



Contribution ID: 126

Type: Oral

Reducing Systematic Differences between Data and Simulation with Generative Models

Wednesday, 13 March 2024 14:50 (20 minutes)

High Energy Physics (HEP) experiments rely on scientific simulation to develop reconstruction algorithms. Despite the remarkable fidelity of modern simulation frameworks, residual discrepancies between simulated and real data introduce a challenging domain shift problem. The existence of this issue raises significant concerns regarding the feasibility of implementing Deep Learning (DL) methods in detector analysis, impeding the adoption of such techniques.

We present our ongoing research in developing new DL strategies to mitigate the differences between simulation and real data. Our approach is based on a combination of Generative Adversarial Networks (GANs) for the translation of simulated samples into the real data domain, reducing the magnitude of domain shift effects.

We discuss our progress made in applying this approach specifically to LArTPC-based particle detectors. We demonstrate the effectiveness of our method on a simplified and cropped LArTPC benchmark dataset. Then we highlight various performance and computational challenges encountered in the process of adapting the method to realistic LArTPC datasets of high-resolution images ($\sim 6000 \times 960$ pixels). By systematically addressing these challenges, our research aims to bridge the gap between the simulated and real data and advance the applicability of DL methods in HEP experiments.

Significance

We successfully applied the proposed method on a toy problem (<https://arxiv.org/abs/2304.12858>). The novelty here is that we managed to make the method work on a realistic dataset of high-resolution images and made the method computationally efficient.

References

<https://arxiv.org/abs/2304.12858>

Experiment context, if any

ProtoDUNE, DUNE

Primary author: TORBUNOV, Dmitrii

Presenter: TORBUNOV, Dmitrii

Session Classification: Track 2: Data Analysis - Algorithms and Tools

Track Classification: Track 2: Data Analysis - Algorithms and Tools