第 54 卷第 1 期
2026 年 1 月

硅 酸 盐 学 报
JOURNAL OF THE CHINESE CERAMIC SOCIETY

Vol. 54，No. 1
January，2026

http://jccsoc.cbpt.cnki.net

DOI：10.14062/j.issn.0454-5648.20250845

综 合 评 述

AI 驱动的材料发现：生成式模型与晶体逆向设计进展

吴克登[1]，陈 炎[2]，邓晓彬[3]，陈 曦[3]，肖 航[3]

(1. 西北大学化工学院，西安 710127；2. 西安交通大学航天航空学院，西安 710049；3. 岭南大学伍絜宜跨学科学院，香港)

摘 要：晶体材料的逆向设计一直是材料科学面临的重大挑战，近年来生成式人工智能技术的发展为解决这一难题带来了新的契机。本文围绕晶体表示方法与生成模型架构两大核心问题，系统综述了该领域的最新研究进展。在晶体表示方法方面，重点评述了文本符号化表示、几何图周期性编码以及像素的实验成像与计算像素化等方法，深入分析了这些方法在可逆性、不变性和对称性编码等关键技术特性上的优劣。在生成模型架构方面，按照技术演进路线，详细介绍了基于潜空间的变分自编码器和生成对抗网络、扩散或流匹配的联合生成模型，以及 Transformer 自回归模型的工作原理与应用现状。最后，针对当前面临的表示方法优化、缺陷材料建模和评估标准统一等关键科学问题，本文展望了面向工业应用的未来发展方向，旨在为晶体材料智能设计提供系统性的理论参考。

关键词：逆向设计；晶体表示法；生成式模型；人工智能；晶体材料
中图分类号：TP183　　**文献标志码**：A　　**文章编号**：0454–5648(2026)01–0090–16
网络出版时间：2026–01–05

近十年来，基于生成模型的分子逆向设计取得了显著进展[1−2]。这一突破主要归功于分子领域已建立起成熟的表示体系，如简化分子线性输入规范(SMILES)[3]、国际化合物标识(InChI)[4]和自引用嵌入式字符串(SELFIES)[5]等。这些表示方法具备 2 个关键特性：可逆性，即材料结构与编码形式之间可实现无损双向转换；不变性，即经过平移、旋转或排列等操作后，表示形式仍对应同一结构。借助这些成功的表示方法，研究者开发出一系列突破性的生成模型，包括用于药物–靶点相互作用建模的交互感知框架[6]、面向结合位点设计的数据与知识驱动 PocketFlow 模型[7]、多模态预训练模型[8]以及有机分子生成模型 G-SchNet[9]等，有力证实了可逆且不变的表示法对推动逆向设计的核心价值。

然而，将生成模型应用于固态材料逆向设计仍面临诸多挑战[10−11]，其根本原因在于尚未建立一种同时满足可逆性、不变性并能编码周期性特征的晶体表示方法，且该方法需涵盖元素周期表中的绝大多数元素。不同于分子体系，开发晶体材料表示法面临以下关键难题[12]：1) 对称性与周期性约束。表示方法需涵盖全部 230 个空间群，准确捕捉平移、旋转和反射等对称操作，同时保持可逆性、不变性和周期性；2) 结构与化学复杂性。晶体结构通常包含更多原子、多样化的缺陷和杂质，其尺寸和排列复杂度远超小分子体系，导致化学空间急剧扩大，显著增加了计算成本；3) 环境响应特性。准确描述晶体材料在不同环境条件下的行为极具挑战，而海量数据的整合与解析进一步加剧了这一难度。

尽管面临上述挑战，近年来研究者已证明：为晶体材料构建兼具可逆性、不变性与周期性编码能力的表示方法是可行的。这些探索催生了多个具有

收稿日期：2025–11–15。　修订日期：2025–12–04。
基金项目：国家重点研发计划(2024YFA1209801)；国家自然科学基金(22203066，12302140)；中国博士后科学基金(2023M732794，2025T180517)；中央高校基本科研业务费专项资金(sxzy012023213)；陕西省科学技术厅人工智能赋能科研专项经费(2025YXYC012)；西安市科学技术协会青年人才托举计划(959202413069)；中国博士后科学基金博士后创新人才支持计划 B 档(GZB20230575)。
第一作者：吴克登(1999—)，男，硕士研究生。
通信作者：肖 航(1989—)，男，博士，助理教授。

Received date: 2025–11–15.　**Revised date:** 2025–12–04.
First author: WU Kedeng (1999–), male, Master candidate.
E-mail: wkd135790_29@163.com
Correspondent author: XIAO Hang (1989–), male, Ph.D., Assistant Professor.
E-mail: hangxiao@ln.edu.hk

里程碑意义的晶体生成模型：MatterGen[13]实现了大规模联合扩散生成，CrystalFormer[14]通过自回归方式生成符合空间群约束的结构，MatterGPT[15]则结合简化线性输入晶体编码系统(SLICES)[16]表示法实现高效的条件生成。这些成果标志着晶体材料逆向设计正从理论探索迈向实际应用。

为系统梳理该领域的研究进展并明确未来发展方向，本文重点探讨晶体材料逆向设计的两大核心问题(见图 1)：1) 晶体表示方法，涵盖文本表示法、几何图表示法和图像表示法；2) 生成模型架构，包括潜空间模型[变分自编码器(VAE)[17]和生成对抗网络(GAN)[18−19]]、扩散/流模型以及自回归模型[20]。文章结合不同材料体系的代表性案例，深入剖析这些方法如何解决晶体材料生成难题及其实际表现。最后，针对晶体表示方法与生成模型面临的瓶颈，从数据增强、表示设计和模型构建等层面提出潜在的解决方案，并展望晶体材料逆向设计的未来发展方向。

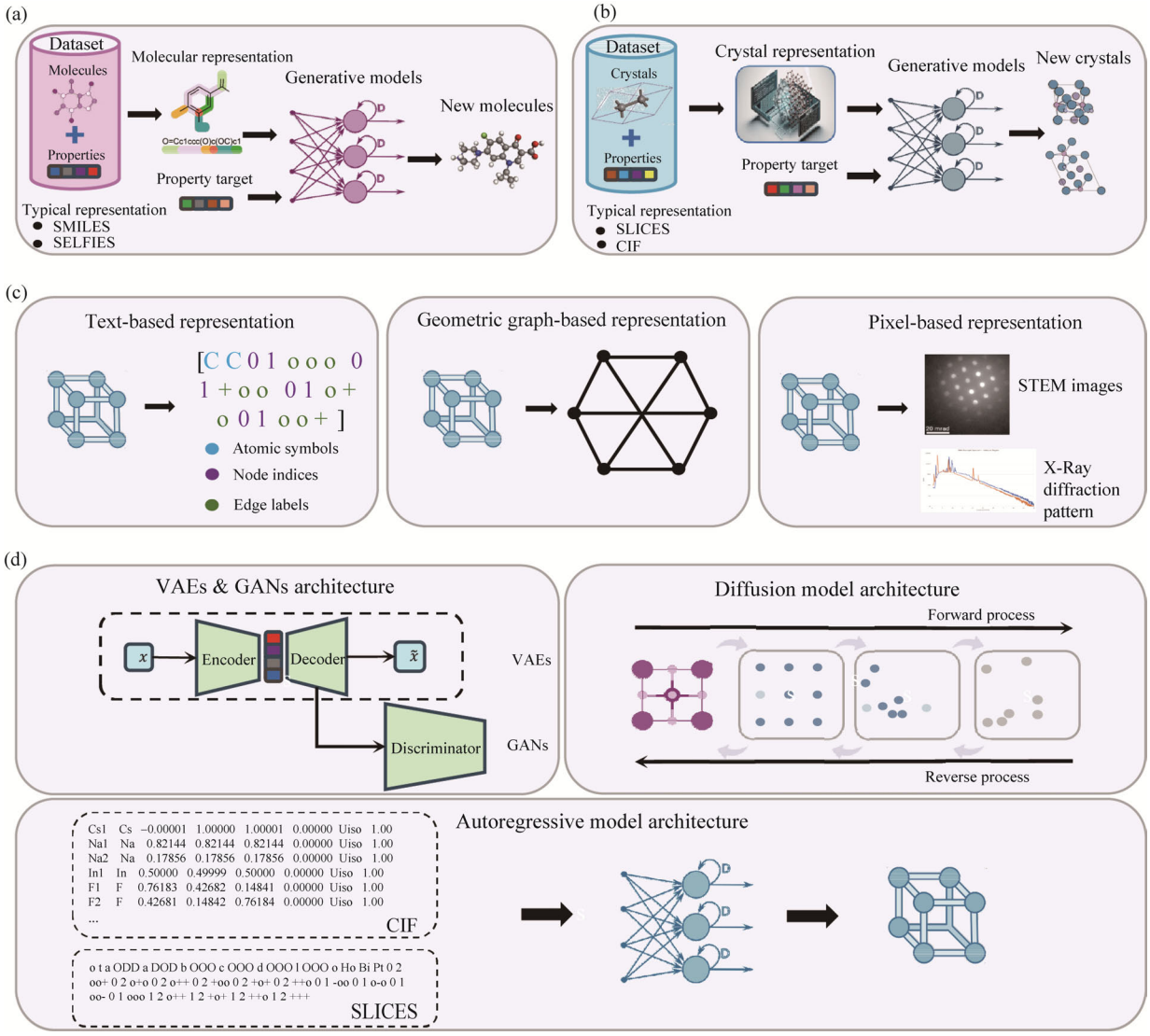


图 1　晶体材料逆向设计的技术框架。(a) 分子材料逆向设计流程：借助 SMILES 和 SELFIES 等典型分子表示方法实现分子生成；(b) 晶体材料逆向设计流程：采用 SLICES 和 CDVAE 等典型晶体表示方法实现晶体生成；(c) 3 类主流晶体表示方法：文本表示法、几何图周期性编码和图像表示法；(d) 3 类生成模型架构：VAE 与 GAN 模型架构、扩散与流模型架构和自回归模型架构

Fig. 1　Technical framework for inverse design of crystalline materials. (a) Molecular materials design: molecular generation based on typical representation methods such as SMILES and SELFIES; (b) Crystal materials design: crystal generation based on typical representation methods such as SLICES and CDVAE; (c) Three mainstream crystal representation methods: text-based symbolic representation, geometry-based periodicity-aware encoding, and pixel-based image representation; (d) Three generative model architectures: latent space-based VAE and GAN models architectures, diffusion and flow models architectures, and autoregressive model architectures

1 晶体表示方法

晶体材料的数字化表示是实现机器学习建模的基础。晶体本质上由原子在三维空间中周期性排列而成，其结构由晶胞、晶格参数、原子坐标及对称操作等要素完整描述。传统上，国际晶体学联合会制定的晶体学信息文件(CIF)作为标准化文本格式[21]，是记录与交换晶体结构信息的通用标准。

然而，CIF 作为一种结构化表格形式，难以被深度学习架构直接处理。与分子领域已建立的 SMILES 等成熟表示方法不同，晶体材料因其周期性和对称性，对表示方法提出了更为严格的物理约束。因此，开发兼具周期性编码能力、可逆性和不变性的晶体表示方法，已成为生成式人工智能驱动晶体逆向设计的核心难题之一。为满足人工智能模型对数据格式与物理约束的双重要求，研究者近年来发展了多种新型晶体表示方法，力求在可逆性、不变性和周期性编码能力之间实现最优平衡。本节将系统介绍晶体材料研究中的主流表示方法，包括文本表示法、几何图表示法以及图像表示法。

1.1 文本表示法

文本表示法借鉴了分子领域 SMILES 等字符串编码的成功经验，将三维晶体结构转化为机器可读的序列形式，使大型语言模型能够直接用于材料设计与发现[22−23]。早期工作尝试直接使用 CIF 作为文本表示形式。CIF 是晶体学的标准归档格式，以文本形式记录晶格参数、原子坐标和空间群等结构信息。已有研究将 CIF 作为输入，通过训练或微调语言模型(如 CrystaLLM[24])实现材料生成。但 CIF 表示(特别是 P1 对称性下)缺乏旋转不变性，需要大量数据增强才能充分覆盖材料空间。

为克服这些局限，科研工作者开发了具有可逆性、不变性和周期性编码能力的新型文本编码方法。SLICES 是一项代表性创新[16]，该方法利用商图作为中介[25]，将晶体结构转化为包含原子符号、节点索引和边标签的紧凑字符串，编码了化学组成与键拓扑信息。SLICES 的突出优势在于高度可逆性，其重建算法在包含超过 40 000 种结构的数据集上，实现了 94.95%的结构重构成功率。

针对空间对称性的编码需求，后续工作提出了改进版本 SLICES-PLUS[26]，该方法通过整合 Wyckoff 记号描述的空间对称性信息(如旋转矩阵和平移向量)，显著增强了对晶体三维空间特征的表征能力。实验表明 SLICES-PLUS 在基于对称性的分类和生成任务中，展现出比原始 SLICES 更高的准确度和鲁棒性。

值得注意的是，SLICES 表示法在晶体性质预测任务中展现出优异性能，该表示法虽仅编码化学组成和键拓扑信息，不包含显式的原子三维坐标，但在 MatText 基准测试[27]中的多项任务上表现突出。实验结果表明：SLICES 的预测精度可与包含完整几何信息的表示法(如 CIF 格式[28])相媲美，甚至在某些任务上更胜一筹。这一发现表明：SLICES 表示法通过简洁的字符串形式，捕捉并编码了决定材料宏观性质的核心本质信息——化学组成与键拓扑信息，为晶体材料信息学提供了高效且轻量化的建模途径。

1.2 几何图表示法

在晶体材料性质预测和生成任务中，基于图的表示法已成为深度学习模型最为广泛采用的方法之一[29−30]，该表示法将晶体结构映射为图结构 $G=(V, E)$，其中节点 $v_i \in V$ 对应晶胞内的原子，边 $e_{ij} \in E$ 表征原子 i 与 j 之间的相互作用或化学键。节点特征用于编码元素固有属性(如原子序数、电负性等)，边特征则承载几何信息，如键长、键角等距离参数。这种表示法的核心优势在于能够有效描述原子的局部化学环境，同时对原子索引的置换操作具有天然的不变性。

在描述原子空间排布时，图表示法通常依据晶格参数(L)和原子在晶胞内的坐标确定原子位置，其中相对于晶格矢量的分数坐标(F)具有良好的周期性编码能力特性，其取值通常限定在[0，1]区间内，能够自然处理晶体的周期性边界。在多数晶体图神经网络的具体实现中，L 和 F 主要用于计算原子间的邻接关系及距离、角度等边特征，而并非作为完整的连续坐标向量显式保存在图中。

然而，这类性质预测导向的传统晶体图表示法侧重于表达连接关系和局部几何特征，并未以“严格可逆”为设计目标。由于距离和角度等几何量往往会经过截断半径、分桶离散或低维基函数展开，不同的三维原子排布在理论上可能被压缩到相同的图结构，导致仅凭图结构难以唯一、精确地重构原始三维坐标。这一缺陷限制了其作为生成模型独立输入的能力，使其无法直接生成精确的原子坐标。受此限制，在生成任务中，图通常不会作为唯一的结构参数化形式，而是与显式的点云/坐标联合使用，或仅用于在扩散、流匹配等生成模型中辅助迭代更新原子位置和晶格参数。

为构建更紧凑且符合物理约束的表示方式，研究者开发了基于对称性的编码策略。这类方法通过晶体空间群 S 和 Wyckoff 位置描述不对称单元(Asymmetric Unit)中的原子信息，仅编码不对称单元的原子信息[即 $M=(A_{asu}, X_{asu}, S, L)$，其中 A_{asu} 表示不对称单元中原子的特征向量矩阵(如原子类型、电负性等)，X_{asu} 为这些原子的笛卡尔坐标矩阵，S 表示空间群类型，L 为晶格参数]。在大幅压缩数据维度的同时，将晶体对称性作为归纳偏置引入表示中，为生成模型提供了更高效且受物理规律约束的生成空间。在此基础上，针对无序和缺陷诱导的对称性破坏问题，Dis-GEN[31]框架进一步提出了一种基于威科夫位点的增强表示策略。该方法不再将晶体视为确定的原子点阵，而是将对称性等价的原子位点聚类，并在节点特征中显式编码原子的部分占有率与空位信息。这种策略使得模型能够在保持晶体位点对称性的前提下，精准描述组分无序和空位缺陷，有效避免了传统超胞方法带来的计算成本高昂及破坏对称性的难题。

对于含缺陷晶体的几何图表示，传统方法面临显著挑战。常规图神经网络通常采用缺陷隐式表示，即不对缺陷位点进行显式标注，模型需从结构中隐式推断缺陷信息，难以准确描述缺陷引起的局部畸变及其复杂相互作用。DefiNet[32]对此提出了缺陷显式图表示策略，通过在节点上附加缺陷类型标记(0 表示完美原子，1 表示替位，2 表示空位)，并结合缺陷感知的消息传递机制，实现了对点缺陷结构几何构型的精确捕捉。该方法在二维材料缺陷结构弛豫任务中，仅需单步推理即可达到接近密度泛函理论(DFT)的预测精度，显著降低了计算成本。此外，自监督学习策略[33]也为缺陷建模提供了新思路：通过随机掩码晶胞中的部分原子并微扰原子位置来构建“受损”结构，迫使模型学习从噪声输入中重构完整结构。这种机制使模型能够隐式地捕获缺陷周围的局部化学环境特征，提升了对非理想结构的鲁棒性。上述针对缺陷与无序结构的表示方法创新，显著提升了模型对真实材料的描述能力，为发展面向工业应用的材料设计模型提供了重要参考。

1.3　图像表示法

基于图像的表示法将晶体材料的结构信息映射为图像或光谱数据，从而利用卷积神经网络(CNN)[34]强大的特征提取能力进行深度学习。根据数据来源和构建方式，这类方法可分为 2 大类别：实验成像数据和计算像素化表示。

1.3.1　实验成像与光谱数据　　第 1 类方法依赖先进的实验表征技术直接获取材料的结构信息。这些技术可根据数据类型分为成像数据和谱学数据两大类。成像数据通过显微技术获得，如扫描透射电子显微镜(STEM)、透射电子显微镜(TEM)和扫描探针显微镜(SPM)等，能够直观反映材料的表面形貌和原子尺度结构特征，对缺陷检测和物相分类等任务具有重要价值[35–38]。

谱学数据则通过测量电磁辐射与材料的相互作用产生可测量的信号。其中，衍射谱(如 X 射线衍射 XRD、电子衍射和中子衍射)以散射强度对角度或波矢的关系呈现，可提供晶面间距、晶粒尺寸等定量结构参数；振动光谱(如拉曼光谱)通过测量散射光的频移，揭示材料的分子振动和晶体结构特征。这些实验数据为结构识别、组成分析和缺陷表征等应用提供了可靠的信息基础。

1.3.2　计算像素化表示　　第 2 类方法采用计算方式将晶体的三维原子结构转化为离散的二维或三维像素化数据。这种表示方式通过将原子信息及其空间分布编码到体素化网格中，使得 CNN 能够有效捕捉局部空间特征和原子间的相互作用关系。例如，iMatGen[39]等模型将晶体结构分解为单元胞图像(通过高斯函数编码晶格参数)和原子坐标的基础图像，实现了结构信息的直观可视化表达。

然而，计算像素化表示在晶体生成任务中面临固有的局限性。首先，该方法通常缺乏旋转和平移不变性：相同晶体结构在不同观察角度或位置下会产生不同的图像张量，模型需要通过数据增强策略(如 Cond-DFC-VAE[40]中采用的随机旋转)来学习这些等价表示，从而缓解不变性问题；其次，从三维连续空间到离散网格的投影过程不可避免地导致几何精度损失，特别是对于复杂的原子排列和键合关系，离散化引入的近似误差以及生成图像的模糊性，会直接削弱重构结构的准确性，甚至导致生成物理上无效的结构。现有的基准测试有力地量化了这些差距。如表 1 所示，在 Perov-5 数据集[41–42]的对比上，采用图像表示的 Cond-DFC-VAE 在生成任务中的结构有效性(Validity)为 73.60%、覆盖精度(COV-P)为 10.13%，显著低于采用几何图表示的 CDVAE[43](分别为 100%和 98.46%)；在重构任务中，Cond- DFC-VAE 的匹配率(Match Rate)为 51.65%，其结构重构的归一化均方根误差(NRMSE)为 0.0217，同样劣于 CDVAE 的 97.52%和 0.0156。此外，另一个图像表示方法 iMatGen 在钒氧化物

数据集上的原子坐标绝对误差约为 0.19 Å。尽管数据集与评价指标略有不同(CDVAE 报告的是归一化均方根误差数值，在 Perov-5 和 MP-20 数据集上分别为 0.015 6 和 0.035 6)，但这仍能反映出基于图像的方法在几何精度上与几何图方法存在数量级的差距。

表 1 不同晶体表示方法的性能对比

Table 1 Performance comparison of different crystal representation methods

Model	Representation	Dataset	Validity/%	COV-P/%	NRMSE	Match rate/%
iMatGen	3D voxel	V–O			0.19 Å	
Cond-DFC-VAE	3D voxel	Perov-5	73.6	10.13	0.021 7	51.65
CDVAE	Periodic graph	Perov-5	100.0	98.46	0.015 6	97.52

因此，基于图像的方法虽然在材料性质预测等判别性任务中表现不错，但在晶体逆向设计这类对几何重构精度要求较高的任务中，仍存在明显短板。其核心问题在于缺乏旋转平移不变性，离散化过程中不可避免的信息损失，以及更大的计算资源消耗。相比之下，几何图表示法和字符串表示法已经发展出较为成熟的技术框架，能够有效编码周期性、保证可逆性并满足不变性要求，而图像表示方法在这些方面还有待突破，目前尚未成为该领域的主流方案。

2 生成模型架构

几何图神经网络、卷积神经网络、语言模型和扩散模型等深度学习模型的发展，为材料科学的进步发展提供了重要支撑[44–46]。晶体材料的逆向设计旨在通过学习材料的构效关系，生成满足目标性质的新型晶体结构。为确保生成结构的可行性，模型需遵循电荷守恒、合理的原子配位和成键规则等基本化学约束。与传统的穷举搜索或试错方法相比，基于深度学习的生成模型能够在更广阔的化学空间中高效探索，从而加速发现稳定且性能优异的新材料。

当前主流的架构通常采用多步流程，通过不同的采样机制、目标函数和生成策略来应对晶体材料固有的结构复杂性与周期性约束。这些模型(例如 MatterGen[13]和 CrystalFormer[14]等)通过集成对物理和化学规律的理解，将复杂的晶体结构生成任务分解为多个子任务，并逐步进行设计与优化，从而实现更高质量的晶体生成。本章将遵循这些架构的发展脉络，重点评述几类主流的生成模型，并根据其模型架构特点和采样机制进行分类阐述。

2.1 潜空间生成模型架构

潜空间生成模型以 VAE 和 GAN 为代表，是生成式模型的早期代表性架构[47]。这两类模型的核心思想在于：通过学习数据从高维空间到低维连续潜空间的映射关系，实现对复杂数据分布的建模与生成。

2.1.1 VAE 架构 VAE 是一种概率生成模型[48]，通过最大化证据下界(ELBO)进行训练，其目标函数可表示为式(1)。

$$L_{\mathrm{VAE}}(\theta,\varnothing)=E_{q_{\varnothing}(z|x)}\times\left[\ln\ p_{\theta}(x|z)\right]-D_{\mathrm{KL}}\times[q_{\varnothing}(z\,|\,x)\,\|\,p(z)] \tag{1}$$

式中：θ 为解码器(生成网络)的可学习参数；$\varnothing$ 为编码器(推理网络)的可学习参数；x 表示观测数据(晶体结构)；z 为潜在变量；$q_{\varnothing}(z\,|\,x)$ 为编码器定义的近似后验分布，通常假设为高斯分布；$p_{\theta}(x|z)$为解码器定义的似然函数；$p(z)$为潜在空间的先验分布，通常设为标准正态分布$\mathcal{N}(0，I)$；D_{KL} 表示 Kullback-Leibler (KL)散度，用于度量近似后验分布与先验分布之间的差异；$E_{q_{\varnothing}(z|x)}$ 表示在分布 $q_{\varnothing}(z\,|\,x)$ 下的期望。

如图 2a 所示，式(1)由 2 部分构成：第 1 项为重构损失，用于确保解码器 $p_{\theta}(x|z)$能够从潜在编码 z 中精确恢复原始数据 x；第 2 项为 KL 散度，作为正则化项约束编码器 $q_{\varnothing}(z\,|\,x)$ 学习到的后验分布逼近一个预设的先验分布 $p(z)$ [通常为$\mathcal{N}(0，I)$]。这种正则化确保了潜空间的连续性和结构性，使其易于进行采样和插值。

在晶体材料生成领域，研究者早期便尝试将 VAE 与不同的晶体表示方法相结合。作为该方向的先驱工作，iMatGen[39]通过体素化图像表示法，在钒氧化物(V–O)体系的逆向设计中展现了其实际应用价值。该模型不仅能够重构已知的实验结构，更在未知的化学空间中发现了 52 种全新的化学组分(如 V_6O_7、V_5O_8等)，这些组分从未被收录于 Materials Project 数据库[49]或训练数据集中。通过对生成结构进行 DFT 验证，研究确认了其中 40 种结构具有热力学相对稳定性($E_{\mathrm{hull}}\leqslant 0.080$ eV，E_{hul} 为每个原子相对于该组分热力学稳定相凸包的能量距离)，验证了模型探索未知材料空间的潜力。Cond-DFC-VAE 则进一步发展了条件生成能力，该模型通过采用随机旋

转等数据增强策略，有效缓解了体素化表示法固有的旋转不变性问题，从而提升了模型的泛化能力[40]。实践表明，该模型能够生成 Lu_2ZnAu 等 14 种新材料，且这些生成结构的晶格参数与实验验证材料高度吻合。

随着研究的深入，研究者逐渐将更多物理知识与晶体学约束融入 VAE 架构。PCVAE[50]是这一方向的重要进展，如图 2b 所示，它将给定化学式的不同可能晶体结构编码为潜空间中的高斯分布。PCVAE 的关键创新在于显式整合了晶体物理知识，包括晶系、布拉维晶格和空间群之间的映射关系，以及对晶格参数的约束条件。这些物理约束不仅引导生成过程产生更符合晶体学规则的结构，更赋予了模型解决“一式多构”(MS4OF)难题的能力。在包含 33 160 个多晶型结构的数据集上，PCVAE 的晶格常数预测 R^2(决定系数，是评估回归模型预测性能的标准指标)分数达 0.79～0.88，显著优于传统方法。对于具有多种稳定构型的体系，模型能通过潜空间采样生成多组候选结构，实现了对多晶型材料的有效建模。

针对特定材料体系，研究者开发了多个面向特定应用的 VAE 模型。MagGen[51]作为一个条件 VAE 模型，专注于稀土无永磁体的逆向设计。该模型采用可逆的实空间晶体学表示，并以形成能和饱和磁化强度为条件对潜空间进行约束。MagGen 的特点是利用图论分析属性嵌入潜空间的结构，使搜索过程能够朝向具有理想磁性能的结构区域收敛，从而实现更高效的候选材料采样。该工作成功发现了 46 种稳定磁性材料，其中 20 种材料的 $E_{hull}\leqslant 0.250$ eV，具备实验合成可能性。尤为重要的是，$Fe_3O_2F_2$ 和 $MnNi_2O_2$ 2 种稀土无永磁候选材料脱颖而出，其单轴磁各向异性常数 K_1 均超过 1 MJ/m³，饱和磁化强度 M_s 均大于 1 T，展现出优异的永磁体应用潜力。WyCryst[52]则将 Wyckoff 位置表示与性质导向的 VAE 相结合，在结构生成及后续密度泛函理论优化过程中显式施加晶体对称性约束，成功获得 Cu_3SnSe_4 等 8 种动力学稳定的新结构，其中 7 种既不在数据库中也未见文献报道，验证了对称性约束生成策略在实际材料发现中的有效性。

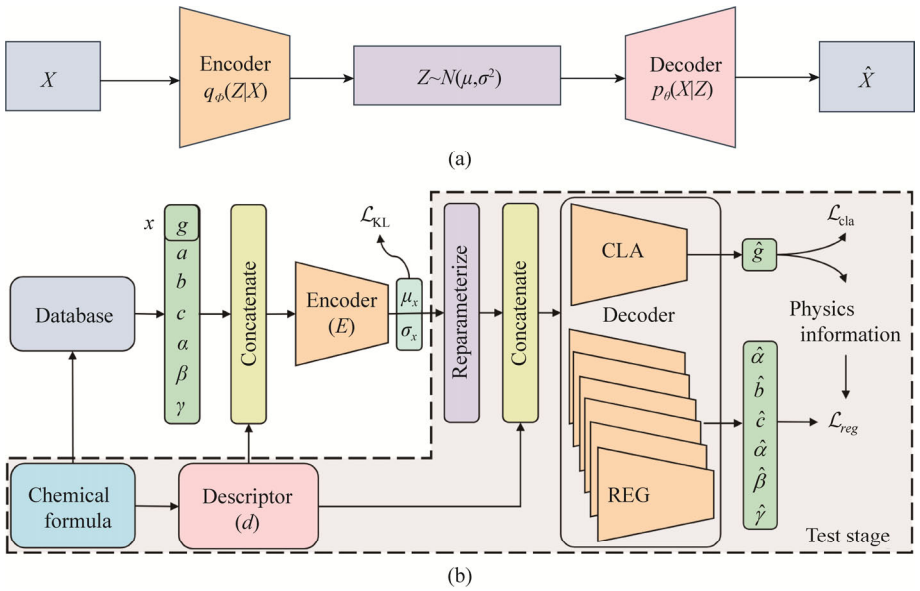


图 2　VAE 生成架构。(a) VAE 的基本框架：通过编码器-潜空间-解码器实现数据重构与生成；(b) PCVAE 的工作流程：融合晶体物理约束(晶系、布拉维晶格、空间群)实现晶体结构的条件生成[50]

Fig. 2　VAE based generative architecture. (a) Basic framework of VAE: data reconstruction and generation via encoder-latent space-decoder; (b) Workflow of PCVAE: conditional generation of crystal structures by integrating physical constraints (crystal system, Bravais lattice, space group)[50]

2.1.2　GAN 架构　　GAN 通过对抗性训练机制来学习数据分布[53]，该网络由生成器和判别器两部分组成，两者在极小极大博弈中进行优化，目标函数如式(2)所示。

$$\min_G \max_D L_{GAN}(D,G) = E_{x\sim p_{data}(x)}\times\left[\ln D(x)\right] + E_{z\sim p_z(z)}\times\left[\ln\left(1-D\left(G(z)\right)\right)\right] \quad (2)$$

式中：$\min_G \max_D$ 表示生成器和判别器的对抗优化过

程(判别器通过最大化目标函数来提升识别真假样本的能力，生成器则通过最小化目标函数来提升生成样本的逼真度；两者交替优化，最终使生成器能够产生与真实数据分布一致的样本)。具体而言，生成器网络 G 将潜在向量 z 映射到数据空间以生成样本 $G(z)$；D 表示判别器网络，输出[0，1]区间内的标量值用于判断输入样本的真实性；x 表示从真实数据分布 $p_{\text{data}}(x)$中采样的真实样本；z 表示从先验分布 $p_z(z)$ [通常为标准正态分布$\mathcal{N}(0, I)$]中采样的潜在向量；$D(x)$表示判别器对真实样本的输出值(接近 1 表示判定为真)；$D(G(z))$表示判别器对生成样本 $G(z)$的输出值(接近 0 表示判定为假)；$E_{x\sim p_{\text{data}}(x)}$和 $E_{z\sim p_z(z)}$ 分别表示在相应分布下的期望；第一项 $E_{x\sim p_{\text{data}}(x)}\times\left[\ln D(x)\right]$衡量判别器正确识别真实样本的能力，第二项 $E_{z\sim p_z(z)}\times\left[\ln\left(1-D\left(G(z)\right)\right)\right]$衡量判别器正确识别生成样本的能力。两者交替优化，最终使生成器能够产生与真实数据分布一致的样本。

如图 3a 所示，生成器 G 学习从简单先验分布 $p_z(z)$中采样的潜在向量 z 到数据分布的映射函数，以生成“伪”样本 $G(z)$；判别器 D 作为二分器，用于区分真实数据分布与生成器产生的数据分布。理想情况下，训练最终达到纳什均衡，此时生成器能够捕获真实数据分布，生成的样本足以“欺骗”判别器。

GAN 在晶体材料生成领域的早期应用主要聚焦于化学组成的生成。CrystalGAN[54]作为最早将 GAN 应用于晶体材料的工作之一，证实了对抗训练框架在材料设计中的可行性。在此基础上，研究者们将 GAN 的应用拓展到更复杂的结构生成任务。ZeoGAN[55]专注于沸石材料的生成，基于 31 713 个已知沸石结构进行训练，成功生成 121 个新型多孔材料候选结构，并实现了定向生成用户所需吸附性能(如甲烷吸附热 18～22 kJ/mol)的沸石骨架设计。PGCGM[56]和 CGWGAN[57]则引入了条件生成机制，使模型能够根据目标性质或对称性约束生成定制化的晶体结构。

为了应对传统 GAN 在处理离散原子类型和连续原子坐标时面临的挑战，研究者们发展了多种改进策略。DeepCSP 将条件 GAN 与图注意力网络 MolGAT 相结合，用于有机晶体结构预测[58]。在该框架中，GAN 负责生成候选结构，而 MolGAT 则通过预测密度等物性参数对候选结构进行排序，从而筛选出最有可能存在的晶体结构。这种生成-评估的混合策略显著提升了结构预测的准确性。针对多目标优化问题，NSGAN[59]将 GAN 与遗传算法 NSGA-II 协同整合，在铝合金设计中实现了强度与延展性等相互冲突性质的同时优化，通过在 GAN 学习的潜空间中高效搜索 Pareto 最优解，在相互制约的性能指标间取得了最佳平衡。

为确保生成的晶体结构满足晶体对称性，如图 3b 所示，CGWGAN[57]采用基于 Wyckoff 位点的模块化生成策略，将生成过程分解为模板生成、原子填充和结构筛选 3 个阶段。该方法通过在生成框架中引入空间群约束，显著提升了生成结构的合理性。在 Ba–Ru–O 三元体系的应用中，CGWGAN 成功发现了 7 个新型稳定晶体，覆盖 4 个晶系、7 个空间群，所有结构的形成能均为负值且经声子谱计算验证具有动力学稳定性，为探索新型氧化物材料提供了有效途径。

GAN 训练过程中的稳定性问题一直是该领域的关键挑战之一。模式崩溃现象，即生成器倾向于生成少数几种相似的结构而忽略数据分布的多样性，在晶体材料生成中尤为突出。为此，GAN-DDLSF[60]提出了一种数据驱动的潜空间融合策略，通过将真实数据统计信息与高斯噪声混合，有效缓解了模式崩溃问题。该方法在氮化镓(GaN)二元体系中生成了 9321 个候选结构，其中 16.59%达到稳定状态、24.21%为亚稳态。经与 Materials Project 数据库[49]比对，仅发现 348 个重合结构，冗余率为 3.73%，显著提升了生成晶体结构的多样性并保证了生成样本与真实材料分布的一致性。此外，Wasserstein-GAN[61]等改进方法通过优化损失函数的设计，有效缓解了梯度消失和训练不稳定问题，显著提升了 GAN 在晶体材料生成任务中的稳定性和可靠性。

2.1.3 潜空间架构的局限与挑战 上述 VAE 和 GAN 的早期工作多采用各自构建的数据集和评估标准，缺乏统一的基准测试体系。例如，iMatGen[39]在 V–O 二元体系上验证生成能力；ZeoGAN[55]在 31 713 个沸石结构上训练，成功生成 121 个新型多孔材料，并实现了甲烷吸附热 18～22 kJ/mol 范围的定向设计；CGWGAN[57]则聚焦于 Ba–Ru–O 三元氧化物体系。这种评估基准的差异使得不同方法间的直接性能(如结构有效性、新颖性等指标)在学术上缺乏严格可比性。直到近年来，研究者才在统一基准数据集(如 Perov-5[41−42]、MP-20[49])上开展系统对比，为客观评估不同架构的性能提

供了可能。尽管如此，从方法论特性、理论基础和现有对比研究中，仍可获得对这两类架构优劣的深刻见解。

作为早期的生成式模型架构，基于潜空间的生成方法能够通过学习数据到连续低维潜空间的映射来建模复杂分布，但这些方法在生成质量和训练稳定性方面存在显著局限。例如，VAE 通过优化似然函数进行显式建模，能够提供清晰的概率解释和连续平滑的潜空间，却因重构损失的平均化特性，导致生成的晶体结构原子位置精度不足、局部配位环境模糊，无法精细捕捉晶格的几何特征；GAN 的隐式对抗训练虽能产生更精确的原子坐标和键长分布，却易受训练不稳定和模式崩溃影响，这限制了其可靠应用[62]。

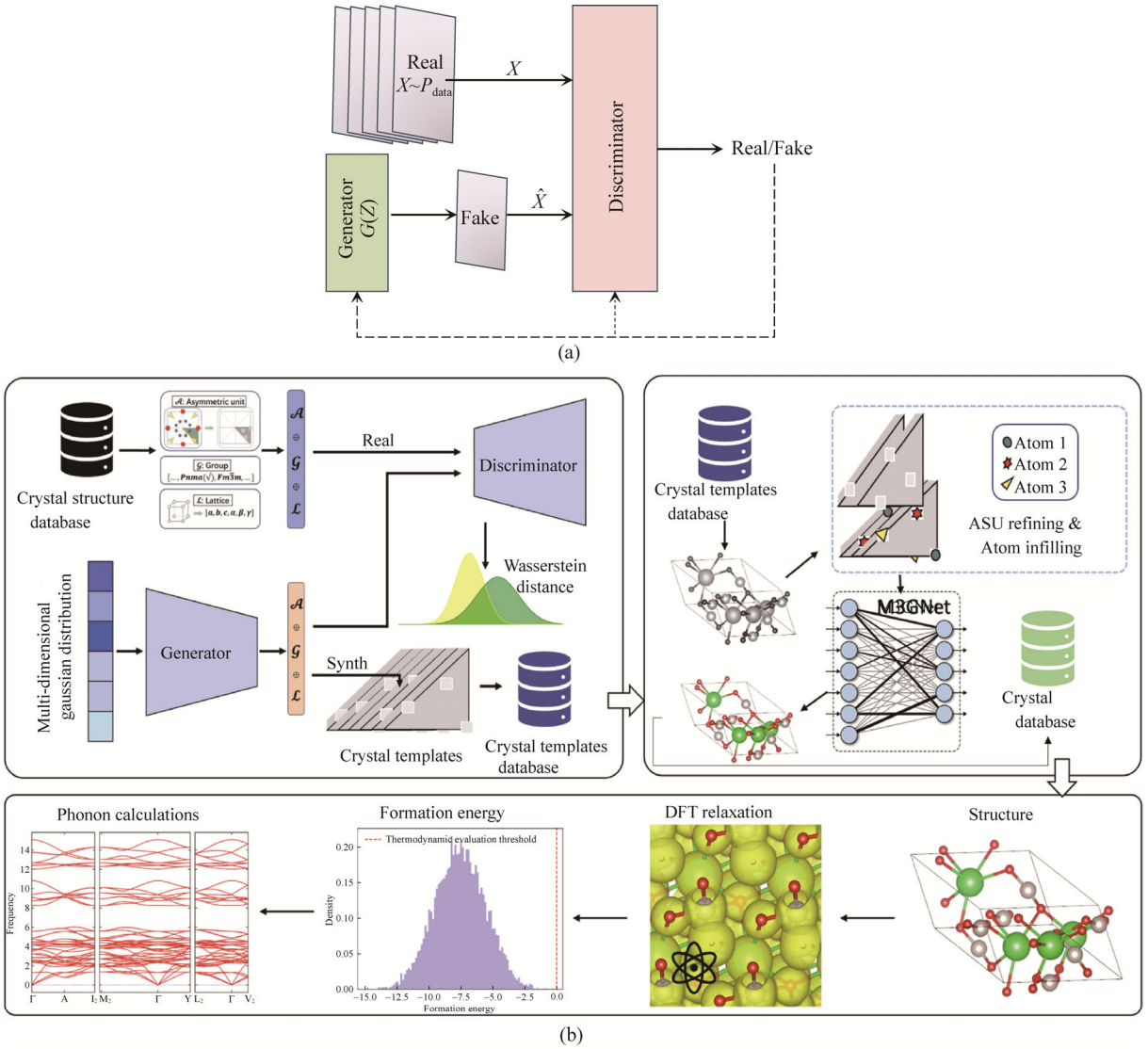


图 3　GAN 架构。(a) GAN 的基本框架：生成器从随机噪声生成候选样本，判别器区分真实样本与生成样本，通过对抗训练优化生成质量；(b) CGWGAN 的模型架构：采用基于 Wyckoff 位点的模块化生成策略，通过模板生成、原子填充和结构筛选三阶段实现空间群约束下的晶体生成[57]

Fig. 3　GAN based generative architecture. (a) Basic framework of GAN: The generator creates candidate samples from random noise while the discriminator distinguishes between real and generated samples. Both networks are optimized through adversarial training to improve generation quality. (b) Architecture of CGWGAN: The model employs a modular generation strategy based on Wyckoff positions, achieving crystal generation under space group constraints through three stages: template generation, atom filling, and structure screening[57]

在晶体材料生成实践中，这两类早期方法在处理晶体固有的周期性、对称性以及保证生成结构的物理合理性方面仍面临诸多挑战。首先，体素化网格或点云等表示法虽便于 VAE 和 GAN 处理，但通常缺乏旋转和平移不变性，需要通过大量数据增强才能学习到这些等价关系。其次，直接在像素空间

或坐标空间生成容易导致原子重叠、键长不合理等物理上不可行的结构。此外，早期模型在满足空间群对称性约束方面能力有限，生成的结构往往需要经过大量处理后才能符合晶体学规则。这些局限性促使研究者发展了基于扩散模型和自回归模型的新一代架构。这些新方法通过更精细的生成过程控制和更强的对称性约束，在很大程度上克服了基于 VAE 和 GAN 的晶体生成架构存在的不足[40, 63]。

2.2 扩散与流模型架构

2.2.1 扩散模型 扩散模型是一类强大的深度生成框架，其核心思想源于非平衡热力学[64]。该模型通过学习逆转一个固定的前向加噪过程来实现数据生成，在前向过程中，真实数据 x_0 经过 T 个离散时间步逐渐添加高斯噪声，直至数据分布 $q(x_T)$收敛到各向同性的高斯分布$\mathcal{N}(0，I)$。

$$q\left(x_t|x_{t-1}\right)=N\left(x_t;\sqrt{1-\beta_t}x_{t-1},\ \beta_tI\right) \tag{3}$$

式中：$q\left(x_t|x_{t-1}\right)$ 表示给定第 $(t-1)$ 步数据 x_{t-1} 条件下第 t 步数据 x_t 的条件概率分布；x_t 和 x_{t-1} 分别表示第 t 和第$(t–1)$时间步的数据状态；$N\left(x_t;\sqrt{1-\beta_t}x_{t-1},\ \beta_tI\right)$ 表示均值为$\sqrt{1-\beta_t}x_{t-1}$，协方差矩阵为 β_tI 的高斯分布；$\beta_t\in(0,1)$ 为第 t 步的噪声调度参数，控制每步添加的噪声量；I 为单位矩阵。等式右边也可表示为 $\mathcal{N}\left(x_t;\sqrt{\alpha_t}x_{t-1},(1-\alpha_t)I\right)$，其中 $\alpha_t=1-\beta_t$。模型的核心任务是学习反向去噪过程 $p_\theta(x_{t-1}\,|\,x_t)$，从纯噪声中迭代重构数据[64, 65]。这一过程通常通过训练去噪神经网络 ε_θ 来实现，该网络用于预测任意时间步 t 所添加的噪声ε。训练目标可简化为式(4)。

$$L^*=E_{t,\ x_0,\ \varepsilon}\times\left[\left\|\varepsilon-\varepsilon_\theta\left(x_t,t\right)\right\|^2\right] \tag{4}$$

式中：L^* 表示训练损失；$E_{t,\ x_0,\ \varepsilon}$ 表示对时间步 t，初始干净数据 x_0 和噪声 ε 的联合期望；$\varepsilon\sim\mathcal{N}(0,I)$ 为从标准正态分布中采样的高斯噪声；$\varepsilon_\theta(x_t,t)$ 为参数化的噪声预测网络，其参数为 θ，用于预测给定噪声数据 x_t 和时间步 t 时添加的噪声；$\left\|\varepsilon-\varepsilon_\theta(x_t,t)\right\|^2$ 表示预测噪声与真实噪声之间的均方误差。其中，$x_t=\sqrt{\overline{\alpha_t}}x_0+\sqrt{1-\overline{\alpha_t}}\varepsilon$，$\varepsilon\sim\mathcal{N}(0,I)$ 是采样的噪声数据，$\overline{\alpha_t}$ 为累积的噪声调度系数，定义为 $\overline{\alpha_t}=\prod_{i=1}^{t}\alpha_i$(即从第 1 步到第 t 步所有 α_i 的连乘积)。生成新样本时，从纯噪声 $x_t\sim\mathcal{N}(0,I)$ 出发，利用学习到的去噪函数 ε_θ 迭代反向推导，最终得到干净样本 x_0。该框架还衍生出多个变体，包括基于分数匹配和朗之万动力学(SMLD)的模型，以及将离散过程推广到连续时间的随机微分方程模型。

扩散模型已迅速成为晶体结构生成领域的主流方法。该领域的开创性工作 CDVAE[43]将变分自编码器与基于分数匹配的扩散解码器相结合，如图 4a 所示，利用周期性图神经网络实现了晶体结构生成。在实验中，该模型针对钙钛矿、碳结构和通用无机晶体材料数据集(MP-20)进行了测试。特别是在 MP-20 数据集上，CDVAE 生成了具有高度真实性的结构，其有效性达到了 100%，覆盖率(COV-R)超过 99%，显著优于之前的 Cond-DFC-VAE[40]等方法，证实了扩散模型在捕捉复杂晶体结构分布方面的能力。

然而，CDVAE 在保持平移不变性和生成高对称性结构方面存在一定局限。为突破这些瓶颈，后续研究开展了多方面改进：LCOM[66]在潜空间中优化形成能以生成更稳定的结构；SyMat[67]增强了平移不变性；GemsDiff[68]允许晶格参数在扩散过程中动态演化；DP-CDVAE[69]采用扩散概率模型(DDPM)替换分数网络以生成更接近基态的结构；TransVAE-CSP 通过 Transformer 增强编码器捕捉长程依赖[70]；Cond-CDVAE[71]等实现了基于组成、压力和属性的条件生成。

另一研究路线则摒弃了两阶段流程，提出联合生成范式，直接对晶体的所有组成部分同时进行扩散。DiffCSP 及其改进版本 DiffCSP++是这一范式的代表：前者利用周期性 $E(3)$(包含旋转、反映和平移的欧几里得群)等变去噪模型对晶格参数、分数坐标和原子类型进行联合扩散，后者进一步通过显式地集成空间群或 Wyckoff 位置约束来增强编码的对称性[72–73]。在此基础上，EquiCSP[74]纳入了晶格置换等变性和周期性质心不变性，SymmCD[75]实现了位点对称性的自动发现。架构创新方面，EH-Diff 采用超图结构以更好地捕捉多原子相互作用[76]；MOFDiff[77]则将扩散模型扩展到金属有机框架(MOF)的生成，在碳捕获材料设计中，该模型生成的前 10 位候选结构中有 9 个性能超越基准 MOF(Al-PMOF)，且筛选效率提升 75%。

MatterGen[13]代表了该领域的集大成之作。如图 4b 所示，该模型在超过 60 万个晶体结构上训练，完全摒弃了变分自编码器框架，实现了对原子类型、坐标和晶格参数的完全联合扩散。通过在分数网络中引入轻量级适配器模块，MatterGen 实现了对化学组成、空间群、力学性质、电子性质和磁性质等多种条件的灵活调控。在实际材料

设计中，该模型展现了跨越不同功能领域的强大搜索能力：在力学领域，MatterGen 成功设计并指导合成了具有无序结构的超硬材料 $TaCr_2O_6$，其实验测得的体模量在 158～169 GPa 的环境下与预测值高度吻合；在磁性材料方面，模型在极少训练数据下发现了 $MnFe_3O_4$ 等低供应链风险的高性能永磁体候选物。实验表明，MatterGen 生成的结构在稳定性、新颖性和多样性方面均显著优于先前方法，其生成结构距局部能量最小值的距离比仅为 CDVAE 的十分之一。

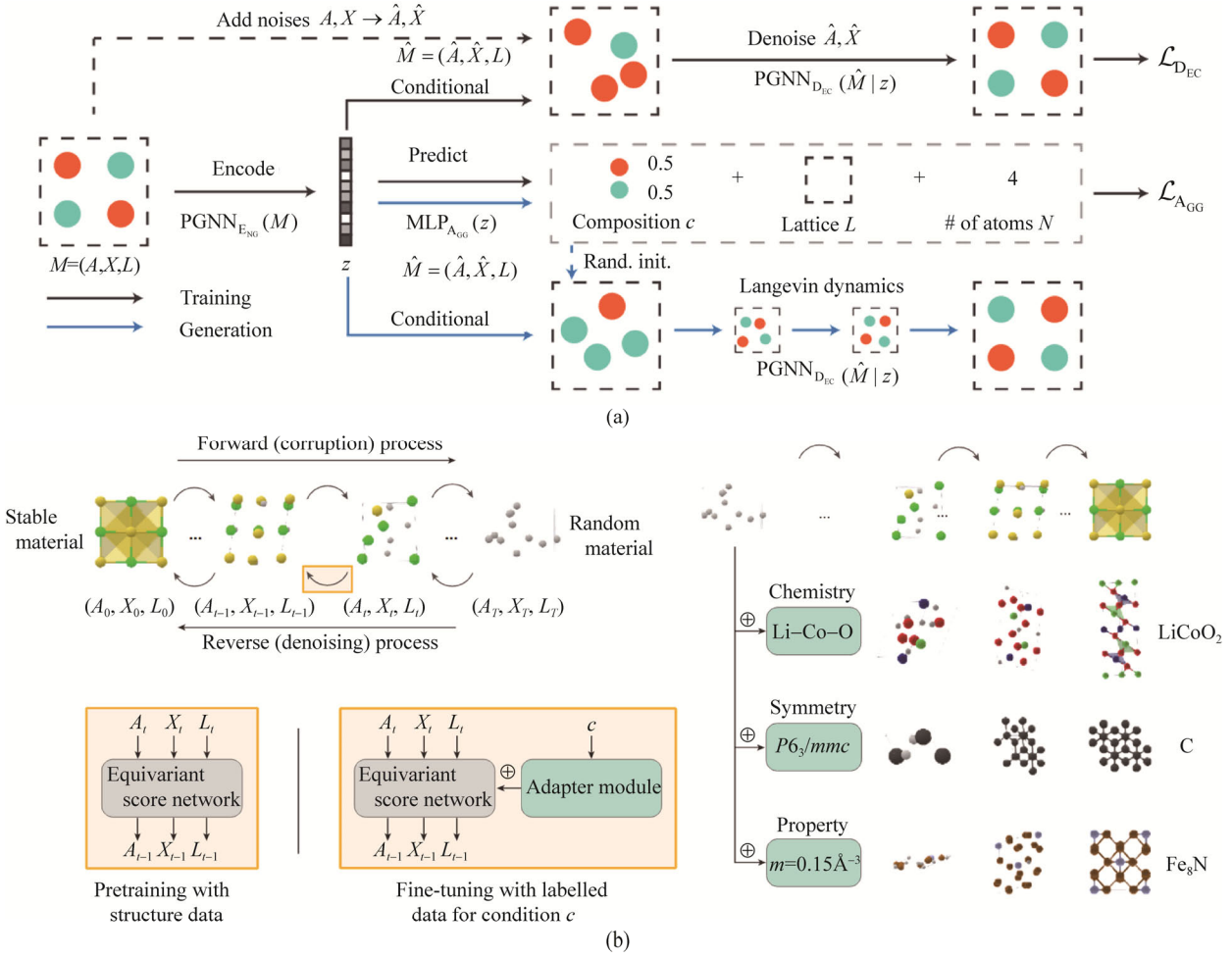


图 4　扩散生成模型架构。(a) CDVAE 的模型架构：结合 VAE 与扩散模型，先通过 VAE 生成晶格参数和初始组成，再通过扩散过程联合优化原子类型与坐标[43]；(b) MatterGen 的条件生成策略：基础扩散模型可通过轻量级适配器进行微调，实现基于化学组成、空间群及材料性质的多约束条件生成[13]

Fig. 4　Typical methods of diffusion models. (a) Architecture of CDVAE: The model combines VAE with diffusion processes, first generating lattice parameters and initial composition *via* VAE, then jointly optimizing atomic types and coordinates through diffusion[43]; (b) Conditional generation strategy of MatterGen: The foundational diffusion model can be fine-tuned with lightweight adapters to achieve multi-constraint conditional generation based on chemical composition, space groups, and material properties[13]

尽管扩散模型在理想晶体结构生成方面取得了巨大成功，但针对含缺陷或组分无序晶体的扩散生成模型仍处于起步阶段。现有的主流扩散模型(如 CDVAE[43]、MatterGen[13]等)通常假设晶体位点被原子完全占据，难以直接对部分占据或空位缺陷进行显式建模。这主要是因为传统的基于图或坐标的表示方法无法同时兼顾对称性保持与部分占据率的精确编码[31]。目前，针对无序结构的生成仍主要依赖于改进的 VAE 架构(如 Dis-GEN[31]，其通过基于 Wyckoff 位点的表示策略显式编码部分占据率)或基于超胞的近似处理。未来，开发能够兼容部分占据率和缺陷特征的扩散模型，将是实现真实材料逆向设计的关键突破口。

2.2.2　流模型　　流模型是一类重要的生成模型，其中标准化流(Normalizing Flow)与扩散模型在连续时间域具有深刻的理论关联。标准化流的核心思想

是通过一系列可逆可微变换，将简单的先验分布精确映射至复杂的数据分布，其对数似然可通过变量替换公式直接计算，如式(5)所示。

$$\ln p_x(x)=\ln p_z\left(f_\theta^{-1}(x)\right)+\ln\left|\det\left(\frac{\partial f_\theta^{-1}(x)}{\partial x}\right)\right| \tag{5}$$

式中：$p_x(x)$表示目标数据空间中 x 的概率密度；p_z 表示基础先验分布；$f_\theta^{-1}(x)$ 表示参数为 θ 的逆变换函数，将数据空间中的样本 x 映射到潜在空间；$\frac{\partial f_\theta^{-1}(x)}{\partial x}$ 表示逆变换 f_θ^{-1} 关于 x 的雅可比矩阵；$\left|\det\left(\frac{\partial f_\theta^{-1}(x)}{\partial x}\right)\right|$ 为雅可比行列式的绝对值，用于校正变换引起的体积变化。通过组合多个简单的可逆变换，标准化流能够在保持精确似然计算能力的前提下构建复杂的数据分布模型。

作为流模型的最新发展，流匹配(Flow Matching)在方法论层面与连续扩散模型中的概率流常微分方程本质上是等价的。流匹配的目标是直接学习一条确定性的连续时间常微分方程轨迹 $\mathrm{d}x/\mathrm{d}t=v_\theta(x,t)$，从而构建从噪声到数据的速度场[78]。其中 $v_\theta(x,t)$为神经网络参数化的速度场函数；θ为神经网络的可学习参数。这使其兼具流模型高效、确定性采样的特点。

流匹配框架的提出推动晶体逆向设计领域涌现出 FlowMM[79]、CrystalFlow[80]等创新模型。FlowMM 是首个将流匹配引入晶体生成的工作，该模型通过在体现晶体几何与周期性的流形上学习对称性编码向量场，使推理速度较传统扩散模型提升数倍。在材料稳定性方面，FlowMM 生成的晶体结构能量更接近局部能量最小值，其 DFT 弛豫后的 E_{hull} 显著低于 CDVAE，且在探索新型稳定材料时展现出更高的计算效率。CrystalFlow 则将连续流匹配与图神经网络相结合，不仅实现了更高效的晶体结构生成，还支持基于压力等外部条件的灵活调控，为极端环境下的材料设计提供了有效工具。这类工作表明，融合流模型与扩散机制的连续时间生成架构既保持了流模型采样的高效性和确定性，又继承了扩散模型训练的稳定性与强大的生成能力，已成为当前高质量晶体结构生成的重要技术路线。

2.3 自回归模型架构

自回归模型将晶体结构生成问题转化为序列预测任务。该架构将数据序列 x (例如，原子类型、晶格参数或原子坐标的序列)的联合概率分布 $p(x)$分解为一系列有序条件概率的乘积，如式(6)所示。

$$p(x)=\prod_{i=1}^{N}p(x_i\mid x_{<i}) \tag{6}$$

式中：x_i 为序列中的第 i 个词元；$x_{<i}$ 为序列 i 之前的序列；N 为序列总长度。训练过程中，模型采用因果掩码机制，词元 x_i 的预测仅依赖于其之前的序列 $x_{<i}$，从而实现自回归特性。

该架构通常以 Transformer 网络为核心。Transformer 通过自注意力机制高效捕获序列元素间的长程依赖关系[81–83]。自注意力机制通过计算查询(Q)、键(K)和值(V)之间的相似度来分配权重，其核心运算为缩放点积注意力，如式(7)所示。

$$\mathrm{Attention}(Q,K,V)=\mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{7}$$

式中：V 为值矩阵，存储实际要传递的信息内容；K^T 表示键矩阵 K 的转置矩阵；d_k 为键向量的维度。

早期工作尝试将 CIF 或 XYZ 坐标文件等标准格式直接作为文本序列进行处理。Pitike[84]等训练了一个 GPT-like 模型，证明了直接在 XYZ、CIF 等格式上训练的语言模型可以生成三维结构。如图 5a 所示，Antunes 等[24]开发的 CrystaLLM 在海量 CIF 数据上训练 Transformer 架构模型。在无条件生成实验中，该模型从 1000 次生成中发现了 102 个训练集中未见的新颖结构，经 DFT 验证后，其中 20 个结构的热力学能量凸包 E_{hull} 在 0.100 eV 内，包括 $Ba_4Na_2Ir_2O_{11}$、$NaAlS_2$ 和 Ca_2YSbO_6 3 个完全稳定的新化合物(E_{hull} = 0)，充分验证了自回归模型生成语法正确晶体结构的可行性。Gruver 等[23]通过微调 Llama 2-70B 模型(CrystalTextLLM)，在文本化点云描述基础上实现了基于自然语言提示的条件生成。然而，CIF 等表示方法本身缺乏平移、旋转和原子重排的物理不变性，这成为制约模型学习效率和泛化能力的关键因素。

为突破这一瓶颈，研究者开始探索能够嵌入物理先验知识的序列化策略。如图 5b 所示，CrystalFormer[14]是其中的代表性工作。该模型不直接生成原子坐标，而是生成由空间群、Wyckoff 位置和晶格参数构成的序列，通过对称性约束显著提升了结构预测准确性。Matra-Genoa[85] 采用类似思路，如图 5c 所示，基于 Wyckoff 位置构建可逆序列表示，并通过稳定性词元实现热力学稳定性的条件生成。在方法验证中，该模型针对 Al–Ca–Cu 三元体系生成 2000 个候选结构，成功预测了 Wang-Botti-Marques (WBM)数据集[86]中 15 个已知稳定结构中的 11 个，预测准确率达 73%，且搜索效率

比随机采样高数个数量级。在大规模应用中，Matra-Genoa 生成了 300 万个晶体结构，经机器学习势优化和筛选后，通过 ALIGNN 模型[87]估算，大量新型结构的 E_{hull} 低于 0.001 eV，展现了高效发现稳定新材料的潜力。这些方法将晶体对称性直接编码到生成流程中，从而保证生成结构严格遵守晶体学规则。

在 SLICES 表示法基础上[16]，研究者利用 RNN 模型[88]成功开展了直接窄带隙半导体的逆向设计。通过设定目标带隙(能带间隙) E_g 为(0.325 ± 0.225) eV、热力学能量凸包 E_{hull}(稳定性指标)小于 0.050 eV 等筛选标准，从约 1000 万个生成的 SLICES 字符串中最终发现了 14 种新型半导体材料，典型案例包括 Na_2Mg_2BiAs(E_g=0.43 eV，E_{hull} = 0)和 $Cs_2Zn_2As_2$(E_g = 0.49 eV，E_{hull} = 0)等。这些材料与训练集中已知结构的差异度均超过 0.75(Materials Project[49]采用的结构新颖性判别阈值)，表明 SLICES 能够生成真正新颖的晶体结构，而非已知材料的简单变体。

尽管取得了上述进展，但 RNN 架构在处理长 SLICES 字符串时仍面临挑战，约 66%的生成字符串因语法错误而无法重构为有效晶体结构。为突破这一瓶颈，MatterGPT[15]模型应运而生。作为基于 Transformer 的自回归模型，MatterGPT 表现出优异的生成性能：结构有效性超过 99%，结构独特性超过 99%，新颖性超过 50%。其核心创新在于通过在 SLICES 序列中拼接属性嵌入，实现了高效的按需逆向设计。由此，MatterGPT 能够直接生成同时满足多种目标属性(如晶格不敏感的形成能 E_f 和晶格敏感的带隙 E_g)的晶体结构，无需复杂的后处理或过滤。

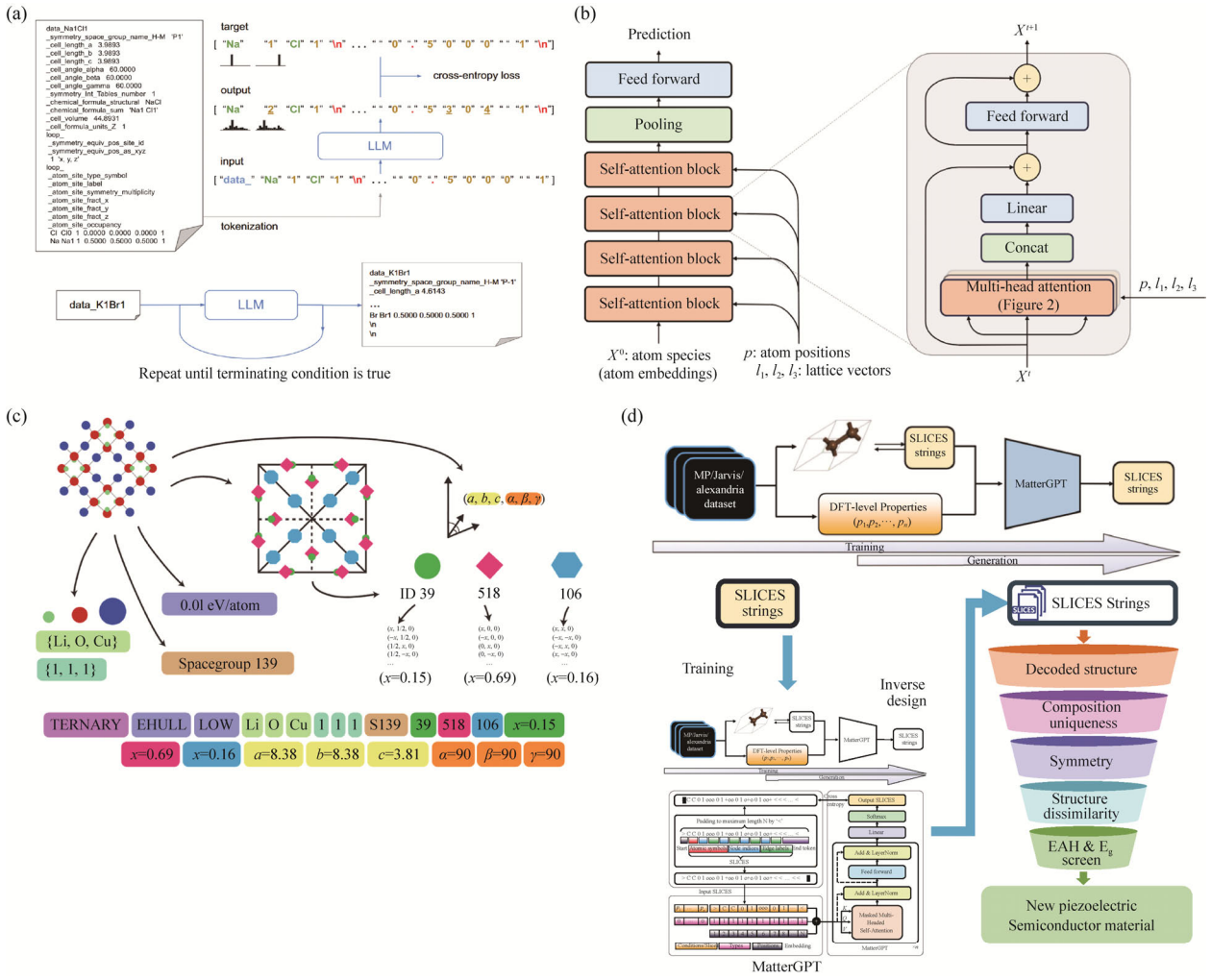


图 5　回归架构的代表性模型：(a) 基于 CIF 文件的大语言模型生成流程[24]；(b) CrystalFormer 模型架构[14]；(c) Matra-Genoa 的可逆序列表示方法[85]；(d) 利用 MatterGPT 生成压电材料的整体流程图[15]

Fig. 5　Representative autoregressive models for crystal generation: (a) Workflow of large language model trained on CIF files[24]; (b) Architecture of the CrystalFormer model[14]; (c) Invertible sequence representation approach of Matra-Genoa[85]; (d) Inverse design workflow for piezoelectric materials based on MatterGPT[15]

如图 5d 所示，MatterGPT 框架已成功应用于特定功能材料的发现。Zhang 等[89]利用该模型在仅含 1556 个样本的压电材料数据集上完成训练，成功实现了压电半导体的逆向设计。通过模型生成与筛选，他们发现了数据库中未记载的候选材料，如 Bi_4O_6，其压电应力系数高达 25.9 C/m^2，带隙为 1.85 eV，充分展示了该框架在小样本条件下发现高性能功能材料的潜力。针对 MatterGPT 在多目标优化中存在的不平衡性以及连续属性值学习的局限性，Deng 等[90]提出了基于属性提示的多尺度数据增强方案。该方案通过可学习模板将形成能、带隙等连续属性值映射至离散提示空间，并通过排列组合不同属性提示实现数据增强。实验结果表明，这种提示工程策略使单一模型在多目标生成任务中的表现更加均衡和鲁棒，特别是在数据稀疏区域，其性能可与独立训练的单目标模型相媲美甚至实现超越。

此外，大型语言模型的快速发展催生了混合架构和智能体系统。GenMS[91]采用分层策略，首先由模型预测化学式，随后由扩散模型生成三维结构；MatLLMSearch[92]将预训练大模型与进化搜索相结合，使模型作为结构生成器执行交叉和突变操作。MatExpert[93]则构建了对话式智能体，通过检索和修改建议来模拟人类专家的设计过程。

自回归模型架构经历了从直接处理 CIF、XYZ 等序列文件到融合 Wyckoff 位置等物理先验的演进历程。随着 SLICES 等可逆不变表示法的提出，MatterGPT 等模型对拓扑结构和材料属性的精确调控得以实现。这些进展标志着材料逆向设计正迈向高效、可控且可解释的新阶段，为计算材料科学的发展开辟了广阔前景。

3 总结与展望

尽管人工智能驱动的晶体材料逆向设计取得了显著进展，但该领域仍处于起步阶段，面临诸多挑战。本文从晶体表示方法、真实材料建模和基准测试 3 个层面，探讨该领域的发展趋势。

晶体表示方法的优化是首要问题。现有表示方法难以同时兼顾可逆性、不变性与周期性编码能力。例如，基于结构文件的表示方法虽易于进行结构重构，但不具备不变性；图神经网络等方法虽满足不变性要求，但可逆性难以保证。针对这一问题，未来研究应重点发展物理信息增强的表示方法，在保证可逆性的前提下，内禀地实现对称性与周期性编码能力。对于 SLICES 等拓扑字符串表示法，关键的改进策略是利用对称性进行信息压缩，通过引入 Wyckoff 位置等晶体学约束实现序列折叠，从而构建具有对称性编码能力的序列化表示方法，提升生成效率和结构的物理合理性。

真实材料的精确建模是第二个挑战。目前的研究主要集中于理想晶体的生成，而实际材料的性能很大程度上取决于缺陷以及服役环境(如温度、压力)等因素[94–96]。因此，未来研究需要将含缺陷和不同服役环境下的晶体结构纳入建模体系。这要求发展基于对称性约束的数据增强技术，并结合分子动力学模拟，构建包含缺陷构象的训练数据集，使模型能够设计出在特定条件下稳定且功能化的材料。

基准测试的标准化是第三个亟待解决的问题。与材料性质预测等正向任务已有成熟基准测试不同，针对逆向设计任务的评估体系尚未建立统一标准。现有评估方法多局限于基本的有效性验证，如检验生成结构是否满足物理或化学约束，而在稳定性、新颖性和计算效率等关键指标上缺乏统一标准。更严重的是，不同研究采用的评估基准各不相同，包括热力学稳定性评估的参考凸包、新颖性检查的参考数据集以及结构匹配算法等，导致模型性能评估比较缺乏客观性。因此，建立社区共同维护的基准测试体系势在必行。该体系应包含标准化的参考凸包、新颖性检查参考数据集以及结构比对算法，评估指标需涵盖多维度性能并明确计算成本，同时扩展至不同化学体系的条件生成任务[97]。

综上所述，本文讨论的晶体表示方法与晶体生成架构创新，最终目标是实现自动化、高效的材料逆向设计。尽管该领域已取得长足进展，但更优的晶体表示法、更精确的材料构效关系建模以及统一的评估基准仍是实现这一目标的关键。有效应对这些挑战，将推动该领域从学术研究走向工业应用。我们期待全球研究者贡献更多原创性成果，共同加速新材料的发现与应用。

参考文献:

[1] WALTERS W P, MURCKO M. Assessing the impact of generative AI on medicinal chemistry[J]. Nat Biotechnol, 2020, 38(2): 143–145.

[2] DU Y Q, JAMASB A R, GUO J, et al. Machine learning-aided generative molecular design[J]. Nat Mach Intell, 2024, 6(6): 589–604.

[3] WEININGER D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules[J]. J Chem Inf Comput Sci, 1988, 28(1): 31–36.

[4] HELLER S R, MCNAUGHT A, PLETNEV I, et al. InChI, the IUPAC international chemical identifier[J]. J Cheminform, 2015, 7: 23.

[5] KRENN M, HÄSE F, NIGAM A, et al. Self-referencing embedded strings (SELFIES): A 100% robust molecular string representation[J].

Mach Learn: Sci Technol, 2020, 1(4): 045024.
[6] ZHUNG W, KIM H, KIM W Y. 3D molecular generative framework for interaction-guided drug design[J]. Nat Commun, 2024, 15(1): 2688.
[7] JIANG Y Y, ZHANG G, YOU J, et al. PocketFlow is a data-and-knowledge-driven structure-based molecular generative model[J]. Nat Mach Intell, 2024, 6(3): 326–337.
[8] CHANG J, YE J C. Bidirectional generation of structure and properties through a single molecular foundation model[J]. Nat Commun, 2024, 15(1): 2323.
[9] WESTERMAYR J, GILKES J, BARRETT R, et al. High-throughput property-driven generative design of functional organic molecules[J]. Nat Comput Sci, 2023, 3(2): 139–148.
[10] PARK H, LI Z Z, WALSH A. Has generative artificial intelligence solved inverse materials design?[J]. Matter, 2024, 7(7): 2355–2367.
[11] BUTLER K T, DAVIES D W, CARTWRIGHT H, et al. Machine learning for molecular and materials science[J]. Nature, 2018, 559(7715): 547–555.
[12] WANG Z L, YOU F Q. Leveraging generative models with periodicity-aware, invertible and invariant representations for crystalline materials design[J]. Nat Comput Sci, 2025, 5(5): 365–376.
[13] ZENI C, PINSLER R, ZÜGNER D, et al. A generative model for inorganic materials design[J]. Nature, 2025, 639(8055): 624–632.
[14] TANIAI T, IGARASHI R, SUZUKI Y, et al. Crystalformer: Infinitely connected attention for periodic structure encoding[J]. arXiv preprint arXiv, 2024,2403.11686.
[15] CHEN Y, WANG X, DENG X, et al. MatterGPT: A generative transformer for multi-property inverse design of solid-state materials[J].arXiv preprint arXiv, 2024,2408.07608
[16] XIAO H, LI R, SHI X Y, et al. An invertible, invariant crystal representation for inverse design of solid-state materials using generative deep learning[J]. Nat Commun, 2023, 14(1): 7027.
[17] JIANG S L, DIENG A B, WEBB M A. Property-guided generation of complex polymer topologies using variational autoencoders[J]. NPJ Comput Mater, 2024, 10: 139.
[18] ANSTINE D M, ISAYEV O. Generative models as an emerging paradigm in the chemical sciences[J]. J Am Chem Soc, 2023, 145(16): 8736–8750.
[19] WANG K, GOU C, DUAN Y, et al. Generative adversarial networks: introduction and outlook[J]. IEEE/CAA J Automatic, 2017, 4(4): 588–598.
[20] KRENN M, POLLICE R, GUO S Y, et al. On scientific understanding with artificial intelligence[J]. Nat Rev Phys, 2022, 4(12): 761–769.
[21] HALL S R, ALLEN F H, BROWN I D. The crystallographic information file (CIF): A new standard archive file for crystallography[J]. Acta Crystallogr Sect A, 1991, 47(6): 655–685.
[22] CAO Z D, LUO X S, LV J, et al. Space group informed transformer for crystalline materials generation[J]. Sci Bull, 2025, 70(21): 3522–3533.
[23] GRUVER N, SRIRAM A, MADOTTO A, et al. Fine-Tuned language models generate stable inorganic materials as text[J]. arXiv E Prints, 2024:arXiv:2402.04379.
[24] ANTUNES L M, BUTLER K T, GRAU-CRESPO R. Crystal structure generation with autoregressive large language modeling[J]. Nat Commun, 2024, 15(1): 10570.
[25] CHUNG S J, HAHN T, KLEE W E. Nomenclature and generation of three-periodic nets: The vector method[J]. Acta Crystallogr A Found Crystallogr, 1984, 40(1): 42–50.
[26] WANG B N, XU Z Y, HAN Z Y, et al. SLICES-PLUS: A crystal representation leveraging spatial symmetry[J]. Mater Des, 2025, 253: 113856.
[27] ALAMPARA N, MIRET S, JABLONKA K M. MatText: Do language models need more than text & scale for materials modeling?[J]. arXiv preprint arXiv: 2406.17295, 2024.
[28] FLAM-SHEPHERD D, ASPURU-GUZIK A. Language models can generate molecules, materials, and protein binding sites directly in three dimensions as XYZ, CIF, and PDB files[J]. arXiv E Prints, 2023: arXiv: 2305.05708.
[29] LIU S, LI Y, LI Z, et al. Symmetry-informed geometric representation for molecules, proteins, and crystalline materials[C]//Advances in Neural Information Processing Systems, 2023, 36: 66084–66101.
[30] DAMEWOOD J, KARAGUESIAN J, LUNGER J R, et al. Representations of materials for machine learning[J]. Annu Rev Mater Res, 2023, 53(1): 399–426.
[31] PETERSEN M H, ZHU R M, DAI H W, et al. Dis-GEN: Disordered crystal structure generation[J]. arXiv: 2507.18275.
[32] YANG Z D, LIU X Q, ZHANG X Y, et al. Modeling crystal defects using defect informed neural networks[J]. NPJ Comput Mater, 2025, 11: 229.
[33] LIU F Z, CHEN Z T, LIU T Y, et al. Self-supervised generative models for crystal structures[J]. iScience, 2024, 27(9): 110672.
[34] YAMASHITA R, NISHIO M, DO R K G, et al. Convolutional neural networks: An overview and application in radiology[J]. Insights Imaging, 2018, 9(4): 611–629.
[35] ZILETTI A, KUMAR D, SCHEFFLER M, et al. Insightful classification of crystal structures using deep learning[J]. Nat Commun, 2018, 9(1): 2775.
[36] YANG S H, CHOI W, CHO B W, et al. Deep learning-assisted quantification of atomic dopants and defects in 2D materials[J]. Adv Sci, 2021, 8(16): 2101099.
[37] CHOUDHARY K, DECOST B, CHEN C, et al. Recent advances and applications of deep learning methods in materials science[J]. NPJ Comput Mater, 2022, 8: 59.
[38] JANGID D K, BRODNIK N R, ECHLIN M P, et al. Q-RBSA: High-resolution 3D EBSD map generation using an efficient quaternion transformer network[J]. NPJ Comput Mater, 2024, 10: 27.
[39] NOH J, KIM J, STEIN H S, et al. Inverse design of solid-state materials *via* a continuous representation[J]. Matter, 2019, 1(5): 1370–1384.
[40] COURT C J, YILDIRIM B, JAIN A, et al. 3-D inorganic crystal structure generation and property prediction *via* representation learning[J]. J Chem Inf Model, 2020, 60(10): 4518–4535.
[41] CASTELLI I E, LANDIS D D, THYGESEN K S, et al. New cubic perovskites for one- and two-photon water splitting using the computational materials repository[J]. Energy Environ Sci, 2012, 5(10): 9034–9043.
[42] CASTELLI I E, OLSEN T, DATTA S, et al. Computational screening of perovskite metal oxides for optimal solar light capture[J]. Energy Environ Sci, 2012, 5(2): 5814–5819.
[43] XIE T, FU X, GANEA O E, et al. Crystal diffusion variational autoencoder for periodic material generation[J]. arXiv E Prints, 2021: arXiv: 2110.06197.
[44] PEI H, WEI B, CHANG K C C, et al. Geom-GCN: Geometric graph convolutional networks[J]. arXiv E Prints, 2020: arXiv: 2002.05287.
[45] ZHAO W X, ZHOU K, LI J, et al. A survey of large language models[J]. arXiv E Prints, 2023: arXiv: 2303.18223, 1(2).
[46] CROITORU F A, HONDRU V, IONESCU R T, et al. Diffusion models in vision: A survey[J]. IEEE Trans Pattern Anal Mach Intell, 2023, 45(9): 10850–10869.
[47] YAN D, SMITH A D, CHEN C C. Structure prediction and materials design with generative neural networks[J]. Nat Comput Sci, 2023, 3(7): 572–574.

[48] KINGMA D P, WELLING M. Auto-encoding variational bayes[J]. arXiv E Prints, 2013: arXiv: 1312.6114.

[49] JAIN A, ONG S P, HAUTIER G, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation[J]. APL Mater, 2013, 1: 011002.

[50] LIU K, GAO S D, YANG K F, et al. PCVAE: A physics-informed neural network for determining the symmetry and geometry of crystals[C]//2023 International Joint Conference on Neural Networks (IJCNN). Gold Coast, Australia. IEEE, 2023: 1–8.

[51] MAL S, SEAL G, SEN P. MagGen: A graph-aided deep generative model for inverse design of permanent magnets[J]. J Phys Chem Lett, 2024, 15(12): 3221–3228.

[52] ZHU R M, NONG W, YAMAZAKI S, et al. WyCryst: Wyckoff inorganic crystal generator framework[J]. Matter, 2024, 7(10): 3469–3488.

[53] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Commun ACM, 2020, 63(11): 139–144.

[54] NOUIRA A, SOKOLOVSKA N, CRIVELLO J C. CrystalGAN: Learning to discover crystallographic structures with generative adversarial networks[J]. arXiv preprint arXiv. 2018, 1810.11203.

[55] KIM B, LEE S, KIM J. Inverse design of porous materials using artificial neural networks[J]. Sci Adv, 2020, 6(1): eaax9324.

[56] ZHAO Y, SIRIWARDANE E M D, WU Z Y, et al. Physics guided deep learning for generative design of crystal materials with symmetry constraints[J]. NPJ Comput Mater, 2023, 9: 38.

[57] SU T H, CAO B, HU S B, et al. CGWGAN: Crystal generative framework based on Wyckoff generative adversarial network[J]. J Mater Inf, 2024, 4(4): 20.

[58] YE Z, WANG N N, ZHOU J T, et al. Organic crystal structure prediction *via* coupled generative adversarial networks and graph convolutional networks[J]. Innovation, 2024, 5(2): 100562.

[59] LI Z, BIRBILIS N. NSGAN: A non-dominant sorting optimisation-based generative adversarial design framework for alloy discovery[J]. NPJ Comput Mater, 2024, 10: 112.

[60] CHEN Z A, LI H C, ZHANG C, et al. Crystal structure prediction using generative adversarial network with data-driven latent space fusion strategy[J]. J Chem Theory Comput, 2024, 20(21): 9627–9641.

[61] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks[C]//Proceedings of the 34th International Conference on Machine Learning – Volume 70. Sydney, NSW, Australia. ACM, 2017: 214–223.

[62] LONG T, FORTUNATO N M, OPAHLE I, et al. Constrained crystals deep convolutional generative adversarial network for the inverse design of crystal structures[J]. NPJ Comput Mater, 2021, 7: 66.

[63] CHEN L T, ZHANG W T, NIE Z W, et al. Generative models for inverse design of inorganic solid materials[J]. J Mater Inform, 2021, 1: 4.

[64] SOHL-DICKSTEIN J, WEISS E A, MAHESWARANATHAN N, et al. Deep unsupervised learning using nonequilibrium thermodynamics[C]// International Conference on Machine Learning, France, 2015: 2256–2265.

[65] HO J, JAIN A, ABBEEL P. Denoising Diffusion Probabilistic Models[C]//Advances in Neural Information Processing Systems, 2020, 33: 6840–6851.

[66] JANMOHAMED H, WOLINSKA M, SURANA S, et al. Multi-objective quality-diversity for crystal structure prediction[C]// Proceedings of the Genetic and Evolutionary Computation Conference, 2024: 1273–1281.

[67] Towards symmetry-aware generation of periodic materials[C]// Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, LA, USA. ACM, 2023: 53308–53329.

[68] KLIPFEL A, FREGIER Y, SAYEDE A, et al. Vector field oriented diffusion model for crystal material generation[J]. Proc AAAI Conf Artif Intell, 2024, 38(20): 22193–22201.

[69] PAKORNCHOTE T, CHOOMPHON-ANOMAKHUN N, ARRERUT S, et al. Diffusion probabilistic models enhance variational autoencoder for crystal structure generative modeling[J]. Sci Rep, 2024, 14(1): 1275.

[70] CHEN Z, YUAN Y, ZHENG S, et al. Transformer-enhanced variational autoencoder for crystal structure prediction[J]. arXiv E Prints, 2025: arXiv: 2502.09423.

[71] LUO X S, WANG Z Y, GAO P Y, et al. Deep learning generative model for crystal structure prediction[J]. NPJ Comput Mater, 2024, 10: 254.

[72] Crystal structure prediction by joint equivariant diffusion[C]// Proceedings of the 37th International Conference on Neural Information Processing Systems. ACM, 2023: 17464–17497.

[73] JIAO R, HUANG W, LIU Y, et al. Space group constrained crystal generation[J]. arXiv E Prints, 2024: arXiv: 2402.03992.

[74] Equivariant diffusion for crystal structure prediction[C]//Proceedings of the 41st International Conference on Machine Learning. ACM, 2024: 29890–29913.

[75] LEVY D, PANIGRAHI S S, KABA S O, et al. SymmCD: Symmetry-preserving crystal generation with diffusion models[J]. arXiv E Prints, 2025: arXiv: 2502.03638.

[76] LIU Y, ZHOU C, ZHANG S, et al. Equivariant hypergraph diffusion for crystal structure prediction[J]. arXiv E Prints, 2025: arXiv: 2501.18850.

[77] FU X, XIE T, ROSEN A S, et al. MOFDiff: Coarse-grained diffusion for metal-organic framework design[J]. arXiv E Prints, 2023: arXiv: 2310.10732.

[78] SONG Y, SOHL-DICKSTEIN J, KINGMA D P, et al. Score-based generative modeling through stochastic differential equations.[J]. arXiv preprint arXiv, 2020, 2011.13456.

[79] MILLER B K, CHEN R T Q, SRIRAM A, et al. FlowMM: Generating materials with Riemannian flow matching[J]. arXiv preprint arXiv, 2024, 2406.04731.

[80] LUO X S, WANG Z Y, WANG Q C, et al. CrystalFlow: A flow-based generative model for crystalline materials[J]. Nat Commun, 2025, 16(1): 9267.

[81] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[C]//North American Chapter of the Association for Computational Linguistics., 2019

[82] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is All you Need[C]//Advances in neural information processing systems, 2017: 30.

[83] HU Y, BUEHLER M J. Deep language models for interpretative and predictive materials science[J]. APL Machine Learning, 2023, 1(1).

[84] PITIKE K C, MACIAS A, EISENBACH M, et al. Computationally accelerated discovery of high entropy pyrochlore oxides[J]. Chem Mater, 2022, 34(4): 1459–1472.

[85] BREUCK P P D, PIRACHA H A, RIGNANESE G M, et al. A generative material transformer using Wyckoff representation[J]. arXiv preprint arXiv, 2025, 2501.16051.

[86] BERGERHOFF G, HUNDT R, SIEVERS R, et al. The inorganic crystal structure data base[J]. J Chem Inf Comput Sci, 1983, 23(2): 66–69.

[87] CHOUDHARY K, DECOST B. Atomistic Line Graph Neural

Network for improved materials property predictions[J]. NPJ Comput Mater, 2021, 7: 185.

[88] ELMAN J L. Finding structure in time[J]. Cognitive science. 1990, 14(2): 179–211.

[89] ZHANG C, LV S Y, GONG H J, et al. Inverse design of high-performance piezoelectric semiconductors *via* advanced crystal representation and large language models[J]. Appl Phys Lett, 2025, 126(11): 111901.

[90] DENG Z, LI S, ZHANG T, et al. An Property-prompted multi-scale data augmentation approach for crystal representation[C]//ICLR 2025 Workshop on Machine Learning Multiscale Processes, 2025.

[91] AYKOL M, BATZNER S, CUBUK E, et al. Generative hierarchical materials search[C]//Advances in Neural Information Processing Systems 37. Vancouver, BC, Canada. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2024: 38799–38819.

[92] GAN J, ZHU Y, SCHWALBE-KODA D, et al. Large language models are innate crystal structure generators[C]//AI for Accelerated Materials Design-ICLR 2025, 2025.

[93] DING Q, MIRET S, LIU B, et al. Matexpert: Decomposing materials discovery by mimicking human experts[J]. arXiv preprint arXiv. 2024, 2410.21317.

[94] ZHONG K, BU R, JIAO F, et al. Toward the defect engineering of energetic materials: A review of the effect of crystal defects on the sensitivity[J]. Chem Eng J, 2022, 429: 132310.

[95] GU Q Q, ZHOUYIN Z, PANDEY S K, et al. Deep learning tight-binding approach for large-scale electronic simulations at finite temperatures with *ab initio* accuracy[J]. Nat Commun, 2024, 15(1): 6772.

[96] YANG M Y, RAUCCI U, PARRINELLO M. Reactant-induced dynamics of lithium imide surfaces during the ammonia decomposition process[J]. Nat Catal, 2023, 6(9): 829–836.

[97] DE BREUCK P P, WANG H C, RIGNANESE G M, et al. Generative AI for crystal structures: A review[J]. NPJ Comput Mater, 2025, 11: 370.

AI-Driven Materials Discovery: Generative Models and Progress in Crystal Inverse Design

WU Kedeng[1], *CHEN Yan*[2], *DENG Xiaobin*[3], *CHEN Xi*[3], *XIAO Hang*[3]

(1. School of Chemical Engineering, Northwest University, Xi'an 710127, China; 2. School of Aerospace Engineering, Xi'an Jiaotong University, Xi'an 710049, China; 3. Wu Jieh Yee School of Interdisciplinary Studies, Lingnan University, Hong Kong, China)

Extended Abstract

The inverse design of crystalline materials aims to generate novel structures with targeted properties. However, achieving this objective remains a long-standing challenge in materials science. Unlike molecular systems that benefit from mature representations such as SMILES, crystalline materials impose stringent physical constraints, requiring representations that are simultaneously periodicity-aware, invertible, and invariant under symmetry operations. Emerging generative artificial intelligence (AI) technologies present transformative opportunities to address these complexities.

This mini-review systematically examines the latest advancements in this field, centering on two critical pillars: crystal representation methods and generative model architectures. In terms of crystal representations, we critically evaluate text-based symbolic encodings (*e.g.*, SLICES), geometry-based periodicity-aware graphs, and pixel-based imaging approaches. These methods are analyzed based on their ability to balance invertibility with symmetry preservation. With respect to generative model architectures, we trace the technological evolution from latent space-based models, such as Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN), to state-of-the-art frameworks including diffusion models, flow-matching, and Transformer-based autoregressive models (*e.g.*, MatterGen, MatterGPT). The review highlights how these architectures integrate physical constraints to navigate the vast chemical space of crystalline materials efficiently.

Summary and Prospects Despite significant progress, the field of AI-driven crystal design faces critical bottlenecks. A primary challenge lies in the absence of a universally accepted representation framework that simultaneously satisfies invertibility, invariance, and periodicity encoding capabilities. Furthermore, current generative models predominantly focus on ideal crystal structures, often neglecting real-world complexities such as defects, impurities, and service environments. To bridge this gap and facilitate the transition from theoretical exploration to practical application, future research should focus on three key directions: 1) Physics-informed representation design: Developing representations that intrinsically embed crystallographic constraints, such as Wyckoff positions and space groups, to ensure symmetry-aware and invertible encodings. 2) High-fidelity modeling of real-world materials: Extending beyond idealized lattice structures to model realistic defects and environmental responses (*e.g.*, temperature, pressure) by integrating symmetry-constrained data augmentation with molecular dynamics simulations. 3) Standardization of benchmarks: Establishing community-wide benchmarks that rigorously evaluate generation stability, novelty, diversity, and computational cost, ensuring objective comparisons across different chemical systems. Addressing these challenges will be pivotal in realizing the automated, intelligent design of functional crystalline materials.

Keywords inverse design; crystal representation; generative models; artificial intelligence; crystal materials