

Article



Learning Representations of Inorganic Materials from Generative Adversarial Networks

Tiantian Hu¹, Hui Song¹, Tao Jiang¹ and Shaobo Li^{1,2,*}

- ¹ College of Computer Science and Technology, Guizhou University, Guiyang 550025, China; gs.tthu18@gzu.edu.cn (T.H.); gs.hsong18@gzu.edu.cn (H.S.); gs.tjiang18@gzu.edu.cn (T.J.)
- ² School of Mechanical Engineering, Guizhou University, Guiyang 550025, China
- * Correspondence: lishaobo@gzu.edu.cn

Received: 4 October 2020; Accepted: 9 November 2020; Published: 17 November 2020



Abstract: The two most important aspects of material research using deep learning (DL) or machine learning (ML) are the characteristics of materials data and learning algorithms, where the proper characterization of materials data is essential for generating accurate models. At present, the characterization of materials based on the molecular composition includes some methods based on feature engineering, such as Magpie and One-hot. Although these characterization methods have achieved significant results in materials research, these methods based on feature engineering cannot guarantee the integrity of materials characterization. One possible approach is to learn the materials characterization via neural networks using the chemical knowledge and implicit composition rules shown in large-scale known materials. This article chooses an adversarial method to learn the composition of atoms using the Generative Adversarial Network (GAN), which makes sense for data symmetry. The total loss value of the discriminator on the test set is reduced from 4.1e¹³ to 0.3194, indicating that the designed GAN network can well capture the combination of atoms in real materials. We then use the trained discriminator weights for material characterization and predict bandgap, formation energy, critical temperature (Tc) of superconductors on the Open Quantum Materials Database (OQMD), Materials Project (MP), and SuperCond datasets. Experiments show that when using the same predictive model, our proposed method performs better than One-hot and Magpie. This article provides an effective method for characterizing materials based on molecular composition in addition to Magpie, One-hot, etc. In addition, the generator learned in this study generates hypothetical materials with the same distribution as known materials, and these hypotheses can be used as a source for new material discovery.

Keywords: material characterization; GAN; deep learning; materials discovery

1. Introduction

Artificial intelligence (AI) has made exciting progress, in which the application of machine learning (ML) and deep learning (DL) technology has brought competitive performance in various fields, including image recognition [1–4], Speech recognition [5–7] and natural language understanding [8–10]. Even in the ancient and complex game of Go, AI players have convincingly defeated the human world champion with or without learning from humans [10]. DL builds the mapping from feature space to target attribute. It does not need to consider the complex internal transformation rules, but trains a set of weights to reflect the transformation rules, so DL can approach any nonlinear transformation in theory. It is this advantage of DL, together with the availability of more and more experimental and/or computational material databases (MP [11], OQMD [12], ICSD [13]), that spurs material scientists to adopt advanced data-driven technology to solve material problems. For example, Zhi et al. [14] used AI technology to rationally design the most energy-saving path to obtain the ideal electronic

band gap materials; Chen et al. [15] developed a MEGNet using graph neural networks to accurately predict the properties of molecules and crystals. it is proved that the MEGNet model is superior to previous ML model in 11 of the 13 attributes of the QM9 [16] molecule dataset; Takahashi et al. [17] used unsupervised and supervised machine learning techniques combined with Gaussian mixture models to understand the data structure of material databases and used random forest (RF) to predict crystal structure. The potential physical laws in materials data will be presented in the form of data, which can be described by a large number of ready-made and relevant information. The collected information variables are developed/trained and used to predict macro characteristics.

The two most important aspects of DL models are data representation and learning algorithms. Proper representation of material data is essential to generate accurate models. Currently, molecular-based material characterization methods include Magpie [18], One-hot [19], etc. The Magpie material characterization method refers to the calculation of a group of element statistical properties of the material, such as the periodic number, group number, atomic number, atomic radius, melting temperature, average fraction of valence electrons from the *s*, *p*, *d* and *f* orbits in all elements, etc. Table 1 shows the statistics of 22 attributes of Magpie. Valen et al. [20] used Magpie to characterize each inorganic compound into a 134-dimensional vector, and used a random forest (RF) algorithm to predict the *T*c of superconductors. Zhuo et al. [21] used the Magpie characterization method to build a model that can predict the band gaps of metal compounds using support vector regression (SVR) algorithm on 3896 metal compound data.

Feature Category	re Category Magpie Features	
Statistical characteristics of positions in the periodic table	Number, MendeleevNumber, Column, Row	
Statistical characteristics of physical properties	AtomicWeight, MeltingT, CovalentRadius, Electronegativity	
Valence electron feature	NsValence, NpValence, NdValence, NfValence, Nvalence, NsUnfilled, NpUnfilled, NdUnfilled, NfUnfilled, Nunfilled	
Atomic stacking	Gsvolume_pa, Gsbandgap, Gsmagmom, SpaceGroupNumber	

Table 1. 22 statistical characteristics of Magpie.

One-hot characterization is to characterize materials by the number of elements in the molecular formula. The characterization vector has non-zero values for all elements present in the compound and zero values for other elements. Assume that all the elements in the periodic table form a fixed-length element vector E^n , where n is the number of elements, the elements present in the compound have a non-zero value and the value is expressed by the number of elements in the compound, and zero elsewhere. Table 2 gives examples of One-hot characterization.

Calfa et al. [19] established a reliable prediction model using One-hot material characterization method and a nuclear ridge regression (KRR) algorithm to predict the total energy, density and band gap of 746 binary metal oxide. Mansouri et al. [22] used the One-hot material characterization method and BP neural network algorithm on the OQMD [23] to establish a prediction model of material formation energy, and in hundreds of millions of possible material spaces (AwBxCyDz: A, B, C, D are elements in the periodic table of the elements, $w + x + y + z \le 10$) searched out hundreds of possible new materials. However, whether it is Magpie or One-hot, which is a feature engineering-based material characterization method, the accuracy and completeness of material characterization cannot be ensured. In response to these problems, we propose to learn the characterization of materials through neural networks based on the chemical knowledge and implicit composition rules shown in large-scale known materials. This paper uses Wasserstein GAN(WGAN) [24] to learn the compositional rules of elements from large-scale known materials and then to characterize the materials, which makes sense for data symmetry.

			Attri	butes			
i	Forr	nula	Energy (eV)		Formation_Ene	Formation_Energy_Atom (eV)	
1	Li	2O	-1-	4.264	-2.	-2.071	
÷			•		:		
11	Na	₂ O ₂	-8.401		-1.312		
:		•	:		:		
214	Ti	0 ₂	-26.902		-3.512		
÷			:		÷		
			One-hot cha	racterization			
i	Н	Li	Na		0		
1		2	0		1		
÷	:	÷	÷	:	÷	:	
11	0	0	2		2		
:	:	:	:	:	:	:	
214	0	0	0	•	2	· 	
:	:	:	:	:	:	:	

Table 2. One-hot coding example.

The generative adversarial network (GAN) consists of a generative model (G) that captures the data distribution and a discriminator model (D) that evaluates whether the sample is from real data or a generative model. It was first proposed by Goodfellow et al. [25] for image generation. Compared with other generative models [26,27], the significant difference of GAN is that it does not directly take the difference between the data distribution and the model distribution as the objective function. Instead, an adversarial approach was adopted. First, learn the difference by discriminating model D, and then guide the generation model G to narrow the difference. WGAN [24] is based on the original GAN proposed by Goodfellow et al. [25] to replace the loss function of the JS divergence [28] definition generator with a Wasserstein distance that better reflects the similarity between two distributions. In the end, the generator model learned how to generate hypothetical materials consistent with the atomic composition law of the training material, and the discriminant model learned to distinguish between real materials and fake materials that differed from the real material atomic composition law. In the process of adversarial learning, the weights of the generated model and the discriminant model store the information of element combination rules, so we can use the network weights of the element combination rules stored in the discriminant model to characterize the materials.

In this work, we first trained a WGAN model on the OQMD. Its generator model can generate hypothetical materials consistent with the atomic combination of the training materials, and its discriminative model's loss value was reduced from $4.1e^{13}$ to 0.3. Using the trained discriminator model, we have created a material characterization method. Supervised experiments such as prediction of bandgap, formation energy, and *T*c on the three public material data sets of OQMD [23], ICSD [29], and SuperCon database [20] show that this method performs better than One-hot and Magpie when the same prediction model is used. The main contributions of this article:

- 1. An effective material characterization method is constructed based on WGAN.
- 2. Compared with Magpie and One-hot, the characterization method proposed in this paper achieved the best results on OQMD, ICSD, SuperCon database and other data sets when using the same prediction model.
- 3. A material generator model capable of generating the same atomic combination law as known materials has been trained, which can be used for efficient sampling from a large space of inorganic materials.

2. Methods

2.1. Data Sources

This article uses OQMD to train WGAN deep learning models. OQMD is a widely used DFT [30–32] database with crystal structures either calculated from high-throughput DFT or obtained from the ICSD database. OQMD is in continuous development. At the time of writing, this article already contains 606,115 compounds, including many compounds with the same molecular formula but different formation energy. This article retains the lowest formation energy data because they represent the most stable compounds. Next, excluding single-element materials and removing data entries with formation energy outside $\pm 5\sigma$ (σ is the standard deviation in the data), Jha et al. [33] also adopted a similar data screening method. After the above data screening, data of 291,884 compounds were finally obtained.

2.2. Material Representation Method for WGAN Input

By performing simple statistical calculations on the materials selected by OQMD, 85 elements were found among the 118 elements in the periodic table of the elements. In any particular compound/formula, each element is usually less than 8 atoms. Each material is then represented as a sparse matrix $X^r \in n \times m$ with n = 85 and m = 8. The matrix has 0/1 cell values, each column represents one of the 85 elements, and the column vector is a thermal encoding of the number of atoms for that particular element (see Figure 1).



Figure 1. Schematic representation of KZn2Ru.

2.3. Structure of the WGAN Model

GAN defines a noise as a prior, which is used by the generative model G to learn the probability distribution on the training data X. The G generates fake samples close to the real samples as much as possible, and the D discriminates as accurately as possible whether the input data comes from the real samples X or from the pseudo samples G(z). In order to win the game, these two processes oppose each other and iteratively optimize, so that the performance of D and G is continuously improved. The goal of the optimization is to find the Nash balance [34] between the D and G. Accordingly, in this mini-maximization optimization problem, the optimization objective functions of the discriminant model and the generated model can be defined as:

$$Loss_{D} = \max_{D} V(D,G) = E_{x:P_{data}}[logD(x)] + E_{x:P_{g}}[log(1 - D(G(x)))]$$
(1)

$$Loss_D = \min_{G} V(D,G) = E_{x:P_g}[log(1 - D(G(x)))]$$
(2)

G(x) in the formula means mapping the input noise *z* into data (for example, generating a molecular map). D(x) represents the probability that *x* comes from the real data distribution. The standard form

of GAN proposed by Goodfellow has problems such as training difficulties, loss functions of generators and discriminators that cannot indicate the training process, and lack of diversity in the generated samples. Arjovsky et al. [24] constructed a WGAN based on the original GAN, replacing the loss function of the JS divergence definition generator with a Wasserstein distance that better reflects the similarity between the two distributions, and generates a meaningful gradient to update *G*, so that the generated distribution is pulled towards the real distribution. The loss function of the generated model and discriminant model is redefined as Formulas (3) and (4). Formula (4) can indicate the training process. The smaller the value, the smaller the Wasserstein distance between the real distribution and the generated distribution, the better the GAN training.

$$Loss_G = -E_{x:P_\sigma}[f_w(x)] \tag{3}$$

$$Loss_D = E_{x:P_g}[f_w(x)] - E_{x:P_r}[f_w(x)]$$

$$\tag{4}$$

In the end, the generator model learned how to generate hypothetical materials consistent with the regular combination of elements in the training materials, and the discriminant model learned to distinguish between real materials and fake materials that differed from the regular combinations of atoms in real materials. In the process of adversarial learning, the weights of the generated models and discriminator models store the information of element combination rules, so we can use the discriminator model network weights that store the element combination rules to perform material characterization.

Neither Magpie, One-hot, or other feature-engineering-based methods or supervised training-based neural transfer characterization methods can guarantee the best characterization. However, the basic physical knowledge of some material composition laws obtained through adversarial learning from large-scale known material data can reflect the nature of material composition. We can use this learned knowledge to characterize materials.

Figure 2 is the architecture of the WGAN model designed in this paper. The generation network consists of a fully connected layer and 4 deconvolution layers [35], and each deconvolution layer is followed by a batch normalization [36] layer. The sigmoid activation function is used because the output layer of the generation network needs to generate only sparse matrices of 0 and 1. The hidden layer uses the rectified linear unit (ReLu) [37] as the activation function after the batch normalization layer. The specific parameters of the generation model are shown in Table 3. Given a set of vectors Z^R that follow a normal distribution, where R = 128, the generator G finally generates a feature map $X^f \in n \times m$.



Figure 2. Architecture of the WGAN model used in this paper.

Layer	Input Shape	Filter	Kernel	Stride	Output Shape
Fc1	[batch, 128]	-	-	-	[batch, $5 \times 4 \times 256$]
Reshape	[batch, $5 \times 4 \times 256$]	-	-	-	[batch, 5, 4, 256]
DeConv1	[batch, 5, 4, 256]	128	(5, 5, 256)	(2, 2)	[batch, 11, 8, 128]
DeConv2	[batch, 11, 8, 128]	64	(3, 3, 128)	(2, 1)	[batch, 22, 8, 64]
DeConv3	[batch, 22, 8, 64]	32	(3, 3, 64)	(2, 1)	[batch, 43, 8, 32]
DeConv4	[batch, 43, 8, 32]	1	(3, 3, 32)	(2, 1)	[batch, 85, 8, 1]

Table 3. Parameters of the WGAN generation model.

In discriminator model, we represent 85 kinds of atoms as an identity matrix $E \in \mathbb{R}^{85 \times 85}$ according to the atomic number. After a fully connected embedding layer, we obtain the embedding representation matrix $T \in \mathbb{R}^{85 \times d}$ of these 85 kinds of atoms, where *d* is the number of neurons in the embedding layer. Then, the hypothetical material X^f generated by the generator and the real material characterization matrix X^r are introduced into the discriminant model, and a matrix operation is performed with the embedded matrix *T* of the atom to generate a *d*-dimensional material characterization vector C. The specific process is as Formulas (5) and (6).

$$\mathbf{C} = \mathbf{SBL}(\mathbf{X}^T(EW) \odot \mathbf{A}) \tag{5}$$

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 2 & 2 & \cdots & 2 \\ \cdots & \cdots & \cdots & \cdots \\ 8 & 8 & 8 & 8 \end{bmatrix}, A \in \mathbb{R}^{m \times d}$$
(6)

In the formula, $W \in n \times d$ is the trainable network weight in the discriminator, the atom embedding matrix T = EW, and A is the matrix we introduced to obtain the number of atoms in the molecular formula. \odot represents the matrix point multiplication. SBL(·) means sum the matrix by rows.

The matrix calculation process is shown in Figure 3: The transposition of the material characterization matrix X is multiplied with the atom embedding matrix T to obtain a material characterization map represented by T. This characterization map does not contain information about the number of each element in the material. Next, we introduce a special auxiliary matrix A to help obtain the number characteristics of the atom. The result of the previous step is multiplied with A, and then the results are summed in rows. The material characterization vector represented by T is obtained. Finally, the characterization vector is input into a fully connected layer to obtain the final score. So as long as the atom embedding matrix T is obtained, any molecular formula can be characterized according to Equations (5) and (6).



Figure 3. Graphical representation of Equation (5).

3. Results

Deep learning algorithm has many hyperparameters, and adjusting these parameters can change the performance of the model. The hyperparameters of neural network mainly include momentum, learning rate, optimization algorithm and batch size, among which learning rate and batch size are the most important. In this paper, the WGAN model is trained from 0.1 to $1e^{-5}$ (10 times less each time), and the number of batches is from 32 to 1024 (increasing by 2 times each time). To obtain the best material characterization model, this paper keeps the model with the smallest loss value of discriminator on the test set. In order to ensure the stability and reliability of the results of the supervised prediction experiment, the follow-up comparative experimental results are obtained by averaging 10 validation calculations. The model in this paper is implemented based on python3, and the neural network model is established by using the Tensrflow 18.0 [38] deep learning framework. All programs run on NVIDIA DIGITS GTX1080Ti.

To evaluate the performance of the supervised regression models, we use the mean absolute error (MAE), the root mean square error (RMSE), and R-Squared (R^2) as the evaluation measures. These performance measures can be calculated as follows:

$$MAE = \frac{1}{m} \sum_{i=1}^{m} |y_i - \hat{y}_i|$$
(7)

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2}$$
(8)

$$R^{2} = 1 - \frac{\sum_{i=1}^{m} (y_{i} - \hat{y}_{i})^{2}}{\sum_{i=1}^{m} (y_{i} - \overline{y})^{2}}$$
(9)

where m is the number of samples, y_i and \hat{y}_i are the true and predicted values of the *i* sample label, \overline{y} is the average of the m sample real labels.

In order to train the WGAN model, we randomly divide the OQMD into a training set and a test set (10%) according to 10%. The training set is used to train the GAN model, and the test set is used to verify the quality of the model. We save the model with the smallest total loss value of the discriminator on the test set in 100,000 iterations as the final trained WGAN model. Experiments show that when the learning rate of the training generator is $1e^{-2}$, the learning rate of the discriminator is $1e^{-2}$, the number of batch size is 256, and the RMSProb optimization algorithm is used, the WGAN model converges best. The total loss value of the discriminator on the test set was reduced from $4.1e^{13}$ to 0.3194 (as shown in Figure 4a). As in Equation (4), the smaller the loss value of the discriminator, the better the network training. The discriminator's loss value converges so well, indicating that the WGAN designed in this study can capture the atomic combination of known materials well.



Figure 4. (a) Changes in discriminator loss values during WGAN training. (b) Dimension reduction using t-SNE after characterizing superconductors.

After obtaining the atomic matrix T from the trained model, we characterize the materials on the SuperCond dataset according to the Formula (5), and use the t-SNE [39] algorithm to reduce the dimension of the characterization vector (see Figure 4b). Red dots represent iron-based superconductors, blue dots represent copper-based superconductors, and black dots represent other types of superconductors. We use only the atom embedding matrix T obtained from the discriminator model to characterize superconductors into vectors, and use t-SNEsne to reduce the characterization vectors to two dimensions, and then display them on the coordinate axis. We find that without using any classification algorithm, simply reducing the size, the three types of superconductors can be automatically separated according to the learned representation, which shows the rationality of the characterization method proposed in this paper.

In order to further illustrate that the material characterization method proposed in this paper can better characterize materials than the existing Magpie and One-hot, we conducted supervised experiments on three data sets: MP, Supercond, and OQMD. The bandgap of non-metallic materials is predicted on MP, the critical temperature (*Tc*) of superconductors is predicted on SuperCond database, and the formation energy of predicted materials is predicted on OQMD. To ensure the fairness of the experiments, we use the same fully connected neural network structure for all three representation methods, and the hidden layer neurons are 256, 128, 64, 32, and 1, respectively. The final results are shown in Table 4 and Figure 5a. The three numbers side by side in Table 2 correspond to the MAE, RMSE, and R^2 values.

Table 4. Results of supervised experiments on three datasets.

	WGAN	Magpie	One-Hot
MP	0.467/0.641/0.828	0.570/0.786/0.745	0.484/0.662/0.810
Supercond	6.05/9.21/0.902	6.47/10.3/0.878	6.11/9.63/0.887
ÔQMD	0.101/0.180/0.907	0.126/0.203/0.841	0.108/0.221/0.866

The characterization method proposed in this paper achieved the best results on all three tasks on three data sets. The material characterization method proposed in this paper achieves the best results on the MP dataset, followed by the One-hot method, and the bad one is Magpie. The One-hot method can also produce good results, which may be related to MP data redundancy. The characterization method proposed in this paper achieves similar results using a simple fully connected neural network and a graph neural network-based method proposed by Goodall et al. [40]. Although the amount of parameters is much smaller than the method proposed by Goodall et al., the results are similar. On the Supercond dataset, Valen et al. used the Magpie characterization method and an RF prediction model to obtain a result of R^2 of 0.88. The Magpie characterization method and RF prediction model used by Valen et al. on the Supercond dataset achieved a result of R^2 of 0.88, and the characterization method proposed in this paper achieved R^2 of 0.902. Our proposed characterization method on the OQMD dataset also achieved the best results. Although our method is only a little better than the other two methods in MP prediction of band gap, the prediction of band gap is a very challenging problem, and the improvement of prediction results is very innovative. Our results are similar to [40], even though the scale of our model is much smaller than that of [40] based on graph neural network. The prediction of Tc on Supercond dataset is close to the best result about Tc at present [34]. Even if a simple model is used, the proposed characterization method can achieve good results both on the large-scale oqmd data set and on the small-scale data such as the Supercond dataset. In general, the neural network-based material characterization method has obvious advantages over the characterization methods such as Magpie, One-hot, etc. constructed by feature engineering, and the advantages of the characterization method in this paper are more obvious on small data sets.



Figure 5. (a) Comparison results of R^2 on the three data sets MP, SuperCond, and OQMD. (b) the percentages of charge-neutral (CN) and electronegativity balance (EN) samples generated from OQMD dataset are very close to those of training set.

Charge neutrality and balanced electronegativity are two fundamental chemical rules of materials. Therefore, the rationality of our model can be verified by examining the hypothetical materials generated by WGAN. We use the charge-neutrality and electronegativity checking procedures proposed in reference [41] to calculate the percentage of samples for training materials and generated materials meeting these two rules. The results are shown in Figure 5b. We find that the percentage of effectively generated materials is very close to that of training set. For OQMD dataset When the charge-neutrality of the training set is 55.8%, the generated sample set is 58.1%. This indicates that our WGAN model has successfully learned the chemical rules implied in the material composition. The detailed experimental procedures and results of charge neutrality and balanced electronegativity check have been discussed in detail in another article [42] of ours. Reference [42] is to learn the combination rules of atoms to generate hypothetical materials, while this article is to learn the characterization methods of materials from large-scale known materials. The generators in these two articles adopt the deconvolution method, but the discriminator of [42] is to judge whether the generator is true or not. In this paper, the material characterization method is designed into the discriminator, so that the discriminator can not only judge the authenticity of the material generated by the generator, but also learn the characterization of the material. All the codes in this paper have been uploaded to https://github.com/Chilitiantian/code/tree/master for researchers to verify or follow-up work.

4. Discussion

The generator model trained in this article has learned how to generate hypothetical materials consistent with the element combination rules of the training materials, and the discriminant model has learned to distinguish between real materials and fake materials that are not the same as the atomic combination rules of real materials. In the process of adversarial learning, the weights of the generated models and discriminant models store the information of element combination rules, so we can use the discriminant model network weights that store the element combination rules to perform material characterization. Neither Magpie, One-hot, or other feature-engineering-based methods or supervised training-based neural transfer characterization methods can guarantee the best characterization. However, the basic physical knowledge of some material composition laws obtained through adversarial learning from large-scale known material data can reflect the nature of material composition. We can use this learned knowledge to characterize materials. A series of experiments prove that the characterization methods.

5. Conclusions

In this work, we first trained a WGAN model on the OQMD. Its generator model can generate hypothetical materials consistent with the atomic combination of the training materials. Using the trained discriminator model, we have created a material characterization method. Supervised experiments such as prediction of bandgap, formation energy, and *T*c on the three public material data sets of OQMD [23], ICSD [29], and SuperCon database [20] show that this method performs better than One-hot and Magpie when the same prediction model is used. This article provides an effective method for characterizing materials based on molecular composition in addition to Magpie, One-hot, etc. In addition, the generator learned in this study can generate hypothetical materials with the same distribution as known materials, and these hypotheses can be used as a source of new material discovery.

Our work can be carried out in various forms. We can use the generation model to generate hypothetical materials, and use the material characterization method proposed in this paper to build a supervised model to screen the generated hypothetical materials, to find new materials with certain properties. We can also combine the generated hypothetical materials with the graph neural network, and use the material characterization method proposed in this paper to construct a semi supervised model for predicting material attributes.

Author Contributions: Conceptualization, T.H., and S.L.; methodology, T.H. and S.L.; software, T.H., and H.S.; validation, T.H., and T.J.; investigation, H.S. and T.J.; resources, S.L.; writing—original draft preparation, T.H., and S.L.; writing—review and editing, T.H., and S.L.; supervision, S.L.; project administration S.L.; funding acquisition S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Major science and technology projects of Guizhou Province (Guizhou kehe major project under Grant [2020] 009, Guizhou kehe major project under Grant [2019] 3003) and High level Innovative Talents Project of Guizhou Province (Guizhou kehe talent under Grant [2015]4011).

Acknowledgments: We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan X Pascal GPU used for this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Choo, K.Y.; Hodge, R.A.; Ramachandran, K.K.; Sivakumar, G. Controlling a Video Capture Device Based on Cognitive Personal Action and Image Identification. U.S. Patent 10,178,294, 8 January 2019.
- 2. Berg, M.J.; Robertson, J.C.; Onderdonk, L.A.; Reiser, J.M.; Corby, K.D. Object Dispenser Having a Variable Orifice and Image Identification. U.S. Patent 9,501,887, 22 November 2016.
- 3. Yang, J.; Li, S.; Gao, Z.; Wang, Z.; Liu, W. Real-time recognition method for 0.8 cm darning needles and KR22 bearings based on convolution neural networks and data increase. *Appl. Sci.* **2018**, *8*, 1857. [CrossRef]
- 4. Dusan, S.V.; Lindahl, A.M.; Watson, R.D. Automatic Speech Recognition Triggering System. U.S. Patent 10,313,782, 4 June 2019.
- Malinowski, L.M.; Majcher, P.J.; Stemmer, G.; Rozen, P.; Hofer, J.; Bauer, J.G. System and Method of Automatic Speech Recognition Using Parallel Processing for Weighted Finite State Transducer-Based Speech Decoding. U.S. Patent 10,255,911, 9 April 2019.
- 6. Juneja, A. Hybridized Client-Server Speech Recognition. U.S. Patent 9,674,328, 6 June 2017.
- 7. Clark, K.; Luong, M.-T.; Khandelwal, U.; Manning, C.D.; Le, Q.V. Bam! born-again multi-task networks for natural language understanding. *arXiv* **2019**, arXiv:1907.04829.
- Thomason, J.; Padmakumar, A.; Sinapov, J.; Walker, N.; Jiang, Y.; Yedidsion, H.; Hart, J.; Stone, P.; Mooney, R.J. Improving grounded natural language understanding through human-robot dialog. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; pp. 6934–6941.
- 9. Marcus, J.N. Initializing a Workspace for Building a Natural Language Understanding System. U.S. Patent 10,229,106, 12 March 2019.
- 10. Wang, F.-Y.; Zhang, J.J.; Zheng, X.; Wang, X.; Yuan, Y.; Dai, X.; Zhang, J.; Yang, L. Where does AlphaGo go: From church-turing thesis to AlphaGo thesis and beyond. *IEEE/CAA J. Autom. Sin.* **2016**, *3*, 113–120.

- Jain, A.; Ong, S.P.; Hautier, G.; Chen, W.; Richards, W.D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *Apl Mater.* 2013, *1*, 011002. [CrossRef]
- 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. [CrossRef]
- 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. [CrossRef]
- 14. Shi, Z.; Tsymbalov, E.; Dao, M.; Suresh, S.; Shapeev, A.; Li, J. Deep elastic strain engineering of bandgap through machine learning. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 4117–4122. [CrossRef]
- 15. Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S.P. Graph networks as a universal machine learning framework for molecules and crystals. *Chem. Mater.* **2019**, *31*, 3564–3572. [CrossRef]
- 16. Ramakrishnan, R.; Dral, P.O.; Rupp, M.; Von Lilienfeld, O.A. Quantum chemistry structures and properties of 134 kilo molecules. *Sci. Data* **2014**, *1*, 140022. [CrossRef]
- 17. Takahashi, K.; Takahashi, L. Creating Machine Learning-Driven Material Recipes Based on Crystal Structure. *J. Phys. Chem. Lett.* **2019**, *10*, 283–288. [CrossRef]
- 18. Ward, L.; Agrawal, A.; Choudhary, A.; Wolverton, C. A general-purpose machine learning framework for predicting properties of inorganic materials. *NPJ Comput. Mater.* **2016**, *2*, 16028. [CrossRef]
- 19. Calfa, B.A.; Kitchin, J.R. Property prediction of crystalline solids from composition and crystal structure. *AICHE J.* **2016**, *62*, 2605–2613. [CrossRef]
- 20. Stanev, V.; Oses, C.; Kusne, A.G.; Rodriguez, E.; Paglione, J.; Curtarolo, S.; Takeuchi, I. Machine learning modeling of superconducting critical temperature. *NPJ Comput. Mater.* **2018**, *4*, 1–14. [CrossRef]
- 21. Zhuo, Y.; Mansouri Tehrani, A.; Brgoch, J. Predicting the band gaps of inorganic solids by machine learning. *J. Phys. Chem. Lett.* **2018**, *9*, 1668–1673. [CrossRef] [PubMed]
- Mansouri Tehrani, A.; Oliynyk, A.O.; Parry, M.; Rizvi, Z.; Couper, S.; Lin, F.; Miyagi, L.; Sparks, T.D.; Brgoch, J. Machine learning directed search for ultraincompressible, superhard materials. *J. Am. Chem. Soc.* 2018, 140, 9844–9853. [CrossRef] [PubMed]
- 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, 1–15. [CrossRef]
- 24. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein gan. arXiv 2017, arXiv:1701.07875.
- Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems (NIPS), Montreal, QC, Canada, 8–13 December 2014; pp. 2672–2680.
- 26. Doersch, C. Tutorial on variational autoencoders. arXiv 2016, arXiv:1606.05908.
- Chen, X.; Duan, Y.; Houthooft, R.; Schulman, J.; Sutskever, I.; Abbeel, P. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems (NIPS), Barcelona, Spain, 5–10 December 2016; pp. 2172–2180.
- 28. Fuglede, B.; Topsoe, F. Jensen-Shannon divergence and Hilbert space embedding. In Proceedings of the International Symposium onInformation Theory, 2004, Chicago, IL, USA, 27 June–2 July 2004; p. 31.
- 29. Brown, I.; Altermatt, D. Bond-valence parameters obtained from a systematic analysis of the inorganic crystal structure database. *Acta Crystallogr. Sect. B Struct. Sci.* **1985**, *41*, 244–247. [CrossRef]
- 30. Sharma, A.; Balasubramanian, G. Dislocation dynamics in Al0. 1CoCrFeNi high-entropy alloy under tensile loading. *Intermetallics* **2017**, *91*, 31–34. [CrossRef]
- 31. Sharma, A.; Deshmukh, S.A.; Liaw, P.K.; Balasubramanian, G. Crystallization kinetics in AlxCrCoFeNi $(0 \le x \le 40)$ high-entropy alloys. *Scr. Mater.* **2017**, *141*, 54–57. [CrossRef]
- 32. Sharma, A.; Singh, P.; Johnson, D.D.; Liaw, P.K.; Balasubramanian, G. Atomistic clustering-ordering and high-strain deformation of an Al 0.1 CrCoFeNi high-entropy alloy. *Sci. Rep.* **2016**, *6*, 31028. [CrossRef] [PubMed]
- 33. Jha, D.; Ward, L.; Paul, A.; Liao, W.-K.; Choudhary, A.; Wolverton, C.; Agrawal, A. Elemnet: Deep learning the chemistry of materials from only elemental composition. *Sci. Rep.* **2018**, *8*, 1–13. [CrossRef] [PubMed]

- Rolla, S.; Alchera, E.; Imarisio, C.; Bardina, V.; Valente, G.; Cappello, P.; Mombello, C.; Follenzi, A.; Novelli, F.; Carini, R. The balance between IL-17 and IL-22 produced by liver-infiltrating T-helper cells critically controls NASH development in mice. *Clin. Sci.* 2016, 130, 193–203. [CrossRef] [PubMed]
- 35. Kundur, D.; Hatzinakos, D. Blind image deconvolution. IEEE Signal Process. Mag. 1996, 13, 43-64. [CrossRef]
- 36. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv* **2015**, arXiv:1502.03167.
- Zhang, Y.-D.; Pan, C.; Chen, X.; Wang, F. Abnormal breast identification by nine-layer convolutional neural network with parametric rectified linear unit and rank-based stochastic pooling. *J. Comput. Sci.* 2018, 27, 57–68. [CrossRef]
- Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv* 2016, arXiv:1603.04467.
- 39. Maaten, L.v.d.; Hinton, G. Visualizing data using t-SNE. J. Mach. Learn. Res. 2008, 9, 2579–2605.
- 40. Goodall, R.E.; Lee, A.A. Predicting materials properties without crystal structure: Deep representation learning from stoichiometry. *arXiv* 2019, arXiv:1910.00617.
- 41. Davies, D.W.; Butler, K.T.; Jackson, A.J.; Morris, A.; Frost, J.M.; Skelton, J.M.; Walsh, A. Computational screening of all stoichiometric inorganic materials. *Chem* **2016**, *1*, 617–627. [CrossRef] [PubMed]
- 42. Dan, Y.; Zhao, Y.; Li, X.; Li, S.; Hu, M.; Hu, J. Generative adversarial networks (GAN) based efficient sampling of chemical space for inverse design of inorganic materials. *arXiv* **2019**, arXiv:1911.05020. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).