Supplementary Information: Scalable Crystal Structure Relaxation Using an Iteration-Free Deep Generative Model with Uncertainty Quantification

Ziduo Yang[1,2†], Yi-Ming Zhao[1†], Xian Wang[3], Xiaoqing Liu[1], Xiuying Zhang[1], Yifan Li[1], Qiujie Lv[1,2], Calvin Yu-Chian Chen[4,5,6,7,8*] and Lei Shen[1,9*]

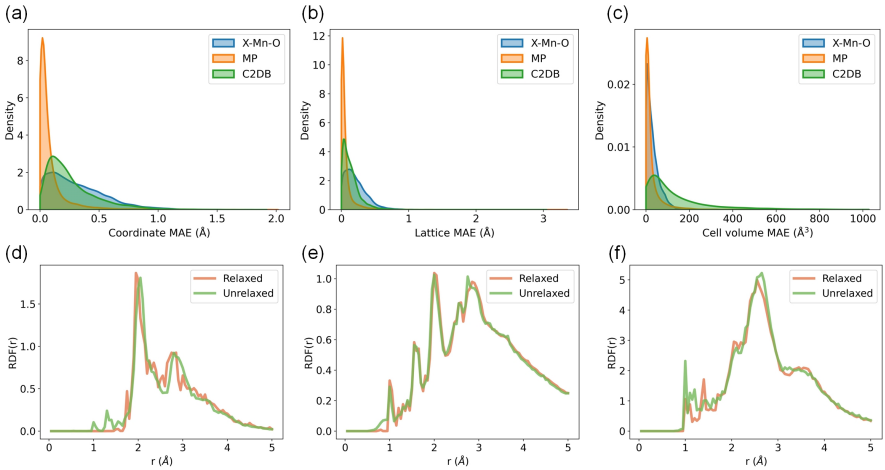
[1]Department of Mechanical Engineering, National University of Singapore, Singapore, 117575, Singapore.
[2]Artificial Intelligence Medical Research Center, School of Intelligent Systems Engineering, Shenzhen Campus of Sun Yat-sen University, Shenzhen, 518107, China.
[3]Department of Physics, National University of Singapore, Singapore, 117551, Singapore.
[4]AI for Science (AI4S)-Preferred Program, School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, Shenzhen, 518055, China.
[5]State Key Laboratory of Chemical Oncogenomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen, 518055, China.
[6]Department of Medical Research, China Medical University Hospital, Taichung, 40447, Taiwan.
[7]Department of Bioinformatics and Medical Engineering, Asia University, Taichung, 41354, Taiwan.
[8]Guangdong L-Med Biotechnology Co., Ltd, Meizhou, Guangdong, 514699, China.
[9]National University of Singapore (Chongqing) Research Institute, Chongqing, 401123, China.

*Corresponding author(s). E-mail(s): cy@pku.edu.cn; shenlei@nus.edu.sg;

[†]These authors contributed equally to this work.

Suppl. Note 1. Dataset distribution

Suppl. Fig. 1(a)-(c) display the distributions of mean absolute errors (MAE) for coordinates, lattice parameters, and cell volumes, comparing initial and final structures across the X-Mn-O, MP, and C2DB datasets. These distributions highlight considerable structural differences within the X-Mn-O dataset, minimal differences within the MP dataset, and moderate differences within the C2DB dataset. Suppl. Fig. 1(d)-(f) show the radial distribution functions (RDFs) for each dataset, indicating narrow interatomic distance ranges in the X-Mn-O dataset, broad ranges in the MP dataset, and intermediate ranges in the C2DB dataset.

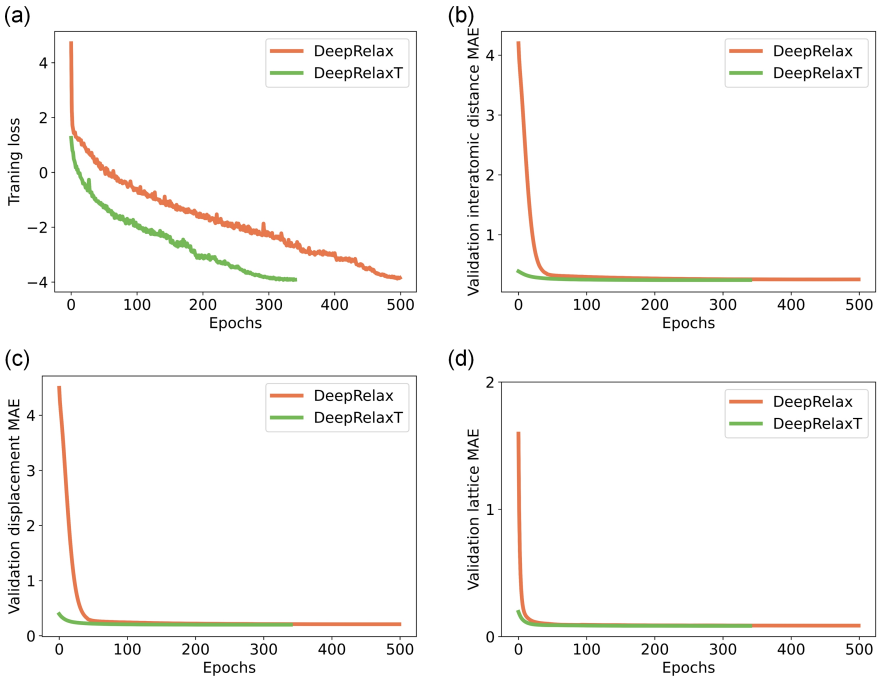


Suppl. Fig. 1 Deviation distribution and radial distribution function (RDF) analysis for X-Mn-O, MP, and C2DB datasets. The deviation distribution is evaluated using (a) coordinate MAE (Å), (b) lattice MAE (Å), and (c) cell volume MAE (Å^3) between initial and DFT-relaxed structures. RDFs are shown for the datasets: (d) X-Mn-O, (e) MP, and (f) C2DB. Source data are provided with this paper.

Suppl. Note 2. Transfer learning

Transfer learning is a technique in machine learning where a model developed for a specific task is reused as the starting point for a model on a second task. Essentially, it involves taking knowledge learned from solving one problem and applying it to a different but related problem.

We extend the application of DeepRelax, initially pre-trained on 3D materials from the MP dataset, to 2D materials through transfer learning. As depicted in Suppl. Fig. 2, DeepRelaxT exhibits a more rapid convergence rate compared to DeepRelax, effectively highlighting the efficacy of the transfer learning approach.



Suppl. Fig. 2 Training and validation metrics for DeepRelax and DeepRelaxT models over epochs. (a) Training loss. (b) Validation MAE of interatomic distance (Å). (c) Validation MAE of displacement (Å). (d) Validation MAE of lattice parameters (Å). DeepRelaxT demonstrates a faster convergence rate in all metrics. Source data are provided with this paper.

Suppl. Table 1 Inter-layer distances (Å) for unrelaxed, DFT-D3-relaxed, and DeepRelax-predicted structures of six vdW crystals from the test set.

Sample ID	1	2	3	4	5	6
Unrelaxed	3.600	3.800	4.120	3.800	3.960	3.600
DFT-D3-relaxed	3.147	3.078	3.855	3.175	2.875	2.794
DeepRelax-predicted	3.143	3.097	3.600	3.263	2.830	2.804

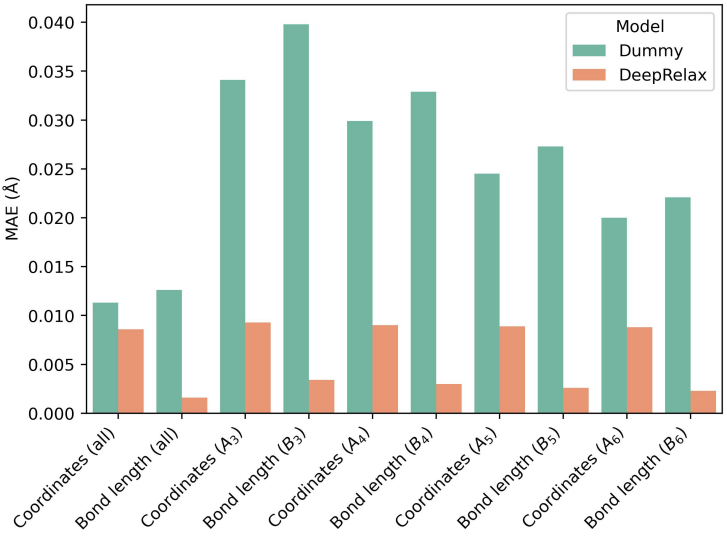
Suppl. Note 3. Application in layered vdW crystals

Suppl. Table 1 shows the inter-layer distances for the unrelaxed, DFT-D3-relaxed, and DeepRelax-predicted structures within the test set. The inter-layer distances of the predicted structures closely match those of the relaxed structures, highlighting the effectiveness of transferred DeepRelax on layered vdW crystals. Further analysis is conducted on the bond length MAE for representative bonding pairs. We define Δ as the bond length difference between the unrelaxed and relaxed structures. Thresholds are set at 0.1 Å,

Suppl. Table 2 Bond length MAE (Å) for representative bonding pairs, with thresholds set at 0.1 Å, 0.2 Å, 0.3 Å, 0.4 Å, and 0.5 Å, filtering out 67%, 77%, 84%, 85%, and 88% of the less changing bonds during relaxation, respectively. The best performance in each metric is highlighted in bold.

Model	$\Delta > 0.1$	$\Delta > 0.2$	$\Delta > 0.3$	$\Delta > 0.4$	$\Delta > 0.5$
Dummy	0.493	0.535	0.685	0.707	0.794
DeepRelax	**0.152**	**0.163**	**0.146**	**0.144**	**0.164**

0.2 Å, 0.3 Å, 0.4 Å, and 0.5 Å. Bonding pairs where Δ exceeds these thresholds are filtered out, effectively removing 67%, 77%, 84%, 85%, and 88% of the bonding pairs that show lesser changes during relaxation, respectively. As shown in Suppl. Table 2, transferred DeepRelax achieves lower bond length MAE across all thresholds compared to the Dummy model, demonstrating its precision in predicting structural changes in layered vdW crystals.



Suppl. Fig. 3 Comparative results of DeepRelax and the Dummy model on defect calculations. The performances are evaluated using the MAE in angstroms (Å) for all-atom coordinates, all-bond lengths, local-atom coordinates, and local-bond lengths. Source data are provided with this paper.

Suppl. Note 4. Application in crystals with defects

Given that structural variations between unrelaxed and relaxed defect structures are mainly localized near the defect sites, we adopt localized MAE

statistics for a more precise evaluation of DeepRelax's performance. Specifically, we use A_x to denote atoms within an x Å radius of the defect sites and B_x for the associated bonds. We set x to 3, 4, 5, and 6 in our experiment. As lattice parameters in superlattices with low defect concentrations remain almost constant during relaxation, they are not predicted for this dataset. Suppl. Fig. 3 presents the experimental results from applying DeepRelax to crystal structures with defects in the test set, which demonstrate a notably lower MAE in both atom coordinates and bond lengths for DeepRelax compared to the Dummy model, thereby underscoring DeepRelax's efficacy in defect structure calculations.

Suppl. Note 5. Comparison with M3GNet

In our study, we further evaluate DeepRelax against M3GNet, a leading iterative ML model. Both models are initially pre-trained on the MPF.2021.2.8 dataset and subsequently assessed using the X-Mn-O dataset. For M3GNet, we utilized the materials graph library and a readily available pre-trained model https://github.com/materialsvirtuallab/matgl. It's important to note that DeepRelax requires only the initial structure for input, whereas M3GNet is trained using complete snapshots, including initial, intermediate, and final structures.

Suppl. Table 3 presents the comparative results. Our observations are twofold: Firstly, compared with the results shown in Table 1 (main text), DeepRelax exhibits relatively lower performance when trained on the MPF.2021.2.8 dataset. This is probably due to the data sparsity in this dataset, especially compared to the X-Mn-O dataset, which contains a substantially higher average number of structures per composition (about 66.6 in the X-Mn-O dataset vs. about 1.2 in the Materials Project dataset) [1]. Secondly, despite utilizing only about a third of the samples compared to M3GNet, DeepRelax still demonstrates competitive performance. More importantly, DeepRelax offers a substantial speed advantage, being approximately 100 times faster than M3GNet. Additionally, the efficiency of DeepRelax can be further enhanced through the implementation of parallel processing.

Suppl. Table 3 Comparative results of pre-trained DeepRelax and M3GNet on the X-Mn-O dataset. The performances are evaluated by the MAE of coordinates (Å), bond length (Å), lattice (Å), cell volume (Å^3) between the predicted and DFT-relaxed structures. The average run time per structure is provided in seconds (s). The best performance in each metric is highlighted in bold.

Model	#Sample	Coordinates	Bond length	Lattice	Cell volume	Time*
Dummy	-	0.314	0.429	0.221	32.8	-
M3GNet	187,687	**0.240**	-	0.227	**9.9**	16.102
DeepRelax	62,724	0.269	**0.325**	**0.167**	14.7	**0.166**

*We measure the time consumption of M3GNet and DeepRelax using NVIDIA RTX A6000.

Suppl. Note 6. Difference between DeepRelax and other direct ML models

Two ML models, DOGSS [2] and Cryslator [1], have been published for predicting relaxed crystal structures directly. Here, we outline the differences and advantages of our DeepRelax model over these two as follows:

DOGSS integrates differentiable optimization with graph neural networks to predict relaxed structures. It learns a simple harmonic force field, utilizing a GNN to predict the parameters: pairwise interatomic distances in the relaxed structure $\boldsymbol{r}_e \in \mathbb{R}^{N\times M\times 1}$, and spring constants $\boldsymbol{k} \in \mathbb{R}^{N\times M\times 1}$ for springs connecting atom pairs. Here, N represents the number of atoms, and M denotes the number of nearest neighbor atoms connected by distinct edges. Once these parameters are learned, the model constructs the harmonic potential energy surface (PES) as

$$\hat{E}(\boldsymbol{r}_o(\boldsymbol{x}_o), \boldsymbol{r}_e, \boldsymbol{k}) = \sum_i \sum_j k_{i,j}(r_{o,(i,j)}(\boldsymbol{x}_o) - r_{e,(i,j)})^2 \tag{1}$$

where $\boldsymbol{r}_o(\boldsymbol{x}_o) \in \mathbb{R}^{N\times M\times 1}$ represents pairwise interatomic distances in the unrelaxed structure. DOGSS solves for optimal structures by minimizing this energy:

$$\boldsymbol{x}_{\text{pred}} = \arg \min_{\boldsymbol{x}_o \notin \boldsymbol{x}_{\text{con}}} \hat{E}(\boldsymbol{r}_o(\boldsymbol{x}_o), \boldsymbol{r}_e, \boldsymbol{k}), \tag{2}$$

where $\boldsymbol{x}_{\text{pred}}$ is a set of atomic positions at local minima in the harmonic PES, and $\boldsymbol{x}_{\text{con}}$ is a set of positions of constrained atoms that are not allowed to move. This method's limitation lies in its inability to predict correct structures in systems without constrained atoms, due to potentially infinite solutions that satisfy the predicted interatomic distances. In contrast, DeepRelax can be applied to systems where all atoms are allowed to move, as it predicts both interatomic distances and displacements simultaneously. This capability, combined with DeepRelax's equivariant representation, explicit handling of periodicity, uncertainty quantification, and ability to predict lattice parameters, significantly extends its applicability and accuracy compared to DOGSS.

Cryslator conceptualizes crystal relaxation as a structure translation task and developed a generative domain translation framework that maps unrelaxed structural domains to the relaxed domains. This model requires training several networks independently, increasing its complexity. Furthermore, the Cryslator's developers argued that Cyslator exhibits low prediction performance on the Materials Project dataset as reported in its original paper [1]. In contrast, DeepRelax maintains its effectiveness on the complex MP dataset (see Table 2 in the main text), demonstrating robust performance across various test scenarios.

Supplementary references

[1] Kim, S., Noh, J., Jin, T., Lee, J., Jung, Y.: A structure translation model for crystal compounds. npj Computational Materials **9**(1), 142 (2023)

[2] Yoon, J., Ulissi, Z.W.: Differentiable optimization for the prediction of ground state structures (dogss). Physical Review Letters **125**(17), 173001 (2020)