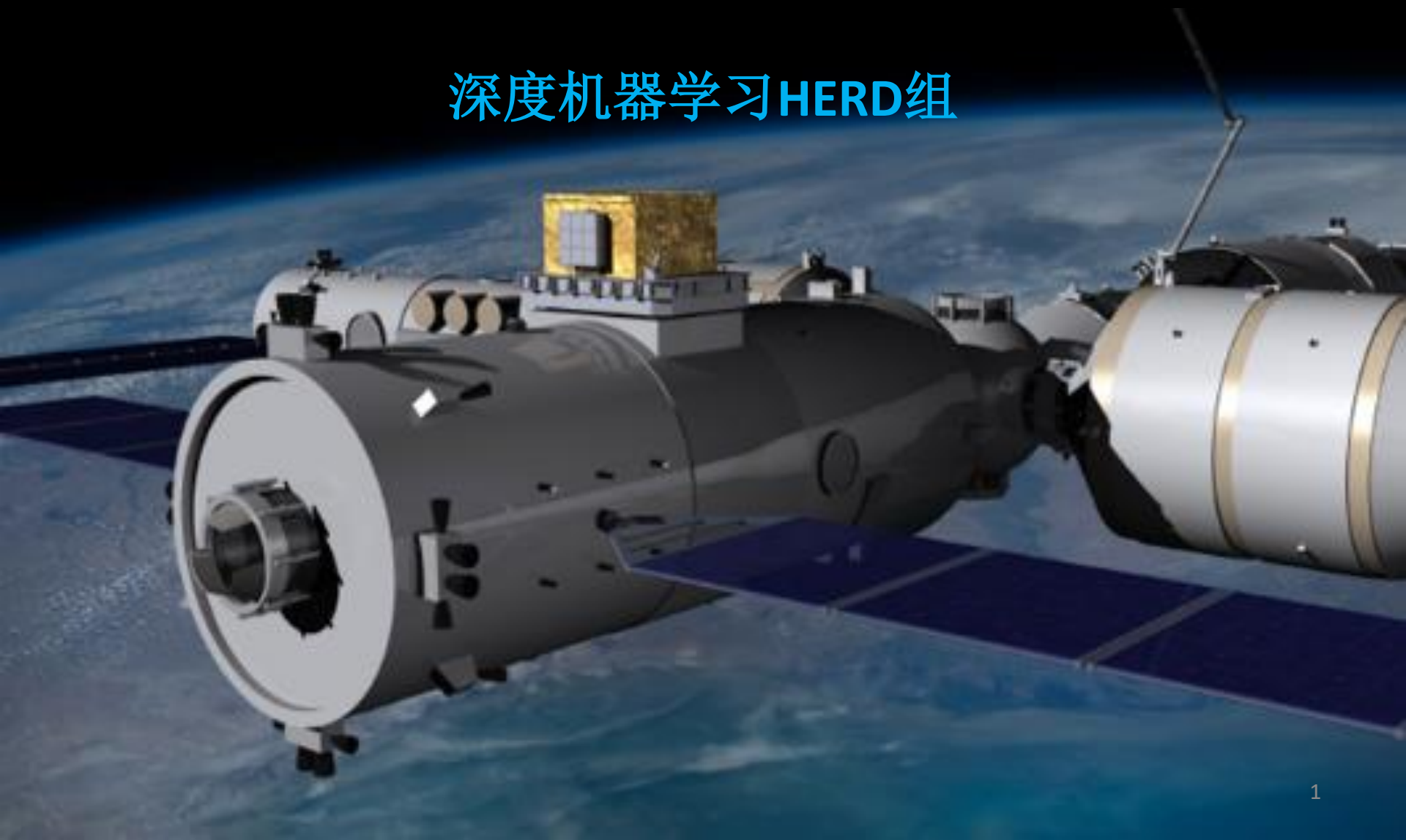


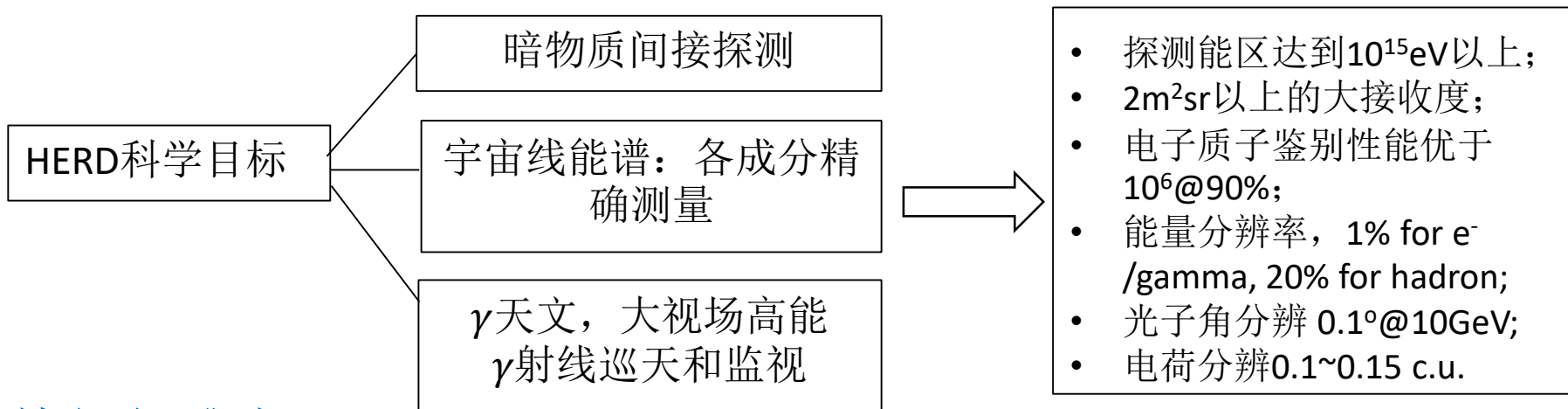
HERD中的机器学习工作

深度机器学习HERD组



HERD中的机器学习工作

高能宇宙辐射探测设施（HERD）是由中国主导的大型国际合作空间天文和粒子天体物理实验，计划安装在中国空间站，在轨运行寿命**10年**。



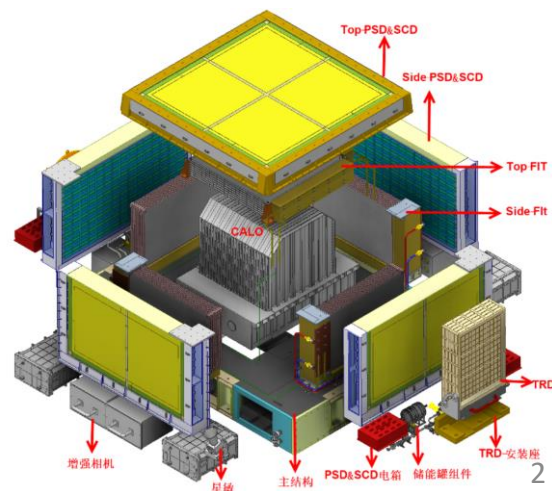
特征与难点：

➤探测器五面灵敏、量能器三维成像

[传统重建方法实现各方向一致的高精度重建难度大]

➤高能宇宙线探测事例少、干净样本获取难度大

[深度机器学习需要考虑MC至实际数据的模型迁移]



HERD中的机器学习工作

开展的深度机器学习工作：

- 量能器的电子径迹重建 **[已完成]**
- 量能器的电子能量重建 **[已完成]**
- 量能器的电子质子鉴别
- 量能器的质子与原子核能量重建
- 量能器的质子径迹重建
- 外围硅探测器的径迹重建

深度机器学习模型和工具：

DNN, CNN, GNN, ResNet, ParticleNet, Transformer,.....

深度机器学习的模型迁移：

Fine-tuning

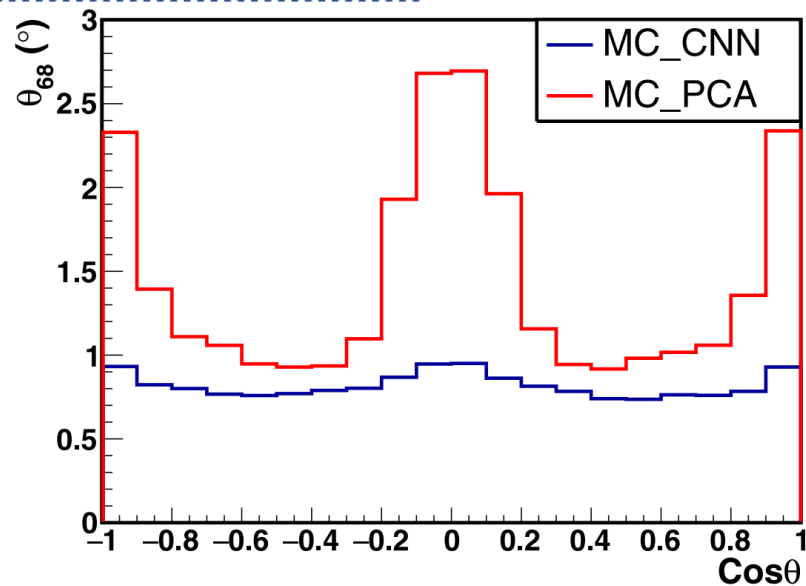
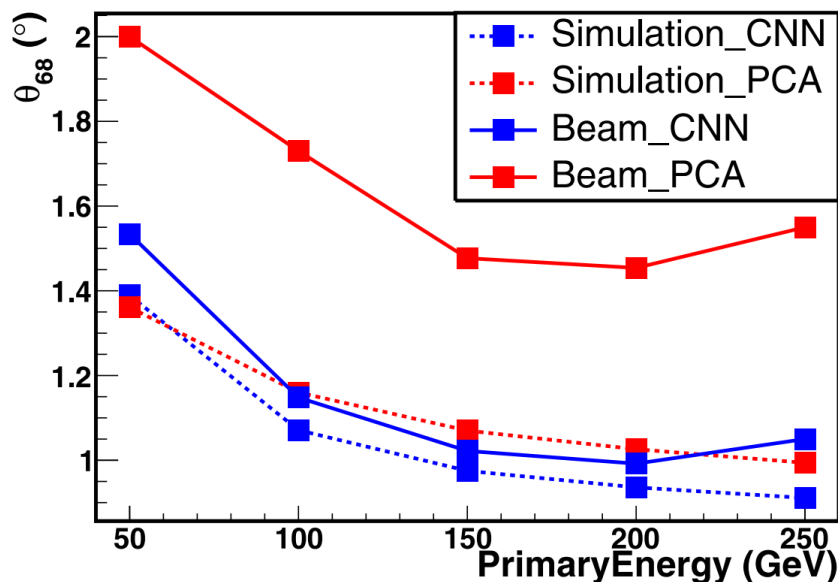
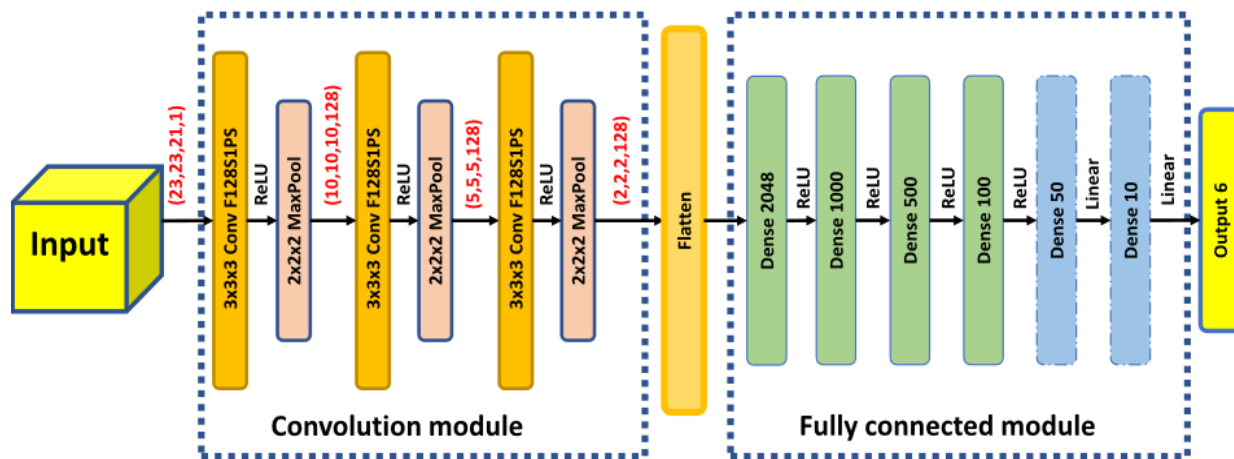
DANN

数据增强

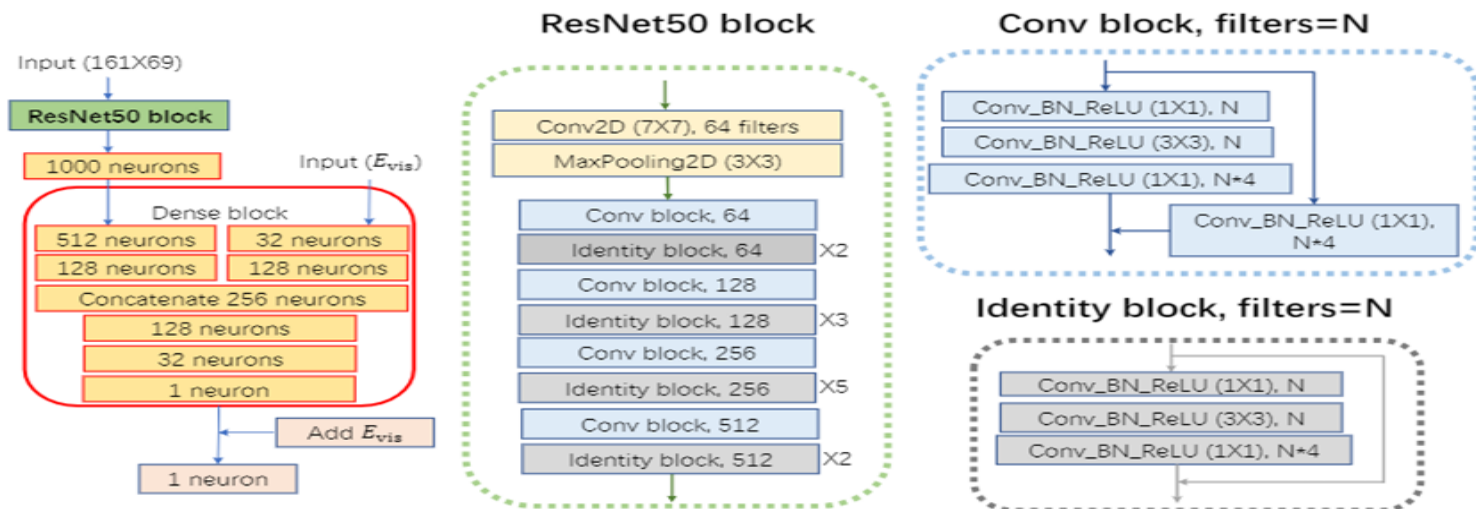
模型优化（加入物理规律的深度机器学习模型）

量能器的电子径迹重建

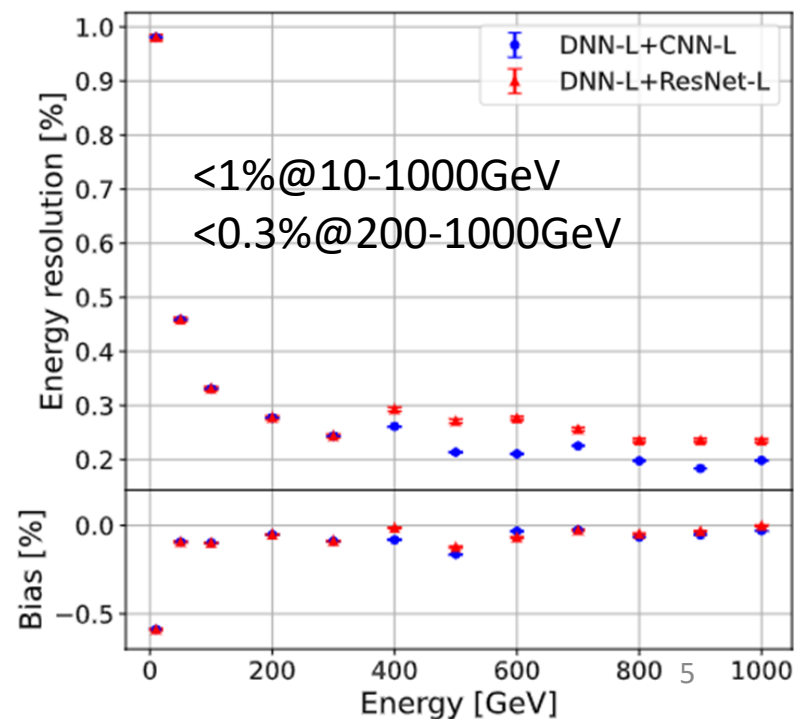
电子径迹重建使用**CNN**深度机器学习模型，算法得到的角分辨比传统重建方法（主成分分析法）提升约**40%@100GeV**，并且有着更小的随角度变化情况。目前该结果已发表文章**NIMA 1066 (2024) 169571**



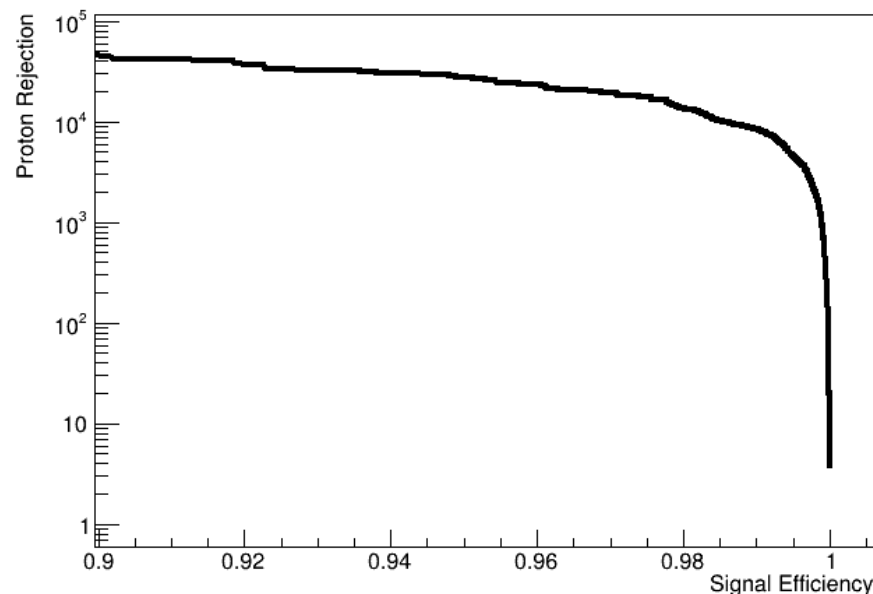
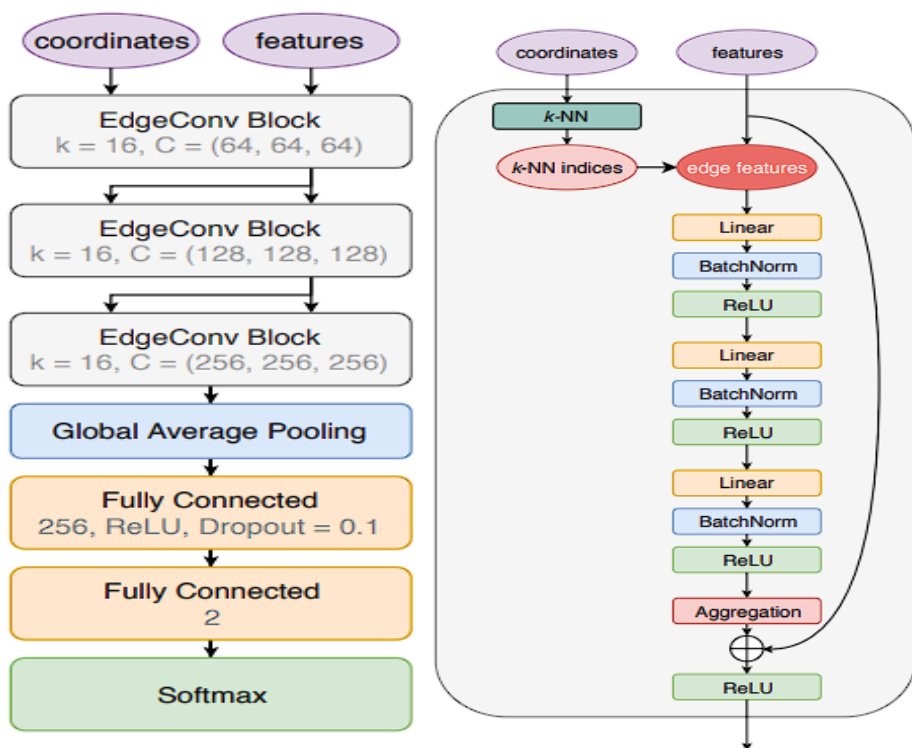
量能器的电子能量重建



- 电子能量重建使用多分类机器学习模型能量分辨 $<1\%$ @10-1000GeV, $<0.3\%$ @200-1000GeV, 重建效果符合预期
- 目前该结果文章《Application of machine learning method for energy reconstruction on space based high granularity calorimeter》已在 Experimental Astronomy 杂志接收



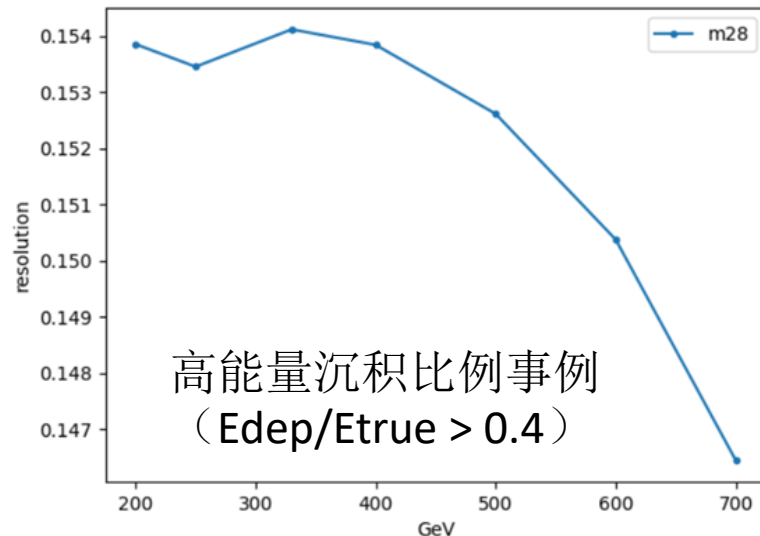
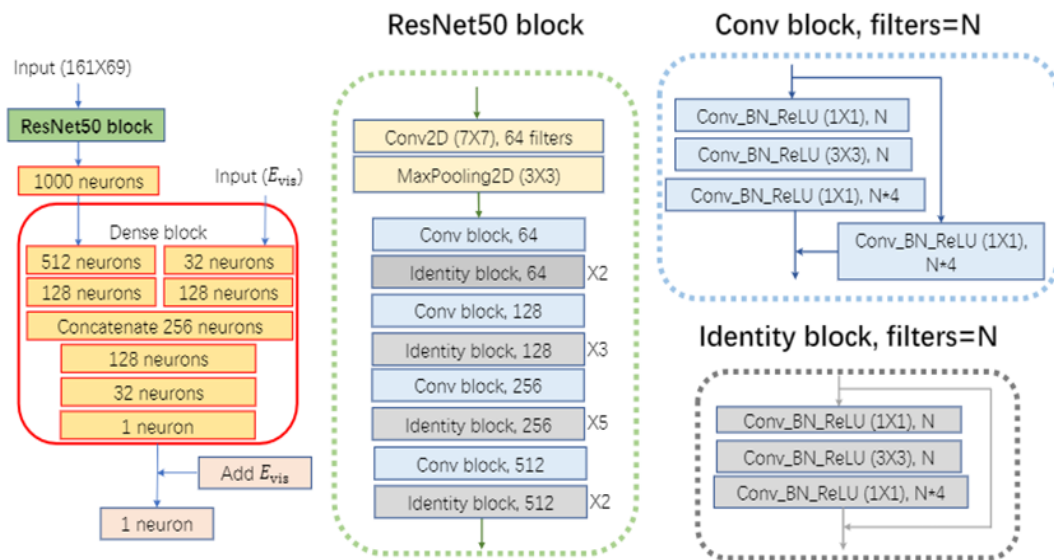
量能器的电子质子鉴别



- 目前使用ParticleNet模型基于MC数据初步训练得到电子质子鉴别结果，对粒子径迹长度大于20cm的事例，在信号效率90%情况下，质子鉴别能力达到 4×10^4 。
- 正在尝试模型优化及ParticleTransformer等模型进行电子质子鉴别的训练。

量能器的质子与原子核能量重建

目前正在进行质子的能量重建深度学习模型选取和训练工作，今后会推广到更高电荷的原子核能量重建。



难点

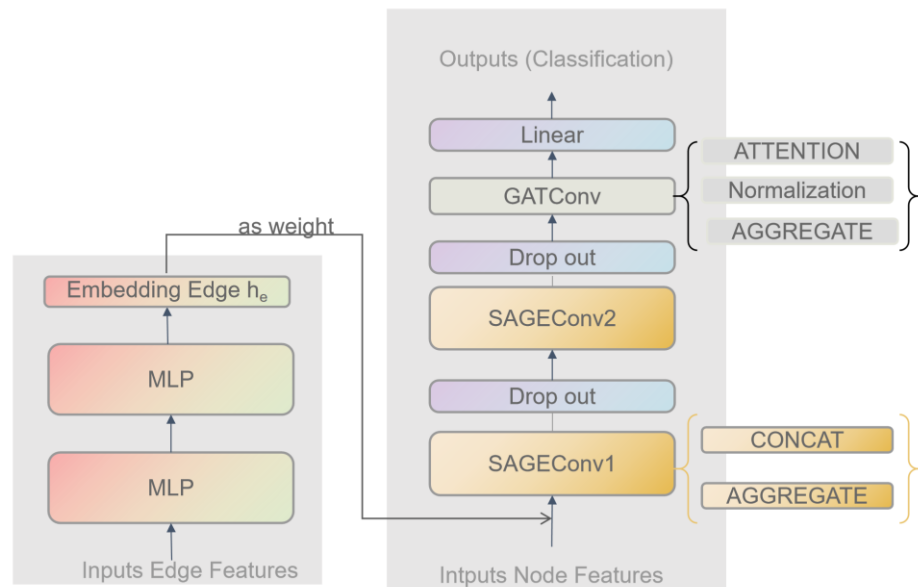
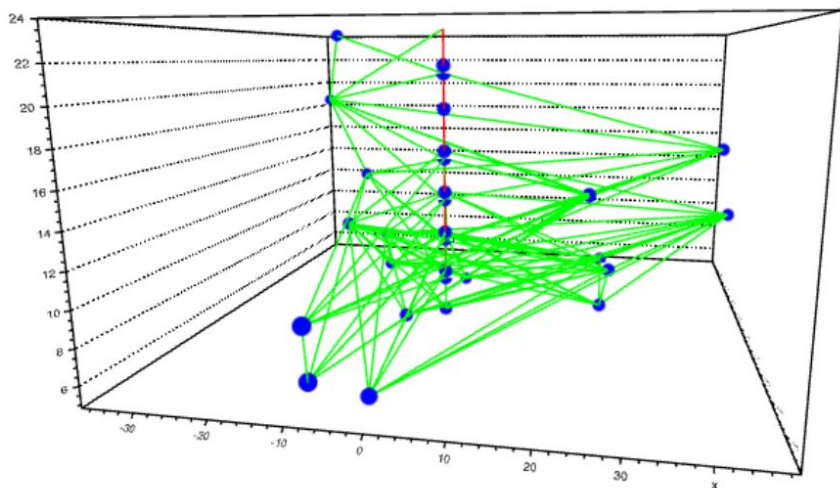
- 逐事例的波动(多样的核过程)
 - 簇射次级粒子种类和数目的多样性
 - 簇射次级粒子能量沉积的空间分布的多样性
 - 不可见能量的波动
- 能量泄露和损失
 - 死物质和大角度斜入射下的簇射形貌损失

下一步

- 低能量沉积比例事例的重建
- 能量重建模型的优化选取

外围硅探测器的径迹重建

目前采用GNN方法进行硅探测器的径迹重建，确定所有Hits点后使用GNN边分类模型进行Track Finding



muon	10GeV	100GeV	500GeV
$\theta_{68}(\text{deg})$ (MC)	0.023	0.0021	0.00041
$\theta_{68}(\text{deg})$ (prediction)	0.021	0.0022	0.0004

量能器的质子径迹重建

- 目前在CALO各向同性电子径迹重建任务上，CNN展现出了非常好的性能。而对于各向同性质子的径迹重建，我们希望在CNN基础上引入Transformer，利用Transformer的注意力机制对数据中的长程依赖性进行建模，进而增强模型的特征提取能力，进一步提升重建分辨率。
- 近年来许多视觉方法在模型中采用了CNN与Transformer融合的设计，相关算法有DeiT、ConViT、ResT、CvT等。

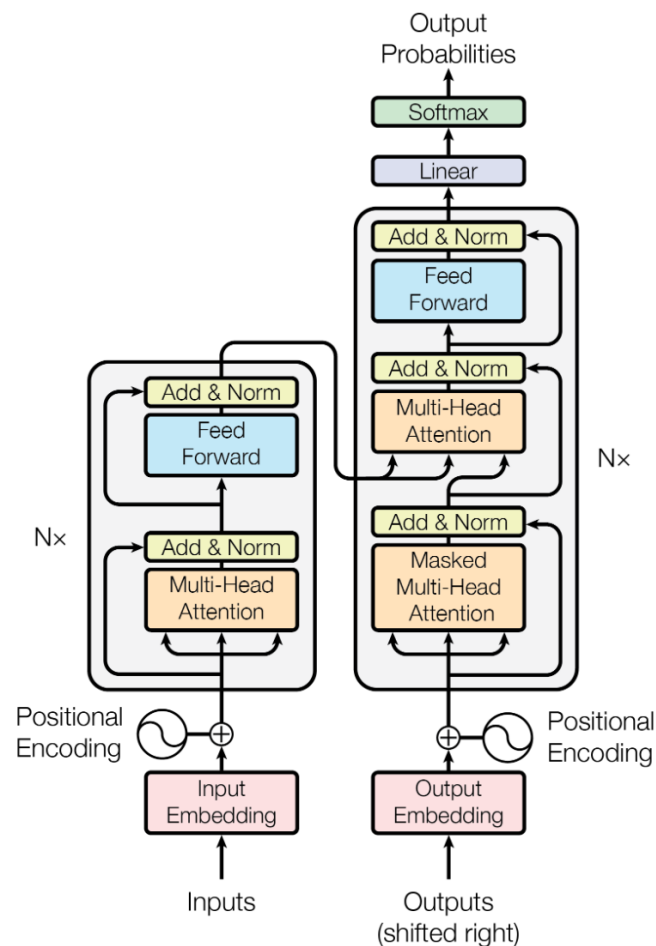


Figure 1: The Transformer - model architecture.

下一步工作计划

继续推进目前正在开展的深度机器学习工作：

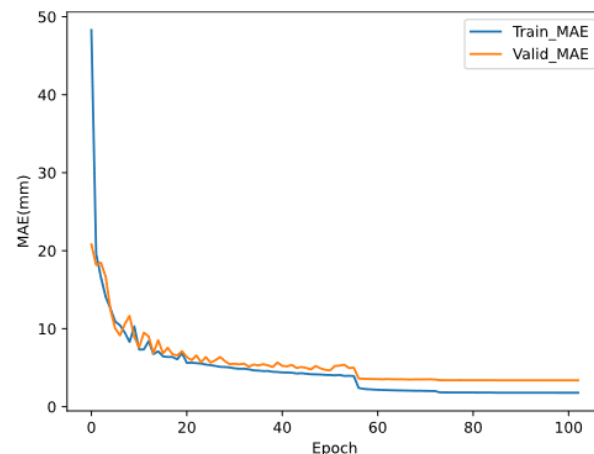
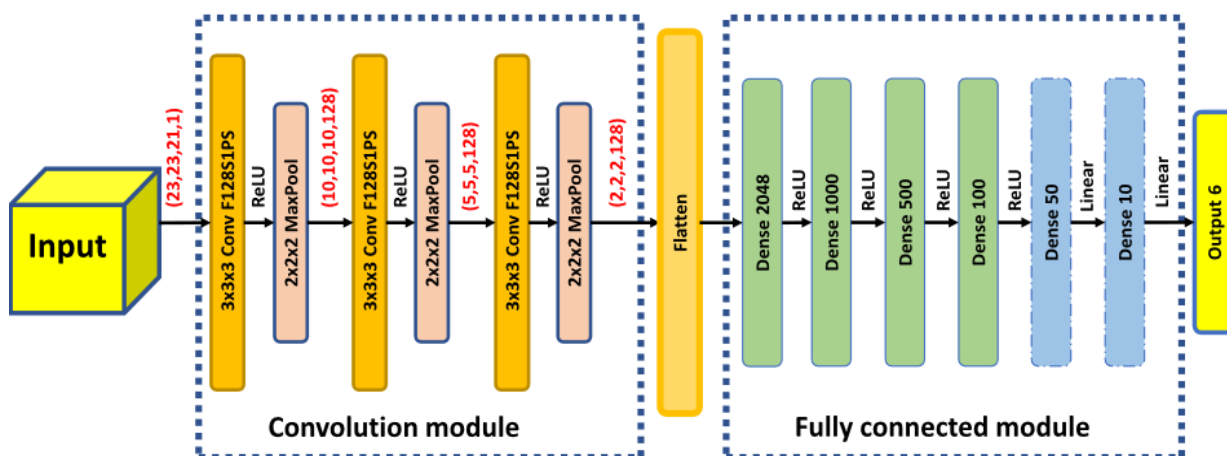
- 优化模型的电子质子鉴别能力及尝试ParticleTransformer等高级模型
- 选取及优化质子能量重建的深度机器学习模型，并推广至更高电荷的宇宙线原子核能量重建
- 基于CNN与Transformer融合的深度机器学习模型，得到分辨率进一步提升的CALO质子径迹重建结果
- 基于GNN及其他深度机器学习模型完善并优化外层硅探测器的径迹重建

优化提升MC至实际探测数据的模型迁移效果（特别是能量重建、粒子鉴别等等）

计划开展更多方面的深度机器学习工作：

- 外层子探测器（SCD）电荷重建
- 簇射团寻找
- 全局智能触发
- 探测器刻度
- Dr. Sai的应用
- ...

电子径迹重建(CNN)



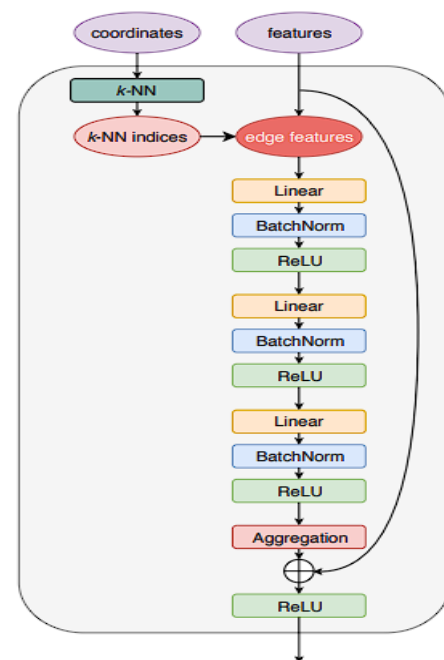
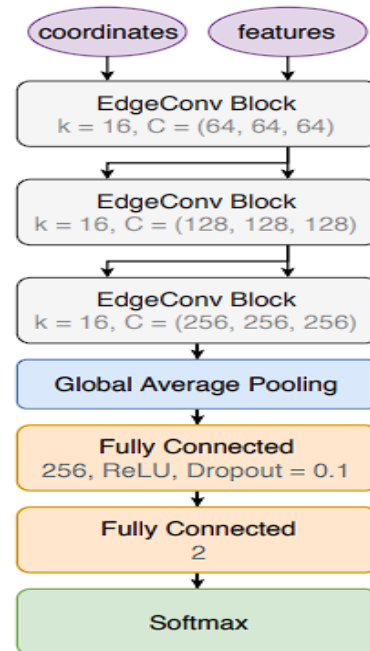
$$MSE(\vec{L}, \vec{L}_{pre}) = \frac{1}{6N} \sum_{i=1}^N (\vec{L} - \vec{L}_{pre})^2$$

$$MAE(\vec{L}, \vec{L}_{pre}) = \frac{1}{6N} \sum_{i=1}^N |\vec{L} - \vec{L}_{pre}|$$

- 输入为23*23*21*1的矩阵，1为通道数，代表沉积能量
- 输出为入射点和出射点组成的6个坐标元素的张量
- 网络包含卷积模块和全连接模块：卷积模块包含卷积层和池化层，卷积层采用3*3*3的卷积核，并设置padding=SAME，确保经过卷积后尺寸与输入尺寸相同，仅通过池化层减小尺寸；卷积模块展开后与全连接层相连，除最后两层激活函数是线性函数外，每一层的激活函数都是ReLU。
- 优化器为Adam，初始学习率为0.001，设置动态学习率，显卡为RTX A5000 GPU，BATCH SIZE为128
- 损失函数为MSE，评价标准为MAE

粒子鉴别(ParticleNet)

- 输入为(F,N)的二维矩阵, F代表晶体特征, N为采样的晶体数
- 输出单元为1(sigmoid)或2(softmax)
- 网络包含两个模块, 分别为边缘卷积模块和全连接模块
- 进入卷积模块前首先在位置空间使用KNN进行近邻点搜寻, 得到矩阵(F,N,K), 综合近邻点与每个节点的的特征之差, 得到矩阵(2F,N,K), 对矩阵进行转换得到矩阵(N,K,2F), 这个矩阵的每个维度对应着二维图片的维度(W,H,C), 采用1*1大小的卷积核对元素进行卷积; 卷积模块可以叠加, 根据卷积后的特征再次搜寻近邻点实现动态图的构建
- 边缘卷积模块得到的输出压缩展平后得到提取的特征, 经过全连接层后输出



$$x'_i = \bigoplus_{j=1}^k h_{\Theta}(x_i, x_{i_j}), \quad (1)$$

$$h_{\Theta}(x_i, x_{i_j}) = \bar{h}_{\Theta}(x_i, x_{i_j} - x_i), \quad (2)$$

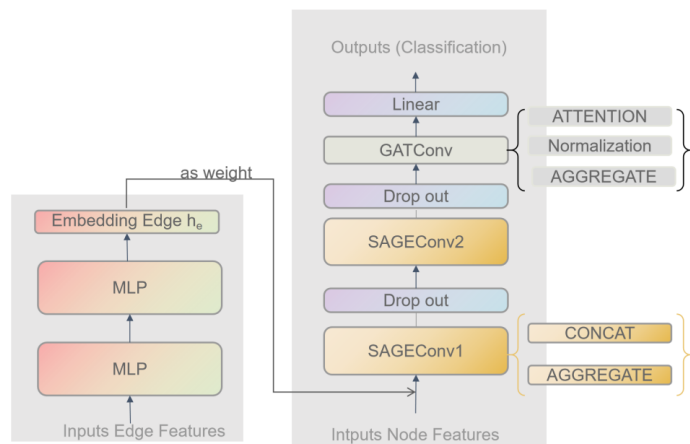
[Huilin Qu. Jet Tagging via Particle Clouds](#)

02.进度汇报

03 边分类模型

在 Hits 图构建完成后 使用GNN边分类模型进行Track Finding

- 其中联合两个node之间的距离能量差异作为 edge features 对 node 进一步embedding
- 卷积部分分别使用SAGEConv对消息进行传递
- 使用GATConv, 利用注意力机制将不同edge赋予不同权重并传递信息
- 组合node最后经过MLP对边分类（二分类）
- 损失函数设置为Focal Loss函数以提升高背景下的Track Finding精度
- 分类出原初径迹的边和非原初径迹的边
- Track Fitting过程采取直线拟合办法



02.进度汇报

04

参数设置

- 将直接读出向量与MC真值拟合后的向量 $\arccos$ 算出方位角,
- 将重建向量与MC真值拟合后的向量 $\arccos$ 算出方位角,

作为角分辨 (angle resolution)

测试验证：输出变量	定义	含义
θ_{pred}	$\theta_{\text{pred}} = \arccos(\mathbf{v}_{\text{pred}} \cdot \mathbf{v}_{\text{gn}})$	角分辨，代表预测结果
θ_{MC}	$\theta_{\text{MC}} = \arccos(\mathbf{v}_{\text{MC}} \cdot \mathbf{v}_{\text{gn}})$	蒙卡角分辨，代表蒙卡数据与gn差距

- 模型预测有效事例 / 排除次级粒子触发的事例 (Event Reconstruction Efficiency)
- 模型预测原初径迹边总数 / MC原初径迹边总数 (Edge Reconstruction Efficiency)

η_{graph}	$\frac{a_2 - a_3}{A - a_1}$	Event级别重建效率	
η_{edge}	$\frac{a}{t_2}$	Edge级别重建效率等价precision	模型识别为正类的样本中，真正为正类的样本所占的比例

- 模型预测原初径迹（正样本）的精度 (Purity)
- 模型预测全部样本的精度 (Accuracy)