



## Graph-based deep learning frameworks for molecules and solid-state materials

Weiyi Gong, Qimin Yan\*

*Department of Physics, Temple University, Philadelphia, PA 19122, USA*



### ARTICLE INFO

**Keywords:**

Graph neural network  
Deep learning for materials science  
Machine learning

### ABSTRACT

Recent years have witnessed the rapid increase in the application of deep learning in atomistic systems including both molecules and solid-state materials. The use of graphs and associated design of message passing strategies have enabled multiple deep learning frameworks to achieve reliable and efficient predictions of materials properties with a much smaller cost compared with first-principles atomistic simulations. In this review, we will focus on recent development of graph-based deep learning frameworks and their applications for both molecules and solid-state material systems. The history of the development of graph-based representations for molecules and crystals will be introduced. Essential learning processes defined by the so-called message passing will be reviewed, based on which the performance of different models will be compared. Furthermore, recent development of graph learning frameworks that incorporate material information beyond atom level will be introduced. Current challenges and future perspectives on this emerging field at the crossroad of material science, physics, chemistry, and computer science will be given, with an emphasize on how multiple tiers of material information can be organized and combined in a graph neural network setup.

### 1. Introduction

Machine learning (ML) techniques [1] provide a novel opportunity to significantly reduce computational costs and speed up the pace of materials discovery and design by utilizing data-driven paradigms [2]. Combined with data-driven technologies, ML has become a powerful tool in materials research [3,4]. Shallow ML models have been adopted to predict a large set of material properties, including phase stability [5], crystal structure [6], electronic structures (such as band gaps) [7], atomization energies [8], as well as effective potentials for molecule dynamics simulations [9]. To facilitate the effective learning of molecules and solid-state systems, representations for bonded atomic structures have been developed in recent years, such as Coulomb matrix [10], bag of bonds [11], and many-body tensor based representation [12].

A notable limitation of shallow learning models lies in their heavy reliance on designing hand-crafted features based on domain knowledge. The renaissance of artificial intelligence (AI) is largely driven by the breakthrough in feature learning (or representation learning) boosted by deep learning techniques [13]. Such learning paradigm enables automatic model parameter estimation (i.e., end-to-end learning) and frees us from tedious and highly non-trivial feature designing

process. In recent years, with the aim to avoid the process of explicit feature design, deep learning has been applied widely to both molecules and solid-state material systems [14,15].

Moving from shallow to deep learning of atomistic systems, the concept of graph has emerged as a material information collector and communicator that synthesizes atomistic systems and deep neural networks. Taking advantage of bonding connections between atoms and many-particle interactions beyond that, graphs are a natural representation for both molecules and crystals. From this point of view, quantum chemistry and solid-state physics are merging in the deep learning paradigm. Graph-based deep learning frameworks have been applied to both molecules and solid-state crystalline material systems and showed promise compared with shallow learning [16].

This review will cover the development of deep learning frameworks that take advantage of atom-based graph data for both molecules and solid-state crystalline systems. Especially, we will discuss multiple message passing algorithms and aggregation procedures that generates the functions of the entire input molecule or crystal graphs. We will introduce the recent development of graph learning frameworks that incorporate material information beyond atom level. Current challenges and future perspectives on this emerging field will be given, with an

\* Corresponding author.

E-mail address: [qiminyan@temple.edu](mailto:qiminyan@temple.edu) (Q. Yan).

emphasize on how multiple tiers of material information can be organized and combined in a graph neural network setup.

## 2. Graphs for atomistic systems and graph-based deep learning

A key bottleneck limiting the applications of traditional deep learning techniques, such as convolutional neural networks (CNN) [17], for materials science is the great diversity in atomic systems that cannot be described by rigid grids. CNNs have shown promising achievements in the property predictions for small molecules and orthogonal-grid systems. However, for nonorthogonal grids and large molecules, CNN-based methods are found to be insufficient. The existence of bonding structures and many-particle interactions in atomic systems naturally encourages the use of graph representations. Graphs for atomic systems was first introduced in the training of small molecules. Before the emergence of graph-based learning frameworks, most ML models can only handle inputs of a fixed size. Off-the-shelf fingerprints [18] are used to compute fixed-dimensional feature vectors as inputs to deep neural networks or other standard ML methods. In the work by D. Duvenaud *et al.*, the function that computes molecular fingerprint vectors was replaced by a convolutional neural network whose input is a graph representing the original molecule [18]. It is shown that the effective features extracted from graphs are more interpretable and the well-trained graph neural networks (GNNs) can achieve outstanding predictive performance on a variety of tasks.

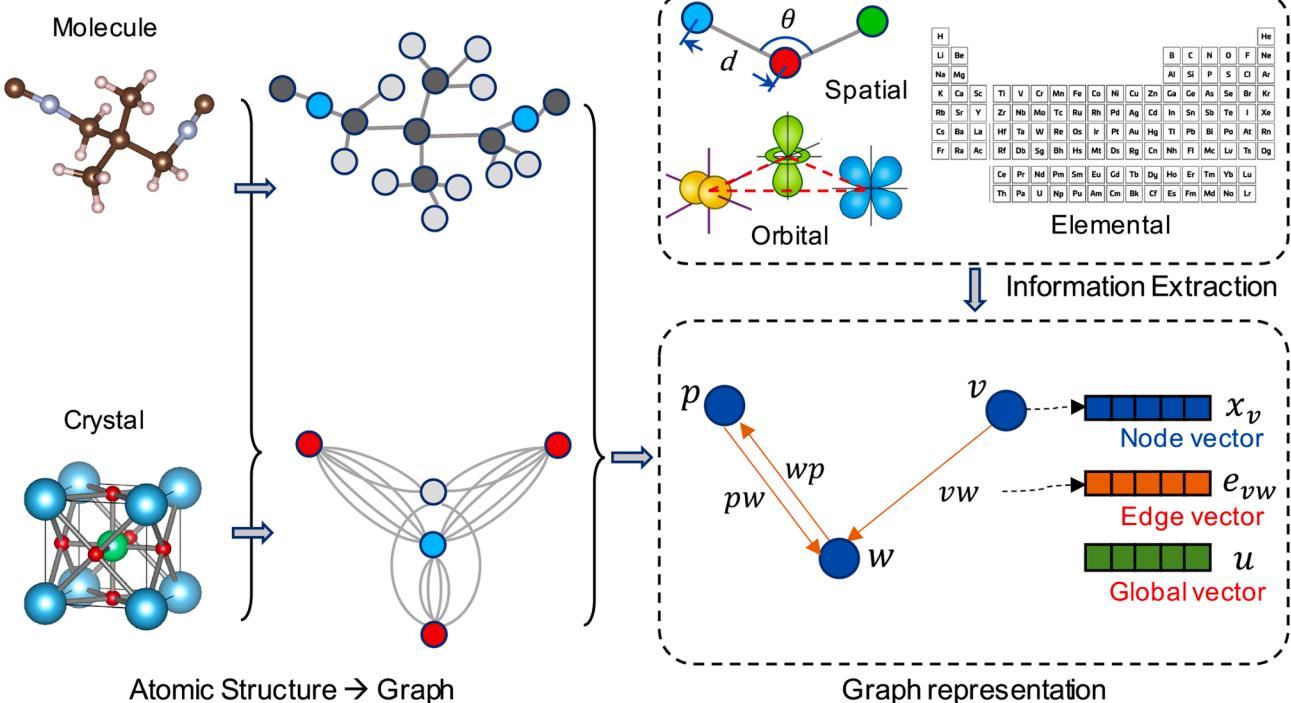
Next we will briefly introduce the graph concept used in the ML models in this review. Generally, a molecule or crystal can be represented by a graph  $G$ , where individual atoms are represented as nodes and bonds are presented as edges, either directed or undirected. The bonds are not necessarily chemical bonds but can be any interaction between two atoms and the choice of bonds depends on the learning task. As shown in Fig. 1, nodes are embedded in a graph (for either a molecule or a crystal) as vectors  $x_v$ , while edges are embedded as vectors  $e_{vw}$ . An optional global environment can be embedded as a vector  $u$ . Information exchange can be repeatedly applied to the graph

representation, leading to updated states  $x_v^t$ ,  $e_{vw}^t$ , and  $u^t$  at time step  $t$ .

The recent development of graph-based deep learning offers a novel tool [19], when combined with domain knowledges, to create an innovative representation of molecule or crystal structures. One great advantage of GNN is that they allow end-to-end learning based on graphs of arbitrary size and shape, which is helpful to remove roadblocks that impede the wide application of deep learning for materials science. With the emerging GNN, any type of grid and atomic structure can be successfully modeled and analyzed within this framework [20–22]. Furthermore, symmetry information of molecule and condensed matter systems can be easily integrated into the graph convolution. Recent progresses in general GNNs [20,23], where convolutional operations are optional instead of mandatory, bring another potential for improving property predictions of complicated material systems. In the following sections, we will discuss the recent development of graph-based deep learning architectures in two major fields, molecules and solid-state crystalline systems.

## 3. Message passing strategies for molecule graphs

To effectively capture the mapping between molecule structures and properties encoded by the complex interactions within a molecule, individual node and edge information in molecule graphs need to be communicated and extracted by designing the appropriate message, update, and output functions. The so-called message passing is therefore an essential component in a graph-based deep learning architecture [24]. In the work by D. Duvenaud *et al.* [18], the message function is chosen as  $M(x_v^t, x_w^t, e_{vw}) = x_w^t \oplus e_{vw}$ , where  $\oplus$  denotes concatenation. The node update function is  $U_t(x_v^t, m_v^t) = \sigma(H_t^{deg(v)} m_v^t)$ , where  $H_t^N$  is a learned matrix at time step  $t$  and node degree  $N$ ,  $\sigma$  is the sigmoid function and  $deg(v)$  is the degree of node  $v$ . The readout function takes hidden states of all time steps into consideration:  $R = f(\sum_{v,t} \text{softmax}(W_t x_v^t))$ , where  $f$  is a neural network and  $W_t$  is a learned matrix for each time step. The message on each node is the



**Fig. 1.** A molecule or crystal can be represented as a graph  $G$ , where nodes ( $v, w, p$  in the graph) represent atoms in the molecule or crystal cell, and edges ( $vw, wp, pw$  in the graph) represent any interatomic information such distance, bond type, solid angle subtended by the face of Voronoi cell, etc. Nodes, edges and the global state ( $u$  in the graph) of the graph are then embedded into vectors based on multiple levels of information extracted from the materials.

concatenation of summed node and edge features separately:  $m_v^t = (\sum_{w \in N(v)} x_w^t) \oplus (\sum_{w \in N(v)} e_{vw})$  [24]. The limitation of this message passing strategy is that it is unable to incorporate correlations between node states and edge states.

Based on several existing models, Gilmer *et al.* from the Google team proposed a general framework for supervised learning on graphs called Message Passing Neural Networks (MPNNs) that abstracts the commonalities among several existing neural models for graph structured data [23–25]. The MPNNs take advantages of the symmetry of molecule systems and can be designed to be invariant to graph isomorphism when applied to molecule-based graph structured data. MPNNs operates on an undirected graph  $G$  with node features  $x_v$  and edge features  $e_{vw}$ . During the message passing process, a message is computed by applying the message function  $M_t$  on each triad  $(x_v^t, x_w^t, e_{vw})$ . The total message on a node is the aggregation of neighboring messages:  $m_v^t = \sum_{w \in N(v)} M_t(x_v^t, x_w^t, e_{vw})$ . Node features are then updated by an updating function  $U_t$ :  $x_v^{t+1} = U_t(x_v^t, m_v^t)$ . The message passing process is repeated for  $T$  steps and a readout function  $R$  computes a graph feature afterwards:  $\hat{y} = R(\{x_v^T | v \in G\})$ .

The training in the above work was focused on the QM9 dataset [31] which has 134 k molecules. The results showed that MPNNs with the appropriate message, update, and output functions achieved state of the art prediction results on all 13 target molecular properties in 4 general types (see Table 1) and achieved chemical accuracy on 11 out of 13 targets, outperforming several strong baselines and eliminating the need for complicated feature engineering (Table 2).

In another work that moves beyond fingerprints and incorporate molecular graph convolutions [29], the authors proposed a graph convolution called Weave module which satisfies the following properties: (i) the model output is invariant to permutations on the input; (ii) the atom and pair representations (or the values computed from this representation) in all time steps are invariant to permutations that the input are invariant to; (iii) pair representation is invariant to exchange of two atoms. Node and edge features are updated as:

$$x_v^{t+1} = \text{NN}\left(\text{NN}(x_v^t) \oplus \sum_{w \in N(v)} \text{NN}(e_{vw}^t)\right)$$

$$e_{vw}^{t+1} = \text{NN}\left(\text{NN}(e_{vw}^t) \oplus \sum_{\substack{(i,j)=(v,w) \\ \text{or } (w,v)}} \text{NN}(x_i^t \oplus x_j^t)\right)$$

After several convolutions, the final atom features are encoded using the gaussian membership function to get a fuzzy histogram for each feature dimension:

$$x_v^{T+1} = [\dots, f_j(x_{vi}^T - a_j), \dots], a_j = \min, \min + \Delta, \dots, \max$$

where  $x_{vi}$  is the  $i^{\text{th}}$  entry of  $x_v$ ,  $a_j$  are bins from  $\min(x_{vi})$  to  $\max(x_{vi})$  and  $f_j$  are a pre-defined sets of Gaussian functions. The final prediction is

**Table 1**  
List of frequently used acronyms in this article.

Acronym	Term
CNN	Convolutional Neural Networks
DFT	Density Functional Theory
GNN	Graph Neural Networks
ML	Machine Learning
DTNN	Deep Tensor Neural Networks
MPNN	Message Passing Neural Networks
D-MPNN	Directed Message Passing Neural Networks
HOMO	Highest Occupied Molecule Orbital
LUMO	Lowest Unoccupied Molecule Orbital
AUC	Area Under the Curve

obtained through a neural network:  $\hat{y} = \text{NN}(x_v^{T+1})$ . Unlike MPNNs, the edge states are updated during the convolutions, which may enhance the information exchange, or in another word, interaction between node-node and node-edge. The model achieves comparable mean Area Under the Curve (AUC) scores to baseline models on three biochemical molecule datasets PCBA [32], MUV [33] and Tox21 [34].

In the work by Schütt *et al.*, the author proposed a model called Deep Tensor Neural Networks (DTNN). The message function is designed as  $M_t = \tanh(W^{fc}((W^{cf}x_w^t + b_1) \odot (W^{df}e_{vw} + b_2)))$ . The update function is  $U_t(x_v^t, m_v^t) = x_v^t + m_v^t$ . The readout function is  $R = \sum_v W^{out_2} \tanh(W^{out_1}x_v^T + b^{out_1}) + b^{out_2}$ , where  $W$  are matrices and  $b$  are bias vectors. The model is applied to different subsets of the GDB database [35] to make predictions on molecular energies and local chemical potentials. The atomic features are randomly initialized with fixed length and the inter-atomic distances are expanded as Gaussian features which are defined on a series of Gaussian functions with means on a range of distances. The DTNN achieved a MAE of 1.0 kJ mol<sup>-1</sup> on energy predictions. Due to the increased amount of data, larger molecules exhibit a lower training MAE. Note that, due to the fact that the training set used in this work is relatively small, there is a significant overfitting for larger molecules.

In 2018, a continuous-filter convolutional neural network (SchNet) was proposed to handle quantum interactions for both molecules and solid-state materials through a graph setup [36]. The architecture of SchNet is shown in Fig. 2, where the message function is  $M_t(x_v^t, x_w^t, e_{vw}) = x_v^t \odot h_\Theta(e^{-\gamma(e_{vw} - \mu)})$  and the readout function is a summation pooling function. This convolution scheme is named by the authors as a “continuous filter”. Atom features are iteratively updated by the interaction between itself and neighboring atoms. This kind of information exchanging process is essentially a generalization of traditional convolutional neural network on graph structured data. A pooling layer with period boundary condition, which is the notable feature of crystal materials, is applied at the end of the network to generate final predictions. As the first graph-based deep learning model that can undertake regression tasks on both molecule and crystal datasets, SchNet achieved benchmark results of MAE (0.035 eV/atom) for formation energy prediction on a dataset containing 69,640 bulk crystals. The model also achieved benchmark results for the prediction of energies, electron, and thermal properties on QM9 dataset [31] and the prediction of potential energy surfaces and force fields on MD17 dataset [37].

In the work by Yang *et al.*, the authors proposed a learning framework called Directed MPNN (D-MPNN) [32], which is a variant of the MPNN framework. Instead of generating messages on nodes, messages are regarded as edges in directed graphs in this new framework. The message function is:  $m_{vw}^{t+1} = \sum_{k \in N(v) \setminus w} M_t(x_v, x_k, e_{kv}^t)$  where the summation is computed for all neighbors of node  $v$  except node  $w$ , and edges are updated by an update function:  $e_{vw}^{t+1} = U_t(e_{vw}^t, m_{vw}^{t+1})$ . Specifically, an update function  $U_t(e_{vw}^t, m_{vw}^{t+1}) = g(h_{vw}^0 + Wm_{vw}^{t+1})$  with  $h_{vw}^0 = g(W[x_v \oplus e_{vw}])$  is used, where  $g$  is a ReLU activation function. After several convolutions, the edge features are summed to node features by the readout function:  $R = \sum_v g(W[x_v \oplus m_v])$ , where  $m_v = \sum_{w \in N(v)} e_{vw}^T$  is the summation of all edge features with one end at the node.

#### 4. Graph-based deep learning for solid-state crystalline materials

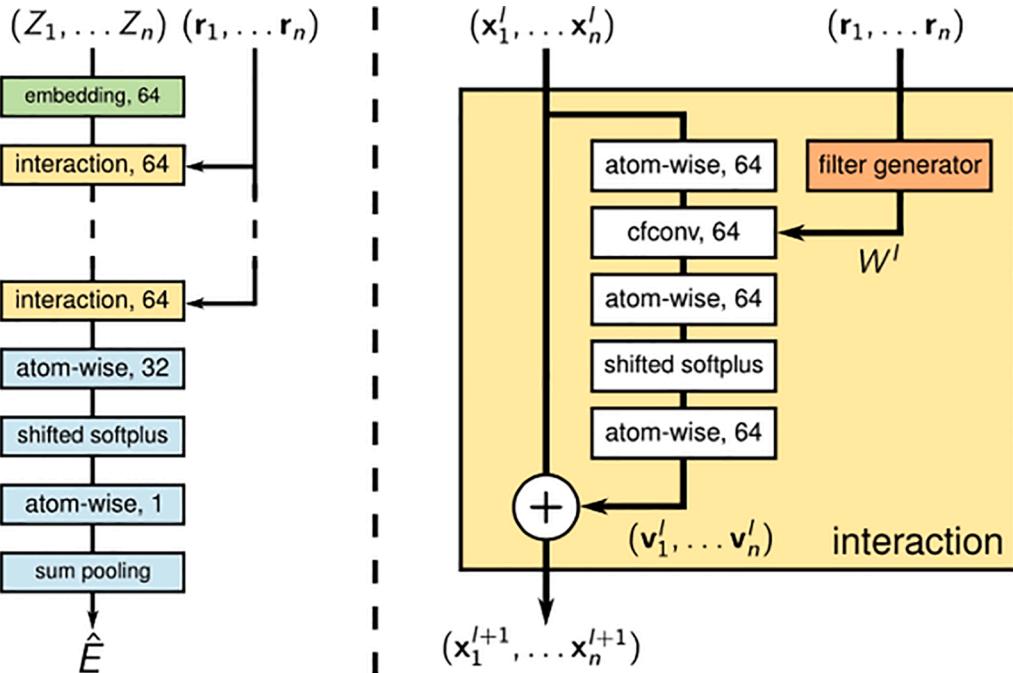
Due to the presence of periodic array of repeated unit cells in crystal structures, solid-state materials are defined by both local and global symmetries. Although there have been ML models such as SchNet [36] that incorporates the periodic boundary condition in the convolution and can be applied for both molecules and solid-state materials, most deep learning models that were developed for molecules cannot be directly applied to solid-state systems.

Very recently, multiple graph-based deep learning models were

**Table 2**

Comparison of different models with MPNN baselines. Reproduced with permission from Ref. [24].

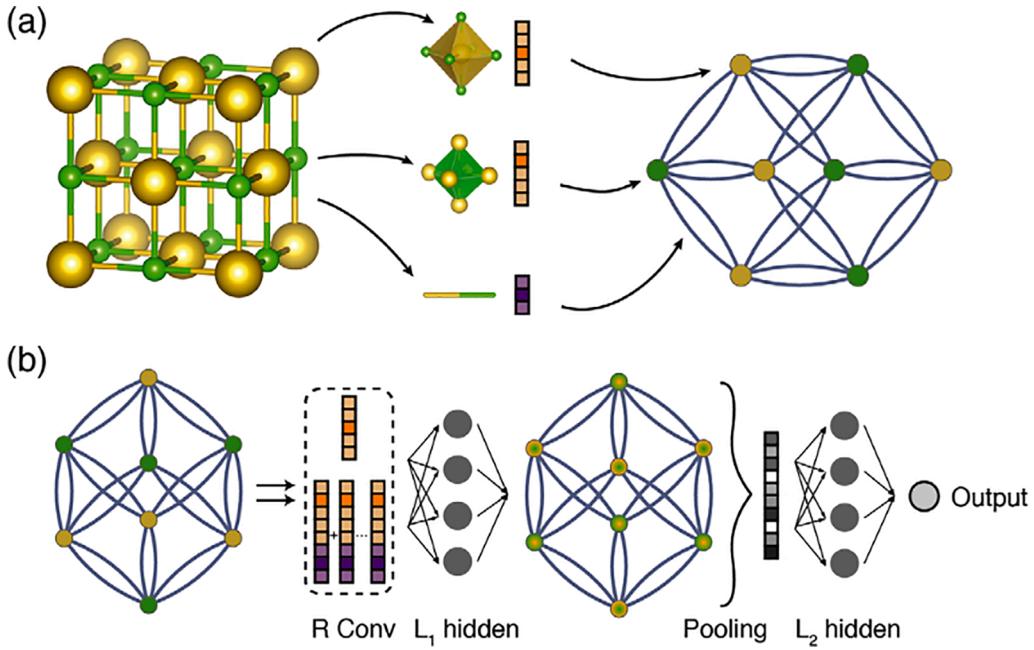
Target	BAML [26]	BOB [11]	CM [10]	ECP4 [27]	HDAD [28]	GC [29]	GG-NN [25]	DTNN [30]	enn-s2s [24]	enn-s2s-ens5 [24]
Mu	4.34	4.23	4.49	4.82	3.34	0.70	1.22	–	0.30	0.20
Alpha	3.01	2.98	4.33	34.54	1.75	2.27	1.55	–	0.92	0.68
HOMO	2.20	2.20	3.09	2.89	1.54	1.18	1.17	–	0.99	0.74
LUMO	2.76	2.74	4.26	3.10	1.96	1.10	1.08	–	0.87	0.65
Gap	3.28	3.41	5.32	3.86	2.49	1.78	1.70	–	1.60	1.23
R2	3.25	0.80	2.83	90.68	1.35	4.73	3.99	–	0.15	0.14
ZPVE	3.31	3.40	4.80	241.58	1.91	9.75	2.52	–	1.27	1.10
U0	1.21	1.43	2.98	85.01	0.58	3.02	0.83	–	0.45	0.33
U	1.22	1.44	2.99	85.59	0.59	3.16	0.86	–	0.45	0.34
H	1.22	1.44	2.99	86.21	0.59	3.19	0.81	–	0.39	0.30
G	1.20	1.42	2.97	78.36	0.59	2.95	0.78	0.84	0.44	0.34
Cv	1.64	1.83	2.36	30.29	0.88	1.45	1.19	–	0.80	0.62
Omega	0.27	0.35	1.32	1.47	0.34	0.32	0.53	–	0.19	0.15
Average	2.17	2.08	3.37	53.97	1.35	2.59	1.36	–	0.68	0.52

**Fig. 2.** Illustration of SchNet architecture (left), which consists of atom embedding layer (green), interaction blocks (yellow, details shown on the right) and property prediction networks (blue). Reproduced with permission from Ref. [36].

developed with a focus on material property predictions of crystal systems [38,39], which is a fundamental and inspiring step toward the large-scale use of GNNs for inorganic crystalline materials. Most of the recent graph networks applied to crystalline materials [38–40] are based on graphs on the atomic level  $G_0^{\text{atom}}$  as input for the network. Using atoms in the unit cell as nodes, distances between atoms as edges, an atom graph for a crystal can be created. Atoms related by a translation are equivalent and treated as one node. Following the notation used in the GNN framework [38], an attributed graph is defined as  $G = (V, E)$ . Several atom graph neural networks have been proposed [38–40] that formulate the task of predicting physical properties of materials as learning a mapping  $f(G; W) \rightarrow y$ , where  $W$  is a set of learnable parameters, and  $y$  is a target property. The nature of such an atom graph for a crystal structure is a small number of nodes combined with a large number of edges (see Figs. 1 and 3). In addition, there are multiple edges between nodes due to translational symmetry and the periodic boundary condition. Only the translational symmetry has been incorporated into the crystal graphs in this atom-based graph representation. Despite the recent advances, the application of graph-based deep learning frameworks in solid-state materials research is still in its infancy, leaving a big

room for further exploration.

The first work that combine crystal graphs with convolutional neural networks was introduced by T. Xie and J. C. Grossman in 2018 [39]. The proposed graph-based learning framework CGCNN takes crystal graph information as the input for a convolutional neural network (see Fig. 3). The framework can be viewed in the general picture of MPNNs [24] with the message function as  $M_t(x_v^t, x_w^t, e_{vw}) = \sigma(z_{ij}W_f + b_f) \odot g(z_{ij}W_s + b_s)$ , where  $z_{ij} = x_v^t \oplus x_w^t \oplus e_{vw}$  is the concatenation of center atom feature vector, neighbor atom feature vector and edge feature vector,  $\sigma$  and  $g$  are sigmoid and softplus functions. The readout function is a global mean pooling function which calculates the mean feature vector of all atoms in a crystal to obtain a single feature vector for each crystal. The model achieved benchmark prediction accuracy at the time for electronic and elastic properties such as formation energy, band gap, and Poisson's ratio. A MAE of 0.039 eV/atom and 0.072 eV/atom is obtained for formation energy and total energy prediction respectively, while a MAE of 0.388 eV and 0.363 eV is obtained for band gap and Fermi energy prediction respectively. Regarding the elastic properties of solid-state compounds, MAEs of 0.054 log(GPa), 0.087 log(GPa) and 0.030 are obtained for bulk moduli, shear moduli and Poisson ratio. The model



**Fig. 3.** Illustration of the CGCNN architecture. (a) To construct the crystal graph, atoms in the cell are represented as nodes and bonds are represented as edges. Note in the figure that translational symmetric atoms are regarded as one node. (b) Architecture of CGCNN. L<sub>1</sub> hidden layers are connected after R convolution layers. After a global pooling layer, L<sub>2</sub> hidden layers are connected to the final output predictions. Reproduced with permission from Ref. [39].

also achieves 80% (95%) accuracy for metal (nonmetal) classification.

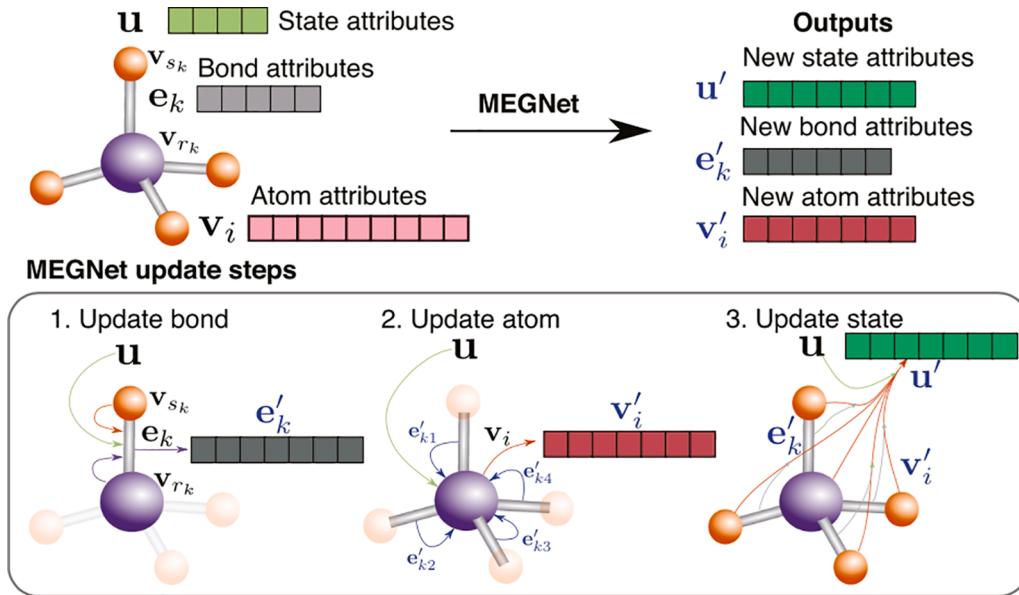
In a more recent work by X. Chen *et al.* published in 2019 [38], the authors proposed a general graph-based deep learning framework (MEGNet) and applied it to both molecule and crystal structures (see Fig. 4). In this new framework, a global feature vector is randomly initiated and optimized in the later training steps, representing the global state of the structure such as temperature or pressure. At each time step, the graph features consist of feature vectors of nodes, edges and the global state. New graph features are generated by interactions between them through convolution functions. The convolution functions are:

$$e_{vw}^{t+1} = \text{MLP}(x_v^t \oplus x_w^t \oplus e_{vw}^t \oplus u^t)$$

$$x_v^{t+1} = \text{MLP}(\bar{e}_{vw}^t \oplus x_v^t \oplus u^t), \quad \bar{e}_{vw}^t = \frac{1}{N_v} \sum_{w \in N(v)} e_{vw}^t$$

$$u^{t+1} = \text{MLP}(\bar{e}^t \oplus \bar{x}^t \oplus u^t),$$

where  $u^t$  is the global state feature vector, MLP is the multi-layer perceptrons with modified softplus function as nonlinear activation function. The bar sign represents the mean value. After several convolution blocks, the output node and edge features are computed from a



**Fig. 4.** Architecture of MEGNet. The molecule graph is represented as a graph with atom attributes  $v$ , bond attributes  $e$  and global state attributes  $u$ , which are then repeatedly updated by information exchanging. After several steps, the latest attributes are used to generate final outputs. Reproduced with permission from Ref. [38].

*set2set* model as a readout function, which is invariant to the order of node and edges. The model outperforms the benchmark models on the QM9 dataset for the prediction of molecular orbital energies and thermal properties. A MAE of 0.038 eV, 0.031 eV and 0.061 eV are obtained for the regression of HOMO/LUMO energy and energy gap. The model outperforms benchmark models for the prediction of different thermal properties. It also outperforms CGCNN [39] and SchNet [36] trained on the Materials Project crystal dataset [41] for the prediction of Fermi energy, band gap, elastic properties, and metal-nonmetal classification. A MAE of 0.028 eV/atom and 0.33 eV is obtained for Fermi energy and band gap prediction respectively. And a MAE of 0.050 log(GPa) and 0.079 log(GPa) is obtained for bulk moduli and shear moduli respectively. The model achieves 78.9% (90.6%) accuracy for metal (non-metal) classification.

As an improved model on top of CGCNN [39], C. W. Park and C. Wolverton proposed the so-called iCGCNN model [40] which incorporates the information of Voronoi tessellation of crystal structures and three-body correlations among atoms. Also, a convolution of edge features was introduced to keep them updated during the atomic convolutions. The model outperforms more than 20% for formation energies and the distances to convex hull compared to CGCNN. Using both CGCNN and iCGCNN, the authors screened 132,600 compounds with elemental decorations of the ThCr<sub>2</sub>Si<sub>2</sub> crystal structure and identified 97 new stable compounds by performing only 757 DFT calculations. This effort demonstrated that iCGCNN and related graph-based learning frameworks can be used to accelerate the discoveries of new materials by quickly and accurately identifying crystalline systems with properties of interest.

## 5. Graph neural networks for non-crystalline complex materials systems

Recently, GNNs have been applied for automated classification and characterization of amorphous materials [42]. In this work, a D-MPNN model [32] with an attention mechanism was proposed to classify and characterize amorphous materials. The attention function is  $\sigma = \text{softmax}(g(HW_{\text{attn}}v_{\text{attn}}))$ , where  $g$  is a ReLU function,  $H$  is a  $n \times h$  matrix with  $n$  the number of edges and  $h$  the dimension of edge features,  $W$  and  $v$  are learned matrix and vector to calculate attention. The weights  $\sigma$  are multiplied with edge features as attention coefficients to yield new edge features:  $H' = \Sigma \odot H$ . With no priori information about interactions between particles, the model outperforms CNN when performing a classification of liquid and glass configurations. It is pointed out in this work that the use of graph structured data has several advantages over the traditional CNN in image classification, as graph permutation and rotation are preserved in the data structure and it's flexible to adjust the graph size.

Another recent work by T. Xie *et al.* developed a graph dynamical network model to understand the dynamical information in amorphous material systems by applying unsupervised learning on a linear Koopman model from molecular dynamics data [43]. For each pair-wise data  $z_{ij} = x_v^t \oplus x_w^t \oplus e_{vw}$ , an exponential attention is calculated:

$$\alpha_{ij} = \frac{\exp(z_{ij}W + b)}{\sum_j \exp(z_{ij}W + b)}$$

The adopted message function is  $M_t(x_v^t, x_w^t, e_{vw}) = \sum_i \alpha_{ij} g(z_{ij}W + b)$ , where  $g$  is a ReLU function. The dynamics learned from two dynamical systems, solid-liquid interfaces and solid polymer electrolytes, indicate the potential of applying graph dynamical network to a wide range of material systems. With the large amounts of molecular dynamics data available in many aspects of materials design, the approach provides a broadly applicable learning tool to understand atomic scale dynamics in material systems that are crucial to their performances.

## 6. Moving beyond atom-based graphs

As discussed above, atom-based graph neural networks have been proposed and applied to predict materials properties of solid-state compounds, showing considerable performance improvement compared with standard neural networks or shallow learning models [38–40]. Although a great performance improvement has been achieved, there are still some fundamental limitations. For instance, compared to energetics and mechanical properties, electronic structure related properties of solid-state materials are found to be much more difficult to be learned and predicted by an atom-based GNN [38,44]. This limitation of atom-based models calls for the incorporation of beyond-atom material information and physical principles in a ML architecture, which is a great challenge and a fundamental step toward the continued development of artificial intelligence for inorganic materials.

As inspired by the Pauling's rule, we recently proposed that structure motifs, building blocks in inorganic crystals, can serve as a central input to a machine learning framework. Structure motifs in crystal structures are higher-level abstraction of structure information beyond the atom level. It is analogous to looking at an image with lower resolution in which a pixel is the aggregation of several small pixels from the original image. Therefore, these structure motifs can be treated as additional fingerprints providing higher-level knowledge of crystal structures.

In our recent work, it is shown that an unsupervised learning algorithm *Atom2Vec* can learn high dimensional vector representations of atoms that encode basic properties of atoms by utilizing only the chemical formulas in the Materials Project database [45]. Moving forward from atom to motif, we demonstrated in a more recent work that the presence of structure motifs and their connections in a large set of crystalline compounds can be converted into unique vector representations using an unsupervised learning algorithm. We proposed an Atom-Motif Dual graph Network (AMDNet) which incorporates crystal structure information to enhance the prediction of electronic structure related material properties [44]. AMDNet consists of two graphs at different information level: atom graph (where nodes are atoms and edges are interatomic distances) and motif graph (where nodes are motifs and edges are binary properties between motifs such as face angles and distances between motif centers). Both graphs are fed into graph convolution blocks as used in the MEGNet model [38]. Node and edge features are then passed to a *set2set* model to extract order-independent features. Finally, atom and motif features are concatenated as total features, which are passed to dense layers to general final output (property prediction). It is demonstrated that a combination of atom-based and motif-based graphs outperforms an atom-only graph for the prediction of electronic structure related properties such as band gaps of complex metal oxides.

Another tier of fundamental material information is atomic orbitals that are highly correlated with electronic structures and other properties. In the field of quantum chemistry, two recent work initiated the innovative use of atomic orbitals in graphs to represent molecule systems [46,47]. In the work of SchNOrb (SchNet for Orbitals) deep learning framework published in 2019, a local basis of atomic orbitals is applied to construct symmetry-adapted pairwise features which are incorporated in a deep neural network [46]. The architecture of SchNOrb is shown in Fig. 5. The model is an extension of the previous model SchNet, in a way that atom features are learned by SchNet in the front layers. Then pair-wise feature matrices are constructed to represent the interaction Hamiltonian between atoms. Rotational symmetry is adapted in the atomic orbital features. The model outputs include total energies, Hamiltonian, and overlap matrices, which are used to evaluate a combined regression loss. Owing to the explicit use of atomic orbital representation, SchNOrb is able to predict the quantum mechanical wavefunction from which all other ground-state properties can be derived. The model is not only able to give good predictions on molecule properties, but also helpful to aid the molecular dynamic simulations while reducing the computational cost.

In a more recent work by Z. Qiao *et al.* [47], the authors proposed a model called OrbNet which takes symmetry adapted atomic orbitals (from low-cost first principle calculations) as input and makes use of MPNN as convolution modules (see Fig. 6). The message function is chosen as  $M(x_v^t, x_w^t, e_{vw}^t) = \sigma(W[x_v^t \odot x_w^t \odot e_{vw}^t] + b)$  and the update function is  $U_t = x_v^t + \sigma(W[\oplus_i \sum_{w \in N(v)} \alpha_{vw} m_{vw}^t] + b)$ , where  $\sigma$  is sigmoid function,  $\odot$  is element-wise product,  $\oplus$  is concatenation,  $W$  is weight matrix and  $b$  is bias vector. An attention mechanism is used by calculating a weight for each pair of nodes:  $\alpha_{vw} = \sigma(\sum (Wx_v^t) \odot \sum (Wx_w^t) \odot e_{vw}^t / n_e)$  where  $n_e$  is the dimension of edge feature vector. After several convolutions, a readout function  $R = \sum_v \sum_v x_v^t$  is used to aggregate learned features from all time steps and all nodes. OrbNet outperforms many previous ML methods in terms of learning efficiency and transferability for the prediction of molecule properties such as energies at a very low computational cost.

The recent development of orbital-graph based learning models in quantum chemistry motivates us to adopt orbital-based graphs for solid-state material systems in a graph neural network learning architecture. To fully take into account of orbital interactions in a crystal, we recently proposed GNNs to effectively learn Hamiltonian of solid-state materials [48]. Representing the interactions among atomic orbitals in any material, a material Hamiltonian provides all the essential elements that control the structure–property correlations in solid-state compounds. In this work, we presented several different graph convolution networks with the aim to predict the band gaps for inorganic materials. By incorporating two different orbital features (the information of each orbital and the interactions between orbitals), our model trained on 530 half-Heusler compounds can achieve a promising prediction accuracy for the metal–nonmetal classification task compared with shallow learning models. The work is a fundamental step toward the continued development of orbital graphs for the deep learning of solid-state materials.

Very recently, an orbital graph convolutional neural network (OGCNN) was proposed to partially take into account of atomic orbital information for crystalline materials [49]. The proposed architecture is shown in Fig. 7. Based on the CGCNN model [39], the model encodes information of bonding between atomic orbitals into the Orbital Field Matrix (OFM) representation [50], in which electron configurations of constituting atoms and solid angles subtended by the neighboring atoms are used as additional information. The nodes and improved edge features are then passed to the CGCNN framework to make single-value predictions. Since the new edge feature vectors in this framework

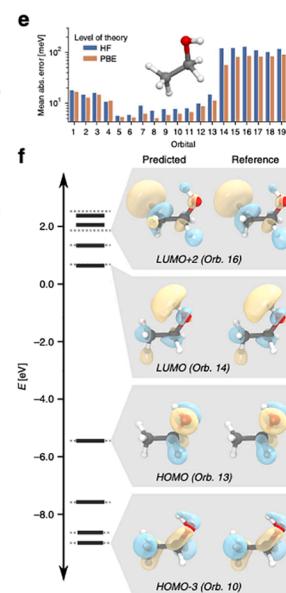
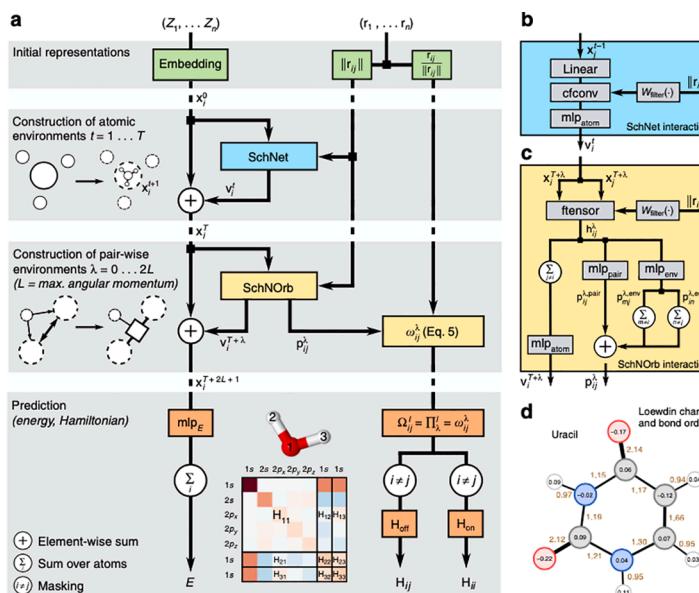
have much larger dimension than those in CGCNN, the front-end-modified networks have a stronger learning ability and a better prediction performance than CGCNN for both formation energetics and band gaps. Also, because more information about bonding is used in the architecture, the communication between nodes is enhanced.

## 7. Summary and outlook

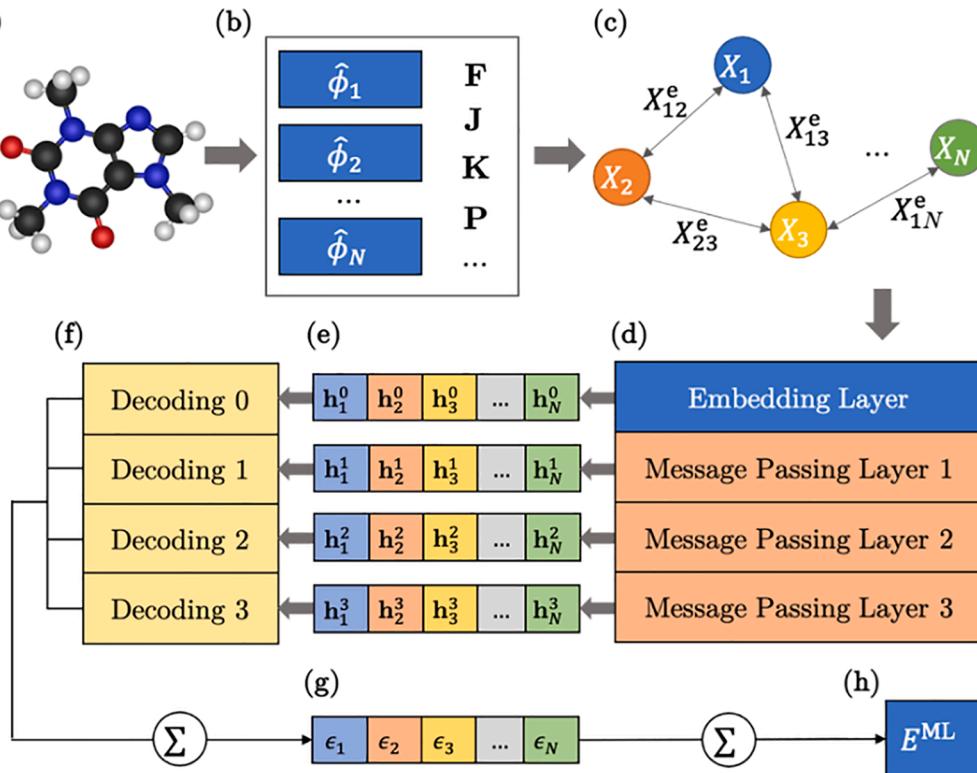
Although in its infancy, graph-based deep learning frameworks have already demonstrated their creative roles in the design and discovery of functional materials by identifying structure–property correlations and making efficient property predictions. By representing material systems in graphs and properly designing message passing strategies, deep learning architectures gain a new level of flexibility, efficiency, and reliability toward physical principle enhanced learning of material systems.

Before reaching the stage that graph-based deep learning frameworks can be readily applied to various types of diverse materials systems, several critical challenges need to be addressed by the community. First of all, endowed from the CNN architecture, the end-to-end learning strategy in GCNNs typically requires a huge amount of training samples. In the field of image recognition, the wide application of deep learning was not possible without the introduction of ImageNet [51], a large dataset containing millions of manually labeled images. For solid-state inorganic materials, unfortunately we do not have this large amount of labeled data (at least for now). To address this issue, a joint effort by the community to create a large amount of labelled material data with standard formats at a high accuracy level is called for. Furthermore, domain knowledges should be incorporated in the design of message passing strategies to control the overfitting of learning models with relatively limited data.

Another challenge is related to the training costs or the scalability of the models for both molecules and solid-state systems. For instance, it is noted that the performance of MPNNs dramatically decrease when dealing with larger molecules [24], which is attributed to the pair-wise interactions and the fully connected information passing channels. A single step of the message passing phase for a dense graph requires  $O(n[2]d^2)$  floating point multiplications, where  $n$  is the number of nodes and  $d$  is the dimensional of node vectors. This is computationally expensive for large material systems that go beyond small molecules and crystals. To overcome this scalability issue, the development of graph-based learning frameworks that can generalize effectively to larger graphs



**Fig. 5.** Architecture of SchNOrb. (a) Illustration of the model. The model consists of embedding layer, interaction block and prediction of Hamiltonian matrix and molecular energy. (b) Illustration of SchNet interaction block. (c) Illustration of SchNOrb interaction block. (d) Loewdin population analysis on the density matrix calculated from Hamiltonian prediction. (e) Comparison of predictions and DFT calculation results. (f) Predicted and DFT calculated orbital energies of molecules. Reproduced with permission from Ref. [46].



**Fig. 6.** Illustration of OrbNet workflow. (a) A low cost calculation is performed on molecular system to generate required quantum mechanical inputs (b) such as density matrix  $P$ , overlap matrix  $S$ , Fock operator  $F$ , Coloumb operator  $J$  and exchange operator  $K$ . (c) A molecule graph is constructed with nodes and edges representing diagonal and off-diagonal matrix elements of the inputs, respectively. (d) The graph input is passed to embedding and message passing layers. (e) The updated node and edge states are passed to (f) decoding layers and outputs (g) node-wise features. (h) A global pooling layer is connected to produce final predictions of energies. Reproduced with permission from Ref. [47].

relies on the construction of novel network architectures enabled by the incorporation of physical principles such as symmetries and topologies. As pointed out in the MPNN work [24], adding an attention mechanism over the incoming message vectors is an interesting direction to explore and there have been some applications on molecular systems along this route [52–54].

In the field of solid-state crystalline materials, a remaining challenge for the community is the reliable predictions of electronic structures and related physical properties. The predictions of electronic structure related material quantities (such as band gaps and band edges) by existing graph learning frameworks are still not satisfactory and need further improvement. Future development of graph-centric deep learning frameworks relies on novel architecture designs that go beyond real space and take into account multiple tiers of material information of atomic systems, such as orbitals, local and global symmetries, topologies, and reciprocal space information.

Last but not least, an important aspect of applying graph-based ML to material science is the feasibility to expand the applications of learning

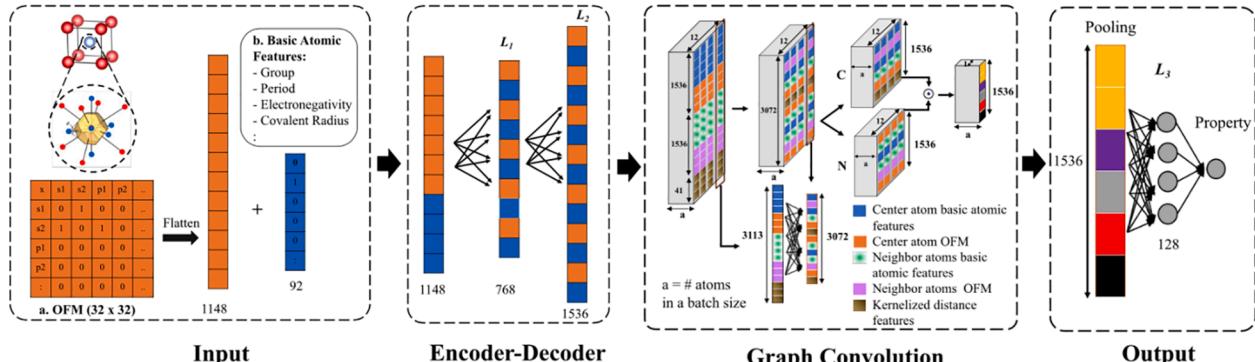
models developed based on small and regular graphs to large and irregular material systems, such as solids with point and line defects, surface and interface structures, molecule-solid hybrid systems and so on. Those technologies that are under active development in other fields of ML, such as transfer learning and graph attention, deserve more attentions to identify their roles in this emerging and exciting research field.

## 8. Data availability

No datasets were generated or analyzed during the current study.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.



**Fig. 7.** Illustration of OGCNN workflow. In the input parts, atomic and orbital data are embedded as vectors, which are then processed by encoding and decoding layers. Afterwards, CG convolution [39] layers are connected, followed by a global pooling layer to generate the final outputs. Reproduced with permission from Ref. [49].

## Acknowledgment

This work was supported by the U.S. Department of Energy, Office of Science, under award number DE-SC0020310.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.commatsci.2021.110332>.

## References

- [1] M.I. Jordan, T.M. Mitchell, *Science* 349 (2015) 255–260.
- [2] R. Ramakrishnan, P.O. Dral, M. Rupp, O.A. von Lilienfeld, *J. Chem. Theory Comput.* 11 (5) (2015) 2087–2096.
- [3] P. Raccuglia, K.C. Elbert, P.D.F. Adler, C. Falk, M.B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, A.J. Norquist, *Nature* 533 (7601) (2016) 73–76.
- [4] G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K.R. Muller, O.A. von Lilienfeld, *New J. Phys.* 15 (2013), 095003.
- [5] A.P. Bartók, S. De, C. Poelking, N. Bernstein, J.R. Kermode, G. Csányi, M. Ceriotti, *Sci. Adv.* 3 (12) (2017), e1701816.
- [6] K. Kim, L. Ward, J. He, A. Krishna, A. Agrawal, C. Wolverton, *Phys. Rev. Mater.* 2 (12) (2018), 123801.
- [7] G. Pilania, J.E. Gubernatis, T. Lookman, *Comput. Mater. Sci.* 129 (2017) 156–163.
- [8] K. Hansen, G. Montavon, F. Biegler, S. Fazli, M. Rupp, M. Scheffler, O.A. von Lilienfeld, A. Tkatchenko, K.R. Muller, *J. Chem. Theory Comput.* 9 (8) (2013) 3404–3419.
- [9] A.P. Bartók, M.C. Payne, R. Kondor, G. Csanyi, *Phys. Rev. Lett.* 104 (13) (2010), 136403.
- [10] M. Rupp, A. Tkatchenko, K.-R. Müller, O.A. von Lilienfeld, *Phys. Rev. Lett.* 108 (5) (2012), 058301.
- [11] K. Hansen, F. Biegler, R. Ramakrishnan, W. Pronobis, O.A. von Lilienfeld, K.-R. Müller, A. Tkatchenko, *J. Phys. Chem. Lett.* 6 (12) (2015) 2326–2331.
- [12] H. Huo and M. Rupp, arXiv:1704.06439 (2017).
- [13] Y. LeCun, Y. Bengio, G. Hinton, *Nature* 521 (2015) 436.
- [14] Z. Shi, E. Tsymbalov, M. Dao, S. Suresh, A. Shapeev, J. Li, *Proc. Natl. Acad. Sci.* 116 (10) (2019) 4117.
- [15] Y. Dong, C. Wu, C. Zhang, Y. Liu, J. Cheng, J. Lin, *NPJ Comput. Mater.* 5 (1) (2019) 26.
- [16] G.B. Goh, N.O. Hodas, A. Vishnu, *J. Comput. Chem.* 38 (16) (2017) 1291–1307.
- [17] K. He, X. Zhang, S. Ren and J. Sun, 2016 IEEE Conference on Computer Vision and Pattern Recognition, 770–778 (2016).
- [18] D. Duvenaud, D. Maclaurin, J. Aguilera-Iparraguirre, R. Gómez-Bombarelli, T. Hirzel, A. Aspuru-Guzik and R. P. Adams, arXiv:1509.09292 (2015).
- [19] Z. Zhang, P. Cui and W. Zhu, arXiv:1812.04202 (2018).
- [20] M.M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, P. Vandergheynst, *IEEE Signal Process. Mag.* 34 (4) (2017) 18–42.
- [21] M. Henaff, J. Bruna and Y. LeCun, arXiv:1506.05163 (2015).
- [22] T. N. Kipf and M. Welling, arXiv:1609.02907 (2016).
- [23] P. W. Battaglia, J. B. Hamrick, V. Bapst, A. Sanchez-Gonzalez, V. Zambaldi, M. Malinowski, A. Tacchetti, D. Raposo, A. Santoro, R. Faulkner, C. Gulcehre, F. Song, A. Ballard, J. Gilmer, G. Dahl, A. Vaswani, K. Allen, C. Nash, V. Langston, C. Dyer, N. Heess, D. Wierstra, P. Kohli, M. Botvinick, O. Vinyals, Y. Li and R. Pascanu, arXiv:1806.01261 (2018).
- [24] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals and G. E. Dahl, arXiv: 1704.01212 (2017).
- [25] Y. Li, D. Tarlow, M. Brockschmidt and R. Zemel, arXiv:1511.05493 (2015).
- [26] B. Huang, O.A. von Lilienfeld, *J. Chem. Phys.* 145 (16) (2016), 161102.
- [27] D. Rogers, M. Hahn, *J. Chem. Inf. Model.* 50 (5) (2010) 742–754.
- [28] F. A. Faber, L. Hutchison, B. Huang, J. Gilmer, S. S. Schoenholz, G. E. Dahl, O. Vinyals, S. Kearnes, P. F. Riley and O. Anatole von Lilienfeld, arXiv:1702.05532 (2017).
- [29] S. Kearnes, K. McCloskey, M. Berndl, V. Pande, P. Riley, *J. Comput. Aided. Mol. Des.* 30 (8) (2016) 595–608.
- [30] K.T. Schütt, F. Arbabzadah, S. Chmiela, K.R. Müller, A. Tkatchenko, *Nat. Commun.* 8 (1) (2017) 13890.
- [31] R. Ramakrishnan, P.O. Dral, M. Rupp, O.A. von Lilienfeld, *Sci. Data* 1 (1) (2014), 140022.
- [32] K. Yang, K. Swanson, W. Jin, C. Coley, P. Eiden, H. Gao, A. Guzman-Perez, T. Hopper, B. Kelley, M. Mathea, A. Palmer, V. Settels, T. Jaakkola, K. Jensen, R. Barzilay, *J. Chem. Inf. Model.* 59 (8) (2019) 3370–3388.
- [33] S.G. Rohrer, K. Baumann, *J. Chem. Inf. Model.* 49 (2) (2009) 169–184.
- [34] A. Mayr, G. Klambauer, T. Unterthiner, S. Hochreiter, *Front. Environ. Sci.* 3 (2016) 80.
- [35] L.C. Blum, J.-L. Reymond, *J. Am. Chem. Soc.* 131 (25) (2009) 8732–8733.
- [36] K.T. Schütt, H.E. Sauceda, P.J. Kindermans, A. Tkatchenko, K.R. Müller, *J. Chem. Phys.* 148 (24) (2018), 241722.
- [37] S. Chmiela, A. Tkatchenko, H.E. Sauceda, I. Poltavsky, K.T. Schütt, K.-R. Müller, *Sci. Adv.* 3 (5) (2017), e1603015.
- [38] C. Chen, W. Ye, Y. Zuo, C. Zheng, S.P. Ong, *Chem. Mater.* 31 (9) (2019) 3564–3572.
- [39] T. Xie, J.C. Grossman, *Phys. Rev. Lett.* 120 (14) (2018), 145301.
- [40] C.W. Park, C. Wolverton, *Phys. Revi. Mater.* 4 (6) (2020), 063801.
- [41] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K.A. Persson, *APL Mater.* 1 (1) (2013), 011002.
- [42] K. Swanson, S. Trivedi, J. Lequieu, K. Swanson, R. Kondor, *Soft Matter* 16 (2) (2020) 435–446.
- [43] T. Xie, A. France-Lanord, Y. Wang, Y. Shao-Horn, J.C. Grossman, *Nat. Commun.* 10 (1) (2019) 2667.
- [44] H. R. Banjade, S. Hauri, S. Zhang, F. Ricci, G. Hautier, S. Vucetic and Q. Yan, arXiv: 2007.04145 (2020).
- [45] Q. Zhou, P. Tang, S. Liu, J. Pan, Q. Yan, S.-C. Zhang, *Proc. Natl. Acad. Sci.* 115 (28) (2018) E6411.
- [46] K.T. Schütt, M. Gastegger, A. Tkatchenko, K.R. Müller, R.J. Maurer, *Nat. Commun.* 10 (1) (2019) 5024.
- [47] Z. Qiao, M. Welborn, A. Anandkumar, F.R. Manby, T.F. Miller, *J. Chem. Phys.* 153 (12) (2020), 124111.
- [48] H. Bai, P. Chu, J.-Y. Tsai, N. Wilson, X. Qian, Q. Yan and H. Ling, arXiv:2005.13352 (2020).
- [49] M. Karamad, R. Magar, Y. Shi, S. Siahrostami, I.D. Gates, A. Barati Farimani, *Phys. Rev. Mater.* 4 (9) (2020), 093801.
- [50] T. Lam Pham, H. Kino, K. Terakura, T. Miyake, K. Tsuda, I. Takigawa, H. Chi Dam, *Sci. Technol. Adv. Mater.* 18 (1) (2017) 756–765.
- [51] J. Deng, W. Dong, R. Socher, L. Li, L. Kai and F.-F. Li, 2009 IEEE Conference on Computer Vision and Pattern Recognition, 248–255 (2009).
- [52] Z. Xiong, D. Wang, X. Liu, F. Zhong, X. Wan, X. Li, Z. Li, X. Luo, K. Chen, H. Jiang, M. Zheng, *J. Med. Chem.* 63 (16) (2020) 8749–8760.
- [53] M. Withnall, E. Lindelöf, O. Engkvist, H. Chen, *J. Cheminform.* 12 (1) (2020) 1.
- [54] Ł. Maziarzka, T. Daniel, S. Mucha, K. Rataj, J. Tabor and S. Jastrzębski, arXiv: 2002.08264 (2020).



**Dr. Qimin Yan** is an Assistant Professor at Temple University, USA. He obtained his Ph.D. degree from University of California, Santa Barbara. During 2013–2016 he held a postdoctoral position at the Lawrence Berkeley National Lab and UC Berkeley. In 2016, he joined the Department of Physics at Temple University. His main research interests are data-driven discovery of quantum materials, physical principle enhanced machine learning for solid-state materials, functional semiconductors for energy conversion, and defects for quantum information technologies.