Multi-Particle PID CNN Performance on Low Energy Events

Andrea Herman, MIT/Nevis Labs
Samantha Sword-Fehlberg, NMSU
Kazuhiro Terao, SLAC
Erica Snider, Fermilab
July 28th, 2017
How does CNN perform on low energy $1e1p$ events?

1. Can network properly identify image contents?

2. Are network failure cases grounded in legitimate PID challenges?
The Low Energy 1e1p Test Events

- 50,000 images containing exactly one proton and one electron generated at the event vertex
- Vertex location generated uniformly in the detector
- Isotropic momentum
- Electron kinetic energy $[30,100]$ MeV and proton kinetic energy $[40,100]$ MeV.
Can network correctly label particles that are present?

For each image, network returns a score between 0 and 1 for each particle class it knows: [e, γ, μ, π, p]

Majority of the time, network is very certain that electron/proton is present in the image (and it is!)
Can network correctly label particles that are present?

For each image, network returns a score between 0 and 1 for each particle class it knows: [e, γ, μ, π, p]

Majority of the time, network is very certain that electron/proton is present in the image (and it is!)
Network labels for absent particles illuminate failure cases

No $\gamma$, $\mu$, or $\pi$ are present in test images, and majority of network labels reflect this...
Network labels for absent particles illuminate failure cases

No $\gamma$, $\mu$, or $\pi$ are present in test images, and majority of network labels reflect this…

But network is occasionally very certain about $\gamma$, $\mu$ false positives.
What cases cause this misidentification?
Examining Muon False Positives
Examining the cause of high $\mu^-$ scores:

Network only gives a high muon score when electron score is low:

electron is being **mistaken** for muon
Examining the cause of high $\mu^-$ scores:

Network only gives a high muon score when electron score is low:

electron is being mistaken for muon

Could this be low energy electrons, which appear more track-like?
Examining the cause of high $\mu^-$ scores:

False positive muon scores are caused by lowest energy electrons.
An exercise for the reader... Can you find the true $\mu^-$ event display?
An exercise for the reader... Can you find the true $\mu^-$ event display?
Examining Gamma False Positives
Relationship between electron and gamma scores shows three event populations
Relationship between electron and gamma scores shows three event populations

Network successfully identifies electrons (majority case)
Relationship between electron and gamma scores shows three event populations

1. Network successfully identifies electrons (majority case)

2. High γ scores correlated with low electron scores – network mistakes electrons for gammas.
Relationship between electron and gamma scores shows three event populations

2. High $\gamma$ scores correlated with low electron scores – network mistakes electrons for gammas.

3. Network is barely able to distinguish electrons from gammas – there is some confusing factor

1. Network successfully identifies electrons (majority case)
Relationship between electron and gamma scores shows three event populations

2. High $\gamma$ scores correlated with low electron scores – network mistakes electrons for gammas.

3. Network is barely able to distinguish electrons from gammas – there is some confusing factor.

1. Network successfully identifies electrons (majority case)
Examining the cause of γ false positives:

Very high γ scores are correlated with low electron scores – **network mistakes electrons for gammas**.

Let’s study this mistaken population:

Gamma score > 0.95
Electron score < 0.05

(90 events in total)
Examining the cause of γ false positives:

Studying low electron/high gamma score images shows three types of network mistakes.

Electron irradiates very early – image **does** contain a gamma!

Proton and electron are co-linear, mimicking a gamma.

Other

There are still some unexplained mistakes, but most high gamma scores cases can be understood as the network facing difficult PID choices.
Nine sample events with high gamma score and low electron score

- Majority $\gamma$ pixels
- Co-linear particles
- Other
Examining the cause of γ/e confusion:

Network is sometimes gives images **moderate scores for both electron and gamma.**
Examining the cause of γ/e confusion:

Network is sometimes gives images moderate scores for both electron and gamma.

If the electron appears as a shower, one of few ways to differentiate it from a gamma is \( \frac{dE}{dX} \) at the start of the shower.

Is this confusion due to angle with respect to wire plane increasing \( \frac{dQ}{dS} \)?
Examining the cause of γ/e confusion:

Events with gamma/electron confusion more frequently have electron angles **parallel to the wire plane** – high $dQ/dS$ creates gamma-like signal.
Wrapping up....

In a majority of cases, CNN is able to correctly answer the question: **What particles are in this image?**

Network failure cases are primarily due to legitimate PID challenges.

Next steps:
- Study performance on general $\nu_\mu$ and $\nu_e$ events
- Real data?!

To follow our footsteps in network architecture development, go to our github repository!
Acknowledgements

• Dr. Kazuhiro Terao, SLAC
• Prof. Michael Shaevitz, Columbia University
• Prof. Georgia Karagiorgi, Columbia University
• Nevis Labs REU Program & National Science Foundation
• Samantha Sword-Fehlberg, NMSU
• Dr. Erica Snider

Thank you!
In a majority of cases, CNN is able to correctly answer the question: **What particles are in this image?**

Network failure cases are primarily due to legitimate PID challenges.

Next steps:
- Study performance on more general numu and nue events
- Real data?!

To follow our footsteps in network architecture development, go to our github repository!
Additional Slides
DL Tensorflow Toolkit on Github!

Tensorflow is an open-source DL software package. On github – a toolkit to learn Tensorflow machine learning applications:

- Python-based image generator
- Example scripts for building and training deep neural networks in Tensorflow
- More documentation pending!

Sample images: Variable shape type and multiplicity available for development of different network architectures.
Network uses the **sigmoid regression algorithm** for multi-particle PID

**Sigmoid (Binomial Logistic) Function**

For each particle type:

\[ P(y = 1) = \frac{e^{(b+mx)}}{1 + e^{(b+mx)}} \]

Score correlating to **whether or not particle is present**.

Network-learned weights, biases and features

**Score 1**: Particle type is present

**Score 0**: Particle type is absent
Network uses the **sigmoid regression algorithm** for multi-particle PID

**Sigmoid (Binomial Logistic) Function**

For each particle type:

\[ P(y = 1) = \frac{e^{(b+mx)}}{1 + e^{(b+mx)}} \]

Score correlating to whether or not particle is present.

Network-learned weights, biases and features

Answers the question: **which particles are in this image?**

**Does not require exclusivity of particle types in image** – a strength over previous network.
Network Training

Train network using Tensorflow open-source DL software

**Training Data: Multi-Particle Generator**
- One vertex per event, random location in TPC
- Random multiplicity between 1 and 4
- Isotropic momentum
- Five particle types: \([e, \gamma, \mu, \pi, p]\)
Proof-of-Principle: **Network can identify the contents of an image!**

An example of Sigmoid regression on toy data:

Introduce four simple shape classes: Network returns an array of scores for each shape type.

Network Scores: [0.999, 1., 0.004, 0.998]

Network can identify which shape types are present and absent.
The Big Picture: where is PID in MicroBooNE LEE event reconstruction?

Event Selection:
1. PMT precuts
2. Cosmic Rejection

Interaction Reconstruction:
3. Track/Shower ID
4. Vertex Reconstruction
5. Particle Identification (PID)

The PID Goal:
Use a convolutional neural net (CNN) to identify multiple particle types in an image
Orientation of electron track with respect to wire plane can be deceptive for $dE/dX$ reconstruction:

A particle will have a higher $dQ/dS$ on the wire plane if it travels parallel to the wires.

Is the network seeing electrons with high $dQ/dS$ and mistaking it for gammas with high $dE/dX$?

Particles have the same $dX$ but very different $dS$. 