Generative Adversarial Networks for Generation of Synthetic High Entropy Alloys

Christian Eike Precker, Andrea Gregores-Coto, Santiago Muíños-Landín

Abstract—High-performance materials are a key tool for several reasons. On the one hand, their use brings obvious progress in the performance of the pieces where they are used in fields such as aeronautics, construction, or biotechnology. On the other hand, high-performance materials also allow more efficient use of energy in industrial processes where the use of such energy becomes intensive with its consequences in terms of environmental and economic sustainability. For these reasons, the emergence of high-performance materials such as high entropy alloys (HEAs) has captured the attention of industry and researchers within the last years. However, the development of these materials requires a large amount of time and money invested in the design, synthesizability evaluation, construction, and characterization of such compounds. The use of artificial intelligence for the design of materials, even in its current infancy status, provides a valuable tool to accelerate the initial phases of materials design and HEAs, where the high number of combinations brings a perfect scenario for the deployment of Machine Learning techniques. In this work, a Generative based approach is used, namely Generative Adversarial Networks (GANs) to generate synthetic HEAs for highly intensive industrial processes. The architecture model of a GAN involves two neural networks. The first one is a generator model for generating chemical compositions of candidate alloys to form the HEAs. The second one is a discriminator model for classifying the generated samples coming from the generator in real or fake compositions. The discriminator learns from a specific data structure that contains data from real samples to classify the generated samples. A GAN extension that conditionally generates the synthetic outputs by the addition of extra inputs was used. This so-called conditional tabular generative adversarial network (CTGAN) was developed to be used with tabular datasets as input. Such data is normally composed of a mix of continuous and discrete columns, making some deep neural network models fail in performing a properly modeling for this kind of inputs. In the present approach, the generated realistic synthetic data was based on the conventional parametric design parameters used for HEAs, i.e., atomic size difference δ, mean atomic radius a, average melting temperature T_m, mixing enthalpy ΔH_{mix}, mixing entropy ΔS_{mix}, electronegativity χ, valence electron concentration (VEC), mean bulk modulus K, and the standard deviation for most of them. As conditioned input data, the chemical composition of the alloys and their phase has been considered. The phase was classified in four classes, namely amorphous, intermetallic, solid solution, and solid solution + intermetallic, which can be used as an indicator for their applicability. The CTGAN provides as output candidates of HEAs, the expected parameters mentioned above, and corresponding phase. The generated data is compared with the calculated data and a verification of novel generated compositions is done in open materials databases available in the literature. Finally, a specific data structure for the CTGAN training and results of the performance of this approach is provided, which was developed in the framework of the European project ACHIEF for the discovery of novel materials to be used in industrial processes.

C. E. Precker, A. Gregores-Coto, and S. Muíños-Landín are with the AIMEN Technology Centre, Smart Systems and Smart Manufacturing–Artificial Intelligence and Data Analytics Laboratory, PI. Cataboi, 36418 Pontevedra, Spain. e-mail: christian.precker@aimen.es.

Manuscript received August 18, 2021; revised August 24, 2021.

Keywords—Artificial Intelligence, High Entropy Alloys, Generative Adversarial Networks, Intensive Energy Processes.

I. Introduction

TECHNOLOGICAL advances lead humans to search for new possibilities in all fields of science. One example was the development of high-entropy alloys (HEAs) by two independent research groups in 2004 [1], [2], which is a class of materials containing multiple principal chemical elements in near-equiatomic proportions. These kinds of materials are of interest to many fields due to their remarkable physical properties, such as superior hardness, strength, and great wear resistance [3]. Before the HEAs, the common alloying approach consisted of using a primary element, e.g., iron, followed by the addition of small amounts of secondary elements, e.g., chromium, to increase corrosion resistance, and carbon to increase the strength. This primary element method makes the combination space of elements limited, whereas, in the case of HEAs, many exploitable combinations are still open for discovery, with improved mechanical and thermodynamic performance.

In HEAs, the presence of multiple chemical elements in near-equiatomic proportions (composed of five or more principal elements, with each of them possessing between 5 and 35 atomic percentage) increases sufficiently the entropy of mixing, overcoming the enthalpy formation of the compounds, giving rise to stable solid solution formations, rather than intermetallic compounds [1]. HEAs can also be defined in terms of the calculated mixing entropy ΔS_{mix} by the equation

$$\Delta S_{\text{mix}} = -R \sum_{i=1}^{n} c_i \ln c_i, \tag{1}$$

where c_i is the stoichiometric ratio of the i-th component in the alloy, and R is the gas constant [4]. The mixing entropy can be written in terms of the gas constant R, and HEAs defined when a composition has $\Delta S_{\text{mix}} \geq 1.5\,R$. For $1\,R \leq \Delta S_{\text{mix}} \leq 1.5\,R$ the compounds are defined as medium entropy alloys (MEAs) and for $\Delta S_{\text{mix}} < 1\,R$, low entropy alloys (LEAs) [5]. From Eq. (1) it is possible to see that with the increase of the number of elements, the entropy also increases, e.g., an alloy containing five and six equiatomic elements has $\Delta S_{\text{mix}} = 1.61\,R$ and $1.79\,R$ respectively.

The high entropy effect in HEAs is important because it can enhance the formation of phases. Among the phases in which HEAs can be found, the alloys can be classified as solid-solution (SS), intermetallic (IM), amorphous (AM), or a mixture of them. The SS phase means a significant or

complete mixing of all constituent elements in the structures of body-centered cubic (BCC), face-centered cubic (FCC), or hexagonal close-packed (HCP). IM phases mean stoichiometric compounds with specific Strukturbericht designation, such as B2 (for example NiAl) and $L1_2$ (for example Ni_3Al) [3], [5]. The phase is an important parameter for HEAs since it determines the physical properties. For example, to achieve high hardness, the SS is indicated, for better elasticity, the AM, and for great wear resistance, IM [6], [7].

With the graphic processor unit (GPU) developments using parallel processing for imaging, e.g., in video games, video processing, and simulations, in 2009, a window was opened to make use of GPUs in neural networks [8]. Since then, artificial intelligence applications have had significant growth in applications in all fields, such as in medicine, identifying metastatic breast cancer tumors, or in transport, with autonomous cars. Furthermore, artificial intelligence and machine learning methods can also be used in the field of materials science to speed up discoveries, saving time and money when compared to traditional methods [9], [10], [11].

In this work, artificial intelligence is used to generate synthetic data based on real HEAs found in the literature, and calculated parameters data, providing as output candidates of possible new alloys and corresponding parameters. For that, a dataset containing a large amount of HEAs and specific calculated design parameters [12] was used in a generative model, called conditional tabular generative adversarial network (CTGAN) [13].

II. Methodology

In this section, the methodology used will be discussed, i.e., how the generation of the dataset used in this work was done, as well as the explanation about the Deep Learning methods addressed. Fig. 1 shows a sketch of the workflow, where

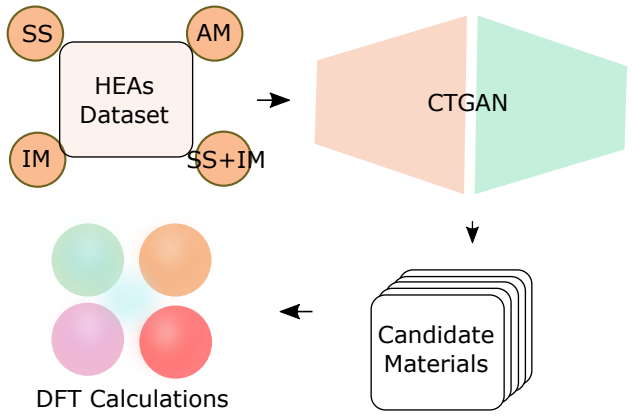

Fig. 1. The approach followed in this work to obtain synthetic HEAs compositions. A specific dataset has been developed and made accessible for this purpose, containing a large amount of HEAs data. The data feeds the CTGAN network architecture, where the Discriminator is used to classify the Generator outcomes until the synthetic compounds become realistic considering real data. Candidate materials that come out from the CTGAN are evaluated via DFT calculations to consider its synthesizability for a real case.

TABLE I
Design parameters.

Parameter	Equation
Mixing entropy	$\Delta S_{mix} = -R\sum_{i=0}^{n} c_i \ln c_i$
Mixing enthalpy	$\Delta H_{\text{mix}} = 4\sum_{i\neq j} c_i c_j H_{ij}$
Standard deviation of mixing enthalpy	$\sigma_{\Delta H} = \sqrt{\sum_{i\neq j} c_i c_j \left(H_{ij} - \Delta H_{mix}\right)^2}$
Mean atomic radius	$a = \sum_{i=0}^{n} c_i r_i$
Atomic size difference	$\delta = \sqrt{\sum_{i=0}^{n} c_i \left(1 - \frac{r_i}{a}\right)^2}$
Electronegativity	$\chi = \sum_{i=0}^{n} c_i \chi_i$
Electronegativity standard deviation	$\Delta\chi = \sqrt{\sum_{i=0}^{n} c_i \left(\chi_i - \chi\right)^2}$
Valence electron concentration (VEC)	$VEC = \sum_{i}^{n} c_i VEC_i$
Standard deviation of VEC	$\sigma_{VEC} = \sqrt{\sum_{i=0}^{n} c_i \left(VEC_i - VEC\right)^2}$
Mean bulk modulus	$K = \sum_{i=0}^{n} c_i K_i$
Standard deviation of bulk modulus	$\sigma_K = \sqrt{\sum_{i=0}^{n} c_i \left(K_i - K\right)^2}$
Averege melting temperature	$T_m = \sum_{i=0}^{n} c_i T_{mi}$
Standard deviation of melting temperature	$\sigma_{T_m} = \sqrt{\sum_{i=0}^{n} c_i \left(1 - \frac{T_{mi}}{T_m}\right)^2}$

the dataset containing the phases feeds the generative model, creating new candidates o HEAs, and finally verified in density functional theory (DFT) based open materials databases.

A. Data Collection

For this work, a dataset containing several HEAs was prepared. The HEAs were collected from works available in the literature and merged [7], [11], [14], [15], [16]. After filtering and removing duplicated compounds, the given dataset ended with 1117 entries [12]. The phases were used as conditional training parameters, and because this information for some

compounds was unknown, they were also removed, and the dataset used to train the CTGAN model had at the end 1103 entries, composed of 195 AM, 362 IM, 350 SS, and 196 SS+IM.

Previous studies on predicting HEAs phases have used parametric approaches, based on the Hume-Rothery rules, which concern the mutual solubility at high temperatures [3], [11], [17], [18], [19]. Based on these works, 13 design parameters were chosen, calculated, and included in the dataset, i.e., the mixing entropy ΔS_{mix}, where $R = 8.314$ J·K^{-1}·mol^{-1} is the gas constant, and c_i is the stoichiometric ratio of the i-th component in the alloy, the mixing enthalpy ΔH_{mix}, its standard deviation $\sigma_{\Delta H}$, where H_{ij} is the binary mixing enthalpy in the liquid phase, the mean atomic radius a, where r_i is the atomic radius of the i-th component in the alloy, the atomic size difference δ, the Pauling electronegativity χ, its standard deviation $\Delta\chi$, the valence electron concentration VEC, its standard deviation σ_{VEC}, the mean bulk modulus K, its standard deviation σ_K, the average melting temperature T_m, and its standard deviation σ_{T_m}, listed in Table I. Finally, columns containing the chemical elements and their corresponding fraction in the alloy were included and used as conditional training parameters.

B. CTGAN

A neural network (NN) can be fed with a dataset, so that it will learn the relationship between the features present in that dataset, i.e., what characterizes an output and classify it accordingly. But would a NN be able to create synthetic data that is very close to the real data? That's exactly what the generative adversarial networks (GANs) do, turning the world's attention to them in the last years due to their ability to generate realistic fake content. A GAN is a generative model first used to create images [20], but now the scope is extended to create other contents, e.g., furniture designs for 3D printing [21].

The GANs work with two NN models, one competing against the other. One of the models is called generator (G), responsible for generating synthetic data from a noisy entry z (a bunch of random values, e.g., randomized values from a normal distribution between 0 and 1), which tries to generate a synthetic sample as close as possible to a real one. The other model is called discriminator (D), which is trained with both real and fake data, learning the difference between them and classifying the data from the generator as real or fake. The results from the discriminator's classification are used as input for the Generator, which learns from these results and calibrates its weights to generate samples that look closer to the real samples. After the generator improves its generated data, the discriminator is also improved, being updated by the new bunch of samples coming from the generator, calibrating its weights, working as a loop, where the discriminator tries to become better at differentiating the generated from the real data.

In this work, the conditional tabular generative adversarial network (CTGAN) [13] was used, one member of the large family of GANs (DCGAN, WGAN, etc.). This specific kind

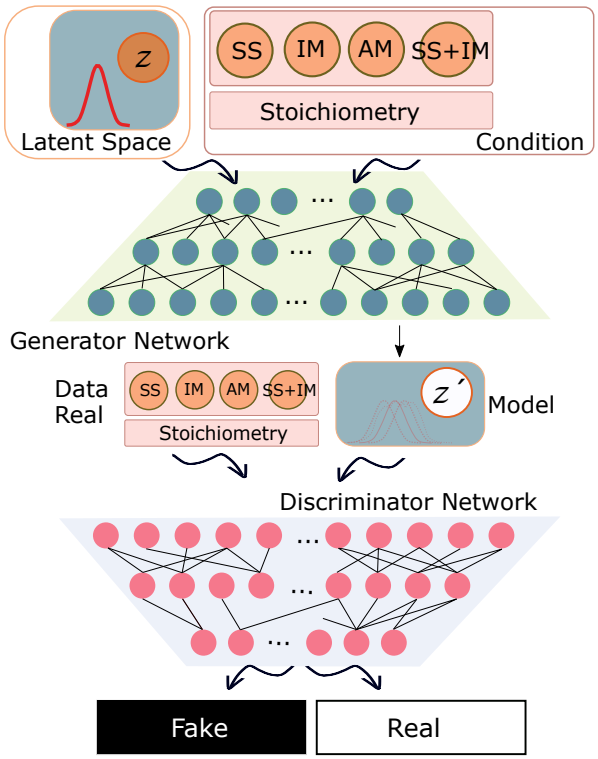

Fig. 2. The main structure of a CTGAN. Random vector and conditions feed the Generator Network, whose outcome feeds the discriminator to classify using also real data to distinguish between real and fake compounds. The architecture of the Networks is described in the main text. The conditions are based on the phases of the compounds and their stoichiometry.

of GAN gives solutions to data problems such as mixed data types, non-Gaussian distributions, multi-modal distributions, learning from sparse one-hot-encoded vectors, and highly imbalanced categorical columns, which normal GANs don't address. Since our HEAs dataset comprises mixed types of data, containing discrete and continuous values, the CTGAN addresses the needs imposed by the dataset, and it can be used to generate new synthetic tabular data.

The CTGAN can be conditioned on some extra information y, which can be any kind of auxiliary information, feeding the network with an additional input layer, e.g., class labels or data from other modalities. In a conditional GAN, the networks G and D are trained and optimized in an adversarial learning framework, called objective function, as follows [22]

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}(x)}[\log D(x|y)] + E_{zp_z(z)}[\log(1 - D(G(z|y)))], \quad (2)$$

where x represents the real data. In the conditional training, the CTGAN encodes the conditioned tabular data columns and categorical variables in condition vectors, using these vectors as generator inputs. This architecture uses recent GAN approaches where the quality and stability of the generated data are improved, e.g., it uses the discriminator of the PacGAN [23], and the loss function of the WGAN-GP [24], defined as

$$L = E_{G(z\sim P_g}[D(G(z))] - E_{x\sim P_r}[D(x)] + \lambda E_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla_{\hat{x}} D(\hat{x})\| - 1)^2], \quad (3)$$

where the two first terms are the original loss of the WGAN [25] and the last term the gradient penalty loss, implemented to control the discriminator's gradient for random samples $\hat{x} \sim P_{\hat{x}}$. $\hat{x}$ represents samples that are interpolated by the real data, λ is the gradient coefficient penalty, and the distribution of the real and generated data are represented by P_r and P_g.

Fig. 2 shows the CTGAN's architecture sketch, comprised of the generator and discriminator models with the conditional entries used in this work, i.e., the phases and stoichiometry to obtain the desired modes from the trained model. Table II summarizes the used architecture. The following parameters were used in both G and D neural networks models: Adam optimizer with a learning rate of 2×10^{-4}, and weight decay of 1×10^{-6}. For G, the ReLU activation was used in the input and hidden layers, and a tanh activation function in the output. For D, the LeakyReLU activation was used in the input and hidden layers and the sigmoid activation function in the output.

TABLE II
CTGAN STRUCTURE USED IN THIS STUDY.

	Generator		**Discriminator**	
Layer	Type	Dimension	Type	Dimension
Input	Latent + Cond.	90	Features + Cond.	71
Hidden 1	Dense layer	256	Dense layer	256
Hidden 2	Dense layer	128	Dense layer	128
Output	Dense layer	71	Dense layer	1

III. DISCUSSION

In this section, the obtained results using the structure shown in Table II and details mentioned in the previous section are discussed.

A. Generation of Data

The CTGAN was fed with real HEAs data as input, containing the stoichiometry, the phase, and 13 design parameters. The loss function for G and D versus the training epochs is showed in Fig. 3. The convergence of the loss function in both G and D occur at approximately epoch 50, which means that from this point, the model reached a limit where G and D stopped to evolve.

Once the model was trained, synthetic data based on the knowledge acquired during the training process was generated. The outputs provided by the CTGAN were the same 71 parameters used as inputs, i.e., the 13 design features, the 4 phases, and the columns containing the chemical elements fraction. Some of the generated compounds were identical to compounds provided in the input dataset, e.g., the compounds AlCo and TiZrNbMoV_2. All generated features were compared with the features in the initial dataset. As result, the values were identical, indicating that the generated data respects the knowledge acquired from the input data.

Within the generation, new compositions were also delivered by the CTGAN. Between the generated compounds, an experimentally known HEA that was not included in the initial dataset was generated, the TiZrCuNiBe, and the correct phase AM was attributed [26], which means that this method really opens the possibility of generating real

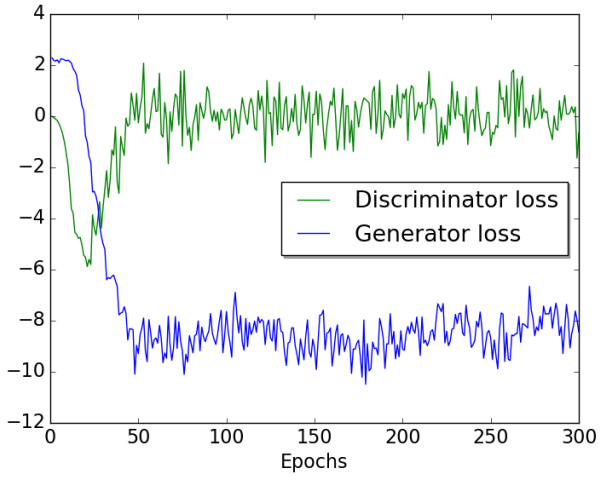

Fig. 3. Generator and Discriminator loss of the CTGAN training progress. The convergence after epoch 50 means that one of the models stopped evolving, and consequently the second model also stops to evolve, since it's a competition between generating fake samples that look real and discriminating these samples as real and fake.

HEAs. Some other examples of possible HEAs candidates (experimentally not proved) are $TiAl_{0.75}CrCo_7Ni$, B_2CoGa_2VZr, $Al_{0.5}BCoCr_{6.3}FeMn$, $AlCoNdNi_{10}Ti$, and $CoCuFeSn_3TiZn_{0.5}$.

The evaluation metrics used to get the score of the model were CSTest, KSTest, KSTestExtended, and ContinuousKL-Divergence, which are statistical metrics found in the ecosystem of libraries of the synthetic data vault (SDV) [27]. The average score value obtained from all these metrics together reached a value of 93 %.

Some of the present compounds in the initial dataset and some generated compounds were taken for evaluation in DFT-based open databases for materials, the Open Quantum Materials Database (OQMD) [28], [29], and Automatic-FLOW for Materials Discovery (AFLOW) [30]. Fig. 4 shows on the upper part four aleatory selected compounds from the HEAs dataset (real compounds), and at the bottom, four compounds selected from the CTGAN generation (synthetic compounds). They are classified according to their phase, i.e., the real phase for the dataset compounds and the expected phase attributed by the CTGAN for the generated compounds. The bars inside the phase areas compare the mixing enthalpy ΔH for the real compounds (in the case of the dataset, calculated from Table I) and synthetic with the mixing enthalpy calculated from the OQMD. When compared, the values are in good agreement in both cases for the real and synthetic data. Note that for comparison purposes, the mixing enthalpy modulus $|\Delta H|$ was used in Fig. 4. Other generated compounds were found in the database of AFLOW, e.g., AlCuTi, $Al_{0.5}CuV$, AlFeNi, and AlCrNiTi, which once more validates the CTGAN as a generative candidate model for the discovery of novel HEAs.

IV. CONCLUSION

With a specific database containing experimentally proved HEAs, a study using the generative model CTGAN was performed to generate new HEAs candidates. The CTGAN

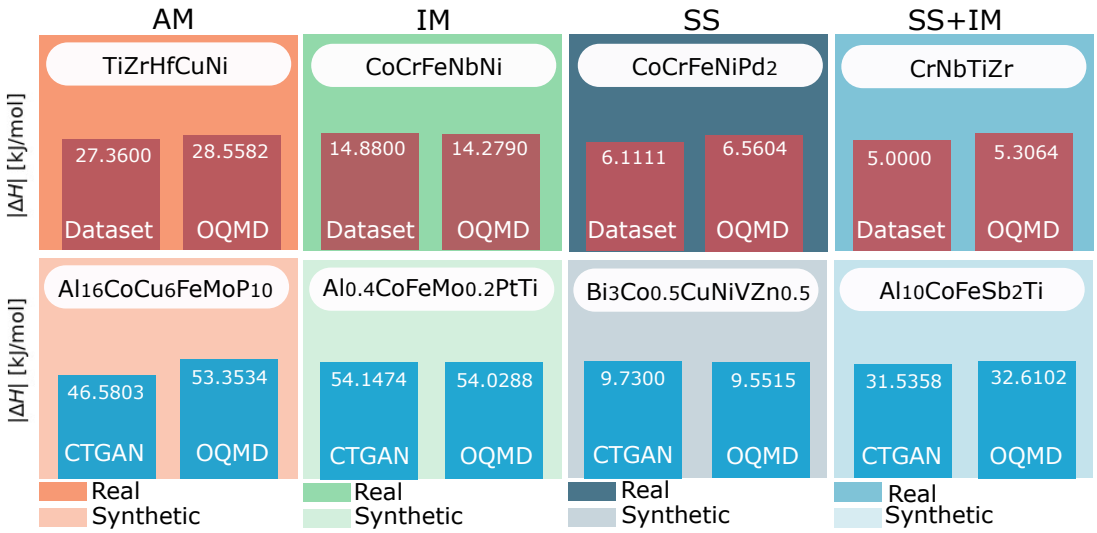

Fig. 4. Main results from the generative approach. Comparison of real values of ΔH for HEAs and those generated by our CTGAN. The figure shows results for the four different phases, AM, IM, SS, and SS+IM, separated in columns. Note that absolute values are shown (all the obtained values for ΔH are negative). The dark boxes contain compounds and ΔH values from real compounds (dark red columns), while those contained in the faded color boxes (blue columns) have been obtained with the CTGAN. The values of ΔH from the dataset and also those generated by the CTGAN are compared with the calculations performed using OQMD.

proved to be a tool that suits this purpose since it can generate experimentally proved compounds. To improve the ability to generate novel HEAs, some filters can be taken into account, e.g., a phase-oriented training, but in that case, more data would be necessary to train the model. Data validation was done using DFT-based open databases, e.g., comparing the mixing entropy calculated by the OQMD and the CTGAN generated values, and finding out that some of the generated compounds were also found in the AFLOW database. The use of artificial intelligence in the materials science field has the potential to improve the industrial sector, e.g., in the case of intensive energy processes by providing alloys with improved thermo-mechanical properties, such as a material with higher strength at high temperatures, being more agile and cost-friendly when compared with traditional methods as trial and error.

Acknowledgment

This research has received funding from the European Union's Horizon 2020 research and innovation programme under the project ACHIEF with Grant Agreement 958374.

The authors want to thank the comments and fruitful discussions with all the members of the Artificial Intelligence and Data Analytics Lab (AIDA-Lab) of the Smart Systems and Smart Manufacturing (S3M) department of AIMEN.

References

[1] J. W. Yeh, S. K. Chen, S. J. Lin, J. Y. Gan, T. S. Chin, T. T. Shun, C. H. Tsau, and S. Y. Chang, *Nanostructured High-Entropy Alloys with Multiple Principal Elements: Novel Alloy Design Concepts and Outcomes*, Advanced Engineering Materials, 2004, vol. 6, no 5, pp. 299-303, DOI: https://doi.org/10.1002/adem.200300567.

[2] B. Cantor, I. T. H. Chang, P. Knight, A. J. B. Vincent, *Microstructural development in equiatomic multicomponent alloys*, Materials Science and Engineering: A, 2004, Vol. 375–377, pp. 213-218, DOI: https://doi.org/10.1016/j.msea.2003.10.257.

[3] X. Wang, W. GUO, Y. FU, *High-entropy alloys: emerging materials for advanced functional applications*, Journal of Materials Chemistry A, 2020, vol. 9, pp. 663-701, DOI: https://doi.org/10.1039/D0TA09601F.

[4] J. W. Yeh, *Physical Metallurgy of High-Entropy Alloys*, JOM 67, 2015, pp. 2254–2261, DOI: https://doi.org/10.1007/s11837-015-1583-5.

[5] J. W. Yeh, *Alloy Design Strategies and Future Trends in High-Entropy Alloys*, JOM 65, 2013, pp. 1759–1771, DOI: https://doi.org/10.1007/s11837-013-0761-6.

[6] Z. Wang, S. Guo, C. T. Liu, *Phase selection in high-entropy alloys: from nonequilibrium to*, Jom, 2014, vol. 66, no 10, p. 1966-1972, DOI: https://doi.org/10.1007/s11837-014-0953-8.

[7] M. H. Tsai, R. C. Tsai, T. Chang, W. F. Huang, *Intermetallic Phases in High-Entropy Alloys: Statistical Analysis of their Prevalence and Structural Inheritance*, Metals, 2019, vol. 9, 247, DOI: https://doi.org/10.3390/met9020247.

[8] R. Raina, A. Madhavan, A. Y. Ng, *Large-scale Deep Unsupervised Learning using Graphics Processors*, ICML '09: Proceedings of the 26th Annual International Conference on Machine Learning, 2009, pp. 873-880, DOI: https://doi.org/10.1145/1553374.1553486.

[9] Y. Dan, Y. Zhao, X. Li, et al. *Generative adversarial networks (GAN) based efficient sampling of chemical composition space for inverse design of inorganic materials*, npj Comput Mater, 2020, vol. 6, 84, DOI: https://doi.org/10.1038/s41524-020-00352-0.

[10] Kaufmann, K., Maryanovsky, D., Mellor, W.M. et al., *Discovery of high-entropy ceramics via machine learning*, npj Comput Mater 6, 42 (2020), DOI: https://doi.org/10.1038/s41524-020-0317-6.

[11] Zhou, Z., Zhou, Y., He, Q. et al., *Machine learning guided appraisal and exploration of phase design for high entropy alloys*, npj Comput Mater 5, 128 (2019), DOI: https://doi.org/10.1038/s41524-019-0265-1.

[12] C. E. Precker, A. Gregores Coto, and S. Muíños-Landín, *Materials for Design Open Repository. High Entropy Alloys*, Zenodo, 2021, DOI: https://doi.org/10.5281/zenodo.5155149.

[13] L. Xu, M. Skoularidou, A. Cuesta-Infante, and K. Veeramachaneni, *Modeling Tabular data using Conditional GAN*, NeurIPS, 2019, Available: https://arxiv.org/abs/1907.00503v2.

[14] D. B. Miracle, O. N. Senkov, *A critical review of high entropy alloys and related concepts*, Acta Materialia, 2017, Vol. 122, pp. 448-511, DOI: https://doi.org/10.1016/j.actamat.2016.08.081.

[15] S. Gorsse, M. H. Nguyen, O. N. Senkov, D. B. Miracle, *Database on the mechanical properties of high entropy alloys and complex con-*

centrated alloys, Data in brief, 2018, vol. 21, p. 2664-2678, DOI: https://doi.org/10.1016/j.dib.2018.11.111.

[16] M. Vaidya, G. M. Muralikrishna, B. S. Murty, *High-entropy alloys by mechanical alloying: A review*, Journal of Materials Research, 2019, vol. 34, pp. 664–686, DOI: https://doi.org/10.1557/jmr.2019.37.

[17] W. Huang, P. Martin, H. L. Zhuang, *Machine-learning phase prediction of high-entropy alloys*, Acta Materialia, 2019, vol. 169, pp. 225-236, DOI: https://doi.org/10.1016/j.actamat.2019.03.012.

[18] Y. Zhang, X. Yang,P. K. Liaw, *Alloy Design and Properties Optimization of High-Entropy Alloys*, JOM, 2012, vol. 64, pp. 830–838, DOI: https://doi.org/10.1007/s11837-012-0366-5.

[19] Y. Zhang, Y. J. Zhou, J. P. Lin, G. L., Chen, P. K. Liaw, *Solid-Solution Phase Formation Rules for Multi-component Alloys*, Advanced Engineering Materials, 2008, vol. 10, issue 6, pp. 534-538, DOI: https://doi.org/10.1002/adem.200700240.

[20] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, *Generative Adversarial Networks*, 2014, Available: https://arxiv.org/abs/1406.2661.

[21] C. González-Val, S. Muíños-Landín, *Generative design for Social Manufacturing*, Proceedings of the Workshop on Applied Deep Generative Networks, 2020, 2692, DOI: https://doi.org/10.5281/zenodo.4597558.

[22] M. MIRZA, S. OSINDERO, *Conditional generative adversarial nets*, 2014, Available: https://arxiv.org/abs/1411.1784.

[23] Z. Lin, A. Khetan, G. Fanti and S. Oh, *PacGAN: The Power of Two Samples in Generative Adversarial Networks*, IEEE Journal on Selected Areas in Information Theory, 2020, vol. 1, no. 1, pp. 324-335, DOI: 10.1109/JSAIT.2020.2983071.

[24] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. C. Courville, *Improved training of Wasserstein GANs*, 2017, Available: https://arxiv.org/abs/1704.00028.

[25] M. Arjovsky, S. Chintala, and L. Bottou, *Wasserstein GAN*, 2017, Available: http://arxiv.org/abs/1701.07875.

[26] H.Y. Ding, K.F. Yao, *High entropy Ti20Zr20Cu20Ni20Be20 bulk metallic glass*, Journal of Non-Crystalline Solids, 2013, vol. 364, pp. 9-12, ISSN 0022-3093, DOI: https://doi.org/10.1016/j.jnoncrysol.2013.01.022.

[27] N. Patki, R. Wedge, K. Veeramachaneni, *The Synthetic Data Vault*, 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2016, pp. 399-410, DOI: https://doi.org/10.1109/DSAA.2016.49.

[28] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, C. Wolverton, *Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD)*, JOM, 2013, vol. 65, pp. 1501-1509, DOI: https://doi.org/10.1007/s11837-013-0755-4.

[29] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, C. Wolverton, *The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies*, npj Computational Materials 1, 2015 15010, DOI: https://doi.org/10.1038/npjcompumats.2015.10.

[30] S. Curtarolo, W. Setyawan, G. L. W. Hart, M. Jahnatek, R. V. Chepulskii, R. H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M. J. Mehl, H. T. Stokes, D. O. Demchenko, D. Morgan, *AFLOW: An automatic framework for high-throughput materials discovery*, Computational Materials Science, 2012, vol. 58, pp. 218-226, ISSN 0927-0256, DOI: https://doi.org/10.1016/j.commatsci.2012.02.005.