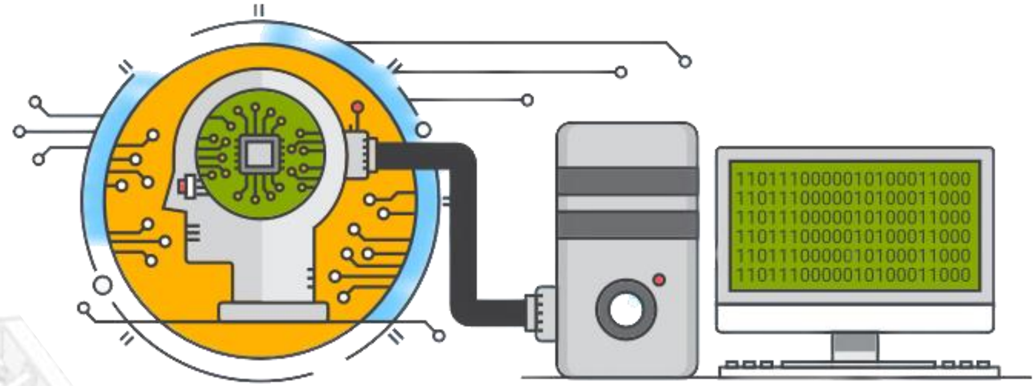
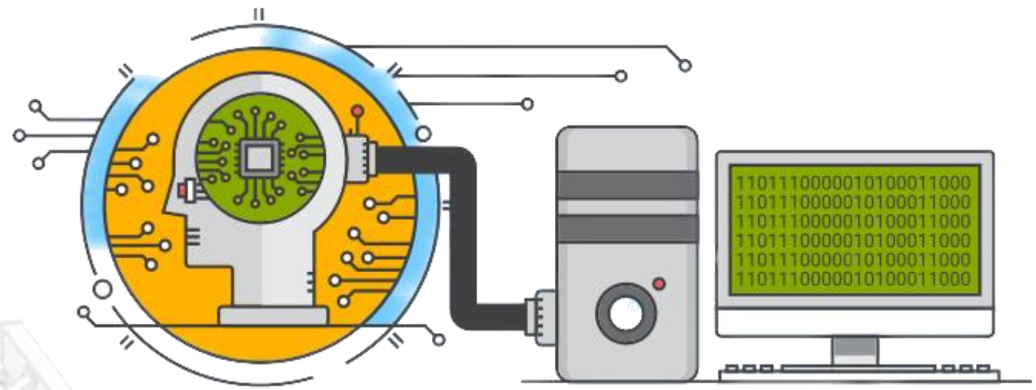
NPTwins, Genova, 16-18 December 2024



# Artificial Intelligence in low-level data reconstruction

Raffaella De Vita (Jefferson Lab)
for the CLAS Collaboration

NPTwins, Genova, 16-18 December 2024



# Artificial Intelligence in low-level data reconstruction at CLAS12

Raffaella De Vita (Jefferson Lab)
for the CLAS Collaboration

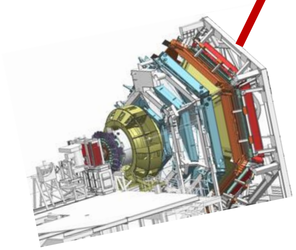
# Introduction & Outline

- In recent years, the use of AI/ML tools in our field has grown progressively, with applications in
  - Simulations
  - Detector design
  - Accelerator operation
  - Detector monitoring and operation
  - Event reconstruction
  - Data analysis
  - Data preservation
  - …


- The CLAS12 experiment at Jefferson Lab has been leveraging AI/ML techniques to enhance its performance, from online data-taking, to offline reconstruction and data analysis
  - Charged particle tracking in high-background conditions to increase detection efficiency and allow high-luminosity operation
  - Fast online event reconstruction for highly selective software trigger
  - Real-time detector monitoring and fault identification
  - Signal-background separation in physics analysis

- A few notes:
  - I am not an AI expert…
  - Results based on the work of many within the CLAS Collaboration and Jlab staff
  - Thanks to G. Gavalian for the presentation material

# Jefferson Lab @ 12 GeV

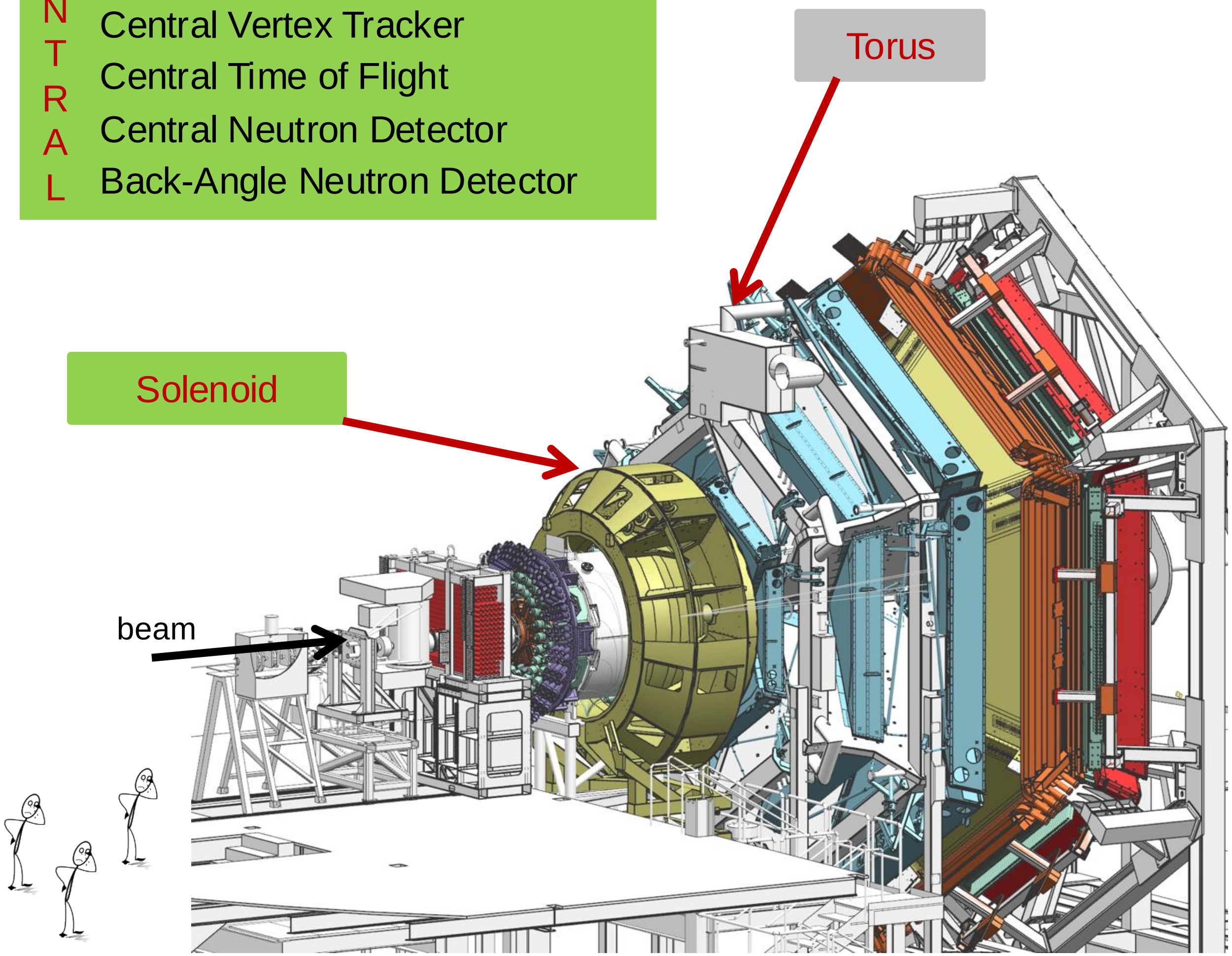



- Accelerator Upgrade completed in September 2017
  - CW electron beam
  - $E_{max}$ = 12 GeV, $I_{max}$ = 90 mA, $Pol_{max}$ ~ 90%
- Physics Operation
  - 4 halls running simultaneously since January 2018

# CLAS12

C Beamline
E Target
N Central Vertex Tracker
T Central Time of Flight
R Central Neutron Detector
A Back-Angle Neutron Detector
L

F High Threshold Cherenkov
O Forward Tagger
R Drift Chambers
W Low Threshold Cherenkov
A Ring Imaging Cherenkov
R Forward Time of Flight
D EM Calorimeter

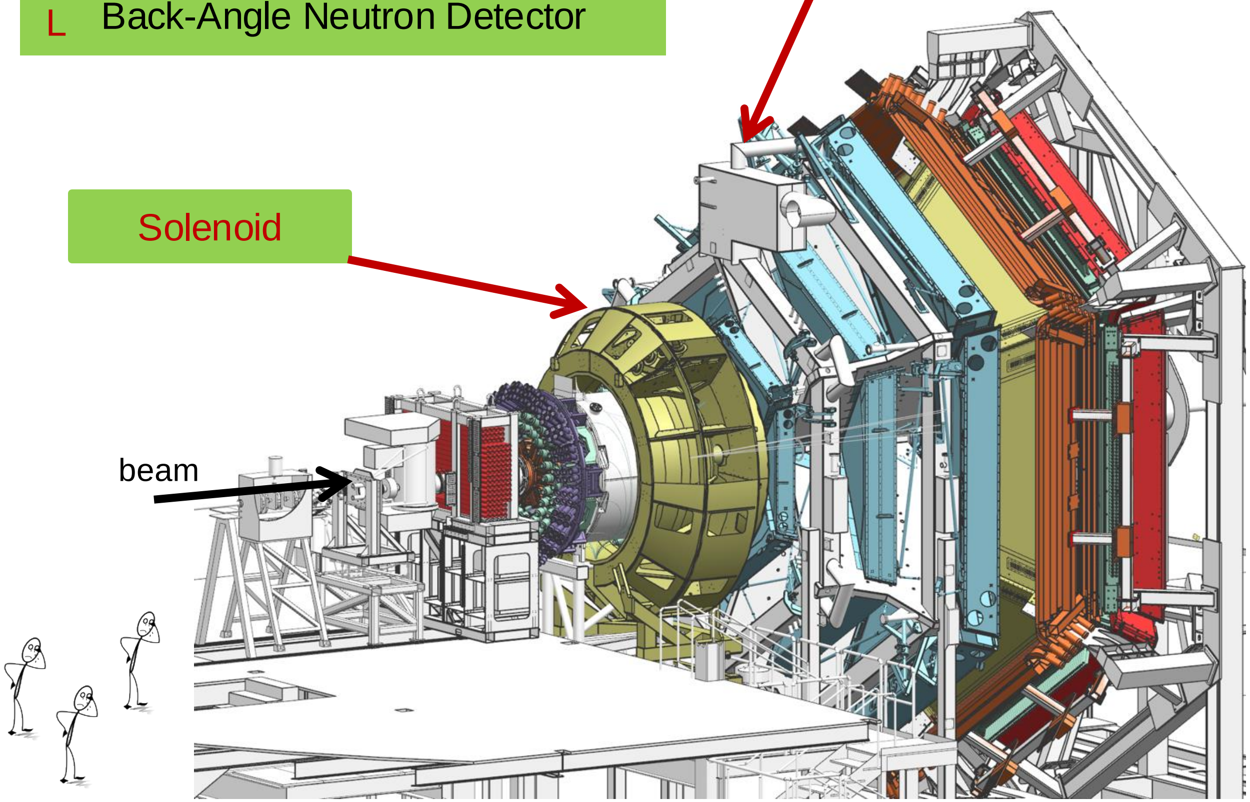



| | Forward | Central |
|---|---|---|
| Angular coverage | 5º – 35º | 35º – 135º |
| Momentum resolution | dp/p < 1% | dp/p < 5% |
| θ resolution | 1 mrad | 5 – 10 mrad |
| ϕ resolution | 1 mrad/sinθ | 5 mrad/sinθ |

# CLAS12

| | |
|---|---|
| C | Beamline |
| E | Target |
| N | Central Vertex Tracker |
| T | Central Time of Flight |
| R | Central Neutron Detector |
| A | Back-Angle Neutron Detector |
| L | |

| | |
|---|---|
| F | High Threshold Cherenkov |
| O | Forward Tagger |
| R | Drift Chambers |
| W | Low Threshold Cherenkov |
| A | Ring Imaging Cherenkov |
| R | Forward Time of Flight |
| D | EM Calorimeter |



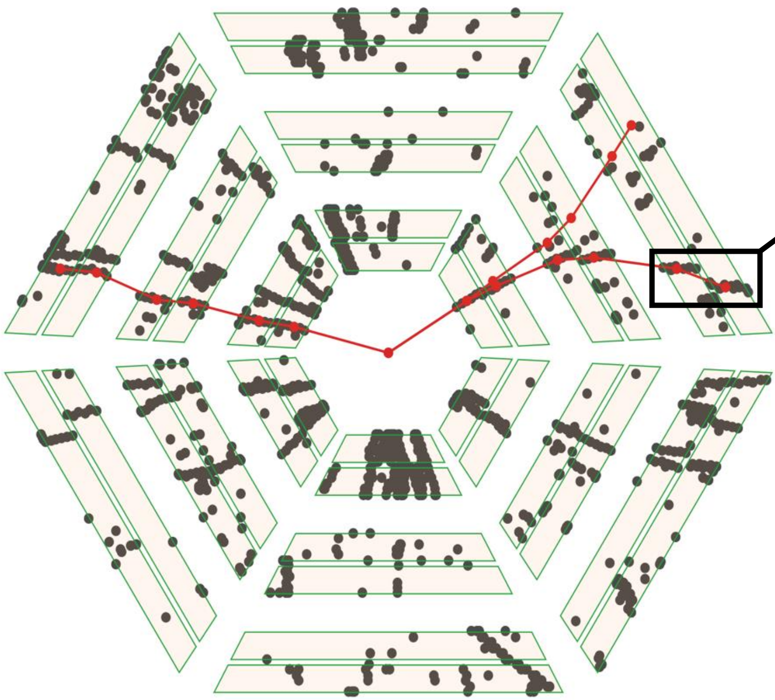

- Readout channels >100000
- Luminosity $10^{35} cm^{-2} s^{-1}$ limited by detector occupancy due to beam-related background
- Trigger rate up to 25 kHz (>> rates of reactions of interest)
- Data rate ~500 MB/s
- Data size ~1 PB/y
- Large acceptance for both charged and neutral particles

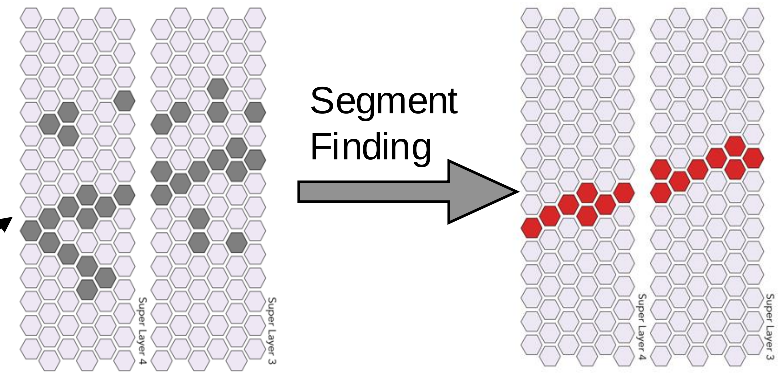
→ **Ideal for studying multiparticle final states with small cross-sections**

# Forward-detector tracking

**Drift chambers:**

- 6 sectors with 3 regions in each sector
- 12 wire planes in each region grouped in 2 superlayers with 6-degree stereo angle
- 112 wires per plane, hexagonal cells






**(Conventional) Tracking:**

- Find segments in each superlayer
- Combine segments into track candidates
- Identify the correct combinations among the candidates
- Fit the candidates to determine the particle 3-momentum (Kalman-Filter)

**Challenges:**

- Separated true hits from background in segment finding
- Limit the number of track candidates that are fitted
- Maximize the efficiency and reduce the processing time

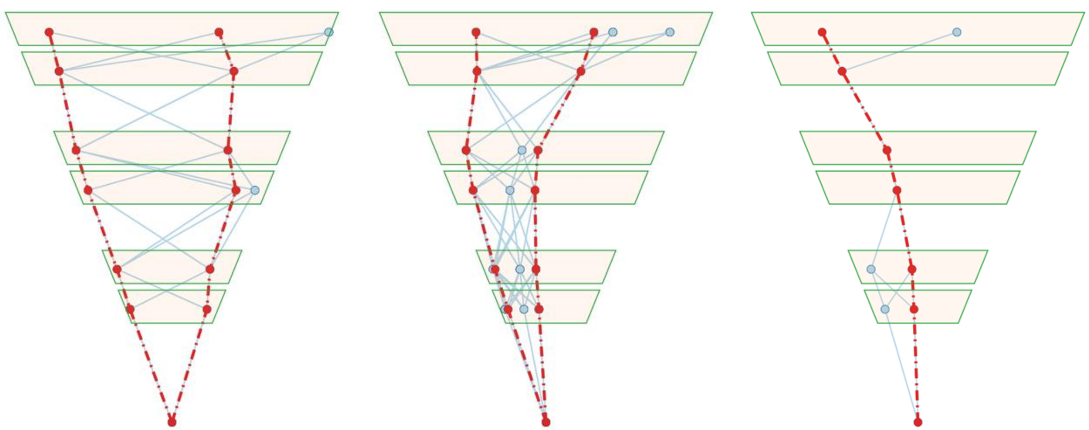
# AI/ML in track finding

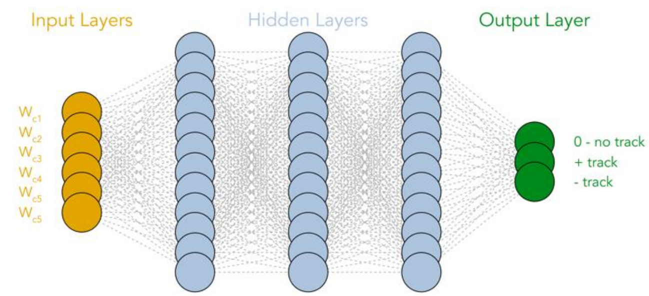
First inefficiency that was addressed is in “track finding”, i.e. linking segments into tracks

- In conventional tracking, done building and fitting all combinations with minimal cuts
- Slow and inaccurate when only wire positions are used

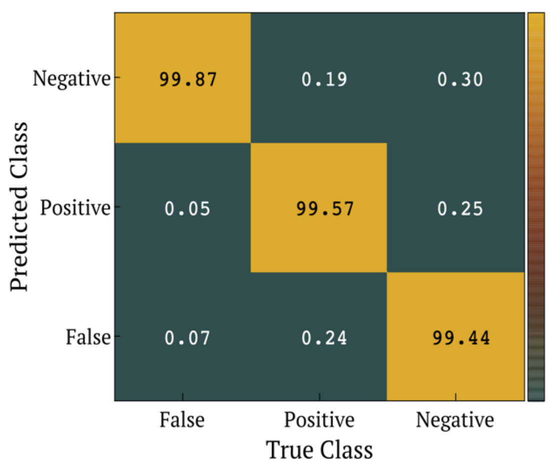
With AI, a neural network is used to recognize segments’ combinations of real tracks:

- The track classifier assigns a probability of the track candidate to be a positive, a negative, or a false track.
- The network is trained on reconstructed data where the right combinations are determined with the conventional algorithm
- False combinations of segments are generated by interchanging clusters from different tracks



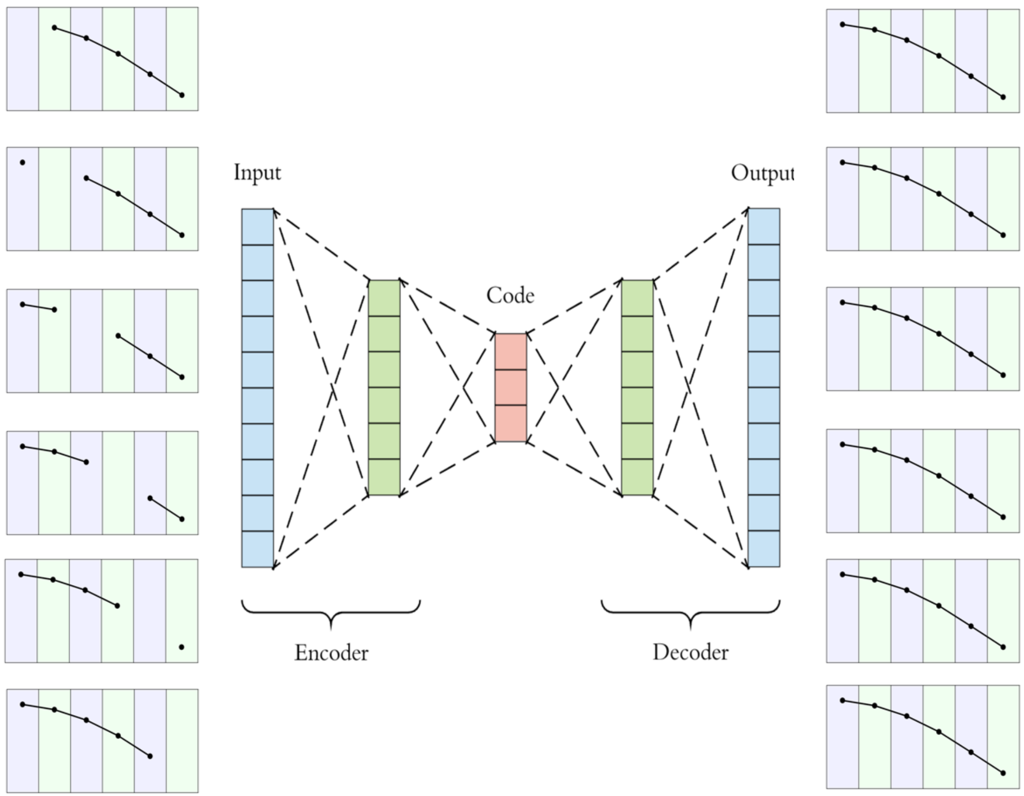



- Input: W [1..6] - average wire position of the segment
- Output: [false track, positive track, negative track]



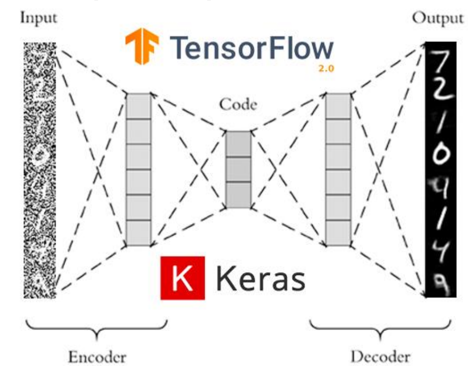
# AI/ML in track finding

- Allow for a **missing superlayer segment** to improve tracking efficiency
- Use **Corruption Auto-Encoders** to find the position of the missing segment

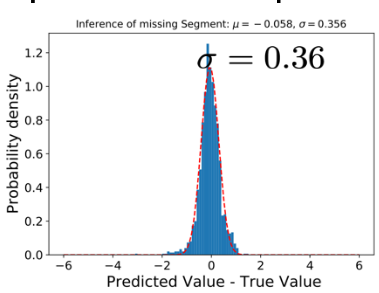


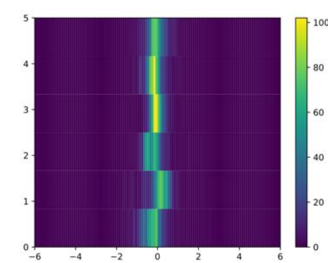

Good, 6-superlayers, reconstructed tracks are used to generate training samples by removing one of the segments

An auto-encoder is composed of an encoder and a decoder sub-models. The encoder compresses the input and the decoder attempts to recreate the input from the compressed version provided by the encoder
**Typically used for de-noising, but can be used for fixing glitches (our case)**




The network predicts the missing cluster position with a precision of 0.36 wires



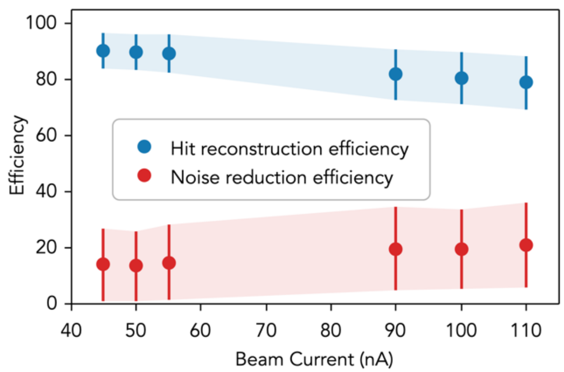



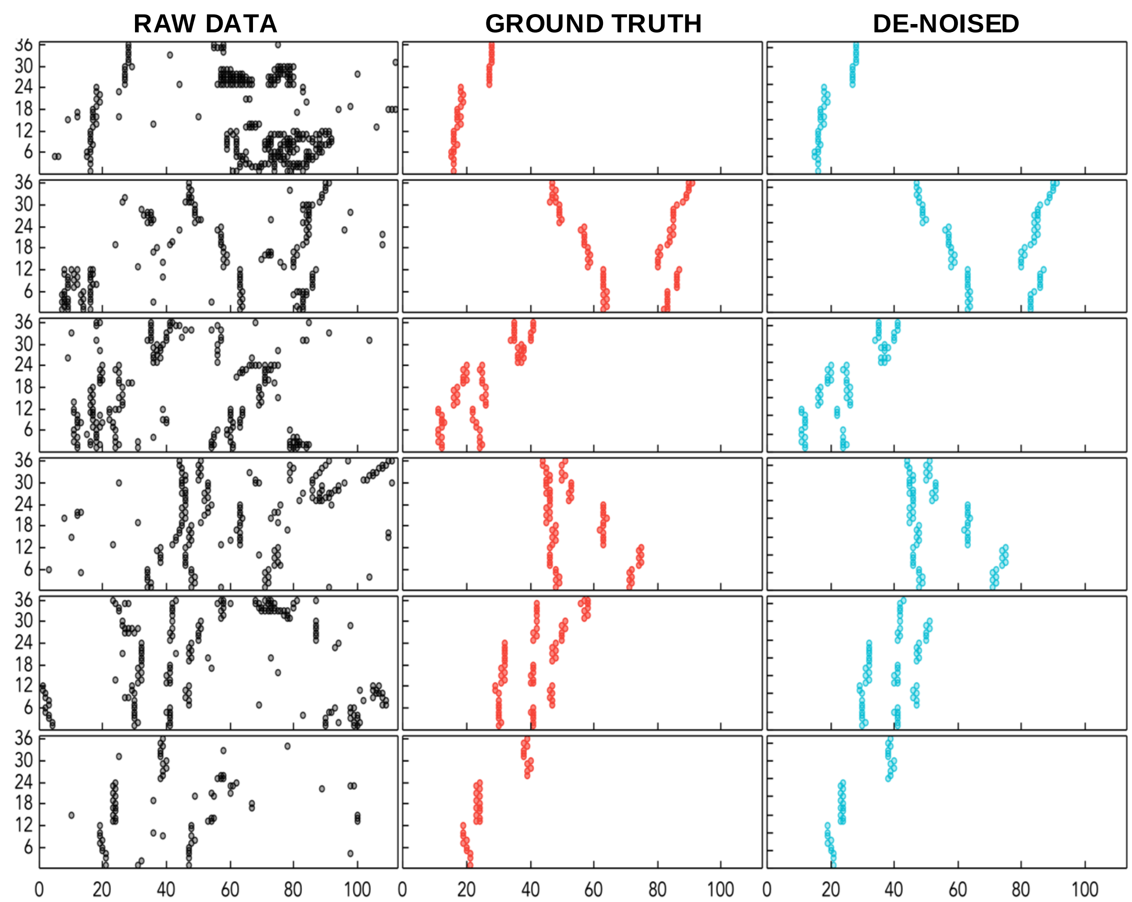
# “De-noising”

A Convolutional Auto-Encoder is also used to **de-noise** drift chamber raw data

- The network is trained on reconstructed data, separating hits-on-track among raw hits
- The resulting model can isolate hits that potentially belong to valid tracks from the background
- Large background reduction at the expense of some hit loss

**Network Performance**



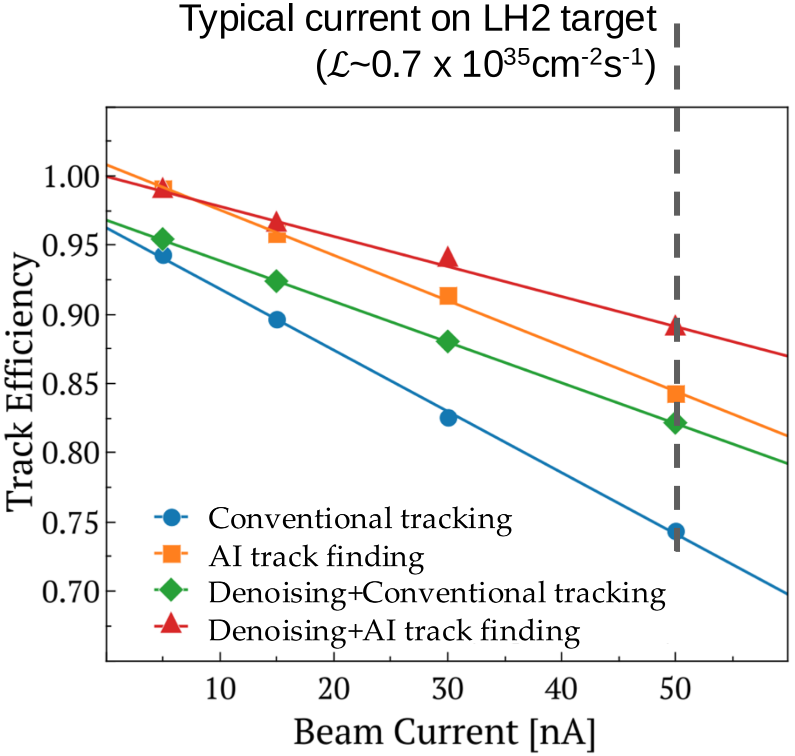
# Performance and impact on physics

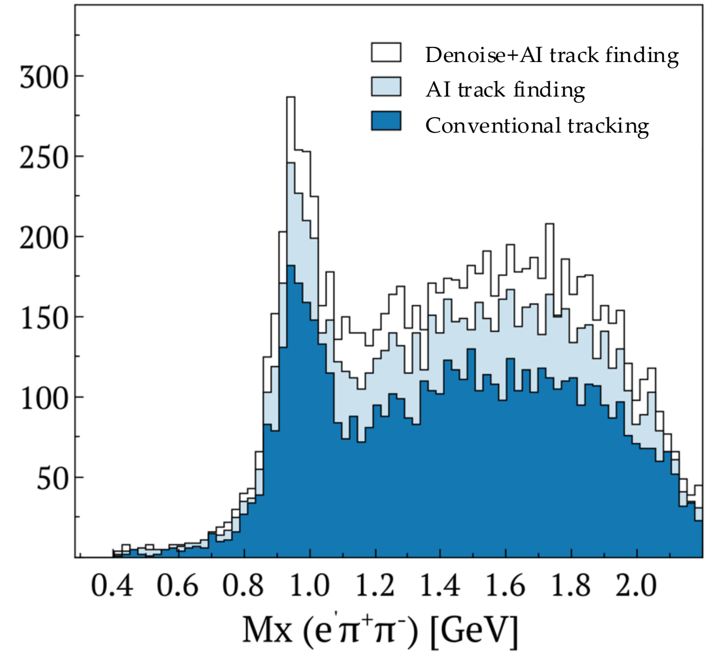
Performance of AI-based vs. conventional tracking algorithms studied in detail:

- Event-by-event comparison of reconstructed tracks to determine the relative efficiency and gain
- Dependence of luminosity of track multiplicities to estimate absolute efficiency
- Processing time

~18% gain per track
>50% gain per event, in 3-tracks final states
>30% reduction in overall processing time

Typical current on LH2 target
($\mathcal{L}$~0.7 x $10^{35}cm^{-2}s^{-1}$)

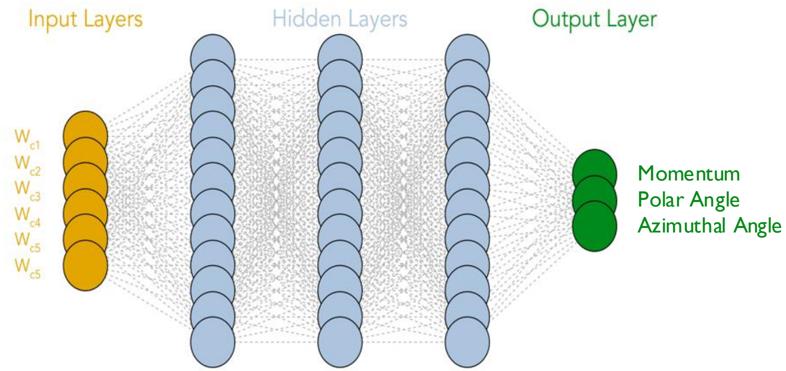


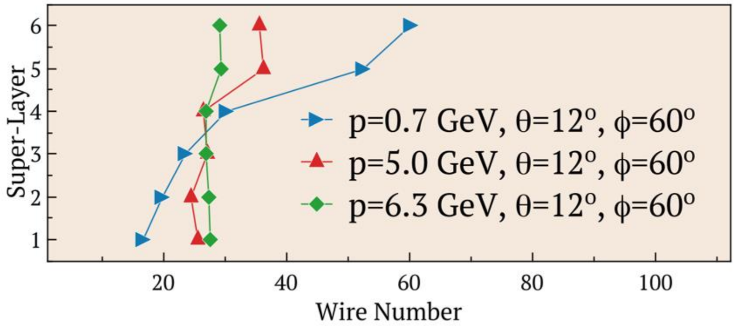
# New developments: InstaRec

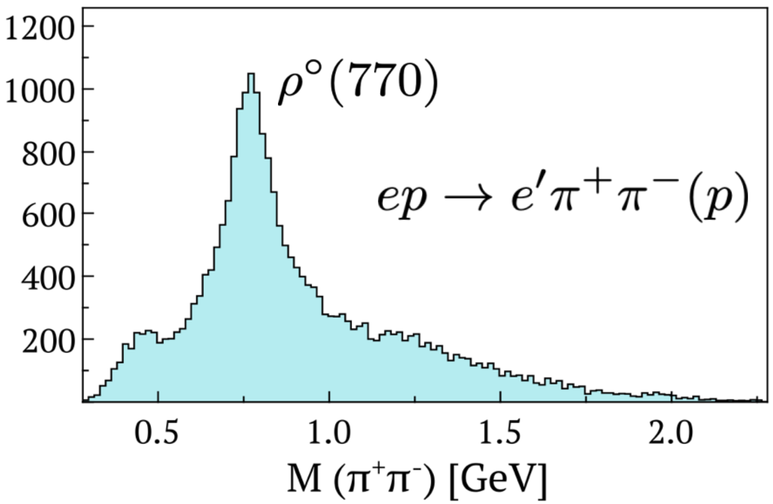
Move towards full event reconstruction:

- **Predict track 3-momentum**
- Link tracks to hits/clusters in Cherenkov detectors, ToF, and calorimeters to determine particle ID




*Input: W [1..6] - average wire position of the segment*
*Output: track momentum and angles*

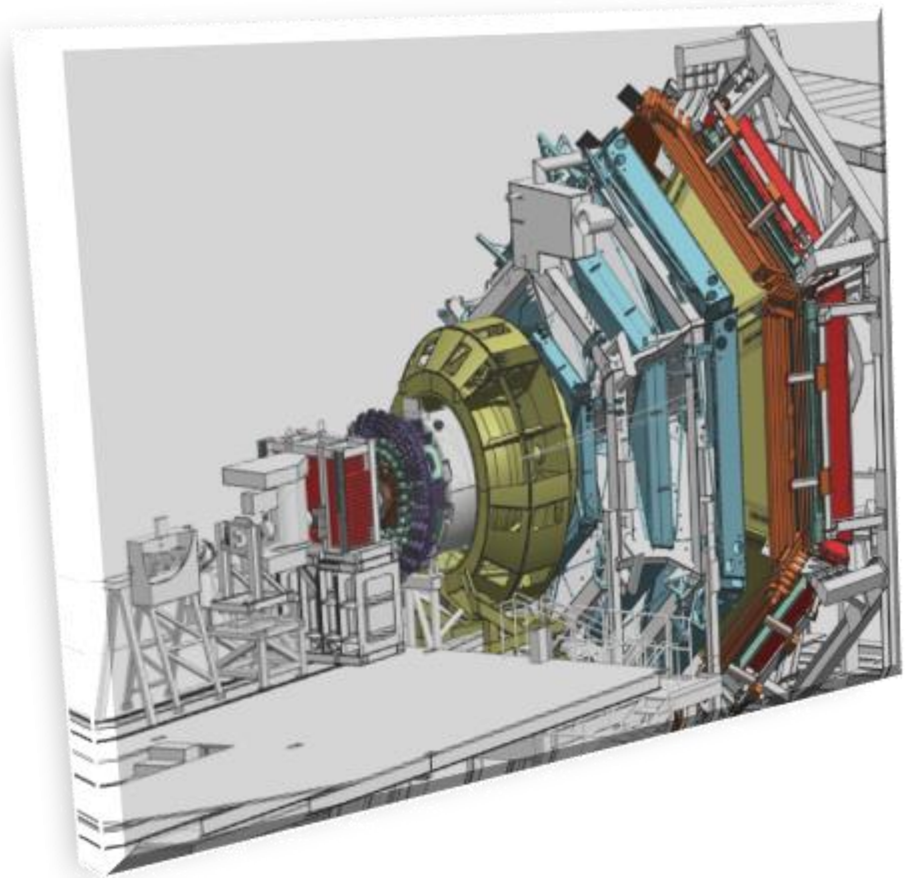


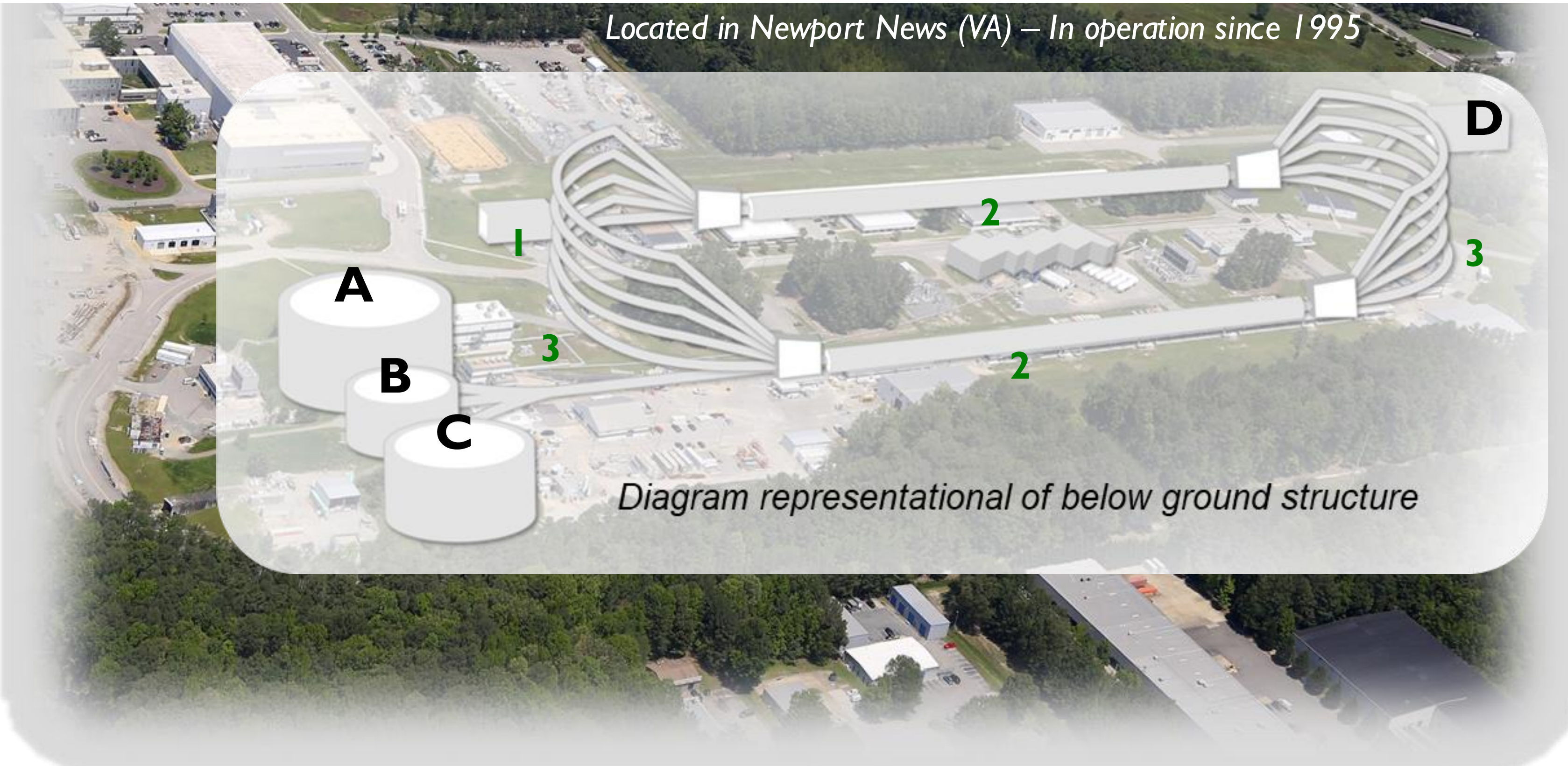

- Reconstruction rates of tens of kHz on a single CPU
  - Comparable to current triggered DAQ rate
- Possibility for:
  - Extensive online data quality monitoring
  - Real-time event filtering for high-rate triggerless DAQ
  - Event tagging for fast data processing/analysis
  - …

# Summary

- AI/ML tools are used in CLAS12 to support data taking, reconstruction, and analysis
- Large impact on experiment performance
- Further development in progress, aiming at real-time event reconstruction for event selection and data reduction in future high-luminosity runs



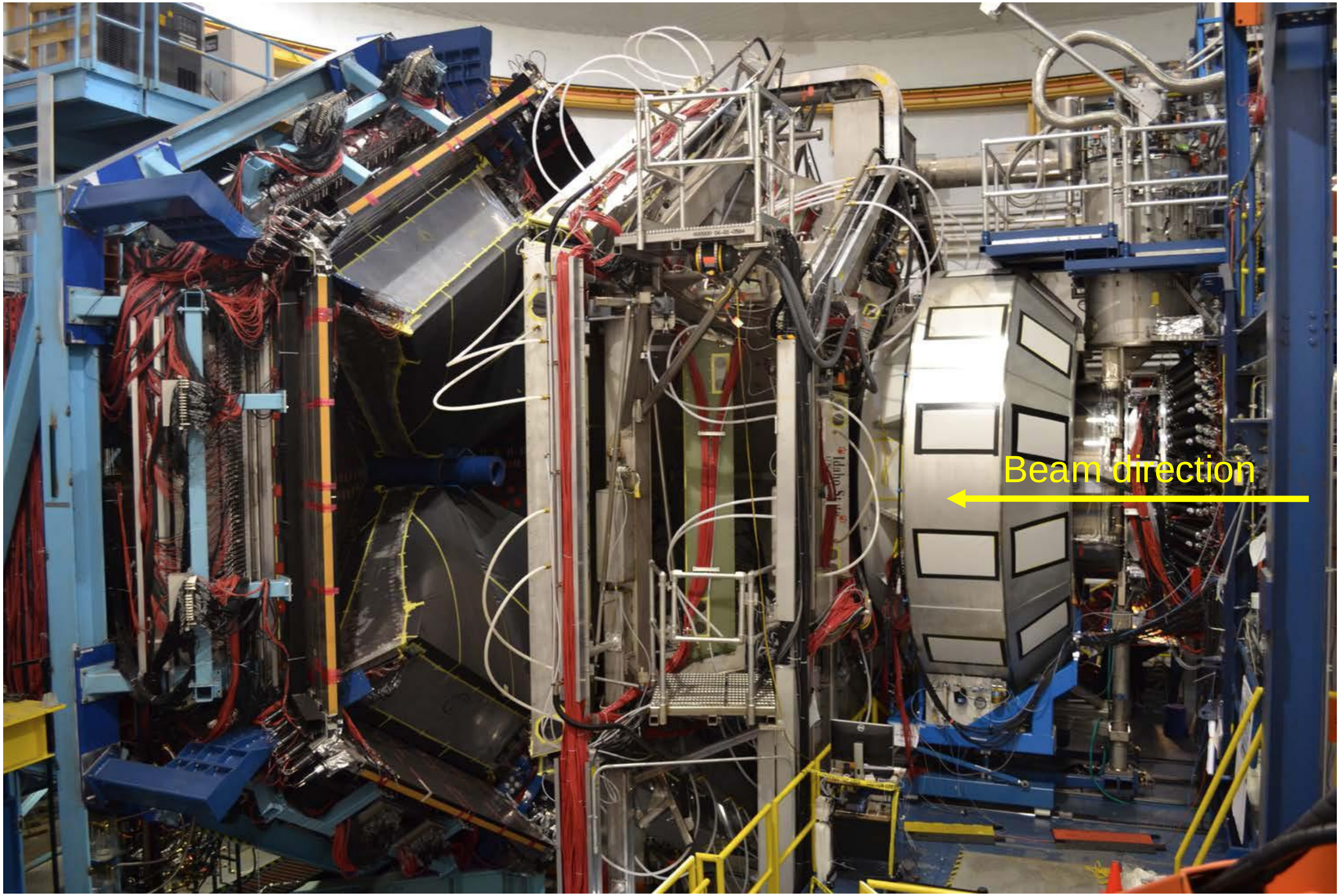
# Jefferson Lab @ 12 GeV



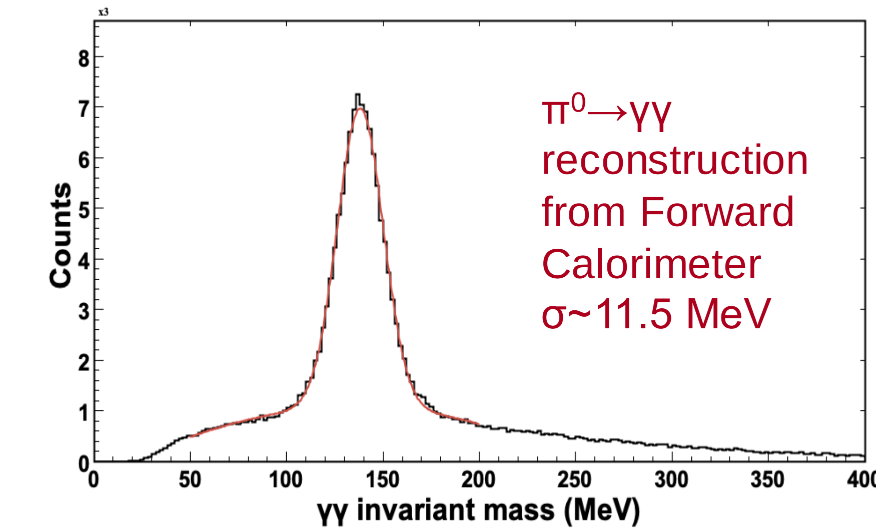

1. INJECTOR
2. LINACS
3. RECIRCULATION ARCS

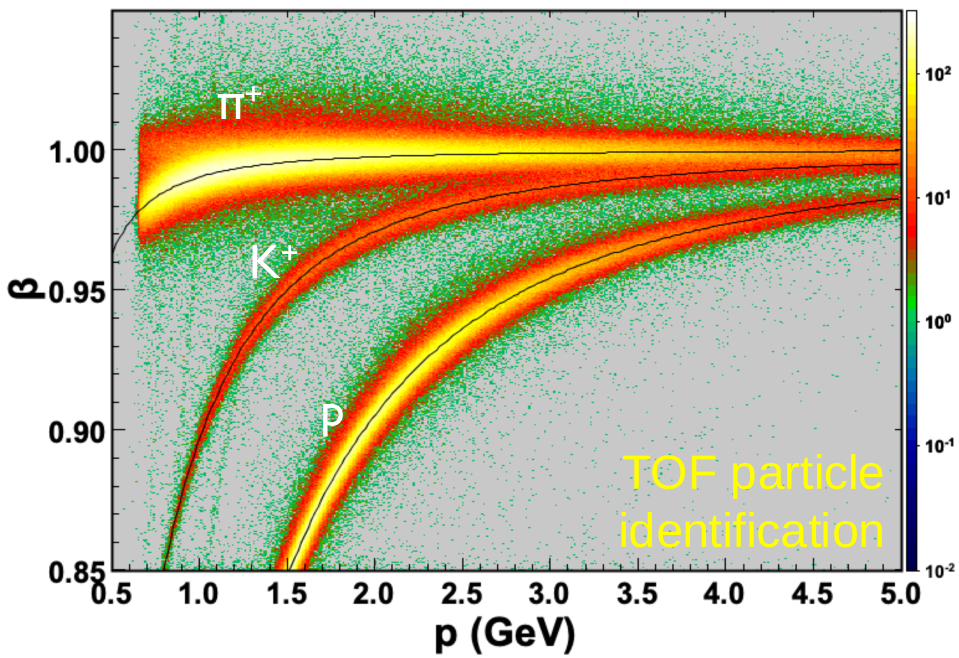
- Accelerator Upgrade completed in September 2017
  - CW electron beam
  - $E_{max}$ = 12 GeV, $I_{max}$ = 90 mA, $Pol_{max}$ ~ 90%
- Physics Operation
  - 4 halls running simultaneously since January 2018

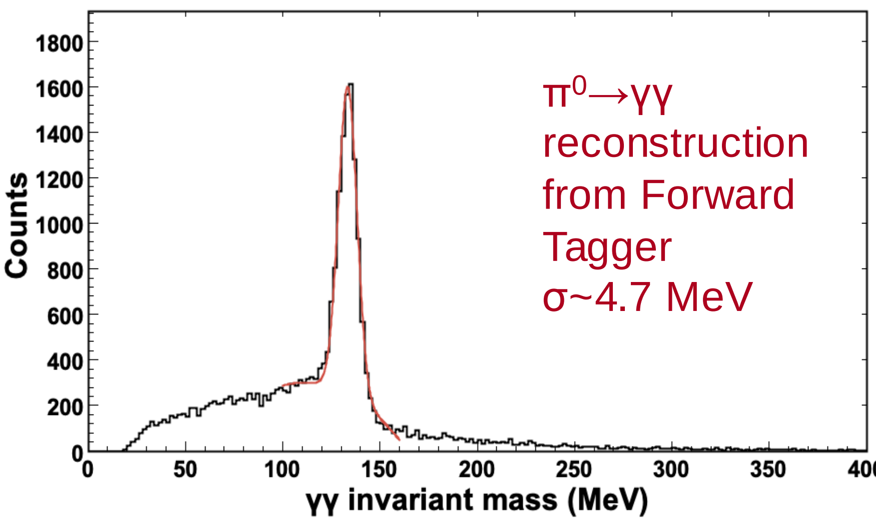
# CLAS12 in Hall B

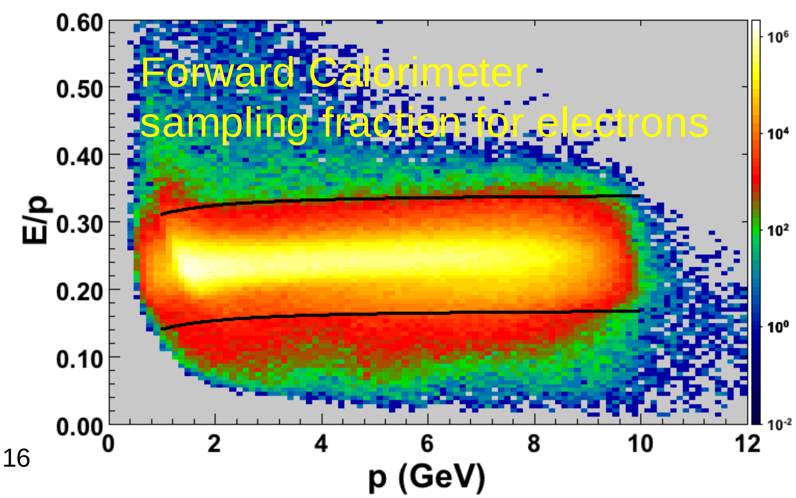


# Event reconstruction



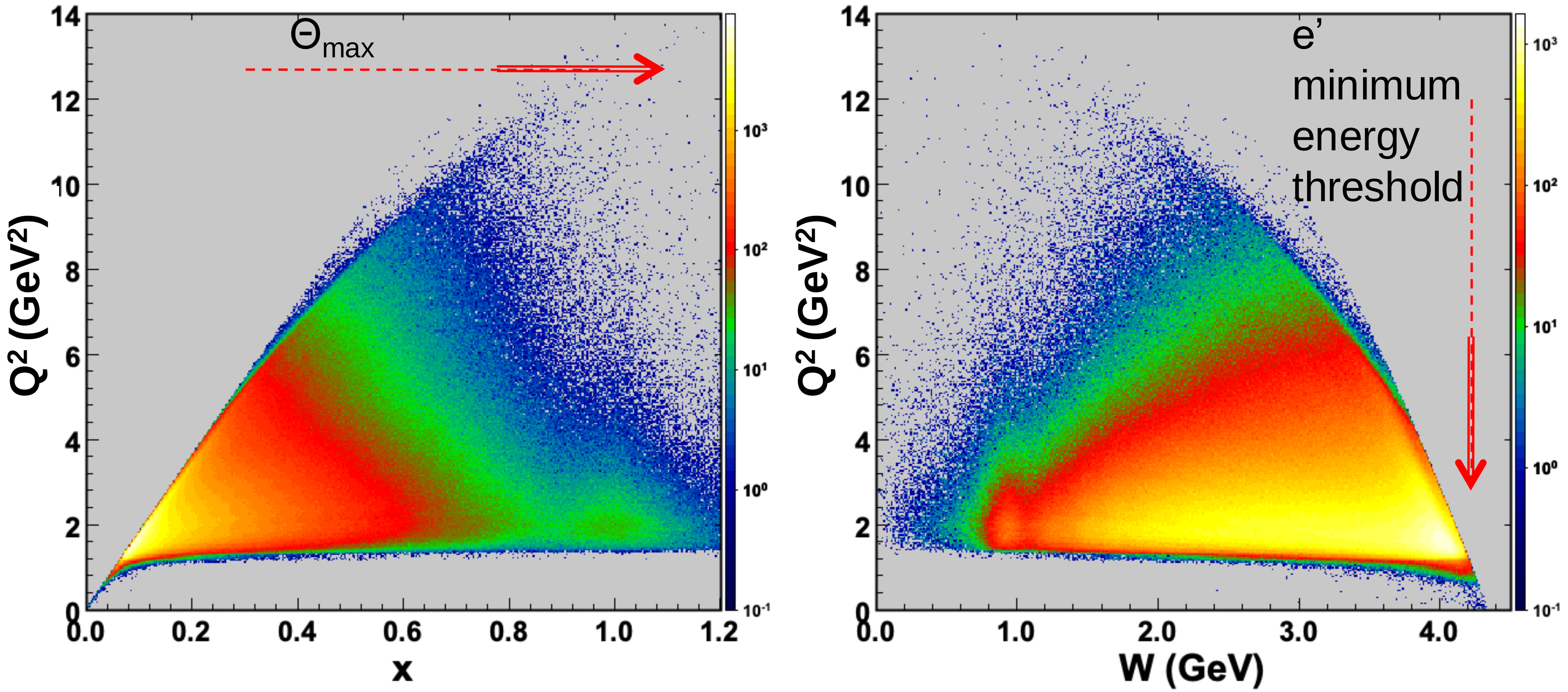



# CLAS12 kinematic reach

Beam energy at 10.6 GeV Torus current 3770 A, electrons in-bending, Solenoid magnet at 2416 A.

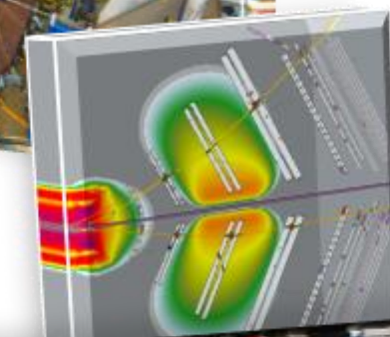
p(e,e')X




Plots based on 200 min. of data taking

# CLAS12 in numbers





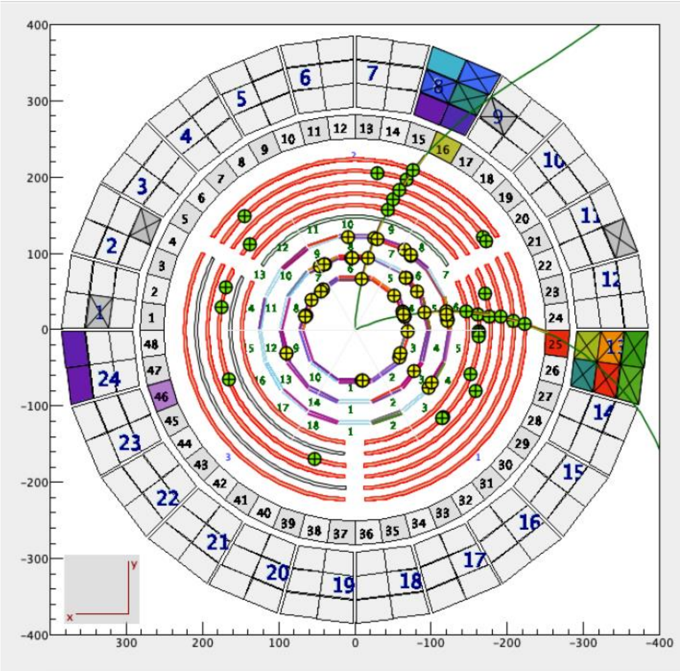

## Collaboration:

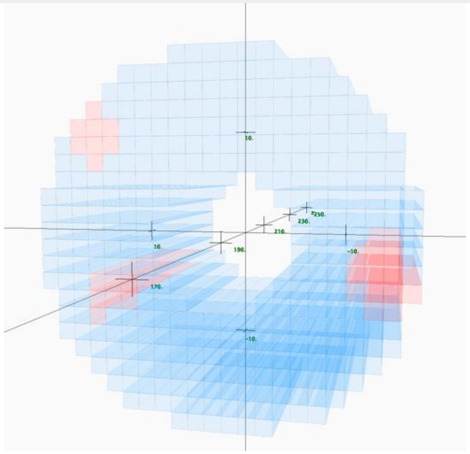
- More than 200 members
- 43 institutions
- 9 countries

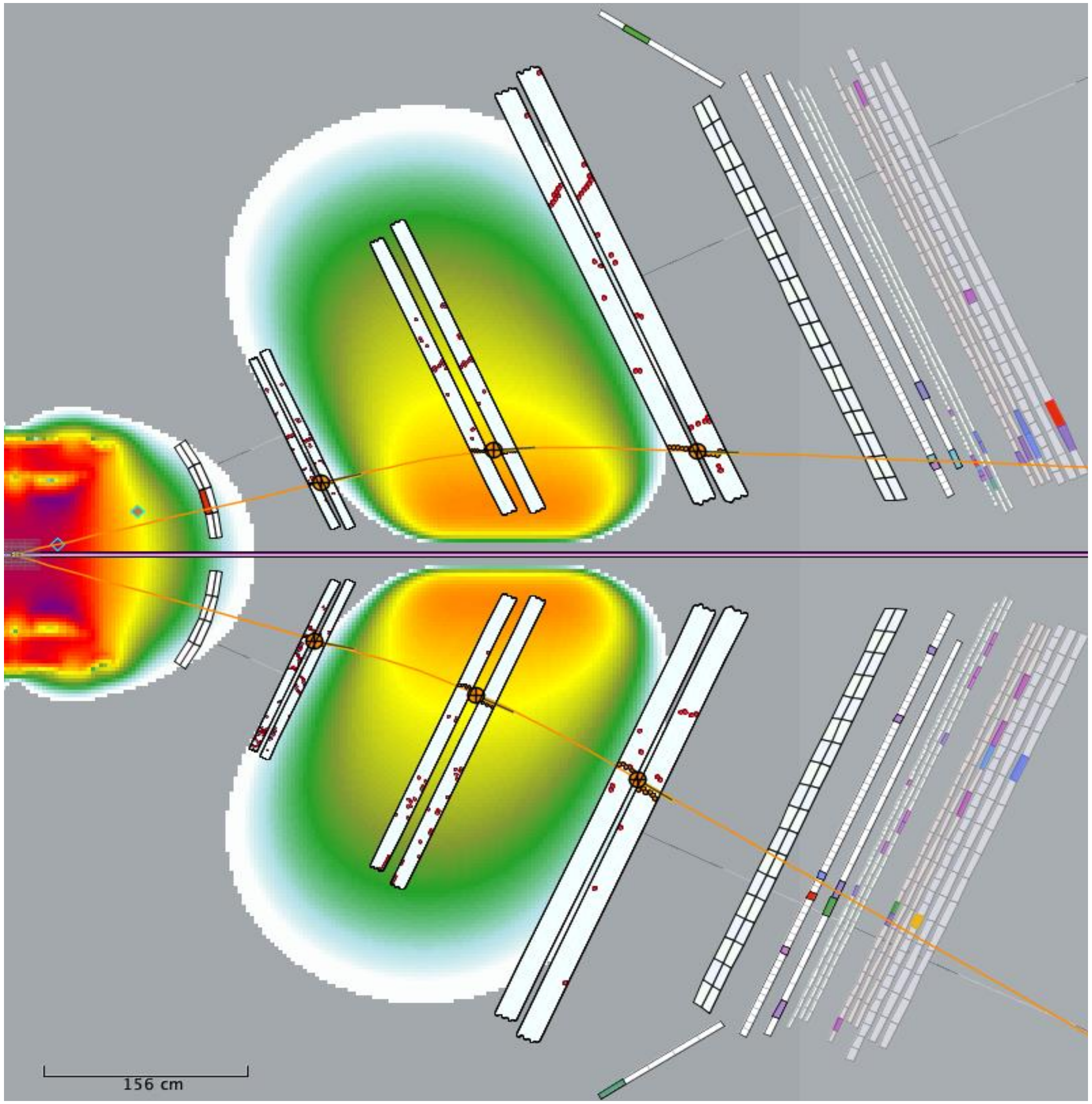
## Experimental program:

- 47 approved proposals:
  - targets:
    - proton, deuteron and nuclei
    - unpolarized, longitudinally and transversally polarized
    - solid, liquid and gas
  - beam:
    - highly polarized electron beam
    - linearly polarized quasi-real photons
  - final states: inclusive, semi-inclusive and exclusive
  - luminosity up to $10^{35}$ $cm^{-2}s^{-1}$
- 3188 PAC days
- 12 Run Groups
- 1171 Run Group days
- **10 years of approved data taking**

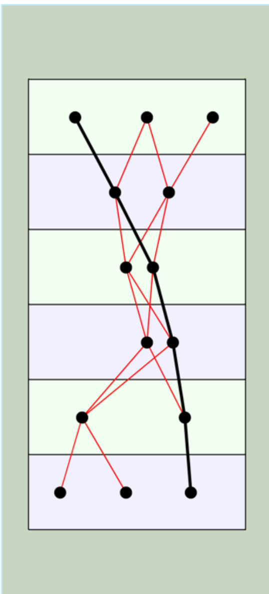
# CLAS12 Event Display



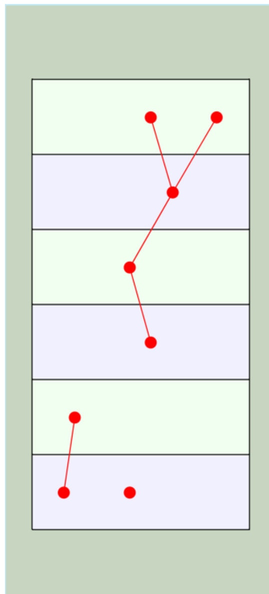




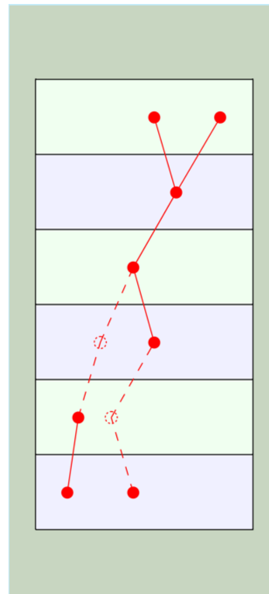
# Putting all together…



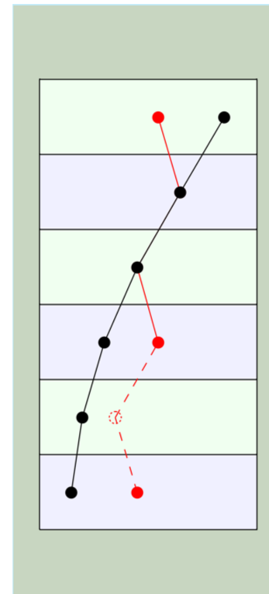
Classifier picks the correct track from 6 superlayer combinations



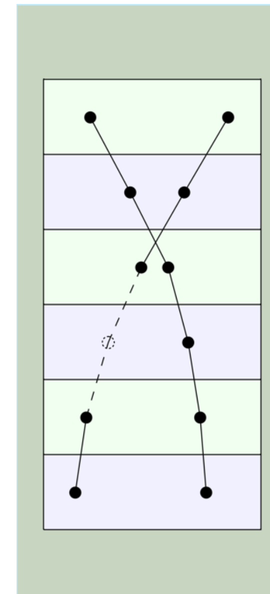
Remove all clusters belonging to identified track



Construct pseudo-clusters for all 5 super layer combinations using Corruption Auto-Encoder



Identify tracks using 6 super layer candidates with pseudo-clusters



Voila!

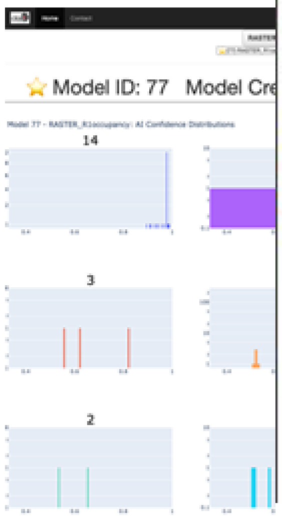
# AI/ML for data monitoring

HYDRA

**Front-End components**

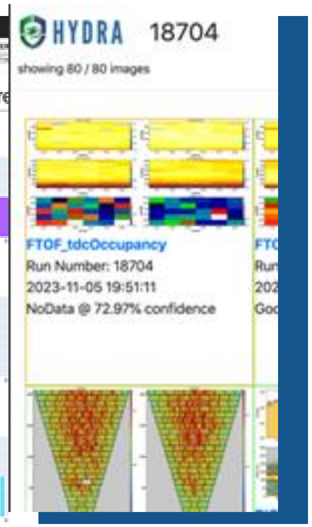
### Data Labeler

Efficiently label hundreds (thousands) of images



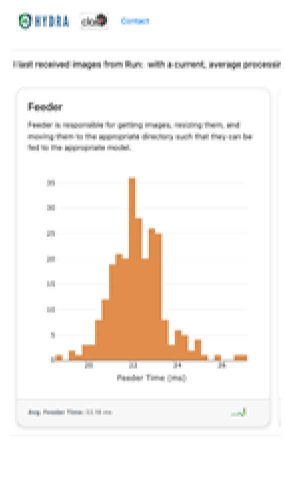

### Library

Contains enhanced confusion matrix, thresholds, active model designations




### Run

See predictions in real time




### Status

Monitor heartbeats for back end processes and image processing time

### Grafana

Dashboard displays all predictions over time

### Log

Display concerning plots sorted by detector from previous day

Extensible framework for real-time data quality monitoring using computer vision

Initially developed for Hall D/Gluex, then adopted by Hall B/CLAS12, now deployed in the 4 Halls

**Supported by JLab EPSCI group**

# AI/ML in data analysis

- Increasing use of AI/ML to solve complex, multiparametric problems in physics analysis
- Some examples:
  - Modeling Dilepton Background using Boosted Decision Trees
  - Lepton Identification using TMVA Methods
  - Gradient Boosted Decision Trees for photon classification
  - Neutron identification in the central detector
  - …
- AI group established within the collaboration to share tools, know-how, …