

MATERIALS SCIENCE

Inverse design of porous materials using artificial neural networks

Baekjun Kim*, Sangwon Lee*, Jihan Kim†

Generating optimal nanomaterials using artificial neural networks can potentially lead to a notable revolution in future materials design. Although progress has been made in creating small and simple molecules, complex materials such as crystalline porous materials have yet to be generated using any of the neural networks. Here, we have implemented a generative adversarial network that uses a training set of 31,713 known zeolites to produce 121 crystalline porous materials. Our neural network takes in inputs in the form of energy and material dimensions, and we show that zeolites with a user-desired range of 4 kJ/mol methane heat of adsorption can be reliably produced using our neural network. The fine-tuning of user-desired capability can potentially accelerate materials development as it demonstrates a successful case of inverse design of porous materials.

INTRODUCTION

The quest to discover new materials using artificial intelligence has spawned a great deal of research in the past few years. Subsequently, significant progress has been made in using various artificial neural networks (ANNs) to generate undiscovered molecules and materials (1–3). Unfortunately, no one has yet successfully used ANNs to create novel crystalline materials as machine learning has only been applied to predict material properties such as compositions, bandgap energy, formation energy, and gas adsorption uptakes (4–10). Among crystalline materials, porous materials contain dense arrangements of microscopic pores that lead to high surface area and pore volume, and as such are viewed as an important class of materials for many different energy- and environmental related applications (11, 12). It is our opinion that these porous materials [e.g., zeolites, metal-organic frameworks (MOFs), and covalent organic frameworks (COFs)] are especially challenging to generate using ANNs due to their relatively complex topologies compared with other crystalline materials. Moreover, various other factors (e.g., nonunique representation of the unit cell, complex chemistry, ambiguous assignment of the lattice parameters, and constraints of periodic boundary conditions) all contribute to the challenges in successfully incorporating ANNs for crystalline materials generation.

Here, we devised an ANN that can successfully generate crystalline porous materials. Specifically, we have targeted a case study problem of generating pure silica zeolite structures due to their relative simplicity (e.g., only two atom types in silicon and oxygen) and wealth of materials available [i.e., 234 International Zeolites Association (IZA) experimental structures and 331,163 Predicted Crystallography Open Database (PCOD) hypothetical zeolites (13–15)] that can be used to train the neural network (see Materials and Methods). It must be noted that Deem and co-workers (13, 14) have already developed a Monte Carlo-based algorithm to generate pure silica zeolites. Moreover, many researchers have used both bottom-up (e.g., self-assembly based on building blocks) and top-down (e.g., topology-based construction of porous materials) approaches to generate porous materials such as MOFs and COFs (16–18). However,

most of these algorithms are deficient in the sense that finely tuned, user-desired properties cannot be targeted during the materials generation stage. As such, while there are few examples recently on using evolutionary algorithm to target material properties (19–22), most of these conventional methods lead to brute-force generation of porous materials, which necessitates a computationally expensive screening procedure to identify the optimal materials for a given application. Moreover, most of the computational screening work conducted on a large database of materials has revealed that the majority of these generated materials have poor properties, leading to inefficient allocation of computational resources (16, 23, 24). By designing the neural network to represent the inputs in both the material and energy dimensions, our algorithm has a unique advantage in that inverse design of materials can be achieved with the ANN by biasing the energy dimension, which correlates with material properties.

RESULTS AND DISCUSSION

Generative adversarial network for zeolites

Among many different choices in ANN, generative adversarial network (GAN) was used to produce crystalline porous materials due to its enhanced capability in generating realistic objects such as human faces (25, 26). GAN consists of both a discriminator and a generator, where the goal of the discriminator is to differentiate between the real and the fake data, whereas the generator acts to deceive the discriminator by progressively creating realistic fake objects. Overall, this setup leads to adversarial learning where realistic objects are generated as a by-product of the improvement in the learning process for both the discriminator and the generator. Recently, Arjovsky *et al.* (27) have suggested an improved version of GAN called Wasserstein GAN (WGAN). In WGAN, the discriminator is trained to estimate the Earth mover distance (EMD) between the data distribution and the generator distribution, and the generator is trained to minimize the EMD by generating realistic samples. In this case, the discriminator is renamed to be the “critic” as the critic does not discriminate anything. The critic has to be Lipschitz-1 continuous, and the constraint on the critic is attained by weight clipping in the original paper. Even more recently, Gulrajani *et al.* (28) modified WGAN with gradient penalty (WGAN-GP, which is the basis of our ANN), and consequently,

Copyright © 2020
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

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

*These authors contribute equally to this work.

†Corresponding author. Email: jihankim@kaist.ac.kr

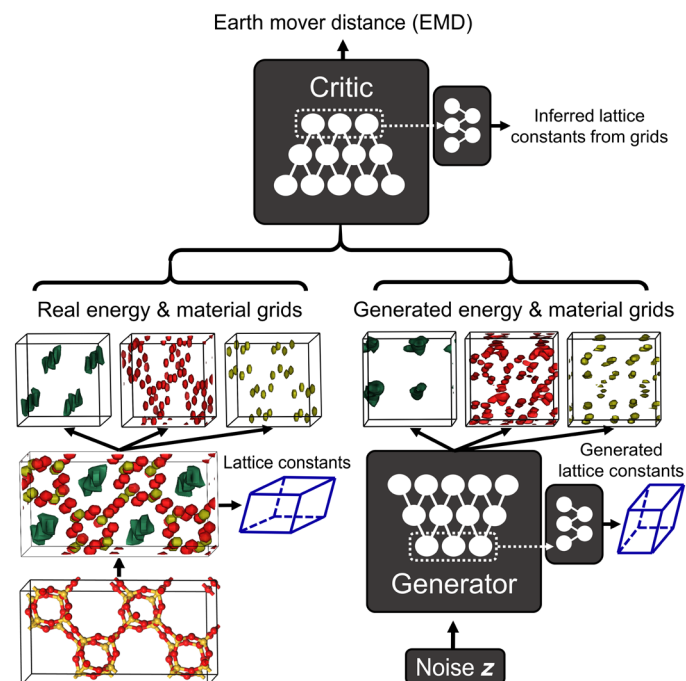


Fig. 1. Overall schematics of the ZeoGAN. Energy (green) in this case refers to methane potential energy, and material grids indicate silicon (red) and oxygen (yellow) atoms.

constraint was implemented by adding gradient penalty on the loss function for the critic.

With the specific goal in mind of generating materials and energy shapes, we developed a new type of GAN named zeolite GAN (ZeoGAN). In previous work, we developed a neural network that had the capability to generate just the energy shapes, but materials generation was impossible (29). Here, as a test case study problem, the energy dimension was designated to be the methane potential energy due to its importance in various methane storage applications (30, 31) and facile creation of methane energy grids (ergo, fast generation of the training set for machine learning purposes) using classical molecular simulations (32). Although restricted to methane in this particular study, our ANN can, in practice, be easily generalized to other gas molecules (e.g., hydrogen, water, and carbon dioxide) given that we only would need to change the selection of the gas molecules within the classical molecular simulations.

The overall schematic of the ZeoGAN is shown in Fig. 1 (details in sections S1-1 and S1-2). The input to the neural network is divided into the materials and the energy grids, with the materials grid further subdivided into the silicon and the oxygen atom grids. The size of each of the three grids is set to be $32 \times 32 \times 32$ points in fractional coordinates, with equal mesh size within a given zeolite unit cell. Fractional coordinates were used to keep the sizes of these grids the same for all of the zeolite materials. The number of grid points was kept small to reduce the memory cost as larger grids lead to a very slow learning process. The positions of the silicon and the oxygen atoms are represented by Gaussian functions, with the peak of the Gaussian corresponding to the position of the zeolite atoms. Since

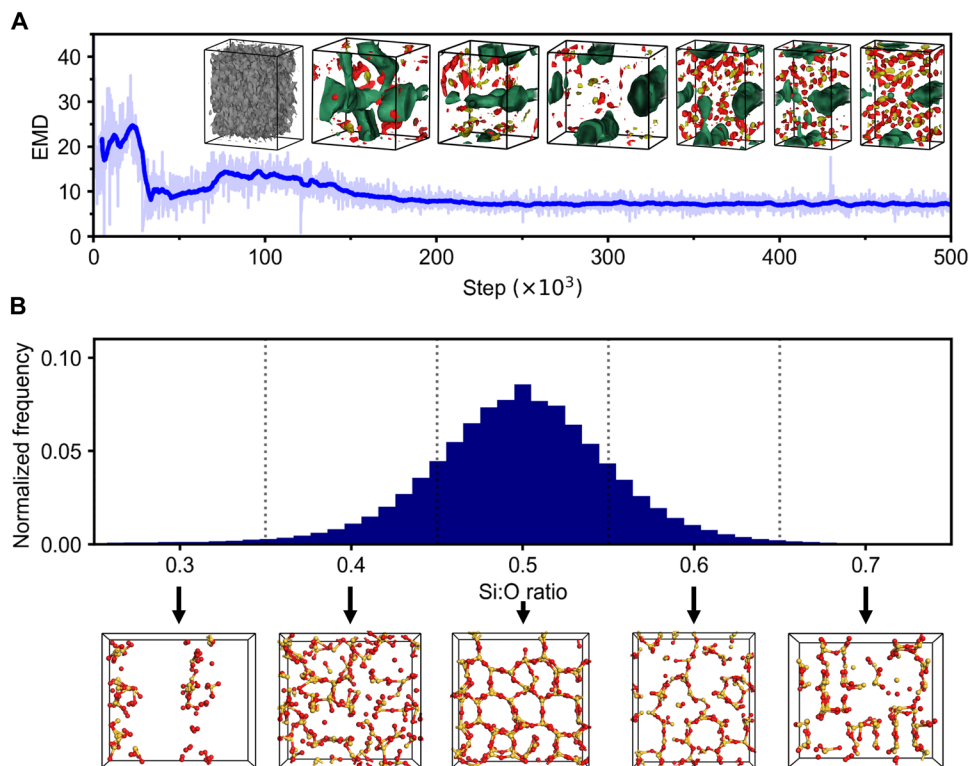


Fig. 2. Learning curve of ZeoGAN and histogram of Si:O ratio values. (A) EMD as a function of ZeoGAN iteration steps. The inset figure shows the evolution of a specific material (red/yellow) and energy (green) shapes. (B) Normalized frequency of Si:O ratio values for 1 million ZeoGAN outputs (top). Representative zeolite structures of the positions extracted from the zeolite shapes generated by the ZeoGAN for the outputs with different Si:O ratios (bottom).

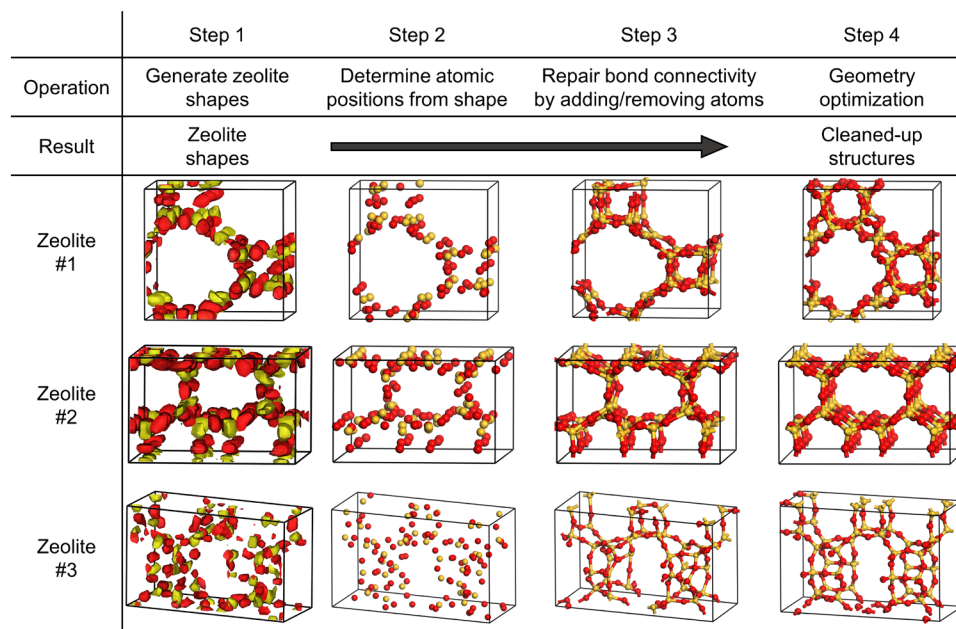


Fig. 3. Evolution of three zeolite shapes that successfully passed through the cleanup operation to yield Si:O = 0.5 and 100% bond connectivity.

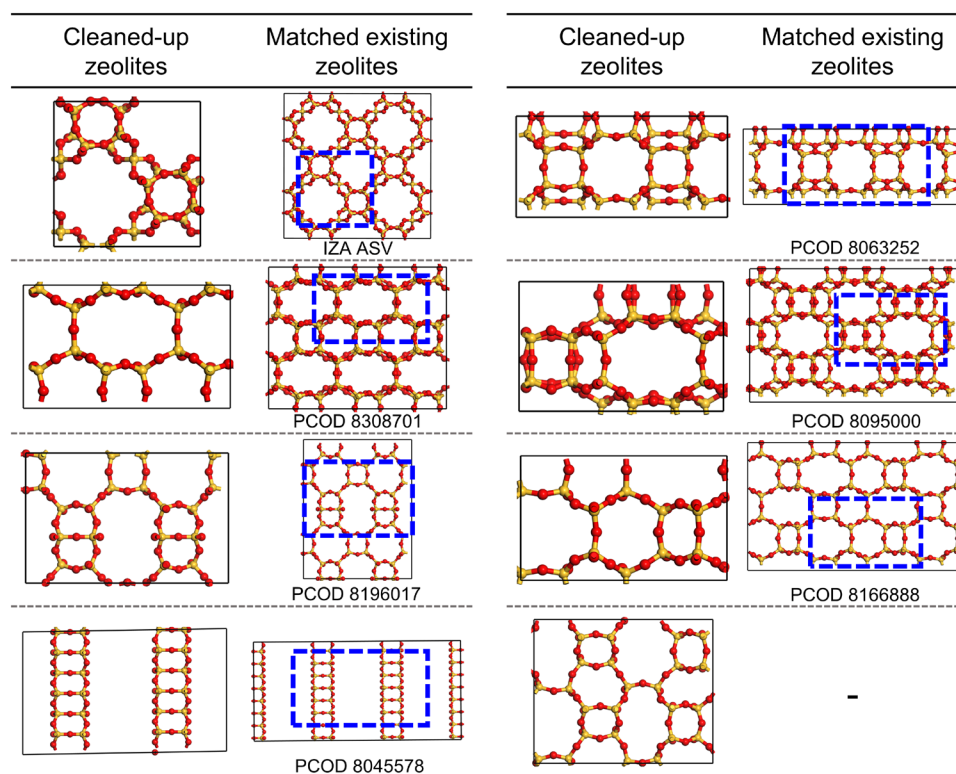


Fig. 4. Eight cleaned-up zeolites generated from ZeoGAN and their corresponding counterpart zeolites in IZA/PCOD. Dashed blue line indicate the matching unit cell portion. Few of the generated zeolites are not three-dimensional structures because the training set of the zeolite database already contained these structures.

the oxygen and silicon grids were separated akin to RGB color channels, both of the Gaussian functions were assigned to the same amplitude of 1.0 and variance of 0.5. Both the methane potential energy grids and the material grids were generated using conventional molecular simulations

(see Materials and Methods). The whole grids (silicon/oxygen/methane potential) are combined into a single tensor, and the tensor is used as an input of ZeoGAN. The ZeoGAN is trained to generate realistic tensors that resemble tensors calculated from the real zeolite.

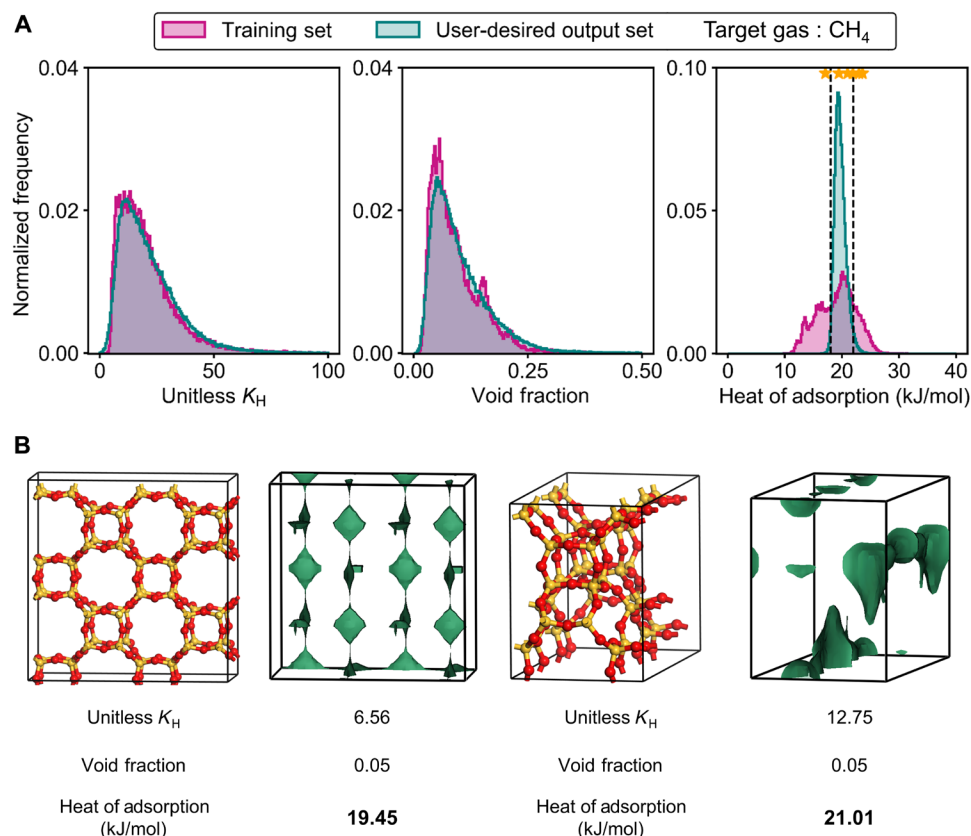


Fig. 5. User-desired generation results. (A) Distributions (methane K_H , methane void fraction, and methane heat of adsorption) for 31,713 training set zeolites (pink), 1 million user-desired zeolite shapes (green), and 6 user-desired zeolites (yellow markers). (B) Two representative structures generated from the user-desired scheme that yielded methane heat of adsorption in the user-desired range of 18 to 22 kJ/mol.

The goal of the generator in the ZeoGAN is to produce realistic zeolite materials and their corresponding energy shapes, and as such, several features were added to the ZeoGAN. Within the critic, periodic padding was implemented to impose a severe penalty for generating nonrealistic shapes (29). Without this feature, most of the materials produced by the ANN have unrealistic bonds formed across the periodic boundaries. To facilitate convergence, feature matching (33) was added to the ZeoGAN for both the materials and the energy shapes. Last, given that lattice constants were not explicitly included in both the materials or the energy grids, a lattice constant-generating network that produces accurate parameters based on the correlation between the lattice constants and materials/energy grids was added (29). The details on the training of ZeoGAN can be found in section S1-2.

Generation of pure silica zeolites

Overall, 31,173 methane accessible zeolites were used to train our neural network. The learning process of ZeoGAN shows the evolution of the material/energy shapes from their initial Gaussian noise distributions (Fig. 2A). Interpreting the distributions as similar to probability distributions, the EMD was selected as a metric to determine convergence in our training, with smaller EMD corresponding to more realistic zeolite shapes (section S1-2). Evolution of EMD from Fig. 2A shows converging behavior, and snapshots at different learning steps were illustrated to show the evolution in the shapes. Specifically, the material/energy shapes initially resemble a typical noise distribution (first picture in the inset of Fig. 2A), but as the

learning progresses, it can be seen that the materials and the energy shapes occupy separate regions in the unit cell space, morphing into shapes that look like a typical zeolite.

Together, 1 million zeolite shapes (both material and energy) were generated from the ZeoGAN (step 1); from these shapes, a simple rule (section S1-4) was used to assign the positions of the oxygen and the silicon atoms (step 2). Next, the Si:O ratio was calculated for each of the outputs of step 2, and the results were tabulated as shown in Fig. 2B (noting that all four connected pure silica zeolites should, in theory, have Si:O = 0.5). To assess the quality of the outputs, representative ZeoGAN outputs are illustrated in Fig. 2B for Si:O = 0.3, 0.4, 0.5, 0.6, and 0.7. While several issues (e.g., wrong Si:O ratio, inaccurate bond lengths and bond angles) arise that prevent most of these outputs from being real zeolites, the majority of them share similar characteristics in terms of distributions of the silicon and the oxygen atoms compared with the real zeolites. Continuing from step 2, a subset of outputs from step 2 with $0.45 < \text{Si:O} < 0.55$ were filtered (65% of the total zeolite shapes), and from this set, the ones with 75% or higher proper bond connectivity (where each Si has four oxygen bonds, and each oxygen has two silicon bonds) were kept, reducing the candidate set to 901. For bond connectivity, Si and O atoms were assumed to form bonds if their distance was less than 2.5 Å. From this reduced set, bond connectivity was repaired by adding/removing atoms via a random-based search (step 3), yielding 674 structures with Si:O ratio = 0.5 and 100% connectivity. From this set, only the structures with a small

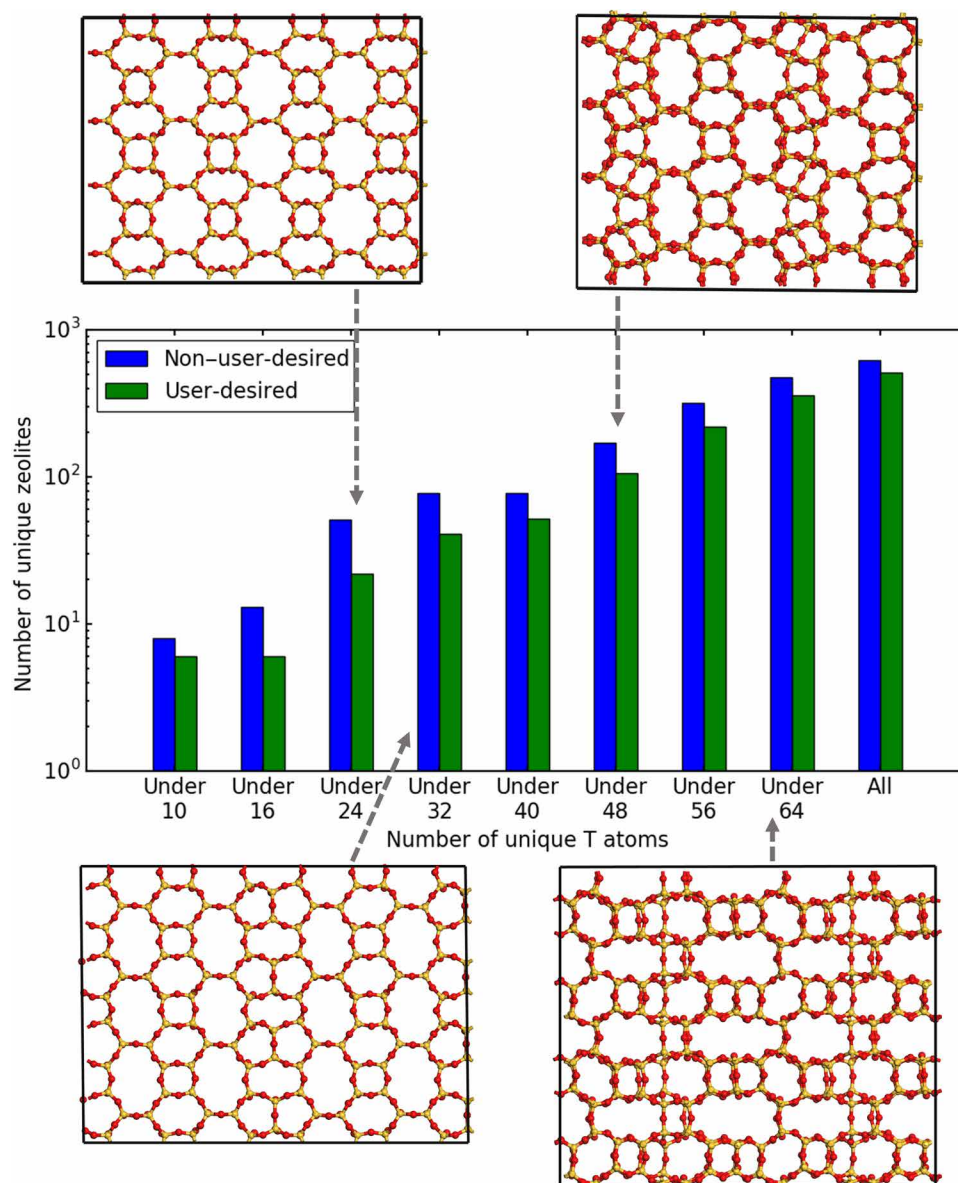


Fig. 6. Number of zeolites versus the number of unique T atoms. Some representative zeolites are shown for different numbers of T atoms: 12 (left top), 28 (left bottom), 48 (right top), and 64 (right bottom).

number of symmetrically unique T atoms (≤ 10) were kept as most of the structures with a high number of unique T atoms tend to be nonsymmetric, and as a result, reducing the total number to eight. Last, these fully connected structures were optimized using classical molecular simulations (section S1-4) with the same parameters used by Deem and co-workers to facilitate comparisons (step 4).

The overall operation explained above is summarized in Fig. 3 for three representative structures that passed through all the criteria. As can be seen from Fig. 3, all of the final relaxed structures resemble their respective initial zeolite shapes, providing evidence that our postprocessing operation does not significantly alter the essence of the zeolite shapes.

For the eight resulting structures, the coordination sequence (34) was used as a simple metric to identify whether these structures exist in either the IZA or the PCOD database. As such, the coordination

sequences for all of these zeolites were obtained, and it was found that seven were in either the IZA or the PCOD test set while one was an unknown zeolite, not found in any of the sets. None of the eight zeolites were in the original training set, indicating that ZeoGAN successfully created new zeolites that were not seen during the learning process. The cleaned-up zeolites and the corresponding zeolites from the IZA/PCOD database are shown in Fig. 4. One of the zeolites identified is ASV, which consists of an experimentally verified topology with 12, 6, and 4 membered rings and one-dimensional channels.

Inverse design of zeolites using ZeoGAN

Thus far, the zeolites generated from the ZeoGAN did not have any user-desired properties. To improve upon our design, the methane heat of adsorption was selected to test the user-desired capability

since it is an important metric that is often targeted for materials design (35) by the experimentalists. Subsequently, the ZeoGAN loss function was altered such that zeolites with the heat of adsorption value between 18 and 22 kJ/mol were targeted for generation (section S1-3). As can be seen from the data for 1 million newly generated zeolite shapes in Fig. 5A, there is a sharp change in the methane heat of adsorption distribution compared with the training set case, indicating the proper functioning of the user-desired criterion. It is also worth mentioning that the distributions for methane K_H and void fraction did not change much for the new loss function, indicating their uncorrelated nature with the selected methane heat of adsorption range.

Next, the same cleanup procedure described in Fig. 3 was conducted for the 1 million user-desired zeolite shapes, and the operation yielded six new zeolites (two from the training set, two from the PCOD/IZA set, and two outside of the dataset) and one zeolite (i.e., PCOD 8308701) that was also produced in the non-user-desired set. Out of these six zeolites, four were found to have methane heat of adsorption between 18 and 22 kJ/mol (with the two other having 17.1 and 23 kJ/mol, respectively), providing a reliable indication of successful inverse design of the zeolites. To the best of our knowledge, none of the existing experimental or computational methods in porous materials can a priori target a property with this specific range. Two representative cleaned-up zeolites are shown in Fig. 5B, and comparison of the coordination sequences reveals that they correspond to zeolite ACO and PCOD 8242361 structures (see section S2 for more detail).

Thus far, the zeolite candidates generated from ANN were restricted to having 10 unique T atoms, as most of the experimentally synthesized zeolites and all of the hypothetical zeolites have this restriction. However, it is conceivable that zeolites have more than this number as evidenced by zeolite UOZ (in IZA database), which has 25 unique T atoms. As such, upon removing the restriction on the number of unique T atoms, we observe a significant increase in the number of zeolites that were generated via the ANN (see Fig. 6). Specifically, adopting the traditional value where zeolites are deemed to be thermodynamically stable (30 kJ/mol Si) (14), the number of feasible zeolite structures was counted to be 121. One can extend the 30 kJ/mol Si limit by adopting the maximum framework energy of zeolite RWY (105.6 kJ/mol Si), and the total number of ANN-generated zeolites jumps to 1138 (although 105.6 kJ/mol Si might be too large of a limit for practical synthesis). Representative zeolites in Fig. 6 indicate that these materials look to be reasonable. Without the restriction on the unique T atoms, the total number of zeolites generated jumps from 14 to 1138, vastly enhancing the set and increasing the efficiency of generated zeolites from 0.0007 to 0.06% from the 2 million generated material/energy shapes. It is interesting to note that 1127 of these structures are not found in any existing database given that these are more complex zeolites, and as such, our ANN has successfully extended the number of new zeolites in the pure silica zeolite material space.

CONCLUSIONS

Here, we successfully demonstrated an inverse design of zeolites using our in-house-developed ANN. This work can potentially pave the way to use ANNs to target user-desired property before material design/synthesis, and the tools can extend to any application in which the energy grid can map onto the properties relevant to the

application itself (e.g., gas storage/separation, catalysis, sensors). Moreover, while being restricted to just the silicon and the oxygen atoms for simplicity purposes, the number of input channels in our ANN can be incremented to cover more complex crystalline materials such as MOFs and COFs, thereby enhancing the scope of the work and affecting the future design of diverse classes of materials.

MATERIALS AND METHODS

Existing zeolite structures

To design zeolites that have sufficient methane accessibility, 217 IZA (four connected and nondisordered) and 331,163 PCOD zeolites were trimmed down to 99,362 structures that have a methane Henry coefficient (K_H) value larger than 10^{-6} mol kg⁻¹ Pa⁻¹. In this methane-accessible zeolite set, 63,426 structures that have orthogonal cell angles (102 IZA and 63,324 PCOD) were used in this study. For the validation of nonexistent zeolite structures generation, the half of orthogonal zeolites was randomly selected as the input of ANNs.

Molecular simulations

The energy grids were constructed using classical molecular simulations. To keep the grid sizes consistent, the three-dimensional lattice vectors were divided by 32. For each of the $32 \times 32 \times 32$ grid points, a methane molecule was placed to calculate the interaction energies between the methane and framework atoms. The interaction energy was computed using the following Lennard-Jones (LJ) 12-6 potential model

$$U_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (1)$$

where r is the distance between methane and the framework atom, and the LJ force field parameters ϵ and σ were chosen from García-Pérez *et al.* (36). The Lorentz-Berthelot mixing rule was used to compute the interaction of different atom types.

By using the energy grids, adsorption properties like the Henry coefficient (K_H) and the isosteric heat of adsorption (Q_{st}) can be directly calculated at the learning stage of the ANNs. The unitless K_H and the Q_{st} of a methane molecule in the property grid can be calculated from following equations

$$\text{unitless } K_H = \frac{1}{N} \sum_i^N e^{-\beta E_i} \quad (2)$$

$$Q_{st} = k_B T - \frac{\sum_i^N E_i e^{-\beta E_i}}{\sum_i^N e^{-\beta E_i}} \quad (3)$$

where N is the number of grid points, k_B is the Boltzmann constant, β is $1/k_B T$, and E_i is the energy value at the i th grid point. Here, all of these adsorption properties were calculated at 298 K.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/1/eaax9324/DC1>

Supplementary Materials and Methods

Section S1. Computational methods

Section S1-1. Details of ZeoGAN

Section S1-2. Training of ZeoGAN

Section S1-3. Training for user-desired properties

Section S1-4. Zeolite cleanup

Section S2. Zeolite generation results

Section S2-1. Generated non-user-desired zeolite structures

Section S2-2. Generated user-desired zeolite structures

Fig. S1. Architecture of ZeoGAN.

Fig. S2. Training results after adding user-desired loss.

Fig. S3. Allowed next structure moves for the connectivity repairing algorithm.

Fig. S4. Distributions (methane K_{st} , methane void fraction, and methane heat of adsorption) for non-user-desired generation.

Fig. S5. Summary of non-user-desired zeolite shapes (both materials and energy shapes) and corresponding cleaned-up structures (both materials and energy shapes).

Fig. S6. Summary of user-desired zeolite shapes (both materials and energy shapes) and corresponding cleaned-up structures (both materials and energy shapes).

Fig. S7. Matching zeolites in IZA/PCOD for the user-desired zeolites.

Table S1. ZeoGAN hyperparameters.

Table S2. Various methane properties of the eight non-user-desired zeolites.

Table S3. Various methane properties of the six user-desired zeolites.

References (37–40)

REFERENCES AND NOTES

- N. De Cao, T. Kipf, MolGAN: An implicit generative model for small molecular graphs (2018); available at <http://arxiv.org/abs/1805.11973>.
- B. Sanchez-Lengeling, A. Aspuru-Guzik, Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **361**, 360–365 (2018).
- R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, A. Aspuru-Guzik, Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent. Sci.* **4**, 268–276 (2018).
- M. Fernandez, P. G. Boyd, T. D. Daff, M. Z. Aghaji, T. K. Woo, Rapid and accurate machine learning recognition of high performing metal organic frameworks for CO₂ capture. *J. Phys. Chem. Lett.* **5**, 3056–3060 (2014).
- Y. Zhuo, A. Mansouri Tehrani, J. Brgoch, Predicting the band gaps of inorganic solids by machine learning. *J. Phys. Chem. Lett.* **9**, 1668–1673 (2018).
- T. Xie, J. C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
- Z. Zhang, J. A. Schott, M. Liu, H. Chen, X. Lu, B. G. Sumpter, J. Fu, S. Dai, Prediction of carbon dioxide adsorption via deep learning. *Angew. Chem. Int. Ed.* **58**, 259–263 (2019).
- Y. Liu, T. Zhao, W. Ju, S. Shi, Materials discovery and design using machine learning. *J. Materiomics* **3**, 159–177 (2017).
- S. Shi, J. Gao, Y. Liu, Y. Zhao, Q. Wu, W. Ju, C. Ouyang, R. Xiao, Multi-scale computation methods: Their applications in lithium-ion battery research and development. *Chin. Phys. B* **25**, 018212 (2016).
- Y. Liu, T. Zhao, G. Yang, W. Ju, S. Shi, The onset temperature (T_g) of As_xSe_{1-x} glasses transition prediction: A comparison of topological and regression analysis methods. *Comput. Mater. Sci.* **140**, 315–321 (2017).
- K. K. Gangu, S. Maddila, S. B. Mukkamala, S. B. Jonnalagadda, A review on contemporary Metal-Organic Framework materials. *Inorganica Chim. Acta* **446**, 61–74 (2016).
- H. Furukawa, K. E. Cordova, M. O’Keeffe, O. M. Yaghi, The Chemistry and Applications of Metal-Organic Frameworks. *Science* **341**, 1230444 (2013).
- D. J. Earl, M. W. Deem, Toward a database of hypothetical zeolite structures. *Ind. Eng. Chem. Res.* **45**, 5449–5454 (2006).
- R. Pophale, P. A. Cheeseman, M. W. Deem, A database of new zeolite-like materials. *Phys. Chem. Chem. Phys.* **13**, 12407–12412 (2011).
- International Zeolites Association (IZA) Home Page; available at www.iza-structure.org/databases.
- C. E. Wilmer, M. Leaf, C. Y. Lee, O. K. Farha, B. G. Hauser, J. T. Hupp, R. Q. Snurr, Large-scale screening of hypothetical metal-organic frameworks. *Nat. Chem.* **4**, 83–89 (2012).
- J. Keupp, R. Schmid, TopoFF: MOF structure prediction using specifically optimized blueprints. *Faraday Discuss.* **211**, 79–101 (2018).
- R. L. Martin, C. M. Simon, B. Medasani, D. K. Britt, B. Smit, M. Haranczyk, In silico design of three-dimensional porous covalent organic frameworks via known synthesis routes and commercially available species. *J. Phys. Chem. C* **118**, 23790–23802 (2014).
- Y. Bao, R. L. Martin, C. M. Simon, M. Haranczyk, B. Smit, M. W. Deem, In silico discovery of high deliverable capacity metal-organic frameworks. *J. Phys. Chem. C* **119**, 186–195 (2014).
- Y. G. Chung, D. A. Gómez-Gualdrón, P. Li, K. T. Leperi, P. Deria, H. Zhang, N. A. Vermeulen, J. F. Stoddart, F. You, J. T. Hupp, O. K. Farha, R. Q. Snurr, In silico discovery of metal-organic frameworks for precombustion CO₂ capture using a genetic algorithm. *Sci. Adv.* **2**, e1600909 (2016).
- Y. Bao, R. L. Martin, M. Haranczyk, M. W. Deem, In silico prediction of MOFs with high deliverable capacity or internal surface area. *Phys. Chem. Chem. Phys.* **17**, 11962–11973 (2015).
- E. Berardo, L. Turcani, M. Miklitz, K. E. Jelfs, An evolutionary algorithm for the discovery of porous organic cages. *Chem. Sci.* **9**, 8513–8527 (2018).
- L.-C. Lin, A. H. Berger, R. L. Martin, J. Kim, J. A. Swisher, K. Jariwala, C. H. Rycroft, A. S. Bhowm, M. W. Deem, M. Haranczyk, B. Smit, In silico screening of carbon-capture materials. *Nat. Mater.* **11**, 633–641 (2012).
- S. Chong, G. Thiele, J. Kim, Excavating hidden adsorption sites in metal-organic frameworks using rational defect engineering. *Nat. Commun.* **8**, 1539 (2017).
- I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative Adversarial Networks (2014); available at <http://arxiv.org/abs/1406.2661>.
- A. Radford, L. Metz, S. Chintala, Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks (2015); available at <http://arxiv.org/abs/1511.06434>.
- M. Arjovsky, S. Chintala, L. Bottou, Wasserstein GAN (2017); available at <http://arxiv.org/abs/1701.07875>.
- I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville, Improved Training of Wasserstein GANs (2017); available at <https://arxiv.org/abs/1704.00028>.
- S. Lee, B. Kim, J. Kim, Predicting performance limits of methane gas storage in zeolites with an artificial neural network. *J. Mater. Chem. A* **7**, 2709–2716 (2019).
- Y. He, W. Zhou, G. Qian, B. Chen, Methane storage in metal-organic frameworks. *Chem. Soc. Rev.* **43**, 5657–5678 (2014).
- T. A. Makal, J.-R. Li, W. Lu, H.-C. Zhou, Methane storage in advanced porous materials. *Chem. Soc. Rev.* **41**, 7761–7779 (2012).
- J. Kim, R. L. Martin, O. Rübel, M. Haranczyk, B. Smit, High-throughput characterization of porous materials using graphics processing units. *J. Chem. Theory Comput.* **8**, 1684–1693 (2012).
- T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, X. Chen, Improved Techniques for Training GANs (2016); available at <http://arxiv.org/abs/1606.03498>.
- W. M. Meier, H. J. Moeck, The topology of three-dimensional 4-connected nets: Classification of zeolite framework types using coordination sequences. *J. Solid State Chem.* **27**, 349–355 (1979).
- W. Jeong, D. W. Lim, S. Kim, A. Harale, M. Yoon, M. P. Suh, J. Kim, Modeling adsorption properties of structurally deformed metal-organic frameworks using structure-property map. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7923–7928 (2017).
- E. García-Pérez, J. B. Parra, C. O. Ania, A. García-Sánchez, J. M. van Baten, R. Krishna, D. Dubbeldam, S. Calero, A computational study of CO₂, N₂, and CH₄ adsorption in zeolites. *Adsorption* **13**, 469–476 (2007).
- D. P. Kingma, J. Ba, Adam: A Method for Stochastic. *Optimization* **2014**, 1–15 (2014).
- A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, W. M. Skiff, UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **114**, 10024–10035 (1992).
- J. D. Gale, A. L. Rohl, The General Utility Lattice Program (GULP). *Mol. Simul.* **29**, 291–341 (2003).
- K. P. Schröder, J. Sauer, M. Leslie, C. Richard, A. Catlow, J. M. Thomas, Bridging hydrolyl groups in zeolitic catalysts: A computer simulation of their structure, vibrational properties and acidity in protonated faujasites (H-Y zeolites). *Chem. Phys. Lett.* **188**, 320–325 (1992).

Acknowledgments

Funding: This work was supported, in part, by the Mid-Career Researcher Program (NRF-2017R1A2B4004029), and, in part, by the Energy Cloud R&D Program (NRF-2019M3F2A1072233) through the NRF (National Research Foundation of Korea), both funded by the Ministry of Science and ICT. This work was also supported by the BK21 Plus Program funded by the Ministry of Education (MOE, Korea). **Author contributions:** J.K. formulated the project. B.K. conducted the molecular and ANN simulations. S.L. designed the neural network. All authors contributed to the writing/revising of the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. ZeoGAN code can be downloaded from <https://github.com/good4488/ZeoGAN>. Additional data related to this paper may be requested from the authors.

Submitted 6 May 2019

Accepted 7 November 2019

Published 3 January 2020

10.1126/sciadv.aax9324

Citation: B. Kim, S. Lee, J. Kim, Inverse design of porous materials using artificial neural networks. *Sci. Adv.* **6**, eaax9324 (2020).

Inverse design of porous materials using artificial neural networks

Baekjun KimSangwon LeeJihan Kim

Sci. Adv., 6 (1), eaax9324. • DOI: 10.1126/sciadv.aax9324

View the article online

<https://www.science.org/doi/10.1126/sciadv.aax9324>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).