Edge Based Architecture for Total Energy Regression Models for Computational Materials Science

Kangmo Yeo[1], Sukmin Jeong[2(✉)], and Soo-Hyung Kim[1(✉)]

[1] Department of Artificial Intelligence Convergence, Chonnam National University, Gwangju 61186, Republic of Korea
shkim@jnu.ac.kr

[2] Department of Physics and Research Institute of Physics and Chemistry, Jeonbuk National University, Jeonju 54896, Republic of Korea
jsm@jbnu.ac.kr

Abstract. It is widely acknowledged that artificial intelligence (AI) technology has been extensively applied and has achieved remarkable advancements in various fields. The field of computational materials science has also embraced AI techniques in diverse ways. Today, computational materials science plays a crucial role in the development of cutting-edge materials, including pharmaceuticals, catalysts, semiconductors, and batteries. One significant task in this field is the regression of the total energy of atomic structures that form various materials. In this study, we propose a modified model architecture aimed at improving the performance of existing total energy regression models. Traditional total energy regression models calculate the total energy by summing the energies of individual nodes represented in the atomic structure graph. However, our approach suggests a modified architecture that not only predicts the energy for nodes but also incorporates energy prediction for edges in the graph. This novel architecture achieved a 3.9% reduction in energy error compared to the base model. Moreover, its simplicity provides the advantage of general applicability to other total energy regression models.

Keywords: Computational materials science · Atomic structure

1 Introduction

Artificial intelligence (AI) has already demonstrated its powerful performance and efficiency through numerous examples. In the field of natural science, AI is rapidly improving efficiency, comparable to the fourth wave. One area where AI algorithms have been introduced and are making a significant impact is computational materials science [1–5]. This field, which primarily simulates the properties and phenomena of materials, is gaining attention across various industries, including semiconductors, batteries, light-emitting devices, chemistry, new drug development, catalysts, and solar cells [6–14].

H. Lu et al. (Eds.): ACPR 2023, LNCS 14408, pp. 106–112, 2023.
https://doi.org/10.1007/978-3-031-47665-5_9

Density Functional Theory (DFT) [15,16], proposed in the 1970s, boasts surprisingly high accuracy and has become a major method in the field of computational science. However, DFT methods require large computational resources at the cost of high accuracy [10,17,18]. Recently, AI technology has been rapidly introduced to solve these problems and has achieved considerable results. However, it still lacks accuracy compared to the DFT method [19].

A fundamental and central task in computational materials science is to calculate the total energy for a given atomic structure. A total energy regression AI model called machine learning potential is based on calculating the total energy by predicting and summing the contributions of each target atom to the total energy [20–24]. This method has a similar structure to the empirical potential, which is a traditional method of calculating total energy.

We propose a new architecture that considers both atoms and bonds. This approach is more physically intuitive as the total energy of a material consists of both the energy of the atoms themselves and the energy of chemical bonds. This architecture is more natural and similar to DFT calculations. If the architecture of the AI model is similar to how DFT calculations work, then it can be expected that the AI model will learn more easily. This effect is especially effective in non-metallic materials with clear chemical bonds rather than metallic materials with ambiguous bonds. Any AI model that interprets atomic structure as a graph and calculates total energy can benefit from this approach and improve its performance.

2 Background and Base Model

Machine Learning Potential. In 2007, Behler and Parrinello first proposed an AI architecture for regressing total energy based on atomic positions [20]. Density functional theory-based first-principles calculations are one of the most widely used methodologies in computational materials science and can calculate various properties for materials. Total energy is one of the most basic and essential physical quantities in property calculations. Total energy is a kind of function that takes atomic structure as a variable, meaning the type and arrangement of atoms that make up an arbitrary material. For example, an atomic structure containing N atoms is expressed as an atomic number $\mathbf{Z} = \{Z_1, Z_2, Z_3, ..., Z_N\}$ and each atom's position $\mathbf{R} = \{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, ..., \mathbf{r}_N\}$ where $\mathbf{r}_i$ is a Cartesian coordinate $\mathbf{r}_i = \{x_i, y_i, z_i\}$. The traditional computational chemical methodology of empirical potential was adopted. According to this methodology, total energy E_{total} is

$$E_{total} = \sum_i E_i \tag{1}$$

where E_i is the contribution of the i_{th} atom to total energy. The architecture of machine learning potential is shown in Fig. 1. Almost all machine learning potentials published so far have a structure like this. In Fig. 1, G_i^{φ} is an input vector of size φ transformed by symmetric function.

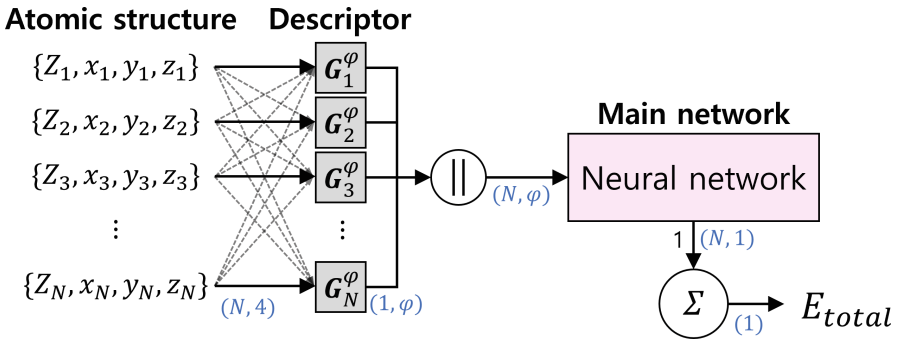

Fig. 1. Basic architecture of a machine learning potential [20]. The gray and pink boxes represent the descriptor and main network, respectively. All black and dotted arrows indicate the flow of data. The round circle represents an operation that calculates the sum of input values. The data shape at each step is indicated by blue numbers. (Color figure online)

Base Model. The model proposed in this study is based on GemNet-OC [25] and uses a similar architecture. GemNet-OC is a graph neural network (GNN) that represents an atomic system as a graph G = (V, E), where the set of graph nodes, V, represents each atom and the set of edges, E, is defined as all pairs of atoms within a certain cutoff distance. The first model, similar to state-of-the-art GNNs, was proposed in 1997 but only gained popularity after several works demonstrated its potential for a wide range of graph-related tasks [26]. GemNet-OC evolves the two-level message passing scheme proposed in MXMNet into an interaction layer and utilizes both edge and node embeddings. GemNet-OC is a model based on Geometric Message Passing Neural Networks (GemNet) [27], which uses a similar architecture and improves the accuracy of forces experienced by atoms.

3 Datasets

We used the Open Catalyst 2022 (OC22) dataset for model training [19]. The OC22 dataset is designed to enable the development of generalizable machine learning (ML) models for catalysts, particularly for oxygen and hydrogen evolution reactions and oxide electrocatalysis. The dataset consists of oxide surface structures combined with constituent elements and oxide surface structures with adsorbed molecules, as well as defects such as atomic substitution and vacancies. By providing a diverse and representative training dataset, OC22 aims to support the development of generalized models that can accurately predict catalytic reactions on oxide surfaces. Models generated from this dataset are expected to accelerate the discovery and design of new catalysts for a wide range of applications.

The primary task of OC20 is to regress the total energy obtained through first-principles calculations based on DFT from the atomic structure. The

dataset is divided into training/validation/test sets and each set includes both material surface structures and surface structures with adsorbed molecules. The dataset includes 19,142 material surface structures and 43,189 surface structures with adsorbed molecules, with a total of 9,854,504 data points. Diversity in surface structure and adsorption structure was prioritized when constructing the dataset to ensure that a generalized model can be built.

4 Edge Based Architecture of GemNet-OC

GemNet-OC is a graph neural network (GNN) that represents atomic systems as graphs, with its architecture being improved to map the energy of the edge embedding using GemNet as a base model. In this architecture, nodes and edges are embedded respectively, with the edge embedding being used to regress the force received by atoms. An architecture similar to the empirical potential for the total energy was proposed, as shown in Eq. (1) and Fig. 1.

However, the total energy can also be described in terms of heat statistics. Specifically, it can be expressed as

$$E_{total} = \Omega + \sum_i \mu_i \tag{2}$$

where Ω is the formation energy representing chemical interaction between atoms and μ_i is the energy of one atom in terms of thermostatistics. The formation energy Ω can be further expressed as $\Omega = \sum_m e_m$, which is the sum of the binding energies e_m of atomic pairs. The binding energies e_m of atomic pairs are mapped to the values from the edges of the atomic structure graph by the main network.

This structure has several advantages. Firstly, μ_i is more consistent about the placement of atoms than E_i, making the model easier to train. Secondly, the total energy naturally regresses by the formation energy calculated as the sum of the binding energies. To reflect these formulas, the architecture was modified to map energy to edges as well. The modified architecture can be seen in Fig. 2.

5 Results

Due to limited computational resources, only 200,000 training data points (1/40 of the total OC22 dataset) were used for training over 9 epochs. However, since the same dataset was used for all models being compared and 200,000 is still a large number of data points, it is still meaningful to test the performance of the models. The training and validation errors were reduced equally for all models, indicating that there was no underfitting or overfitting (Fig. 3). The inset of Fig. 3-b shows that the validation error of GemNet-EB (shown in red) is about 3.9% smaller than that of the base model GemNet-OC (shown in green).

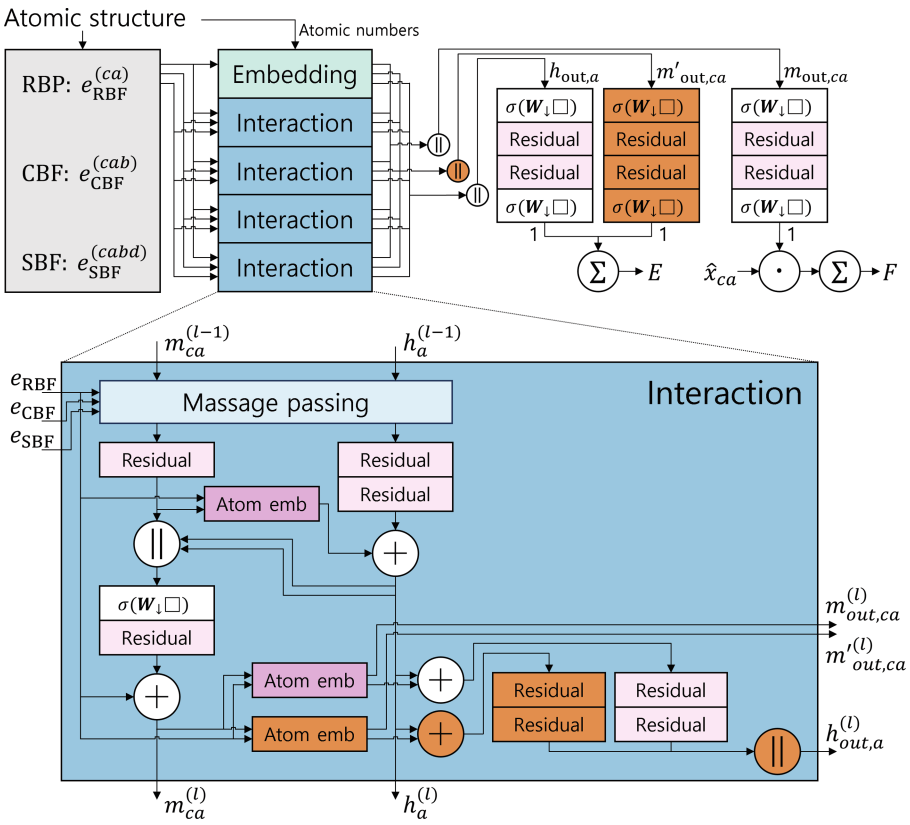

Fig. 2. Main network of the GemNet-EB architecture. Changes are highlighted in orange. $\square$ denotes the layer's input, || concatenation, σ a non-linearity. The massage passing block and Embedding block have same architecture with the GemNet-OC [25]. (Color figure online)

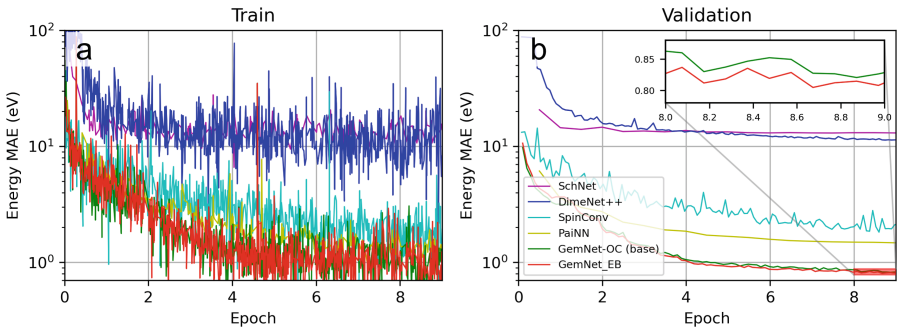

Fig. 3. Energy mean absolute error (MAE) of train (a) and validation (b) respectively. The x-axis represents the epoch and the y-axis represents the MAE on a log scale. The inset in (b) is a zoom-in of the red box zone at the bottom right. (Color figure online)

6 Conclusion

We proposed a new and improved architecture for regressing the total energy from atomic structures. Our newly proposed GemNet-EB model achieved a 3.9% lower validation error than the base model. Of course, other experiments are possible, but further studies are needed because it is difficult to test more than this due to the limitation of computer resources. However, this study is still significant because we applied the new concept of the total energy prediction model, as shown in Eq. (2), and achieved a low validation error rate with this method. There is a significant difference in errors between the other models and the GemNet base model, indicating that the edge embedding of the GemNet-OC model also plays an important role in total energy regression. By directly mapping bond energy using edge embedding, we were able to improve the accuracy of the model. While edge embedding is indirectly reflected in node embedding through the interaction block, we were able to improve performance by directly connecting it to total energy regression.

However, considering that the accuracy of DFT calculations used for surface structure and surface adsorption energy studies is less than 0.01 eV, there is still room for improvement. Since our new method is not limited to any specific model or architecture and can be applied to any GNN-based model for atomic structures, our proposed new architecture can serve as a foundation for advancing machine learning potential.

Acknowledgement. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1I1A3A04036408) and also supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-02068, Artificial Intelligence Innovation Hub).

References

1. Butler, K.: Machine learning for molecular and materials science. Nature **559**, 547–555 (2018)
2. Schmidt, J.: Recent advances and applications of machine learning in solid-state materials science. npj Comput. Mater. **5**, 83 (2019)
3. Wei, J.: Machine learning in materials science. InfoMat **1**, 338–358 (2019)
4. Morgan, D.: Opportunities and challenges for machine learning in materials science. Ann. Rev. **50**, 71–103 (2020)
5. Fiedler, L.: Deep dive into machine learning density functional theory for materials science and chemistry. Phys. Rev. Mater. **6**, 040301 (2022)
6. Ladha, D.: A review on density functional theory-based study on two-dimensional materials used in batteries. Mater. Today. Chem. **11**, 94–111 (2019)
7. Adekoya, D.: DFT-guided design and fabrication of carbon-nitride-based materials for energy storage devices: a review. Nanomicro Lett. **13**, 13 (2021)
8. Liang, Q.: Transition metal compounds family for Li-S batteries: the DFT-guide for suppressing polysulfides shuttle. Adv. Funct. Mater. **33**, 2300825 (2023)

9. Neugebauer, J.: Density functional theory in materials science. WIREs Comput. Mol. Sci. **3**, 438–448 (2013)
10. Mattsson, A.: Designing meaningful density functional theory calculations in materials science-a primer. Model. Simul. Mat. Sci. Eng. **13**, R1 (2005)
11. Tandon, H.: A brief review on importance of DFT in drug design. Res. Med. Eng. Sci. **7**, 791–795 (2019)
12. Sade, V.: Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: a review. Eur. J. Med. Chem. **224**, 113705 (2021)
13. Weijing, D.: The application of DFT in catalysis and adsorption reaction system. Energy Procedia **152**, 997–1002 (2018)
14. Liao, X.: Density functional theory for electrocatalysis. Energy Environ. Mater. **5**, 157–185 (2021)
15. Hohenberg, P.: Inhomogeneous electron gas. Phys. Rev. **136**, B864 (1964)
16. Kohn, W.: Self-consistent equations including exchange and correlation effects. Phys. Rev. **140**, A1133 (1965)
17. Zhang, L.: Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. Phys. Rev. Lett. **120**, 143001 (2018)
18. Chandrasekaran, A.: Solving the electronic structure problem with machine learning. npj Comput. Mater. **5**, 22 (2019)
19. Tran, R.: The Open Catalyst 2022 (OC22) dataset and challenges for oxide electrocatalysts. ACS Catal. **13**(5), 3066–3084 (2023)
20. Behler, J.: Generalied neural-network representation of high-dimensional potential-energy surface. Phys. Rev. Lett. **98**, 146401 (2007)
21. Bartók, A.: Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. Phys. Rev. Lett. **104**, 136403 (2010)
22. Artith, N.: An implementation of artificial neural-network potentials for atomistic materials simulations: performance for TiO_2. Comput. Mater. Sci. **114**, 135–150 (2016)
23. Schütt, K.: SchNet - a deep learning architecture for molecules and materials. J. Chem. Phys. **148**, 24 (2018)
24. Lee, K.: SIMPLE-NN: an efficient package for training and executing neural-network interatomic potentials. Comput. Phys. Comm. **242**, 95–103 (2019)
25. Gasteiger, J.: GemNet-OC: developing graph neural networks for large and diverse molecular simulation datasets. arXiv preprint arXiv:2204.02782 (2022)
26. Zelinsky, N.: A neural device for searching direct correlations between structures and properties of chemical compounds. J. Chem. Inf. Comput. Sci. **37**, 715–721 (1997)
27. Gasteiger, J., Becker, F., Günnemann, S.: GemNet: universal directional graph neural networks for molecules. In: 35th Conference on Neural Information Processing Systems, vol. 34, pp. 6790–6802 (2021)