

DEEP LEARNING METHODS AS A TOOL FOR OVERCOMING THE CRISIS OF PARTICLE TRACKING IN HIGH LUMINOSITY HEP EXPERIMENTS

G. A. Ososkov^{1, *}

¹ Joint Institute for Nuclear Research, Dubna, Russia

A key stage in offline processing of the experimental HEP data is the reconstruction of trajectories (tracks) of the interacting particles from measurement data. For modern high-luminosity collider experiments, such as HL-LHC and NICA, a particular challenge for tracking is the very high, megahertz frequency of interactions, leading to an order-of-magnitude increase in the intensity of the data stream to be processed and, in addition, to a significant overlap of event track data when they are registered in track detectors. All these circumstances, recognized by physicists as the “tracking crisis”, have shown that the tracking algorithms already in use are not efficient, accurate, and scalable enough to handle data obtained in high-luminosity collider experiments.

To overcome this crisis, in 2018, a group of physicists from CERN and other physics centers in the HEPTrkX project staged a TrackML competition to develop new solutions to tracking problems using deep neural networks. A data set for their training and testing was prepared and published on the Kaggle platform. The TrackML competition stimulated a lot of important research leading to the development of effective tracking algorithms based on graph neural networks, transformers, as well as the reanimation of tracking based on Hopfield neural networks, enhanced with computational means of adiabatic quantum computers.

The experience in the development of tracking algorithms based on machine learning methods, accumulated during the last decade by the specialists from MLIT JINR, allowed them to actively engage in research on overcoming the problems of the “tracking crisis” not only by using the information from already published results, but also through original innovations, taking into account the specificity of domestic detectors in the high-luminosity experiments of the NICA megaproject at JINR. We make a brief review of the ongoing work and discuss its prospects.

Ключевым этапом офлайн-обработки экспериментальных данных НЕР является реконструкция траекторий (треков) взаимодействующих частиц по данным измерений. Для современных экспериментов на коллайдерах высокой светимости, таких как HL-LHC и NICA, особой проблемой для отслеживания является очень высокая, мегагерцовая, частота взаимодействий, что приводит к увеличению интенсивности потока обрабатываемых данных на порядок и, кроме того,

* E-mail: gososkov@gmail.com

к значительному перекрытию данных о треках событий при их регистрации в трековых детекторах. Все эти обстоятельства, признанные физиками «кризисом отслеживания», показали, что используемые алгоритмы отслеживания недостаточно эффективны, точны и масштабируемы для обработки данных, полученных в экспериментах на коллайдерах высокой светимости.

Для преодоления этого кризиса в 2018 г. группа физиков из ЦЕРН и других физических центров в проекте NEPTrkX организовала конкурс TrackML для разработки новых решений задач отслеживания с использованием глубоких нейронных сетей. Подготовлен и опубликован на платформе Kaggle набор данных для их обучения и тестирования. Конкурс TrackML стимулировал множество важных исследований, приведших к разработке эффективных алгоритмов трекинга на основе графовых нейронных сетей, трансформаторов, а также к реанимации трекинга на основе нейронных сетей Хопфилда, улучшенных вычислительными средствами адиабатических квантовых компьютеров.

Накопленный за последнее десятилетие специалистами ЛИТ ОИЯИ опыт разработки алгоритмов трекинга на основе методов машинного обучения позволил активно включиться в исследования по преодолению проблем трекингового кризиса за счет не только использования информации из уже опубликованных результатов, но и оригинальных инноваций, учитывающих специфику отечественных детекторов в экспериментах на коллайдерах высокой светимости мегапроекта NICA в ОИЯИ. Представлен краткий обзор текущей работы, и обсуждаются ее перспективы.

PACS: 02.70.-с; 07.05.-t