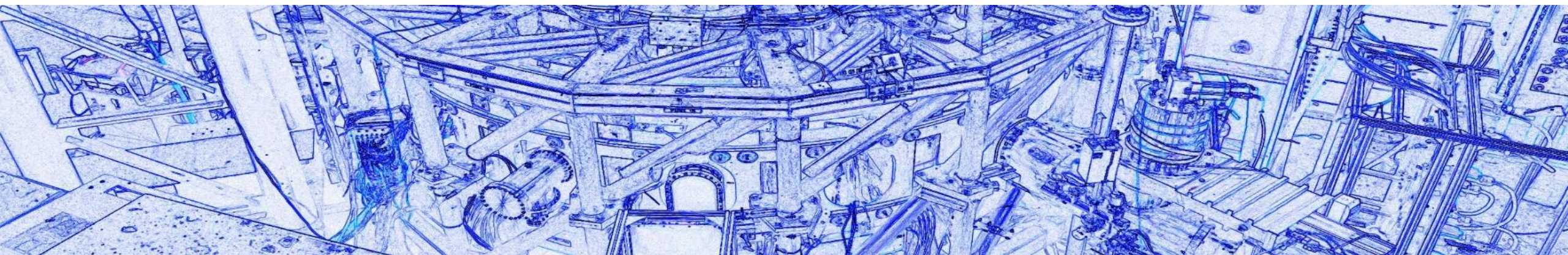


人工智能等离子体破裂预测

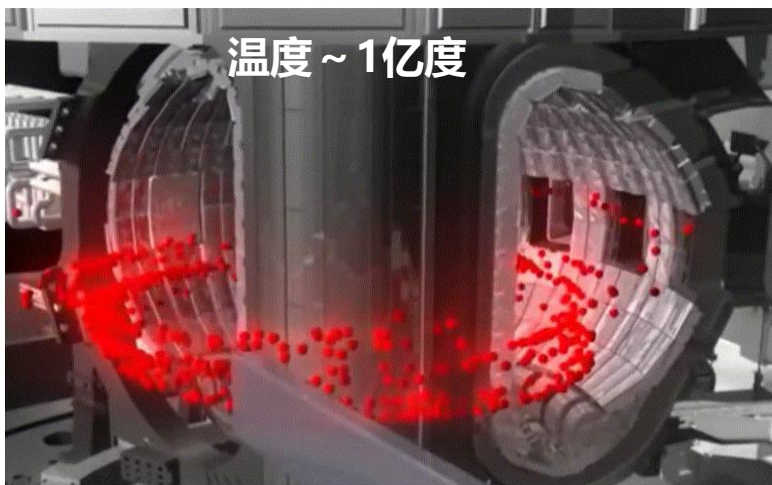
华中科技大学

郑玮



- 为什么要做破裂预测，为什么要用人工智能做破裂预测
- 破裂预测as监督学习问题
- 未来破裂预测面临的问题

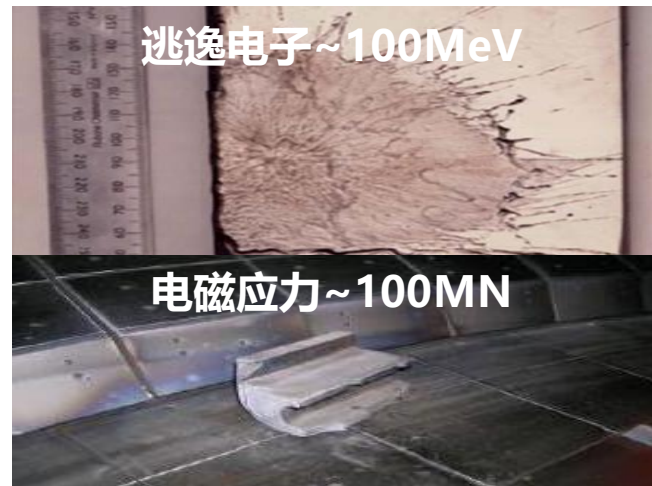
- **破裂**是等离子体失去约束、瞬间释放全部能量的现象
- 等离子体大破裂将可能导致未来聚变堆严重损毁



托卡马克等离子体破裂



ITER破裂能量约百公斤TNT当量

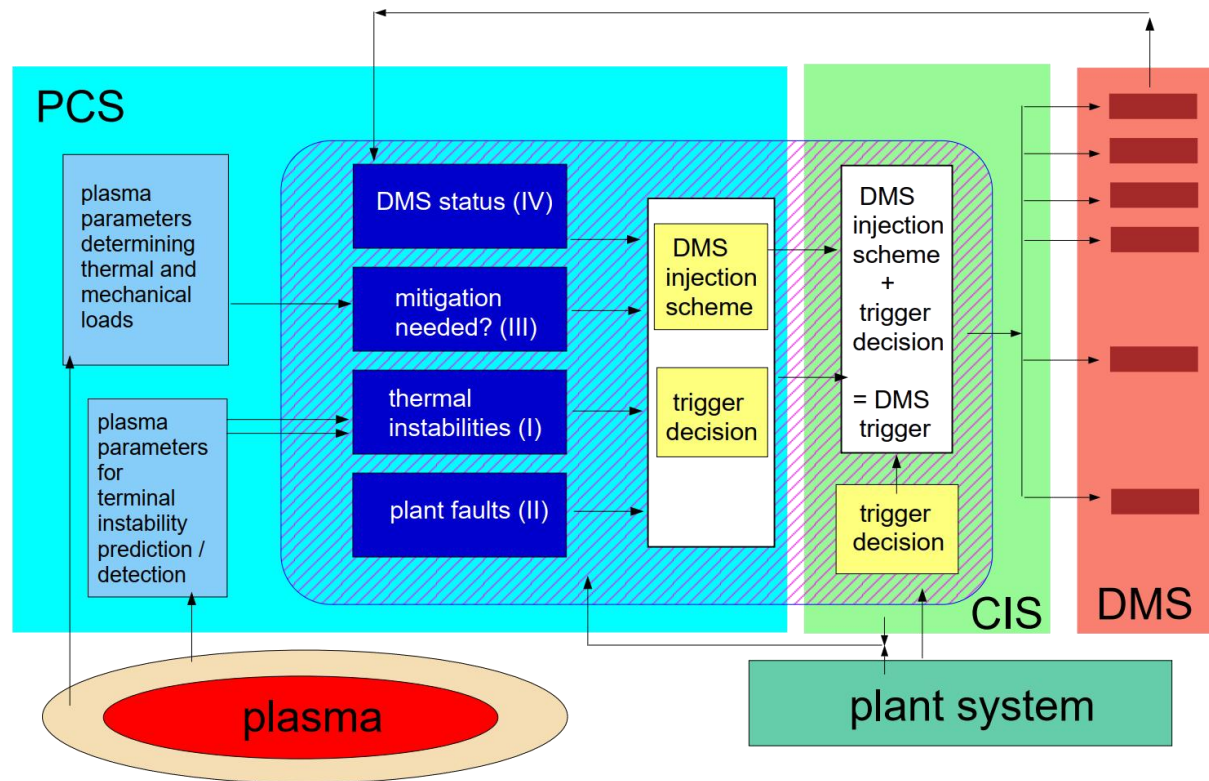
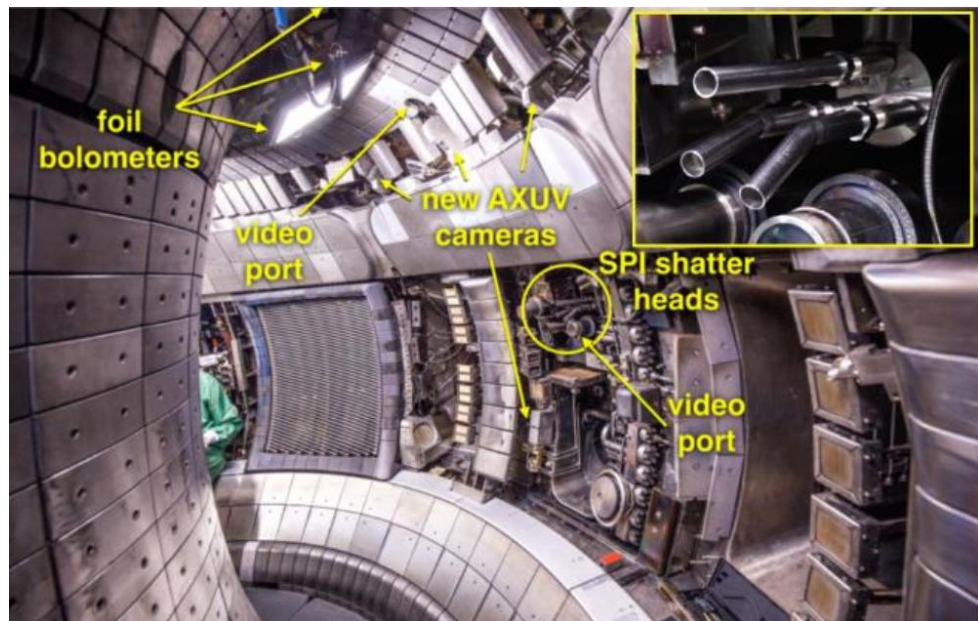


破裂造成聚变堆严重损毁

聚变堆运行需要安全气囊系统，缓解破裂带来的危害

为什么要做破裂预测

- DMS 是 ITER 的“安全气囊”，使用SPI系统缓解破裂的危害
- ITPA MDC-22以高准确率、低误报率的实时破裂预测触发为目标

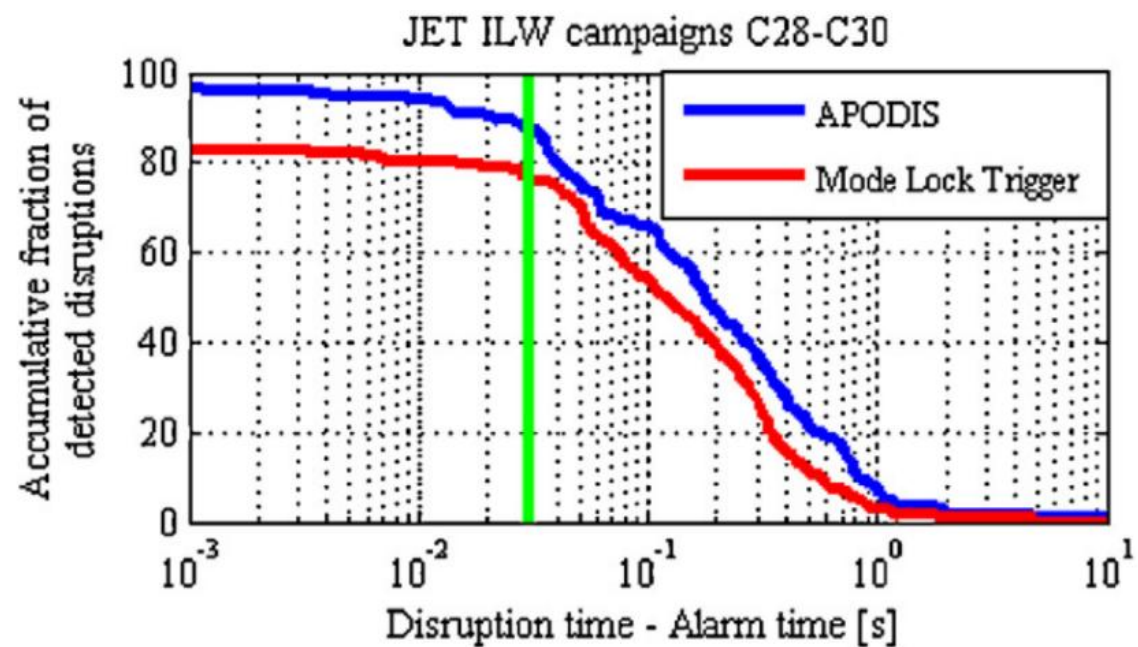
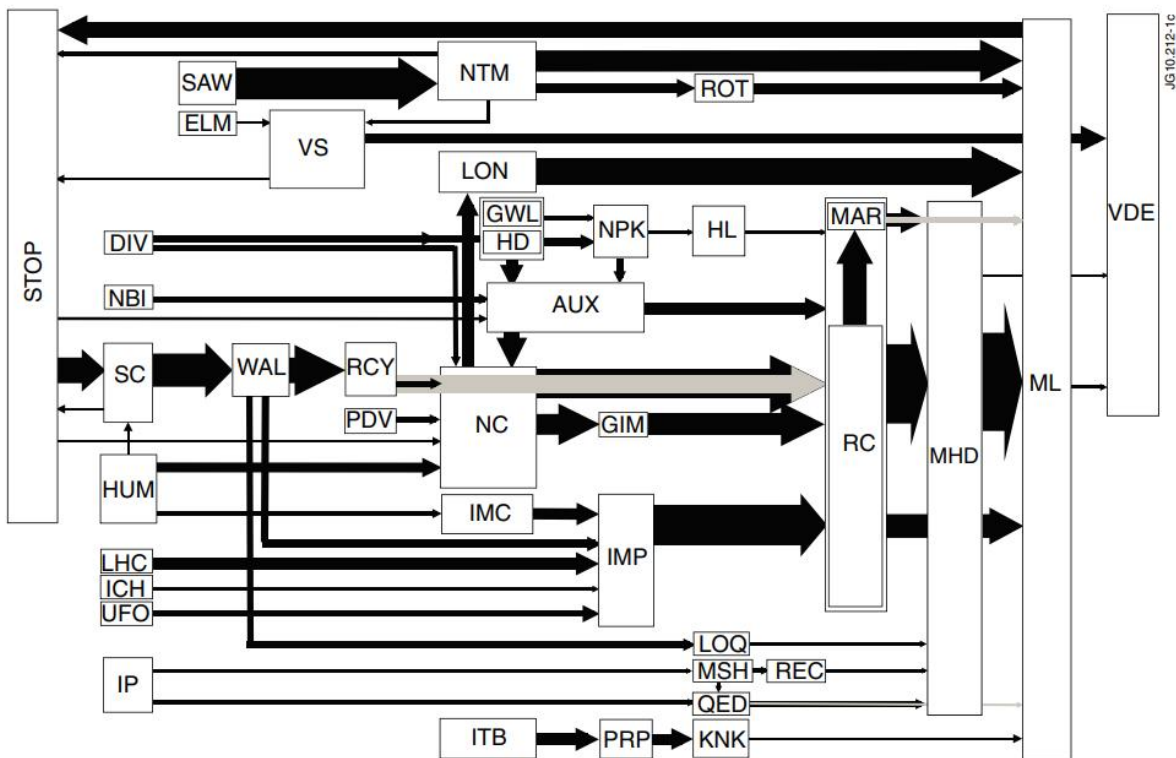


破裂预测是一切DMS的前提

The ITER disruption mitigation trigger: developing its preliminary design.

为什么要用人工智能做破裂预测

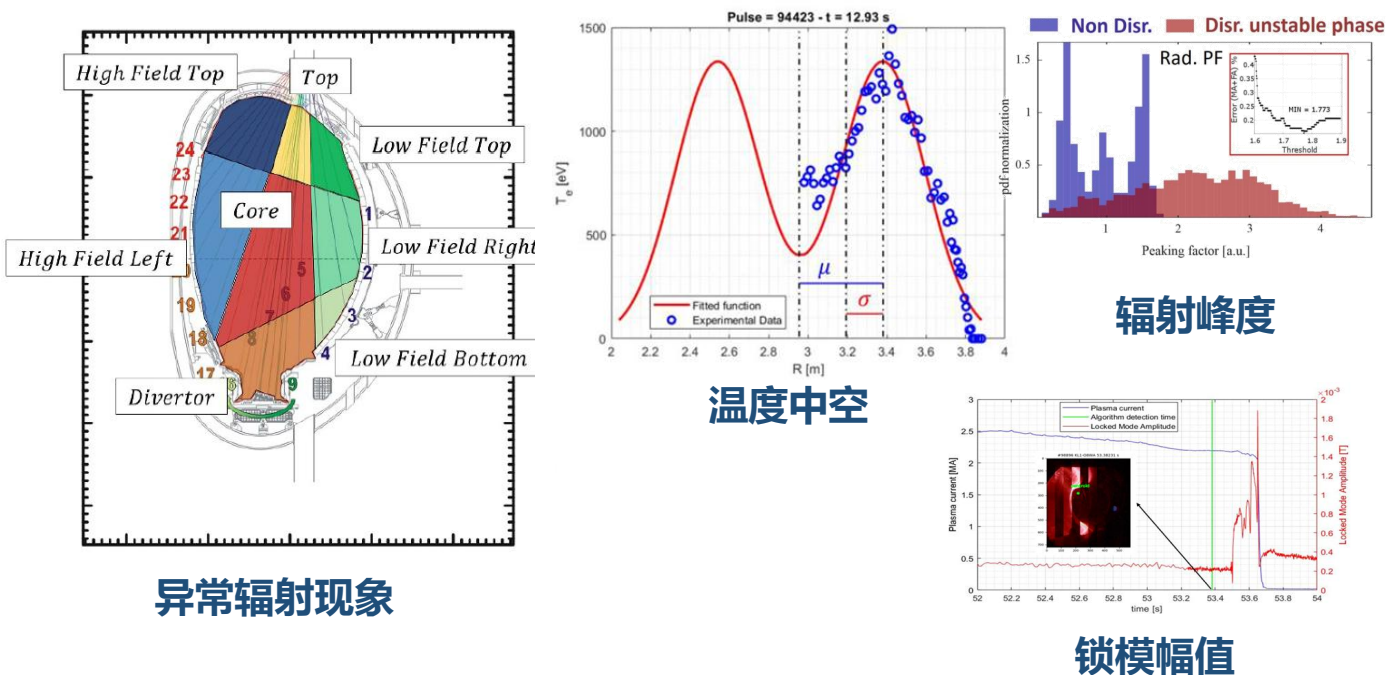
- 破裂的物理过程非常复杂，但是大部分破裂最终都会经历锁模这一阶段
- 通过检测锁模信号强度，超过阈值出发DMS， $B_r > LM_\theta$



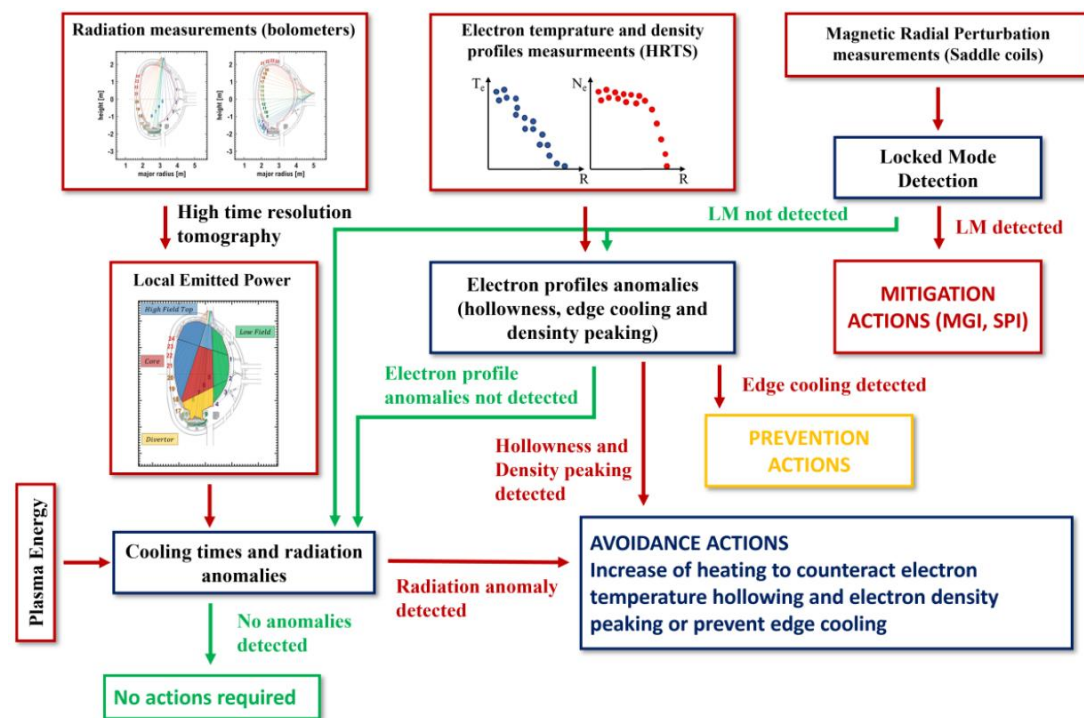
J. Vega, S., et al. Fusion Eng. Des., 2013

通过锁模阈值实现的破裂预测，准确率仍旧过低

根据实验数据与经验，手动编写现象特征识别算法，根据实验数据统计，设定不同现象出现的阈值



根据领域知识提取不同阶段前兆特征



Murari, A., et al, Nature Communication, 2023

- **破裂机理不明**
 - 难以通过第一性原理进行推导
- **破裂前兆难以描述**
 - 由于破裂机理不明，破裂前兆复杂（例如锁模，MARFE等），现有诊断难以完整描述完整的破裂前兆
- **托卡马克实验产生了大量数据**
 - J-TEXT, EAST, HL-2A, 均超过10万炮实验
- **最适合人工智能做的任务**
 - 人类还没有相关专长——破裂机理不明
 - 人类专常难以描述——破裂前兆难以描述
 - 从海量数据总找规律——大量实验数据

□ 监督学习

- 使用有标签的数据训练
- 回归
- 分类

□ 无监督学习

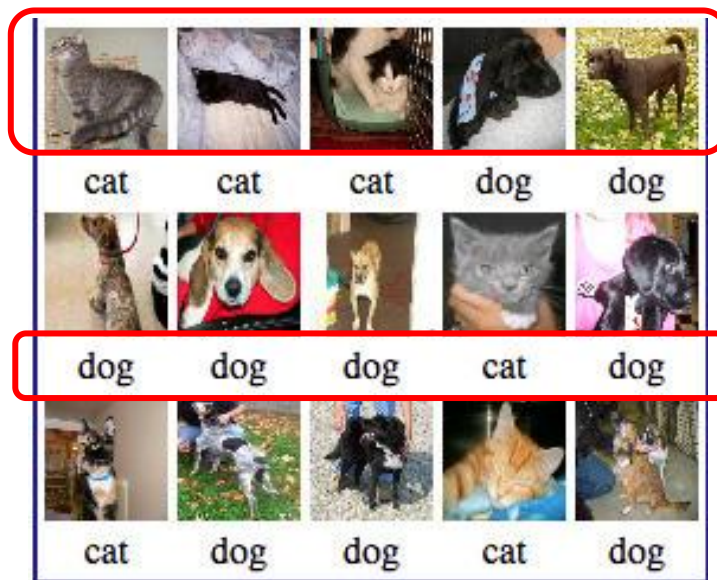
- 使用无标签的数据训练
- 聚类, 异常检测

□ 其他

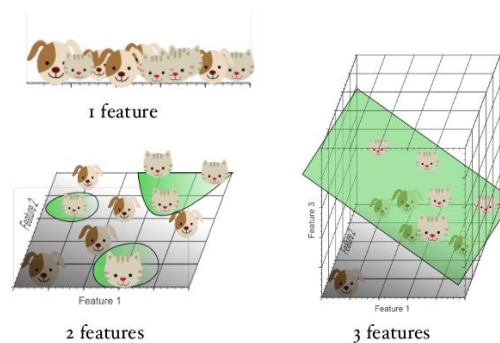
- 自监督学习、半监督学习、强化学习.....

特征X

标签: Y



监督学习: 分类



无监督学习: 聚类

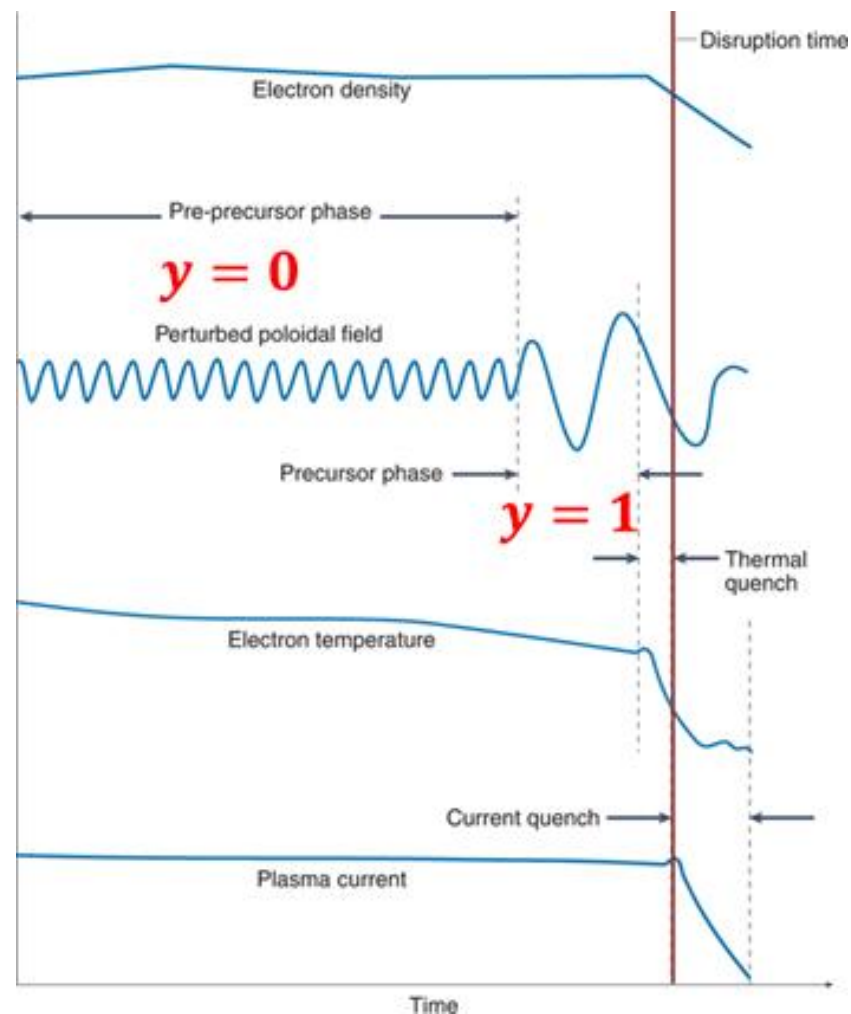
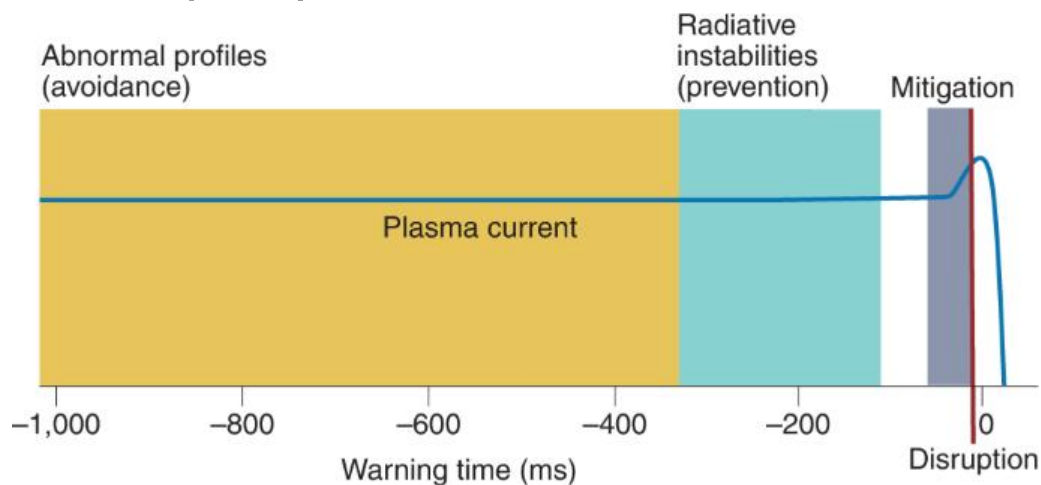
- 破裂预测实际是一个二分类问题， X （特征） $\rightarrow$ 诊断信号， Y （标签） $\rightarrow$ 这个信号是破裂破裂前兆的概率
- $Y=0$ 表示没有破裂风险， $Y=1$ 表示明显破裂前兆，即将破裂
- 破裂预测是找到一个函数 $h_{\theta}(x)$ 拟合 $y=1$ 的概率

$$P(Y = y|x) = \begin{cases} h_{\theta}(x) & : y = 1 \\ 1 - h_{\theta}(x) & : y = 0 \end{cases} \longrightarrow P(Y = y|x) = (h_{\theta}(x))^y (1 - h_{\theta}(x))^{1-y}$$

最大似然估计，一切人工智能的基础

$$\hat{\theta} = \operatorname{argmax}_{\theta} P(Y|X; \theta)$$

- 如何做人工智能破裂预测
- 准备X：选合适的诊断，如Mirnov coils, Bolometer, Thomson Scattering...；数据处理，特征提取（可选）
- 准备Y：选择一个破裂前兆提前时间，之后的标记为1，之前的标记为0
- 选择、设计算法：SVM，随机森林，各种神经网络， $\text{Model.Train}(Y, X)$ ，以上。

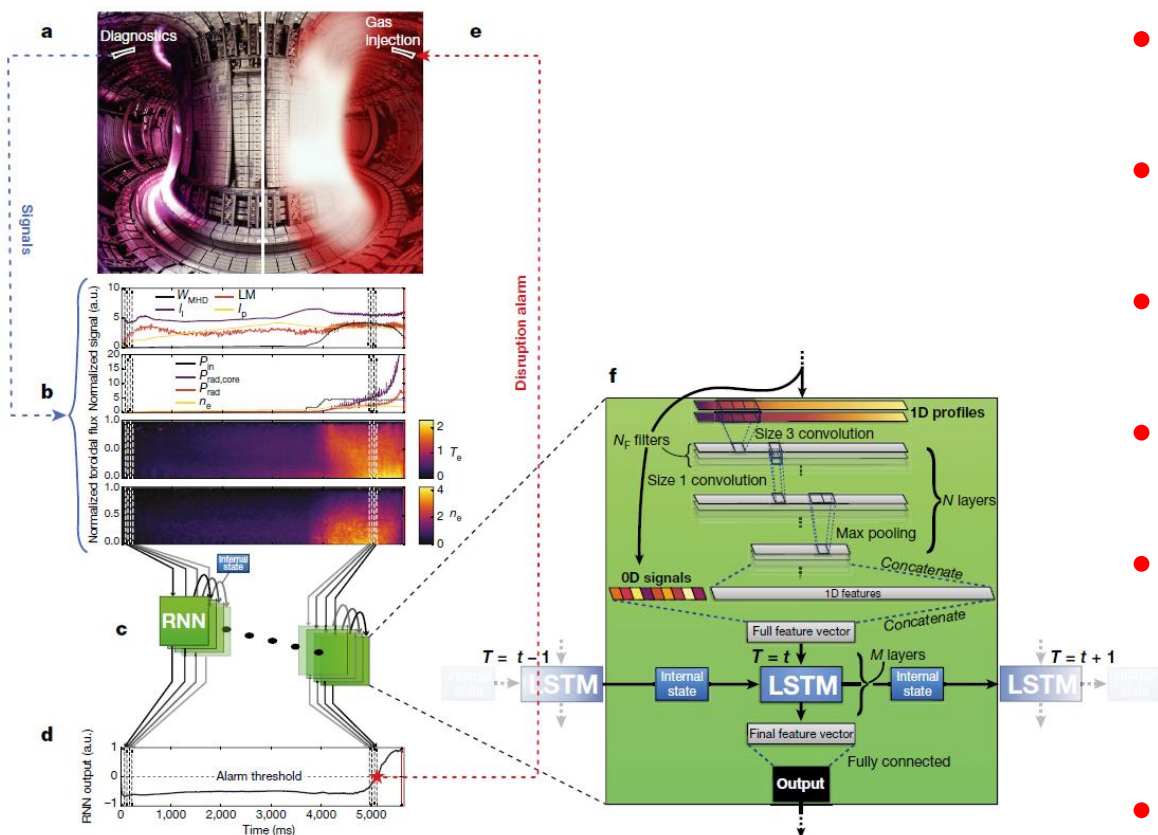


- 锁模信号 x 大于阈值出发破裂预警 LM_θ
- 从大量数据中找到锁模信号的阈值 LM_θ

$$P(Y = y|x) = \begin{cases} h_\theta(x) & : y = 1 \\ 1 - h_\theta(x) & : y = 0 \end{cases} \quad h_\theta(x) = \begin{cases} 1 & : x \geq LM_\theta \\ 0 & : x < LM_\theta \end{cases}$$

$$L\hat{M}_\theta = \operatorname{argmax}_{LM_\theta} P(Y|x; LM_\theta)$$

• 典型的深度学习破裂预测：FRNN

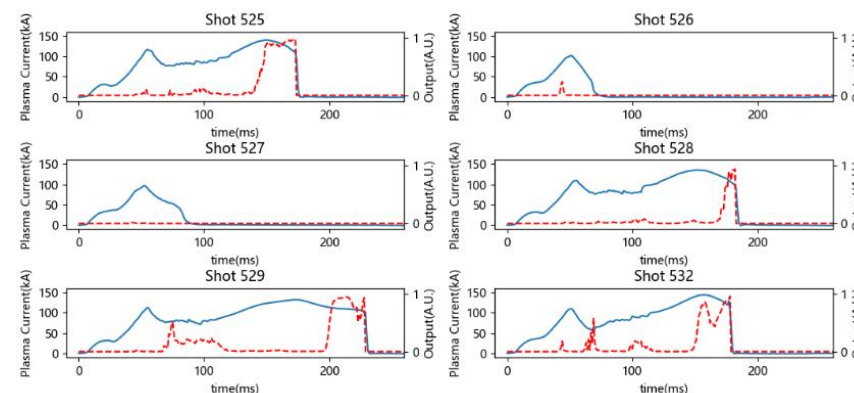
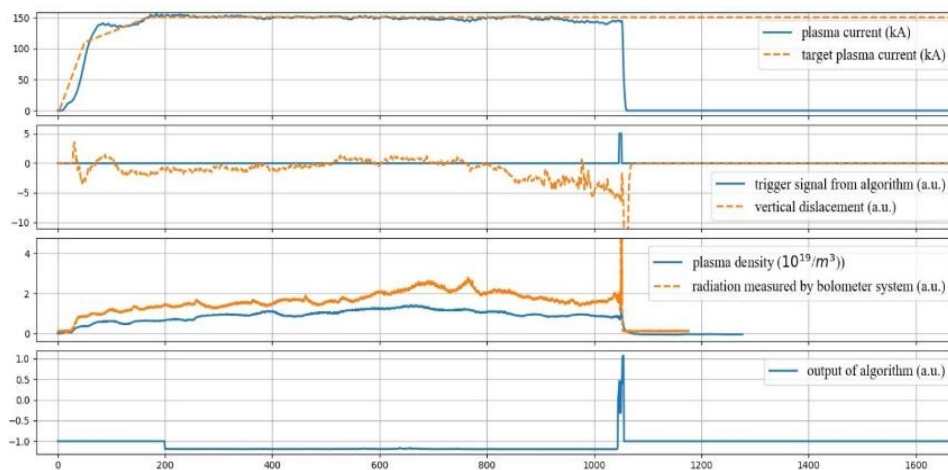
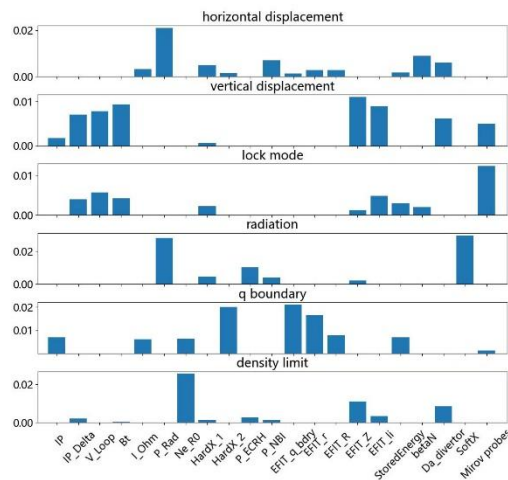


- FRNN: Fusion Recurrent Neural Network
- 输入：1D剖面+0D信号（平衡量&不稳定性）
- 输出：二分类（破裂 vs. 非破裂）
- 用CNN提取1D剖面的空间局部特征
- 1D局部特征与0D信号拼接后送入LSTM提取长时间尺度演化特征
- 输入步长/推理周期：1ms

HL-2A/HL-3上的破裂预测



- 基本框架：1.5D CNN + LSTM
- 在HL-2A上实现提前30ms, 96.1%的TPR
- 利用J-TEXT和HL-2A数据做领域自适应，使用44炮HL-3数据实现HL-3初步保护

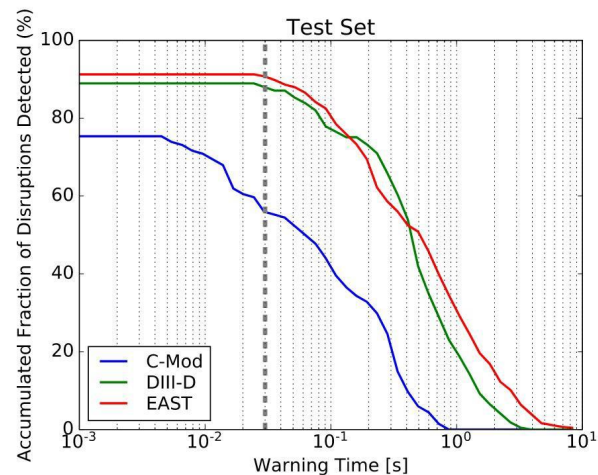
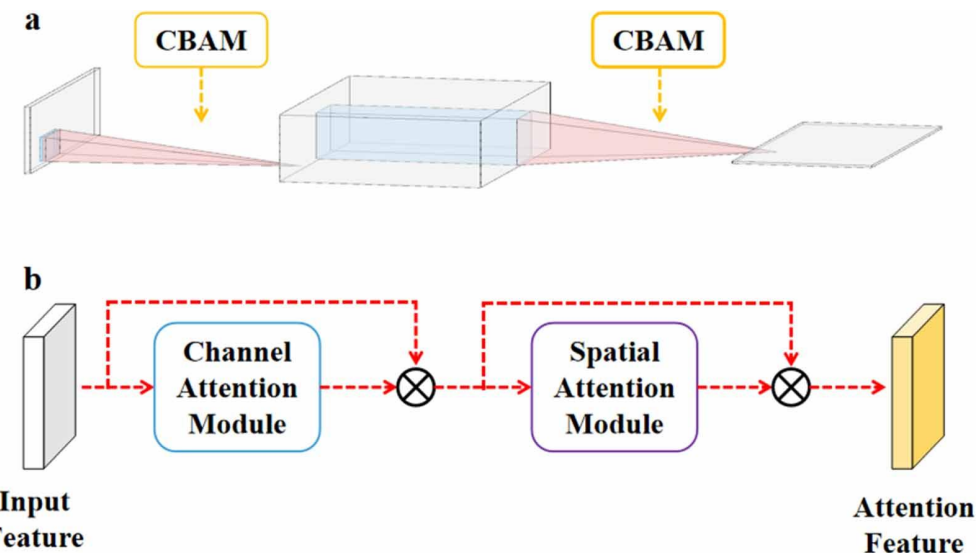


可解释性: HL-2A上分析6种破裂诱因的输入信号与破裂间的统计相关度

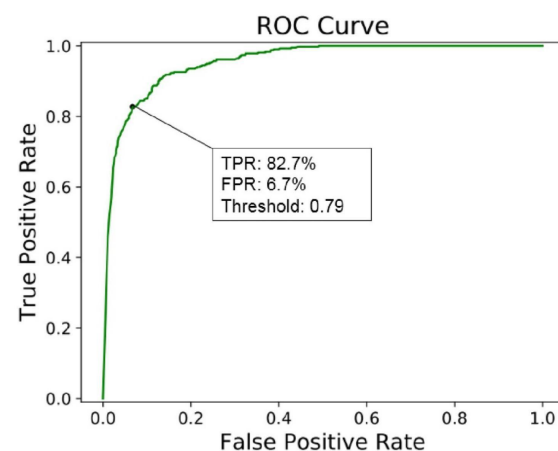
实时化: 闭环测试中成功预测并缓解垂直位移和芯部冷却造成的破裂

跨装置外推: 仅用44炮HL-3数据训练得到的破裂预测效果

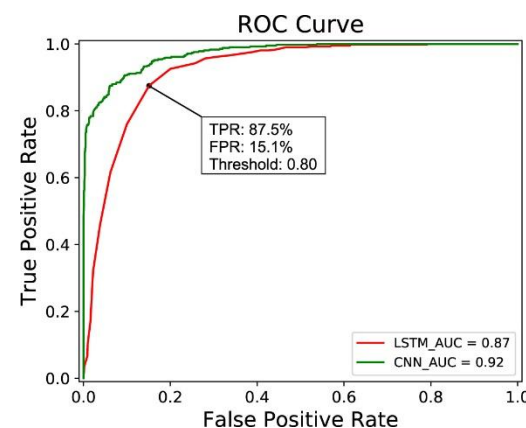
- EAST上的破裂预测经历了Random Forest, CNN, LSTM, CNN+LSTM混合几个阶段
- 预警准确率逐步从75%提升到95%
- 通过引入注意力机制和领域自适应方法, 非全金属壁->全金属壁准确率提升至89%



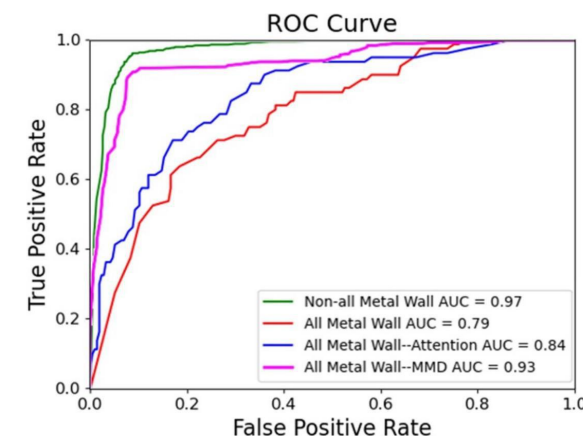
Random Forest



CNN



LSTM



CNN+LSTM+Attention+ $\mathcal{L}_{MMD}$

- 未来装置没有数据

- 机器学习是数据驱动的
- 未来聚变堆无法在损坏自身前积累足够多的数据

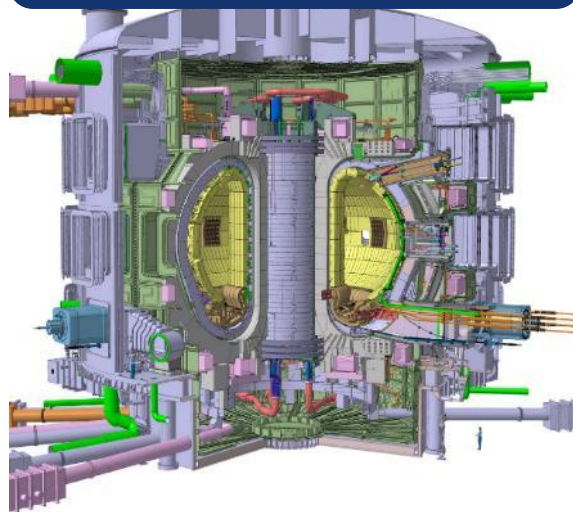
破裂是日常情况 😊



破裂危害在
可接受范围内



一炮未缓解的高参数
破裂将损坏装置 😬



训练装置: J-TEXT



测试装置: J-TEXT 测试装置: EAST



AUC: 0.98

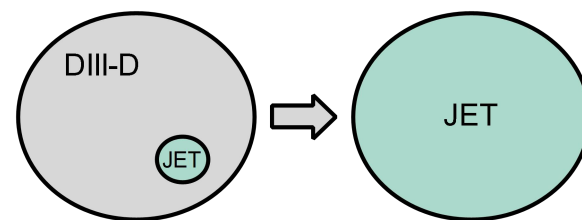
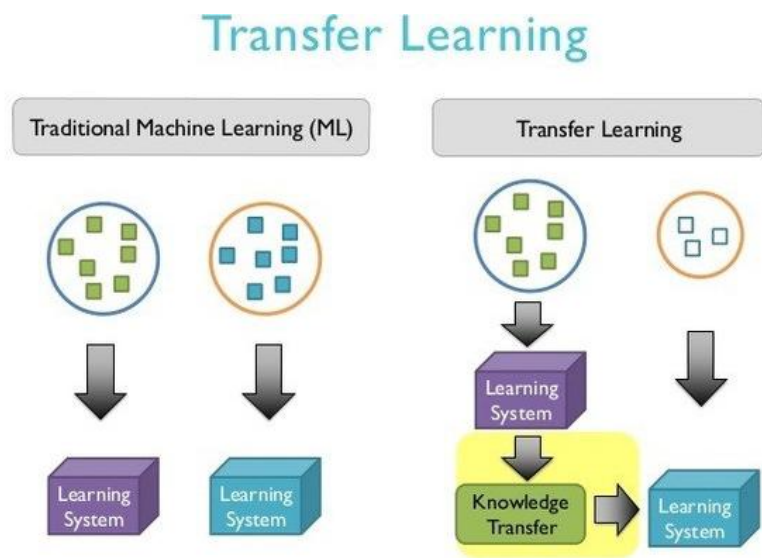


AUC: 0.64

- 如何利用已有装置数据建立对未来聚变堆有效的破裂预测系统？——迁移学习

- 训练集混合目标装置数据的“一瞥”

- 训练集主要由现有装置数据构建
- 从目标装置中混合了一点点数据
- 在深度学习的作用下，“奇迹”发生了



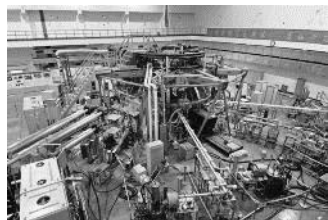
“一瞥” = 5炮 破裂炮
+
16炮 非破裂炮

	Cross-machine	Cross-Machine with “glimpse”
Training Set	DIII-D	DIII-D + δ
Testing Set	JET (ILW late)	JET (ILW late) - δ
Best classical Model	0.483	0.899
FRNN 0D	0.793	0.907
FRNN 1D	0.824	0.922

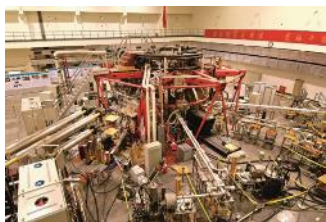
J. Kates-Harbeck, et al., Nature, 2019.

• 简单的混合就够了吗？

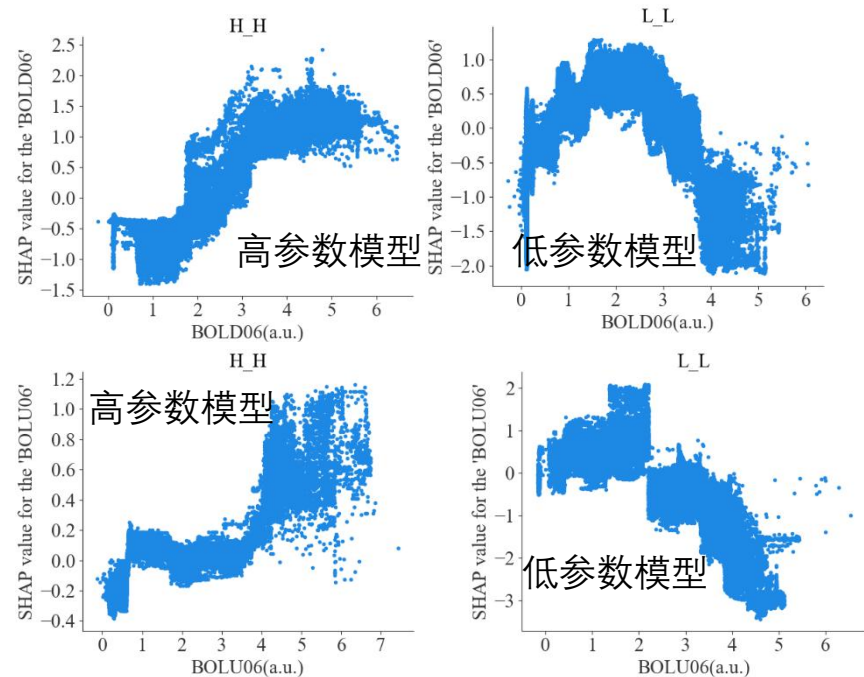
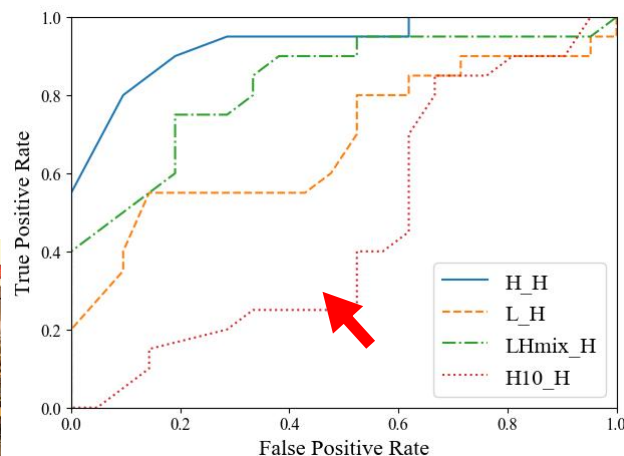
- 训练集：大量低参数放电+10炮高参数放电
- 测试集：只有高参数放电



低参数放电



高参数放电



HL-2A上的大量低参数数据中混入小部分高参数数据（10炮）显著提升了模型在高参数区间的预测性能
但不同参数区间的特征的贡献相互矛盾的情况
混合数据需要尽可能的小心，以减少负面效果

$$\varepsilon_t(h) \leq \varepsilon_s(h) + d_{\tilde{\mathcal{H}}}(\mathcal{D}_s, \mathcal{D}_t) + \min\{\mathbb{E}_{\mathcal{D}_s}[l_d(f_s, f_t)], \mathbb{E}_{\mathcal{D}_t}[l_d(f_s, f_t)]\}$$

J. Wang, et al. arXiv:2103.03097, 01 3 2021

减小源域误差

减小源域与目标域的差异

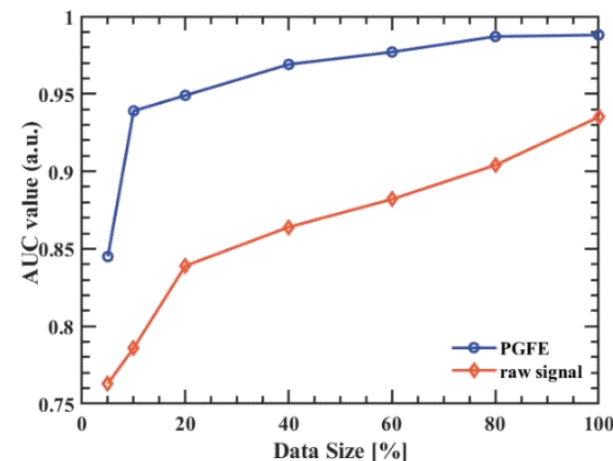
- 要提升未来聚变堆的破裂预测模型性能，要提升现有装置上的模型性能，并减少现有装置与未来装置之间的差异

- **首先进行特征提取:**

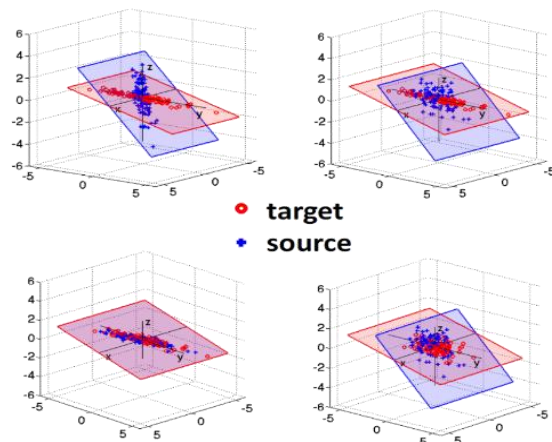
- 提取装置无关的物理特征
- 尽可能减小现有装置和目标装置之间的差异

- **其次采用领域自适应算法—CORAL:**

- 利用目标装置上的少量数据，可以把原装置特征映射到目标装置的特征空间中，进一步减少两个装置区别



➤ 用更少的数据能够达到比使用原始诊断更好的效果

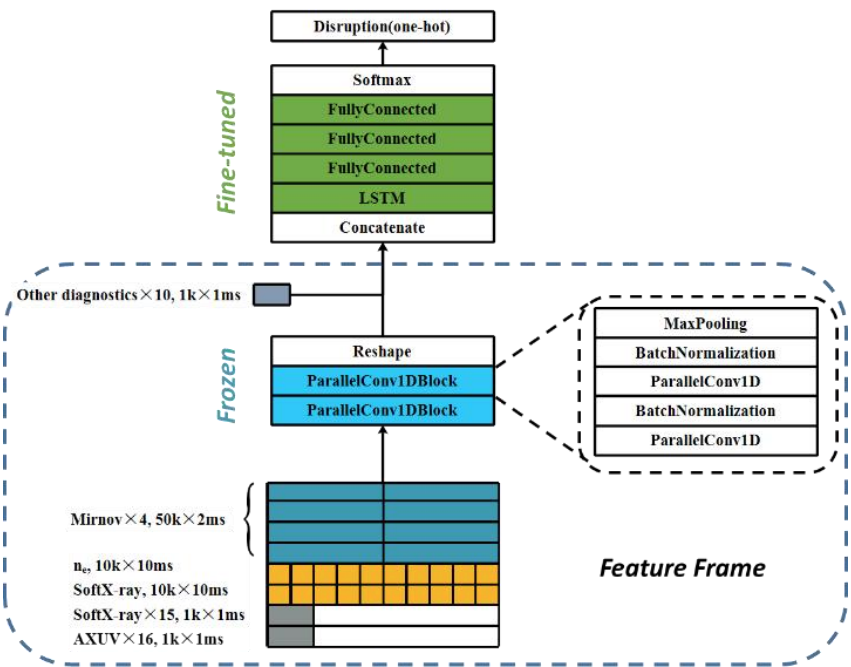


➤ 相比直接混合10炮数据性能有明显提升- 0.764->0.890

	EAST Data	Data Strategy	AUC
1	None	/	0.642
2	1896 (355)	All EAST	0.936
3	110 (10)	Mixing	0.764
4	110 (10)	CORAL	0.890

深度迁移学习：

- 神经网络底层，能够学习到低级的，共性的破裂前兆规律
- 可以冻结这些底层的网络，保留共性的破裂前兆规律
- 利用少量目标装置微调模型，适应目标装置特征

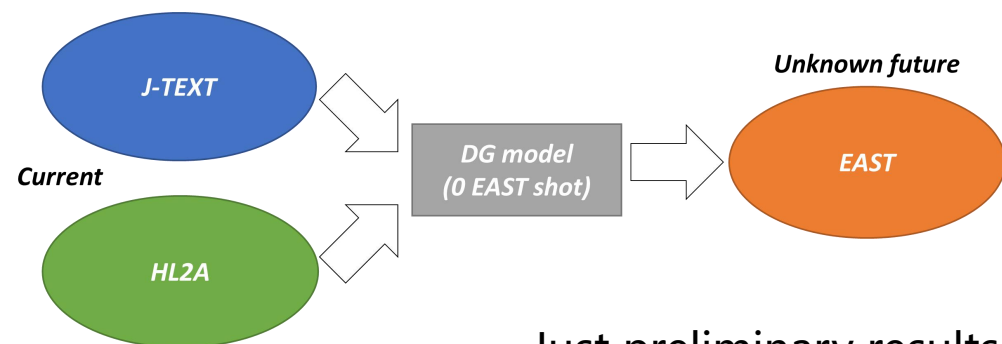
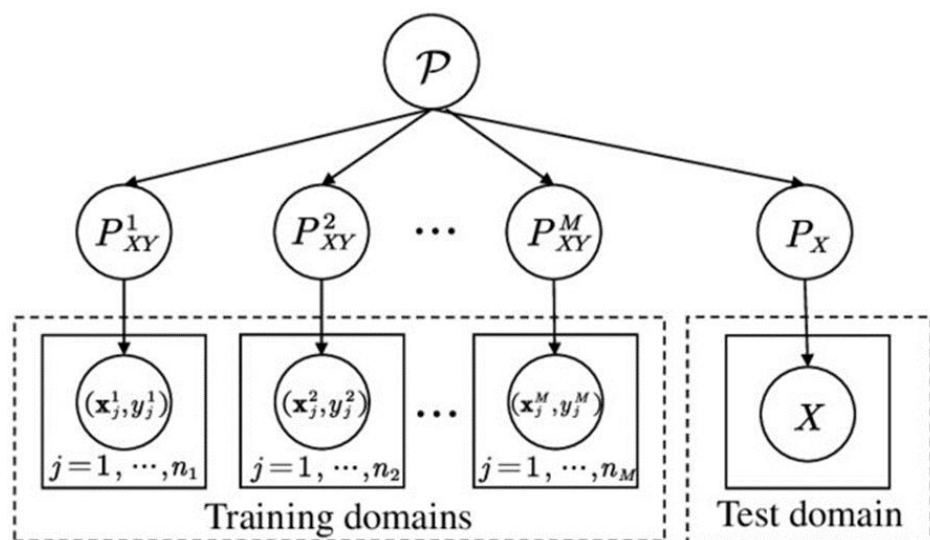


训练集		AUC
完整EAST	目标装置基准	0.821
J-TEXT+少量EAST	深度迁移	0.811
J-TEXT+少量EAST	仅混合	0.661
仅少量EAST	仅少量数据	0.615
仅J-TEXT	仅原装置数据	0.611

- 仅用少量EAST数据实现和完整EAST数据模型类似的性能

• 领域泛化破裂预测

- 假设:存在一个对于所有托卡马克都适用的破裂原因
- 如果装置更多, 是否可以探索出更加泛化的模型
- 模型需要实现两个任务, 一是需要识别出破裂及非破裂样本, 二是需要无法分辨出装置间的数据差异, 从而找到共性前兆



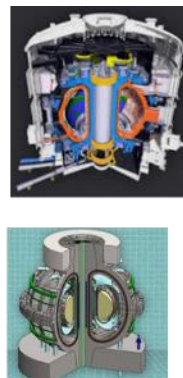
Just preliminary results,

Cases (0 shot)	TPR	FPR	AUC
DG1 Mixing	60.00%	11.39%	0.8009
DG3 DANN	90.00%	31.01%	0.8343

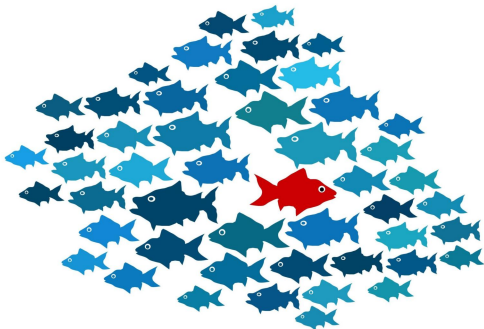
• 跨装置破裂数据库

- 与西物院，MIT，IAEA和ITU合作开展“人工智能加速聚变能研发”项目，
- 基于自主开发数据处理工具JDDB，开源了包含J-TEXT，HL-2A， Alcator C-Mod的公开数据库；
- 组织了ITU旗下的**首届国际破裂预测挑战赛**，吸引了众多非聚变社区的团队参赛

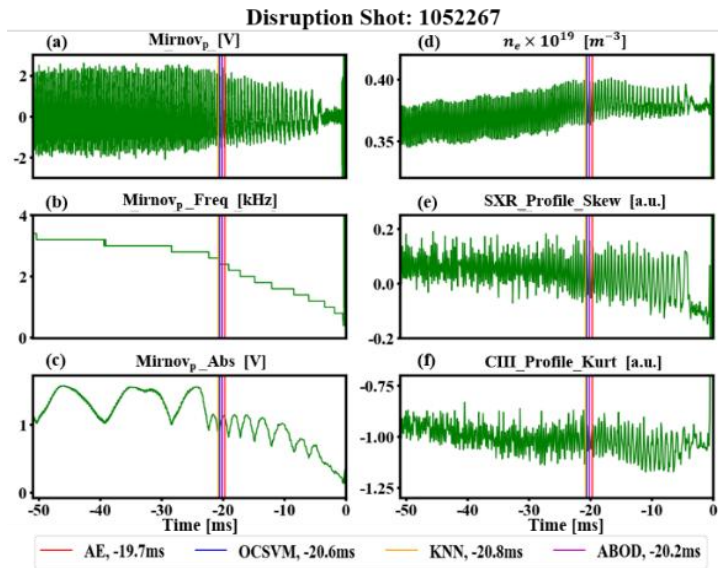
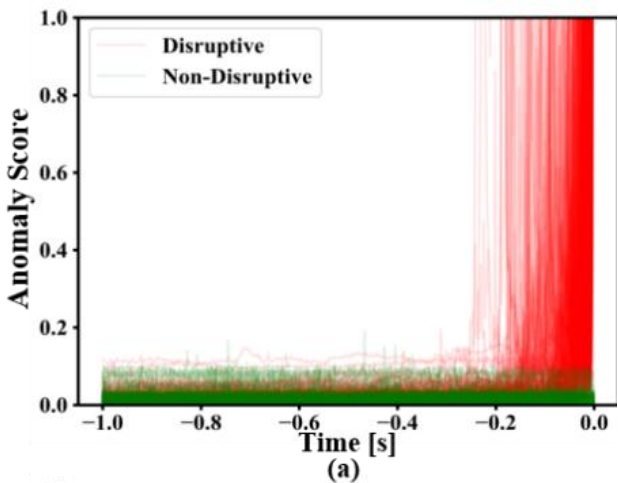
<https://zindi.africa/competitions/multi-machine-disruption-prediction-challenge>



- 未来聚变堆破裂危害大，无法提供破裂数据
- 那我们就仅用非破裂数据构建破裂预测模型——异常检测



- 但也不可能在不出现任何破裂的情况下获取大量非破裂数据
- **发现: 不同破裂炮前兆出现时间不同**
- **采用固定破裂前兆时间标注的方法并不是最优的**



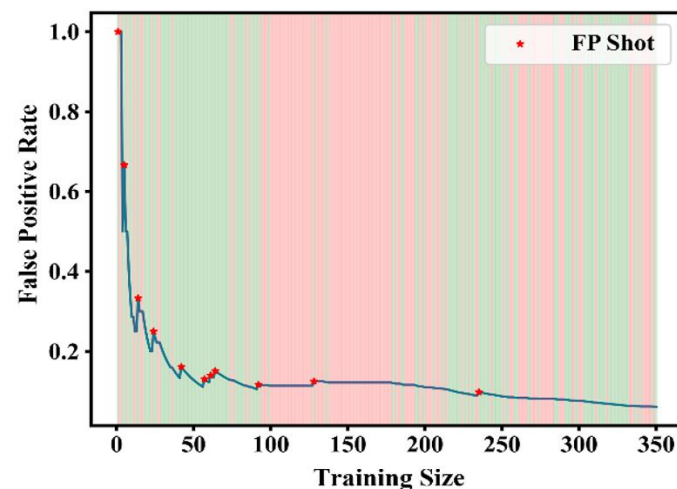
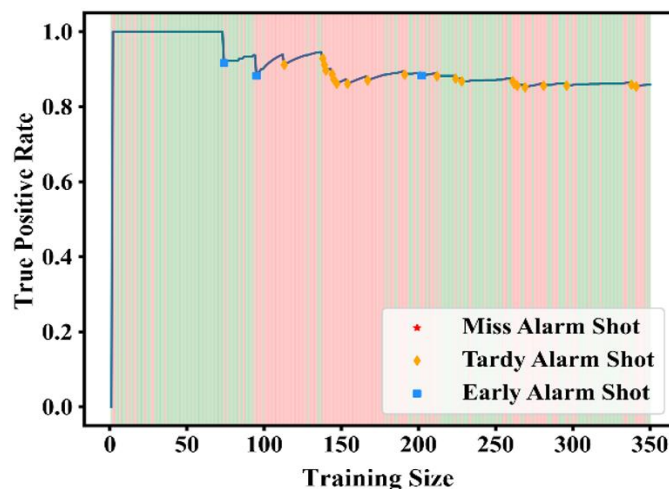
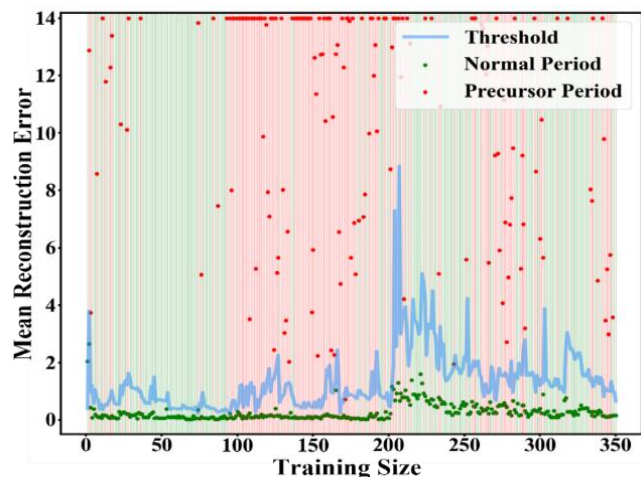
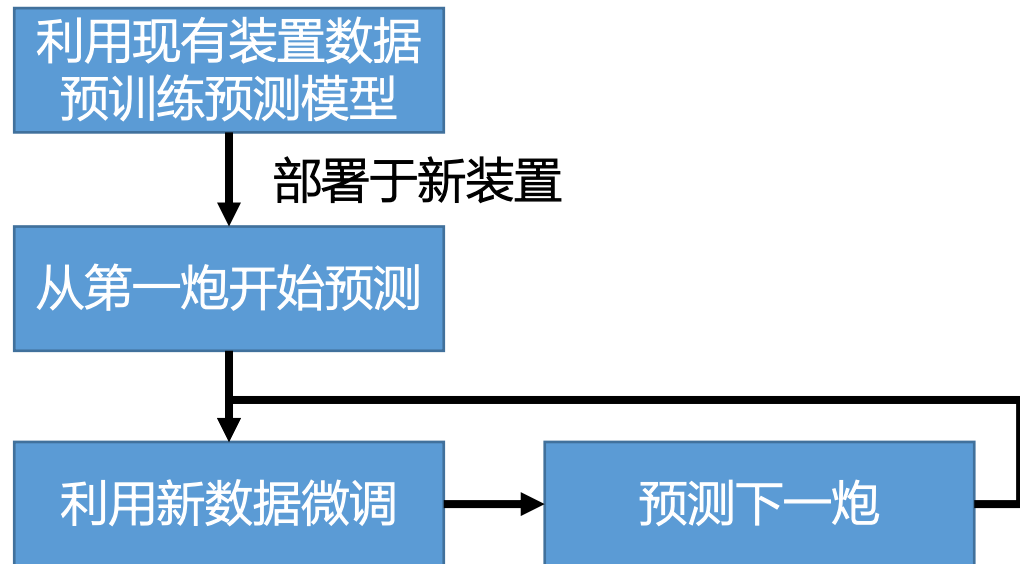
- 用异常检测标注破裂前兆训练前述的破裂预测模型，可以提高预测性能

Model ◦	Accuracy [%] ◦	TPR [%] ◦	FPR[%] ◦
Optimized ◦	93.75 ◦	92.50 ◦	5.00 ◦
25ms ◦	93.50 ◦	92.50 ◦	5.50 ◦
50ms ◦	93.00 ◦	92.50 ◦	6.50 ◦
100ms ◦	92.25 ◦	92.50 ◦	8.00 ◦

第一炮怎么办？装置运行区间变化

- 从零开始的自适应模型

- 破裂预测模型不需要准确预测没一炮破裂
- 只要准确预测下一炮就行



- 首先前面的都是未解之谜
- 目前的数据依然不足以解决破裂预测问题
- 模型是黑盒无法解释
- 在没有未来装置数据的情况下，如何评估在未来装置上的性能？

- [1] W. Zheng, F.R. Hu, M. Zhang, et al., Hybrid neural network for density limit disruption prediction and avoidance on J-TEXT tokamak, Nucl. Fusion, 58 (2018) 056016.
- [2] M. Zhang, Q. Wu, W. Zheng, et al., A database for developing machine learning based disruption predictors, Fusion Engineering and Design, 160 (2020) 111981.
- [3] W. Zheng, Q.Q. Wu, M. Zhang, et al., Disruption predictor based on neural network and anomaly detection on J-TEXT, Plasma Physics and Controlled Fusion, 62 (2020) 045012.
- [4] Y. Zhong, W. Zheng, Z.Y. Chen, et al., Disruption prediction and model analysis using LightGBM on J-TEXT and HL-2A, Plasma Physics and Controlled Fusion, 63 (2021) 075008.
- [5] C. Shen, W. Zheng, Y. Ding, et al., IDP-PGFE: An Interpretable Disruption Predictor based on Physics-Guided Feature Extraction, in, 2022, pp. arXiv:2208.13197.
- [6] W. Zheng, F.M. Xue, C.S. Shen, et al., Overview of machine learning applications in fusion plasma experiments on J-TEXT tokamak, Plasma Sci. Technol., 24 (2022) 12.
- [7] X.K. Ai, W. Zheng, M. Zhang, et al., Tokamak plasma disruption precursor onset time study based on semi-supervised anomaly detection, Nuclear Engineering and Technology, (2023).
- [8] W. Zheng, F. Xue, Z. Chen, et al., Disruption prediction for future tokamaks using parameter-based transfer learning, Communications Physics, 6 (2023) 181.
- [9] X. Ai, W. Zheng, M. Zhang, et al., Cross-tokamak deployment study of plasma disruption predictors based on convolutional autoencoder, Plasma Physics and Controlled Fusion, (2024).
- [10] C. Shen, W. Zheng, B. Guo, et al., Cross-tokamak disruption prediction based on domain adaptation, Nucl. Fusion, 64 (2024) 066036.
- [11] Y. Zhong, W. Zheng, Z. Chen, et al., High-beta disruption prediction study on HL-2A with instance-based transfer learning, Nucl. Fusion, (2024).
- [12] X.K. Ai, W. Zheng, M. Zhang, et al., Adaptive anomaly detection disruption prediction starting from first discharge on tokamak, Nucl. Fusion, 65 (2025) 036011.

- [1] C. Rea, R.S. Granetz, K. Montes, et al. Disruption prediction investigations using machine learning tools on DIII-D and Alcator C-Mod. Plasma Physics and Controlled Fusion, 2018, 60(8), 084004.
- [2] J. Kates-Harbeck, A. Svyatkovskiy and W. Tang. Predicting disruptive instabilities in controlled fusion plasmas through deep learning. Nature, 2019, 568: 526-531.
- [3] Z. Yang, F. Xia, X. Song, et al. A disruption predictor based on a 1.5-dimensional convolutional neural network in HL-2A. Nucl. Fusion, 2020, 60(1), 016017.
- [4] B. Guo, B. Shen, D. Chen, et al. Disruption prediction using a full convolutional neural network on EAST. Plasma Physics and Controlled Fusion, 2021, 63(2), 025008.
- [5] Z. Yang, F. Xia, X. Song, et al. In-depth research on the interpretable disruption predictor in HL-2A. Nuclear Fusion, 2021, 61(12), 126042.
- [6] W.H. Hu, C. Rea, Q.P. Yuan et al., Real-time prediction of high-density EAST disruptions using random forest, Nucl. Fusion 61, 066034 (2021).
- [7] J.X. Zhu, C. Rea, K. Montes, et al. Hybrid deep-learning architecture for general disruption prediction across multiple tokamaks. Nucl. Fusion, 2021, 61(2) , 026007.
- [8] H. Li, Y. Fu, J. Li, et al., Simulation Prediction of Heat Transport with Machine Learning in Tokamak Plasmas, Chinese Physics Letters 40, 125201 (2023).
- [9] Chenguang Wan, Zhi Yu, Alessandro Pau, et al. A machine-learning-based tool for last closed-flux surface reconstruction on tokamaks, 2023 Nucl. Fusion 63 056019
- [10] B. Guo, D. Chen, C. Rea, et al. Disruption prediction on EAST with different wall conditions based on a multi-scale deep hybrid neural network Nuclear Fusion, 2023, 63(9), 094001.
- [11] L. Tong, Z. Wang, L. Wei, et al., Enhancement of ECCD by the current condensation effect for stabilizing large magnetic islands caused by neoclassical tearing modes in tokamak plasmas, Nucl. Fusion, 64 (2024) 036001.
- [12] Zongyu Yang, Wulyu Zhong*, Fan Xia, Zhe Gao, Xiaobo Zhu, Jiyuan Li, Liwen Hu, Zhaohe Xu, Da Li, Guohui Zheng et al “Implementing deep learning-based disruption prediction in a drifting data environment of new tokamak: HL-3” 2025 Nucl. Fusion 65 026030.

• 谢谢

