Target-learning the latent space of a variational autoencoder model for the inverse design of stable perovskites

Tetteh Chenebuah, Ericsson; Nganbe, Michel; Tchagang, Alain

Publisher's version / Version de l'éditeur:
https://doi.org/10.21428/594757db.07402193
Proceedings of the Canadian Conference on Artificial Intelligence, 2023-06-05

Target-learning the Latent Space of a Variational Autoencoder model for the Inverse Design of Stable Perovskites

Ericsson Tetteh Chenebuah †¥*, Michel Nganbe †, Alain Beaudelaire Tchagang †¥

† Department of Mechanical Engineering, University of Ottawa, 161 Louis-Pasteur, Ottawa, ON, K1N 6N5, Canada.

¥ Digital Technologies Research Centre, National Research Council of Canada, 1200 Montréal Road, Ottawa, ON, K1A 0R6, Canada.

Abstract

At the forefront of discoverable materials are perovskites that stand out as some of the most chemically diverse and multifunctional energy materials. Theoretically, the estimated number of ternary perovskites exceeds a hundred thousand distinct compounds, notwithstanding that only a small fraction of this estimate has been reported in existing crystal databases. Therefore, the study takes advantage of the reliable, inexpensive and rapid opportunity offered by deep generative modeling for accelerating the search for unknown perovskites. In the process of making such findings, an inverse design-modeling scheme is resolved, which aims at assimilating deterministic target properties with their corresponding perovskite structure. The inverse design pipeline is architectured by combining a generative Variational AutoEncoder (VAE) model with Target-Learning (TL) feed-forward neural networks to form the TL-VAE perovskite generator, thereby making the complete modeling process semi-supervisory. The TL feed-forward neural network model serves the purpose of organizing the non-linear latent space of the VAE model and further assists in isolating deterministic target properties that are of interest to the core objective of the study. The property to be target-learned in the latent space is the formation energy, which is a crucial indicator for calibrating perovskite stability. The results report the discovery of promising new perovskite candidates, which are unique and polymorphic material variants. Upon conducting Density Functional Theory (DFT) validation on the newly identified perovskites, candidates that undergo full geometrical relaxation are recommended for further investigation and/or synthesization. In conclusion, the study demonstrates the efficacy of the inverse design TL-VAE model for the generation of stable ternary perovskites.

Keywords: Perovskite discovery, Inverse design, Variational autoencoder, Latent space, Semi-supervised

1. Introduction

Traditionally, the discovery of materials has been conducted by experimental synthesization and/or direct first-principle (ab initio) calculations. Despite both techniques having achieved major successes in the past, certain limitations exist, which can hinder their generalizability or applicability. For instance, direct first-principle deterministic methods strongly rely on the systematic screening of promising candidates by performing noble ionic substitutions on new chemical combinations. Such screening processes are typically high-throughput quantum-mechanically demanding and, as such, strictly involve solving the costly Schrodinger equation on a many-body Hamiltonian system to obtain the ground-state material properties [1]. On the other hand, experimental synthesization is generally a trial-and-error approach necessitating considerable resources that make it unpractical when a large variety of materials possibilities are to be explored. The challenges associated with these tried-and-tested means open up opportunities for other innovative materials discovery routes. Specifically, AI-driven techniques are currently gaining popularity due to their advantageous qualities over their contemporary counterparts. With advancements in supercomputing and verifiable databases, AI is now applied in material informatics to provide

* echen013@uottawa.ca

novel solutions due to their reliability, high speed and low-cost [2-6]. In particular, by solving an inverse design scheme, unsupervised AI that is driven by Deep Learning (DL) models, can identify certain underlying features associated with the latent space of a dataset. The goal of the inverse design approach is therefore to regenerate unknown materials given their initialized target properties through a sequential deconstruction of DL algorithms. At the forefront of Deep Generative Models (DGM) frequently used to solve the inverse design scheme are Variational AutoEncoders (VAE) and Generative Adversarial Networks (GAN). Although fundamentally different, both operate by transforming the original training data into a valuable latent (or noisy) space with emerging materials exhibiting definitive properties that are of significant chemical variation to the originals. In some prior arts moreover, the latent space has been treated methodically by performing explorative and exploitative analysis, which has led to several material findings of experimental importance. For example, the discovery of new metastable Vanadium Oxide (VO) materials in their polymorphic forms [3] was made possible by exploring the continuous latent space of a VAE model. A property-structured VAE model was engineered for discovering general inorganic materials that can be applied in solar cells [4]. In a different study moreover, the efficient sampling of a chemical composition latent space using a Wasserstein GAN produced new materials that were considerably different from the training dataset [5]. Likewise, the development of a constrained deep convolutional GAN was architectured specifically for the discovery of Bismuth Selenide (Bi-Se) [6]. A common denominator with all aforementioned studies (and many more) is in the conceptualization of a well-structured latent space, knowing fully well that the novel generation of new properties of interest will emerge from there. Moreover, the latent space can be influenced by performing sampling operations (e.g. semantic vector interpolation), and/or by constraining the mapping of embedding features usually in the form of a semi-supervisory learning approach [7]. Constrained target-learning can be achieved with the incorporation of a supervisory model to the unsupervised training process. By doing so, the general cost function of the overall model is modified to include the inter-dependable training losses from the different learning modes. For materials discovery, target-learning the latent space of a DGM can therefore generate novel materials that possess deterministic properties of interest, e.g. stable vs unstable formation energy, finite vs infinite bandgaps, crystal system class, etc.

Hence, the present study develops a semi-supervisory VAE model that learns the latent space based on a predefined target towards the generation of novel stable perovskite energy materials. Unlike organic drug or protein molecules that have greatly been the subject of many discoveries using AI-driven simulation [8], the perovskite material class is a crystalline structure with long-range atomic ordering. Moreover, perovskites are among the most versatile energy materials with multifunctional properties extending to superconductivity, photocatalysis, ferroelectricity, piezoelectricity, optoelectronics, photovoltaic, biomaterials, and so on [9, 10]. Their multifunctionality is credited to the overwhelming wealth of design space afforded by their stoichiometry as multiple chemical elements across the periodic table can occupy distinctive ionic sites within the crystal structure. Considering the ideal and most prevalent ternary ABX_3 perovskite stoichiometry for example, the A-site and B-site chemical elements are normally cations (positively charged) and are coordinated by twelve and six X-site anions (negatively charged), respectively, to form a crystal lattice [11]. Taking into account the 94 naturally occurring chemical elements on the periodic table, while mindful of ionic-swapping forms (i.e. anti- or inverse- perovskites), the potential number of unique ABX_3 compounds could be estimated at $\mathbb{C}_3^{94} = 134{,}044$. This rough estimate of the perovskite design space could even climb further with the consideration of other polymorphic or allotropic stoichiometrical forms with significantly different physical phases. With only a fraction of this estimate presently discovered (as reported in several crystal databases [12-13]), DGM processes can accelerate novel discoveries of such chemical combinations for applications in several industries related to energy, healthcare, biological, defense,

agricultural, etc. In the process of searching for new perovskites using ML, the formation energy will serve as a learnable target for organizing the non-linear latent space. By definition, the formation energy $(\boldsymbol{E_f})$ is the energy required to form a chemical structure from its disintegrated form and is necessary for developing phase diagrams [14]. Perovskite materials with $\boldsymbol{E_f}$ values below a predefined threshold, can therefore be considered physically stable, and be recommended for further material analysis and/or synthesization. To ensure the generated candidates meet formability and geometrical relaxation standards, Density Functional Theory (DFT) [1, 14] first-principle calculations are implemented on the novel perovskite candidates using the Quantum Espresso package [15].

This current paper is therefore organized as follows. First, the methodology used to represent a perovskite material in the training set for the modeling exercise is explained. Second, procedure details are provided on how the Target-learning VAE (TL-VAE) model was developed consequential to perovskite discovery. The effective sampling strategy carried-out for performing semantic vectorial operations in hyperdimensional latent space is then presented. Finally, the results are demonstrated for new material findings and are further validated using DFT computational means.

2. Methods

2.1. Invertible Descriptor Design

Unlike proteins and drug-like molecules, crystals in general do not have absolute descriptor designs due to their periodic boundary conditions and diverse range of stable stoichiometries. Therefore, the present study customizes a descriptor design specifically for representing perovskites. The descriptor design is inspired by the Fourier Transformed Crystal Property (FTCP) representation and takes advantage of the reciprocal space of a crystal lattice [4]. However, unlike the FTCP representation that is built for general inorganic crystal materials, the present study focusses on the ABX_3 perovskite material class. Therefore, the current descriptor design is remodeled to capture materials of the perovskite stoichiometry and incorporates additional features that are fundamentally missing in the FTCP descriptor. As such, the present descriptor design consists of two main feature blocks, namely: (1) Direct Space Features (DSF) and (2) Periodic Reciprocal Features (PRF). The DSF block accommodates explicit crystal properties that wholly describe the real content of the crystal and are crucial for reconstructive purposes. In addition, the DSF incorporates thirteen discretized chemical properties (one-hot encoded vectors) that characterize the respective chemical element occupying a specific crystallographic site in the perovskite material. The selected property features represent thermochemistry qualities and significantly aid in the target-learning of the desired properties. On the other hand, the PRF block serves the purpose of modeling the long-range atomic ordering of a periodic crystal material and further assists in accurately predicting the target properties. By implementing a Fourier transforming operation [4, 16-17], all discretized vectors in the DSF are projected into the reciprocal space of a crystal lattice as described in Equation-1:

$$\boldsymbol{S_{hkl}} = \sum_j \left\{ -\mathbf{0.5}\boldsymbol{i}\left(\boldsymbol{Q_j}\, \boldsymbol{ln}\left[\boldsymbol{e}^{\mathbf{2\pi i}\left(\boldsymbol{hx_j}+\boldsymbol{ky_j}+\boldsymbol{lz_j}\right)}\right]\right)\right\} \tag{1}$$

Whereby for a specific set of $\boldsymbol{hkl}$ miller indices (i.e. crystallographic plane in 3D space), $\boldsymbol{Q_j}$ is the discretized chemical vector (e.g. electronegativity) that is Fourier-transformed using their respective atomic coordinates $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$. As a result, $\boldsymbol{S_{hkl}}$ represents a vector-wise summation over all $\boldsymbol{j}$ atoms that make up the perovskite crystal. Note that the $\boldsymbol{hkl}$ planes used for projection are limited to $|\boldsymbol{h}| + |\boldsymbol{k}| + |\boldsymbol{l}| \leq \mathbf{3}$, and in the present study, they are 57 in total. Supplementary to the Fourier-transformed features, the PRF block also contains other reciprocal properties such as: the transformed reciprocal lattice vectors, the magnitude of the

reciprocal vector normal to a crystal plane $|\boldsymbol{G_{min}}|$, and the shortest distance between similar crystal planes, $\boldsymbol{d_{hkl}}$. Overall, both the DSF and PRF blocks are adjoined along their adjacent column arrays. Prior to adjoining however, the PRF side is zero-padded to fit the matrix organization. The invertible descriptor design is reorganized into a $(\mathbf{60 \times 88 \times 3})$ high-dimensional input (RGB image). Fig. 1(a) reveals a ball-and-stick chemical representation of a ferroelectric KNO_3 perovskite sample from the training set (Inorganic Crystal Structure Database (ICSD) ID: 384). The corresponding grayscale and reshaped RGB feature-engineered images are visualized using Fig. 1(b) and Fig. 1(c), respectively. From the normalized image representations, it can be observed that denser pixel values cluster at the lower regions due to the prominent effect of the Fourier-transformed reciprocal space. Fig. 1(d) illustrates the adjoined feature-engineering array that is transformed into a high-dimensional input image.

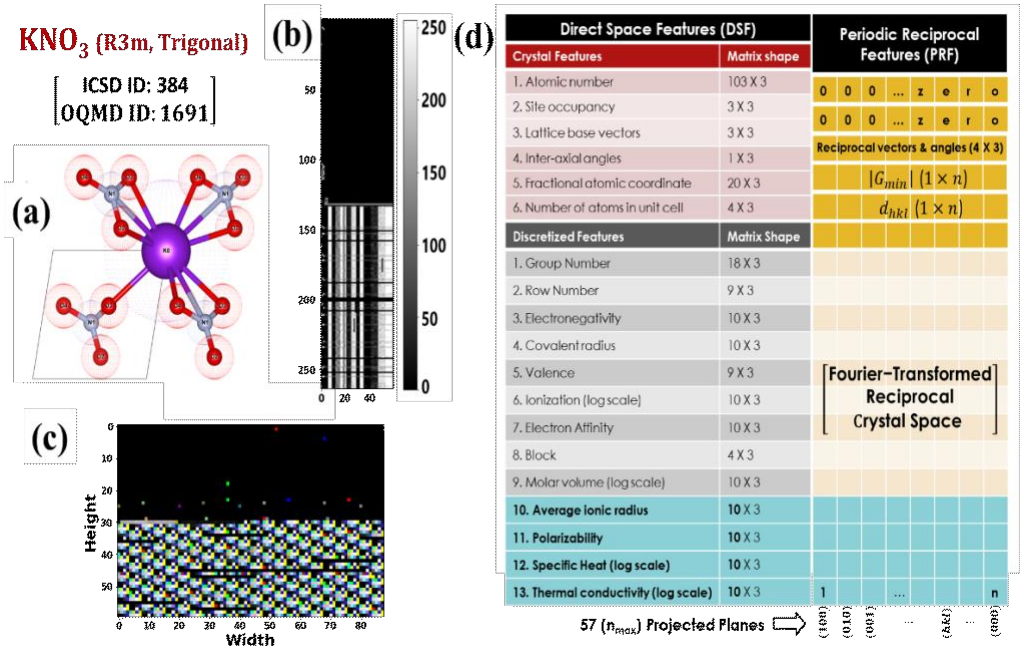

Figure 1. (a) Ball-and-stick KNO_3 perovskite in the training set; (b) corresponding gray-scale image; (c) reshaped $(\mathbf{60 \times 88 \times 3})$ RGB image; and (d) Invertible descriptor design with all feature embedding. Relative to the FTCP representation [4], the additional features incorporated in the present descriptor design include average ionic radius, polarizability, specific heat, and thermal conductivity, as well as, the reciprocal lattice vectors, the magnitude of the reciprocal vector normal to a crystallographic plane $|\boldsymbol{G_{min}}|$, and the shortest distance between similar planes $\boldsymbol{d_{hkl}}$.

2.2. Target-learning Variational AutoEncoder (TL-VAE) model architecture

The inverse design approach can be formulated as: Given the desired properties, find the perovskite crystal structure, i.e. $chemical\ composition + conformational\ structure = \boldsymbol{f^I}(properties)$, where $\boldsymbol{f^I}$ is the inverse function or sequence of deconstructive Deep Learning (DL) algorithm that regenerates the crystal structure from the predefined target properties. In the present study, a semi-supervised VAE model [18-19] is used to solve the inverse design challenge for the generation of novel perovskites. Given a known set of perovskite samples (i.e. real data points in training set) $\{\boldsymbol{x_i}\} \subseteq \boldsymbol{X}$, the developed VAE model encodes all perovskite inputs using a convolutional neural network. In contrast to some VAE materials discovery applications that encode using a one-dimensional convolutional neural network (e.g. FTCP), the current study rather applies two-dimensional convolution encoding layers (Conv2D). This therefore enables a thorough screening of all contributing pixel

embedding by 2D kernel filter size, given that the current descriptor image has three channels. The encoded perovskite inputs are mapped via a distribution $\boldsymbol{P_\theta}$ onto a hyperdimensional latent space $\{\boldsymbol{z_i}\} \subseteq \boldsymbol{Z}$, which is parameterized by $\boldsymbol{\theta}$. Emerging from the re-parameterization technique, the new data points generated by the model are thus of close variance to the original $\boldsymbol{x_i}$. The reparameterization produces distinguished latent sampling vectors for each perovskite feature representation, as expressed in Equation-2:

$$\boldsymbol{z = \mu + \sigma \odot \epsilon}, \qquad \textbf{where } \boldsymbol{\epsilon \sim \mathcal{N}(0, I)} \tag{2}$$

Whereby $\boldsymbol{z}$ is the sampled perovskite latent vector; $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are deterministic vectors denoting mean and standard deviation, respectively; and $\boldsymbol{\epsilon}$ is a random variable, which can be generated using a standard Gaussian (normal) distribution $\boldsymbol{\mathcal{N}}$. As such, VAE reconstructions are statistical variants of original perovskite inputs that potentially possess definitive properties from the unsupervisory training process. In general, the training losses for a VAE model are of two types, namely: reconstruction loss and Kullback-Leibler (KL) loss. The reconstruction loss (i.e. Mean Squared Error (MSE) or L2 loss) ensures that perovskite outputs emerging from the decoding phase are in close similarity to their original encoded forms. The KL divergence loss function measures the divergence (or distance) between two probability distributions and is responsible for the smooth overlap of perovskite data points within the latent space. Through a sequence of gradient descent and back-propagation, the goal of a VAE model is therefore to minimize both MSE and KL losses. Optimizing both reconstruction and KL losses results in the generation of a well-structured latent space that maintains similarity to neighboring perovskite points through clustering and dense-packing. Moreover, the latent space can be further optimized by including an additional loss function from a supervised learning arm, which further assists in distinguishing specific regions of interest in the latent space. As illustrated in Fig. 2, the present study develops such approach by combining both the unsupervisory learning architecture from the VAE model with two supervisory Multi-layer Perceptron (MLP) models (i.e. feed-forward neural networks). The two interdependent MLP models (i.e. $\boldsymbol{f(z) \mapsto E_f}$) are structured for target learning the formation energy based on a dual regression and classification analysis. For regressive analysis, the MLP function $\boldsymbol{f(.)}$ maps the known perovskite inputs to the unknown continuous target variables, while for classification, the function simply distinguishes predefined binary labels. For classification purposes, perovskites with formation energy values $\boldsymbol{E_f} \leq -1.5$ eV/atom [4] are labeled as stable (1). Otherwise, they are considered unstable (0). The study utilizes both regressive analysis and classification based on the rationale that the latent space can be non-linearly separated into two distinct regions (i.e. classification) with both regions attempting to reveal the order in continuous values (i.e. regression). As a result, the loss functions from both MLPs that are incorporated into the general VAE model are the mean squared error and the binary cross-entropy (log-loss) for regression and classification, respectively. The overall losses to be minimized can therefore be summarized using a simplified expressional form, as given in the Equation-3:

$$\boldsymbol{\mathcal{L}} = \underbrace{\frac{\boldsymbol{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}\left(\boldsymbol{X_i - \hat{X}_i}\right)^{\boldsymbol{2}}}_{\boldsymbol{reconstruction}} + \underbrace{\mathbb{E}_{\boldsymbol{q(z)}}\left[\ln \boldsymbol{q(z)}/\boldsymbol{p(z|x)}\right]}_{\boldsymbol{KL\left(q(z)|p(z|x)\right)}} + \boldsymbol{MLP\ losses} \tag{3}$$

Whereby $\boldsymbol{\hat{X}}$ denotes the decoded perovskite input features that are in close similarity to their originally encoded forms $\boldsymbol{X}$ over $\boldsymbol{n}$ number of perovskite samples, i.e. $i = 1,2, \dots, n$. For the KL loss, $\boldsymbol{q(z)}$ is the probabilistic distribution of the continuous prior model characterizing the latent space $\boldsymbol{z}$; $\boldsymbol{p(z|x)}$ is the decoder (i.e. recognition) probabilistic distribution; and $\mathbb{E}_{\boldsymbol{q(z)}}$ is the expectation with respect to the prior model [19].

2.3. Latent space sampling strategy

Several sampling strategies for conducting latent space operation have been proposed in past studies such as random sampling, local perturbation, global perturbation, and spherical linear interpolation. The FTCP inverse design for example utilizes local perturbation to search for unknown compounds that are within a fixed Gaussian noise scale of 0.4 around the training dataset [4]. Such confined search is suggested to promote higher exploitation (i.e. validity) at the expense of exploration (i.e. variety). The present study rather utilizes the Spherical Linear Interpolation (SLERP) [20] technique for extensively exploring the latent space due to its preferred functionality for carrying-out complex vector interpolations in the hyperdimensional space. Given the sampling perovskite data points $\{\boldsymbol{z_i}\}_{\boldsymbol{i=1}}^{\boldsymbol{n}}$ that are within the region of interest in the hyperdimensional latent space, SLERP is formulated as:

$$\vec{\boldsymbol{Z}}_{\boldsymbol{12}}(\boldsymbol{z_1}, \boldsymbol{z_2}; \boldsymbol{t}) = \boldsymbol{z_1}\frac{\sin(\boldsymbol{1}-\boldsymbol{t})\boldsymbol{\theta}}{\sin\boldsymbol{\theta}} + \boldsymbol{z_2}\frac{\sin \boldsymbol{t\theta}}{\sin\boldsymbol{\theta}} \tag{4}$$

Whereby $\vec{\boldsymbol{Z}}_{\boldsymbol{12}}$ is the interpolated vector between two fundamentally different perovskite points $\boldsymbol{z_1}$ and $\boldsymbol{z_2}$ along finite length $\boldsymbol{t} \in [\boldsymbol{0}, \boldsymbol{1}]$ and $\boldsymbol{\theta}$ is the angle between both vectors. As a result, $\vec{\boldsymbol{Z}}_{\boldsymbol{12}}$ contains intrinsic and beneficial properties hereditary to both $\boldsymbol{z_1}$ and $\boldsymbol{z_2}$ perovskite data points.

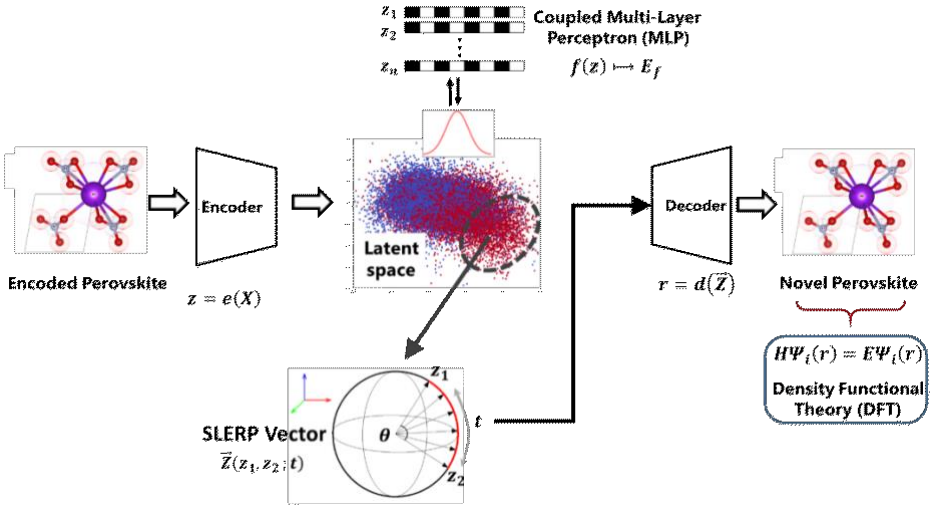

Figure 2. Target-learning Variational AutoEncoder (TL-VAE) pipeline for novel perovskite discovery.

3. Results

3.1. Perovskite Dataset

For deep learning, the study harnesses 27,587 inorganic perovskite samples of the ABX_3 stoichiometry from the Open Quantum Materials Database (OQMD) [12]. The OQMD platform is a high-throughput materials database with more than one million density functional theory (DFT) compounds in total energy calculations, and has been utilized in many scientific data mining studies. For data screening and preprocessing measures, the study implements similar method previously described in our past work [17]. The extracted and preprocessed dataset contains about 5% of International Crystal Structure Database perovskites (i.e. mostly experimentally verified), while the rest are DFT theoretical perovskites. Based on the predefined formation energy threshold ($\boldsymbol{E_f \leq -1.5\ eV/atom}$), 55.5% of all samples are confirmed to be stable and can hence be potentially synthesized. In addition, the dataset

consists of about 30% of non-oxide perovskites, which makes it considerably diverse, given that a vast amount of discoverable perovskites are limited to oxides (O^{2-}). The diverse compounds that are non-oxide perovskites include Br^-, Cl^-, F^-, H^-, I^-, N^{3-}, S^{2-}, Se^{2-}, and Te^{2-} anions. These anions are situated at the X-site of the ABX_3 stoichiometry, and they are all with a sizable amount of data representation.

3.2. Target prediction and inverse reconstruction

Generally, good invertible descriptors are expected to be *two-way* in form and application. The two-way analogy refers to their applicability for dual purposes, as it relates to forward target-mapping and inverse generative modeling. The study therefore examines the present descriptor capabilities by performing preliminary investigations in order to confirm such target-mapping and recoverability characteristics. As a result, the target-mapping efficacy of the descriptor is visualized using Fig. 3(a), which graphs the kernel density regressive fitting on the predicted versus actual formation energy. The prediction errors are evaluated on a 20% held-out test set, and are estimated at 0.125 eV/atom, 0.222 eV/atom, and 96.04% R^2, corresponding to the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and coefficient of determination, respectively. Furthermore, Fig. 3(b) illustrates the result of the binary classification analysis using a heat map (confusion) matrix. Considering the predefined formation energy threshold on stability, the F1-score and Receiver Operating Characteristic (ROC) scores are determined at 95% and 98.5% on test set, respectively. Note that for both cases involving regressive and classification analysis, similar deep learning architectures are utilized in the forward design approach. However, modeling modifications are enabled at the final outputted layer by using linear and sigmoidal activation functions for regression and classification analysis, respectively. Moreover, the descriptor's capability for use in the case of generative inverse design modeling is likewise investigated. The important features considered for such analysis include all crystal properties that are located in the DSF block of the descriptor (Fig. 1(d)). Besides, these important features are crucial for further DFT verification on newly generated perovskites. Using an autoencoder, Table 1 reports the crystal features reconstruction errors on 60% training, 20% validation, and 20% testing sets. The results realized in the demonstration confirms the good recoverability quality of the developed inverse descriptor design. Considering some previous benchmark evaluations for general materials design using machine learning [4, 17], the current invertible descriptor design is seen to be effective for generative perovskite modeling that incorporates conditional target-learning operation.

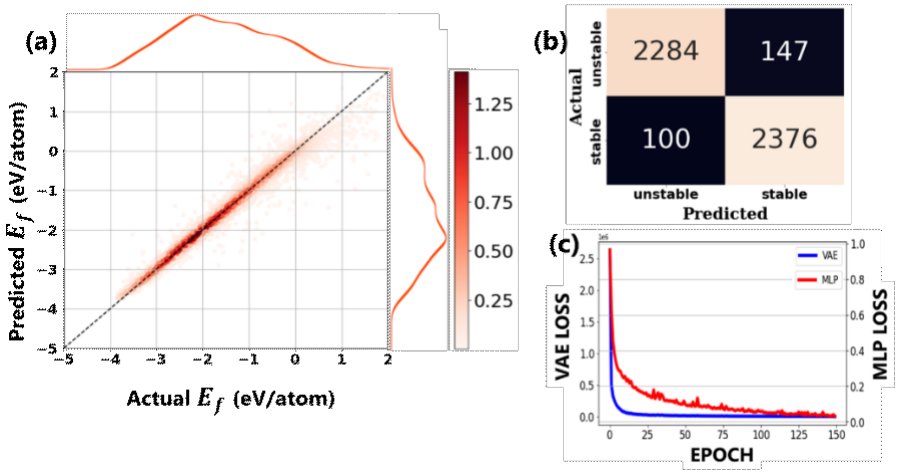

Figure 3. (a) Test set regression fitting on $\boldsymbol{E_f}$; (b) confusion matrix on test set with respect to stable vs unstable $\boldsymbol{E_f}$; and (c) VAE and MLP learning curves.

Table 1. Reconstruction error of important feature (descriptor) embedding on the training, validation and test sets at 60%, 20%, and 20%, respectively.

Features	Atomic number	Atomic site occupancy	Lattice vectors (Å)	Angle (°)	Fractional atomic coordinate (Å/Å)	Number of atoms in a cell
Training set						
MAE	0.016	0.237	0.764	4.809	0.135	0.036
RMSE	0.128	0.487	1.574	9.986	0.206	0.190
Validation set						
MAE	0.015	0.242	0.762	4.886	0.137	0.038
RMSE	0.123	0.492	1.583	9.970	0.209	0.194
Testing set						
MAE	0.016	0.241	0.820	5.289	0.136	0.038
RMSE	0.127	0.491	1.656	10.795	0.207	0.196

3.3. Sampling novel perovskites

The sampling process begins by identifying the region in hyperdimensional latent space where stable perovskites are most likely to cluster. Using Keras with a TensorFlow backend, the present TL-VAE model architecture dimensionally reduces each encoded perovskite in the training set into a latent-space vector length of $\mathbb{R}^{256}$: 1×256. The dimensionally reduced space is preliminarily explored using data visualization technique, as it relates to the Principal Component Analysis (PCA). By using the best factorial axis that produces the largest variance, the hyperdimensional space is further reduced into two principal components for observing the latent space. The present study preferably applies t-Distributed Stochastic Neighbor Embedding (t-SNE) PCA over other decomposition or clustering algorithms, which is strongly due to their ability in capturing non-linear data distributions in hyperdimension [21]. As a result, Fig. 4(a) demonstrates the results from the PCA investigation, revealing data agglomeration of marginally higher formation energy values towards the center (lighter-colored region). The clustering, as observed using the PCA study, is substantially due to the coupled target-learning MLP network, which assists in the organization of the latent space. For sampling and sequel interpolation, the real latent space values are used instead of the PCA values. This is because PCA applications are irreversible due to loss of information that accompanies data projection. Considering the unaltered $\mathbb{R}^{256}$ values of vector length, the top two axis that adequately capture displacement of data points into visible stable and unstable regions are identified. Hence, Fig. 4(b) displays the 2D plot using the identified top two axes. The red and blue points represent stable and unstable perovskites, respectively. Fig. 4(c) closely shows the concentration of stable versus unstable data points within a cut-section in the sampling region. Based on 868 distinctive data points, the isolated region statistically consists of about 84.7% of stable perovskites, and therefore constitutes the most-likely region for producing stable perovskites. Prior to SLERP however, the unstable points within the region of interest are discarded, and stable perovskite points are crosschecked to ensure they maintain their stability even after decoding.

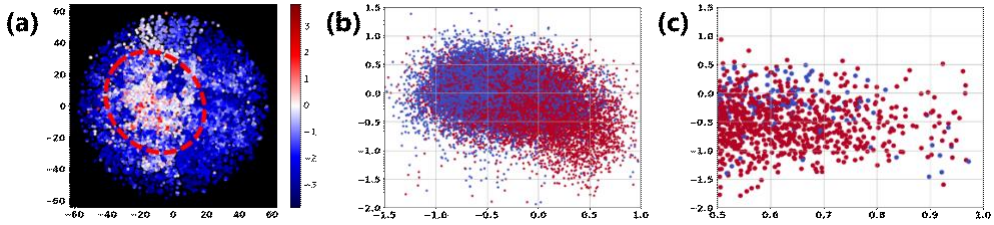

Figure 4. Target latent space: (a) Compressed t-SNE PCA; (b) Real hyperdimensional plane of latent space; (c) Region of interest in real latent space. For Figs. (b) & (c) only, red dots and blue dots are stable and unstable perovskites, respectively.

With the implementation of SLERP, each stable data point is iteratively interpolated against one another to produce $(\Delta t^{-1} - 1) \times \mathbb{C}_2^Z$ new data points, where Δt is the finite spherical spacing or interval for creating new data points (i.e. line-space $\boldsymbol{t} \in [\mathbf{0}, \mathbf{1}]$) and $Z = \{\boldsymbol{z_i}\}_{\boldsymbol{i=1}}^{\boldsymbol{n}}$ are the *stable* sampling points. Moreover, the main challenge between interpolation and decoding is the possibility of generating perovskites that have overlapping atomic configuration in a unit cell. The present study tackles this challenge by comparing all decoded data points with some proven perovskite configurations. Using a similarity test, all decoded chemical compounds are checked with three standard structures to confirm if their unit cell geometry are in close agreement with standardized perovskites. The similarity tests compares the relative position differences in atomic coordinates between decoded perovskites and conventional standard perovskite forms. The standard perovskites chosen for similarity analysis include trigonal $NdAlO_3$ (ICSD ID: 10333), tetragonal $NaNbO_3$ (ICSD ID: 23563), and tetragonal $CaTiO_3$ (ICSD ID: 162919) [22]. As a result, decoded interpolated points with atomic configurations far away from standards are screened out and not considered further for DFT analysis. A potential downside with the TL-VAE model is that the similarity analysis is performed after the decoding process. To address this limitation, future studies will aim at directly integrating the similarity analysis within the inverse pipeline in order to make it more efficient for geometrical screening and pre-DFT calculations.

4. Discussion

4.1. First principle validation using the Density Functional Theory (DFT)

Using first-principle DFT technique, novel perovskites emerging from the inverse design (TL-VAE) pipeline are validated to ascertain if they meet formability. In general, DFT techniques are computational quantum-mechanical modeling methods that aim at solving the Schrödinger equation for obtaining ground-state electronic properties of many-body systems. As such, perovskites that successfully undergo full geometrical DFT relaxation or optimization can be recommended for further scientific investigation and potential synthesization. The present study applies standard DFT practice by using the Quantum Espresso (QE) software package [15]. For all newly generated perovskites, QE performs plane-wave spin-polarized Perdew-Burke-Ernzerhof (PBE) [23] DFT calculations using projector-augmented wave (PAW)-PBE pseudopotentials [24]. For relaxing atomic and unit cell properties, the convergence criteria for energy and force are set at $1.0e^{-7}$ and $1.0e^{-3}$, respectively. A relatively sparse K-point grid mesh size is used for sampling the primary Brillouin zones in the reciprocal space of the crystal lattice. As such, Fig. 5 reveals ball-and-stick symmetrical chemical images of the four most promising candidates generated in this study. The candidates successfully underwent DFT validation based on total energy and force convergence, as well as predicted stability. The promising candidates are determined to be *$AlPtS_3$*, *$AlPtO_3$*, *$GaOPd_3$*, and *$GaPdO_3$*. Comparison with the initial dataset reveals $AlPtS_3$ and $GaOPd_3$ as unique, whereas $AlPtO_3$ and $GaPdO_3$ are polymorphic duplicate variants. This suggests that the developed TL-VAE model can equally generate both unique and non-unique perovskite compounds, depending on the application or functionality of interest. Interestingly, the most promising candidates are seen to coincide with A-sites Aluminum (Al) and Gallium (Ga) perovskites, and B-sites Platinum (Pt) and Palladium (Pd) perovskites, according to the ABX_3 distinctive ionic sites in a three dimensional polyhedral. This is a consequence of the sampling region isolated in the SLERP process, as most perovskite structures that are encompassed within this region are predominately of a similar chemical combination. Besides, Al and Ga metals have the same group numbers (i.e. group number 13) and the same number of valence electrons in their outermost shells (+3 ionic charge). Likewise, Pt and Pd share the same group number on the periodic table of elements (i.e. group number 10). Moreover, both group number and valence of constitutive chemical elements were used in the descriptor engineering process (i.e. for computing the direct discretized

features and the Fourier-transformed reciprocal features). Therefore, the results suggest that the current TL-VAE model does not only perform random chemical or ionic substitutions, but also intuitively searches for new materials that have similar chemical traits. As a result, generating other diverse forms of the perovskite material class will involve modified operations that possibly promote the isolation of targeted chemical combinations or properties of interest. Furthermore, it is worth mentioning that the TL-VAE model can equally search for uncommon forms of interesting ABX_3 stoichiometries such as the $GaOPd_3$ anti-/inverse-perovskite. Chemically, inverse perovskites are of profound interest to the material science community, thanks to their eccentric behavior for applications related to diverse engineering functionalities [25]. In summary, Table 2 outlines the most important crystallographic features of the generated perovskite candidates, as well as the DFT-determined energy bandgap and predicted formation energy of the optimized structures. The predicted formation energy values were determined using the initially saved modeling weights, as previously applied in assessing the forward-mapping efficiency of the developed descriptor (i.e. Fig. 3(a)). Taking into account a safe buffer threshold of $\boldsymbol{E_f} \leq 0.5$ eV/atom, it can thus be inferred that all predicted $\boldsymbol{E_f}$ values are within metastability. Besides, the safe buffer region is a tolerable estimate in order not to neglect any intriguing, but potentially stable perovskite [3], notwithstanding the possibility of inherent errors that may be associated with the DFT software [1]. All newly discovered perovskite compounds consist of ten atoms in their conventional unit cells (i.e. Z=2). The Crystallographic Information Files (CIF) and other supplementary information, as related to the new compounds, are openly made available on our GitHub website (see data availability statement).

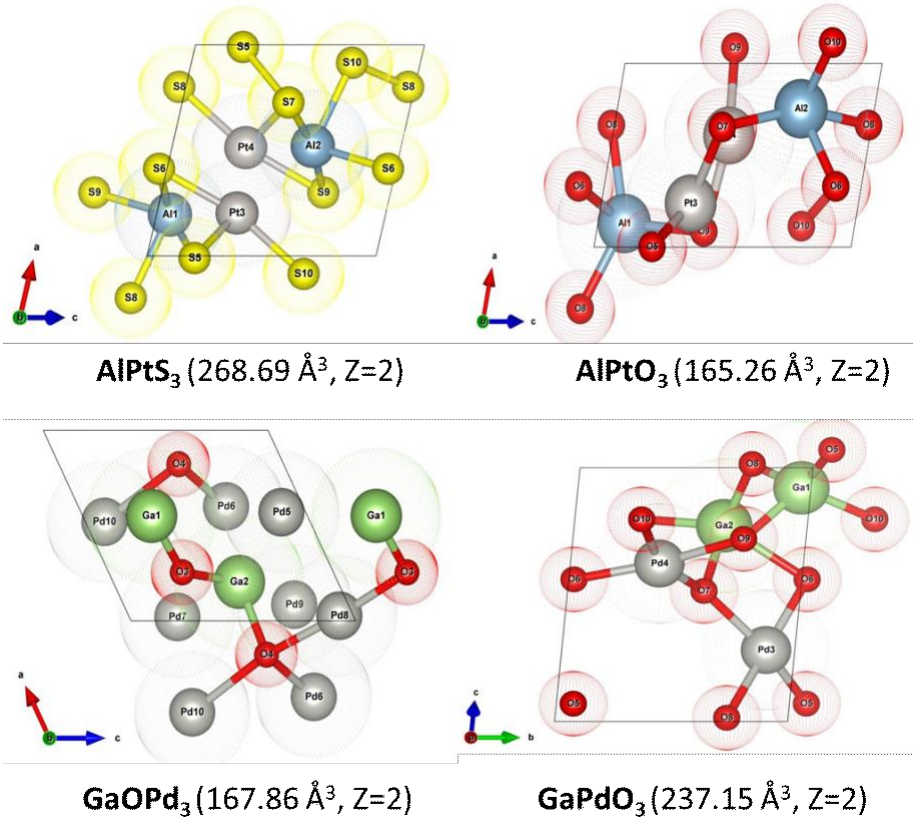

Figure 5. DFT verified perovskite candidates emerging from the TL-VAE pipeline. In brackets are the unit cell volume (in cubic-Angstroms, Å^3) and the number of formula units, Z.

Table 2. Crystallographic features and properties of the generated novel perovskites.

Perovskite	A (Å)	B (Å)	C (Å)	Alpha (°)	Beta (°)	Gamma (°)	Predicted formation energy (eV/atom)	Energy Bandgap (eV)
$AlPtS_3$	5.706	5.879	8.010	132.57	72.13	102.45	0.009	0.000
$AlPtO_3$	3.653	6.749	6.703	124.72	74.93	101.10	-0.607	0.304
$GaOPd_3$	4.906	6.333	5.402	80.56	117.18	103.55	-0.575	0.000
$GaPdO_3$	5.178	5.635	8.129	68.15	134.30	116.48	-0.854	0.106

5. Conclusion

The current study demonstrates an inverse design pipeline for the discovery of novel stable ternary perovskites. The TL-VAE model explores the latent space of a Variational Autoencoder (VAE) in combination with Target-Learning (TL) feed-forward neural networks. The TL networks effectively learns the formation energy target in hyperdimension using both regression and classification loss metrics. As a result, the latent space is demonstrated to be segregated into two regions that qualify stability, thereby enabling a smooth sampling operation for generating new data points. The new data points are obtained by implementing Spherical Linear Interpolation (SLERP) and are further screened by comparing their decoded atomic configurations to standards (i.e. similarity test). To verify if potential perovskites from the TL-VAE pipeline are feasible materials, quantum-mechanical Density Functional Theory (DFT) simulation is implemented on some of the emerging candidates. Using the Quantum Espresso (QE) DFT software, the study reveals four novel perovskites that showed full geometrical relaxation. The promising candidates include $AlPtS_3$, $AlPtO_3$, $GaOPd_3$, and $GaPdO_3$, and they are all successfully predicted to be metastable. Overall, the study illustrates the potential of the TL-VAE perovskite generator and open up novel modeling opportunities for multi-objective latent-space optimization, which is the area of future studies.

Data Availability

The source code used to develop the TL-VAE model, in addition to the crystallographic information files (CIF) of the newly generated perovskites, are made available at: https://github.com/chenebuah/TL-VAE

Acknowledgements

This research was supported by the National Research Council of Canada (NRC) through its Artificial Intelligence for Design Program led by the Digital Technologies Research Centre.

References

[1] A.J. Cohen, P. Mori- Sánchez, W. Yang, Challenges for Density Functional Theory, Chem. Rev., 112(1), (2012), 289-320. https://doi.org/10.1021/cr200107z

[2] E.T. Chenebuah, M. Nganbe, A.B. Tchagang, Comparative analysis of machine learning approaches on the prediction of the electronic properties of perovskites: A case study of ABX3 and A2BB'X6, Mater. Today Commun., 27, (2021), 102462, https://doi.org/10.1016/j.mtcomm.2021.102462

[3] J. Noh, J. Kim, H.S. Stein, et al., Inverse Design of Solid-State Materials via a Continuous Representation, Matter, 1(5), (2019), 1370-1384. https://doi.org/10.1016/j.matt.2019.08.017

[4] Z. Ren, et al., An invertible crystallographic representation for general inverse design of inorganic crystals with targeted properties, Matter, 5(1), (2022), 314-335. https://doi.org/10.1016/j.matt.2021.11.032

[5] 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, 6, 84, (2020). https://doi.org/10.1038/s41524-020-00352-0

[6] T. Long, N.M. Fortunato, I. Opahle, et al., Constrained crystals deep convolutional generative adversarial network for the inverse design of crystal structures, npj Comput. Mater, 7, 66, (2021). https://doi.org/10.1038/s41524-021-00526-4

[7] K. Kamnitsas, D.C. Castro, L. Le Folgoc, et al., Semi-Supervised Learning via Compact Latent Space Clustering, arXiv:1806.02679v2 [cs.LG], (2018). https://doi.org/10.48550/arXiv.1806.02679

[8] J. Vamathevan, D. Clark, P. Czodrowski, et al., Applications of machine learning in drug discovery and development. Nat Rev Drug Discov 18, 463–477 (2019). https://doi.org/10.1038/s41573-019-0024-5

[9] N.S. Kumar, K.C.B. Naidu, A review on perovskite solar cells (PSCs), materials and applications, J Materiomics, 7(5), (2021), 940-956. https://doi.org/10.1016/j.jmat.2021.04.002

[10] K. Uchino, Glory of piezoelectric perovskites, Sci. Technol. Adv. Mater. 16, (2015), 046001. http://doi.org/10.1088/1468-6996/16/4/046001

[11] M. Johnsson, P. Lemmens, Crystallography and Chemistry of Perovskites. In Handbook of Magnetism and Advanced Magnetic Materials (eds H. Kronmüller, S. Parkin, M. Coey, A. Inoue and H. Kronmüller), (2007). https://doi.org/10.1002/9780470022184.hmm411

[12] J.E. Saal, S. Kirklin, M. Aykol, et al., Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD), JOM, 65, (2013), 1501-1509. https://doi.org/10.1007/s11837-013-0755-4

[13] A. Jain, S.P. et al., The Materials Project: A materials genome approach to accelerating materials innovation, APL Materials, 1(1), (2013), 011002. https://doi.org/10.1063/1.4812323

[14] A. Emery, C. Wolverton, High-throughput DFT calculations of formation energy, stability and oxygen vacancy formation energy of ABO3 perovskites. Sci. Data, 4, 170153 (2017). https://doi.org/10.1038/sdata.2017.153

[15] P. Giannozzi, et al., QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials, J. Phys.:Condens. Matter, 21(39), 395502, (2009). https://doi.org/10.1088/0953-8984/21/39/395502

[16] S.H. Simon, The Oxford Solid State Basics, Oxford University Press, (2013).

[17] E.T. Chenebuah, M. Nganbe, A.B. Tchagang, A Fourier-transformed feature engineering design for predicting ternary perovskite properties by coupling a two-dimensional convolutional neural network with a support vector machine (Conv2D-SVM), Mater. Res. Express., (2023). https://doi.org/10.1088/2053-1591/acb683

[18] D.P. Kingma, M. Welling, An Introduction to Variational Autoencoders, arXiv:1906.02691v3 [cs.LG], (2019). https://doi.org/10.48550/arXiv.1906.02691

[19] D.P. Kingma, N. Welling, Auto-Encoding Variational Bayes, arXiv:1312.6114v11 [stat.ML], (2013). https://doi.org/10.48550/arXiv.1312.6114

[20] K. Shoemake, Animating rotation with quaternion curves, SIGGRAPH Comput. Graph, 19(3), (1985), 245–254. https://doi.org/10.1145/325165.325242

[21] L. van der Maaten, G. Hinton, Visualizing Data using t-SNE, JMLR, 9, (2008), 2579-2605.

[22] A. Belsky, M. Hellenbrandt, V.L. Karen, P. Luksch, New developments in the Inorganic Crystal Structure Database (ICSD): accessibility in support of materials research and design, Acta Cryst., B58, (2002), 364-369. https://doi.org/10.1107/S0108768102006948

[23] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. Phys. Rev. Lett., 77(18), (1996), 3865-3868. https://doi.org/10.1103/PhysRevLett.77.3865

[24] P.E. Blöchl, Projector augmented-wave method. Phys. Rev. B 50(24), (1994), 17953-17979. https://doi.org/10.1103/PhysRevB.50.17953

[25] J. Zheng, B. Perry, Y. Wu, Antiperovskite Superionic Conductors: A Critical Review, ACS Mater. Au., 1(2), (2021), 92-106. https://doi.org/10.1021/acsmaterialsau.1c00026