Reconstruction of Electromagnetic Shower Axis in the ECAL using Deep Learning Method

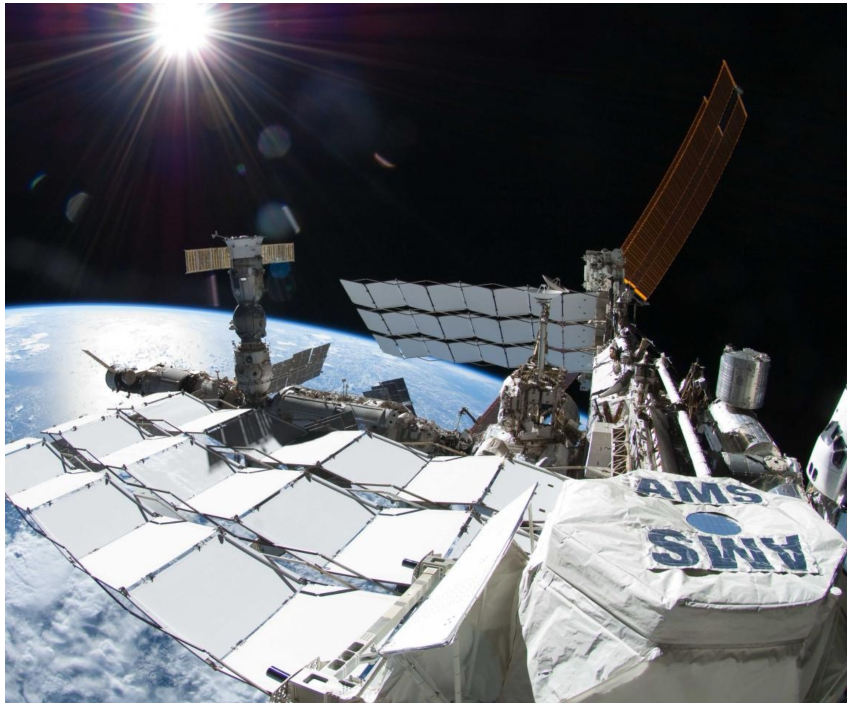
Yaozu Xiong, Hai Chen (ZJU)

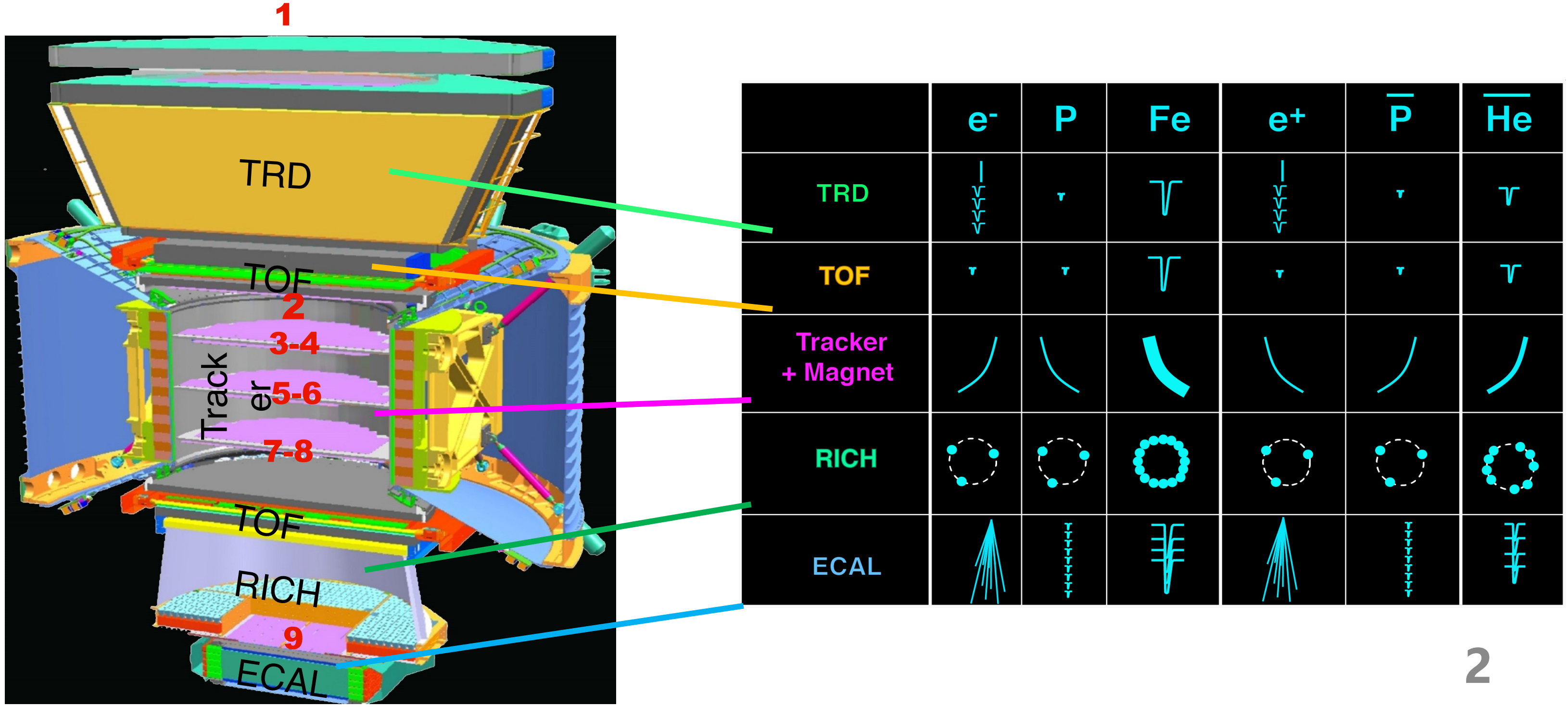
Quantum Computing and Machine Learning Workshop

2025/08/23, Qingdao

AMS-02 Experiment

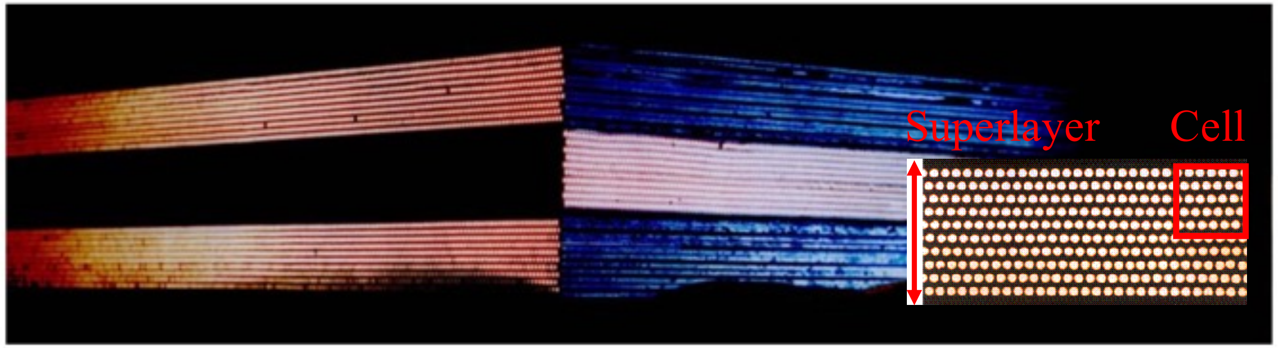
- The Alpha Magnetic Spectrometer is a high energy particle detector which was successfully deployed on the International Space Station since 2011 and has collected over 250 billion cosmic ray particles to date.
- Through precise measurements of the energy, charge, direction, and spectrum of cosmic ray particles, AMS-02 have revealed a series of new phenomena beyond current theoretical expectations and sparked extensive discussions on the origin, acceleration, propagation mechanisms of cosmic rays, and indirect detection of dark matter.

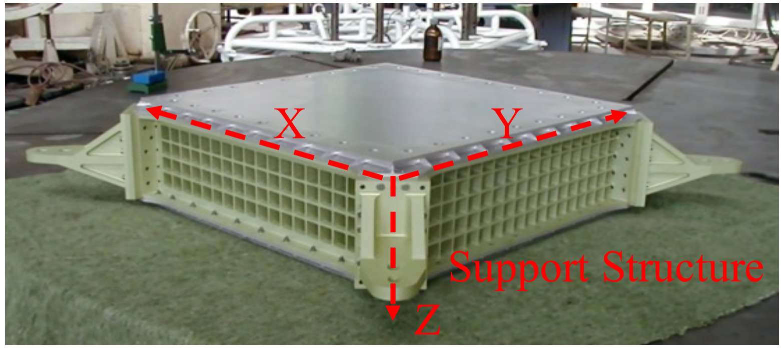


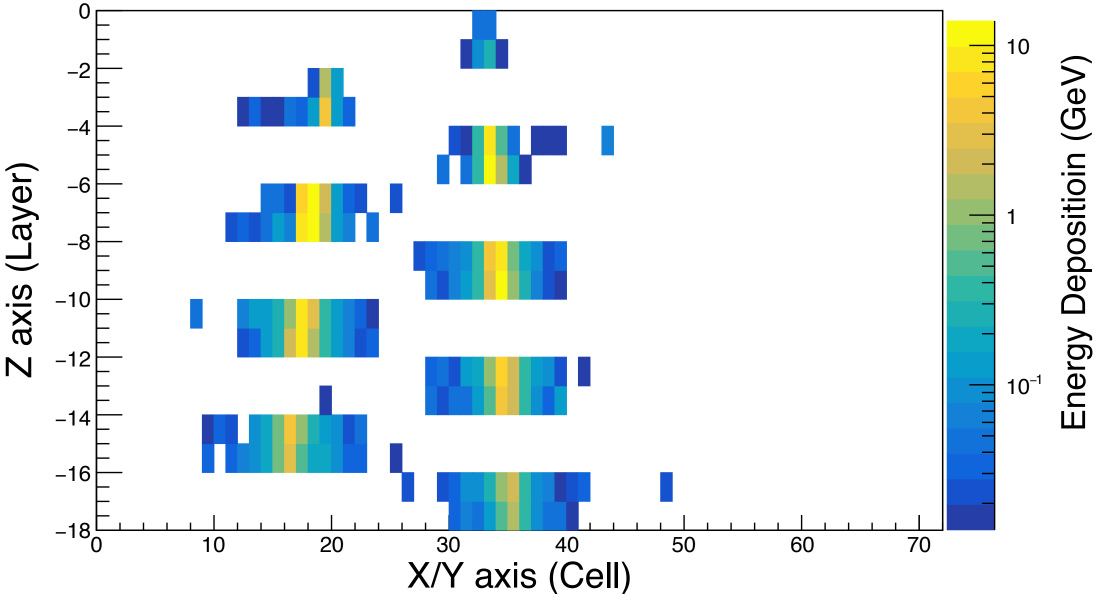


Electromagnetic Calorimeter

- The Electromagnetic Calorimeter (ECAL) in the AMS experiment is a 3D imaging detector;
- The detector consists of 9 superlayers with lead foils and scintillating fibers alternatively parallel to the x-axis (5 superlayers) and y-axis (4 superlayers);
- Read out by 324 PMTs. Each PMT has four anodes, and each anode covers an active area of 9×9 mm² (18x72 cell).



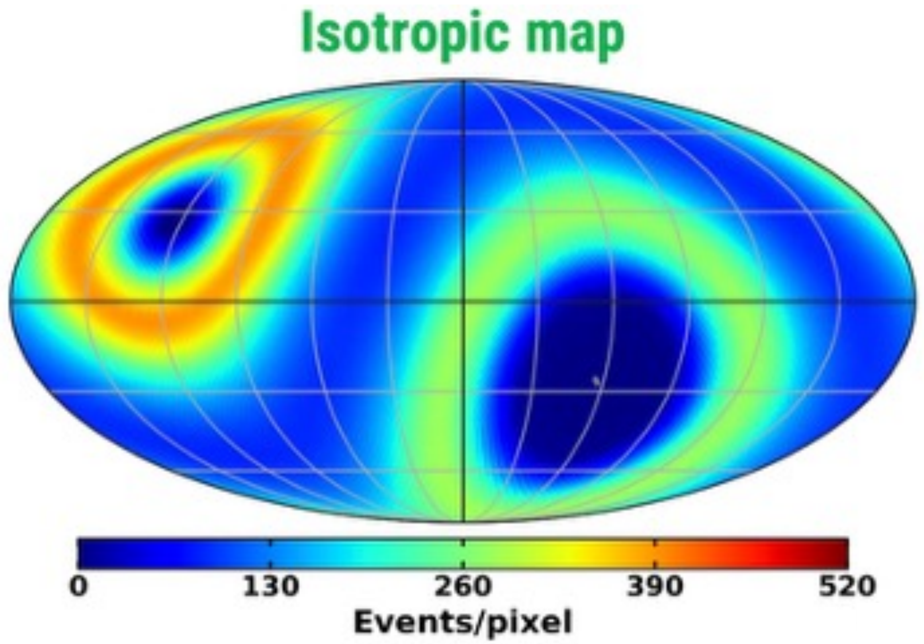


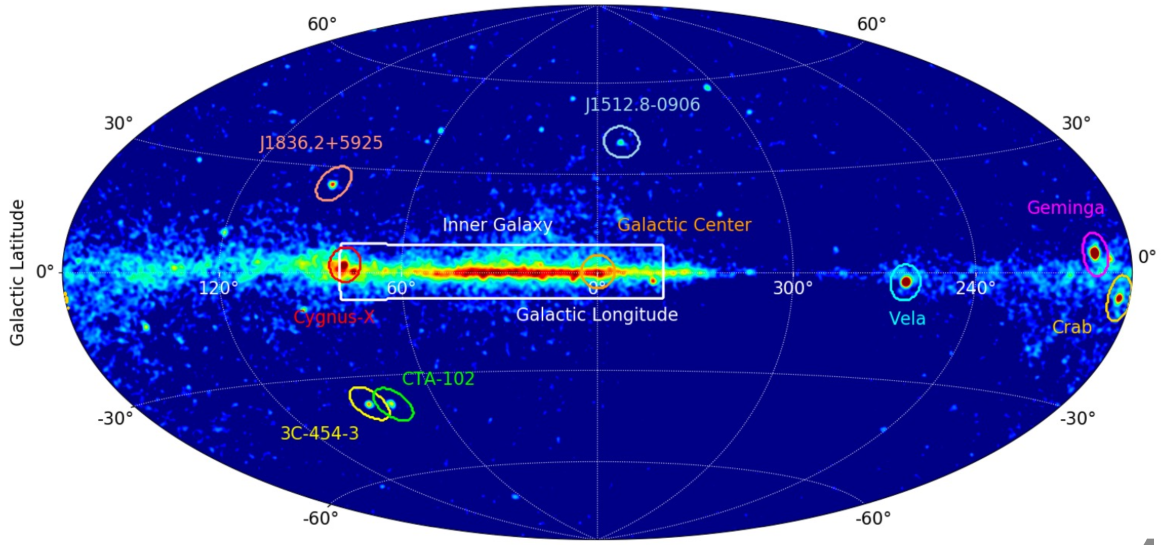


A 1.1 TeV electron event display in the ECAL

Physics Motivation

- Anisotropy of Electrons and Positrons: The direction reconstruction from the ECAL helps in better matching the track direction and position in the Tracker.
- High-Energy Gamma Analysis: For non-converting photons, ECAL is the only detector that can be used to measure the energy and direction of high-energy photons.
- By analyzing gamma ray from different directions, we can get the sky map of diffuse emission and catalog of γ-ray sources.



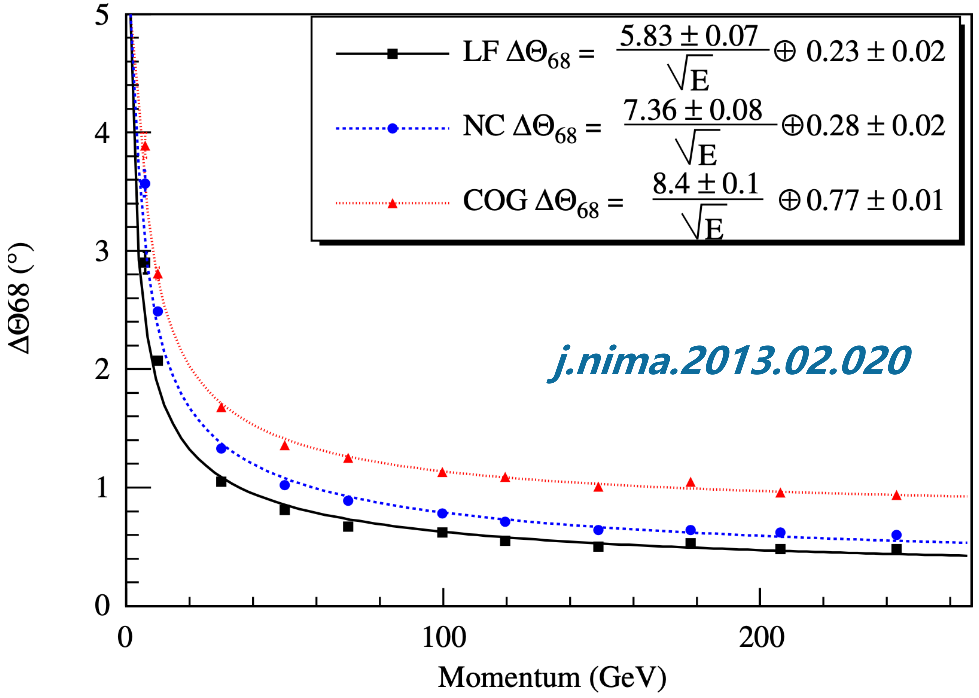


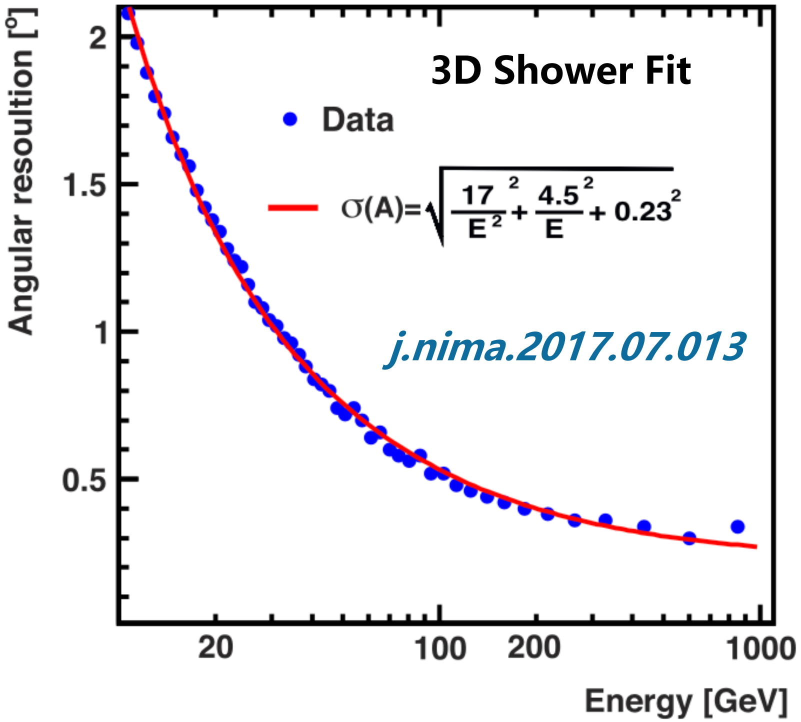
Shower Axis Reconstruction

- There are four conventional methods to reconstruct the shower axis in the ECAL:

  1. Center of Gravity 2. Neighbor Cell Ratios 3. Lateral Shower Fit 4. 3D Shower Fit

- With Deep Learning method, we aim to reconstruct the angular and spatial information of electrons, positrons and gammas in ECAL standalone.

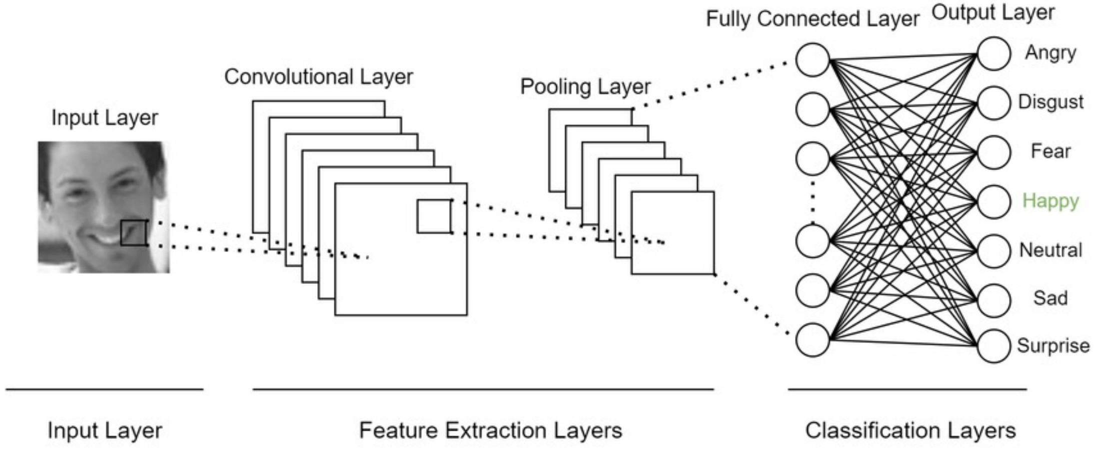




Deep Learning Model

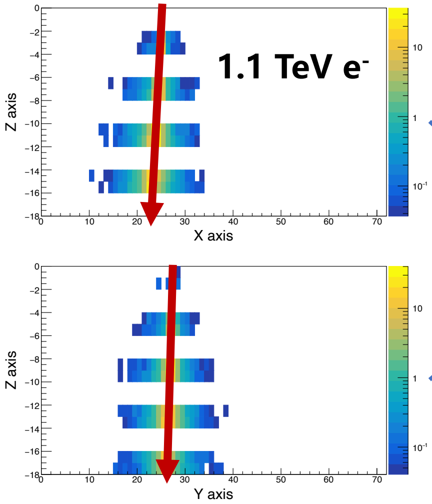
If we treat the energy deposition distribution of a electron/positron shower as a 3D image, can we benefit from computer vision / deep learning methods in classification or regression?

Input:
A face image

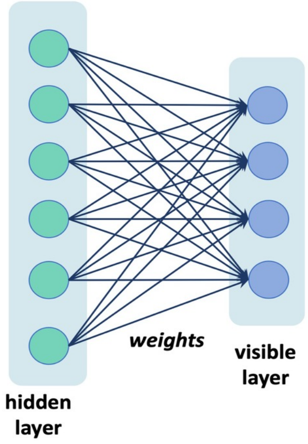


Output:
Facial emotion classification

Input:
ECAL image



Shower Feature Extraction Layers



output

Output:
Shower axis parameters

ResNet Architecture

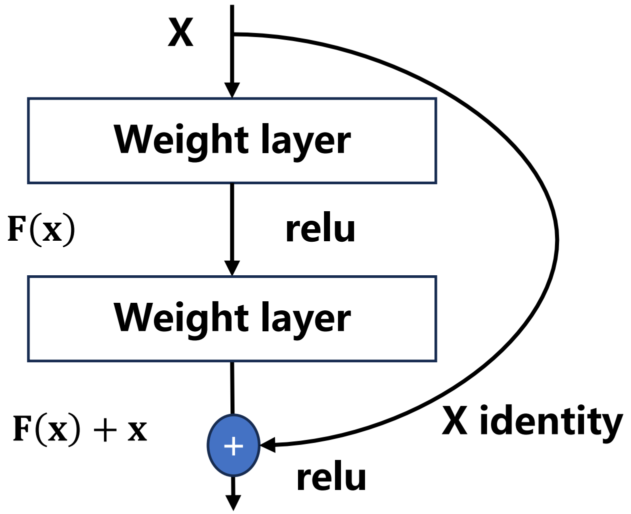
What is the ResNet?

Residual neural network (ResNet), a state-of-the-art architecture in the field of computer vision , was proposed by Kaiming He (MIT) and others in 2016. The main improvement of residual neural networks is the invention of "shortcut connections" to address the degradation problem, greatly eliminating the difficulty of training neural networks with excessive depth.

This model has been widely applied in various fields such as computer vision, particle physics, engineering, etc., with more than 268650 total citations.

Kaiming He (MIT)

$$\mathbf{H(x) = F(x) + x}$$

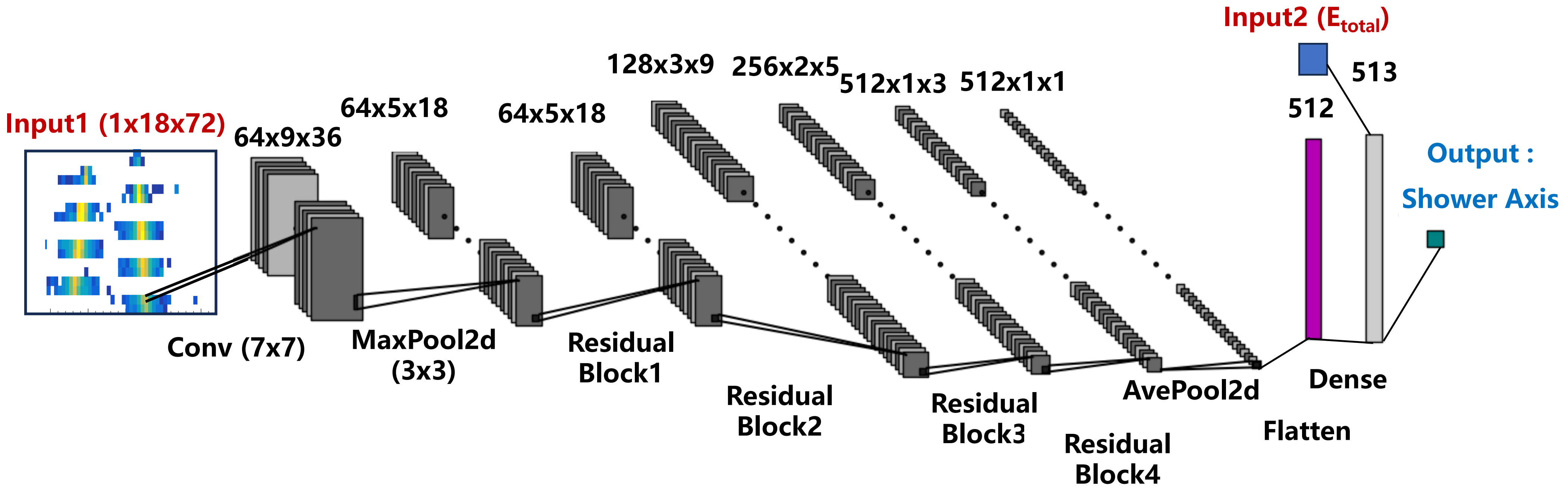


shortcut connections

Deep Residual Learning for Image Recognition K He, X Zhang, S Ren, J Sun Computer Vision and Pattern Recognition (CVPR), 2016, 2016	268650
Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks S Ren, K He, R Girshick, J Sun Neural Information Processing Systems (NIPS), 2015, 2015	88201
Mask R-CNN K He, G Gkioxari, P Dollár, R Girshick International Conference on Computer Vision (ICCV), 2017, 2017	43626

ResNet18 Model

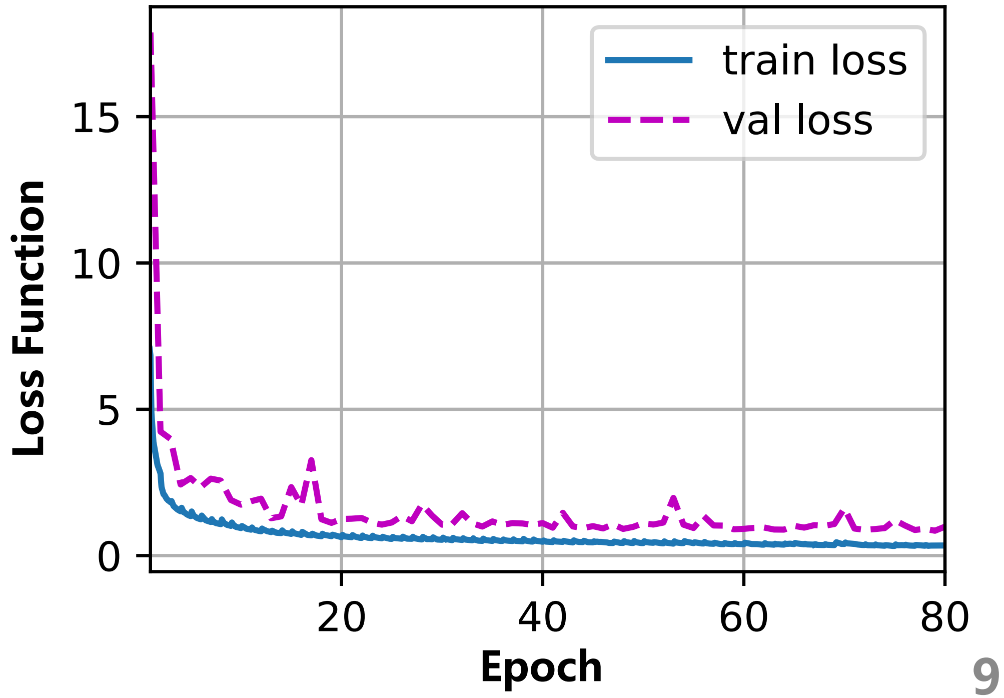
- ResNet18, 34, 50 are the most common ResNet architectures. The main difference lies in the number of residual blocks and network depth.
- ResNet18 is chosen in this study due to its lightweight architecture. It requires less computing resources, and is sufficient to our 18x72 ECAL input format. Our model is composed of 17 convolutional layers and a single dense layer.
- The total deposited energy E_{total} is used as an additional input, to learn and constrain the energy dependence of the model's output, ensuring self-consistency in physics.



Model Training

- We normalize the inclination angle $\theta_{\text{true}} = \frac{\theta_{tag}}{180}$ for ResNet learning, where θ_{tag} is the target value of θ_{resnet};
- To bolster model robustness, we choose data augmentation method by adding Gaussian noise (1%, due to calibration precision) in the 18x72 ECAL image.

- **Loss Function:** $\frac{1}{N} \times \sum_{i=1}^{N} \frac{(\theta_{\text{prediction},i} - \theta_{\text{true},i})^2}{\sigma_{\text{true}}^2}$;
- σ_{true} is the resolution vs energy obtained by iteration;
- Learning rate(stepsize): 0.001; batch size(block): 128;
- Optimizer(algorithm): Adam; training epoch: 80;
- GPU: NVIDIA RTX A4000, CPU: 8 × Intel(R) Xeon(R) CPU E5-2686 v4 @ 2.30GHz;
- Time cost: about 8 hours for training and validation.



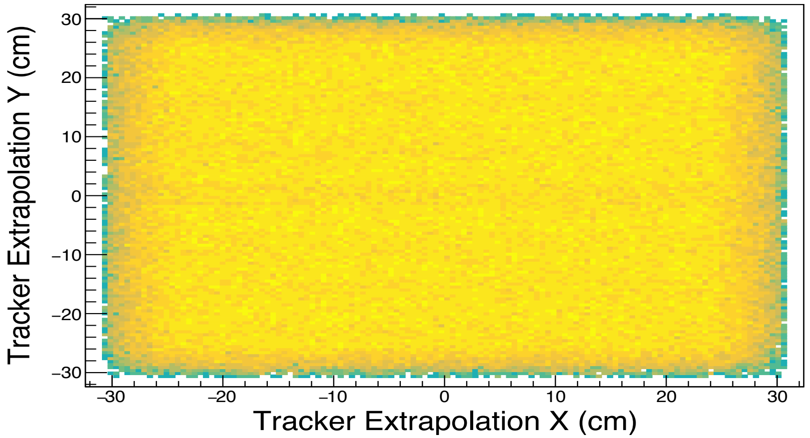
Dataset for ResNet

- We train the model using MC electron sample, then validate it with MC positron sample and ISS data;
- The Tracker extrapolation angle on the surface of ECAL as the reference (truth) direction;

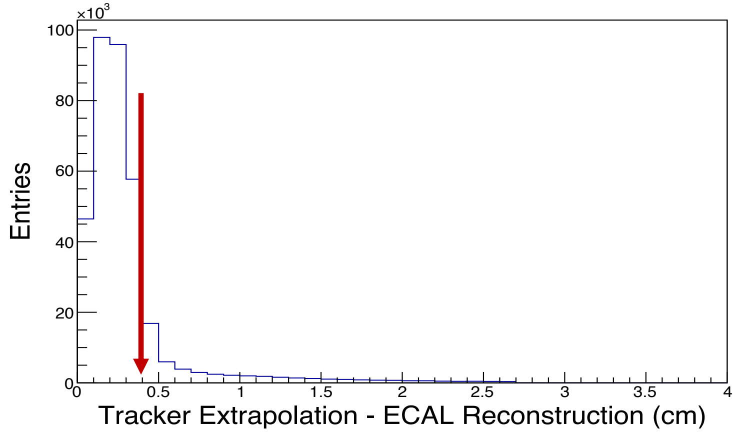
Data Type	Electron MC (el.B1236)	Positron MC (pos.B1236)
Truth Energy	2.5-4000 GeV with power law flux E^{-1}	2.5-4000 GeV with power law flux E^{-1}
Incident Angle	θ: 155-180°	θ: 155-180°
Total Event Number	$3*10^6$, 60% Training, 10% Validation, 30% Test	$9*10^5$

- To improve training data quality:
  - Shower entry & exit with 2 cells away from border;
  - Loose cuts on Tracker-ECAL matching (within half cell size)

| Shower Entry & Exit | < 30.6 cm

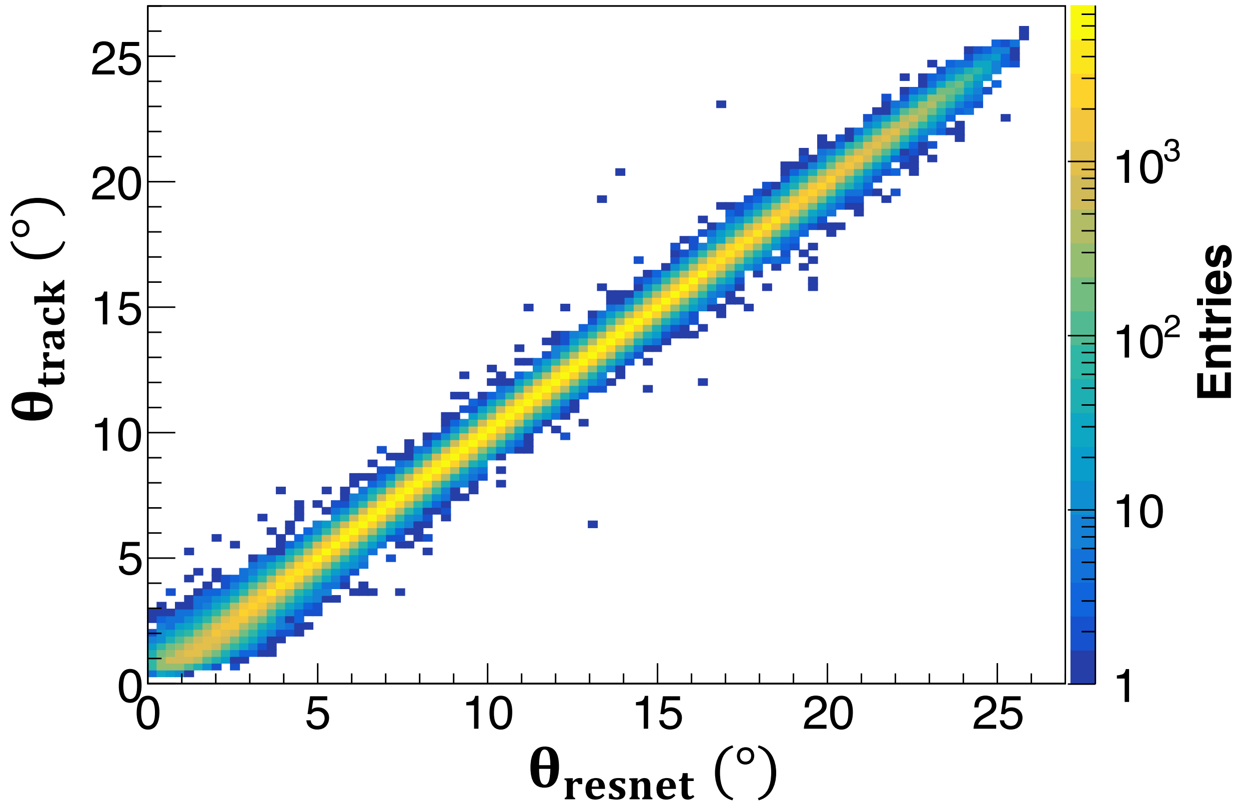


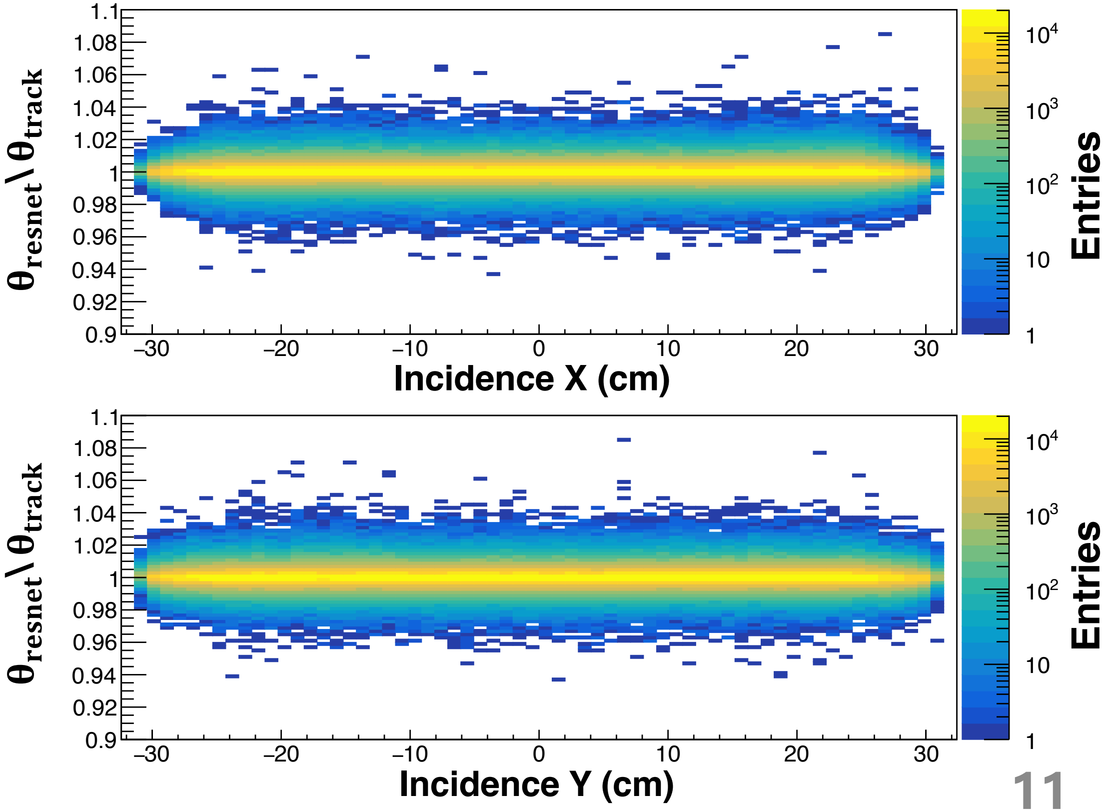
| Tracker - ECAL | < 0.45 cm



Performance in MC Electron

- ResNet model can effectively reconstruct the showers ranging from vertical incidence up to 25° inclination angle within Tracker acceptance;
- Furthermore, the dependence on incident position (X/Y) of tracker extrapolation are verified.



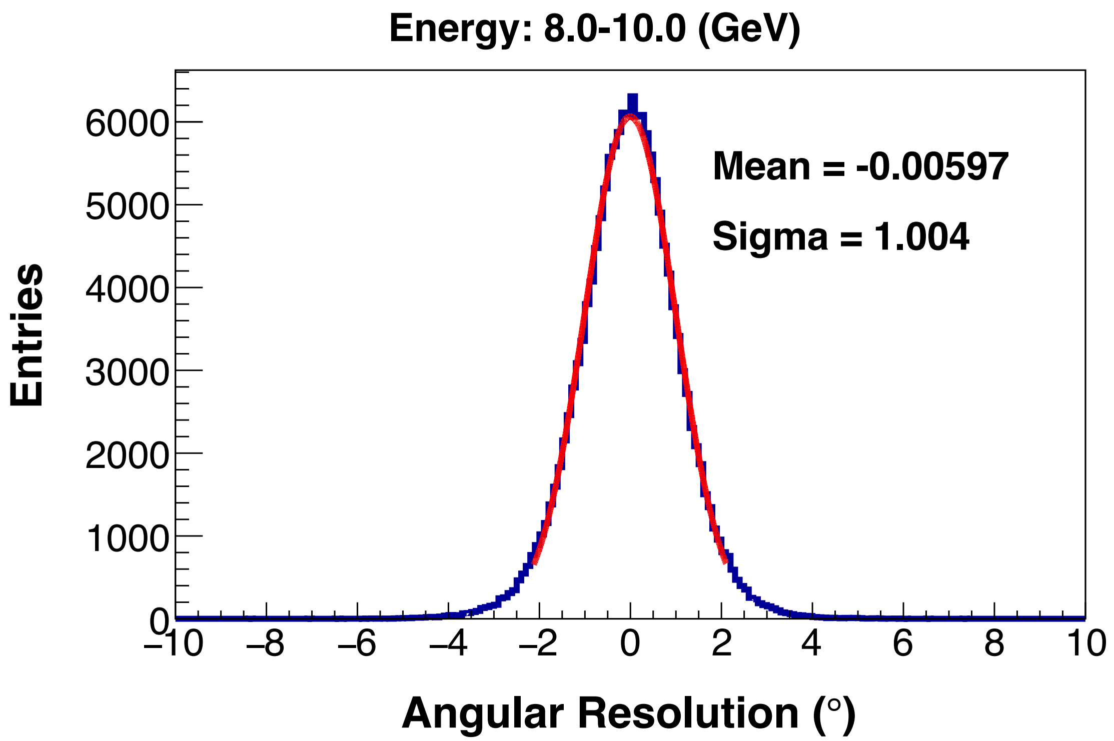


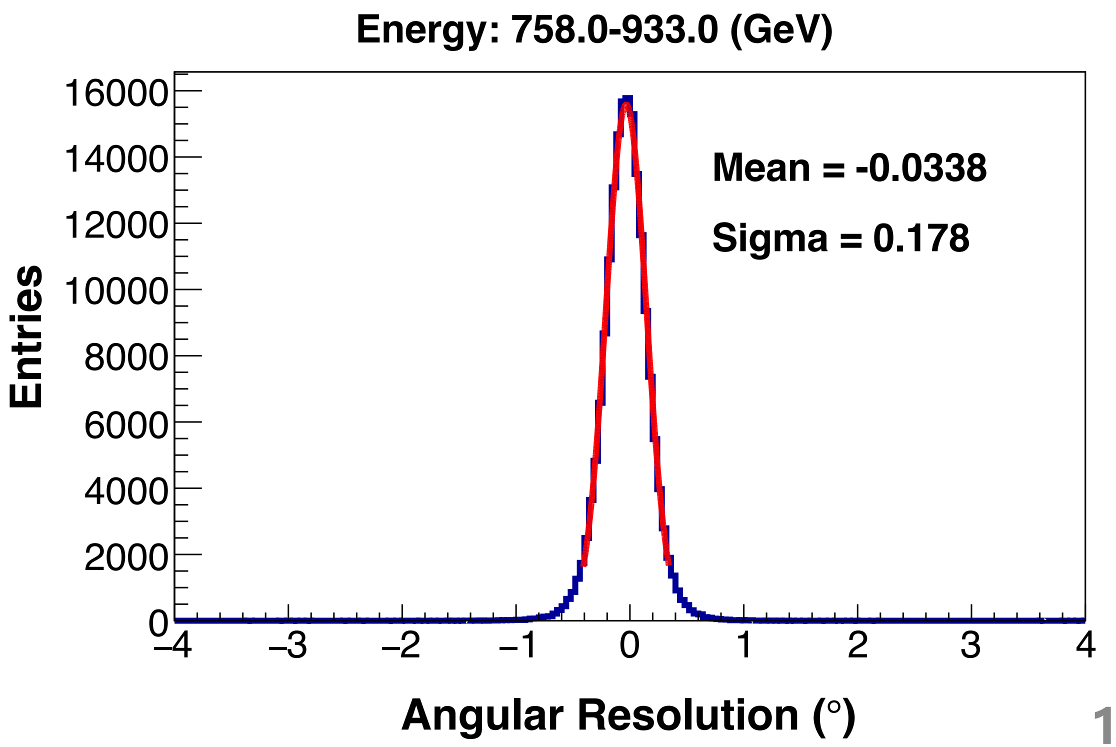
Residual on MC Electron

- **Defining the residual as the difference between the model's reconstructed result and the regression target:**

$$\sigma = \theta_{resnet} - \theta_{tag}$$

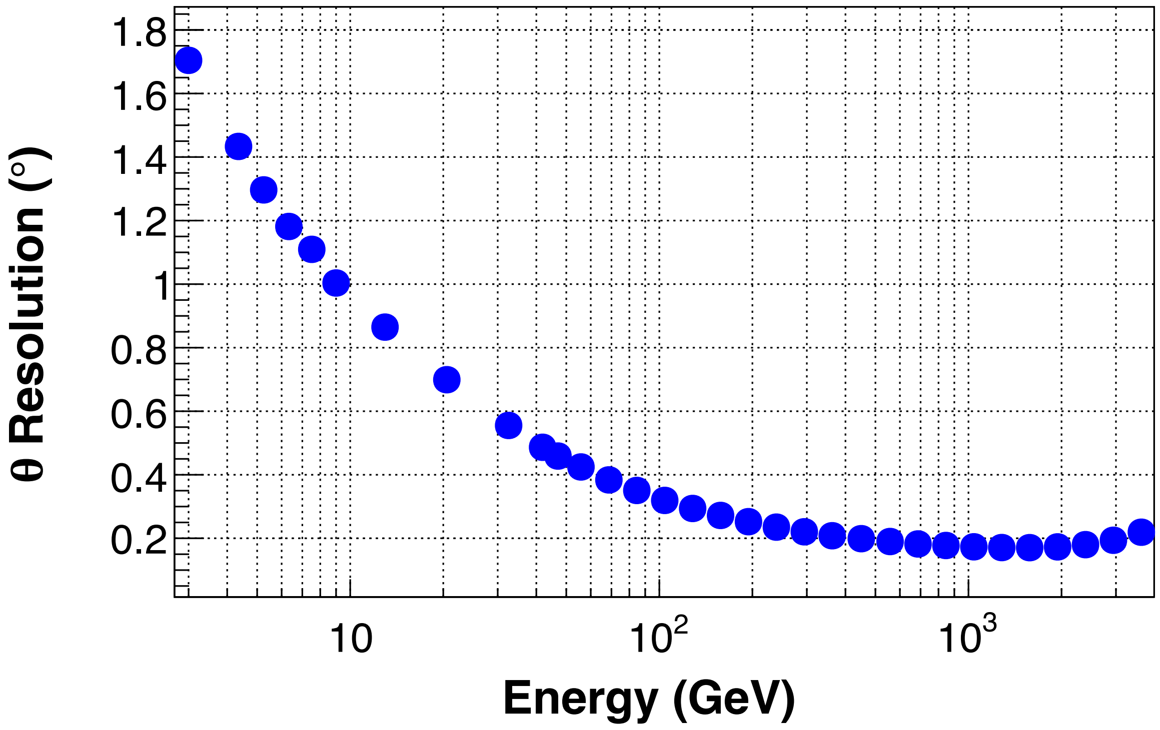
- **Below are two examples of Gaussian fit to the residual in low energy and high energy range.**

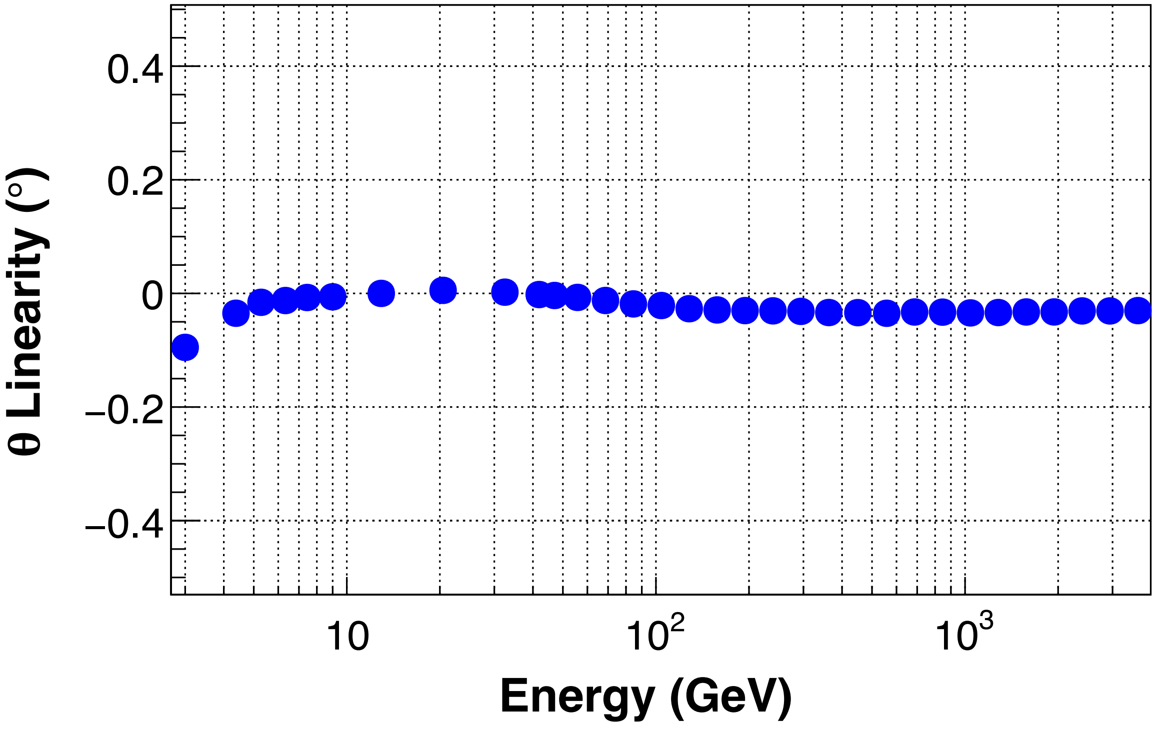




Angular Resolution and Linearity

- The ResNet resolution reaches about 1° at 10 GeV and 0.2° above 300GeV;
- Meanwhile, the linearity is stable and almost within 0.05° in the full energy range.

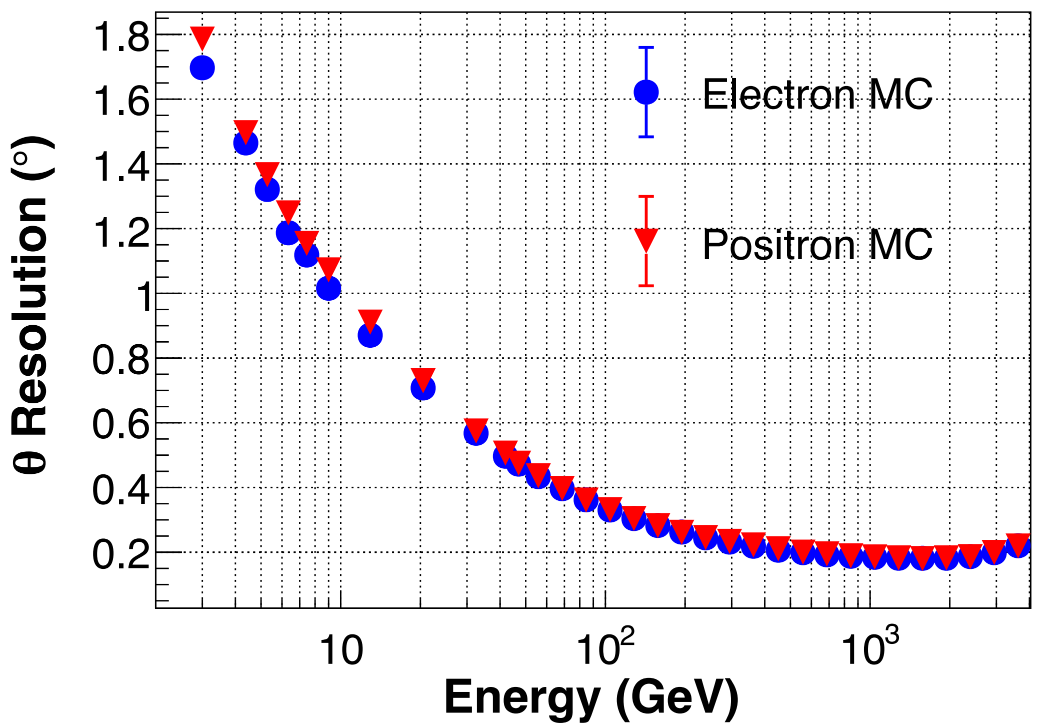


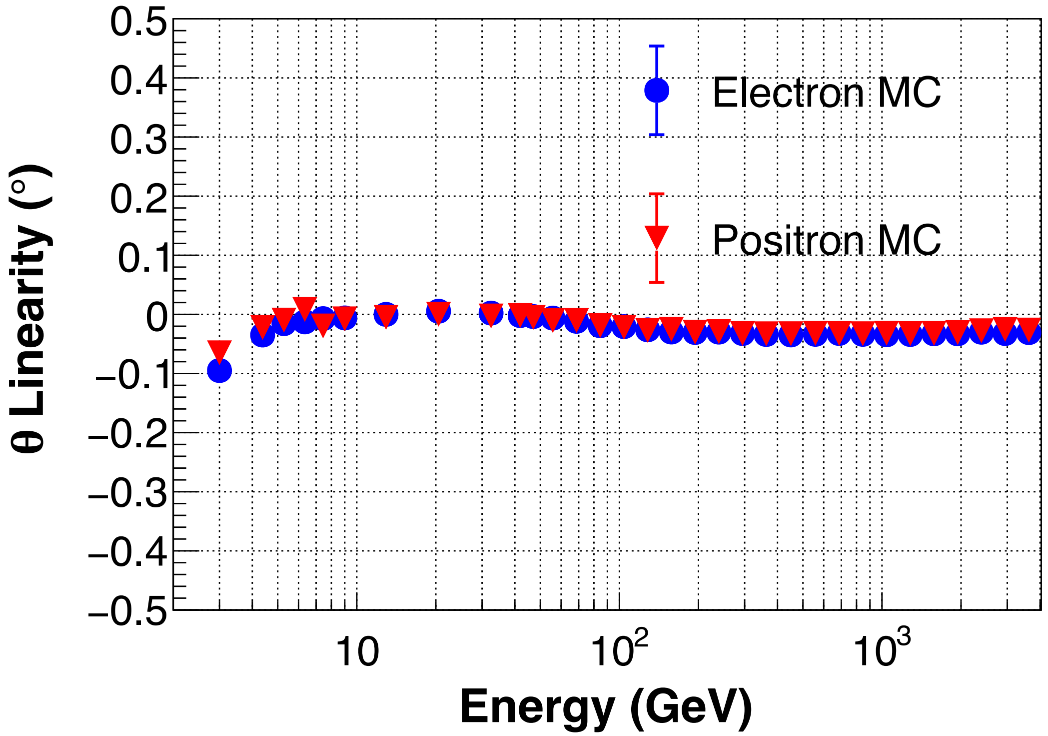


Positron Sample Validation

- As an independent crosscheck, we apply the training weights on the the positron sample for verification;
- The reconstruction result is almost identical to the electron sample.

Positron MC (pos.B1236)





ISS Data Sample Selection

- AMS has a unique and high purity ISS electron sample after 14 years operation on the International Space Station.
- In order to verify the performance of angular model on the ISS electron data sample, we apply the following event selections with ECAL, Tracker and TRD.

ECAL

- Number of hit > 100
- Number of shower > 0
- BDT > 0.5
- Likelihood < 3.5

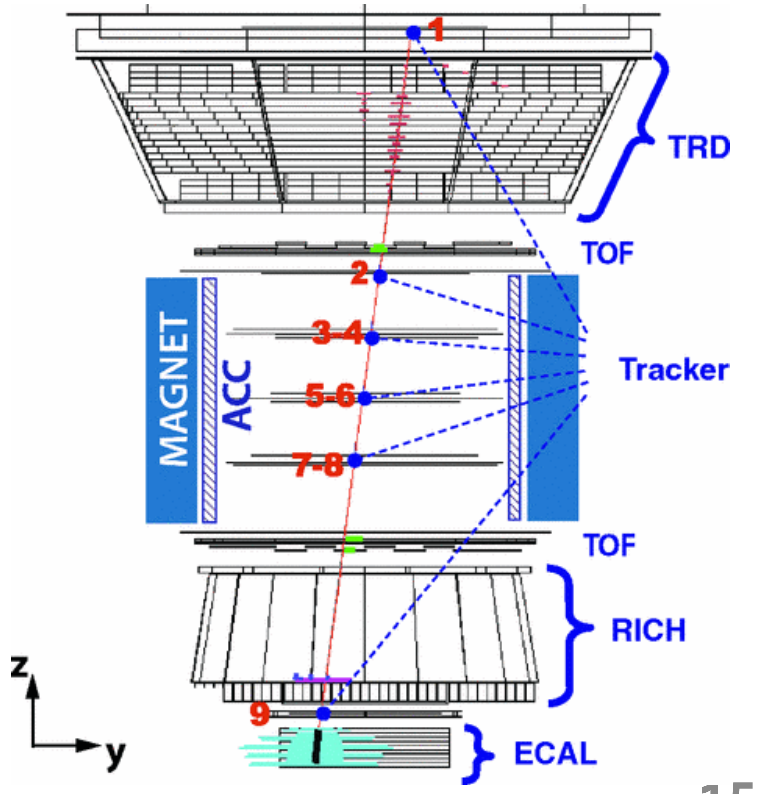
TRD

- Number of track > 0
- LikelihoodRatio <0.6

Tracker

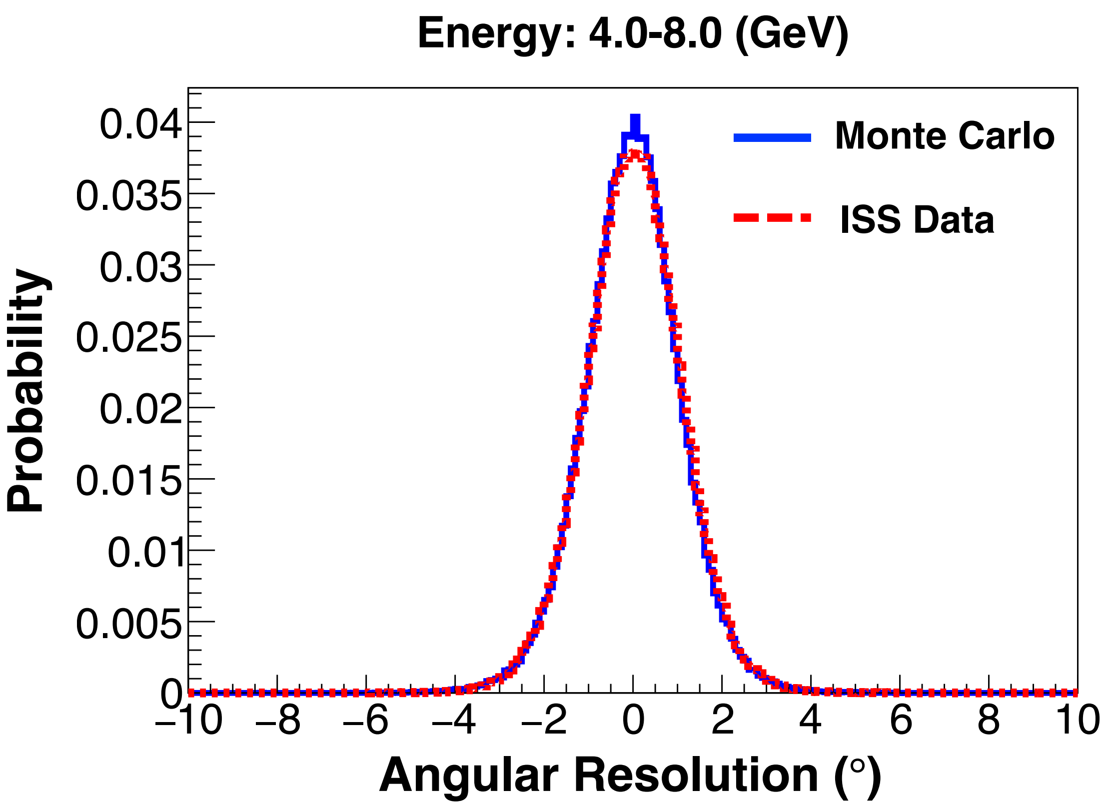
- Number of track = 1
- $Rigidity_{inner+L1} < 0$
- $0.8 < Charge_{inner+L1} < 1.2$
- $| Tracker_{inner+L1} - ECAL | < 0.4$ cm
- Shower incidence point < 30.6
- Shower exit point < 30.6
- $Log(|Energy/ Rigidity_{inner+L1}|) > -2$ and < 4

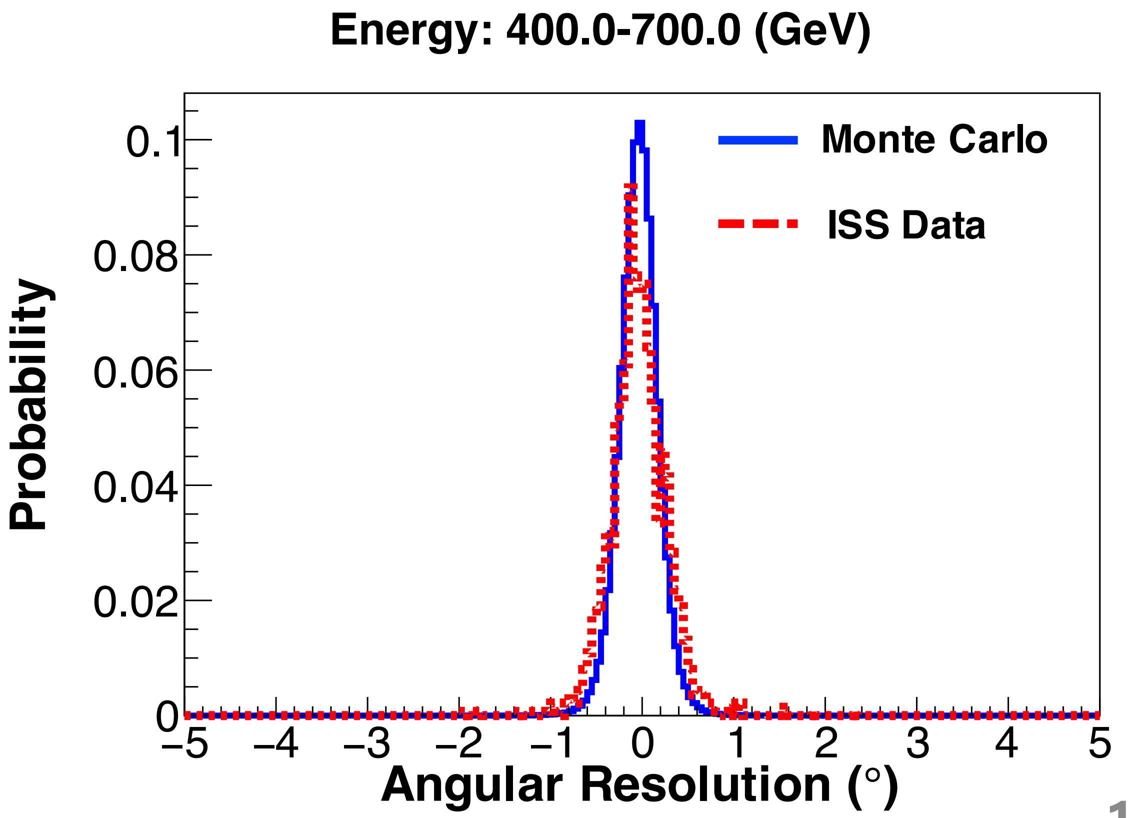
These selections are widely used for electron and positron analysis within AMS Collaboration.



Residual Comparison of ISS and MC

- An example of Gaussian fitting for the difference between the reconstructed angle and the regression target is shown.
- Good consistency between ISS data and Monte Carlos simulation is observed.





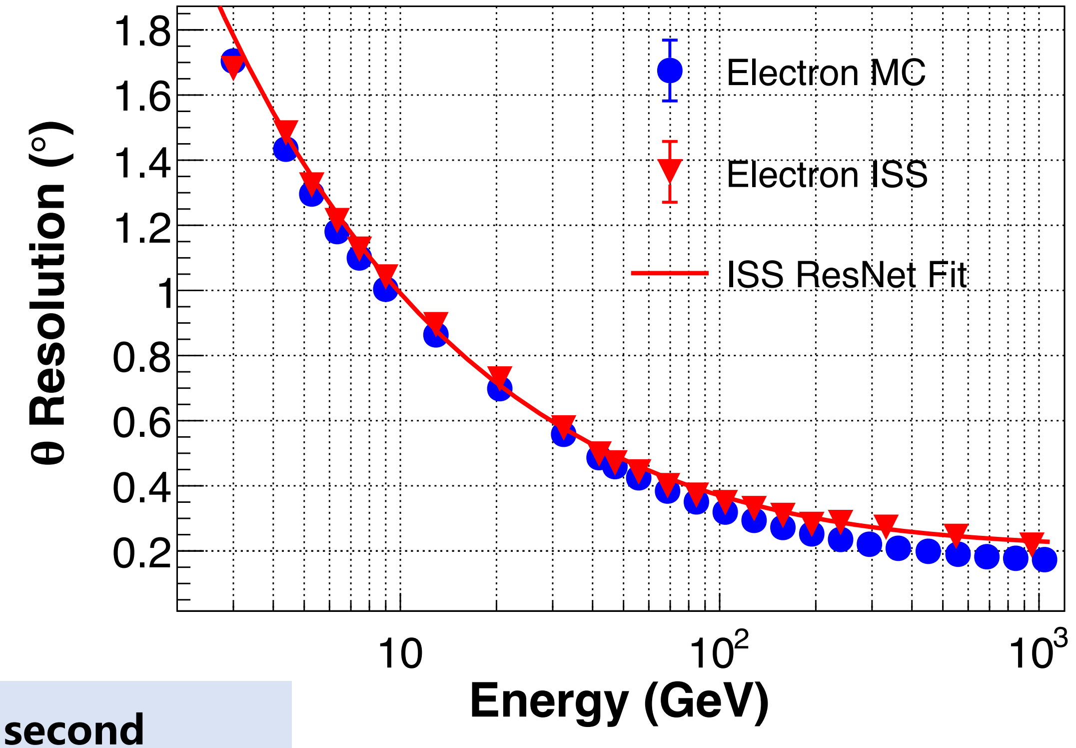
Angular Resolution on ISS

- Consistence between the ISS data and MC is observed in the full energy range from 2 GeV to 1.0 TeV.

- The slight discrepancy in the high energy region (above 200 GeV) could be attributed to detector calibration non-uniformities and proton contamination in the ISS sample;

- The angular resolution function of ResNet model in ISS data can be fitted with:

$$\sqrt{\frac{3.0764^2}{E} + 0.199^2}$$

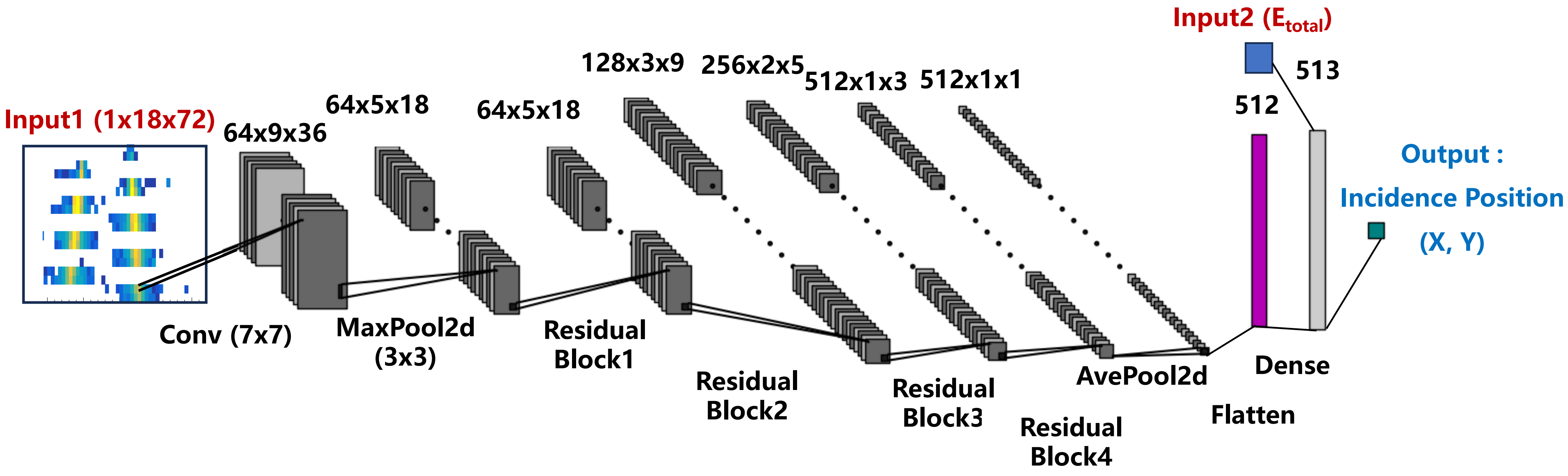


- The first term originates from the shower fluctuation, and the second constant term is mainly related to the granularity of ECAL PMTs.

ResNet Model for Incident Position

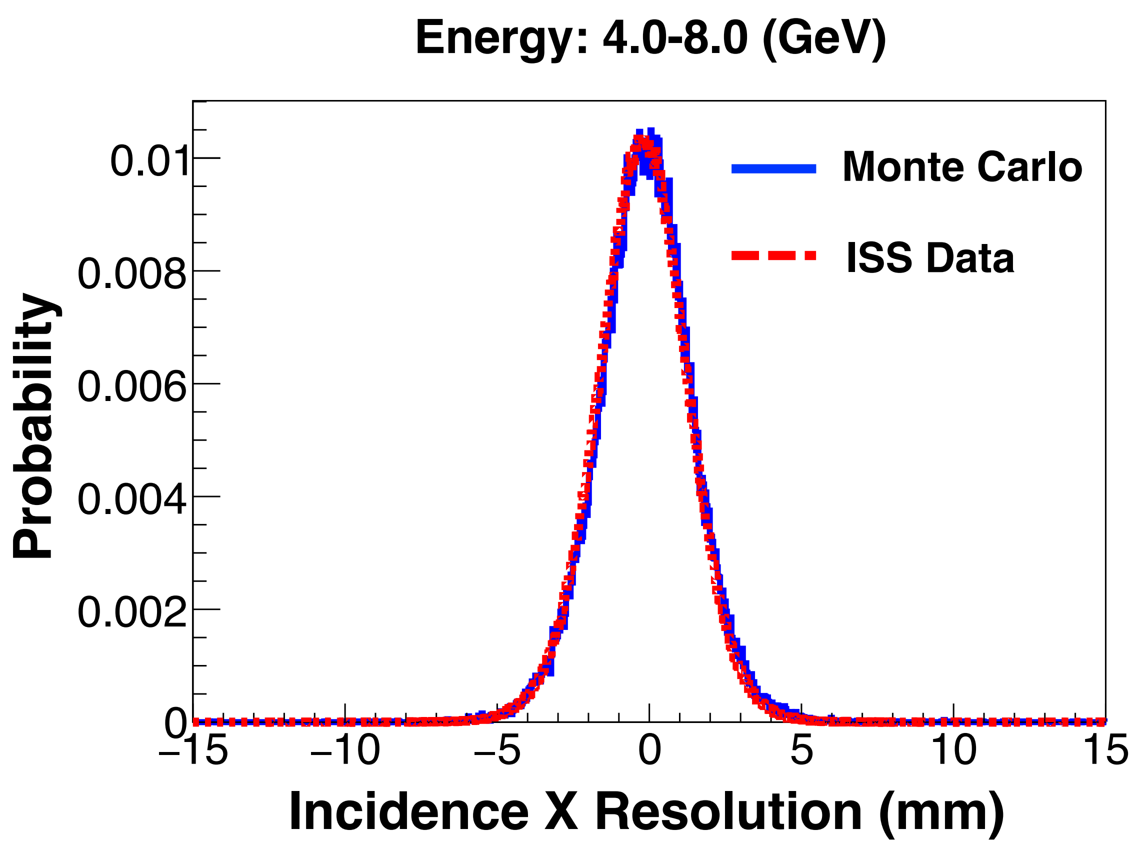
- By training the model with the same network structure, we can reconstruct the incident position as well;
- The minor difference lies in the output form, which is a 2D vector array of (X_0, Y_0).

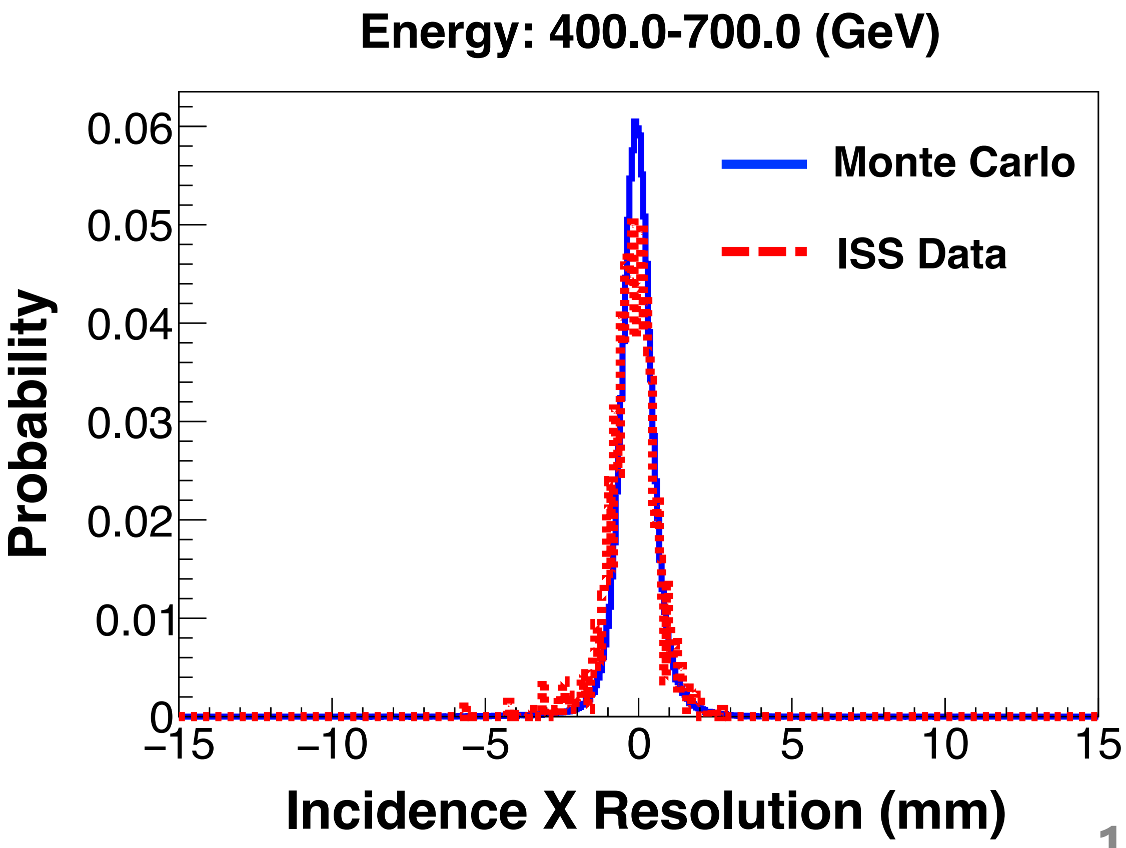
Loss function: $$\frac{1}{N} \times \sum_{i=1}^{N} \frac{\left(X_{\text{prediction,i}} - X_{\text{true,i}}\right)^2}{\sigma_{X,\text{true}}{}^2} + \frac{\left(Y_{\text{prediction,i}} - Y_{\text{true,i}}\right)^2}{\sigma_{Y,\text{true}}{}^2}$$



Residual Comparison of ISS and MC

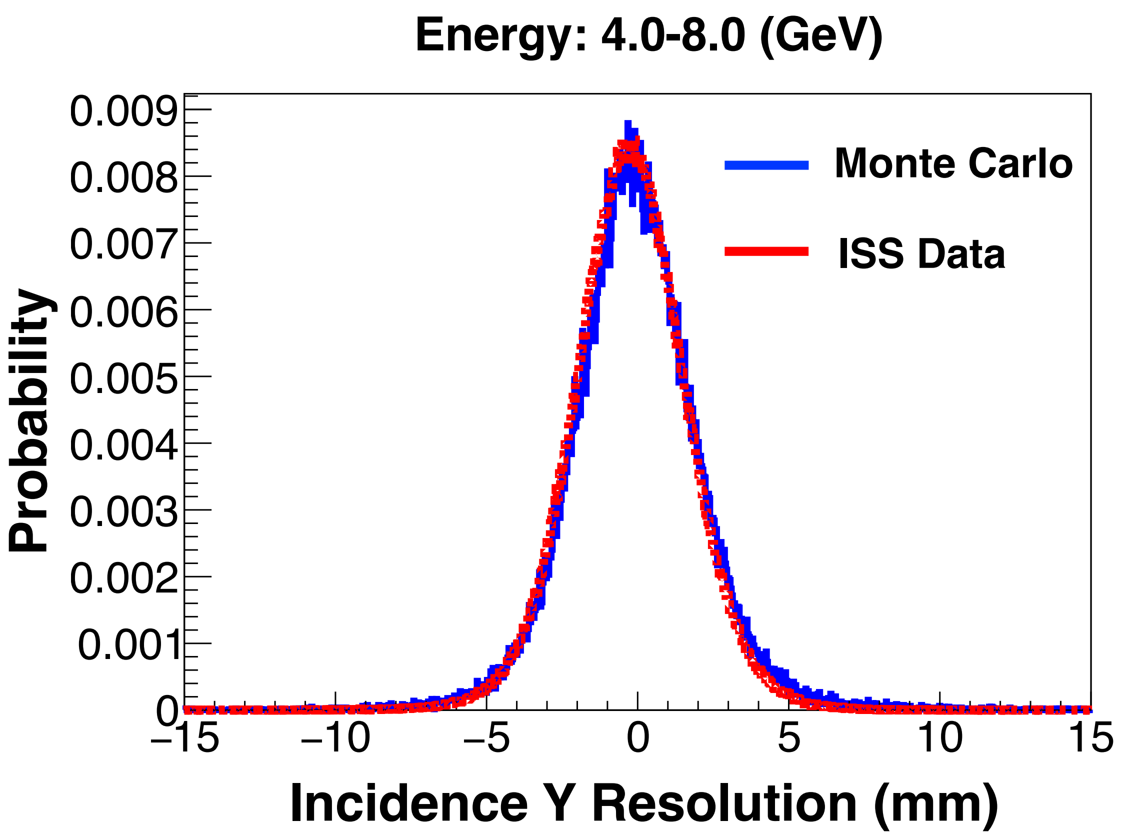
- **In the X-direction, an example of Gaussian fitting for the difference between the reconstructed coordinates and the regression targets is shown.**

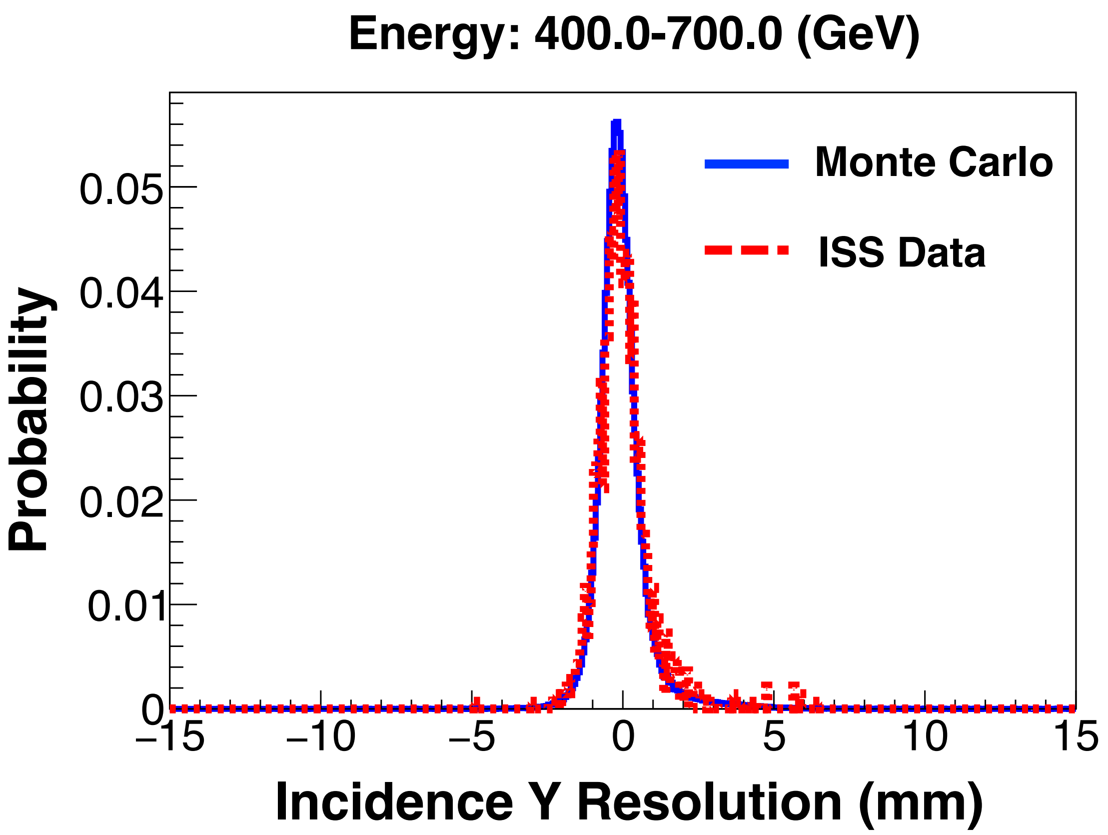




Residual Comparison of ISS and MC

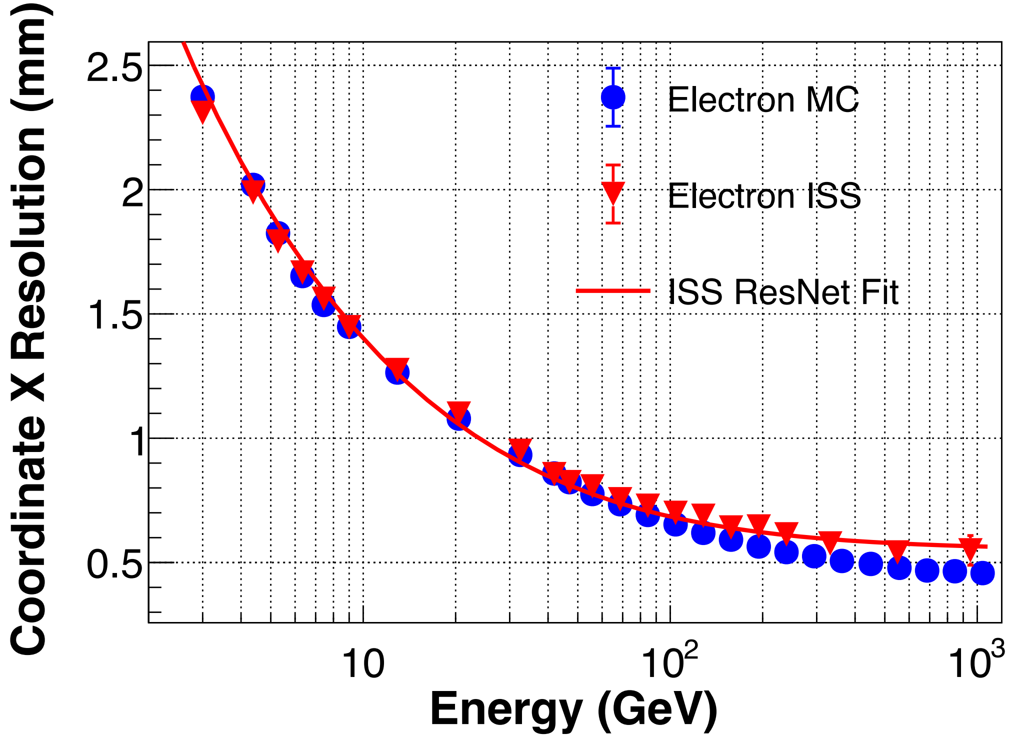
- In the Y-direction, an example of Gaussian fitting for the difference between the reconstructed coordinates and the regression targets is shown.

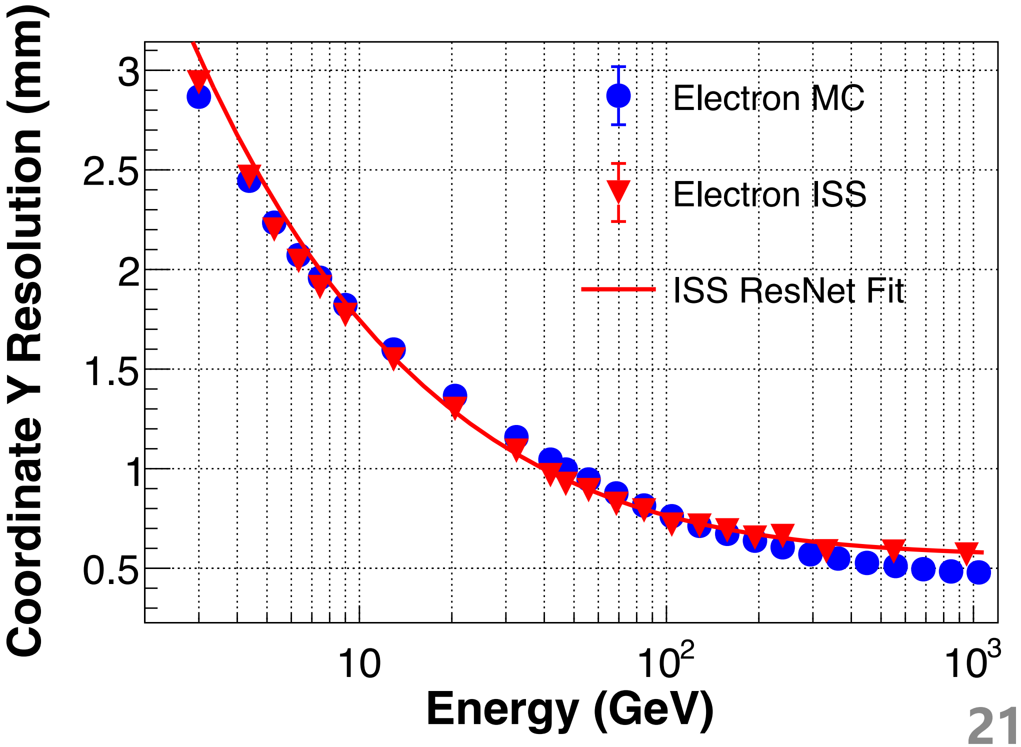




Incidence Resolution on ISS

- Spatial resolution is better than 1.5 mm in X direction and 2.0 mm in Y direction for electrons at 10 GeV;
- The resolution reaches ~ 0.5 mm above 300 GeV;
- The resolution in X is better than that in Y, primarily due to more number of layers along X (10/18).





Summary

- In this work, we utilize Deep Learning technique (ResNet Model), to reconstruct the angular and incident position of electrons and positrons.
  - ✓ The inclination angle θ shows a resolution of 1° at 10 GeV and 0.2° at 1TeV in Monte Carlo sample.
  - ✓ The performance is further validated with unique AMS ISS electron data.
  - ✓ The incident position (x_0, y_0) is reconstructed by the dual-output model, reaching 0.5 mm precision above 300 GeV.
- Compared to existing methods, the ResNet model demonstrates significant improvement, which is of great importance for the positrons, electrons, and high-energy gamma rays analysis in AMS experiments.
- Additionally, this Deep Learning technique can be applied into the physics reconstruction of other high granularity calorimeters (HERD, CMS…).