Check for updates

Design guidance and band gap prediction of two-dimensional hybrid organic–inorganic perovskites by ensemble learning and graph convolutional neural networks

Jianfei Liu,[a] Xia Cai, [ID] *[b] Lin Wang[a] and Yiqiang Zhan [ID] [b]

Two-dimensional (2D) hybrid organic–inorganic perovskites (HOIPs) are promising materials for addressing the stability challenges in perovskite solar cells due to their exceptional environmental stability, exciton dynamics, and broad-band emission. However, research on the structure-directing role of organic cations in 2D HOIPs remains limited, and existing models for predicting their band gap lack sufficient accuracy. Here, we develop ensemble learning models and a site-attention-based graph convolutional neural network (SATGNN) to predict the dimensionality of lead iodide-based HOIPs and the band gap of 2D HOIPs, respectively. The ensemble learning models, leveraging molecular descriptors with eXtreme Gradient Boosting, achieve 88% cross-validation accuracy, and MaxAbsEStateIndex, Chi2n and Kappa2 are identified as critical features for dimensionality determination. The SATGNN model incorporates a convolution function tailored to the unique layered structure of 2D HOIPs and a novel site-attention mechanism to prioritize elemental contributions. The SATGNN significantly outperforms existing approaches by accurately capturing structural interactions and spatial configurations. Furthermore, the visualization of SATGNN confirms model's ability to identify structural features of 2D HOIPs and distinguishes the effects of different elemental types on material properties. By revealing interpretable molecular descriptors that govern 2D HOIP formation and integrating accurate band gap prediction, this two-stage framework offers both actionable design guidelines for organic cation selection and a scalable tool for the accelerated discovery of 2D HOIPs with targeted optoelectronic properties.

1   Introduction

Hybrid organic–inorganic perovskites (HOIPs) as promising next-generation photovoltaic materials have attracted great attention in recent years. Since the first successful report of HOIP solar cells ($CH_3NH_3PbX_3$) by Kojima *et al.* in 2009,[1] they have been investigated extensively. The power conversion efficiency (PCE) of HOIP-based photovoltaic systems has increased from 3.8% to 26% in only 10 years.[2] Although HOIP-based solar cells have made great progress, most HOIPs suffer from poor stability. Fortunately, compared with three-dimensional (3D) HOIPs represented by $CH_3NH_3PbX_3$, two-dimensional (2D) HOIPs exhibit improved environmental stability due to the incorporation of bulky organic cations, which form hydrophobic layers and sterically shield the inorganic framework from moisture and oxygen ingress. The different structures of perovskites are shown in Fig. 1. The structure of 3D HOIPs
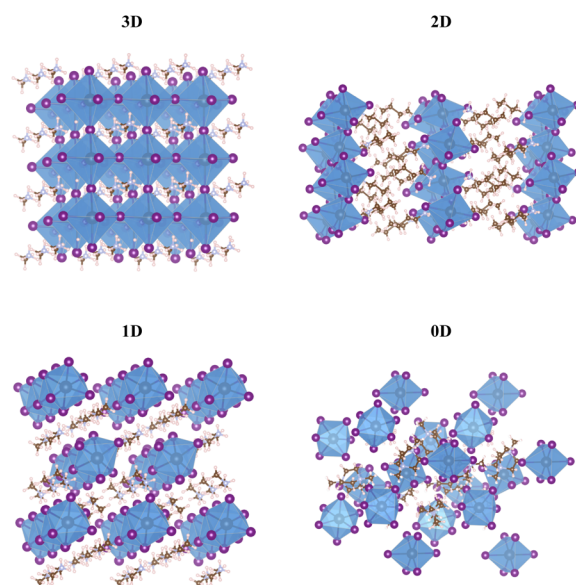
[a]College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai 200234, China

[b]School of Information Science and Technology, Fudan University, Shanghai 200433, China. E-mail: xcai17@fudan.edu.cn

Fig. 1   Structures of 3D, 2D, 1D, and 0D perovskites.

consists of a 3D network of corner-sharing metal halide octahedra, with organic cations occupying the 12-fold coordinated sites between the octahedra. The 2D structures are formed by taking *n*-layer thick cuts along a particular crystallographic plane of the 3D structure and stacking these slabs in alternation with organic cation layers. 1D structures are formed from chains or ribbons of metal halide octahedra surrounded by organic cations, and 0D structures consist of isolated octahedra or clusters of connected octahedra. Compared to 3D HOIPs, 2D HOIPs offer the advantages of compositional diversity, quantum-well electronic structure, broad-band emission, and layer-tunable photoelectronic properties.[3,4] However, a major limitation remains: achieving 2D HOIPs with a band gap close to the theoretical Shockley–Queisser limit (1.34 eV),[5] a critical parameter for their optoelectronic applications, is challenging. In order to achieve 2D HOIPs with a specific band gap, researchers must continually synthesize new materials through trial-and-error experiments until perovskites with the desired band gap are identified, or employ density functional theory (DFT) calculations to predict and screen potential candidates, both of which are very expensive in terms of time and cost.

Recently, high-throughput computational material design based on machine learning (ML) methods has emerged as an efficient approach for discovering new materials. The main advantage of the ML method is that it bypasses complex calculations involved in solving quantum mechanics equations, allowing it to learn the relationships between material structures and properties from material data, and can rapidly predict target properties with less computational resources. There are currently a few published studies that use ML to guide the design of perovskites with specific dimensions. Lyu *et al.*[6] developed a linear regression (LR) model with an L1 penalty to predict whether low-dimensional HOIPs belong to the 2D group, achieving 82% accuracy on the test set, and subsequently analyzed four features of organic cations that influence the formation of 2D HOIPs. However, their model construction relied on a limited set of features that could not accurately capture the characteristics of organic cations, resulting in a lower prediction accuracy and adversely affecting the feature analysis outcomes. Yuan *et al.*[7] developed a K-Nearest Neighbour (KNN) model to classify the dimension of low-dimensional HOIPs into 0D, 1D and 2D, and then analyzed two key features that influenced the dimension of these materials. However, the KNN model employed in their work is an instance-based non-parametric model, which can capture some nonlinear relationships but has limitations when dealing with high-dimensional data, potentially missing features that significantly impact the dimensions of HOIPs. The properties of organic cations in HOIPs are important factors affecting their dimension, but there is a lack of studies on the structure-directing effect of organic cations on 2D HOIPs.

In addition, the predictions of material band gap have been made using both traditional ML methods and deep learning (DL) models based on graph neural networks (GNNs). Marchenko *et al.*[8] used a ML model to predict the band gap of 2D perovskites with the database they developed, resulting in a mean absolute error (MAE) of 0.103 eV on the test set.

Traditional ML methods, when dealing with 2D HOIPs with complex structures, fail to adequately capture the mapping relationships between their structures and properties. In contrast, DL methods such as GNN demonstrate excellent performance in dealing with these intricate relationships. Xie *et al.*[9] were among the first researchers to apply GNN to material property prediction, proposing a crystal graph convolutional neural network framework that learns material properties directly from the connections of atoms in crystals. Their model demonstrated excellent prediction performance across various material properties, including formation energy, absolute energy, and band gap. Louis *et al.*[10] developed a graph convolutional neural network composed of augmented graph-attention layers and a global attention layer, improving the prediction accuracy on a broad array of material properties. Moreover, Chen *et al.*[11] developed universal GNN models for accurate property prediction in molecules and crystals. Choudhary *et al.*[12] presented a GNN architecture that performs message passing on both the interatomic bond graph and its line graph corresponding to bond angles. Recent work has also advanced perovskite-specific band gap prediction. For example, Gao *et al.*[13] utilized cluster-level descriptors for electronic property estimation, and complementary experimental efforts have demonstrated effective band gap tuning *via* compositional engineering,[14,15] offering practical pathways toward optimized materials. However, existing GNN models are not specifically developed for the structural characteristics of 2D HOIPs, making it difficult to accurately capture their unique structural properties and resulting in lower prediction accuracy. Moreover, few studies have sought to combine dimensionality classification with band gap prediction into a unified framework specifically designed for 2D HOIPs.

Based on these issues, we report a two-stage strategy for both design guidance and band gap prediction of 2D HOIPs. Rather than merely reproducing established findings, this work aims to establish a generalizable framework that can accelerate the discovery of 2D HOIPs beyond the limitations of existing datasets. Firstly, we present a classification strategy assisted by supervised ML to explore how organic cations affect the dimensionality of HOIPs. The organic cation data are obtained from an open-access database and existing publications, and classification models are used to fit the data based on 2D and low-dimensional (1D and 0D) categories. By analyzing the decision-making process of the model, fundamental chemical and structural insights into the dimensional effects of organic cations can be obtained, which will be useful for the design of 2D HOIPs. Secondly, we propose a site-attention-based graph convolutional neural network (SATGNN) tailored to the structural characteristics of 2D HOIPs for improved band gap prediction. In this model, we perform convolution separately on atoms at different positions in 2D HOIPs to more effectively capture their respective interaction strengths and spatial configurations. Additionally, we propose a site-attention mechanism to enable the model to capture the distinct contributions of different layer components in 2D HOIPs to the material properties. Visualization of the learned element representations confirms that the SATGNN captures key
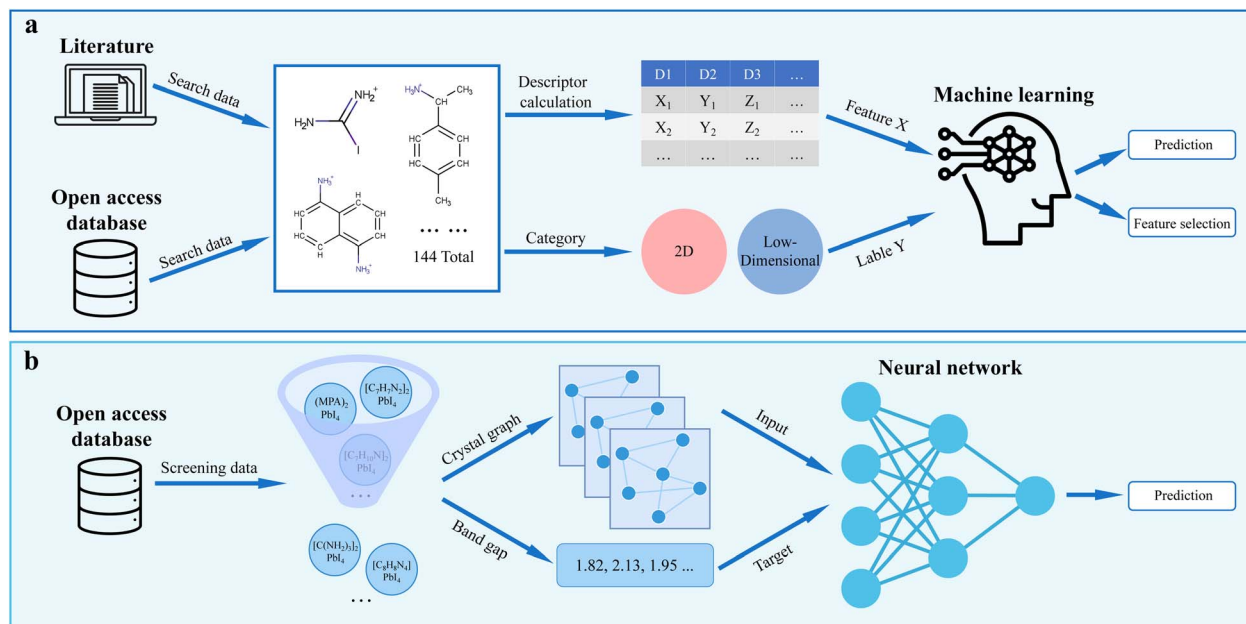
Fig. 2 (a) Machine learning-assisted workflow for predicting the dimensionality of lead iodide-based HOIPs. (b) Graph neural network-assisted workflow for predicting the band gap of 2D HOIPs.

structural features and differentiates the effects of various elemental types on material behavior. Overall, this approach delivers a fast, interpretable, and transferable toolset for screening and designing 2D HOIPs, offering practical value for data-driven materials discovery that extends beyond conventional DFT-based workflows.

2 Methods and results

2.1 Exploration of 2D HOIPs

The dimensionality of HOIP structures is strongly influenced by the size and shape of the organic A-site cation (A^+). When the cation becomes sufficiently bulky, it hinders the 3D connectivity of the inorganic octahedra, thereby favoring the formation of low-dimensional perovskite structures. This phenomenon, commonly referred to as the structure-directing role of organic cation, arises because its molecular geometry and steric constraints govern the stacking, spacing, and connectivity of the inorganic layers. In addition to the A^+ cation, the type of $[BX]^-$ octahedra also affects the dimensionality of HOIP structures due to electrostatic interactions between the A^+ organic cations and $[BX]^-$ octahedra.[7] Hence, the study of how the structure of organic cations affects the dimensionality of HOIPs needs to be based on a single type of $[BX]^-$ HOIPs. In this work, we select the most widely studied lead iodide perovskites. To ensure that the classification model learns to distinguish between structurally similar HOIPs, 0D and 1D structures are included in the training set as low-dimensional (LD) references. In contrast, 3D HOIPs are excluded because their fundamentally different structural characteristics and formation mechanisms introduce excessive heterogeneity. Including them could obscure key trends across the 0D, 1D, and 2D series and reduce the model's ability to identify relevant structural features.

The workflow for exploring 2D HOIPs is shown in Fig. 2a. We first review the literature and open-access database to obtain all reported lead iodide perovskites. Organic cations in perovskites are extracted and classified into "2D" and "LD" on the basis of the dimensionality of the formed perovskites. The molecular features of these organic cations (input X) and their classification (output Y) constitute the training data and are used to build predictive models. Then, the prediction capability of the optimal model is tested with several untested organic cations. The details are shown below.

The organic cations of lead iodide perovskites are obtained from the literature[6,16–18] and the open-access 2D perovskites database.[8] In the process of collecting data, we remove the organic cations that can form different perovskite structures due to different synthesis methods. Finally, a total of 144 organic cations are collected, with 42 classified into the LD group (0D and 1D perovskites) and 102 classified into the 2D group (2D perovskites). Based on this classification, the label Y is set to 0 and 1, where 0 represents the LD group and 1 represents the 2D group.

For a ML method that targets specific material properties, it usually relies on a certain number of feature descriptors. The Python package RDKit,[19] an open source toolkit for chemical informatics, is used to create feature descriptors in this work. We first convert the collected organic cations into the corresponding molecular files (.mol). Subsequently, various descriptor functions, including solid geometry, hydrogen bonding, charge, elemental analysis, topology analysis, *etc.*, are called in RDKit to calculate the features. These descriptors capture steric, topological, and electronic characteristics without requiring explicit crystal structure information, enabling efficient pre-screening across large chemical spaces. However, selecting a broad range of features can increase

computational costs and potentially lead to the curse of dimensionality.[20] Therefore, it is important to screen the key features for constructing the model, which directly affects the accuracy of the model. We screen the key feature in three steps as follows: (I) remove features with variance less than 1.0. (II) Remove features with a correlation higher than 80% by checking the Pearson correlation coefficients between features. (III) Employ the Recursive Feature Elimination (RFE) method to further eliminate unimportant features. Upon completion of feature preprocessing, the retained numerical descriptors are standardized using z-score normalization to ensure zero mean and unit variance. Finally, a total of 11 features is selected and used as inputs for building the ML models.

After obtaining features and labels, different classification models are trained and evaluated. We note that there is less data in the LD group compared to the 2D group. When the number of categories in the classification task is imbalanced, the model tends to predict the class that appears more frequently and ignore the minority class, making the model perform poorly when predicting the minority class. To address this issue, class weights are automatically assigned based on the inverse class frequency, allowing the model to emphasize the minority class during training without manual intervention. Moreover, ensemble learning (*e.g.* boosting methods) is useful for dealing with the class imbalance problem. Finally, K-Nearest Neighbor (KNN), Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), eXtreme Gradient Boosting (XGBoost), Gradient Boosting Decision Tree (GBDT), and Adaptive Boosting (AdaBoost) are employed in this work. For each ML algorithm, hyperparameter tuning is performed using grid search with 5-fold cross-validation on the training set. The optimized parameters are then used to evaluate model performance on the test data. The dataset is randomly split into training and test sets in an 80 : 20 ratio. To prevent data leakage and ensure robust evaluation, the test samples are kept completely separate from the training set, with strict separation to avoid data leakage and ensure a fair evaluation. We use classification accuracy, which is defined as the number of correct predictions divided by the total number of predictions, as the metric for evaluating the models.

Table 1 shows the accuracy of 5-fold cross-validation for different ML models. The XGBoost model achieves the highest accuracy of 0.88, indicating its superior efficiency in learning the mapping relationship between features and labels. Therefore, the XGBoost model is finally chosen for this study. To

evaluate the performance of the XGBoost model on unseen data, we assess its predictive accuracy using the organic cations from the test set and present the results through a confusion matrix, as shown in Fig. 3a. Only one LD sample among the nine samples is incorrectly classified as a 2D sample, which demonstrates the feasibility of our model in practical applications. The proposed dimensionality classifier enables rapid prediction of structural dimensionality solely from molecular descriptors of A-site organic cations, thereby bypassing the need for computationally expensive quantum mechanical calculations. This approach can serve as a high-throughput pre-screening tool for navigating vast chemical spaces in future materials design.

In order to further confirm the key features that affect the structural dimensionality of HOIPs, we use SHapley Additive exPlanations (SHAP)[21] to analyze the XGBoost model. SHAP is a game-theoretic method to explain the output of a ML model by considering the contribution of each feature to the predictions, helping researchers clearly understand how much importance the model gives to different features, where the ranking of feature importance is determined by the SHAP value. The 11 features are ranked by the calculated SHAP value based on the optimal XGBoost model, as shown in Fig. 3b. A point in Fig. 3b corresponds to a sample, with red and blue colors indicating high and low values of a particular feature, respectively. The x-axis labeled as the SHAP value represents the impact of features on the HOIP dimensionality, where positive and negative SHAP values represent positive and negative effects on the prediction results, respectively. We find that the top three features (MaxAbsEStateIndex, Chi2n and Kappa2) play a decisive role in the prediction results of the model, as shown in the inset of Fig. 3b.

MaxAbsEStateIndex is a derivation based on the electrotopological state (E-state) proposed by Hall and Mohney.[22] The E-state is developed from chemical graph theory as an index of the graph vertex (or skeletal group). This index combines both the electronic character and the topological environment of each skeletal atom in a molecule. The E-state of a skeletal atom is formulated as an intrinsic value I_i plus a perturbation term ΔI_i, which arises from the electronic interaction and is modified by the molecular topological environment of each atom in the molecule. The E-state, S_i, for atom i is defined as $S_i = I_i + \Delta I_i$. The MaxAbsEStateIndex descriptor calculates the maximum absolute value of the E-state index for all atoms in the molecule, which reflects the most significant electronic and topological properties of the molecule. It can be seen from Fig. 3b that the larger value of MaxAbsEStateIndex, the easier it is to form 2D HOIPs.

Chi2n is a derivation based on the chi indexes proposed by Hall and Kier.[23] The chi index is a weighted count of a given type of subgraph, where each subgraph is a fragment decomposed from the molecular graph. There are two attributes of the chi index: the order and the type. The order of a chi index is the number of graph edges in the corresponding subgraph. The type refers to the particular arrangement of the edges in the subgraph. The descriptor Chi2n selected in this study represents the second-order chi index. For the first three lower-order indexes (order <3), there is only one type of subgraph.

Table 1 Accuracy of 5-fold cross-validation for different ML models

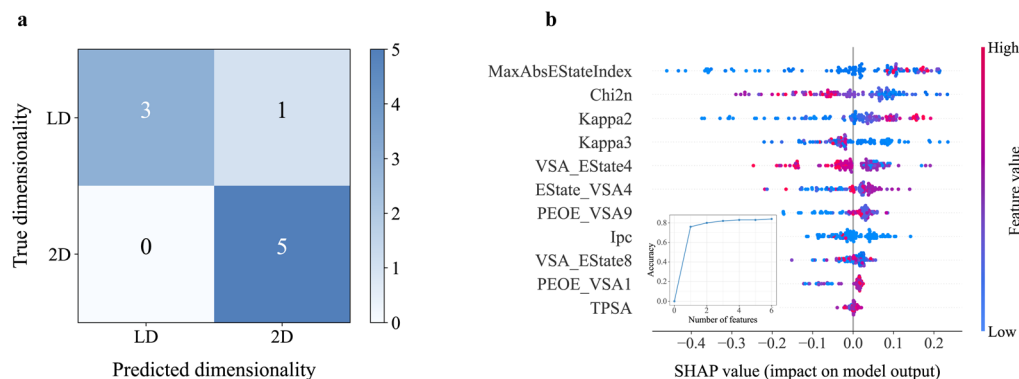| ML model	Accuracy	Hyperparameter
LR	0.66	C: 1.6, penalty: l1
KNN	0.85	leaf_size: 10, n_neighbors: 9
SVM	0.82	C: 8.5, gamma: 0.1
RF	0.85	min_samples_leaf: 3
XGBoost	**0.88**	Subsample: 0.6, learning_rate: 0.2
GBDT	0.87	Subsample: 1.0, learning_rate: 0.2
AdaBoost	0.85	n_estimators: 100, learning_rate: 0.01

Fig. 3 (a) Confusion matrix for prediction results of the XGBoost model on the test set. (b) The feature importance ranking generated by the SHAP library displays the XGBoost model's features in descending order of importance. The inset shows the prediction accuracy of the model as features are incrementally added from the top-ranked to the sixth feature.

Therefore, the Chi2n descriptor is not given a type designation, and its calculation formula is given by $\sum {}^2c_s$, where ${}^2c_s$ denotes all subgraphs with two edges. For ${}^2c_s$, the node information in the subgraphs is calculated using $\Pi(\delta_i)_s^{-1/2}$, where δ_i encodes the electronic identity of the atom i in terms of both valence electron count and core electron count, and the subscript s designates atoms that belong to the subgraph. In the subgraphs with two edges, there should be 3 δ terms. Therefore, the Chi2n descriptor encodes the structural information resident in the entire molecular skeleton, reflecting the constitutive nature of the molecular structure. As shown in Fig. 3b, we find that the smaller the value of Chi2n, the easier it is to form 2D HOIPs.

The Kappa index is calculated from the molecular graph[23] and is used to describe the shape of the molecule. It is calculated by quantifying the structural characteristics of the molecule. Kappa2 is the second-order Kappa index, which reflects the spatial density of atoms within the molecule and is related to the degree of star pattern or linear pattern of the molecule. Kappa2 is calculated as $(A + \alpha - 1)(A + \alpha - 2)^2/({}^2P_i + \alpha)^2$, where A represents the number of atoms in the molecule, ${}^2P_i$ represents the number of two-path fragments, i.e., two adjacent bonds, in the molecular graph, and α encodes the atom identity. And the atom identity is represented by the ratio of the covalent radii. It can be seen from Fig. 3b that the larger the value of Kappa2, the easier it is to form 2D HOIPs.

These three key features, which influence the dimensionality of HOIPs, characterize the topology and electronic states of the organic cation. Therefore, it can be concluded that the topology and electronic states of the organic cation are critical factors determining the dimensionality of HOIPs, which is consistent with literature reports.[24,25] Based on these findings, we provide preliminary design guidelines for screening and selecting organic cations likely to form 2D HOIPs. Specifically, organic cations meeting the following criteria favor 2D HOIP formation: (I) MaxAbsEStateIndex between 5.68 and 13.04, (II) Chi2n between 0.33 and 2.50, and (III) Kappa2 between 5.06 and 17.96. These descriptors capture essential aspects of molecular topology and electronic structure and can serve as effective filters in high-throughput screening workflows. This preliminary guidance complements our subsequent band gap prediction model, together enabling a two-stage strategy for the efficient discovery and rational design of 2D HOIPs. Unlike conventional DFT approaches, which offer accurate but case-specific predictions, our model extracts interpretable structure–property relationships from a chemically diverse dataset. Notably, the identification of MaxAbsEStateIndex, Chi2n, and Kappa2 as key descriptors governing 2D structural formation provides valuable, generalizable insights for the rational design of organic cations.

2.2 Band gap prediction of 2D HOIPs

The band gap is a key parameter that describes the optoelectronic properties of materials, and an accurate prediction of the band gap can accelerate the application of 2D HOIPs in optoelectronics. GNNs, specifically designed for handling graph-structured data, have demonstrated excellent performance in various materials property prediction tasks. GNN converts the crystal structure into graph-structured data, which can flexibly represent chemical bonds between atoms, lattice structures, spatial symmetry, etc. Consequently, we develop a graph convolutional neural network based on the structural properties of 2D HOIPs for accurate band gap prediction, with the workflow illustrated in Fig. 2b. We first screen 2D HOIPs with DFT calculated band gap from an open-access database of 2D perovskites. Subsequently, these 2D HOIPs are transformed into the corresponding crystal graph. The graph data, along with their band gap values, are then input into the model for training and subsequent band gap prediction.

The crystal graph created in this study is an undirected multigraph that allows multiple edges between the same pair of nodes, which is characteristic of crystals due to their periodicity. The crystal graph G is defined as a tuple (V, E), where V defines a set of nodes and E defines a set of edges. Here, $v \in V$ and $e \in E$ are the feature vectors corresponding to atoms and the edges connecting atoms in the crystal, respectively. Each atom i is represented by a feature vector v_i, which encodes properties of the atom, such as electronegativity, covalent radius, atomic volume, etc. Each edge $(i,j)_k$ is represented by a feature vector
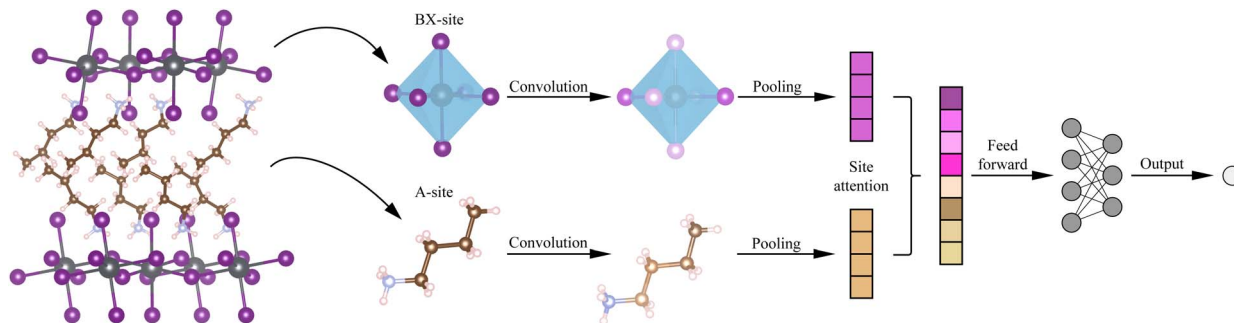
Fig. 4 Architecture of the SATGNN model. Molecules at the A-site and BX-site are convolved separately, allowing each atom to encode information about its local environment. The pooling layer is then performed independently for the molecules at each positions, generating two vectors that represent the A-site and BX-site molecule, respectively. The site-attention layer subsequently updates these vectors and then concatenates them to from a single vector that represents the entire crystal. This vector is then fed into the fully connected layers, followed by the output layer to provide the prediction.

$e_{(i,j)_k}$, which corresponds to the k-th bond connecting atom i and atom j, and encodes the distance between these atoms. For the band gap prediction task, where the target property is highly sensitive to the full crystal geometry and the local bonding environment at all lattice sites (A, B, and X), we employed a site-attention-based graph convolutional neural network (SATGNN) to learn directly from the crystal graph. This architecture enables the model to effectively capture complex interatomic interactions and local distortions critical for accurate band gap estimation. As illustrated in Fig. 4, SATGNN comprises three main components: convolution layers, a pooling layer, and a site-attention layer.

Based on the crystal graph G, we design a site-attention-based graph convolutional neural network, named SATGNN, which consists of three major components: convolution layers, a pooling layer, and a site-attention layer, as shown in Fig. 4.

In the convolution layer, atomic feature vectors are iteratively updated by convolution with surrounding atoms and bonds, enabling the capture of complex interactions within the material. Notably, the structure of 2D HOIPs is different from that of 3D materials, in which the A-site organic cations and the BX-site octahedra are stacked alternately in a layered structure to form 2D HOIPs. We utilize ^{A}v and ^{BX}v to distinguish atoms in 2D HOIPs into A-site atoms and BX-site atoms. In order to accurately capture the features and interactions of the various layers within the 2D HOIPs, we perform convolution separately on the atoms located at different positions, which can be represented by the following equation.

$$
\begin{aligned}
^{A}\boldsymbol{v}_i^{(t+1)} &= \mathrm{Conv}\left(^{A}\boldsymbol{v}_i^{(t)},\, ^{A}\boldsymbol{v}_j^{(t)},\, \boldsymbol{e}_{(i,j)_k}\right)\, ^{A}v_i,\, ^{A}\boldsymbol{v}_j \in\, ^{A}\boldsymbol{v} \\
^{BX}\boldsymbol{v}_i^{(t+1)} &= \mathrm{Conv}\left(^{B}\boldsymbol{v}_i^{(t)},\, ^{X}\boldsymbol{v}_j^{(t)},\, \boldsymbol{e}_{(i,j)_k}\right)\, ^{B}v_i,\, ^{X}\boldsymbol{v}_j \in\, ^{BX}\boldsymbol{v}
\end{aligned}
\tag{1}
$$

The choice of convolution functions in neural network architecture has a great impact on prediction performance. Previous graph convolution functions failed to account for atoms at different positions in 2D HOIPs, resulting in inadequate representation of local features and loss of spatial information, thereby reducing the accuracy of the model. Inspired by

the CGCNN model developed by Xie $et\ al.$,[9] we perform convolution using

$$
\begin{aligned}
^{A}\boldsymbol{v}_i^{(t+1)},\ ^{BX}\boldsymbol{v}_i^{(t+1)} =& \sum_{j,k} \boldsymbol{\sigma}\left(\left(^{A}\boldsymbol{x}_{(i,j)_k}^{(t)},\ ^{BX}\boldsymbol{x}_{(i,j)_k}^{(t)}\right) \boldsymbol{W}_1^{(t)} + \boldsymbol{b}_1^{(t)}\right) \\
& \odot g\left(\left(^{A}\boldsymbol{x}_{(i,j)}^{(t)},\ ^{BX}\boldsymbol{x}_{(i,j)}^{(t)}\right) \boldsymbol{W}_2^{(t)} + \boldsymbol{b}_2^{(t)}\right) \\
& + \left(^{A}\boldsymbol{v}_i^{(t)},\ ^{BX}\boldsymbol{v}_i^{(t)}\right)
\end{aligned}
\tag{2}
$$

where $^{A}\boldsymbol{x}_{(i,j)_k}^{(t)}$ and $^{BX}\boldsymbol{x}_{(i,j)_k}^{(t)}$ are constructed by concatenating atoms at the A-site and BX-site in 2D HOIPs along with their neighboring atoms that belong to the same position and the edge vectors between them. $\boldsymbol{W}_1^{(t)}$ and $\boldsymbol{W}_2^{(t)}$ are the learnable weight matrices of the t-th layer. $\boldsymbol{b}_1^{(t)}$ and $\boldsymbol{b}_1^{(t)}$ are the learnable biases of the t-th layer, and $\boldsymbol{\sigma}$ and $\boldsymbol{g}$ are nonlinear activation functions. In eqn (2), the $\boldsymbol{\sigma}(\cdot)$ function serves as a learned weight matrix to differentiate interactions between neighbors, and adding $(^{A}\boldsymbol{v}_i^{(t)},\ ^{BX}\boldsymbol{v}_i^{(t)})$ makes learning deeper networks easier. After multiple convolutions, the feature vector of each atom incorporates the surrounding environmental information corresponding to its spatial position within the 2D HOIP structure.

The pooling layer is then used to generate the overall feature vectors ^{A}v and ^{BX}v of the A-site and BX-site of the 2D HOIPs. Pooling is also applied to atoms separately according to their different positions in 2D HOIPs, which is represented by the following equation,

$$
\begin{aligned}
^{A}\boldsymbol{v} &= \mathrm{Pool}\left(\boldsymbol{v}_0^{(T)}, \boldsymbol{v}_1^{(T)}, \ldots, \boldsymbol{v}_N^{(T)}\right) \boldsymbol{v}_{[0,1,\ldots,N]} \in\, ^{A}\boldsymbol{v} \\
^{BX}\boldsymbol{v} &= \mathrm{Pool}\left(\boldsymbol{v}_0^{(T)}, \boldsymbol{v}_1^{(T)}, \ldots, \boldsymbol{v}_N^{(T)}\right) \boldsymbol{v}_{[0,1,\ldots,N]} \in\, ^{BX}\boldsymbol{v}
\end{aligned}
\tag{3}
$$

In this work, the mean function is employed as the pooling operation to aggregate information from different atomic positions in 2D HOIPs. This approach can effectively reduce the dimensionality of the feature space while preserving the fundamental characteristics of atomic configurations.

In 2D HOIPs, the A-site organic cations influence the band gap of the material through steric hindrance and electronic effects, while the BX-site octahedra modulate the band gap via

octahedral distortions and the selection of different elements.[26],[27] Since the A-site organic cations and BX-site octahedra affect the band gap of 2D HOIPs through distinct mechanisms, significant differences exist in their effects on the band gap of the material. To effectively capture the varying degrees of contribution from the A-site organic cations and the BX-site octahedra to the material properties, we design the site-attention layer based on the attention mechanism. The attention mechanism was originally proposed in the field of Natural Language Processing (NLP) to improve neural network translation models.[28] Subsequently, its application has expanded to various tasks beyond NLP.[29–32] Here, we introduce the attention mechanism for the first time to predict the band gap of 2D HOIPs, allowing the model to more effectively capture the distinct contributions of different layer components to the material properties, thereby enhancing the model's understanding of the material structure. The site-attention layer updates the vectors $^A v$ and ^{BX}v based on the different contributions of A-site and BX-site components to material properties, generating new feature vectors $^A v'$ and $^{BX}v'$. Subsequently, we concatenate these two vectors to obtain the overall feature vector v_g for the crystal. This process is represented by the following equation.

$$\left(^A v',\ ^{BX}v'\right) = \mathbf{Attention}\left(^A v,\ ^{BX}v\right)\ v_g = {}^A v' \parallel {}^{BX}v' \tag{4}$$

The specific formula for the attention function is as follows,

$$\mathbf{Attention}\left(^A v, ^{BX}v\right) = \mathbf{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V$$
$$Q, K, V = W^Q, W^K, W^V\left(^A v, ^{BX}v\right) \tag{5}$$

where $\left(^A v, ^{BX}v\right)$ represents a vector with the components $^A v$ in the first row and ^{BX}v in the second row. W^Q, W^K, and W^V are trainable parameter matrices. Q, K, and V represent the vectors obtained from $\left(^A v, ^{BX}v\right)$ after transformation by these matrices, maintaining the same dimensionality. Additionally, $\sqrt{d}$ is the scaling factor, where d is equal to the dimension of $\left(^A v, ^{BX}v\right)$. In eqn (5), the $\mathbf{softmax}(\cdot)$ function is used to calculate the attention coefficient between vectors $^A v$ and ^{BX}v. In addition to the convolution layer, pooling layer, and site-attention layer, we add a fully connected layer to capture the complex mapping between the crystal structure and band gap. Finally, an output layer is used to predict the band gap ŷ.

The objective of training the neural network model is to find a set of parameters that minimizes the difference between the predicted band gap value yŷ^ and the DFT calculated band gap value y, defined by the loss function. During the training process, the model parameters are updated through multiple iterations until the loss function reaches a satisfactory level. Once trained, our SATGNN model enables rapid and generalizable band gap predictions across a wide range of 2D HOIP compositions. This capability not only accelerates property estimation for novel compounds but also facilitates extrapolation to hypothetical molecules beyond those previously reported or characterized by DFT.

The dataset used in this study is obtained from an open-access 2D perovskites database developed by Marchenko et al.,[8] containing a total of 849 2D perovskites. Accessed in March 2024, approximately 70% of these entries have DFT-calculated band gap values available. We screened the data to maximize usable samples by excluding entries without DFT band gap calculations or those not classified as 2D HOIPs. After screening, 491 2D HOIPs remained for analysis. All band gap data for these 491 HOIPs are calculated consistently within this database using the DMol3 module of Materials Studio with the DNP + atomic basis set and explicit spin–orbit coupling, ensuring methodological uniformity and reliable agreement with experimental results. This uniform dataset maintains sufficient chemical and structural diversity to support robust and generalizable structure–property relationships. The filtered dataset is then split into a training and test set, with 80% of the data used for model training and the remaining 20% reserved for independent performance evaluation. All test samples are strictly excluded from the training process, ensuring no data overlap and preventing leakage, thereby enabling a rigorous and unbiased performance assessment.

During the training process, the model exhibiting the best performance on the test set is selected to construct the final model. Fig. 5a shows the decline curve of the train loss and the test loss throughout the training process. It can be clearly seen that overfitting does not occur. Moreover, our model, SATGNN, achieves a MAE of 0.058 eV on the test set. Fig. 5b presents the corresponding parity plot, and approximately 90% of the crystals are predicted within 0.025 and 0.06 eV errors. In comparison, Marchenko et al. developed a ML model based on the 2D perovskites database to predict their band gap, achieving a MAE of 0.103 eV. Our model performance improves by 43.7%. We also compare other excellent GNN models that have been reported for material property prediction, as shown in Table 2. Our model outperforms all other models on the band gap prediction task, showing an improvement of 15.9% over CGCNN and 25.6% over GATGNN in terms of MAE. Given the comparison, our SATGNN model demonstrates reliable band gap predictions and has potential applications in predicting other properties of 2D HOIPs. Additionally, it is important to note that the dataset used in this study is relatively small for constructing neural network models. Furthermore, the structures of 2D HOIPs are very similar, and the range of band gap values calculated by DFT in the dataset is very small. Thus, we argue that our model can effectively learn useful patterns from this dataset. This inference is supported by the experimental results presented above, as our model demonstrates good predictive performance without overfitting.

We next investigate which aspects of our model contribute to performance improvement. For this purpose, we compare the performance of different models in predicting band gap, including a variant of our model, 2D-GCNN, which excludes the site-attention mechanism, as shown in Table 3. It is evident that without employing the convolution method specifically designed for the structural characteristics of 2D HOIPs, CGCNN shows significantly lower performance in band gap prediction. In contrast, the 2D-GCNN model, which employs this
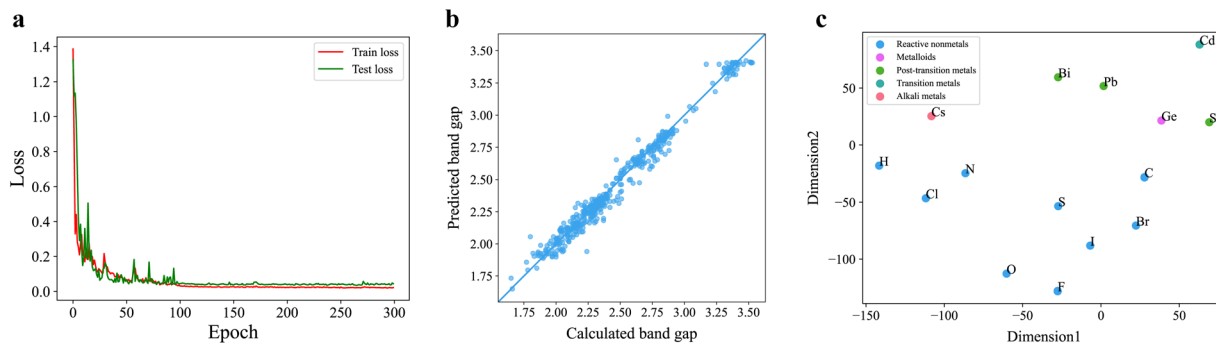
Fig. 5 (a) Trend of loss decline during training for the SATGNN model. (b) Parity plot showing the predicted band gap against the DFT calculated band gap. (c) Visualization of the element representations learned from the 2D HOIP dataset, with elements colored according to their elemental groups.

Table 2 Performance comparison over the band gap prediction problem of our model compared to other models

Model	MAE	Units
Marchenko *et al.*	0.103	eV
CGCNN	0.069	eV
GATGNN	0.078	eV
SATGNN	**0.058**	eV

Table 3 Performance comparison of band gap predictions for CGCNN, 2D-GCNN, and SATGNN

Model	MAE	Units
CGCNN	0.069–0.073	eV
2D-GCNN	0.062–0.065	eV
SATGNN	**0.058–0.060**	eV

convolution method, achieves approximately 10% improvement in performance compared to CGCNN. This phenomenon indicates that the convolution method we designed effectively addresses the layered structure of 2D HOIPs and accurately captures and utilizes features from different sites within these materials. Moreover, SATGNN, which integrates the convolution method specifically designed for 2D HOIPs and the site-attention mechanism, demonstrates approximately 6% improvement in performance compared to 2D-GCNN. The incorporation of the site-attention mechanism leads to further improvements in model performance, as it is consistent with the physical intuition that the A-site cation and BX-site octahedra in 2D HOIPs contribute differently to the material properties. In the band gap prediction problem for 2D HOIPs, our convolution method combined with the site-attention mechanism yields significant improvements in model performance, achieving superior results compared to other excellent models.

Model interpretability is highly desirable in materials science, especially for complex neural networks, as it provides valuable insights to guide material design. In our model, each atom is initially represented by a feature vector $\mathbf{v}_i$, which encodes basic elemental properties such as group number,

electronegativity, covalent radius, and other relevant descriptors.[9] This vector corresponds to the atom's element type and is further refined during training *via* the embedding layer. Since 2D HOIPs share similar crystal structures but vary widely in composition, understanding how the model encodes and organizes elemental information is crucial. To probe the model's learned chemical representation, we extract the atom vectors $\mathbf{v}_i^{(0)}$ from the output of the embedding layer prior to any convolutional operations. At this stage, these vectors depend solely on element identity without incorporating structural context, thereby reflecting the model-inferred similarity between elements based on their roles in forming 2D HOIPs.

The dataset that we used in this study contains 491 different 2D HOIPs, comprising a total of 15 elements: H, C, N, Pb, I, Sn, Ge, Br, Cl, O, S, F, Bi, Cs, and Cd. After training with the 2D HOIP dataset, we employ *t*-SNE[33] to project these element representations onto the 2D plane, as shown in Fig. 5c. From the top to the bottom of Fig. 5c, the electronegativity of elements increases progressively, which aligns with the trends observed in the periodic table. Meanwhile, it can be seen that elements are grouped according to their position in the structure of 2D HOIPs. To quantitatively assess the local organization of these embeddings, we conduct a *k*-nearest neighbor ($k = 3$) purity analysis[34] within the 2D *t*-SNE projection space, grouping elements by their crystallographic sites (A-, B-, and X-sites). This analysis yields an average purity of 75%, demonstrating that the learned embeddings meaningfully preserve local coherence with respect to the elemental site roles in the crystal structure. Specifically, the elements at the B-site (Pb, Sn, Ge, Bi, and Cd) are clustered in the upper right part of Fig. 5c. The primary components of the organic cation at the A-site (C, N, and H) are concentrated in the central part of Fig. 5c, while the halogen elements at the X-site (Cl, I, F, and Br) are predominantly located in the lower part. In addition, as a Group 12 element, Cd is positioned relatively closer to Group 14 elements (Pb, Ge, and Sn), while being more distantly located from the Group 15 element Bi. Similarly, the element Cs can serve as an A-site element in 2D HOIPs, and since both Cs and H belong to Group 1, the model places them in close proximity. The X-site elements (Br, I, and F) from Group 17 are also observed to be

near the A-site elements (O and S) from Group 16. The above phenomenon indicates that the element representations learned from our model incorporate the structural information from 2D HOIPs while preserving the information related to periodic trends of elements in the periodic table, which can further demonstrate the model's reliability in band gap prediction.

3. Conclusions

We report an ensemble learning model and a graph convolutional neural network model for different tasks. An ensemble learning model is employed to classify lead iodide-based HOIPs into LD or 2D categories. The optimal XGBoost classifier achieves 88% classification accuracy, identifying Max-AbsEStateIndex (5.68–13.04), Chi2n (0.33–2.50), and Kappa2 (5.06–17.96) as key topology- and electronic-structure-related features. These insights provide practical guidelines for selecting organic components that promote 2D HOIP formation and effectively narrow the search space in high-throughput screening. Additionally, we propose a graph convolutional neural network model called SATGNN for predicting the band gap of 2D HOIPs. This model outperforms baseline approaches by at least 15.9% and generalizes well to unseen compositions, owing to a convolution function tailored for 2D HOIPs that captures position-dependent interactions and a site-attention mechanism that distinguishes the contributions of different layer components. Visualization of the learned element embeddings further reveals that SATGNN can discriminate element positions within the 2D HOIP framework while retaining periodic table trends, thereby enabling reliable property prediction to guide materials design. Although trained on published DFT data, both models support high-throughput screening, generalization to novel compositions, and extraction of interpretable design rules, collectively establishing a predictive framework for the discovery of new 2D HOIPs with targeted optoelectronic properties.

Conflicts of interest

The authors have no conflicts to disclose.

Data availability

The primary data for this article, including 2D perovskite data files (.cif), are available in the "2D Perovskites Database" at http://pdb.nmse-lab.ru/. Based on this database, we systematically organized the perovskite data. The datasets and source code supporting this study are openly available on GitHub and have also been archived on Zenodo. The archived version associated with this manuscript can be accessed at https://doi.org/10.5281/zenodo.17018718.

Acknowledgements

References

1 A. Kojima, K. Teshima, Y. Shirai and T. Miyasaka, Organometal halide perovskites as visible-light sensitizers for photovoltaic cells, *J. Am. Chem. Soc.*, 2009, **131**(17), 6050–6051.

2 NREL, *Best research-cell efficiencies*, 2024, online available, https://www.nrel.gov/pv/cell-efficiency.html.

3 L. Dou, A. B. Wong, Y. Yu, M. Lai, N. Kornienko, S. W. Eaton, A. Fu, C. G. Bischak, J. Ma, T. Ding, *et al.*, Atomically thin two-dimensional organic-inorganic hybrid perovskites, *Science*, 2015, **349**(6255), 1518–1521.

4 H. Tsai, W. Nie, J.-C. Blancon, C. C. Stoumpos, R. Asadpour, B. Harutyunyan, A. J. Neukirch, R. Verduzco, J. J. Crochet, S. Tretiak, *et al.*, High-efficiency two-dimensional ruddlesden–popper perovskite solar cells, *Nature*, 2016, **536**(7616), 312–316.

5 N. Ashari-Astani, F. Jahanbakhshi, M. Mladenovic, A. Q. Alanazi, I. Ahmadabadi, M. R. Ejtehadi, M. I. Dar, *et al.*, Ruddlesden–popper phases of methylammonium-based two-dimensional perovskites with 5-ammonium valeric acid $AVA_2MA_{n-1}Pb_nI_{3n+1}$ with n= 1, 2, and 3, *J. Phys. Chem. Lett.*, 2019, **10**(13), 3543–3549.

6 R. Lyu, C. E. Moore, T. Liu, Y. Yu and Y. Wu, Predictive design model for low-dimensional organic–inorganic halide perovskites assisted by machine learning, *J. Am. Chem. Soc.*, 2021, **143**(32), 12766–12776.

7 S. Yuan, Y. Liu, J. Lan, W. Yang, H. Xiong, W. Li and J. Fan, Accurate dimension prediction for low-dimensional organic–inorganic halide perovskites *via* a self-established machine learning strategy, *J. Phys. Chem. Lett.*, 2023, **14**(32), 7323–7330.

8 E. I. Marchenko, S. A. Fateev, A. A. Petrov, V. V. Korolev, A. Mitrofanov, A. V. Petrov, E. A. Goodilin and A. B. Tarasov, Database of two-dimensional hybrid perovskite materials: open-access collection of crystal structures, band gaps, and atomic partial charges predicted by machine learning, *Chem. Mater.*, 2020, **32**(17), 7383–7388.

9 T. Xie and J. C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, *Phys. Rev. Lett.*, 2018, **120**(14), 145301.

10 S.-Y. Louis, Y. Zhao, A. Nasiri, X. Wang, Y. Song, F. Liu and J. Hu, Graph convolutional neural networks with global attention for improved materials property prediction, *Phys. Chem. Chem. Phys.*, 2020, **22**(32), 18141–18148.

11 C. Chen, W. Ye, Y. Zuo, C. Zheng and S. P. Ong, Graph networks as a universal machine learning framework for molecules and crystals, *Chem. Mater.*, 2019, **31**(9), 3564–3572.

12 K. Choudhary and B. DeCost, Atomistic line graph neural network for improved materials property predictions, *npj Comput. Mater.*, 2021, **7**(1), 185.

13 Z. Gao, Y. Bai, M. Wang, G. Mao, X. Liu, P. Gao, W. Yang, X. Ding and J. Yao, Novel prediction model of band gap in

organic–inorganic hybrid perovskites based on a simple cluster model database, *J. Phys. Chem. C*, 2022, **126**(31), 13409–13415.

14 X. Luo, Y. Hu, Z. Lin, X. Guo, S. Zhang, C. Shou, Z. Hu, X. Zhao, Y. Hao and J. Chang, Theoretical analysis of all-inorganic wide bandgap perovskite/sn-based narrow bandgap perovskite tandem solar cells, *Sol. RRL*, 2023, **7**(10), 2300081.

15 C. W. Ahn, J. H. Jo, J. S. Choi, Y. H. Hwang, I. W. Kim and T. H. Kim, Heteroanionic lead-free double-perovskite halides for bandgap engineering, *Adv. Eng. Mater.*, 2023, **25**(1), 2201119.

16 J.-H. Im, J. Chung, S.-J. Kim and N.-G. Park, Synthesis, structure, and photovoltaic property of a nanocrystalline 2H perovskite-type novel sensitizer $(CH_3CH_2NH_3)PbI_3$, *Nanoscale Res. Lett.*, 2012, **7**, 353.

17 A. F. Xu, R. T. Wang, L. W. Yang, N. Liu, Q. Chen, R. LaPierre, N. I. Goktas and G. Xu, Pyrrolidinium containing perovskites with thermal stability and water resistance for photovoltaics, *J. Mater. Chem. C*, 2019, **7**(36), 11104–11108.

18 M.-Y. Zhu, L.-X. Zhang, J. Yin, J.-J. Chen and L.-J. Bie, A fluorescence quenching sensor for Fe^{3+} detection using $(C_6H_5NH_3)_2Pb_3I_8\cdot 2H_2O$ hybrid perovskite, *Inorg. Chem. Commun.*, 2019, **109**, 107562.

19 RDKit, *Open-source cheminformatics*, online available, https://www.rdkit.org.

20 C. M. Bishop, *Pattern Recognition and Machine Learning*, 1st edn, Springer, New York, 2006.

21 S. M. Lundberg and S.-I. Lee, A unified approach to interpreting model predictions, in *Advances in Neural Information Processing Systems* vol. 30, 2017, pp. 4765–4774.

22 L. H. Hall, B. Mohney and L. B. Kier, The electrotopological state: structure information at the atomic level for molecular graphs, *J. Chem. Inf. Comput. Sci.*, 1991, **31**(1), 76–82.

23 L. H. Hall and L. B. Kier, *The Molecular Connectivity Chi Indexes and Kappa Shape Indexes in Structure-Property Modeling*, 1st edn John Wiley & Sons, Ltd, 1991.

24 A. Lemmerer and D. G. Billing, Synthesis, characterization and phase transitions of the inorganic–organic layered perovskite-type hybrids $[(C_nH_{2n+1}NH_3)_2PbI_4]$, n = 7, 8, 9 and 10, *Dalton Trans.*, 2012, **41**, 1146–1157.

25 D. G. Billing and A. Lemmerer, Inorganic–organic hybrid materials incorporating primary cyclic ammonium cations: The lead iodide series, *CrystEngComm*, 2007, **9**, 236–244.

26 M. Azeem, Y. Qin, Z.-G. Li and W. Li, Cooperative B-site octahedral tilting, distortion and A-site conformational change induced phase transitions of a 2D lead halide perovskite, *Mater. Chem. Front.*, 2021, **5**(20), 7587–7594.

27 M. A. Kuddus Sheikh, F. Maddalena, D. Kowal, M. Makowski, S. Mahato, R. Jedrzejewski, R. Bhattarai, W. Drozdowski, *et al.*, Effect of dual-organic cations on the structure and properties of 2D hybrid perovskites as scintillators, *ACS Appl. Mater. Interfaces*, 2024, **16**(19), 25529–25539.

28 A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, Attention is all you need, in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

29 A. Gulati, J. Qin, C.-C. Chiu, N. Parmar, Y. Zhang, J. Yu, W. Han, S. Wang, Z. Zhang, Y. Wu, and R. Pang, *"Conformer: Convolution-augmented transformer for speech recognition"*, 2020. online available, https://arxiv.org/abs/2005.08100.

30 W. Yu, M. Luo, P. Zhou, C. Si, Y. Zhou, X. Wang, J. Feng, and S. Yan, "Metaformer is actually what you need for vision", in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 10809–10819.

31 L. Dong, N. Yang, W. Wang, F. Wei, X. Liu, Y. Wang, J. Gao, M. Zhou, and H.-W. Hon, "Unified language model pre-training for natural language understanding and generation", in *Advances in Neural Information Processing Systems*, vol. 32, 2019.

32 U. Naseem, I. Razzak, K. Musial and M. Imran, Transformer based deep intelligent contextual embedding for twitter sentiment analysis, *Future Gener. Comput. Syst.*, 2020, **113**, 58–69.

33 L. van der Maaten and G. Hinton, Visualizing data using t-SNE, *J. Mach. Learn. Res.*, 2008, **9**(86), 2579–2605.

34 J. Chen, H.-r. Fang and Y. Saad, Fast approximate knn graph construction for high dimensional data *via* recursive lanczos bisection, *J. Mach. Learn. Res.*, 2009, **10**, 1989–2012.