EGMOF: Efficient Generation of Metal-Organic Frameworks Using a Hybrid Diffusion-Transformer Architecture

Seunghee Han[1‡], Yeonghun Kang[1,2,3‡], Taeun Bae[1], Varinia Bernales[2,4], Alan Aspuru-Guzik[2,3,4,5,6,7,8,9,10] and Jihan Kim[1*]*

1 Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea.

2 Department of Chemistry, University of Toronto, 80 St. George St., Toronto, ON M5S 3H6, Canada

3 Vector Institute for Artificial Intelligence, W1140-108 College St., Schwartz Reisman Innovation Campus, Toronto, ON M5G 0C6, Canada

4 Acceleration Consortium, 700 University Ave., Toronto, ON M7A 2S4, Canada

5 Department of Computer Science, University of Toronto, 40 St George St., Toronto, ON M5S 2E4, Canada

6 Department of Materials Science & Engineering, University of Toronto, 184 College St., Toronto, ON M5S 3E4, Canada

7 Department of Chemical Engineering & Applied Chemistry, University of Toronto, 200 College St., Toronto, ON M5S 3E5, Canada

8 Institute of Medical Science, 1 King's College Circle, Medical Sciences Building, Room 2374, Toronto, ON M5S 1A8, Canada

9 Canadian Institute for Advanced Research (CIFAR), 661 University Ave., Toronto, ON M5G 1M1, Canada

10 NVIDIA, 431 King St. W #6th, Toronto, ON M5V 1K4, Canada

‡ These authors contributed equally

*Corresponding author: jihankim@kaist.ac.kr and alan@aspuru.com

ABSTRACT

Designing materials with targeted properties remains challenging due to the vastness of chemical space and the scarcity of property-labeled data. While recent advances in generative models offer a promising way for inverse design, most approaches require large datasets and must be retrained for every new target property. Here, we introduce the EGMOF (Efficient Generation of MOFs), a hybrid diffusion-transformer framework that overcomes these limitations through a modular, descriptor-mediated workflow. EGMOF decomposes inverse design into two steps: (1) a one-dimensional diffusion model (Prop2Desc) that maps desired properties to chemically meaningful descriptors followed by (2) a transformer model (Desc2MOF) that generates structures from these descriptors. This modular hybrid design enables minimal retraining and maintains high accuracy even under small-data conditions. On a hydrogen uptake dataset, EGMOF achieved over 95% validity and 84% hit rate, representing significant improvements of up to 57% in validity and 14% in hit rate compared to existing methods, while remaining effective with only 1,000 training samples. Moreover, our model successfully performed conditional generation across 29 diverse property datasets, including CoREMOF, QMOF, and text-mined experimental datasets, whereas previous models have not. This work presents a data-efficient, generalizable approach to the inverse design of diverse MOFs and highlights the potential of modular inverse design workflows for broader materials discovery.

TOC