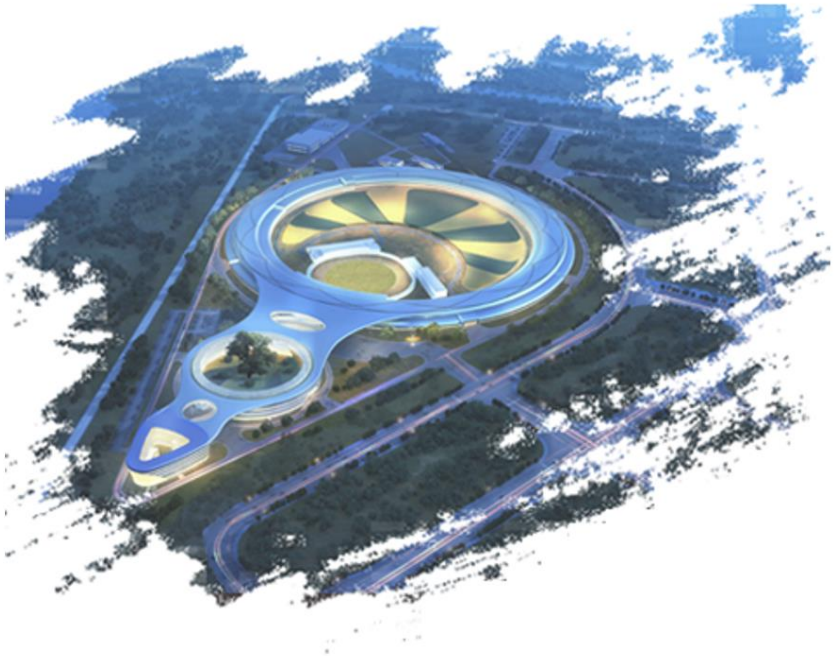
X-ray Diffraction Data Analysis Assistant Powered by Large Language Model

Yihe Pang

Institute of High Energy Physics, Chinese Academy of Sciences

2025-8-12



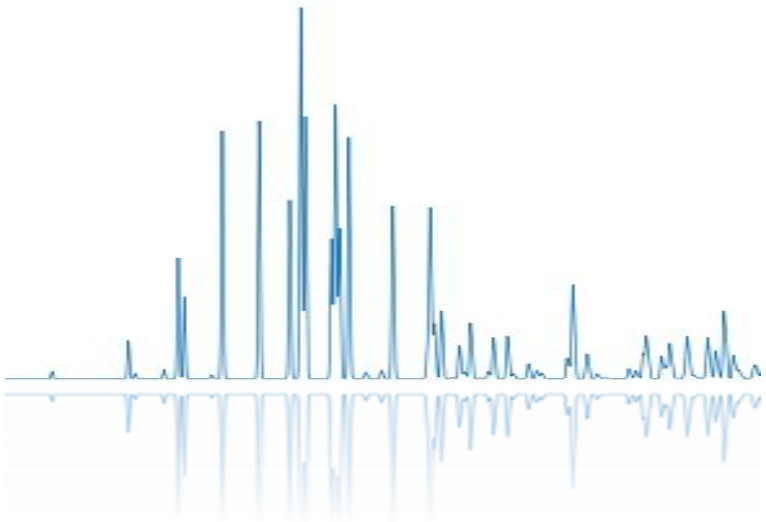
Outline

- A Deep Learning Methods for Crystal Symmetry Identification from XRD
- Automatic XRD Data Analysis Powered by LLM

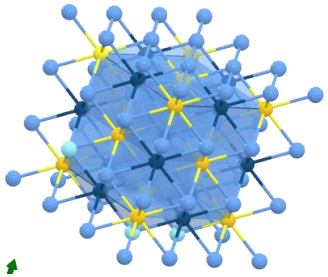
Outline

- A Deep Learning Methods for Crystal Symmetry Identification from XRD
- Automatic XRD Data Analysis Powered by LLM

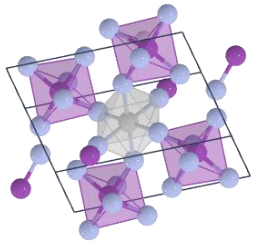
➢ XRD: a powerful technique for characterizing the symmetry of crystal structure



X-ray Diffraction



Crystal System
(7 categories)



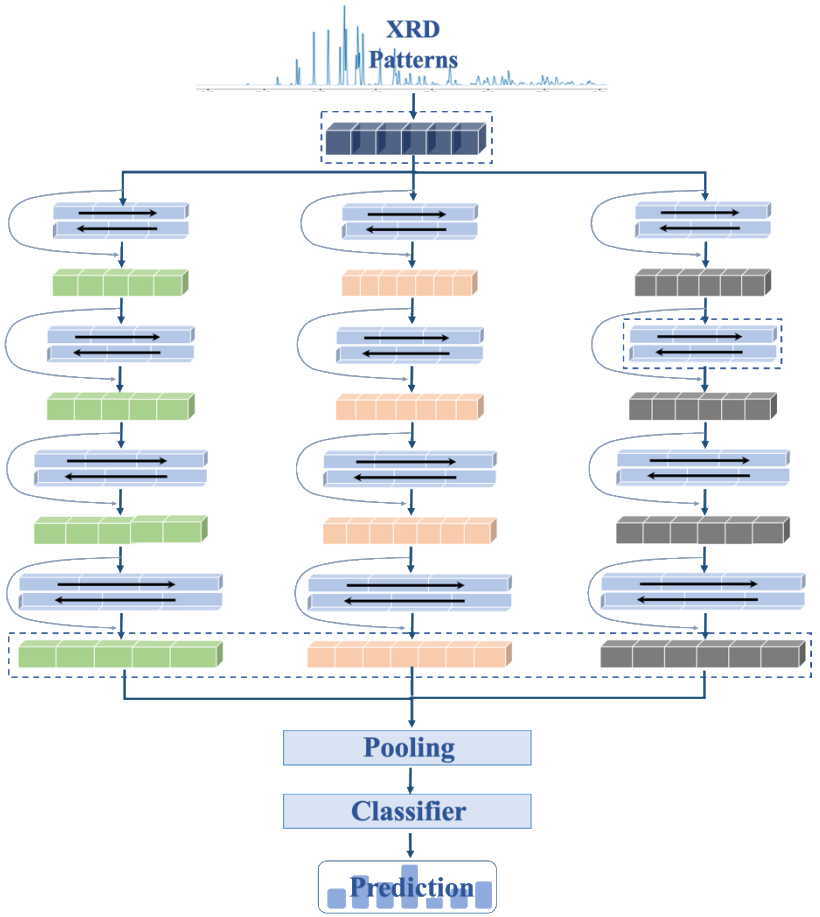
Space Group
(230 types)

Deep learning enables high-throughput and efficient XRD pattern analysis

➢ Challenges

- Imbalance in numbers between symmetry categories
- XRD pattern have limited information
- Large gaps between experimental data and simulated data

➢ A Hybrid Convolutional Neural Network for Crystal Symmetry Identification from XRD



a Depth-wised Separable CNN

b Multi-Scaled CNN

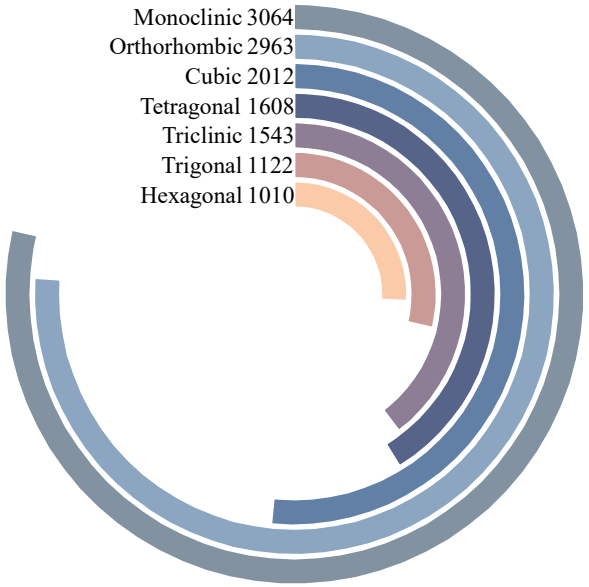
➢ HCNN achieves the best performance

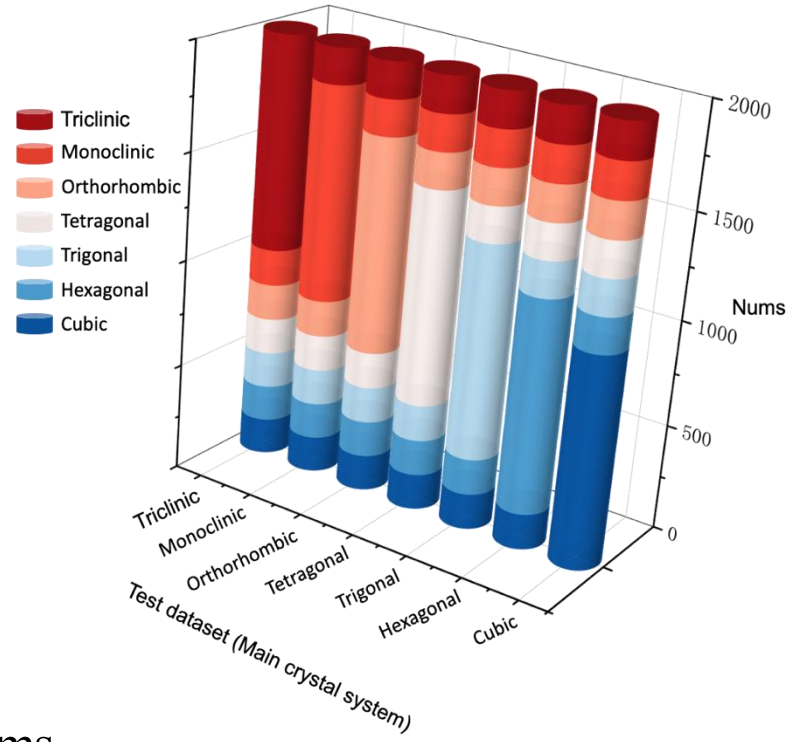
Table 1. Performance comparation between HCNN and other methods

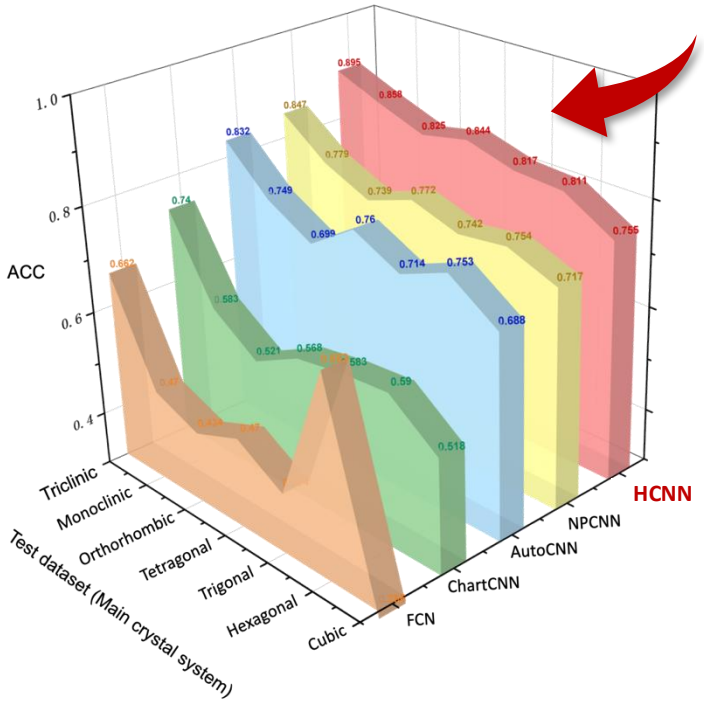
Methods	Crystal System				Space Group			
	Accuracy	Macro Precision	Macro F1	Macro PR-AUC	Accuracy	Macro Precision	Macro F1	Macro PR-AUC
FCN	54.249	56.734	49.020	54.183	13.244	0.264	0.363	5.556
ChartCNN	60.494	60.669	58.922	63.341	34.937	20.933	16.121	18.069
AutoCNN	74.756	76.664	75.156	77.220	38.016	7.197	7.040	12.699
NPCNN	76.325	77.615	76.972	83.804	57.492	43.395	41.757	44.853
HCNN(ours)	**82.428**	**84.308**	**83.508**	**89.216**	**65.271**	**54.932**	**50.487**	**55.162**

The Materials Project : https://next-gen.materialsproject.org/

➢ HCNN performs the most stable performance

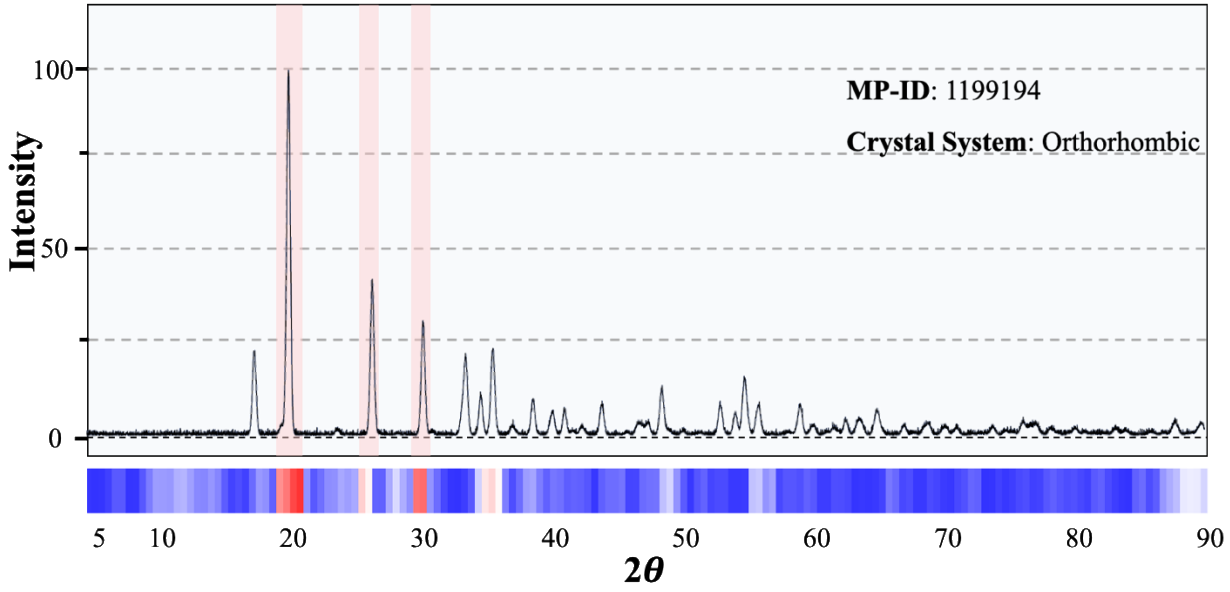


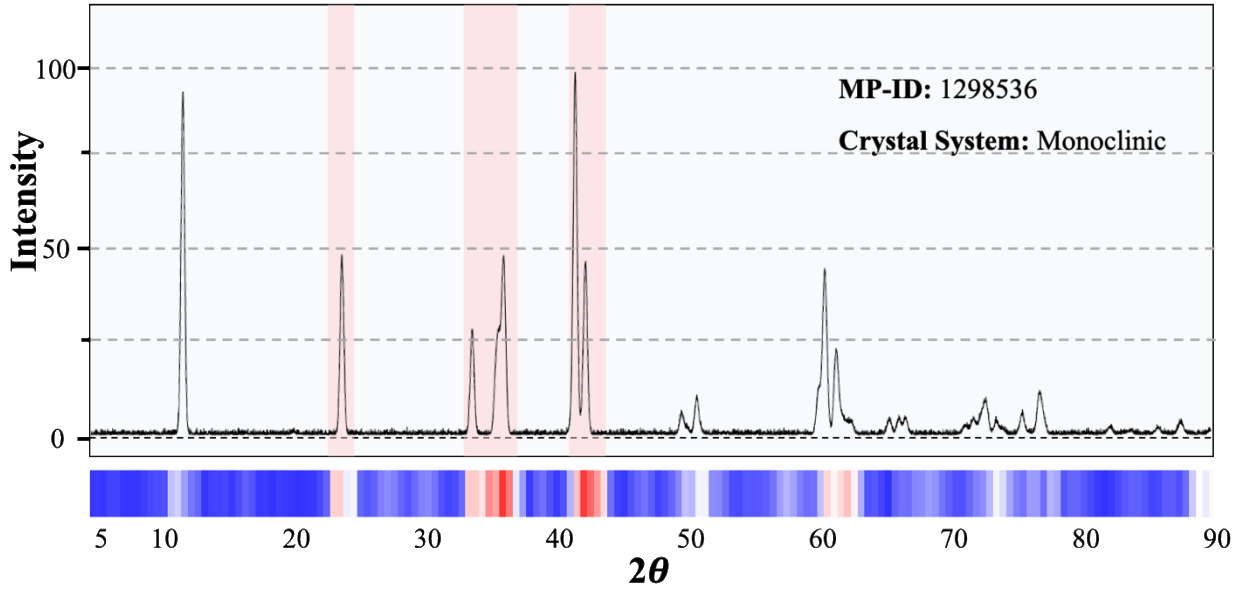




Unbalanced data in terms of crystal systems

➢ HCNN is physically explainable





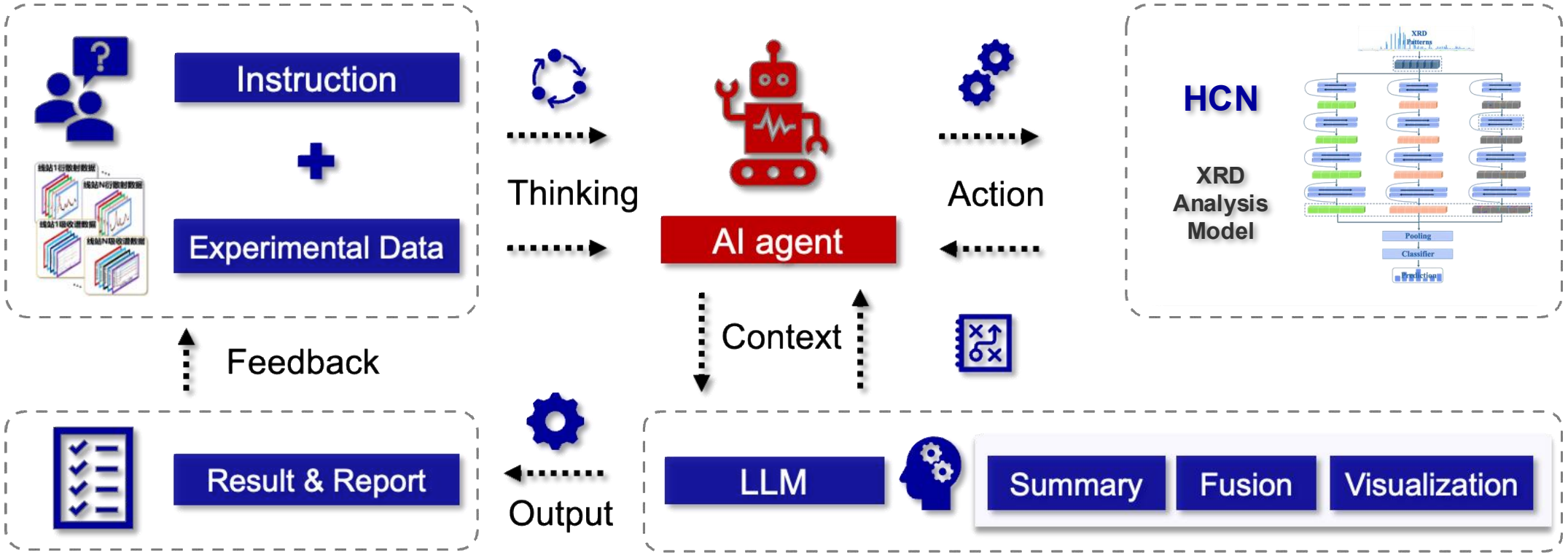
Outline

- A Deep Learning Methods for Crystal Symmetry Identification from XRD
- Automatic XRD Data Analysis Powered by LLM

➢ LLM: powerful AI for understanding and generating natural language

Automated analysis of XRD with text explanation

Workflow: Task decomposition, action decision, and analysis tools scheduling




同步辐射实验数据智能解析平台
新对话设置
请上传实验数据
从本地上传
开始对话
和机器人聊天

- **Acknowledgements**

  - postdoctoral co-advisor : Prof. Zhao
  - Research groups : AI molecules group
  - Student from the collaborating university : Jing Li (Minzu University of China，MUC)