Latent Diffusion Pretraining for Crystal Property Prediction

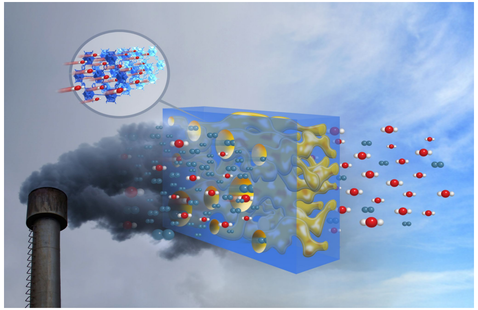
Shrimon Mukherjee[1*], Kishalay Das[2*], Partha Basuchowdhuri[1], Pawan Goyal[2], Niloy Ganguly[2]

[1]Indian Association for the Cultivation of Science, India
[2]Indian Institute of Technology Kharagpur, India

CNeRG

3D Materials Design forms the Foundation of Modern Technology

- Many of the problems we are facing today, are really **bottlenecked** by finding new materials.



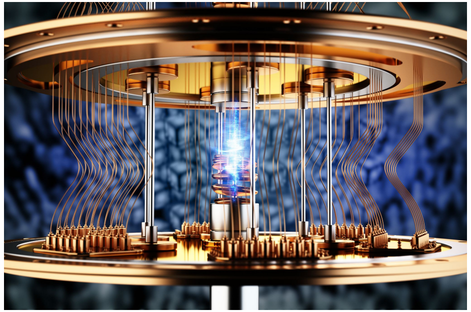
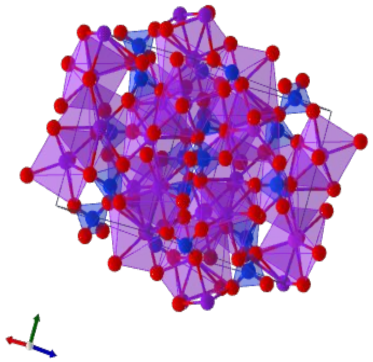


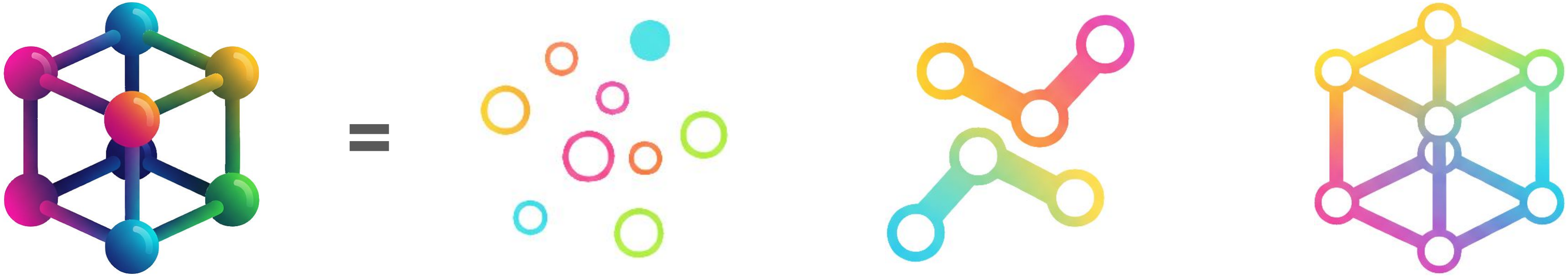
Image: Tian Xie

Crystal Material

- ❑ Crystals are represented by a **Minimal Unit Cell.**
- ❑ Constituent **atoms** in different **coordinates.**
- ❑ **Repeated** infinite times in **3D space on a regular lattice.**
- ❑ Material structures are **periodic in nature**.



Material Representation



M = (A,X,L) A: Atom Types X: Fractional Coordinates L: Lattice Structure

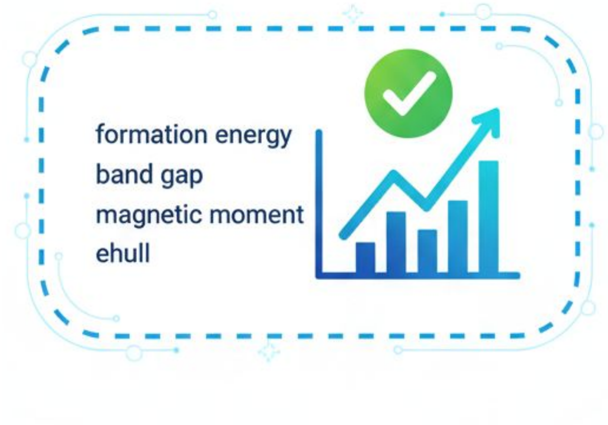
- Represent **Infinite Periodic Structure** as:

$$\hat{\mathbf{X}} = \{\hat{\mathbf{x}}_i | \hat{\mathbf{x}}_i = \mathbf{x}_i + \sum_{j=1}^{3} k_j l_j\};\ \hat{\mathbf{A}} = \{\hat{\mathbf{a}}_i | \hat{\mathbf{a}}_i = \mathbf{a}_i\}$$

Material Property Prediction

Input Material Structure

Neural Network (GNN, Transformer ..)

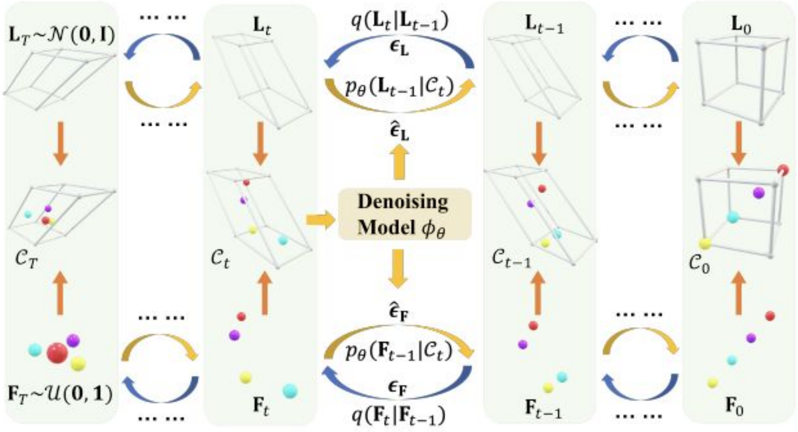


Predicted Property

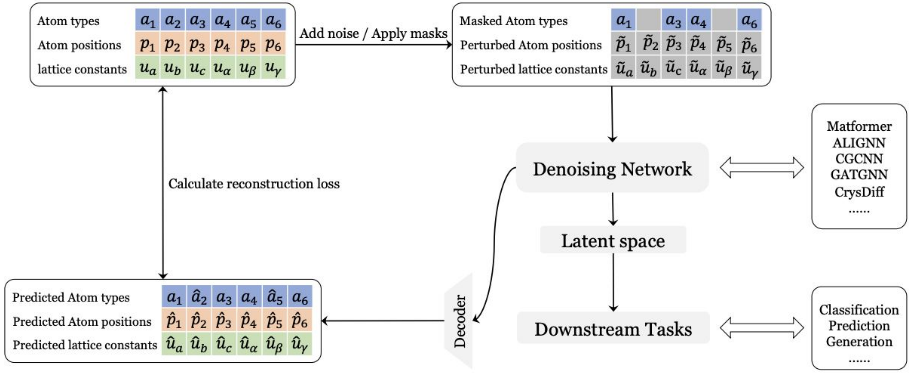
Limitations of Existing Works

- ❑ Scarcity of **Labeled Data**
- ❑ **DFT Error Bias**
- ❑ Lack of **Pre-trained Graph Model**
- ❑ Limited Capture of **Global Structural Knowledge**

Diffusion based Pre Training Frameworks



CrysDiff (Song et al., 2024)[1]



DPF (Shen et al., 2025)[2]

[1] Song, Zixing, Ziqiao Meng, and Irwin King. "A diffusion-based pre-training framework for crystal property prediction." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 38. No. 8. 2024.

[2] Shen, Shuaike, et al. "A denoising pre-training framework for accelerating novel material discovery." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 39. No. 27. 2025.

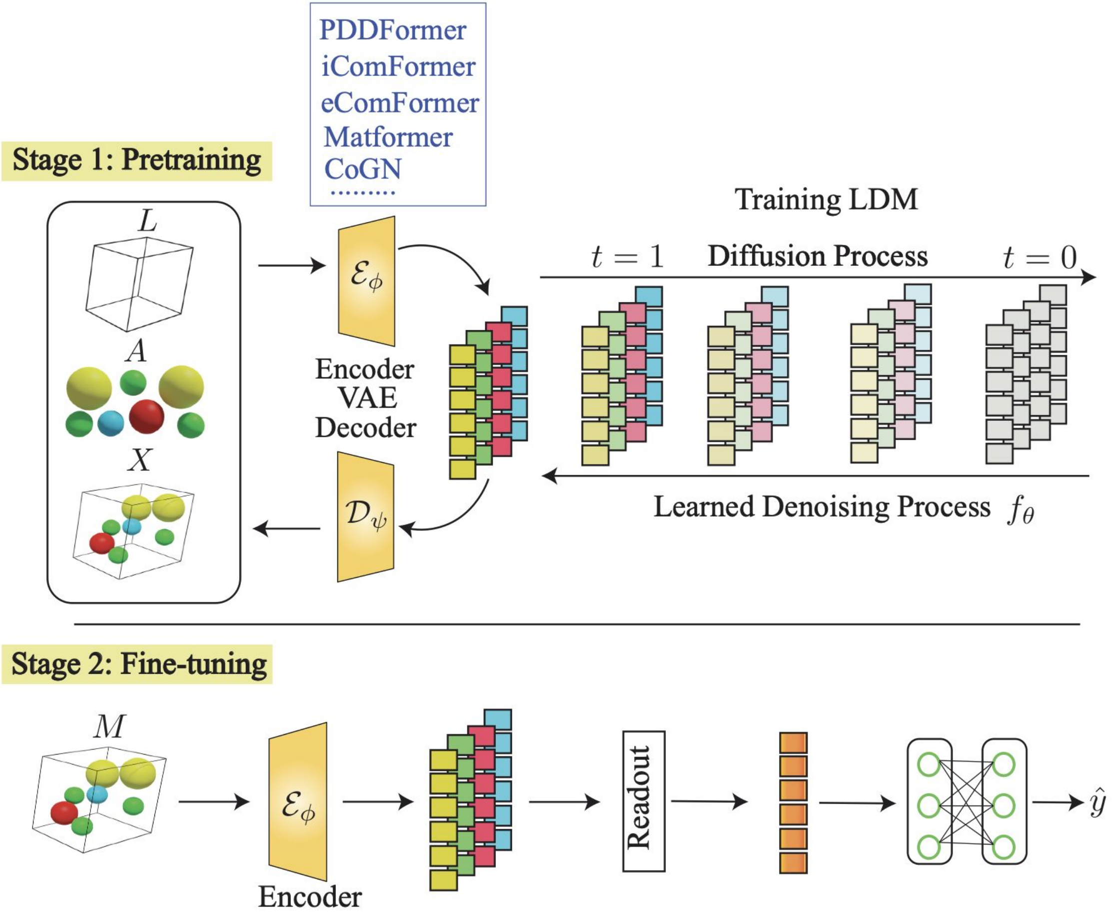
Diffusion based Pre Training Frameworks

- ❏ Operates directly on **Heterogeneous, High-Dimensional** Feature Space
- ❏ Different components require **Different Diffusion Formulations.**
  - ➢ Atom types → Discrete Diffusion (D3PM)
  - ➢ Fractional coordinates → Score-based Diffusion
  - ➢ Lattice parameters → DDPM
- ❏ **Complex Denoising Architectures** and **large numbers of diffusion steps.**
- ❏ learned representations are often **less expressive and suboptimal**.

Latent Diffusion based Pre Training Frameworks

- **Latent diffusion–based pretraining framework**
- learn robust and expressive crystal representations from large-scale **unlabeled structural data**
- overcome the **scarcity of labeled property data**
- Perform diffusion in a **smooth, lower-dimensional latent space** — via a VAE encoder.
- Richer and more transferable representations than applying diffusion directly in the **heterogeneous high-dimensional feature space** of atom types, coordinates, and lattice parameters?
- Can the pretrained representations, when finetuned with **limited experimental data**, effectively correct **DFT error bias** and maintain strong performance across diverse downstream property prediction tasks in **low-data regimes**?

CrysLDNet: Latent Diffusion Pre Training



Results - Downstream Task Evaluation

	Model	JARVIS-DFT (Choudhary et al., 2020)									Materials Project (Chen et al., 2019)			
		Formation Energy	Bandgap (OPT)	Total Energy	Ehull	Bandgap (MBJ)	Bulk Modulus	Shear Modulus	SLME (%)	Spillage	Formation Energy	Bandgap (OPT)	Bulk Modulus	Shear Modulus
Supervised Models (Train-from-scratch)	CGCNN	0.063	0.200	0.078	0.170	0.410	14.47	11.75	8.022	0.454	0.031	0.292	0.047	0.077
	SchNet	0.045	0.190	0.047	0.140	0.430	13.25	11.12	7.431	0.409	0.033	0.345	0.066	0.099
	MEGNet	0.047	0.145	0.058	0.084	0.340	14.20	12.25	7.213	0.445	0.030	0.307	0.060	0.099
	GATGNN	0.047	0.170	0.056	0.120	0.510	14.32	12.48	7.504	0.431	0.033	0.280	0.045	0.075
	CoGN	0.027	0.122	0.029	0.047	0.264	9.382	8.982	4.546	0.367	0.050	0.204	0.046	0.070
	DimeNet++	0.059	0.239	0.074	0.142	0.394	10.50	10.00	5.291	0.374	0.049	0.392	0.041	0.068
	Equiformer	0.191	0.265	0.486	0.286	0.649	12.54	14.77	6.133	0.361	0.405	0.565	0.055	0.075
	ALIGNN	0.033	0.142	0.037	0.076	0.310	10.40	9.481	5.146	0.389	0.022	0.218	0.051	0.078
	Matformer	0.033	0.137	0.035	0.064	0.300	11.21	10.76	5.260	0.398	0.021	0.211	0.043	0.073
	PotNet	0.029	0.127	0.032	0.055	0.270	10.11	9.232	4.570	0.361	0.019	0.204	0.040	0.065
	eComFormer	0.028	0.124	0.032	0.047	0.282	10.79	9.826	4.610	0.373	0.018	0.202	0.042	0.073
	iComFormer	0.027	0.122	0.029	0.044	0.261	9.617	9.098	4.583	0.360	0.018	0.193	0.038	0.064
	PDDFormer	0.027	0.120	0.028	0.033	0.251	9.546	8.808	4.300	0.358	0.016	0.189	0.034	0.062
Pretrain–Finetune	CrysXPP	0.062	0.190	0.072	0.139	0.378	13.61	11.20	5.110	0.363	0.034	0.269	0.055	0.084
	Crystal Twins	0.042	0.160	0.050	0.132	0.374	13.41	11.18	4.967	0.393	0.034	0.269	0.051	0.082
	CrysGNN	0.056	0.183	0.069	0.130	0.371	13.42	11.07	5.452	0.374	0.033	0.266	0.043	0.076
	CrysDiff	0.029	0.131	0.034	0.062	0.287	9.875	9.191	5.030	0.358	–	–	–	–
	DPF	0.029	0.122	0.032	0.059	0.311	10.43	9.596	5.129	0.358	0.020	0.203	0.042	0.073
	CrysLDNet	**0.026**	**0.118**	**0.027**	**0.032**	**0.238**	**8.817**	**8.428**	**4.120**	**0.340**	**0.015**	**0.184**	**0.032**	**0.059**

Table 1. Summary of MAE results for various properties on JARVIS-DFT (left block) and Materials Project (right block). For CrysDiff on MP, results are unavailable and shown as "–". Best and second-best are in bold and underlined, respectively.

Results - Limited Training Setup

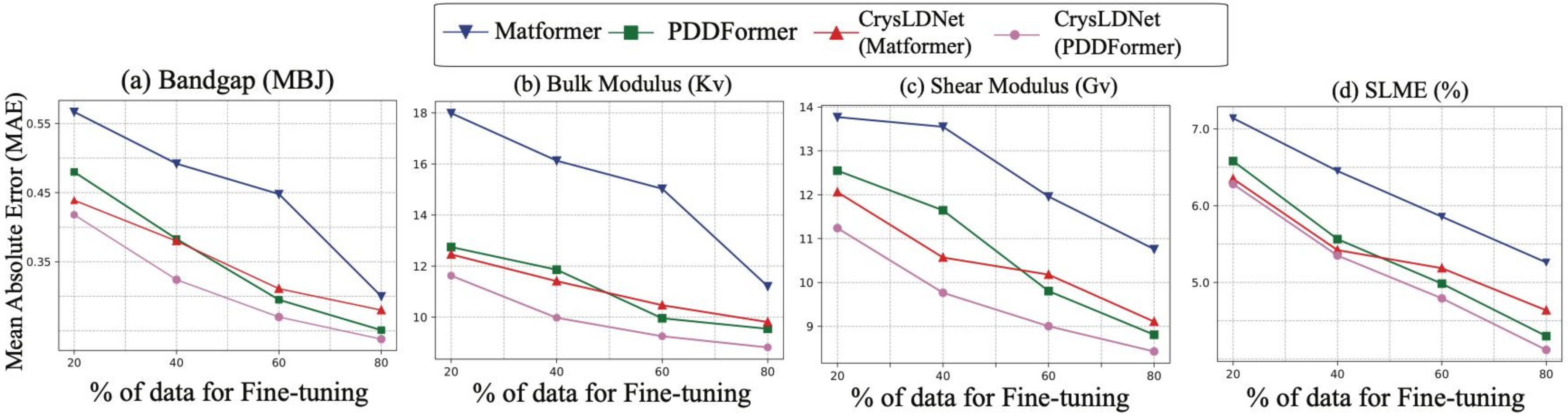


Figure 2. Performance comparison (MAE) under limited training data. MAE on four JARVIS properties using 20%, 40%, 60%, and 80% of finetuning data, comparing supervised baselines (PDDFormer, Matformer) with their corresponding CrysLDNet variants.

Conclusion

- ❑ CrysLDNet pretrain **crystal encoders** via **latent diffusion** — no labels needed, state-of-the-art properties.
- ❑ CrysLDNet combines a **VAE with latent diffusion** pretraining to learn transferable crystal representations from unlabeled data
- ❑ CrysLDNet outperforms all baselines on **Materials Project and JARVIS-DFT — especially with limited labeled data**

Github Repo:
https://github.com/shrimonmuke0202/CrysLDNet.git