Article | Received xx; Accepted xx; Published xx
https://doi.org/xx/xx

Supporting Information

Data-efficient and interpretable inverse materials design using a disentangled variational autoencoder

Cheng Zeng[1,2,*,†], Zulqarnain Khan[1,*,†], and Nathan L. Post[2]

[1]Institute for Experiential AI, Northeastern University, Boston, MA 02115, United States
[2]The Roux Institute, Northeastern University, Portland, ME 04101, United States

† These authors contribute equally: Cheng Zeng, Zulqarnain Khan.

* Correspondence author; E-mail: c.zeng@northeastern.edu, z.khan@northeastern.edu

This Supporting Information (SI) includes a case study for design on multiple materials properties.

S-1 A workflow to design high-entropy alloys with multiple properties using disentangled variational autoencoder

Although the method is only demonstrated for single property in this work. It can be readily extended to multiple materials properties because of the disentanglement of target properties from the latent space. Here we present a scenario where we could use the methodology for multiple materials properties. Note that the exact demonstration for tuning multiple properties is not available until all relevant datasets are curated.

The inverse design of high-entropy alloys targeting three materials properties is shown in Figure S1. We suppose we would like to design high-entropy alloys with desirable phase formation, surface energies and stacking fault energies. For the phase formation, we could use the experimental data as used in this work. For the surface energies and stacking fault energies, we could use machine learning accelerated high-throughput atomistic simulations to create relevant configurations and obtain the energetics. The core idea is similar, but this time we have another two independent routes for the respective two additional materials properties. On the generative model, the new alloy will be generated given the latent space and the three materials properties. On the recognition model, the latent space will be generated conditioned on three materials properties. Since we anticipate the surface energy and stacking fault energy may be more dependent on composition itself, the alloy composition can be directly related to those two properties, rather than going through a feature engineering step, e.g. f(x) for single phase formation.

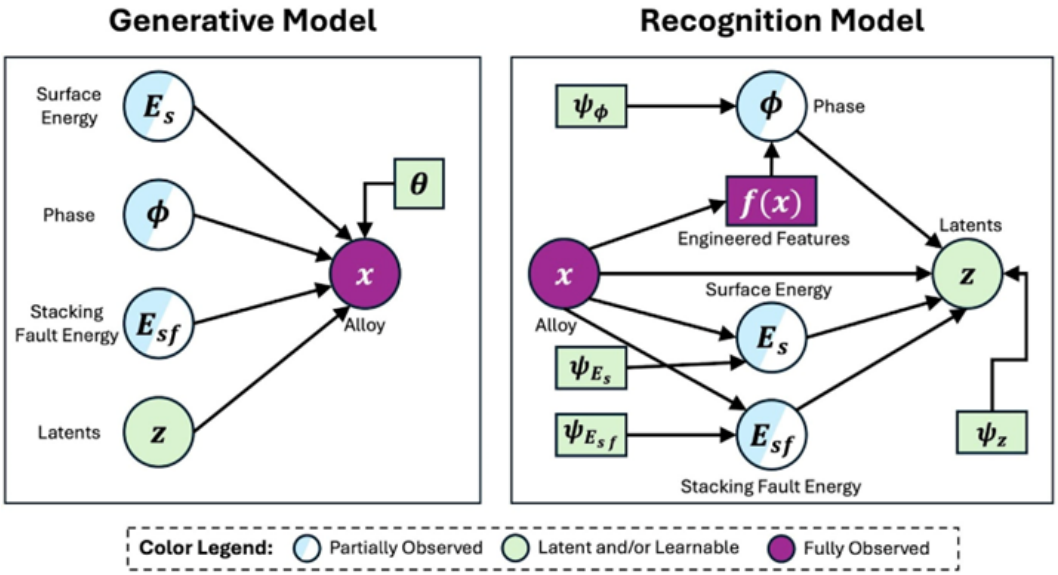

Figure S1. A case study for design of high-entropy alloys with three target properties, including single-phase formation, surface energies and stacking fault energies.