ARTICLE　　OPEN

Check for updates

Constrained crystals deep convolutional generative adversarial network for the inverse design of crystal structures

Teng Long [ID][1✉], Nuno M. Fortunato[1], Ingo Opahle[1], Yixuan Zhang[1], Ilias Samathrakis[1], Chen Shen[1], Oliver Gutfleisch[1] and Hongbin Zhang [ID][1✉]

Autonomous materials discovery with desired properties is one of the ultimate goals for materials science, and the current studies have been focusing mostly on high-throughput screening based on density functional theory calculations and forward modeling of physical properties using machine learning. Applying the deep learning techniques, we have developed a generative model, which can predict distinct stable crystal structures by optimizing the formation energy in the latent space. It is demonstrated that the optimization of physical properties can be integrated into the generative model as on-top screening or backward propagator, both with their own advantages. Applying the generative models on the binary Bi-Se system reveals that distinct crystal structures can be obtained covering the whole composition range, and the phases on the convex hull can be reproduced after the generated structures are fully relaxed to the equilibrium. The method can be extended to multicomponent systems for multi-objective optimization, which paves the way to achieve the inverse design of materials with optimal properties.

INTRODUCTION

The discovery and exploitation of materials have enormous benefits for the welfare of society and technological revolutions[1], which motivates the launch of the Materials Genome Initiative in 2011[2,3]. Till now, high-throughput (HTP) workflows based on density functional theory (DFT) enable massive calculations on existing and hypothetical compounds, accelerating materials discovery dramatically[4]. For instance, crystal structure predictions can be performed based on brute force substitution of the known prototypes or the evolutionary algorithms as implemented in CALYPSO[5] and USPEX[6]. Nevertheless, the soaring computation cost prevents exhaustive screening over immense phase space, limiting the application of such methods and calling for more efficient solutions. Global optimization methods added a target function to guide the searching and therefore reduce the computational cost[7], e.g., crystal structure prediction with Bayesian optimization can efficiently select the most stable structure from a large number of candidate structures with a lower number of searching trials and has been successfully applied to predict structures of NaCl and Y_2Co_{17}[8]. The random search method has also demonstrated its potential to find the global minimum with highly parallel implementation, as manifested by the applications on elemental boron, silicon cluster, and lithium hydrides[9,10]. Last but not least, the evolutionary algorithm is promising in predicting thermodynamically stable crystal structures, such as demonstrated for tungsten borides, carbon, and TiO_2[11,12].

The emergent machine learning techniques have become the fourth paradigm for materials science[13], as exemplified by the rapid developments and applications in the last decade[14]. Particularly, the generative machine learning methods have been utilized to discover distinct materials, e.g., applying neural networks to assist molecule design[15,16]. Two particularly powerful methods therein are the variational autoencoder (VAE)[17] and generative adversarial network (GAN)[18], resulting in significant progresses on molecule design[19]. Combined with the successful

forward modeling of physical properties[20], generative machine learning on the existed structures enables inverse design, i.e., prediction of distinct structures with the desired properties[1,21]. Unfortunately, unlike molecules that can be represented using simplified molecular input line entry specification (SMILES)[22], the inverse design of three-dimensional (3D) crystal structures has been rare due to the challenges in obtaining a continuous representation in the latent space[20].

Recently, Nouira et al. developed a CrystalGAN model to design ternary A-B-H phases starting from binary A-H and B-H structures using a vector-based representation[23], and successfully applied it on the Ni-Pd-H system. Another promising representation is the two-dimensional (2D) crystal graphs constructed based on the voxel and autoencoder methods[24–26]. For instance, Noh et al. developed the iMatGen model to represent the image-based crystals continuously, and successfully generated distinct V_xO_y using VAE[27]. The ZeoGAN scheme developed by B. Kim et al. is based on the same representation applied on zeolites, where Wasserstein GAN (WGAN) is used to generate porous materials[28]. S. Kim et al. also developed the Crystal-WGAN to design Mn-Mg-O system using such a continuous representation[29]. These models exhibit excellent abilities in generating distinct structures, but not all predicted phases host the desired functionality. For example, only ZeoGAN integrates the optimization of the heat of adsorption, whereas the other schemes exert mostly constraints to filter the generated structures. It is noted that by nature, inverse design entails the optimization of properties in the latent space, and thus internalizes the desired property as a joint objective of the generator[30].

In this work, we developed a GAN-based inverse design framework for crystal structure prediction with target properties and applied it to the binary Bi-Se system. It is first demonstrated that our deep convolutional generative adversarial network (DCGAN) can be applied to generate distinct crystal structures[31]. Taking formation energy as the target property, its optimization is integrated into the DCGAN model in two schemes: DCGAN +

[1]Institute of Materials Science, Technische Universität Darmstadt, Darmstadt, Germany. ✉email: tenglong@tmm.tu-darmstadt.de; hongbin.zhang@tu-darmstadt.de
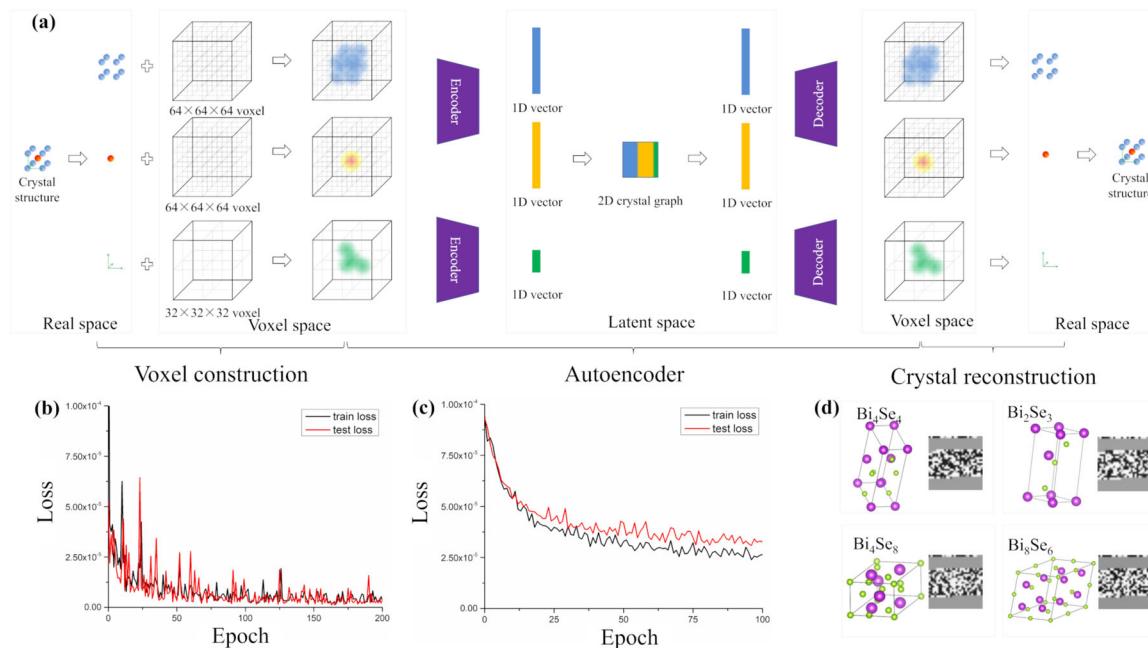
Fig. 1 **Construction of 2D crystal graphs. (a)** Schematic diagram of generating 2D crystal graphs; **(b)** Learning curve of lattice autoencoder, where the red line represents the test loss and the black line represents the training loss; **(c)** Learning curve of sites autoencoder, where the red line represents the test loss and the black line represents the training loss; **(d)** Four typical crystal structures (Bi_4Se_4, Bi_2Se_3, Bi_4Se_8, Bi_8Se_6) and their corresponding 2D crystal graphs.

constraint to select structures following the conventional screening method, and constrained crystals deep convolutional generative adversarial network (CCDCGAN) with an extra feedback loop for automatic optimization. The performance of the DCGAN + constraint and CCDCGAN models are comparatively evaluated, and it is demonstrated CCDCGAN is more efficient in generating stable structures, while the DCGAN + constraint model has a higher generation rate of meta-stable structures. Both schemes can be generalized to perform multi-objective inverse design for multicomponent systems[32], hence our work paves the way to achieve autonomous inverse design of distinct crystal phases with desired properties.

RESULTS

Data for generative adversarial network
Typically, DCGAN requires 10^3 known crystal structures as positive examples during training. However, there are mostly tens of known crystalline phases for a given binary system, e.g., Materials Projects (MP)[33] provides only 17 already known Bi_xSe_y materials, which are not sufficient for deep learning. In order to get enough training data, DFT calculations are performed on the prototype binary structures used in Ref. [27], where are selected based on two criteria: (1) there are no >20 atoms in the unit cell and (2) the maximum lattice constant is smaller than 10 Angstrom. Starting from 10981 prototype structures, we managed to converge 9810 cases after substituting Bi and Se atoms[34], which are taken as training databases. Hereafter, such a database is referred to as the Bi-Se database. Out of the 15 experimentally achievable phases[35], 8 structures (i.e., $BiSe_3$, Bi_2Se_3, Bi_4Se_5, Bi_3Se_2, Bi_5Se_3, Bi_2Se, Bi_7Se_3, and Bi_3Se) are included in the database, whereas the other 7 compounds (i.e., $BiSe_2$, Bi_3Se_4, Bi_8Se_9, $BiSe$, Bi_8Se_7, Bi_6Se_5, and Bi_4Se_3) are excluded by the selection criteria, as demonstrated in Supplementary Fig. 1. Furthermore, out of the Bi-Se database, 155 (1.56%) structures of the database are stable (i.e., formation energy lower than 0 eV/atom and distance to the convex hull smaller than 100 meV/atom) and 707 (7.2%) cases are meta-stable

(i.e., with negative formation energy and distance to the convex hull smaller than 150 meV/atom).

Continuous representation for crystal structures
Although the crystallographic and chemical information is stored in the crystallographic information file (CIF), the discontinuous and heterogeneous formats are not a suitable choice of representation in a generative model, thus a continuous and homogeneous representation including both the chemical and structural information is required. Following Ref. [27], the lattice constants and atomic positions are translated into the voxel space, followed by encoding into a 2D crystal graph through autoencoder, as demonstrated in Fig. 1(a). In this way, a continuous latent space is constructed. The whole process is reversible, i.e., a random 2D crystal graph can be reconstructed into a crystal structure in real space, which is essential for a generative model. When applied to the Bi-Se database, 9420 of 9810 (i.e., 96%) crystal structures are successfully reconstructed. Such a high ratio is not caused by overfitting, illustrated by the learning curve of both autoencoders as shown in Fig. 1(b, c). The differences of training set loss and test set loss are both negligible, suggesting that the 2D crystal graphs in the latent space contain adequate information to reconstruct crystal structures. This is confirmed by the diversity of twelve typical 2D crystal graphs in the latent space (cf. Fig. 1(d) for four cases). Please refer to "Methods" as well as Supplementary Information for more details.

Construction and prediction of DCGAN
Turning now to the realization of generative models, VAE[17] and GAN[18] are two most popular algorithms. VAE is a mutation of the autoencoder discussed above, which assumes a specific (such as Gaussian) distribution of data (in our case 2D crystal graphs) in the latent space. To be generative, such a distribution function should be defined properly, i.e., to be consistent with the distribution of 2D crystal graphs of the known crystal structures in the latent space. In addition, the distribution function has to be specified prior to the training process and its form determines the
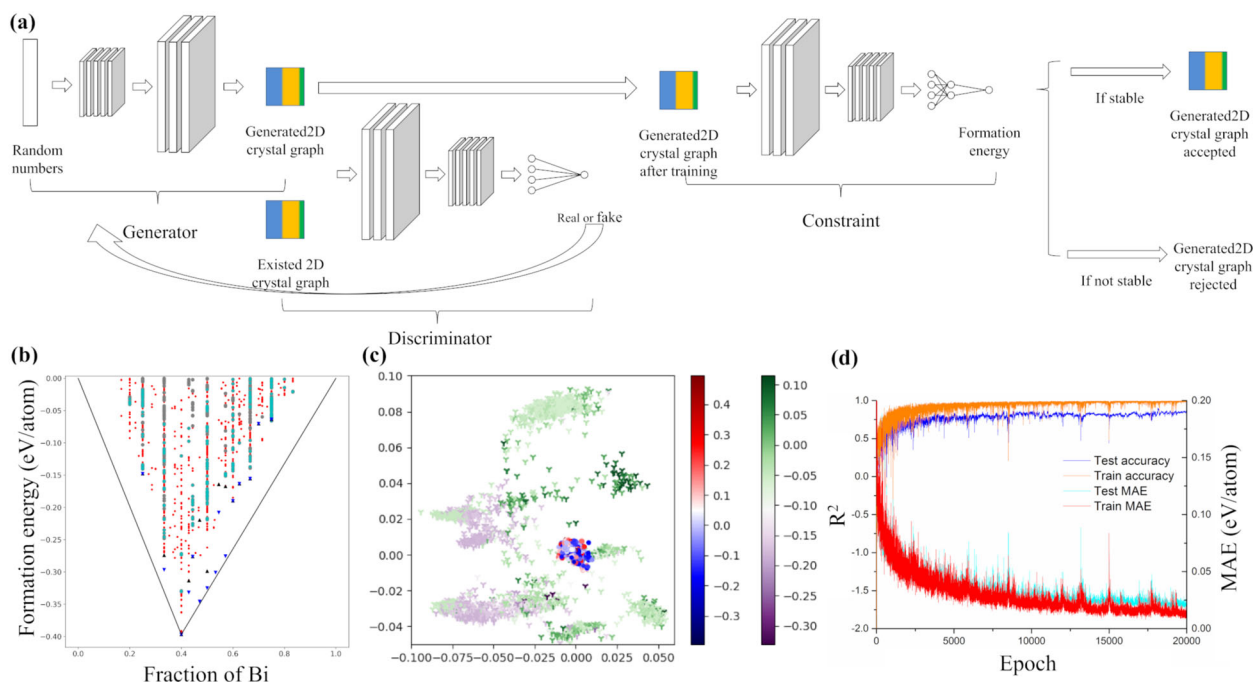
Fig. 2 Illustration on the DCGAN + constraint model. (a) Schematic diagram of the DCGAN + constraint model; (b) Scatter plot of formation energy vs. composition for structures generated by DCGAN and DCGAN + constraint models, where the black line represents the convex hull, red points denote the crystal structures considered in our machine learning database, gray circles indicate the generated structures, cyan stars denote the DCGAN + constraint structures, blue triangles indicate the experimentally achievable phases (cf. Supplementary Table 7), black triangles mark the crystal structures of the available Bi-Se phases in the Bi-Se database; (c) Distribution of the crystal structures in the latent space, the circles denote the structures in the Bi-Se database, while the stars denote the generated structures. The color maps denote the corresponding formation energy; (d) Learning curves of the training process of the DCGAN model, red and orange represent the MAE and accuracy of the training set, while blue and cyan represent the accuracy and MAE of the test set.

performance of the generative model, which demands domain expertise in statistics and profound understanding of the input data. For instance, a recent work by Ren et al. used sigmoid function in VAE to design inorganic crystals and find structures that do not exist in training set[36]. Court et al. implemented VAE to generate crystal structures and successfully applied it to a broad class of materials including binary and ternary alloys, perovskite oxides, and Heusler compounds[37]. In contrast, GAN trains two mutual-competitive neural networks (i.e., generator and discriminator) to generate data statistically the same as the training data without assuming the distribution function[18]. That is, the discriminator tries to distinguish the generated data from the training data, while the generator attempts to fool the discriminator by generating data similar to the training data. Mathematically the objective of GAN is defined by the following equation:

$$\max_{D}\left(\min_{G}\left(\frac{1}{2}\cdot E_{\mathbf{x}\sim p_x}[1-D(\mathbf{x})]+\frac{1}{2}\cdot E_{\mathbf{x}\sim p_g}[D(\mathbf{x})]\right)\right) \quad (1)$$

where D is the discriminator, G is the generator, E means expectation value, $\mathbf{x}$ is the original data, $D(\mathbf{x})$ is the output of discriminator with $\mathbf{x}$ as input, p_x is the possibility density function of the original data, while p_g is the possibility density function of the generated data. Through the competition between the generator and discriminator, the distribution of generated data becomes hardly distinguishable with the distribution of training data. Compared with VAE, GAN does not require specifying a distribution function in the latent space (p_g), which makes the generation process more robust, and thus is adopted in this work[19].

Such an implementation of DCGAN can generate crystal structures with a high success rate (defined as the ratio of the number of generated crystals over the number of generated 2D crystal graphs), e.g., 2832 crystal structures are transformed from 13,000 generated 2D crystal graphs. The generated structures cover a large composition range as shown in Fig. 2(b), where the red points denote the original data in the Bi-Se database and the gray circles mark the generated structures by DCGAN. One interesting observation is that the generated structures are mostly with negative formation energies, on average lower than that of the original database (Supplementary Fig. 2(a)). For instance, 2548 structures out of 2832 have negative formation energies after follow-up DFT relaxations, which consists of 89.9% of the generated structures, whereas only 46.8% (4588 out of 9810) of the Bi-Se database have negative formation energies. Out of 2832 generated structures, it is found that 1233 of them are meta-stable, and 58 are stable.

Detailed analysis reveals that 476 out of the 2832 structures are considered to be distinct structures, i.e., different from those in the Bi-Se database and different from each other, leading to 73 (15) meta-stable (stable) structures, respectively. For instance, Bi_2Se_4 and $BiSe_3$ shown in Supplementary Fig. 2(b, c) are the automatic generated structures, whose formation energies (distances to the convex hull) are −0.192 (0.136) eV/atom and −0.135 (0.111) eV/atom, respectively. Furthermore, DCGAN explores a much larger phase space when generating structures. As illustrated in Fig. 2(c), the structures in the Bi-Se database are concentrated in a small region in the latent space, whereas much larger phase space is covered by DCGAN. Thus, DCGAN can generate distinct crystal structures beyond the phase space of the known crystal structures.

Training of constraint model

As mentioned above, the objective of inverse design is to design compounds with desired properties, including thermodynamical, mechanical, and functional properties. In this work, we take the formation energy (i.e., thermodynamic stability) as the target

property. In order to be able to optimize the formation energy in the latent space, another convolutional neural network (CNN) model is trained taking 2D crystal graphs in the latent space as descriptors and formation energies as the output physical property. The training is carried out by a 90%/10% training / test-set ratio over the Bi-Se database, with the resulting R^2 score and mean absolute error (MAE) being about 85% and 0.019 eV/ atom (Supplementary Fig. 3(a)), respectively. Such a MAE is even smaller than that (0.021 eV/atom) obtained using the state-of-the-art CGCNN model[38]. That is, the 2D crystal graphs in the latent space can be considered as effective descriptors to perform forward prediction of physical properties. We note that overfitting is a marginal issue for such high accuracy, as seen in the learning curve of the training process in Fig. 2(d). Details on the CNN model are described in "Methods" and Supplementary Information.

Construction and prediction of DCGAN + constraint

A straightforward way to implement the optimization of physical properties, e.g., formation energy in this work, is to apply an add-on constraint on the generated structures using DCGAN, as sketched in Fig. 2(a). The advantage of doing so is that there is no need to train another model, which saves training time. In addition, all the existing machine learning models can be transplanted no matter whether such forward predictions can take 2D crystal graphs in the latent space or vector-based chemical and structural information as descriptors. This allows the optimization of a wide spectrum of physical properties[20]. Nevertheless, this method is essentially a selection of the structures generated by DCGAN, thus it cannot search automatically for a specific region in the latent space to reach the local optimal values. More details are described in Supplementary Information.

The DCGAN + constraint model demonstrates its advantage in generating metastable structures. Applying the constraint on the formation energy (with a tolerance of 0.3 eV/atom) on 2832 generated structures by DCGAN, 2148 crystal structures are selected and optimized by further DFT calculations. Such a screening procedure does not affect the diversity of the composition as shown in Supplementary Fig. 2(d). The application of such a constraint has evident effect in the latent space, as demonstrated in Supplementary Movie 1 where the constraint screens the unwanted points out leading to a shrunk region in the latent space. With such a constraint applied, the ratio of structures with negative formation energy is higher, increasing from 89.9% for DCGAN to 94.8% (2037/2148). The ratio of meta-stable structures also becomes higher compared to the DCGAN case, i.e., 1145 meta-stable structures are generated from 2037 structures, reaching the ratio of 56.2% which is better than 43.5% in the DCGAN model. However, the number of generated stable structures reduces to 36 due to the selection process.

An interesting observation is that applying the constraint jeopardizes the generation of distinct structures. After performing DFT calculations, only 247 distinct structures remain, which are significantly reduced compared with the previous 476 cases obtained by DCGAN. The reason is two-fold. On the one hand, the generated crystal structures by DCGAN are not guaranteed to be at the mechanical and dynamical equilibrium, i.e., the lattice constants and atomic positions will change during DFT relaxation. On the other hand, the CNN model applied as constraint is not good enough in predicting formation energy, though the MAE for the formation energy is only about 0.019 eV/atom. Thus, 629 candidates have been screened out by considering a tolerance of 0.3 eV/atom. After comparing with the known crystal structures in the Bi-Se database, 67 distinct meta-stable structures and 12 distinct stable structures are obtained, e.g., two of them Bi_3Se and Bi_2Se_2 are demonstrated in Supplementary Fig. 2(e, f), with the formation energies (distances to the convex hull) being −0.065 (0.099) eV/atom and −0.202 (0.126) eV/atom, respectively. Overall,

generated structures via DCGAN + constraint have a higher ratio of meta-stable structures.

Construction and prediction of CCDCGAN

The constraint can also be integrated into DCGAN as a back propagator, as illustrated as CCDCGAN in Fig. 3(a), to realize automated optimization in the latent space so that inverse design can be realized. In this way, the constraint gives rise to an additional optimization objective of the generator, which can be mathematically described as:

$$\min_{G} \left(\frac{1}{2} \cdot E_{\mathbf{z} \sim p_x}[D(\mathbf{z})] + \omega \cdot E_f(\mathbf{z}) \right), \mathbf{z} \sim p_g, \quad (2)$$

where E_f is the formation energy predicted by the constraint model, $\mathbf{z}$ is the generated 2D crystal graph, and ω is defined as the weight of formation energy loss. Note that such an additional optimization objective cannot outweigh the primary objective, leading to lower weight for the formation energy loss (0.1 in this work) than the discriminator loss. Unlike the DCGAN + constraint model, CCDCGAN can accomplish automated searching for the local minima in the latent space and thus improve the efficiency of discovering distinct stable structures. It is noted that CCDCGAN requires training from scratch, which takes 18 h on a Quadro P2000 GPU machine, which is 2 h longer than the time needed for training the DCGAN + constraint model. Please refer to "Methods" and Supplementary Information for more details.

In comparison with the other two models (i.e., DCGAN and DCGAN + constraint models), CCDCGAN has a higher success rate of generating crystal structures, i.e., 3743 crystal structures are transformed successfully from 13,000 generated 2D crystal graphs. Correspondingly, the number of structures that are with negative formation energy and meta-stable is 3556 and 1910, respectively. Furthermore, CCDCGAN has a higher generation efficiency of stable structure (307 out of 3743) when compared with the other two models, which suggests that the back propagation indeed causes optimization in the latent space. This is also proved by the average formation energy of generated structures, where the average formation energy (−0.11 eV/atom) for the CCDCGAN generated structures is lower than that from the DCGAN (−0.08 eV/atom) and DCGAN + constraint (−0.09 eV/atom).

The most intriguing result is that the number of distinct structures is bigger than that from DCGAN or DCGAN + constraint, where 511 distinct structures have been identified after performing DFT relaxations on 3743 structures. Correspondingly, the number of distinct meta-stable and stable structures is 187 and 42, respectively. Such numbers are similar to those obtained using iMatGen and Crystal-WGAN[27,29], indicating comparable performance of our DCGAN-based models with the existing ones. Three particularly promising structures with formation energy lower than their counterparts in Bi-Se database are Bi_2Se_4, Bi_4Se_2, and Bi_6Se_2, whose formation energies (distance to the convex hull) are −0.287 (0.041) eV/atom, −0.139 (0.069) eV/atom, and −0.063 (0.101) eV/atom, respectively. In this regard, CCDCGAN performs better than both the DCGAN + constraint and DCGAN models in generating distinct structures.

To explore the full competence of the CCDCGAN model, we generated 100,000 crystal structures and performed DFT calculation on the generated structures. The CCDCGAN model is able to regenerate most (11 out of 15) experimentally achievable phases. As seen from Fig. 3(c) and Supplementary Table 7, all the 8 experimentally achievable phases in the Bi-Se database can be regenerated using our CCDCGAN model, i.e., exactly the same crystal structures as the experimentally achievable phases of $BiSe_3$, Bi_2Se_3, Bi_4Se_5, Bi_3Se_2, Bi_5Se_3, Bi_2Se, Bi_7Se_3, and Bi_3Se can be obtained. Furthermore, our CCDCGAN model can also generate three out of seven non-included experimentally achievable phases, namely, $BiSe_2$, Bi_3Se_4, and Bi_6Se_5, with exactly the same crystal structures as
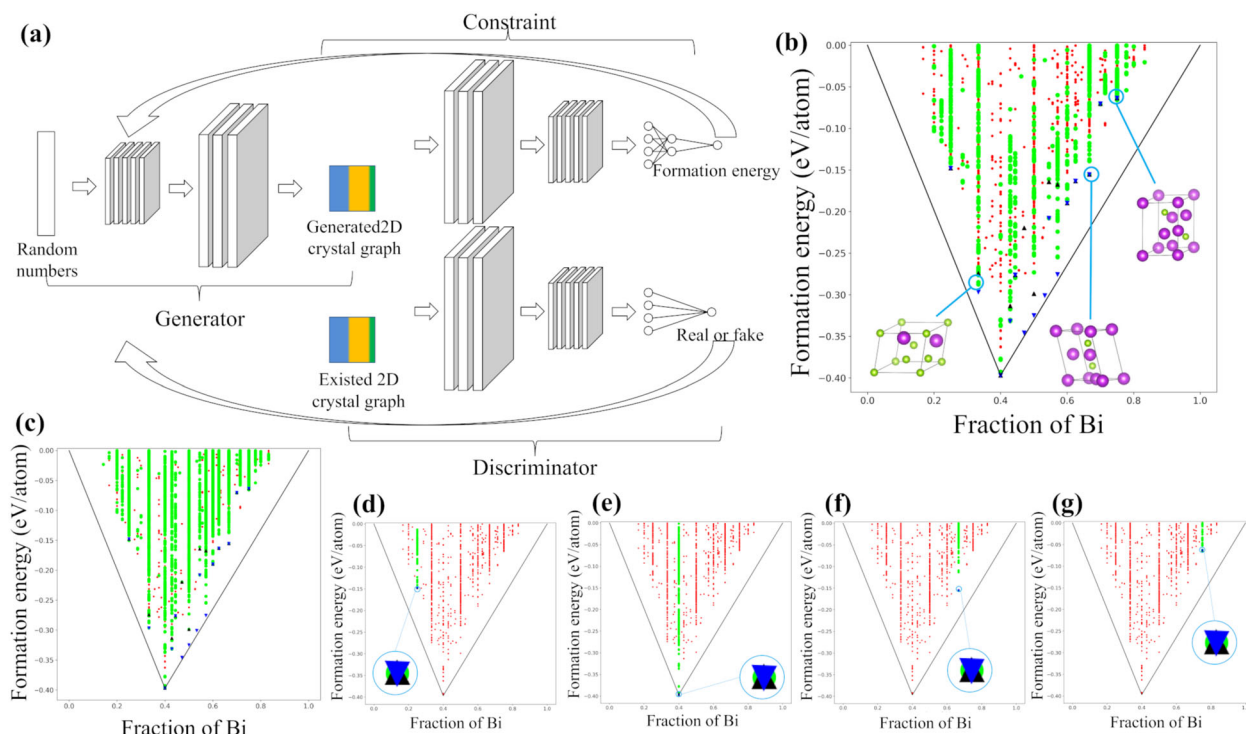
Fig. 3 Illustration on the CCDCGAN model. (a) Schematic diagram of the CCDCGAN model; (b) Scatter plot of formation energy vs. composition of CCDCGAN generated structures, where the black line represents the convex hull, red points denote the crystal structures considered in our machine learning database, green circles represent the generated structures, blue triangles indicate the experimentally achievable phases, black triangles are the crystal structures of the available Bi-Se phases in the Bi-Se database; and crystal structure of generated Bi_2Se_4, Bi_4Se_2, and Bi_6Se_2; (c) Scatter plot of formation energy vs. composition of 100,000 structures generated by CCDCGAN; (d) The $BiSe_3$ phases regenerated by CCDCGAN, which are trained with the $BiSe_3$ composition removed; (e) The Bi_2Se_3 phases regenerated by CCDCGAN, which are trained with the Bi_2Se_3 composition removed; (f) The Bi_2Se phases regenerated by CCDCGAN, which are trained with the Bi_2Se composition removed; (g) The Bi_3Se phases regenerated by CCDCGAN, which are trained with the Bi_3Se composition removed.

reported. Unfortunately, the other four phases, namely, Bi_8Se_9, $BiSe$, Bi_8Se_7, and Bi_4Se_3, out of seven non-included phases cannot be reproduced, where the generated structures of $BiSe$ and Bi_4Se_3 have comparable formation energy as the experimentally achievable structures, while the formation energies of generated Bi_8Se_9 and Bi_8Se_7 structures are higher than those reported phases. It is suspected that this can be attributed to the constraints (i.e., unit cell dimension and number of atoms) we applied to select the training database, which will be extended in the future.

To further test the predictive capability of the CCDCGAN model, we deliberately removed the crystal structures of 4 specific compositions (i.e., $BiSe_3$, Bi_2Se_3, Bi_2Se, and Bi_3Se) from the training database, and observed that the experimental crystal structures of such four phases can still be successfully regenerated, as shown in Fig. 3(d, g). This proves that the CCDCGAN model can generate crystal structures for unknown compositions, and thus it could accelerate the discovery of distinct crystal phases.

DISCUSSION

We have developed an inverse design framework CCDCGAN consisting of a generator, discriminator, and constraint and successfully applied it to design unreported crystal structures with low formation energy for the binary Bi-Se system. It is demonstrated that 2D crystal graphs can be used to construct a latent space with the continuous representation of the known crystal structures, which serve as effective descriptors for modeling the physical properties (e.g., formation energy) and can be decoded into real space crystal structures enabling the generation of distinct crystal structures. Such an inverse design model can be easily generalized to the multicomponent cases, as

demonstrated by Ref. [29]. Importantly, we elucidate that the optimization of physical properties (e.g., formation energy) can be integrated into the generative deep learning model as explicit constraints or back propagators. This allows further development of a multi-objective inverse design framework to optimize other physical properties by modifying the objective function, as demonstrated by Ref. [28]. One apparent challenge is how to get the generated structures into their mechanical and dynamical equilibria, entailing further exploration such as relaxation using the machine learning interatomic potential[39].

METHODS

Data for generative adversarial network

The Bi-Se database is based on the substitution of all binaries of MP, in the spirit of high-throughput calculation, previous research has selected 10981 structures of them by eliminating the large unit cell, i.e. only select structures with the maximum number of atoms in unit cell < 20 and the maximum length of unit cell smaller than 10 Angstrom, and we do the substitution based on their result[27,29]. Following this criteria, several stable phases like $BiSe$, Bi_4Se_3, Bi_8Se_7 close to the convex hull are screened out. We use the high-throughput environment to optimize the structure and determine the thermodynamical stability of Bi-Se database and the generated structures[40], where thermodynamical stability is evaluated by calculation of the formation energy. All data in the Bi-Se database are considered to calculate the convex hull. For VASP setting[41], Perdew–Burke–Ernzerhof (PBE) is adopted, the energy cutoff is 500 eV, and K space density is 40 in each direction for the Brillouin zone. The Bi-Se database is in Supplementary Data 1.

Continuous representation for crystal structures

Specifically, 3D voxel grids are used for a typical binary compound: two grids to record the atomic positions of two elements separately and the

third one to store the lattice constants, i.e., the lengths and angles of/between them. Hereafter such grids are referred to as site voxel and lattice voxel, respectively. Using the probability density based on Gaussian functions, the lattice voxel is defined as

$$p_{x,y,z} = e^{-\frac{r_{x,y,z}^2}{2\sigma^2}}, \tag{3}$$

where x, y, z denotes the grid index, p is the probability density of the denoted grid, $\mathbf{r}$ is the real space distance between this grid to the center of the unit cell, and σ denotes the standard deviation of the Gaussian function. Similarly, sites voxel is defined as

$$p_{x,y,z} = \sum_{i=1}^{n} e^{-\frac{r_{n,x,y,z}^2}{2\sigma^2}}, \tag{4}$$

where n is number of atoms and r is the real space distance between this grid point and the atom. In this way, the inverse transformation is trivial for the lattice voxel while that for the sites voxel relies on the image filter technique[27].

Autoencoder is a type of artificial neural network, which is able to encode inputs into low dimension vectors, as well as decode the vectors back[42]. In this work, the autoencoder is applied to encode 3D voxel data into one-dimensional (1D) vectors in the latent space, which is realized based on 3D convolutional neural network (CNN) autoencoder. The loss function of autoencoder consists of two parts: the first is the loss of information, i.e., the difference between inputs and outputs, and the second is the regulation term to prevent overfitting, which yields

$$\text{loss} = \|\mathbf{x} - \mathbf{y}\|^2 + \frac{\lambda}{2} \cdot \sum \boldsymbol{\omega}^2, \tag{5}$$

where $\mathbf{x}$ and $\mathbf{y}$ are the input and output of the autoencoder, λ is regulation coefficient, and ω is weight in the autoencoder[17]. The construction of autoencoder for both site and lattice voxel are the same, where both are trained using a 90%/10% training/test set ratio. Both site and lattice voxel autoencoders have high reconstruction ratio. All of them are considered to be identical by the crystal structure comparison subroutine available in pymatgen[43], with fractional length tolerance being 0.2, sites tolerance 0.3, and angle tolerance 5 degrees.

For data transformation between real space and voxel space, the voxel used for atomic positions is $64 \times 64 \times 64$, while the voxel for lattice parameters is $32 \times 32 \times 32$ and the σ used in Gaussian function is always 0.15, function "gaussian_filter" from scipy package is used[27]. The autoencoders constructed by tensorflow in python. The sites autoencoder consists of 10 3D convolutional layers, i.e. 5 for encoder and 5 for decoder, strides except the first and last on are $2 \times 2 \times 2$ and activation function is leakyRELU, Adam optimizer is used, learning rate is 0.003, λ is 0.000001, β_1 is 0.9 and β_2 is 0.99[27]. Similar, the lattice autoencoder consists of only 8 3D convolutional layers (4 for encoder and 4 for decoder), other parameters are the same as the previous one[27]. Detailed design of these two models can be found in Supplementary Tables 1 and 2.

Training of DCGAN
The left half of Fig. 2(a) illustrates our implementation of DCGAN to generate crystal structures. First, both the generated 2D crystal graphs by the generator and original 2D crystal graphs are fed into the discriminator, which is trained to distinguish such graphs. Afterward the generator is further trained by the feedback from the discriminator through back propagation to generate structures more similar to the original structures. By repeating the processes, the generator will evolve to be able to eventually generate structures that are indistinguishable (by the discriminator) to the original structures. Details to construct DCGAN can be found in "Methods" and Supplementary Information.

The DCGAN model consists of generator and discriminator, they both use Adam optimizer with 0.002 learning rate, β_1 is 0.5 and β_2 is 0.999. 2D convolutional layer, dropout and batch normalization are used in both model, while RELU, leakyRELU, tanh, and sigmoid are used as activation functions. 1,000,000 steps are trained for this model for 16 h on Quadro P2000 GPU. The design of the generator and discriminator are listed in Supplementary Tables 3 and 4, respectively. The latent space of the GAN model has 200 dimensions.

Training of constraint model
The constraint model consists of 4 convolutional 2D layers and 6 fully connected layers, conducted by keras[44]. It also uses Adam optimizer,

learning rate 0.002, β_1 is 0.5 and β_2 is 0.999, optimization loss is mean square error, activation function is leakyRELU. To prevent overfitting, dropout, and batch normalization are used after each convolutional layer[45]. The detailed design of this model is described in Supplementary Table 5.

Training of CCDCGAN
The generator, discriminator, and constraint have exactly the same design as the previous model, all parameters are same as parameters in DCGAN. The only difference is the optimization objective becomes Eq. (2), where the weight of the formation energy loss is 0.1. The training time is about 18 h under the same condition as DCGAN. Typical generated structures are listed in Supplementary Table 6, more data is available in Supplementary Data 2.

REFERENCES
1. Sanchez-Lengeling, B. & Aspuru-Guzik, A. Inverse molecular design using machine learning: generative models for matter engineering. *Science* **361**, 360–365 (2018).
2. de Pablo, J. J. et al. New frontiers for the materials genome initiative. *npj Comput. Mater.* **5**, 1–23 (2019).
3. de Pablo, J. J. et al. The materials genome initiative, the interplay of experiment, theory and computation. *Curr. Opin. Solid State Mater. Sci.* **18**, 99–117 (2014).
4. 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* **65**, 1501–1509 (2013).
5. Wang, Y., Lv, J., Zhu, L. & Ma, Y. CALYPSO: a method for crystal structure prediction. *Comput. Phys. Commun.* **183**, 2063–2070 (2012).
6. Glass, C. W., Oganov, A. R. & Hansen, N. USPEX—Evolutionary crystal structure prediction. *Comput. Phys. Commun.* **175**, 713–720 (2006).
7. Noh, J., Ho, Gu,G., Kim, S. & Jung, Y. Machine-enabled inverse design of inorganic solid materials: promises and challenges. *Chem. Sci.* **11**, 4871–4881 (2020).
8. Yamashita, T. et al. Crystal structure prediction accelerated by Bayesian optimization. *Phys. Rev. Mater.* **2**, 013803 (2018).
9. Deringer, V. L., Pickard, C. J. & Csányi, G. Data-driven learning of total and local energies in elemental boron. *Phys. Rev. Lett.* **120**, 156001 (2018).
10. Pickard, C. J. & Needs, R. J. Ab initio random structure searching. *J. Condens. Matter Phys.* **23**, 053201 (2011).
11. Li, Q., Zhou, D., Zheng, W., Ma, Y. & Chen, C. Global structural optimization of tungsten borides. *Phys. Rev. Lett.* **110**, 136403 (2013).
12. Lyakhov, A. O. & Oganov, A. R. Evolutionary search for superhard materials: methodology and applications to forms of carbon and TiO₂. *Phys. Rev. B* **84**, 092103 (2011).
13. Tolle, K. M., Tansley, D. S. W. & Hey, A. J. The fourth paradigm: data-intensive scientific discovery [point of view]. *Proc. IEEE* **99**, 1334–1337 (2011).
14. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
15. Pyzer-Knapp, E. O., Li, K. & Aspuru-Guzik, A. Learning from the Harvard Clean Energy Project: the use of neural networks to accelerate materials discovery. *Adv. Funct. Mater.* **25**, 6495–6502 (2015).
16. Ryan, K., Lengyel, J. & Shatruk, M. Crystal structure prediction via deep learning. *J. Am. Chem. Soc.* **140**, 10158–10168 (2018).
17. Kingma, D. P. & Welling, M. Auto-encoding variational bayes. Preprint at https://arxiv.org/abs/1312.6114 (2013).
18. Goodfellow, I. et al. Generative adversarial networks. Preprint at https://arxiv.org/abs/1406.2661 (2014).
19. Schwalbe-Koda, D. & Gómez-Bombarelli, R. Generative models for automatic chemical design. *Mach. Learn. Meets Quantum. Phys.* 445–467 (2020).
20. Schmidt, J., Marques, M. R. G., Botti, S. & Marques, M. A. L. Recent advances and applications of machine learning in solid-state materials science. *npj Comput. Mater.* **5**, 1–36 (2019).

21. Gómez-Bombarelli, R. et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent. Sci.* **4**, 268–276 (2018).

22. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **28**, 31–36 (1988).

23. Nouira, A., Sokolovska, N. & Crivello, J.-C. CrystalGAN: Learning to discover crystallographic structures with generative adversarial networks. Preprint at https://arxiv.org/abs/1810.11203 (2018).

24. Hoffmann, J. et al. Data-driven approach to encoding and decoding 3-D crystal structures. Preprint at https://arxiv.org/abs/1909.00949 (2019).

25. De, S., Bartók, A. P., Csányi, G. & Ceriotti, M. Comparing molecules and solids across structural and alchemical space. *Phys. Chem. Chem. Phys.* **18**, 13754–13769 (2016).

26. Kaufmann, K. et al. Paradigm shift in electron-based crystallography via machine learning. Preprint at https://arxiv.org/abs/1902.03682 (2019).

27. Noh, J. et al. Inverse design of solid-state materials via a continuous representation. *Matter* **1**, 1370–1384 (2019).

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

29. Kim, S., Noh, J., Gu, G. H., Aspuru-Guzik, A. & Jung, Y. Generative adversarial networks for crystal structure prediction. *ACS Cent. Sci.* **6**, 1412–1420 (2020).

30. Bojanowski, P., Joulin, A., Lopez-Paz, D. & Szlam, A. Optimizing the latent space of generative networks. Preprint at https://arxiv.org/abs/1707.05776 (2019).

31. Zhang, H. et al. Topological insulators in Bi_2Se_3 Bi_2Te_3 and Sb_2Te_3 with a single Dirac cone on the surface. *Nat. Phys.* **5**, 438–442 (2009).

32. Heim, E. Constrained generative adversarial networks for interactive image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10753–10761 (2019).

33. Jain, A. et al. Commentary: the materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).

34. Singh, H. K. et al. High-throughput screening of magnetic antiperovskites. *Chem. Mater.* **30**, 6983–6991 (2018).

35. Binary Alloy Phase Diagram—an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/engineering/binary-alloy-phase-diagram.

36. Ren, Z. et al. Inverse design of crystals using generalized invertible crystallographic representation. Preprint at https://arxiv.org/abs/2005.07609 (2020).

37. Court, C. J., Yildirim, B., Jain, A. & Cole, J. M. 3-D inorganic crystal structure generation and property prediction via representation learning. *J. Chem. Inf. Model.* **60**, 4518–4535 (2020).

38. Xie, T. & Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).

39. Deringer, V. L., Caro, M. A. & Csányi, G. Machine learning interatomic potentials as emerging tools for materials science. *Adv. Mater.* **31**, 1902765 (2019).

40. Opahle, I., Madsen, G. K. H. & Drautz, R. High throughput density functional investigations of the stability, electronic structure and thermoelectric properties of binary silicides. *Phys. Chem. Chem. Phys.* **14**, 16197–16202 (2012).

41. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).

42. Kramer, M. A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE J.* **37**, 233–243 (1991).

43. Ong, S. P. et al. Python materials genomics (pymatgen): a robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).

44. Chollet, F. et al. Keras. https://github.com/fchollet/keras (2015).

45. Gron, A. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (O'Reilly Media, 2017).

AUTHOR CONTRIBUTIONS

This work originated from the discussion of H.Z., T.L., and Y.Z. H.Z. and O.G. supervised the research. T.L. worked on the machine learning model. T.L., N.F., I.O., and C.S. worked on the DFT calculations. T.L., I.S., and Y.Z. worked on data analysis. All authors contributed in the writing.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-021-00526-4.

Correspondence and requests for materials should be addressed to T.L. or H.Z.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.