

深度学习在核物理中的前沿进展

张靖宗^{1,2,3} 郭 爽^{2,3} 朱励霖¹ 王凌霄^{3,4} 马国亮^{2,3}

1(四川大学 物理学院 成都 610064)

2(复旦大学 现代物理研究所 上海 200433)

3(理论物理专款上海核物理理论研究中心 上海 200438)

4(日本理化学研究所RIKEN 和光市 351-0198, 日本)

摘要 随着高能核物理研究进入多维度、高复杂度数据分析阶段,深度学习技术正逐步成为理解极端条件下核物质行为的关键工具,并推动研究范式从经验驱动向数据驱动的根本转变。本文简要梳理了机器学习在该领域的演进,并着重介绍了深度学习方法在其中的前沿进展:早期(20 世纪末至 21 世纪 10 年代)研究主要采用人工神经网络和支持向量机等传统算法,通过核质量预测、相变识别等任务验证了机器学习处理核物理问题的可行性,但受限于人工特征提取和计算能力的制约,尚未触及物理特征的自主挖掘;深度学习时代(21 世纪 10 年代至今),研究者创新性地引入点云网络架构,通过直接处理末态粒子四动量数据,不仅突破了传统方法依赖人工构造统计观测量的局限,更开启了从数据表象到物理实在认知跃迁的进程。与此同时,无监督学习方法推动研究重心从假设验证转向数据驱动的物理规律自主发现,不仅实现了异常信号的敏锐捕捉,更催生出物理现象涌现性研究的新思路。展望未来,从发展包含物理先验的深度学习算法以提升模型的物理含义,到元学习与自监督框架深化稀有事件分析;从量子机器学习加速提取高维数据特征,到生成式模型重构物理数据生态,这些发展或将推动高能核物理从观测数据的被动解释转向物理规律的主动发现,从局部特征的碎片化分析转向系统行为的整体性认知,最终有可能构建具有自主知识发现能力的智能物理研究体系。

关键词 机器学习, 深度学习, 重离子碰撞, 科学智能

中图分类号 TL99

DOI: 10.11889/j.0253-3219.2025.hjs.48.250130

CSTR: 32193.14.hjs.CN31-1342/TL.2025.48.250130

The development and application of deep learning in high-energy nuclear physics

ZHANG Jingzong^{1,2,3} GUO Shuang^{2,3} ZHU Lilin¹ WANG Lingxiao^{3,4} MA Guoliang^{2,3}

1(Department of Physics, Sichuan University, Chengdu 610064, China)

2(Key Laboratory of Nuclear Physics and Ion-beam Application (MOE), Institute of Modern Physics, Fudan University, Shanghai 200433, China)

3(Shanghai Research Center for Theoretical Nuclear Physics, NSFC and Fudan University, Shanghai 200438, China)

4(Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), RIKEN, Wako, Saitama 351-0198, Japan)

Abstract As high-energy nuclear physics research enters a phase characterized by multi-dimensional and highly complex data analysis, deep learning techniques are gradually becoming essential tools for understanding nuclear

国家自然科学基金(No.12147101, No.12325507)、国家重点研发计划(No.2022YFA1604900)、广东省基础与应用基础研究重大项目(No.2020B0301030008)资助

第一作者: 张靖宗, 男, 2003 年出生, 2025 年毕业于四川大学, 研究领域为机器学习与核物理交叉研究

通信作者: 王凌霄, E-mail: lingxiao.wang@riken.jp; 马国亮, E-mail: glma@fudan.edu.cn

收稿日期: 2025-03-24, 修回日期: 2025-05-01

Supported by National Natural Science Foundation of China (No.12147101, No.12325507), the National Key Research and Development Program of China (No.2022YFA1604900), the Guangdong Major Project of Basic and Applied Basic Research (No.2020B0301030008)

First author: ZHANG Jingzong, male, born in 2003, graduated from Sichuan University in 2025, focusing on machine learning and its applications in nuclear physics

Corresponding author: WANG Lingxiao, E-mail: lingxiao.wang@riken.jp; MA Guoliang, E-mail: glma@fudan.edu.cn

Received date: 2025-03-24, revised date: 2025-05-01

matter behavior under extreme conditions. This shift is driving a fundamental transformation in research paradigms from experience-driven approaches toward data-driven methodologies. This article briefly reviews the evolution of machine learning in this field, emphasizing recent advancements involving deep learning techniques. Early research (from the late 20th century to the 2010s) primarily employed traditional algorithms such as artificial neural networks and support vector machines. These studies validated the feasibility of machine learning approaches in nuclear physics through tasks like nuclear mass prediction and phase transition identification. However, due to limitations in manual feature extraction and computational capabilities, such methods did not yet extend to autonomous exploration of physical features. In the deep learning era (2010s to present), researchers have innovatively introduced point-cloud neural network architectures, enabling direct processing of final-state particle four-momentum data. This advancement has overcome the constraints of traditional methods that relied heavily on manually constructed statistical observables and initiated a conceptual leap from superficial data representations toward intrinsic physical insights. Simultaneously, unsupervised learning methods have shifted research focus from hypothesis validation to autonomous, data-driven discovery of physical laws, facilitating not only sensitive detection of anomalous signals but also opening new avenues for investigating emergent physical phenomena. Looking ahead, from developing deep learning algorithms incorporating physical priors to enhance the model physical interpretation, to meta-learning and self-supervised frameworks deepening rare event analysis; from quantum machine learning accelerating high-dimensional feature extraction, to generative models reconstructing the physical data ecosystem, these advancements will potentially propel high-energy nuclear physics research from the passive interpretation of observational data toward active discovery of physical laws, shifting analysis from fragmented, local feature exploration toward holistic comprehension of systemic behaviors. Ultimately, this progression may pave the way toward constructing an intelligent physics research system capable of autonomous knowledge discovery.

Key words Machine learning, Deep learning, Heavy-ion collisions, AI for science

相对论重离子碰撞研究旨在探讨核物质在极端温度和密度条件下的行为,重点关注夸克-胶子等离子体(Quark-Gluon Plasma, QGP)的形成与性质,并探索强相互作用的相图^[1-2]。高能粒子在碰撞过程中经历剧烈的演化,从初始的高温高密子系统逐渐冷却,并最终通过强子化形成末态粒子。这一过程中涉及的非平衡动力学机制十分复杂,既包括集体流、喷注淬火等强相互作用效应^[3-6],也涉及量子色动力学(Quantum Chromodynamics, QCD)在高温状态下的表现。深入研究 QGP 的特性不仅有助于理解极端条件下的核物质性质,也能为早期宇宙的物理过程提供重要的理论依据^[7-9]。

该领域的研究通常依赖于末态粒子的观测来反演碰撞过程,从中提取关键的物理信息。然而,传统的分析方法主要基于手工构造的统计观测量,例如横向动量分布、椭圆流(v_2)和高阶流(v_3 、 v_4)等。这些方法虽然成功地探测到 QGP 的集体性质、流体动力学响应、淬火效应等物理现象^[10-13],但在从高维度、多变量相关性强的数据中抓取敏感物理特征时往往显得力不从心。一方面,由于 QGP 的演化涉及强相互作用物质非平衡态的输运过程,模型参数空间的维度极高,导致传统数值方法(如格点 QCD、流体动力学模型)计算成本巨大。另一方面,实验数据

的复杂性也在不断提高,例如 ALICE、STAR 等实验所产生的碰撞数据包含数十亿的最低偏差事件(Minimum Bias Event, NMB),大部分事件都涉及数百个粒子的动量、角度、种类等信息,如何从海量数据中提取关键的物理模式成为一大挑战^[14]。

近年来,机器学习特别是深度学习技术取得了长足的发展,极大地推动了数据驱动科学的发展。一系列基于 Transformer^[15]结构的大语言模型展现出了卓越的泛化能力,而深度学习在计算机视觉、医学影像分析、自动驾驶等领域的成功应用也证明了其强大的数据处理能力。在此背景下,研究者开始探索如何将先进的深度学习技术引入高能核物理研究^[16-18],机器学习,特别是深度学习,正逐步成为应对该领域高维度、大数据和复杂关联挑战的工具,推动研究范式向数据驱动转变,多篇权威综述已经详细阐述机器学习方法在核物理理论、实验、加速器科学乃至基础 QCD 问题中的广泛应用^[19-22]。基于深度学习的点云分析方法(如 PointNet、Point Transformer)已在三维数据处理任务中展现了强大的特征提取能力^[23-26],而高能核物理实验所产生的末态粒子数据本质上也是一种无序点云数据。因此,借助点云网络,研究者能够直接从粒子级数据中学习碰撞系统的特征,从而绕开传统统计观测量的

局限性,为QGP研究提供全新的思路^[27]。

1 机器学习的发展

机器学习技术的发展经历了从简单到复杂的演变过程。早期的研究主要依赖人工设定的规则系统,需要研究人员手动提取数据特征并编写判断逻辑。随着数据量的快速增长和计算能力的提升,现代机器学习逐渐转向数据驱动模式,通过算法自动从数据中学习规律,而深度学习则是其中最为成功的分支。这种转变使得统计学、数据科学和深度学习技术成为解析复杂物理系统的重要工具,它们不再局限于验证已有理论,而是能够主动发现数据中隐藏的新规律^[15]。特别是在高能核物理领域,重离子碰撞实验每天产生的TB级多维数据(包括粒子动量、能量、角度等信息)^[28],传统分析方法已难以有效处理。深度学习技术的出现提供了突破口:通过构建多层神经网络,计算机可以自动识别数据中的复杂模式,例如从数百万个粒子轨迹中提取集体流特征,或是发现不同碰撞能量下的相变信号,这些能力显著扩展了人类探索极端核物质行为的技术手段。

根据学习方式的不同,机器学习主要分为三大类:监督学习(Supervised Learning)、无监督学习(Unsupervised Learning)和强化学习(Reinforcement Learning)^[15]。监督学习需要预先标注的数据作为“参考答案”,类似于学生在老师指导下学习。在高能核物理中,这种方法被广泛应用于关键物理现象的识别任务。而无监督学习则不需要预先标注,它更擅长从原始数据中发现潜在结构。当研究人员面对全新的实验现象时,这类方法能自动将相似特征的事件归类,帮助发现可能的物理信号。常用的技术包括主成分分析(Principal Component Analysis, PCA)和自编码器:PCA通过线性变换找到数据的主要差异方向,将上千维的粒子动量数据压缩到几个关键维度;自编码器则通过神经网络重建数据的过程,自动提取出最本质的特征。强化学习虽然当前应用较少,但其“试错优化”的特点在实验控制调节等场景展现出潜力,例如使用强化学习方法,自动调整相对论重离子对撞机(Relativistic Heavy Ion Collider, RHIC)中加速器设置,以实现轨迹调节和优化,提高加速器操作效率^[29]。因此,解决高能核物理中的不同问题可以利用不同的训练方式,例如利用监督学习预测核性质或分类碰撞事件^[30],借助无监督学习分析数据内在结构或进行异常信号检测,相关综述已经对这些机器学习方法及其应用场景进行了系统性的梳理^[31-32]。

这些方法的组合应用正在改变传统研究流程。

过去需要数月时间的人工数据分析工作,现在可以通过监督学习模型在数小时内完成初步分类;曾经依赖物理直觉的假设验证,现在能够借助无监督学习进行系统性探索。更重要的是,机器学习不仅提高了数据处理效率,还帮助研究人员注意到那些容易被人类忽视的微弱关联,例如不同碰撞几何条件下集体流模式的细微差异,或是粒子关联中暗示的临界涨落特征。这种从“人工筛检”到“智能挖掘”的转变,标志着高能核物理研究正在进入数据驱动的新阶段。

2 机器学习在核物理研究中的早期探索

早在20世纪90年代,研究者便开始尝试利用人工神经网络(Artificial Neural Network, ANN)研究核物理问题。例如在实验中,Bass等^[33]使用人工神经网络,以末态重子的动量分布作为输入,进行碰撞参数重构,其性能比经典技术提高近两倍,并发现对于该任务的输入表示,简单的网络架构,甚至无隐藏层的感知机就足以有效重构;而在理论研究上,St. Louis-Urbana合作组在计算实验中成功应用人工神经网络,预测核物理性质,包括原子质量、分离能以及壳稳定超重核的存在范围^[34]。人工神经网络展现出极高的准确率让研究人员确信,ANN具备显著的预测能力,将为核性质计算及多体系统的唯象分析提供了全新的工具。这些探索作为机器学习与核物理结合的先驱代表之一,通过与传统的核物理模型(如液滴模型)对比,验证了神经网络在预测核稳定性方面的优势,并作为传统核物理模型的有力补充。但对于极端条件下的核稳定性(如高温、高压环境),模型的适用性及其泛化能力仍需进一步验证和完善。这些成果构成了机器学习与高能核物理交叉发展的重要起点,并被回顾性综述视为开创性工作^[19]。

受此影响,后续研究进一步探索并拓展了ANN在核自旋与宇称系统学中的应用^[35]。Costiris等^[36]通过神经网络实现了对复杂核系统学数据的高效建模,优化了核系统学的参数预测,提升了模型的拟合精度,提供了一种新的分析工具,但如何在其他核物理领域(如粒子物理、强相互作用等)推广这些方法仍待进一步探索。

除此之外,其他的传统机器学习算法,例如图1所示的支持向量机(Support Vector Machine, SVM)也被应用于核物理理论分析中。Clark等^[37]采用支持向量机研究核质量、 β 衰变以及基态自旋/宇称的系统性规律,进一步验证了机器学习方法在核物理研究中的可行性。Clark等通过该方法对核性质进行全局预测,克服传统方法在预测核物质特性方面

的局限性,成功地将支持向量机应用于核物理中的多种预测任务,并且提供了比传统方法更为精确的结果。但是支持向量机相比于深度学习算法等,在处理多分类问题时,往往需要构建多个二分类器,训练复杂度较高;其次,其性能常深度依赖核函数、惩罚因子等参数,需要多次实验调整参数;再次,传统的支持向量机需要人工设计的特征作为输入,无法自动从数据中学习多层次的特征表示。

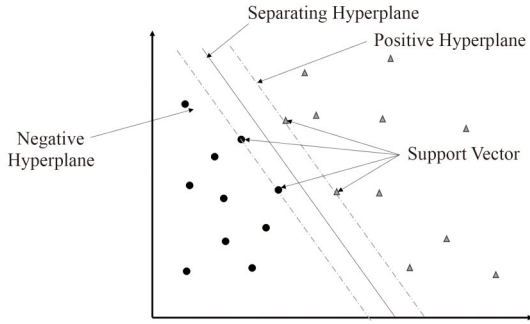


图1 支持向量机基本原理示意图

Fig.1 Schematic illustration of the fundamental principles of support vector machines

3 点云深度学习的发展

尽管早期传统机器学习方法在核物理研究中取得了初步成功,面对高能重离子碰撞产生的海量数据与多维度特征提取需求,传统算法逐渐暴露出表征能力不足、高度依赖人工设计的特征等瓶颈。这一现实需求推动了诸如卷积神经网络(Convolutional Neural Network, CNN)、点云神经网络(Point Cloud Network, PCN)等深度学习技术在核物理中的应用。事实上,重离子对撞(Heavy Ion Collision, HIC)产生的实验数据主要由粒子的四动量及其他属性构成,而早期研究通常将数据转换为2D图像,以便CNN处理^[38]。如Pang等^[38]将重离子碰撞末态带电粒子数据处理为横动量与方位角组成的二维图像,通过卷积神经网络实现了高精度区分不同状态方程的流体力学演化,并通过预测差异分析(Prediction Difference Analysis, PDA)将CNN模型中对分类决策最重要的输入特征区域;Du等^[39]在此基础上,利用CNN从pion谱中识别QCD状态方程,系统评估了不同数据处理方式对识别性能的影响,证明了在复杂混合模型框架下,CNN仍能从末态pion谱中解码状态方程信息,准确率表现优于全连接深度神经网络(Deep Neural Network, DNN),同时明确指出统计平均是抑制涨落显著提升CNN识别精度的有效手段。

随着点云分析技术的发展,基于点云的深度学

习模型逐渐应用于相对论重离子碰撞研究。相比于传统方法,点云网络能够直接处理粒子数据,无须额外的数据预处理,提高了分析效率的同时保留了粒子间高维度特征的关联^[40]。值得注意的是,各种点云算法的引入与高能重离子碰撞模拟数据的形式密不可分。以多相输运(A Multi-Phase Transport, AMPT)模型为例,其作为一种多阶段输运模型,不仅能够很好地符合实验数据,模拟高能重离子碰撞,还可以为深度学习方法提供可信的数据来源。AMPT模型的核心由4个主要部分组成:初始条件、部分子级联、强子化过程和强子级联,其模拟结果可以有效地描述粒子的快度分布、横向动量谱、椭圆流,并与实验数据高度一致^[41]。特别是弦熔化(string-melting)版本的AMPT模型,能够同时描述小系统和大系统碰撞中产生的集体流与粒子关联^[42]。如图2所示,AMPT模型生成的末态数据格式本质上就是点云数据,包含末态粒子的四动量(能量和动量分量)以及粒子种类等特征。这种数据结构与机器学习算法的训练和测试所要求的数据结果完美契合,尤其是涉及粒子物理和复杂系统分析的任务。

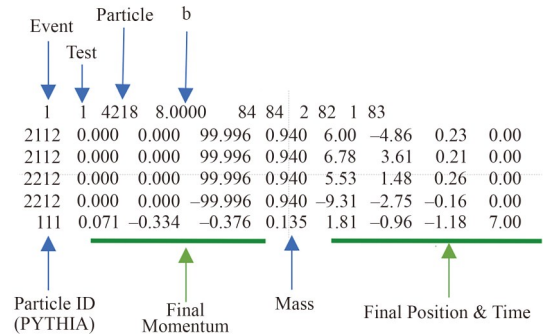


图2 AMPT输出结果示意图

Fig.2 A schematic representation of AMPT output results

早期的点云模型如PointNet解决了点云数据的置换不变性问题^[43]。如图3所示,Guo等^[44]利用PointNet模型,通过监督学习技术,成功实现了碰撞系统的特征辨识。在AMPT模型所生成的p+Pb与Pb+Pb碰撞事件中,PointNet有效捕捉到小系统(p+A)与大系统(A+A)碰撞中的差异。但是仍然需要进一步研究,完善对于大小碰撞系统的流的研究,而且PointNet网络本身相比于其他深度学习网络,局部特征提取能力相对较弱,无法捕捉局部的复杂几何关系。为提升点云网络提取复杂局部特征的能力,Huang等^[45]利用JAM模型和Critical Monte-Carlo模型生成模拟数据,并转换到 (p_x, p_y) 平面生成训练数据,采用动态边缘卷积神经网络,利用监督学习技

术,建立粒子间的关联拓扑,成功探测到高能重离子碰撞中可能存在的临界现象特征;Qu等^[46]设计并使用了一种名为ParticleNet的点云网络架构,通过学习粒子间的邻近关系和层级化特征,有效提取喷流信息,其性能表现超越了深度卷积网络(ResNeXt-50)、一维卷积网络(P-CNN)或集合网络(PFN)在内的其他深度学习方法。近年来,PointNet++也进一步引入了局部特征提取机制,使其能够更精准地刻画点云系统的局部细节^[47];基于掩码自编码器的点

云重建方法为高维数据的特征学习提供了新的途径^[48];Point Transformer通过注意力机制增强了点云数据的全局建模能力^[49-52]。由于Transformer架构在特征提取领域展现出的优秀性能,Qu等^[53]构造了一个大规模的喷注数据集JETCLASS,作为推动领域发展的重要资源,并在此基础上设计了Particle Transformer架构,同时利用单个粒子特征和粒子对的相互作用信息,在JETCLASS上超越ParticleNet,并通过预训练+微调显著提升了现有的性能纪录。

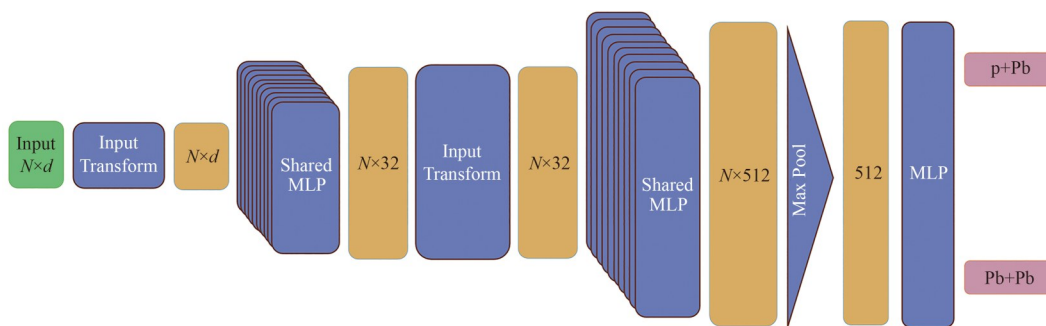


图3 PointNet基本结构示意图^[44]
Fig.3 The fundamental architecture of PointNet^[44]

将这些算法引入高能核物理的背后,离不开多元模拟数据生态的支撑。除了AMPT模型外,其余唯象模型也在高能重离子碰撞与人工智能结合的研究中发挥重要作用。如VISH2+1模型输出的连续场数据(如能量密度分布)^[54-55]以及iEBE-VISHNU模型生成的事件-事件波动数据^[39,56]JAM模型和Critical Monte-Carlo模型提供的临界现象模拟数据^[45,57-58],这些模拟方法与AMPT形成互补,共同构成点云算法在相变探测与系统分类等多任务中丰富而可靠的训练数据来源。然而,面对如此多元复杂的物理数据体系,如何有效提取其内在物理特征是理解数据背后物理含义的重要问题。

4 无监督学习驱动的特征提取

相较于前文提到的基于监督学习的识别分类方式,高维物理数据的内在规律挖掘需要更有效的特征提取方式,因为高能重离子碰撞得到的模拟数据与实验数据通常具有高维度和复杂的特征,例如粒子的能量分布、动量分布、角度信息等,因此这往往需要使用无监督学习驱动的降维方法进行特征提取。例如,借助主成分分析,Bożek等^[59]研究了不同微分流之间的耦合,揭示了其对粒子分布和流体行为的影响。该方法为未来的高能碰撞实验提供了有效的数据分析工具,尤其在提取重要物理信息方面具有很大的潜力。不过,目前仍需在更高能量范围

内验证其对非线性耦合效应的普适性,并进一步与更复杂的实验数据做精确对比,这是该研究尚未解决的主要问题。此外,Liu等^[54]通过主成分分析法研究高能重离子碰撞中的集体流,探索了如何从原始数据中自动发现重离子碰撞中的流动特征,而不依赖于传统的人为设置的傅里叶变换。结论表明:PCA方法加强了流动谐波和初始偏心度之间的线性关系,尤其对于高阶流动成分($n \geq 4$),这与传统认为的强非线性耦合效应相对立^[54]。但是,该方法虽然PCA在VISH2+1流体动力学模拟数据中取得了显著成果,但其在其他模拟数据与真实实验数据中的适用性仍需要进一步验证。

此外,自编码器(Auto-Encoder,AE)也在高能核物理中得到了广泛应用,尤其是在异常检测任务中。面对碰撞得到的高维度复杂实验数据,传统的降维方法一般依赖有限的线性操作来提取特征,但在处理复杂的非线性关系时存在局限性。自编码器是一种以深度神经网络为基础的无监督学习模型,它通过学习数据的低维潜在表示,在保持重要特征的同时减少数据冗余。例如,在寻找新物理信号时,Finke等^[60]借助如图4所示的基于图像自编码器网络(Image-based autoencoder network),通过重构误差衡量异常程度,处理了模型无关的喷注(jet)分类任务,并探索了改进自编码器结构的方法。具体而言,这项研究所使用的自编码器网络是一个典型的卷积自编码器,由编码器(encoder)与解码器

(decoder) 组成。在编码器中, 多层卷积层(convolution)使用 4×4 的卷积核进行局部特征提取, 多层平均池化层(average pooling)则使用 2×2 的池化层进行降维, 最后通过全连接层进一步压缩维度, 将数据映射到 32 维的潜在空间中; 而解码器则采用与解码器对称点反卷积(transpose convolution)与上采样(upsampling)层, 将低维潜在表示重构为原始输入数据。这项研究通过自编码器(autoencoders)方法实现了高能物理中的无监督异常检测, 展现了其在大规模实验数据中的强大潜力。

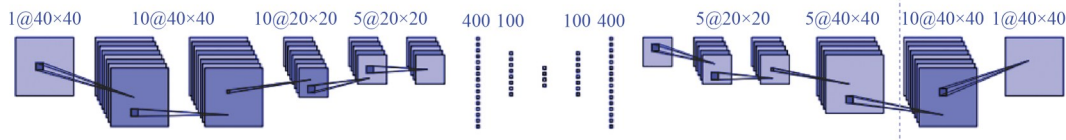


图4 基于图像的自编码器网络基本结构^[60]
Fig.4 The fundamental architecture of an image-based autoencoder network^[60]

近期, Wang 等^[61]也采用自编码器网络, 直接从高能重离子碰撞的原始末态数据中分析相变, 确定液气相变的临界温度。通过自编码器对核液-气相变进行了研究, 展示了该方法在复杂核物理问题中的应用潜力。通过学习实验数据中的隐含物理规律, 深度神经网络能够准确预测核物质的相变行为, 并为传统物理模型提供了有效的补充。其采用的 MLP 自编码器只由全连接层组成, 没有卷积、注意力机制等复杂操作, 参数量相对较少, 对小规模数据集更为友好, 能有效避免卷积网络或 transformer 等架构在小规模数据上常出现的过拟合问题, 但也正因为如此, 这种简单的自编码器只能学习全局特征, 不具备局部感受野(local receptive field), 无法像卷积神经网络那样有效提取局部特征, 也无法像 transformer 架构那样捕捉长距离依赖。

5 机器学习在高能核物理中的前景

机器学习技术的发展正在深刻影响高能核物理的研究。随着计算硬件(如 GPU、TPU)的升级以及深度学习算法(如 Transformer、扩散模型、图神经网络)的不断优化, 深度学习正逐步从数据分析的辅助工具, 演变为提供核心科学洞见的关键方法。在高能核物理研究中, 实验数据往往具有极高的维度和复杂的特征, 包括末态粒子的动量分布、角度信息、能量损失、集体流效应等。传统分析方法依赖于手工设计的统计观测量, 而深度学习的引入使得研究者能够直接从原始数据中自动提取潜在的物理特性, 并提高物理模型的拟合精度。

近年来, 深度学习已经在多个高能核物理研究

自编码器能够有效地识别出与正常模式显著不同的异常事件, 尤其在处理复杂、高维的高能物理数据时, 展现出较高的准确性和鲁棒性。该方法为高能物理实验中的异常事件识别提供了一种新的有效工具, 并且具有广泛的应用前景。但是, 这种基于卷积的图像自编码器, 在一定程度上, 更加擅长局部信息的提取, 而非全局特征的理解, 尤其是相较于 transformer 或扩散模型等神经网络而言, 可能无法充分捕捉长距离依赖关系。此外, 当图像分辨率增加时, 卷积运算往往需要极大的计算与存储开销。

方向展现了显著优势, 但其可解释性问题仍然是当前研究的重要挑战^[62]。神经网络存在“黑箱”问题, 研究者难以理解网络在不同层级学习到的物理模式。为了解决这一问题, 未来研究需要探索结合物理规律的可解释深度学习方法。例如, 基于 SHAP (Shapley Additive Explanations) 方法^[63]的特征分析模型能够量化不同输入变量(如粒子动量、碰撞中心度)对最终预测结果的贡献, 从而揭示深度网络所学习到的集体流信号或临界涨落的形成机理^[64]。此外, 符号回归方法结合神经网络的策略也值得探索, 它可以自动学习解析公式, 将神经网络提取的模式转换为物理学可解释的数学表达式, 提高模型的透明度和可信度^[65-66]。

另一方面, 高能核物理研究中常面临数据稀缺和小样本学习的挑战。由于实验运行成本高昂, 某些关键物理现象(如 QGP 形成、喷注淬火)仅在特定能区和碰撞条件下出现, 可用于训练深度神经网络模型的数据极为有限。针对这一问题, 可以借鉴计算机视觉领域的元学习(Meta-Learning)策略^[67], 构建预训练-微调(Pretraining-Finetuning)框架。具体而言, 研究者可以先使用 AMPT 等物理模型模拟数据训练一个基础特征提取网络, 使其学习粒子动量分布、碰撞系统特征等基本模式, 然后结合真实实验数据(如 ALICE 观测数据)进行微调, 以筛选关键物理特征, 同时降低对大量实验数据的依赖。此外, 结合自注意力机制(Self-Attention)可以让神经网络更加关注稀有现象, 如高 p_T 事件或高阶流信号, 提高模型在小样本环境下的泛化能力。

除了元学习, 数据增强(Data Augmentation)和

自监督学习(Self-Supervised Learning, SSL)也是提高模型健壮性的重要策略^[68-70]。例如,在实验数据有限的情况下,可以通过扰动粒子动量分布、旋转几何结构等方式生成更多样本,以增强模型的泛化能力。同时,自监督学习可以利用无标签数据进行预训练,例如采用对比学习方法,让神经网络自主学习粒子分布的内在模式,减少对人工标注数据的依赖。这些方法的结合将有助于攻克小样本学习瓶颈,使深度学习模型能够在有限的实验数据条件下依然保持高精度的物理分析能力。

展望未来,随着计算机科学与高能核物理的进一步交叉融合,机器学习有望在更多方向上拓展应用边界。例如,量子机器学习(Quantum Machine Learning, QML)可以利用量子计算的加速能力,优化高维数据的特征提取,提高QGP模型的计算效率^[71-73];扩散模型(Diffusion Model)在粒子生成中的应用也值得探索^[74-75],研究者可以使用扩散模型生成更符合实验分布的事件;此外,强化学习(Reinforcement Learning, RL)还可以用于优化高能核物理实验的探测器参数设置,从而提高数据采集的效率^[76-77]。

综上所述,深度学习的引入不仅提供了更强大的数据分析工具,还可能在未来为高能核物理研究提供全新的研究视角。为了充分发挥机器学习的潜力,研究者需要进一步探索深度学习与物理规律的结合方式,并不断优化模型的可解释性和泛化能力,以推动该领域的深入发展。

作者贡献声明 张靖宗负责文章的初稿写作;郭爽、朱励霖、王凌霄、马国亮负责文章整体设计,理论研究指导,对文章进行审阅与修订。

参考文献

- Adams J, Aggarwal M M, Ahammed Z. Experimental and theoretical challenges in the search for the quark-gluon plasma: The STAR Collaboration's critical assessment of the evidence from RHIC collisions[J]. *Nuclear Physics A*, 2005, **757**: 102 - 183. DOI: 10.1016/j.nuclphysa.2005.03.085.
- Adcox K, Adler S S, Afanasiev S. Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the PHENIX collaboration[J]. *Nuclear Physics A*, 2005, **757**: 184 - 283. DOI: 10.1016/j.nuclphysa.2005.03.086.
- Schenke B, Jeon S, Gale C. Elliptic and triangular flow in event-by-event $D=3+1$ viscous hydrodynamics[J]. *Physical Review Letters*, 2011, **106**(4): 042301. DOI: 10.1103/physrevlett.106.042301.
- Heinz U, Snellings R. Collective flow and viscosity in relativistic heavy-ion collisions[J]. *Annual Review of Nuclear and Particle Science*, 2013, **63**: 123 - 151. DOI: 10.1146/annurev-nucl-102212-170540.
- Qin G Y. Theory of jet quenching in ultra-relativistic nuclear collisions[J]. *Nuclear Physics A*, 2014, **931**: 165 - 175. DOI: 10.1016/j.nuclphysa.2014.09.004.
- Qin G Y, Wang X N. Jet quenching in high-energy heavy-ion collisions[J]. *International Journal of Modern Physics E*, 2015, **24**(11): 1530014. DOI: 10.1142/s0218301315300143.
- Song H C, Bass S A, Heinz U, *et al.* 200 AGeV Au+Au collisions serve a nearly perfect quark-gluon liquid[J]. *Physical Review Letters*, 2011, **106**(19): 192301. DOI: 10.1103/physrevlett.106.192301.
- Müller B, Nagle J L. Results from the relativistic heavy ion collider[J]. *Annual Review of Nuclear and Particle Science*, 2006, **56**: 93 - 135. DOI: 10.1146/annurev-nucl.56.080805.140556.
- Gale C, Jeon S, Schenke B. Hydrodynamic modeling of heavy-ion collisions[J]. *International Journal of Modern Physics A*, 2013, **28**(11): 1340011. DOI: 10.1142/s0217751x13400113.
- Karsch F, Bazavov A, Ding H T, *et al.* Conserved charge fluctuations from lattice QCD and the beam energy scan [J]. *Nuclear Physics A*, 2016, **956**: 352 - 355. DOI: 10.1016/j.nuclphysa.2016.01.008.
- Xu H J, Li Z P, Song H C. High-order flow harmonics of identified hadrons in 2.76 ATeV Pb+Pb collisions[J]. *Physical Review C*, 2016, **93**(6): 064905. DOI: 10.1103/physrevc.93.064905.
- Zhu X, Zhou Y, Xu H. Correlations of flow harmonics in 2.76 ATeV Pb-Pb collisions[J]. *Physical Review C*, 2017, **95**: 044902. DOI: 10.1103/PhysRevC.95.044902.
- Chang N B, Cao S S, Qin G Y. Probing medium-induced jet splitting and energy loss in heavy-ion collisions[J]. *Physics Letters B*, 2018, **781**: 423 - 432. DOI: 10.1016/j.physletb.2018.04.019.
- Acharya S, Adamová D, Adler A, *et al.* (ALICE Collaboration). The ALICE experiment: a journey through QCD[J]. *The European Physical Journal C*, 2024, **84**(8): 813. DOI: 10.1140/epjc/s10052-024-12935-y.
- Bishop C M, Bishop H. Deep learning: Foundations and concepts[M]. Springer Nature, 2023. DOI: 10.1007/978-3-031-45468-4_6.

- 16 He J J, He W B, Ma Y G, *et al.* Machine-learning-based identification for initial clustering structure in relativistic heavy-ion collisions[J]. *Physical Review C*, 2021, **104**(4): 044902. DOI: 10.1103/physrevc.104.044902.
- 17 王凌霄, 庞龙刚, 周凯. 深度学习在核物理中的应用[J]. *中国科学: 物理学 力学 天文学*, 2022, **52**(5): 252003. DOI: 10.1360/SSPMA-2021-0300.
WANG Lingxiao, PANG Longgang, ZHOU Kai. Applications of deep learning in high energy nuclear physics[J]. *Scientia Sinica Physica, Mechanica & Astronomica*, 2022, **52**(5): 252003. DOI: 10.1360/SSPMA-2021-0300.
- 18 Zhou K, Wang L X, Pang L G, *et al.* Exploring QCD matter in extreme conditions with Machine Learning[J]. *Progress in Particle and Nuclear Physics*, 2024, **135**: 104084. DOI: 10.1016/j.ppnp.2023.104084.
- 19 Boehnlein A, Diefenthaler M, Sato N, *et al.* *Colloquium*: Machine learning in nuclear physics[J]. *Reviews of Modern Physics*, 2022, **94**(3): 031003. DOI: 10.1103/revmodphys.94.031003.
- 20 He W B, Ma Y G, Pang L G, *et al.* High-energy nuclear physics meets machine learning[J]. *Nuclear Science and Techniques*, 2023, **34**(6): 88. DOI: 10.1007/s41365-023-01233-z.
- 21 Aarts G, Fukushima K, Hatsuda T, *et al.* Physics-driven learning for inverse problems in quantum chromodynamics[J]. *Nature Reviews Physics*, 2025, **7**: 154 – 163. DOI: 10.1038/s42254-024-00798-x.
- 22 Ma Y G, Pang L G, Wang R, *et al.* Phase transition study meets machine learning[J]. *Chinese Physics Letters*, 2023, **40**(12): 122101. DOI: 10.1088/0256-307X/40/12/122101.
- 23 Zhang R R, Guo Z Y, Gao P, *et al.* Point-M2AE: multi-scale masked autoencoders for hierarchical point cloud pre-training[J]. *Advances in Neural Information Processing Systems*, 2022, **35**: 27061 – 27074. DOI: 10.48550/arXiv.2205.14401.
- 24 Zhang R R, Wang L H, Qiao Y, *et al.* Learning 3D representations from 2D pre-trained models via image-to-point masked autoencoders[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 17-24, 2023, Vancouver, BC, Canada. IEEE, 2023: 21769 – 21780. DOI: 10.1109/CVPR52729.2023.02085.
- 25 Yu X M, Tang L L, Rao Y M, *et al.* Point-BERT: pre-training 3D point cloud transformers with masked point modeling[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 19291 – 19300. DOI: 10.1109/CVPR52688.2022.01871.
- 26 Liu H T, Cai M, Lee Y J. Masked discrimination for self-supervised learning on point clouds[M]. *Computer Vision – ECCV 2022*. Cham: Springer Nature Switzerland, 2022: 657 – 675. DOI: 10.1007/978-3-031-20086-1_38.
- 27 Mehta P, Bukov M, Wang C H, *et al.* A high-bias, low-variance introduction to machine learning for physicists [J]. *Physics Reports*, 2019, **810**: 1 – 124. DOI: 10.1016/j.physrep.2019.03.001.
- 28 Buncic P, Krzewicki M, Vyvře P. Technical design report for the upgrade of the online-offline computing system [R]. 2015.
- 29 Kain V, Hirlander S, Goddard B, *et al.* Sample-efficient reinforcement learning for CERN accelerator control[J]. *Physical Review Accelerators and Beams*, 2020, **23**(12): 124801. DOI: 10.1103/physrevaccelbeams.23.124801.
- 30 高泽鹏, 李庆峰. 利用机器学习方法对几个核物理问题的深入研究[J]. *核技术*, 2023, **46**(8): 080009. DOI: 10.11889/j.0253-3219.2023.hjs.46.080009.
GAO Zepeng, LI Qingfeng. Studies on several problems in nuclear physics by using machine learning[J]. *Nuclear Techniques*, 2023, **46**(8): 080009. DOI: 10.11889/j.0253-3219.2023.hjs.46.080009.
- 31 李甫鹏, 庞龙刚, 王新年. 基于机器学习的重离子碰撞中 QCD 相变的研究[J]. *核技术*, 2023, **46**(4): 040014. DOI: 10.11889/j.0253-3219.2023.hjs.46.040014.
LI Fupeng, PANG Longgang, WANG Xinnian. Application of machine learning to the study of QCD transition in heavy ion collisions[J]. *Nuclear Techniques*, 2023, **46**(4): 040014. DOI: 10.11889/j.0253-3219.2023.hjs.46.040014.
- 32 He W B, Li Q F, Ma Y G, *et al.* Machine learning in nuclear physics at low and intermediate energies[J]. *Science China Physics, Mechanics & Astronomy*, 2023, **66** (8): 282001. DOI: 10.1007/s11433-023-2116-0.
- 33 Bass S A, Bischoff A, Maruhn J A, *et al.* Neural networks for impact parameter determination[J]. *Physical Review C*, 1996, **53**(5): 2358 – 2363. DOI: 10.1103/physrevc.53.2358.
- 34 Gazula S, Clark J W, Bohr H. Learning and prediction of nuclear stability by neural networks[J]. *Nuclear Physics A*, 1992, **540**(1 – 2): 1 – 26. DOI: 10.1016/0375-9474(92)90191-L.
- 35 Gernoth K A, Clark J W, Prater J S, *et al.* Neural network

- models of nuclear systematics[J]. *Physics Letters B*, 1993, **300**(1 - 2): 0370269393907384. DOI: 10.1016/0370-2693(93)90738-4.
- 36 Costiris N, Mavrommatis E, Gernoth K, *et al.* Decoding beta-decay systematics: a global statistical model for beta-half-lives[J]. *Physical Review C*, 2009, **80**: 044332. DOI: 10.1103/PhysRevC.80.044332.
- 37 Clark J W, Li H C. Application of support vector machines to global prediction of nuclear properties[J]. *International Journal of Modern Physics B*, 2006, **20**(30n31): 5015 - 5029. DOI: 10.1142/s0217979206036053.
- 38 Pang L G, Zhou K, Su N, *et al.* An equation-of-state-meter of quantum chromodynamics transition from deep learning[J]. *Nature Communications*, 2018, **9**(1): 210. DOI: 10.1038/s41467-017-02726-3.
- 39 Du Y L, Zhou K, Steinheimer J, *et al.* Identifying the nature of the QCD transition in relativistic collision of heavy nuclei with deep learning[J]. *The European Physical Journal C*, 2020, **80**(6): 516. DOI: 10.1140/epjc/s10052-020-8030-7.
- 40 Mikuni V, Canelli F. Point cloud transformers applied to collider physics[J]. *Machine Learning: Science and Technology*, 2021, **2**(3): 035027. DOI: 10.1088/2632-2153/ac07f6.
- 41 Lin Z W, Ko C M, Li B A, *et al.* Multiphase transport model for relativistic heavy ion collisions[J]. *Physical Review C*, 2005, **72**(6): 064901. DOI: 10.1103/physrevc.72.064901.
- 42 Lin Z W, Zheng L. Further developments of a multi-phase transport model for relativistic nuclear collisions[J]. *Nuclear Science and Techniques*, 2021, **32**(10): 113. DOI: 10.1007/s41365-021-00944-5.
- 43 Charles R Q, Hao S, Mo K C, *et al.* PointNet: deep learning on point sets for 3D classification and segmentation[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 77 - 85. DOI: 10.1109/CVPR.2017.16.
- 44 Guo S, Wang H S, Zhou K, *et al.* Machine learning study to identify collective flow in small and large colliding systems[J]. *Physical Review C*, 2024, **110**(2): 024910. DOI: 10.1103/physrevc.110.024910.
- 45 Huang Y G, Pang L G, Luo X F, *et al.* Probing criticality with deep learning in relativistic heavy-ion collisions[J]. *Physics Letters B*, 2022, **827**: 137001. DOI: 10.1016/j.physletb.2022.137001.
- 46 Qu H L, Gouskos L. Jet tagging *via* particle clouds[J]. *Physical Review D*, 2020, **101**(5): 056019. DOI: 10.1103/physrevd.101.056019.
- 47 Qi C R, Yi L, Su H, *et al.* PointNet++: deep hierarchical feature learning on point sets in a metric space[J]. *Advances in Neural Information Processing Systems*, 2017, **30**. DOI: 10.48550/arXiv.1706.02413.
- 48 Sohail S S, Himeur Y, Kheddar H, *et al.* Advancing 3D point cloud understanding through deep transfer learning: a comprehensive survey[J]. *Information Fusion*, 2025, **113**: 102601. DOI: 10.1016/j.inffus.2024.102601.
- 49 Zhao H S, Jiang L, Jia J Y, *et al.* Point transformer[C]. 2021 IEEE/CVF International Conference on Computer Vision (ICCV). October 10-17, 2021, Montreal, QC, Canada. IEEE, 2021: 16239 - 16248. DOI: 10.1109/ICCV48922.2021.01595.
- 50 Wu X, Lao Y, Jiang L, *et al.* Point transformer v2: Grouped vector attention and partition-based pooling[J]. *Advances in Neural Information Processing Systems*, 2022, **35**: 33330 - 33342. DOI: 10.48550/arXiv.2210.05666.
- 51 Wu X Y, Jiang L, Wang P S, *et al.* Point transformer V3: simpler, faster, stronger[C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 16-22, 2024, Seattle, WA, USA. IEEE, 2024: 4840 - 4851. DOI: 10.1109/CVPR52733.2024.00463.
- 52 Wu X Y, Xu X, Kong L D, *et al.* Point transformer V3 extreme: 1st place solution for 2024 waymo open dataset challenge in semantic segmentation[EB/OL]. 2024: 2407.15282. <https://arxiv.org/abs/2407.15282v1>.
- 53 Qu H, Li C, Qian S. Particle transformer for jet tagging [C]. *International Conference on Machine Learning*. PMLR, 2022: 18281 - 18292. DOI: 10.48550/arXiv.2202.03772.
- 54 Liu Z M, Zhao W B, Song H C. Principal component analysis of collective flow in relativistic heavy-ion collisions[J]. *The European Physical Journal C*, 2019, **79**(10): 870. DOI: 10.1140/epjc/s10052-019-7379-y.
- 55 Song H C, Heinz U. Causal viscous hydrodynamics in 2+1 dimensions for relativistic heavy-ion collisions[J]. *Physical Review C*, 2008, **77**(6): 064901. DOI: 10.1103/physrevc.77.064901.
- 56 Shen C, Qiu Z, Song H C, *et al.* The iEBE-VISHNU code package for relativistic heavy-ion collisions[J]. *Computer Physics Communications*, 2016, **199**: 61 - 85. DOI:

- 10.1016/j.cpc.2015.08.039.
- 57 Nara Y. JAM: an event generator for high energy nuclear collisions[J]. EPJ Web of Conferences, 2019, **208**: 11004. DOI: 10.1051/epjconf/201920811004.
 - 58 Wu J, Lin Y F, Wu Y F, *et al.* Probing QCD critical fluctuations from intermittency analysis in relativistic heavy-ion collisions[J]. Physics Letters B, 2020, **801**: 135186. DOI: 10.1016/j.physletb.2019.135186.
 - 59 Bożek P. Principal component analysis of the nonlinear coupling of harmonic modes in heavy-ion collisions[J]. Physical Review C, 2018, **97**(3): 034905. DOI: 10.1103/physrevc.97.034905.
 - 60 Finke T, Krämer M, Morandini A, *et al.* Autoencoders for unsupervised anomaly detection in high energy physics [J]. Journal of High Energy Physics, 2021, **2021**(6): 161. DOI: 10.1007/JHEP06(2021)161.
 - 61 Wang R, Ma Y G, Wada R, *et al.* Nuclear liquid-gas phase transition with machine learning[J]. Physical Review Research, 2020, **2**(4): 043202. DOI: 10.1103/physrevresearch.2.043202.
 - 62 Tanaka A, Tomiya A, Hashimoto K. Deep learning and physics[M]. Singapore: Springer, 2021. DOI: 10.1007/978-981-33-6108-9.
 - 63 Lundberg S M, Lee S I. A unified approach to interpreting model predictions[J]. Advances in Neural Information Processing Systems, 2017, **30**. DOI: 10.48550/arXiv.1705.07874.
 - 64 Pezoa R, Salinas L, Torres C. Explainability of high energy physics events classification using SHAP[J]. Journal of Physics: Conference Series, 2023, **2438**(1): 012082. DOI: 10.1088/1742-6596/2438/1/012082.
 - 65 Angelis D, Sofos F, Karakasidis T E. Artificial intelligence in physical sciences: symbolic regression trends and perspectives[J]. Archives of Computational Methods in Engineering, 2023, **30**(6): 3845 – 3865. DOI: 10.1007/s11831-023-09922-z.
 - 66 Udrescu S M, Tegmark M. AI Feynman: a physics-inspired method for symbolic regression[J]. Science Advances, 2020, **6**(16): eaay2631. DOI: 10.1126/sciadv.aay2631.
 - 67 Vilalta R, Drissi Y. A perspective view and survey of meta-learning[J]. Artificial Intelligence Review, 2002, **18** (2): 77 – 95. DOI: 10.1023/A:1019956318069.
 - 68 Maharana K, Mondal S, Nemade B. A review: Data pre-processing and data augmentation techniques[J]. Global Transitions Proceedings, 2022, **3**(1): 91 – 99. DOI: 10.1016/j.gltp.2022.04.020.
 - 69 Shorten C, Khoshgoftaar T M. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, **6**(1): 60. DOI: 10.1186/s40537-019-0197-0.
 - 70 Mumuni A, Mumuni F. Data augmentation: a comprehensive survey of modern approaches[J]. Array, 2022, **16**: 100258. DOI: 10.1016/j.array.2022.100258.
 - 71 Biamonte J, Wittek P, Pancotti N, *et al.* Quantum machine learning[J]. Nature, 2017, **549**(7671): 195 – 202. DOI: 10.1038/nature23474.
 - 72 Schuld M, Sinayskiy I, Petruccione F. An introduction to quantum machine learning[J]. Contemporary Physics, 2015, **56**(2): 172 – 185. DOI: 10.1080/00107514.2014.964942.
 - 73 Cerezo M, Verdon G, Huang H Y, *et al.* Challenges and opportunities in quantum machine learning[J]. Nature Computational Science, 2022, **2**(9): 567 – 576. DOI: 10.1038/s43588-022-00311-3.
 - 74 Mikuni V, Nachman B. Score-based generative models for calorimeter shower simulation[J]. Physical Review D, 2022, **106**(9): 092009. DOI: 10.1103/physrevd.106.092009.
 - 75 Devlin P, Qiu J W, Ringer F, *et al.* Diffusion model approach to simulating electron-proton scattering events [J]. Physical Review D, 2024, **110**: 016030. DOI: 10.1103/physrevd.110.016030.
 - 76 Wiering M A, Van Otterlo M. Reinforcement learning[J]. Adaptation, Learning and Optimization, 2012, **12**(3): 729. DOI: 10.1007/978-3-642-27645-3.
 - 77 Li Y X. Deep reinforcement learning: an overview[EB/OL]. 2017: 1701.07274. <https://arxiv.org/abs/1701.07274v6>.