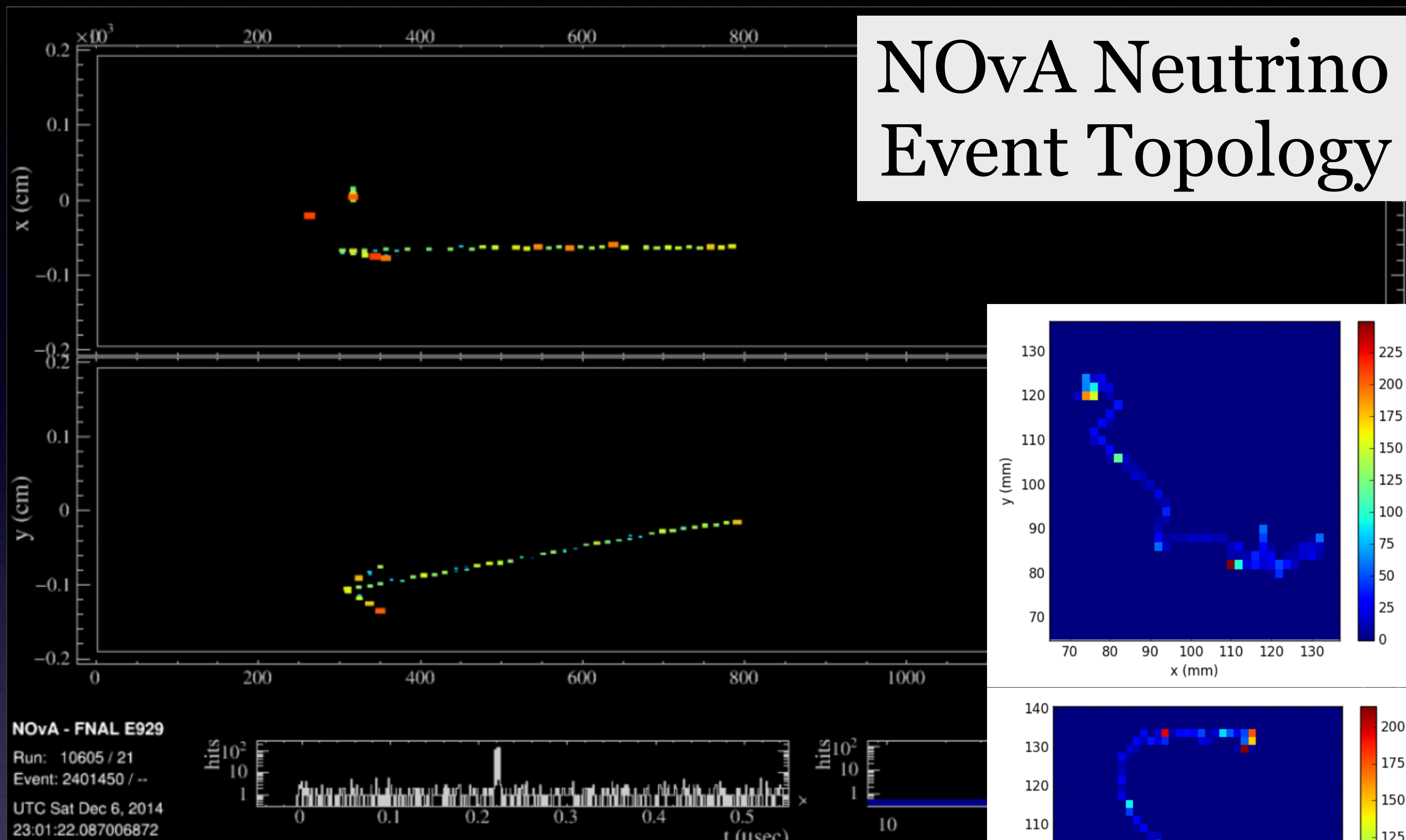
Leopard



Jaguar

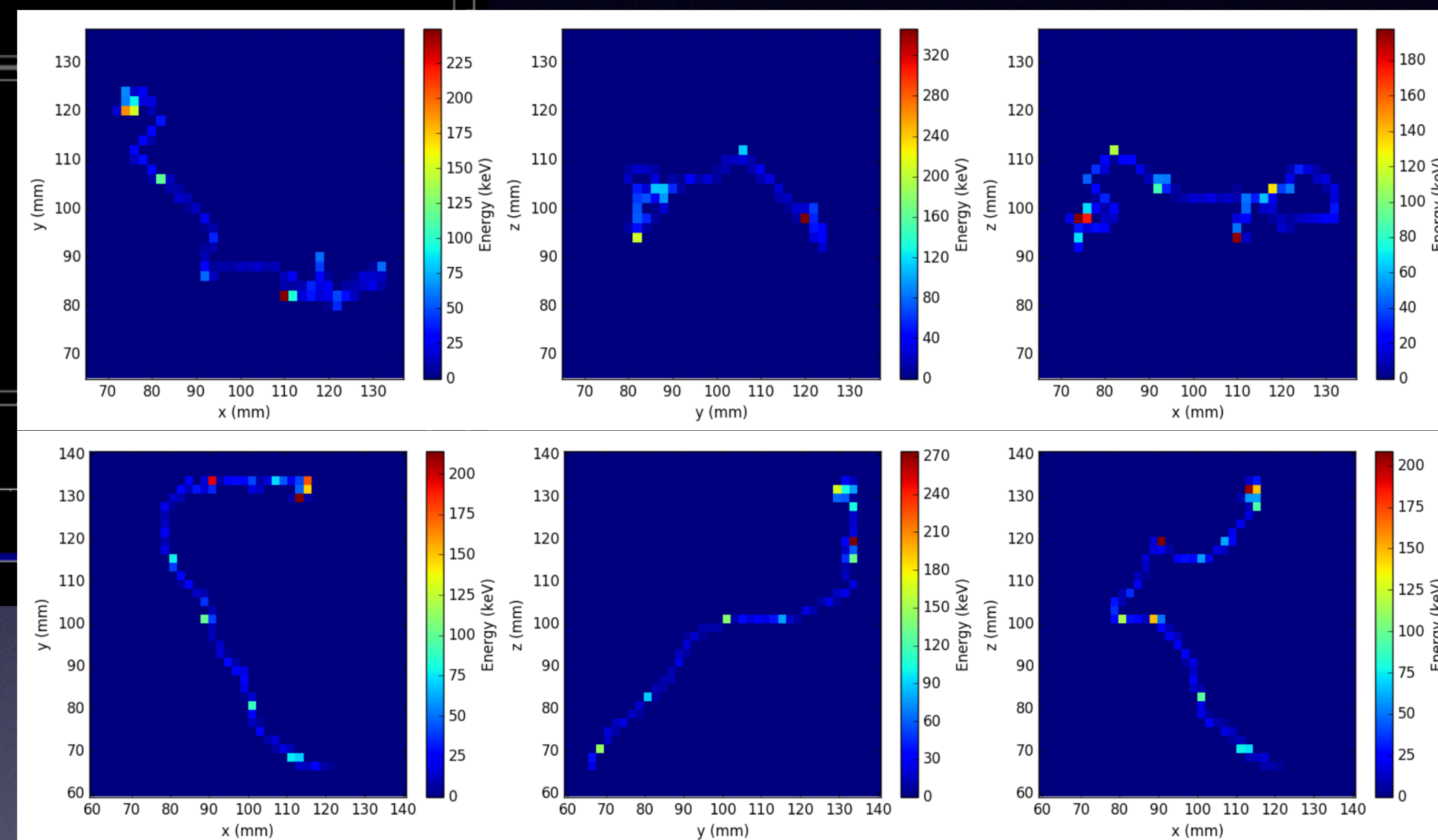
From 2012 AlexNet
Today's results much better

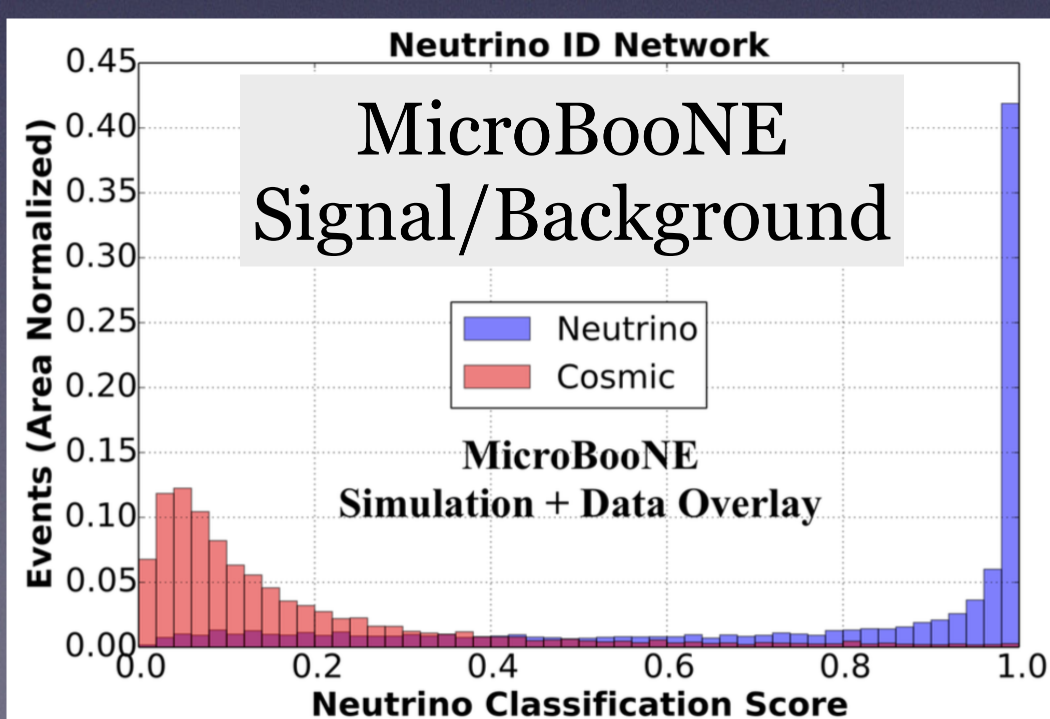
Image Classification for Physics Analysis



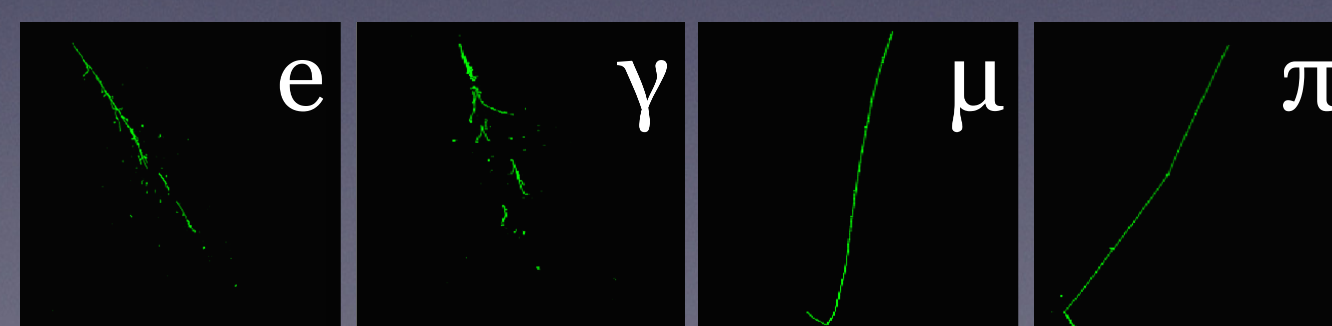
NOvA Neutrino
Event Topology

NEXT
Signal vs. Background





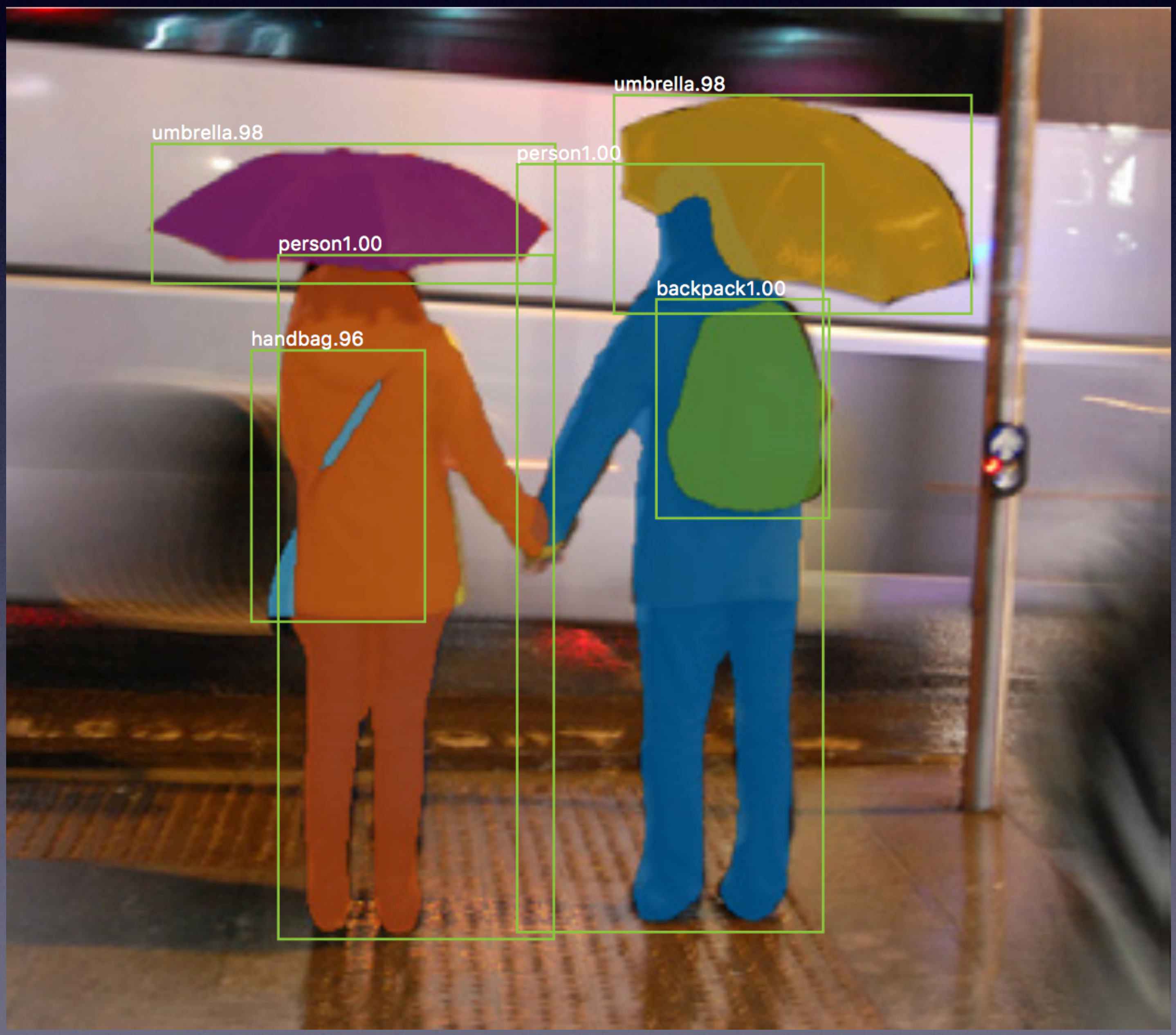
MicroBooNE
Signal/Background

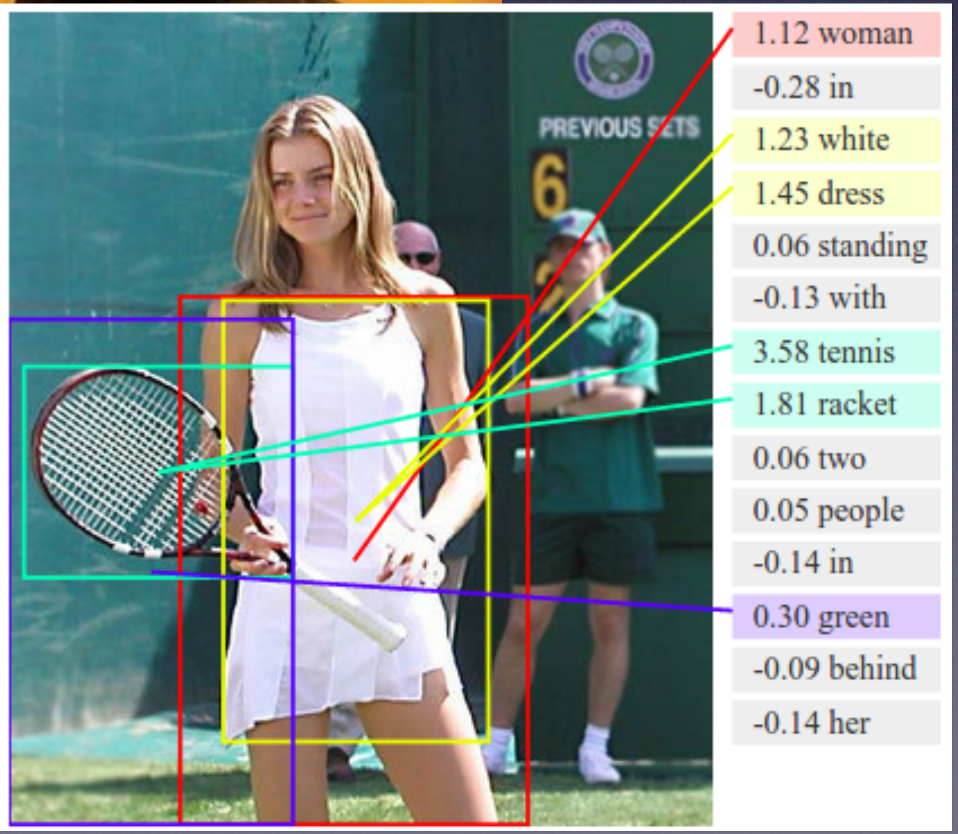


MicroBooNE Particle ID

Beyond Image Classification

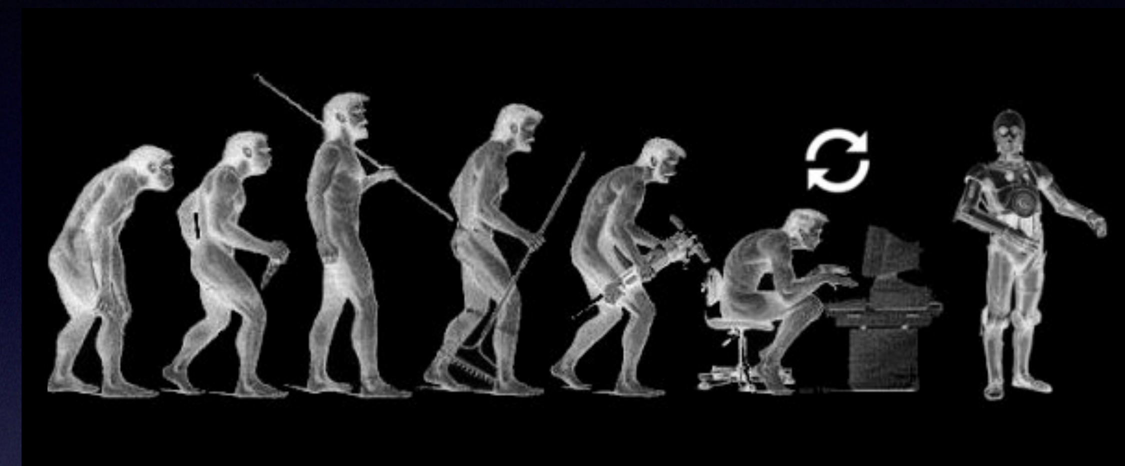
Wide variety of applications: classification, detection, pixel-level component analysis, natural language processing



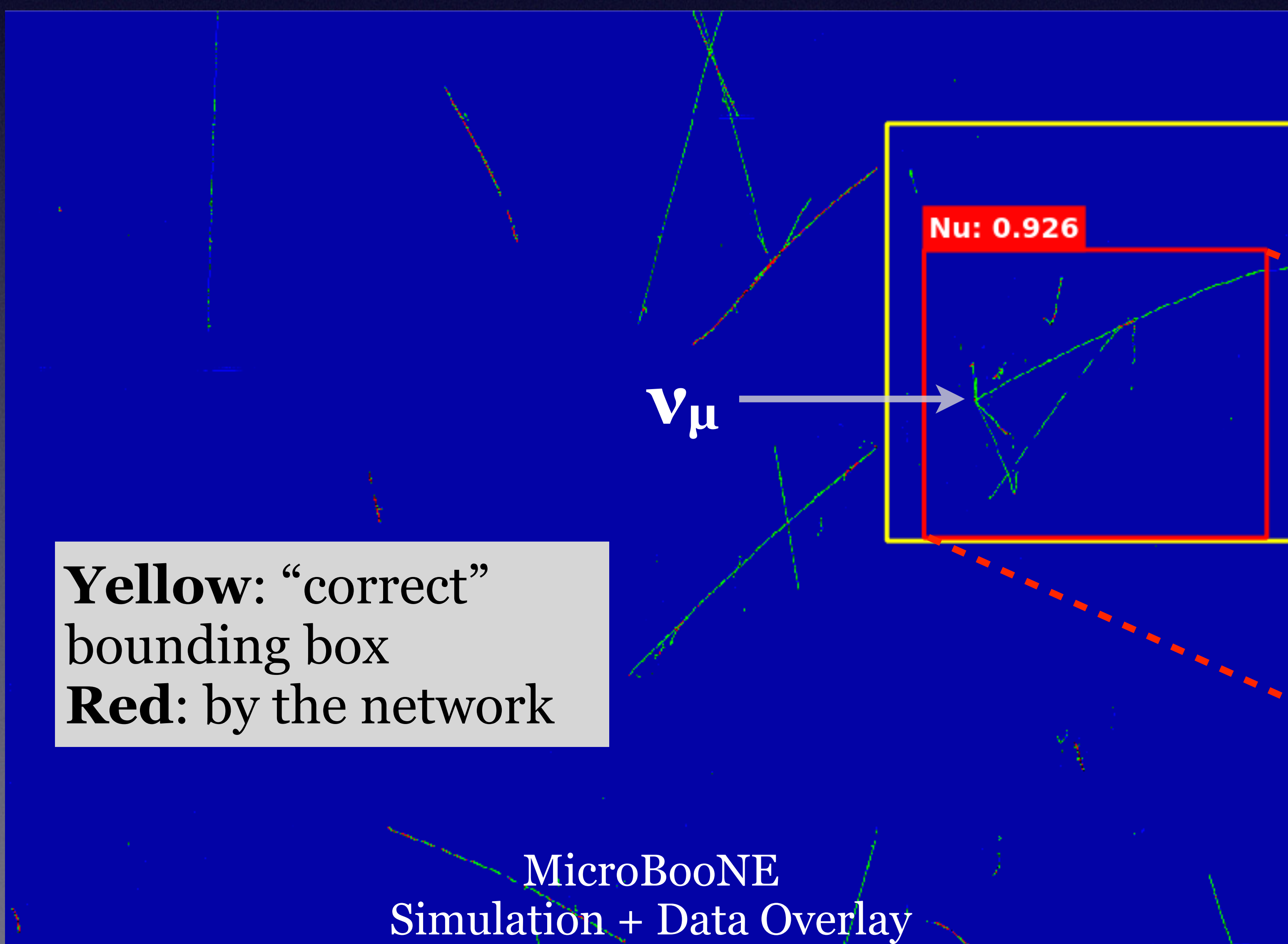


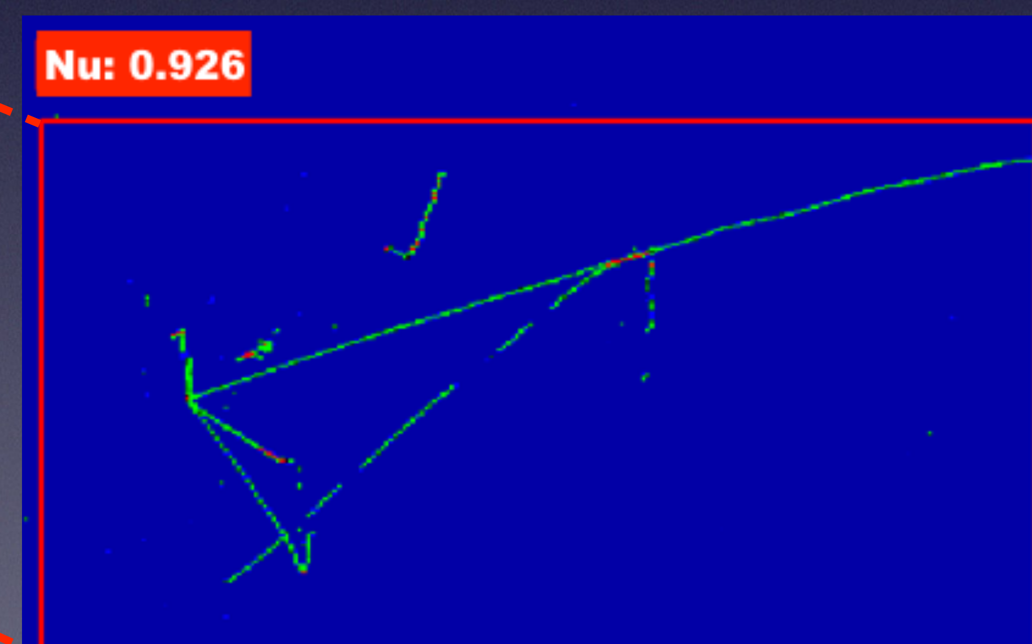
Application for Physics Analysis

- Object detection technique applied to localize neutrino interaction region in MicroBooNE data
- DNNs for “feature mining”



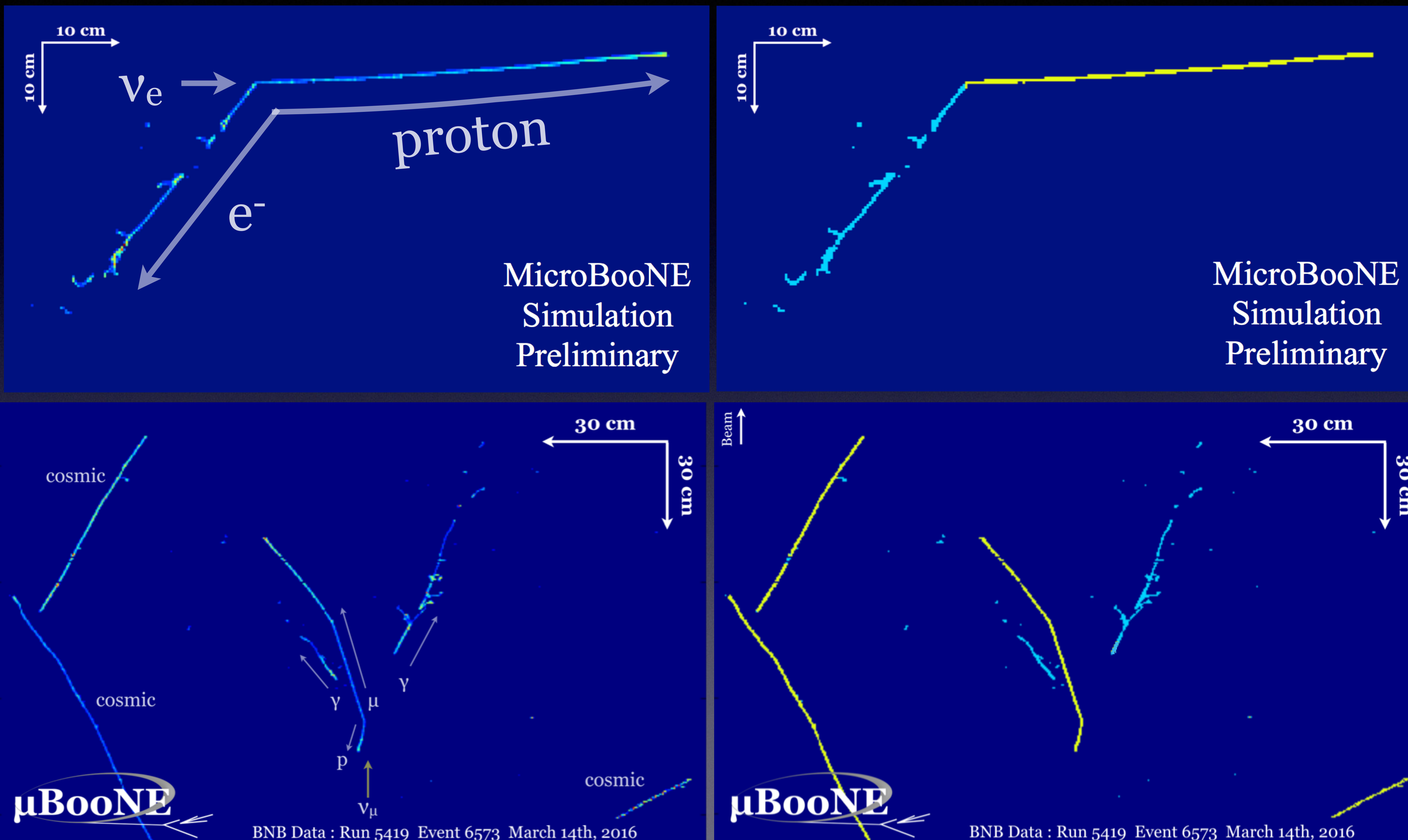
arxiv:1611.05531





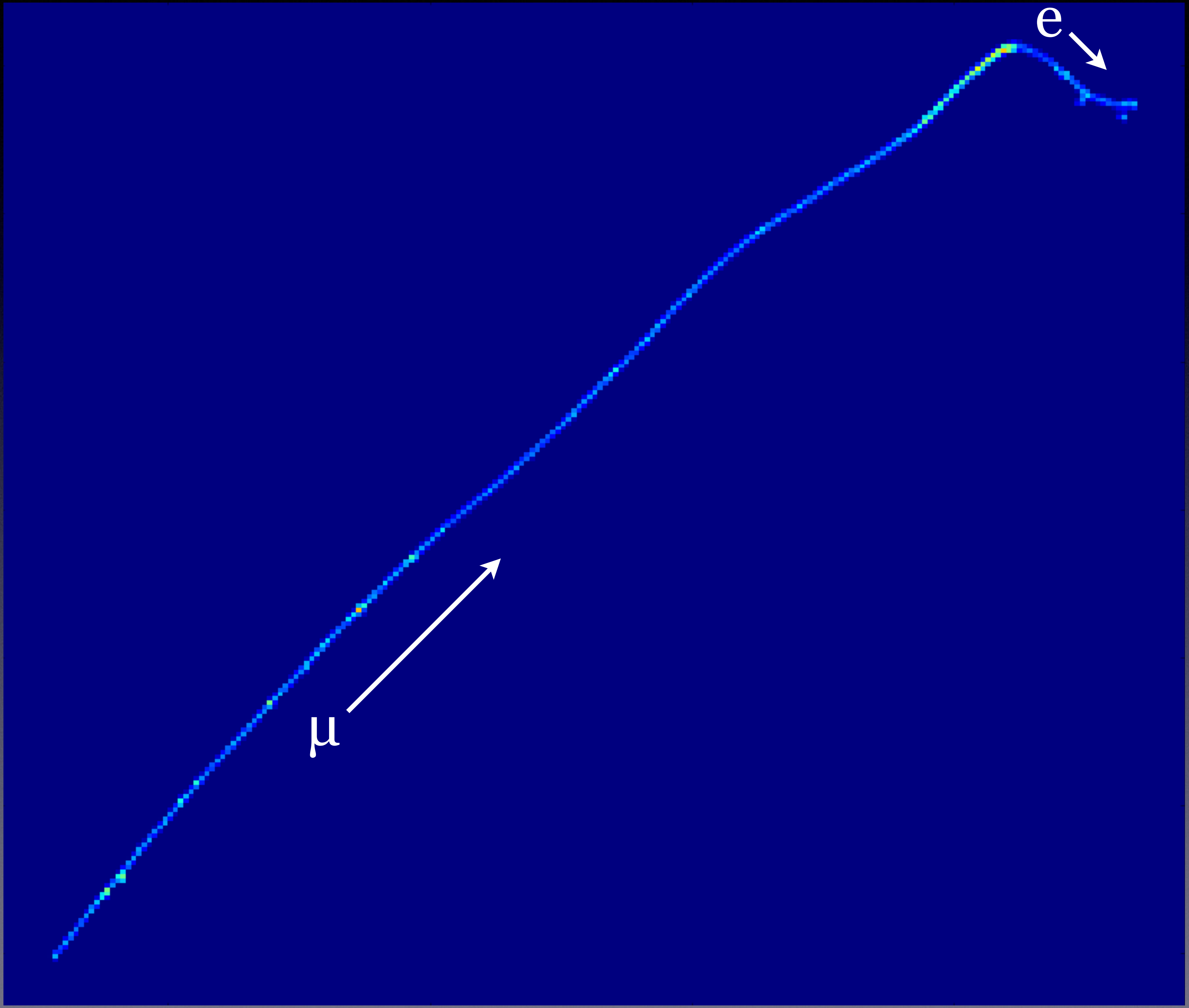
Network Output
≃ 2.6m (width) x 1 m (height)

Pixel-level Identification of Electrons

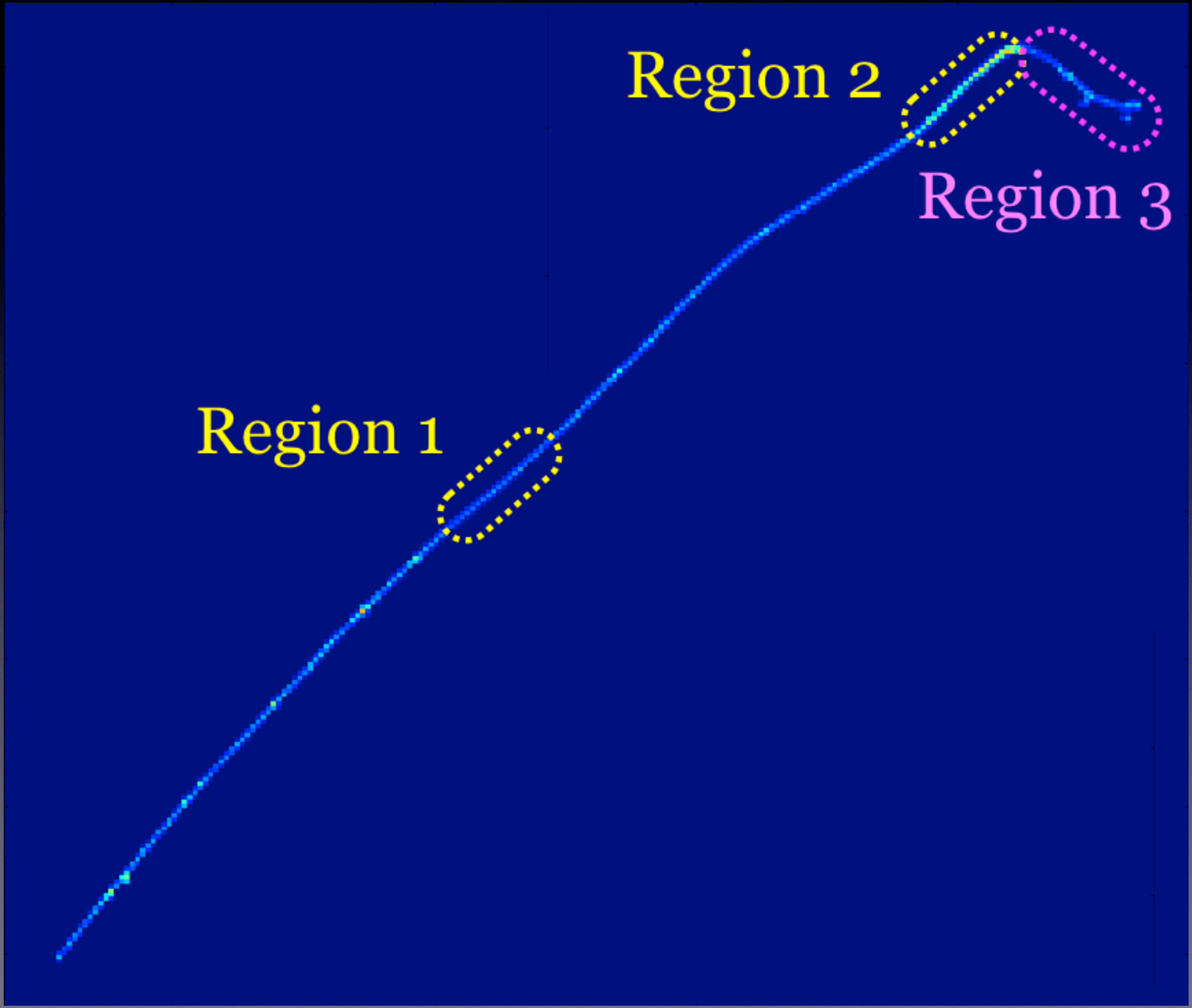


Network Input **Network Output**

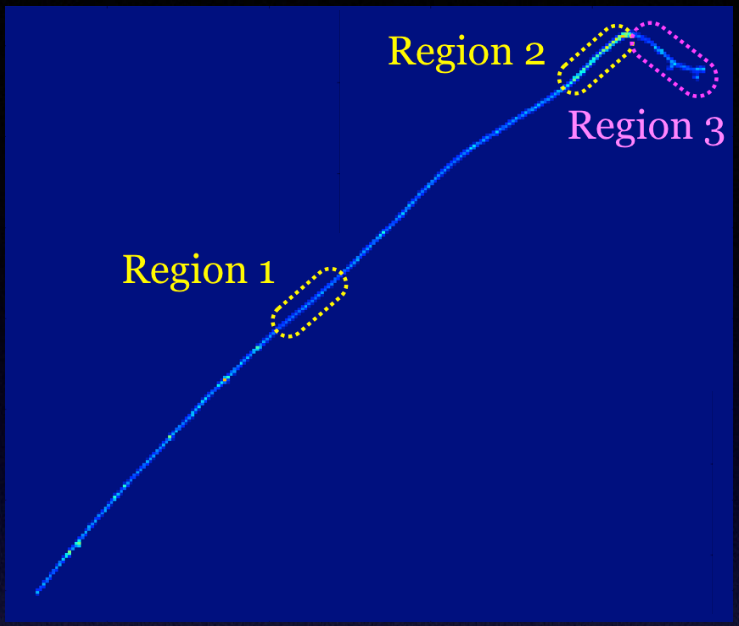
Extra: Qualitative Analysis



Extra: Qualitative Analysis

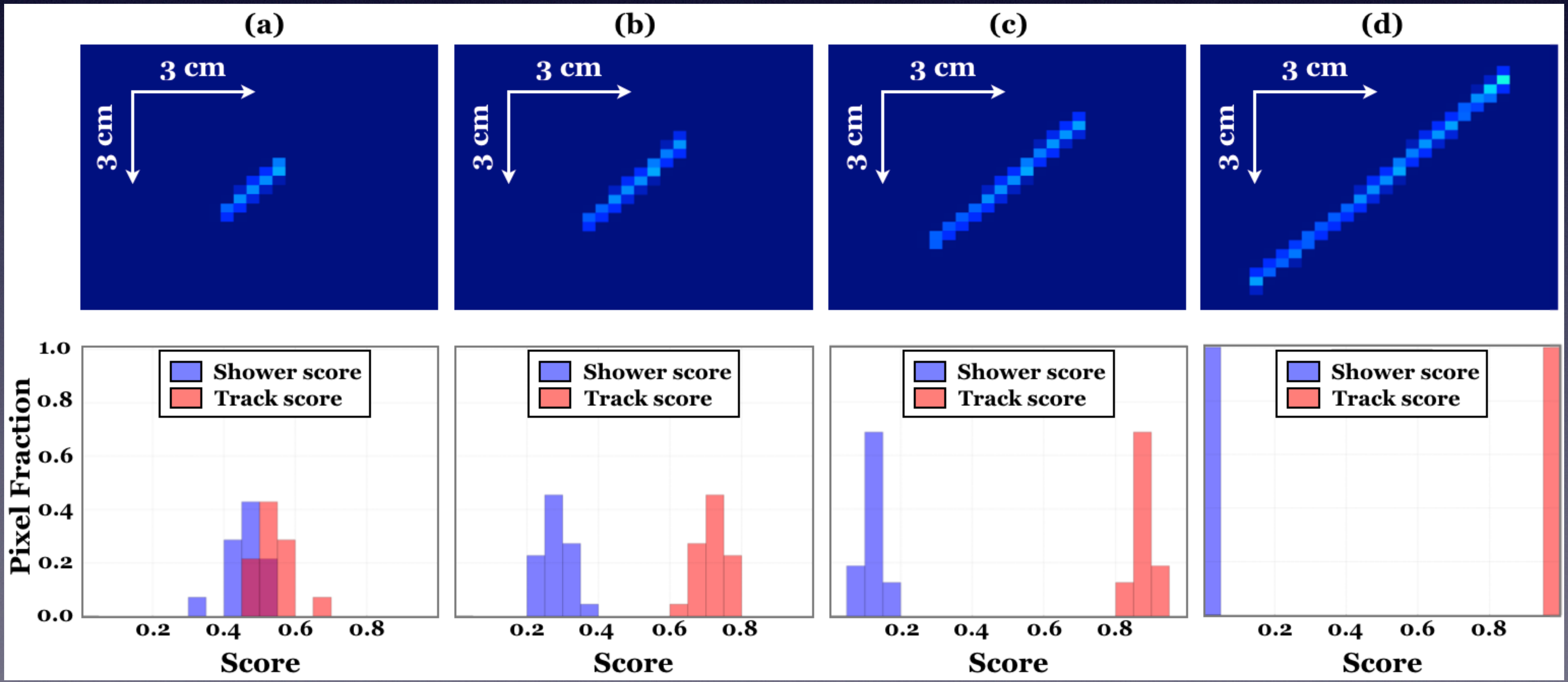


Extra: Qualitative Analysis

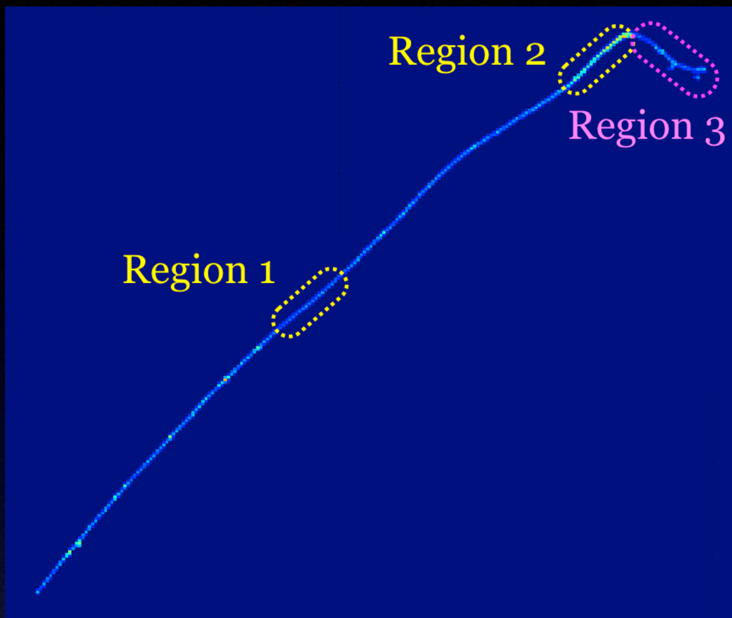


Region 1

- Mask all pixels but a small fraction of trajectory. Inspected how the network response changes as a function of trajectory length

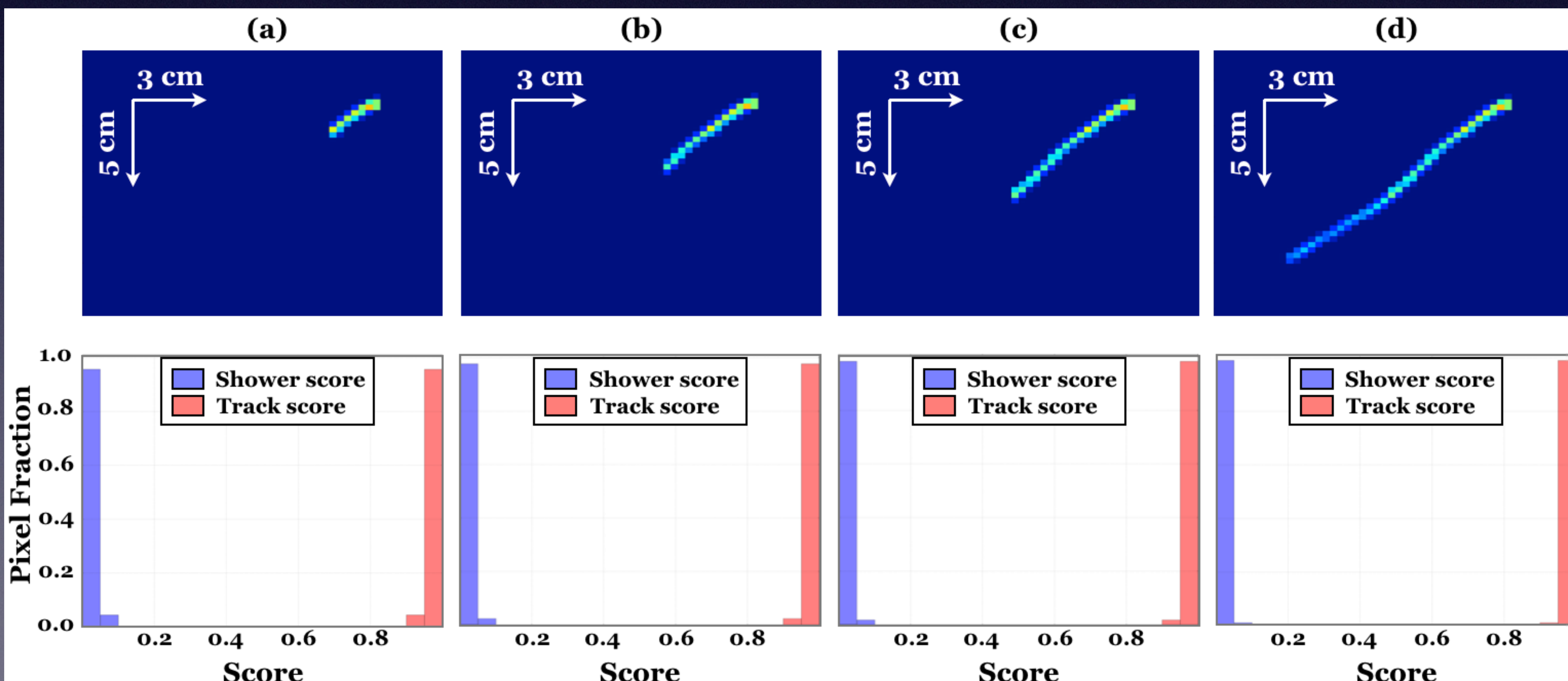


Extra: Qualitative Analysis

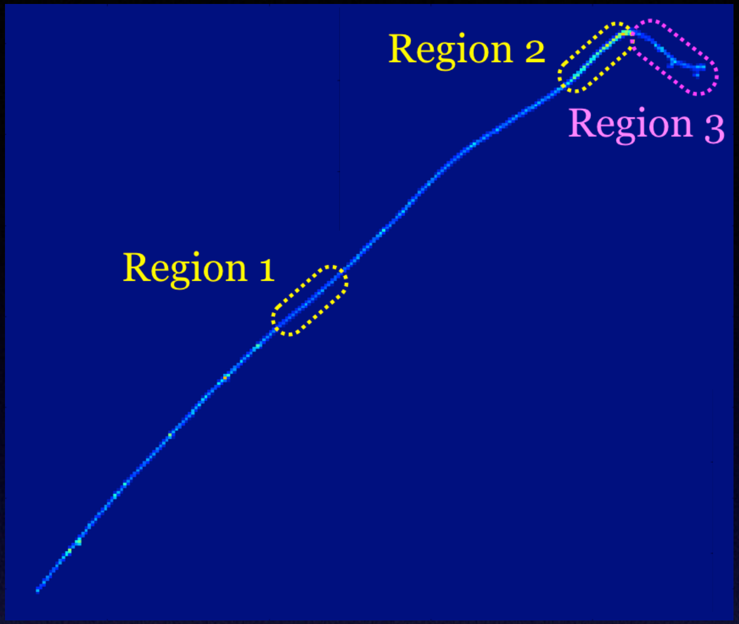


Region 2

- Mask all pixels but a small fraction of trajectory. Inspected how the network response changes as a function of trajectory length where dE/dX is high.

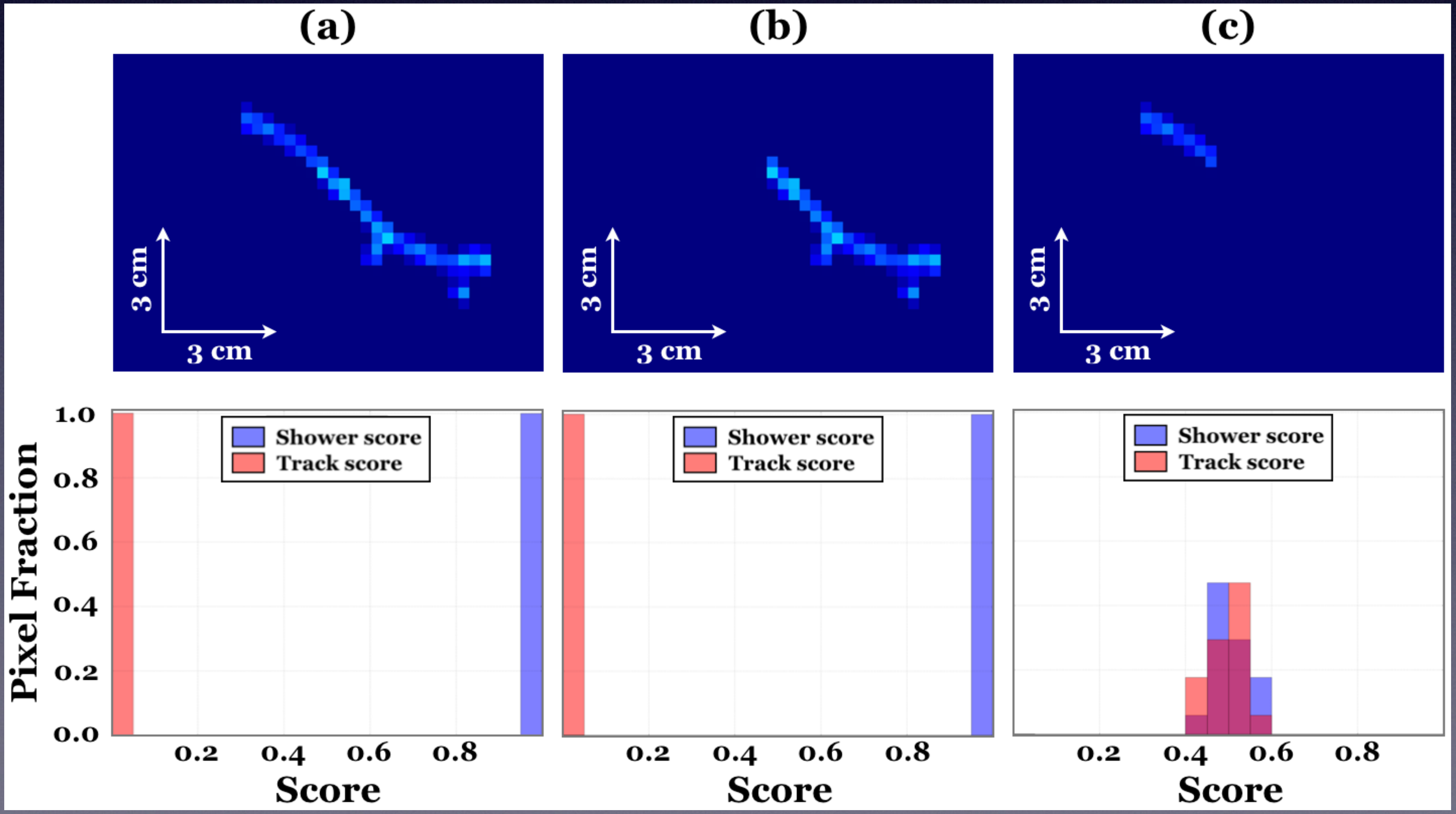


Extra: Qualitative Analysis

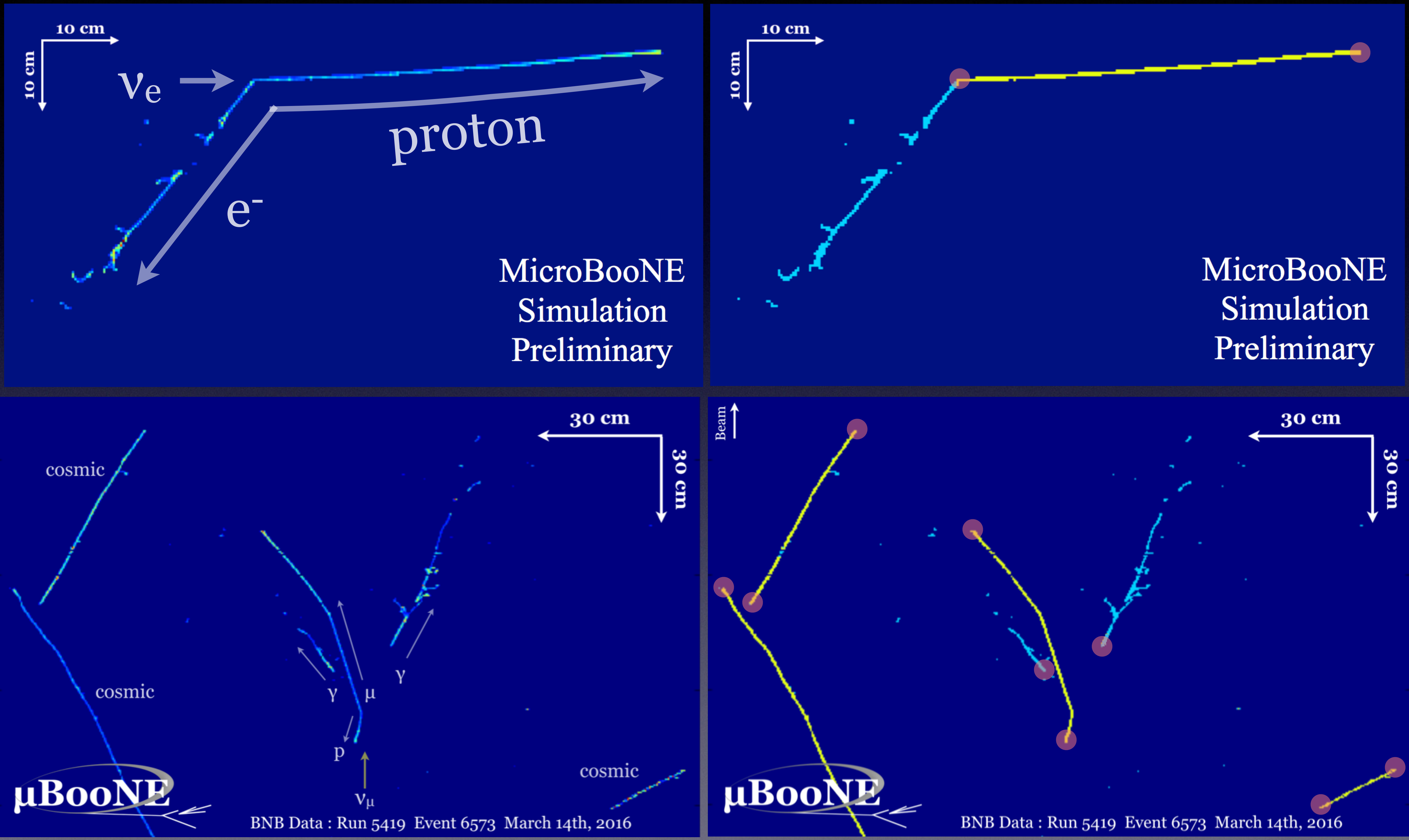


Region 3

- Mask all pixels but a small fraction of trajectory. Inspected how topological difference affect the pixel score, and correlation to neighboring pixels.

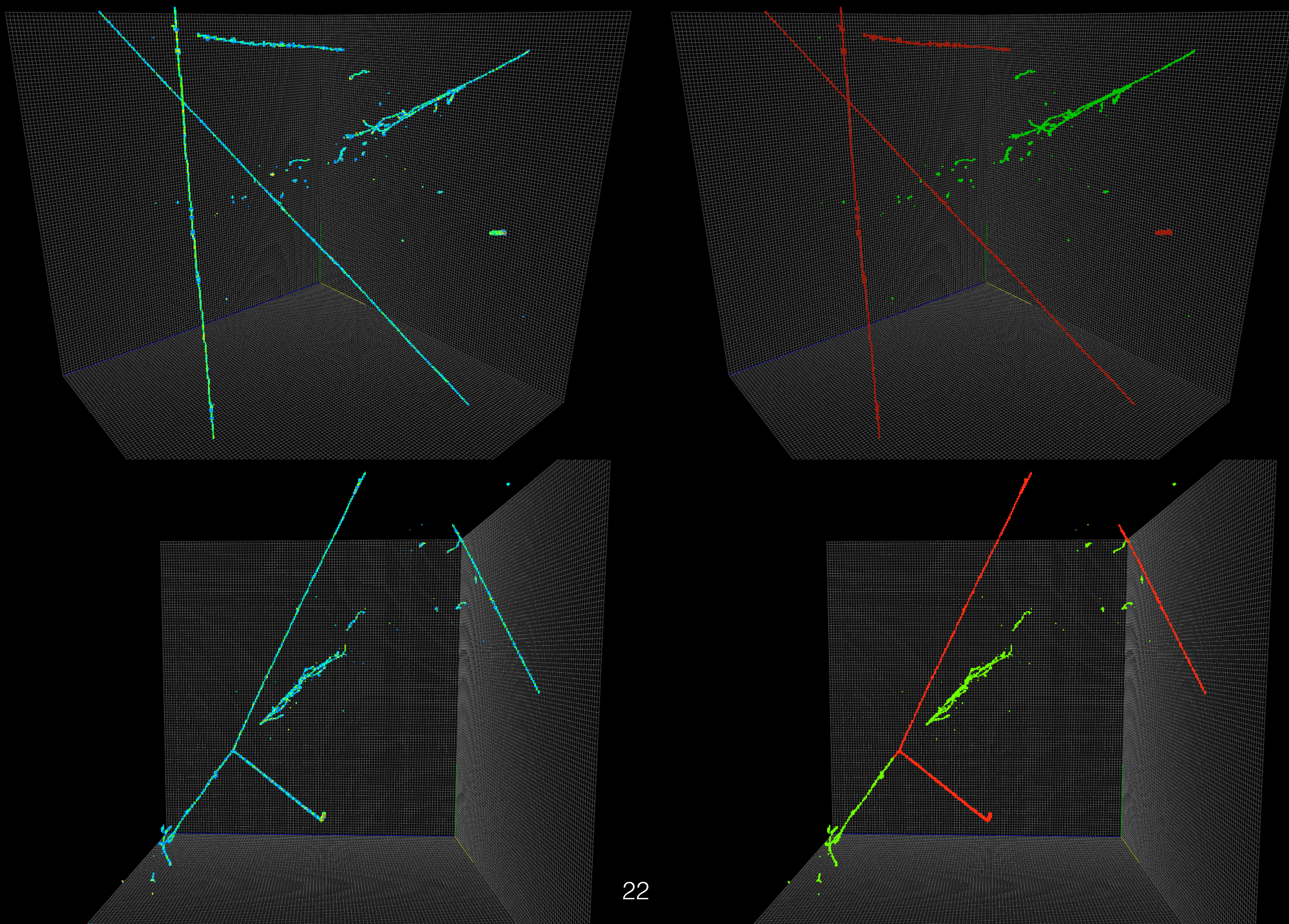


Feature Space Point Finding



Network Input **Network Output**

Application for 3D



Where We're Heading Toward

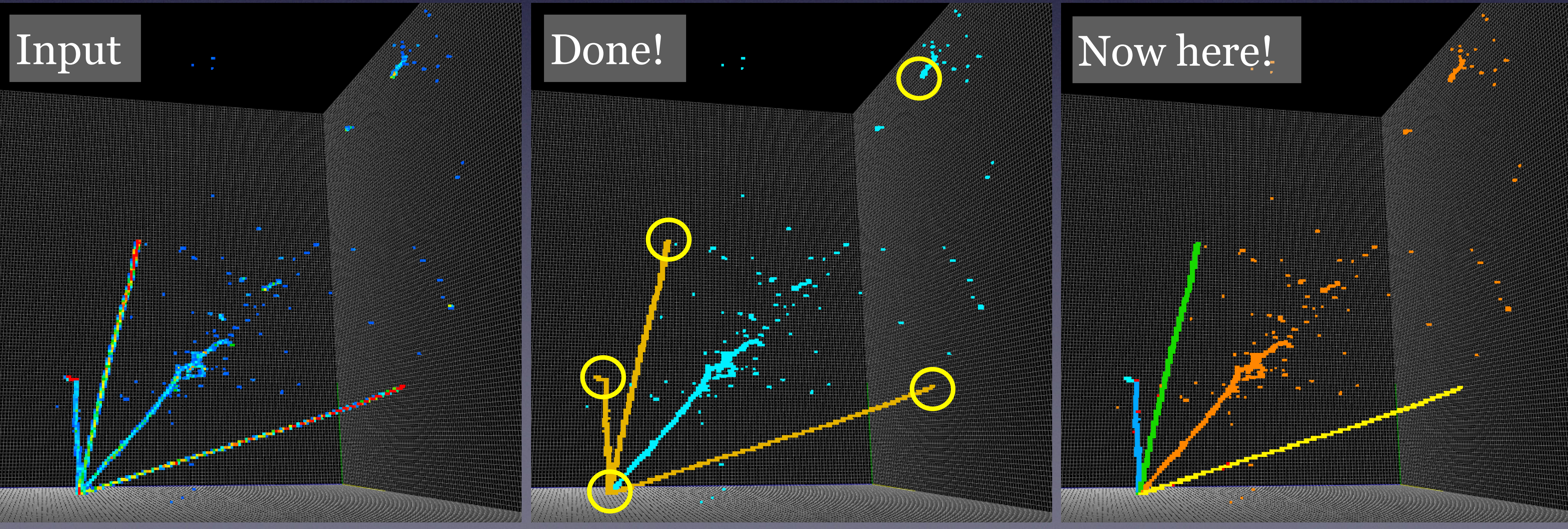
Full Reconstruction Chain

- Individual particle clustering
- Trajectory reconstruction
- Topology classification
- Particle hierarchy analysis


Four SLAC Scientists Awarded Prestigious DOE Early Career Research Grants

Tais Gorkhover, Michael Kagan, Kazuhiro Terao and Joshua Turner will each receive $2.5 million for research that studies fundamental particles, nanoscale objects, quantum materials and machine learning.

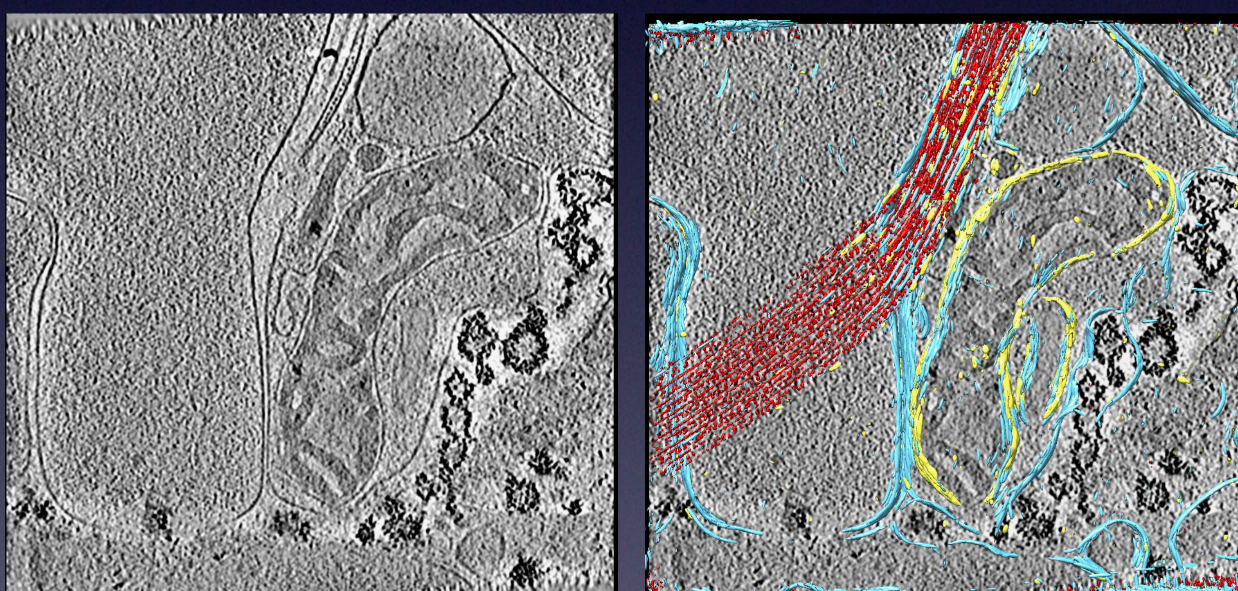
Support from U.S. DOE/NSF
- DOE Early Career Award
- DOE ML@SLAC pilot program
- Many data science initiatives



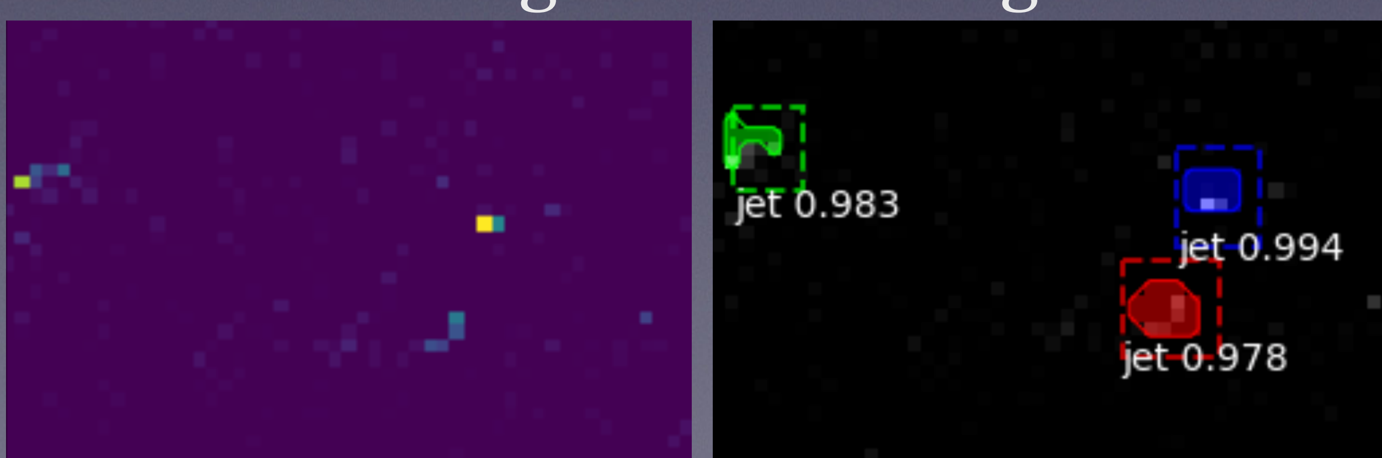
Interdisciplinarity / Synergy

Accelerator Operation/Maintenance

Cryo-EM 3D data labeling

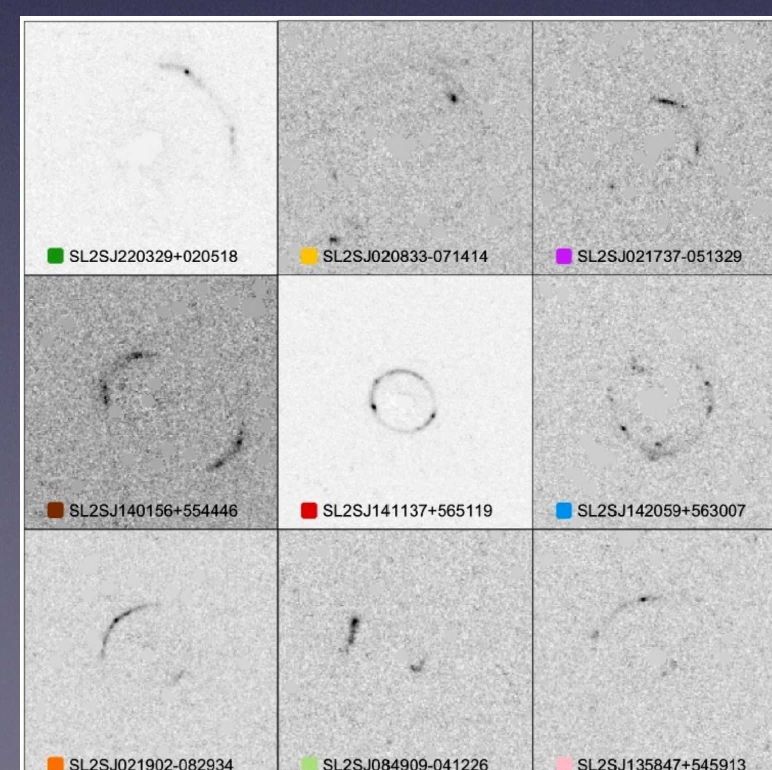


ATLAS Jet Image Instance-Segmentation



Lots of collaborative effort to develop and share tools, experience, and holding workshops for training across national labs & universities.

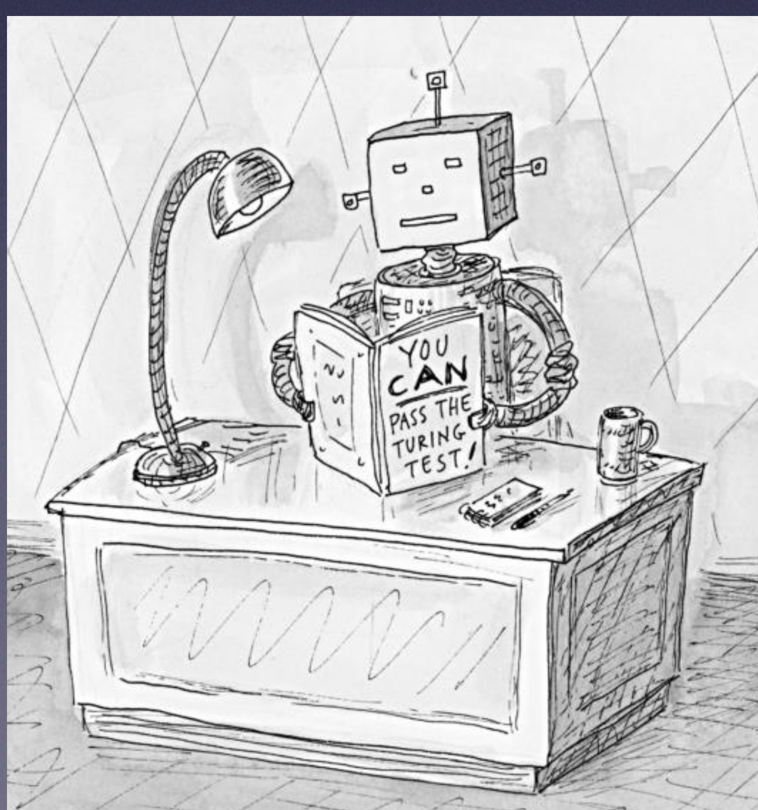
This workshop is great :)



Fast Analysis Gravitational Lensing (LSST)

… Wrapping Up …

- **Particle imaging detectors** are in **the core** of experimental accelerator neutrino physics program
- **Computer vision techniques** are strong tools
- Collaborative development of applications based on **machine learning**, in particular **deep learning**, is active across and beyond particle physics community.



Thank you!

Please come and talk to me if you have questions / etc.

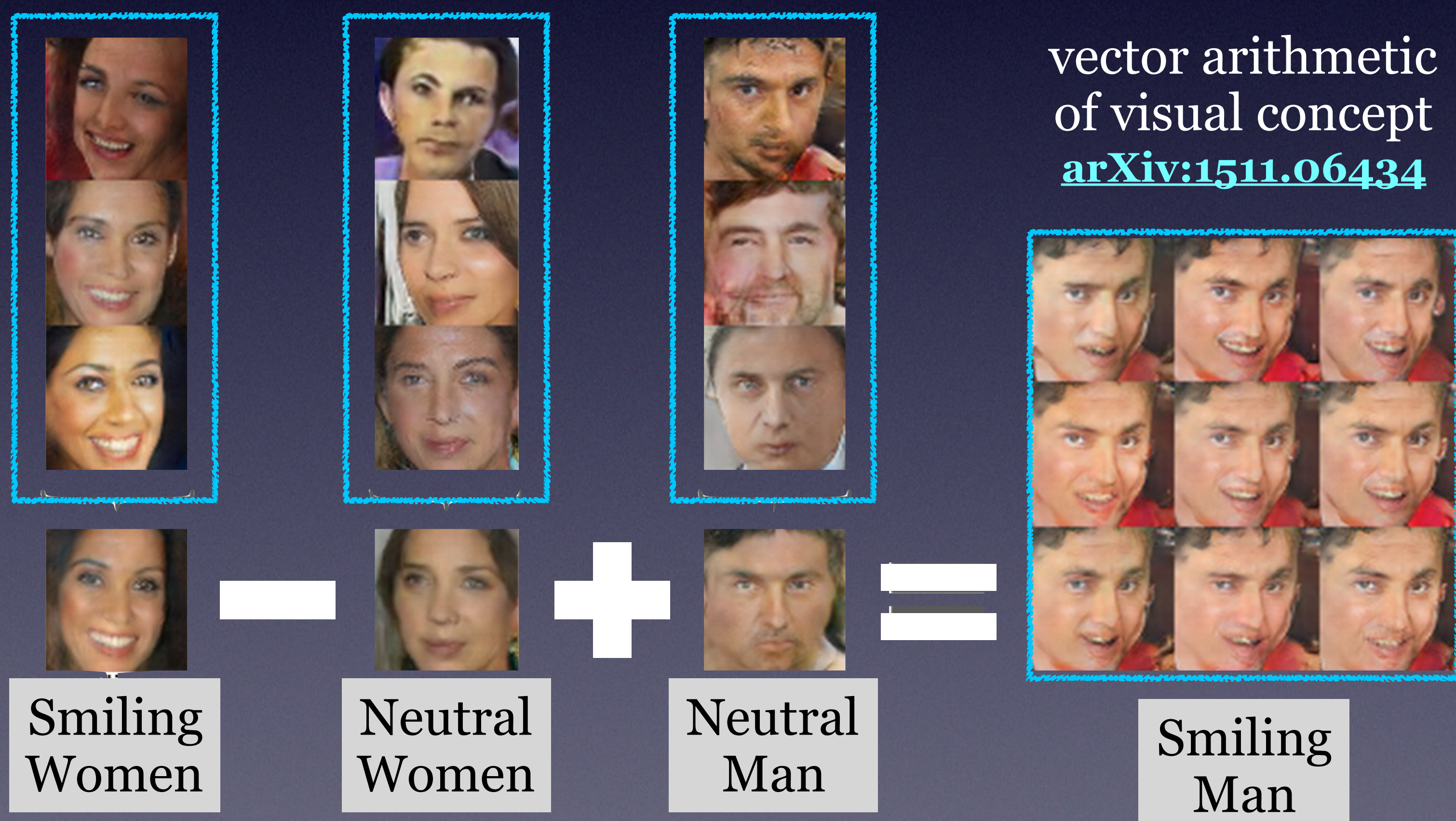
Some technical jargons

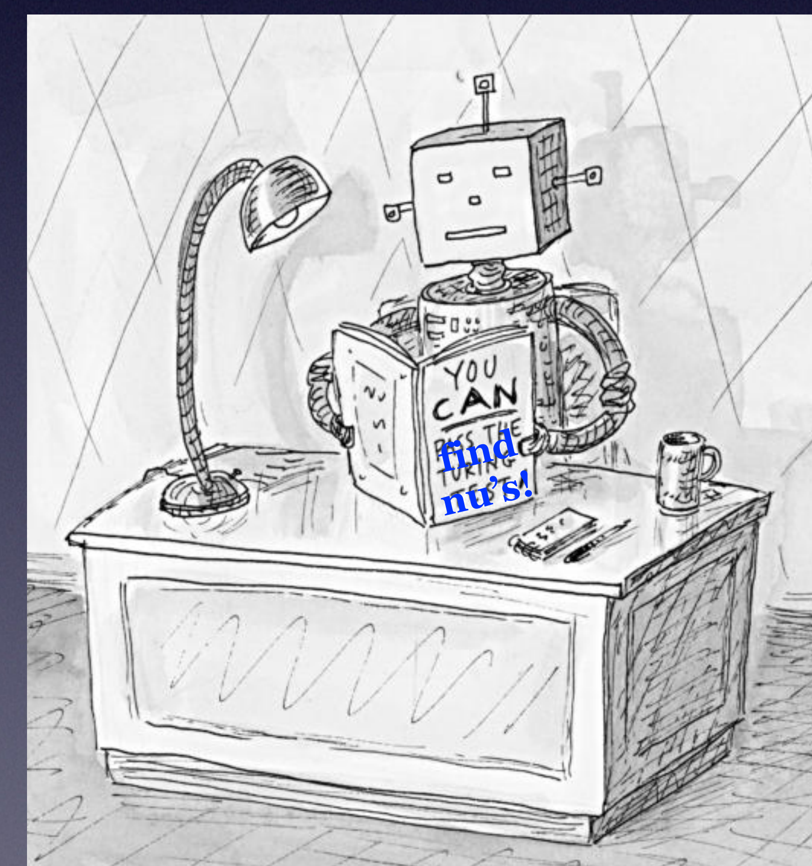
CNN, RNN, GAN, Graph-CNN,
Mask R-CNN, U-Net, ResNet,
Reinforcement Learning, MXNet, PyTorch, TensorFlow

Back Up Slides

… more exciting projects …

vector arithmetic
of visual concept
[arXiv:1511.06434]





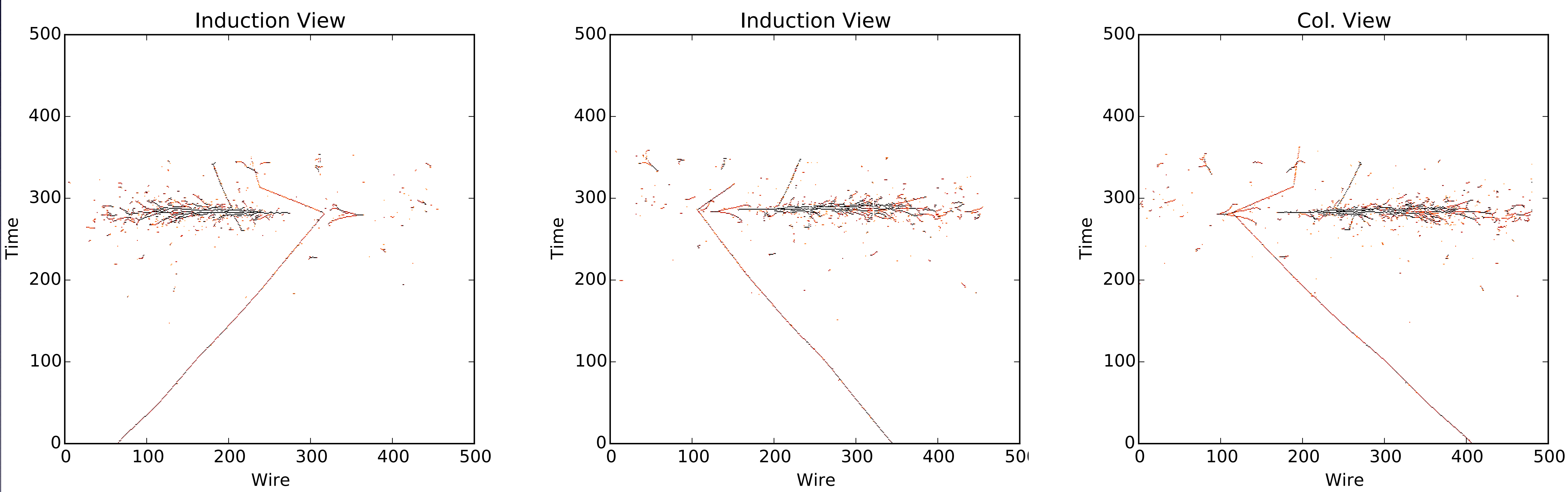
SBND Cosmic Rejection w/ U-ResNet

Collection plane view, similar performance on induction planes (from C. Adams)

Our Input

DL @ DUNE FD Analysis

Each “pixel” is the integrated ADC response in that time/space slice. These maps are chosen to be 500 wires long and 1.2ms wide (split into 500 time chunks).

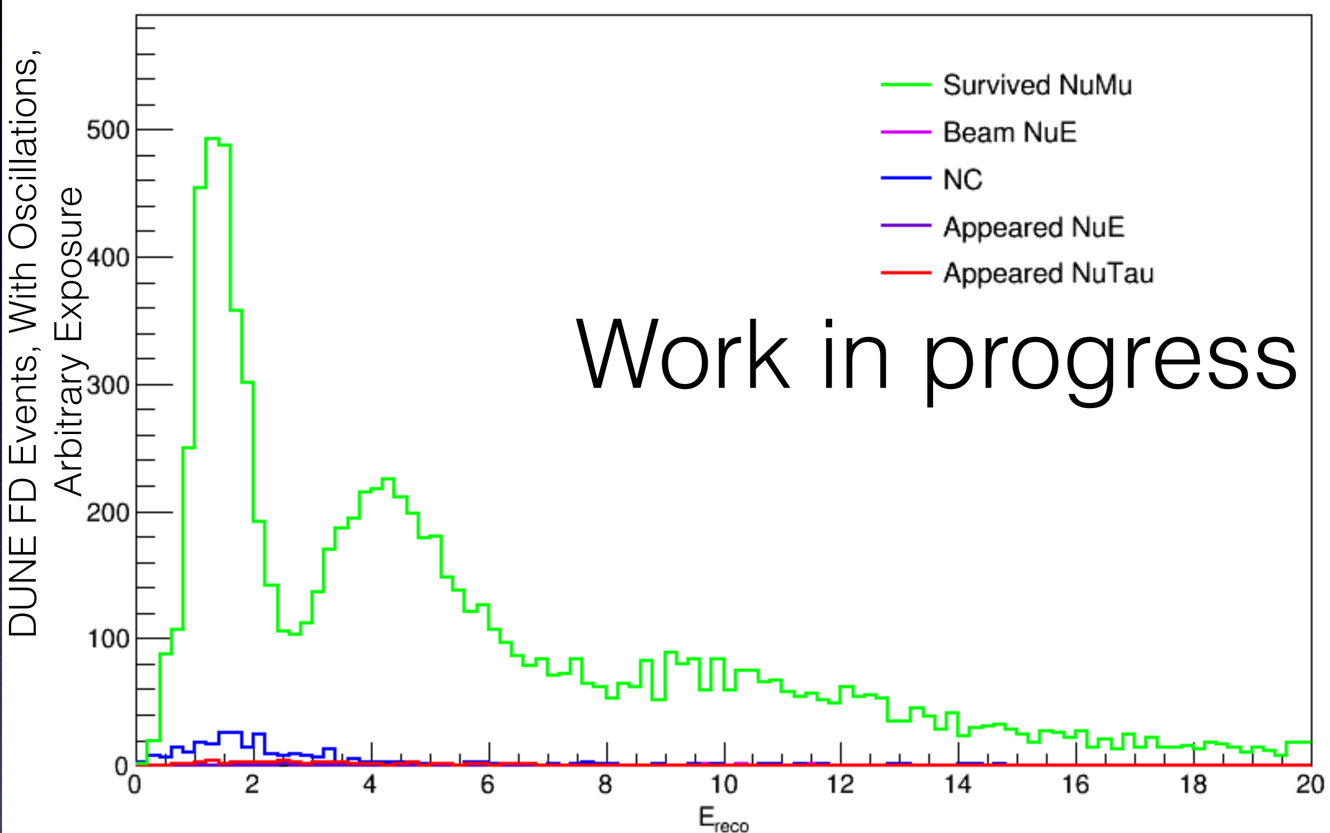


Slide 1/3 from A. Radovic

NuMu Selected Events Reconstructed Energy Spectr

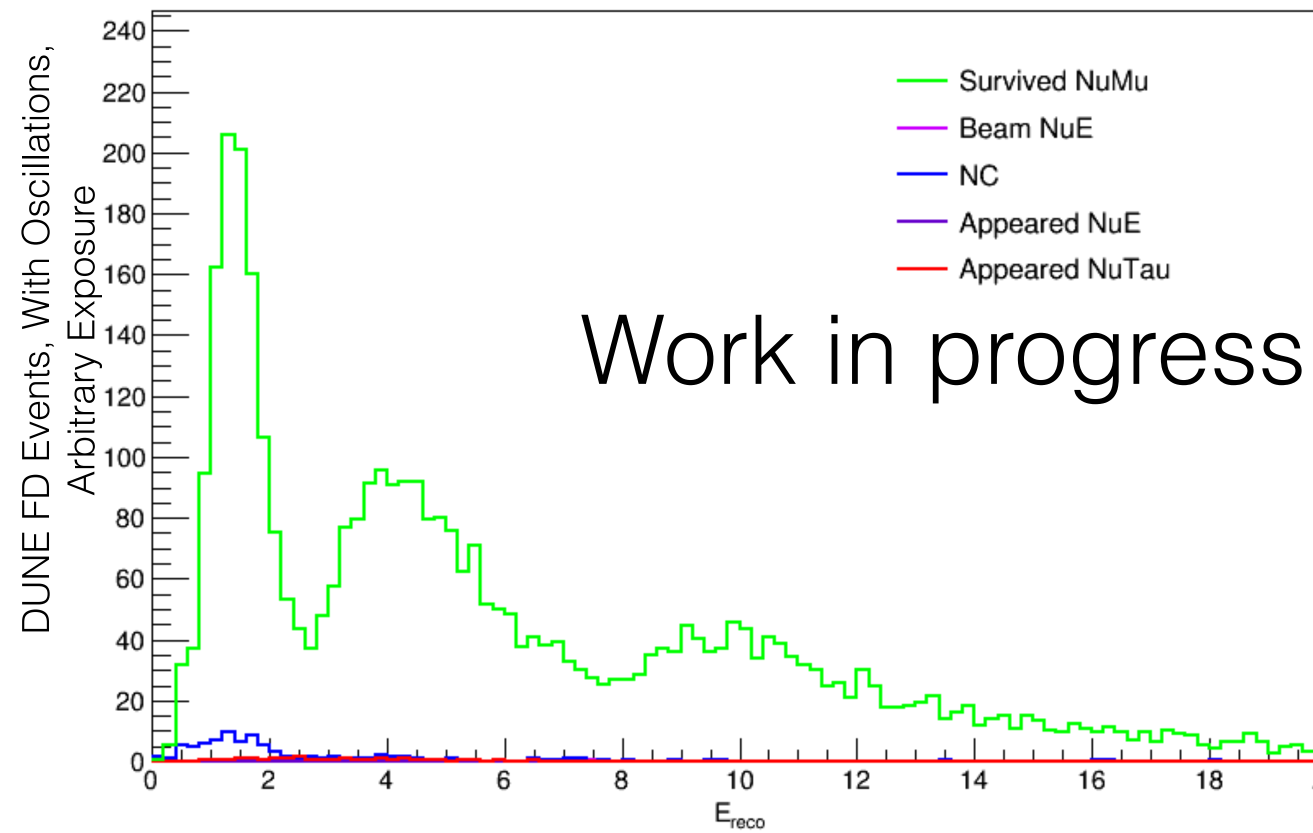
DL @ DUNE FD Analysis

Neutrino Beam



	NuMu	Appeared NuE	Beam NuE	NC	NuTau
Efficiency	**80.6**				
Rejection		99.0	98.7	**97.6**	81.5

Anti-Neutrino Beam

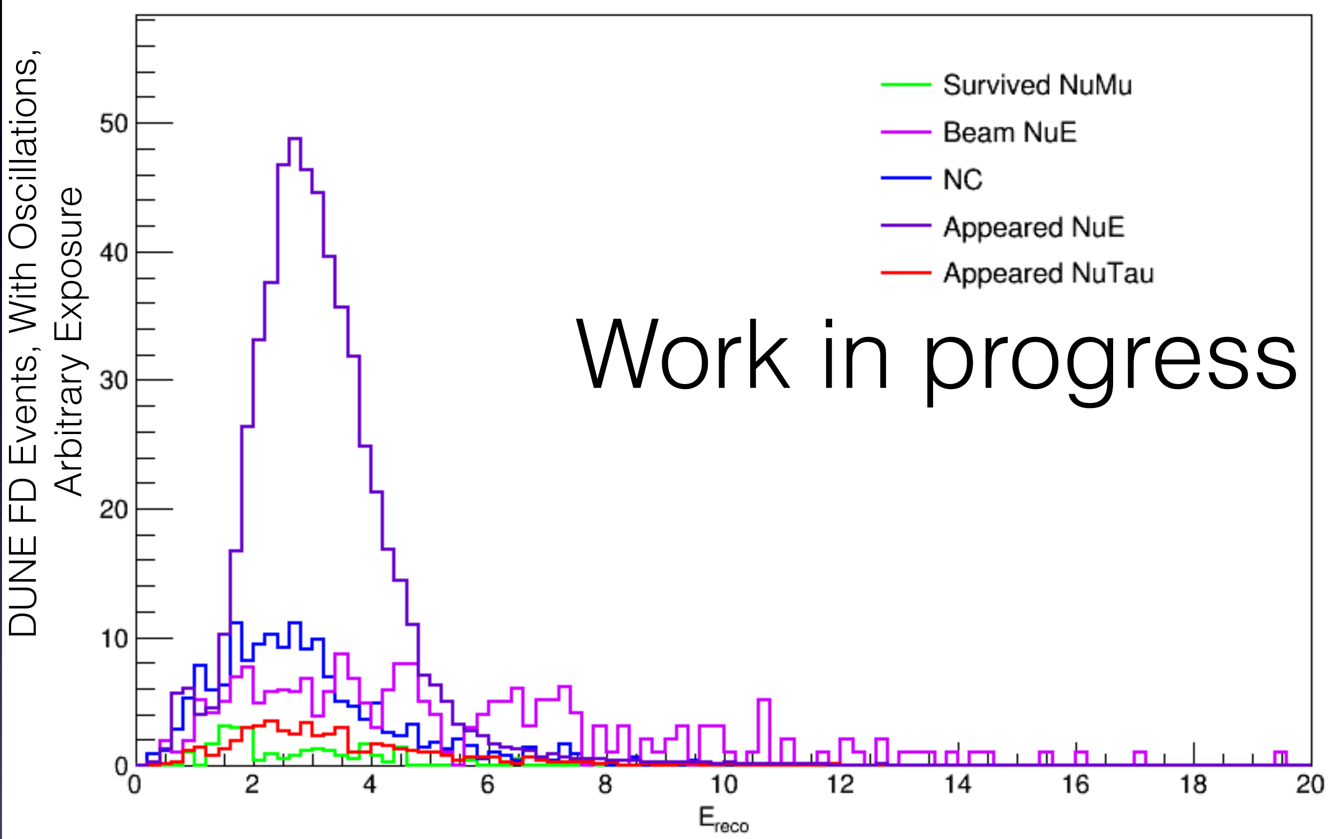


	NuMu	Appeared NuE	Beam NuE	NC	NuTau
Efficiency	**87.7**				
Rejection		99.6	99.3	**98.3**	81.4

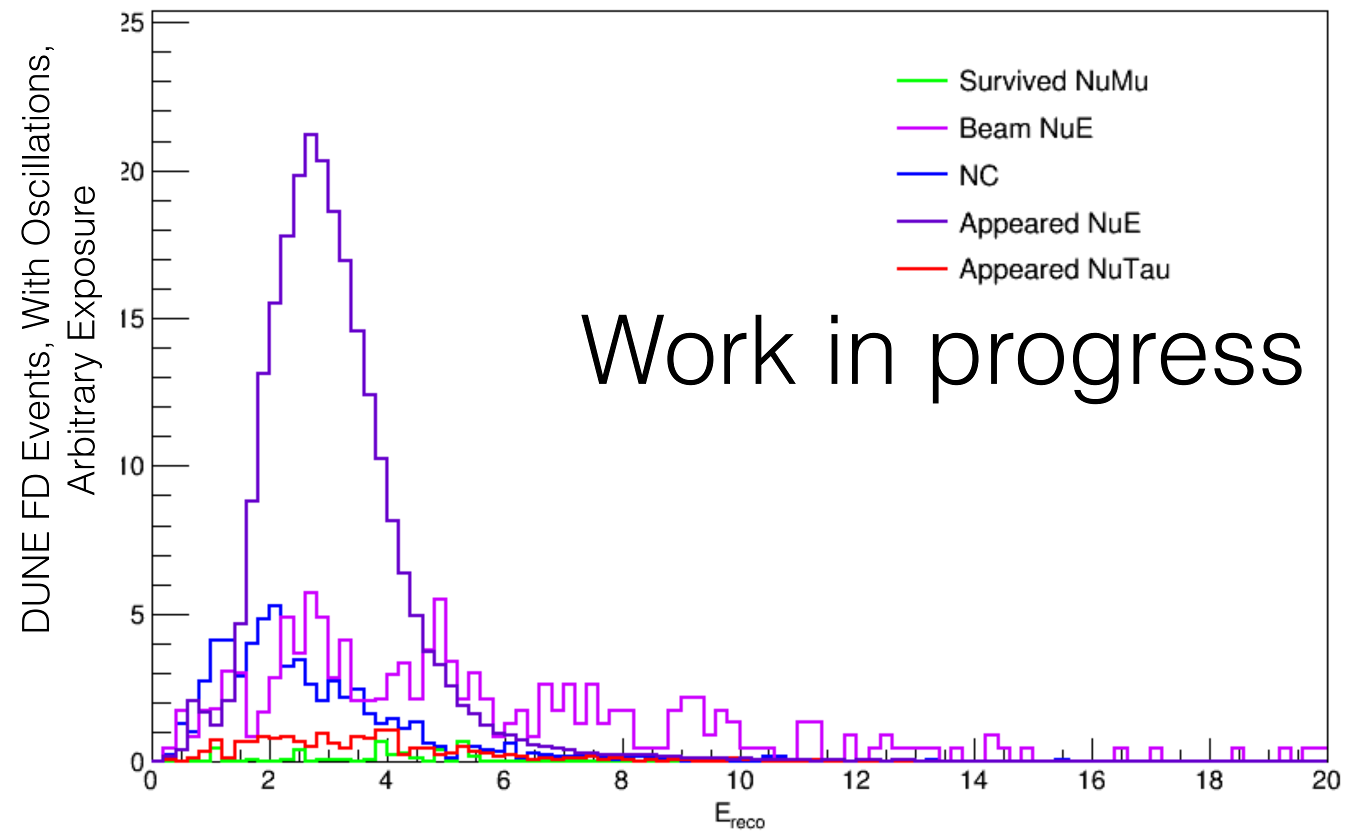
NuE Selected Events, Reconstructed Energy Spectra

DL @ DUNE FD Analysis

Neutrino Beam



Anti-Neutrino Beam



Neutrino Beam

	Appeared NuE	NuMu	Beam NuE	NC	NuTau
Efficiency	**67.5**				
Rejection		99.8	52.1	**98.6**	85.8

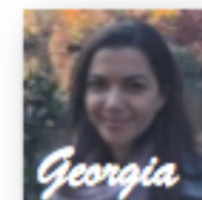
Anti-Neutrino Beam

	Appeared NuE	NuMu	Beam NuE	NC	NuTau
Efficiency	**79.3**				
Rejection		99.9	48.2	**98.8**	87.6

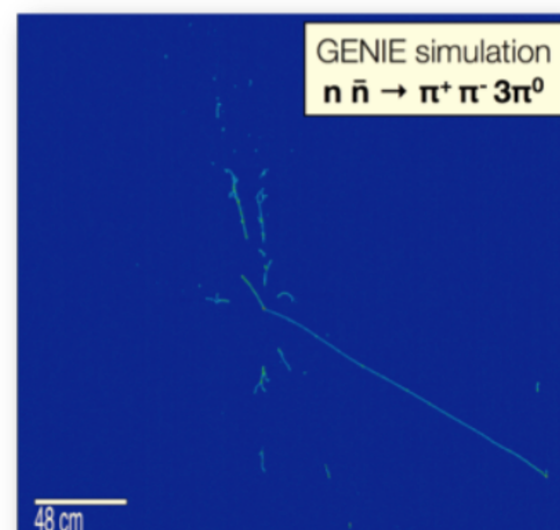
n-nbar Search in DUNE FD

Deep Learning application for rare event searches (and more) in DUNE

Group: Georgia Karargiorgi (Columbia U/U Manchester), Jeremy Hewes (formerly U Manchester), Yuyang Zhou (Columbia U)



CNN application in DUNE: originally developed as a DL-based analysis for a search for **rare neutron-antineutron oscillation events** (B-violating signature) in DUNE.



Simulated n-nbar event in DUNE; striking ("star event") topology

MANCHESTER 1824
The University of Manchester

Nevis Laboratories
Columbia University

n-nbar Search in DUNE FD

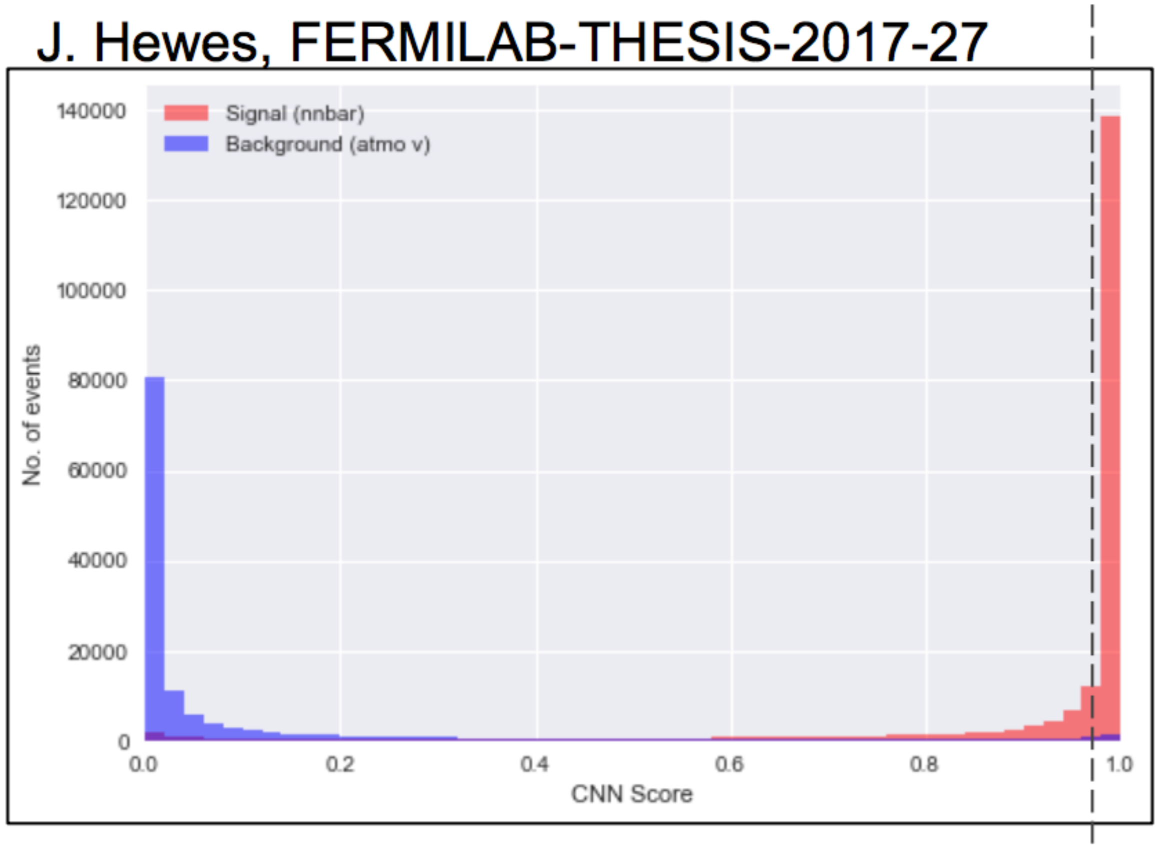
CNN-based search for n-nbar in DUNE

vgg16 network

Trained to differentiate n-nbar events from atmospheric neutrino events* (training samples of 50k events), and tested (samples of 200k events).

*atmospheric neutrino events expected to be the dominant background in DUNE

J. Hewes, FERMILAB-THESIS-2017-27



An optimized cut on CNN score yields signal efficiency of 14% background mis-ID rate of 0.003%

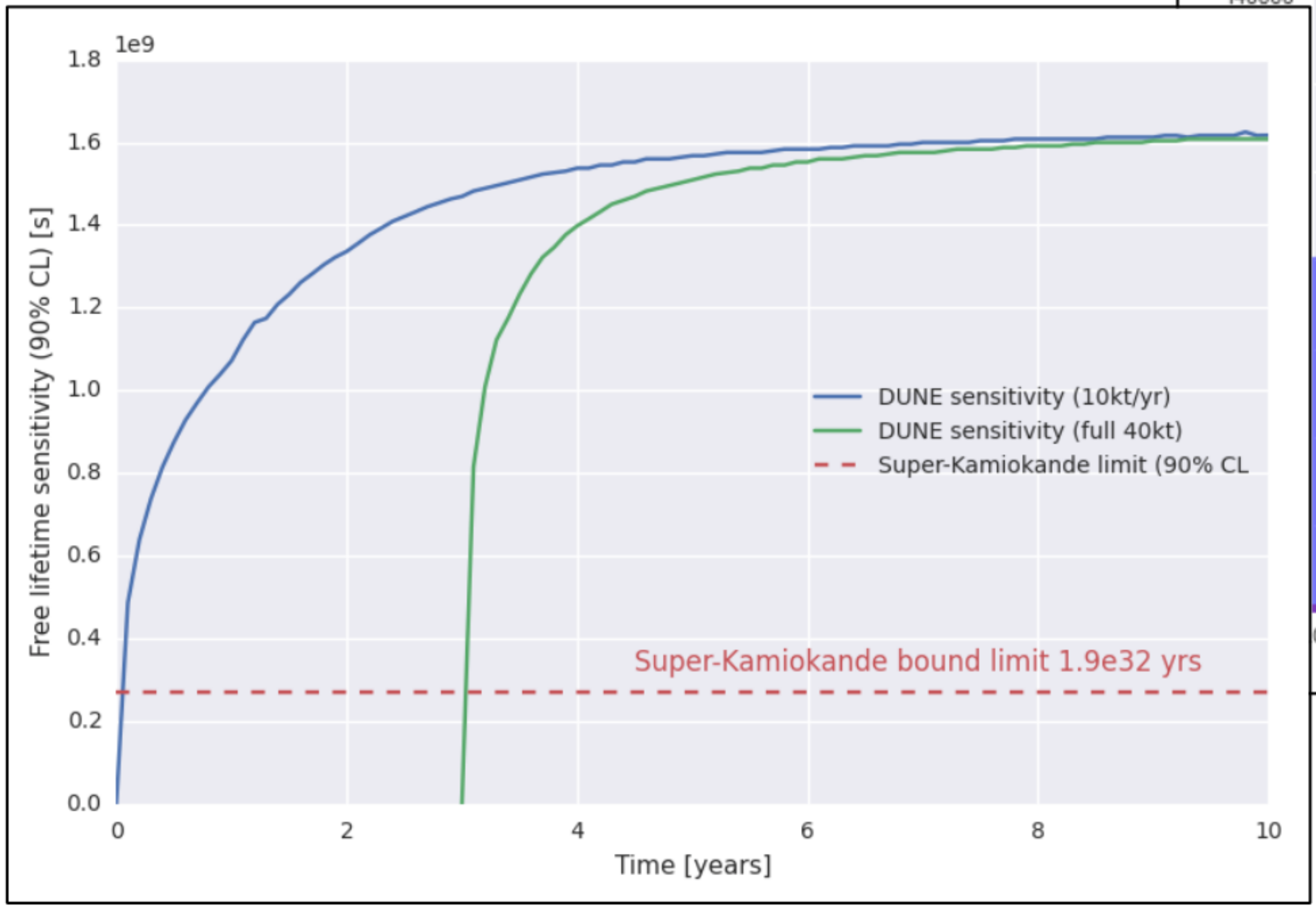
MANCHESTER 1824 The University of Manchester

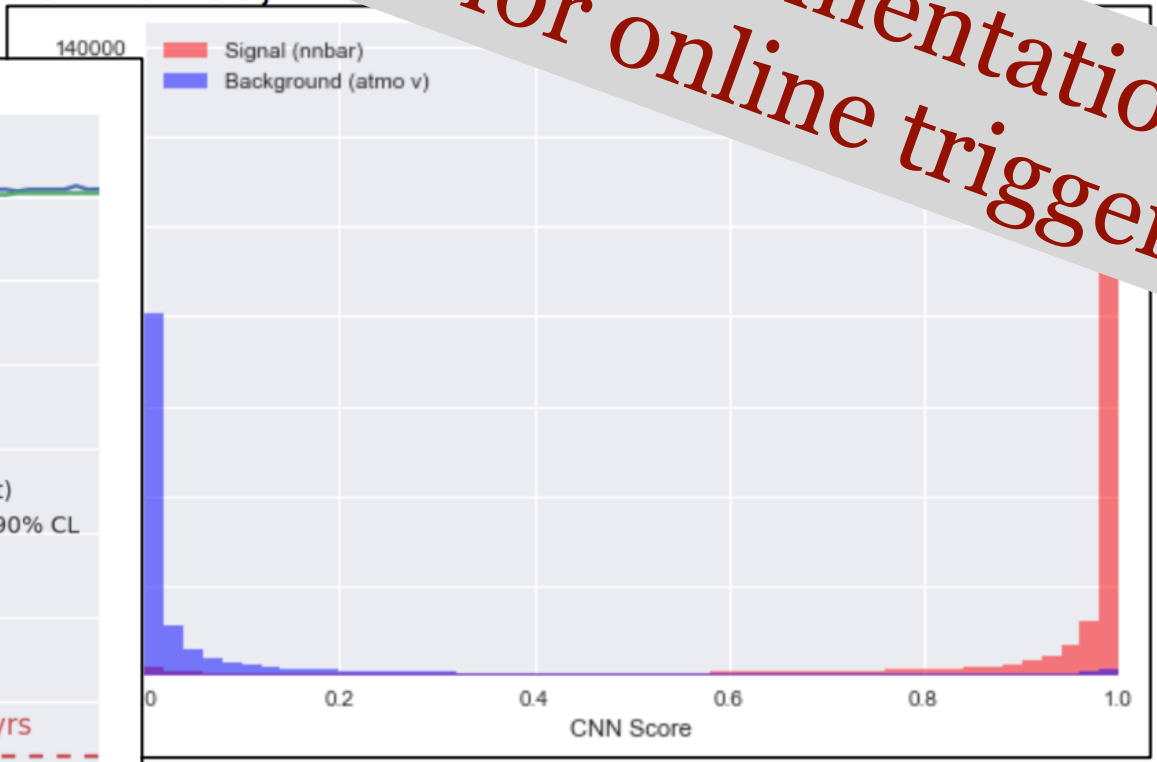
NEVIS LABORATORIES COLUMBIA UNIVERSITY

n-nbar Search in DUNE FD

CNN-based search for n-nbar in DUNE

J. Hewes, FD





FPGA Implementation
R&D for online trigger

Resulting projected sensitivity of DUNE for given efficiency and mis-ID rate, as a function of run time.
Sensitivity shows 5x improvement over current Super-K limit.

MANCHESTER 1824
The University of Manchester
Nevis Laboratories
Columbia University

Distributed CNN Training at PNNL

E. Church, J. Daily, C. Siegel, M. Schram, J. Strube, K. Wierman

- Full event image: **3600 wires x 3600 time bins x 3 planes x 4 Bytes**
  - MicroBooNE simulated single particle events
  - **~150 MB / event**
- Even a moderately small network only leaves room for a mini-batch size of 1-2 events on a modern GPU, for full event fidelity
  - This is smaller than required given the latent space of the CNN → slow development. Distributed scaling of compute resources will help significantly.
  - Scaling allows increase in network depth too (if required)
- For deep learning, one wants large training samples.
  - Training may become quickly I/O bound and hence prohibitively slow
  - Even a dedicated ”large-mem” node cannot fit more than a few thousand samples into memory, at best.

→ We are studying PNNL’s MaTEx for distributed training

Easier to ”drop in” than say the uber solution, and locally supported!

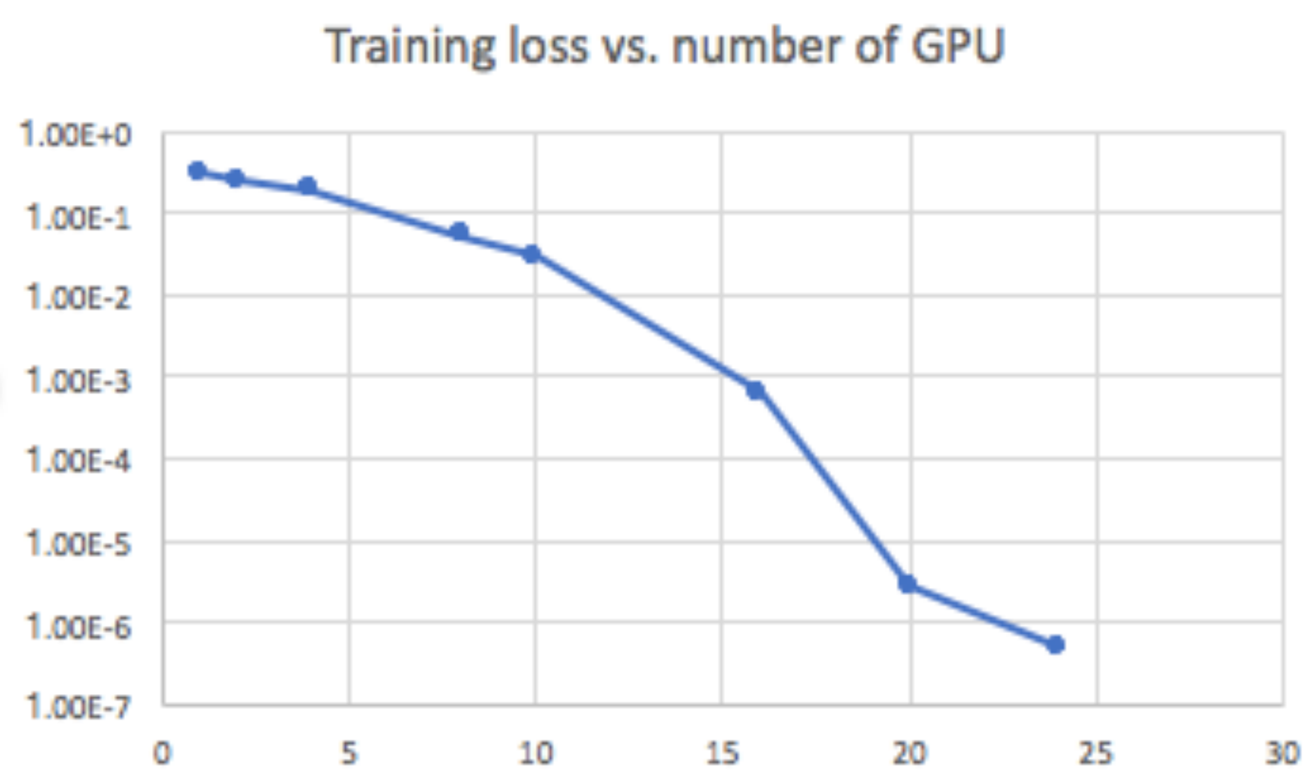
→ And using in-memory loss-less image compression

Slide 1/2 from E. Church

DL Software @PNNL
Framework Development

Current status (preliminary)

Training time: mini-batch size = 2, 10000 steps per GPU … 10 epochs
Identical networks, loss functions, optimizers and input data
➞ MaTEx does not currently introduce noticeable overhead at this scale



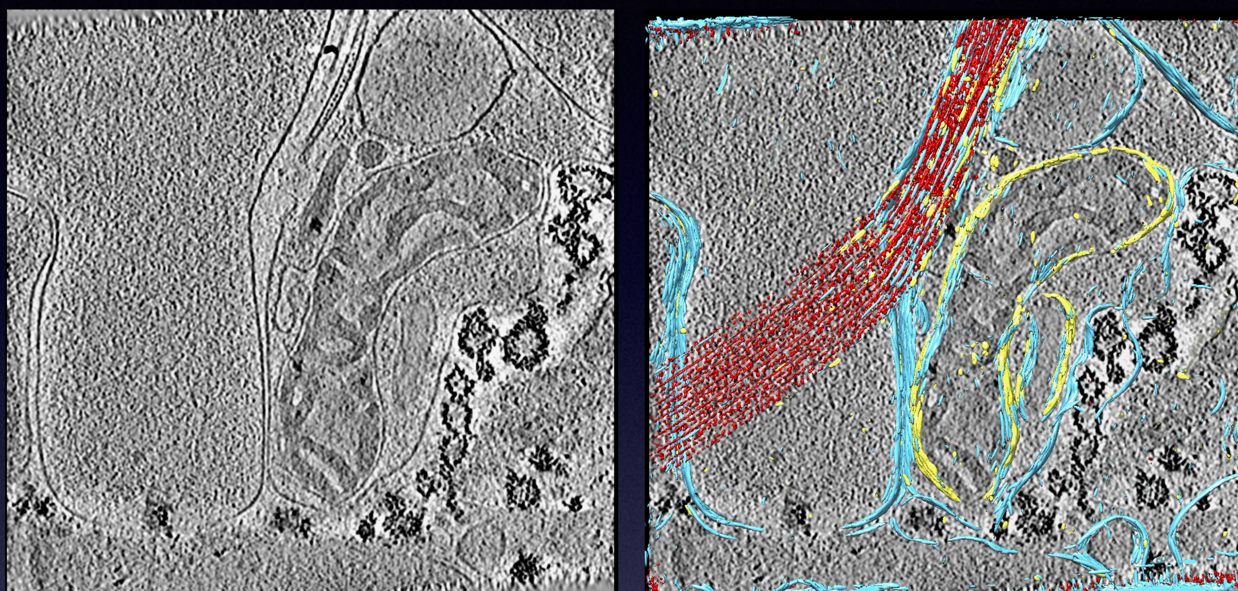
For the same wall time, training improves with number of GPUs
➞ Studies ongoing, significant updates planned for CHEP2018

Slide 2/2 from E. Church

DL Software @PNNL
Framework Development

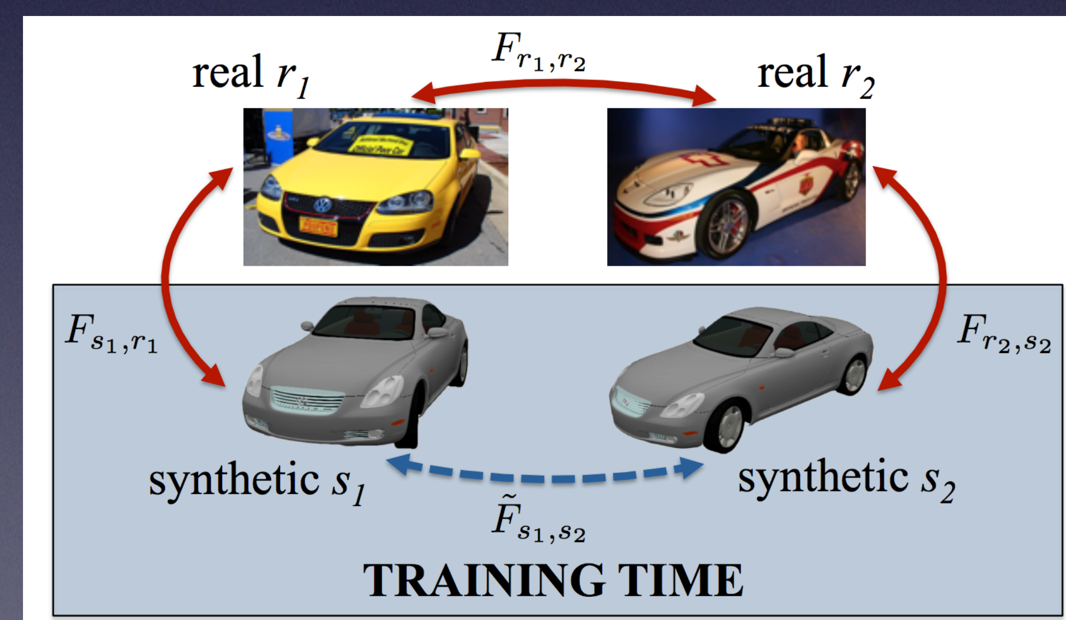
More Exciting Stuffs … come chat w/ me :)

3D voxel labeling of Cryo-EM image
(below: mitochondrion detection)



Multi-network Training
Techniques R&D

Detection + Clustering (Mask R-CNN)
of ATLAS jet images
(w/ SLAC ATLAS group)



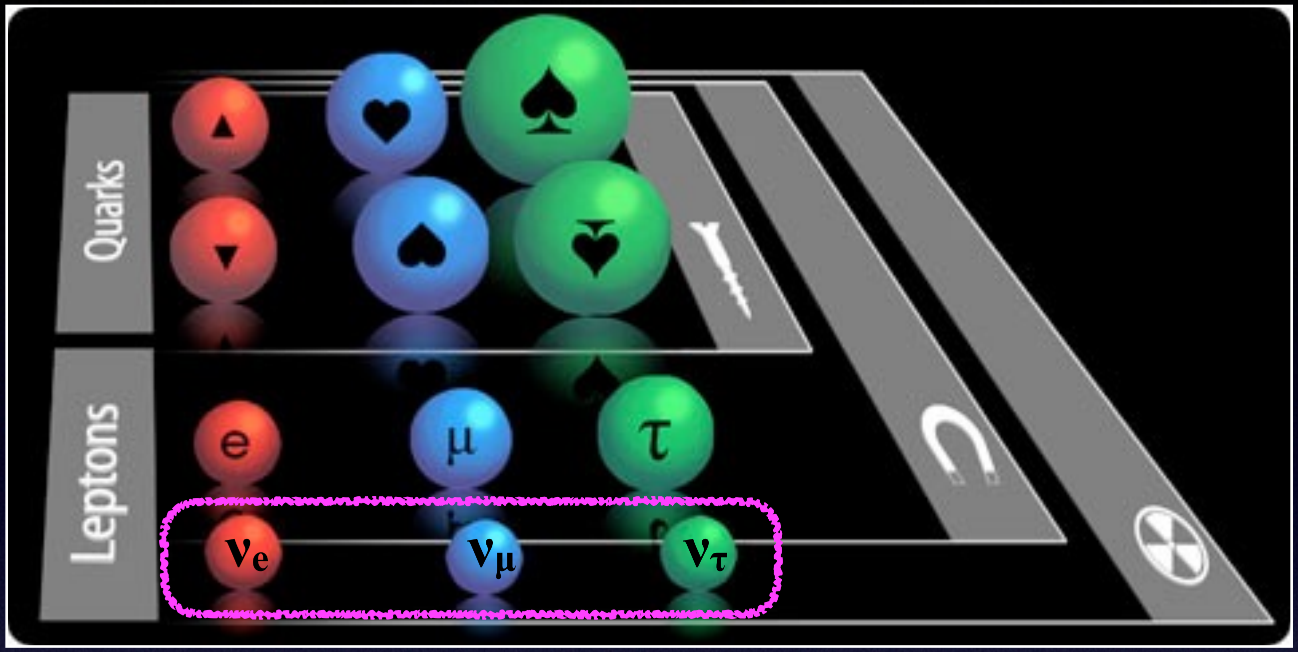
Pixel-Flow network for 3D track reco
(via cross-plane pixel correlation)

Why Neutrinos

Why Neutrino Physics? (I)

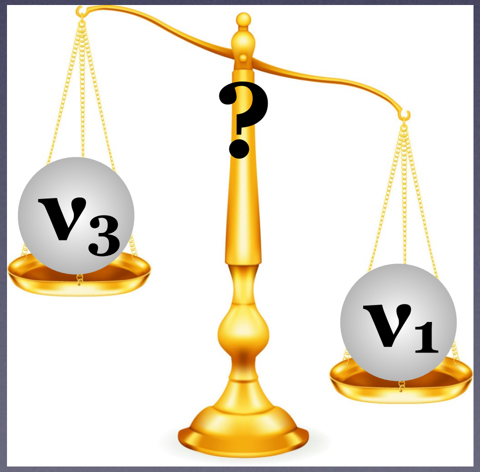
Standard Model (SM)

Successful description of how we know particles interact in nature … but **not so much on neutrinos!**



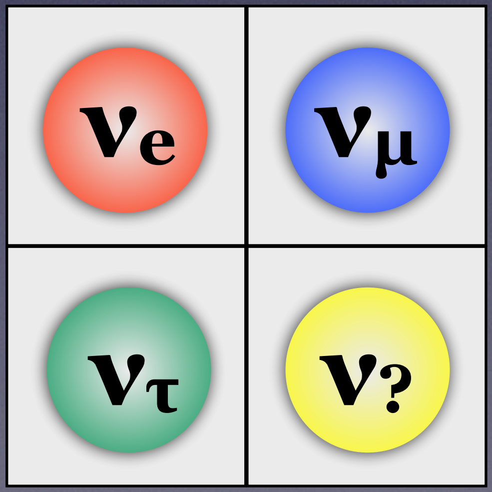
Neutrinos *beyond* SM

With **neutrino oscillations** firmly in place, we know at least there are 3 mass eigenstates. But there is **much more to learn**…



Mass hierarchy
$m_1 > m_3$?

CP violation

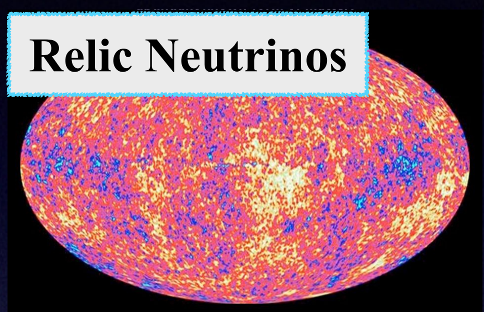


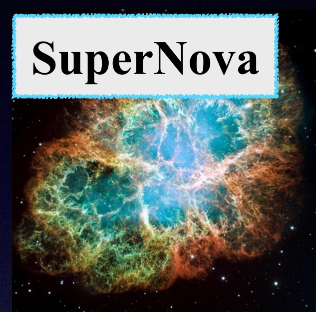
Sterile neutrino?

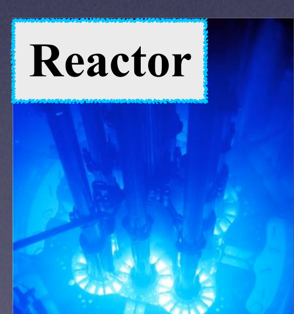
Why Neutrino Physics? (II)

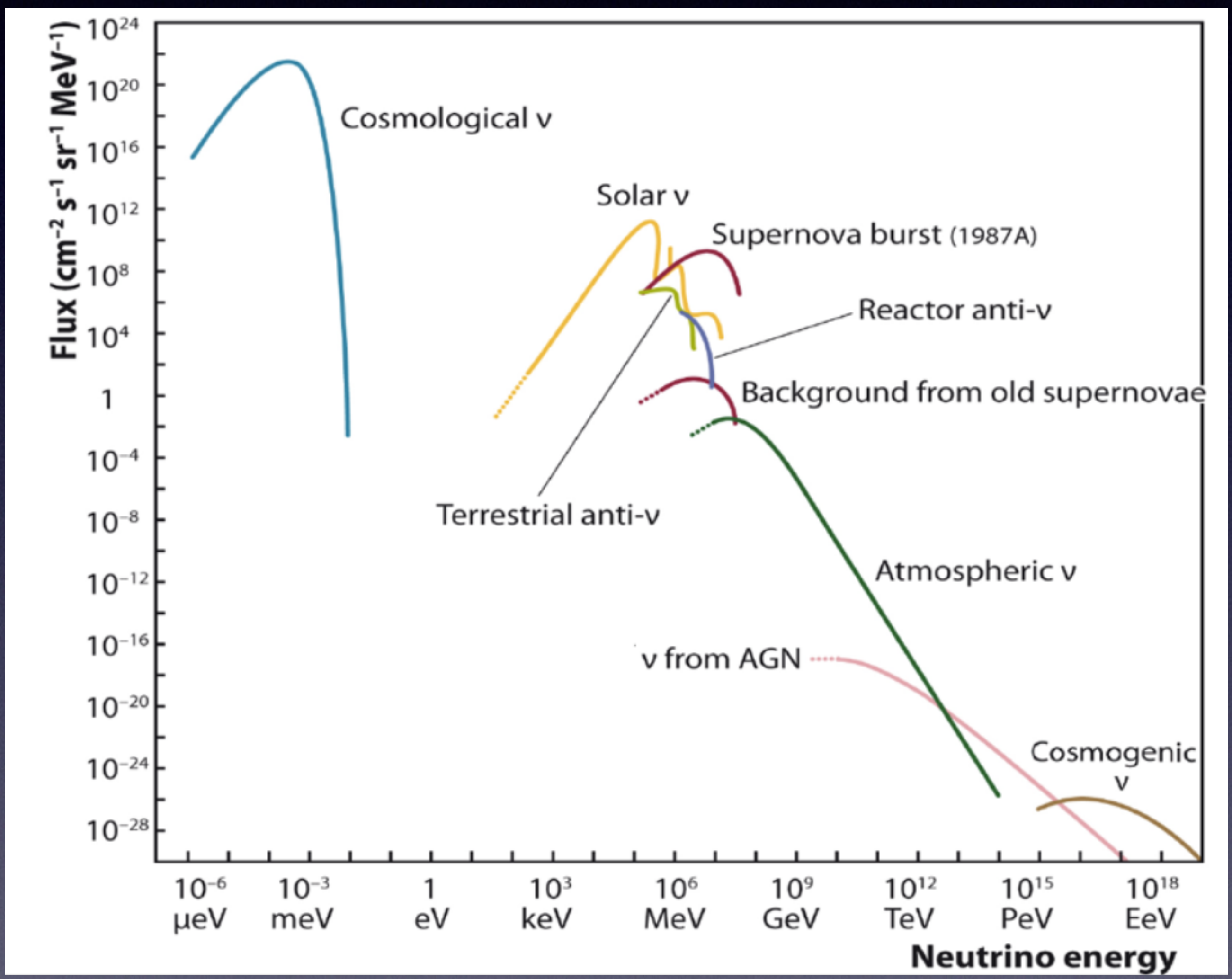
Neutrinos are everywhere

Which makes them **natural probes to the universe and its history**









EPJ H37 (2012) 3:515-565

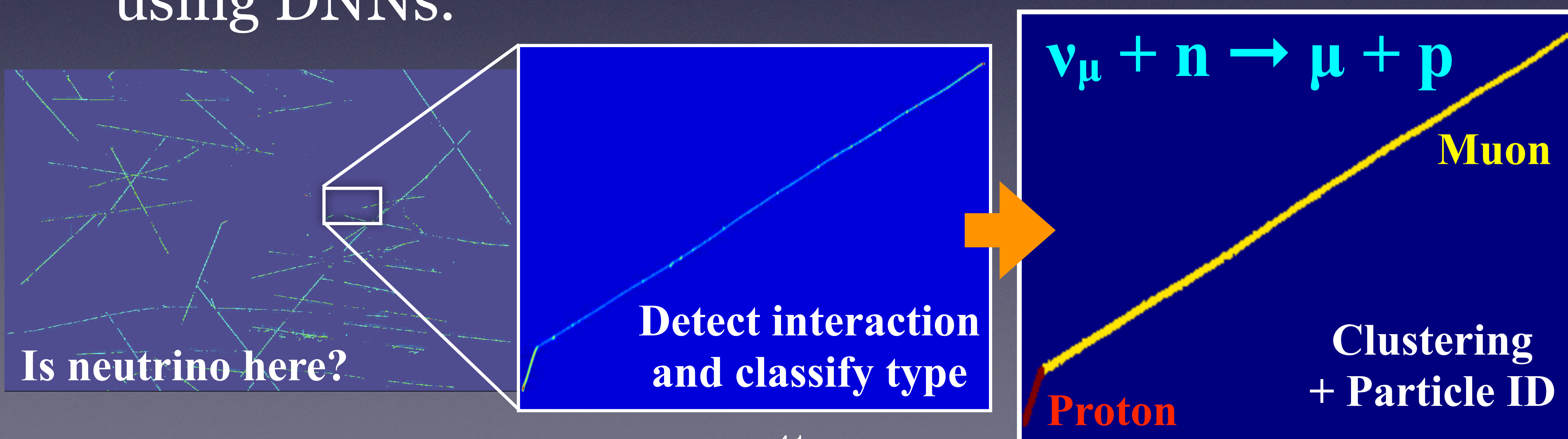
Need to understand more about them!

Oscillation physics has taught us a lot, but still much to learn…

My Interest: ML Applications

Reconstruction chain using DNNs

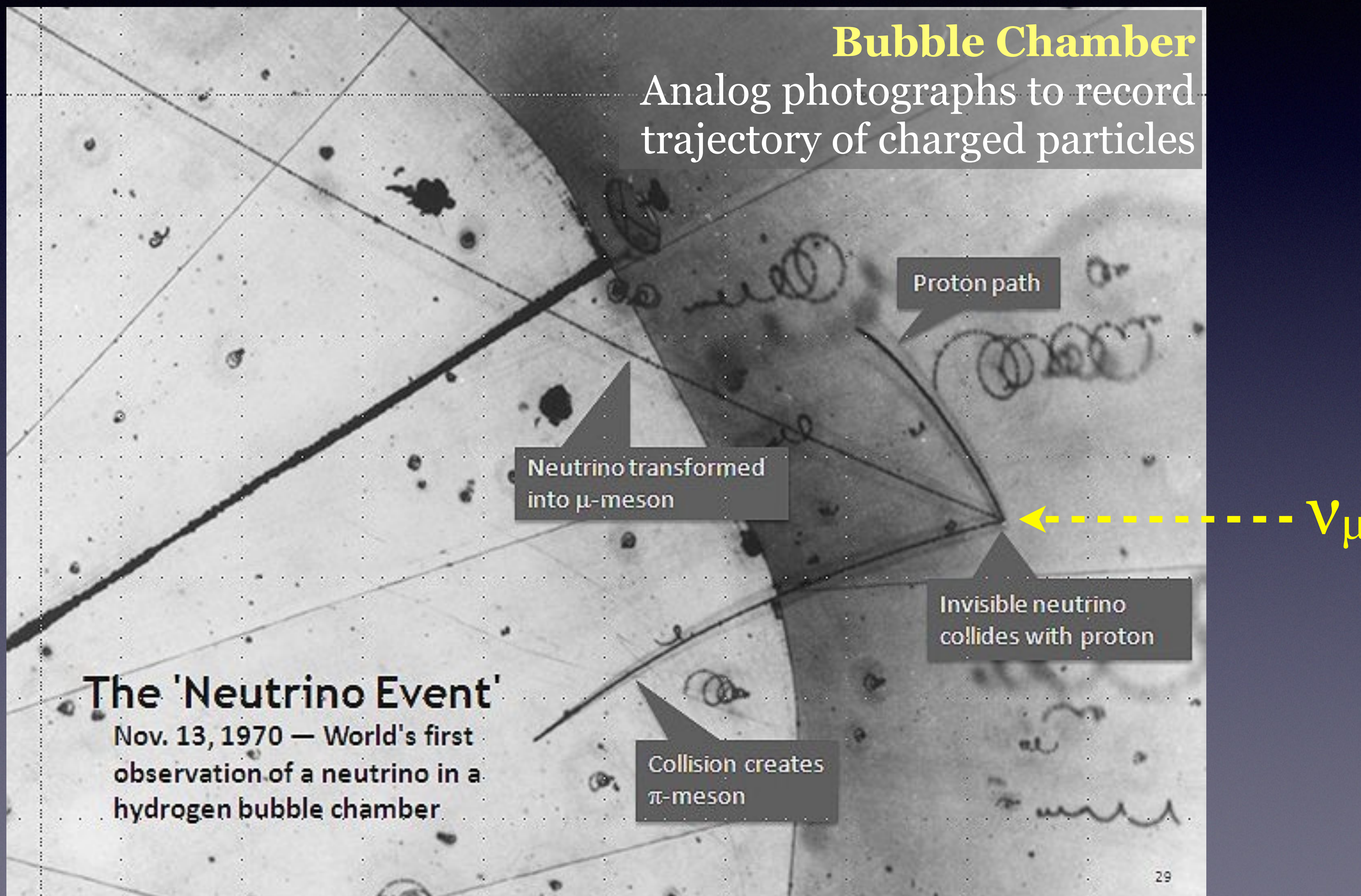
- Design **DNNs for key feature extraction**
  - Interaction vertex, particle clustering, type identification, hierarchy reconstruction, etc. …
- **Chain them up**: optimize the whole process
  - Still extracts key individual features.
  - Leaves flexibility to implement some tasks without using DNNs.



Detectors

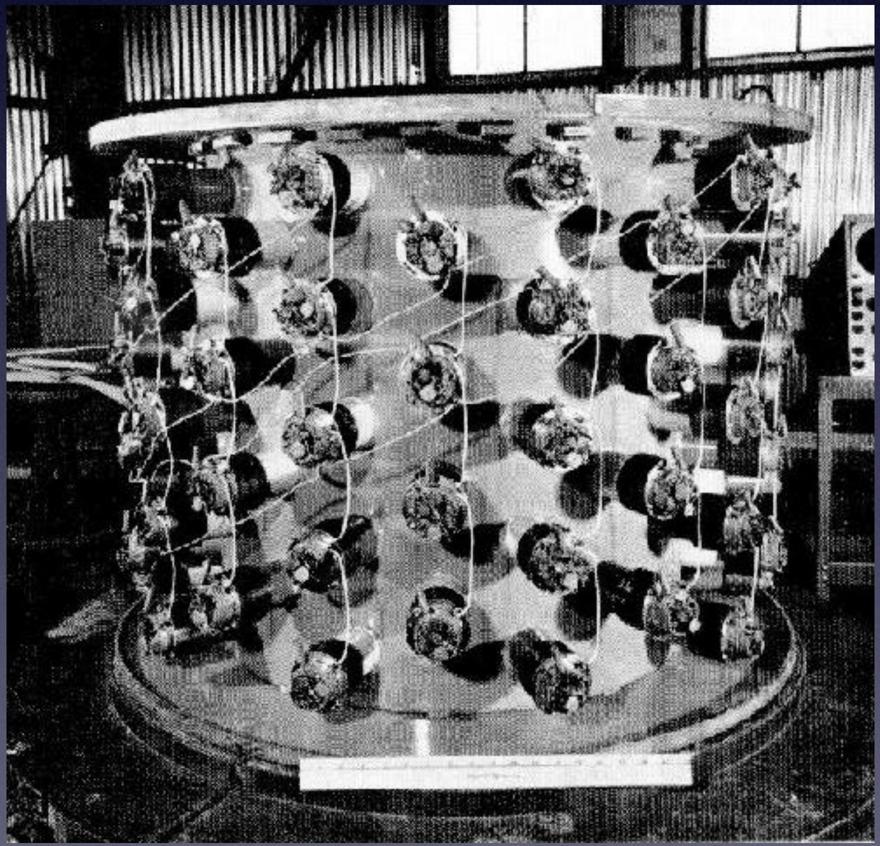
Detecting Neutrinos: BMB

We cannot observe neutrinos, but we can detect particles that come out of a neutrino interaction.

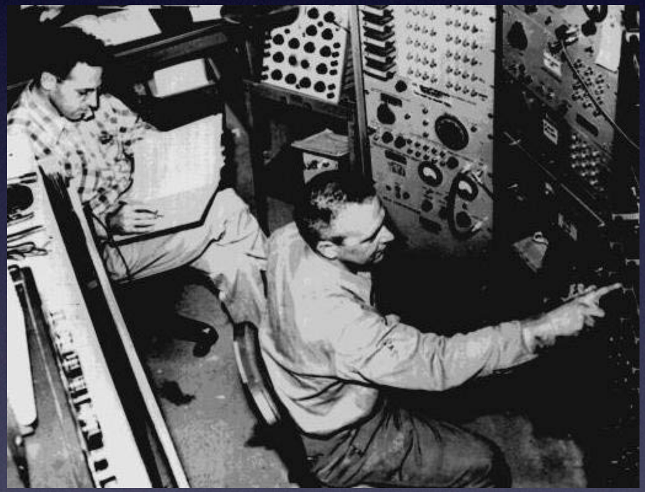


Neutrino Oscillation Experiments (I)

Evolution of Detectors



Cd-doped water
0.4 ton, 100 PMTs

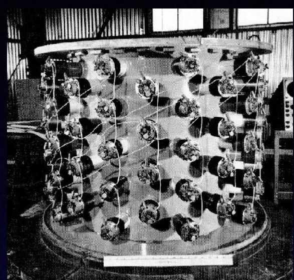


Inverse Beta Decay (IBD)
$\overline{\nu}_e + p \rightarrow e^+ + n$
by Reines & Cowan (Nobel Prize 1995)

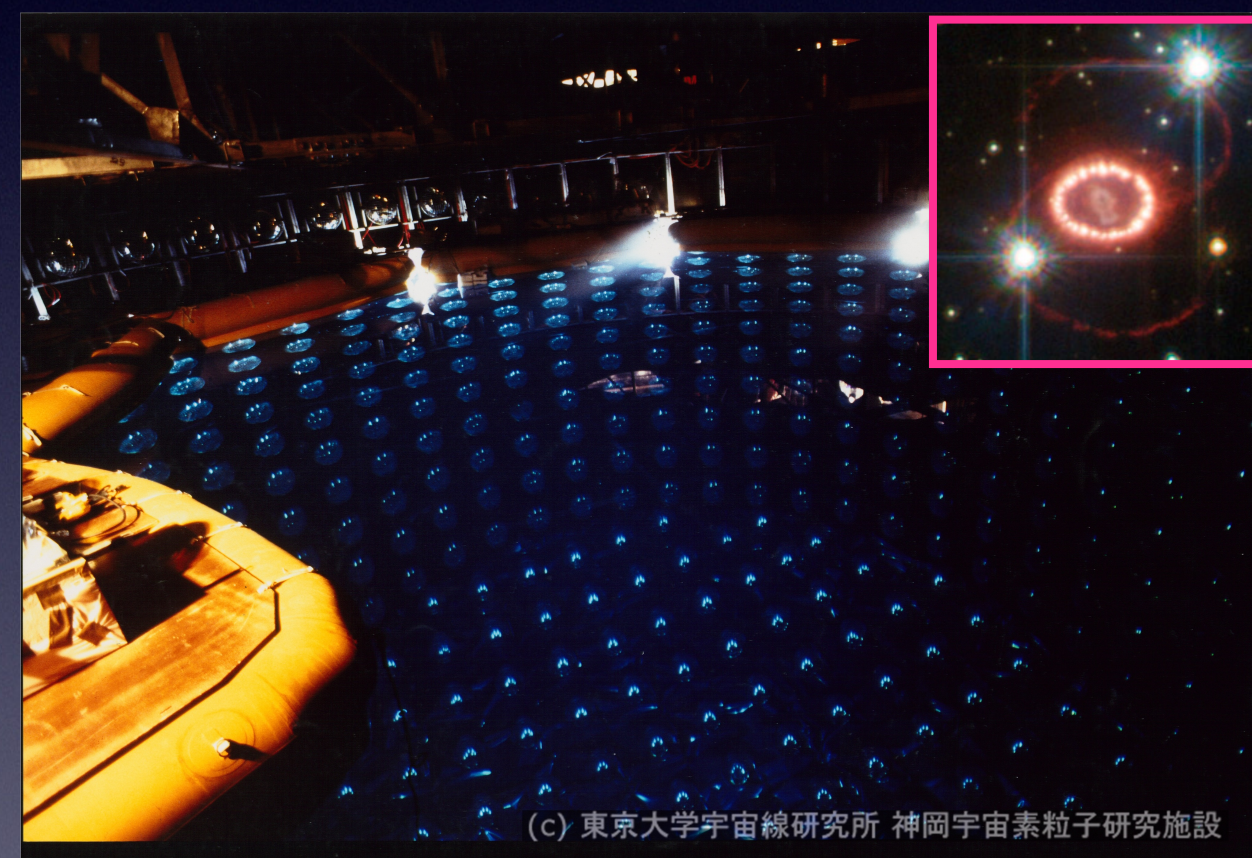
First neutrino detection

Neutrino Oscillation Experiments (I)

Evolution of Detectors



Cd-doped water
0.4 ton, 100 PMTs
(1956)

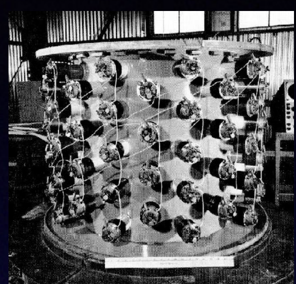


KamiokaNDE Detector
3 kton ultra-pure water, 1000 20” PMTs
(shared Nobel Prize 2002)

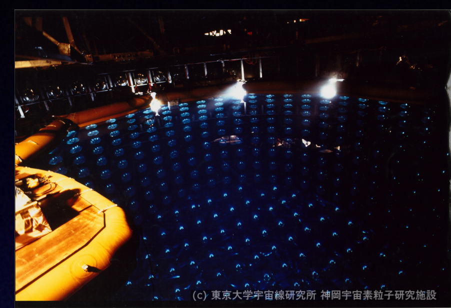
Birth of neutrino astrophysics!

Neutrino Oscillation Experiments (I)

Evolution of Detectors



Cd-doped water
0.4 ton, 100 PMTs
(1956)



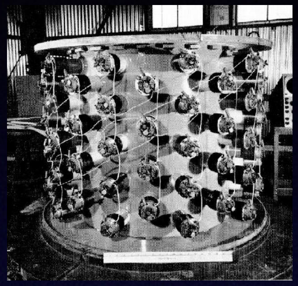
Ultra-pure water
3 kton, 1000 PMTs
(1983)

Discovery of ν_{atmo} oscillation!

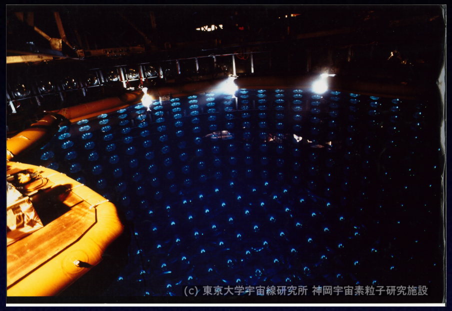
Super-KamiokaNDE
50 kton ultra-pure water,
11000 PMTs
(shared Nobel Prize 2015)

Neutrino Oscillation Experiments (I)

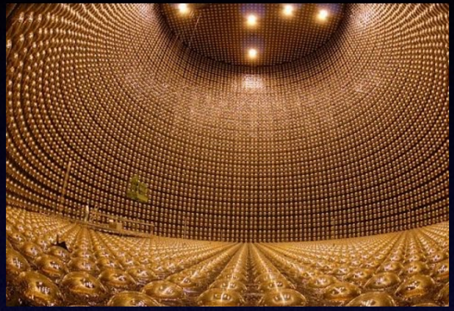
Evolution of Detectors



Cd-doped water
0.4 ton, 100 PMTs
(1956)

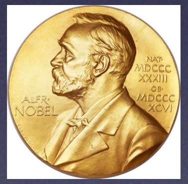


Ultra-pure water
3 kton, 1000 PMTs
(1983)

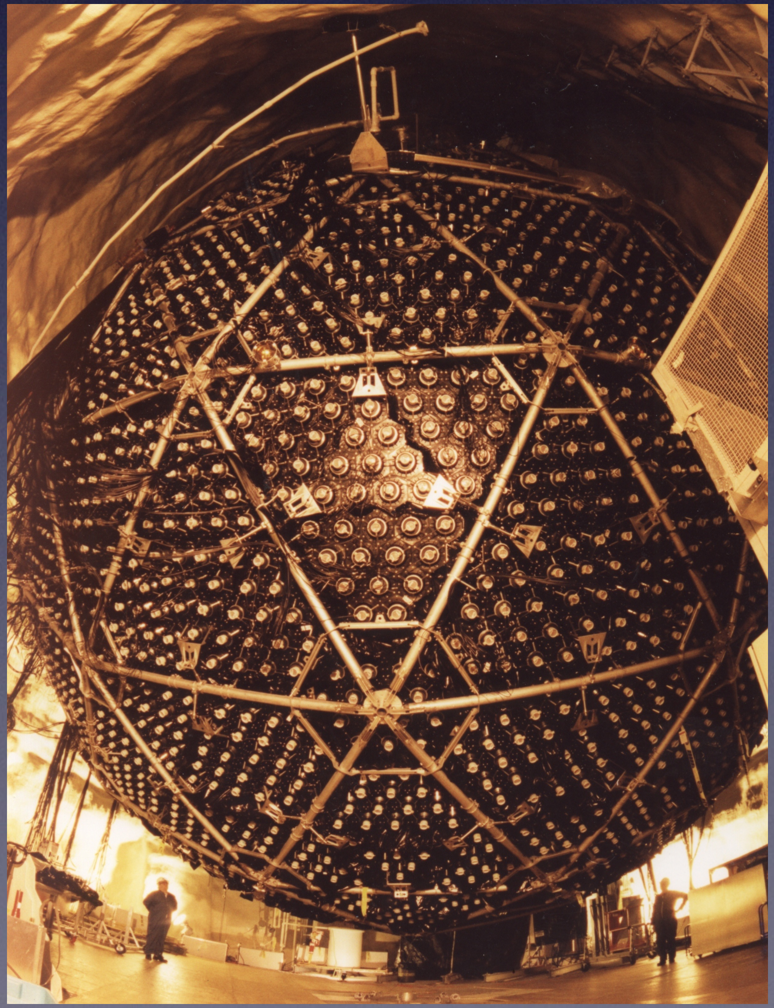


Ultra-pure water
50 kton, 11000 PMTs
(1996)

Discovery of ν_{solar} oscillation!

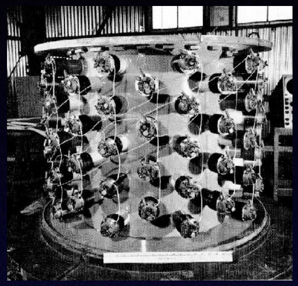


SNO
1 kton heavy-water Cherenkov,
9600 PMTs
(shared Nobel Prize 2015)

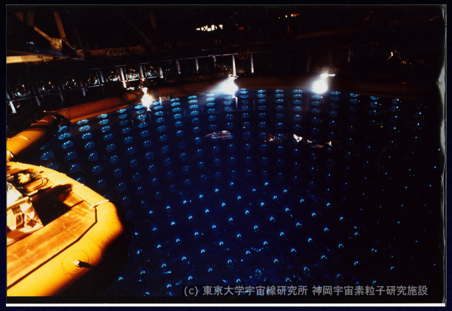


Neutrino Oscillation Experiments (I)

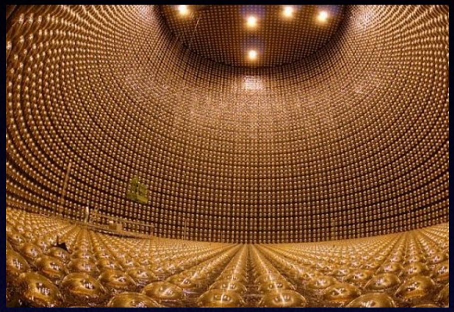
Evolution of Detectors



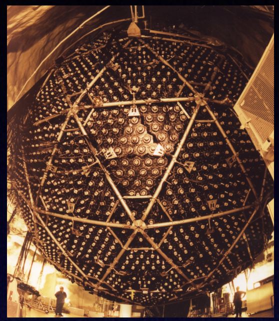
Cd-doped water
0.4 ton, 100 PMTs
(1956)



Ultra-pure water
3 kton, 1000 PMTs
(1983)



Ultra-pure water
50 kton, 11000 PMTs
(1996)



Heavy water
1 kton, 9600 PMTs
(1999)

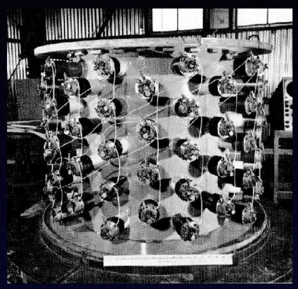
Reactor neutrino oscillation!

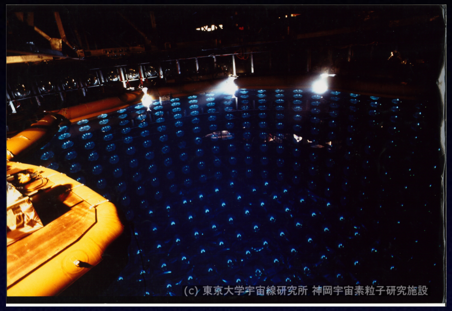
(the solar model is right!)

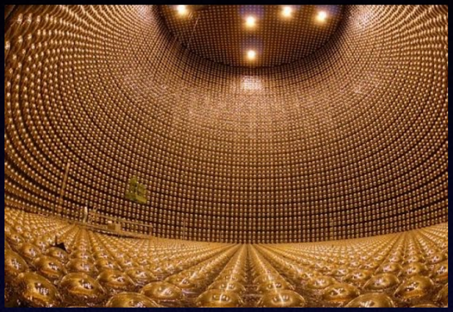
KamLAND
1 kton liquid scintillator, 1900 PMTs
My first neutrino experiment
(undergraduate RA @ UC Berkeley)

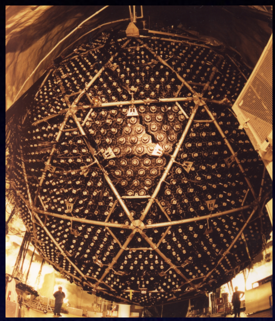
Neutrino Oscillation Experiments (I)

Evolution of Detectors

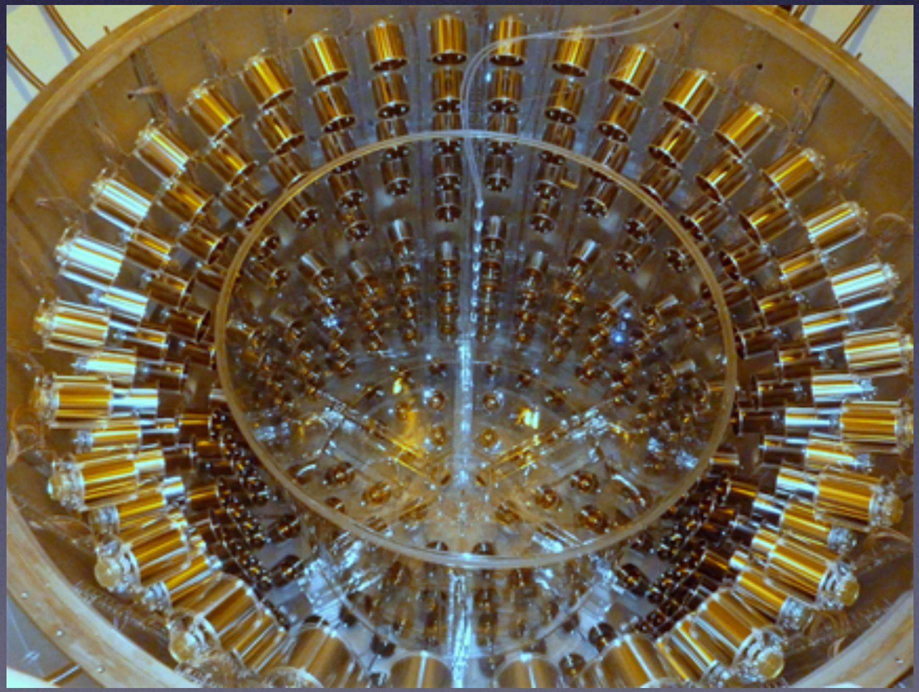

Cd-doped water
0.4 ton, 100 PMTs
(1956)


Ultra-pure water
3 kton, 1000 PMTs
(1983)


Ultra-pure water
50 kton, 11000 PMTs
(1996)


Heavy water
1 kton, 9600 PMTs
(1999)


Liquid Scintillator
1 kton, 1900 PMTs
(2002)

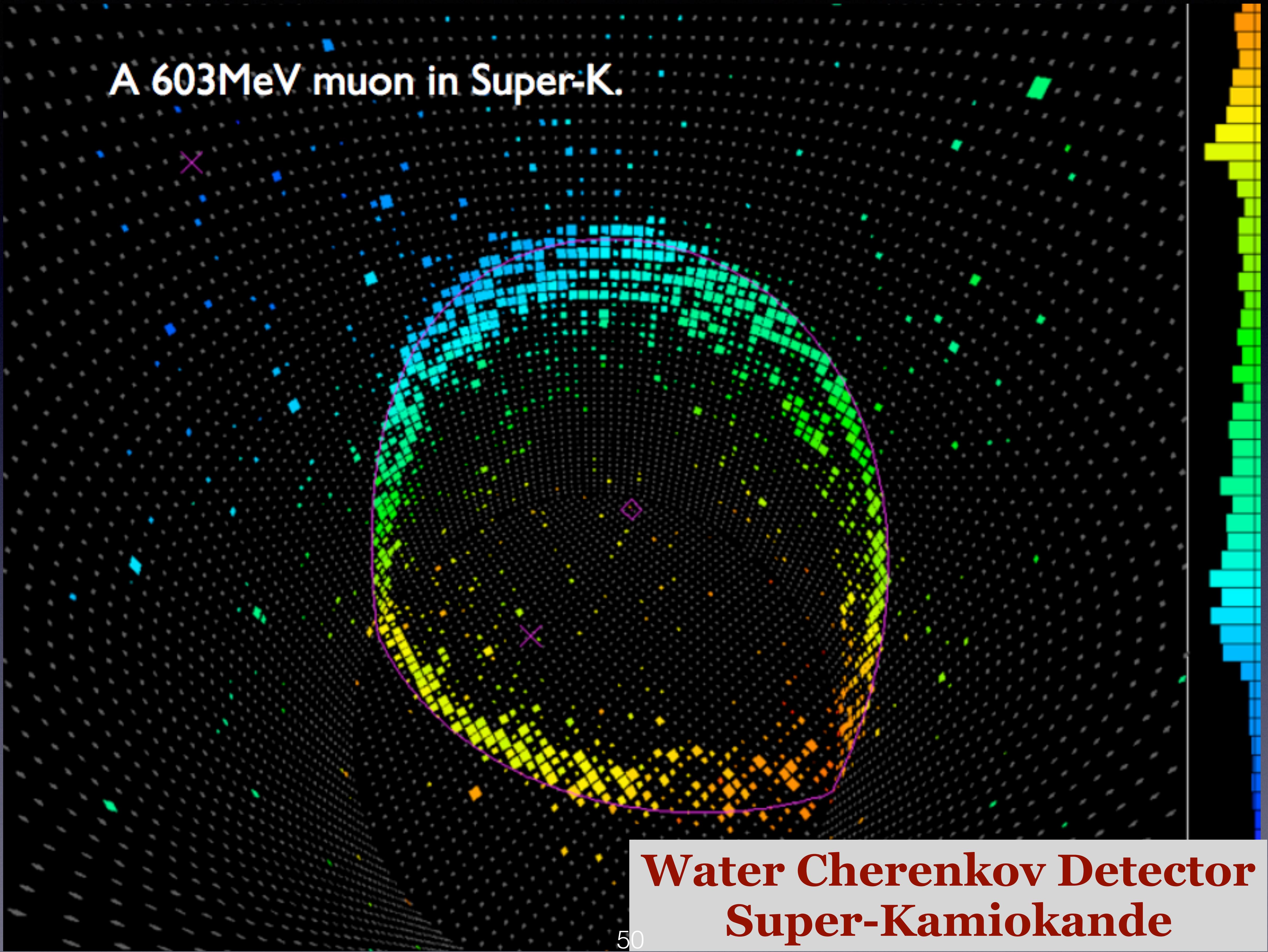

Gd-doped liquid scintillator
RENO, Daya Bay, Double Chooz

"Near" & "Far" design
2 x 16 ton detectors with 400 PMTs each (Double Chooz)

My Ph.D thesis! (MIT)

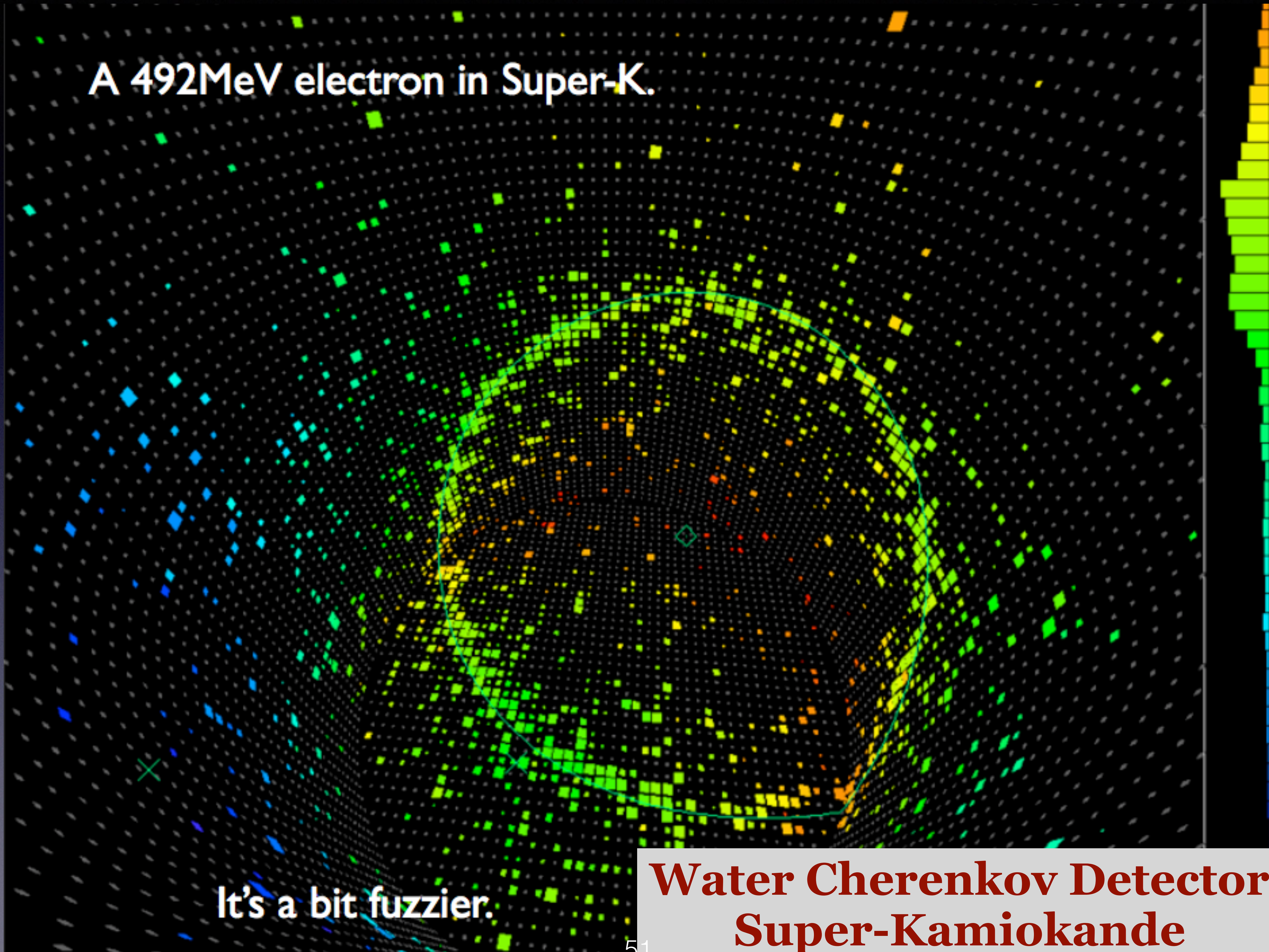
"Last mixing angle"
θ_{13} Experiments!

Neutrino Oscillation Experiments (I)



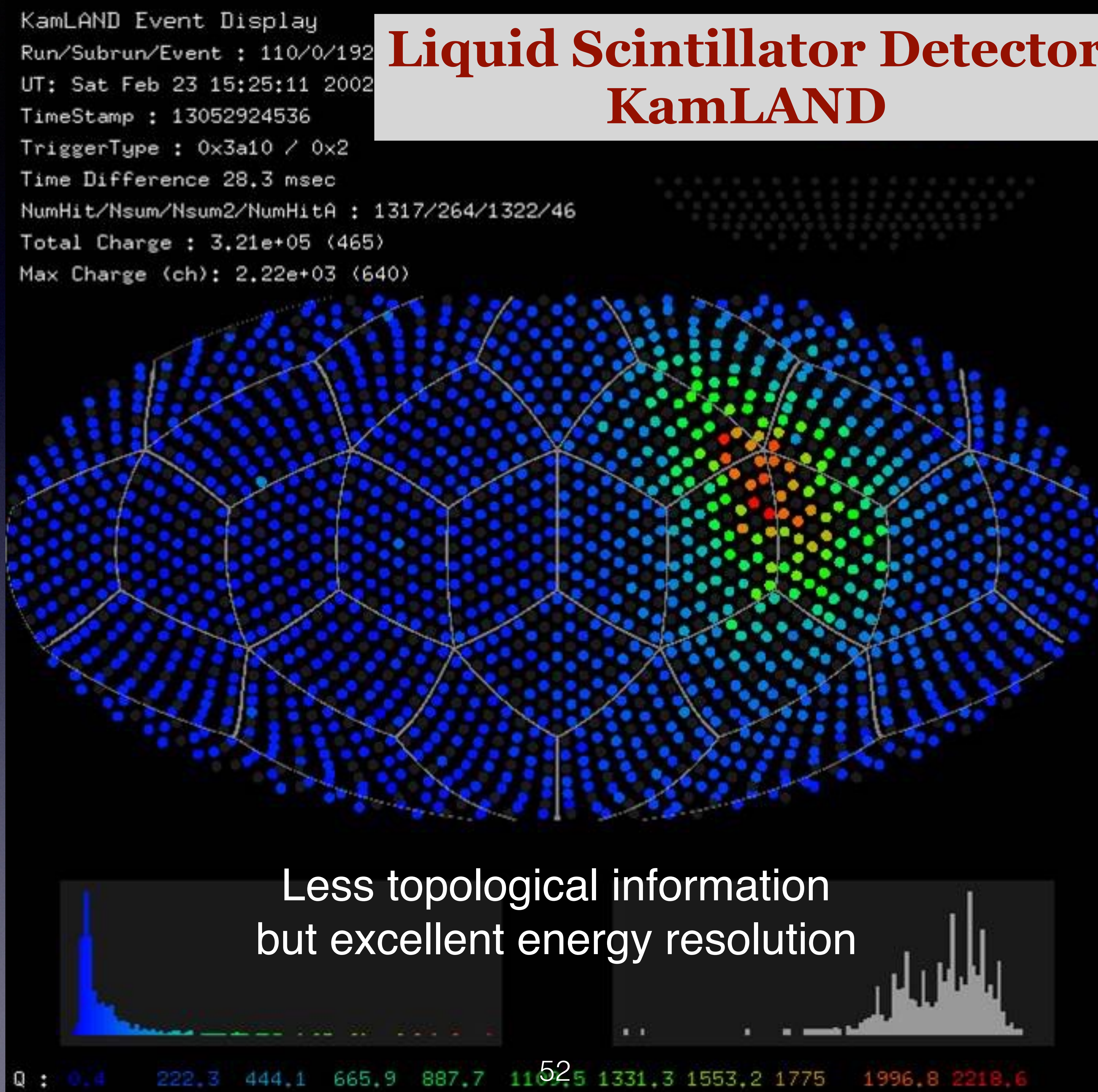
Water Cherenkov Detector
Super-Kamiokande

Neutrino Oscillation Experiments (I)



Water Cherenkov Detector
Super-Kamiokande

Neutrino Oscillation Experiments (I)



Liquid Scintillator Detector
KamLAND

Less topological information
but excellent energy resolution

Neutrino Oscillation Experiments (II)

How can we do better?

Three important detector features for oscillation measurement

$$P(\nu_\mu \to \nu_{\mathrm{e}}) = \sin^2 2\theta \sin^2\left(\frac{1.27\, \Delta m^2\, L}{E_\nu}\right)$$

Good Energy Resolution

Precise E_ν reduce oscillation uncertainty

Large Mass (scalable)

"More" statistics to measure rare physics process

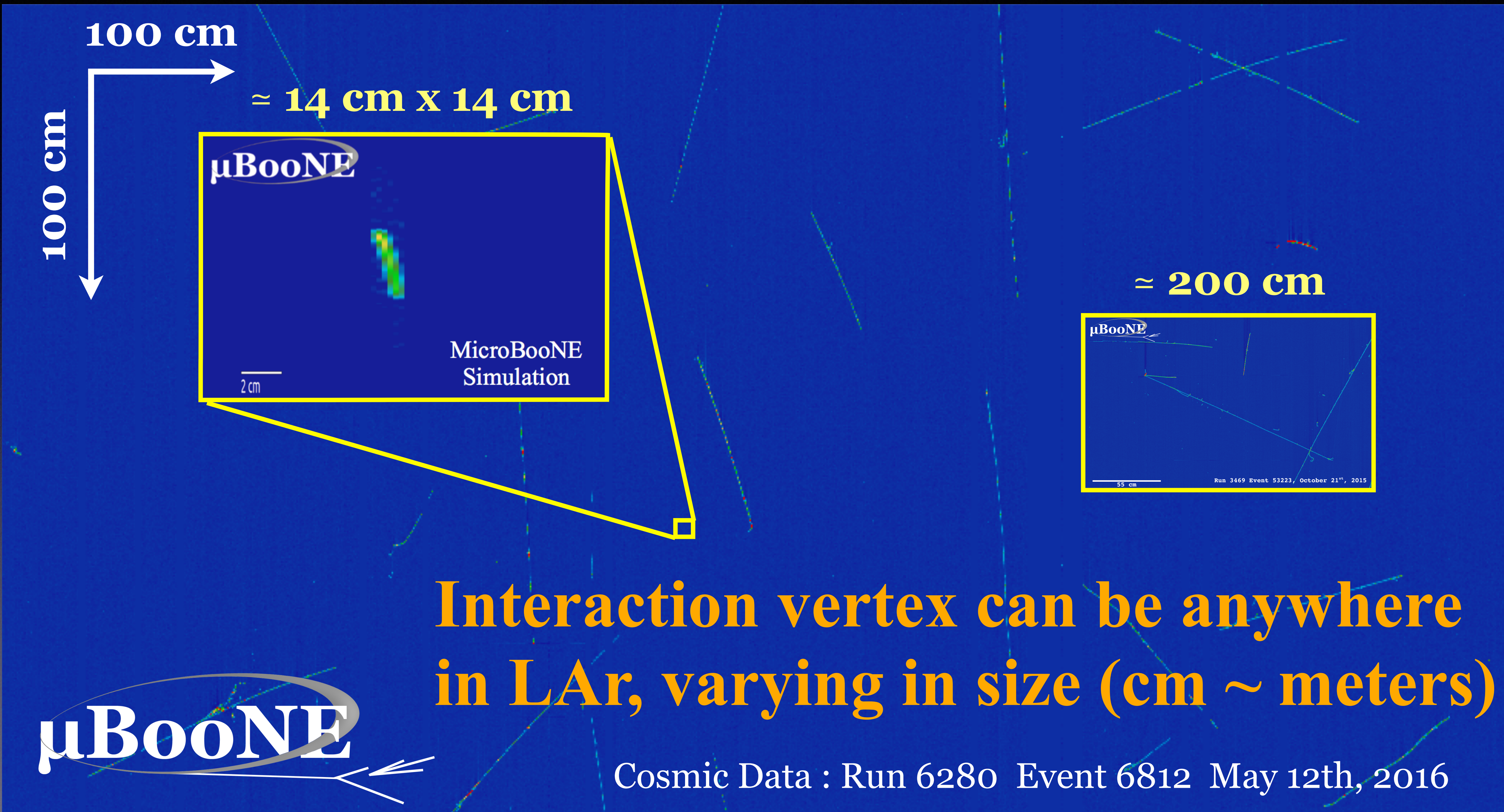
Particle ID Capability

Better ν identification background rejection

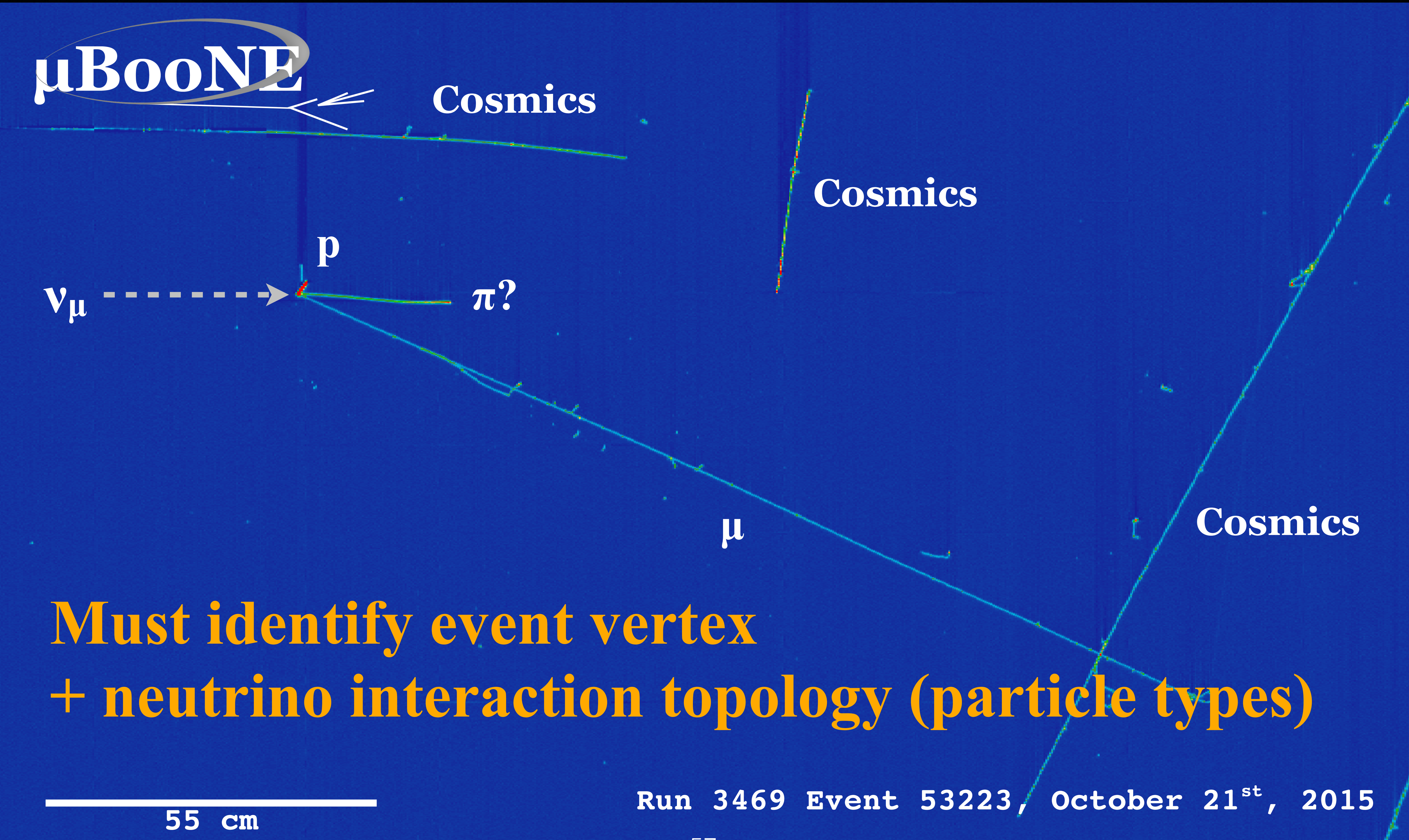
Challenges

Analysis Challenges

Analysis Challenges

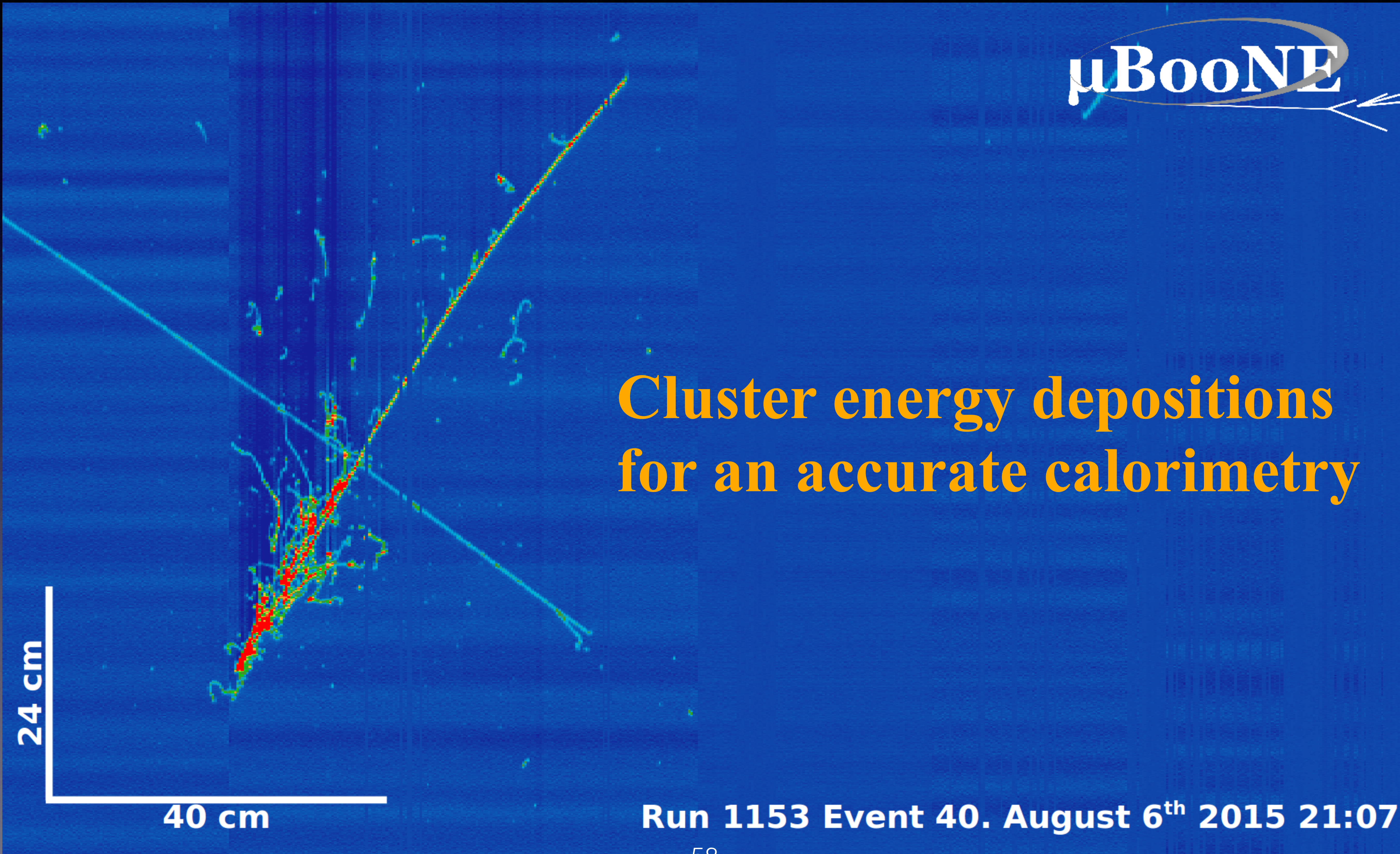


Analysis Challenges



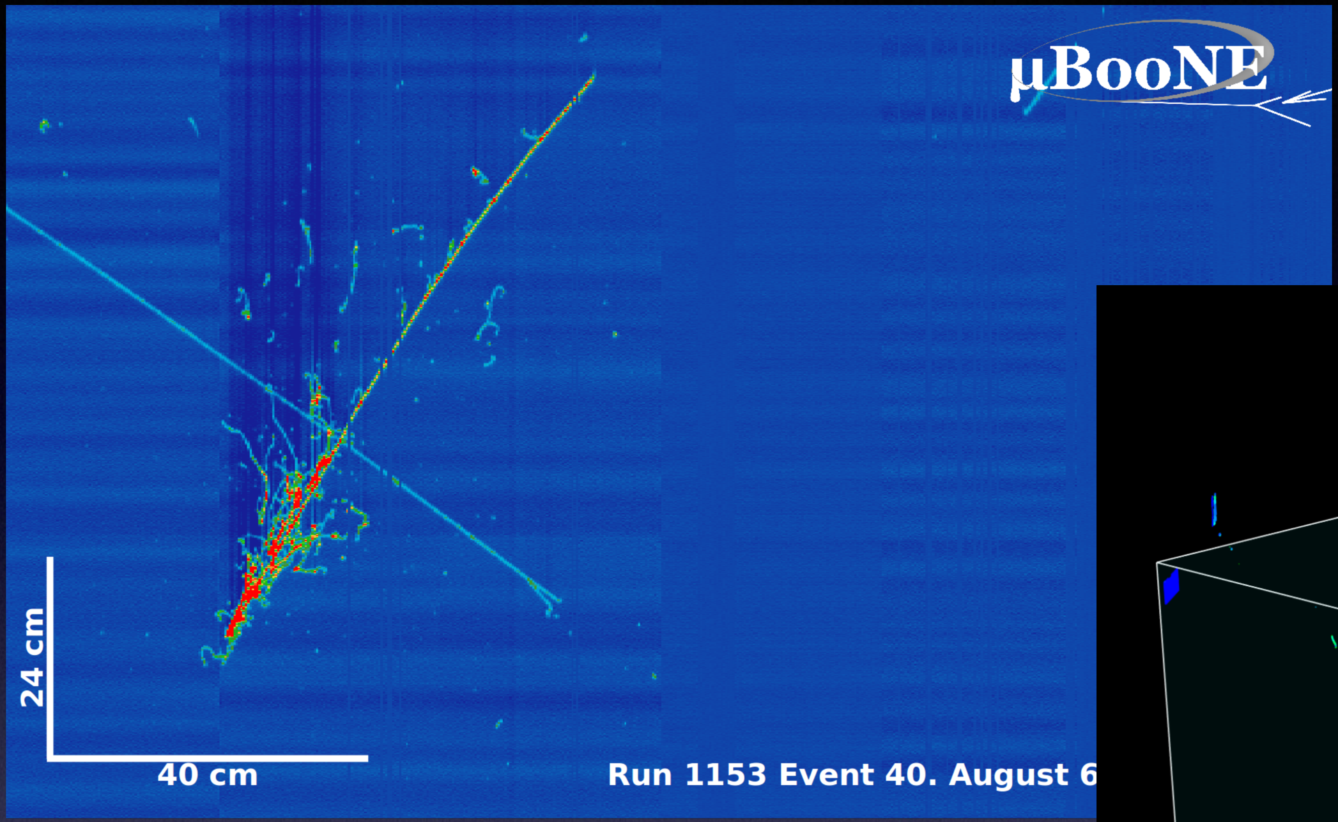
Must identify event vertex
+ neutrino interaction topology (particle types)

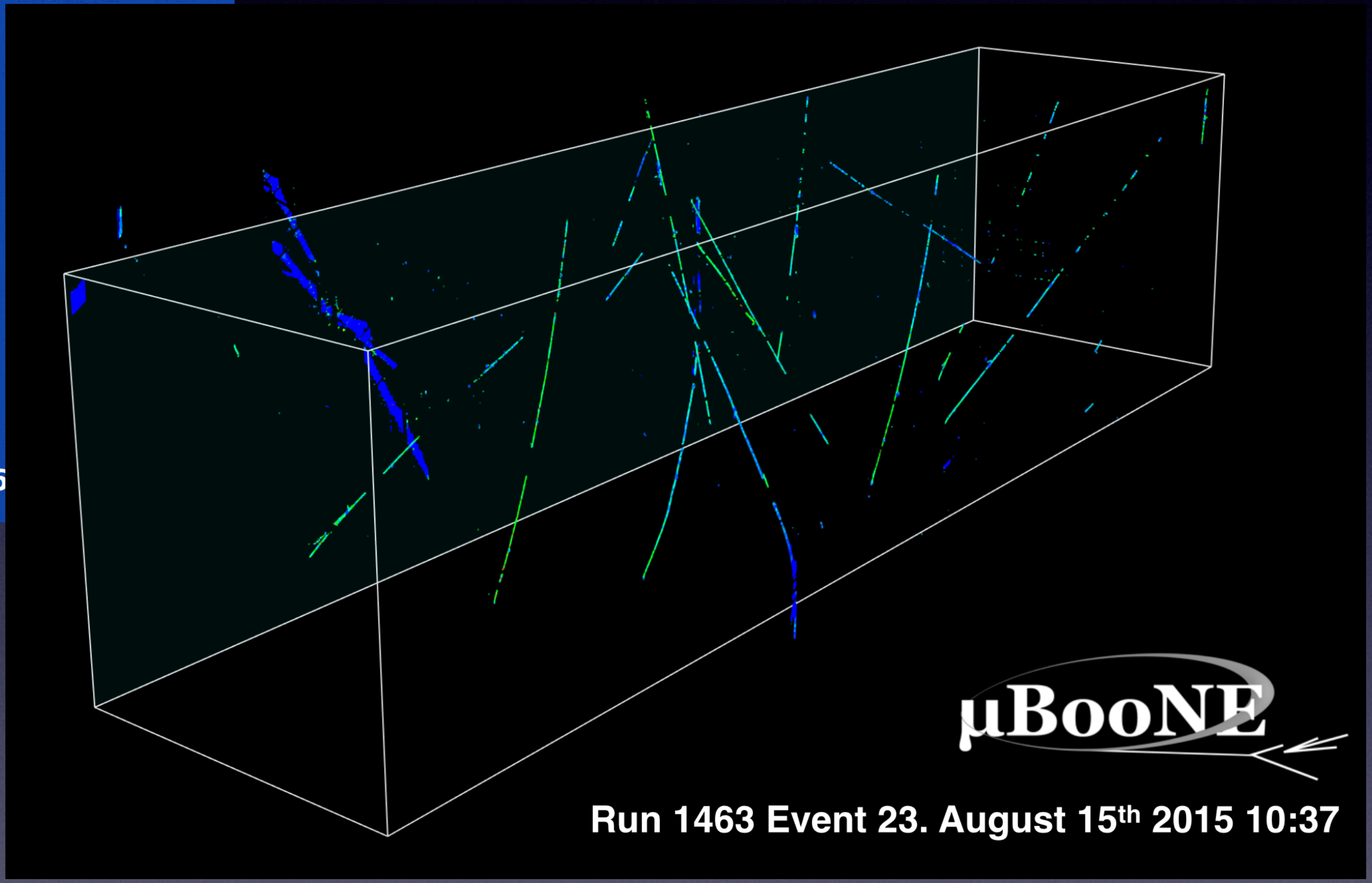
Analysis Challenges



Cluster energy depositions for an accurate calorimetry

Analysis Challenges





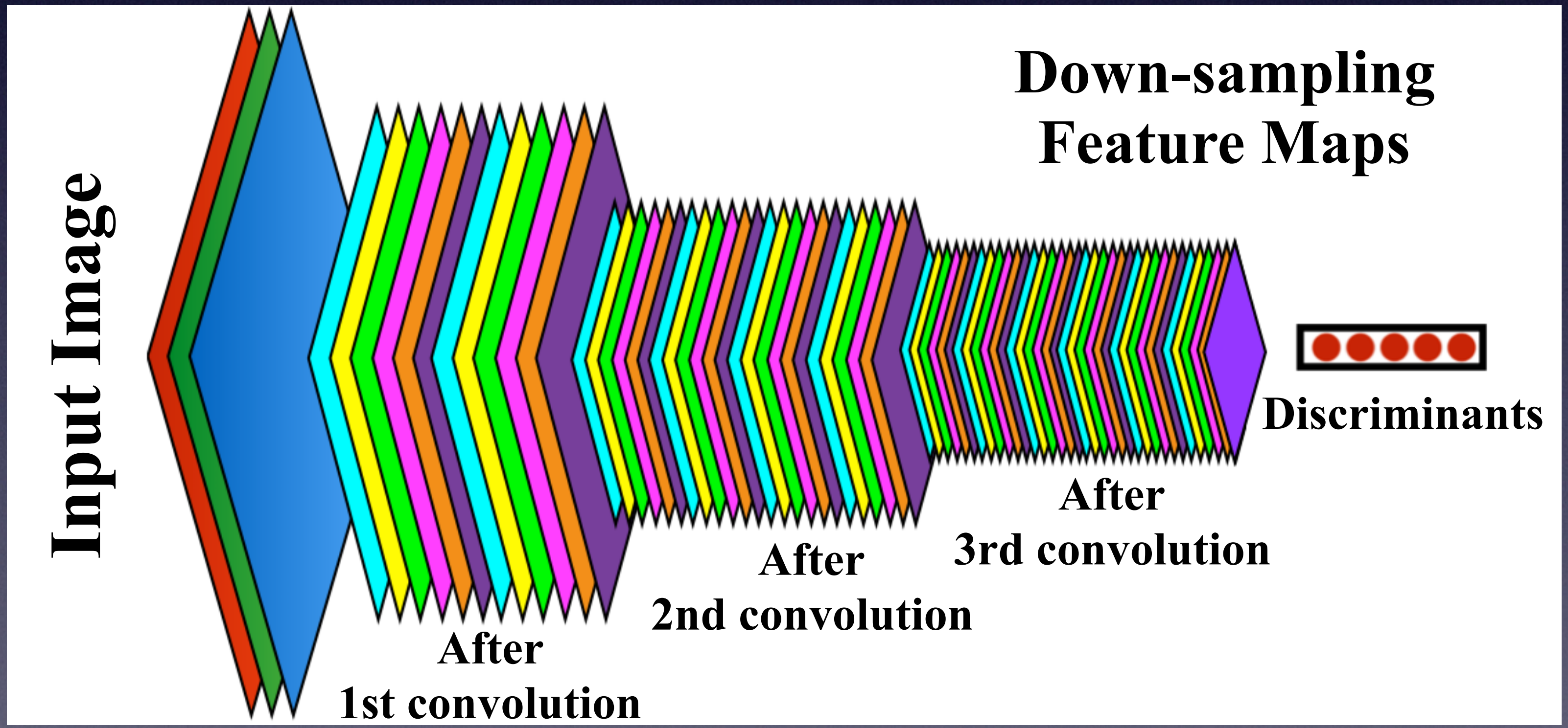
Deal with optical illusions in 2D projections + 3D pattern recognitions

NN & CNN
Basics
~ How Does It Work? ~

How Image Classification Networks Work

Goal: extract features to give “single label” to an image

1. **Convolution operation**
2. **Down-sampling**

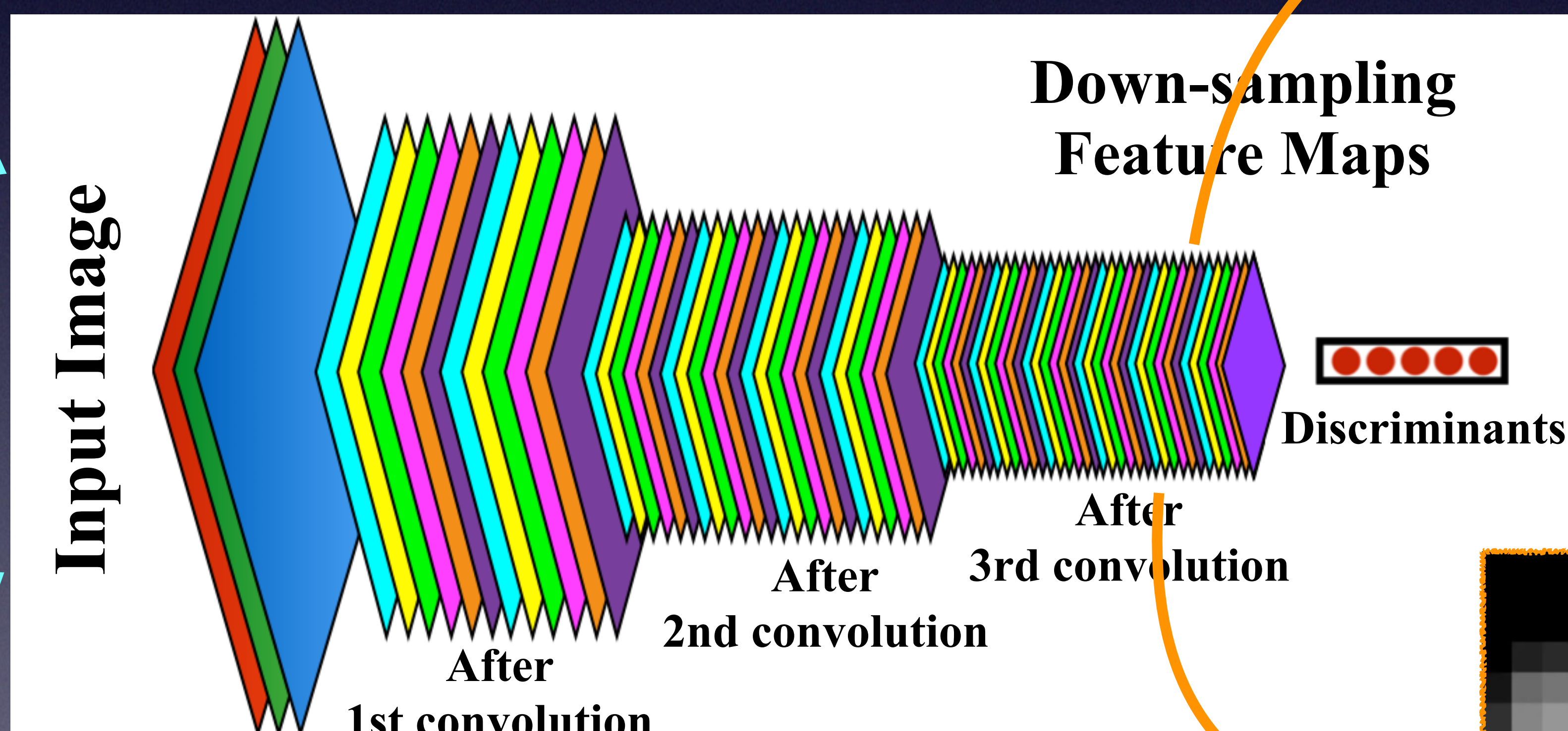


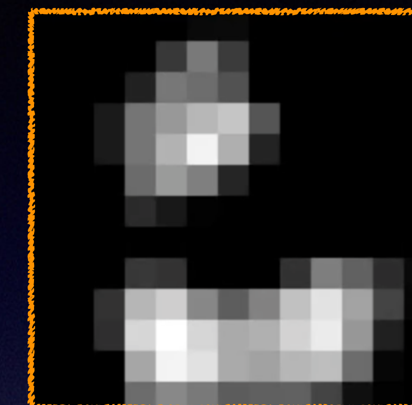
Series of convolutions
+ down-sampling

How Image Classification Networks Work

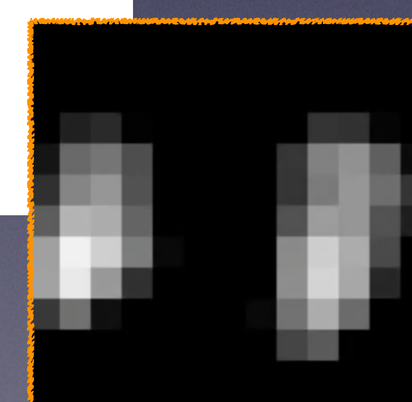
Goal: extract features to give “single label” to an image

1. **Convolution operation**
2. **Down-sampling**





“Written Texts” feature map



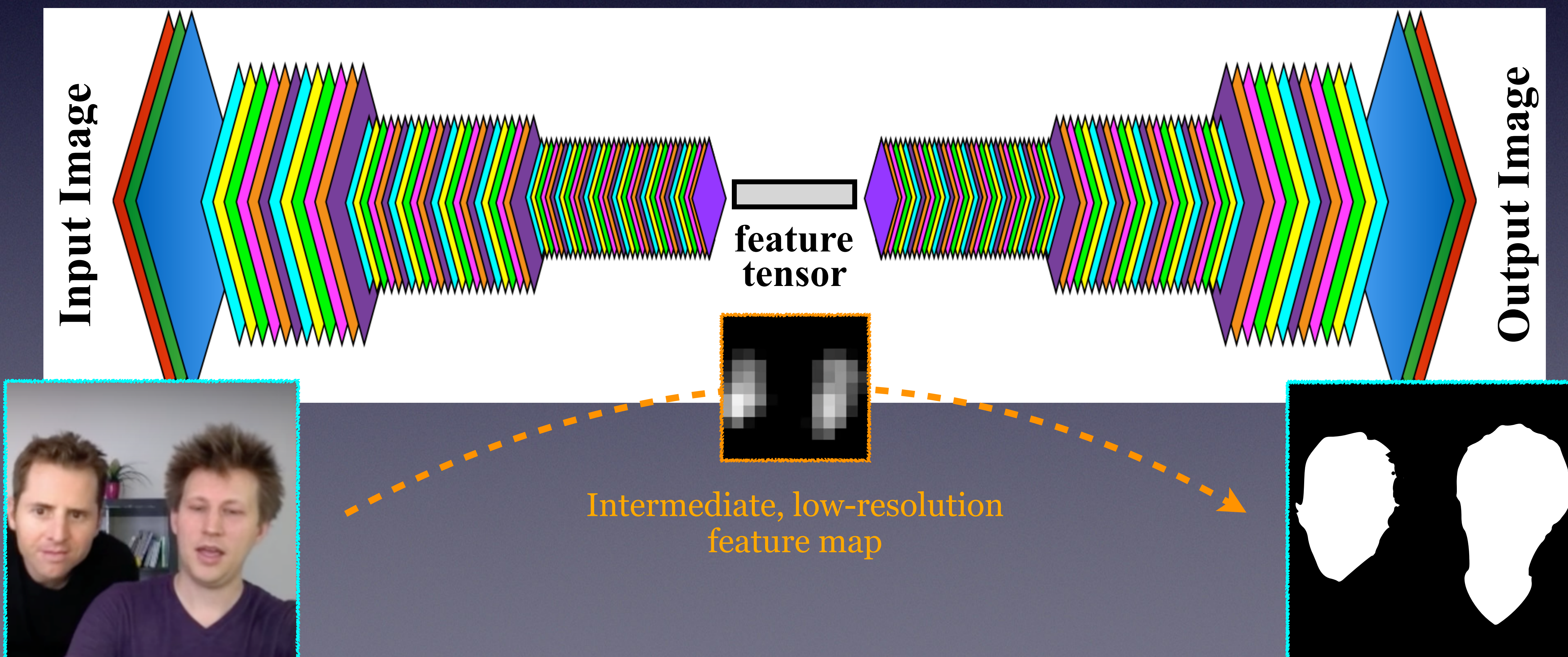
“Human Face” feature map

Series of convolutions
+ down-sampling

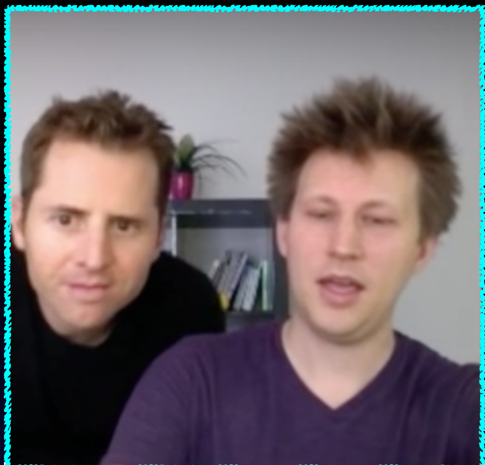
How SSNet Works

Goal: recover precise, pixel-level location of objects

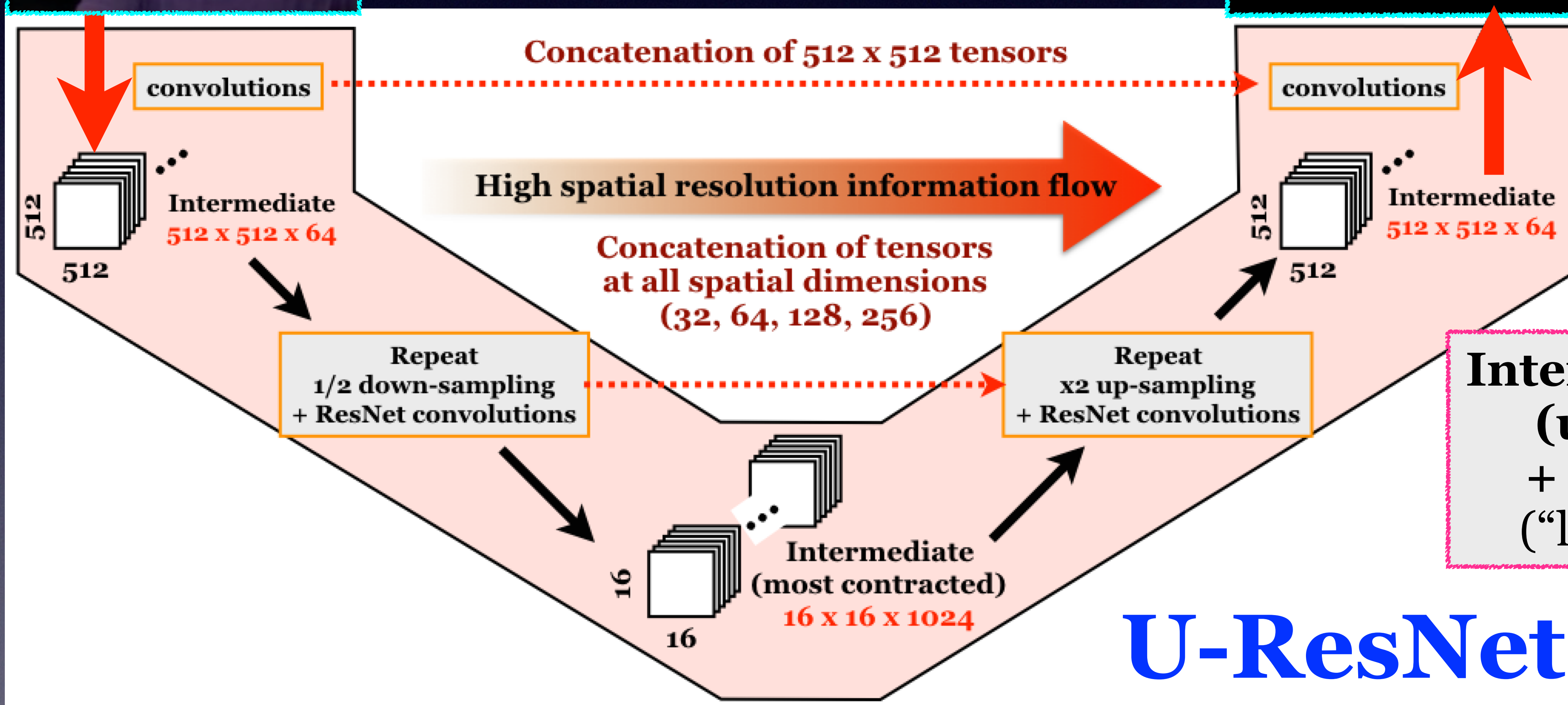
1. **Up-sampling**
   - Expand spatial dimensions of feature maps
2. **Convolution**
   - Smoothing (interpolation) of up-sampled feature maps



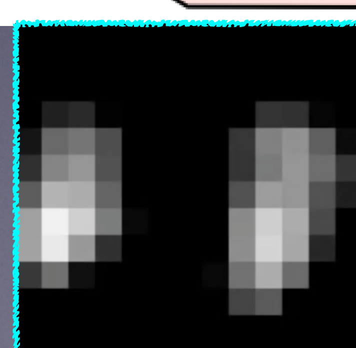
DNN for LArTPC Data Reconstruction



How does U-ResNet Work?



Interpolation filters (up-sampling) + Convolutions (“learnable” filter)



Down sampling + Convolutions to identify highly abstract features (e.g. “human face”)

Misc.

Response Study on Real Data

- Physicist labeled pixels, compared with DNN
- Repeated procedure for real detector data and simulation sample. Validated response on real data