An interpretable, transferable and real-time disruption predictor in HL-2A based on deep learning

Z. Y. Yang[1&2], F. Xia[1], X. M. Song[1], Z. Gao[2], Y. X. Li[1], Y. B. Dong[1], Y. P. Zhang[1], W. Zheng[3], S. Wang[1], B. Li[1], X. Q. Ji[1], Z. Y. Chen[3], X. B. Zhu[1] and Y. H. Liu

[1] *Southwestern Institute of Physics, Chengdu, China*

[2] *Tsinghua University, Beijing, China*

[3] *Huazhong University of Science and Technology, Wuhan, China*

Email address of the main author: zy-yang@swip.ac.cn

Abstract

For future tokamak's disruption predictor, many quanlities are needed besides the high accuracy, including interpretability, real-time capacity and transferability and so on. In this research, two breakthroughs are made on the top of the interpretable disruption predictor in HL-2A, focusing on its real-time capacity and transferability. For the real-time capacity, the algorithm is accelerated to deal with an input slice within 0.3ms by some adjustments on the algorithm and the TFLite framework. It is implemented into HL-2A's plasma control system(PCS) and gets an accuracy of 89.0% during online test. Some demo shots are also got where the algorithm predicted the impending disruptions and triggered the SMBI or MGI to mitigate them. For the transferability, a preliminary disruption predictor is successfully developed in HL-2M, a newly built tokamak in China. Although only 31 and 23 shots are used as the training and validation set, respectively, it still gives reasonable outputs on testing set with the help of data from HL-2A and J-TEXT. In general, HL-2A's disruption predictor has proved that deep learning has enough flexibility to meet all kinds of demands along with a high accuracy and is a good potential choice for future tokamak's disruption prediction.

1. INTRODUCTION

Disruptions in large-scale future tokamaks are of concern due to their potential harmful effects on the devices, specifically, electromagnetic loading, thermal loading and runaway electrons[1]. To eliminate the risks, disruption prevention and mitigation techniques are needed[2]. And disruption predictor takes the responsibility to trigger these techniques on appropriate time[3].

In the past decades, lots of machine learning-based disruption predictors are developed in different tokamaks[4]. New algorithms, feature extraction methods and large datasets are introduced to get better performance on accuracies, prediction advance time and some other aspects. The feasibility of utilizing machine learning-based disruption predictor to handle the problem of disruption has been basically validated.

However, as part of the control system of a large-scale scientific construction project like ITER, the disruption predictor calls for all-around excellent performance. For example, high accuracy, long prediction advance time and real-time capacity are required to satisfy the primary need of triggering the disruption mitigation system[5, 6, 7]. Good interpretability is also important to ensure that the algorithm will be reliable and easy to debug in application[8, 9]. The capacity of cross-tokamak disruption prediction, or the capacity to be well-trained with limited data is essential as well, since the algorithm is going to be implemented on a newly-built tokamak[10, 11]. There is still a lot of work to do in the field of disruption prediction.

In this research, a series of further investigations are implemented on top of the first version of HL-2A's disruption predictor. The aim is to evaluate its potential to perform well in more aspects besides accuracy. And the result shows that by fully utilizing the flexibility of deep learning paradigm, the algorithm can perform well on all the four aspects as follows, accuracy, interpretability, real-time capacity and transferability.

The rest of this paper consists of 4 parts. Section 2 will briefly introduce the first version of HL-2A's disruption predictor and the method to interpret its output, which is already described in previous researches[12, 13]. Section 3 will introduce the process of implementing the algorithm into HL-2A's PCS and give the result of online testing. Section 4 will try to train a transferable disruption predictor on a multi-device dataset. A preliminary disruption predictor is successfully developed on HL-2M, a newly-built tokamak with very limited data. Section 5 is a brief summary.

2. HL-2A'S DISRUPTION PREDICTOR AND ITS INTERPRETABILITY

The first version of HL-2A's disruption predictor is proposed in [12]. It has a true positive rate(TPR) of 92.2% and a true negative rate(TNR) of 97.5% on the testing set, which consists of 475 disruptive shots and 1271 non-disruptive shots. A novel 1.5-D CNN + LSTM structure is used in this algorithm and proves quite helpful for the accuracy.

As for the interpretability, a special node is found in the 1.5D CNN structure. Before this node, signals from different input channels are dealt with separately for some neural network layers to eliminate the difference of their temporal structures and statistical distributions. While after this node, they will be merged into an array and be mixed in subsequent layers. Therefore, on this node each input channel of algorithm can be evenly disturbed by a gaussian noise and the corresponding offset of the algorithm's output will indicate the importance of each input channel. Both the result of single shot visualization and statistical analysis on a disruption causes dataset show good coherence with the cause of disruption. More detailed descriptions of this method can be found in [13].

3. REAL-TIME CAPACITY

The first version of HL-2A's disruption predictor takes 17ms to analyse an input slice, exceeding the limitation set by working cycle of HL-2A's PCS. Therefore, a reduced version is proposed in [13], which takes 2ms instead. A further promotion is obtained by utilizing TFLite, an inference framework developed by Google[14]. TFLite accelerates the computation by customized optimization for a determined neural network structure and quantization techniques[15]. With the help of TFLite, the algorithm can deal with an input slice within 0.3ms.

To test the algorithm's performance in online environment, an integrated disruption prediction and mitigation system is established in HL-2A. Figure 1 gives the framework of this system. The data acquisition system (DAS) gathers all the needed diagnostic signals and sends them to the PCS. These data are first used to do the position and shape control by the corresponding module, in which process some secondary signals are produced. Then these secondary signals, together with the raw diagnostic signals, are sent to the disruption prediction module. In HL-2A, the PCS is mainly developed in C language while the disruption prediction algorithm is developed in Python, so a cross language interaction is required. Thus, there is a C-based disruption prediction module, which organizes the input data and calls the Python-based disruption prediction module. Finally, the prediction result is sent back to the C-based module to decide whether or not a trigger signal should be sent to the disruption mitigation system.

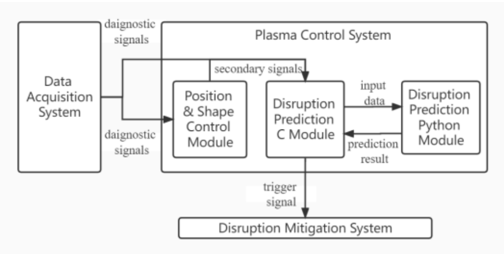

Figure 1 Design of the integrated system

Figure 2 Vertical displacement induced disruption mitigated by SMBI

The system is firstly validated in open-loop online testing, where the algorithm keeps running during every discharge in Shot Nos. 38650-39347 of HL-2A. 230 of 240 disruptive shots are correctly predicted and 32 false alarms are triggered in 142 non-disruptive shots. The corresponding TPR and TNR are 0.958 and 0.775, respectively. The accuracy is lower than in the offline testing, majorly due to two restrictions. For one thing, some input signals are not available in real-time environment due to some engineering problems. And for the other, the simplification on neural network structure to realize the real-time computation also brings degradations.

Closed-loop online testing is also tried in a few shots. Figure 2 gives a demo shot where a vertical displacement induced disruption is predicted and then mitigated by SMBI.

4. TRANSFERABILITY

Since the researches on disruption prediction aims to provide a reliable algorithm for future tokamaks, there is an important issue on how to develop the algorithm on a newly-built tokamak with very limited data available. In this section, HL-2M is selected as the testing platform, which is a newly-built tokamak with only 81 shots suitable for disruption prediction[16].

To solve the problem of poor training data, J-TEXT and HL-2A's data are introduced to provide auxiliary constraints. This mixed dataset makes it possible to train a reliable neural network on HL-2M. Table 1 and Table 2 give the detailed information about the train, validation and testing set. The network structure and training strategy are basically same as the version in [12]. During testing, the algorithm trained on this mixed dataset predicted 17 of 18 disruptive shots and triggered 2 false alarms in 9 non-disruptive shots. Figure 3 shows the output of algorithm during some example shots. The result confirms that data from exists devices are helpful to develop the algorithm on a new device.

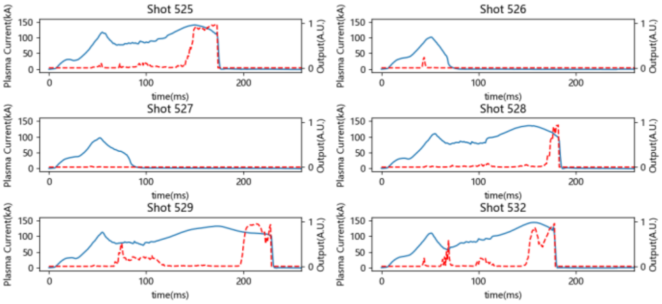

Figure 3 Output of HL-2M's disruption predictor

Table 1 Shot numbers for training, validation, and testing of the HL-2M's disruption predictor. D means disruptive and ND means non-disruptive

Device	Training		Validation		Testing	
	D	ND	D	ND	D	ND
HL-2A	1532	4317	364	470	-	-
J-TEXT	1477	4091	415	1221	-	-
HL-2M	25	6	17	6	18	9

Table 2 Detailed information about the input of the HL-2M's disruption predictor

Signal name	Sample rate(kHz)	Physical meanings	signal tag		
			HL-2M	J-TEXT	HL-2A
Ip	1	plasma current	IP_2M	ip	IP
Target_Ip	1	target plasma current	ccIP_2M	oh_set	ccIP
Bt	1	toroidal magnetic field	BT1_2M	bt	Bt
Density	1	density of electrons	AMW_INT1_02	LIN_DEN_CH06	Density1
Dh	1	horizontal displacement	dh_2M	dx	FluxDh
Bolometer	1	power of bolometer radiation	DBOL_10	AXUV_CA_08	BOLU10
Mir_Probe_A	10	a pair of symmetric poloidal Mirnov probes	Mpol_17_24	MA_POL_CA01T	Mpol_04
Mir_Probe_B	10		Mpol_17_51	MA_POL_CA19T	Mpol_13

5. SUMMARY

In this research, a series of updates are implemented on HL-2A's disruption predictor, bringing it interpretability, real-time capacity and transferability. Since the disruption predictor in large-scale scientific construction project like ITER calls for all-around excellent performance. Deep learning seems to be a good potential choice, which has enough flexibility to adapt to all kinds of demand during future implementation.

References

[1] ITER Physics Basis Editors et al. Nuclear Fusion, 1999, 39(12): 2137.
[2] M. Lehnen, et al. Journal of Nuclear Materials, 2015, 463: 39-48
[3] P. C. de Vries, et al. Fusion Science and Technology, 2016, 69(2): 471-484.
[4] Jesus Vega, et al. Nature Physics, 2022, https://doi.org/10.1038/s41567-022-01602-2
[5] A. Murari, et al. Nuclear fusion, 2020, 60(5): 056003.
[6] Ge Dong, et al. Journal of Machine Learning for Modeling and Computing, 2021, 2(1): 49-64.
[7] B. H. Guo, et al. Plasma Physics and Controlled Fusion, 2021, 63(11) 115007.
[8] C. Rea, et al. Nuclear Fusion, 2019, 59(9): 096016.
[9] William Tang, et al. 28th International Fusion Energy Conference Proceedings, 2021, ID: 1416.
[10] Julian Kates-Harbeck, et al. Nature, 2019, 568: 526-531.
[11] G. A. Ratta, et al. Fusion Science and Technology, 2018, 74(1-2): 13-22.
[12] Zongyu Yang, et al. Nuclear Fusion, 2020, 60(1): 016017.
[13] Zongyu Yang, et al. Nuclear Fusion, 2021, 61(12): 126042.
[14] Juhyun Lee, et al. 2019, arXiv.1907.01989.
[15] Benoit Jacob, et al. 2018, arXiv.1712.05877
[16] X. Song, et al. Nuclear Fusion, 2021, 61(8): 086010.