Implementing deep learning-based disruption prediction in a shifting data environment of new tokamak: HL-3

Zongyu Yang[1], Wulyu Zhong[1], Fan Xia[1], Zhe Gao[2], Xiaobo Zhu[1], Jiyuan Li[1], Liwen Hu[1], Zhaohe Xu[1], Da Li[1], Guohui Zheng[3], Yihang Chen[1], Junzhao Zhang[1], Bo Li[1], Xiaolong Zhang[1], Yiren Zhu[1], Ruihai Tong[1], Yunbo Dong[1], Yipo Zhang[1], Boda Yuan[1], Xin Yu[1], Zongyuhui He[1], Wenjing Tian[1], Xinwen Gong[1] and Min Xu[1].

1 Southwestern Institute of Physics, Chengdu, China

2 Department of Engineering Physics, Tsinghua University, Beijing 100084, China

3 School of Physics, Nankai University, Tianjin 300071, China

e-mail (speaker): zy-yang@swip.ac.cn

Disruption occurs when the plasma confinement is suddenly lost during tokamak discharges, resulting in a rapid release of thermal and electromagnetic energy onto the plasma-facing components. [1]. In future large-scale tokamaks, disruptions could pose a safety threat to the device [2].In the past decades, numerous algorithms have been developed using physical criterions and data-driven methods to predict the disruptions on many tokamaks and have achieved high accuracies [3]. However, most of the algorithms are developed and tested in stable data environments provided by devices that have been operating for a long time. There is still a lack of experience about the implementation on new tokamaks, with the consideration of limited and unstable data available on these devices. In this research, a deep learning-based disruption predictor is developed in a new tokamak, HL-3. A series of techniques are developed to overcome the obstacles introduced by the unstable data environment. Finally, an algorithm with an Area Under receiver-operator characteristic Curve (AUC) of 0.940 is realized.

A comprehensive dataset is collected during the first several experimental campaigns of HL-3, ranging from Shot 1241 to Shot 4077. The dataset includes pre-disruption plasma parameters, disruption causes, operational strategies, and the harmful effects of disruptions. A systematic analysis of these data is conducted, providing valuable insights for the development of disruption prediction algorithms. The obstacles brought by the data environment of new device could be summarized into 3 aspects. Firstly, the amount of accumulated data is quite limited. Secondly, the plasma parameters in this new tokamak shift rapidly and could occasionally exceed the valid range of diagnostic, leading to frequent adjustments of the systems and breaks the availability and consistency of dataset. Last but not least, the distribution of disruption causes might frequently shift due to the adjustment of plasma parameters and operation strategies.

The technical paradigm of disruption prediction is based on our previous research in HL-2A [4]. A series of novel modules are developed to address the mentions three issues.

(1) Predict-first neural network (PFNN). The backbone of disruption prediction algorithm consists of 5 parts, as shown in figure 1. The left two parts are a surrogate model of offline EFIT and a physical feature extractor, which calculates some essential input features. The middle two parts are plasma current and shape (Ip & shape) predictor and plasma density predictor. They are two NN models trained to predict the evolution of plasma current, shape and density according to the previous status and control schedules. The predicted parameters will be used as supplementary inputs in the right part, namely, the disruption predictor. These new features aim to estimate if the plasma is under control and to introduce an inductive bias, i.e., uncontrollable plasma is more likely to disrupt, into the model.

(2) Data augmentation. Since the evolution of plasma might vary in speed according to complex experimental factors without violating the intrinsic physics, the input of algorithm can be slightly scaled in or out on the time axis. This transformation generates new data to confine the NN and alleviates the problem of lacking data.

(3) Pseudo data placeholder. Instead of simply inputting an array of diagnostic data into the algorithm, the HL-3 disruption predictor supplements another array to input. The new array consists of Boolean flags for the validity of each channel. A Pseudo data array is generated in NN and will replace the original data if the channel is not valid in a certain shot.

(4) Feature dropout. A more effective strategy is to actively adapt NN to missing input channels. Therefore, one can randomly select some of the diagnostics and set the validity flag as 0 in every training step. And NN will be expected to get used to missing input channels during training. Experimental result shows that intentional feature dropout can promote the accuracy of algorithm.

(5) Cross channel reconstruction. Some diagnostics are under active commissioning among the training shots and are missing in almost half of the shots, yet they are quite important for disruption prediction. Therefore, a cross channel reconstruction model is developed, to fill in the missing diagnostics according to other existing related diagnostics.

If all of the optimizations are removed, an ordinary deep learning model can only get an AUC of 0.741 in HL-3, proving the complexity of the data environment in this new tokamak. However, with the help of all the five optimizations, an AUC of 0.940 has been achieved, which demonstrates the feasibility of deep learning-based algorithms to accurately predict the disruptions in future fusion devices.

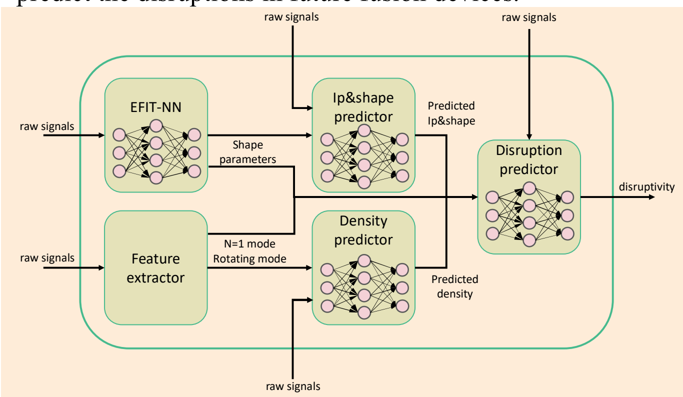

Fig.1. The overarching paradigm of disruption prediction algorithm in HL-3, encompassing 5 modules.

References

[1] T.C. Hender, J.C. Wesley, J. Bialek, et al. 2007 Progress in the ITER physics basis: chapter 3: MHD stability, operational limits and disruptions Nuclear Fusion 47: S128–202.

[2] P.C. de Vries, G. Pautasso, D.A. Humphreys et al. et al. Fusion Science and Technology 69:471-484.

[3] J. Vega, A. Murari, S. Dormido-Canto, et al. 2022 Nature Physics 18: 741–750

[4] Z.Y. Yang, F. Xia, X.M. Song, et al. 2020 Nuclear Fusion 60: 016017.