Article

3-D Inorganic Crystal Structure Generation and Property Prediction via Representation Learning

Callum J. Court,[#] Batuhan Yildirim,[#] Apoorv Jain, and Jacqueline M. Cole*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Generative models have been successfully used to synthesize completely novel images, text, music, and speech. As such, they present an exciting opportunity for the design of new materials for functional applications. So far, generative deep-learning methods applied to molecular and drug discovery have yet to produce stable and novel 3-D crystal structures across multiple material classes. To that end, we, herein, present an autoencoder-based generative deep-representation learning pipeline for geometrically optimized 3-D crystal structures that simultaneously predicts the values of eight target properties. The system is highly general, as demonstrated through creation of novel materials from three separate material classes: binary alloys, ternary perovskites, and Heusler compounds. Comparison of these generated structures to those optimized via electronic-structure calculations shows that our generated materials are valid and geometrically optimized.
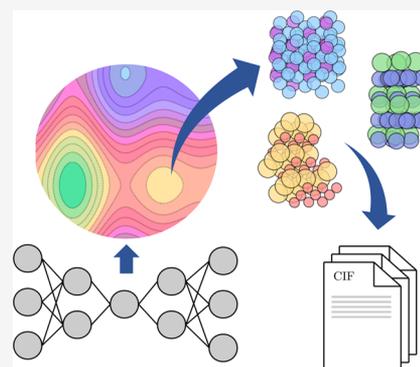
INTRODUCTION

Experimental research has long been the backbone of materials science and discovery, but the cost, from both financial and time perspectives, creates a bottleneck in the "design-to-device" workflow.[1,2] Materials research ultimately aims to employ new materials for functional applications. A key part of that process is the characterization of materials structure, upon which properties and, therefore, applications are heavily dependent. The ability to predict material structure from basic first-principles information, such as the chemical composition, is a long outstanding goal that has yet to be achieved.[3] The vast space of possible chemical compositions and aforementioned cost in characterizing structures experimentally makes it impossible to fully explore the composition-structure space. To narrow down the search space of candidate materials, researchers often employ ab initio methods, such as Density Functional Theory (DFT).[4] This allows for computational simulation of materials and their properties, thus ensuring only the most promising candidate materials need to be synthesized experimentally. These ab initio approaches have had great success in advancing modern computational materials discovery.

Efforts to harmonize computational materials science, by compiling structures and properties into databases, have led to a number of large public data repositories. The Materials Project,[5] Open Quantum Materials Database,[6,7] Novel Materials Database (NoMaD),[8] Inorganic Crystal Structure Database (ICSD),[9] Cambridge Crystal Structure Data (CCSD),[10] and the Crystallography Open Database (COD)[11−14] all provide vast data sets of structures and their properties. Exploration of these data could uncover structure−property relationships and enable rapid progress across multiple domains. However, with each containing hundreds of thousands of data points, the abundance of data on electronic structures now makes it impossible to perform manual analysis of the entire space.

Fortunately, data-science methods are well suited to analyze such large aggregations of data. Within this scope, the exploration of high-dimensional data sets is a task highly suited to machine learning. The rapid rise of machine learning and deep learning in recent years, along with advances in computational capability, has led to a number of projects focused on applications in materials science.[15,16] In particular, various techniques have been employed toward data-driven materials discovery, such as high-throughput computation,[17,18] natural language processing,[19−21] "design-to-device" pipelines[22,23] and deep learning.[24,25]

Outside of the scientific domain, deep-generative models have successfully created novel instances of videos,[26] images,[27] text,[28] and audio.[29] These have created a swell of excitement around the potential use cases in materials science and biomedicine. Deep-learning models trained on existing materials could produce new chemicals, molecules, and drugs at a fraction of the cost of experimental and ab initio research. Furthermore, the ability of Artificial Neural Networks (ANNs)

to find patterns in highly complex feature spaces allows for exploration of structure−property relationships far outside the capability of human analysis.

Recently, several deep-learning methods for generating plausible molecules and drugs have emerged. A typical generative deep-learning pipeline involves learning a representative data distribution and, subsequently, sampling from it to generate novel examples. One popular approach is to use a variational autoencoder (VAE),[30] in which a log-likelihood is optimized to approximate a posterior distribution of the data. Other methods utilize Generative Adversarial Networks (GANs). These circumvent the need to evaluate a log-likelihood and instead optimize a game-theoretic mini-max objective by adversarially training two competing neural networks.[31]

Molecular Generative Models. All molecular generative models follow the same approximate pattern. Molecular structures are input to a deep-learning pipeline in any number of plausible formats. The model is then trained to learn a latent representation of molecules that can be sampled from to produce new crystals. The variation in methods usually surrounds the choice of representation and how to encode meaningful properties into the latent space.

Using a VAE, Gómez-Bombarelli et al.[32] learned a mapping between text-string representations of molecules, in simplified molecular-input line-entry system (SMILES) format, and a continuous latent space. By sampling from regions that were far enough from the training data, they were able to produce a large fraction of novel molecules with a high drug-likeness. More recent methods improve on the shortcomings of this work, particularly the use of SMILES representations as inputs. For example, researchers have employed graph representations that enforce chemical rules to generate molecules with specifically tailored properties.[33−35] Similarly, Lim et al.[36] traded the vanilla VAE used by Gómez-Bombarelli et al. for a conditional-VAE, which conditions the encoded representation on specific properties. As a result, their latent representation consists of a combination of regions for molecular properties and structures. By embedding the target properties in the latent space together with molecular structures, they were able to sample molecules with desired properties from specific regions of the latent space. Kearnes et al.[37] employed a graph encoder but replaced the vanilla decoder with a reinforcement learning-based graph decoder. Specifically, they utilized a Deep Q-Network,[38] which guarantees that molecules afforded by their model are chemically valid.

Other studies have employed GAN architectures to avoid the need to approximate intractable likelihoods during optimization. A typical GAN architecture contains a generator network, that produces novel samples, and a discriminator network that attempts to distinguish between "real" samples and "fake" samples produced by the generator. These are trained jointly using an adversarial loss function to find an equilibrium between the two networks. Such a method has been used to generated two-dimensional graphene hybrids.[39] Another notable example is MolGAN,[40] in which De Cao and Kipf use a generator network to produce graph representations of molecules. Accordingly, a graph-convolutional neural network attempts to discriminate between samples from the training data and samples produced by the generator. They also optimize the validity and novelty of the generated molecules by jointly training a reinforcement-learning-based reward network that assigns zero reward to molecules which

are invalid. While this mostly guarantees valid molecules, an undesirable result is a constrained latent space with little sample variability. Hybrid models, which combine the latent space of autoencoders with the training procedure of GANs, have also recently emerged. Prykhodko et al.[41] trained an autoencoder on SMILES inputs and used a GAN to approximate a VAE latent representation.

Crystal-Structure Generative Models. While the literature related to the creation of novel molecules using generative models has proliferated, similar works that strive to do the same for inorganic crystal structures are less common, but are on the rise. A notable variational method to generate crystal structures is iMatGen,[25] which uses 3-D image inputs to learn a latent space of inorganic structures. This latent space is further enhanced by training a binary-classification formation-energy model on latent vectors of the input crystals, to distinguish stable structures from unstable ones. By performing a case study on V_xO_y compounds, iMatGen has demonstrated its capabilities by rediscovering existing structures when these compounds are not included in the training data, as well as generating novel compounds which are found to be stable, as validated by DFT. Hoffmann et al.[42] extended this work to include a UNet segmentation architecture to generalize the method to multiple material classes but they were unable to produce stable crystal structures.

The first approach to do so via a GAN was CrystalGAN,[43] which employed a CycleGAN[44] model to generate novel ternary structures from known binaries. Although the authors successfully demonstrated that CrystalGAN can generate novel ternary compounds, it is unclear whether their method can be generalized to more complex crystal structures. Kim and Noh et al.[45] similarly employed a GAN using point clouds as inputs, to create a model which can generate structures conditioned on crystal composition. Their model addresses the limitation of iMatGen, which is limited by the requirement of a high-dimensional 3-D image for each elemental type in the class of structures being generated. They demonstrated that their model could generate novel and stable compounds of a single kind (Mg−Mn−O ternary compounds).

Nonetheless, the primary hurdle for production of stable and chemically valid 3-D inorganic crystal structures is the choice of computational representation. Ideally, we require a continuous encoding of crystals that maintains periodicity, conserves rotational and translation symmetry and is agnostic to the unit-cell lengths or the number of atoms therein. It is also desirable that such representations are reversible and therefore easily converted to formats more widely used by the scientific community. Although encodings and descriptors of crystal structures do exist,[25,42,46,47] there is, as yet, no single encoding that meets all of the criteria above while providing a convenient input for ANNs.

Material-Property Prediction. In addition to crystal-structure prediction, material-property prediction is a vital part of computational materials discovery. Without the ability to predict the properties of new structures, existing tools are of limited use since they still require existing property prediction methods, such as DFT, to validate their materials. Modeling structure−property relationships has recently been influenced by the success of graph-neural networks (GNNs), in which molecules or crystals are represented by undirected graphs. Graphs, which consist of nodes and edges, are particularly suited to representing molecules and crystal structures, since
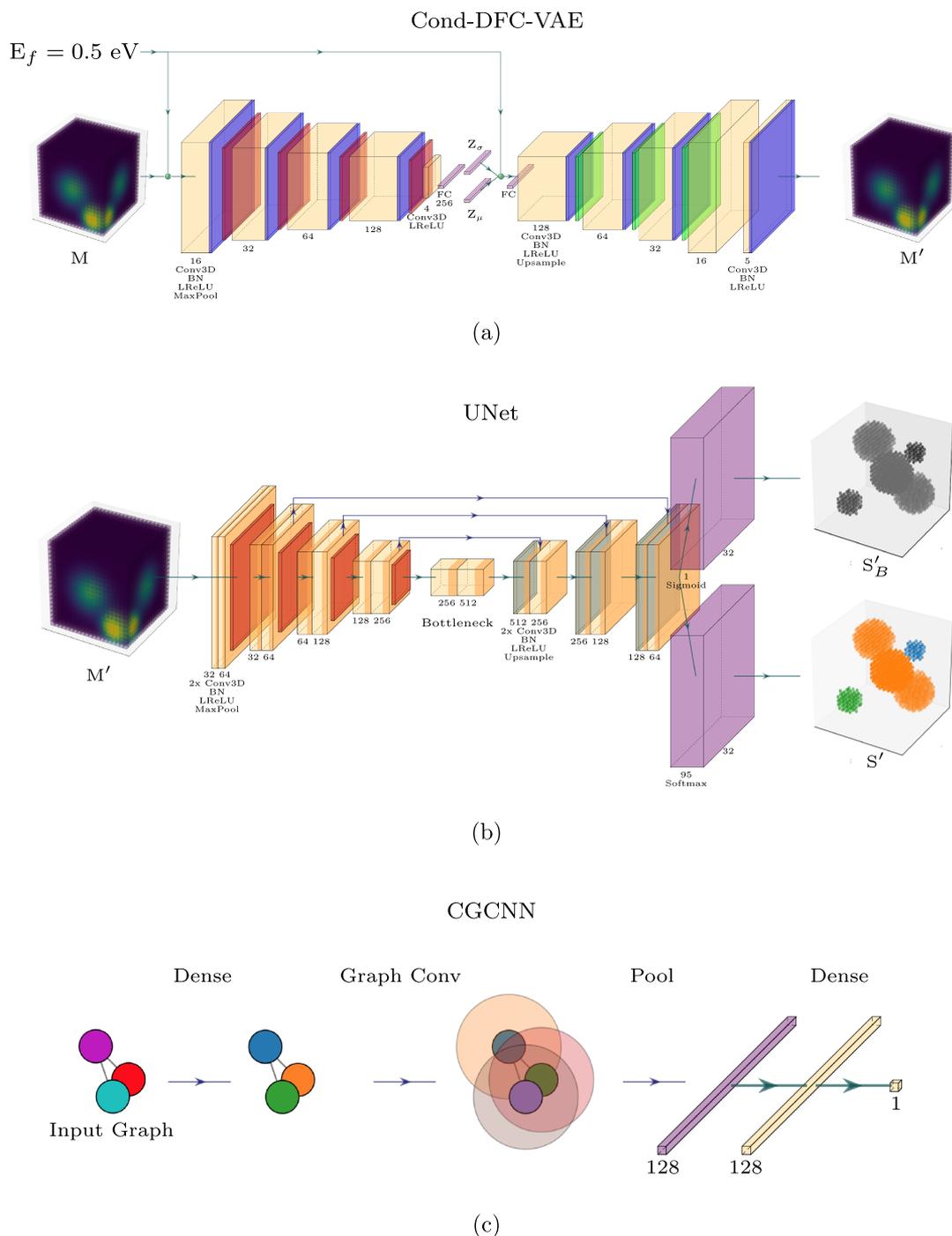
Figure 1. Architecture of our Model. (a) Conditional deep-feature-consistent variational autoencoder (Cond-DFC-VAE). The network takes in electron-density maps, M, with a corresponding property and produces reconstructed maps, M'. The encoder (decoder) architectures consist of repeating Conv3D, BatchNorm, LeakyReLU, and MaxPool (Upsample) operations. (b) UNet converts the electron-density maps to segmented species matrices. The architecture uses Conv3D, ReLU, and Batchnorm blocks with maxpooling or upsampling. (c) CGCNN starts with a single dense layer and is succeeded by a graph convolution to project the input crystal graphs such that the underlying structure of the graph is preserved. Following this, the input graphs are pooled, and projected by two dense layers to output target properties.

relationships between nodes (atoms) can be explicitly encoded by the presence of edges (bonds). Similarly, the absence of an interaction between two nodes is made explicit by a lack of an edge between the two nodes. This representation provides these powerful inductive biases which have recently been exploited to prove the superiority of graph-neural network

models for many tasks including structure–property prediction.[48]

GNNs operate under the assumption that nodes which are connected via an edge, should propagate their information to each other. Many examples of GNN methods have recently emerged which successfully implement this process as learnable transformations.[49−58] Accordingly, variants of
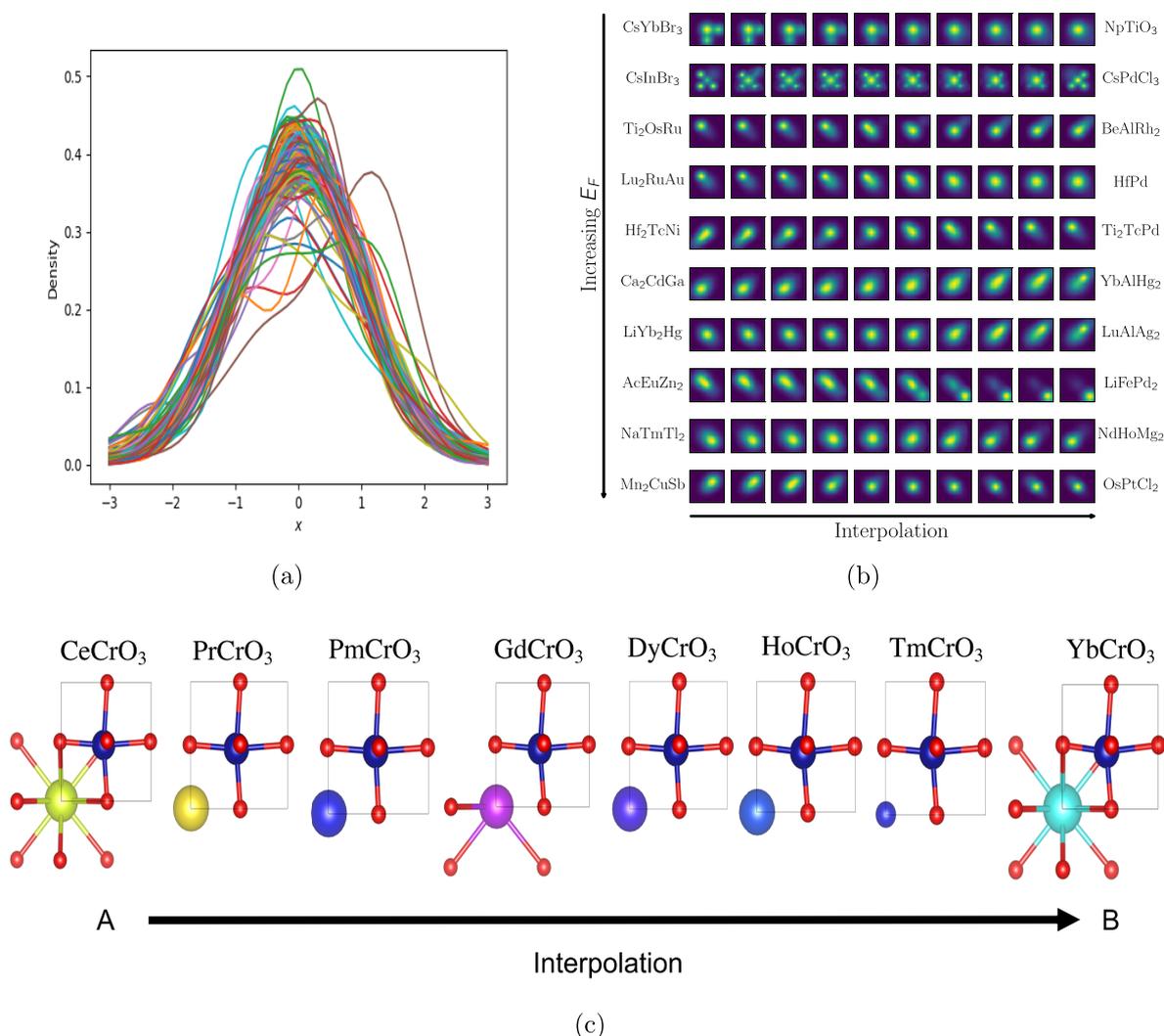
Figure 2. Latent encoding of 3-D crystal structures. (a) Kernel density estimate (KDE) plot of the VAE latent dimensions after training. The majority of dimensions show approximately normal distributions. (b) Latent-space interpolations between electron-density maps. Each row corresponds to a different formation-energy condition imposed on the latent space. (c) Crystals resulting from interpolation between rare-earth chromite perovskites. Linear interpolation between $CeCrO_3$ and $YbCrO_3$ shows traversal of A-site atom along the lanthanides with consistent crystal structures.

GNNs have successfully been applied to estimate the properties of molecules and crystals.[59−63] Although previous work on generative models for materials has employed conditional or reinforcement-based models for properties, as far as we are aware, no previous work combines the generative models with dedicated property prediction that goes beyond standard formation-energy calculations.

Overall, existing methods for crystal-structure generation via representation learning are limited by the need for a priori information. For instance, iMatGen requires a user-defined chemical composition. This limits the generalizability of the model and means that the size of the inputs scale linearly with the number of distinct elements in the input crystals. Furthermore, for materials design purposes, we feel that it is desirable instead to condition the generation on properties rather than the composition, thus allowing for targeted generation of materials for functional applications.

Scope of this Work. We, herein, present a variational deep-representation learning pipeline for the creation of novel 3-D inorganic crystal structures that also predicts the values of eight associated properties. Using a voxelized crystal representation based on iMatGen[25] and Hoffmann et al.,[42] we train a conditional deep-feature-consistent variational autoencoder and UNet segmentation network to learn representations of cubic binary alloys, ternary perovskites and Heusler compounds. By using a conditional autoencoder, that encodes the electron-density maps alongside the formation energy per atom of the associated crystals, the VAE learns to encode both structure and properties simultaneously. Therefore, randomly sampling from the encoded space, subject to a user-defined formation energy condition, produces new examples of crystal structures. Furthermore, for each generated crystal we predict eight associated properties using a GNN. We validate our VAE-generated structures and GNN-generated predictions by comparing them to those that are computed with electronic-structure calculations. Overall, this enables researchers to generate high-quality candidate materials orders-of-magnitude faster than with experimental or ab initio methods.
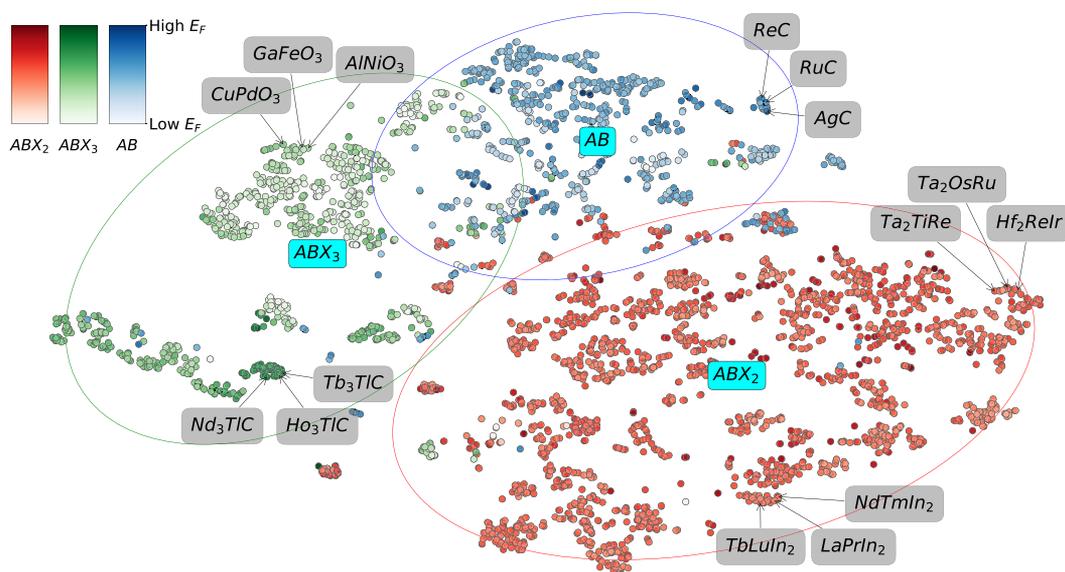
Figure 3. t-SNE embedding of the Cond-DFC-VAE latent space. 5000 training structures were encoded with the Cond-DFC-VAE. The latent space shows clustering according to the three structural types. The points in each cluster are colored according to the formation energy per atom, with more intense colors corresponding to high formation energies. Distinctive clustering by both chemical composition and formation energy is observed.

RESULTS AND DISCUSSION

The operational pipeline of our model architecture for this work is given in Figure 1. It consists of three main system components that together predict 3-D crystal structures and their eight associated properties. We now briefly describe these three main system components in sequence. We first represent crystallographic unit-cells as voxelized electron-density maps. The maps are used to train a conditional deep-feature-consistent VAE (Cond-DFC-VAE) that learns a latent encoding of the crystal structures and their properties. Sampling of the latent space, with a user-provided property condition, produces novel electron-density maps. The density maps are then converted to atomic sites via a combination of a UNet semantic segmentation network and morphological transformations. Finally, a crystal-graph convolutional-neural network (CGCNN) is used to predict eight associated properties of these materials, namely, their formation energy per atom, energy per atom, band gap, bulk and shear moduli, Poisson ratio, refractive index, and dielectric constant.

We trained the VAE and UNet models independently on 78 750 crystallographic information files (CIFs) of computationally generated crystal structures of ternary perovskites, binary alloys, and Heusler compounds which were obtained from the Materials Project. We retained 14 000 of them for out-of-sample validation. The CGCNN model was trained on a general set of structures, not limited to the aforementioned classes. A full description of the architecture and training procedures are given in the Methods section.

The next four sections present the results of various validation steps that were applied to verify this pipeline operation. To this end, we demonstrate that the VAE latent space is smooth, interpretable and can, therefore, be sampled to produce high-quality electron-density maps. Second, out-of-sample validation shows that the pipeline accurately reconstructs atomic positions and unit-cell parameters. We next show that our CGCNN implementation is able to accurately predict DFT-calculated properties. In the fourth section, we compare crystal structures produced by our

operational pipeline to pre-existing materials and use DFT to geometrically optimize crystal structures of selected compounds. The results confirm that the VAE-generated crystals are valid and highly optimized.

Encoding of 3-D Electron-Density Maps: Interpretation of the Latent Space. The Cond-DFC-VAE aims to learn a smooth and compressed encoding of electron-density map features. For all tests herein, the Cond-DFC-VAE is trained on the electron-density maps, while simultaneously being conditioned on their formation energy per atom. From a practical standpoint, this means that the formation energy per atom of each crystal structure is quantized into deciles, converted to a one-hot encoding and then concatenated onto the VAE layers at the input and bottleneck as shown in Figure 1a.

The one-hot encoding of the property-condition vector was chosen, herein, as it ensures the magnitudes of the vector elements are in the $[0, 1]$ interval, thereby matching the scales of the 256-dimensional electron-density map encoding. Furthermore, by having 10 dimensions for the property-condition vector, greater emphasis in the latent-space is placed on the properties of the crystal-structures (10 dimensions in 266), compared to encoding the property as a single value (1 dimension in 257). However, we note that the architecture can be easily modified to encode the property-condition vector in any form, including discrete or continuous representations. Furthermore, the VAE can be trained on any property for which data exist, including categorical variables. The formation energy per atom was chosen, herein, since this property is most readily available in our training set.

A smooth VAE encoding is required for the latent space to be easily sampled and thereby produce realistic and novel instances. Once trained, the dimensions of the latent space were found to be sufficiently smooth, as evidenced via the kernel density estimate (KDE) plot in Figure 2a: all 256 dimensions of the latent vectors show approximately unit Gaussian profiles with slight variations in mean and variance (the 256 latent dimensions show average mean and variances
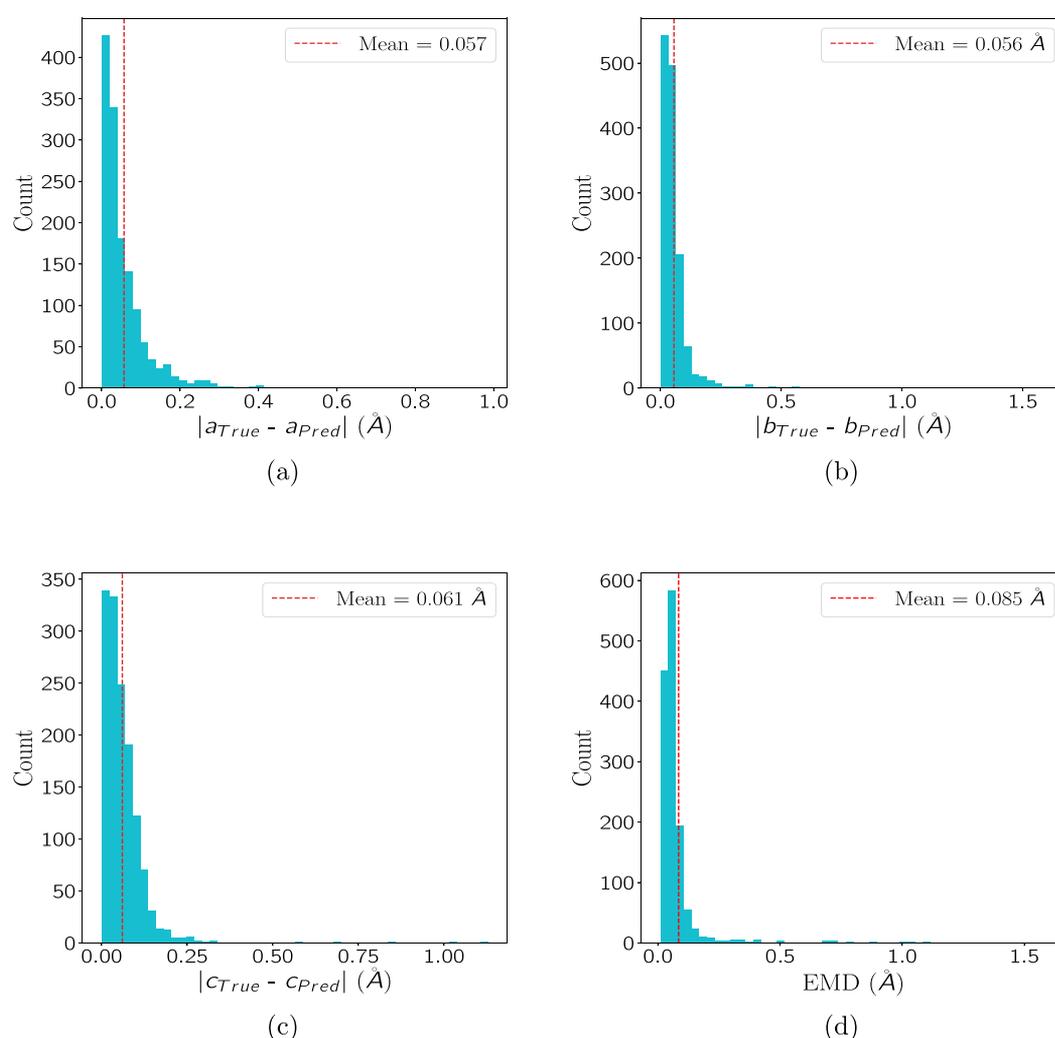
Figure 4. Evaluation of unit-cell parameters and atomic coordinate predictions. (a–c) Distributions of the absolute errors of prediction for the a, b, and c unit-cell parameters on the out-of-sample data set. The red dashed line indicates the mean absolute error (MAE). (d) Average earth-mover distance (EMD) between predicted atomic sites and the ground truth, on the out-of-sample data set.

of 0 and 0.99, respectively). This helps to ensure that the training samples are confined to high-probability regions of the latent space, thereby reducing the chance of sampling unrealistic or blurry samples.

By enforcing continuity on the learned representation, it is possible to interpolate between points in the latent space. This permits an exploration into how different latent dimensions affect the resulting samples. If the VAE has learned a meaningful and smooth representation of the crystals, then interpolations between points in the latent space will show smooth transitions between states without crossing through low-probability regions. Figure 2(b) shows midslices through electron-density maps that result from linear interpolations between 10 pairs of training compounds (with real samples at both ends). Each row corresponds to a different condition of formation energy per atom, with the lowest formation energy pair (CsYbBr$_3$ and NpTiO$_3$), along the top and the highest energy pair (Mn$_2$CuSb and OsPtCl$_2$) along the bottom. All examples show smooth transitions and qualitatively reasonable intermediate samples. This indicates that the latent space is smooth across all property dimensions.

Most importantly, the latent space should be in some way interpretable concerning the features of the crystal structures that it represents. This allows for targeted sampling of the latent space to produce specific materials. Figure 2c shows the result of interpolating between two similar end members of the rare-earth chromites, CeCrO$_3$ and YbCrO$_3$, whose A-site atoms sit at opposite ends of the lanthanides. 1000 samples were generated along the interpolation vector and segmented to form new crystal structures. It is important to note that the latent vectors produce intermediate chromite members, such as HoCrO$_3$, DyCrO$_3$, and PmCrO$_3$. These intermediate samples have identical cubic perovskite structures to the end members, except that the A-site material increases in atomic number. This demonstrates the ability for ions to "traverse" the periodic table between structural isotypes. Crucially, the interpolations also produce chromites that do not appear in the training set, revealing the ability to generalize patterns to produce new materials.

The interpretability of the latent space is further highlighted in Figure 3. The plot shows a t-SNE plot of 5000 crystal structures encoded with the Cond-DFC-VAE. The latent space

shows a distinctive clustering of the structural types corresponding to the binary alloys, ternary perovskites, and Heusler compounds. The color intensity in Figure 3 illustrates the formation energy per atom of individual structures, with brighter colors showing higher formation energies and vice versa. Accordingly, within clusters we find that similar compounds are closer together. For instance, the carbon binary compounds are clearly distinguished from the other binaries, and the high-formation energy perovskites are well separated from the low-energy examples. This partitioning of the latent space therefore allows us to target specific regions of it when generating new materials.

Evaluation of Predicted Unit-Cell Parameters and Atomic Positions. VAEs have previously been shown to have difficulty in determining Cartesian coordinates from images.[64] For crystal-structure prediction, the dimensions of the unit-cell, and associated interatomic distances, are vital in determining chemical stability. Previous works in this domain have largely neglected this aspect of crystal-structure generation. In this work, a coordinate convolution technique, as introduced by Liu et al.,[64] is employed to resolve this problem. Before input into the VAE, each voxel of the electron-density map is concatenated with its 3-D Cartesian coordinates in the original crystal geometry. As a result, the reconstructed electron-density maps (M' in Figure 1) contain implicit knowledge of the unit-cell dimensions. With a simple deterministic transform, it is possible to derive the unit-cell parameters, a, b, and c. This means that the unit-cell parameters of the generated crystals can be calculated solely from the electron-density map, prior to segmentation by the UNet.

The efficacy of our coordinate-convolution technique implementation was evaluated by comparing the predicted unit-cell parameters of the out-of-sample data against the ground-truth unit-cells, using the mean absolute error (MAE). The results are shown in Figure 4a−c. The validation set contains 1,300 unique crystal structures with unit-cell lengths ranging in size from 2 to 10 Å. The average MAE values for a, b, and c are 0.06, 0.06, and 0.06 Å, respectively. This is highly encouraging as it reveals that the system is able to reconstruct the unit-cell lengths accurately over the full range of sizes. The pipeline has a slight tendency to overestimate the true unit-cell lengths; however, the bias is consistently present across all magnitudes. We attribute this bias to the finite grid resolution that leads to slight rounding errors in the exact location of each voxel. For instance, for a unit-cell of approximately 5 Å, the 32 × 32 × 32 grid resolution gives a voxel size of 0.15 Å; therefore, any given coordinate can be offset from the voxel center by this distance. Increasing the grid resolution would likely reduce this bias and improve the resulting unit-cell reconstructions. However, the finer resolution would also greatly increase the computational requirements in terms of both memory and model size.

The ability of the combined VAE-UNet pipeline to reconstruct atomic positions was also evaluated. This verification step considered the average earth-mover distance (EMD) between the ground-truth atomic sites and predicted atomic sites after encoding, decoding and segmentation. As shown in Figure 4d, we see that, on average, atomic sites are 0.09 Å from their true locations. It is, therefore, clear that we are able to reconstruct the unit-cells and atomic positions of the crystals to a high degree of accuracy.

These results are better than those achieved by the iMatGen system,[25] which was trained on 10 980 real and augmented

V_xO_y-type materials. iMatGen achieved a root mean squared error of 0.1 Å for the unit-cell parameters and 0.2 Å for the atomic coordinate reconstructions.

Evaluation of Property Predictions. The CGCNN was trained independently from the VAE and UNet architectures on a data set of CIFs from the Materials Project to predict eight DFT-calculated properties of each material. These are the formation energy per atom, energy per atom, bulk modulus, shear modulus, refractive index, dielectric constant, Poisson ratio and the band gap. These properties were chosen as they were the properties for which enough data points existed in the Materials Project database. The predictive capabilities were evaluated using out-of-sample test sets for each property in terms of the MAE and the symmetric mean absolute percentage error (SMAPE). The performance of each model is outlined in Table 1.
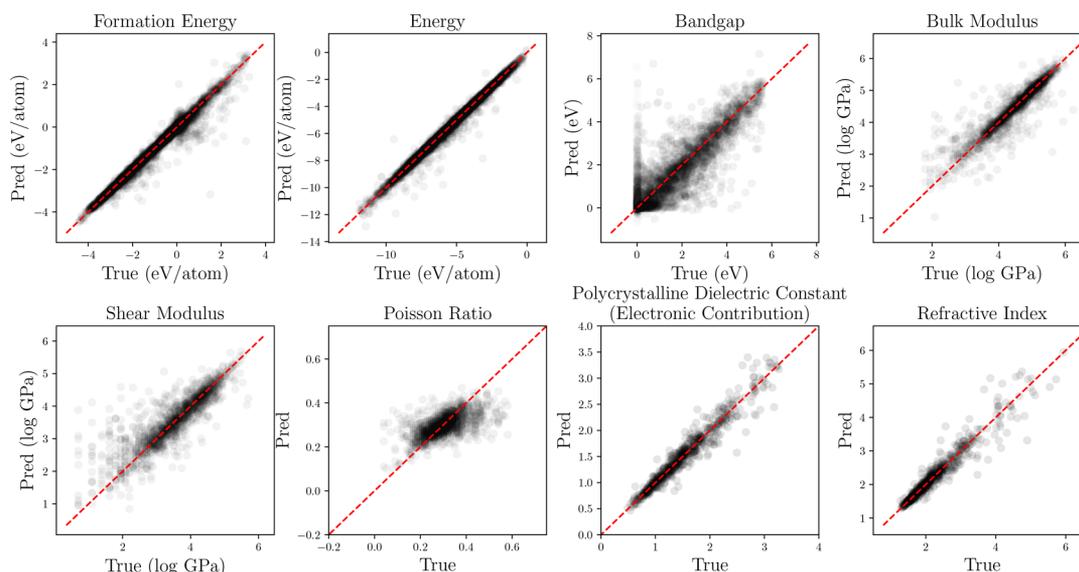
Table 1. CGCNN Property Prediction[a]

	MAE	SMAPE (%)	train no.	test no.
formation energy	0.102 (eV/atom)	27.70	35734	6700
total energy	0.114 (eV/atom)	2.51	35734	6700
bandgap	0.299 (eV)	137.73	35734	6700
bulk modulus*	0.186 (log(GPa))	5.23	8724	1635
shear modulus*	0.306 (log(GPa))	10.72	8192	1536
poisson ratio	0.051	16.72	8448	1585
refractive index	0.155	5.34	3744	702
dielectric constant*	0.115	6.49	3744	702

[a]Mean absolute error (MAE) and symmetric mean absolute percentage error (SMAPE) of each property, as well as the sample sizes used to train and evaluate the models. Asterisk (*) indicates log-scale.

In general, the models perform very well at predicting DFT properties. The relatively high MAE and SMAPE for the band gap property is visually evident in Figure 5, where the spike at 0 eV on the "True" axis indicates that several crystal structures are predicted to have a band gap, while they are not estimated to have a band gap in the Materials Project, cf. the "Pred" axis. The errors in the DFT-calculated band gap data, used to train our band gap model, are well documented in the Materials Project wiki;[5,65] they are primarily due to a derivative discontinuity term in the true density functional, as well as other approximations in the exchange-correlation functional. In principle, it is possible to solve these issues and calculate more accurate band gaps with DFT, as outlined by the Materials Project.[65] However, methods to do so are currently not implemented by the Materials Project, and they stress that the current computed band gaps "should be interpreted with caution".[65]

Generation of New Crystal Structures. It is clear from the evaluation steps described above that the pipeline is able to create novel electron-density maps, accurately transform them back to atomic coordinates and predict their associated properties to a high-level of accuracy. We now explore how new crystal structures are generated by sampling from the VAE latent space. By virtue of implementing a conditional variational autoencoder architecture, whereby the latent space of the VAE is concatenated with the decile-quantized formation energy per atom of the individual crystal structures, we are able to sample new crystal structures by providing a 10-dimensional one-hot encoded property-condition vector and a

Figure 5. Property prediction evaluation. Ground-truth property values versus CGCNN predicted values for all property-prediction models.

256-dimensional latent vector drawn from a standard normal distribution, $\mathcal{N}(0, 1)$. These are concatenated at the input of the decoder (cf., Figure 1a), decoded to form a new electron-density map and then segmented with the UNet to create a new crystal structure.

We first examined crystal structures generated via pure random sampling of the VAE latent space. For this, we sampled 1000 latent vectors from $\mathcal{N}(0, 1)$ (100 samples per formation energy condition) and passed them through the decoder-UNet segmentation pipeline. This yielded 760 (76%) crystal structures that were valid in terms of interatomic distances, as defined by the PyMatGen python package.[65] This algorithm checks the pairwise Euclidean distances of all sites in the crystal structure and deems it as valid if all sites are separated by more than a minimum distance threshold (taken to be 0.5 Å herein). Figure 6a shows the distribution of the top-10 most commonly generated crystal-structure types. We observe that the system produces crystal structures beyond those provided in the training set, such as A, AB_2, ABC, ABCD, and A_2B_3-type structures (examples of some of these crystal structures are shown in Figure 6b). This demonstrates that the system is able to generalize beyond the structures provided in the training set.

Taking formation energy per atom as a crude estimate of crystal structure stability, we found that of the 760 valid crystal structures, 450 (59%) had a CGCNN-predicted formation energy per atom below 0.0 eV/atom. Figure 6c shows boxplots of the corresponding CGCNN-derived formation energy per atom predictions for each crystal structure. Figure 6d also shows the distribution of the top-15 most common individual chemical elements in the VAE-generated set. The system has a tendency to produce compounds that contain chemical elements, which are most common in the training set, such as Li, O, and Mg. However, in general, the system shows good generalizability, producing compounds with rarely occurring chemical elements, such as Ac, Pu, and Mo.

The validity of the randomly generated compounds was explored by matching the chemical formulas against existing cubic structures in the Materials Project database. Of the 760 valid crystal structures, we found 75 matches. For each, we evaluated the MAE in predicted unit-cell lengths and formation energy per atom. Overall, we found an MAE of 0.6 eV/atom for the formation energy values and 0.85 Å for the unit-cell lengths, a, b, and c, respectively. This shows that the system produces unseen chemical compositions that are known to be chemically valid to a high degree of accuracy.

This result shows that the system is able to generalize to new crystal-structure types that are not present in the training set, despite being trained on a small set of different crystal-structure types. However, training the system on a more diverse set of structures would further improve the generalizability of the VAE samples. In this work, we limited our training set to the three structure types owing to computational constraints. Extension to more structure-types forms part of the future work presented in the Conclusions and Future Work section.

Ab Initio Validation of VAE-Generated Crystal Structures. Although the VAE-generated compounds can be sampled randomly from the entire VAE latent space (as shown above), we believe that it is more useful and intuitive for the user to sample the latent space around or between an existing compound. This reflects more accurately the materials-discovery process, where researchers often aim to find materials that have similar characteristics to other known materials. This follows from the VAE latent-space plot in Figure 3, where we observe distinctive clustering of similar materials by crystal-structure type, chemical composition and formation energy per atom. This kind of targeted sampling has also been shown to yield a higher success rate of generation in the iMatGen system.[25]

To test the targeted sampling, we benchmarked VAE-generated crystal structures against DFT-calculated structures. A test set for DFT validation was generated by randomly choosing 12 base compounds (4 per structure type) from the VAE training set to form the latent-space points around which novel samples would be generated. The latent space surrounding the encoded form of each of the base compounds was sampled, in turn, with a variance of 0.5, drawing 1000 samples from $\mathcal{N}(\mu_{base}, 0.5)$ where μ_{base} is the latent vector of each base compound. In each instance, the formation energy was conditioned on the one-hot encoded formation energy per atom of the base compounds. The quality of these VAE-
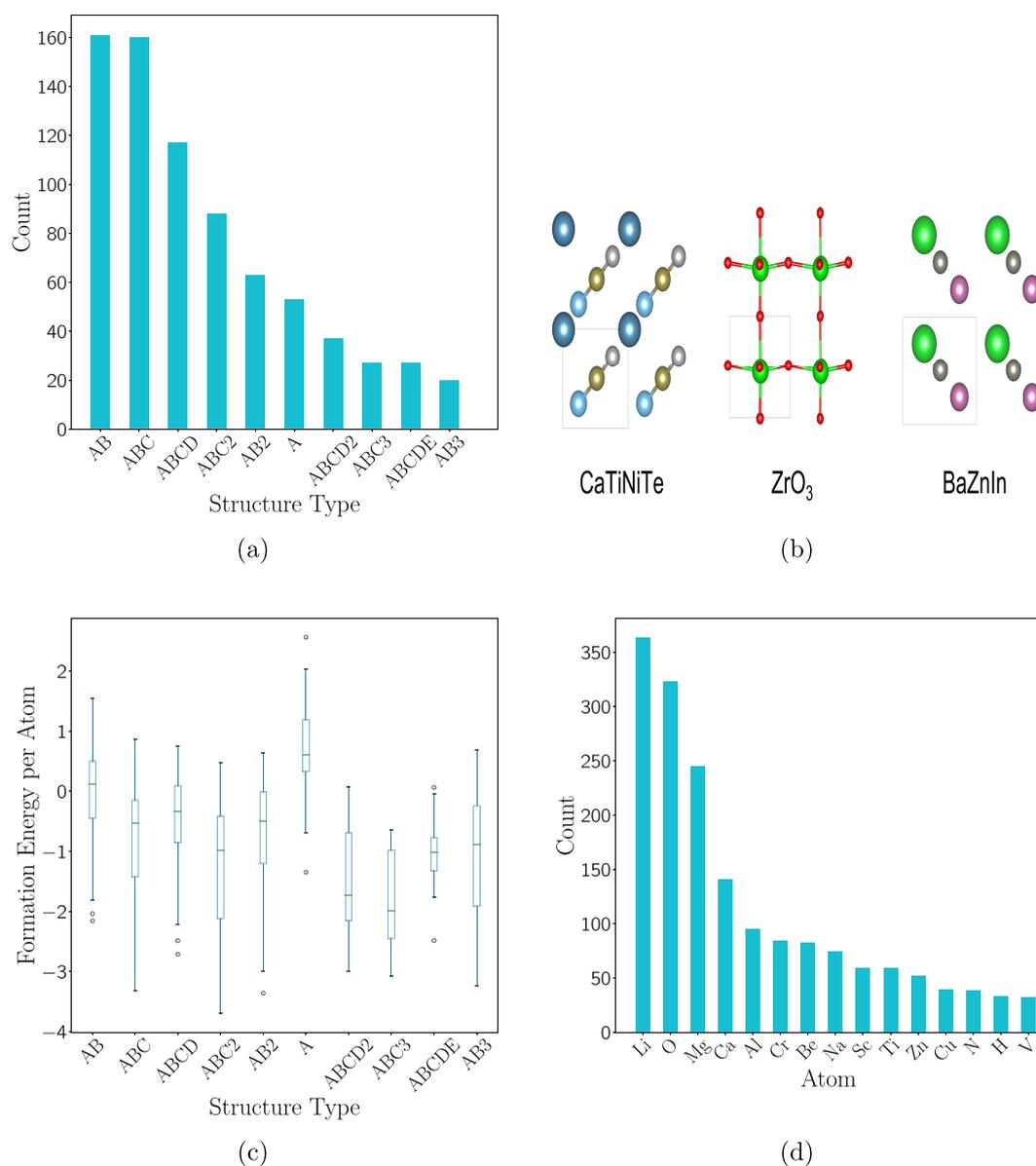
Figure 6. Results of crystal-structure generation via random sampling. (a) Distribution of the top-10 crystal structure types in the VAE-generated set. (b) Examples of some generated crystals with structural types not present in the training set. (c) Box plots of the formation energy per atom of the generated crystal structures by the top-10 most common structure types. Circular dots indicate outliers. (d) Distribution of the top-15 individual elements occurring in the VAE-generated set.

generated structures was verified by benchmarking a subset of them against crystal structures that have been geometrically optimized using DFT. The quality control manifested as the extent by which each DFT-generated crystal structure differed from its cognate VAE-generated crystal structure.

Since 1000 VAE-generated crystal structures result from the data sampling of each base compound, there may be a large number of material "candidates" upon which to apply DFT. The computational cost of performing DFT calculations on all candidates would be too great. Therefore, a selection procedure was needed to filter down the number of candidates to a tractable effort of DFT calculations. A subset of the candidates were selected according to the following criteria, whereby each compound must

1. be present as a VAE-generated crystal structure in the data sampled about the base compound

2. be valid in terms of interatomic distances
3. have the same stoichiometric formula as the base compound
4. not exist in the VAE-UNet training set
5. have a formation energy within 20% of the base compound

After performing these steps, duplicate samples were removed and the sample compound with the lowest formation energy per atom was retained. In cases where multiple candidates passed the above criteria for a given base compound, the 10 candidates closest to the base compound in terms of formation energy per atom were selected to be optimized via DFT. Because of the limitations of DFT when applied to the electron-rich actinide elements, we were unable to perform DFT on those candidates containing such elements. This was no hardship since the radioactivity of the actinide

elements makes them impractical for materials discovery that leads to most functional applications. The filtering process yielded a total of 76 candidates for DFT optimization. The full list of candidates is given in Table S1.

Note that we opted for this filtering process to reflect the case in which a user wishes to generate materials that have a particular structural type and a particular property within a desired range (formation energy per atom). This is the opposite of the fully random-sampling process given above.

Table 2 shows the average number of candidates that meet the above criteria at each stage of the filtering process. On

Table 2. Success Rate of Crystal Structure Generation from a VAE, Sampled around 12 Randomly Selected Base Compounds

filtering step	average no. samples	cumulative total (%)
initial samples	1000	100
interatomic distance validation	890	89
stoichiometric formula	350	35
not in training set	250	25
within 20% of target E_f	60	6
unique composition	20	2

average, each run produces 20 unique and valid sample compounds that meet the above criteria. We observe that the targeted sampling yields 89% valid crystals in terms of interatomic distances, compared to 76% for the random-sampling process.

Examples of some VAE-generated crystal structures, one for each material class, are illustrated in Figure 7. In general, we
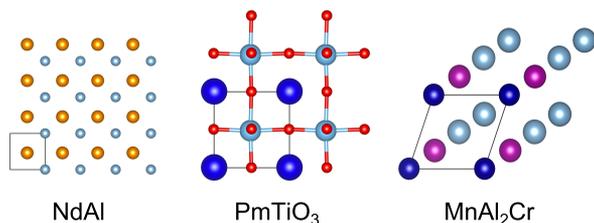
Figure 7. VAE-generated crystal structure from each of the three material classes: binary alloy, NdAl; ternary perovskite, PmTiO$_3$; Heusler compound, MnAl$_2$Cr.

observed the desired behavior that by sampling around a particular base compound, the system generates crystal structures that are chemically and structurally similar. For example, sampling around CeCrO$_3$ yields other perovskite oxides. Accordingly, increasing or decreasing the sample variance decreases or increases the probability of drawing crystal structures that are similar to the base compound. This is important as it allows us to target particular areas of the VAE latent space and, thereby, increase the probability of generating likely materials.

We performed DFT optimization on all of the 76 candidates and compared the geometry-optimized crystal structures to the VAE-generated crystal structures via three metrics: (1) the percentage change in bond lengths as a percentage of the unit-cell, (2) the absolute change in unit-cell parameters, and (3) the difference between the CGCNN-predicted and DFT calculated formation energy per atom. Table 3 shows these results broken down by structural type.

Table 3. Comparison of DFT Calculations Performed on Our 76 Material Candidates by Structural Type^a

type	no. candidates	$|\delta bonds|$ (%)	$|\Delta E_f|$ (eV/atom)	$|\delta cell|$ (%)
perovskites	30	5.23	0.57	0
Heusler	22	1.44	3.34	0
binaries	24	10.88	2.65	0

^{a}The geometry-optimized DFT results were compared against their cognate VAE-generated crystal structures. Results are analyzed in terms of difference in the absolute in formation energy, $|\Delta E_f|$, mean absolute percentage change in bond lengths, $|\Delta bonds|$, and unit-cell parameters, $|\Delta cell|$.

First, we see that across all the candidates, the geometry-optimization process engenders a modest change in the atomic bond lengths of any VAE-generated material candidate; cf. the overall average changes are 5.23%, 1.44%, and 10.88% for the perovskite, Heusler, and binary candidates, respectively, affording a total average percentage change of 6%. However, as shown in Table S1 there is quite a bit of variation between materials, from essentially 0.0% change to cases of binary compounds that show bond-length changes of around 30%. We attribute these large changes to the delocalized nature of binary materials and the relative difficulty by which these materials are geometry optimized owing to their tendency to contain particularly heavy elements, such as Au and Sb.

All the candidates converged, but two candidates (Ho$_2$AgAu and Ho$_2$CdPd) required too large a k-point grid to be optimized using our architecture. None of the candidates showed a change in unit-cell parameters (i.e., no change in the lattice constants or angles were observed). For the vast majority of candidates, the unit-cell relaxations converged in a single iteration with no resulting change in unit-cell dimensions. Three candidates, CoAs, Ho$_2$YbCd, and PrMgO$_3$, required two iterations but yielded no change in unit-cell parameters. Cell relaxation in DFT is traditionally used to determine the quality of convergence of the key parameters, such as the plane-wave cutoff and the k-point mesh being used. It serves as a sanity check for the chosen parameters and indicates that the stresses within the crystal-structure have been minimized. This gives a strong indication that the DFT-models are well optimized. Although there were some instances of larger than average bond length changes, almost all bond length changes were reductions in bond length. This meant there was no added strain on the unit-cell parameters resulting in no changes to the lattice constants, angles, or cell volume. This is a particularly encouraging result since it indicates that the coordinate-convolution method employed in our Cond-DFC-VAE successfully allows for the accurate generation of unit-cells which are locally stable. A full description of the electronic-structure calculations can be found in the Methods section.

The CGCNN-predicted formation energies also match closely to the DFT calculated values. On average the perovskites, Heuslers, and binaries show differences of 0.57, 3.34, and 2.65 eV/atom, respectively (see Table 3). This gives an overall average of 1.99 eV/atom across all candidates (see Table S1). There is again a significant variance with compounds containing heavy elements showing the largest divergence between the predicted and calculated values. This is expected since the DFT-optimization of bond lengths and unit-cell parameters can have a large effect on the resulting energies. Furthermore, determination of formation energies in bulk rare-

Table 4. Evaluation of MAE between VAE-Generated Candidates and Pre-existing Records in the Materials Project Database

candidate	a (Å)	b (Å)	c (Å)	E_f (eV/atom)	energy (eV/atom)	bulk modulus (GPa)	shear modulus (GPa)	poisson ratio	ICSD ID
Lu_2ZnAu	0.19	0.30	0.20	0.02	0.16	84.39	2.36	0.09	
Ho_2ZnAu	0.21	0.18	0.06	0.04	0.19				
$YbSmPd_2$	0.26	0.21	0.17	0.00	0.15	69.57	2.43	0.11	
KWO_3	0.05	0.06	0.15	0.03	0.11	184.78	57.78	0.06	
$CaBeO_3$	0.12	0.28	0.21	0.35	0.70				
$AlSiO_3$	0.22	0.05	0.07	0.45	0.46				
$TiAlO_3$	0.01	0.03	0.10	0.19	0.05	228.63	4.60	0.07	
Sc_2RuRh	0.47	0.34	0.20	0.16	0.05	137.26	55.71	0.00	
$PaSiO_3$	0.45	0.57	0.16	0.03	0.41				
$ZnAu$	1.86	1.79	1.77	0.26	0.28	113.78	13.56	0.20	108022
$NpAlO_3$	0.27	0.33	0.37	0.32	0.89				
$AcMgO_3$	0.20	0.10	0.46	0.12	0.40				
$LiCd$	0.95	0.90	1.49	0.02	0.80	34.67	0.67	0.09	620101
$NaTl$	1.33	1.27	1.53	0.21	0.91				657524
MAE	**0.47**	**0.46**	**0.50**	**0.16**	**0.40**	**121.87**	**19.59**	**0.09**	

earth metals and chalcogens can be difficult (as demonstrated in Table S3). Regardless, an average absolute error of 1.99 eV/atom across all material candidates shows that our pipeline is able to produce VAE-generated crystal structures that are comparable to DFT-generated crystal structures.

Most importantly, the system was able to produce 12 000 candidate materials in 12 h on a typical desktop workstation with a single GPU. In comparison, DFT optimization on the candidate structures took 20 000 h of CPU time using HPC resources on the Intel Haswell cores of Cooley machines at Argonne Leadership Computing Facility, IL, USA.

Comparison to Other Computationally Determined Crystal Structures. In addition to the DFT validation above, we found that 14 of our VAE-generated matched cubic crystal structures that had been computed previously via DFT in the Materials Project (and were not part of any model training set). Accordingly, the VAE-generated structures and their cognate CGCNN property predictions were compared to those reported by the Materials Project. The results are shown in Table 4. In general, we find very good correspondence between the two sets of crystal structures, with unit-cell parameters showing a MAE of approximately 0.45 Å. Similarly, the property predictions for the formation energy per atom, energy per atom, shear modulus, and Poisson ratio all show very accurate results with average MAE values 0.16 eV/atom, 0.40 eV/atom, 19.6 GPa, and 0.09, respectively. The bulk moduli predictions are less good, which we attribute to the small training set size used to train the corresponding CGCNN model. No calculated values were reported in the Materials Project database for the band gap and dielectric constant properties. We also note that the three worst candidates in terms unit-cell reconstruction are all binary compounds as noted earlier, and the unit-cell predictions of these are significantly worse than the perovskites and Heuslers. This is likely due to the fact that the binary compounds are the smallest group of materials in the training set, and are therefore more difficult to generate.

A further 16 candidates from our filtered VAE-generated set of crystal structures were found to have noncubic polymorphs in the Materials Project database (see Table S2). By constraining our training set and thereby limiting the output of our VAE-system, to cubic crystals, we are currently unable to generate noncubic polymorphs which may be more physically valid for a given chemical composition. However,

it is important to note that our VAE-UNet architecture has correctly generated valid chemical compositions that have been shown to exist either computationally or experimentally (cf., ICSD ID entries in Table 4).

■ CONCLUSIONS AND FUTURE WORK

In this work, we have demonstrated a full representation-learning pipeline for 3-D inorganic crystal-structure generation and property prediction. By employing several methods from the computer-vision literature, such as a Deep Feature Consistent VAE with coordinate convolutions, we have successfully enabled meaningful encoded representations of crystal structures and thus have afforded the ability to generate crystals with desired properties using a conditional VAE. Comparison of our VAE-generated crystal structures to those geometrically optimized via DFT calculations shows that they are highly optimized and quantitatively similar to experimentally verified results. Geometry-optimized DFT calculations converge in very few iterations, with little change to bond lengths and unit-cell parameters.

The key distinction between this work and previous efforts by others on crystal-structure generation is that our VAE-UNet pipeline does not require any user-defined knowledge of the constituent elements of a crystal structure or need to be trained on all known structure types in order to generalize. We also apply no postprocessing steps to the crystal structures in order to increase the success rate. We have achieved this by incorporating the UNet segmentation pipeline introduced by Hoffmann et al.[42] and adding element-specific parameters into the electron-density map encoding. As shown in the Methods section, this affords a considerable improvement in segmentation accuracy and coordinate reconstruction. Furthermore, utilizing the DFC model greatly reduces the characteristic "blurring" of samples that is often seen with VAE output. This improves the accuracy of the atomic coordinate predictions and therefore increases the likelihood of generating valid and stable crystals.

There are still limitations to our approach that we aim to resolve with future work. First, because of computational limitations, we have confined the resolution of voxelization to $32 \times 32 \times 32$ and limited our training data to crystal structures with less than 40 atoms per unit-cell. Extension of this to higher dimensions will allow for greater unit-cell complexity, and improve the reconstruction accuracy. With greater

computational capacity, it would also become possible to train the VAE-UNet pipeline in an end-to-end fashion, thereby mitigating the need for a DFC model.

We also restricted our study to cubic crystal systems. It was found here, and previously,[25] that electron-density maps of noncubic systems are distorted by the mapping of their electron-density matrix to a cubic box. In iMatGen, this distortion was used to accurately retrieve unit-cell angles by inverting the density-map function. In this work, we are unable to recover the unit-cell angles in the noncubic case due to the differing radius of each site in the electron-density map. However, by including an element-specific radius on each site, we dramatically improve the segmentation accuracy and therefore reconstruct atomic positions without the need for multiple element-specific channels in our inputs (see the Comparison of Model Performance section). Further work will look at modifying the electron-density map encoding to better handle noncubic crystals.

We have focused herein on training our pipeline on a limited set of crystal-structure types from the Heusler, perovskite, and binary alloy families of materials. This was primarily due to the computational resources required to train the large VAE and UNet models. Although, as demonstrated here, it is not necessary to train a new VAE-UNet pipeline for each new crystal-structure type, doing so may improve the quality of generated samples that are outside of these structure types. Our operational system also makes no distinction between stable and unstable materials beyond basic analysis of formation energies. Therefore, to produce materials for functional applications, it will likely be necessary to incorporate analysis of phase stability.

A key limitation of all crystal-generation methodologies is the reliance on DFT-generated data. Although widely accepted as a high-quality model for materials and their properties, DFT is known to have difficulty in certain specific cases. These quirks manifest themselves in the results presented above. For instance, DFT tends to underestimate band gap energies and struggles to geometrically optimize ionic materials or compounds containing heavy-elements or chalcogens. By training on DFT-generated data that contain these biases, our model also reflects the same problems. This highlights the need for large repositories of experimentally verified crystal structures that are concerted with their cognate properties. Future work will explore the use of transfer learning on experimental data to refine our models in these cases.

Overall, this work presents an important step forward in the way that materials discovery of crystalline compounds can be performed using deep-learning methodologies. With the method presented herein, tens of thousands of new samples of VAE-generated crystal structures can be generated in an order of minutes, with their associated DFT-calculated property predictions issuing a high degree of confidence. As a result, end users are able to generate multiple new candidate materials for potential device applications orders-of-magnitude faster than methods which adopt adaptive optimization or high-throughput experimental or computational approaches. Thus, our methods will enable users to better guide their research, whereby they could more efficiently employ the more expensive techniques, such as DFT and experiments, on only the most viable material candidates. Conversely, our method could form the basis of an adaptive-optimization scheme, whereby VAE-generated crystal structures of material candidates are automatically input to DFT calculations, with

the results fed back to the VAE via a reinforcement-learning feedback loop that is conditioned in a fashion that tailors user-desired properties. Overall, our advances contribute to the prospect of improving the capabilities of materials-discovery platforms, and dramatically improving their "design-to-device" timeline.

■ METHODS

Data Preparation and Formatting. The core data used for this work are crystallographic information files (CIFs) containing atomic positions and unit-cell parameters of 3-D crystalline materials. There are multiple open-source repositories of experimentally and computationally derived crystal structures. For all results presented herein, we use crystal structures that were obtained from The Materials Project.[5]

The Materials Project API was used to source 7189 CIFs for three material classes of interest, namely, cubic binary alloys (AB), ternary perovskites (general formula ABX_3), and Heusler compounds (ABX_2). The crystal structures were encoded into a format based on the works presented by Noh et al.[25] and Hoffmann et al.[42] Thereby, each crystal structure is represented by a $32 \times 32 \times 32$ voxelized electron-density map. We chose this representation as it is agnostic to the number of atoms and can easily be augmented to handle rotational symmetries.

For each crystal structure, five matrices were created: (i) a $32 \times 32 \times 32$ density matrix, M, representing the local electron-density of atoms, (ii) a $32 \times 32 \times 32$ species matrix, S, that assigns to each voxel the atomic number of the contained atom (including a zero-class for those containing no atoms), (iii) the corresponding binary species matrix, S_B, that labels each voxel as being occupied or not, (iv) a 1×3 lattice-parameter vector, $\vec{l}$, representing the crystallographic unit-cell lengths (a, b, c), and finally, (v) a $32 \times 32 \times 32 \times 3$ coordinate grid, C, giving the Cartesian (x,y,z) coordinate of each voxel.

M is computed using eq 1, whereby the value of the voxel, v, is given by

$$M_v = \frac{1}{(2\pi)^{3/2}} \sum_i \frac{Z_i}{r_i^3} \exp\left(-\frac{d(\vec{i}, \vec{v})^2}{2r_i^2}\right) \tag{1}$$

where Z_i is the atomic number of site i, r_i is the corresponding ionic radius and $d(\vec{i}, \vec{v})$ is the Euclidean distance between site i and voxel v. Thus, the value of each voxel is a sum of Gaussian electron-densities, where the contribution is dependent on nearby ions. This is a modified form of the implementations by Noh et al.[25] and Hoffmann et al.[42] such that the standard deviation of each Gaussian distribution is represented by the corresponding ionic radius rather than a fixed parameter. This is to assist the segmentation pipeline in identifying the atomic species. Furthermore, by encoding all ions into a single electron-density map, we negate the need for multiple element-specific channels in our input data. As such, the encoding is agnostic to the material class and does not scale-up linearly with the number of distinct elements.

The corresponding species matrix is defined by

$$S_v = \begin{cases} Z_i, & \text{if } d(i, v) \leq r_i \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where each voxel is labeled with the atomic number of site i provided v is within the ionic radius. In cases where the voxel lies within the label radius of multiple ions, it is assigned to the

closest ion. In addition, the binary mask of the species matrix, S_B, was created according to

$$S_B = \begin{cases} 1, & \text{if } d(i, v) \leq r_i \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

We augmented all these matrices by applying a number of random 90° rotations around the x-, y-, and z-axes to better learn the rotational invariance of the crystal structures. To reduce the complexity of the model, we focused only on cubic unit-cells that contain ≤40 atoms. We attempted to train our model on noncubic systems; however, distortions to the electron-density maps occurred in these cases which reduced performance dramatically. Extending the model to noncubic systems is a goal for further work on this project. In total, this processing pipeline results in 78 588 data points for the three aforementioned material classes. These data were split: 80% were used for training and 20% were retained for out-of-sample validation.

For the property prediction CGCNN, we trained on structures from a wide variety of crystal classes (including all structures used to train the VAE and UNet). To learn a general purpose structure−property prediction model. We outline the number of training samples used for each property model in Table 1. The crystal structures are prepared as per Xie and Grossman[60] (full details are provided in the SI).

Model Architecture. The architecture of our overall model can be broken down into three main components, as summarized in Figure 1. In step 1, we learned a latent space of crystal structures from electron-density maps using a Conditional Deep Feature Consistent Variational Autoencoder (Cond-DFC-VAE).[66,67] In step 2, a 3D UNet was used to segment electron-density maps into atomic-number segmentation maps, from which we recovered the atomic coordinates of a given structure. Finally, a Crystal Graph Convolutional Neural Network (CGCNN) was used to predict properties from structure.

During training, all three models were trained independently. When generating new structures, we first sampled the autoencoder latent space (conditioned on a desired target as outlined in the following section). The sampled latent vector and condition vector were concatenated and decoded by the decoder to produce an electron-density map. Subsequently, the UNet was employed to convert this electron-density map into an atom segmentation map, from which the Cartesian coordinates of the atoms were obtained via morphological transformations. Finally, the generated crystal structures were passed through a multitarget CGCNN to predict the values of eight target properties (formation energy, energy per atom, band gap, bulk modulus, shear modulus, Poisson ratio, refractive index, and dielectric constant). Further details of each component are provided below.

Conditional Deep-Feature-Consistent Variational Autoencoder for 3-D Crystal Structures. A vanilla VAE is a probabilistic latent variable model that estimates an intractable posterior distribution using a deep-neural network.[30] The architecture consists of an encoder network, E, and decoder network, D, that are jointly trained to encode and reconstruct the inputs. In our case, the combined model was trained to reconstruct input electron-density maps, M, such that the latent vector, $\mathbf{z}$, at the bottleneck, encoded semantically meaningful characteristics of the crystal structures.

Instead of estimating the posterior $p(\mathbf{z}|\mathbf{M})$ directly, the encoder network approximates it via a Gaussian distribution. This approximate posterior $q(\mathbf{z}|\mathbf{M})$ is therefore paramaterized by a mean μ and diagonal covariance Σ. It follows that to generate novel samples, $\epsilon \sim \mathcal{N}(0, 1)$ is sampled, and reparametrized by $z = \mu + \epsilon\Sigma$. The decoder network unravels the result to produce a new electron-density matrix.

During training, an attempt is made to minimize the reconstruction error between the input samples, M, and reconstructions, M', while simultaneously encouraging the latent vector, z, to be normally distributed. The latter is controlled via incorporation of the Kullback−Liebler divergence (KLD)[68] between the approximate posterior and the unit normal. In our model, the trade-off between continuity of the latent space and quality of reconstructions was controlled via a β-weighting on the KLD loss term.[69] Thus, the β-VAE loss is given by

$$L_{\beta\text{-VAE}} = \text{MSE}(M, M') + \beta L_{\text{KLD}} \tag{4}$$

where the MSE is the mean-squared error between the real and reconstructed electron-density maps and L_{KLD} is the KLD loss. The β-weighting parameter regularizes the degree of correlation between samples in the latent space.

The goal was to train a VAE that is able to produce samples of new crystal structures, which exhibit properties that are tailored to a given functional application. This goal was achieved by using a method introduced by Sohn et al.,[67] whereby the VAE was conditioned on a desired property. In the work presented herein, we conditioned the VAE on the formation energy per atom as retrieved from the Materials Project via its Application Programming Interface (API). This energy metric was chosen since it is present for all data points and it is a rough indicator of material stability.

In practice, conditioning on a target property value was achieved as follows. Since the properties that our model predicts are continuous, we discretized them into deciles, and created a 10-dimensional one-hot condition vector that is zero everywhere except for the index corresponding to the property decile. During training of the VAE, the property value is known for each training sample (due to supervised learning). We therefore created the condition vector and concatenated it to the latent vector produced by the encoder. We then allowed the decoder to reconstruct the electron-density map using this augmented latent vector. This is shown in Figure 1a. Adding these extra dimensions had the effect of manually partitioning the latent space of the VAE by the property values. During generation of new crystal structures, we constructed the condition vector based on the property value that we desired, and concatenated this to the sampled latent vector. Decoding this augmented latent vector thus produced an electron-density map corresponding to a structure which possessed the desired property value.

Vanilla VAEs (with or without β-weighting) have been shown to produce samples in a variety of domains.[70−72] However, owing to the enforced continuity of the latent space, generated samples tend to be visually blurry, or suffer from edge artifacts. Similarly, VAEs have been shown to exhibit poor performance when attempting to reconstruct Cartesian coordinates from images.[64] Owing to these problems, the β-VAE is unsuitable for use in 3-D crystal-structure generation where high-fidelity structures with defined Cartesian unit-cell lengths are required.

Two adaptations to the β-VAE implementation were employed in this work, in order to overcome its unsuitability for 3-D crystal-structure generation. First, the blurriness is reduced through the use of a deep-feature-consistent VAE (DFC-VAE) technique.[66] In a DFC-VAE, not only are M and M' compared via the MSE, they are also compared at individual layers of a separate perceptual model. The aim is to minimize the difference between the features embedded in the perceptual model and the overall image. In our model architecture, the hidden layers of the UNet are used (vide infra). At each layer, ϕ_p of the UNet, the electron-density matrices are encoded to a feature embedding. By comparing M and M' at each of these layers, good reconstructions are enforced across all features, thus reducing the blurriness of the final image. The DFC-VAE loss function incorporates a third term that describes the degree of difference between the real and reconstructed samples across multiple feature spaces. The overall loss function is then given by

$$L_{\text{DFCV}} = L_{\beta\text{-VAE}} + \alpha \sum_p \text{MSE}(\phi_p(M), \phi_p(M')) \tag{5}$$

where the coefficient α controls the degree of dependence on the perceptual model.

The second adaptation responds to the fact that previous work on creating stable and physical crystal structures has neglected the importance of the crystal-lattice parameters and relative positions of atoms in Cartesian space. We employed the coordinate-convolution method described by Liu et al.[64] to encode these features into our model. Thereby, at input, the electron-density matrix, M, is concatenated with a coordinate grid, C, that maps each voxel to Cartesian coordinates. This means the input and output matrices of the Cond-DFC-VAE have four channels, corresponding to the electron-density and three coordinates at each voxel.

We trained the Cond-DFC-VAE for 50 epochs and observed that a KLD loss of $O(10^2)$ provides a good balance between high-quality samples and reconstructions. The minimum out-of-sample MSE was 0.013, compared to 0.05 without the DFC model (see below for a comparison of the various models).

3-D Multiclass Atom Segmentation. A traditional UNet architecture was employed to determine the atomic species from the electron-density maps and convert them back to atomic sites in Cartesian coordinates. This problem was treated as one of semantic segmentation,[73] whereby each voxel of the electron-density map is individually classified as pertaining to a certain atomic site. Given an electron-density map M', the UNet generates a species matrix, S', with N_{atoms} channels, where N_{atoms} is the number of unique atoms in the data set, and also a binary segmentation matrix, S'_{B}.

$$S', S'_{\text{B}} = \text{UNet}(M') \tag{6}$$

The class (species) of each voxel is then taken as the argmax of the output layer activation function. During training, the UNet attempts to minimize the weighted categorical cross-entropy loss and binary cross-entropy loss defined by

$$L_U = \sum_v \overline{w}S \log(S') - S\log(S') - (1-S)\log(1-S') \tag{7}$$

where $\overline{w}$ is a vector of weights on each atom species and the sum is over all voxels. Incorporation of class-dependent weights is due to the imbalanced class distribution within the species matrices. For example, it is common in our study on 3-

D crystal structures for the background class (class-zero) to make up approximately 90% of the matrix. For all experiments presented herein, each class was weighted inversely proportional to the number of voxels of each in the training set. The weight of the zero-class was set to zero to mitigate the large class imbalance.

We trained the UNet for 50 epochs and achieved an out-of-sample Categorical Crossentropy (CCE) loss of 2.75, corresponding to an F1 classification score of 99%. Owing to the sparsity of the data, the recall for nonzero classes peaks at 87% on the out-of-sample data, indicating that we correctly classify atomic species on a per-voxel basis in nearly 90% of instances.

After segmentation, S' contains regions of labeled voxels corresponding to the approximate location of atomic sites. These need to be converted into atomic coordinates, which was achieved by finding the centroids of these regions using morphological clustering based on the Watershed algorithm[74] (for a full description of the algorithm see Supporting Information). This segmentation process overcomes the need for any user-defined knowledge in inversion of the crystal-structure representation.

Comparison of Model Performance. To summarize, we have herein employed three adaptations to a vanilla VAE-UNet pipeline. First, we have included the Cartesian coordinates of the crystal structures into the VAE inputs via the coordinate convolution method.[64]

As shown in the Evaluation of Predicted Unit-Cell Parameters and Atomic Positions section, this implementation allowed us to accurately reconstruct the unit-cell lengths of the crystal structures up to the resolution limit of voxelization. Second, we employed a DFC perceptual model during training of the VAE, by comparing the input crystal structures and the corresponding reconstructions at separate layers of the UNet. This enabled a reduction in the blurring of the VAE output samples. Third, during creation of the electron-density maps, we used the ionic radius of the atoms as the variance of the Gaussians centered on each atomic site (cf., eq 1).

Table 5 highlights the efficacy of these adaptations, by comparing three different models: (1) A vanilla VAE-UNet

Table 5. Comparison of VAE-UNet Performance before and after Inclusion of the Deep-Feature-Consistent (DFC) Model and Ionic Radius (r_i) in Electrondensity Maps[a]

system	VAE MSE (Å)	UNet nonzero recall (%)	Mean EMD (Å)	Mean Δ_{nsites}
VAE without DFC with $\sigma = 1.0$	0.012	55	0.150	1.0
DFC-VAE with $\sigma = 1.0$	0.013	55	0.112	0.25
DFC-VAE with $\sigma = r_i$	**0.013**	**87**	**0.090**	**0.198**

[a]The performance of our full Cond-DFC-VAE-UNet model is shown in bold formatting.

pipeline (without a DFC perceptual model) with the electron-density Gaussians fixed to a variance of $\sigma^2 = 1.0$, (2) A DFC-VAE-UNet pipeline with $\sigma^2 = 1.0$ and (3) the DFC-VAE-UNet pipeline with σ^2 set to be the ionic radius of each atom, r_i. The models were compared in terms of four evaluation metrics, the MSE of VAE reconstruction, the nonzero classification recall of the UNet, the average EMD between the ground truth and predicted atomic sites and the average difference in true and predicted number of sites per unit-cell, $|\Delta N_{\text{sites}}|$. For each test,

the VAE and UNet were trained until losses ceased to improve using the same hyperparameters given in Table 6.

Table 6. Hyperparameters of Our Model Used for All Training Procedures and Validation Steps

parameter	value
α	1.0
β	0.0005
VAE learning rate	0.0001
VAE batch size	20
UNet learning rate	0.000003
UNet batch size	10
CGCNN learning rate	0.001
CGCNN batch size	256

First, we note that inclusion of the ionic radius dependence improved classification recall by over 30%, increasing from 55% for the $\sigma^2 = 1.0$ case to 87% for the $\sigma^2 = r_i$ case. This confirmed our assumption that, by including the ionic radius into the electron-density map, the UNet is able to more easily identify the correct atom species. Similarly, we saw that inclusion of the DFC model improved the ability of the system to determine the atomic sites in the unit-cell. The average EMD between predicted atomic sites improved from 0.198 Å for the no-DFC model to 0.15 Å for the DFC-VAE. Accordingly a similar improvement was also seen in the predicted number of sites per unit-cell, with the no-DFC model mispredicting by approximately one site per unit-cell, whereas the DFC model improved this to 0.2 sites per cell. This is consistent with the notion that the DFC model reduces the blurriness of the generated samples, allowing the UNet to better identify individual atomic sites.

Crystal-Property Prediction via Graph Neural Networks. We employed GNNs to estimate eight physical properties of the crystal structures that were generated by our VAE. The inputs to the GNN were computed directly from CIFs, and are a node-feature matrix, an edge-feature tensor, and a node-neighbor index matrix. Each model began with a single fully connected layer, followed by a graph convolution layer, which we present in the framework of message passing.[61] The graph convolution layer consisted of the application of message and node-aggregation/update functions, constituting the message-passing graph-convolution phase of the model. This was followed by a readout function which, from the updated node-feature vectors, computes a single global feature-vector, producing the final graph representation of the input crystal structures.

In more detail, our model produced a graph representation of each crystal structure through a series of graph convolutions. We used the node message and update functions of Xie and Grossman[60] to update the node features of our input graphs.

The message function in layer l takes as input each node v, its neighbors denoted by $v_j \in \mathcal{N}(v_i)$, and the edge features between those nodes $e_{i,j}$. The indices i and j denote the index of a node and the index of its neighbor in the crystal graph, respectively. These features are concatenated (denoted by $\|$) to form the feature vector of a node which contains information about itself, its neighbors and the connections between them

$$z_{i,j}^{(l)} = \text{message}(v_i^{(l)}, v_j^{(l)}, e_{i,j}) = v_i^{(l)} \big\| v_j^{(l)} \big\| e_{i,j} \tag{8}$$

The update function subsequently updates each node by

$$\begin{aligned} v_i^{(l+1)} &= \text{update}(v_i^{(l)}, z_{i,j}^{(l)}) \\ &= v_i^{(l)} + \sum_j [\sigma(z_{i,j}^{(l)} W + b) \odot g(z_{i,j}^{(l)} W_s)] \end{aligned} \tag{9}$$

where $\sigma(x) = \frac{1}{1 + e^{-x}}$ is the sigmoid function, W, b are learnable weight parameters, W_s is a self-weight matrix,[60] and $g(x) = \ln(1 + e^x)$ is the Softplus activation function.

By chaining several of these crystal-graph convolution layers, each taking as input the node-feature matrix from the output of the previous layer, a graph representation is produced in which each node is some abstract representation of a neighborhood of projected atomic- and bond-feature vectors. In our models, we used a single graph-convolution layer, as we empirically found that this produced the best overall predictive capability and minimized overfitting when training on properties with fewer training examples. As our readout function, we applied node-wise average-pooling which served to reduce the dimensionality and produced the final graph representation $h \in \mathbb{R}^F$.

$$h = \text{readout}(v^{(L)}) = \frac{1}{N} \sum_i^N v_i^{(L)} \tag{10}$$

where v_i^L is the output of the final graph-aggregation layer with layer index L. Eq 10 is simply the node-wise average over all projected node features in the input graph, and produces a single global node-feature vector as output. Two linear layers are then used to map the graph representation h to the property being estimated. While each property is estimated using a separate model, all models (except formation energy and energy per atom) are trained by transfer-learning and fine-tuning, whereby a single GNN layer which has been pretrained on the formation energy property is used to improve performance and prevent overfitting. Details of the transfer-learning and fine-tuning procedures are outlined in the Supporting Information. This method allowed us to produce a robust and property-agnostic crystal-graph feature-extractor in advance, which can be fine-tuned for subsequent structure–property mapping tasks.

During prediction, all eight properties were initially predicted. However, in instances where a zero band gap was predicted for a crystal structure, the dielectric constant and refractive index predictions were discarded and set to NaN, considering that a nonzero band gap is necessary for a crystal structure to exhibit these properties.[75,76]

Model Parameters. All of the Cond-DFC-VAE, UNet, and CGCNN models were implemented and trained using the Keras[77] python package with the Tensorflow backend.[78] Training procedures were performed using the ADAM optimizer and used the hyper-parameter values outlined in Table 6.

Electronic-Structure Calculations. Periodic plane-wave DFT calculations were performed using the Quantum Espresso[79] suite of programs with norm-conserving pseudopotentials used throughout. All potentials were fully relativistic except for the rare-earth metals where scalar-relativistic ones in the 3+ configuration were used instead with the f-electrons frozen in the core. All calculations employed the PBE functional[80] and the Grimme DFT-D2 method.[81] Initially, the wave function cutoffs were optimized, followed by a

Brillouin zone sampling using the Monkhorst−Pack scheme. The calculations were considered to be optimized when the total energy converged to within a tolerance of 0.00005 eV per atom. The atomic geometry and unit-cell relaxations were considered converged when the energy between successive optimization steps was within 10^{-4} Ry and the forces within 10^{-3} Ry/Bohr. The unit-cell optimizations allowed for the relaxation of the unit-cell dimensions. Gaussian smearing was used throughout the calculations. The formation energy, E_{f}, was calculated by finding the difference between the total energy of the crystal and the sum of the energies of its constituent atoms in their bulk state. This was achieved following the same optimization procedure described above. Calculations required 20 000 CPU-hours on the Intel Haswell cores of Cooley machines at Argonne Leadership Computing Facility, IL, USA.

■ ASSOCIATED CONTENT

ⓈⒾ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.0c00464.

> Details of morphological segmentation algorithms, property prediction model implementation details, a full list of generated material candidates, comparison of generated materials to known polymorphs, and DFT optimization of delocalized compounds (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jacqueline M. Cole − *Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, U.K.; ISIS Neutron and Muon Source, STFC Rutherford Appleton Laboratory, Didcot, Oxfordshire OX11 0QX, U.K.; Department of Chemical Engineering and Biotechnology, University of Cambridge,, Cambridge CB3 0AS, U.K.;* ⓞ orcid.org/0000-0002-1552-8743; Email: jmc61@cam.ac.uk

Authors

Callum J. Court − *Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, U.K.*

Batuhan Yildirim − *Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, U.K.;* ⓞ orcid.org/0000-0001-7741-0516

Apoorv Jain − *Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, U.K.; Department of Chemical Engineering and Biotechnology, University of Cambridge,, Cambridge CB3 0AS, U.K.*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.0c00464

Author Contributions

[#]C.J.C. and B.Y. contributed equally.

Notes

The authors declare no competing financial interest.
The data set used in this study is available from the Materials Project via their API (https://materialsproject.org). All of the code required to train and use the model presented herein is available at https://github.com/by256/icsg3d.

■ REFERENCES

(1) Rajan, K. Materials Informatics. *Mater. Today* **2005**, *8*, 38−45.
(2) Jain, A.; Hautier, G.; Ong, S. P.; Persson, K. New Opportunities for Materials Informatics: Resources and Data Mining Techniques for Uncovering Hidden Relationships. *J. Mater. Res.* **2016**, *31*, 977−994.
(3) Maddox, J. Crystals from First Principles. *Nature* **1988**, *335*, 201−201.
(4) Kohn, W. Nobel Lecture: Electronic Structure of Matter—wave Functions and Density Functionals. *Rev. Mod. Phys.* **1999**, *71*, 1253.
(5) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; et al. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1*, No. 011002.
(6) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials Design and Discovery with High-throughput Density Functional Theory: The Open Quantum Materials Database (OQMD). *JOM* **2013**, *65*, 1501−1509.
(7) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): Assessing the Accuracy of DFT Formation Energies. *npj Comput. Mater.* **2015**, *1*, 15010.
(8) Draxl, C.; Scheffler, M. The NOMAD Laboratory: From Data Sharing to Artificial Intelligence. *J. Phys. Mater.* **2019**, *2*, No. 036001.
(9) Belsky, A.; Hellenbrandt, M.; Karen, V. L.; Luksch, P. New developments in the Inorganic Crystal Structure Database (ICSD): Accessibility in Support of Materials Research and Design. *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58*, 364−369.
(10) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2016**, *72*, 171−179.
(11) Merkys, A.; Vaitkus, A.; Butkus, J.; Okulic-Kazarinas, M.; Kairys, V.; Grazulis, S. *COD::CIF::Parser*: An error-correcting CIF Parser for the Perl Language. *J. Appl. Crystallogr.* **2016**, *49*, 292.
(12) Gražulis, S.; Daškevič, A.; Merkys, A.; Chateigner, D.; Lutterotti, L.; Quirós, M.; Serebryanaya, N. R.; Moeck, P.; Downs, R. T.; Le Bail, A. Crystallography Open Database (COD): An Open-access Collection of Crystal Structures and Platform for World-wide Collaboration. *Nucleic Acids Res.* **2012**, *40*, D420−D427.
(13) Grazulis, S.; Chateigner, D.; Downs, R. T.; Yokochi, A. F. T.; Quirós, M.; Lutterotti, L.; Manakova, E.; Butkus, J.; Moeck, P.; Le Bail, A. Crystallography Open Database − an Open-access Collection of Crystal Structures. *J. Appl. Crystallogr.* **2009**, *42*, 726−729.
(14) Downs, R. T.; Hall-Wallace, M. The American Mineralogist Crystal Structure Database. *Am. Mineral.* **2003**, *88*, 247−250.
(15) Ramprasad, R.; Batra, R.; Pilania, G.; Mannodi-Kanakkithodi, A.; Kim, C. Machine Learning in Materials Informatics: Recent Applications and Prospects. *npj Computational Materials* **2017**, *3*, 54.
(16) Mater, A. C.; Coote, M. L. Deep Learning in Chemistry. *J. Chem. Inf. Model.* **2019**, *59*, 2545−2559.
(17) Oliynyk, A. O.; Antono, E.; Sparks, T. D.; Ghadbeigi, L.; Gaultois, M. W.; Meredig, B.; Mar, A. High-throughput Machine-learning-driven Synthesis of Full-Heusler Compounds. *Chem. Mater.* **2016**, *28*, 7324−7331.
(18) Back, S.; Yoon, J.; Tian, N.; Zhong, W.; Tran, K.; Ulissi, Z. W. Convolutional Neural Network of Atomic Surface Structures to

Predict Binding Energies for High-throughput Screening of Catalysts. *J. Phys. Chem. Lett.* **2019**, *10*, 4401−4408.

(19) Court, C. J.; Cole, J. M. Auto-generated Materials Database of Curie and Néel Temperatures via Semi-supervised Relationship Extraction. *Sci. Data* **2018**, *5*, 180111.

(20) Beard, E. J.; Sivaraman, G.; Vázquez-Mayagoitia, Á.; Vishwanath, V.; Cole, J. M. Comparative Dataset of Experimental and Computational Attributes of UV/Vis Absorption Spectra. *Sci. Data* **2019**, *6*, 307.

(21) Tshitoyan, V.; Dagdelen, J.; Weston, L.; Dunn, A.; Rong, Z.; Kononova, O.; Persson, K. A.; Ceder, G.; Jain, A. Unsupervised Word Embeddings Capture Latent Knowledge from Materials Science Literature. *Nature* **2019**, *571*, 95−98.

(22) Cooper, C. B.; Beard, E. J.; Vázquez-Mayagoitia, Á.; Stan, L.; Stenning, G. B.; Nye, D. W.; Vigil, J. A.; Tomar, T.; Jia, J.; Bodedla, G. B.; et al. Design-to-Device Approach Affords Panchromatic Co-Sensitized Solar Cells. *Adv. Energy Mater.* **2019**, *9*, 1802820.

(23) Cole, J. M. A Design-to-Device Pipeline for Data-Driven Materials Discovery. *Acc. Chem. Res.* **2020**, *53*, 599−610.

(24) Zhou, Q.; Tang, P.; Liu, S.; Pan, J.; Yan, Q.; Zhang, S.-C. Learning Atoms for Materials Discovery. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E6411−E6417.

(25) Noh, J.; Kim, J.; Stein, H. S.; Sanchez-Lengeling, B.; Gregoire, J. M.; Aspuru-Guzik, A.; Jung, Y. Inverse Design of Solid-State Materials via a Continuous Representation. *Matter* **2019**, *1*, 1370−1384.

(26) Cai, H.; Bai, C.; Tai, Y.-W.; Tang, C.-K. Deep Video Generation, Prediction and Completion of Human Action Sequences. *Proceedings of the European Conference on Computer Vision (ECCV)* **2018**, 366−382.

(27) Bao, J.; Chen, D.; Wen, F.; Li, H.; Hua, G. CVAE-GAN: Fine-grained Image Generation Through Asymmetric Training. *Proceedings of the IEEE International Conference on Computer Vision* **2017**, 2745−2754.

(28) Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I. Language Models are Unsupervised Multitask Learners. *OpenAI Blog* **2019**, *1*, 9.

(29) Vasquez, S.; Lewis, M. MELNET: A Generative Model for Audio in the Frequency Domain. *arXiv (Audio and Speech Processing (eess.AS))*, 2019, arXiv:1906.01083. https://arxiv.org/abs/1906.01083.

(30) Kingma, D. P.; Welling, M. Auto-Encoding Variational Bayes. *arXiv (Machine Learning (stat.ML))*, 2013, arxiv:1312.6114. https://arxiv.org/abs/1312.6114.

(31) Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. In *Advances in Neural Information Processing Systems 27*; Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D., Weinberger, K. Q., Eds.; Curran Associates, Inc., 2014; pp 2672−2680.

(32) Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A. Automatic Chemical Design Using a Data-driven Continuous Representation of Molecules. *ACS Cent. Sci.* **2018**, *4*, 268−276.

(33) Samanta, B.; DE, A.; Jana, G.; Chattaraj, P.; Ganguly, N.; Rodriguez, M. NeVAE: A Deep Generative Model for Molecular Graphs. *Proceedings of the AAAI Conference on Artificial Intelligence* **2019**, *33*, 1110−1117.

(34) Jin, W.; Barzilay, R.; Jaakkola, T. S. Junction Tree Variational Autoencoder for Molecular Graph Generation. *Proceedings of the 35th International Conference on Machine Learning, ICML 2018*, Stockholmsmässan, Stockholm, Sweden, July 10−15, 2018; 2018; pp 2328−2337.

(35) Liu, Q.; Allamanis, M.; Brockschmidt, M.; Gaunt, A. In *Advances in Neural Information Processing Systems 31*; Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R., Eds.; Curran Associates, Inc., 2018; pp 7795−7804.

(36) Lim, J.; Ryu, S.; Kim, J. W.; Kim, W. Y. Molecular Generative Model Based on Conditional Variational Autoencoder for De-novo Molecular Design. *J. Cheminf.* **2018**, *10*, 31.

(37) Kearnes, S.; Li, L.; Riley, P. Decoding Molecular Graph Embeddings with Reinforcement Learning. *arXiv (Machine Learning (cs.LG))*, 2019, arXiv:1904.08915. https://arxiv.org/abs/1904.08915.

(38) Mnih, V.; et al. Human-level Control Through Deep Reinforcement Learning. *Nature* **2015**, *518*, 529−533.

(39) Dong, Y.; Li, D.; Zhang, C.; Wu, C.; Wang, H.; Xin, M.; Cheng, J.; Lin, J. Inverse design of two-dimensional graphene/h-BN hybrids by a regressional and conditional GAN. *Carbon* **2020**, *169*, 9.

(40) De Cao, N.; Kipf, T. MolGAN: An Implicit Generative Model for Small Molecular Graphs. *arXiv (Machine Learning (stat.ML))*, 2018, arXiv:1805.11973. https://arxiv.org/abs/1805.11973.

(41) Prykhodko, O.; Johansson, S. V.; Kotsias, P.-C.; Arús-Pous, J.; Bjerrum, E. J.; Engkvist, O.; Chen, H. A De Novo Molecular Generation Method Using Latent Vector Based Generative Adversarial Network. *J. Cheminf.* **2019**, *11*, 74.

(42) Hoffmann, J.; Maestrati, L.; Sawada, Y.; Tang, J.; Sellier, J. M.; Bengio, Y. Data-Driven Approach to Encoding and Decoding 3-D Crystal Structures. *arXiv (Machine Learning (cs.LG))*, 2019, arXiv:1909.00949. https://arxiv.org/abs/1909.00949.

(43) Nouira, A.; Sokolovska, N.; Crivello, J.-C. CrystalGAN: Learning to Discover Crystallographic Structures with Generative Adversarial Networks. *arXiv (Machine Learning (cs.LG))*, 2018, arXiv:1810.11203. https://arxiv.org/abs/1810.11203.

(44) Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A. A. Unpaired Image-to-image Translation Using Cycle-consistent Adversarial Networks. *Proceedings of the IEEE International Conference on Computer Vision* **2017**, 2223−2232.

(45) Kim, S.; Noh, J.; Gu, G. H.; Aspuru-Guzik, A.; Jung, Y. Generative Adversarial Networks for Crystal Structure Prediction. *arXiv (Materials Science (cond-mat.mtrl-sci))*, 2020, arXiv:2004.01396. https://arxiv.org/abs/2004.01396.

(46) De, S.; Bartók, A. P.; Csányi, G.; Ceriotti, M. Comparing Molecules and Solids Across Structural and Alchemical Space. *Phys. Chem. Chem. Phys.* **2016**, *18*, 13754−13769.

(47) Seko, A.; Togo, A.; Tanaka, I. Descriptors for Machine Learning of Materials Data. *Nanoinformatics* **2018**, 3−23.

(48) Battaglia, P. et al. Relational Inductive Biases, Deep Learning, and Graph Networks. *arXiv (Machine Learning (cs.LG))*, 2018, arXiv:1806.01261. https://arxiv.org/abs/1806.01261.

(49) Bruna, J.; Zaremba, W.; Szlam, A.; LeCun, Y. Spectral Networks and Locally Connected Networks on Graphs. *arXiv (Machine Learning (cs.LG))*, 2013, arXiv:1312.6203. https://arxiv.org/abs/1312.6203.

(50) Li, Y.; Tarlow, D.; Brockschmidt, M.; Zemel, R. Gated Graph Sequence Neural Networks. *arXiv (Machine Learning (cs.LG))*, 2015, arXiv:1511.05493. https://arxiv.org/abs/1511.05493.

(51) Battaglia, P.; Pascanu, R.; Lai, M.; Rezende, D. J.; Kavukcuoglu, K. Interaction Networks for Learning about Objects, Relations and Physics. *Proceedings of the 30th International Conference on Neural Information Processing Systems*, Red Hook, NY, USA, 2016; p 4509−4517.

(52) Defferrard, M.; Bresson, X.; Vandergheynst, P. In *Advances in Neural Information Processing Systems 29*; Lee, D. D., Sugiyama, M., Luxburg, U. V., Guyon, I., Garnett, R., Eds.; Curran Associates, Inc., 2016; pp 3844−3852.

(53) Kipf, T. N.; Welling, M. Semi-supervised Classification with Graph Convolutional Networks. *arXiv (Machine Learning (cs.LG))*, 2016, arXiv:1609.02907. https://arxiv.org/abs/1609.02907.

(54) Hamilton, W.; Ying, Z.; Leskovec, J. In *Advances in Neural Information Processing Systems 30*; Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc., 2017; pp 1024−1034.

(55) Bianchi, F. M.; Grattarola, D.; Alippi, C.; Livi, L. Graph Neural Networks with Convolutional ARMA Filters. *arXiv (Machine Learning (cs.LG))*, 2019, arXiv:1901.01343. https://arxiv.org/abs/1901.01343.

(56) Simonovsky, M.; Komodakis, N. Dynamic Edge-Conditioned Filters in Convolutional Neural Networks on Graphs. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* **2017**, 3693−3702.

(57) Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph Attention Networks. *arXiv preprint (Machine Learning (stat.ML))*, 2017, arXiv:1710.10903. https://arxiv.org/abs/1710.10903.

(58) Xu, K.; Hu, W.; Leskovec, J.; Jegelka, S. How Powerful are Graph Neural Networks? *arXiv (Machine Learning (cs.LG))*, 2018, arXiv:1810.00826. https://arxiv.org/abs/1810.00826.

(59) Kearnes, S.; McCloskey, K.; Berndl, M.; Pande, V.; Riley, P. MolecularGgraph Convolutions: Moving Beyond Fingerprints. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 595.

(60) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **2018**, *120*, 145301.

(61) Gilmer, J.; Schoenholz, S. S.; Riley, P. F.; Vinyals, O.; Dahl, G. E. Neural Message Passing for Quantum Chemistry. *Proceedings of the 34th International Conference on Machine Learning* **2017**, 1263−1272.

(62) Duvenaud, D. K.; Maclaurin, D.; Iparraguirre, J.; Bombarell, R.; Hirzel, T.; Aspuru-Guzik, A.; Adams, R. P. In *Advances in Neural Information Processing Systems 28*; Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M., Garnett, R., Eds.; Curran Associates, Inc., 2015; pp 2224−2232.

(63) Schüt, K. T.; Arbabzadah, F.; Chmiela, S.; Müler, K. R.; Tkatchenko, A. Quantum-chemical Insights from Deep Tensor Neural Networks. *Nat. Commun.* **2017**, *8*, 13890.

(64) Liu, R.; Lehman, J.; Molino, P.; Such, F. P.; Frank, E.; Sergeev, A.; Yosinski, J. An Intriguing Failing of Convolutional Neural Networks and the Coordconv Solution. *Advances in Neural Information Processing Systems* **2018**, 9605−9616.

(65) *Calculations Manual—Wiki*, 2020. https://wiki.materialsproject.org/Calculations_Manual#Band_Structure_and_DOS_computations_details [accessed 2020-04-01].

(66) Hou, X.; Shen, L.; Sun, K.; Qiu, G. Deep Feature Consistent Variational Autoencoder. *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)* **2017**, 1133−1141.

(67) Sohn, K.; Lee, H.; Yan, X. Learning Structured Output Representation Using Deep Conditional Generative Models. *Advances in Neural Information Processing Systems* **2015**, 3483−3491.

(68) Kullback, S. Letter to the Editor: The Kullback-Leibler Distance. *American Statistician* **1987**, *1*, 1.

(69) Higgins, I.; Matthey, L.; Pal, A.; Burgess, C.; Glorot, X.; Botvinick, M.; Mohamed, S.; Lerchner, A. Beta-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework. *International Conference on Learning Representations (ICLR '17)* **2017**, *2*, 6.

(70) Gupta, A.; Agarwal, A.; Singh, P.; Rai, P. A Deep Generative Framework for Paraphrase Generation. *Thirty-Second AAAI Conference on Artificial Intelligence* **2018**, 5149−5156.

(71) VanRullen, R.; Reddy, L. Reconstructing Faces from FMRI Patterns Using Deep Generative Neural Networks. *Commun. Biol.* **2019**, *2*, 193.

(72) Brunner, G.; Konrad, A.; Wang, Y.; Wattenhofer, R. MIDI-VAE: Modeling Dynamics and Instrumentation of Music with Applications to Style Transfer. *arXiv (Sound (cs.SD))*, 2018, arXiv:1809.07600. https://arxiv.org/abs/1809.07600.

(73) Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention* **2015**, 234−241.

(74) Abdolhoseini, M.; Kluge, M. G.; Walker, F. R.; Johnson, S. J. Segmentation of Heavily Clustered Nuclei from Histopathological Images. *Sci. Rep.* **2019**, *9*, 4551.

(75) Petousis, I.; Chen, W.; Hautier, G.; Graf, T.; Schladt, T. D.; Persson, K. A.; Prinz, F. B. Benchmarking Density Functional Perturbation Theory to Enable High-throughput Screening of Materials for Dielectric Constant and Refractive Index. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2016**, *93*, 115151.

(76) de Jong, M.; Chen, W.; Geerlings, H.; Asta, M.; Persson, K. A. *A Database to Enable Discovery and Design of Piezoelectric Materials*, 2015.

(77) Chollet, F. *Keras*, 2015. https://keras.io [accessed 2020-05-01].

(78) Abadi, M. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, 2015. https://www.tensorflow.org/ [accessed 2020-05-01].

(79) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; et al. QUANTUM ESPRESSO: A Modular and Open-source Software Project for Quantum Simulations of Materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.

(80) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(81) Grimme, S. Semiempirical GGA-type Density Functional Constructed with a Long-range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.