

# Rethinking Crystal Symmetry Prediction: A Decoupled Perspective

Liheng Yu<sup>1</sup>, Zhe Zhao<sup>1,3</sup>, Xucong Wang<sup>1</sup>, Di Wu<sup>1\*</sup>, Pengkun Wang<sup>1,2\*</sup>

<sup>1</sup> University of Science and Technology of China, Hefei 230236, China

<sup>2</sup> Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou 215123, China

<sup>3</sup> City University of Hong Kong, Hong Kong

{yuliheng, zz4543, xuco, wdcxy}@mail.ustc.edu.cn, pengkun@ustc.edu.cn,

## Abstract

Efficiently and accurately determining the symmetry is a crucial step in the structural analysis of crystalline materials. Existing methods usually mindlessly apply deep learning models while ignoring the underlying chemical rules. More importantly, experiments show that they face a serious sub-property confusion (SPC) problem. To address the above challenges, from a decoupled perspective, we introduce the XRDecoupler framework, a problem-solving arsenal specifically designed to tackle the SPC problem. Imitating the thinking process of chemists, we innovatively incorporate multidimensional crystal symmetry information as super-class guidance to ensure that the model's prediction process aligns with chemical intuition. We further design a hierarchical PXRD pattern learning model and a multi-objective optimization approach to achieve high-quality representation and balanced optimization. Comprehensive evaluations on three mainstream databases (e.g., CCDC, CoREMOF, and InorganicData) demonstrate that XRDecoupler excels in performance, interpretability, and generalization.

**Code** — <https://github.com/baigeiguai/XRDecoupler>

**Extended version** — <https://arxiv.org/abs/2511.06976>

## Introduction

Symmetry determination is a crucial step in powder X-ray diffraction (PXRD) based crystal structure prediction (Altomare et al. 2004; Spence and Zuo 1992). The space group (Koster 1957) defines the symmetry characteristics of the crystal, including rotation, reflection, and inversion operations. These symmetries are fundamental to understanding the crystal structure and its properties (O’Keeffe and Hyde 2020; Bhagavantam and Suryanarayana 1949). Once the space group is established, scientists can build and optimize crystal structure models based on that symmetry (Han et al. 2025; Evans 2011). Incorrect selection of space groups can lead to inaccurate or unreasonable structural models. Therefore, *efficiently and accurately identifying space group types remains a significant challenge.*

As an early typical approach, many researchers attempted to apply a simple neural network for symmetry recognition

\*Corresponding author.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

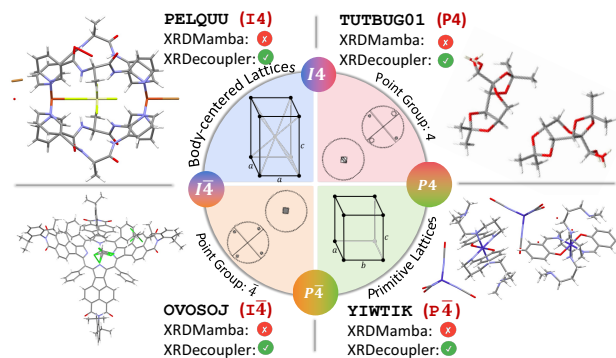


Figure 1: The SPC problem in the space group identification. Different colored blocks represent various symmetry classification systems, such as lattice types and point group types. We illustrate four space groups that current methods often confuse:  $I4$ ,  $I\bar{4}$ ,  $P4$ , and  $P\bar{4}$ . These space groups are intertwined and may belong to a coarser classification. We also present four representative crystal samples, i.e., PELQUU, OVOSOJ, TUTBUG01, and YIWTIK, demonstrating the capability of our method to decouple these confusions.

within specific ranges, advancing the application of machine learning in this task (Park et al. 2017; Ziletti et al. 2018; Oviedo et al. 2019; Vecsei et al. 2019; Dong et al. 2021). However, these methods often remain limited to a single space group or a small subset of space groups, or specific material categories, which restricts their applicability and prevents broader generalization across different materials or space groups. More importantly, simply imposing deep learning models on this task does not align with chemical intuition. Although the latest methods attempt to enhance model architectures from a theoretical perspective of pattern (Yu et al. 2024), they still face a challenging dilemma: *models tend to mistake two crystalline materials with common sub-properties (such as lattice types) for belonging to the same space group (as shown in Figure 1), i.e., the sub-property confusion (SPC) problem.*

To clarify the essence of SPC, we systematically rethought existing methods and found that the causes of SPC are multifaceted: ❶ Previous methods directly process PXRD patterns in a conventional sequential manner, lack-

ing the necessary chemical knowledge to guide the process. ② The optimization direction of these models often tends to favor latent sub-properties. ③ The information overlap among sub-properties makes it difficult for the model to distinguish between closely related structures. To illustrate the SPC problem faced in space group prediction tasks more intuitively, we provide a clear diagram, as shown in Figure 1. Based on crystallography, the space groups  $I4$ ,  $I\bar{4}$ ,  $P4$ , and  $P\bar{4}$  are categorized into two lattice types: I (body-centered) and P (primitive). They are also divided into two point group types: 4 (fourfold rotation axis) and  $\bar{4}$  (fourfold roto-inversion axis). As a result, these four space groups share common sub-properties, making them particularly prone to confusion during classification. We also present the recognition results of different methods on these easily confused samples. *The SOTA model* (XRDMamba (Yu et al. 2024)) **appears to be hindered by the SPC, while our method successfully predicts the space group types for each crystal.**

Our method, *XRDecoupler*, is a problem-solving arsenal specifically designed to tackle the SPC problem in space group prediction. We rethink SPC faced in previous research according to the thinking process of chemists and innovatively incorporate crystal systems, Bravais lattice types, and point groups as superclasses guidance, ensuring that the model’s predictions align with chemical rules and enhancing its ability to distinguish between easily confused samples (► **solving Cause ①**). Additionally, different superclasses may focus on different aspects of the PXRD pattern. To address this, we propose a hierarchical PXRD pattern learning model that captures both local and global pattern information inherent in PXRD patterns, enabling efficient multi-superclass learning (► **solving Cause ③**). Furthermore, we utilize a multi-objective optimization approach to ensure that the model does not favor specific superclass tasks at the expense of others, guiding it toward an optimal training path (► **solving Cause ②**). To validate the effectiveness of XRDecoupler, we conducted comprehensive experimental evaluations on the well-known Cambridge Crystallographic Data Centre (CCDC) database (Allen et al. 1979), CoREMOF (Chung et al. 2019), and InorganicData (Salgado et al. 2023). The results indicate that XRDecoupler significantly outperforms other state-of-the-art baselines.

Our contributions are summarized as follows:

- ① **Crucial Problem and Fresh Perspective:** For the first time, we rethink existing methods and analyze the SPC problem faced in space group prediction tasks. From a decoupled perspective, we introduce the XRDecoupler framework to decouple it.
- ② **Novel Mechanism:** Imitating the thinking process of chemists, we innovatively incorporate multidimensional crystal symmetry information as superclasses guidance, ensuring predictions align with chemical intuition.
- ③ **Reasonable Pattern Learner:** We proposed a hierarchical PXRD Pattern learning model that explicitly models the local and global pattern information in PXRD.
- ④ **Brilliant Performance:** Evaluations on well-known databases demonstrated the superior performance, interpretability, and generalization of XRDecoupler.

## Related Works

► **Crystalline Space Group Prediction.** Identifying crystal space groups is vital for structure prediction. While early work relied on computational methods (Werner, Eriksson, and Westdahl 1985), recent efforts have focused on deep learning. Initial convolutional neural networks showed promise but were often trained or tested on limited datasets (Park et al. 2017; Ziletti et al. 2018; Oviedo et al. 2019; Vecsei et al. 2019; Dong et al. 2021). Subsequent models like NPCNN (Salgado et al. 2023) utilized more comprehensive data but struggled with accuracy. In response, RCNet (Chen et al. 2024) improved performance by customizing categories and using residual structures, while XRDMamba (Yu et al. 2024) pioneered the integration of chemical knowledge and the Mamba architecture (Gu and Dao 2023; Gu 2023). However, these methods often neglect relevant chemical principles and suffer from high confusion between similar space groups. Our approach addresses this by designing a hierarchical framework and introducing multidimensional superclass knowledge to guide model optimization, thereby mitigating confusion and enhancing generalization.

► **Superclass Learning.** Superclass learning improves model performance by incorporating high-level class groupings as intermediate supervision. This technique has proven effective across diverse domains, such as improving feature differentiation in image classification (Dehkordi et al. 2022; Wang et al. 2022), mitigating data imbalance (Zhou, Hu, and Wang 2018; Zhao et al. 2025; Wang et al. 2024; Zhao et al. 2024), and learning high-level relationships in graph neural networks (Du et al. 2023). The benefits of this approach are well-documented. Given the inherent hierarchical classification of space groups, we introduce superclass learning to this task for the first time. This guides our model to learn more detailed structural knowledge about crystals, significantly enhancing its recognition performance. A more comprehensive review is provided in [Supplement H](#).

## Motivation

► **Space Group Prediction Task.** The PXRD data can be viewed as two vectors,  $A$  and  $I$ , both of length  $L$ , representing the diffraction angles and diffraction intensities on the PXRD pattern, respectively. In crystallography, the space group is a description of the symmetry of crystals, with a total of 230 theoretically existing space groups. We define  $D = \{X_i, Y_i\}_{i \in [n]}$  as a crystal dataset, where  $X_i = (A_i, I_i)$  represents the PXRD pattern data of a crystal, and  $Y_i \in \{0, 1, \dots, 229\}$  denotes the space group label of that crystal. We represent the process of space group prediction as a mapping relationship:  $f(X_i) \rightarrow Y_i$ . We provide definitions about space groups in [Supplement A.1](#) for understanding.

► **Multidimensional Symmetry as Superclasses.** The superclass refers to a more general class in a hierarchical classification structure that contains other subclasses. For crystal structures, aside from space groups, there are many coarse-grained partition rules (Nespolo, Aroyo, and Souvignier 2018). As shown in [Figure 1](#) in [Supplement A](#), based on the symmetry of geometric morphology, crystals can be

divided into 7 crystal systems (O’Keeffe and Hyde 2020); based on primitive point symmetry, they can be classified into 32 point groups (Bradley and Cracknell 2009); according to the point symmetry of Bravais lattices, there are 7 lattice systems; based on the spatial symmetry of Bravais lattices, there are 14 types of Bravais lattices (Pitteri and Zanzotto 1996); and based on combinations of point symmetry and translational symmetry, they can be classified into 73 algebraic crystal classes (Wilson 1990), among others (Hahn, Shmueli, and Arthur 1983; Prince 2004). *The partition rules mentioned above can all be viewed as superclasses of space groups, describing only the symmetric properties of a certain part of the crystal.* We also provide an explanation of the superclass mechanism in [Supplement A.3](#).

### Rethinking ‘Culprits’ Behind SPC

► **Culprit 1: Lack of Chemical Knowledge.** Existing models are typically adapted directly from other deep learning tasks, predominantly utilizing rudimentary convolutional neural networks. Although these models have yielded initial results, they often fall short in enhancing performance further. *The current model setup fails to consider chemical knowledge integral to material structure analysis, making it challenging for the model to genuinely comprehend the intricate details of the structural patterns within the data.* Confusion often arises when the model only captures the coarse-grained aspects of the data. Hence, a model tailored with integrated chemical knowledge becomes indispensable to enhance the efficacy of space group identification.

► **Culprit 2: Bias in Optimization.** Successful space group prediction requires accurately identifying the sample’s sub-symmetry properties (e.g., lattice, point, and translation symmetry). During optimization, supervisory signals for each property must be transmitted to guide learning. *When the unique space group label is the sole supervisory signal, it can cause insufficient signals and imbiased gradients across sub-properties, leading the model to learn only some of them effectively.* Consequently, the model struggles to distinguish between samples sharing these learned sub-properties but belonging to different space groups, causing confusion. For example, Figure 2 shows our analysis of a SOTA method’s misclassifications and per-sub-attribute performance. The model evidently learned lattice types and crystal systems well but performed poorly on point group characteristics. Thus, *the model exhibits significant confusion with samples that share the same lattice type and crystal system but differ in their point groups.*

► **Culprit 3: Information Overlap Between Labels.** The space group category serves as a fine-grained structural classification criterion, encompassing numerous sub-properties of crystal structures. For two distinct categories, they often share certain structural sub-properties, a phenomenon we refer to as information overlaps between labels. In the following, we will analyze how this information overlap can lead to confusion from the perspective of information theory (Ash 2012; Batina et al. 2011).

For a given sample  $X$ , we denote its true confidence label as  $Y_{truth}$  and any other space group class label as  $Y_{other}$ . In

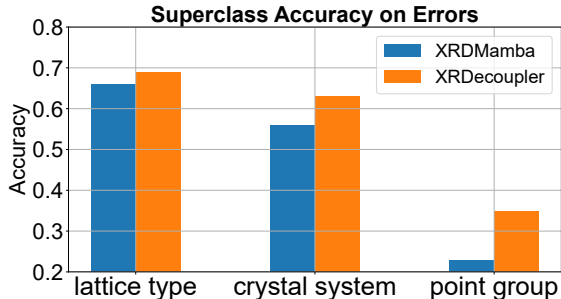


Figure 2: Accuracy statistics of samples with misidentified space groups on three sub-properties (e.g., lattice type, crystal system, and point group) on the SOTA method (XRDMamba (Yu et al. 2024)) and our proposed XRDecoupler.

this context, for sample  $X$ , label  $Y_{other}$  represents an incorrect label.

**Assumption 1.** *For the space group class label  $Y$ , it can be uniquely represented by  $k$  independent structure sub-properties, i.e.,  $Y = (y_1, y_2, \dots, y_i, \dots, y_k)$ , where  $y_i$  denotes  $i$ -th sub-properties.*

**Proposition 1.** *For  $Y_{truth} = (y_1^{truth}, y_2^{truth}, \dots, y_k^{truth})$  and  $Y_{other} = (y_1^{other}, y_2^{other}, \dots, y_k^{other})$ , there are some same structure sub-properties. That is, there exists  $s, t$ , such that*

$$\begin{cases} y_{p_d}^{truth} = y_{p_d}^{other}, d = [1, 2, \dots, s] \\ y_{q_d}^{truth} \neq y_{q_d}^{other}, d = [1, 2, \dots, t] \end{cases} \quad (1)$$

*s.t.  $s + t = k, \{[p_i], [q_i]\} = \{1, 2, 3, \dots, k\}$*

For simplicity, let  $M = \{y_{p_d}^{truth}\}, d \in [1, 2, \dots, s]$  represent the overlap information between  $Y_{truth}$  and  $Y_{other}$ . Let  $y_{truth} = \{y_{q_d}^{truth}\}, d \in [1, 2, \dots, t]$  and  $y_{other} = \{y_{q_d}^{other}\}, d \in [1, 2, \dots, t]$  denote the non-overlapping information between the two labels. Then, we have  $Y_{truth} = (M, y_{truth})$  and  $Y_{other} = (M, y_{other})$ . Usually, our aim is to have the model maximize the mutual information  $I(E; Y_{truth})$  between sample representation and truth label during training, where  $E$  denotes the representation obtained from the encoder.

**Proposition 2.** *If there is significant overlap information between labels, maximizing the mutual information between sample representation and truth label will result in the mutual information of sample representation and overlap label being maximized as well, specifically:*

$$\begin{aligned} I(E; Y_{truth}) \uparrow &\Rightarrow I(E; Y_{other}) \uparrow \\ \text{s.t. } I(M; Y_{truth}) &\gg I(y_{truth}; Y_{truth}). \end{aligned} \quad (2)$$

We provide a theoretical analysis in [Supplement B.1](#). Proposition 2 shows, when there is significant overlap in information between labels, simply maximizing the mutual information between the representation and the truth label is insufficient. Thus, we observe the difference between the mutual information, which serves as a reliable indicator of the model’s confusion in executing a classification.

**Definition 1** (Difference *Diff* in mutual information). *The difference *Diff* in mutual information between two labels*

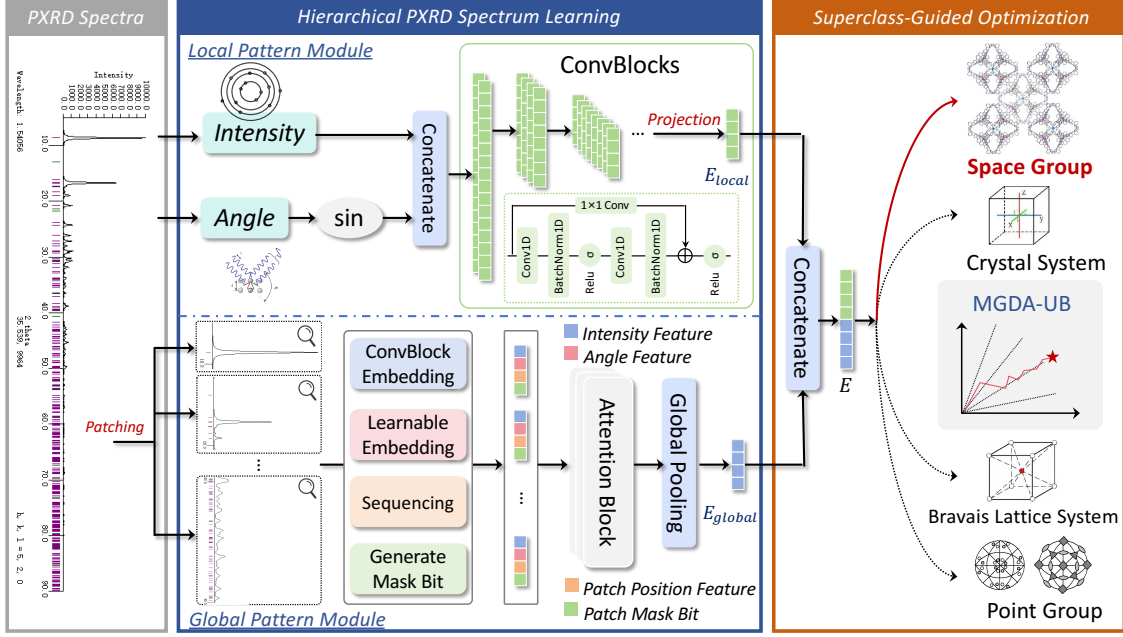


Figure 3: Overview of XRDecoupler.

is equal to the mutual information of their non-overlapping parts.

$$\begin{aligned} Diff &= I(E; Y_{truth}) - I(E; Y_{other}) \\ &= I(E; y_{truth}) - I(E; y_{other}) \end{aligned} \quad (3)$$

We provide the corresponding proof in **Supplement B.2**.

### XRDecoupler is A ‘Nemesis’ of SPC

Here, we propose a novel method for determining space groups from PXRD, called XRDecoupler. This method includes a superclass-guided optimization framework (for **Culprit 2 & 3**) and a hierarchical PXRD pattern learning model (for **Culprit 1**), effectively addressing the significant SPC issues associated with previous methods. An overview of XRDecoupler is illustrated in Figure 3.

#### Superclass-Guided Optimization (► Culprit 2 & 3)

► **Mitigating Information Overlap.** Considering significant information overlap between space group labels, we expect that the model focuses more on label-specific features and devotes less attention to overlapping information during the classification. Thus, we also aim to maximize  $Diff$ .

**Proposition 3.** *The process of maximizing  $Diff$  can be interpreted as maximizing the mutual information between the sample and each structured sub-property.*

$$\begin{aligned} Diff &= I(E; y_{truth}) - I(E; y_{other}) \\ &= \sum_{i=1}^t [I(E; y_i^{truth}) - I(E; y_i^{other})] \end{aligned} \quad (4)$$

We provide the corresponding proof in **Supplement B.3**. From this proposition, we can optimize the model’s ability to learn each sub-property to alleviate the confusion phenomenon. Let’s discuss how to optimize them.

► **Superclass Guidance.** Superclasses of space groups describe a broad property of crystal structures, which aligns with the idea of structured sub-property mentioned earlier. Therefore, it is logical to introduce various superclasses of space groups to represent the sub-properties above, and optimize the sub-properties in the same way as optimizing the superclasses. Through the oversight of these superclasses, the model’s focus on overlapping information can be minimized, thereby enhancing the efficacy of space group identification. Specifically, we focus on the  $T$  types of superclasses of space groups. We use  $C$  to represent the space group category,  $C_i^{Sup}$  to denote the  $i$ -th type of superclass category of space groups,  $y_i$  to indicate the space group category corresponding to the  $i$ -th sample, and  $y_i^{Sup_j}$  to represent the category of the  $j$ -th superclass corresponding to the  $i$ -th sample. We define  $Y_i = \{y_i\} \cup \{y_i^{Sup_j}\}_{j \in [T]}$  to represent the set of all categories to which the  $i$ -th sample belongs. Then, we assume that the classifier for space group classification is  $Cls$ , and the classifiers for the various superclasses are denoted as  $\{Cls^{Sup_t}\}_{t \in [T]}$ . Our optimization objective is given by:

$$\begin{aligned} &\min -I(E; y) - \beta \sum_{t \in [T]} I(E; y^{Sup_t}) \\ &\Leftrightarrow \min \mathbb{E}_{(X, Y)} - \log p(y|E) - \sum_{t \in [T]} \log p(y^{Sup_t}|E) \\ &\Leftrightarrow \min L_{sp}(Cls(f(X)), y) + \beta \sum_{t \in [T]} L^t(Cls^{Sup_t}(f(X)), y^{Sup_t}) \end{aligned} \quad (5)$$

where  $L_{sp}(\cdot)$  and  $L^t(\cdot)$  are the cross-entropy loss,  $f$  is an encoder satisfying  $E = f(X)$ , and  $\beta$  is a hyperparameter.

► **Example.** In the confusion cases illustrated in Figure 1, we clarify the representations of samples from the  $\mathbb{I}4$  &  $\mathbb{P}4$  categories, as well as  $\mathbb{I}\bar{4}$  &  $\mathbb{P}\bar{4}$ , by introducing the Bravais lattice type superclass as supervision. Additionally, we use the point group type superclass to differentiate between  $\mathbb{I}4$  &  $\mathbb{I}\bar{4}$ , and  $\mathbb{P}4$  &  $\mathbb{P}\bar{4}$  samples, thereby enhancing the discriminative capability of these four space groups.

► **Optimization Process.** Following the introduction of superclasses, each one provides a multitude of detailed supervision signals for the model’s optimization process. Nevertheless, as illustrated in Figure 4(left), Culprit 2 remains unresolved, with a persistent skew in the optimization process. In this process, the model’s loss on certain superclasses initially decreases, only to be followed by the gradual optimization of the remaining superclasses once these have been refined. This skewed optimization process is detrimental to enhancing the model’s generalization performance and also impacts the benefits derived from the superclasses. Therefore, we have transformed the optimization process into a multi-objective optimization process. We consider the space group and the  $T$  superclasses as  $T + 1$  independent classification objectives, thereby optimizing the model’s learning as a multi-objective optimization, i.e.,

$$\begin{aligned} \min_W \hat{L}(W^{Enc}, W_1^{Sup}, \dots, W_{T+1}^{Sup}) \\ = \min_W (L^1(W^{Enc}, W_1^{Sup}), \dots, L^{T+1}(W^{Enc}, W_{T+1}^{Sup})) \end{aligned} \quad (6)$$

where  $W^{Enc}$  represents the parameters of the encoder,  $W_t^{Sup}$  represents the parameters of the classifier for each class and  $W = \{W^{Enc}\} \cup \{W_t^{Sup}\}_{t \in [T+1]}$  represents all parameters. The goal of multi-objective optimization is achieving Pareto optimality.

**Definition 2** (Solution  $W$ ). A solution  $W$  dominates a solution  $\bar{W}$  if  $\hat{L}(W^{Enc}, W_t^{Sup}) \leq \hat{L}(\bar{W}^{Enc}, \bar{W}_t^{Sup})$  for all objectives  $t \in [T + 1]$  and  $\hat{L}(W^{Enc}, W_1^{Sup}, W_2^{Sup}, \dots, W_{T+1}^{Sup}) \neq \hat{L}(\bar{W}^{Enc}, \bar{W}_1^{Sup}, \bar{W}_2^{Sup}, \dots, \bar{W}_{T+1}^{Sup})$ .

**Definition 3** (Solution  $W^*$ ). A solution  $W^*$  is Pareto optimal if there exists no solution  $W$  that dominates  $W^*$ .

Therefore, the optimization process of space group and superclass is transformed into finding a Pareto optimal solution  $W^* = (W^{Enc*}, W_1^{Sup*}, \dots, W_t^{Sup*})$ . Inspired by MGDA-UB(Sener and Koltun 2018), we further transform the optimization problem into solving a set of weights  $\{\alpha_i\}$ :

$$\begin{aligned} \min_{\alpha_1, \alpha_2, \dots, \alpha_{T+1}} \left\| \sum_{t=1}^{T+1} \alpha_t \nabla_{W^{Enc}} L^t(W^{Enc}, W_t^{Sup}) \right\|_2^2, \\ \text{s.t.} \quad \sum_{t=1}^{T+1} \alpha_t = 1, \alpha_t \geq 0 \quad \forall t. \end{aligned} \quad (7)$$

where  $\alpha_t$  denotes the weight of the  $t$ -th objective.

Then, we use the FRANK-WOLFESOLVER algorithm (FW) to obtain a gradient direction that improves all  $T + 1$

classification tasks. The optimization process is as follows.

$$\begin{aligned} (1) W_t^{Sup} &= W_t^{Sup} - \eta \nabla_{W_t^{Sup}} L(W^{Enc}, W_t^{Sup}) \forall t \in [T + 1] \\ (2) \alpha_1, \dots, \alpha_t &= FW(W^{Enc}, \{W_t^{Sup}\}_{t \in [T+1]}) \\ (3) W^{Enc} &= W^{Enc} - \eta \sum_{t=1}^{T+1} \alpha_t \nabla_{W^{Enc}} L^t(W^{Enc}, W_t^{Sup}) \end{aligned} \quad (8)$$

where  $\eta$  represents the learning rate.

## Hierarchical PXR D Pattern Learning (► Culprit 1)

The introduction of superclasses brings a wealth of structural knowledge to the model, including point symmetry, crystal structure periodicity, lattice structures, and other rich information. This encompasses global and local information about the crystal structure, raising the bar for the encoder to learn more refined knowledge. As Culprit 3 mentioned, the original models struggle to capture such fine-grained knowledge and are unable to learn detailed representations that incorporate these specific insights.

Therefore, we propose a new hierarchical Pattern learning model tailored to the characteristics of space groups. This model consists of two main components: (i) Local Pattern module (LP), which captures local information between adjacent peaks in the PXR D pattern, outputting a  $k_{local}$ -dimensional representation  $E_{local}$ ; (ii) Global Pattern module (GP), which captures global information from the PXR D pattern, outputting a  $k_{global}$ -dimensional representation  $E_{global}$ . We combine these two representations to obtain the final output representation of the model:

$$E = \text{Concat}(E_{global}, E_{local}) \quad (9)$$

This representation is used for the subsequent superclass-guided optimization process.

► **Local Pattern Module.** The input vectors of the local pattern module are  $A$  and  $I$ , which represent the diffraction angles and intensities on the PXR D pattern, respectively. According to Bragg’s law(Pope 1997),  $2d \sin(\theta) = n\lambda$ , the interplanar spacing is inversely related to  $\sin(\theta)$ . Thus, we replace  $A$  with  $\sin(A)$  and concatenate it with the peak intensity  $I$  to form an input sequence with channel = 2.

Next, we employ several residual 1D convolution blocks (a kernel size of 3 and a stride of 1) to capture the correlations between adjacent peaks in the PXR D pattern. After the convolution process, we flatten the features and project them to obtain the local pattern representation  $E_{local}$ . Therefore, the representation process of the local pattern module can be formalized as:

$$E_{local} = \text{Projection}(\text{Flatten}(\text{ConvBlocks}(\sin(A) \oplus I))) \quad (10)$$

where Projection refers to a linear projection, Flatten denotes the flattening operation, and ConvBlock is a submodule composed of convolutional layers, while  $\oplus$  represents the concatenation operation.

► **Global Pattern Module.** The global pattern module first segments the peak intensity data  $I$  into  $L_p$  consecutive patches of length  $p$ . Each patch  $I_p$  is then processed by

Method	Accuracy (%) on MOF			F1 Score (%)	Recall (%)	Accuracy (%) on MOF-Balanced		
	Top-1	Top-2	Top-5			Top-1	Top-2	Top-5
MLP	9.10 (-29.90)	15.10 (-41.30)	30.10 (-47.48)	6.43 (-16.17)	5.24 (-16.26)	4.10 (-18.8)	5.40 (-27.00)	8.21 (-38.79)
CNN	39.00 (+0.00)	56.40 (+0.00)	77.58 (+0.00)	22.60 (+0.00)	21.50 (+0.00)	22.90 (+0.00)	32.40 (+0.00)	47.00 (+0.00)
NoPoolCNN	38.20 (-0.80)	51.80 (-4.60)	71.12 (-6.46)	34.47 (+11.87)	31.84 (+10.34)	33.80 (+10.90)	40.70 (+8.30)	50.97 (+3.97)
RCNet	59.00 (+20.00)	73.70 (+17.30)	88.37 (+10.79)	41.29 (+18.69)	40.38 (+18.80)	44.50 (+21.60)	55.50 (+23.10)	69.12 (+22.12)
XRDMamba	<u>72.20</u> (+33.20)	<u>85.20</u> (+28.80)	<u>93.42</u> (+15.84)	47.59 (+24.99)	46.00 (+24.50)	48.70 (+25.80)	61.70 (+29.3)	74.83 (+27.83)
XRDecoupler	<b>80.09</b> (+41.09)	<b>90.11</b> (+33.71)	<b>96.26</b> (+18.68)	<b>56.72</b> (+34.12)	<b>55.18</b> (+33.68)	<b>58.87</b> (+35.97)	<b>72.42</b> (+40.02)	<b>85.22</b> (+38.22)

Table 1: Evaluation on MOF subset (left) and MOF-Balanced subset (right) of CCDC dataset with SOTA methods. **Bold** indicates the best performance while underline indicates the second best. (+) and (-) indicate the relative gain with CNN.

convolution blocks to capture local peak correlations, yielding a peak intensity representation,  $e_{intensity} \in \mathbb{R}^{k_{intensity}}$ . Next, we assign a learnable feature  $e_{learnable} \in \mathbb{R}^{k_{learnable}}$  for each patch, allowing the model to adaptively learn representations that characterize the patch. After setting the positional encoding  $e_{position}$  for each patch, we concatenate the peak intensity features and the learnable features, and then add  $e_{position}$  to form the final representation of the patch. Subsequently, we use several attention modules to learn these features, enabling the model to capture the correlations between any two patches and providing more global information. After that, we apply a global pooling layer to obtain the final global representation  $E_{global}$ .

However, we experimentally found that some patches obtained from the PXRD pattern have peak intensities that are entirely zero. These patches hold no value for the model during the learning process and may even have a negative effect on other patches. Therefore, during the calculation of attention, we apply a masking process to these patches, defined as:

$$e_{mask} = (\max(I_p) == 0). \quad (11)$$

This ensures that patches with all zero intensities do not contribute to the attention calculations.

Therefore, the representation process of the global pattern module can be formalized as:

$$E_{global} = \text{GlobalPooling}(\text{AttentionBlocks}(e_{intensity} \oplus e_{learnable}) + e_{position}, e_{mask}) \quad (12)$$

Here,  $e_{intensity}$  represents the peak intensity features of the patch,  $e_{learnable}$  denotes the learnable features for the patch,  $e_{position}$  represents the positional features of the patch, and  $e_{mask}$  indicates whether each patch is masked.

## Experiments

For more detailed information, please refer to **supplementary material in the extended version** of our paper.

► **Dataset and Baselines.** We use the MOF dataset as our main dataset as (Yu et al. 2024), which consists of over 280,000 metal-organic frameworks (MOFs) (Furukawa et al. 2013; James 2003; Zhou, Long, and Yaghi 2012) from the Cambridge Crystallographic Data Centre (CCDC) (Allen et al. 1979) for our experiments. Furthermore, we were successful in acquiring two additional datasets, CoREMOF (Chung et al. 2019) and InorganicData (Salgado et al. 2023), to assist in verifying the effectiveness of our method. We provide the details of datasets and processing

in **Supplement C**. To ensure fairness in the experiments, we selected several SOTA space group prediction models as our baselines, including MLP (Salgado et al. 2023), CNN (Salgado et al. 2023), NoPoolCNN (Salgado et al. 2023), RCNet (Chen et al. 2024), and XRDMamba (Yu et al. 2024). Detailed descriptions are provided in **Supplement D**.

► **Benchmark Results.** As shown in Table 1, we outperform all baselines on both the MOF (left) and MOF-Balanced (right) test sets, surpassing the previous SOTA (XRDMamba). These results demonstrate our method effectively alleviates SPC and establishes a new SOTA. More results on CoREMOF and InorganicData are provided in **Supplement F.1**.

► **Effectiveness of Gradient-based Optimization.** Figure 4(a) shows the loss descent curves for space group and superclasses when the gradient-based multi-objective optimization method is not employed. We can observe that the model initially optimizes the coarse-grained Bravais lattice types and crystal system types. Only after the losses for these superclasses drop to a low level does the model begin to optimize the space groups and point groups. Figure 4(b) presents the loss descent curves for each superclass and space group after introducing the gradient-based multi-objective optimization method. Here, we can see that the optimization directions for each class become consistent, allowing the model to simultaneously learn structural knowledge across multiple dimensions.

► **T-SNE Visualization Analysis of Confusion Phenomena.** Figure 4(right) displays the t-SNE plot of over 3,000 samples randomly sampled from 60 classes in the training set. The figure clearly shows a distinct clustering of features, with noticeable spacing between samples from different space group categories, indicating the superior performance of our model on the training set. We also provide more t-SNE visualization analysis in **Supplement F.2**.

► **Dependencies between Representations and Superclasses.** To validate the roles of global and local representations, we explored their interdependencies with superclass tasks during classification. We analyzed the classifier weights for space groups and each superclass, observing the importance of different representation locations for the classifier’s decision-making by examining the distribution of these weights. The relevant visualizations are shown in Figure 5. Therefore, we can conclude that in XRDecoupler, both the global and local modules play significant roles and

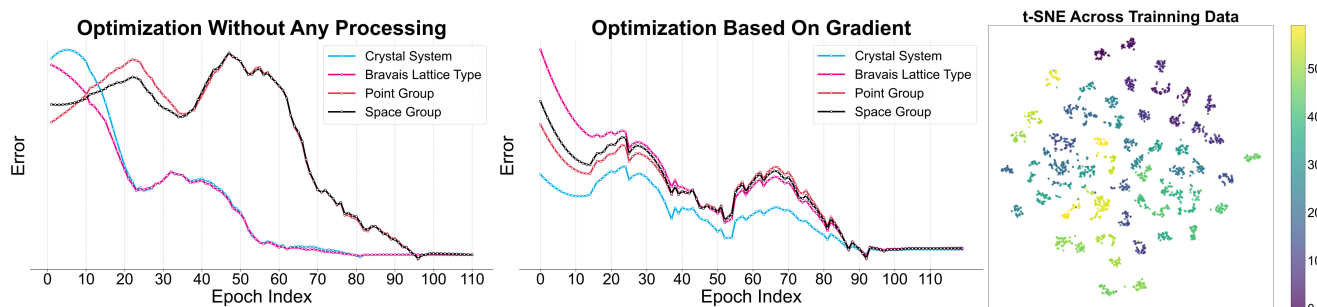


Figure 4: Trend of the model’s training loss. (left) The conventional optimization process of the model on the space group and superclass. (middle) The optimization process of the model on the space group and superclass after introducing the gradient-based optimization method. (right) T-SNE Visualization Analysis of XRDecoupler in the training set.

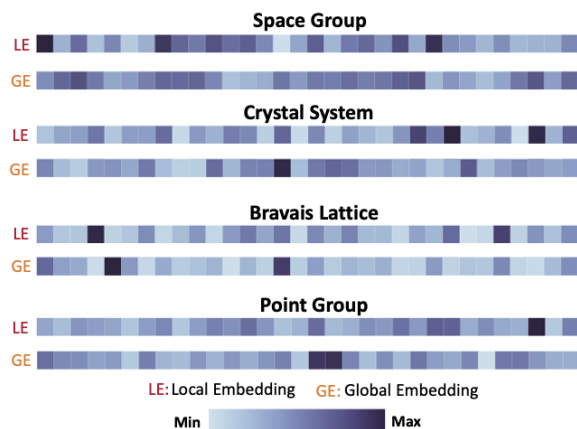


Figure 5: Visualization of the impact of global and local representations on the decision of each superclass.

exhibit notable interdependencies across superclass tasks.

#### ► Focus Tendency of Global & Local Pattern Modules.

To explore the roles of the global and local modules, we visualized their attention levels in Figure 6. Specifically, we used Grad-CAM (Selvaraju et al. 2017; Zeiler 2014; Pope et al. 2019) for the local module’s focus and analyzed the self-attention matrix for the global module. Figure 6 shows the local module focuses on local variations between peaks (red dashed box), while the global module attends to patches with high peaks (orange dashed box). Thus, the two modules capture distinct fine-grained information from the PXRD pattern, which is key to XRDecoupler’s ability to learn detailed representations and achieve superior performance.

► **Generalization Analysis.** We conducted a generalization test using 8,000 inorganic crystal data obtained from (Salgado et al. 2023), which encompasses 178 space group categories. According to the theoretical logic of symmetry classification, inorganic crystals and MOF crystals are analogous, differing primarily in their building units, types of chemical bonds, topological structures, and pore structures. Therefore, inorganic crystals represent out-of-domain data for us, and the model’s performance on this data serves as a good measure of its generalization capability. As shown in

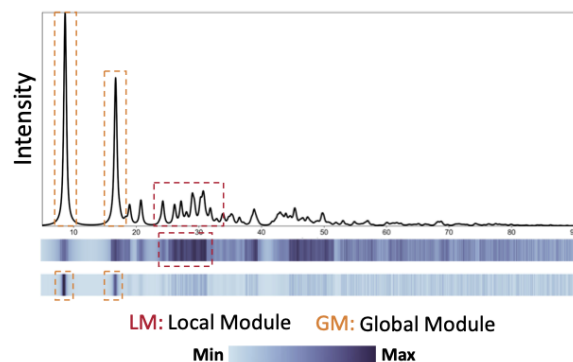


Figure 6: Visualization of the attention of the global and local pattern module to each region in the input PXRD pattern.

Method	Accuracy (%)		
	Top-1	Top-2	F1 Score
MLP	15.50 (−14.10)	21.4 (−23.20)	8.5 (+0.8)
CNN	29.60 (+0.00)	44.60 (+0.00)	7.70 (+0.00)
NoPoolCNN	30.40 (+0.8)	41.60 (−3.00)	15.90 (+8.2)
RCNet	41.70 (+12.10)	52.40 (+7.80)	19.40 (+11.70)
XRDMamba	54.50 (+24.90)	64.70 (+20.10)	24.10 (+16.40)
<b>XRDecoupler</b>	<b>60.22 (+30.62)</b>	<b>70.45 (+25.85)</b>	<b>29.09 (+21.39)</b>

Table 2: Generalization analysis on the inorganic dataset with SOTA methods. **Bold** indicates the best performance while (+) and (−) indicate the the relative gain with CNN.

Table 2, our method outperforms the SOTA methods, which indicates that XRDecoupler significantly enhances generalization performance on out-of-domain data.

► **Further analysis.** Please see **Supplement F** to find more analysis, including ablation study, crystal scale adaptability, and case analyses.

## Conclusion

In this paper, we present the XRDecoupler framework, which efficiently determines the symmetry of crystalline materials. This work advances the application of deep learning in crystalline materials analysis, providing a novel methodology for symmetry identification.

## Acknowledgements

The authors gratefully acknowledge the support from the National Natural Science Foundation of China (NSFC) under Grant Nos. 62402472, and 12227901. This work was also supported by the Natural Science Foundation of Jiangsu Province (No. BK20240461), the Project of Stable Support for Youth Team in Basic Research Field, CAS (No. YSBR-005), and the Academic Leaders Cultivation Program at USTC. The AI-driven experiments, simulations and model training were performed on the robotic AI-Scientist platform of Chinese Academy of Sciences.

## References

- Allen, F. H.; Bellard, S.; Brice, M.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B.; Kennard, O.; Motherwell, W.; et al. 1979. The Cambridge Crystallographic Data Centre: computer-based search, retrieval, analysis and display of information. *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, 35(10): 2331–2339.
- Altomare, A.; Caliandro, R.; Camalli, M.; Cuocci, C.; Silva, I. d.; Giacovazzo, C.; Moliterni, A. G.; and Spagna, R. 2004. Space-group determination from powder diffraction data: a probabilistic approach. *Journal of applied crystallography*, 37(6): 957–966.
- Ash, R. B. 2012. *Information theory*. Courier Corporation.
- Batina, L.; Gierlichs, B.; Prouff, E.; Rivain, M.; Standaert, F.-X.; and Veyrat-Charvillon, N. 2011. Mutual information analysis: a comprehensive study. *Journal of Cryptology*, 24(2): 269–291.
- Bhagavantam, S. t.; and Suryanarayana, D. 1949. Crystal symmetry and physical properties: application of group theory. *Acta Crystallographica*, 2(1): 21–26.
- Bradley, C.; and Cracknell, A. 2009. *The mathematical theory of symmetry in solids: representation theory for point groups and space groups*. Oxford University Press.
- Chen, L.; Wang, B.; Zhang, W.; Zheng, S.; Chen, Z.; Zhang, M.; Dong, C.; Pan, F.; and Li, S. 2024. Crystal Structure Assignment for Unknown Compounds from X-ray Diffraction Patterns with Deep Learning. *Journal of the American Chemical Society*, 146(12): 8098–8109.
- Chung, Y. G.; Haldoupis, E.; Bucior, B. J.; Haranczyk, M.; Lee, S.; Zhang, H.; Vogiatzis, K. D.; Milisavljevic, M.; Ling, S.; Camp, J. S.; et al. 2019. Advances, updates, and analytics for the computation-ready, experimental metal–organic framework database: CoRE MOF 2019. *Journal of Chemical & Engineering Data*, 64(12): 5985–5998.
- Dehkordi, H. A.; Nezhad, A. S.; Kashiani, H.; Shokouhi, S. B.; and Ayatollahi, A. 2022. Multi-expert human action recognition with hierarchical super-class learning. *Knowledge-Based Systems*, 250: 109091.
- Dong, H.; Butler, K. T.; Matras, D.; Price, S. W.; Odarchenko, Y.; Khatri, R.; Thompson, A.; Middelkoop, V.; Jacques, S. D.; Beale, A. M.; et al. 2021. A deep convolutional neural network for real-time full profile analysis of big powder diffraction data. *NPJ Computational Materials*, 7(1): 74.
- Du, Y.; Shen, J.; Zhen, X.; and Snoek, C. G. 2023. Superdisco: Super-class discovery improves visual recognition for the long-tail. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19944–19954.
- Evans, P. R. 2011. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Biological crystallography*, 67(4): 282–292.
- Furukawa, H.; Cordova, K. E.; O’Keeffe, M.; and Yaghi, O. M. 2013. The chemistry and applications of metal-organic frameworks. *Science*, 341(6149): 1230444.
- Gu, A. 2023. *Modeling Sequences with Structured State Spaces*. Stanford University.
- Gu, A.; and Dao, T. 2023. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*.
- Hahn, T.; Shmueli, U.; and Arthur, J. W. 1983. *International tables for crystallography*, volume 1. Reidel Dordrecht.
- Han, Y.; Ding, C.; Wang, J.; Gao, H.; Shi, J.; Yu, S.; Jia, Q.; Pan, S.; and Sun, J. 2025. Efficient crystal structure prediction based on the symmetry principle. *Nature Computational Science*, 1–13.
- James, S. L. 2003. Metal-organic frameworks. *Chemical Society Reviews*, 32(5): 276–288.
- Koster, G. F. 1957. Space groups and their representations. In *Solid state physics*, volume 5, 173–256. Elsevier.
- Nespolo, M.; Aroyo, M. I.; and Souvignier, B. 2018. Crystallographic shelves: space-group hierarchy explained. *Journal of Applied Crystallography*, 51(5): 1481–1491.
- O’Keeffe, M.; and Hyde, B. G. 2020. *Crystal structures*. Courier Dover Publications.
- Oviedo, F.; Ren, Z.; Sun, S.; Settens, C.; Liu, Z.; Hartono, N. T. P.; Ramasamy, S.; DeCost, B. L.; Tian, S. I.; Romano, G.; et al. 2019. Fast and interpretable classification of small X-ray diffraction datasets using data augmentation and deep neural networks. *npj Computational Materials*, 5(1): 60.
- Park, W. B.; Chung, J.; Jung, J.; Sohn, K.; Singh, S. P.; Pyo, M.; Shin, N.; and Sohn, K.-S. 2017. Classification of crystal structure using a convolutional neural network. *IUCrJ*, 4(4): 486–494.
- Pitteri, M.; and Zanzotto, G. 1996. On the definition and classification of Bravais lattices. *Acta Crystallographica Section A: Foundations of Crystallography*, 52(6): 830–838.
- Pope, C. G. 1997. X-ray diffraction and the Bragg equation. *Journal of chemical education*, 74(1): 129.
- Pope, P. E.; Kolouri, S.; Rostami, M.; Martin, C. E.; and Hoffmann, H. 2019. Explainability methods for graph convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10772–10781.
- Prince, E. 2004. *International Tables for Crystallography, Volume C: Mathematical, physical and chemical tables*. Springer Science & Business Media.
- Salgado, J. E.; Lerman, S.; Du, Z.; Xu, C.; and Abdolrahim, N. 2023. Automated classification of big X-ray diffraction

- data using deep learning models. *npj Computational Materials*, 9(1): 214.
- Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; and Batra, D. 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, 618–626.
- Sener, O.; and Koltun, V. 2018. Multi-task learning as multi-objective optimization. *Advances in neural information processing systems*, 31.
- Spence, J.; and Zuo, J. 1992. Symmetry Determination. In *Electron Microdiffraction*, 145–167. Springer.
- Vecsei, P. M.; Choo, K.; Chang, J.; and Neupert, T. 2019. Neural network based classification of crystal symmetries from x-ray diffraction patterns. *Physical Review B*, 99(24): 245120.
- Wang, B.; Wang, P.; Xu, W.; Wang, X.; Zhang, Y.; Wang, K.; and Wang, Y. 2024. Kill two birds with one stone: Rethinking data augmentation for deep long-tailed learning. In *The Twelfth International Conference on Learning Representations*.
- Wang, D.; Song, Y.; Huang, J.; An, D.; and Chen, L. 2022. SAR target classification based on multiscale attention super-class network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15: 9004–9019.
- Werner, P.-E.; Eriksson, L.; and Westdahl, M. 1985. TREOR, a semi-exhaustive trial-and-error powder indexing program for all symmetries. *Journal of Applied Crystallography*, 18(5): 367–370.
- Wilson, A. 1990. Space groups rare for organic structures. II. Analysis by arithmetic crystal class. *Foundations of Crystallography*, 46(9): 742–754.
- Yu, L.; Wang, P.; Zhao, Z.; Yi, Z.; Nan, S.; Wu, D.; and Wang, Y. 2024. XRDMamba: Large-scale Crystal Material Space Group Identification with Selective State Space Model. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, 4233–4237.
- Zeiler, M. 2014. Visualizing and Understanding Convolutional Networks. In *European conference on computer vision/arXiv*, volume 1311.
- Zhao, Z.; Wen, H.; Wang, P.; Wang, Z.; Zhang, Q.; Wang, Y.; et al. 2025. Balancing Model Efficiency and Performance: Adaptive Pruner for Long-tailed Data. In *Forty-second International Conference on Machine Learning*.
- Zhao, Z.; Wen, H.; Wang, Z.; Wang, P.; Wang, F.; Lai, S.; Zhang, Q.; and Wang, Y. 2024. Breaking Long-Tailed Learning Bottlenecks: A Controllable Paradigm with Hypernetwork-Generated Diverse Experts. *Advances in Neural Information Processing Systems*, 37: 7493–7520.
- Zhou, H.-C.; Long, J. R.; and Yaghi, O. M. 2012. Introduction to metal–organic frameworks.
- Zhou, Y.; Hu, Q.; and Wang, Y. 2018. Deep super-class learning for long-tail distributed image classification. *Pattern Recognition*, 80: 118–128.
- Ziletti, A.; Kumar, D.; Scheffler, M.; and Ghiringhelli, L. M. 2018. Insightful classification of crystal structures using deep learning. *Nature communications*, 9(1): 2775.