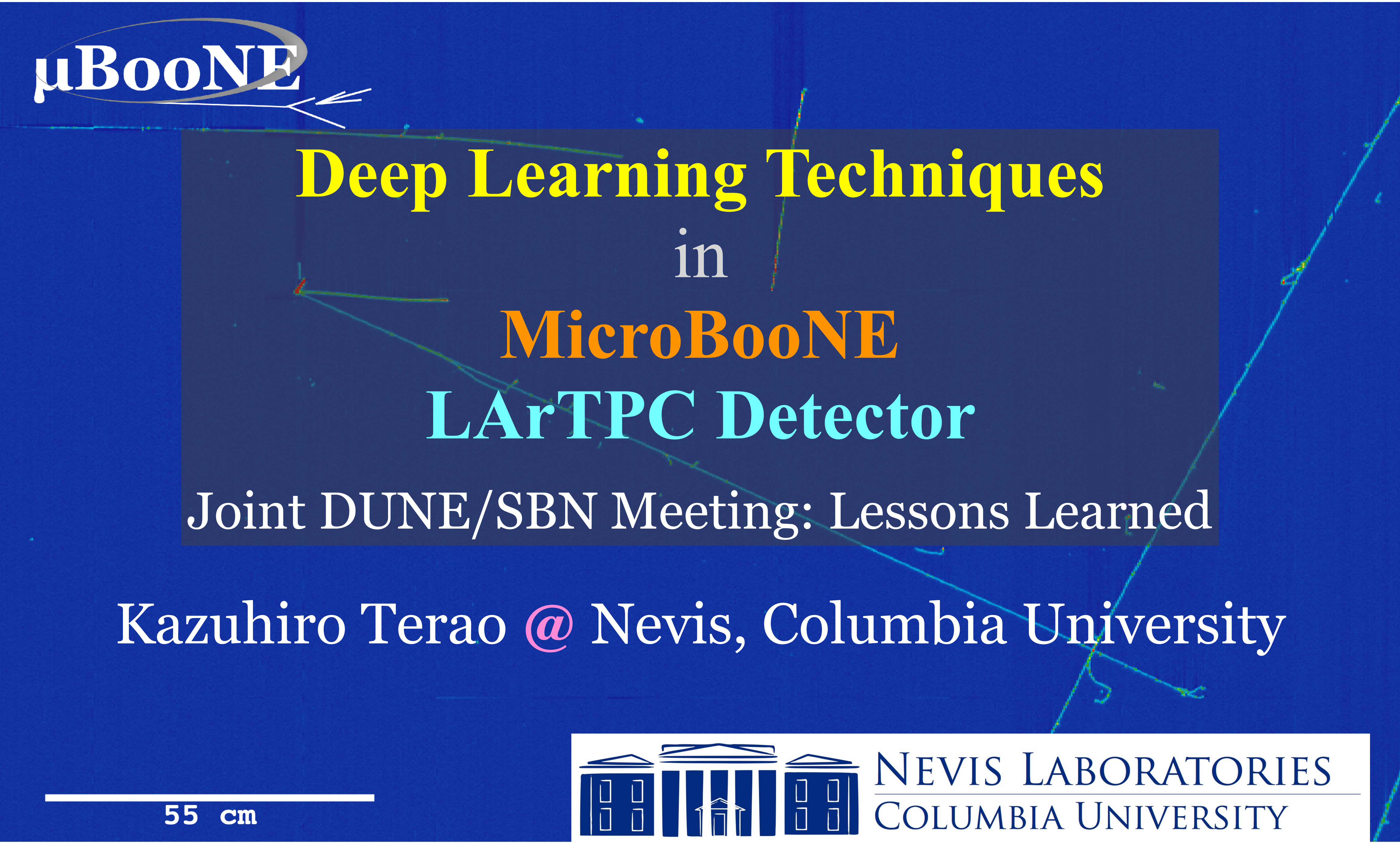
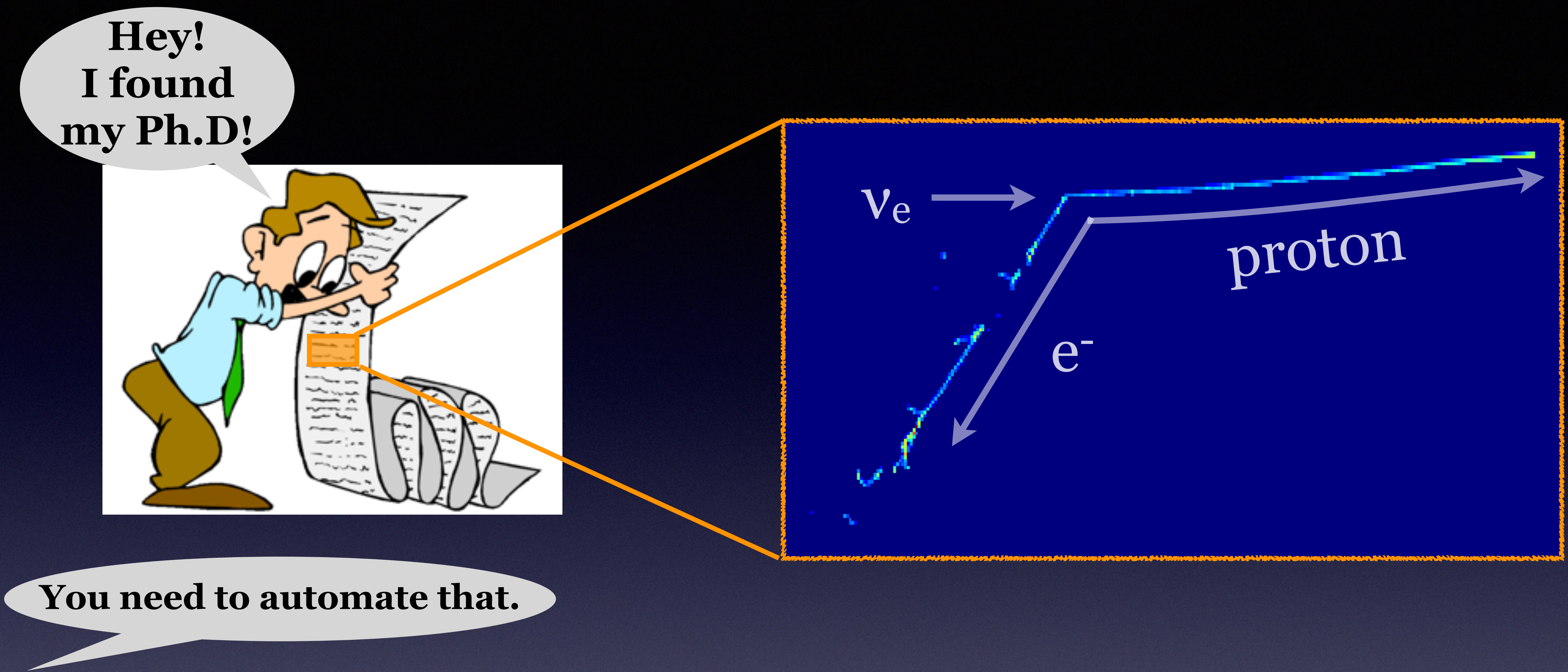
Deep Learning Techniques
in
MicroBooNE
LArTPC Detector

Joint DUNE/SBN Meeting: Lessons Learned

Kazuhiro Terao @ Nevis, Columbia University



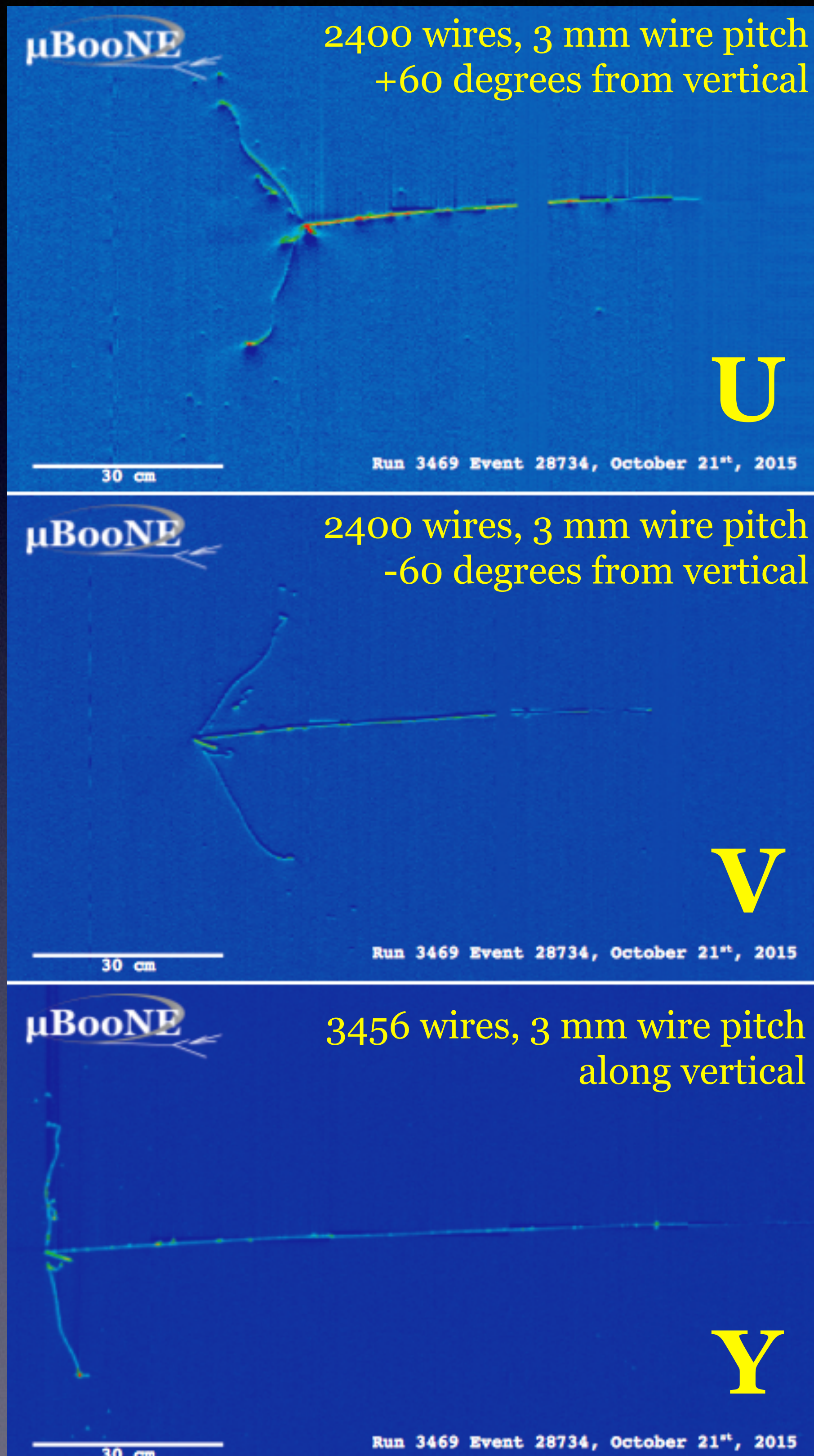
Nevis Laboratories
Columbia University



Outline

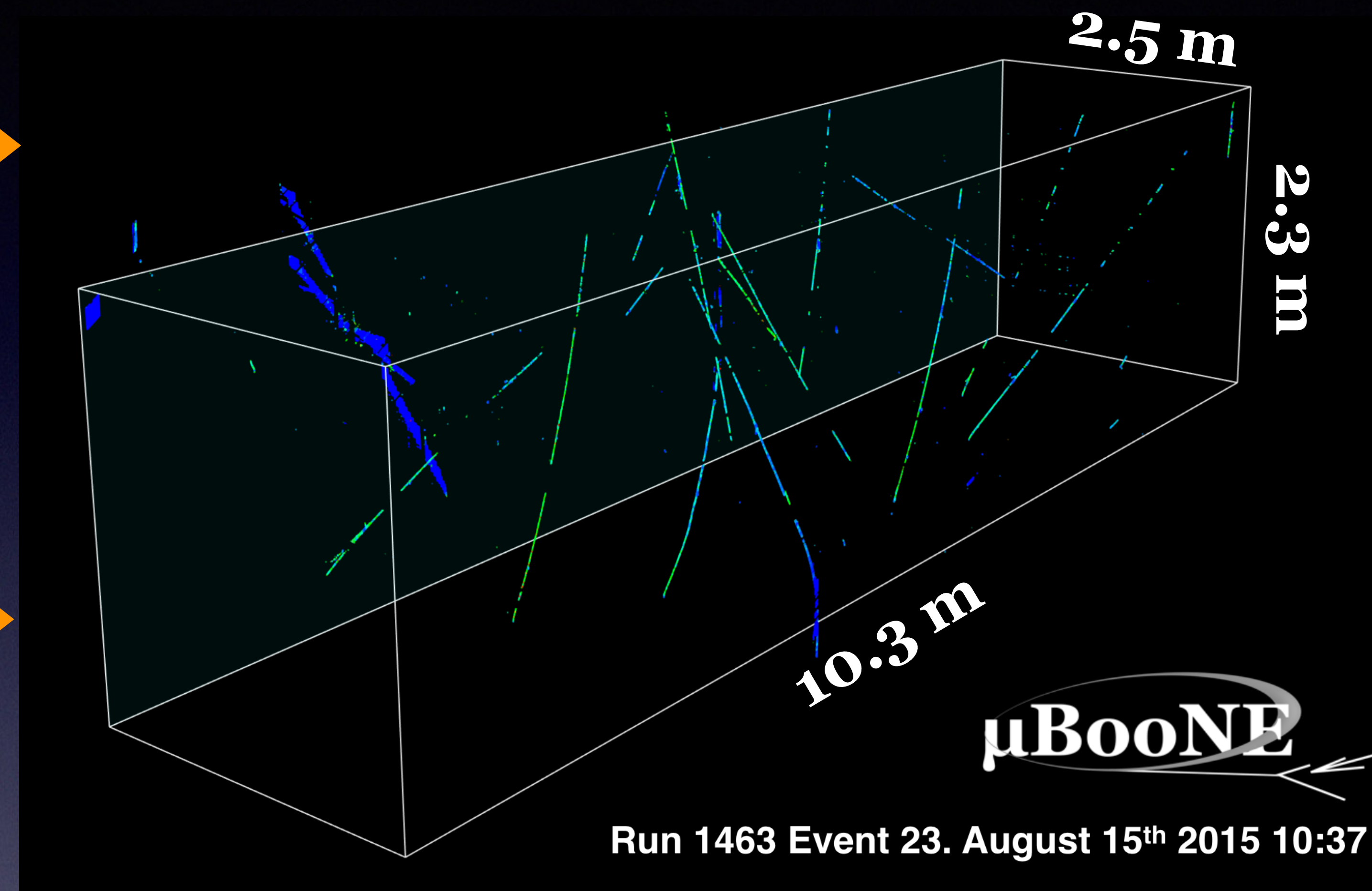
- **MicroBooNE and Deep Neural Networks**
- Deep Learning “lessons learned”
- Deep Learning “lessons learning”
- Summary

LArTPC: Particle Imaging Machine



Three 2D Views

Reconstructed 3D View

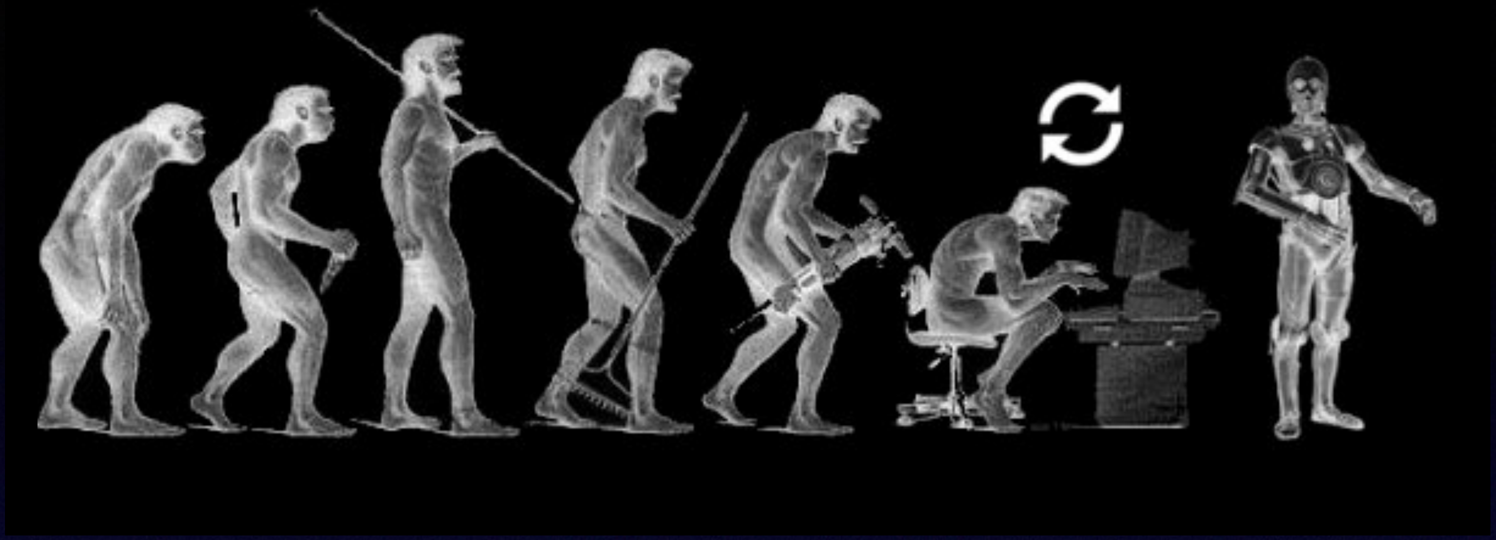


WireCell 3D reconstruction

Reconstruction is a challenging task…

… so is analysis!

Data Reconstruction / Analysis Challenge



Solutions?

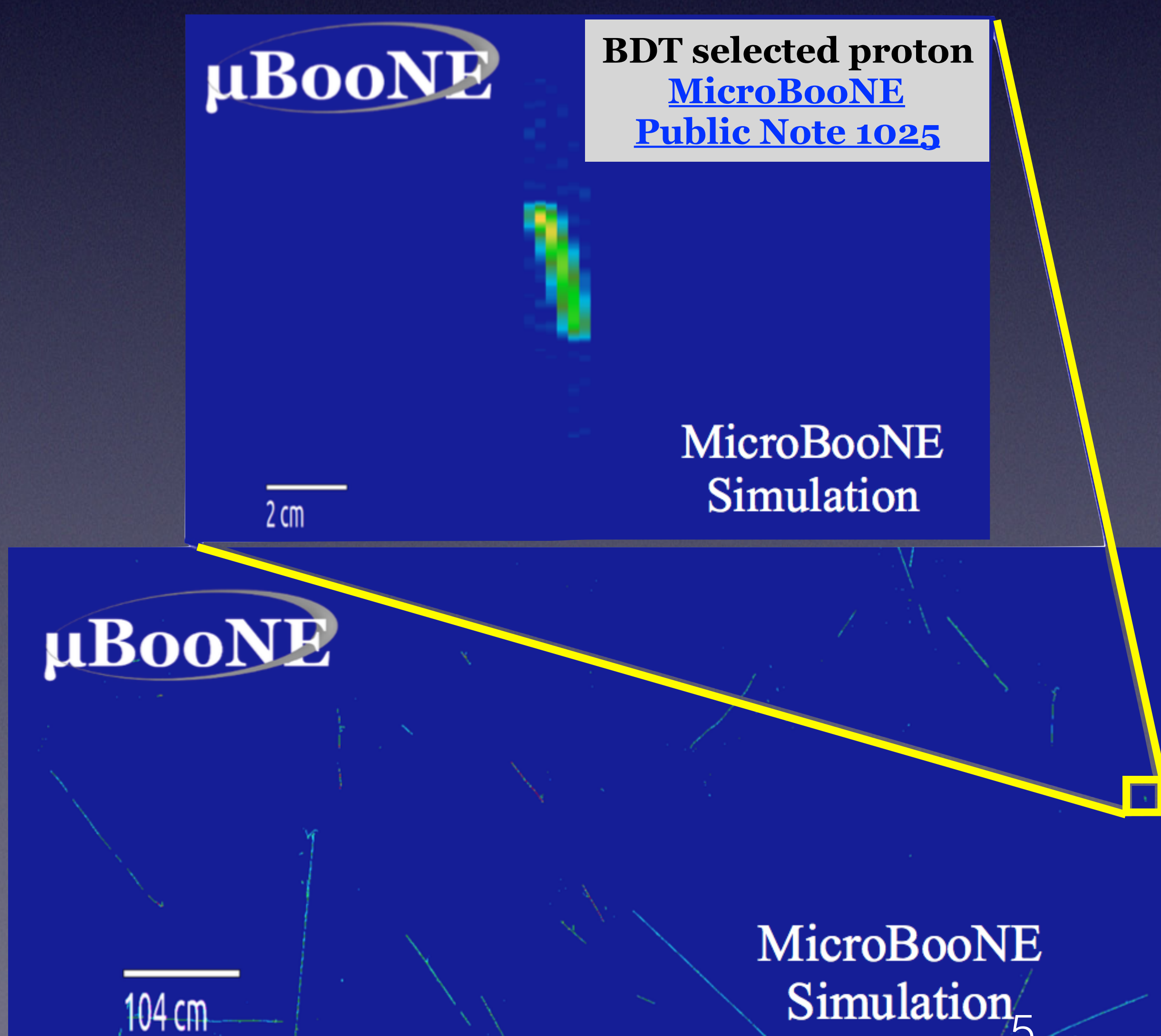
- **Path A**: "**traditional path**"
  - Hand-engineered reconstruction algorithms
- **Path B**: **machine learning**
  - "**Deep Learning**"
    - In particular...
      Convolutional Neural Networks (CNNs)
    - Scalable technique, generalizable to various tasks
    - Superb performance on image data analysis

Machine Learning

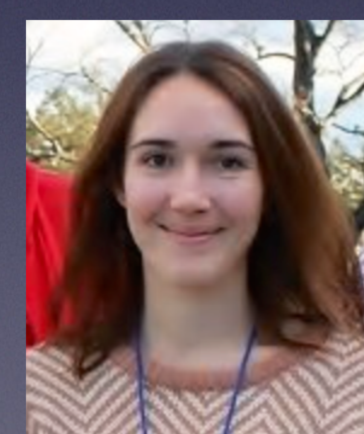
Deep ...ing Techniques in UB

Boosted Decision Tree

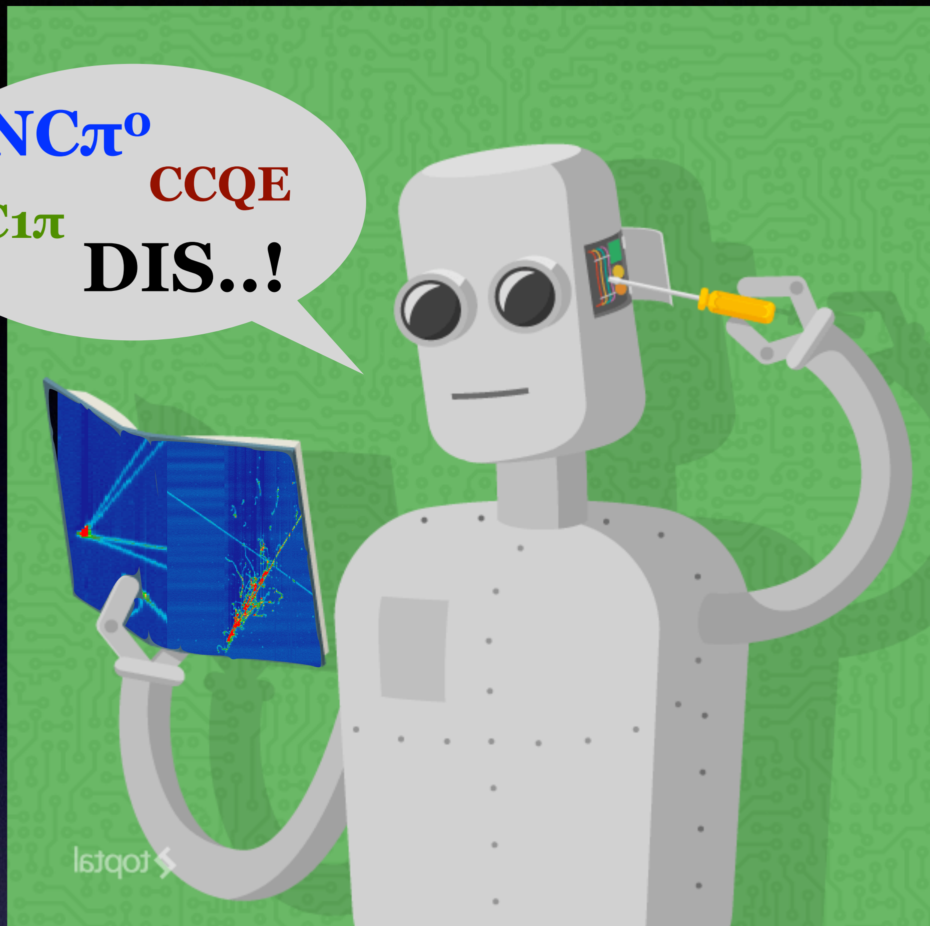
- Used for low energy (>40 MeV) single proton search
- Input: reconstructed parameters (length, angle, etc...)
- Analysis details available in [UB public note page]

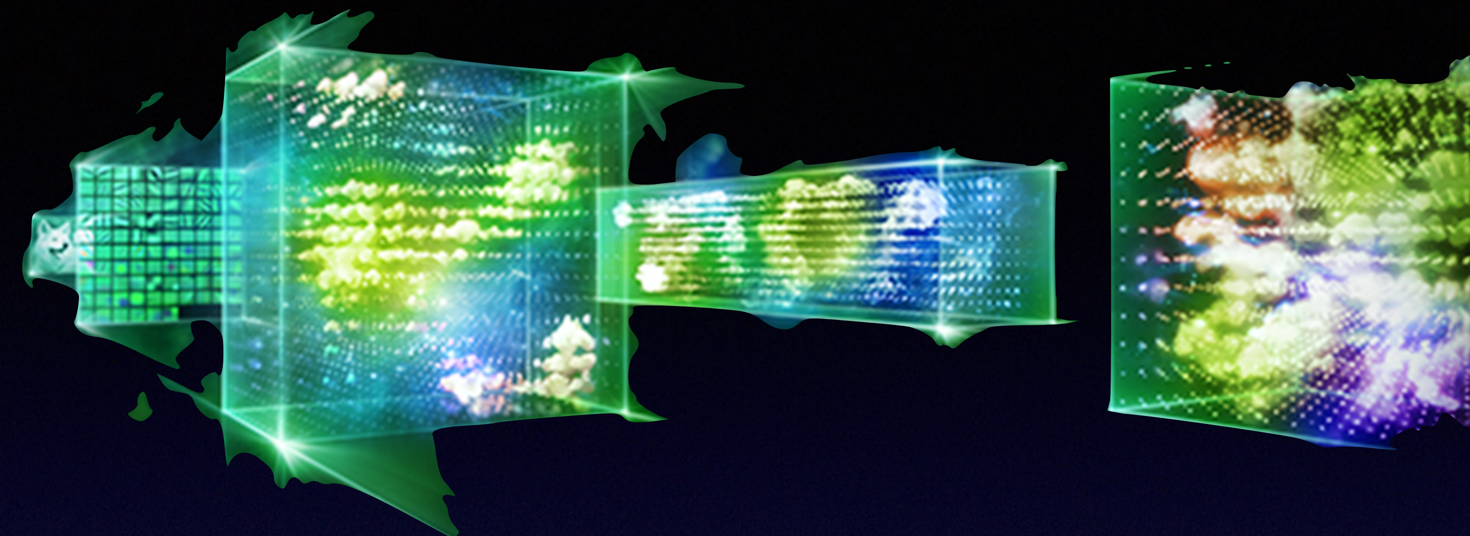


Developed by



Katherine Woodruff
(NMSU)



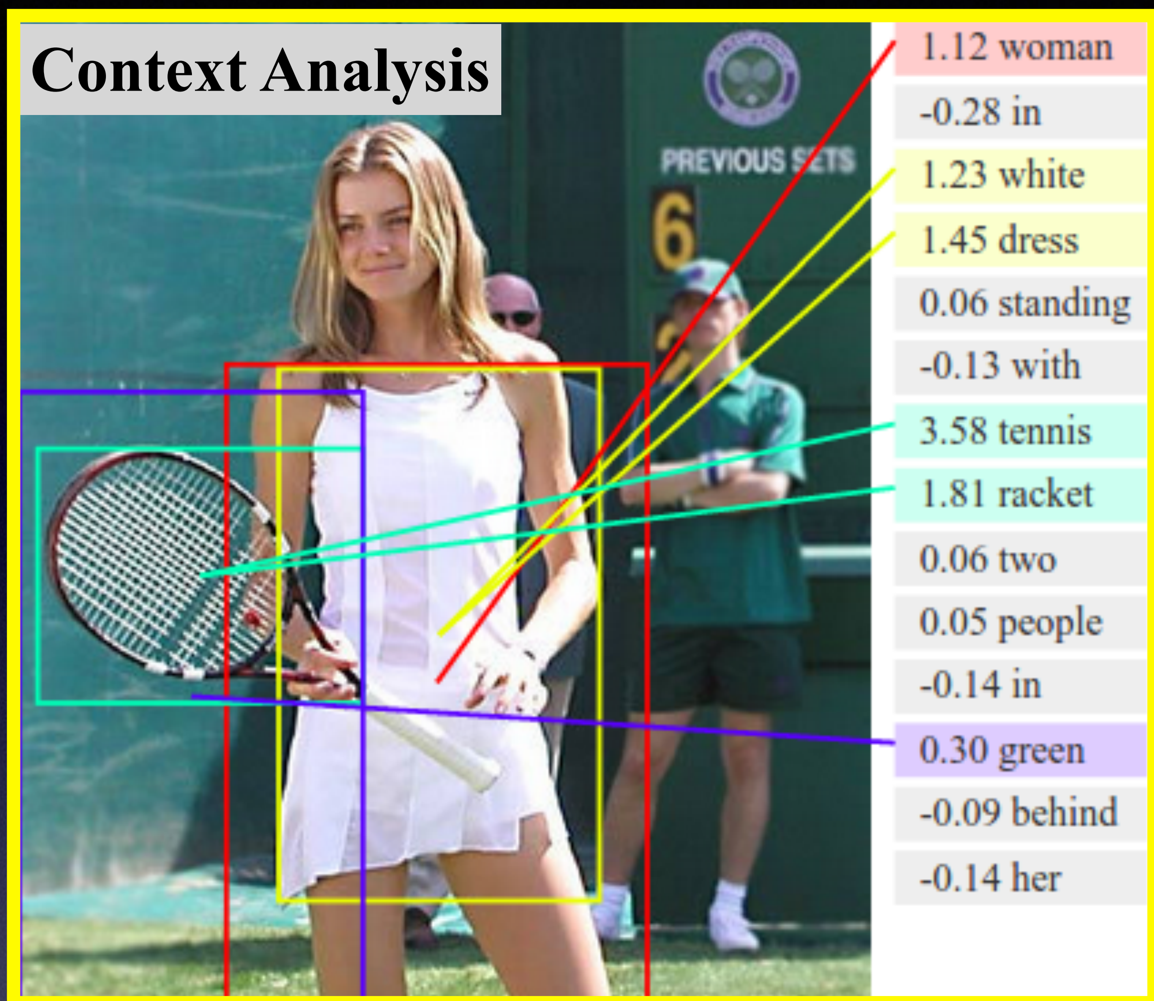


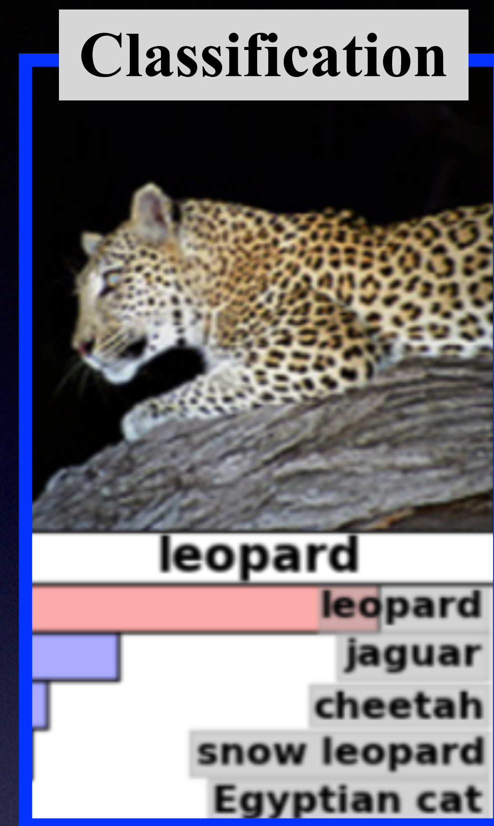
Convolutional Neural Networks for **LArTPC Analysis**

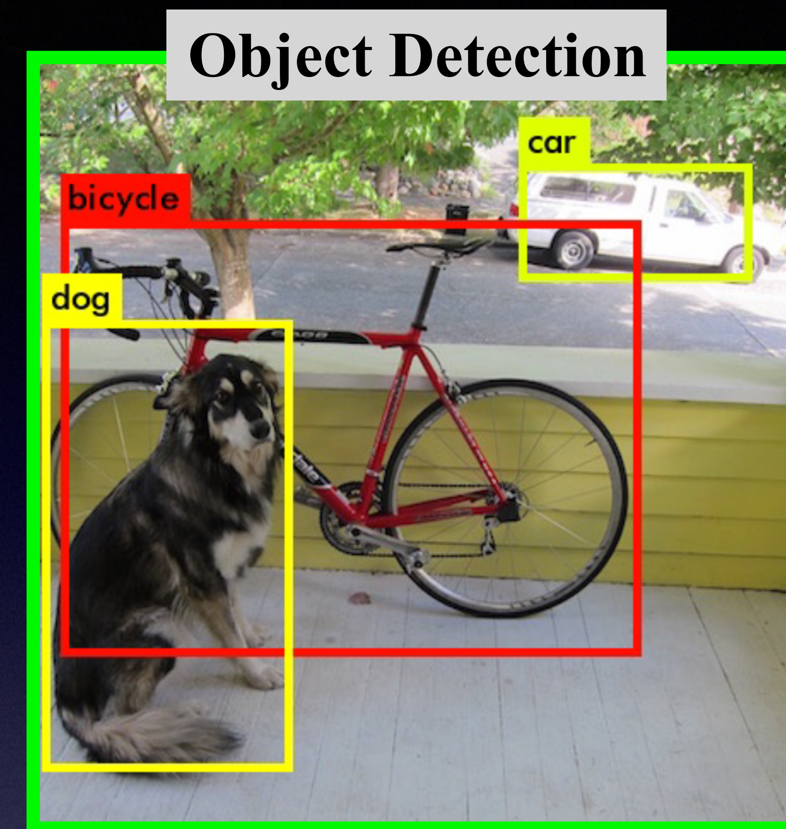
Outline

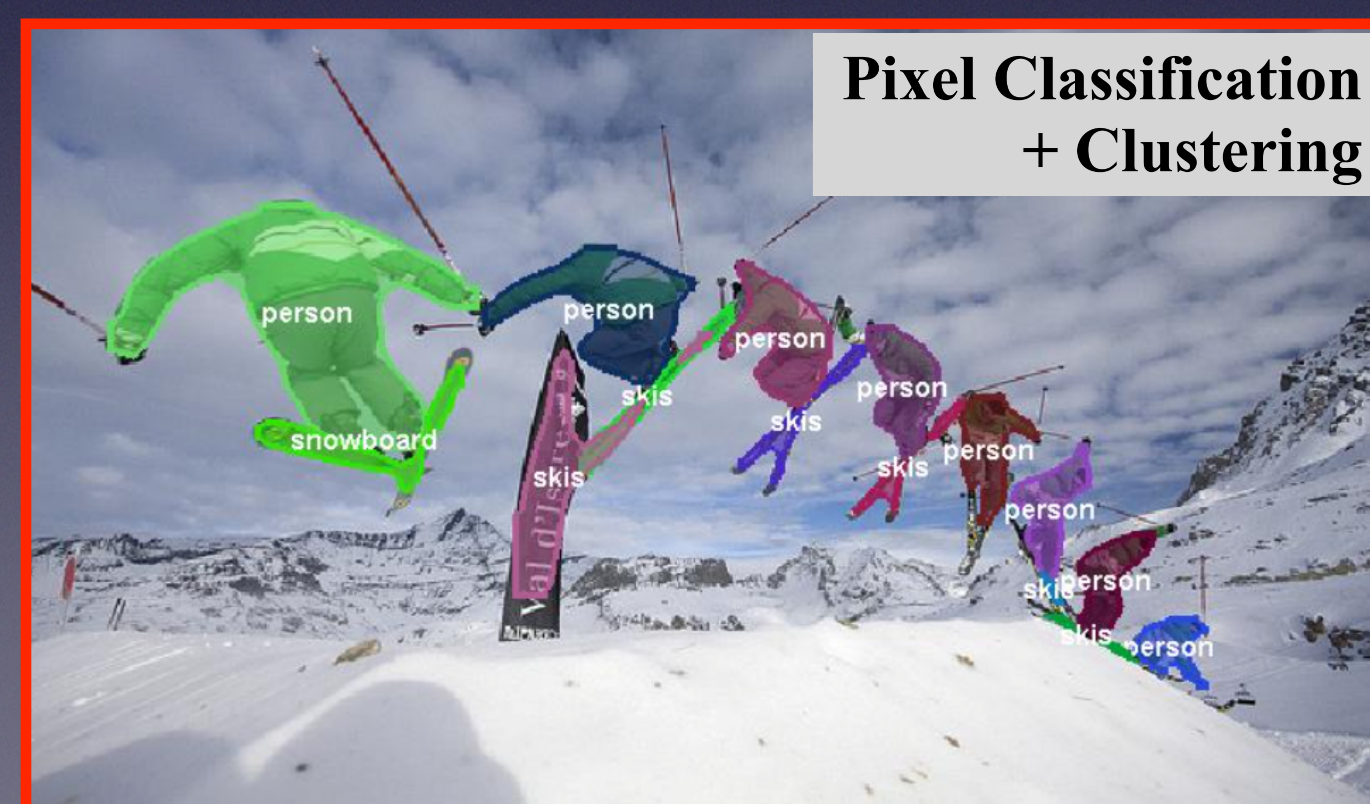
- MicroBooNE and Deep Neural Networks
- **Deep Learning “lessons learned”**
- Deep Learning “lessons learning”
- Summary

CNNs for Image Analysis







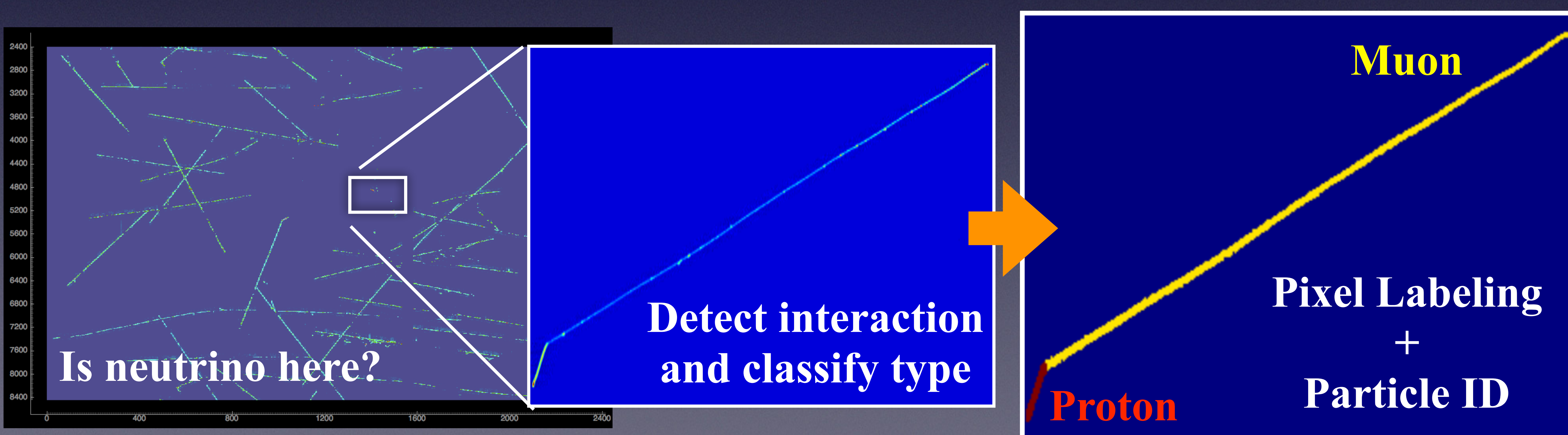


- Superb image analysis capabilities
- Trainable from raw data (large tensor)

CNN for Event Reconstruction

CNN-based reconstruction tools in MicroBooNE

- **Event selection** (image classification)
- **Vertex finding** (object detection)
- **Clustering** (semantic segmentation)
- **Particle identification** (image classification)

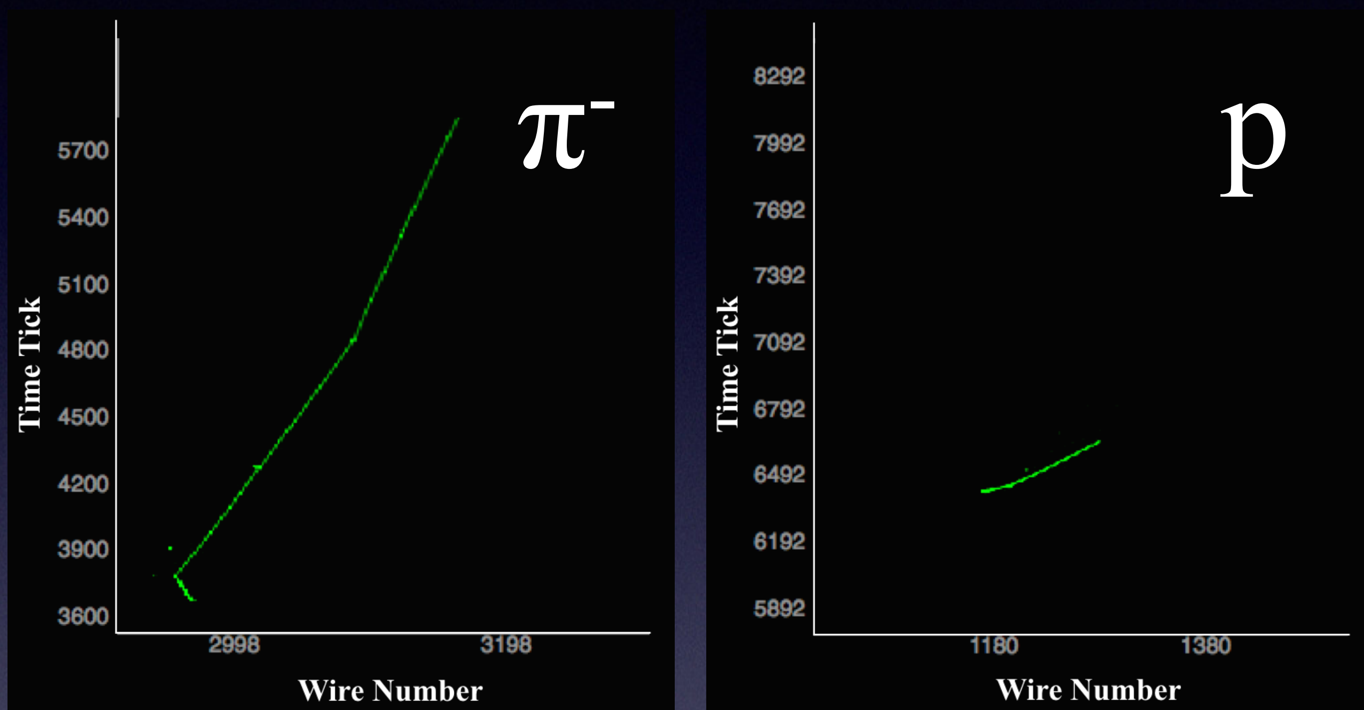


$$\nu_\mu + n \rightarrow \mu + p$$

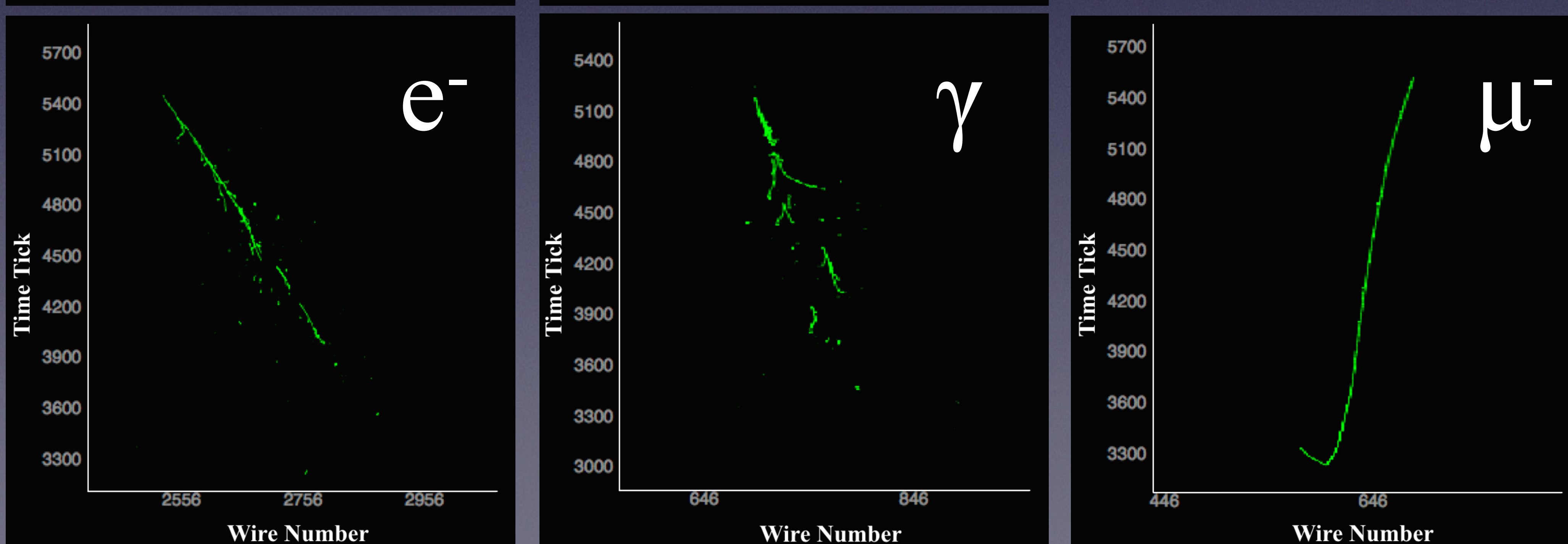
CNN in UB: Image Classification (I)

Particle identification

Trained a network to distinguish 5 particle types



- **Simulated particles**
  - using 1 (collection) plane
- **Supervised training**
  - 22,000 images / type
- **Flat momentum dist.**
  - Uniform position
  - Isotropic [100, 1000] MeV/c



CNN in UB: Image Classification (I)

Particle identification

Trained a network to distinguish 5 particle types

Particle	Correct Fraction	Typical Mis-ID
e^-	0.778	γ … 0.20
γ	0.834	e^- … 0.15
μ^-	0.897	π^- … 0.054
π^-	0.710	μ^- … 0.226
proton	0.912	μ^- … 0.046

JINST 10.1088/1748-9221

Further improvement?

- ~5 to 10% improvement by exploring network architectures
  - network width, effective depth
- more improvement by combining 3 plane information

Resource Usage

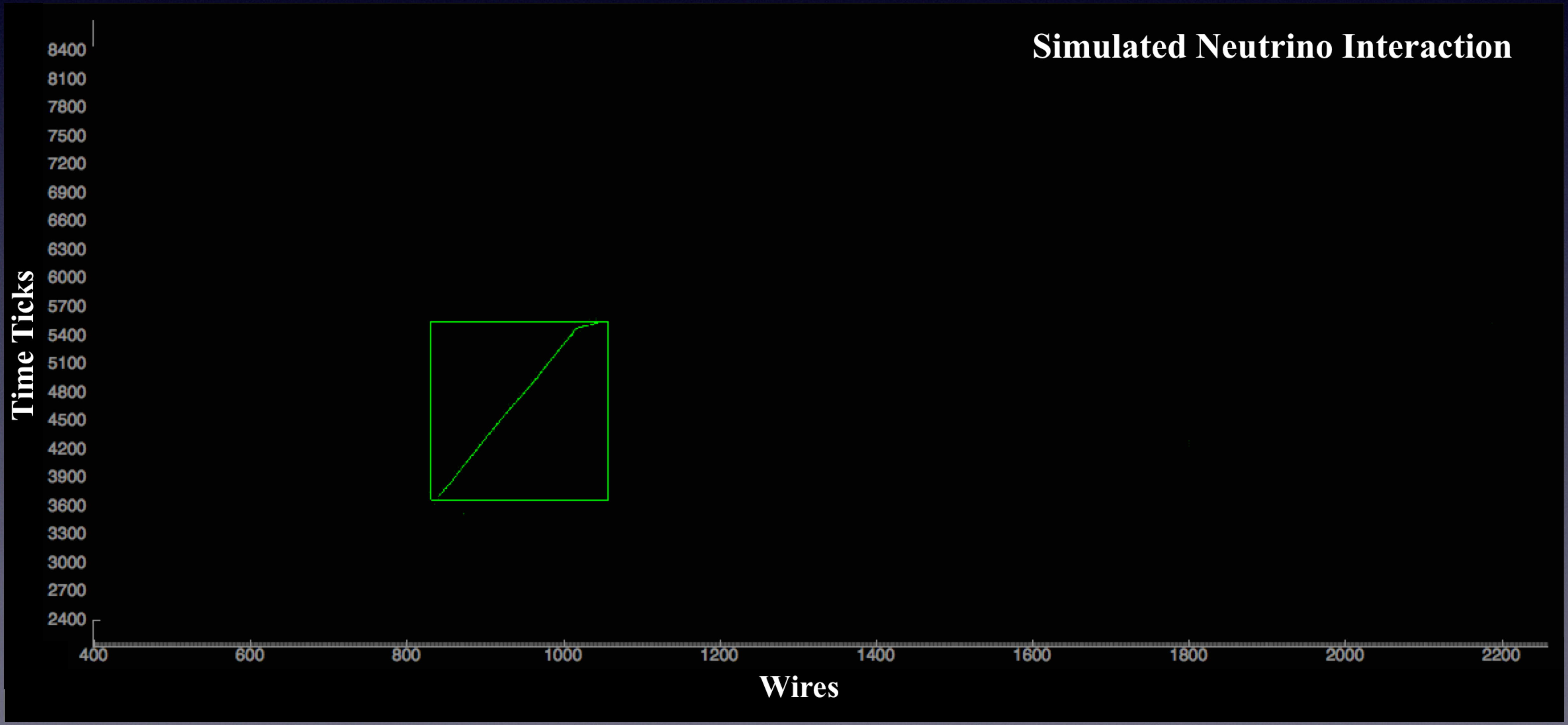
Architecture study include performance vs. speed!
Current architecture choice ~7 ms/image @ Titan X GPU)

CNN in UB: Image Classification (II)

Neutrino event selection

Distinguish neutrino+cosmic vs. cosmic-only events

- Training sample uses simulated neutrino + cosmic data image

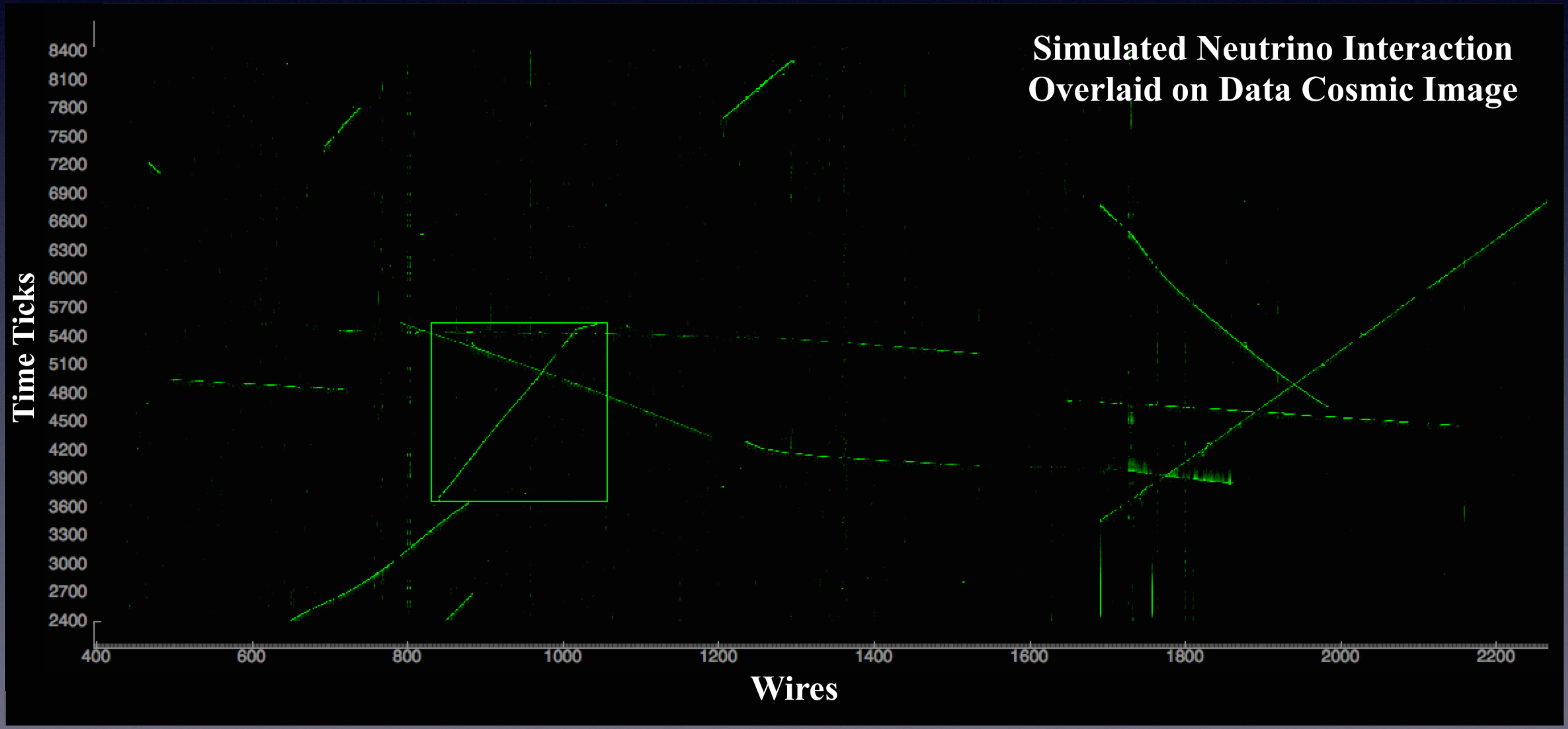


CNN in UB: Image Classification (II)

Neutrino event selection

Distinguish neutrino+cosmic vs. cosmic-only events

- Training sample uses simulated neutrino + cosmic data image

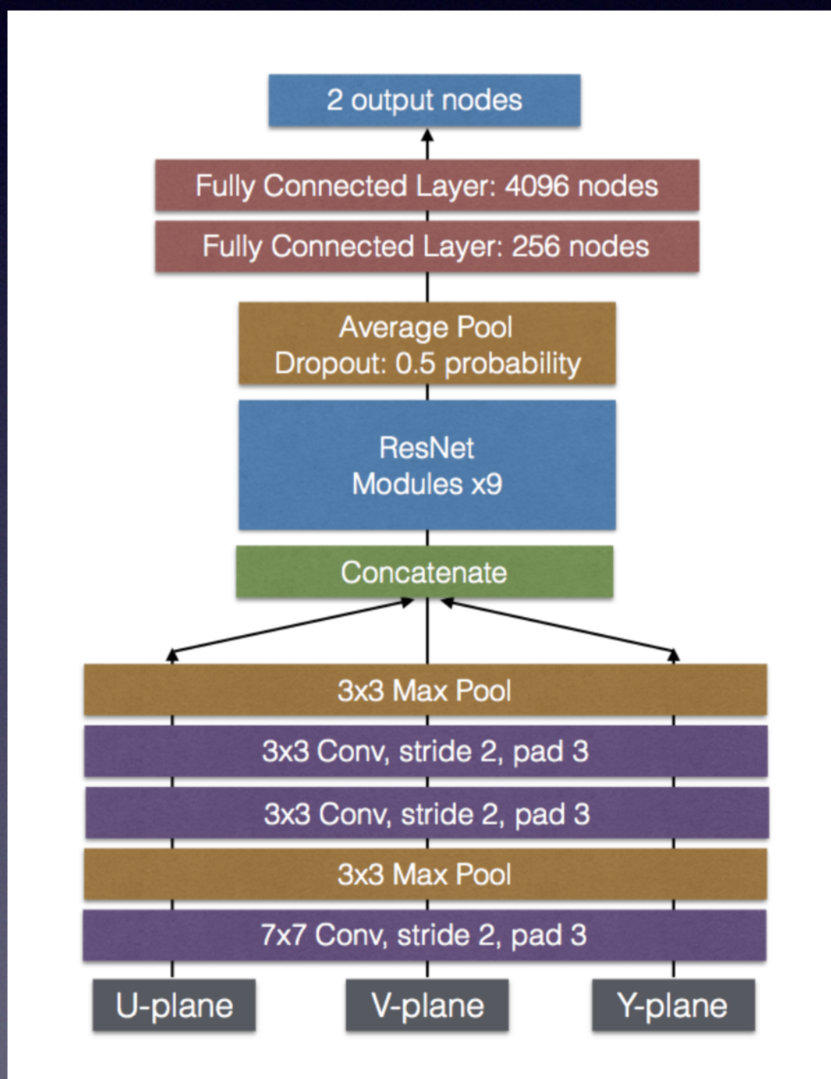


CNN in UB: Image Classification (II)

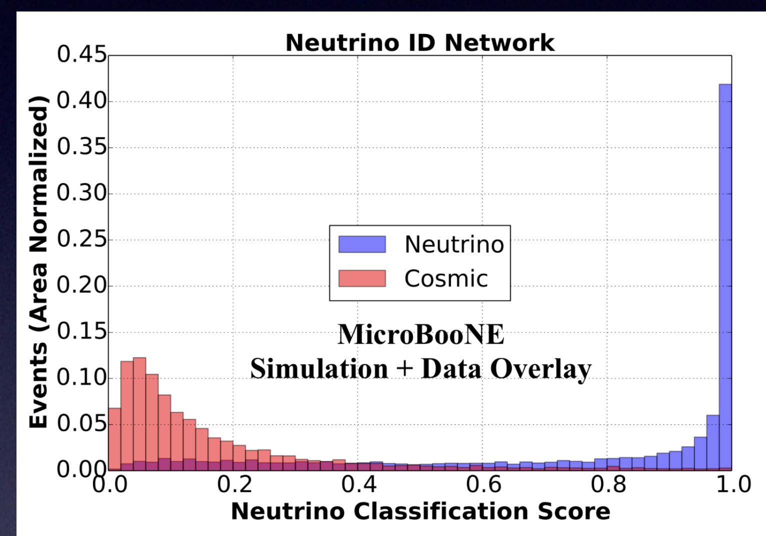
Neutrino event selection

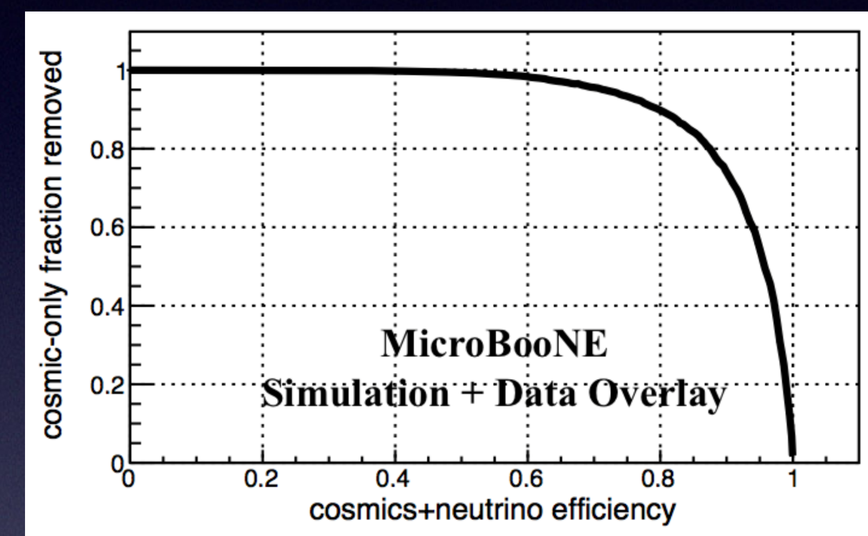
Distinguish neutrino+cosmic vs. cosmic-only events

- Training sample uses simulated neutrino + cosmic data image



Siamese Architecture for 3 plane analysis





Take aways

JINST 10.1088/1748-9221

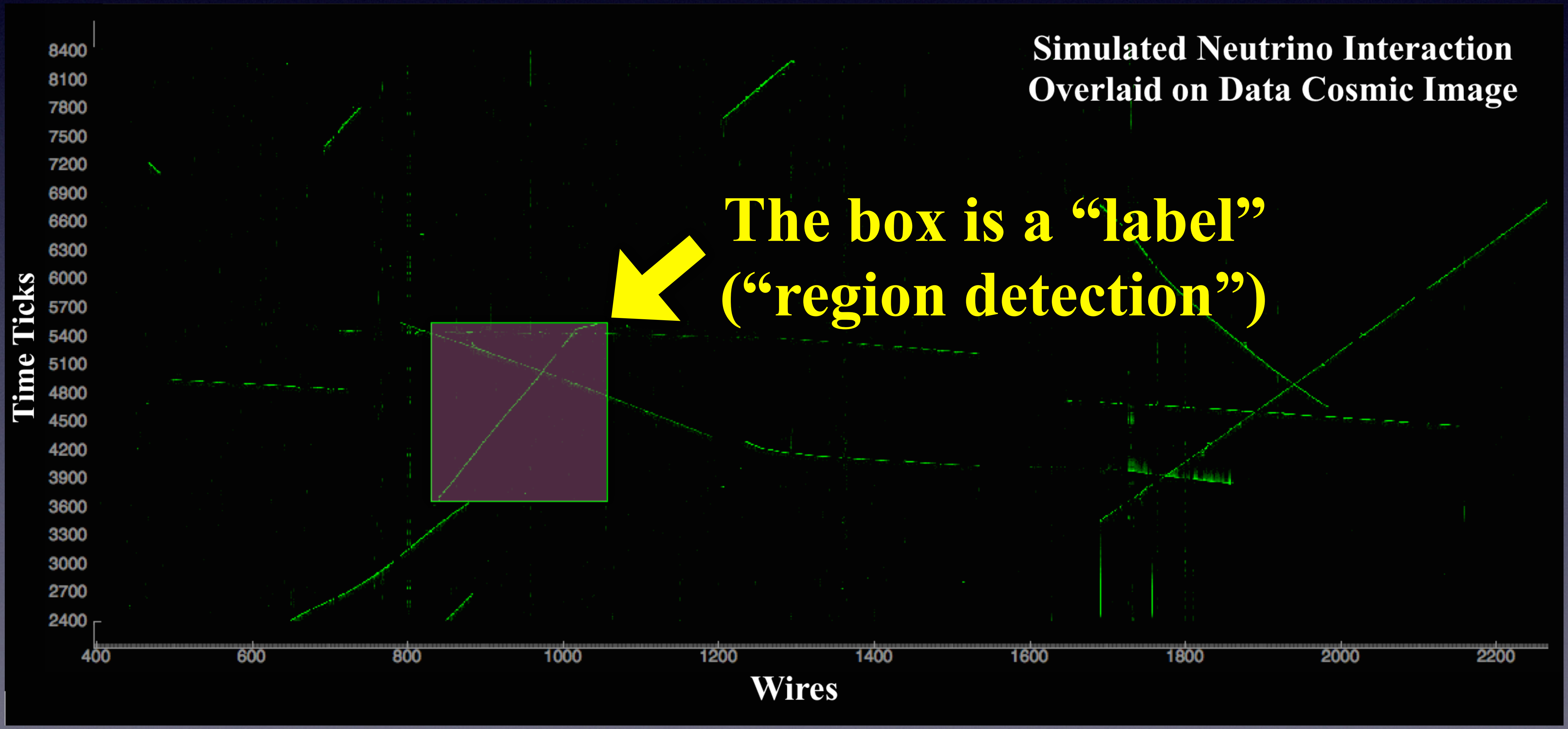
- Successfully combined 3 planes
- Poorer performance on real data
  - Tested with CC-inclusive selection sample from traditional reco
  - Importance to test/study with real data

CNN in UB: Object Detection

Event vertex detection

Trained a network to find neutrino interaction region

- Training sample uses simulated neutrino + cosmic data image
  - Supervised training using ≃100,000 collection plane images (1-plane)

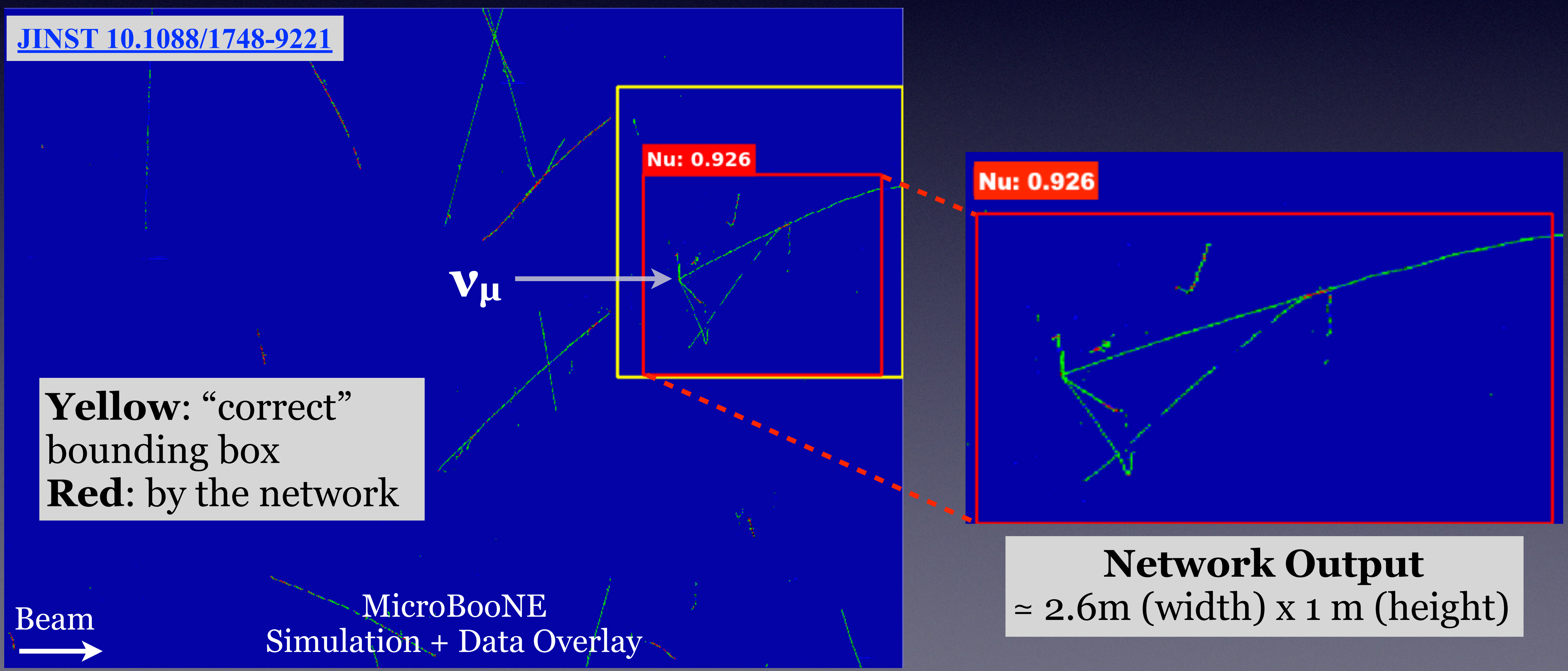


CNN in UB: Object Detection

Event vertex detection

Trained a network to find neutrino interaction region

- Training sample uses simulated neutrino + cosmic data image
  - Supervised training using ≃100,000 collection plane images (1-plane)

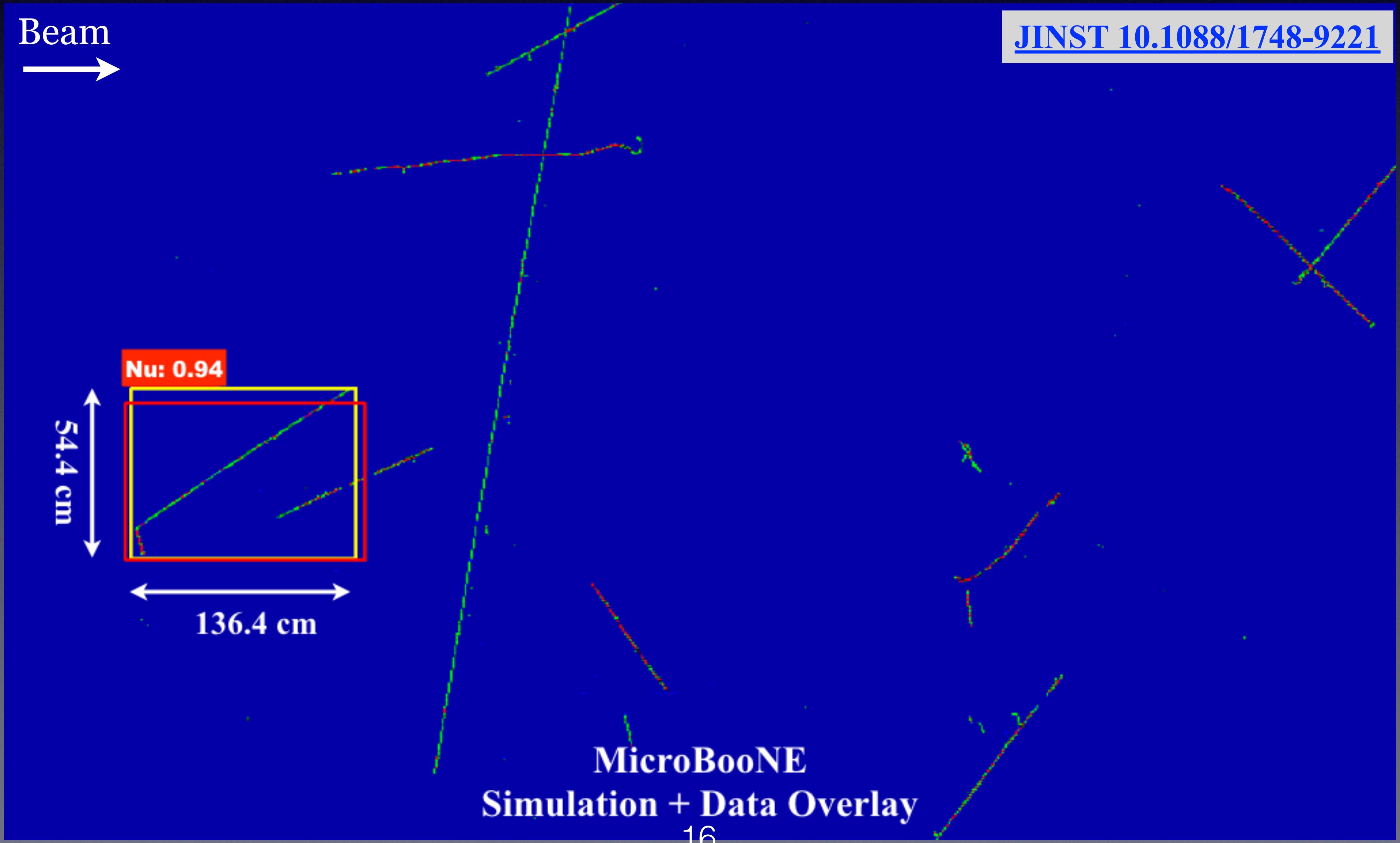


CNN in UB: Object Detection

Event vertex detection

Trained a network to find neutrino interaction region

- Training sample uses simulated neutrino + cosmic data image
  - Supervised training using ≃100,000 collection plane images (1-plane)

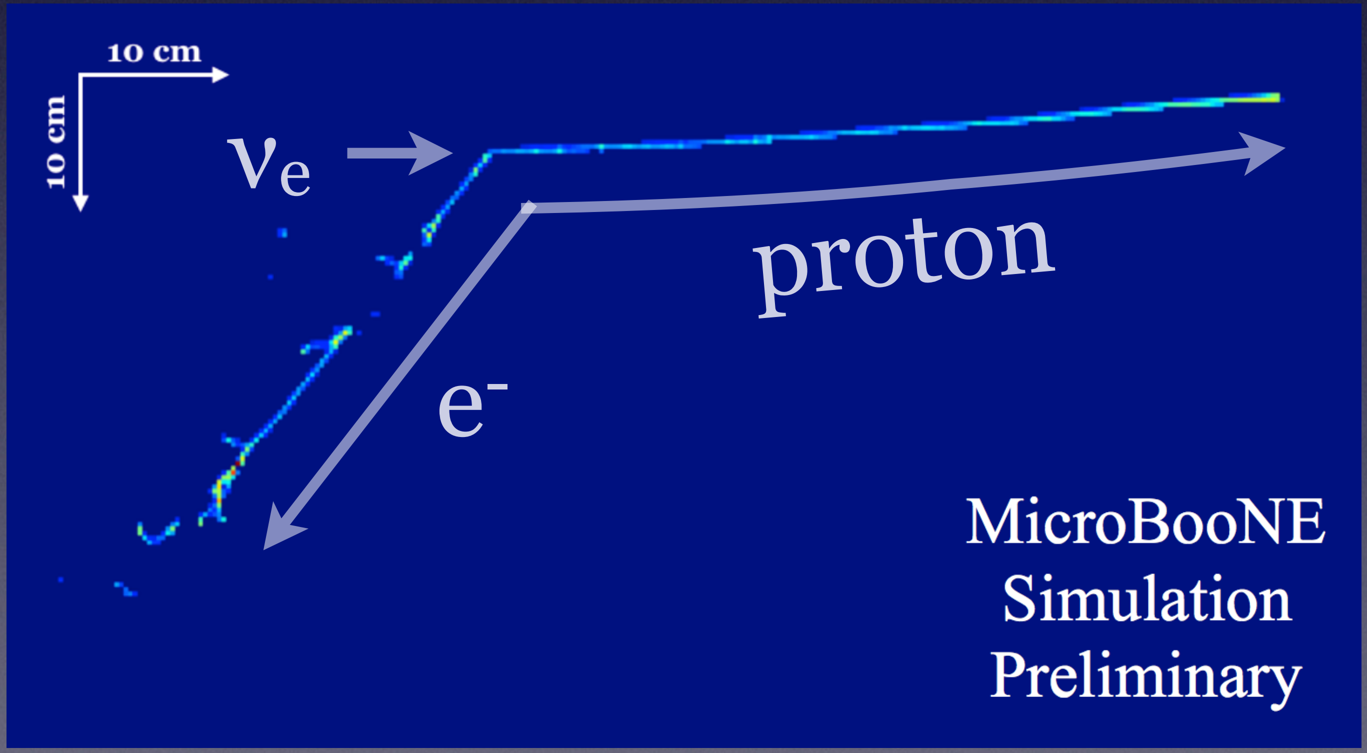


CNN in UB: Semantic Segmentation

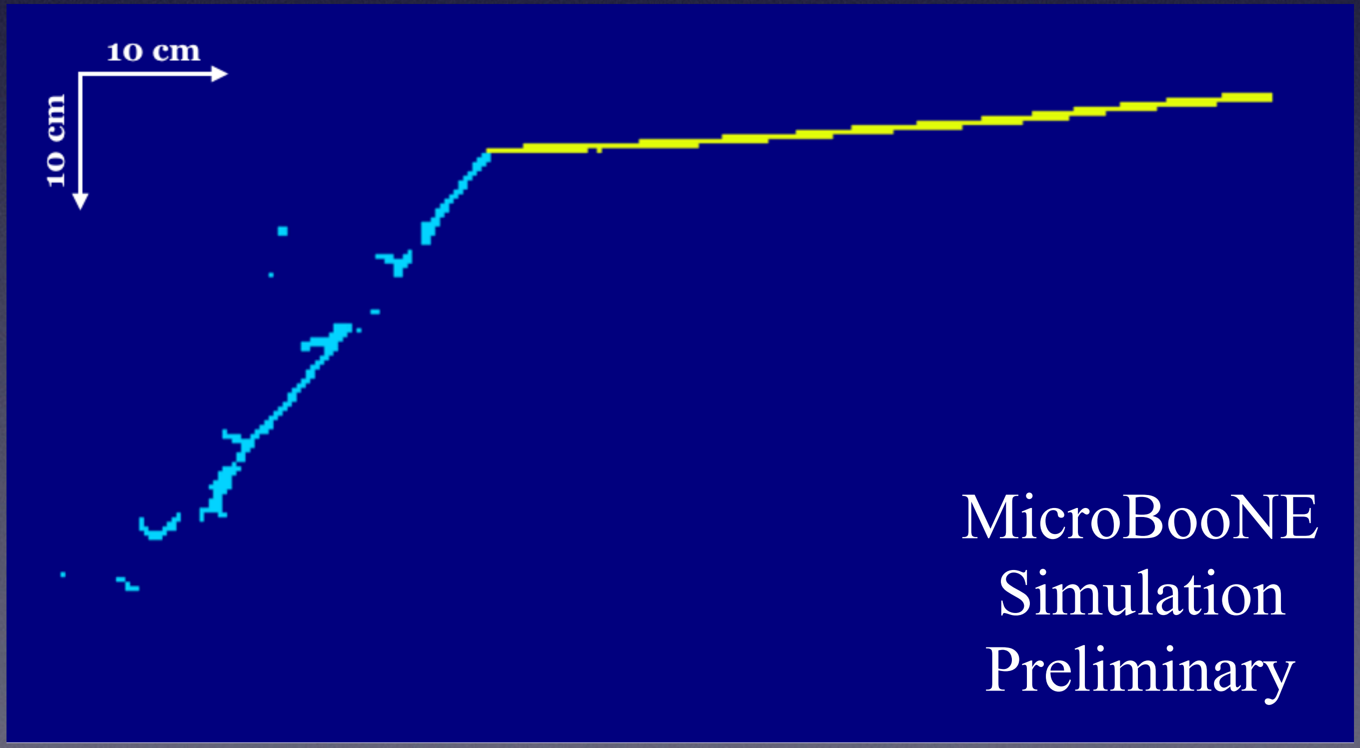
Particle clustering using a network

CNN designed to **segment pixels by predefined semantics**

- Current semantics: [background, shower, track]
- Supervised training on purely simulated images
  - **Custom training technique** to improve performance
  - On-going work: particle-wise pixel clustering



ADC Image



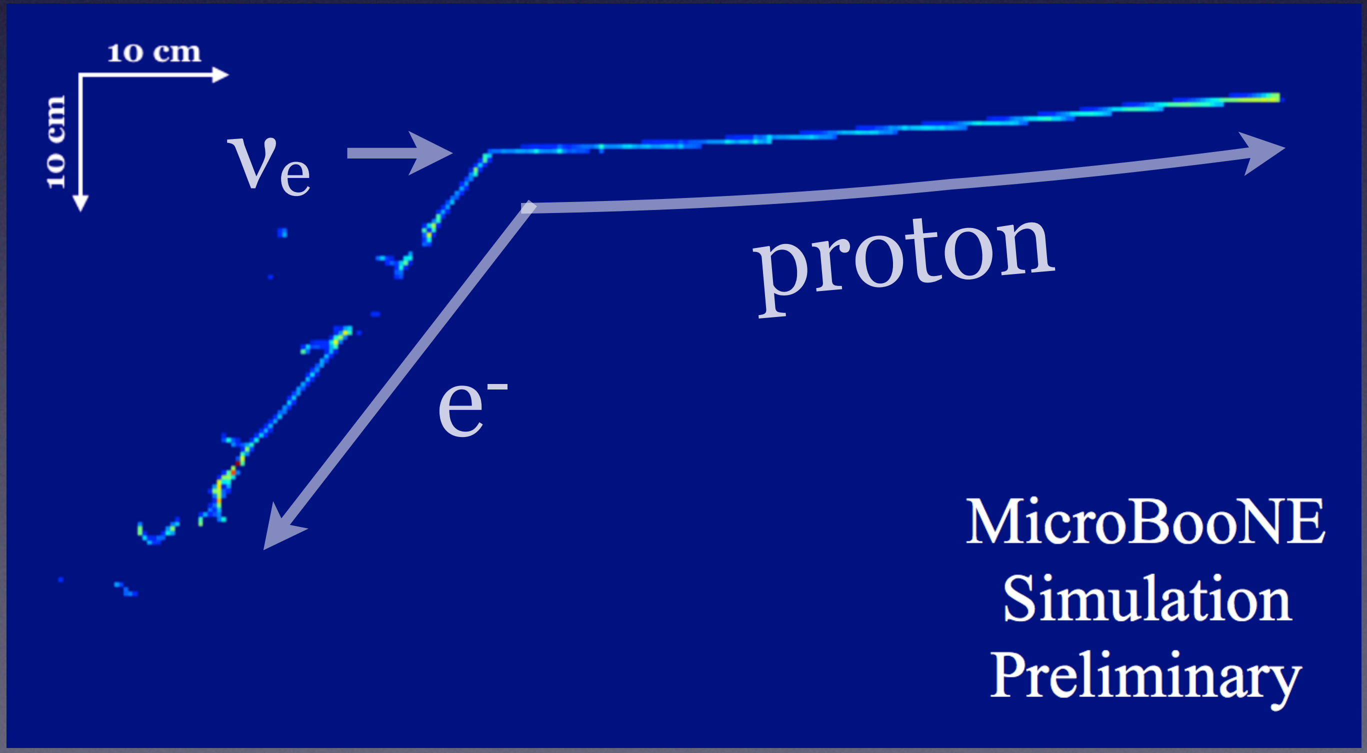
Label Image

CNN in UB: Semantic Segmentation

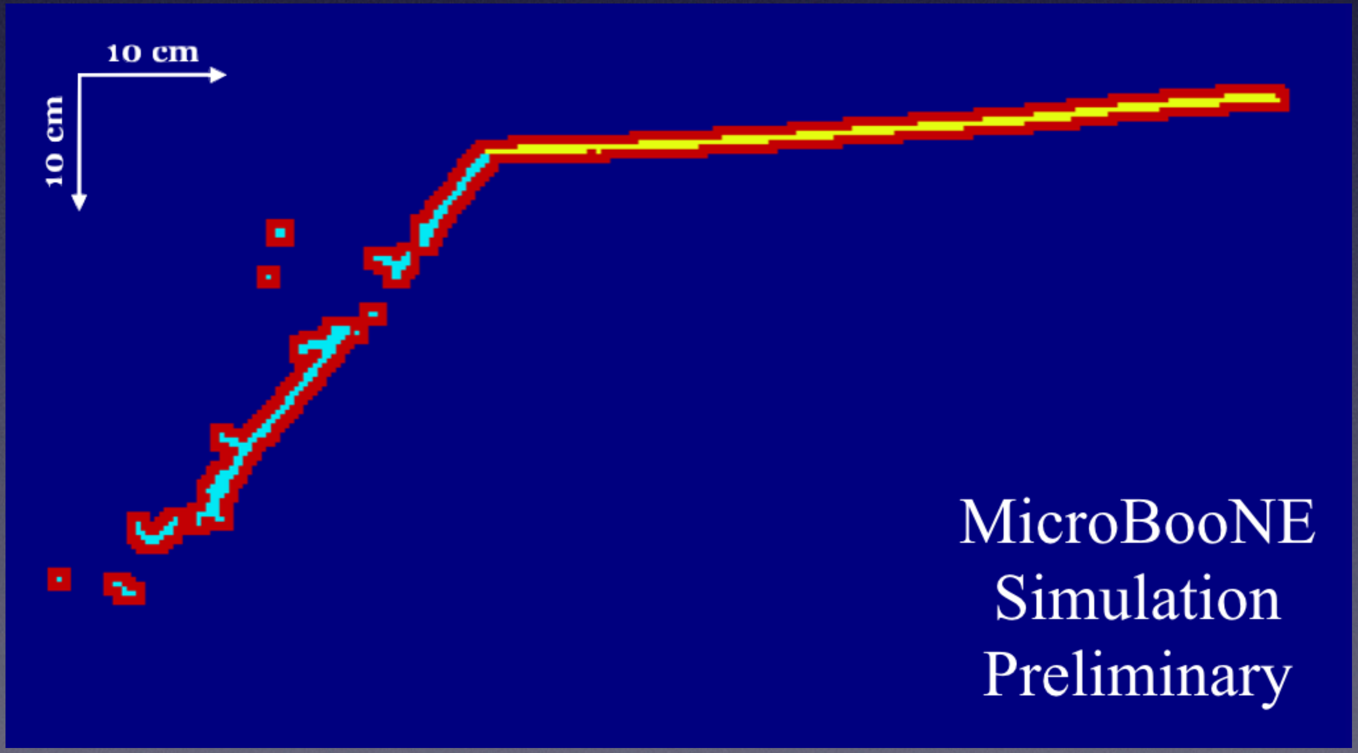
Particle clustering using a network

CNN designed to **segment pixels by predefined semantics**

- Current semantics: [background, shower, track]
- Supervised training on purely simulated images
  - **Custom training technique** to improve performance
  - On-going work: particle-wise pixel clustering



ADC Image



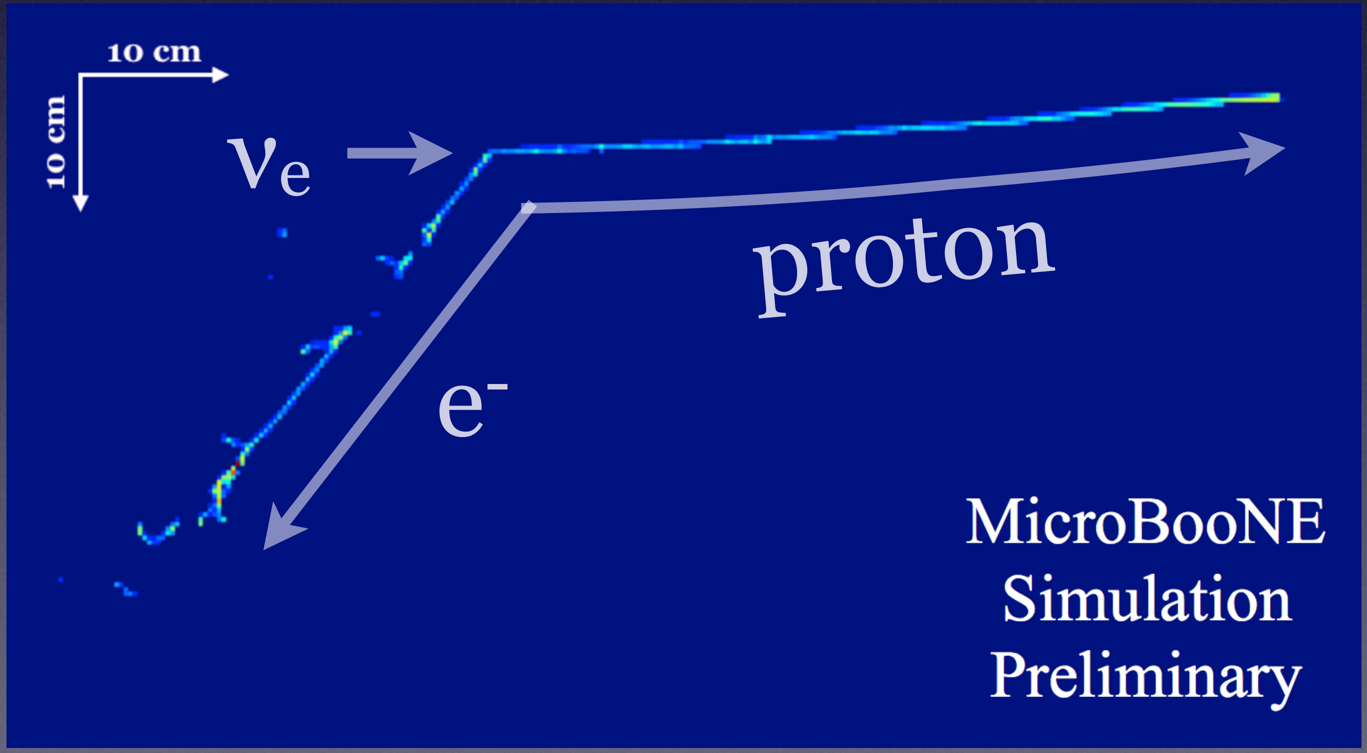
Error Weight Image

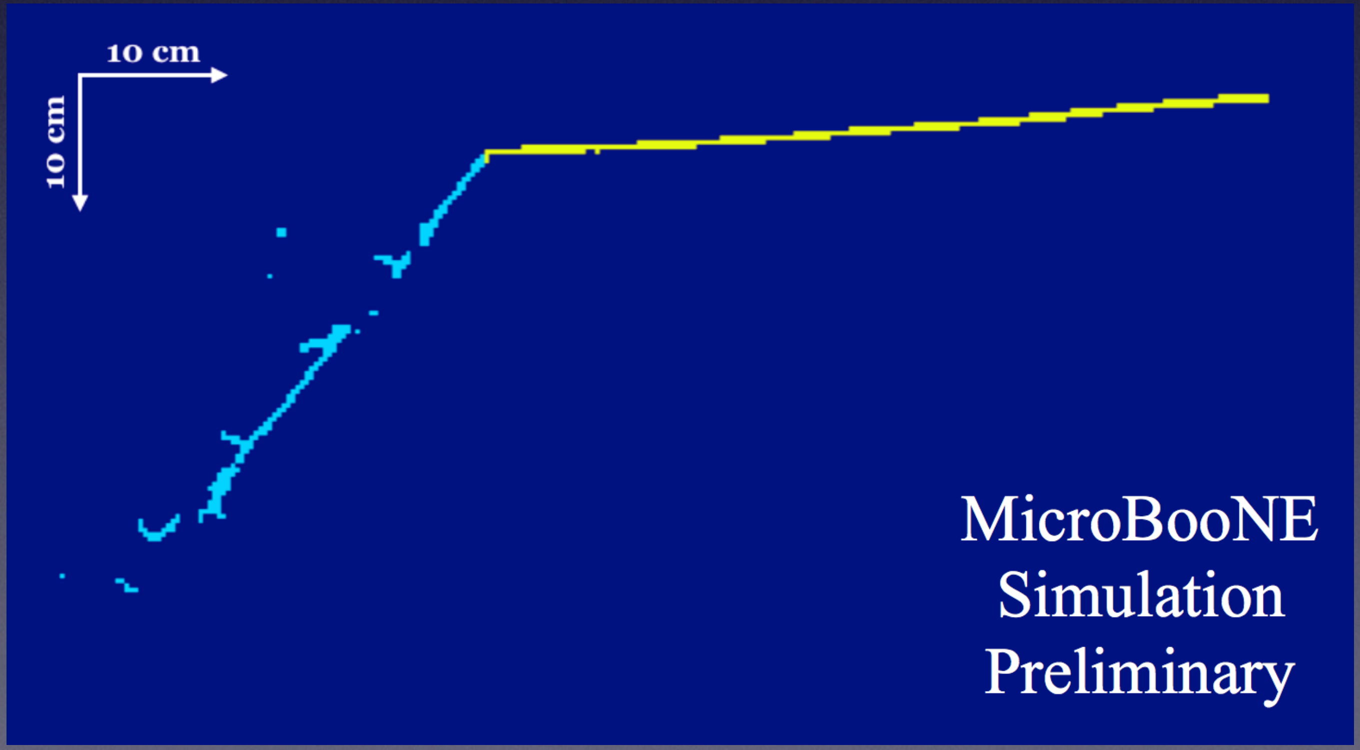
CNN in UB: Semantic Segmentation

Particle clustering using a network

CNN designed to **segment pixels by predefined semantics**

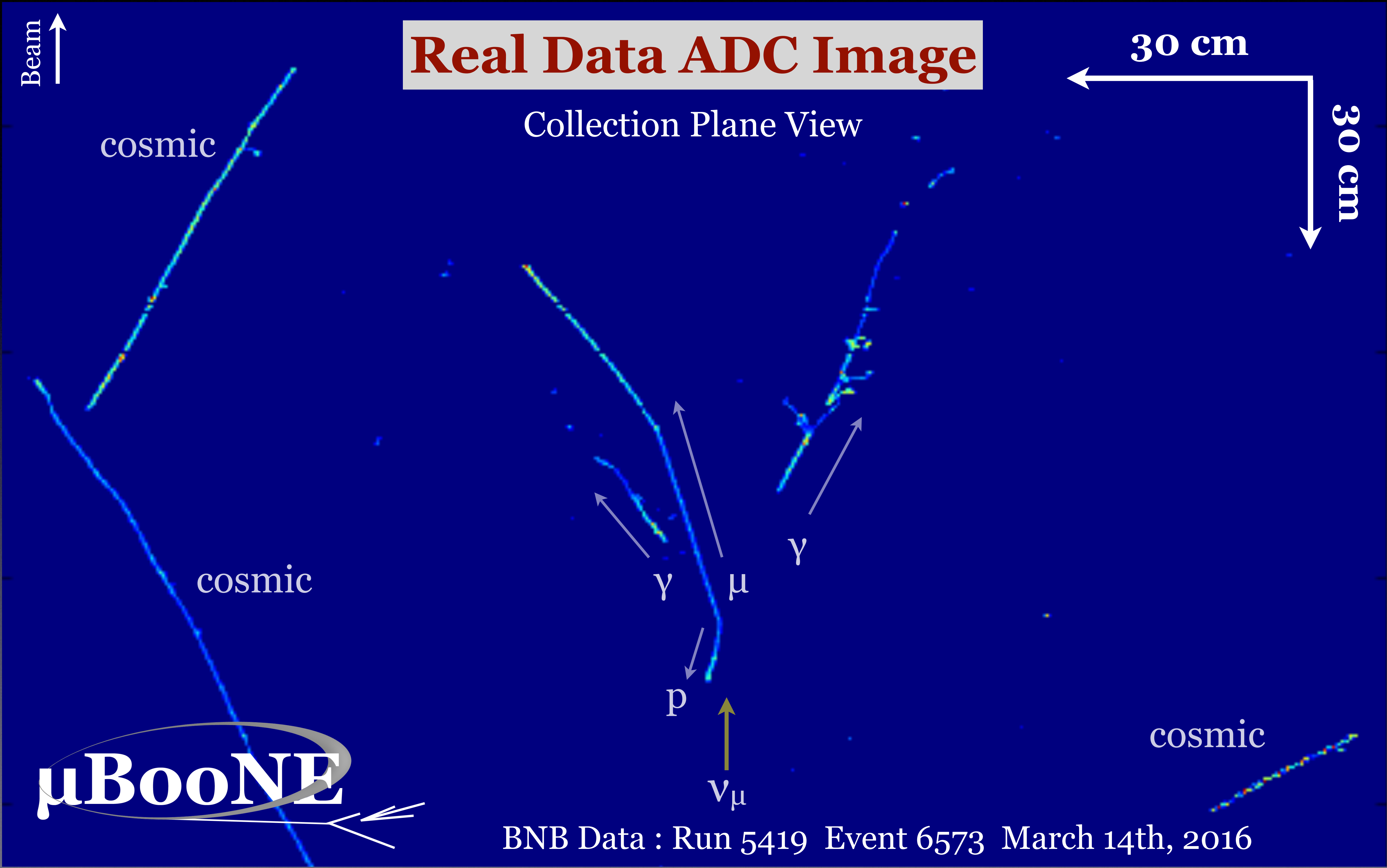
- Current semantics: [background, shower, track]
- Supervised training on purely simulated images
  - **Custom training technique** to improve performance
  - On-going work: particle-wise pixel clustering



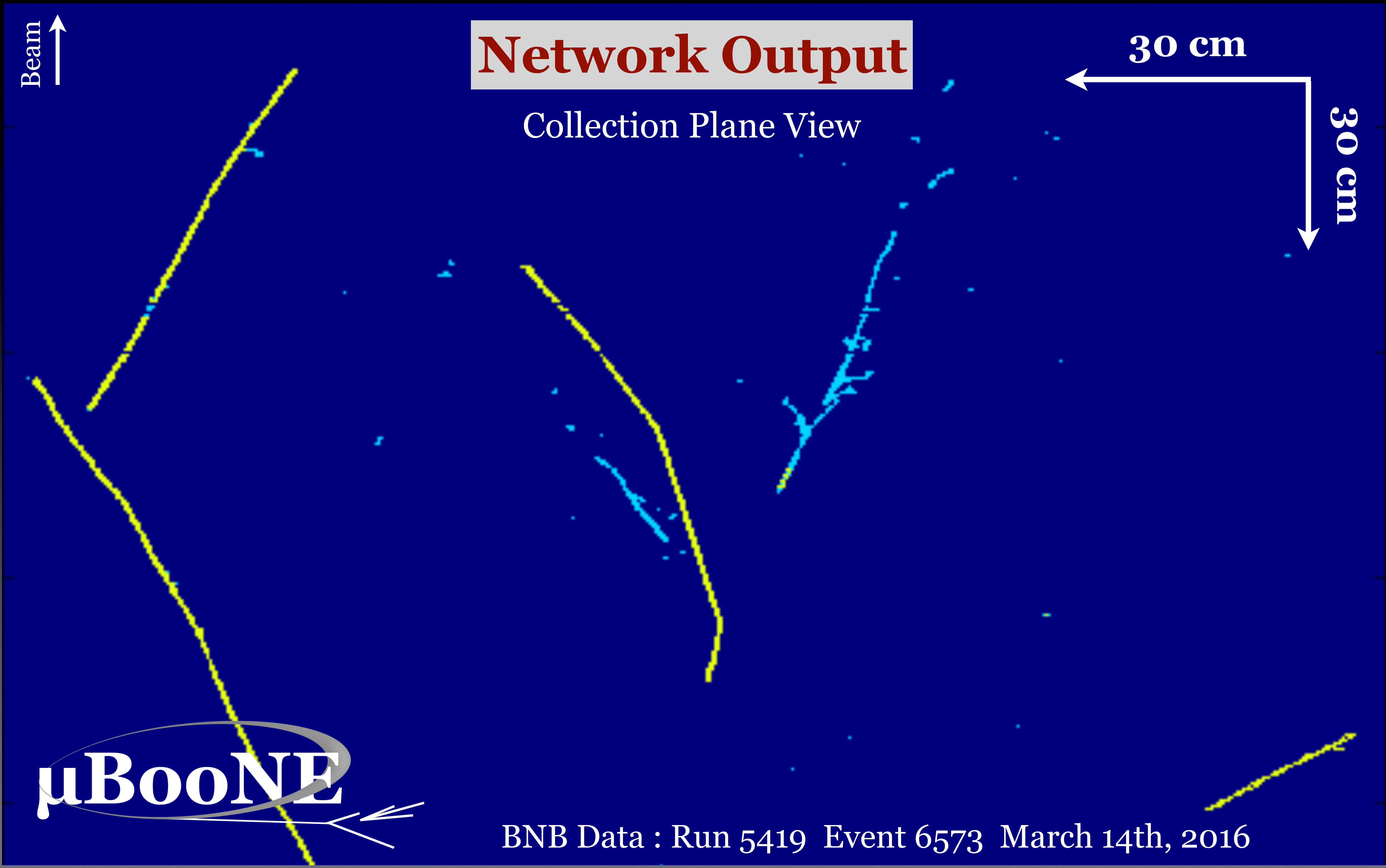


ADC Image | **Network Output**

CNN in UB: Semantic Segmentation



CNN in UB: Semantic Segmentation



End-to-End Reconstruction Training

Optimize multiple tasks together

“Multi-task Network Cascade” can introduce task dependencies

- Allows to optimize the whole chain together

… sorry for my parenthood …

CNN in MicroBooNE

Some studies published!

- Event selection (image classification)
- Vertex finding (object detection)
- Clustering (semantic segmentation)
- Particle identification (image classification)

MicroBooNE's 1st paper
JINST 10.1088/1748-9221
arXiv 1611.05531

Feel free to contact us for details!

Cornell University Library

arXiv.org > physics > arXiv:1611.05531

Physics > Instrumentation and Detectors

Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber

MicroBooNE collaboration: R. Acciarri, C. Adams, R. An, J. Asaadi, M. Auger, L. Bagby, B. Baller, G. Barr, M. Bass, F. Bay, M. Bishai, A. Blake, T. Bolton, L. Bugel, L. Camilleri, D. Caratelli, B. Carls, R. Castillo Fernandez, F. Cavanna, H. Chen, E. Church, D. Cianci, G. H. Collin, J. M. Conrad, M. Convery, J. I. Crespo-Anadón, M. Del Tutto, D. Devitt, S. Dytman, B. Eberly, A. Ereditato, L. Escudero Sanchez, J. Esquivel, B. T. Fleming, W. Foreman, A. P. Furmanski, G. T. Garvey, V. Genty, D. Goeldi, S. Gollapinni, N. Graf, E. Gramellini, H. Greenlee, R. Grosso, R. Guenette, A. Hackenburg, P. Hamilton, O. Hen, J. Hewes, C. Hill, J. Ho, G. Horton-Smith, C. James, J. Jan de Vries, C.-M. Jen, L. Jiang, R. A. Johnson, B. J. P. Jones, J. Joshi, H. Jostlein, D. Kaleko, G. Karagiorgi, W. Ketchum, et al. (75 additional authors not shown)

(Submitted on 17 Nov 2016)

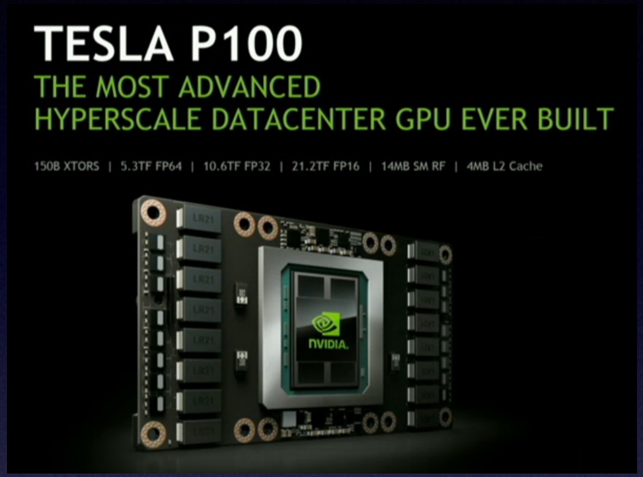
We present several studies of convolutional neural networks applied to data coming from the MicroBooNE detector, a liquid argon time projection chamber (LArTPC). The algorithms studied include the classification of single particle images, the localization of single particle and neutrino interactions in an image, and the detection of a simulated neutrino event overlaid with cosmic ray backgrounds taken from real detector data. These studies demonstrate the potential of convolutional neural networks for particle identification or event detection on simulated neutrino interactions. We also address technical issues that arise when applying this technique to data from a large LArTPC at or near ground level.

CNN in MicroBooNE

Hardware resource!

PIs responded our voice to **expand GPU resource for R&D**

- MIT, Columbia, Yale, UM Ann Arbor, PNNL
- Want more! (and more!) (and more!)



DL Interface Software

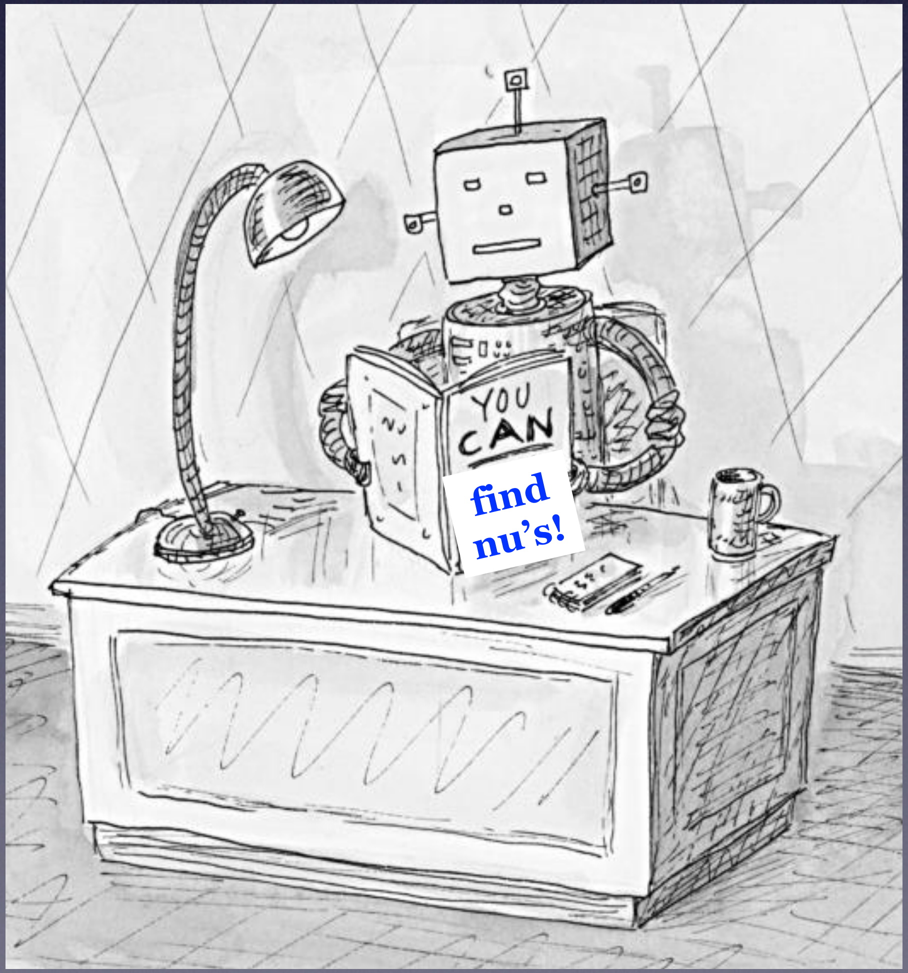
Generic i**mage processing software** (no need to be LArTPC)

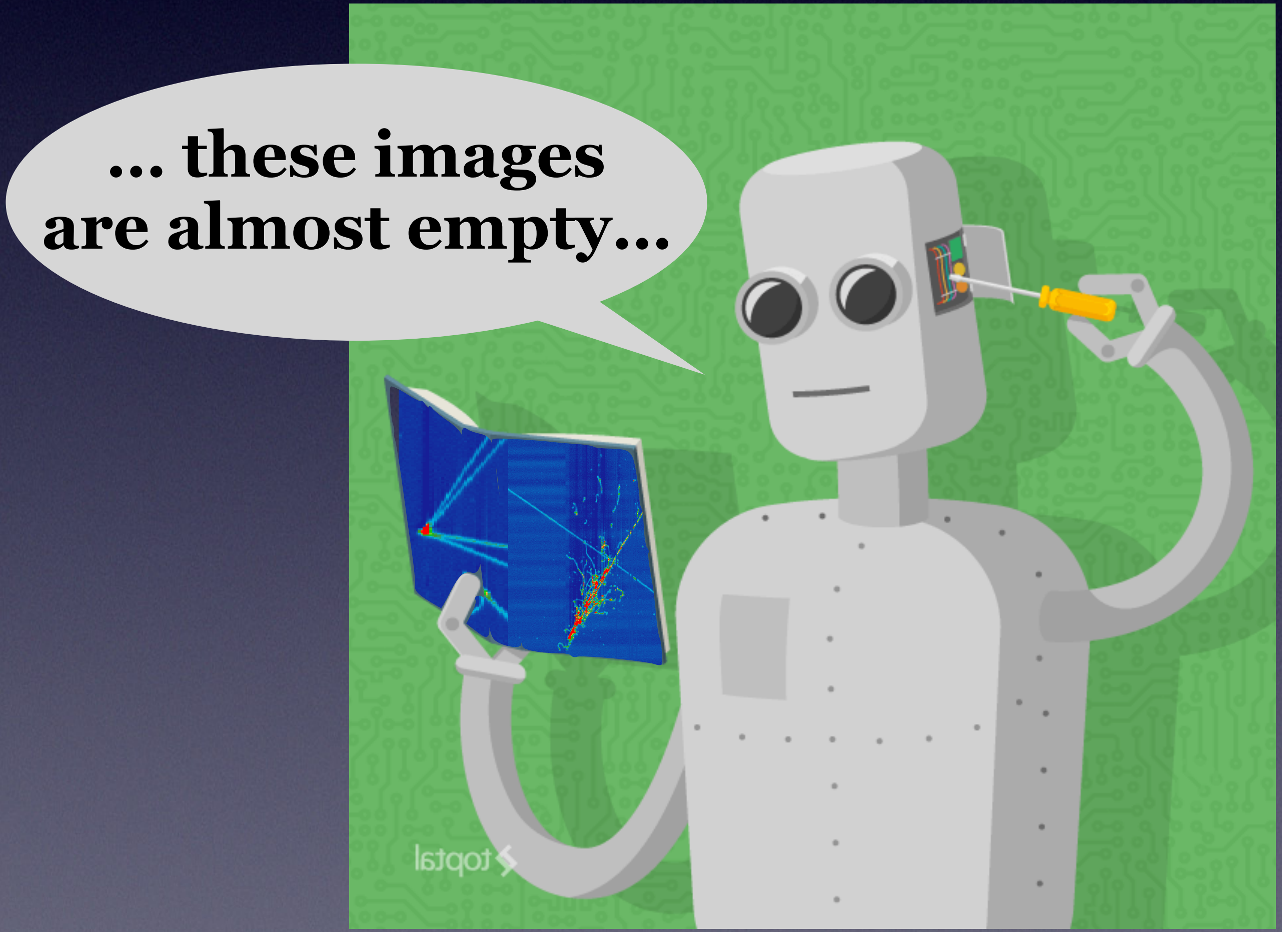
- Written in C++, extensive Python support
- Interface to C++/Python DL softwares (caffe, TensorFlow, etc.)
  - Fast, threaded IO to maximally utilize GPUs
  - Can bridge LArSoft (or any std::vector<float>) and DL software w/o file format conversion for running inference.

Lessons Learned

- CNNs can perform **reconstruction tasks**
  - Classification, object detection, and pixel clustering
- **CNNs are promising** techniques for LArTPC
  - Low information density: custom techniques can be helpful
- Important to **analyze response on real data**
  - Topological feature learning seems more immune
  - Building labeled image database from our data
  - Explore weak-supervision training & adversarial network
- **Initial challenges** = software & hardware (GPU)
  - Happy to advise on your GPU needs ($4k~)
  - Happy to share our software (public github)
  - Planning software workshops (please request!)

Lessons We're Learning
DeepLarning Projects
for
LArTPC Analysis





ν Reconstruction

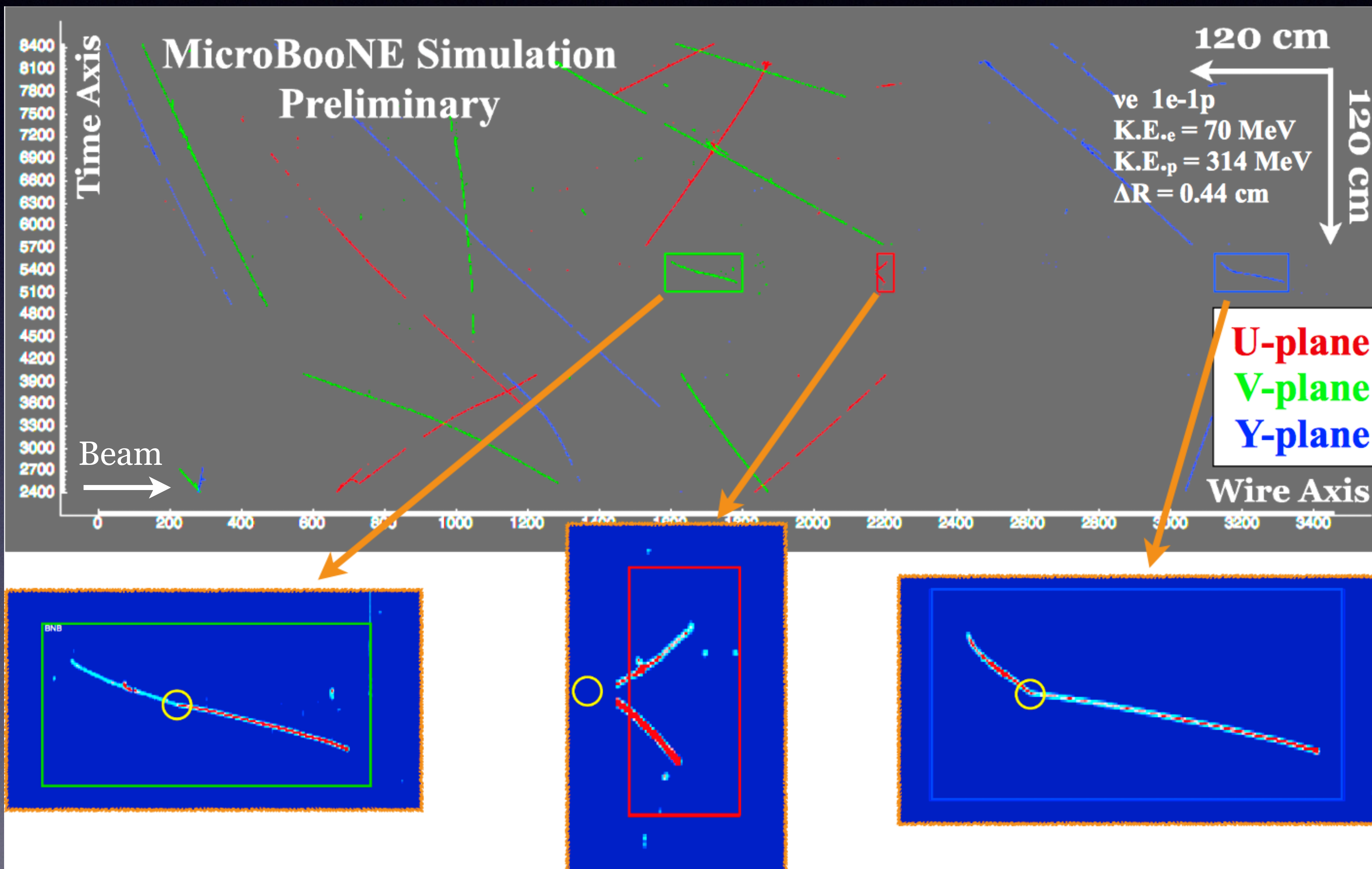

Taritree W.
MIT


Adrien H.
MIT


Jarret M.
MIT


Lauren Y.
MIT


Victor G.
Columbia U.




Rui A.
IIT


Christopher B.
U. Michigan


Jessica E.
Syracuse U.

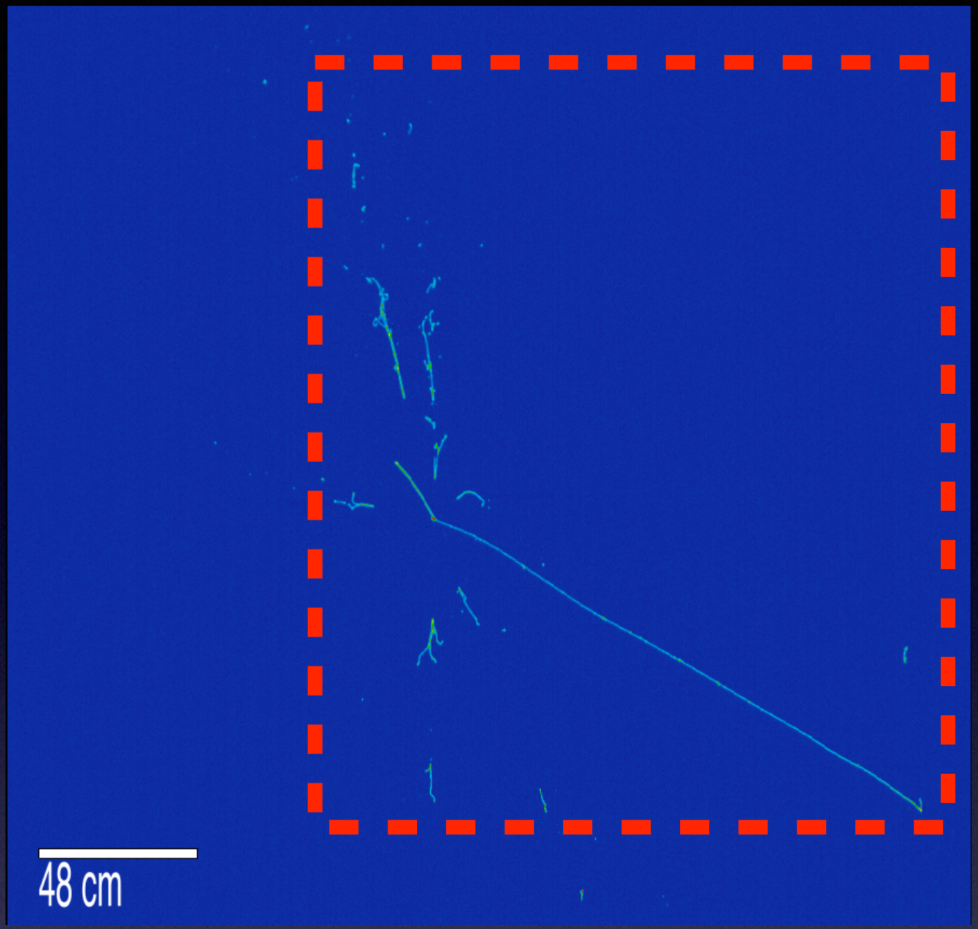
ν_e reconstruction (courtesy of Adrien Hourlier @ IPA)

n-$\bar{n}$ Oscillation

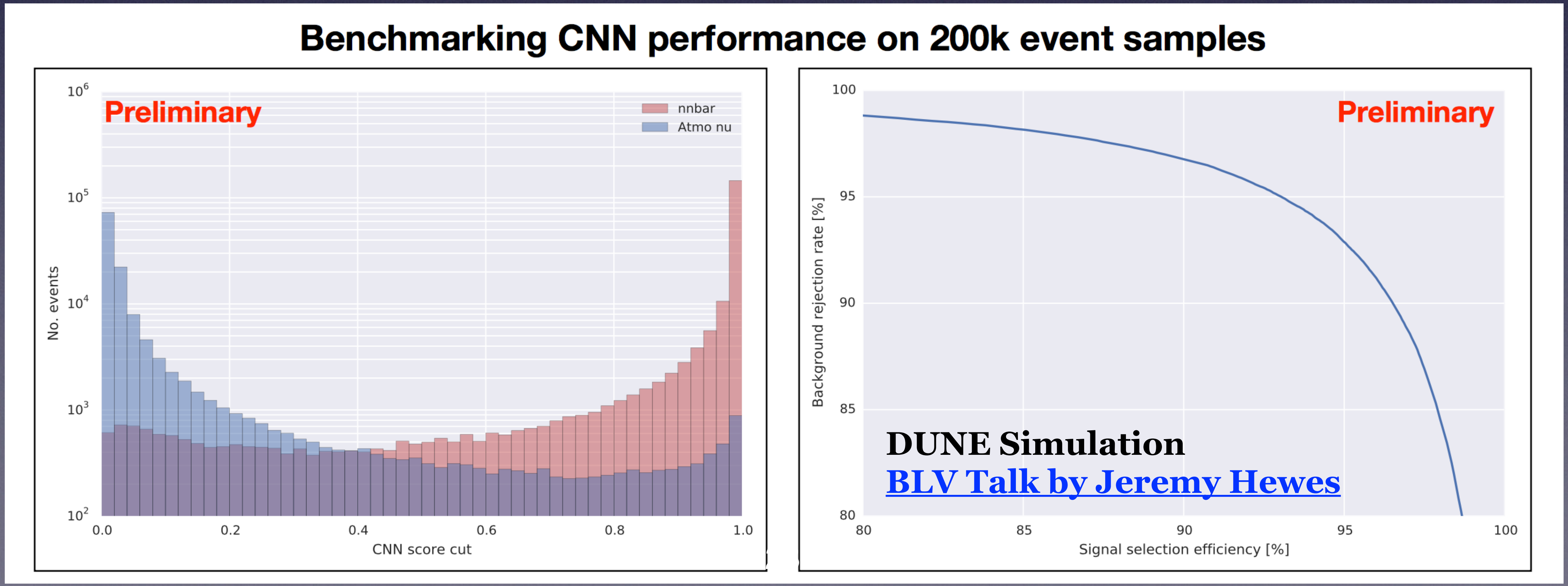
Jeremy Hewes
U. Manchester

Georgia Karagiorgi
Columbia U.

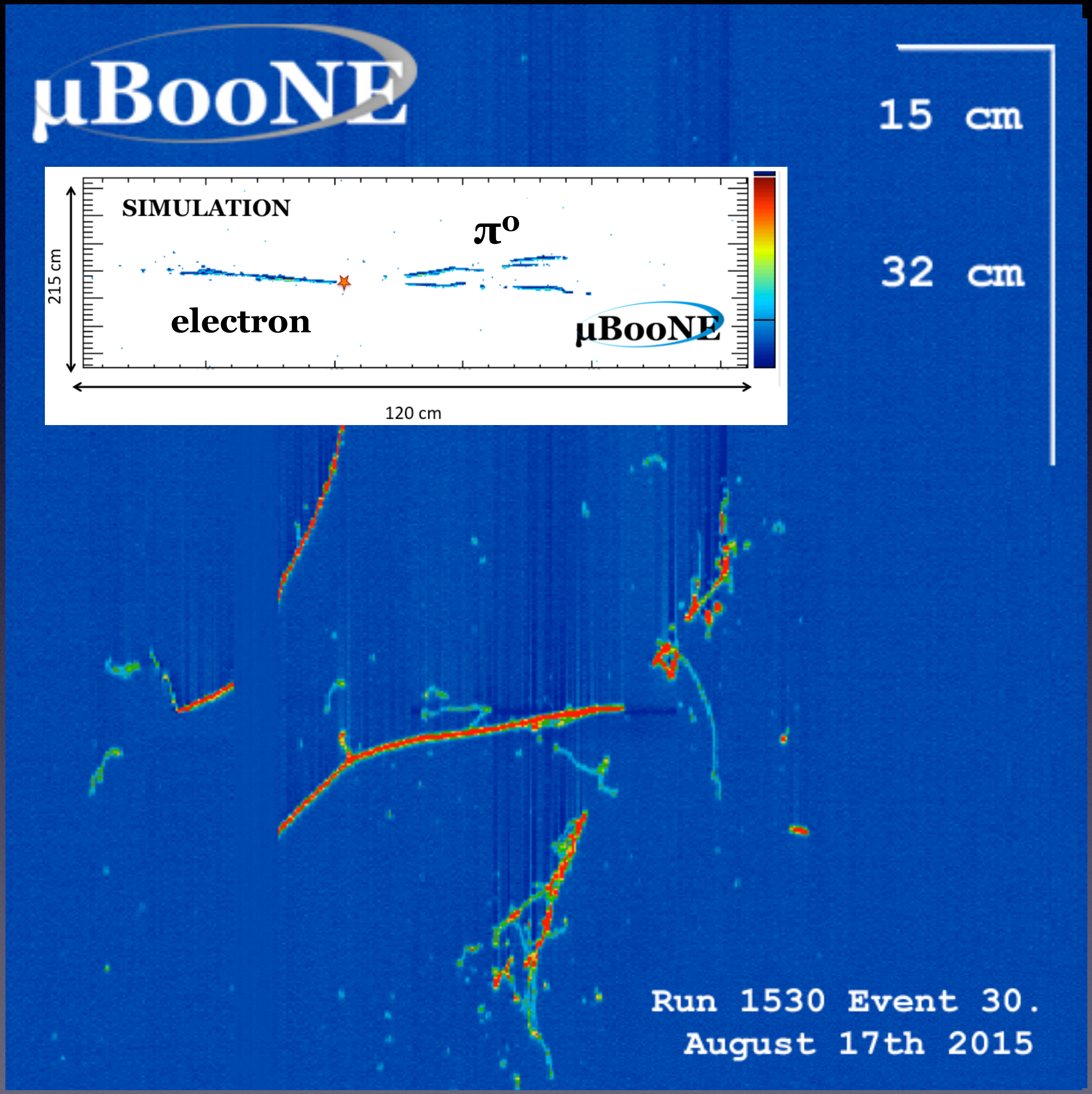
DUNE Simulation
BLV Talk by Jeremy Hewes



- New physics!
- Signal vs. background (atm. ν's)
- Developed from UB for DUNE
- Rich event topology, suited for CNN pattern recognition power



Proton Decay



PDK background study in UB
(Elena G. @ TAUP 2015)

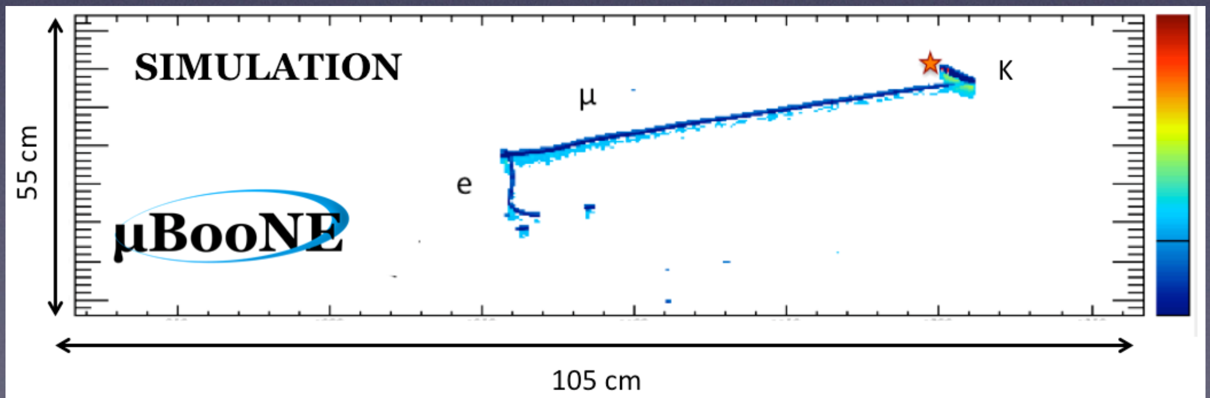


Elena Gramellini
Yale U.

Kevin Wierman
PNNL

Eric Church
PNNL

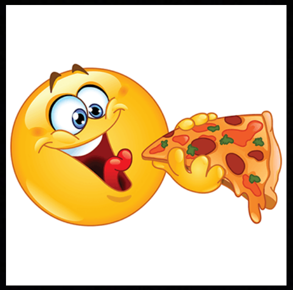
- New physics!
- Starting from UB work, real application @ DUNE
- Current focus on K^+/π^+ decay channel (PNNL)
- Topology classification



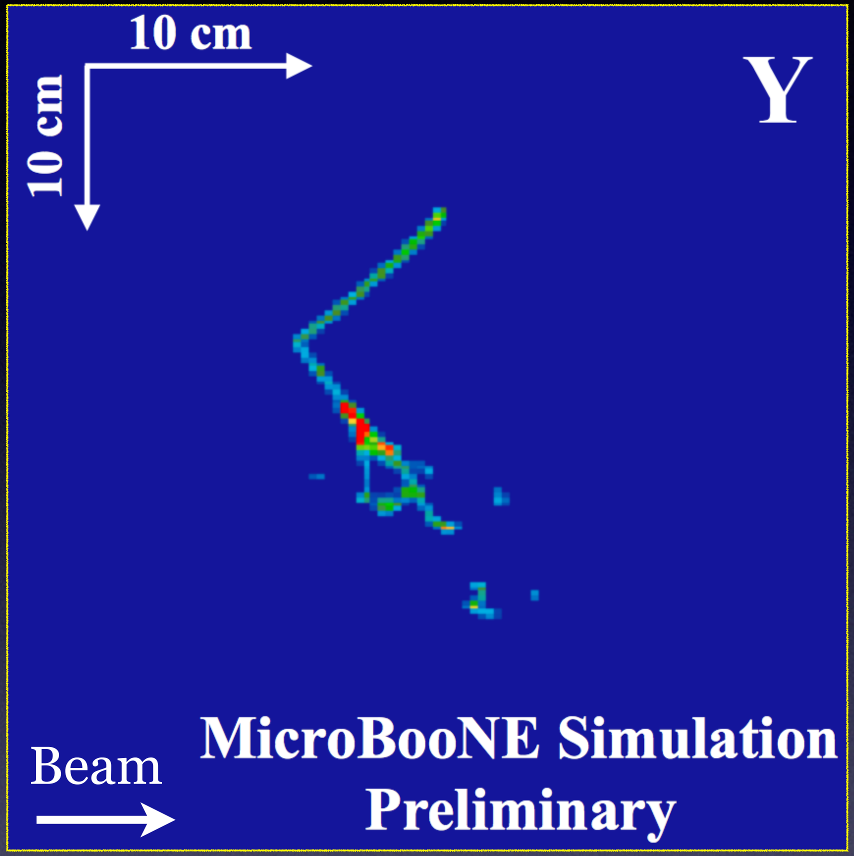
Example Kaon decay channel
(Kevin W./Eric C. @ PNNL)

PID for Neutrino Analysis

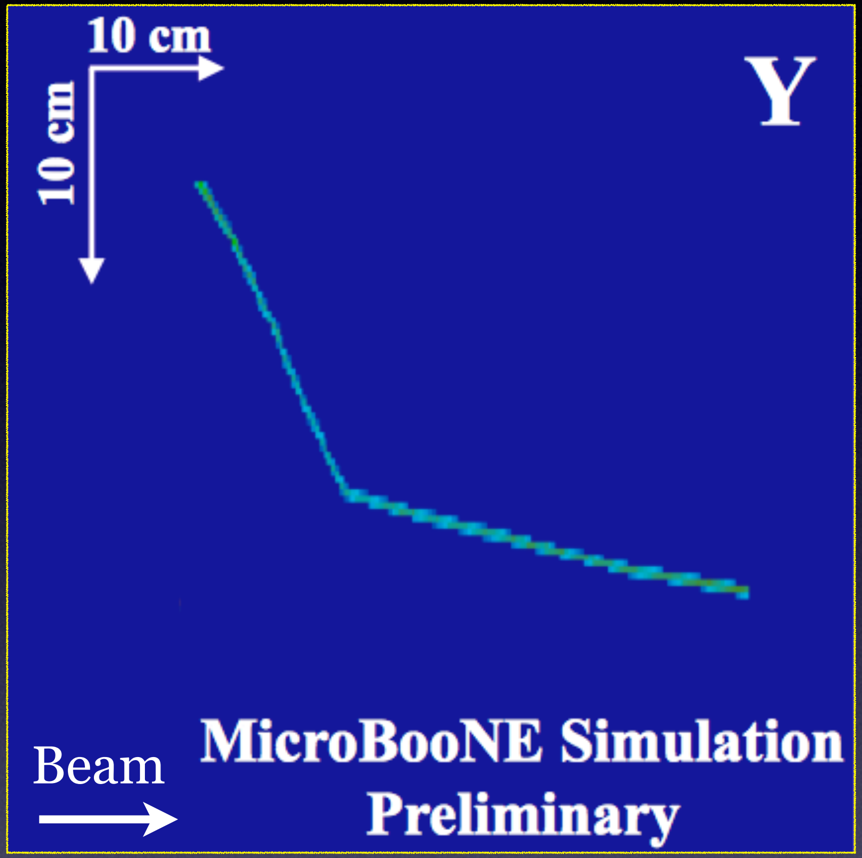
Summer Student A



Summer Student B



1 e^- & 1 proton



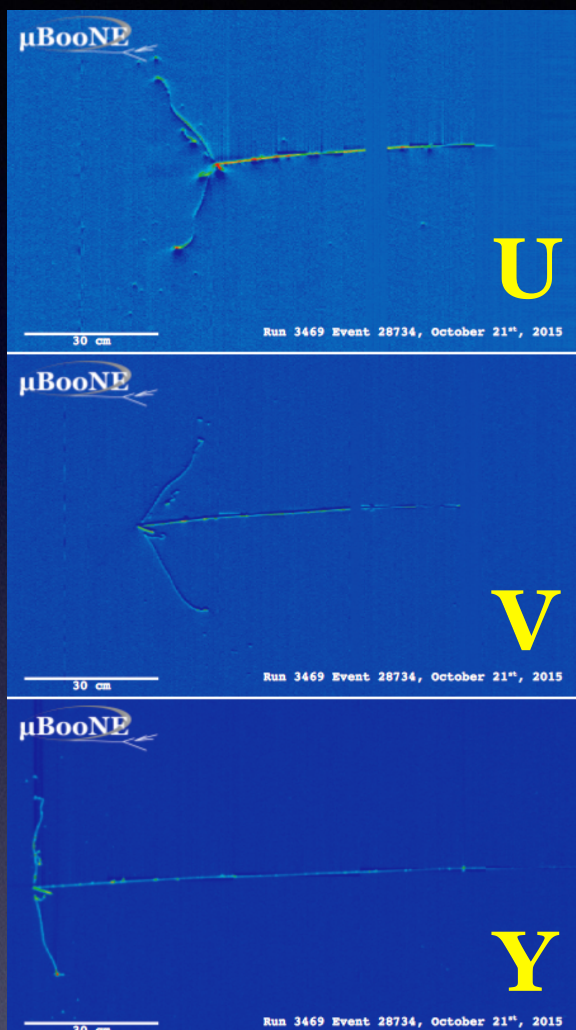
1 μ^- & 1 proton

- Predict multiplicity and types of particles involved
- Train on randomly generated multi-particle images
  - Avoid using an event generator with a certain model

3D Point Prediction

More Summer 2017 Projects!

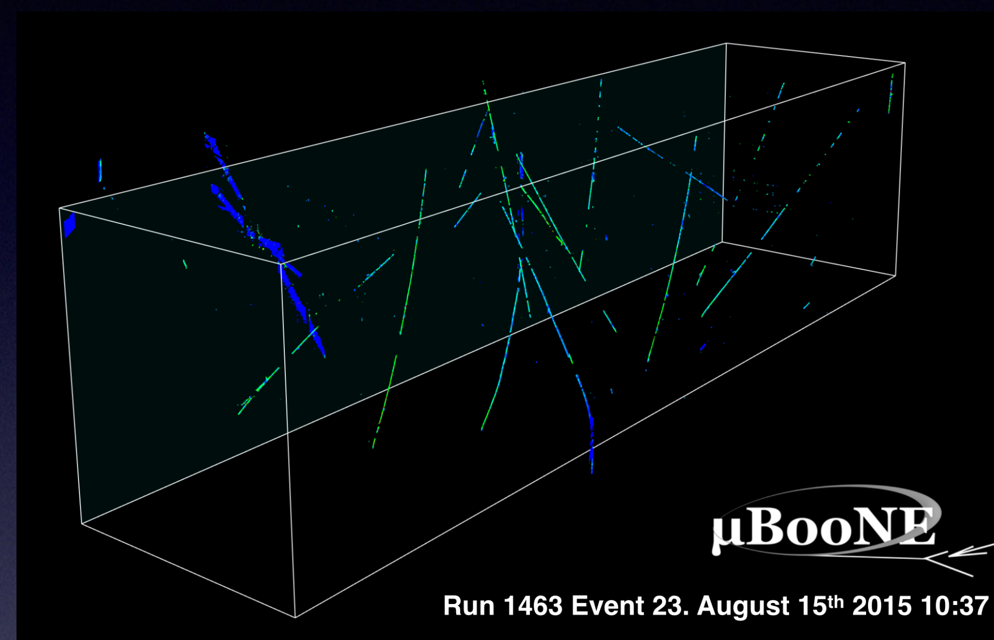
μBooNE



2D Views

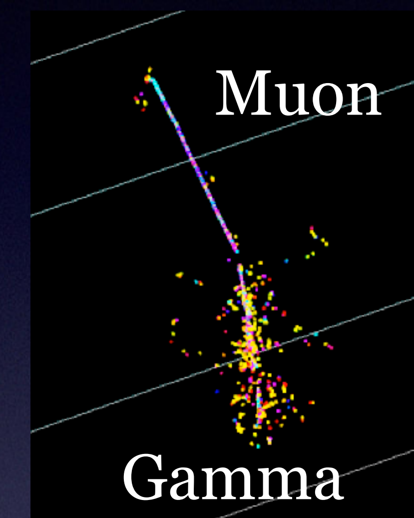
WireCell

Network



Reconstructed 3D View

Network



3D Clustering + PID

- Predict 3D point with charge from 2D plane views
- 3D feature recognition (3D point clustering + PID)

Lessons Learning

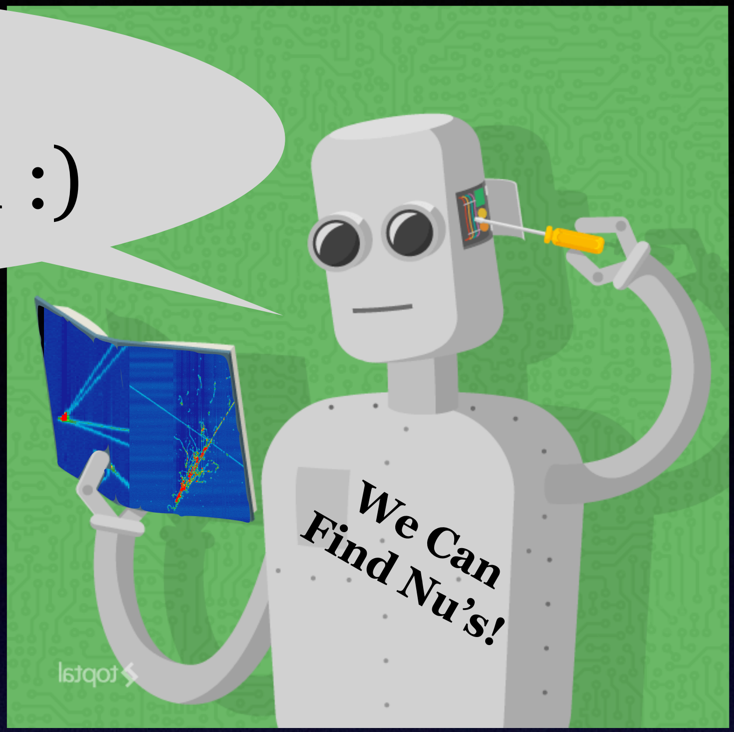
- **Fully CNN-based reconstruction**
  - similar to staged LArSoft reconstruction steps
  - allows stage-by-stage comparison
  - WireCell-like 3D reconstruction + analysis path
- Applying CNNs for **physics analysis**
  - neutrino analysis
  - rare event search: nn-bar oscillation, proton decay

… wrapping up …

Outline
- MicroBooNE and Deep Neural Networks
- Deep Learning “lessons learned”
- Deep Learning “lessons learning”
- **Summary**

I want Ph.D!
I want a job!

Thank you!
for your attention :)



Take Away Messages

1. **LArTPCs** need **advanced pattern recognition algorithms**

2. **MicroBooNE** develops **CNN-based reconstruction tools**

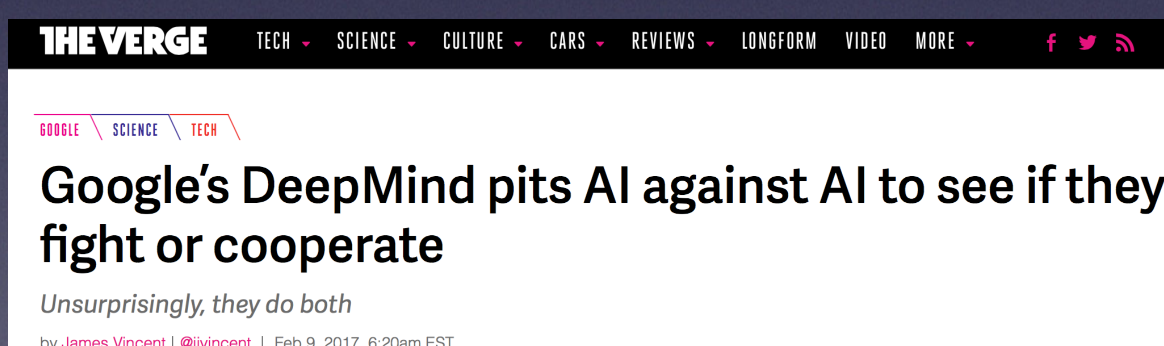
3. **MicroBooNE** applies **CNN techniques to physics analysis**

4. **MicroBooNE** studies **network response on real data**

5. **MicroBooNE** shares **tools developed and knowledge learnt**

Extracurricular Lessons Learned

Remember what happen"ed" with AI



THE VERGE TECH SCIENCE CULTURE CARS REVIEWS LONGFORM VIDEO MORE

GOOGLE \ SCIENCE \ TECH

Google's DeepMind pits AI against AI to see if they fight or cooperate

Unsurprisingly, they do both

by James Vincent | @jjvincent | Feb 9, 2017, 6:20am EST

"... the cleverer AI decided it was better to be aggressive in all situations."

We try to be careful :)

Back up

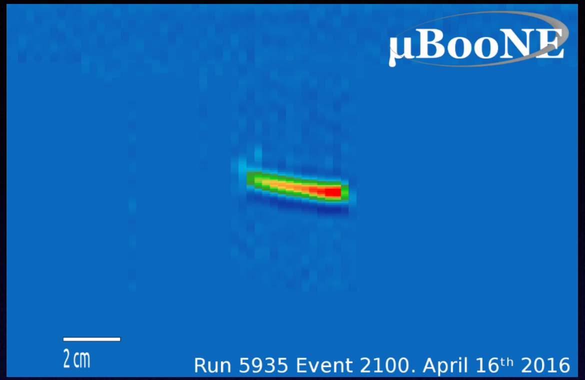
More Projects?

µBooNE

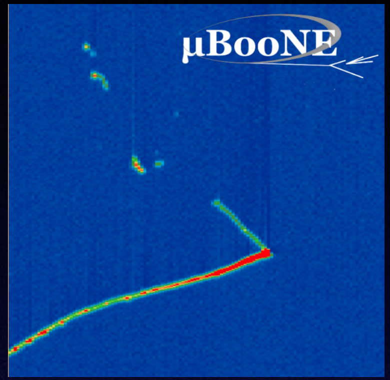
More Summer 2017 Projects!

Training on Data

- Labeled image database
  - Labeling software tools
  - "Chimera" image maker
- Weakly supervised training



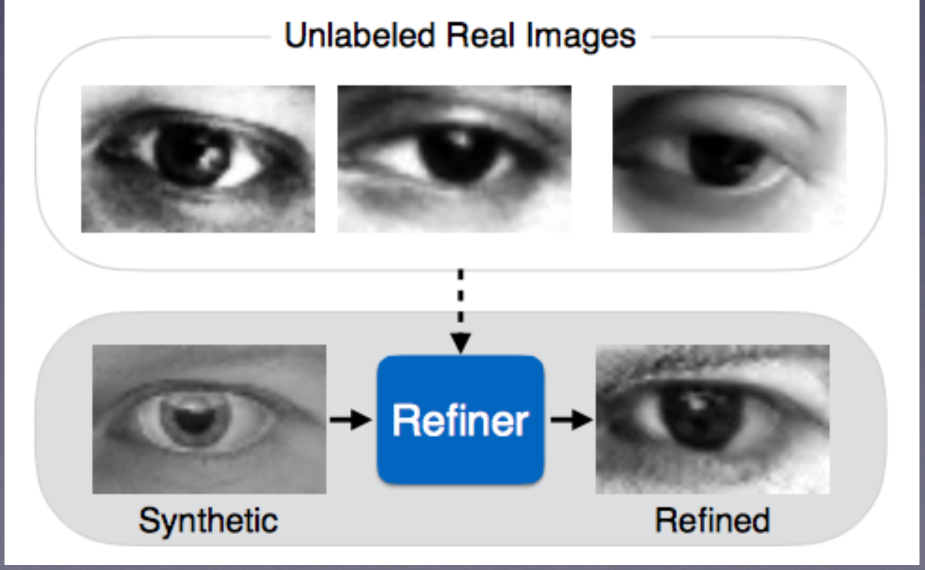
Proton



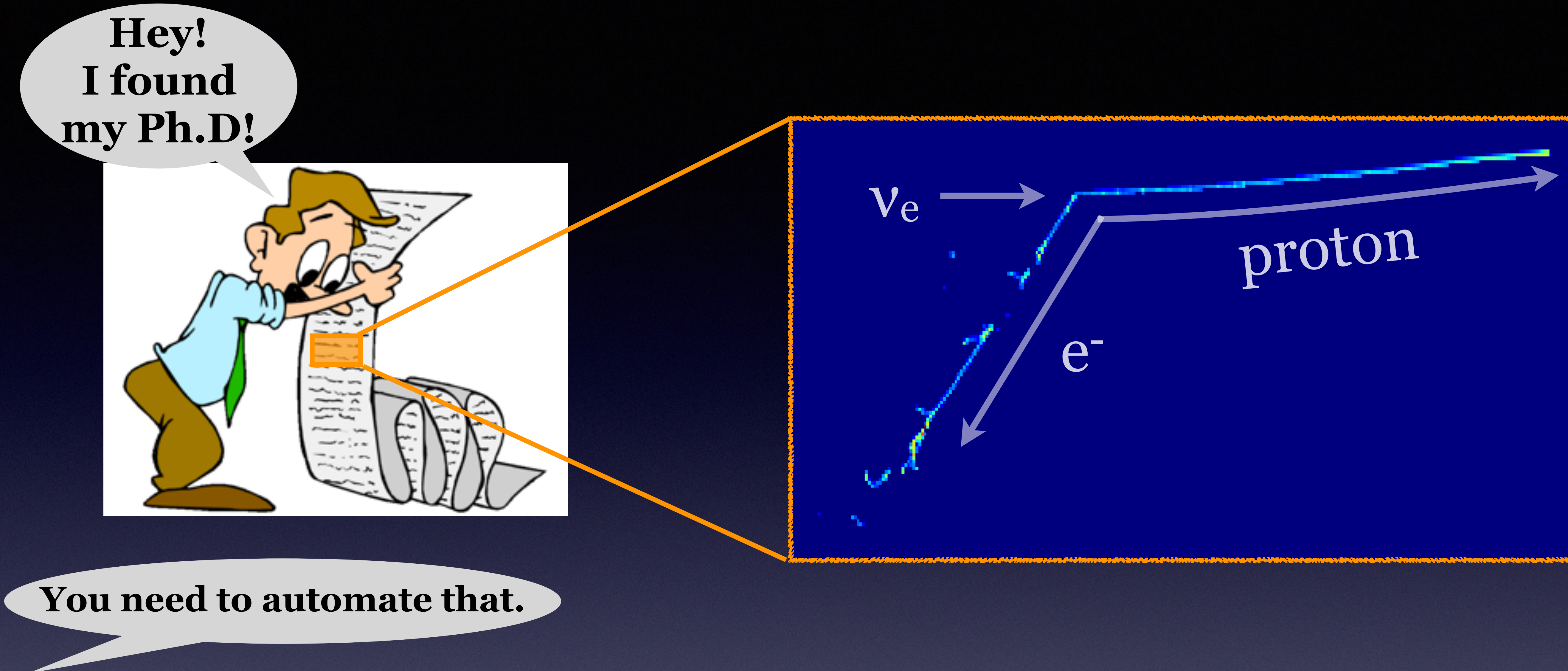
Michel

Data/Sim. Discrepancy

- Train discriminator, study the cause
- Generative Adversarial Network
  - Refine MC image to look like data
  - Train analysis CNN on refined sim.



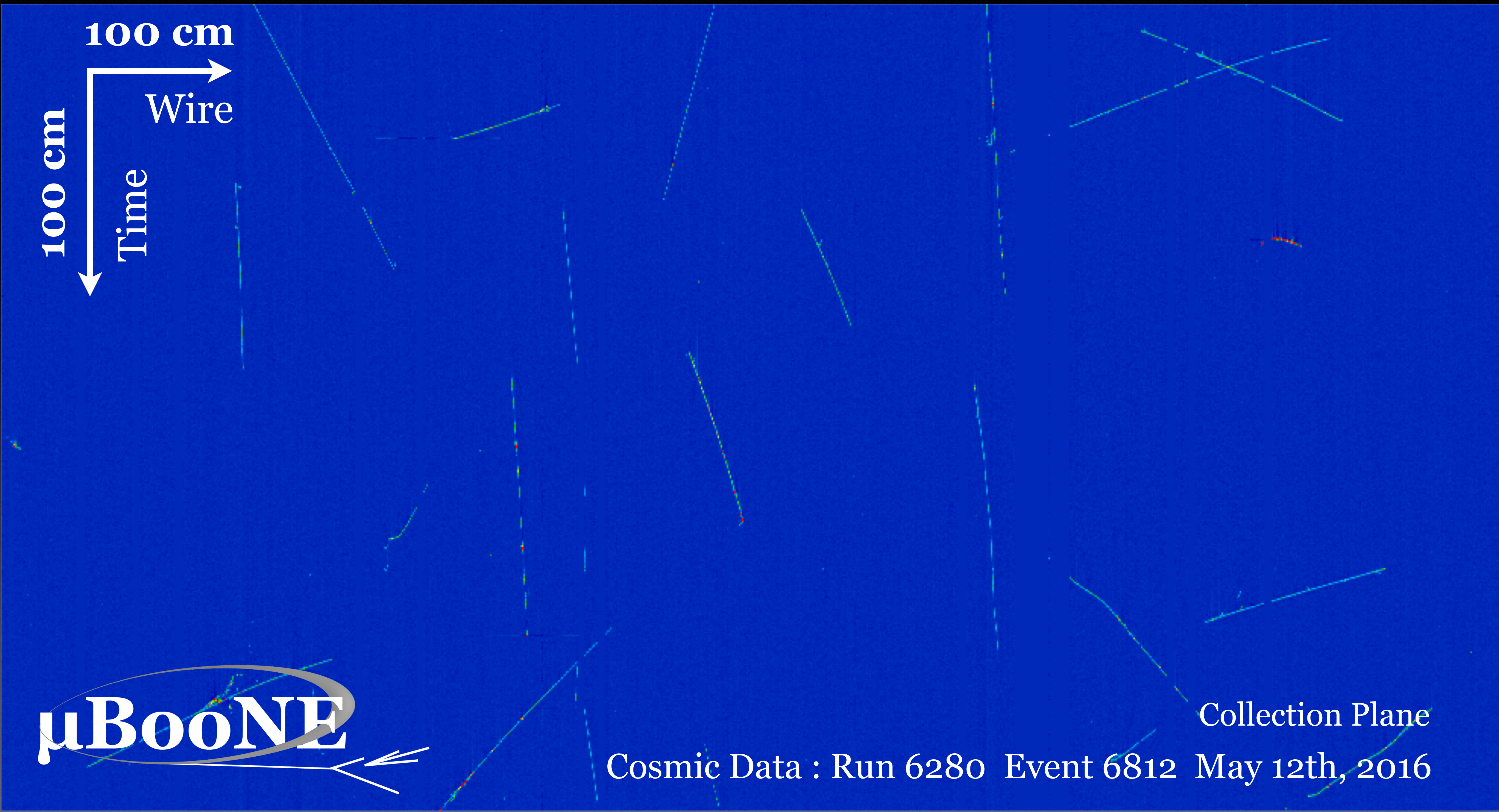
"**encoder**" for human eye illustration by Apple research team [arXiv:1612.07828](https://arxiv.org/abs/1612.07828)



Outline

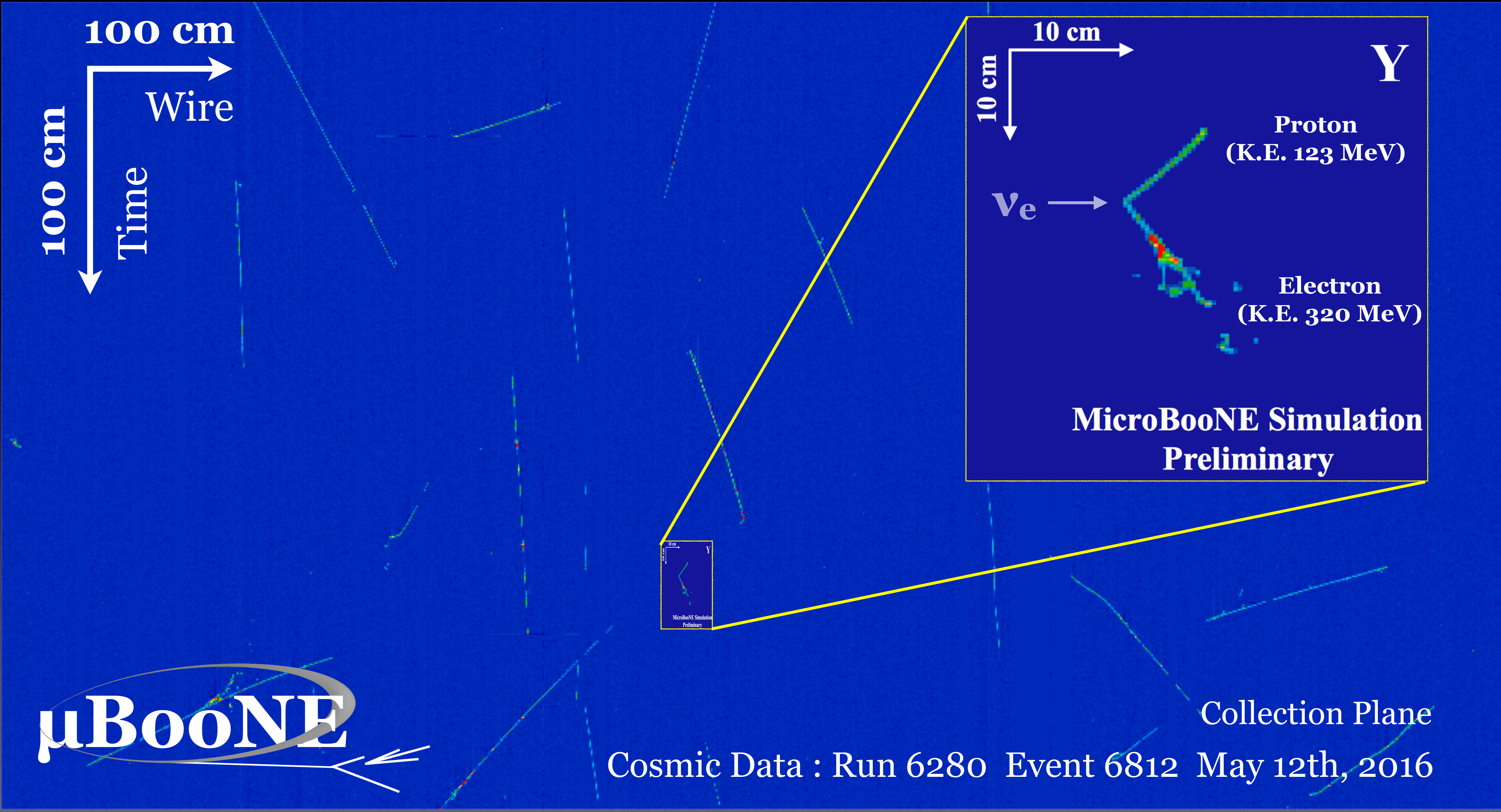
- Intro: what is deep learning?
- **Event reconstruction + analysis challenges**
- Deep neural network applications
- Summary

Challenges for Neutrino Analysis (I)



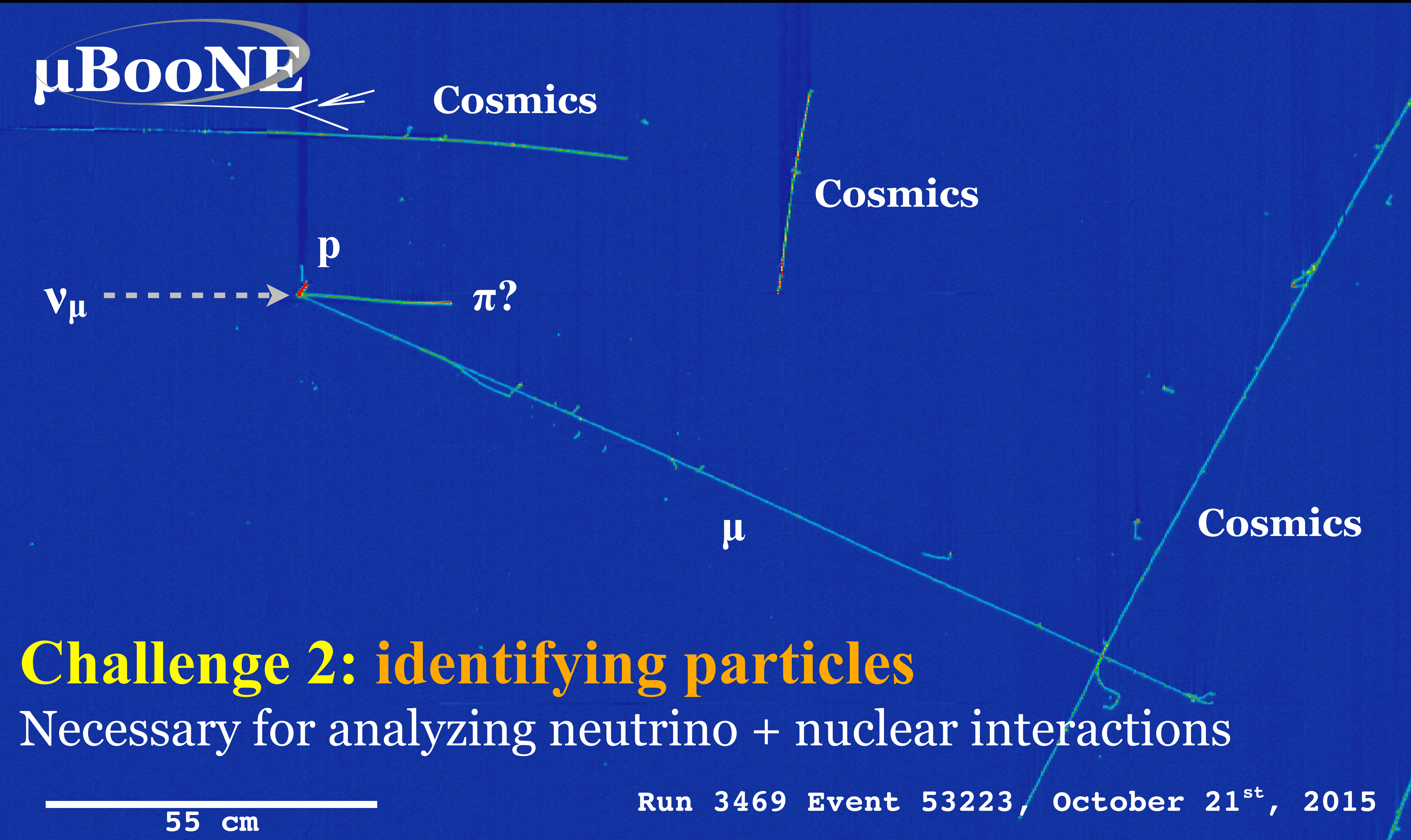
Challenge 1: our detector is **filled with cosmics** and **neutrino is rare**

Challenges for Neutrino Analysis (I)



Challenge 1: our detector is **filled with cosmics** and **neutrino is rare** … and **signal is small**

Challenges for Neutrino Analysis (II)



Challenge 2: identifying particles

Necessary for analyzing neutrino + nuclear interactions

Challenges for Neutrino Analysis (III)

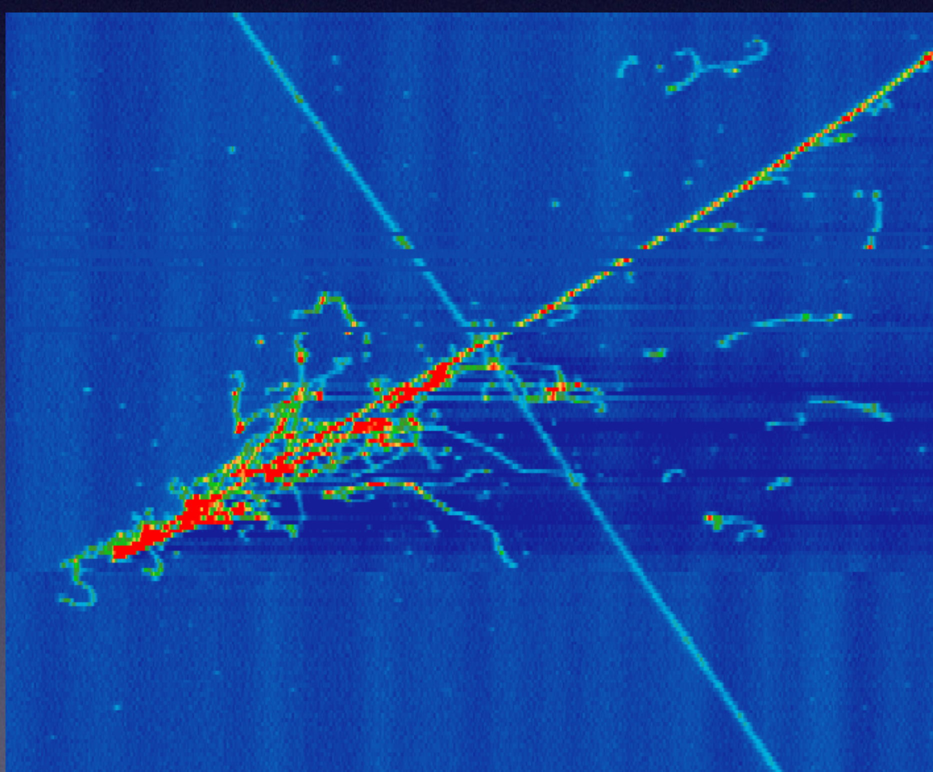
Challenge 3: Clustering

Reconstruction is already hard, and one must cluster all scattered charges

Challenges for Neutrino Analysis (IV)

Challenge 4: programming is not easy

Need efficient, fast pattern recognition algorithms and a framework to run a chain (or multiple chains) of them



Our data is an "image", a matrix of numbers

we wish

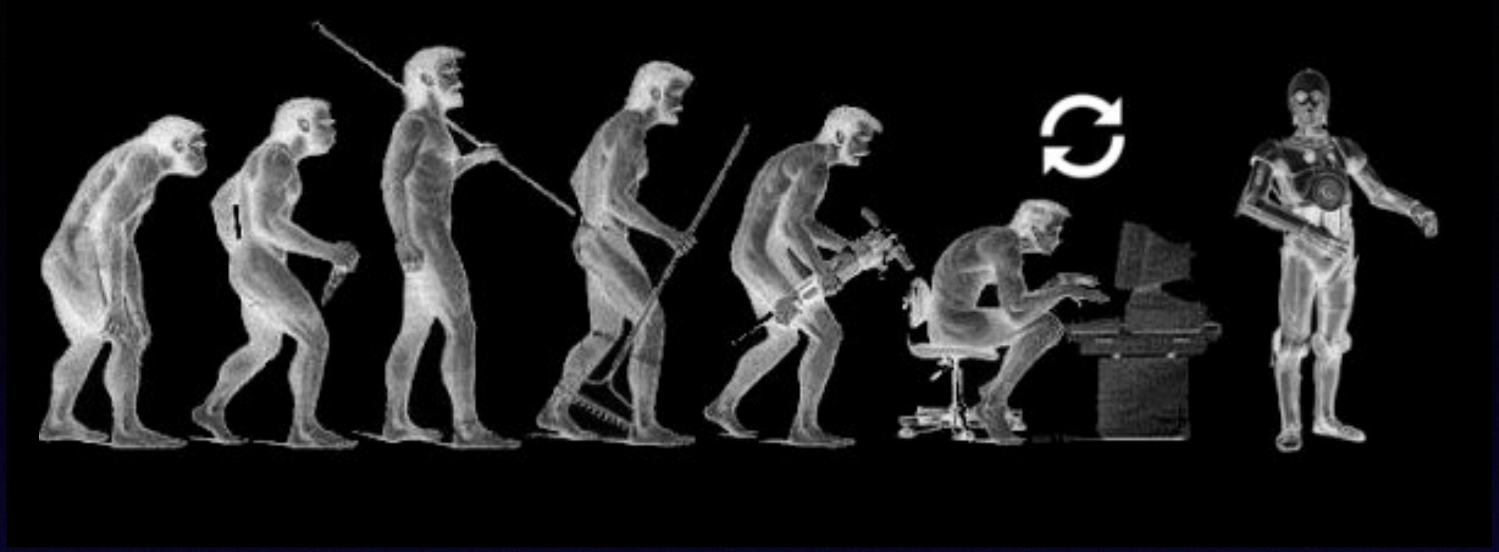
Not how it looks in C++

in reality

01101010100101011010101001011010
10111010101001010100010010101101
0101001011010101001010110101010
01011010101001010110101010101101
0101001010110101010010110101010
01011010101001010110101010010110
10101001010110101010101101010100
1010110101010011010110101010100101

This is how it looks in C++

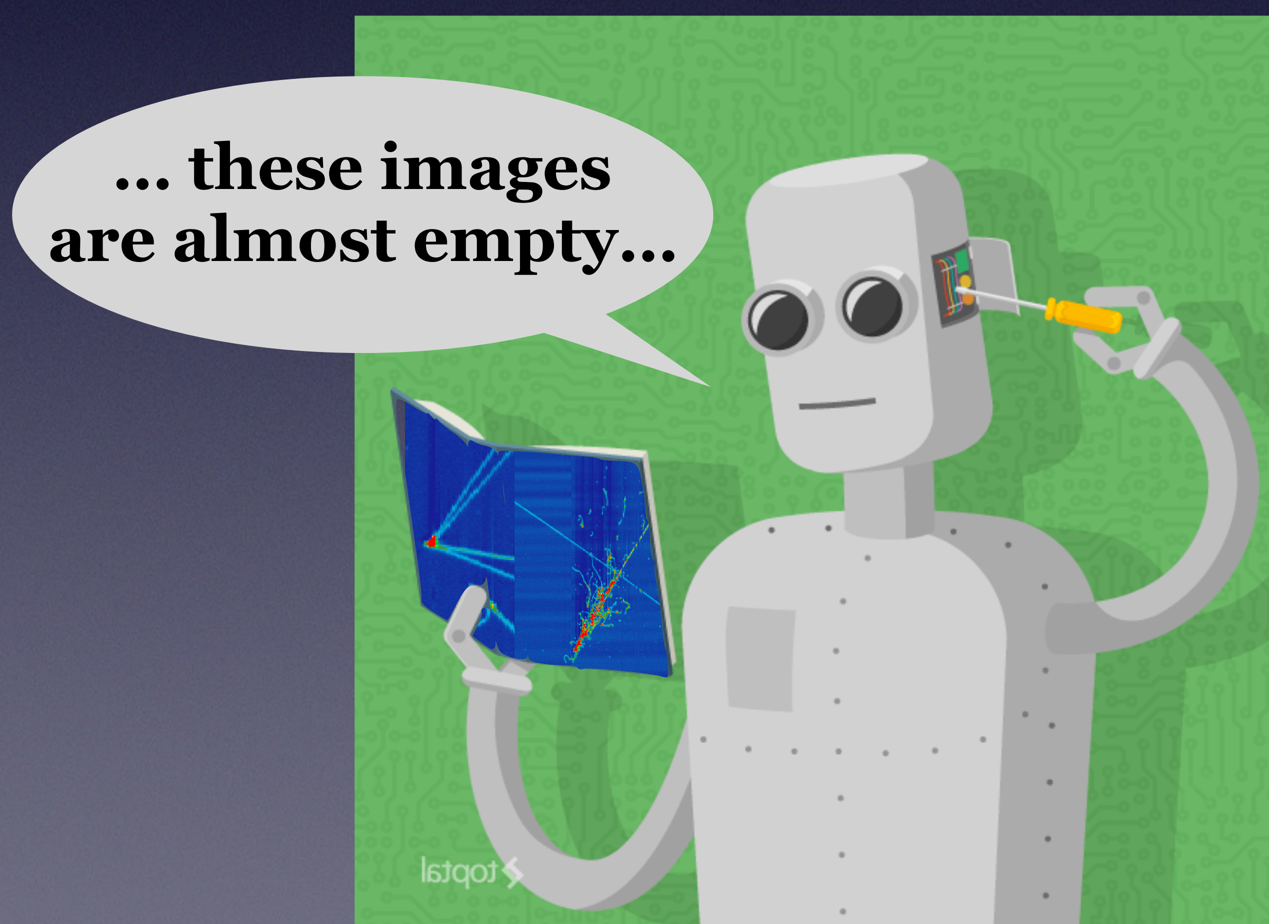
… enough challenges …



Solutions?

- **Path A**: "**traditional path**"
  - Hand-engineered reconstruction algorithms
- **Path B**: **machine learning**
  - "**Deep Learning**"
    - In particular…
      Convolutional Neural Networks (CNNs)
    - Scalable technique, generalizable to various tasks
    - Superb performance on image data analysis

CNN for LArTPC
Image Analysis

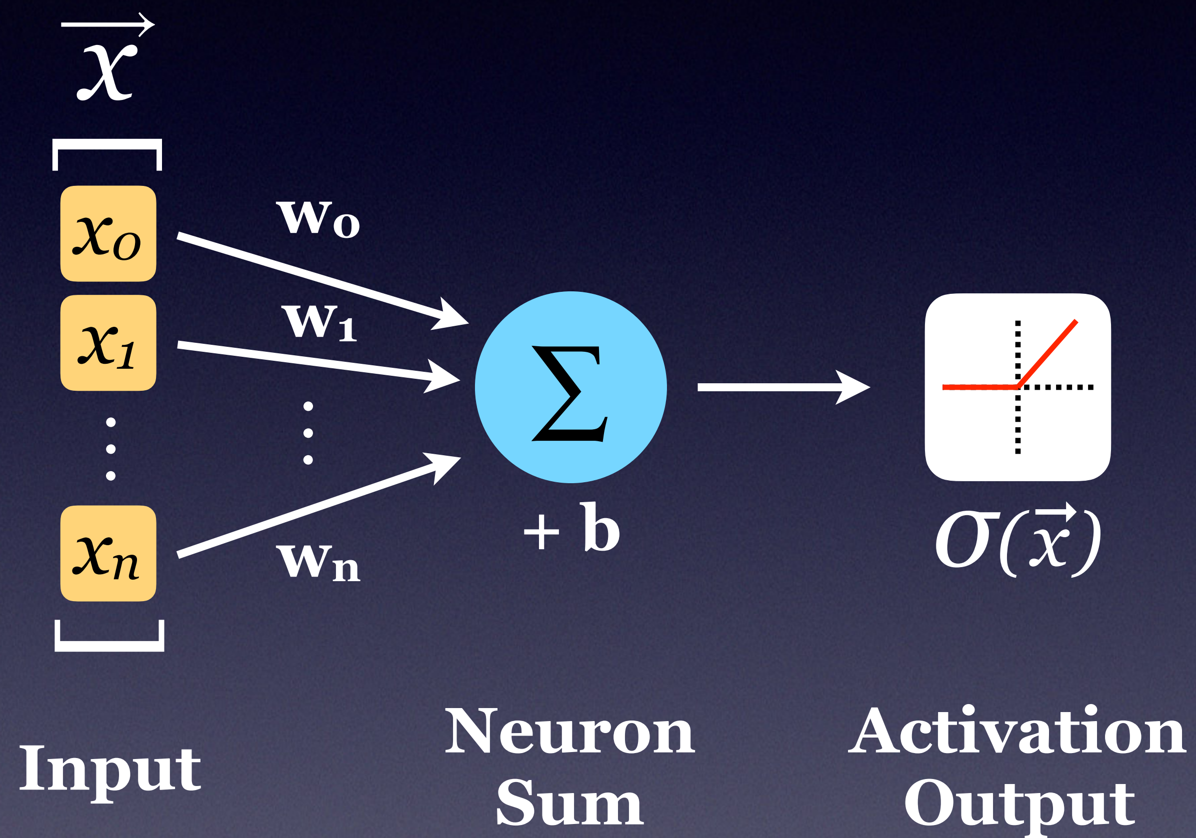


Introduction to CNNs (II)

Background: Neural Net

The basic unit of a neural net is the ***perceptron*** (loosely based on a real neuron)

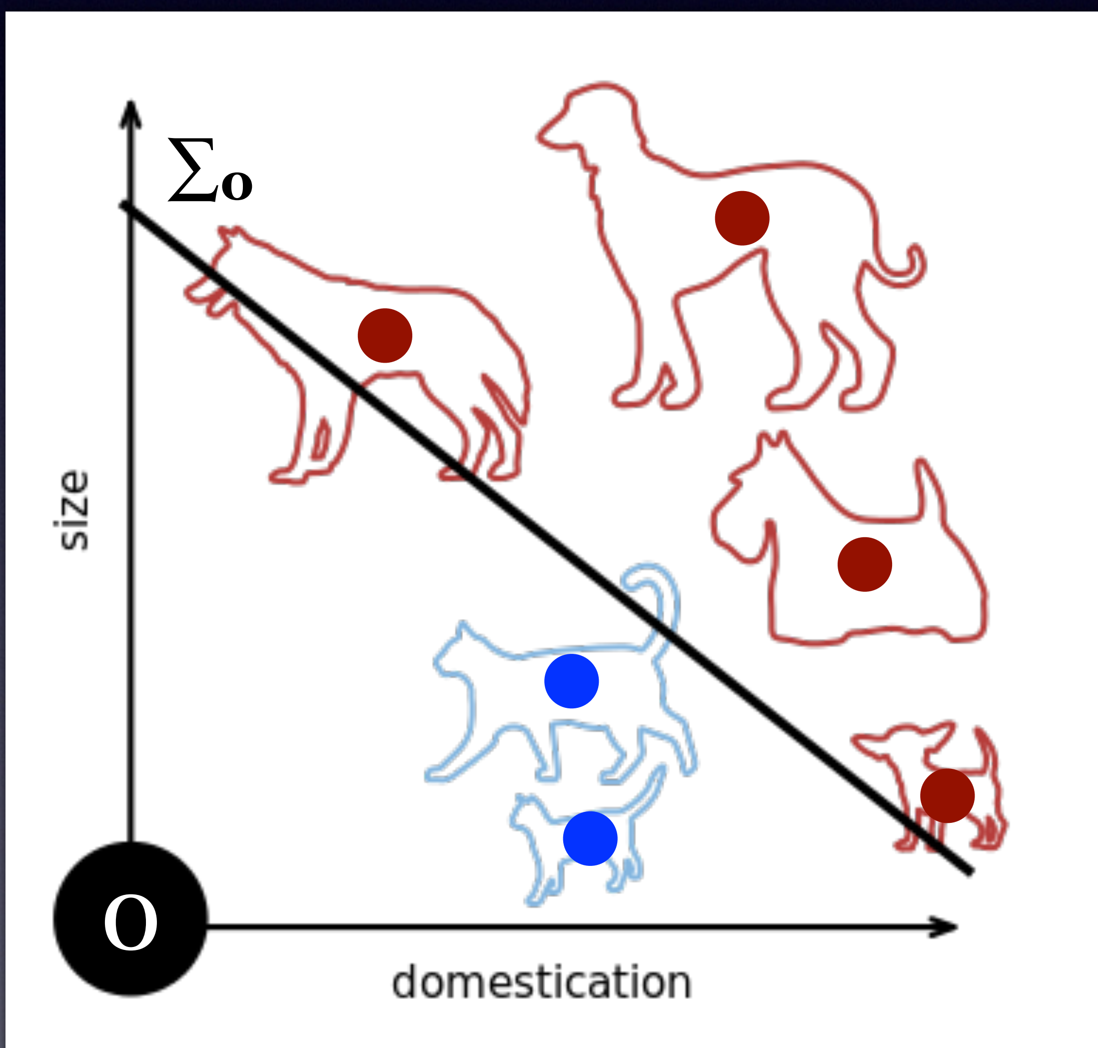
Takes in a vector of inputs (*x*). Commonly inputs are summed with weights (*w*) and offset (*b*) then run through activation.



$$\sigma(\vec{x}) = \begin{cases} \vec{w}_i \cdot \vec{x} + b_i & \vec{w}_i \cdot \vec{x} + b_i \geq 0 \\ 0 & \vec{w}_i \cdot \vec{x} + b_i < 0. \end{cases}$$

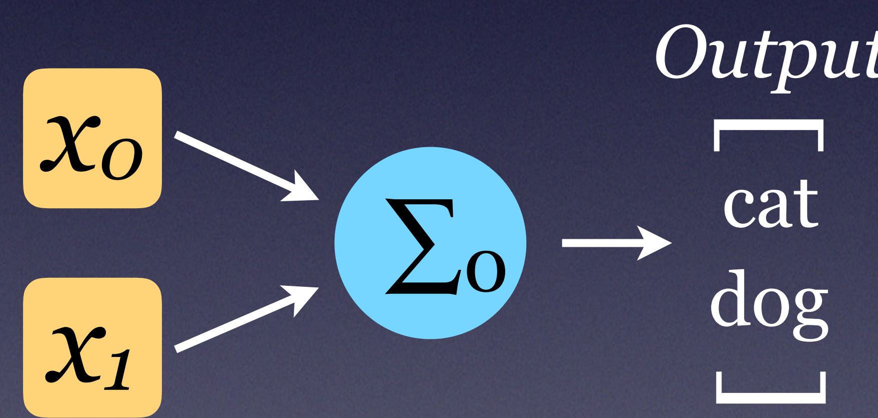
Introduction to CNNs (II)
Perceptron 2D Classification

Imagine using two features to separate cats and dogs



from [wikipedia]

$$\sigma(\vec{x}) = \begin{cases} \vec{w}_i \cdot \vec{x} + b_i & \vec{w}_i \cdot \vec{x} + b_i \geq 0 \\ 0 & \vec{w}_i \cdot \vec{x} + b_i < 0. \end{cases}$$

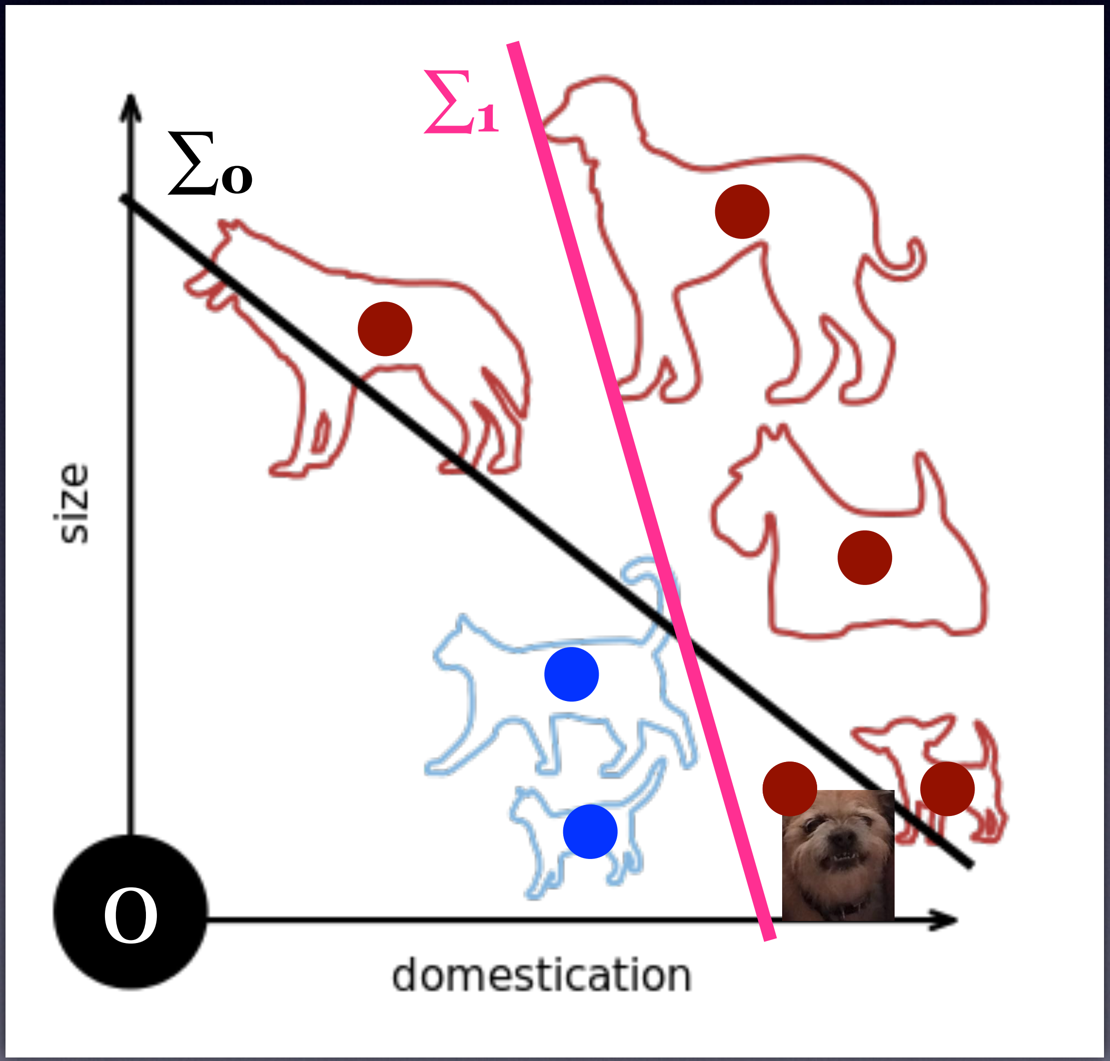


By picking a value for w and b, we define a boundary between the two sets of data

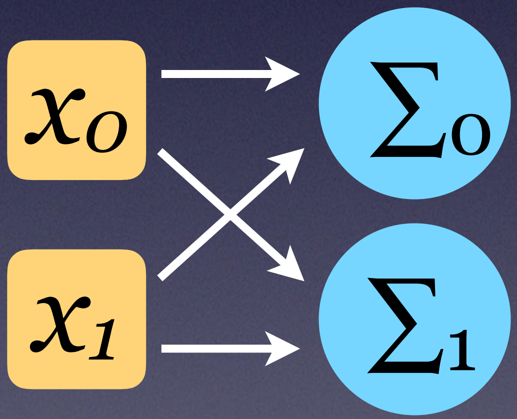
Introduction to CNNs (II)

Perceptron 2D Classification

Maybe we need to do better: assume new data point
(My friend's dog — small but not as well behaved)



from [wikipedia]

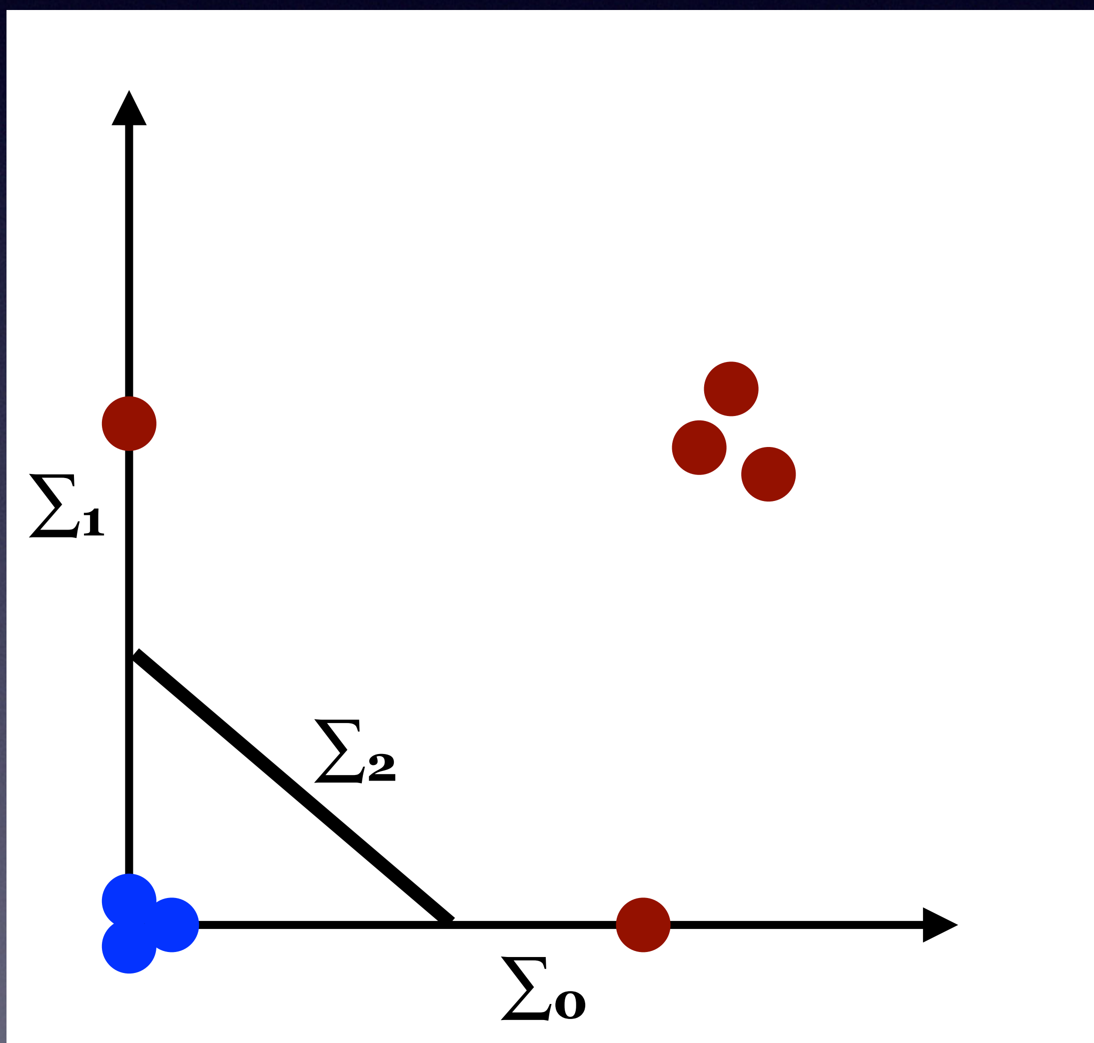


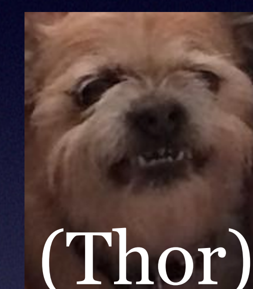
We can add another perceptron
to help classify better

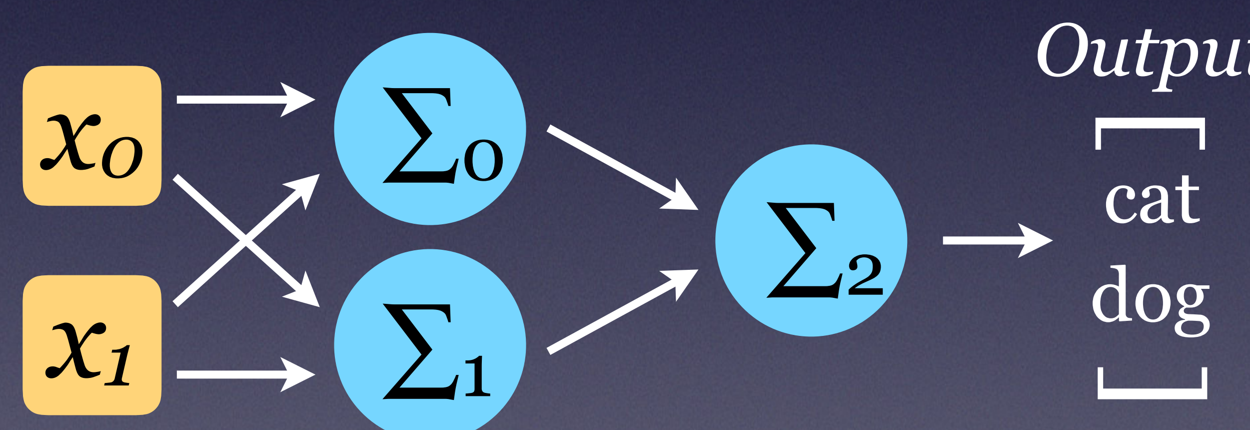
Introduction to CNNs (II)

Perceptron 2D Classification

Maybe we need to do better: assume new data point
(My friend's dog — small but not as well behaved)





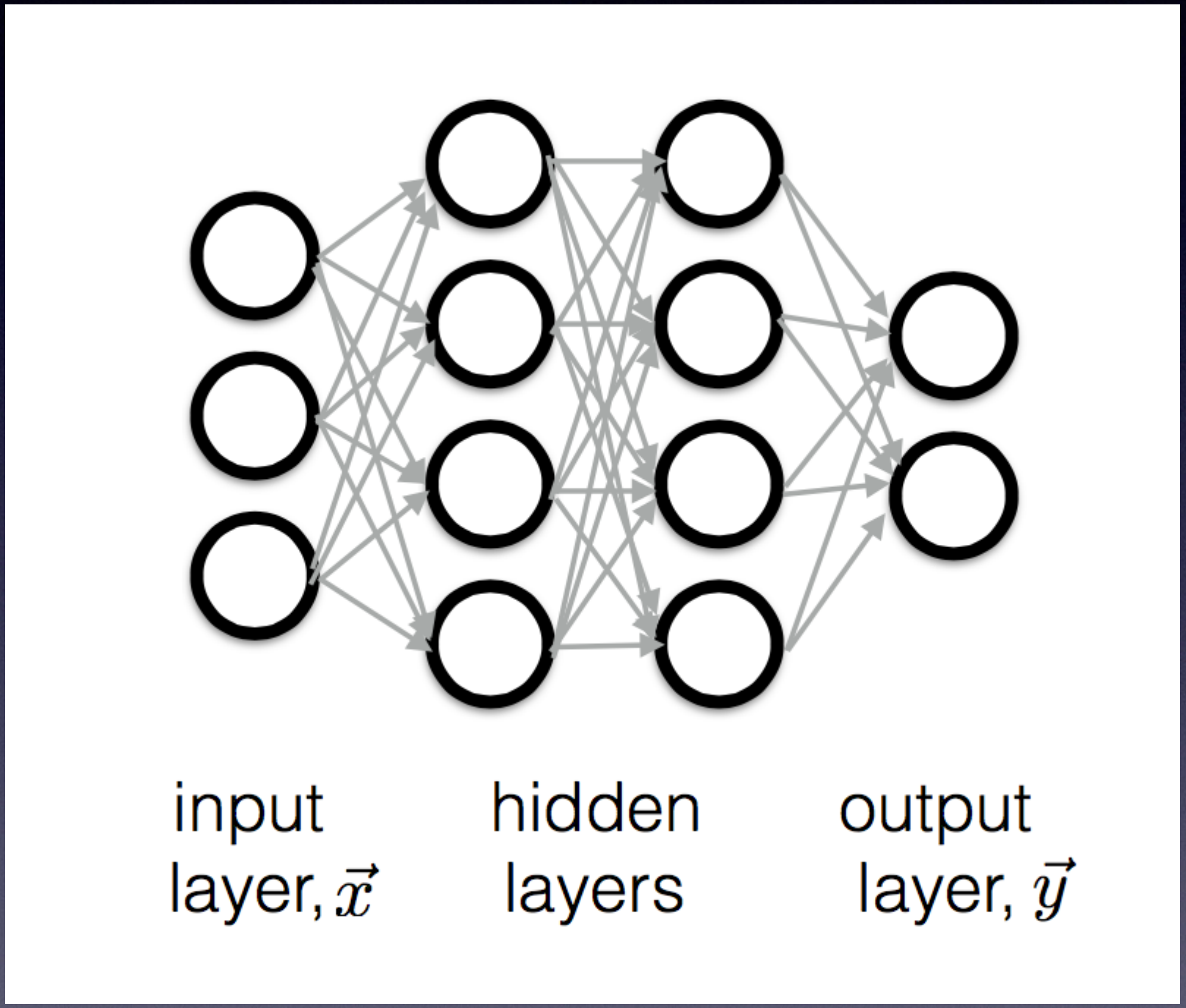


Another layer can classify based on preceding feature layer output

Introduction to CNNs (III)

"Traditional neural net" in HEP

Fully-Connected Multi-Layer Perceptrons



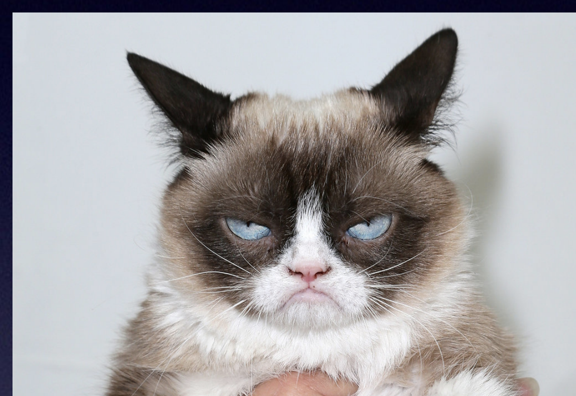
A traditional neural network consists of a stack of layers of such neurons where each neuron is ***fully connected*** to other neurons of the neighbor layers

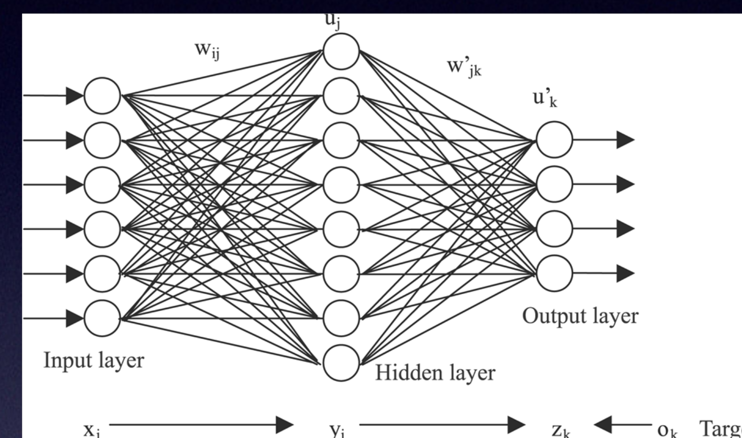
Introduction to CNNs (III)

“Traditional neural net” in HEP

Problems with it…

Feed in entire image

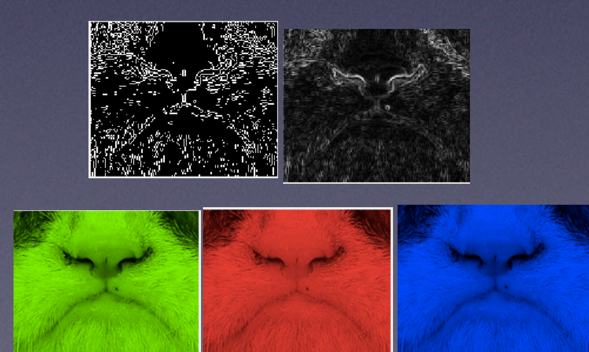


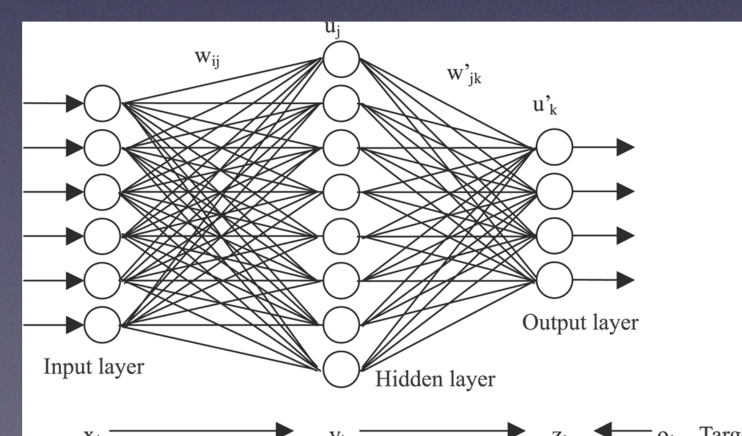


Cat?

Problem: scalability

Use pre-determined features



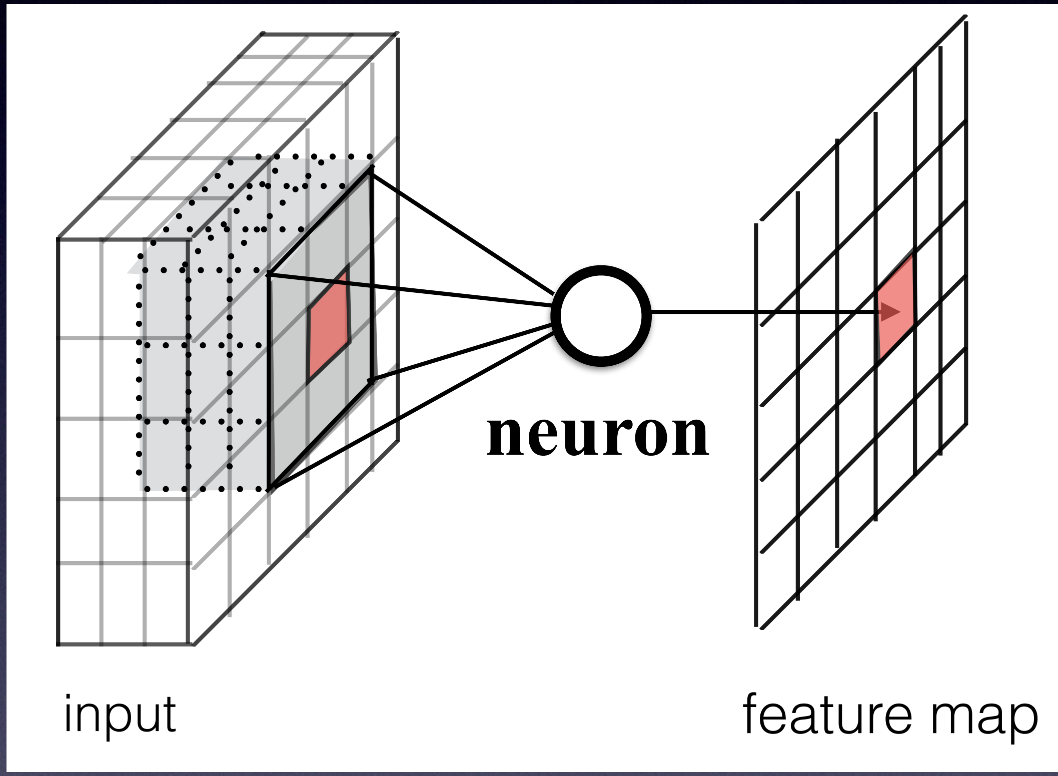


Cat?

Problem: generalization

Introduction to CNNs (III)

CNN introduce a ***limitation*** by forcing the network to look at only local, translation invariant features

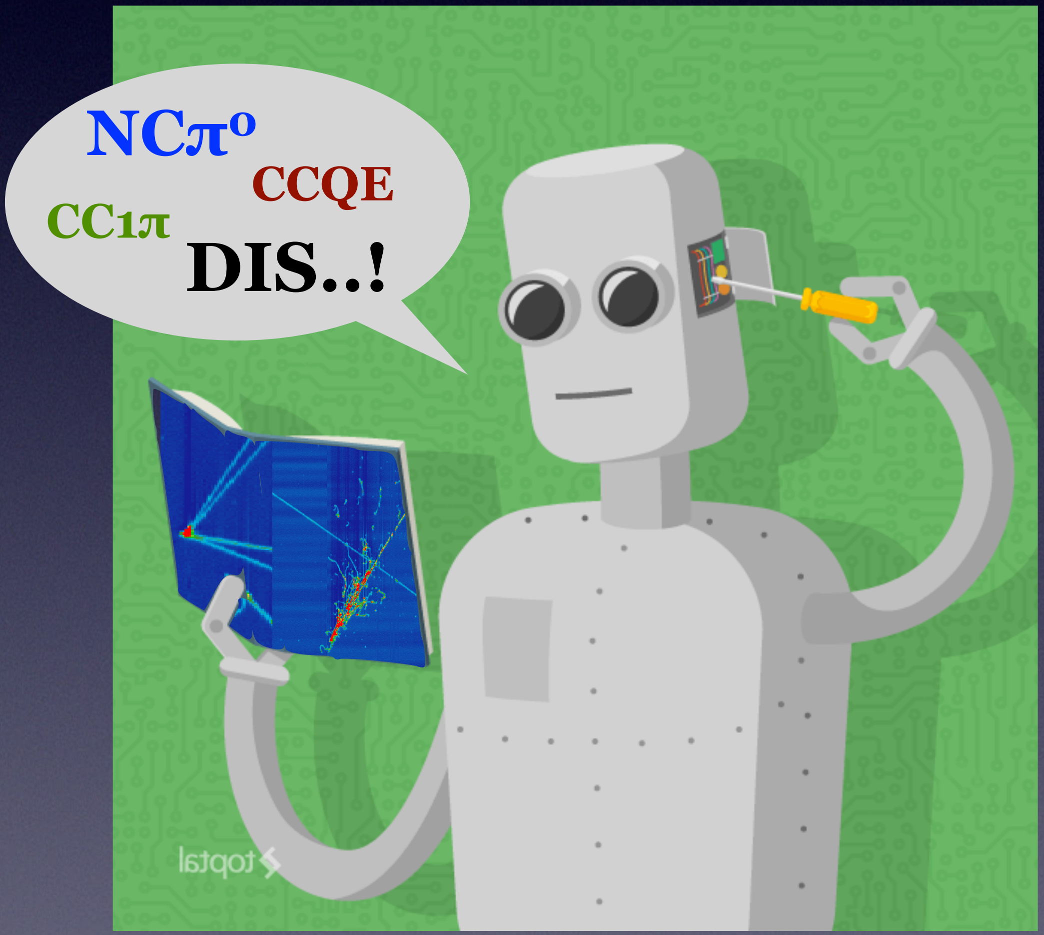


$$f_{i,j}(X) = \sigma\left(W_i \cdot X_j + b_i\right),$$

Activation of a neuron depends on the element-wise product of 3D weight tensor with 3D input data and a bias term

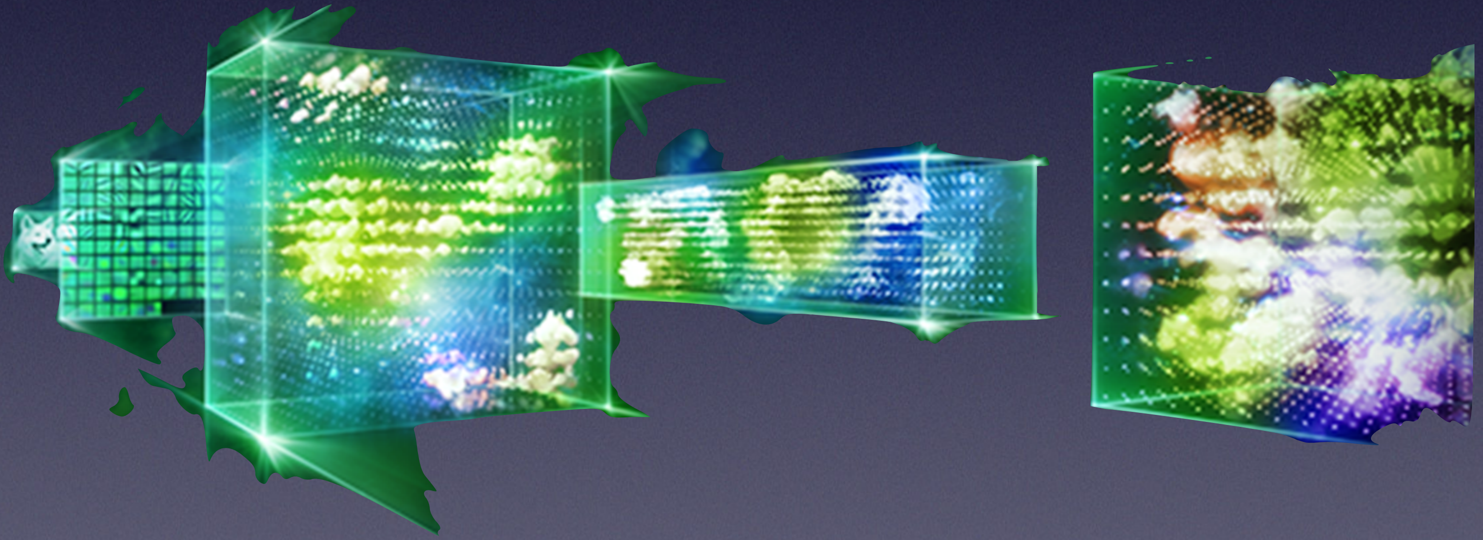
- Translate over 2D space to process the whole input
- Neuron **learns translation-invariant features**
- Applicable for a "homogeneous" detector like LArTPC

Want more details? Feel free to ask me later!



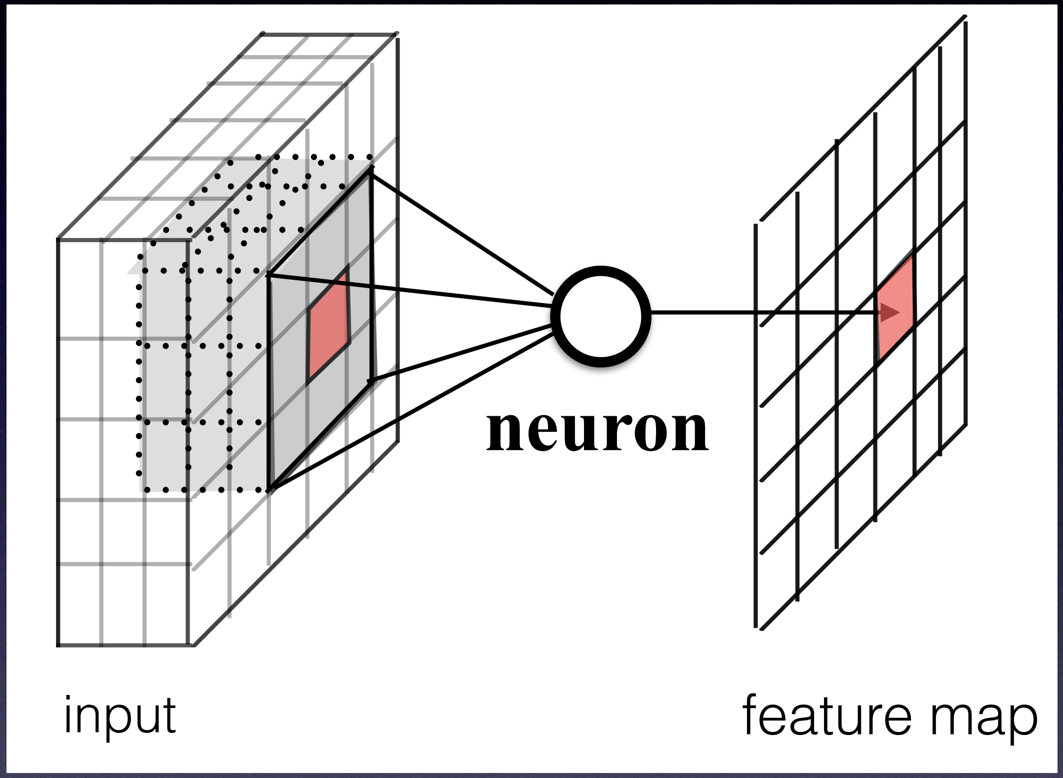
Track/Shower Pixel Labeling

~ How Does SSNet Work? ~



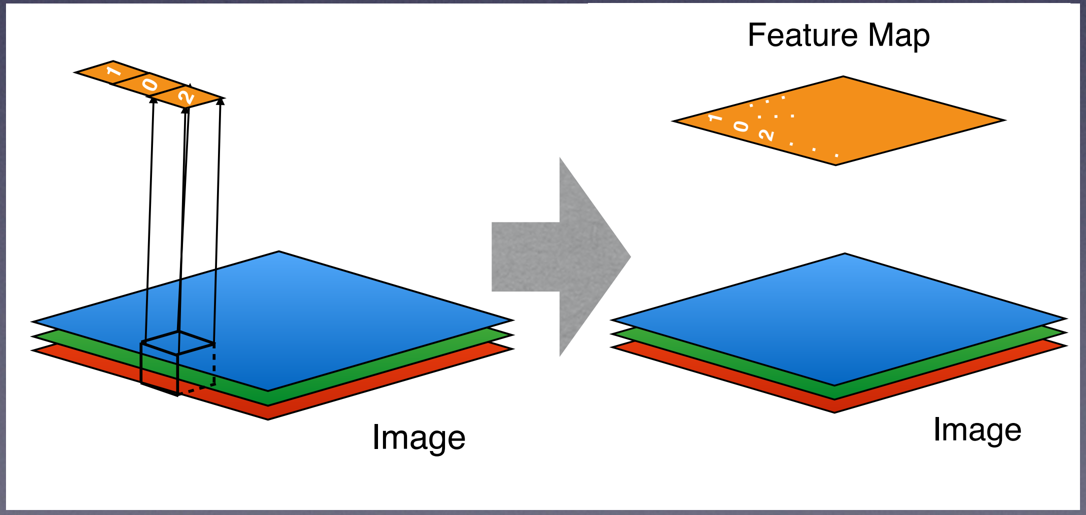
Quick Recap on CNN

CNN is a neural network formed with multiple convolution layers of neurons



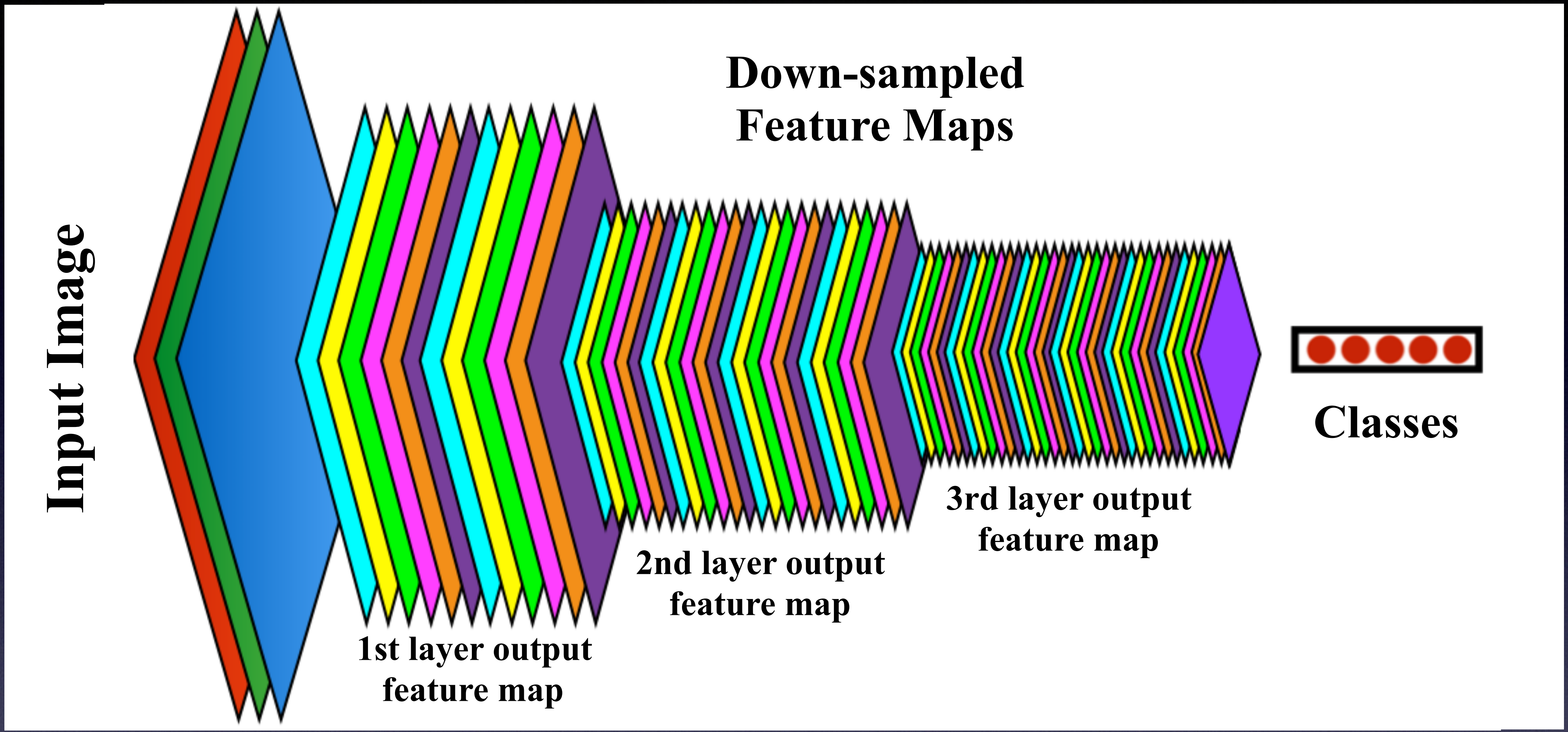
$$f_{i,j}(X) = \sigma\left(W_i \cdot X_j + b_i\right),$$

Activation of a neuron depends on the element-wise product of 3D weight tensor with 3D input data and a bias term



Each **filter** (neuron) translates over 2D space to process the whole input, producing a "***feature map***".

Quick Recap on CNN



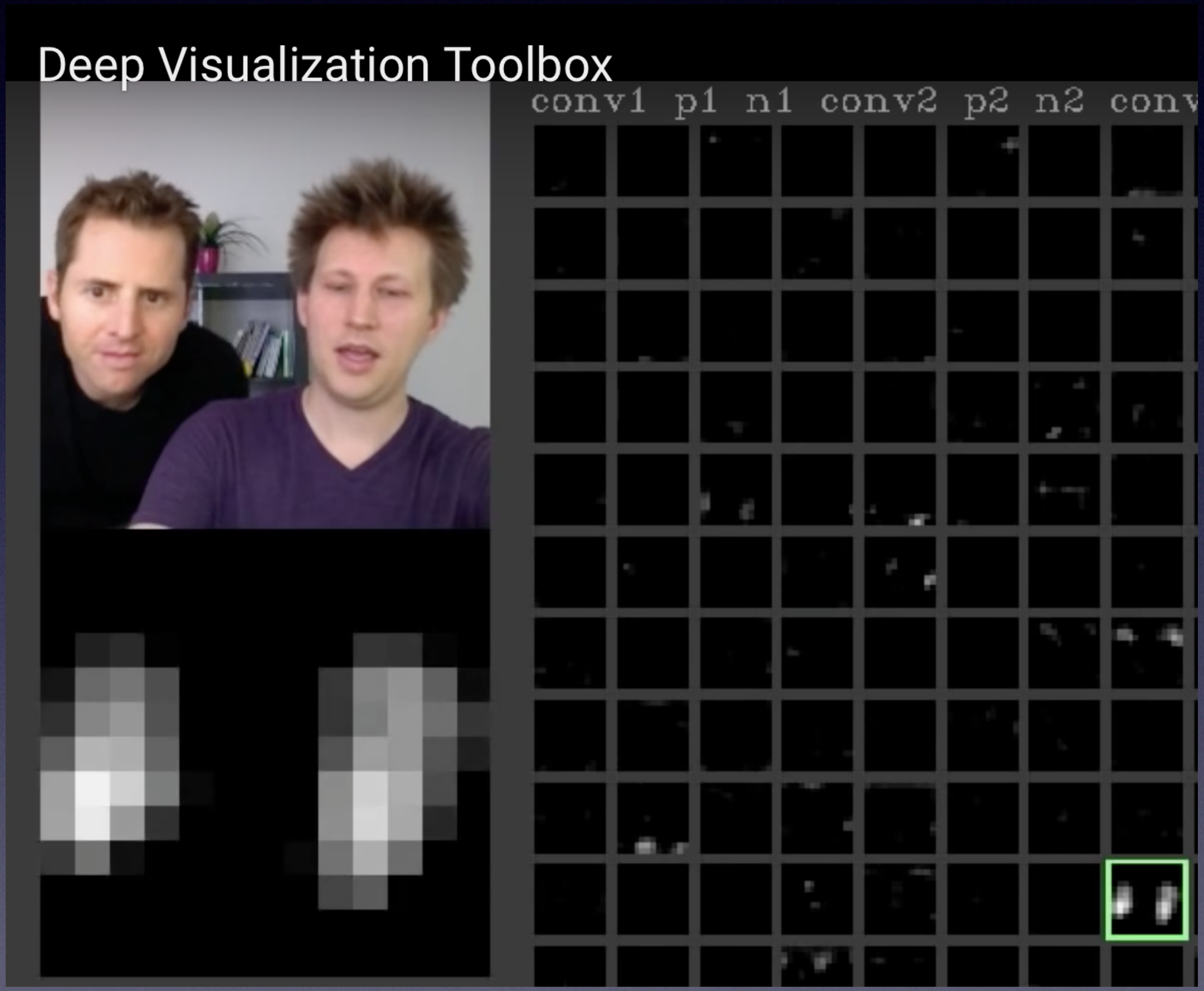
CNN for image classification

- **Goal**: provide a single label for the whole image
- **How**: transform the higher spatial resolution input (i.e. image) into a vector of image features, ultimately a 1D array of feature parameters useful for the whole image labeling, by a successful chain of convolutional and pooling operations.

Quick Recap on CNN

Feature map visualization example

- https://www.youtube.com/watch?v=AgkfIQ4IGaM



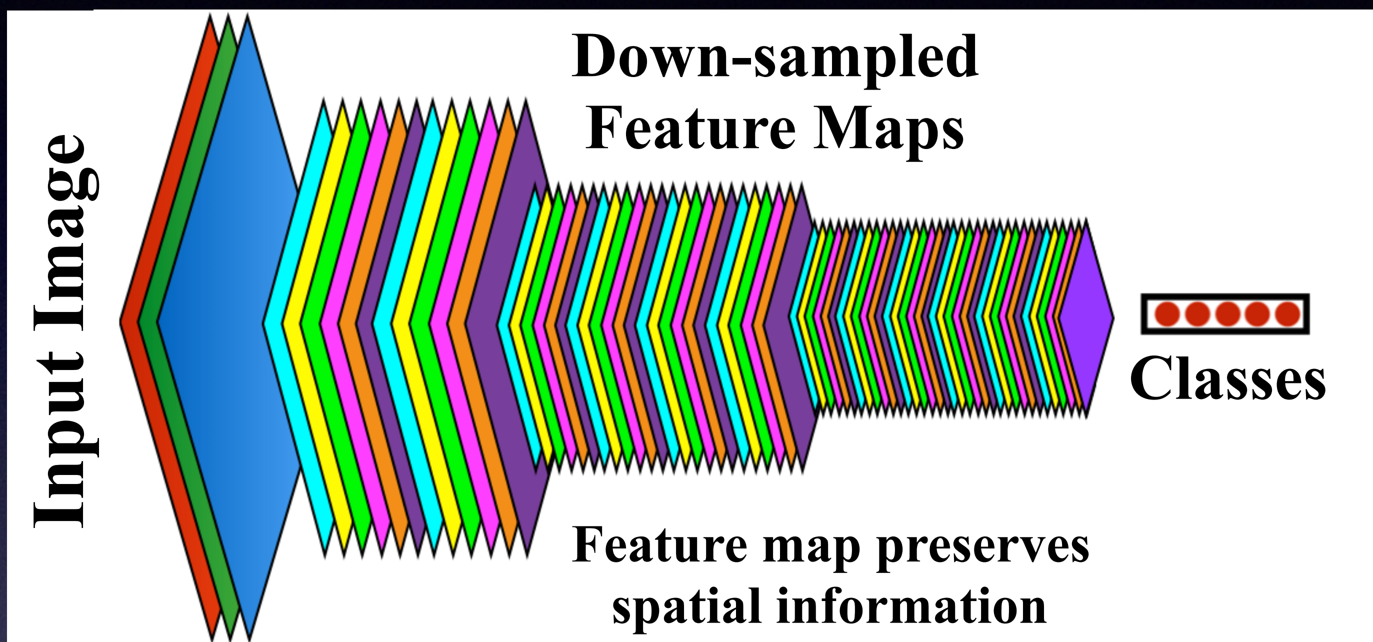
Neuron concerning face

Neuron loving texts

Semantic Segmentation Network

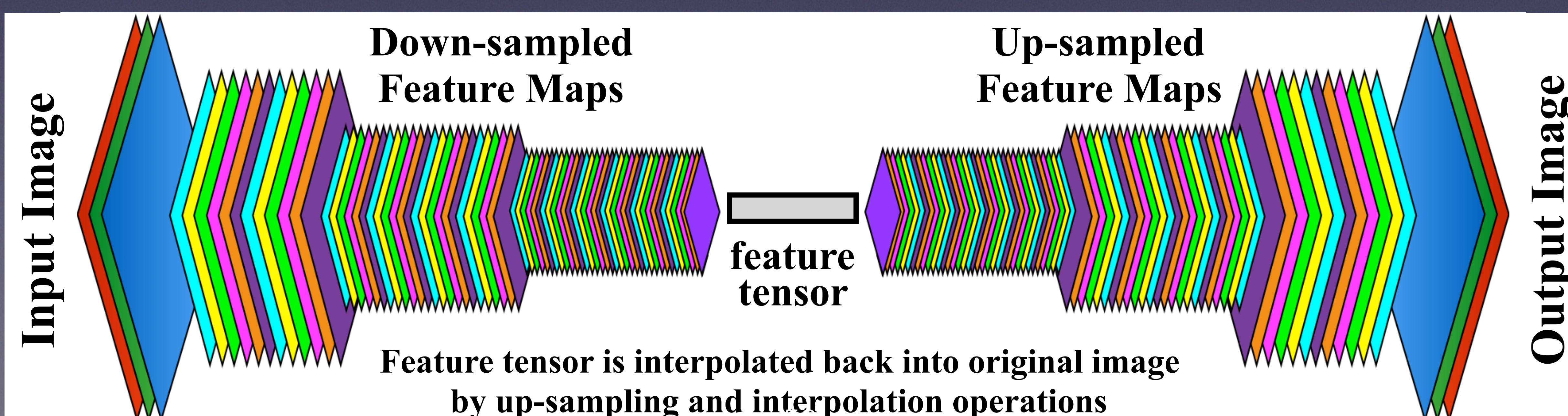
How is it different from ***Image Classification***?

Example CNN for Image Classification



- Classification network reduces the whole image into final "class" 1D aray
- SSNet, after extracting class feature tensor, interpolates back into original image size

Example CNN for Semantic Segmentation

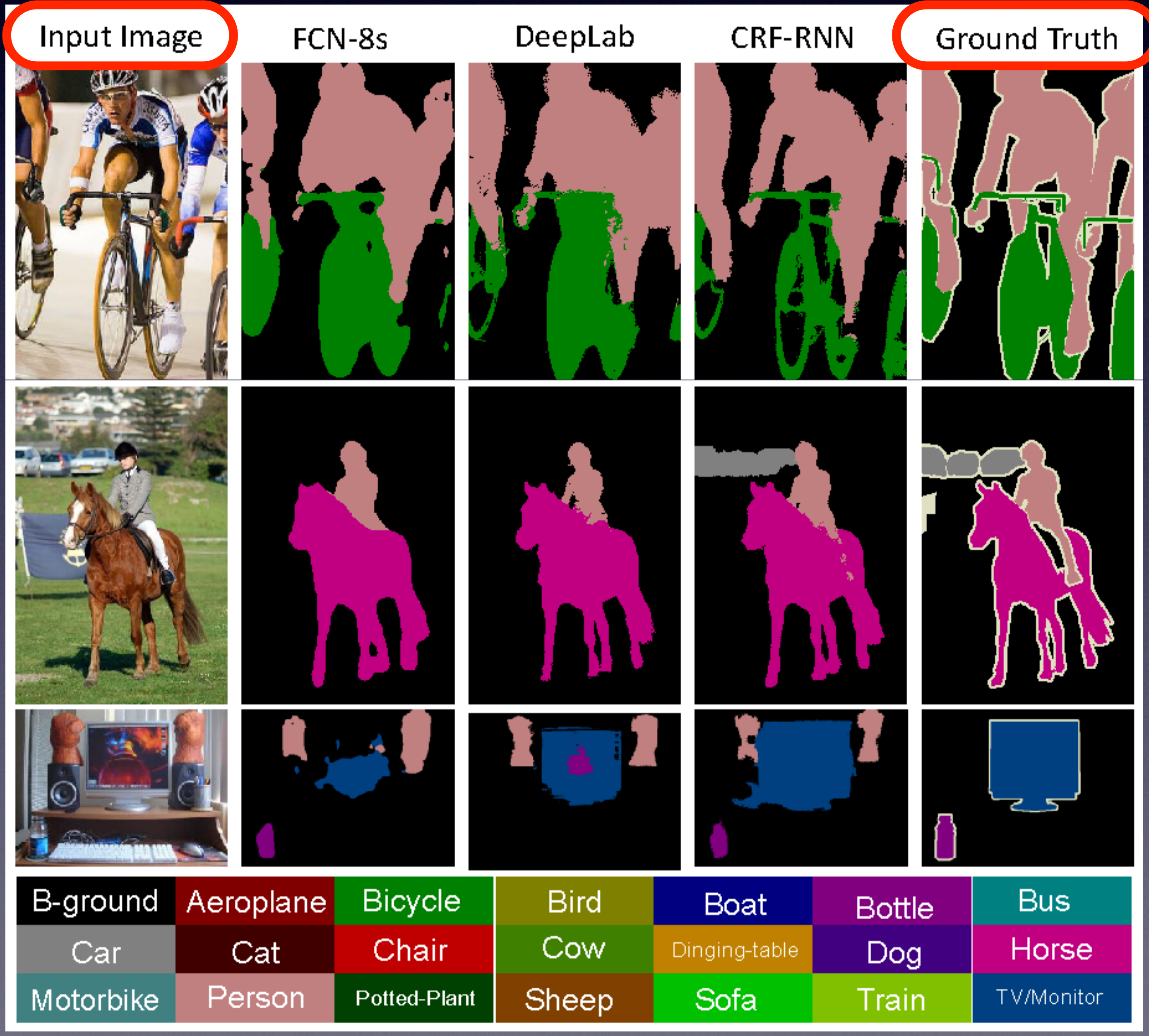


Semantic Segmentation Network

How to train SSNet?

Supervised training, like image classification

But the ***labels (and errors) are pixel-wise***



Semantic Segmentation Network

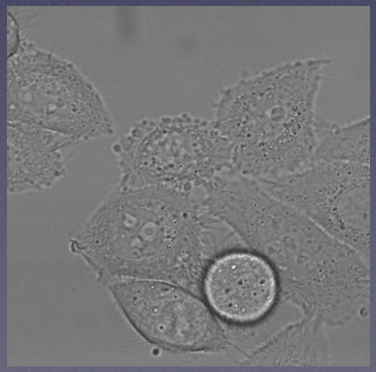
SSNet UB Analysis

U-Net + **ResNet** module design

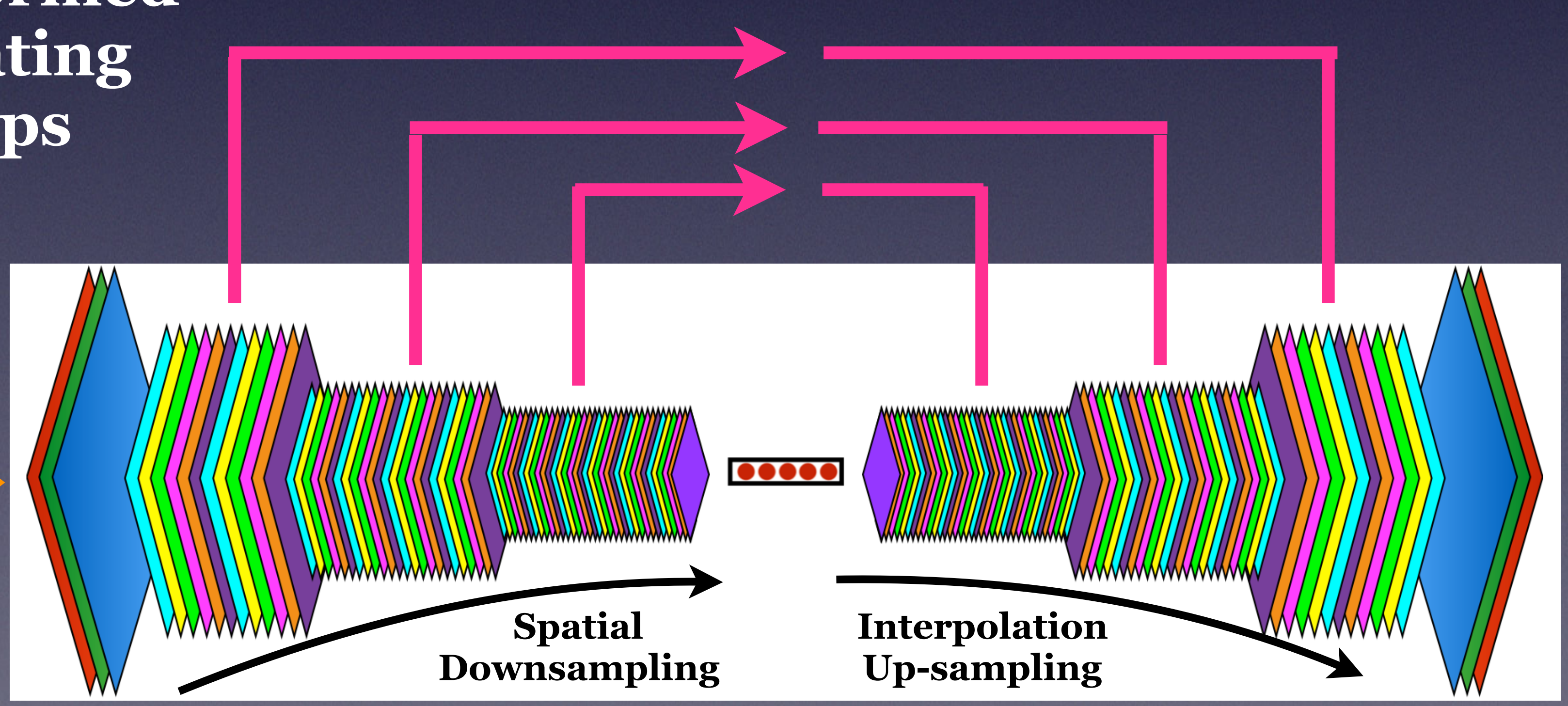
- Developed for bio-medical research
  - … to mask pixels of living cells (for automatized image analysis)
  - Designed for better spatial accuracy to get cells' boundary correct
- Use ResNet architecture for convolution layers

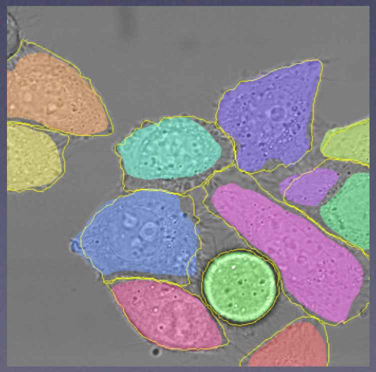
"U" shape if formed by concatenating feature maps

Segmented pixels of living cells (yellow boarder is truth label) arXiv:1505.04597



Biocell "Raw" Image





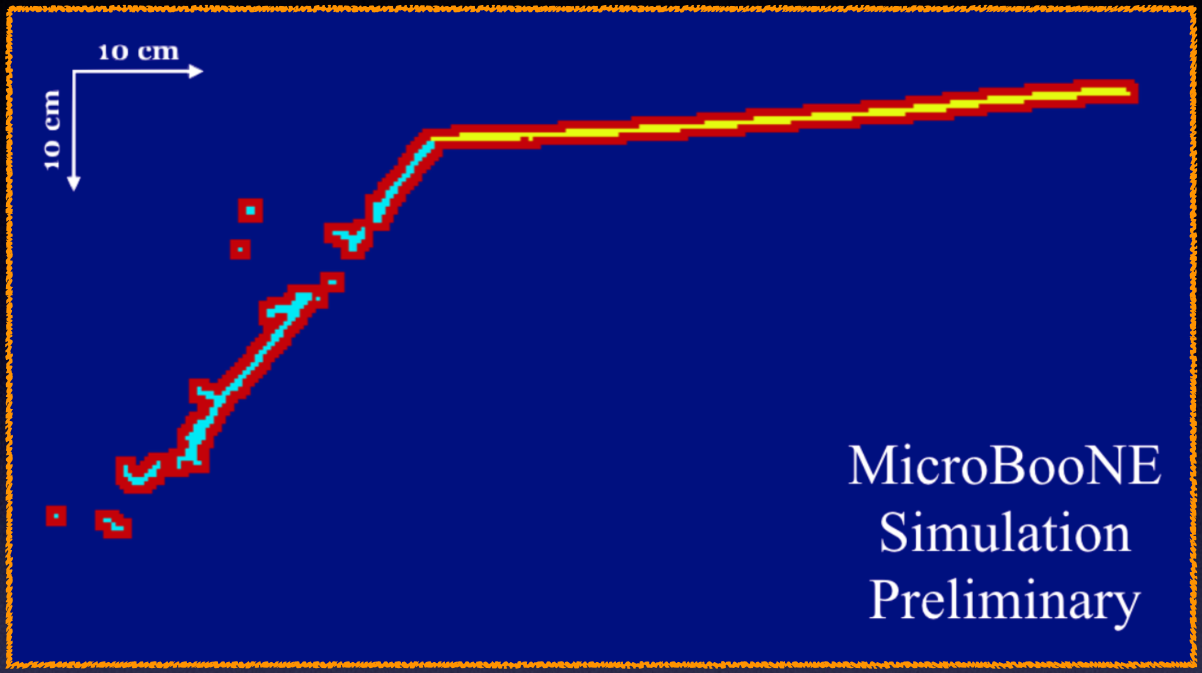
Biocell Segmented

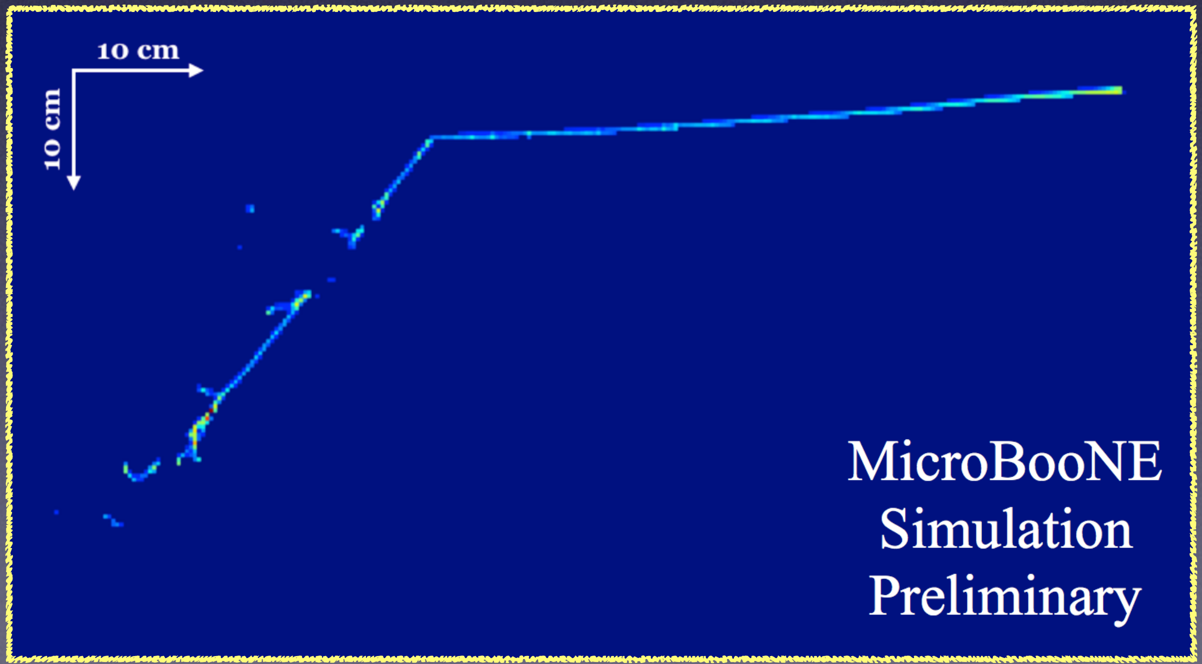
Training SSNet

"Pixel Weight" for training

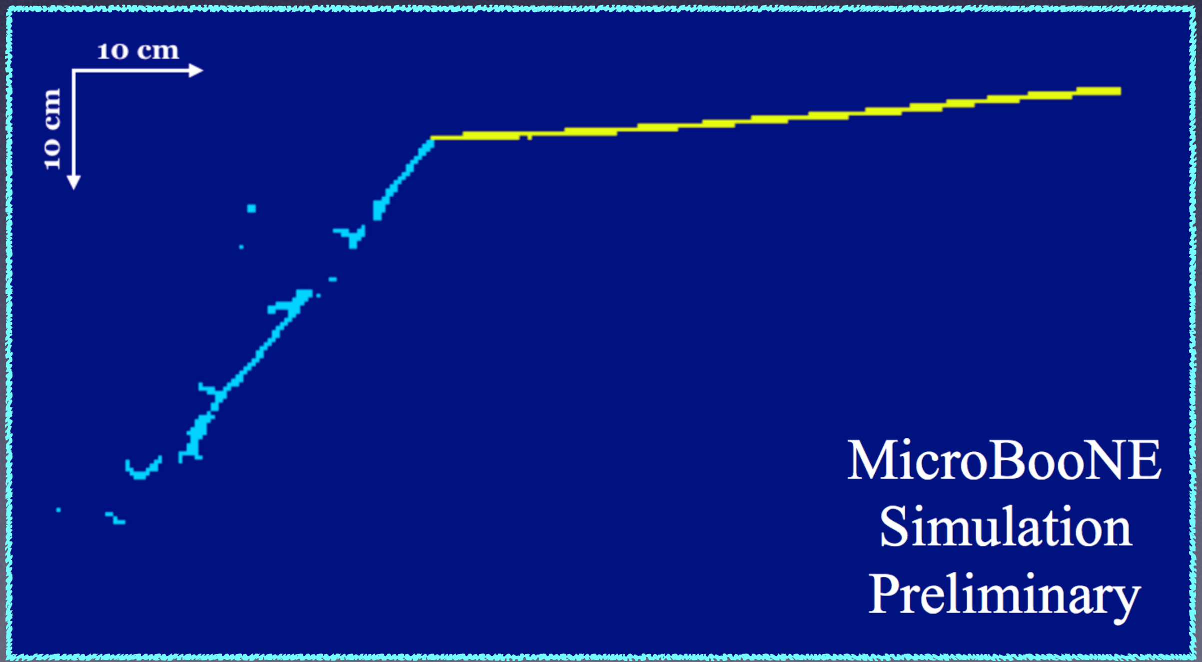
- Assign pixel-wise "weight" to penalize mistakes
- Weights inversely proportional to each type of pixel count
- Useful for LArTPC images (low information density)

"Weight" Image
(for training)





Input Image

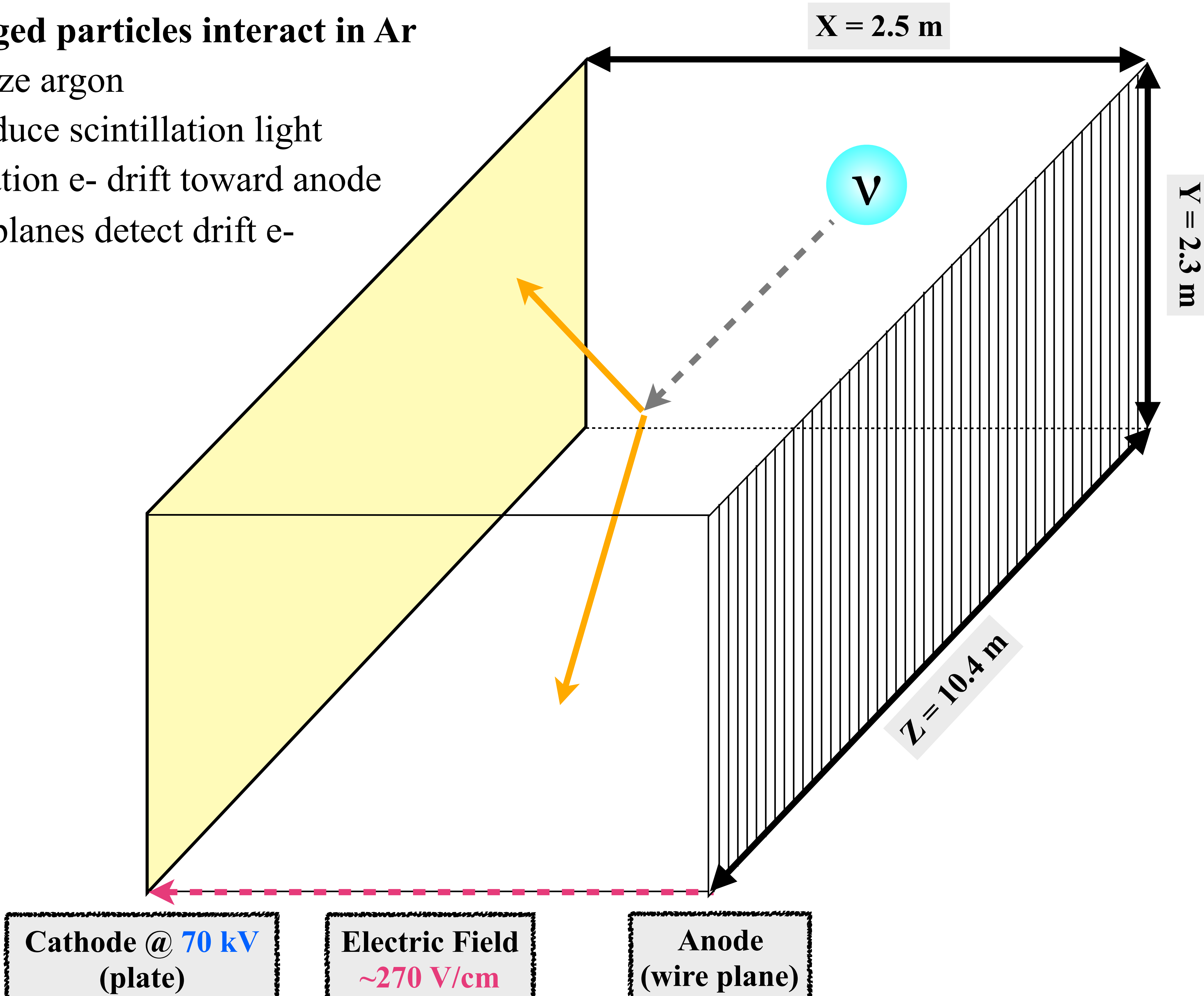


"Label" Image
(for training)

MicroBooNE LArTPC Detector Quick Guide

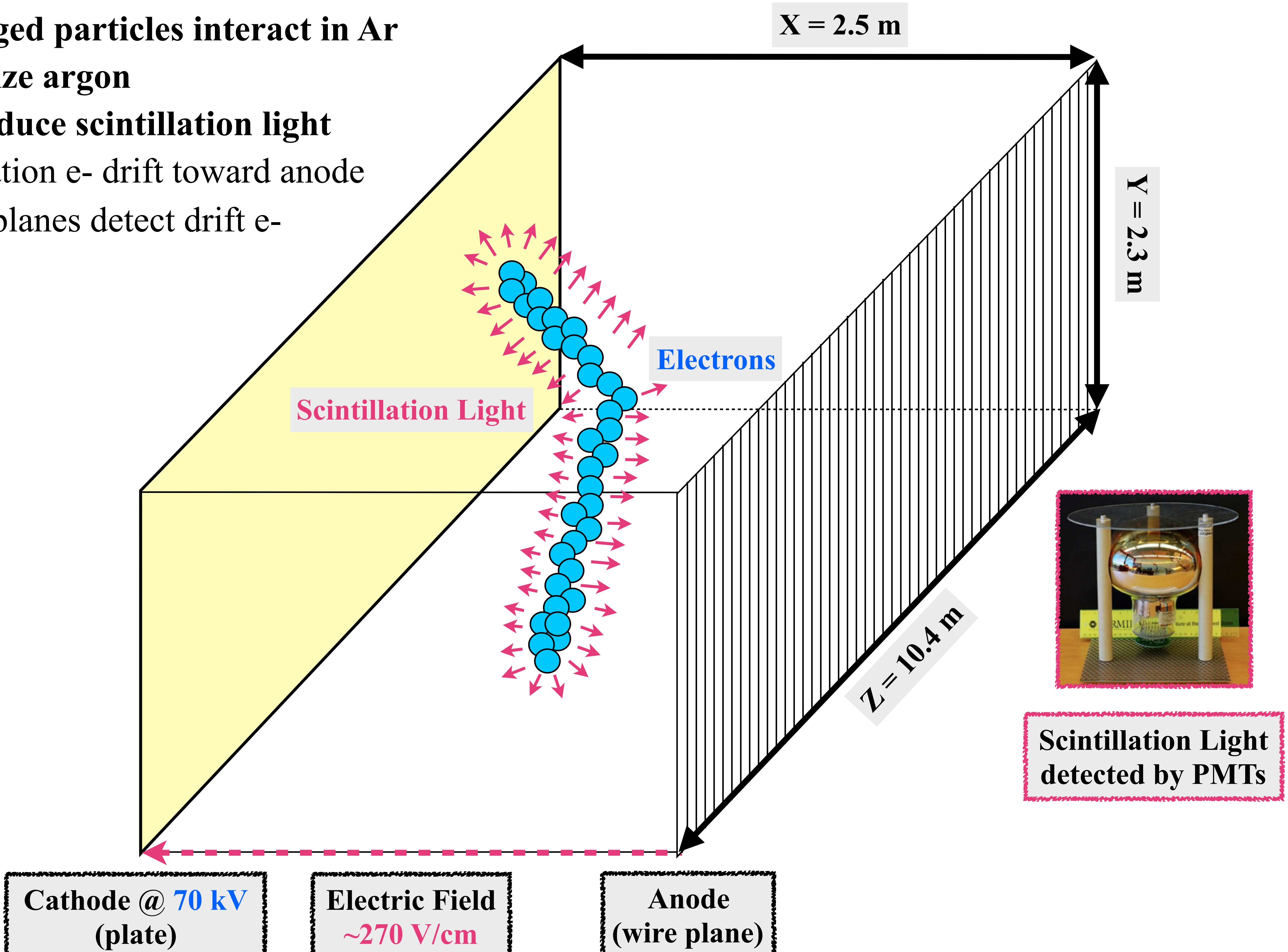
TPC Working Principle (I)

1. **Charged particles interact in Ar**
    - Ionize argon
    - Produce scintillation light
2. Ionization e- drift toward anode
3. Wire planes detect drift e-



TPC Working Principle (II)

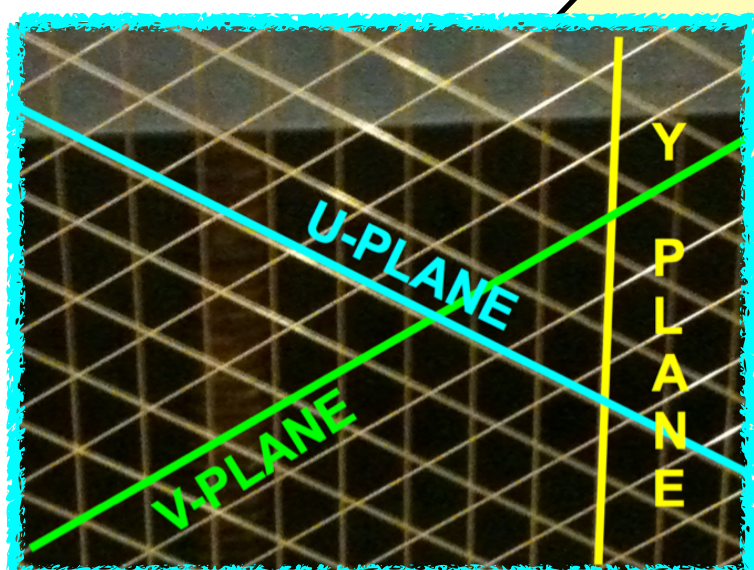
1. **Charged particles interact in Ar**
   - **Ionize argon**
   - **Produce scintillation light**
2. Ionization e- drift toward anode
3. Wire planes detect drift e-

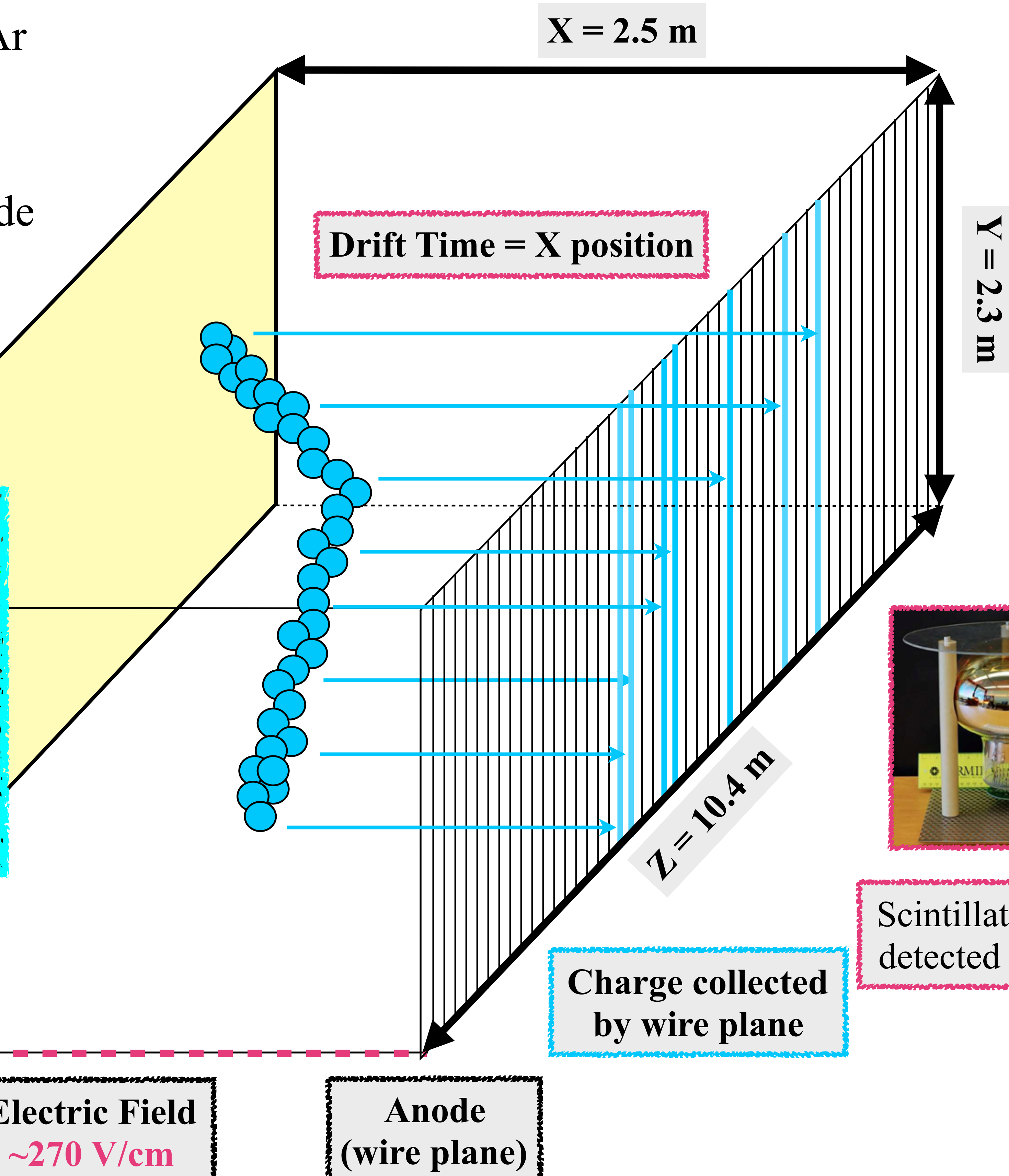


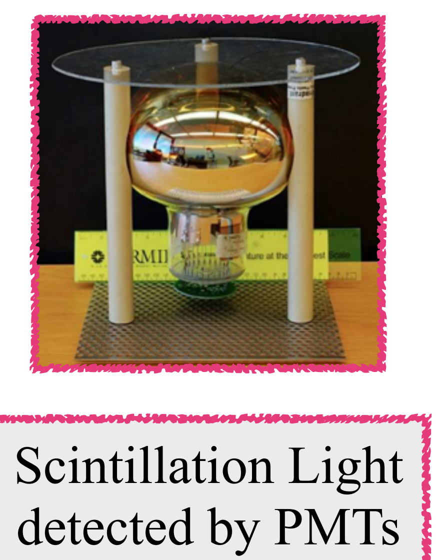
TPC Working Principle (IV)

1. Charged particles interact in Ar
   - Ionize argon
   - Produce scintillation light
2. Ionization e- drift toward anode
3. **Wire planes detect drift e-**

Three Wire Planes







MicroBooNE TPC & Cryostat

Anode Wire Plane

Cathode Plate

MicroBooNE TPC & Cryostat

Anode Wir

ode Plate

MicroBooNE TPC & Cryostat

What's Deep Learning?

Deep Learning … What & Why

What is ***Deep Learning***?

- A buzz word to gain attention from job recruiters

Dear Kazuhiro Terao,

Happy Friday! I wanted to reach out to you because my client,a Proprietary Global Market Maker, trading on major financial markets around the world, is looking for a Data Scientist to join their team. They are looking for an experienced Data Scientist to join their Chicago Office. They would like someone who has 3 + years of work or post-doc experience applying Reinforcement Learning or Deep Learning techniques and is proficient with Python or C.

Hey, lessons learned!

Deep Learning … What & Why

What is ***Deep Learning***?

- ~~A buzz word to gain attention from job recruiters~~
- Collective term for neural network (NN) architectures
  - Consists of large number of layers ***(deep)***
  - Breakthrough in computer vision (2012), now AI and more…
- It is a non-linear functional approximation
  - NN with millions of parameters to map *input* to *output* space

Not in the talk:
What's different from "traditional" NN?
How is it better?
(Feel free to ask later or during questions)

Deep Learning … What & Why

What is ***Deep Learning***?

- ~~A buzz word to gain attention from job recruiters~~
- Collective term for neural network (NN) architectures
  - Consists of large number of layers ***(deep)***
  - Breakthrough in computer vision (2012), now AI and more…
- It is a non-linear functional approximation
  - NN with millions of parameters to map *input* to *output* space

Why ***Deep Learning***?

- LArTPC image data analysis = feature recognition
- Explore the technique for reconstruction/analysis
  - 1st MicroBooNE collaboration paper on CNN