Electron Energy Reconstruction in HERD with Machine Learning Method

Liao Chuanli

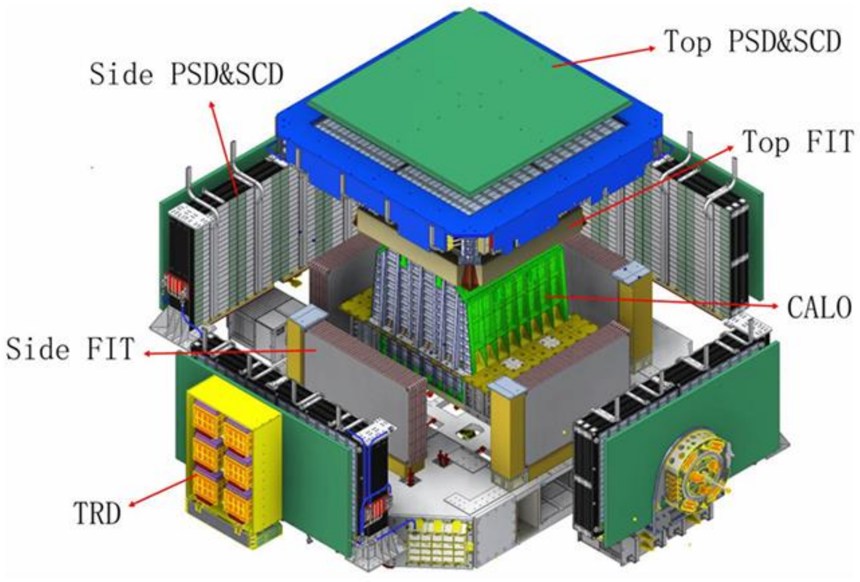
2025-08-22

Outline

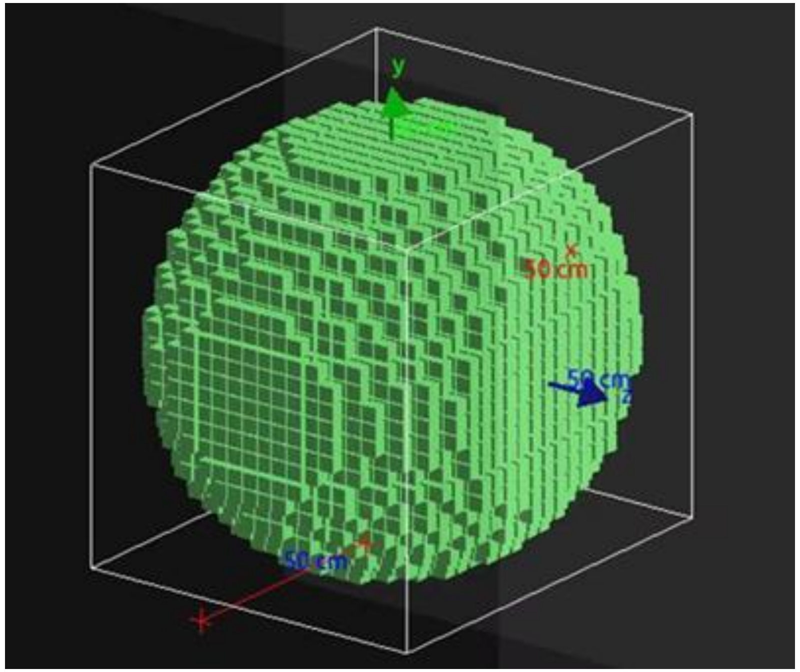
- Brief introduction of HERD
- Energy reconstruction algorithm based on deep learning
  - Motivation
  - Dataset
  - ML Model
  - Multi-class multi-prediction model
  - Reconstruction results
- Transform learning
  - data augmentation
  - beam data validation
- Summary

Brief introduction of HERD

- The High Energy Cosmic-Radiation Detection Facility (HERD) is scheduled to be installed on the China Space Station in 2028, operating in orbit for at least 10 years。
- Scientific Objectives: 1) search for dark matter; 2) precisely measure the cosmic ray spectrum and composition; 3) survey the high-energy gamma-ray sky.
- The HERD calorimeter features an innovative design with three-dimensional imaging and five-sided sensitivity. HERD is expected to achieve breakthroughs in both dark matter and cosmic ray research.



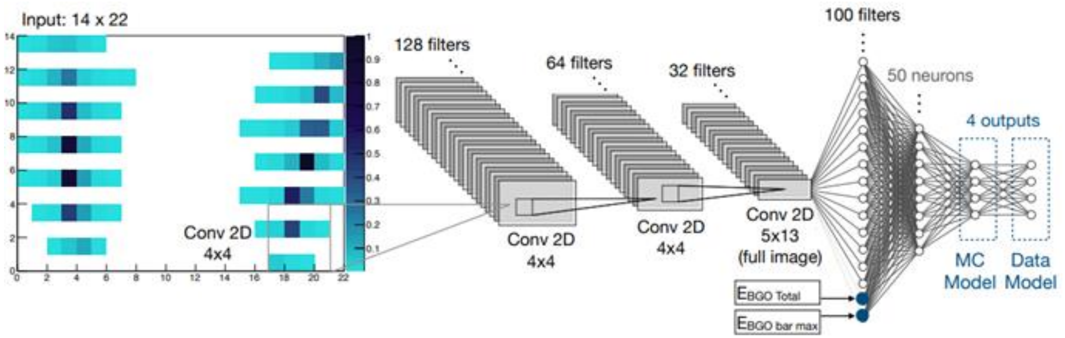
The payload of HERD

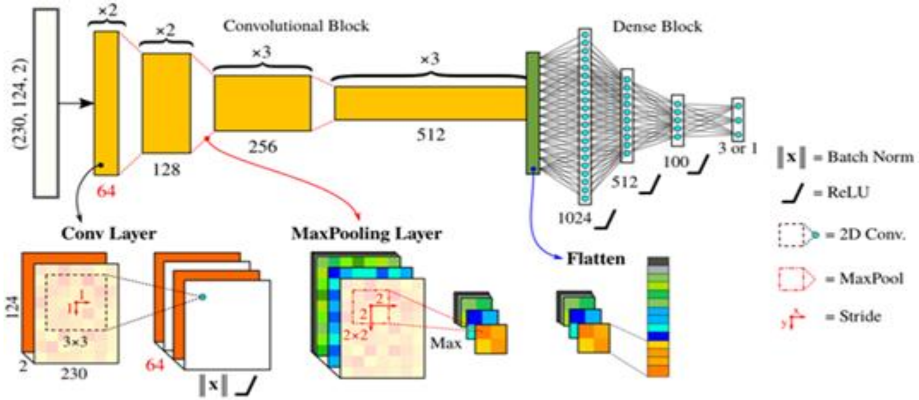


HERD CALO 3D simulation

Motivation

◆ ML is widely used in HEP





◆ particles Incident from top with small angle → all directions

- Traditional methods
  - Reconstruct shower principal axis
  - Iterative fitting parameters
  - Correct shower leakage
- Machine learning
  - Intelligence and automation
  - Improve reconstruction efficiency

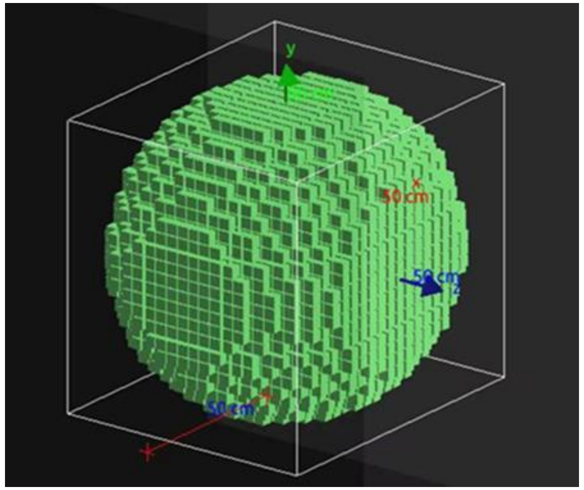
Item	Traditional Methods	Machine Learning
Incident Axis Reconstruction	Center of Gravity Method (COG)	Principal Component Analysis (PCA) Convolutional Neural Network (CNN)
Particle Identification	TMVA (BDT, MLP, FISHER)	ParticleNet GNN
Energy Reconstruction	Shower Correction 3D parametric fitting method	Deep Neural Network (DNN) Convolutional Neural Network (CNN) Residual Neural Network (ResNet)

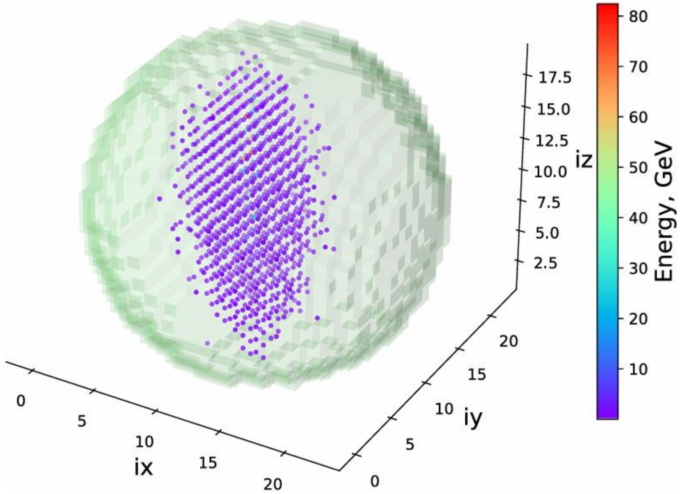
Dataset

- Simulation data based on HERD offline software
- Train：
  - 600,000 electrons uniformly distributed from 0.1 to 1000GeV and incident from all directions
- Test：
  - Fixed energy electron incident from all directions at energies of 10, 50, 100, 200, ..., 1000 GeV. Each energy point contains 10,000 test samples.
- Preprocess
  - Poisson fluctuate
  - Response threshold
  - Event selection
  - Feature extraction
  - Normalization

Table 1 The extracted feature information from the simulation includes crystal cube information and integrated information

Parameter	Name	Type[× size]
Monte Carlo truth		
Event ID		int
Simulated energy	E_{p}	float
Crystal cell information		
Number of crystal cell	N_{p}	int×7489
Energy deposition of crystal cell	E_{xyz}	float×7489
Position of crystal cell	x, y, z	int×3×7489
Fired crystal	B_{xyz}	Bool×7489
Integrated information		
Energy of projection	E_x, E_y, E_z	float×[23+23+21]
Number of fired crystal of projection	N_x, N_y, N_z	int×[23+23+21]
Energy deposition	E_{vis}	float

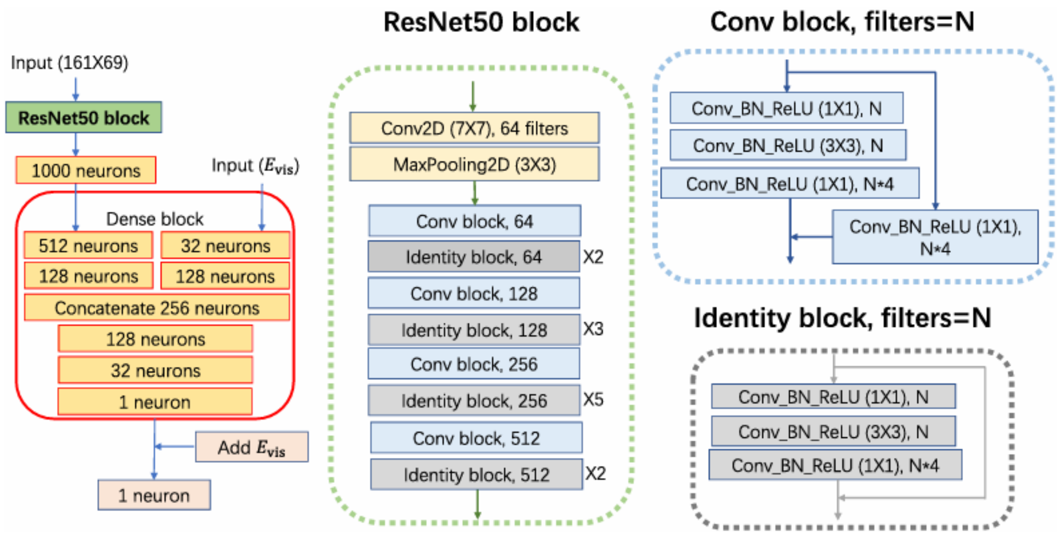




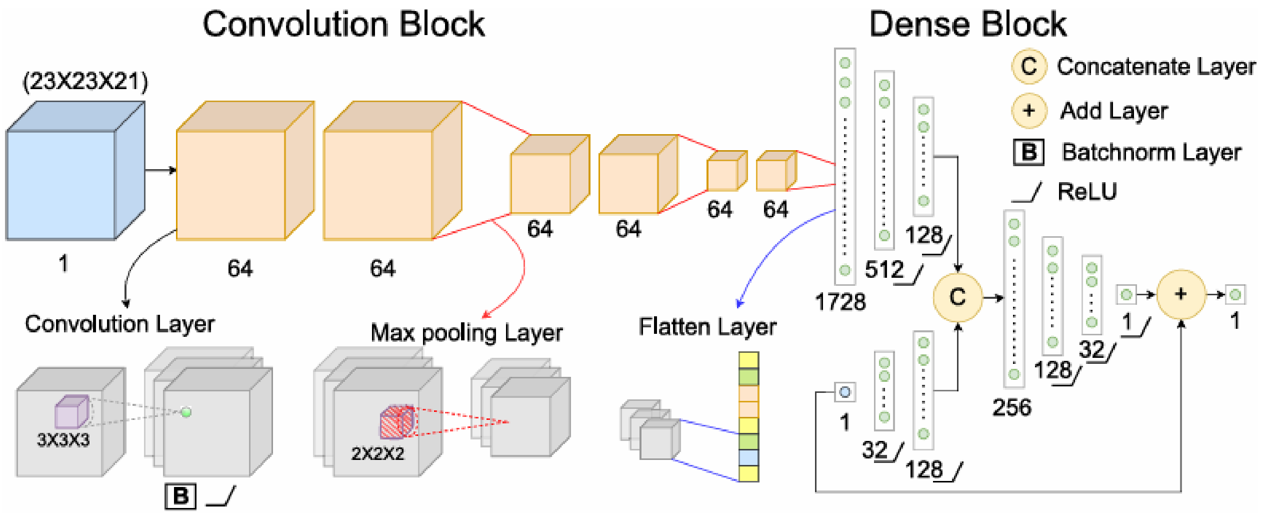
ML Models

- **Input:**
  - DNN: integrated features
  - CNN: 3D energy deposition images
  - ResNet: the cropped and spliced versions of the 3D images.
- **Output:** indirect prediction with end-layer visible energy correction. The main part of the models predicts the energy loss.

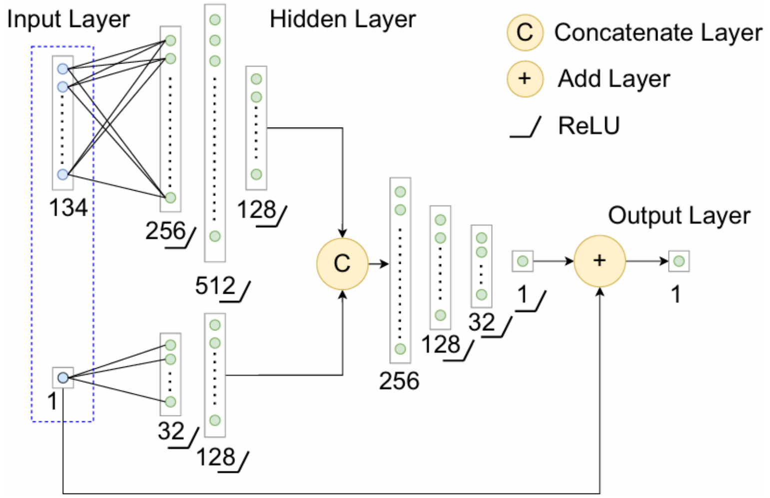
- ResNet



- CNN



- DNN



Multi-class multi-prediction model

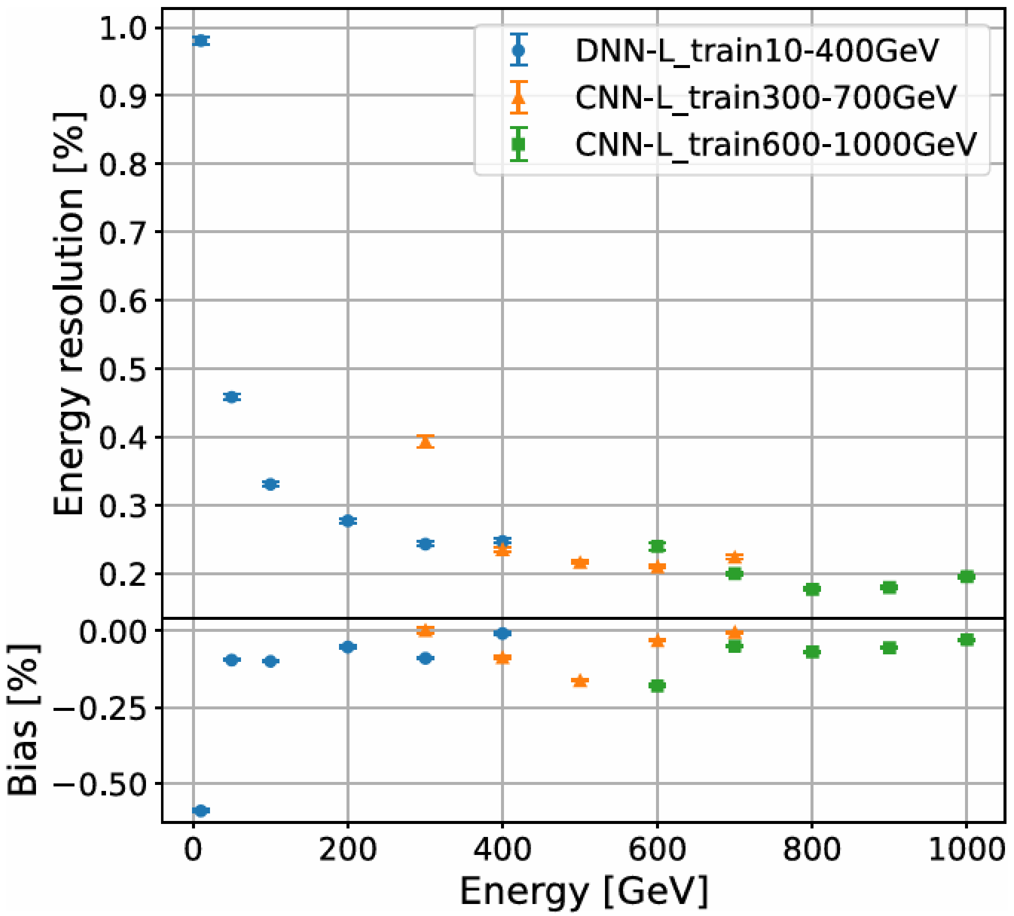
- Combined model: make full use of the advantages of different models
  - An appropriate range of training energy can improve model performance
  - The performance of different models varies at different energy range
- Classification + regression
  - coarse prediction: approximate energy range
  - Fine prediction: precisely predict energy

➢ Training strategy：

Table 2 The strategy of Multi-class multi-prediction model

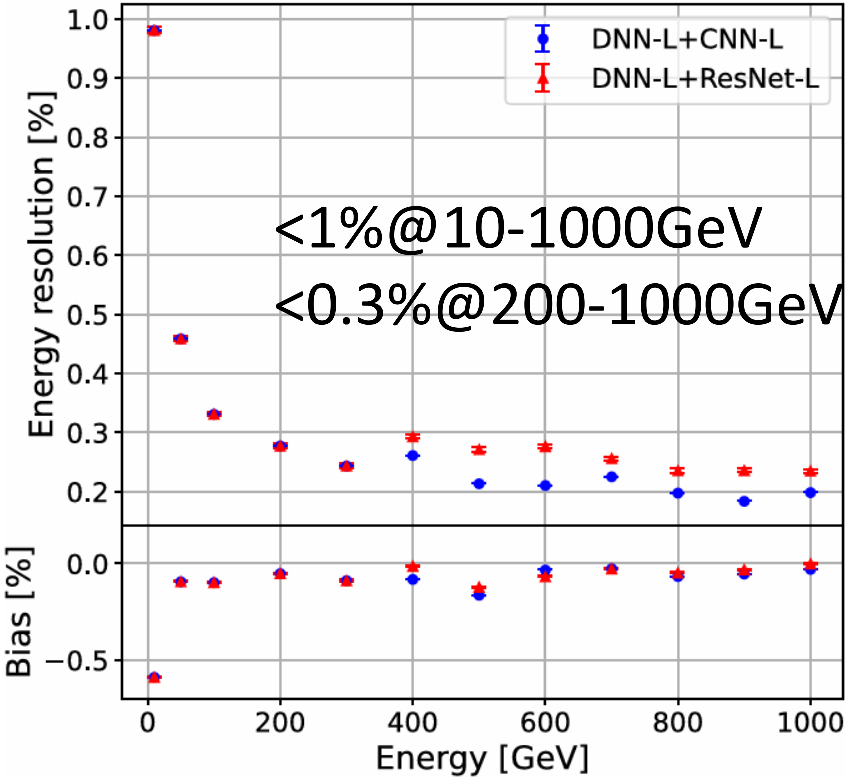
Model	Energy range	Loss
DNN-L	10−400 GeV	Mean square of relative error
CNN-L	300−700 GeV 600−1000 GeV	Mean square error
ResNet50-L	300−700 GeV 600−1000 GeV	Mean square error

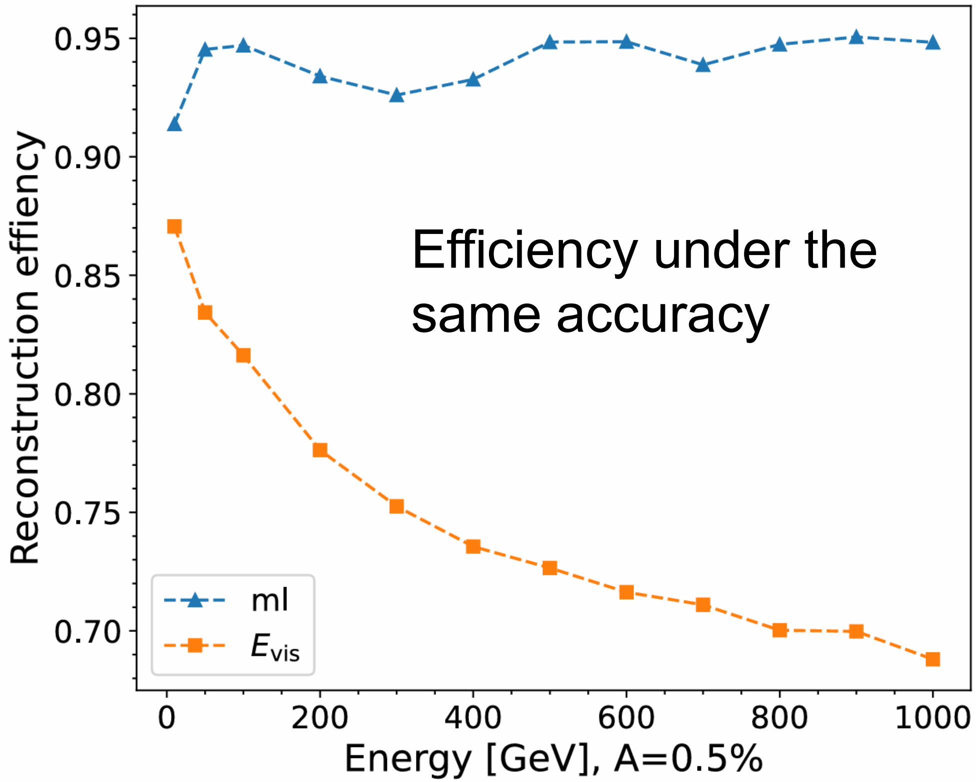
➢ DNN+CNN：



Results

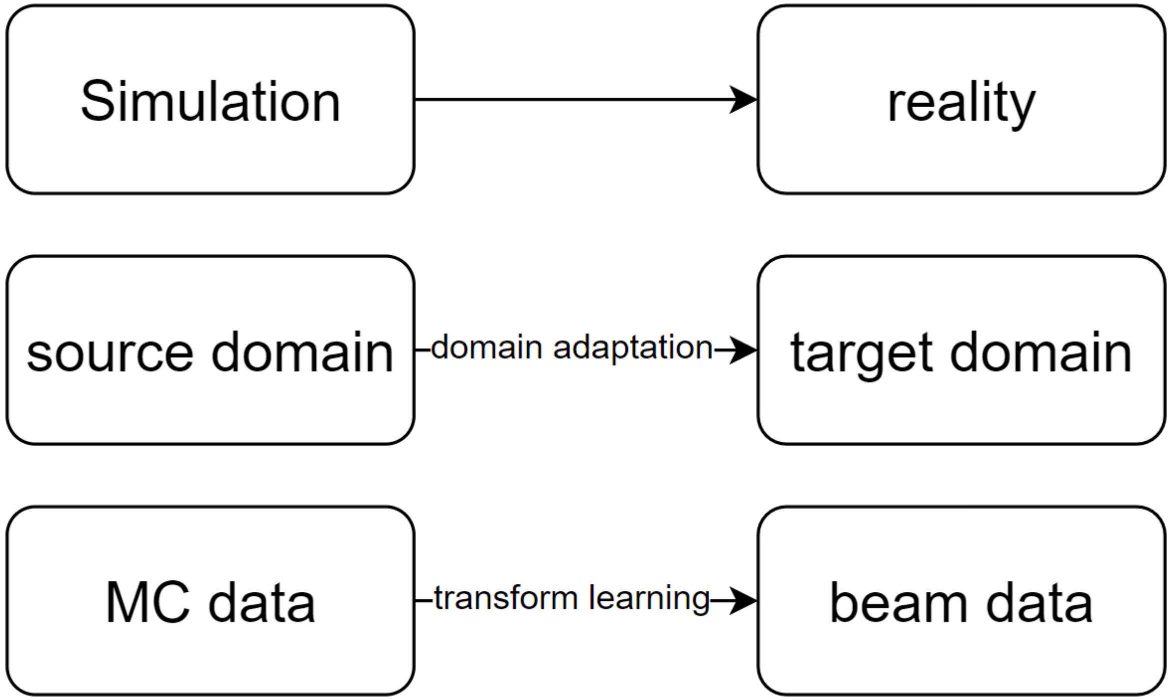
- The multi-class muti-prediction machine learning model achieves an energy resolution <1%@10-1000GeV， <0.3%@200-1000GeV
- The ResNet model slightly underperforms compared to the CNN in the high-energy range.
- Multi-class multi-prediction model improves the reconstruction efficiency while maintaining high resolution

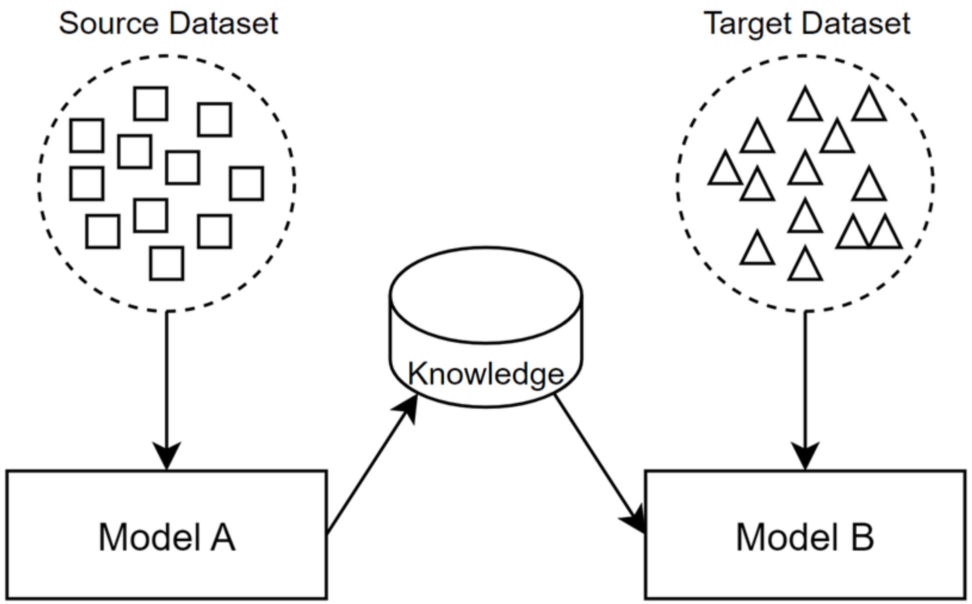




Transform learning

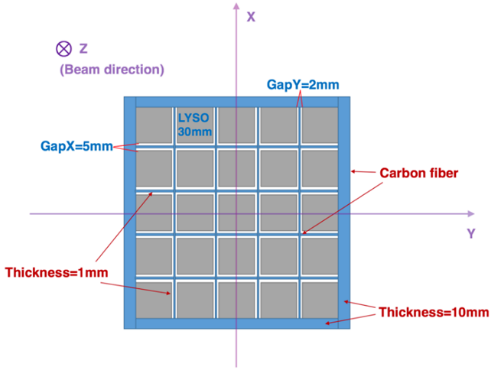
- Difference between simulation and reality
- Domain Adaptation: Adapting models trained on the source domain to the target domain.
- Transfer Learning: Applying knowledge or patterns learned from one domain or task to a different but related domain or problem.
- Objective: To enable models trained on simulated data to be effectively applied to beam data.

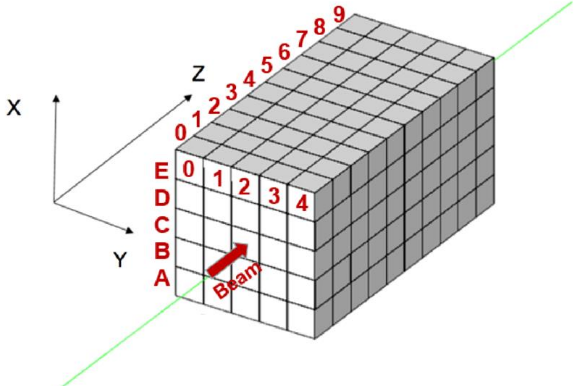




Data augmentation

- Beam test：fixed energy electron vertically incident the 5*5*10 crystal array



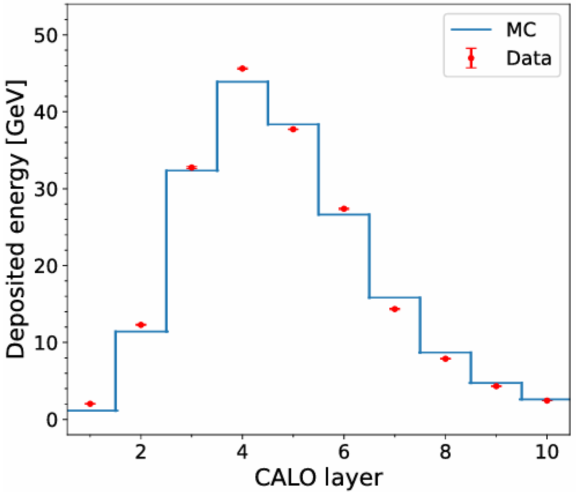


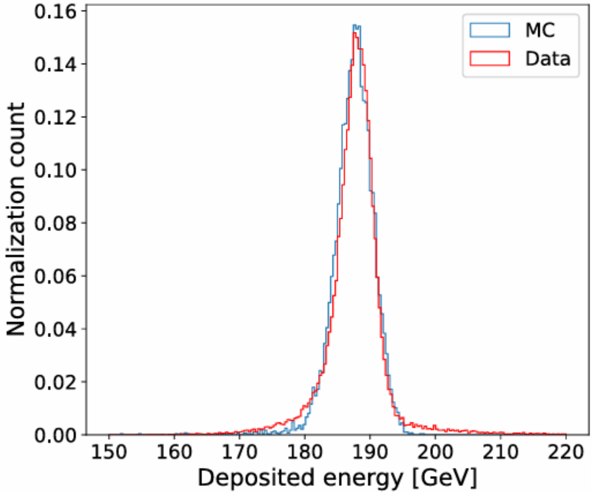
- Data augmentation：By increasing the quantity and diversity of training data, the generalization ability and robustness of the model are improved.

$$\widehat{E}_{\text{cell}} = F\left(\mu = E_{\text{cell}}, \sigma = \alpha E_{\text{cell}}\right) - \beta E_{\text{cell}},$$

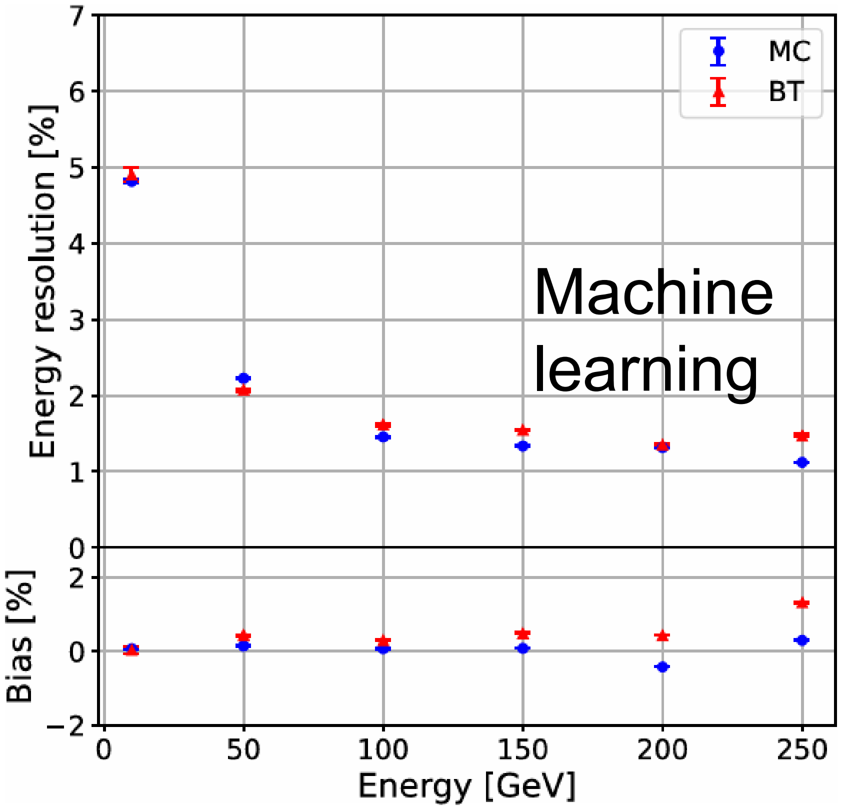
Gauss noise

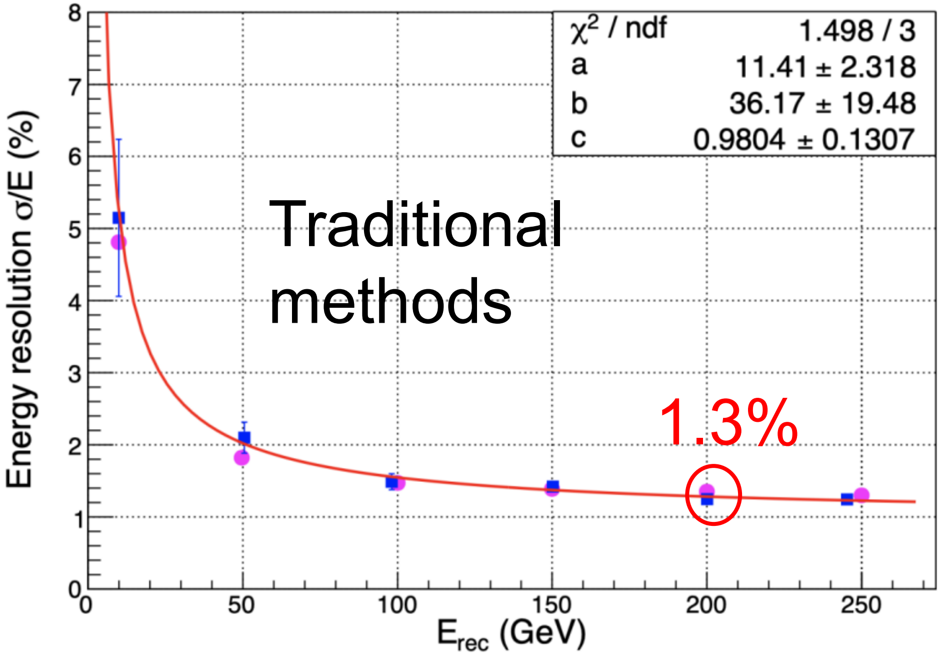
- Comparison of augmentation MC data and beam test data:





Beam data validation





- ML performs similarly on both simulated and beam test data.
- The energy resolution of the machine learning method is comparable to that of traditional methods.

Summary

- In MC simulated data, the optimized machine learning model achieves significantly high energy resolution <1%@10-1000GeV, and greatly improving reconstruction efficiency.
- In beam test data, the use of transfer learning enables models trained on MC data to be effectively applied to beam test data. The energy resolution achieved by the ML is comparable to that of traditional methods, which validates the reliability of the ML method.

Thank you