(Affiliated with the CYGNUS Proto-collaboration)

(Affiliated with the MIGDAL Collaboration)

Machine Learning (ML) for Improved Analyses of High Resolution Gaseous Detector Data

Dr. Jeff Schueler*

CPAD 2023: Nov 9th, 2023

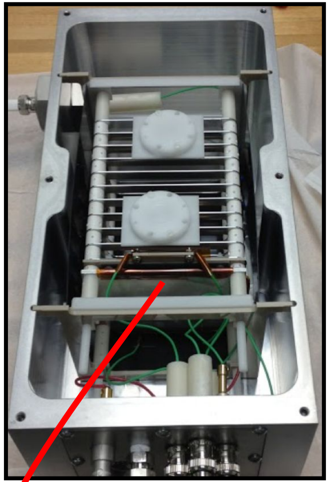
*Email: schuel93@gmail.com

MIGDAL

Migdal In Galactic Dark mAtter expLoration

Detector technology highlighted today

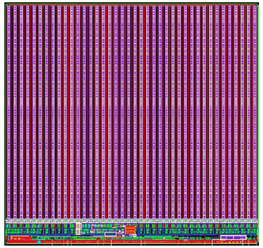
BEAST TPCs: 40 cm³ pixel readout *CYGNUS prototype directional detection vessels*



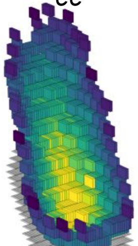
- 70:30 He:CO_2 1 atm
- Double thin GEM amplification
- 2cm x 1.68cm pixel readout with 250μm x 50μm segmentation

Goal: Demonstrate directionality at the lowest possible NR energies

ATLAS FE-I4b



Pixel-level 3D reconstruction ~300 keV$_{ee}$ He recoil

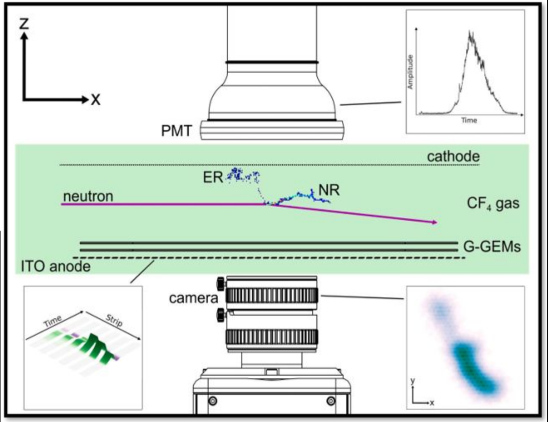


MIGDAL TPC: 108 cm³ with CMOS camera, strip, and PMT*

MIGDAL

Migdal In Galactic Dark mAtter expLoration

- 50 torr CF_4
- Double glass GEM
- High dynamic range
- 8cm x 4.5cm optical readout, 39μm seg.



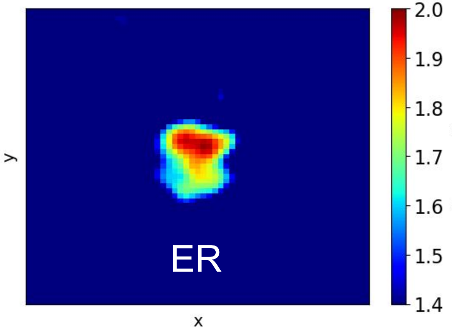
Goal: Directly measure the Migdal effect in nuclear scattering

****This talk only focuses on the camera readout!***

The ML techniques described in these slides are not restricted to pixel and optical readouts!

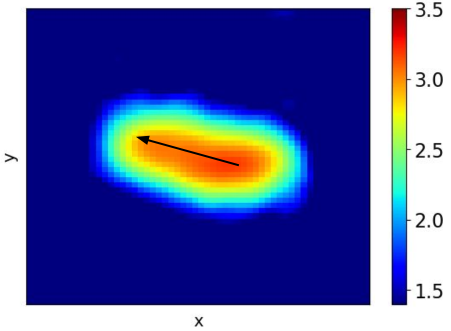
Applications of deep learning in gas detectors

Classification



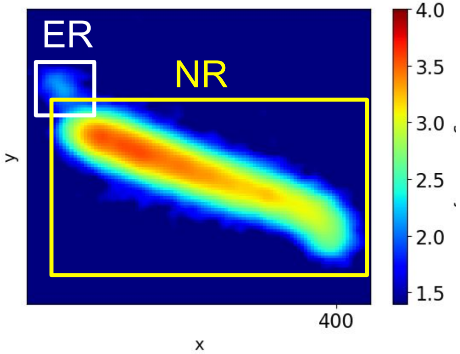
Application:
Particle ID

Regression



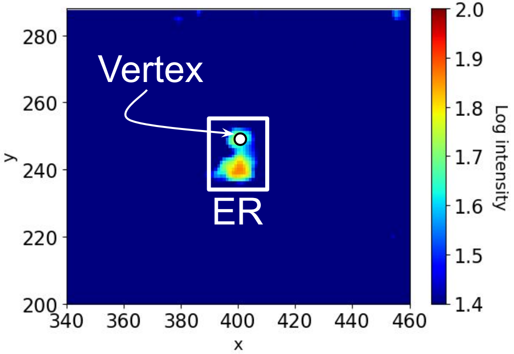
Applications:
*1. **Directional reconstruction***
2. Energy reconstruction

Object detection



Application:
Rare event searches (Migdal effect)

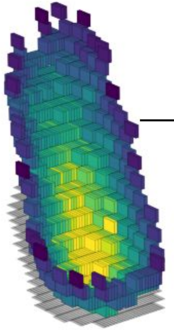
Key point(s) detection

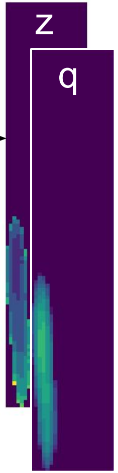


Applications:
1. Vertex detection
2. Head/tail identification
3. Trajectory fitting

Many implementations of algorithms for object detection and key point detection already exist (see here and here), and there are many tools for developing custom classification and regression models for 2D and 3D image data (PyTorch tutorial, TensorFlow tutorial, spconv for sparse 3D convolutional neural networks).
Today I'll highlight examples of (1) directional reconstruction for a CYGNUS prototype BEAST TPC and (2) object detection for the rare event Migdal search

Deep learning to assign principal axis and head/tail direction (3D reconstruction)





Training Data
X

Labels
$\mathbf{y}_{truth,1}$, $y_{truth,2}$

Model

Custom ResNet 50
(Architecture in backup)

Output 1: Recoil axis

$\mathbf{y}_{pred,1}$
(1 x 3) axis

$L_1(\mathbf{y}_{pred,1}, \mathbf{y}_{truth,1})$

$\mathbf{y}_{truth,1}$
Truth axis

Output 2: Head/tail assignment

$p_{pred,2}$
Probability

$L_2(p_{pred,2}, y_{truth,2})$

$y_{truth,2}$
0 or 1

Train model to minimize *total loss,* L (encodes the difference between prediction and truth)

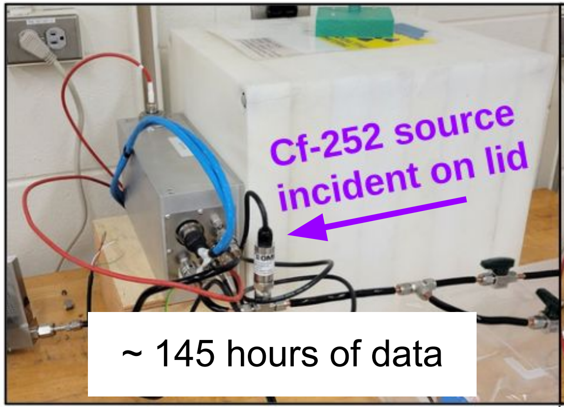
GOOD CHOICE OF LOSS FUNCTION IS ESSENTIAL!

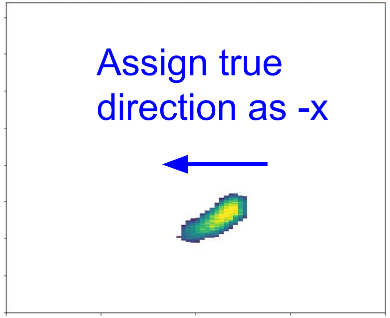
$$L \equiv L_1(\mathbf{y}_{pred,1}, \mathbf{y}_{truth,1}) + \alpha * L_2(p_{pred,2}, y_{truth,2})$$

- We set α to 1
- $L_1 \equiv 1 - |\cos(\mathbf{y}_{pred,1}, \mathbf{y}_{truth,1})|$ **(Ang. resolution loss)**
- $L_2 \equiv -(y_{truth,2} \log(p_{pred,2}) + (1-y_{truth,2}) \log(1-p_{pred,2}))$ **(Head/tail loss)**

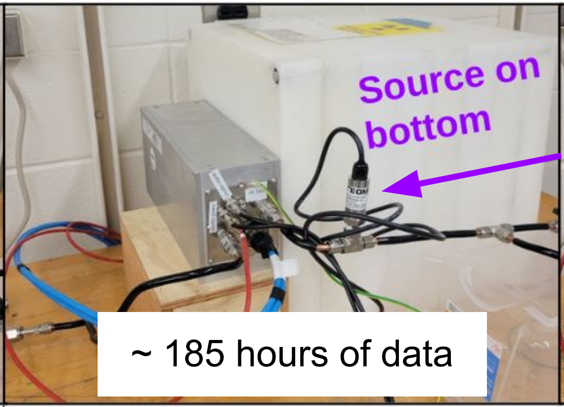
Testing our model on head/tail recognition

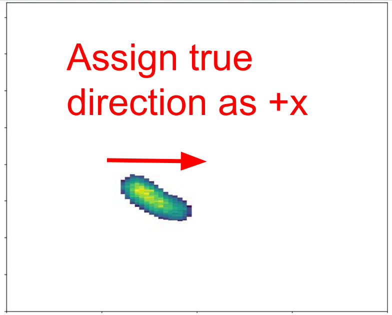
“-x” sample



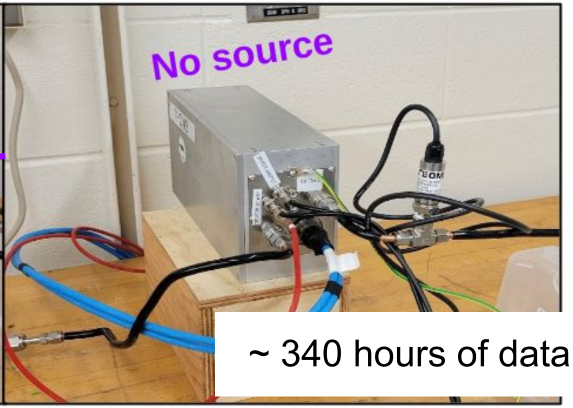


“+x” sample





Background sample

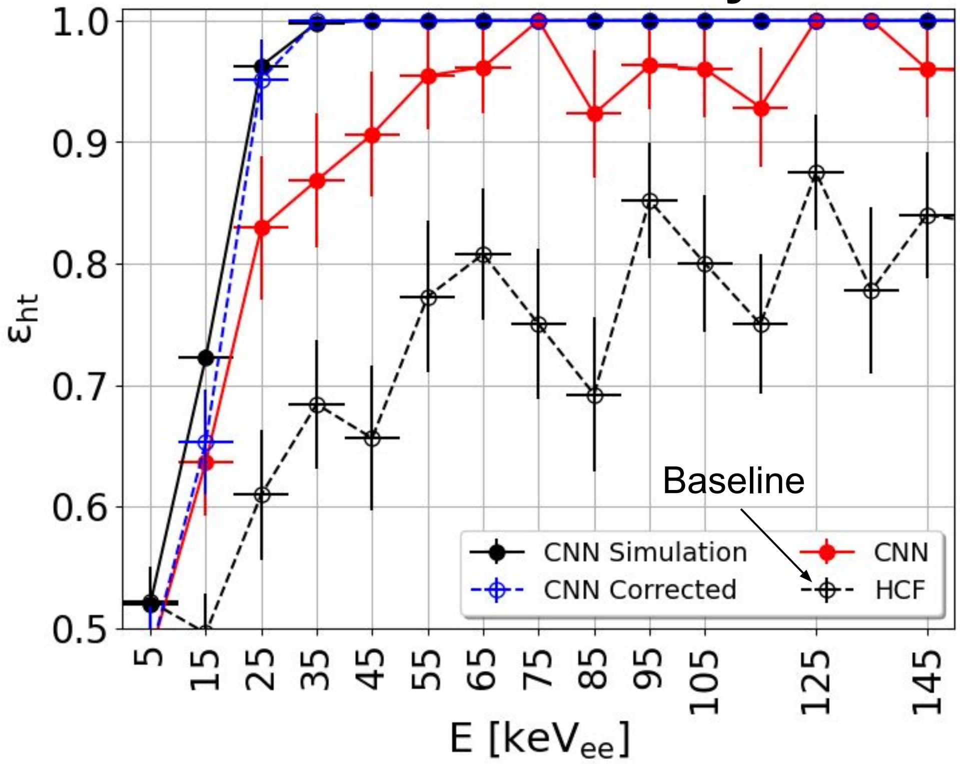


ResNet is trained on simulation, we evaluate its head/tail recognition on measured data. *We don’t have angular resolution measurements to test on**

***angular resolution results on simulation in backup**

Results on He recoils

Measurement, 50% He-recoil efficiency

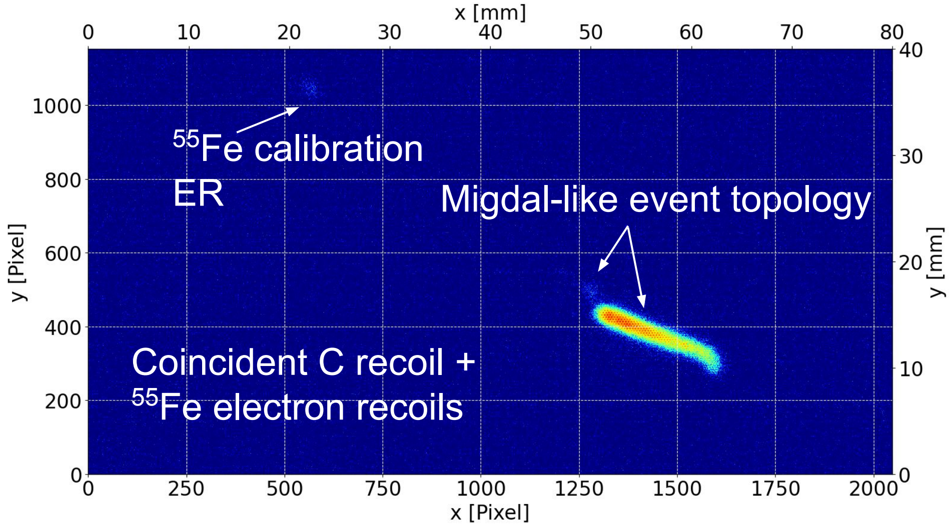


ResNet yields significant improvement in $\boldsymbol{\varepsilon_{ht}}$ on measurement →First significant $\boldsymbol{\varepsilon_{ht}}$ below 20 keV$_{ee}$ in an atmospheric pressure gas mixture!

Rare event searches: The Migdal effect

MIGDAL
Migdal In Galactic Dark mAtter expLoration

- Migdal effect is rare and has never been observed in nuclear scattering
    - ***Migdal signal with 5-15 keV ER expected in ~1 in 50,000 nuclear recoils (E > 100 keV) induced by 2.5 MeV neutrons in 50 torr CF_4 (details here)***

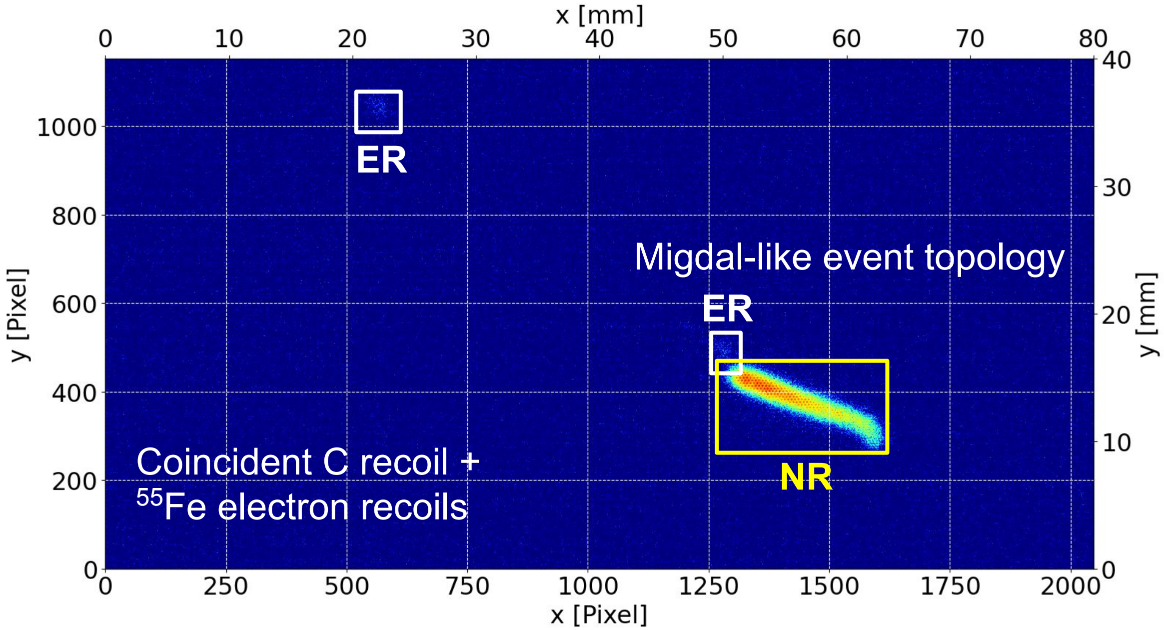


See E. Tilly's talk for a more thorough intro to the MIGDAL experiment

The MIGDAL experiment is designed to record (2,048 x 1,152) images at a rate of up to 120 Hz
*→**Over 1 PB of data per month!***
→Analyzing this amount of data in a timely manner is challenging, especially in a small experiment, but machine learning can help us!

Enter YOLO (You only look once)

YOLOv8 is a state of the art *object detection* algorithm that simultaneously locates (draws a bounding box) and identifies objects of interest in an image



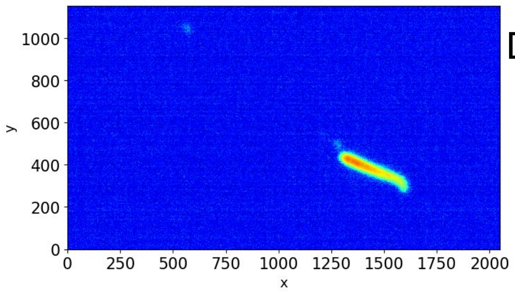
***We train YOLOv8 on measured data** to identify ERs, NRs, protons, alphas, sparks, camera afterglow, rolling shutter, etc.*

Benefits:

1. *Can identify multiple particle species within a continuous cluster*
2. *Not trained specifically to find Migdal candidates →robust and doesn't need to be trained on simulation!*
3. *Single-shot identification and analysis of tracks*
4. *Enables real time ^{55}Fe calibrations and ER/NR event rate counting*

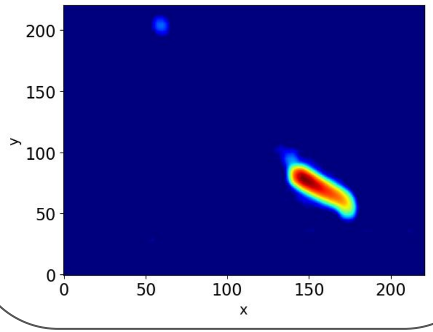
The data processing pipeline

(1) Dark subtracted image



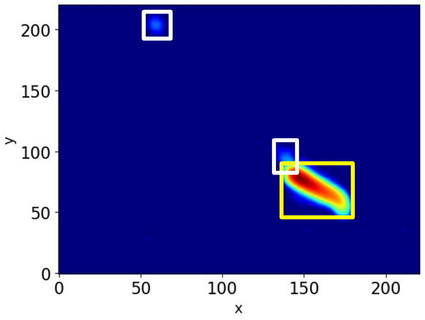
(2) Downsample

At this stage we train YOLOv8 by hand labeling bounding boxes



(3) YOLOv8 predicts bounding boxes

Retrain as needed



(4) Perform analysis on each bounding box, computing qtys such as: Intensity, track length, angle (with head/tail), bounding box centroid

(5) Save coordinates of each bounding box, as well as extracted physics information

This entire pipeline runs at 200 fps on a consumer desktop GPU, reduces data size by a factor of ~5,000 and is integrated with the MIGDAL DAQ → Real time feedback!

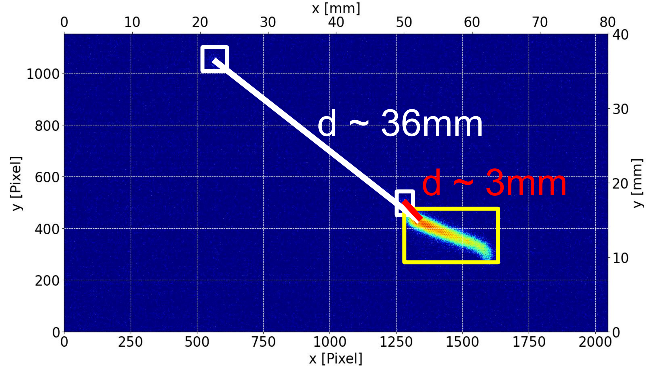
Application to MIGDAL searches

- Initial science run recorded from July 17th 2023 - August 3rd 2023
- Collected an unblinded dataset consisting of 10 million 2,048 x 1,152 images
  - **20ms exposure → We expect coincidences in the camera frames**

(1) Initial sample of ***(10 million frames)***

(2) Find frames containing at least 1 ER and 1 NR ***(~25,000 frames)***

(3) Map bounding boxes to raw image and compute the distances between NR vertices and centers of ER bounding boxes

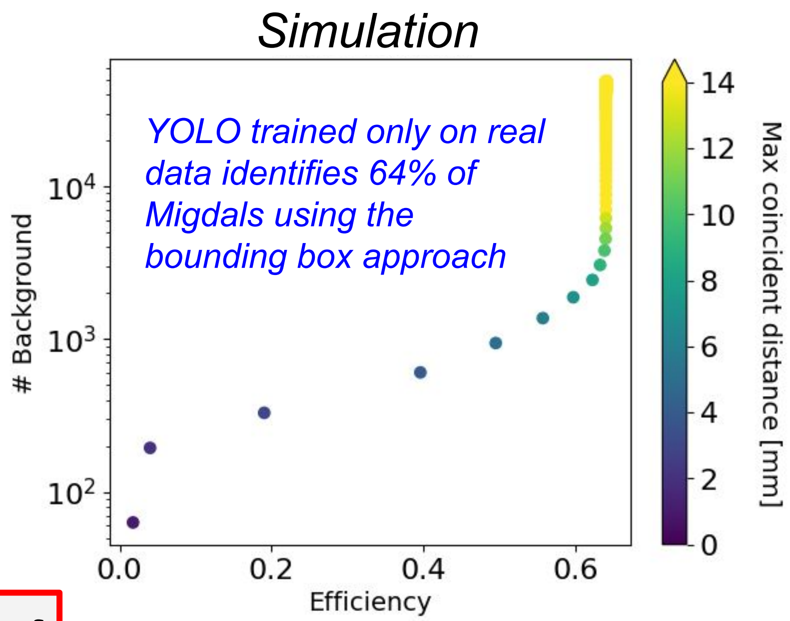


(4) Filter frames based on distance between ER and NR ***(~1,000 frames*** with d < 7.5mm and NR > 100 keV_{ee}*)*

Migdal identification performance

- Evaluate YOLOv8 (trained on real data) on simulation to determine a Migdal detection efficiency
  - **Background sample:** ~50,000 frames containing an ER and NR with random uniform separation
  - **Signal sample:** ~2,000 frames with simulated Migdals that contain ER pixels outside of the NR

- We expect the 64% Migdal efficiency evaluated on simulation to be a lower limit of performance (it should do better when applied to real data)



YOLOv8 bounding box analysis takes us from 10's of millions of frames to a few thousand frames while maintaining at least 64% of the Migdal signal → No longer a rare event search! We can spend the rest of our resources optimizing signal purity!

Summary

- ResNets improve low energy directional reconstruction in a pixel-readout TPC
  - He-recoil head/tail sensitivities below 20 keV_{ee} in 1atm $He{:}CO_2$
- YOLOv8 enables real time analysis of high throughput, high resolution image data
  - Able to analyze 100TB of data per day on a consumer desktop PC with a single graphics card
  - **Reduction from 10’s of millions to thousands of image frames turns a rare event search into an ordinary event search, opening up many avenues for signal purity optimization!**

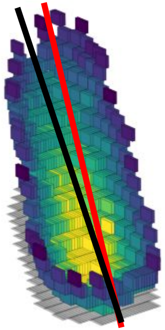
Modern deep learning approaches are more accessible and user friendly than ever for improving analyses of gas detector data!

Backup

Quantifying 3D directional reconstruction (BEAST TPCs)

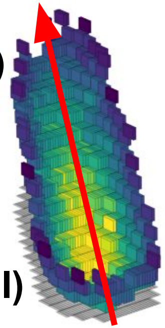
Angular resolution: The mean difference in 3D angle, θ, between the true recoil ***axis***, $\mathbf{v}_1$, and the measured recoil axis, $\mathbf{v}_2$.

- Perfect angular resolution: 0°
- No angular resolution: 57°



Head/tail recognition efficiency ($\boldsymbol{\varepsilon}_{ht}$): Fraction of events where the dot product of the reconstructed vector and the true initial recoil direction is positive

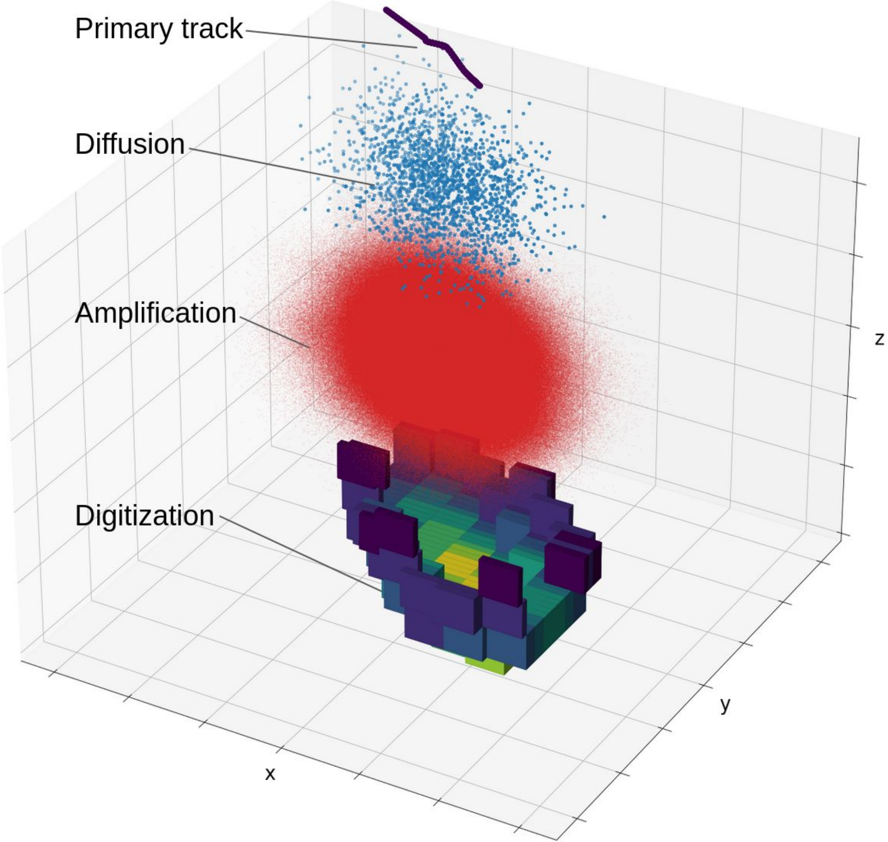
$\boldsymbol{\varepsilon}_{ht}$ **= 0.5 ⟸⟹ No head/tail sensitivity**
$\boldsymbol{\varepsilon}_{ht}$ **= 1 ⟸⟹ Perfect head/tail sensitivity**

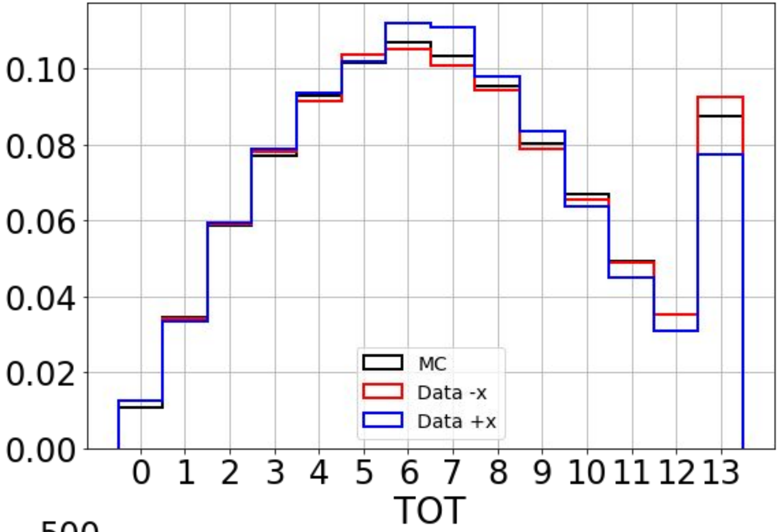


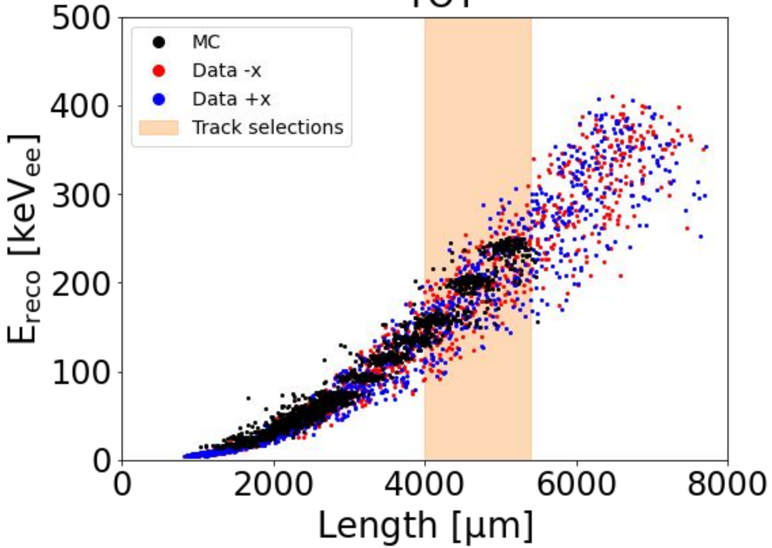
We train a custom 50 layer ResNet, to predict (1) a recoil axis (angular resolution), and (2) a vector (head/tail) direction. To train it, we assign two labels to each piece of training data:
Label 1 (Angular resolution): Truth recoil axis
Label 2 (Head/tail): 1 if track points in +x direction, 0 if track points in –x direction

Data vs simulation in BEAST TPCs

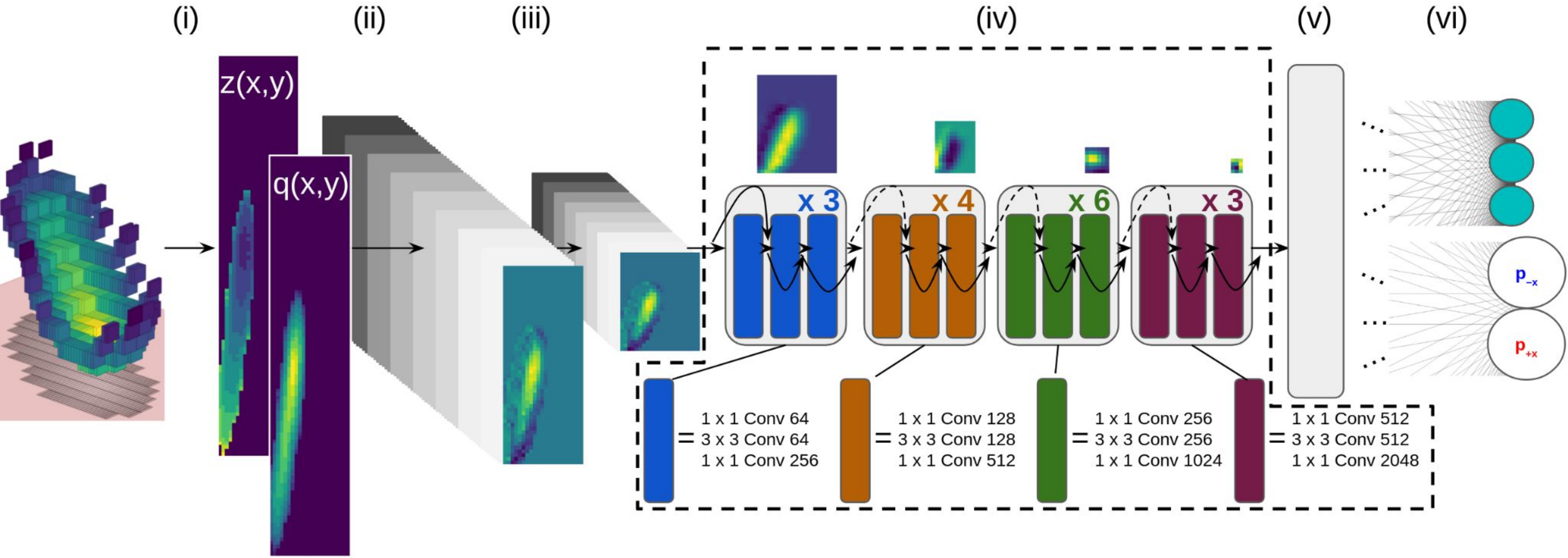
Simulation schematic



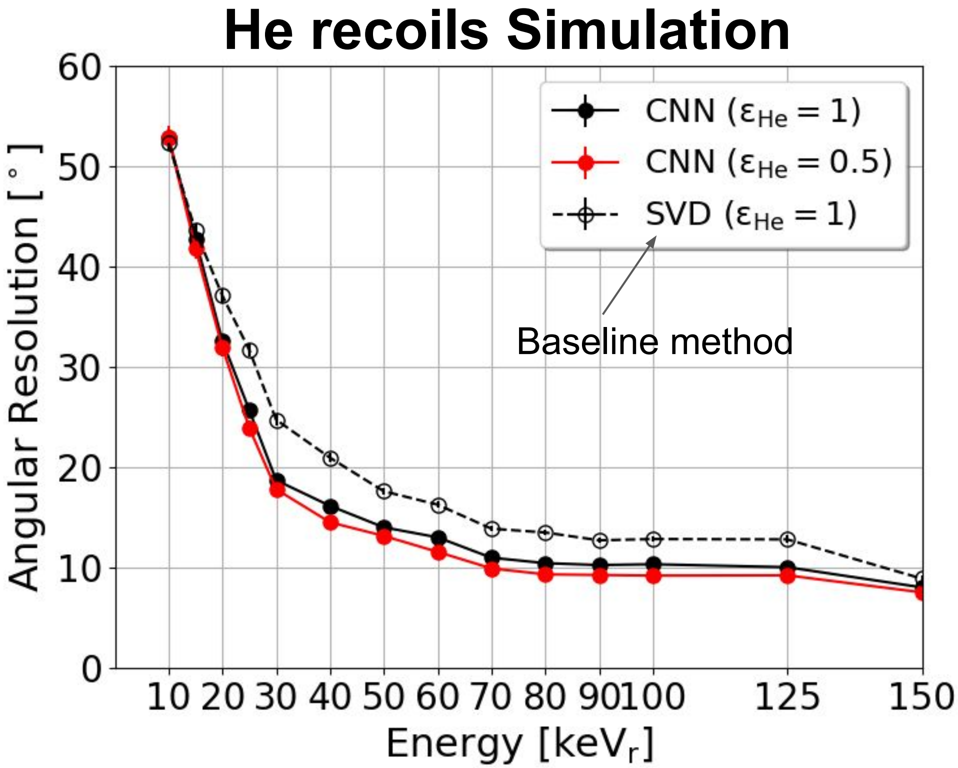




3D angular and directional reconstruction in the ResNet 50 architecture



Angular resolution results (ResNet trained on simulation)

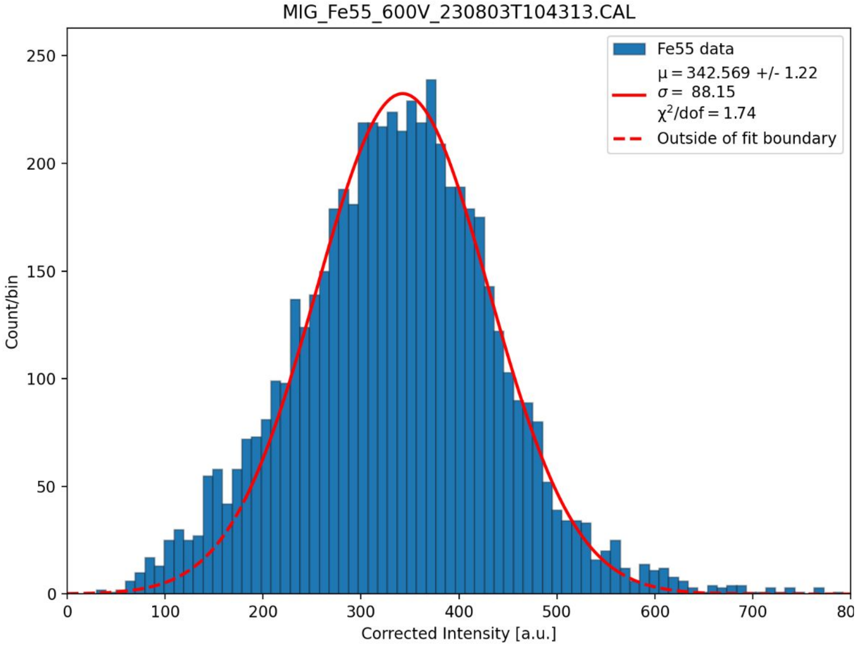


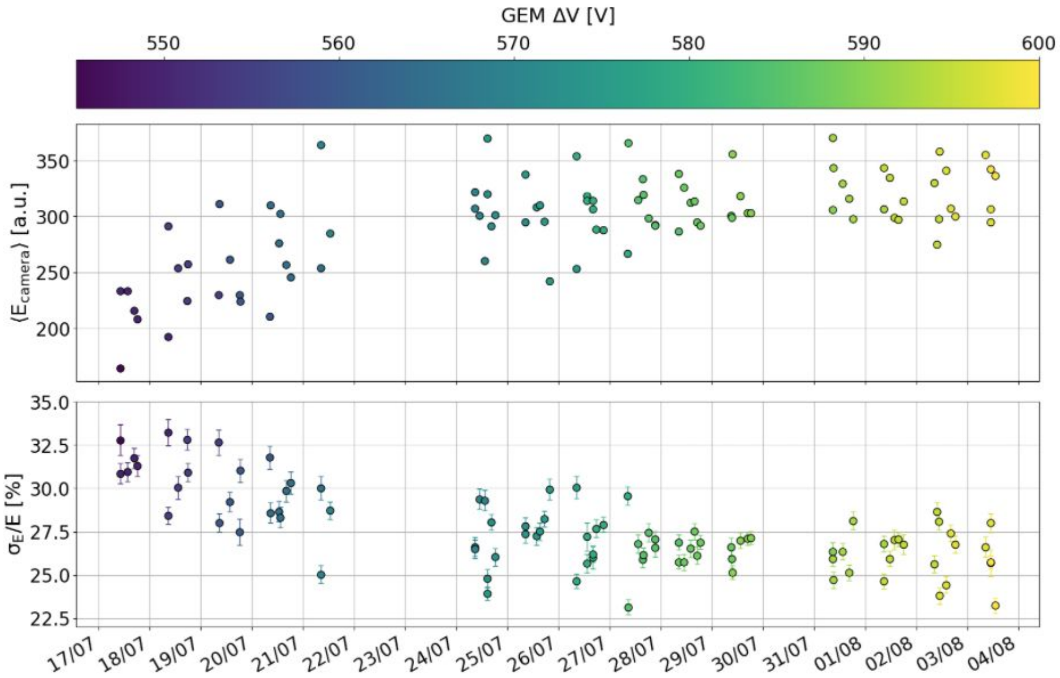
Noticeable improvement in angular resolution evaluated on simulation. We don't have measurements to test angular resolution on data, but we can test head/tail on data!

Expected Migdal backgrounds per 1 million DD-induced nuclear recoils with E > 100 keV

Component	Topology	D–D neutrons	
		>0.5	5–15 keV
Recoil-induced δ-rays	Delta electron from NR track origin	≈0	0
Particle-Induced X-ray Emission (PIXE)			
X-ray emission	Photoelectron near NR track origin	1.8	0
Auger electrons	Auger electron from NR track origin	19.6	0
Bremsstrahlung processes[a]			
Quasi-Free Electron Br. (QFEB)	Photoelectron near NR track origin	112	≈0
Secondary Electron Br. (SEB)	Photoelectron near NR track origin	115	≈0
Atomic Br. (AB)	Photoelectron near NR track origin	70	≈0
Nuclear Br. (NB)	Photoelectron near NR track origin	≈0	≈0
Neutron inelastic γ-rays	Compton electron near NR track origin	1.6	0.47
Random track coincidences			
External γ- and X-rays	Photo-/Compton electron near NR track	≈0	≈0
Trace radioisotopes (gas)	Electron from decay near NR track origin	0.2	0.01
Neutron activation (gas)	Electron from decay near NR track origin	0	0
Muon-induced δ-rays	Delta electron near NR track origin	≈0	≈0
Secondary nuclear recoil fork	NR track fork near track origin	–	≈1
Total background	Sum of the above components		1.5
Migdal signal	Migdal electron from NR track origin		32.6

Fe55 spectra with YOLO





Real time energy vs length output from the object detection pipeline