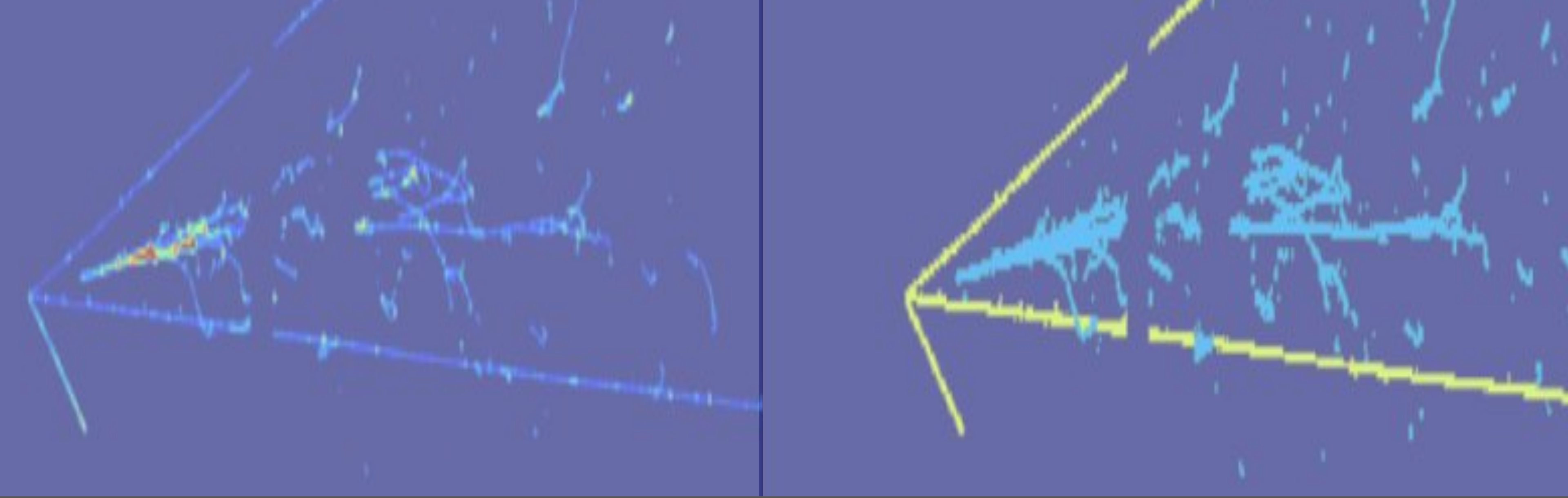


Machine Learning on MicroBooNE

T. Wongjirad for the MicroBooNE Collaboration

NNN18 Workshop

2018/11/01

Outline

- Present how MicroBooNE is using deep convolutional neural networks for reconstruction and analysis
- Focus on recent results on data
- Future directions

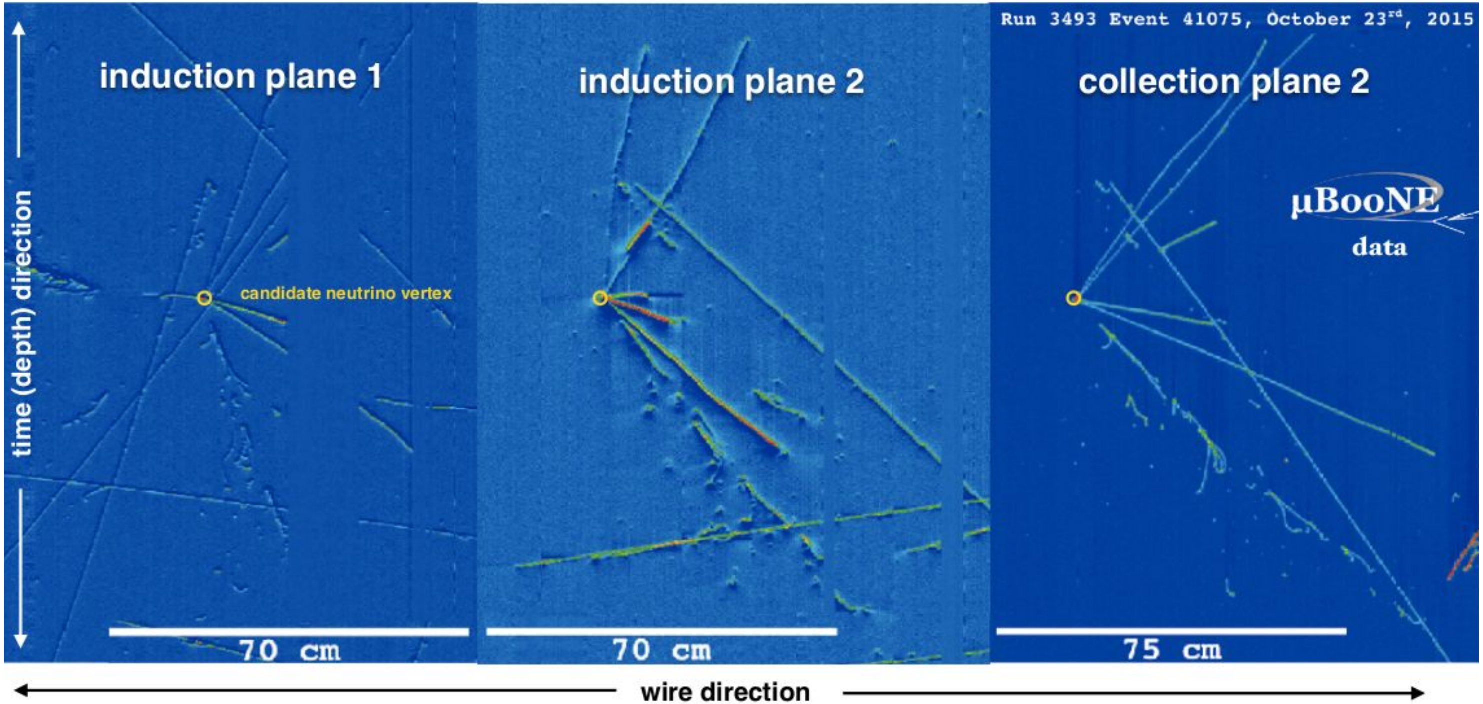
MicroBooNE Experiment

- MicroBooNE is a neutrino experiment using a liquid argon time projection chamber (LArTPC)
- Goals of MicroBooNE:
  - To investigate the MiniBooNE excess – to confirm or deny potential evidence for sterile neutrinos
  - To measure neutrino-argon cross sections around 1 GeV
  - R&D for future LArTPCs like DUNE

the detector in the pit during construction

MicroBooNE Data

Information about 3D trajectories encoded in a set of 3 2D images
Images are projections from wire planes: 2D → 3D not trivial



MicroBooNE Data

Information about 3D trajectories encoded in a set of 3 2D images
Images are projections from wire planes: 2D → 3D not trivial

For this talk, focusing on machine learning techniques on these images

Please see other MicroBooNE and other LArTPC talks for details on detector and how images are made

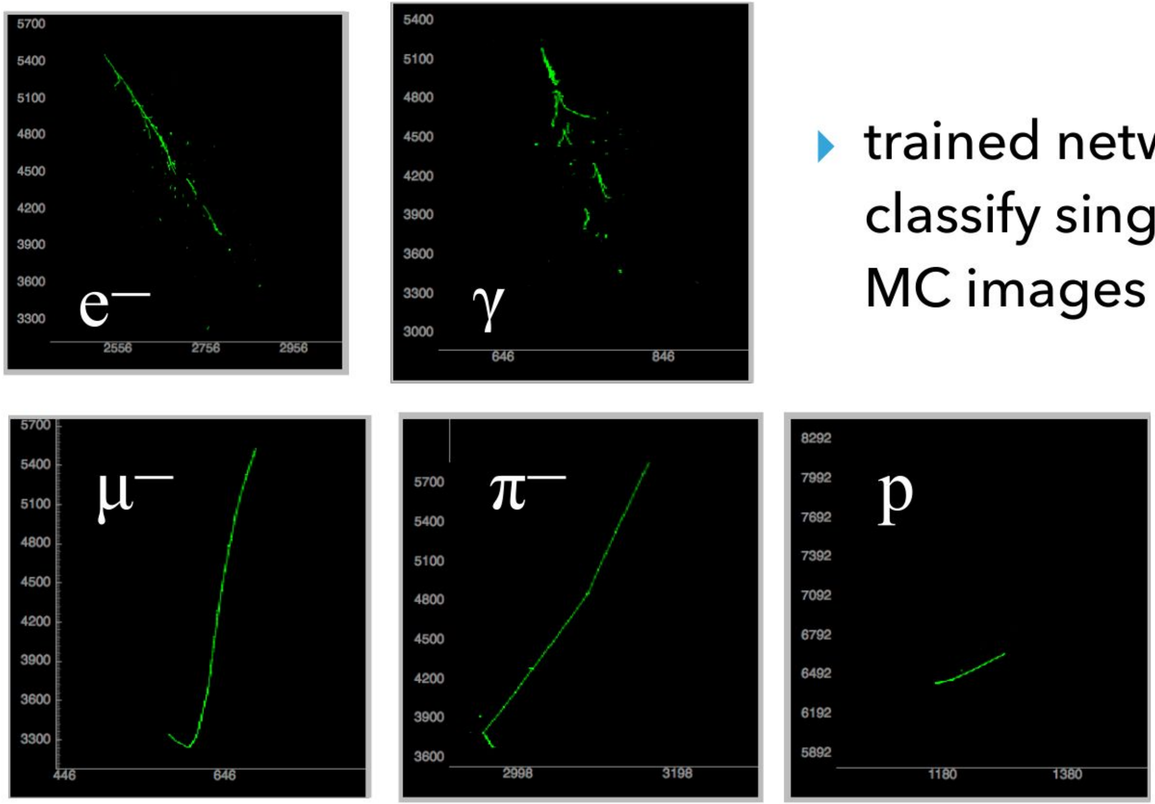
<u>MicroBooNE/Short-Basline Neutrino Program Talks</u>
B. Russel on SBN Program (Saturday Morning plenary)
S. Porzio on MicroBooNE analysis and systematics (Friday afternoon parallel)
J. Crespo-Anadon on Astroparticle physics on MicroBooNE (Friday afternoon parallel)
H. Rogers on ICARUS (Friday afternoon parallel)

First Studies for MicroBooNE

- Testing a number of different CNN tasks
  - Classification for single particles
  - Whole Event Classification as background or neutrino
  - Neutrino Detection Network -- draw a box
- Goals
  - Is there enough information for CNNs to work?
  - What is the performance level (on simulated images)?
  - Technical infrastructure (available to community on github)
    - Deeplearnphysics.org for tutorials and generic LArTPC datasets
    - github.com/deeplearnphysics/larcv2: library that handles image IO, physics-metadata, and interfaces to networks

MicroBooNE: JINST 12 (03), P03011

First Studies: single particle classifier



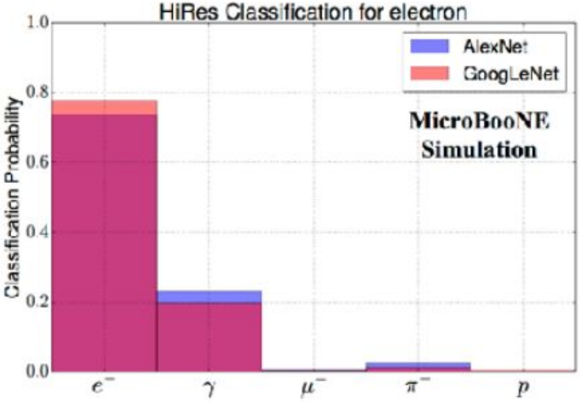
- trained network to classify single particle MC images

(simulated images)

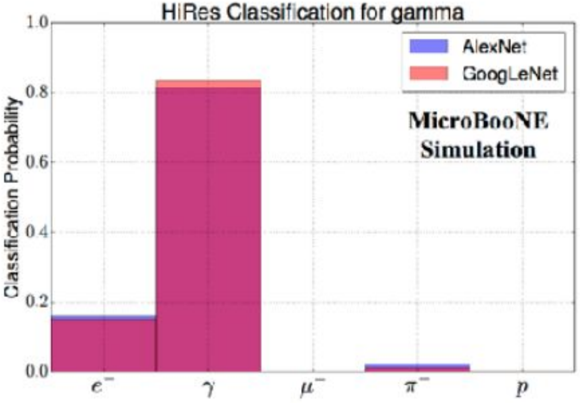
MicroBooNE: JINST 12 (03), P03011

First Studies: single particle classifier

- Trained the networks to categorize a simulated single particle
- Uniform position and momentum from 100 MeV to 1 GeV in isotropic direction



e⁻ classification performance

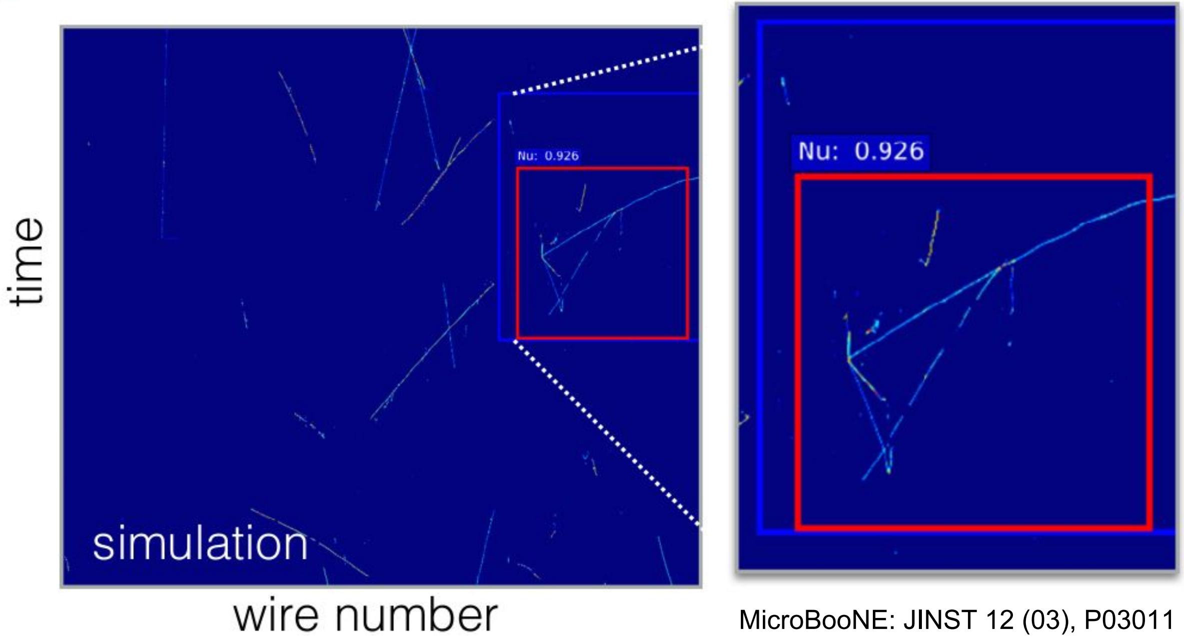


γ classification performance

MicroBooNE: JINST 12 (03), P03011

First Studies: Neutrino Detection

- Trained a network to place a bounding box around a neutrino interaction within a whole event view
- Note: cosmic off-beam data + MC neutrino overlaid



MicroBooNE: JINST 12 (03), P03011

Used faster-RCNN network
Ren, He, Girshick, Sun
NIPS2015

Simulating MicroBooNE Data

- With promising first results, started to build analysis around CNNs to find neutrinos
- Potential systematic issues strongly guided strategy
- For good simulated data for training, must produce
  - Model of how charge induces signals on the wires
  - Model of the electronics response
  - Model of the ionization
  - Model of the physics of final state particles particle physics (e.g. scattering, decay)
  - Model of the neutrino interactions on nuclei (e.g. NEUT, GENIE)

Low-level detector response

Physics of interest

Simulating MicroBooNE Data

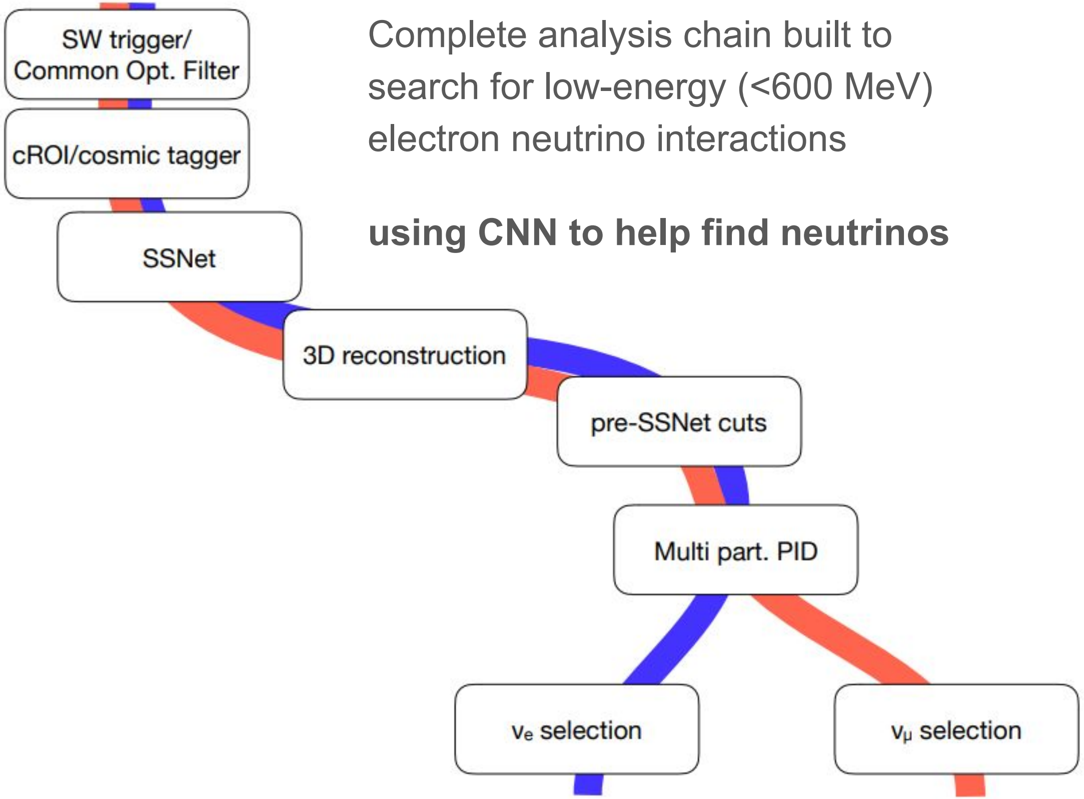
- With promising first results, started to build analysis around CNNs to find neutrinos
- Potential systematic issues strongly guided strategy
- For good simulated data for training, must produce
  - Model of how charge induces signals on the wires
  - Model of the electronics response
  - Model of the ionization
  - Model of the physics of final state particles particle physics (e.g. scattering, decay)
  - Model of the neutrino interactions on nuclei (e.g. NEUT, GENIE)

Low-level detector response

Physics of interest

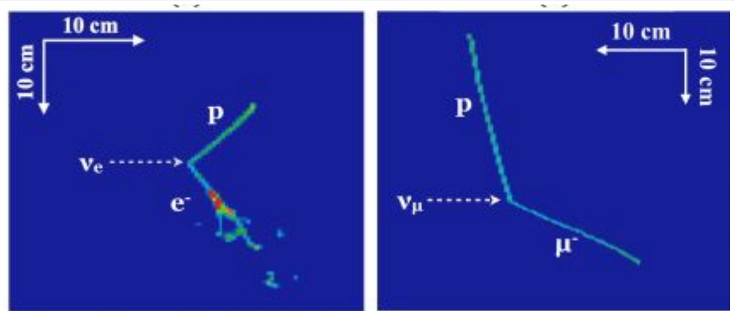
In principle, could target high-level objects (e.g. neutrino detection). But limited ourselves to topological features we can check with independent data sets (e.g. off-beam).
No near-detector (until SBN program) to provide neutrino interactions for validation

MicroBooNE DL Low Energy Nue Search



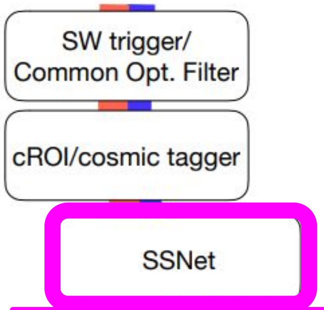
Complete analysis chain built to search for low-energy (<600 MeV) electron neutrino interactions

using CNN to help find neutrinos



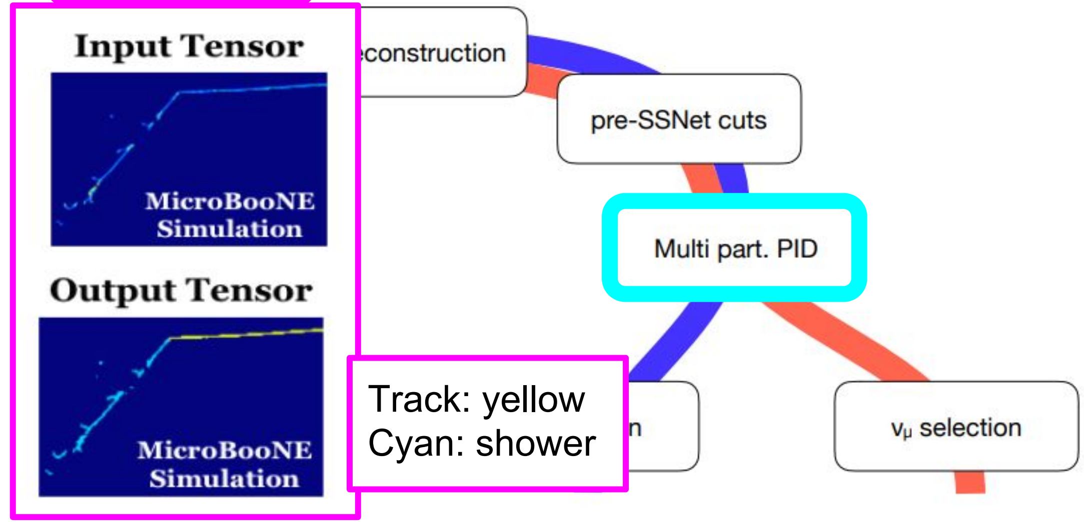
- **Target**: 1 lepton 1 proton final state
- Proton provides handle to more easily reject backgrounds
  (at cost of statistics, 45% compared to inclusive. At low energies, rest are almost all single electron events)

MicroBooNE DL Low Energy Nue Search



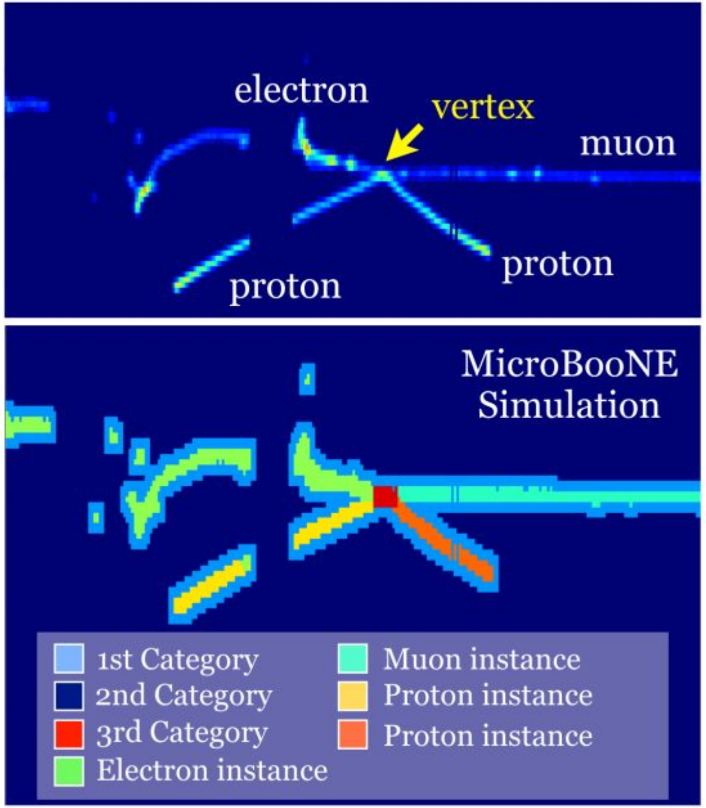
Complete analysis chain built to search for low-energy (<600 MeV) electron neutrino interactions

using CNN to help find neutrinos



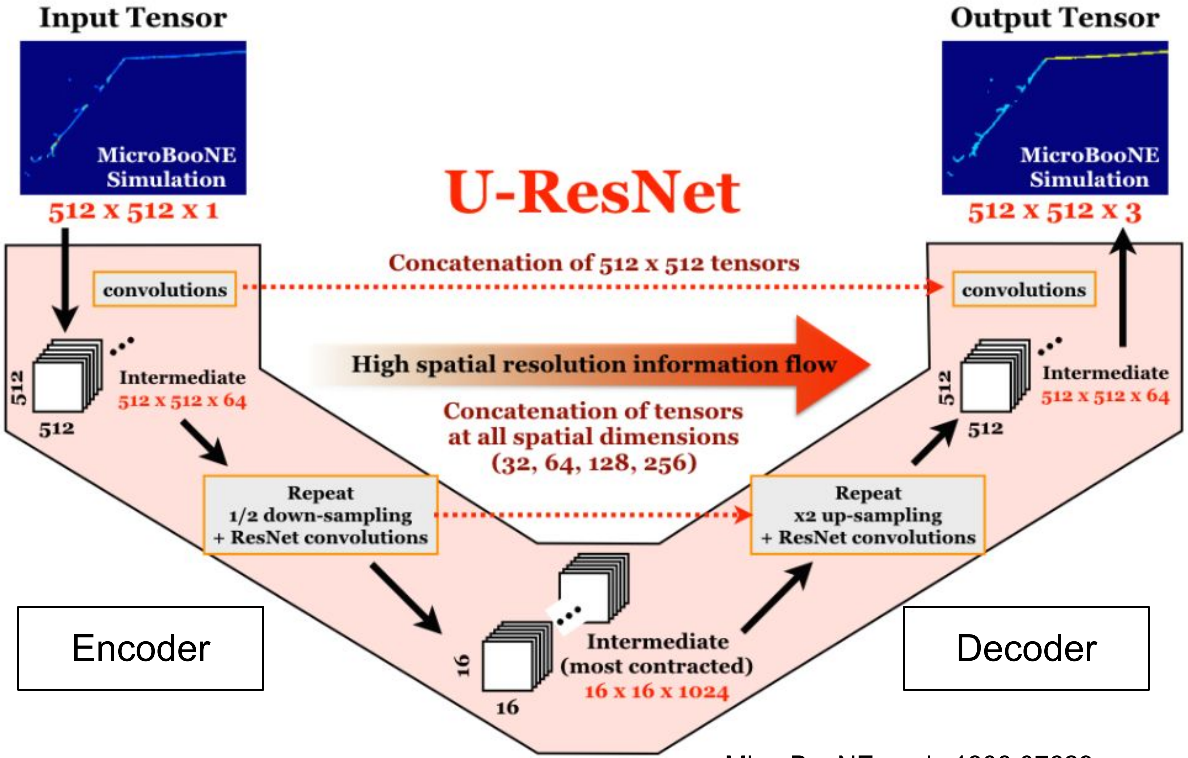
- CNNs applied in two places
    - **Track/shower pixel labeling**
    - Particle ID
- Neighboring track/shower cluster is vertex seed
- Rest of algorithms “traditional” algorithms
- Output we can produce independent samples for: no explicit neutrino-interaction predictions

Training Data



- Trained on “particle-bomb” events
- Uniform distribution in number of particles
- Uniform distribution of particle momenta
- Isotropic direction
- Pixel weighting important for training
- Goal is to ***overcover*** *signal sample*
  - Dealing with training domains an interesting topic
  - One way to handle described in paper by Minerva! (arXiv:1808.08332)

Network architecture



MicroBooNE: arxiv:1808.07629

Technical Description

- Auto-encoder with skip-connections: “U-Net” (Ronneberger, Fisher, Brox. arXiv:1505.04597)
- Using Residual Conv. layers (He at al. arXiv:1512.03385)

Short summary

- Two halves: encoder-decoder
- Encoder finds features
- Decoder projects back to original image resolution for pixel-level classifier

SSNet Performance

Sample	ICPF mean	ICPF 90%	Shower	Track
Test	1.9	4.6	4.1	2.6
ν_e	6.0	13.8	7.6	13.8
ν_μ	3.9	4.5	14.2	4.3
1e1p	2.2	5.7	2.8	4.0
1μ1p-LE	2.3	2.2	6.2	2.4
1e1p-LE	3.9	11.5	3.8	8.0

MicroBooNE: arxiv:1808.07629

- Quantify performance using **ICPF:** incorrect pixel fraction (per image)
- ICPF average per image are few percent of pixels
- Performance various over signal domain -- as one might expect
- Sufficient performance for 1L1P search

Training Data

Sample	ICPF mean	ICPF 90%	Shower	Track
Test	1.9	4.6	4.1	2.6
ν_e	6.0	13.8	7.6	13.8
ν_μ	3.9	4.5	14.2	4.3
1e1p	2.2	5.7	2.8	4.0
1μ1p-LE	2.3	2.2	6.2	2.4
1e1p-LE	3.9	11.5	3.8	8.0

- Quantify performance using **ICPF:** incorrect pixel fraction (per image)
- ICPF average per image are few percent of pixels
- Performance various over signal domain -- as one might expect
- Sufficient performance for 1L1P search

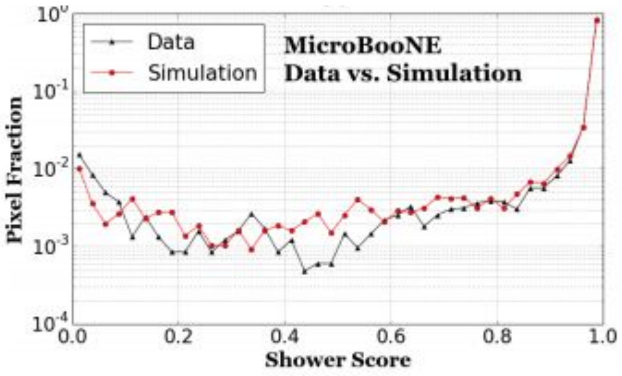
What about data versus MC behavior?

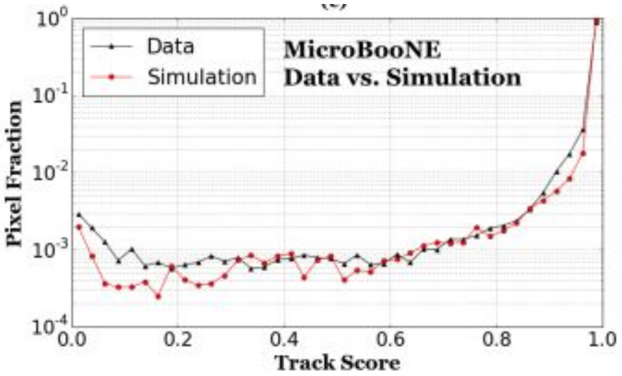
SSNet data versus MC tests

- Real data do not have truth labels, of course
- Measure indistinguishable data vs MC behavior between human versus network -- a kind of Turing test

- Two tests (beyond comparing various distributions for data versus MC)
  - Measure disagreement in labels for stopping muons and CC numu pi0 events selected using completely independent pattern recognition algorithm (Pandora)
  - Qualitative look at how scores changes when parts of images removed -- looking to see that network output behaves in ways similar to human analyzer
  - **Work published in “A Deep Neural Network for Pixel-Level Electromagnetic Particle Identification in the MicroBooNE Liquid Argon Time Projection Chamber” arxiv:1808.07629**

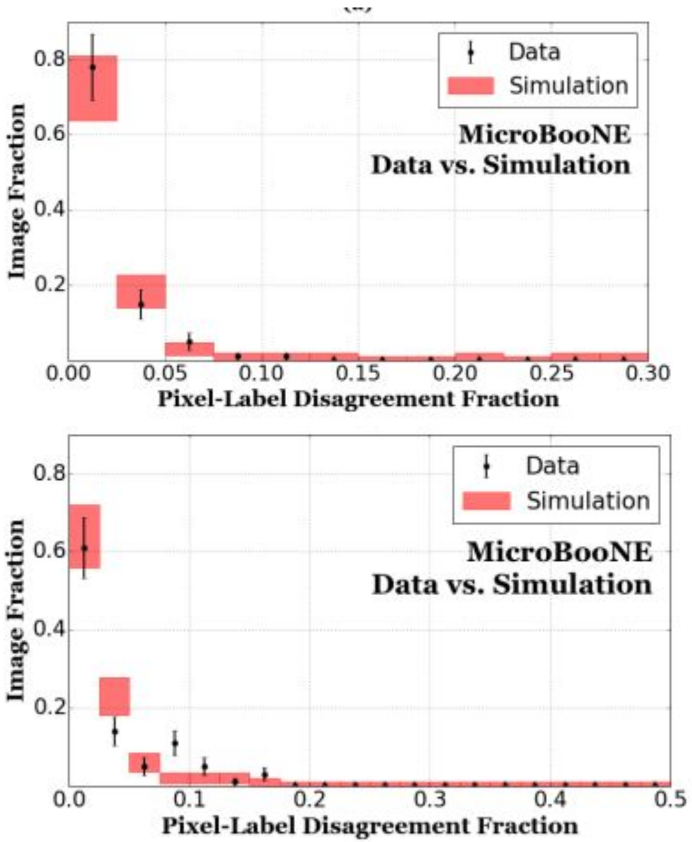
SSNet data versus MC score distributions





- Sample: stopping muons (tests on numu CC pi0 events similar result)
- Score distributions similar
- Mean ICPF at the few percent level

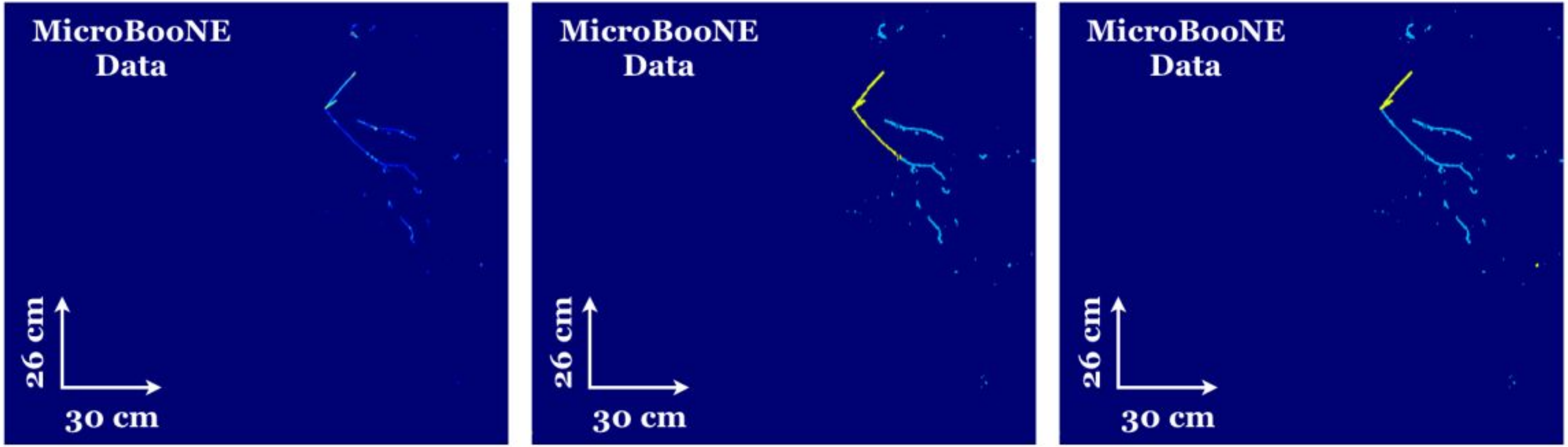
Human-Network Disagreement



- Pixel-level disagreement rate for data versus MC
- Comparing network versus human
- For both stopping muons and numu CCPi0 sample, disagreement rate similar in data and MC

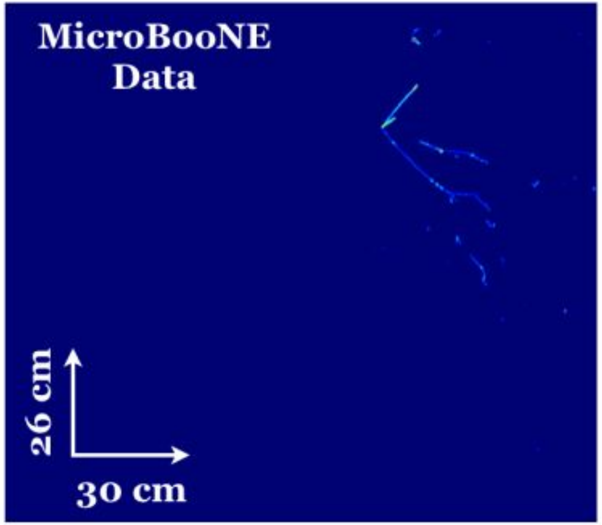
Human-Network Disagreement

Input

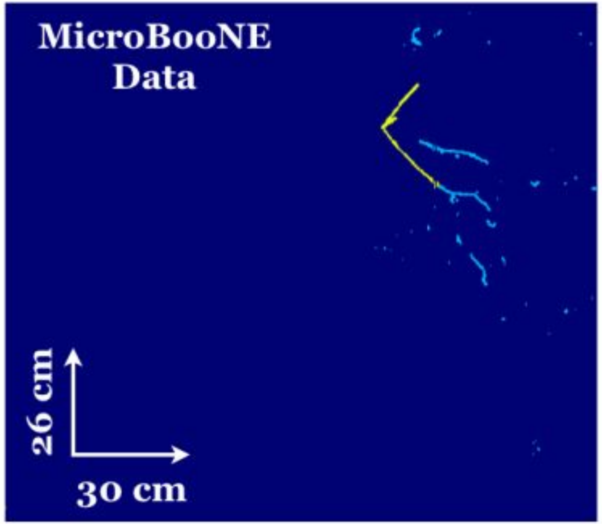


Human-Network Disagreement

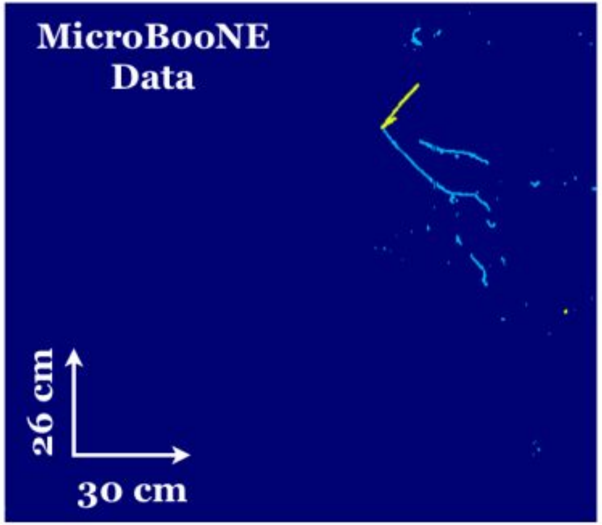
Input



Physicist



Network

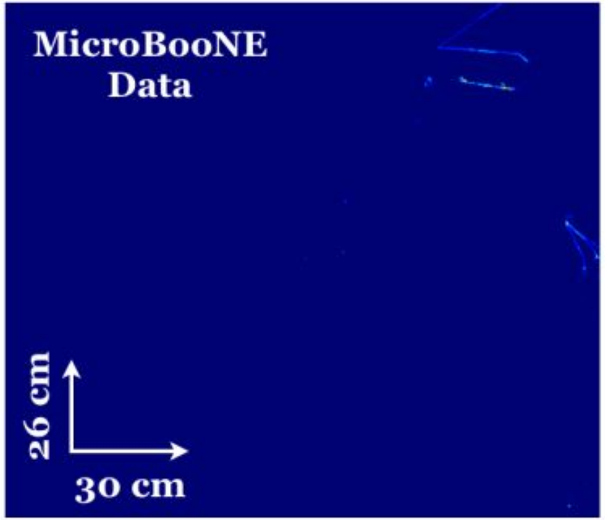


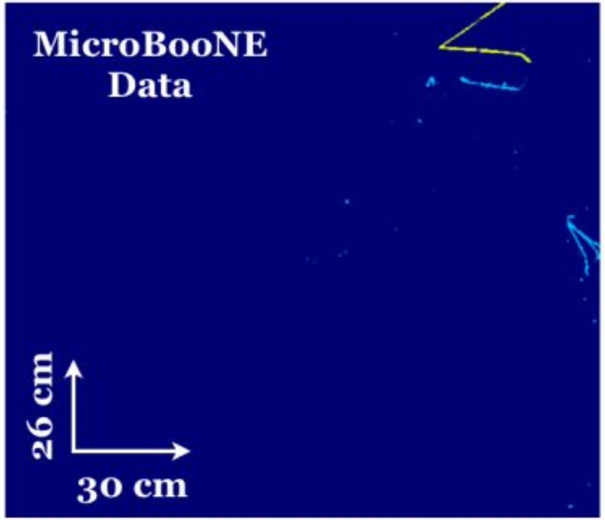
Human-Network Disagreement

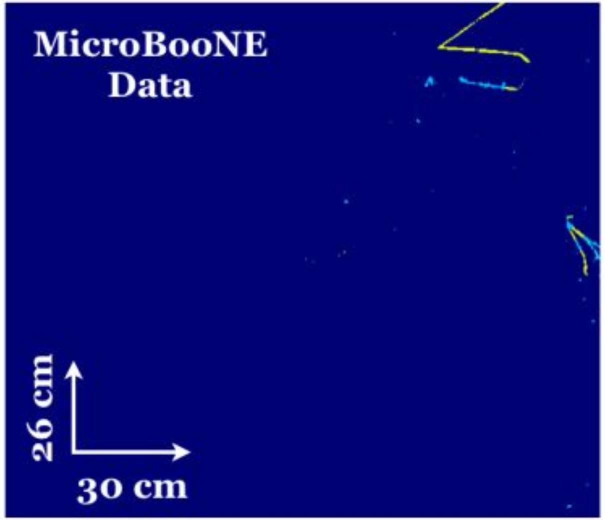
Input

Physicist

Network

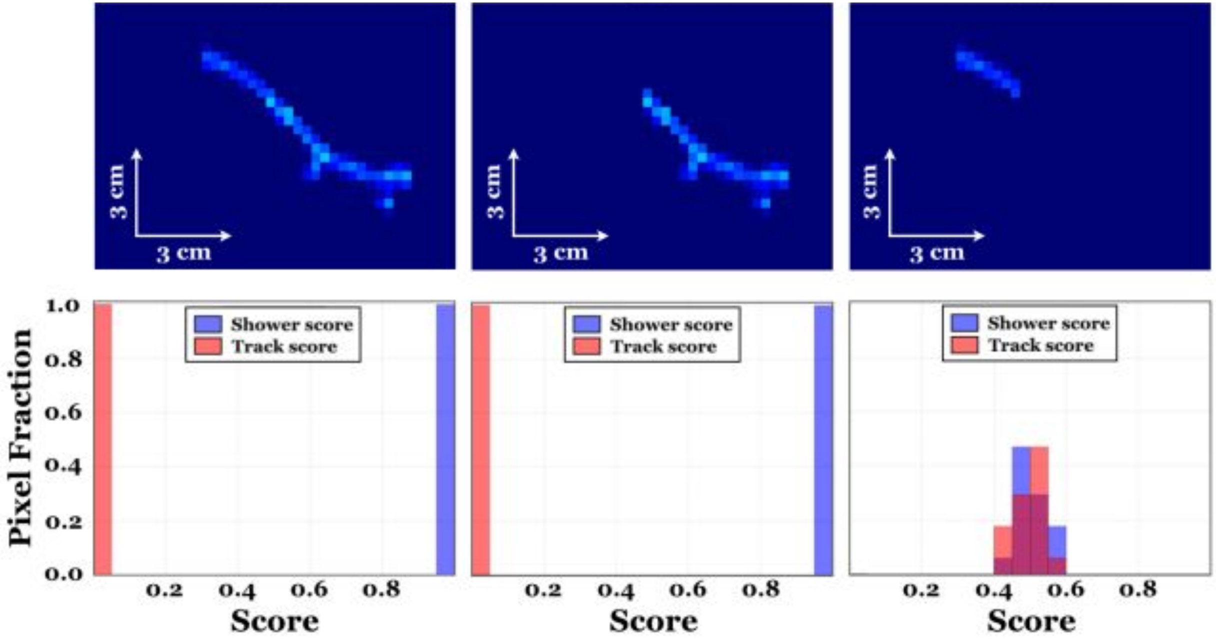






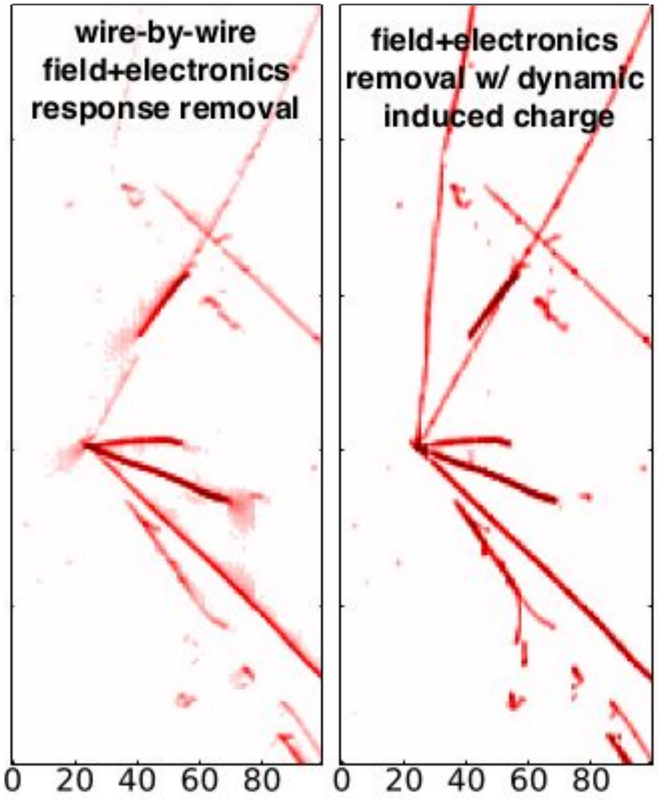
Qualitative Behavior Tests

- One example: As shower reduced to low-energy line, network becomes less certain



Collection vs. Induction Planes

- Study gives us confidence networks trained on MC will behave similarly on data -- ***for the collection plane***
- We do not use the induction plane to make selection cuts as network behaves differently
    - Clear data and MC feature differences seen
    - However, vast improvements in induction plane quality using “2D deconvolution”, which accounts for signals induced on neighboring wires **(MicroBooNE: arxiv:1804.02583,1802.08709)**



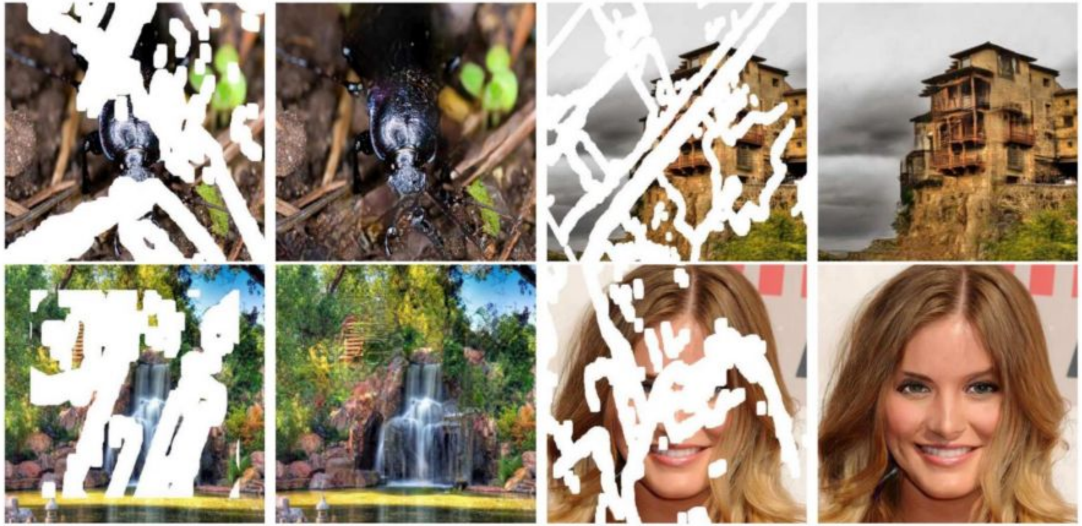
Future Directions

- Pursuing new applications
  - Extend SSNet labels to more useful features, e.g. track ends, shower starts, noise
  - Instance-aware semantic segmentation: combines clustering and particle ID
  - Predict trajectories in dead regions
  - Reconstruct 3D charge deposition from 2D image
- Systematic studies
  - Qualitative visualization of the features that correlate with activation
  - Closer look at domain dependency

Inferring cosmic muon trajectories in dead regions

NVIDIA has created a CNN to fill in dead region in pictures.

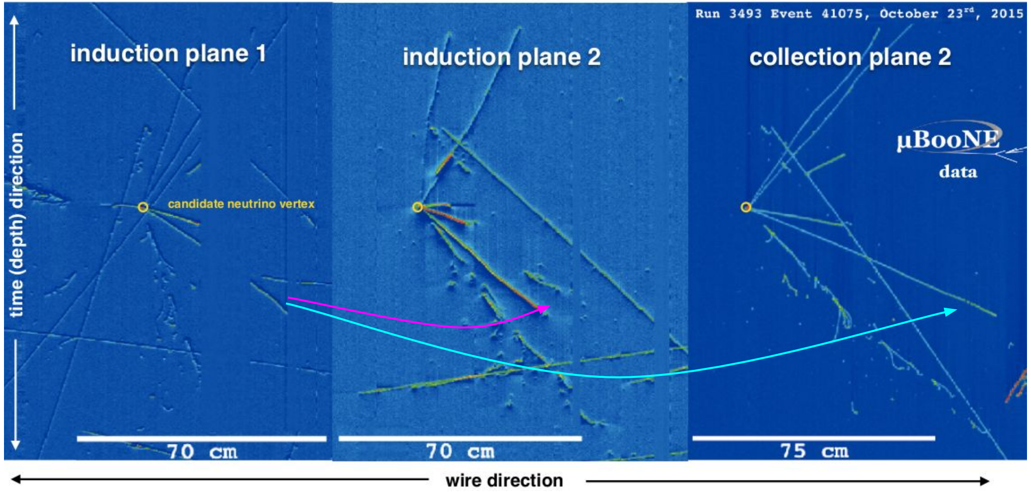
http://arxiv.org/pdf/1804.07723.pdf NVIDIA Corporation



- About 10% of MicroBooNE wires are unresponsive (but 3-plane redundancy means that only few percent of detector is unreconstructable
- Use network to project muon trajectories in dead regions
- Target is for cosmic muon tagging (note: only for position, not for calorimetry)

3D Space-points from Feature Correspondence

- Matching features across planes = locating 3D position
- Have two correspondences (per starting plane)
- Predictions should be 3D consistent, incorporated into loss function
- Early development looks promising



Zhou, Krähenbühl et al. “Learning Dense Correspondence via 3D-guided Cycle Consistency” CPVR 2016

Conclusions

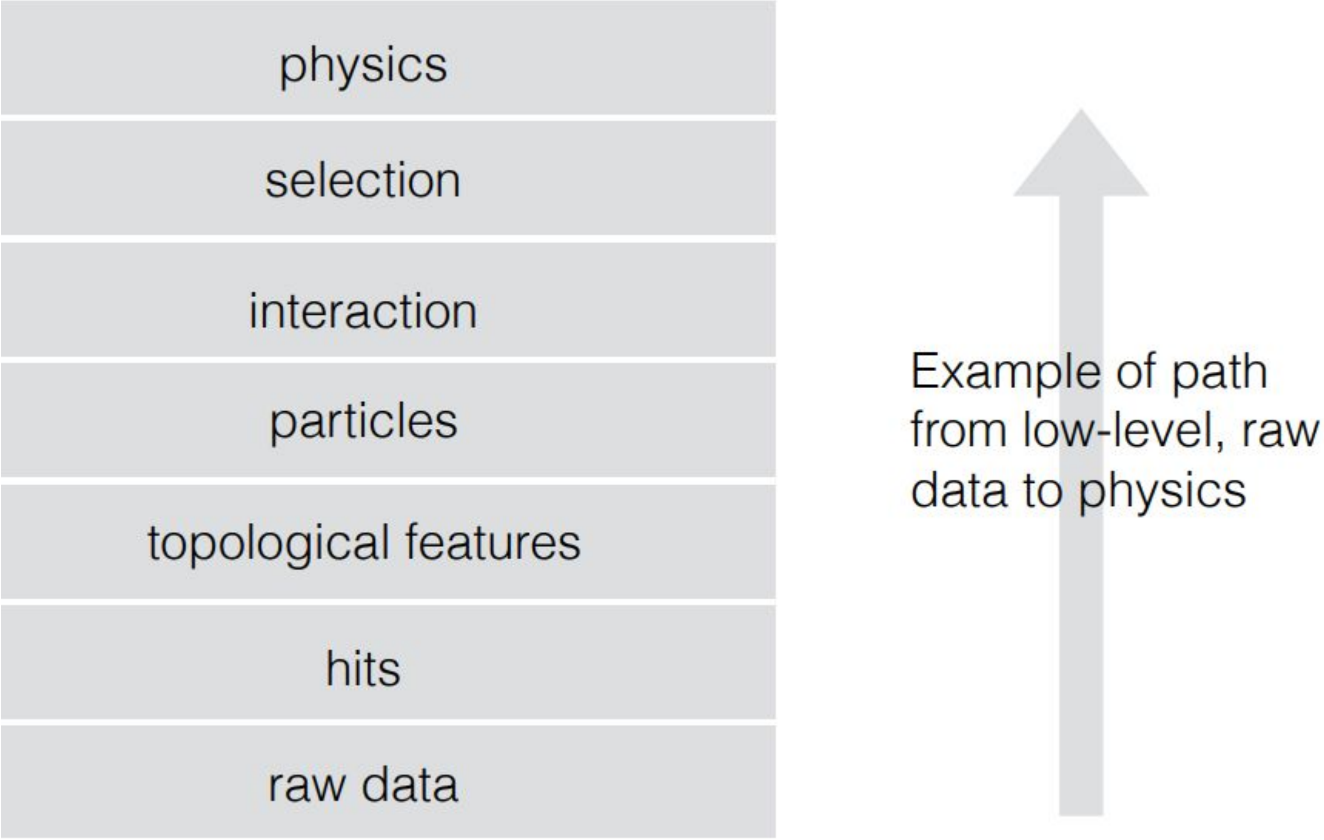
- MicroBooNE is using Deep Convolutional Neural Networks for physics analyses
- Started with a low-level topological feature finder: *track versus shower*
- Passed important milestone confirming network behavior does not diverge significantly on data and MC on the collection plane
    - Studies on quantities more directly related to the analyses are on-going
- More network applications are being developed
    - Optimistic that major improvements in wire signal processing and simulation will support new applications <u>on all three planes</u>
    - Still avoiding explicit targeting of neutrino-interaction level information (for now)

Backups

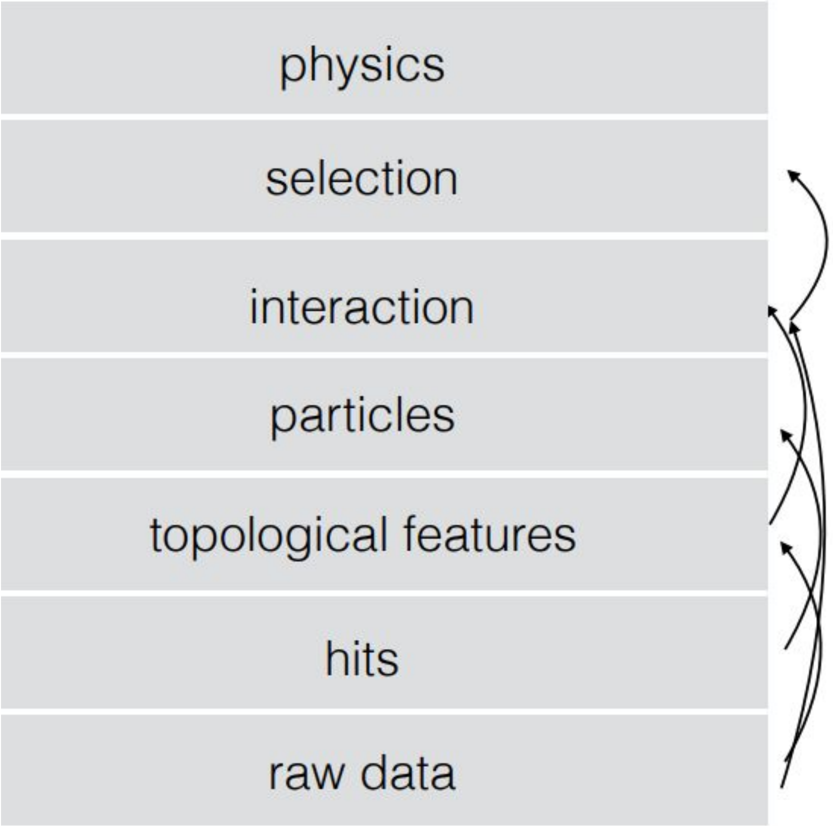
Semi-supervision

- Much development in DL research towards semi-supervision
- ***Does not require truth-labeled data***
    - No longer have to worry about simulated versus real data behavior difference
- Semi-supervised learning has network using data to reproduce data
- Can use certain constraints to tell networks what data features to learn
- Can we incorporate data into training?
    - can network learn features only on data (e.g. through image completion task), then use these features for typical fully-supervised tasks utilizing simulation?
    - Can we first train using labeled simulated data, then fine tune network on data only given self-consistency constraints (plane correspondence task)
    - Provide constraints on features learned

Hierarchy of reconstruction/analysis products



Having simulations: a double-edged sword



Today's techniques still heavily rely on large training data, which provides examples of the correlations you want to learn

HEP has an **advantage** here. Our simulations can produce almost unlimited simulated training data — connect any level of information we want

But simulation data built on models, which much be sufficiently correct