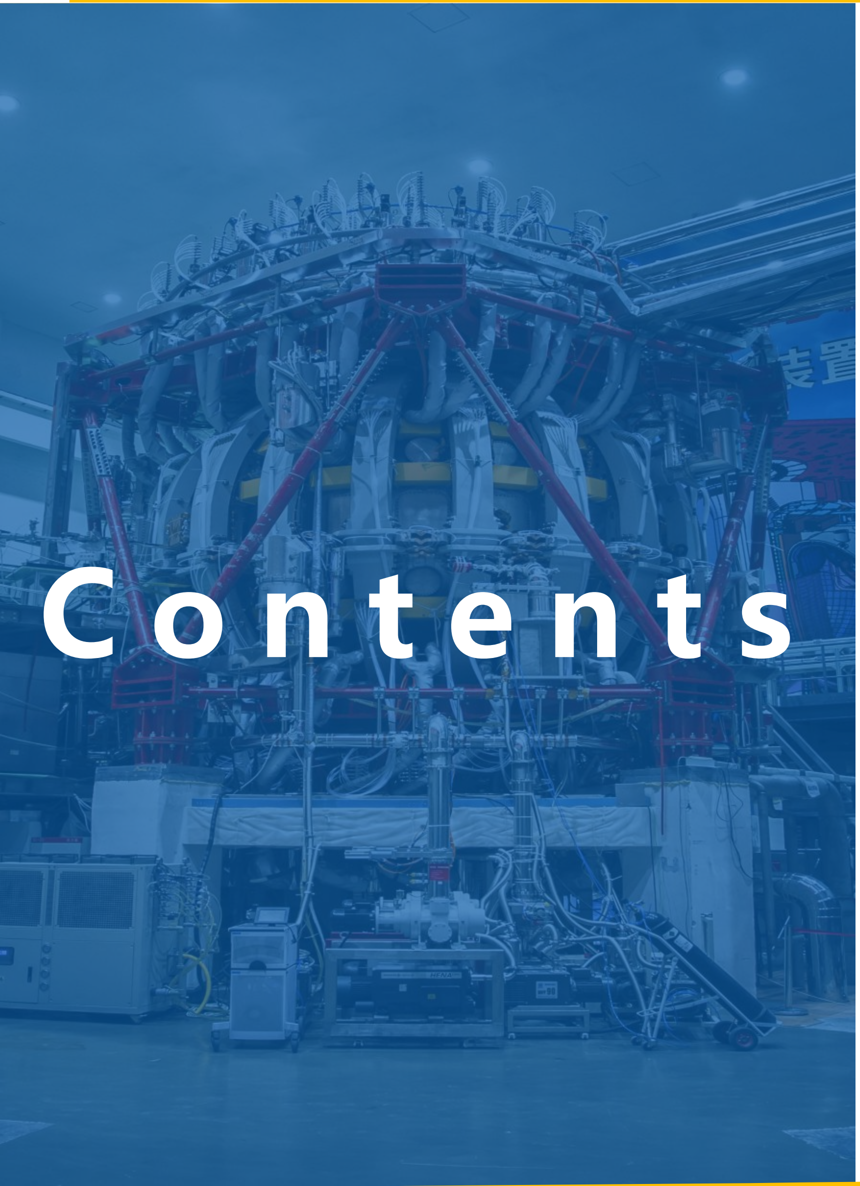
A Neural Network Model for Rapid Analysis and Extrapolation in Spectral Diagnostics

Wenjing Tian, Zongyu Yang, Shuosu Yang, Min Xu, Xiaoxue He, Deliang Yu, Zhe Gao[2], Zhongbing Shi

Southwestern Institute of Physics

2 Tsinghua University



Contents

Significance of real-time measurement

■ Essential for Future Reactors

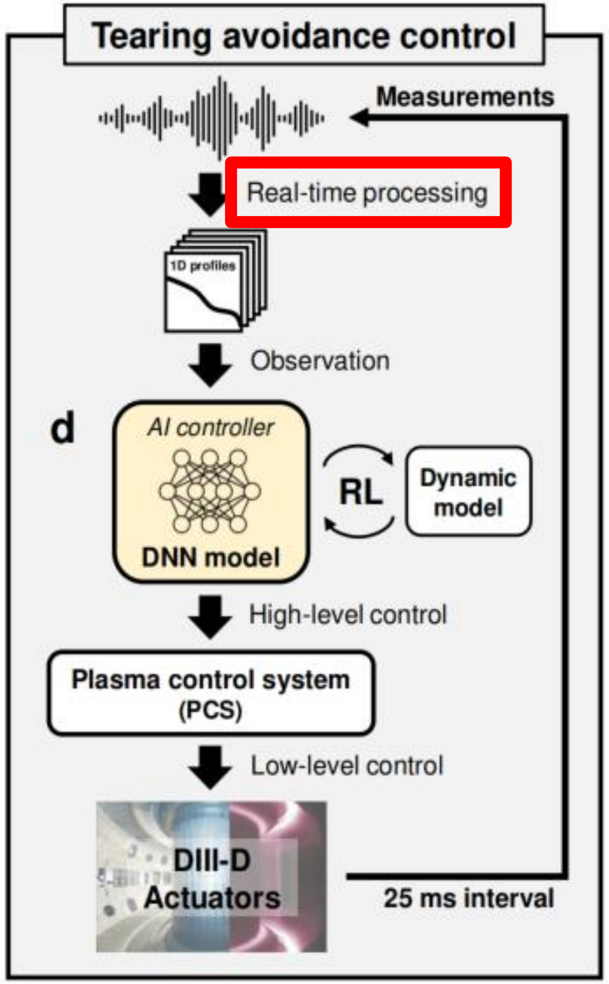
Real-time plasma control is essential for the operation of future fusion reactors.

■ Key Physics Parameters

Ion temperature (T_i) and rotation velocity (v_t) are among the most critical parameters.

■ The Bottleneck

Roles of T_i and v_t have been missing in most real-time control scenarios.



Real-time measurement of plasma parameters is the input for the control loop

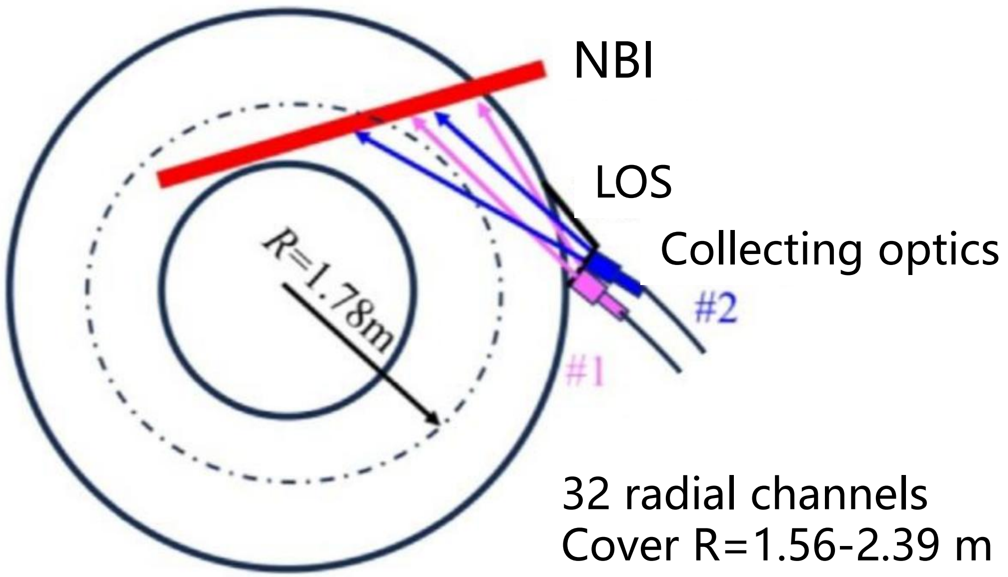
Jaemin Seo et al. *Nature* (2023):626.

CXRS diagnostic on HL-3

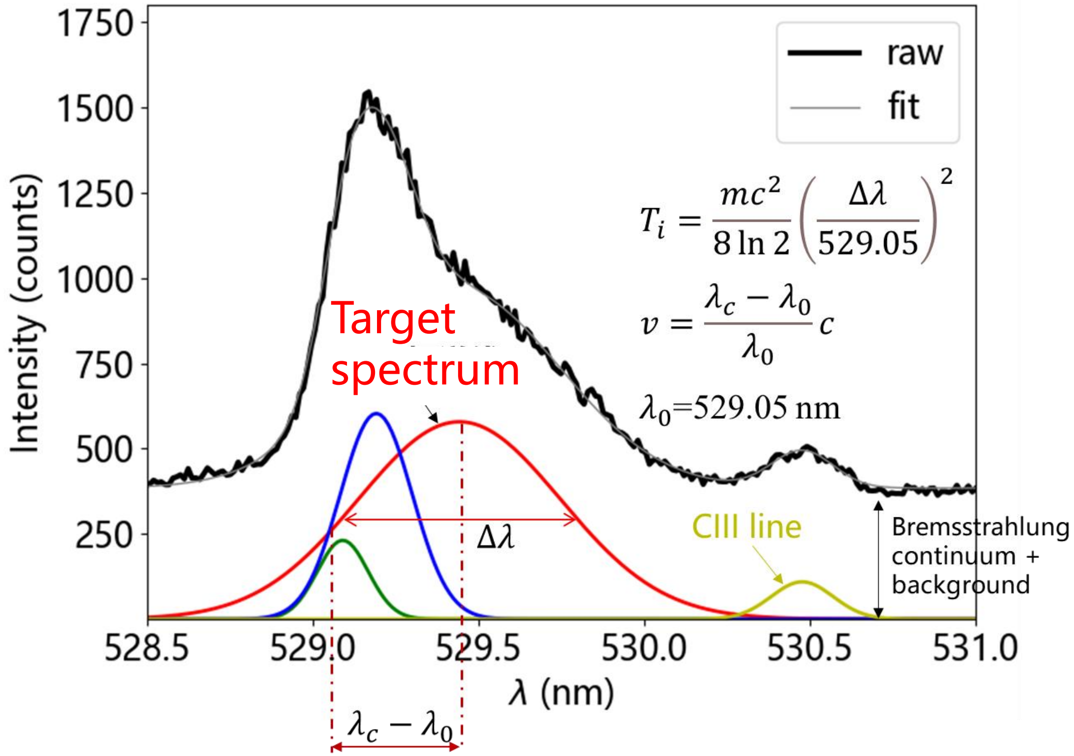
■ **Primary Diagnostic：** Charge Exchange Recombination Spectroscopy (CXRS)

CXRS diagnostic measures light (hν) emitted by ionized impurities (e.g., C^{6+}) following charge exchange with neutral beam atoms.

$$D^0 + C^{6+} \rightarrow D^+ + C^{5+*} \rightarrow D^+ + C^{5+} + h\nu$$

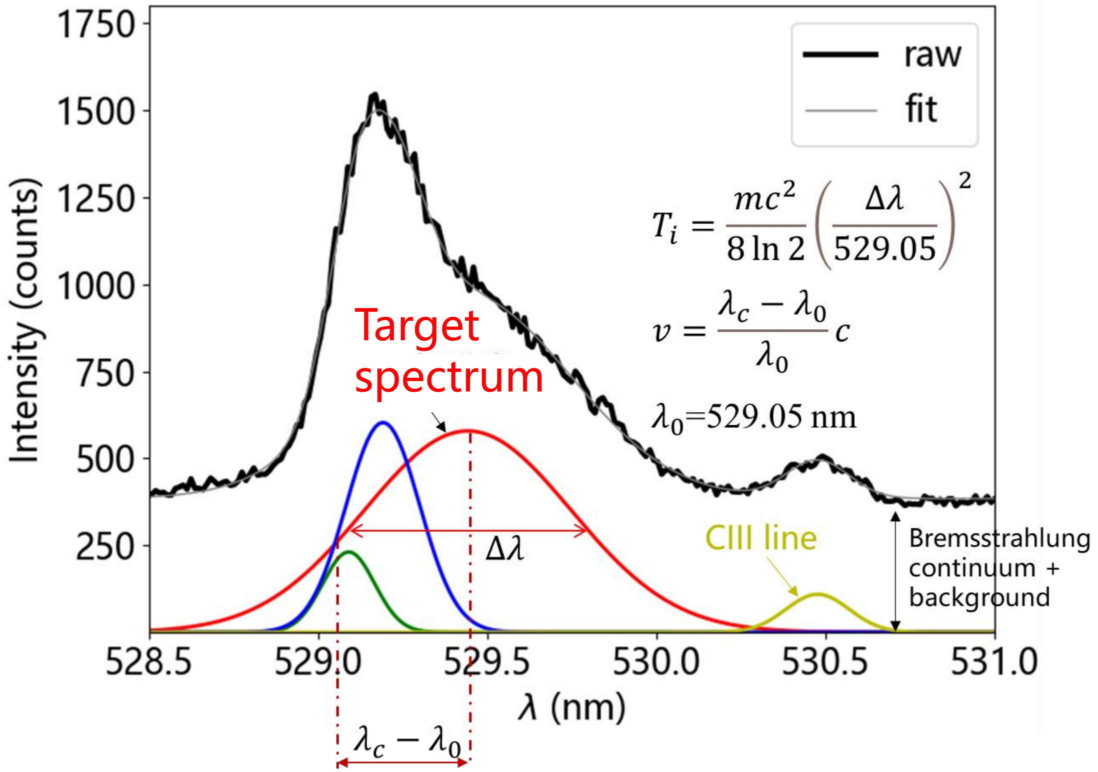


Schematic of CXRS diagnostic on HL-3 tokamak



Example of the collected CXRS spectra

Limitations of Conventional Spectral Analysis



Example of the collected CXRS spectra and its fit results

Non-linear Least-Squares Fitting

$$I(\lambda) = I_0 + \sum_{i=1}^{N} a_i \exp\left(-\frac{(\lambda-\lambda_i)^2}{2\sigma_i^2}\right)$$

$$E = \sum_{j=1}^{M} \frac{\left(I(\lambda) - y_j\right)^2}{\sigma_j^2}$$

⚠️ **High Computational Latency**

~ 100 - 1000 ms per spectrum

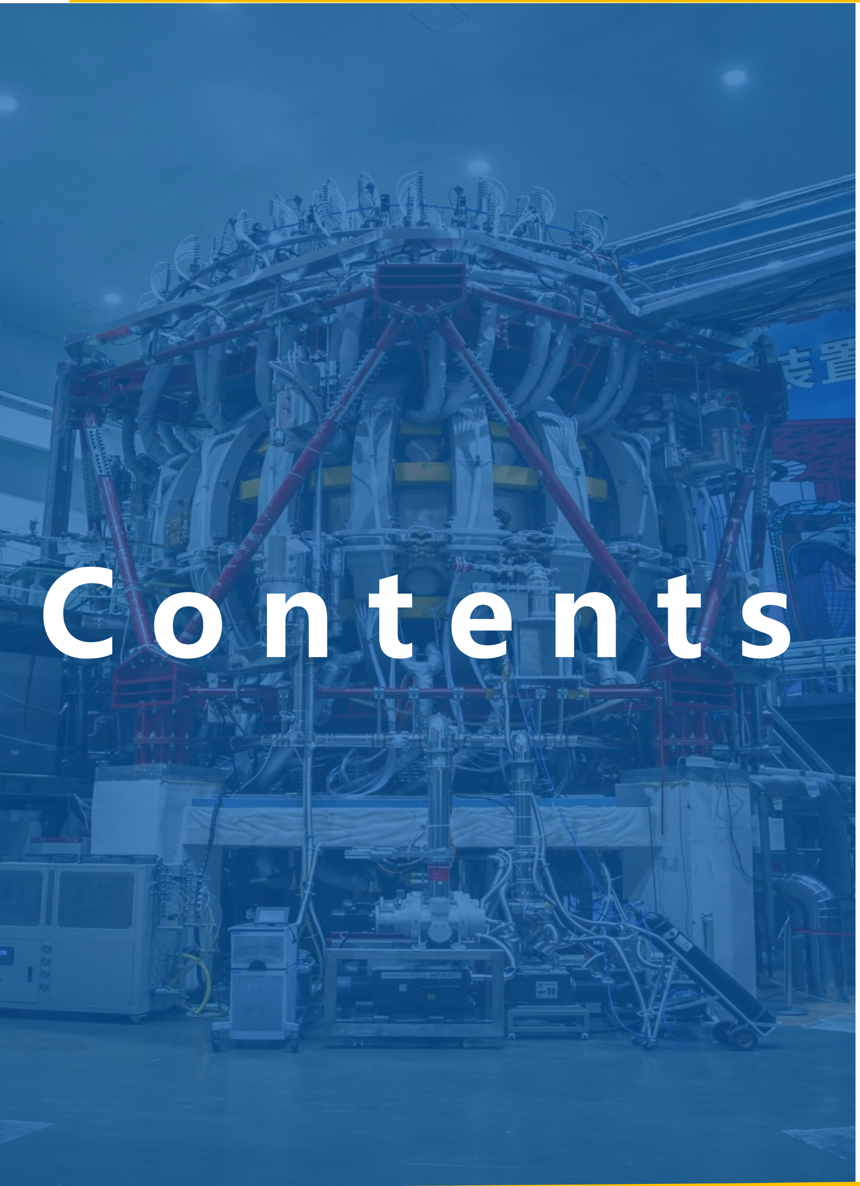
⚠️ **Dependence on Expert Knowledge**

Requires accurate initial guesses for parameters

Our solution： Neural network spectrum analysis

Objectives

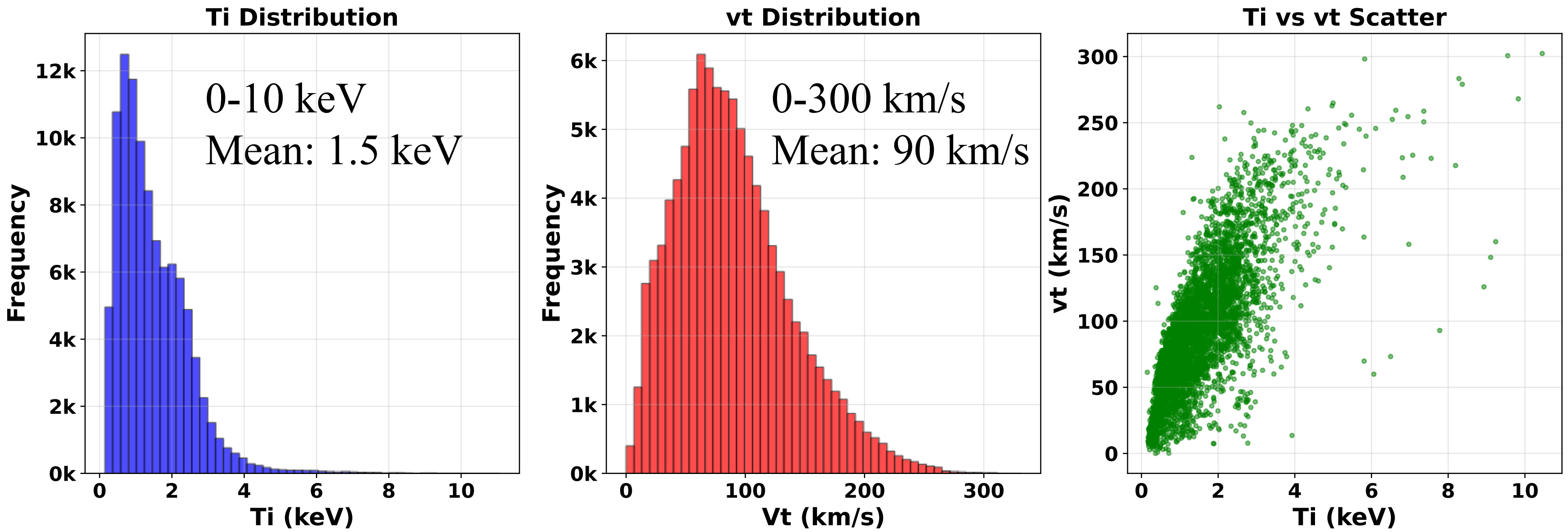
- Explore the feasibility of using NNs for rapid spectral analysis
- Explore Interpretability of the 'black-box' model: attribution analysis
- For application in future devices: cross-device investigation (HL-2A and HL-3 tokamaks)



Contents

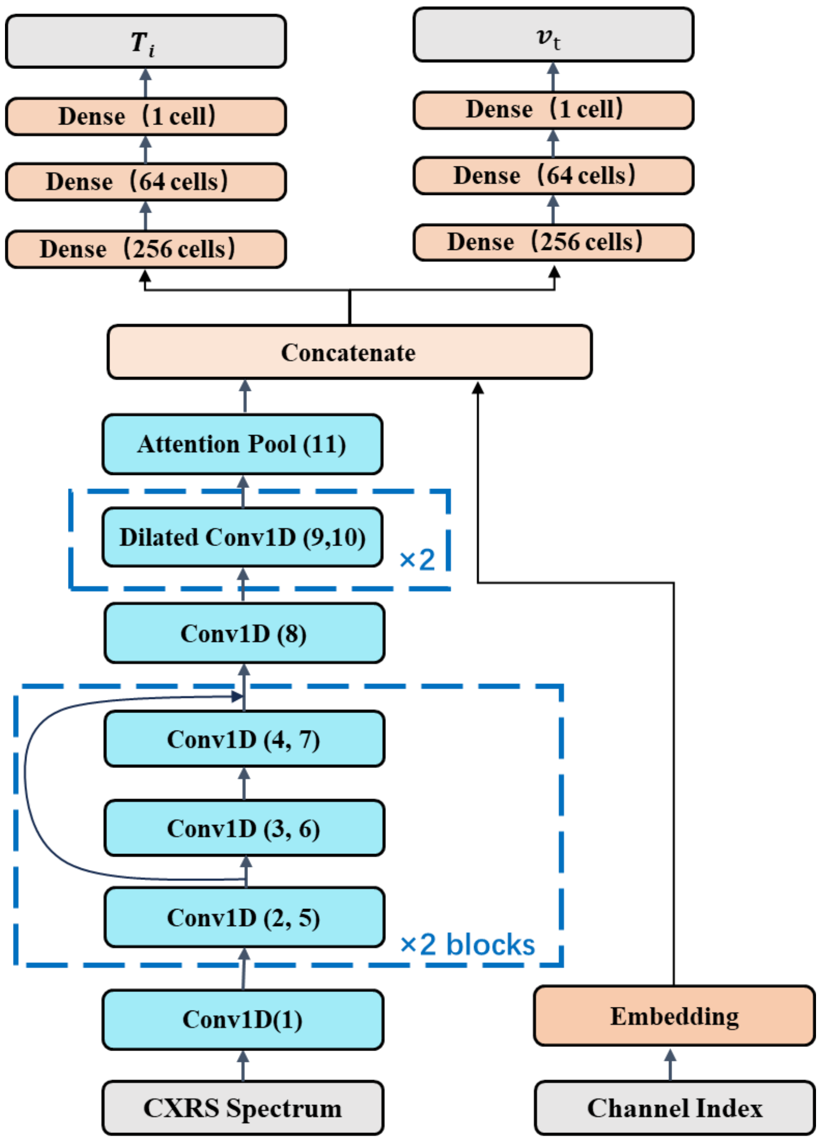
■ Dataset

- Input： spectrum （1×200） ， channel index （1×1）
- Output (label): T_i and v_t (1×1, 1×1) from Least-Squares Fitting
- Size: >100k spectra from 190 discharges on HL-3 tokamak
- Splitting rule: 5-fold cross validation. Stratified by shot number



Distributions of Ti and vt in the dataset

Architecture of model



CNN backbone: For extracting localized, structured features from spectral signals.

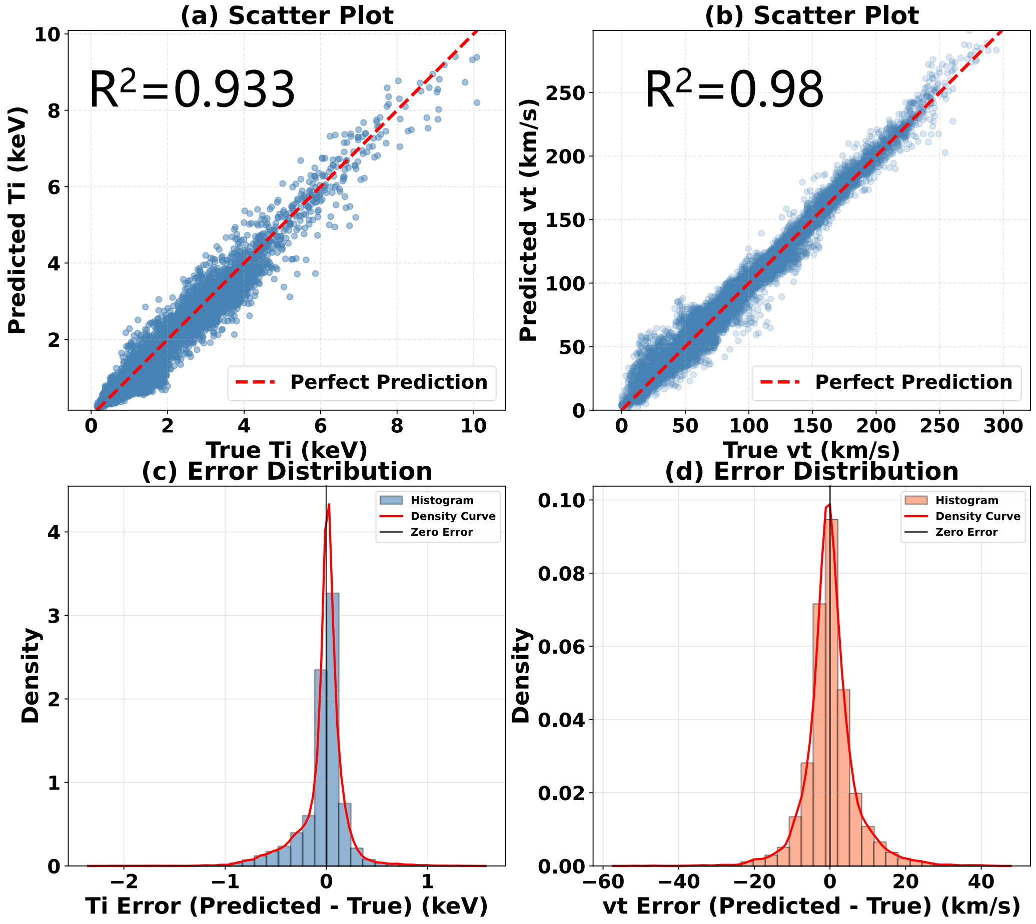
Channel embedding: Provides spatial awareness by encoding the channel/position index.

Attention pooling: Replaces global pooling to adaptively focus on critical spectral regions.

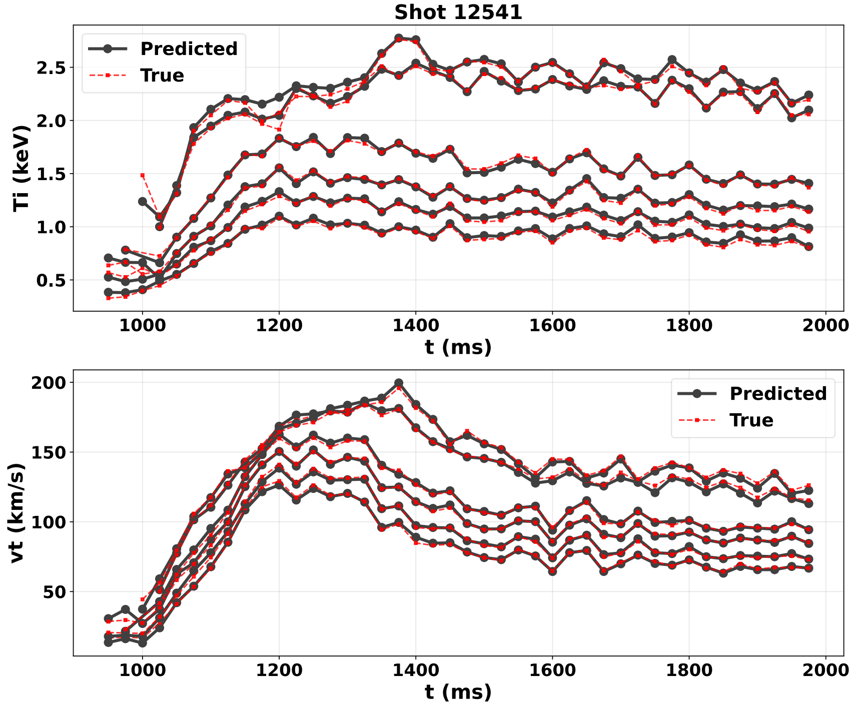
Multi-task learning: Shares representations to jointly predict T_i and v_t efficiently.

Results

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$



Parameter	R^2	MAE	MRE	Inference time
Ti	0.933	0.14keV	9.8%	0.56 ms/ 32 channels
vt	0.980	4.50km/s	——	



- ✅ Achieves real-time capability
- ✅ High fidelity across time evolution

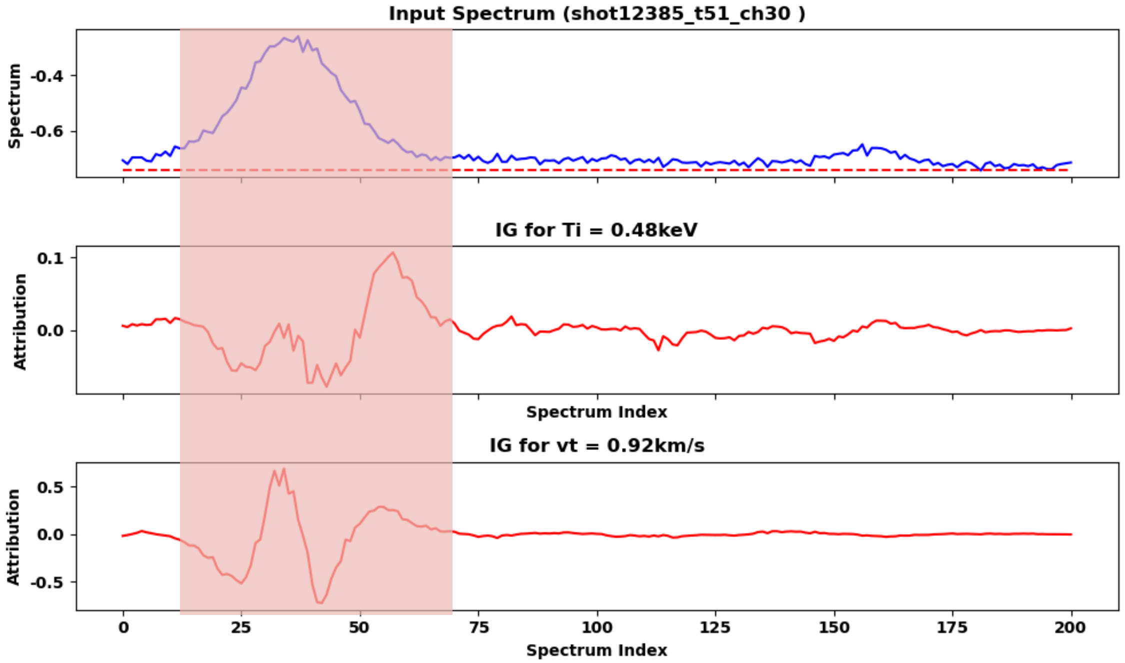
Interpretability

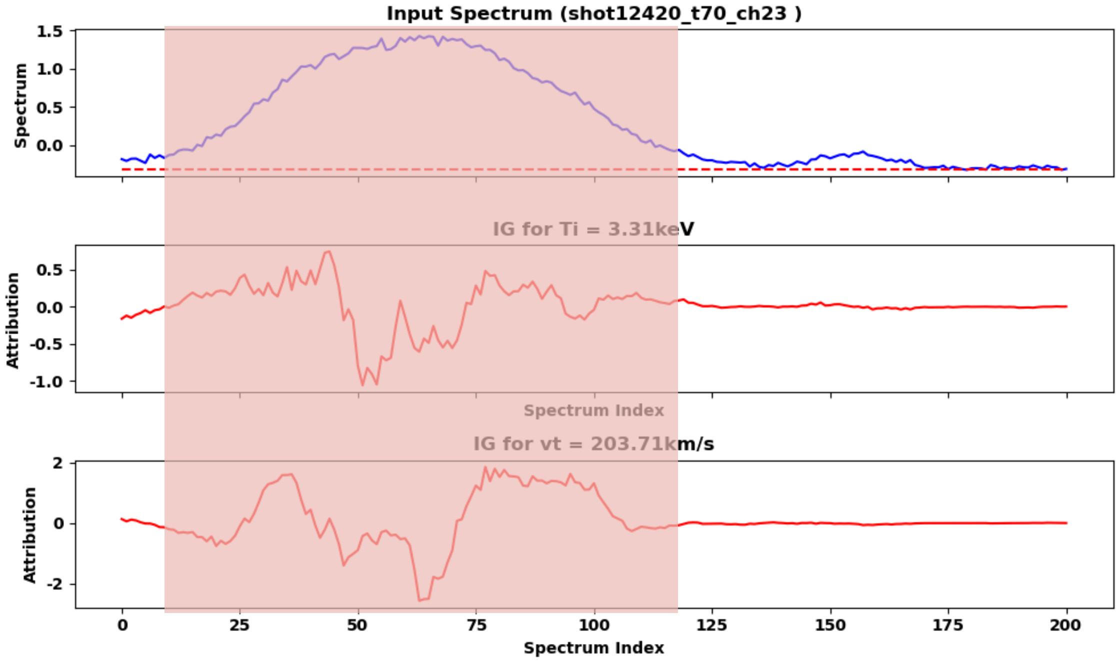
■ Why interpretability?

To understand and trust the NN model

To guide model optimization

$$\text{Integrated Gradient:}\quad IG_i(x) = (x_i - x_i^{'}) \times \int_0^1 \frac{\partial F\left(x^{'} + \alpha \times (x - x^{'})\right)}{\partial x_i} d\alpha$$





Interpretability

- **Why interpretability?**

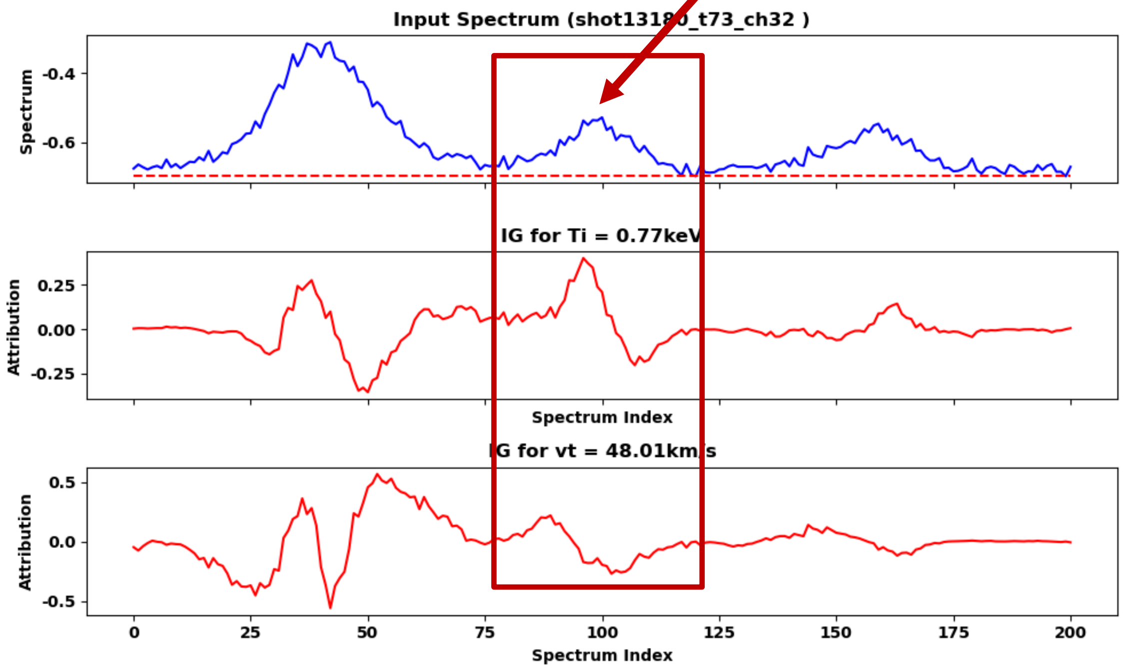
To understand and trust the NN model

To guide model optimization

Interpretability Diagnoses a Failure：

True Ti： 0.77 keV
Pred Ti： 2.70 keV

Affected by 'less-seen' interfering spectral peak



- Potential solution: synthetic spectra as data augmentation[1].

[1] Wenjing Tian, et al. Acta Physica Sinica, 74, 078901 (2025)

Brief Summary

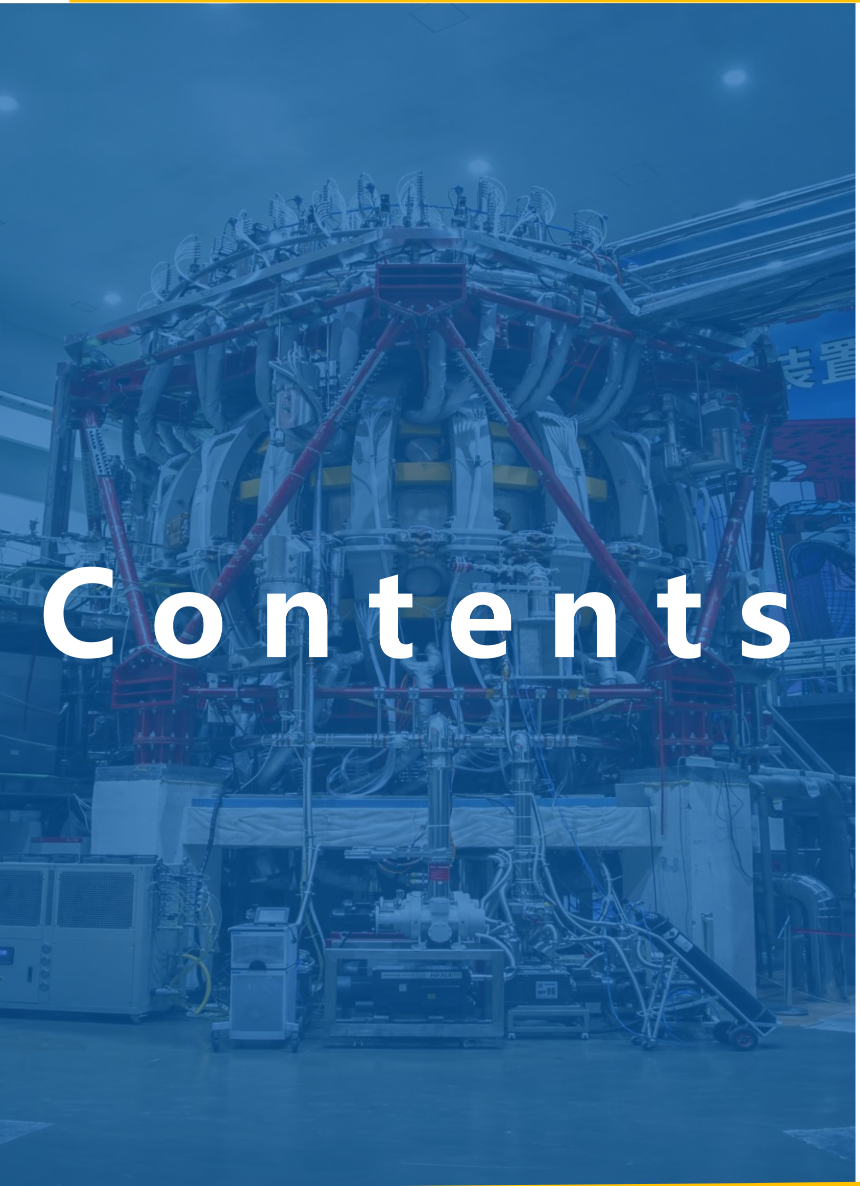
- We developed an NN model for CXRS spectral analysis.

High accuracy: reaches R^2 of 0.93 for Ti and 0.98 for vt

Real-time speed: provide a profile of 32 channels in <1 ms

Interpretable & reliable: Makes decisions based on physically meaningful features.

Parameter	R^2	MAE	MRE	Inference time
Ti	0.933	0.14	9.8%	0.56 ms/ 32 channels
vt	0.980	4.50	——	

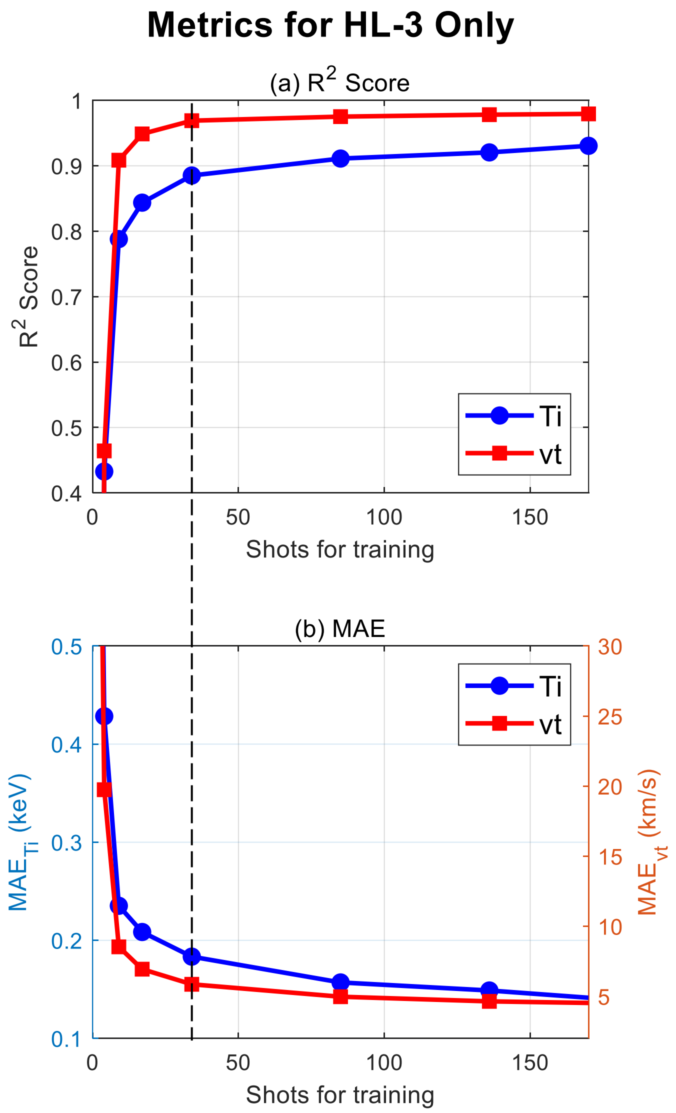


Contents

The Data Dilemma

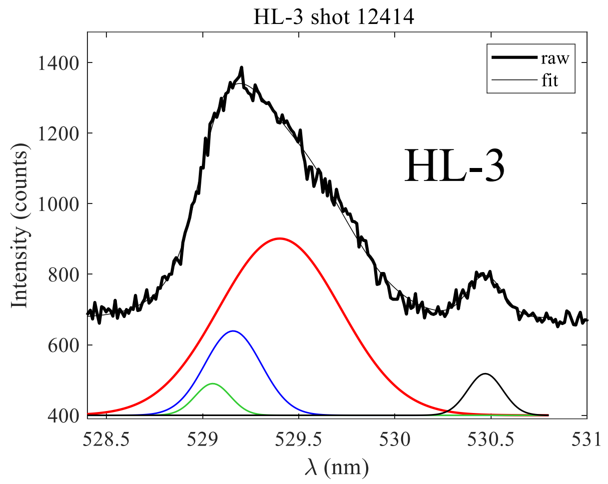
■ Scan the number of shots for training

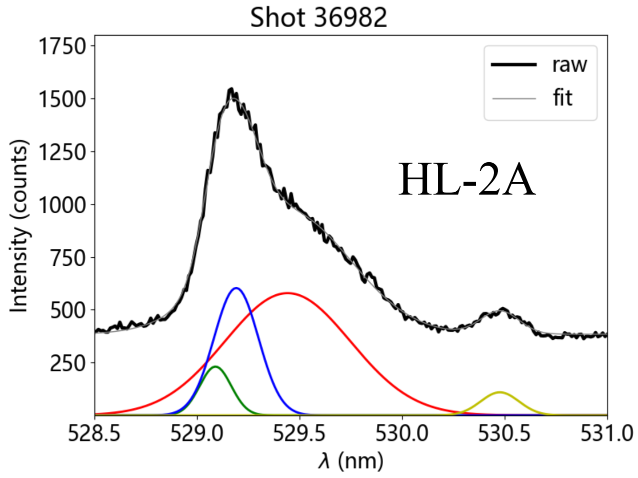
- ~30 shots
- The more data, the better.



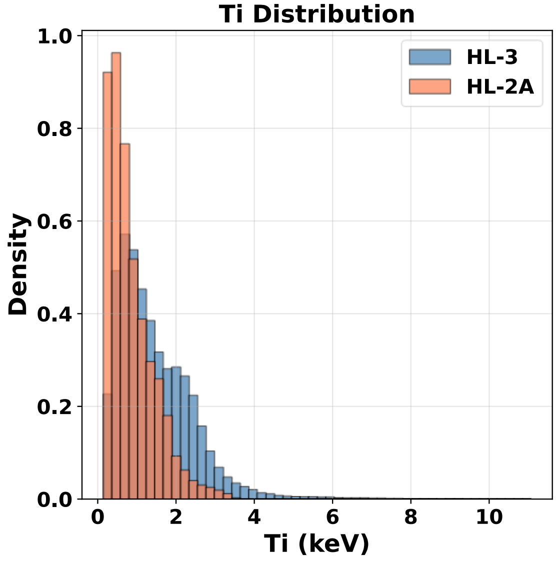
Could spectral data from previous device be of help?

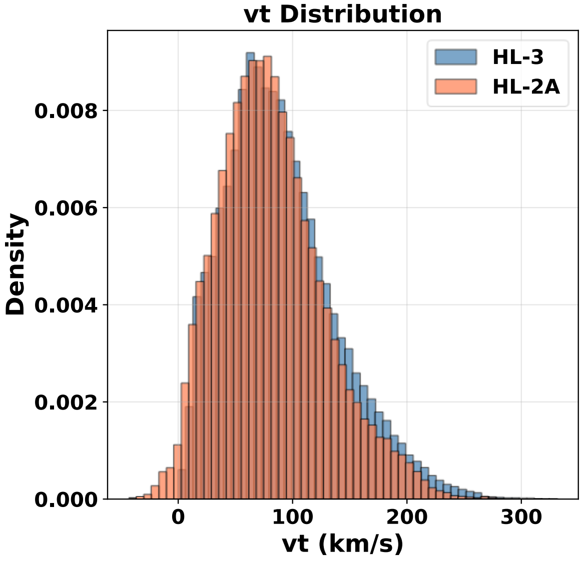
HL-2A v.s. HL-3 dataset

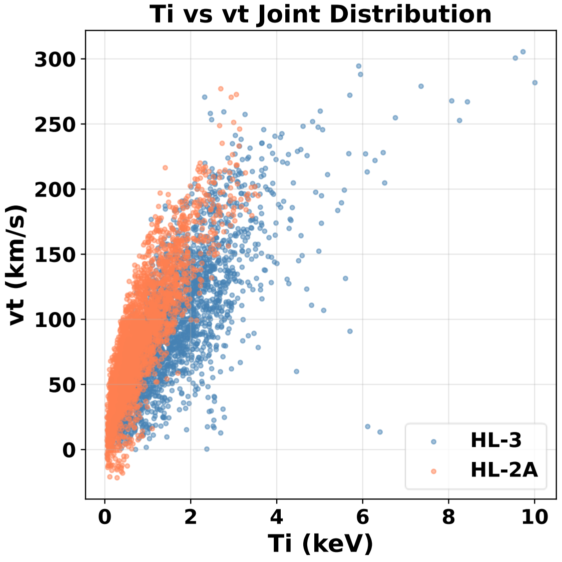




Similar components,
Different Spectral structures







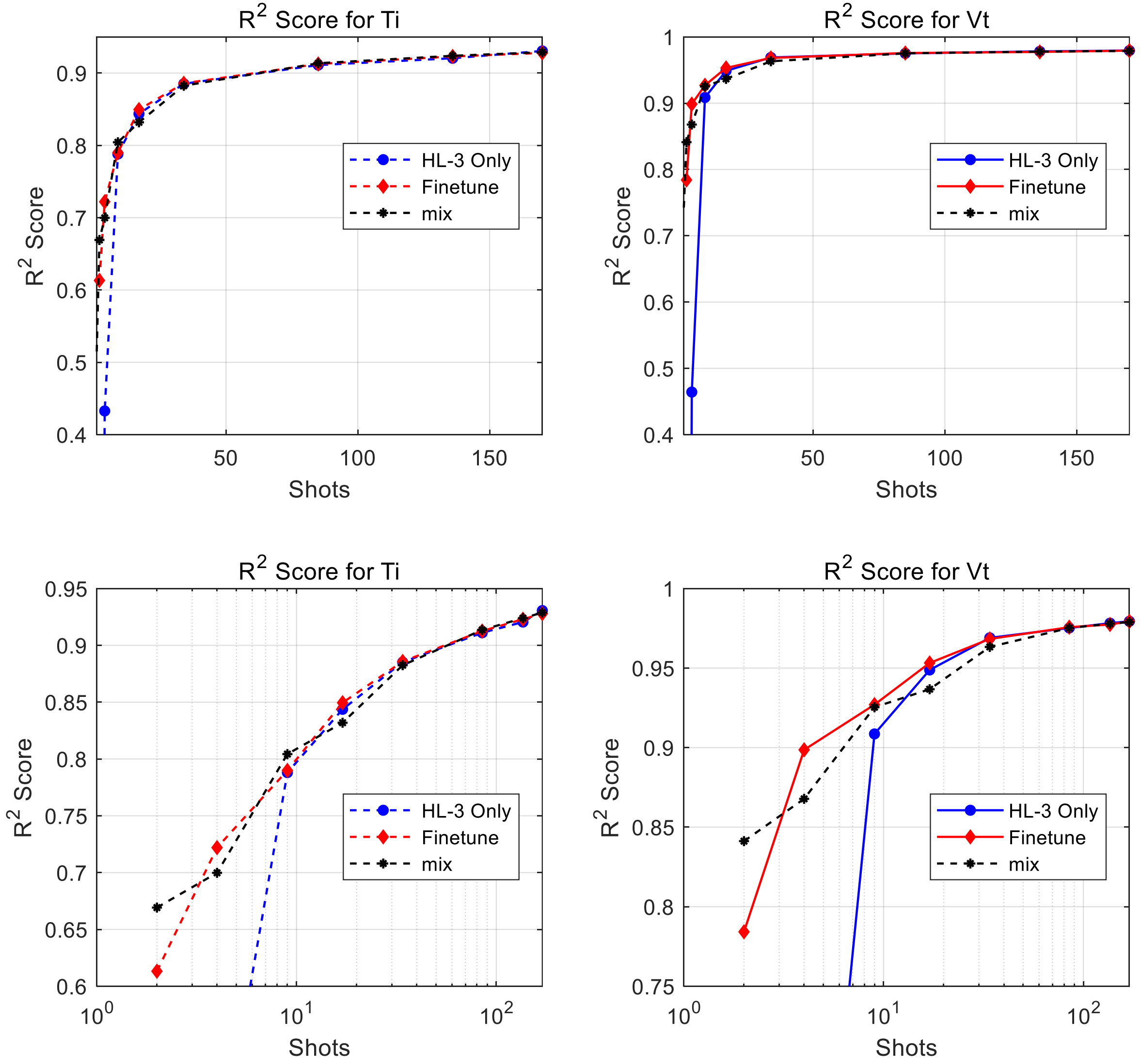
HL-3: 100k spectra from 190 shots.
HL-2A: 122k spectra from 89 shots.

Approaches

Mix: Simply mix HL-2A data into training set.
Finetune: Pretrain with HL-2A data, finetune with HL-3 data.

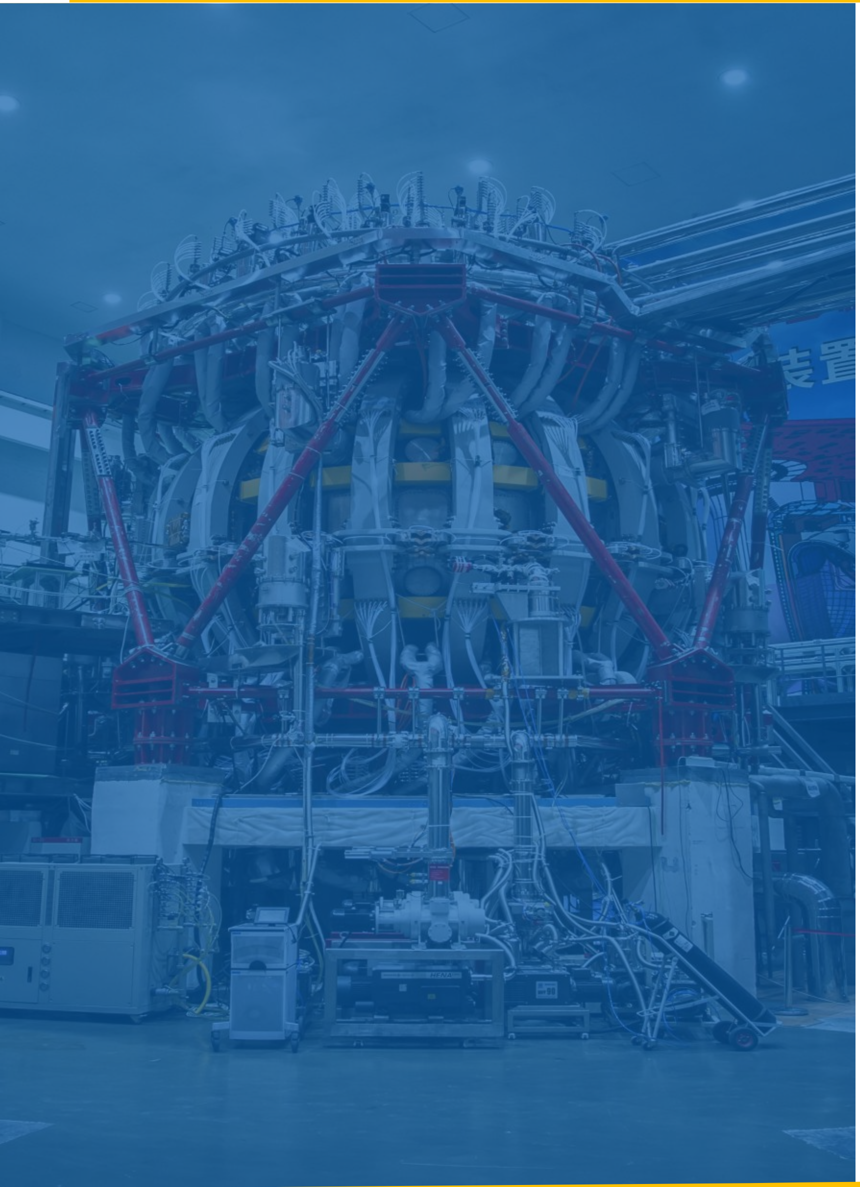
Scan the number of HL-3 shots for training.

Data from HL-2A helps to get a head start.



Summary

- We developed an NN model for CXRS spectral analysis with high accuracy, real-time speed and is interpretable.
- Spectral data from previous device help to get a head start for the NN model in a new device.
- Contribute to the development of AI-assisted plasma diagnostics analysis and real-time control in fusion reactors



Thanks

Wenjing Tian
E-mail:
tianwenjing@swip.ac.cn