AUGMENTING THE EXTRAPOLATION CAPABILITY OF DISRUPTION PREDICTION TO EXTENDED PARAMETER REGIMES BY PREDICT-FIRST NEURAL NETWORK

[1]Z. Y. Yang, [1*]W. L. Zhong, [1]J. Y. Li, [1]Y. H. Chen, [1]D. Li, [1]Z. J. Zhang, [1&2]G. H. Zheng [1]X. Fan, [1]B. Li, [1]Y. P. Zhang, [1]Y. B. Dong, [1]T. F. Sun, [1]Z. H. Xu, [1]R. S. Qiu [1]X. Sun and [1]J. R. Wen

[1]*Southwestern Institute of Physics, Chengdu 610041, China*
[2]*School of Physics, Nankai University, Tianjin 300071, China*
Email: zhongwl@swip.ac.cn

1. INTRODUCTION

Disruption prediction is a crucial topic, especially for large-scale tokamaks, given the catastrophic consequences of plasma instabilities on reactor components. Recent progresses have proved that deep neural network can accurately predict the coming disruptions by learning from historical data, which becomes a potential solution for the disruption prediction in future devices [1]. This methodology has also been proved in HL-2A and HL-3 tokamak, both in offline testing and online experiment [2, 3]. But for future devices like ITER, these devices will initially operate with lower plasma current, stored energy and aspect pressure and progressively accessing higher parameter spaces, creating a critical need for models with robust extrapolation capabilities.[4].
The HL-3 tokamak provides an ideal testbed for this investigation. The device starts to operate with 100kA, 100ms plasmas in 2020, and has accessed the plasma current of 1.6MA, the stored energy of more than 400kJ, and the aspect ratio higher than 3 in recent years. In this research, the extrapolation capability of deep learning-based disruption prediction algorithm is evaluated by data from HL-3. It has been observed that the performance of conventional algorithm degrades with the promotion of key plasma parameters, as expected. To address this limitation, predict-first neural network (PFNN) is applied, which fundamentally reorients the disruption prediction approach from direct parameter regression to plasma evolution pattern recognition.

2. PREDICT-FIRST NEURAL NETWORK

The PFNN architecture operates through a dual-stage prediction process. In the first stage, a series of temporal convolutional networks generate time-resolved predictions of key plasma parameters, including plasma current, density, shape, and stored energy. These predictions are based on the current plasma state, control targets, and actuator inputs. The second stage of the network then compares these predicted trajectories with the actual experimental measurements, quantifying the deviations and identifying anomalous plasma behaviour that may precede disruptions. This two-stage approach provides several critical advantages over conventional methods. First, by focusing on the evolution of plasma parameters rather than their absolute values, the model becomes less sensitive to shifts in parameter distributions. Second, the prediction-first paradigm inherently incorporates fundamental plasma physics principles, including force balance, energy balance, and particle balance, through the design of its prediction tasks. This physics-informed approach enhances the model's ability to generalize to new parameter regimes.

3. THE EXTRAPOLATION CAPABILITY

To evaluate the extrapolation capabilities of the PFNN, we conducted extensive testing on a dataset comprising 479 HL-3 discharges, including 214 disruptive shots and 265 non-disruptive shots. The dataset spans a wide range of plasma parameters, with plasma currents ranging from 0.1 to 1.6 MA, stored energies from 40 to 420 kJ, and aspect ratios from 0.3 to 3.2. The performance of the PFNN was compared against a baseline TCN model [5] trained on the same dataset. The results demonstrate a clear advantage for the PFNN approach. While the baseline model exhibits a progressive decline in prediction accuracy as the plasma parameters extend beyond the training range, the PFNN maintains consistent performance across all parameter regimes. This stability is particularly evident in high-performance conditions, where the baseline model's accuracy drops by around 10%, while the PFNN shows only minimal variation in performance.
The implications of this research extend beyond immediate performance improvements. The success of the PFNN architecture highlights the importance of incorporating physical principles into machine learning models for fusion applications. As we move toward reactors like ITER and beyond, where operational experience will be limited and parameter ranges will be significantly extended, such physics-informed approaches may prove

essential for developing reliable prediction and control systems. Future work will focus on further refining the PFNN architecture and exploring its application to other challenging scenarios in tokamak operation, such as the prediction of specific disruption precursors and the integration with real-time control systems.

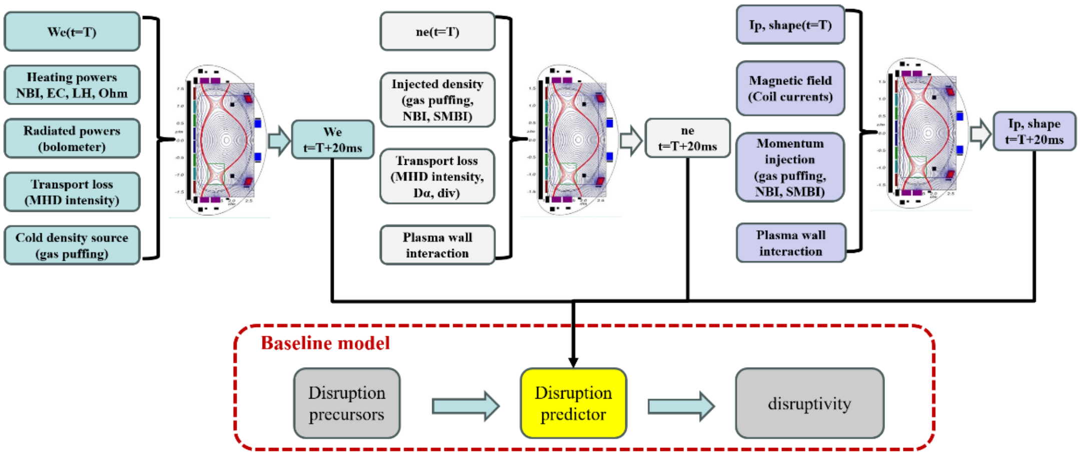

Figure 1 The structure of predict-first neural network. Three TCNs are developed to predict the evolution of plasma stored energy (W_e), density(n_e) and current (I_p) & shape. The predicted trajectories are incorporated into a conventional baseline model to consist a PFNN algorithm.

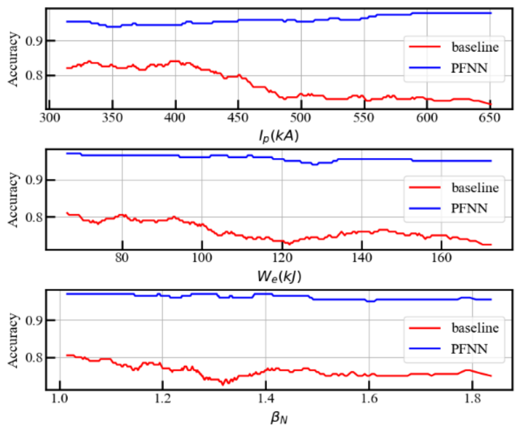

Figure 2 The accuracy of baseline algorithm and PFNN on shots with different ranges of plasma current (I_p), stored energy (W_e) and aspect pressure (β_N). The curve has been smoothed with a window of 100 shots and x-axis represents the mean value of the 100 maximum parameters for the shots. Therefore the range of x-axis is not totally same as described in text.

ACKNOWLEDGEMENTS

This work is supported by National MCF R&D program of China under Grant No.2019YFE03010003, National Natural Science Foundation of China under Grant No. U21A20440 and Sichuan Province Innovative Talent Funding Project for Postdoctoral fellows under Grant No. BX202222. The authors wish to thank all the members at Southwestern Institute of Physics for doing their best to co-operate during the collection of dataset and development of algorithm.

REFERENCES

[1] J. Vega, A. Murari, S. Dormido-Canto, et al. Disruption prediction with artificial intelligence techniques in tokamak plasmas. 2022 Nat. Phys. 18 741–50

[2] Zongyu Yang, Fan Xia, et al. Recent progress on deep learning-based disruption prediction algorithm in HL-2A tokamak. Chinese Physics B. 2023, 32(7) 075202

[3] Zongyu Yang, Fan Xia et al. Implementing deep learning-based disruption prediction in a drifting data environment of new tokamak: HL-3. Nuclear Fusion. 2025, 65(2) 026030

[4] P. C. de Vries, G. Pautasso, D. Humphreys, et al. Requirements for Triggering the ITER Disruption Mitigation System.

[5] J. Gehring, M. Auli, D.Grangier, D. Yarats and Y. N. Dauphin. Convolutional Sequence to Sequence Learning Proc. 34th Int. Conf. on Machine Learning (Sydney, Australia, 6–11 August, 2017) (PMLR) p 70 (available at: https://icml.cc/Conferences/2017/AcceptedPapers)