

생성 모델을 활용한 물성 기반 결정 구조 생성 기법 조사

황영서, 황효석, 석준희*
고려대학교 전기전자공학과

youngseoh6@korea.ac.kr, shdhkj960@korea.ac.kr, *jseok14@korea.ac.kr

A Survey on Property Guided Crystal Structure Generation via Generative Models

YoungSeo Hwang, Hyo-Seok Hwang, Junhee Seok*
Korea Univ. Electrical Engineering

요약

최근 소재의 설계 및 개발은 다양한 산업 분야에서 핵심 기술로 주목받고 있으며, 특정 물성을 갖는 재료를 효율적으로 탐색하고 확보할 수 있는지가 기술 경쟁력을 좌우하고 있다. 그러나 전통적인 실험 기반 탐색 방식은 물리적 시뮬레이션에 많은 자원이 소요되기 때문에 여전히 한계가 존재한다. 이러한 문제를 극복하기 위한 대안으로, 최근에는 생성 모델을 활용하여 원하는 물성 조건을 만족하는 결정 구조를 직접 생성하는 접근이 활발히 연구되고 있다. 본 연구는 결정 구조 생성을 위한 대표적인 생성 모델인 variational autoencoder(VAE)와 확산 기반 모델(diffusion model)의 구조적 특징과 학습 방식을 살펴보고, 조건 기반 생성을 위한 임베딩 기법과 물성 조건 반영 방식에 대해 정리한다. 특히, 결정 구조의 핵심 요소에 노이즈를 추가하고 이를 복원하는 확산 기반 접근과, 조건 임베딩을 구조 생성 과정에 통합하는 다양한 전략을 중심으로, 물성 조건 기반 결정 구조 생성 기술의 가능성을 제시한다.

1. 서론

더 나은 재료 소재를 얼마나 빠르게 발견할 수 있는지는 반도체 설계 및 기술의 발전 속도에 큰 영향을 미친다. 전통적으로는 실험과 사람의 직관에 기반하여 새로운 소재를 탐색해왔으나, 이러한 방식은 탐색 가능한 후보 수의 한계와 긴 소요 시간이라는 한계점을 지닌다.

이를 보완하기 위해 최근에는 수십만 개의 후보 물질을 컴퓨터 시뮬레이션으로 자동 평가하는 고속 탐색 (high-throughput screening) 기법이 활용되고 있다. 하지만 이러한 방식조차도 현재까지 밝혀진 안정한 소재 구조는 약 100만개 수준에 불과하며, 이는 이론적으로 가능할 수 많은 조합에 비해 극히 일부에 해당한다.

특히 반도체와 같은 첨단 산업 분야에서는 아주 미세한 물성 차이가 제품의 성능을 좌우하기 때문에, 단순히 새로운 소재 구조를 생성하는 것을 넘어 목표 물성을 만족하는 구조를 직접 설계하는 역설계적 접근의 중요성이 더욱 커지고 있다.

이러한 필요성을 바탕으로, 최근에는 사람의 직관이나

반복적인 수작업에 의존하던 기존 소재 발굴 방식의 한계를 극복하고, 구조 생성 과정을 자동화할 수 있는 인공지능 기반 접근으로서 생성 모델이 주목받고 있으며, 이를 활용한 연구가 활발히 진행되고 있다. 생성 모델은 완전히 새로운 소재 구조를 효율적으로 탐색할 수 있을 뿐 아니라, 전이 학습 혹은 조건 기반 생성을 통해, 특정 화학 조성이나 물성 조건을 만족하는 구조를 직접 생성할 수 있는 가능성을 제시한다.

2. 예비 지식

variational autoencoder는 데이터를 인코더에 입력하여 잠재 벡터로 매핑하고, 디코더는 이 잠재벡터를 기반으로 새로운 샘플을 복원하는 모델이다. diffusion 모델은 확률 생성 모델로, 데이터에 점진적으로 노이즈를 추가한 후, 이를 역으로 제거하면서 새로운 샘플을 생성하는 모델이다.

결정 구조 생성을 위한 대표적인 시도로는 crystal diffusion variational autoencoder(CDVAE) [1]가 있다.

이는 VAE 기반의 구조로, 디코더는 노이즈가 추가된 구조로부터 원래 구조를 복원하는 denoising 학습 방식으로, 보다 안정적이고 일반화된 결정 구조 생성을 가능하게 한다.

3. 물성 기반 결정 구조 생성 기법

최근에는 물성 조건을 반영하여 결정 구조를 생성하는 다양한 모델들이 제안되고 있으며, 그 접근 방식에 따라 네트워크 구조와 조건 반영 방법이 다양하게 나타난다. 다음에서는 세 가지 연구 사례를 소개한다.

3.1. Conditional Crystal Diffusion Variational Auto-encoder

[2]에서는 CDVAE 구조를 확장한 모델을 제안한다. 연속형과 이산형 물성 조건은 각각 gaussian basis 함수와 학습 가능한 벡터를 통해 임베딩되며, 이들은 multi-layer perceptron을 통해 통합되어 조건 임베딩으로 사용된다. 해당 조건 임베딩은 잠재 벡터와 결합되어 디코더에 전달되며, 동시에 속성 예측을 위한 predictor 모듈을 통해 잠재 공간이 물성 조건을 반영하도록 학습된다. 이후, 주어진 물성 조건에 따라 잠재 벡터를 직접 생성하기 위해 denoising diffusion probabilistic models (DDPM) 기반의 prior 모듈이 학습되며, 이를 통해 조건에 따라 다양한 결정 구조를 생성할 수 있다.

3.2. MatterGen

[3]에서는 구조의 대칭성과 주기성을 반영하기 위해, 격자 벡터, 원자 좌표, 원자 종류에 독립적인 노이즈를 독립적으로 추가하고, 이를 복원하는 equivariant score network를 사전 학습한다. 이 네트워크는 물리적 대칭을 보존하면서 각 요소의 denoising 방향을 예측하도록 설계된다. 이후, 물성 정보가 포함된 데이터셋을 활용하여 adapter module을 통한 fine-tuning 되며, 이는 classifier-free-guidance 기법과 결합되어, 목표 물성 조건을 더욱 강하게 반영하는 구조 생성을 가능하게 한다.

3.3. Generative Inverse Design of Crystal Structures via Diffusion Models with Transformers

[4]에서는 결정 구조의 역설계를 위해 격자 벡터, 원자 좌표, 원자 종류에 각각 확산 모델 DDPM, score-matching based models, discrete DDPM을 적용한 joint diffusion framework를 제안한다. 복원 과정은 transformer 기반의 self-attention을 통해 세 가지 요

소의 정보를 합쳐 통합하여 수행된다. 또한, 원하는 물성 값을 만족하는 구조 생성을 위해 조건 임베딩을 도입하고, 이를 입력 임베딩에 통합하는 방식(CDI)과 self-attention 계산에 더하는 방식(CDS) 두 가지 조건 반영 방법을 제안하여, 목표 물성에 맞는 결정구조의 생성을 가능하게 한다.

4. 결론

생성모델 기반 물성 생성 접근은, 특성 물성 조건을 만족하는 새로운 구조를 직접 설계할 수 있는 가능성을 제시한다. 하지만, 공간군 같은 정보를 제대로 반영하지 못해 대칭성이 거의 없는 구조를 생성하는 경향이 있으며, 생성된 구조가 겉보기에는 타당해 보일지라도 DFT등의 물리적 시뮬레이션을 통해 안정화를 할 경우 원자 위치나 격자 벡터가 크게 변화하여, 결과적으로 물리적으로 신뢰하기 어려운 구조를 생성되는 한계가 존재한다. 이를 해결하기 위해서, 공간군 정보를 명시적으로 조건에 포함하거나, 대칭성과 에너지 안정성을 보존하도록 설계될 수 있도록 다양한 학습 전략이 필요하다.

ACKNOWLEDGMENT

본 논문은 2025년도 한국연구재단 (NRF-2022R1A2C2004003)의 지원을 받아 수행된 연구임.

참 고 문 헌

- [1] Xie, T., Fu, X., Ganea, O. E., Barzilay, R., & Jaakkola, T. (2021). Crystal diffusion variational autoencoder for periodic material generation. *arXiv preprint arXiv:2110.06197*.
- [2] Ye, C. Y., Weng, H. M., & Wu, Q. S. (2024). Con-CDVAE: A method for the conditional generation of crystal structures. *Computational Materials Today*, 1, 100003.
- [3] Zeni, C., Pinsler, R., Zügner, D., Fowler, A., Horton, M., Fu, X., ... & Xie, T. (2023). Mattergen: a generative model for inorganic materials design. *arXiv preprint arXiv:2312.03687*.
- [4] Takahara, I., Shibata, K., & Mizoguchi, T. (2024). Generative Inverse Design of Crystal Structures via Diffusion Models with Transformers. *arXiv preprint arXiv:2406.09263*.