

当粒子物理遇见深度学习：BESIII实验如何“看见”中子

宋昀轩 李彦谷 吕晓睿

(中国科学院大学 100049; 北京大学 100871)

2025年1月15日,北京谱仪Ⅲ(BESⅢ)实验合作组在《自然·通讯》杂志发表题为“Observation of a rare beta decay of the charmed baryon with a Graph Neural Network”的文章,宣布在世界范围内首次观测到粲重子 Λ_c^+ 的卡比玻压低半轻衰变 $\Lambda_c^+ \rightarrow ne^+\nu_e$,信号显著度超过10倍标准偏差,并为弱相互作用基本常数 $|V_{cd}|$ 提供了新的独立约束。在该研究中,团队创新性地提出并实践了一种基于深度学习的中子重建方法,对高精度测量结果的获得起到了核心作用。该方法克服了高能物理实验探测器在中子直接探测方面的性能限制,有望系统性地增强实验上对含长寿命中性强子末态的各类衰变的测量能力。

1. 中性强子的重建挑战

粒子物理学致力于研究物质世界的最小组成单元——基本粒子及决定其行为的基本相互作用。然而,大部分由基本粒子构成的亚原子粒子并不天然地存在于人们身边。高能粒子对撞机于是成为了当代粒子物理学最重要的实验工具之一,通过让两束粒子在极高的能量下迎头相撞,可以产生海量且种类丰富的亚原子粒子,并测量它们的物理性质。对撞过程中产生的中间态粒子普遍寿命极短,在形成后迅速衰变,经过级联的衰变过程,最终转化为可以稳定飞行一段距离的末态粒子,并被围绕着对撞点的各类探测器所记录。根据探测器输出的电子学信号,研究人员可以确定末态粒子的种类、能动量、顶点等信息,并进而还原出包括中间态粒子在内的整个对撞事例。这一方法被称为粒子与事例的重建,是对撞机数据分析的核心环节之一。

在 高能粒子对撞机上,带有电荷的电子、 μ 子、

质子、 π^+ 介子与 K^+ 介子可以通过电离效应在专用的径迹探测器进行重建,而中性光子则可以通过电磁量能器上的能量沉积进行重建。然而,中子和 K_L^0 介子等中性长寿命强子的重建则相对困难,这是由其与探测器材料相互作用的独特机制决定的。如图1所示,当高能光子进入量能器时,主要发生级联的电子对生成与韧致辐射过程,最终形成一个电磁簇射;而当中性强子进入量能器时,主要与原子核发生散射、裂变、吸收与湮灭等过程,最终形成一个强子簇射。强子簇射相对电磁簇射的空间发展范围更广,沉积能量的统计涨落更大,且有部分能量以中微子等形式逃逸,因此能量测量精度也更差。进一步地,若要在BESIII等没有专门强子量能器的对撞机实验中,借用电磁量能器重建中性强子则具有更大的挑战。由于材料与尺寸上的不匹配,电磁量能器上的强子簇射信号相当弥散,且具有可观比例的能量泄露。

中性长寿命强子测量与重建能力的不足,对于BESIII实验在粲强子衰变、轻强子谱、超子衰变、奇特强子态等诸多方向上研究含中性强子末态的衰变过程构成了挑战。例如粲重子 Λ_c^+ 有 $(32.6 \pm 1.6)\%$ 的概率衰变至含中子的末态,其中约20%的衰变过程尚未被测量过。目前,大部分相关实验研究主要依靠对撞过程能动量守恒的反冲质量方法,绕过了使用量能器信息的直接重建。然而,这仅适用于事例中其他粒子均能精确重建的部分情形,而对含有中微子的 $\Lambda_c^+ \rightarrow ne^+\nu_e$ 、 $D^+ \rightarrow K_L^0 e^+\nu_e$ 等半轻衰变,以及含有多个中子的 $e^+e^- \rightarrow n\bar{n}$ 等过程则不适用。虽然相关研究也各自发展了更为复杂的分析策略,并达成了各自的物理目标,但仍缺乏普适的可移植到更

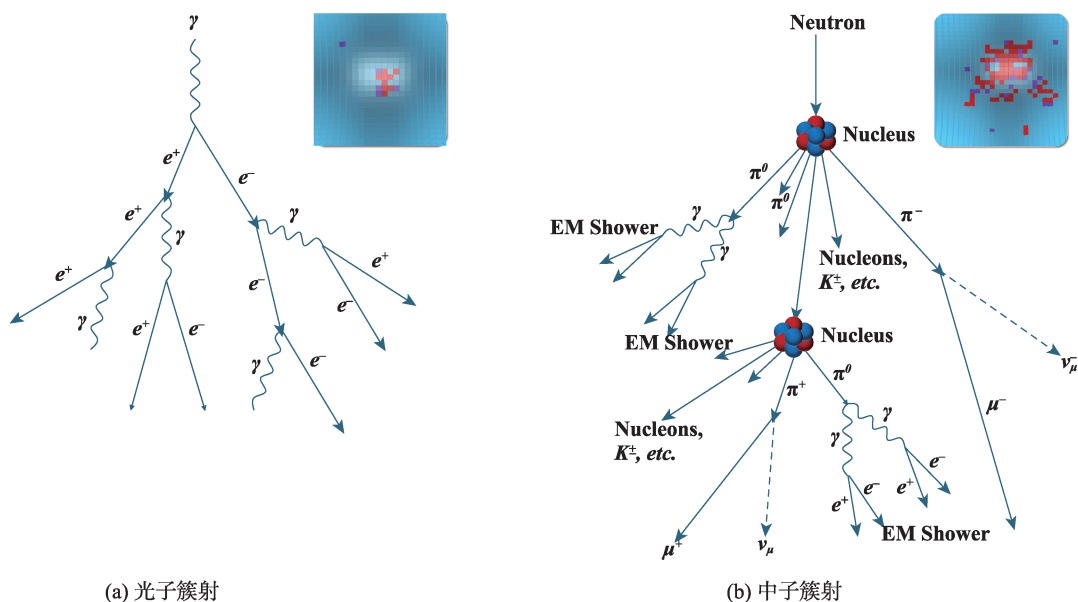


图1 光子与中子在材料中形成簇射的主要机制,及其在 BESIII 电磁量能器中着火晶体的分布示意图

广泛末态情况的中性强子重建方案。

2. 深度学习开启中性强子鉴别新途径

尽管中性强子可以在量能器中留下一些信息,但对撞机实验传统的数据处理方法难以充分挖掘与利用这些信息,导致了重建的困难。以粒子探测中最基础的粒子鉴别任务为例,人们知道强子簇射与电磁簇射的空间发展范围不同,其会影响量能器上相邻着火晶体连接而成的簇团形状。通过构造量化这一形状差异的特征变量并进行阈值判选,或者将一系列特征变量输入以提升决策树(Boosted Decision Trees, BDT)为代表的多变量分析工具中,人们确实可以在一定程度上鉴别对撞事例中的中性强子与光子。但这些由探测器底层输入人为地聚合、抽象而来的高层次变量带来了信息的涂抹与丢失,使得此类方法的鉴别性能较为有限,难以处理包含各类复杂本底成分与探测器噪声的真实场景,对于要从海量本底事例中挑选出稀有信号事例的任务往往力不从心。

过去十余年来,深度学习技术的飞跃式发展深刻改变了包括基础科学、工业界与商业界在内的人类社

会各个领域。通过构建包含多层神经元的深度神经网络(Deep Neural Network, DNN),可以在不同抽象层次上自动学习数据的潜在模式,在许多任务中表现出超越传统方法的能力,也为解决中性强子的重建困难提供了新的可能性。从深度学习的视角出发,如果将量能器的底层信息——着火晶体的位置与能量分布——直接输入深度神经网络,并训练其区分包含中性强子的信号与各类本底事例样本,就有可能实现对中性强子的高效率和低误判率鉴别。

值得强调的是,引入深度学习并不意味着研究者们放弃了物理上的洞察,反而更加需要将人的经验知识与计算技术相结合。粒子物理学界高度重视实验结果的严谨性与可靠性,对实验流程的充分验证与各项误差的准确评估往往与结果本身同等乃至更加重要。深度神经网络在可解释性方面的欠缺、对模拟与真实数据间差异的敏感性以及输入信息中不同误差来源的耦合等问题,都需要研究者们发挥才智,找到相应的解决方法。

3. 深度学习在 BESIII 实验研究中的首次实践

粲强子弱衰变的精确测量是 BESIII 实验的主要

物理目标之一,为探索标准模型中强相互作用与弱相互作用的耦合机制提供了理想的平台,相关实验研究可以推进量子色动力学(Quantum Chromodynamics, QCD)在微扰与非微扰交界能区有效理论的发展,并精确测量电弱理论中卡比玻-小林-益川(CKM)幺正矩阵等基本参数。基态粲重子 Λ_c^+ 的半轻衰变模式包括粲夸克到奇异夸克与粲夸克到下夸克的转变,分别对应卡比玻允许的 $\Lambda_c^+ \rightarrow \Lambda l^+ \nu_l$ 和卡比玻压低的 $\Lambda_c^+ \rightarrow n l^+ \nu_l$ 衰变过程。BESIII实验在2015年报道了 $\Lambda_c^+ \rightarrow \Lambda e^+ \nu_e$ 的绝对分支比,并在2022年精确测量了其微分衰变率与形状因子,揭示了实验结果与格点QCD计算间的一定差异。然而, $\Lambda_c^+ \rightarrow n e^+ \nu_e$ 的衰变模式未曾在实验上被观测到。人们预期其分支比相对 $\Lambda_c^+ \rightarrow \Lambda e^+ \nu_e$ 小一个数量级,各种有效理论唯象模型和格点QCD理论对该衰变的分支比给出了不同的理论计算,通过实验检验与修正这些理论预言是BESIII的紧要任务之一。

从实验角度来看,寻找 $\Lambda_c^+ \rightarrow n e^+ \nu_e$ 存在着明显的技术挑战。由于中子与中微子都存在于衰变末态中,这使得我们无法通过传统的反冲质量的方法甄别中微子信号,必须使用电磁量能器尝试直接探测中子。更直接的问题是来自 $\Lambda_c^+ \rightarrow \Lambda e^+ \nu_e$, $\Lambda \rightarrow n \pi^0$ 过程的本底,它与信号衰变末态的区别仅为 π^0 衰变产物的两个额外光子(如图2所示),而统计量高达信号过程的十倍左右。因此,我们需要高效地鉴别中子与 Λ 超子中性衰变末态在电磁量能器上的能量

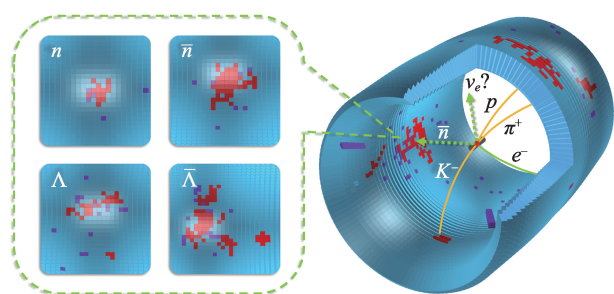


图2 (右) $\Lambda_c^+ \rightarrow p K^- \pi^+$, $\bar{\Lambda}_c^- \rightarrow \bar{n} e^- \bar{\nu}_e$ 事例在BESIII探测器中的可视化显示。蓝色的圆柱体表示桶部电磁量能器的结构,红色和紫色的像素代表着火的量能器晶体;
(左)虚线框中的区域分别例举了中子、反中子、 Λ 超子和 $\bar{\Lambda}$ 超子在电磁量能器上的典型击中模式

沉积分布,并构造一个观测变量以提取 $\Lambda_c^+ \rightarrow n e^+ \nu_e$ 信号。进一步地,考虑到当前基于GEANT4的中子与量能器相互作用的模拟与实际物理过程仍存在一定偏差,基于蒙特卡罗(Monte Carlo, MC)模拟方法估计的探测效率的需要校准这一偏差才可能得到可靠的结果。上述任务超出了BESIII实验传统数据分析技术的能力,导致 $\Lambda_c^+ \rightarrow n e^+ \nu_e$ 的测量迟迟未能实现突破。

为了解决以上问题,本项目团队研究创新性地使用了图神经网络(Graph Neural Network, GNN)来区分含中子与 Λ 超子的物理事例。在GNN训练中,电磁量能器上与这些中性强子相关联的能量沉积簇团被表示为点云的数据结构,并进一步抽象为图的节点与边。由于点云和图具有转动不变性,意味着得到的物理结果不会随坐标系改变。针对模拟与真实数据的偏差,我们提出了一套不依赖于MC模拟的数据驱动的分析策略。我们利用BESIII实验采集的、世界范围内最大的100亿 J/ψ 衰变数据,可以使用反冲策略筛选出高统计量、高纯度的中性强子控制样本,包括来自 $J/\psi \rightarrow \bar{p} \pi^+ n + c.c.$ 等过程的中子和反中子样本,以及来自 $J/\psi \rightarrow \bar{p} K^+ \Lambda + c.c.$ 等过程的 Λ 和 $\bar{\Lambda}$ 样本。这些控制样本中包含了中性强子与电磁量能器的相互作用信息,能够如实反映实际物理的中子与探测器相互作用模式。利用这些控制样本,团队实现了GNN模型的训练与校准、物理结果的检验与相关系统误差的估计。

图3展示了GNN模型对 Λ_c^+ 数据中 n/Λ 分类评分的分布,可以看到 $\Lambda_c^+ \rightarrow n e^+ \nu_e$ 信号与 $\Lambda_c^+ \rightarrow \Lambda e^+ \nu_e$ 本底得到了很好的区分,且数据中的分布与修正后的模拟样本分布符合得很好。通过拟合这一分布提取信号,研究团队成功以超过10倍标准偏差的信号显著度直接观测到了 $\Lambda_c^+ \rightarrow n e^+ \nu_e$ 及其反粒子的衰变过程,并以大约10%的精度测量了其分支比为 $(3.57 \pm 0.37) \times 10^{-3}$ 。这一结果为十余种理论模型的预言提供了检验与校准,并结合格点QCD理论对衰变形状因子的计算结果,给出了CKM矩阵元 $|V_{cd}|$ 首个来自粲重子半轻衰变的独立实验约束,与目前

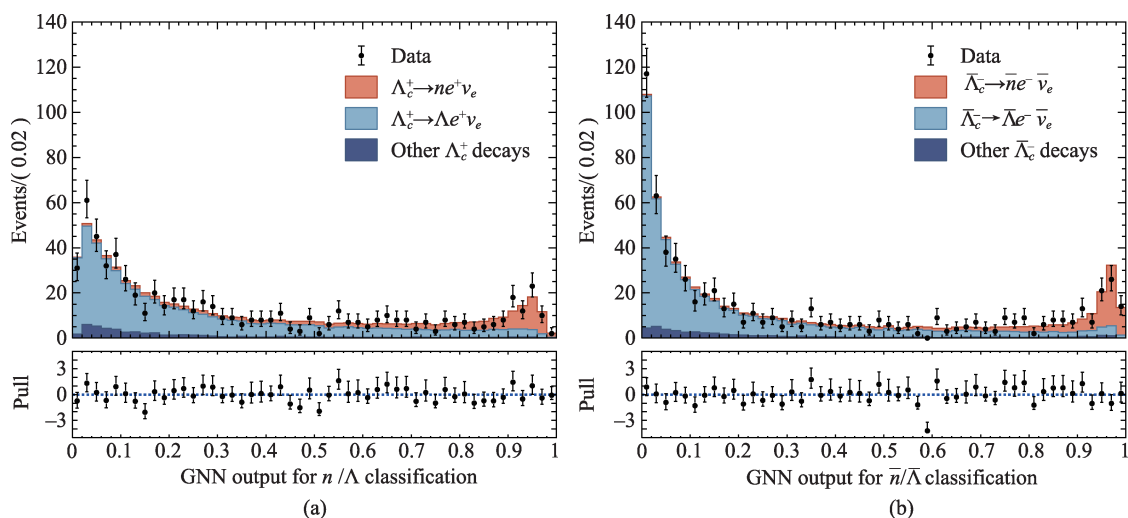


图3 GNN模型对 Λ_c 数据样本中分类评分的分布,其中黑色点代表数据,彩色的直方图代表蒙特卡罗模拟的各衰变成分

的世界平均值在误差范围内相符。

4. 对反中子完整重建的进一步探索

实现电磁量能器对含中子末态事例的鉴别只是研究的起点,对于完整的粒子重建任务,人们还需要知道它在量能器上确切的入射位置与动量大小,从而获得其运动学信息。这些信息将支持更精细的物理测量工作,包括测量重子半轻衰变的形状因子、弱相互作用的轻子普适性检验以及 CP 破坏寻找等。由于中性强子在电磁量能器中可观的能量泄露与次级散射效应,这一任务相较中子鉴别的难度更大。

BESIII实验在这方面已经取得了初步进展。来自中国科学院大学与北京大学的研究人员与计

算机视觉领域的专家合作,使用该领域的前沿技术联合开发了视觉模型架构Vision Calorimeter(ViC)。这一算法创新性地将反中子在电磁量能器上的重建任务转化为视觉目标检测问题,即在图像中识别与定位感兴趣的目标,并用边界框选定目标区域。具体而言,反中子在电磁量能器上的能量沉积分布首先被映射为二维图像,并围绕反中子入射位置的真值生成边界框的标注。训练后的深度神经网络使用边界框的中心作为入射位置的预测,并通过分支网络同时实现粒子鉴别与动量预测的下游任务,如图4所示。初步结果表明,ViC在入射位置测量上的精度相对BESIII的传统方法提升约80%,平均角度误差小于 10° ,并且首次在动量预测方面建立了预测值与真值的正向关联,平均相对误差低至20%。在BESIII实验的实际应用场景中,ViC重建出的反

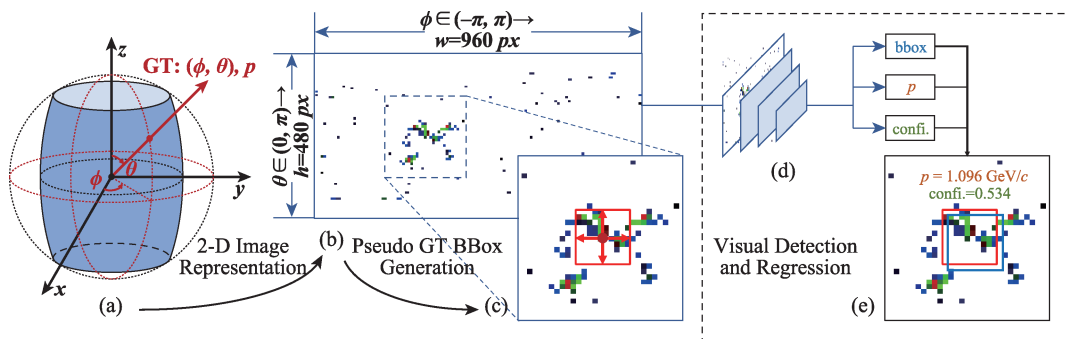


图4 ViC实现反中子重建的工作流程

中子四动量能够用于计算均值基本无偏且分辨良好的母粒子不变质量峰。目前研究团队正着手对其进行优化与全面的物理表现评估,有望在未来将其应用于 BESIII 实验的实际分析工作中。

5. 展望

BESIII 实验对中性强子重建的成功实践,是人工智能赋能基础科学研究的一个缩影。深度学习等 AI 技术正迅速融入包括粒子物理在内的诸多基础科学领域,极大地拓展了研究者分析问题的能力

和工具箱。在对撞机实验的装置设计与运行、数据采集与刻度、事例模拟与重建、物理分析与计算等各个环节中,由深度学习驱动的技术相比传统方法都具有潜在的优势。不过我们也需要看到,深度学习在对撞机实验上的大规模应用仍尚需时日,造成这一现象的因素是多方面的,包括人力与资源投入、实验运行周期、社群共识普及,以及学界对实验结果严谨性与可靠性的审慎要求等。因此,需要进一步促进粒子物理与计算机科学的深度交流与合作,培养跨学科的复合型人才,通过 AI 技术与物理领域知识的融合形成学科之间的双向赋能。

她用物理的情趣,引我们科苑揽胜;
她用知识的力量,助我们奋起攀登!

欢迎投稿,欢迎订阅

《现代物理知识》杂志隶属于中国物理学会,由中国科学院高能物理研究所主办,是我国物理学领域的中、高级科普性期刊。

科技创新、科学普及是实现创新发展的两翼。《现代物理知识》旨在普及科学知识、弘扬科学精神,设有物理知识、物理前沿、科技经纬、科学源流、教学参考、中学园地、科学书屋和科苑快讯等栏目。诚邀在物理学及相关领域工作的科技、教育和科普等方面的专家学者,以公众喜闻乐见的文字,深入浅出、图文并茂地与读者分享现代物理知识、科学前沿成果和大科学装置进展等精彩故事,共襄“两翼齐飞”之盛举。投稿时请将稿件的 Word 文档发送至本刊电子信箱 mp@mail.ihep.ac.cn,并将联系人姓名、详细地址、邮政编码,以及电话、电子信箱等联系方式附于文章末尾。

所投稿件一经本刊录用,作者须将该篇论文各种介质、媒体的版权转让给编辑部所有,并签署《现代物理知识》版权转让协议书(全部作者签名),如不接受此协议,请在投稿时予以声明。来稿一经发表,将一次性酌情付酬,以后不再支付其他报酬。

2025 年《现代物理知识》每期定价 15 元,全年 6 期 90 元,欢迎新老读者订阅。

需要往期杂志的读者,请按下列价格付款:

2010~2021 年单行本每期 10 元;2022~2024 年单行本每期 15 元;2010~2021 年合订本每本 60 元,2022~2023 年合订本每本 90 元。

订阅方式

(1) 邮局订阅 邮发代号:2-824。

(2) 编辑部订阅(请通过银行转账到以下账号,并在附言中说明“现代物理知识**年**期”)

名称:中国科学院高能物理研究所

开户行:工商银行北京永定路支行

账号:0200004909014451557

(3) 科学出版社期刊发行部:联系电话 010-64017032 64017539;

(4) 网上购买:搜淘宝店、微店店铺名称:中科期刊;淘宝、微信购买二维码:



淘宝网购刊



微信购刊