Fast and Accurate Artificial Neural Network Potential Model for MAPbI₃ Perovskite Materials

Hsin-An Chen and Chun-Wei Pao *

Research Center for Applied Sciences, Academia Sinica, Taipei 11529, Taiwan

ABSTRACT: Hybrid organic–inorganic perovskite materials are promising materials for photovoltaic and optoelectronic applications. Nevertheless, the construction of a computationally efficient potential model for atomistic simulations of perovskite with high fidelity to ab initio calculations is not a trivial task given the chemically complex nature of perovskite in terms of its chemical components and interatomic interactions. In the present study, we demonstrate that artificial neural network (ANN) models can be employed for efficient and accurate potential energy evaluation of MAPbI₃ perovskite materials. The ANN models were trained using training sets composed of thousands of atomic images of tetragonal MAPbI₃ crystals, with their respective energies and atomic forces obtained from ab initio calculations. The trained ANN models were validated by predicting the lattice parameters and energies/atomic forces of cubic MAPbI₃ perovskite and had excellent agreement with ab initio calculations. The phonon modes could also be extracted using the trained ANN model with good agreement with ab initio calculations, provided that the atomic forces were incorporated into the training processes. Finally, we demonstrate that artificial neural network (ANN) models can be employed for exhaustively sampling the configuration spaces of chemically complex materials for predictions of thermodynamic properties and phase stabilities.

INTRODUCTION

Hybrid organic–inorganic perovskites have recently stimulated significant research effort worldwide because of their extraordinary optical and optoelectronic properties, including their high optical absorption coefficient in the visible light region, tunable band gap energy, low exciton binding energy, and long, balanced carrier diffusion lengths. Perovskite materials have been widely employed as the active material in solar cells, light-emitting diodes, lasers, photodetectors, or photocatalysts. Solar cells made from perovskites have reached remarkably high power conversion efficiencies of more than 20%, making them promising photovoltaic materials for replacing silicon and reducing both device fabrication costs and weight. Among all perovskites, methylammonium lead iodide (CH₃NH₃PbI₃ or MAPbI₃) is one of the most-widely studied materials and has a band gap energy of ~1.6 eV and absorption spectrum closely matched to the solar spectrum. In devices fabricated from perovskite materials, such as solar cells, the performance of the device critically relies on the morphology and grain size of perovskite materials. The morphologies of perovskite films are sensitive to their fabrication protocols, hence, elucidation of the processing–structure–property (PSP) relationships for perovskite films is the key to controlling the morphology of the film for commercial deployment of perovskite-based devices. Atomistic scale simulation is a powerful tool for elucidating the PSP relationship of perovskite materials. Ab initio calculations are considered to be the most accurate of atomistic simulation methods; however, the system size of interest for perovskites is of the order of at least several tens of nanometers, which is well beyond the reach of ab initio calculations.

On the other hand, classical molecular dynamics (MD) simulation can handle systems of tens/hundreds of thousands atoms, which is ideal for studying the structural and morphological properties of perovskite materials. However, the reliability of classical MD simulation depends on the accuracy of the classical interatomic force fields, and achieving such accuracy becomes challenging with the increasing complexity of the chemical space. In the MAPbI₃ perovskite, five elements are present in the system (Pb, I, C, N, and H atoms) with a mixture of covalent, van der Waals, ionic, and hydrogen bonding interactions. Several earlier works have parametrized classical force fields for MAPbI₃ crystals by decomposing interatomic interactions into analytical functions describing aforementioned interactions. The accuracy of force fields therefore critically relies on the interaction functions selected, and contributions from effects such as atomic polarization or hydrogen bonding were not taken into accounts in these classical force fields. As a result, despite that these force fields in general show reasonable agreements with experiments or first-principle calculations in certain perovskite
properties such as lattice parameters, there still is little room of improvement after benchmark testing with first-principle calculations.25 Furthermore, these force fields cannot be utilized to study complex structural transitions such as two-step processes of perovskite21,22 or perovskite materials with complex compositions such as mixed ion perovskite alloying FA, MA, and Cs anions and Br or Cl cations.27−29 Hence, the development of an atomistic force field that can describe the interactions of a chemically complex system with high fidelity to first-principle calculations is mandatory for understanding the structure and morphology of perovskite materials.

By harnessing the power of machine learning, interatomic force fields based on artificial neural networks (ANN) have drawn increasing attention.30 It has been demonstrated that a well-trained ANN model can successfully replicate the system energies of given structures obtained from first-principles calculations. In contrast to aforementioned classical force fields, the ANN potential models backed by the universal approximation theorem allows encapsulation of all interatomic interactions into one neural network model, offering force fields with high fidelity to the first-principle calculations. To train an ANN model of given material, a training set composed of thousands of images of a given material with atomistic configurations labeled with respective energies from first-principles calculations is required. Datasets such as the Materials Project33 compile a massive amount of data of materials from first-principles calculations and have been successfully utilized to train models to predict material properties such as lattice parameters, with good agreement to those obtained from density functional theory (DFT) calculations. Hence, the present study demonstrates that the ANN-based energy/force evaluator is efficient and reliable and can be readily implemented in systems with complex chemical compositions, such as mixed-ion perovskite materials.

Figure 1. (a) Schematic of the network architecture in the ANN scheme. (b) Structure of tetragonal perovskite MAPbI3 (dark gray: Pb, brown: I, light gray: C, white: H, and blue: N).

**Materials and Methods**

**Neural Network Function and Descriptor.** The ANN is a computational model inspired by the biological neural network.39 In the present work, we employed the energy partitioning scheme proposed by Behler and Parrinello to construct an ANN model for efficiently computing system energies.30 One of the advantages of this energy partition scheme is the transferability of the ANN model, namely, the trained ANN model can, in principle, be utilized for atomistic simulations of systems with identical chemical compositions and an arbitrary number of atoms. Figure 1a depicts the scheme of the ANN model employed in the present study. The objective of the ANN model is to predict the system energy from given atomic coordinates with high fidelity to the corresponding ab initio calculations. As depicted in the upper panel of Figure 1a, in the first step, the atomic coordinates of all of the atoms \(A_i(x_i)\) are transformed into a set of rotationally/translationally invariant descriptor (or fingerprint) functions as the input feature vector for the atomic neural network function. The output of the atomic neural network function of atom \(i\) is its atomic energy \(E_i\) and the summation of the atomic energies is the overall system energy \(E\), namely, \(E = \sum_i E_i\). The architecture of each atomic neural network is schematically displayed in the lower panel of Figure 1a. The atomic neural network is composed of the input layer, the hidden layers, and the output layer. The input layer is composed of a series of atomic descriptor functions transformed from the Cartesian coordinates of the individual atoms and their neighbors that depict their atomistic chemical environments. The output layer will output the atomic energy of each atom. Each of the layers in the neural network is composed of a finite number of nodes (neurons), and all of the nodes are connected (see the arrows in the lower panel of Figure 1a) via a set of weighting parameters \(w\), so that the vector of all of the nodes in the \(i\)th layer can be expressed as

\[
\sigma_i = f_w(\sigma_{i-1})
\]
where $W_i$ is the weighting matrix connecting the nodes in layer $i$ and layer $i - 1$. The atomic energy of the $i$th atom obtained for an atomic neural network (the lower panel of Figure 1a) with $M$ hidden layers can be written as

$$E_i = N(I_i, \{W\}) = f_o(W_{M} \cdot \sum_{j \neq i} f_i(W_{M-1} \cdot \cdots f_2(WI) \cdots))$$

where $N(I_i, \{W\})$ is called the neural network function and $I_i$ is the feature vector of the $i$th atom. We note that each chemical species in the system (e.g., Pb, I, C, N, and H in the present study) should have their own atomic neural network function $N$ (namely, $N_{Pb}$, $N_{I}$, $N_{C}$, $N_{N}$, and $N_{H}$ in the present study). In the present study, we selected the hyperbolic tangent function $\tanh(\sigma)$ as the activation function $f_i$. Note that the rectifier linear unit (ReLU) activation function has been extensively employed as the activation function of deep neural network because of effective suppression of gradient vanishing upon training; however, in the present study, the hyperbolic tangent function was employed to prevent discontinuous derivatives of neural network function. In performing atomistic simulations, the derivatives of the ANN potential model with respect to atomic coordinates are required to be evaluated as atomic forces exerted to each atom. The ReLU function may induce discontinuity in atomic forces, which is unfavorable for atomistic simulations. Continuous ReLU alternatives such as the softplus function can avoid the discontinuous force problems; however, in the Behler and Parrinello energy partitioning scheme, the output of atomic neural network function $N(I_i, \{W\})$ are atomic energies, which are usually negative in magnitude; as a result, it would be impossible to implement the ReLU activation functions directly. To implement ReLU activation functions in the energy partitioning scheme employed in the present study, substantial modifications in the overall neural network architectures would be required, such as including a separate sum pooling layer or linear regression employed in SchNet and HIP-NN while utilizing ReLU activation functions in the atom-wise or interaction layers. This could potentially further promote the training of ANN models for complex systems and is a promising future research direction. The weighting matrix set $\{W\}$ connecting the hidden layers in the neural network are key for efficient energy/force predictions, and these parameters must be trained to become a valid model that can be utilized for subsequent atomistic simulations.

The feature vector $I_i$ of the $i$th atom is composed of a series of descriptor functions that transform the Cartesian atomic coordinates of the $i$th atom into translationally/rotationally invariant fingerprints specifying its chemical environment. In this work, the Gaussian descriptor functions proposed by Behler were employed as the descriptor functions for the input layer of the atomic neural network functions $N$. In the present study, the Gaussian descriptor functions were divided into two categories, namely, the radial descriptor $G^{II}$ and the angular descriptor $G^{IV}$. The radial and angular descriptors can be expressed as

$$G^{II}_{i} = \sum_{i \neq j} \exp \left( \frac{-\eta R_{ij}^2}{\zeta c_i^2} \right) f_{c}(R_{ij})$$

$$G^{IV}_{i} = 2^{1-i} \sum_{j \neq i} (1 + \lambda \cos \theta_{ij}) \exp \left( \frac{-R_{ij}^2 + \lambda^2 R_{ij}^2 + R_{ij}^2}{\zeta c_i^2} \right) f_{c}(R_{ij}) f_{\lambda}(\theta_{ij})$$

where $R_{ij}$ is the cutoff distance for the descriptor functions and $\eta$ and $\zeta$ are predefined parameters for the descriptors. The cutoff function $f_{c}$ can be expressed as

$$f_{c}(R) = \begin{cases} 
0.5 \left[ 1 + \cos \left( \frac{\pi R}{R_c} \right) \right], & R < R_c \\
0, & R > R_c 
\end{cases}$$

The parameters of the Gaussian descriptors employed in the present study are compiled in Tables S1 (radial $G^{II}$-type descriptor) and S2 (angular $G^{IV}$-type descriptor) in the Supporting Information. The cut-off radius for the potential training was set to be 6.5 Å, which is approximately twice the Pb–I bond length to ensure the second-nearest neighbors in the backbone part to be included for fingerprint evaluations. The feature vector for the atomic neural network is composed of twenty $G^{II}$-type descriptor functions and sixty $G^{IV}$-type descriptor functions. The feature vectors $I_i$ were computed by accumulating contributions to all eighty descriptors from each neighboring atoms within the cut-off sphere of atom $i$ compiled in Tables S1 and S2 in the Supporting Information. Note that in Tables S1 and S2, parameters with identical parameters were to specify contributions from neighboring atoms or atom pairs of different atomic species (for radial descriptors) or pair combinations (for angular descriptors, because the atom $i$ serves as the center atom of the atomic triplet consisting the angle). Take a radial descriptor (Table S1) as an example, descriptor no. 1 (2) in Table S1 accumulates contributions from all neighboring C (H) atoms (within the cut-off radius) with specified parameters ($\eta = 5$) and will return zero value if there are no neighboring C (H) atoms inside the cut-off sphere. Because there were five elements for MAPbI$_3$ (C, N, H, Pb, and I), for each set of radial Gaussian descriptor, there were five possible elements as neighbors. This is the reason why there were five radial descriptors with identical parameters. Similarly, angular descriptors with identical parameters were to accumulate contribution from atom pairs specified in Table S2 in the Supporting Information.

The ANN model must be trained to enable its use for predicting system energies/forces with high accuracy. During the training processes, the weighting parameter set $\{W\}$ is optimized by minimizing the quadratic error function.

10952
where $\sigma$ is a structure within the training set, $E_{r}^{\text{ref}}$ is the respective reference total energy, $E_{o,k}^{\text{ref}}$ is the reference component $k$ of the atomic force of atom $i$, and $N_{\sigma}$ is the number of atoms in the structure $\sigma$, respectively. We note that the first and second terms on the right-hand side refer to the quadratic errors in the system energies and atomic forces, respectively. The parameter $\alpha$ is the weighting coefficient of the force, which can be set to zero if only system energies are used as the training targets. The weighting parameter sets $\{W\}$ were obtained by minimizing the quadratic error function $\varepsilon$. In the present study, we employed the limited-memory BFGS for bound-constrained optimization (l-BFGS-b) minimizer implemented in the Atomistic Machine-Learning Package\textsuperscript{43} to train the ANN model. The energy and force convergence criteria for the l-BFGS-b minimization processes are described in the following paragraph.

In the present manuscript, tetragonal perovskite MAPbI$_3$ was selected as the structure to train the ANN model (Figure 1b). The lattice parameters of the optimized tetragonal cell were $a = 12.26$ Å and $c = 12.77$ Å, close to the results of previous ab initio calculations\textsuperscript{17,44} and experimental results.\textsuperscript{45,46} The atomic structures in the training set were generated by performing ab initio MD (AIMD) simulations under the NVT ensemble using the Langevin thermostat. Because the neural network function as well as its derivatives with respect to atomic coordinates needs to be evaluated for each atom (up to $10^3$ atoms) every atomistic simulation step (up to $10^3$ to $10^5$ steps), the neural network architectures (i.e., number of hidden layers and number of nodes per layer) were chosen to ensure balance between computational efficiency and model accuracy. In the present study, we trained neural network models of incrementally reduced network size from a network of [10, 10, 10] architecture (namely, three hidden layers and ten neurons per hidden layer), and the architectures specified in Table 1 were chosen because they were among the simplest architectures with accuracy compatible with the [10, 10, 10] model. Note that three different settings for training were considered (Table 1). In settings 1 and 2, we tested the effects of the network architecture on the trained ANN model, whereas in sets 2 and 3, we fixed the neural network architectures and compared training the ANN model using only system energies (setting 2) and training the ANN model using both energies and atomic forces (setting 3). The energy convergence criteria for settings 1–3 were set to 0.001 eV/atom, and the force convergence criterion was set to 0.1 eV/Å in setting 3. To perform atomistic simulations, we implemented the ANN potential model into the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS).\textsuperscript{37} The LAMMPS input files for the ANN potential models can be found in the Supporting Information and the code for the implemented ANN pair style is available on request.

**Training Set Preparation.** Twenty-two separate AIMD simulations of tetragonal MAPbI$_3$ subjected to hydrostatic strains that incrementally increased from $-5$ to $+5\%$ were performed to expand the span of the atomic feature vectors to ensure the transferability of the ANN model. For each AIMD simulation, the system temperature and step size were set to 300 K and 0.5 fs, respectively, and a trajectory of 1000 AIMD steps was collected. Hence, a total number of 22 000 images were collected. Two thousand images were randomly selected from the pool of 22 000 images for the training set for potential training, and another two thousand images were selected for validation. Note that despite the images in the validation sets were extracted from the same pool as the training set, there were literally no or only weak correlation between images in the training and validation sets because validation images from one particular AIMD trajectory were not correlated with training images from the rest 21 AIMD trajectories. In addition, despite validation images from one AIMD trajectory were not totally uncorrelated with training images from the same trajectory, their atomic coordinates deviate from those of the training images with atomic displacements of the order of angstroms, thereby allowing testing the trained ANN models using inputs outside the images selected for the training processes for the model validation purpose.

**Ab Initio Molecular Simulations.** All AIMD calculations and structural minimization were performed using the Vienna Ab initio Simulation Package (VASP).\textsuperscript{48} The generalized gradient approximation was used with the Perdew–Burke–Ernzerhof exchange–correlation functional\textsuperscript{51,52} and the projector augmented wave\textsuperscript{53,54} pseudopotentials. The cutoff energy was set to 400 eV, and a $3 \times 3 \times 3$ Monkhorst–Pack $k$-point mesh was employed. The self-consistent field convergence criterion was set to $\Delta E < 10^{-5}$ eV, where $\Delta E$ is the total energy difference between two steps. The DFT-D3 with Becke–Johnson damping\textsuperscript{55,56} was employed to incorporate the van der Waals interactions between the atoms.

### RESULTS AND DISCUSSION

**Validation and Testing of the Trained ANN Potentials.** The ANN models were trained based on three different training settings (referring to Table 1) and the training set composed of 2000 randomly picked system images. Then, the trained ANN models were validated using the validation set, which consisted of another 2000 randomly picked system images. Figure 2a displays the correlations (or parities, left panels) and errors (right panels) between the potential energies from three trained ANN models and the VASP calculations in the validation set. In the left panels of Figure 1a (the parity plots), the potential energies from the trained ANN models lie on the diagonal line with very small fitting errors. The RMSEs were $\sim 0.001$ eV/atom and the maximum absolute

### Table 1. Neural Network Architectures and Training Targets Employed in the Present Study\textsuperscript{46}

| # | Pb, I | C, N, H | energy training | force training |
|---|---|---|---|---|
| 1 | 10, 5, 5 | 5, 5, 5 | O | X |
| 2 | 10, 5, 5 | 5, 5 | O | X |
| 3 | 10, 5, 5 | 5, 5, 5 | O | O |

“The atomic neural network architectures of Pb, I atoms (2nd column) and C, N, H atoms (3rd column) are specified, respectively. Note that the numbers separated by comma in the network architectures denote the number of nodes in each hidden layer.
errors (MAE) were less than 0.005 eV/atom in the three training settings—much lower than the atomic thermal energies at 300 K (ca. 0.026 eV/atom), demonstrating that these trained ANN models can successfully predict the energies of tetragonal MAPbI₃ crystals without performing computationally expensive DFT calculations. The learning curves of the energy RMSE with respect to the validation set as the function of training set sizes for three training settings can be found in Figure S1 in the Supporting Information. From the learning curve, we can see that for all three training settings a training set of 2000 images already gives very low RMSE in the validation set (∼0.001 eV/atom) with no sign of overfitting.

To test the transferability of the trained ANN models' for different crystal polymorphs of MAPbI₃, we performed separate DFT calculations for MAPbI₃ cubic and orthorhombic perovskite (the ANN models were trained using a tetragonal structure). We first applied the trained ANN model to evaluate the potential energy curves of cubic and orthorhombic perovskite MAPbI₃ against incremental hydrostatic strain of 0.5% from −5 to +5% and compared the respective potential energy curves with those obtained from VASP calculations. Figure 2bc displays the potential energies of cubic (Figure 2b) and orthorhombic (Figure 2c) from VASP (with fitted potential energy curve) and the potential energies from the trained ANN models with the changing hydrostatic strains. We observed that the potential energies from VASP and the trained ANN models almost overlapped, demonstrating that the ANN models provide excellent predictions for potential energies with high fidelity to the corresponding VASP calculations.

Next, we performed separate testing of the AIMD simulations of cubic perovskite MAPbI₃ with a lattice parameter of 12.5 Å and compared the potential energies and atomic forces obtained from the AIMD simulations with those from the trained ANN models. In the testing of the AIMD simulations, the system temperature was set to 300 K, and the step size and the total number of steps of the testing runs were 0.5 fs and 2000 steps (1.0 ps in total), respectively. The stepwise system energy evolution and respective error (between trained ANN models and VASP) are displayed in the left panels of Figure 3a–c. Note that the potential energies and atomic forces from the trained ANN models were obtained by feeding atomic coordinates from respective AIMD trajectory into the trained ANN models. The advantage of this approach is to ensure the input of both VASP and ANN models to be identical, thereby allowing direct comparison between VASP and ANN potential models. The right panels in Figure 3a–c display the corresponding atomic force parities. We note that all three components of each atomic force were included in the atomic force parity plots. It is observed from the left panels of Figure 3a–c that the system energies predicted using the trained ANN models (red lines in the left panels) are in good agreement with those from the AIMD simulations (black lines). The RMSE values for the energy per atom are 0.00123, 0.00174, and 0.00181 eV/atom in Sets 1, 2, and 3, respectively, with corresponding MAEs in the energy per atom of less than 0.005 eV. We note that, in all cases, the RMSE values for the energy per atom were much smaller than the thermal energy at 300 K (∼0.026 eV/atom), demonstrating the transferability of the trained ANN models for predicting system energies. The RMSE for the energy per atom of Set 1 is slightly smaller than those of Set 2 and Set 3, which can be attributed to the more complicated network architecture in Set 1 for the C, N, and H atoms relative to those of Sets 2 and 3. On the other hand, for the prediction of atomic forces, the RMSE value of the atomic force components is much smaller.

Figure 2. (a) Validations of the trained ANN models with different training settings; the left panels display the parity plots and the right panels display the corresponding errors. (b,c) Potential energies of (b) cubic and (c) orthorhombic perovskite MAPbI₃ with respect to the imposed hydrostatic strains obtained using VASP and the trained ANN models, where the lowest potential energy was set to be 0. Note that the dashed line is fitted with the potential energies from VASP calculations.
for Set 3 (0.105 eV/Å) than those for Set 1 (0.75 eV/Å) and Set 2 (0.81 eV/Å), as is also evident from an examination of the parity plots of the atomic force components in the right panels of Figure 3a–c. The better atomic force predictions for Set 3 can be attributed to the inclusion of atomic force training during the training process. The present study hence demonstrated that simple neural network architectures can predict energies/forces of MAPbI₃ with high fidelity to DFT calculations. We do anticipate that more complicated neural network architectures should further promote accuracy; however, we have to balance the computational accuracy and efficiency because energy/force evaluation is the most computationally intensive part in atomistic simulations. It must be noted that despite the ANN potential models presented can successfully predict energies of both tetragonal and cubic MAPbI₃, at this moment these ANN models cannot be extended to mixed-ion perovskite materials such as FAₓMA₁₋ₓPbI₃, because the images of FAₓMA₁₋ₓPbI₃ from AIMD simulations were not included in the training sets in the present study. The present study already demonstrated the capability of the ANN potential model in prediction of energy of organometal halide perovskite with mixed covalent, ionic, 

Figure 3. (a–c) Left panels: potential energy evolution from both the trained ANN model and AIMD trajectory and the respective errors of Sets 1–3; right panels: respective atomic force component parity plots. (d) Parity plot of the atomic force magnitude and distribution of the error of the atomic force direction for Set 3.
and vdW interactions between atoms, and we are actively working on extending the ANN potential model to FA\textsubscript{x}MA\textsubscript{1−x}PbI\textsubscript{y}Br\textsubscript{1−y} mixed-ion perovskite material.

The atomic force parities between the trained ANN and the respective VASP calculations were further examined for Set 3 (energy/force training). Figure 3d displays the parities of the atomic force magnitudes and the error distribution of the atomic force directions of the ANN model trained using training setting 3 (energy/force training). The RMSE of the atomic force magnitudes is 0.0746 eV/Å with an MAE of 0.836 eV/Å. For the force direction parities, 91.5% of the atomic force directions predicted from the ANN model (Set 3) deviate from those obtained by the VASP calculations by less than 20°. Hence, the atomic forces from the ANN model trained using both energy and forces (Set 3) are in very good agreement with those obtained from the VASP calculations.

**PDOS Obtained Using the Trained ANN Model.** We computed the PDOS and extracted the corresponding normal modes of tetragonal MAPbI\textsubscript{3} using the ANN model trained using both energies and atomic forces (Set 3 in Table 1). The finite displacement method\textsuperscript{57} were employed for PDOS calculations, and the supercell size in VASP calculation was 2 × 2 × 2, whereas both 2 × 2 × 2 and 6 × 6 × 6 supercells were used for PDOS calculations using the ANN potential model. Figure 4a shows the PDOS of the tetragonal perovskite for phonon frequencies of up to 20 meV computed from both trained ANN model of different supercell sizes and DFT calculations using VASP. Characteristic peaks at 4, 10, 15, and 20 meV and valleys at 8 and 18 meV can be observed in the PDOS diagram obtained using the ANN models, and these characteristic frequencies were in good agreements with VASP calculations as well as earlier works.\textsuperscript{56,59} Note that the characteristic peak heights between VASP and ANN models are in good agreements when the supercell sizes were identical (2 × 2 × 2) for both VASP and ANN model; hence, the discrepancies between VASP and ANN model (6 × 6 × 6 supercell) can be attributed to the difference in simulation cell sizes. Normal modes with a negative or very low frequency correspond to the translations and rotations of MA cations with Pb and I atoms nearly fixed. Such low frequencies result in low activation energies for the vibrations and rotations of MA cations in octahedral cages of perovskites, creating degenerate states. Normal modes with frequencies of approximately 4 meV (≈30 cm\textsuperscript{-1}) can be attributed to the off-axis displacements of I atoms with modest coupling vibrations of MA cations, as shown in Figure 4b and Movie M1 in the Supporting Information. Normal modes with frequencies of approximately 10 meV (≈80 cm\textsuperscript{-1}) are due to the on-axis displacements of the Pb atoms and I atoms of the backbone, as shown in Figure 4c (and Movie M2 in the Supporting Information). The strong coupling of MA cation libration with the out-of-axis displacements of I atoms (≈80 cm\textsuperscript{-1}) can be attributed to the off-axis displacements of I atoms with modest coupling vibrations of MA cations, as shown in Figure 4d and Movie M3 in Supporting Information. The strong coupling of MA cation libration with the out-of-axis displacements of I atoms (≈80 cm\textsuperscript{-1}) can be attributed to the off-axis displacements of I atoms with modest coupling vibrations of MA cations, as shown in Figure 4e and Movie M4 in the Supporting Information. Hence, the trained ANN model can be utilized to extract the phonon modes that are in good agreement with the phonon modes obtained by the DFT calculations.

We compare the PDOS computed with the trained ANN models obtained using energy training (Set 2 in Table 1) and using energy/force training (Set 3 in Table 1), as shown in Figure 4f.
We note that even though both trained ANN models have identical neural network architectures, the PDOS computed from the two training settings yields distinct results. Relative to the energy/force trained ANN model, the ANN model trained without force training underestimates the frequencies of the low-frequency modes (the backbone modes), while it overestimates the frequencies of the high-frequency modes (MA modes or strongly coupled MA modes). The discrepancies between the PDOSs obtained from the trained ANN models can be attributed to the relatively poor agreements of atomic forces with the DFT calculations in the ANN model trained without forces. Hence, even though the training using system energies leads to an ANN model that yields accurate predictions of energies, the PDOS is in poor agreement with the PDOS obtained from DFT calculations. Therefore, to obtain an accurate evaluation of PDOS and thermal transport properties, the atomic forces must be included into the training sets during the training processes.

Benchmark of Time Consumption in Energy Evaluations with Ab Initio Calculations. We demonstrated that the trained ANN models can predict system energies with high fidelity to the corresponding ab initio calculations. To benchmark the efficiency of the trained ANN models in time consumption per energy evaluation relative to ab initio calculations, we carried out a series of atomistic calculations for tetragonal perovskite crystals with different system sizes using both the trained ANN models (setting 1 and 2) and VASP. Figure 5 displays the comparison of the time consumed per energy evaluation (the vertical axis) using the ANN model (Set 1 in Table 1) and VASP calculations. We note that the ANN model calculations were carried out only using 16 CPU cores (Intel Xeon E5-2630), whereas the VASP benchmark calculations were performed using both CPUs (16 and 64 cores) and GPUs (1 and 4 NVIDIA Tesla P100). Note that for ANN models, neural network Set 2 is about 10 to 20% faster than Set 1, which can clearly be attributed to one less hidden layer for C, N, and H atoms in Set 2 than in Set 1 (see Table 1). For a system of fewer than 10⁴ atoms, the trained ANN model offers a 10⁴ to 10⁵ increase in speed relative to VASP calculations. We note that in the present study, the system size for VASP calculations was limited to 2592 atoms due to computational resource limitations, whereas the largest system size for ANN calculations was near 10⁸ atoms. Hence, the ANN model not only yields high accuracy with ab initio calculations but also provides a significant increase in computation, allowing efficient sampling of the configurational spaces for thermodynamic property calculations and the construction of phase diagrams using atomistic Monte Carlo simulations. We anticipate that further computational boost can be reached by implementing the ANN models to GPUs, which can potentially facilitate simulation of models of hundreds of thousands atoms; however, it is beyond the scope of the present study.

■ CONCLUSIONS

In summary, in the present study, we successfully trained ANN models of MAPbI₃ perovskite materials by feeding a training set composed of thousands of images of tetragonal perovskite crystals obtained from ab initio calculations using VASP into the ANN model. We demonstrated that despite the complex nature of both the chemical composition (Pb, I, C, N, and H atoms) and interatomic interactions (covalent, van der Waals, and ionic interactions), the trained ANN model can successfully predict the energies of cubic perovskite crystals for several different ANN network architectures and training schemes. Furthermore, the phonon modes can be computed using the trained ANN models, and good agreements with PDOS obtained by ab initio calculations can be reached by training the ANN using both system energies and atomic forces. We benchmarked the time consumption for both the trained ANN model and VASP calculations and demonstrated that for systems of given sizes, the ANN model offers an increase in speed of a factor of 10⁴ to 10⁵ in energy evaluations, thereby allowing exhaustive sampling of the configurational spaces of perovskite for thermodynamic property calculations. Hence, the present study demonstrates that the ANN model can be trained to model chemically complex systems with high fidelity with respect to ab initio calculations as well as provide a significant acceleration in computational efficiencies and can be utilized for predictions of the thermodynamic and structural properties of chemically complex materials, such as mixed-ion perovskites or high entropy alloys.

■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.9b00378.

Parameters of Gaussian descriptors and learning curve as the function of training set sizes (PDF)

Normal modes with frequencies of approximately 4 meV (∼30 cm⁻¹) attributed to the off-axis displacements of I atoms with modest coupling vibrations of MA cations (AVI)

Normal modes with frequencies of approximately 10 meV (∼80 cm⁻¹) attributed to the on-axis displacements of the Pb atoms and I atoms of the backbone (AVI)

Normal modes with frequencies of approximately 15 meV (∼120 cm⁻¹) attributed to the strong coupling of MA cation libration with the out-of-axis displacements of I atoms (AVI)
Normal modes with frequencies of approximately 20 meV (~160 cm⁻¹) attributed to the liberation of MA cations without the displacement of the backbone atoms (AVI).

LAMMPS potential file for three ANN potential models (ANN_Potential_File.txt) (TXT)

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: cwpaogate.sinica.edu.tw. Phone: +886-2-27873145.

ORCID

Chun-Wei Pao: 0000-0003-0821-7856

Author Contributions

H.-A.C. contributed to generation of the results and subsequent analysis. C.-W.P. contributed to project oversight. All the authors contributed to preparation of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Academia Sinica Career Development Award, grant no. 2317-1050100, Academia Sinica Sustainability Science Project, grant no. AS-SS-106-02-4, and Ministry of Science and Technology, Taiwan, grant no. MOST 105-2112-M-001-009-MY3 for financial support and the National Center for High-performance Computing for computational support.

■ REFERENCES

(1) De Wolf, S.; Holovskyy, J.; Moon, S.-J.; Löper, P.; Niesen, B.; Ledinsky, M.; Haug, F.-J.; Yum, J.-H.; Ballif, C. Organometallic Halide Perovskites: Sharp Optical Absorption Edge and Its Relation to Photovoltaic Performance. J. Phys. Chem. Lett. 2014, 5, 1035–1039.

(2) Collavini, S.; Völker, S. F.; Delgado, J. L. Understanding the Outstanding Power Conversion Efficiency of Perovskite-Based Solar Cells. Angew. Chem., Int. Ed. Engl. 2015, 54, 9757–9759.

(3) Cui, D.; Yang, Z.; Yang, D.; Ren, X.; Liu, Y.; Wei, Q.; Fan, H.; Zeng, J.; Liu, S. Color-Tuned Perovskite Films Prepared for Efficient Solar Cell Applications. J. Phys. Chem. C 2016, 120, 42–47.

(4) Kulkarni, S. A.; Baikie, T. P.; Boix, P. P.; Yantara, N.; Mathews, N.; Mhaisalkar, S. Band-Gap Tuning of Lead Halide Perovskites Using a Sequential Deposition Process. J. Mater. Chem. A 2014, 2, 9221–9225.

(5) Stranks, S. D.; Eperon, G. E.; Grancini, G.; Menelaou, C.; Alcocer, M. J. P.; Leijtens, T.; Herz, L. M.; Petrozza, A.; Snaith, H. J. Electron-Hole Diffusion Lengths Exceeding 1 Micrometer in an Organometal Trihalide Perovskite Absorber. Science 2013, 342, 341–344.

(6) Wehrenfennig, C.; Eperon, G. E.; Johnston, M. B.; Snaith, H. J.; Herz, L. M. High Charge Carrier Mobilities and Lifetimes in Organolead Trihalide Perovskites. Adv. Mater. 2014, 26, 1584–1589.

(7) Kojima, A.; Teshima, K.; Shirai, Y.; Miyasaka, T. Organometal Halide Perovskites as Visible-Light Sensitizers for Photovoltaic Cells. J. Am. Chem. Soc. 2009, 131, 6050–6051.

(8) Lee, M. M.; Teuscher, J.; Miyasaka, T.; Murakami, T. N.; Snaith, H. J. Efficient Hybrid Solar Cells Based on Meso-Superstructured Organometal Halide Perovskites. Science 2012, 338, 643–647.

(9) Liu, M.; Johnston, M. B.; Snaith, H. J. Efficient Planar Heterojunction Perovskite Solar Cells by Vapour Deposition. Nature 2013, 501, 395–398.

(10) Xiao, Z.; Kerner, R. A.; Zhao, L.; Tran, N. L.; Lee, K. M.; Koh, T.-W.; Scholes, G. D.; Rand, B. P. Efficient Perovskite Light-Emitting Diodes Featuring Nanometre-Sized Crystallites. Nat. Photonics 2017, 11, 108–115.

(11) Deschler, F.; et al. High Photoluminescence Efficiency and Optically Pumped Lasing in Solution-Processed Mixed Halide Perovskite Semiconductors. J. Phys. Chem. Lett. 2014, 5, 1421–1426.

(12) Dong, L. T.; Yang, Y.; You, J. B.; Hong, Z. R.; Chang, W. H.; Li, G.; Yang, Y. Solution-Processed Hybrid Perovskite Photodetectors with High Detectivity. Nat. Commun. 2014, 5, 5404.

(13) Park, S.; Chang, W. J.; Lee, C. W.; Park, S.; Ahn, H. Y.; Nam, K. T. Photocatalytic Hydrogen Generation from Hydriodic Acid Using Methylammonium Lead Iodide in Dynamic Equilibrium with Aqueous Solution. Nat. Energy 2017, 2, 16185.

(14) Yang, W. S.; Noh, J. H.; Jeon, N. J.; Kim, Y. C.; Ryu, S.; Seo, J.; Seok, S. I. High-Performance Photovoltaic Perovskite Layers Fabricated through Intramolecular Exchange. Science 2015, 348, 1234–1237.

(15) Yang, W. S.; et al. Iodide Management in Formamidinium-Lead Halide-Based Perovskite Layers for Efficient Solar Cells. Science 2017, 356, 1376–1379.

(16) Noh, J. H.; Im, S. H.; Heo, J. H.; Mandal, T. N.; Seok, S. I. Chemical Management for Colorful, Efficient, and Stable Inorganic–Organic Hybrid Nanostructured Solar Cells. Nano Lett. 2013, 13, 1764–1769.

(17) Umari, P.; Mosconi, E.; De Angelis, F. Relativistic Gw Calculations on Ch3nh3pbi3 and Ch3nh3sni3 Perovskites for Solar Cell Applications. Sci. Rep. 2015, 4, 4467.

(18) Gao, P.; Grätzel, M.; Nazeeruddin, M. K. Organohalide Lead Perovskites for Photovoltaic Applications. Energy Environ. Sci. 2014, 7, 2448–2463.

(19) Jeon, N. J.; Noh, J. H.; Yang, W. S.; Kim, Y. C.; Ryu, S.; Seo, J.; Seok, S. I. Compositional Engineering of Perovskite Materials for High-Performance Solar Cells. Nature 2015, 517, 476–480.

(20) Ahn, N.; Son, D. Y.; Jang, I. H.; Kang, S. M.; Choi, M.; Park, N. G. Highly Reproducible Perovskite Solar Cells with Average Efficiency of 18.3% and Best Efficiency of 19.7% Fabricated Via Lewis Base Adduct of Lead(ii) Iodide. J. Am. Chem. Soc. 2015, 137, 8696–8699.

(21) Im, J.-H.; Jang, I.-H.; Pellet, N.; Grätzel, M.; Park, N.-G. Growth of Ch3nh3pbi3 Cuboids with Controlled Size for High-Efficiency Perovskite Solar Cells. Nat. Nanotech. 2014, 9, 927–932.

(22) Burschka, J.; Pellet, N.; Moon, S.-J.; Humphrey-Baker, R.; Gao, P.; Nazeeruddin, M. K.; Grätzel, M. Sequential Deposition as a Route to High-Performance Perovskite-Sensitized Solar Cells. Nature 2013, 499, 316–319.

(23) Mattoni, A.; Filippetti, A.; Saba, M. I.; Delugas, P. Methylammonium Rotational Dynamics in Lead Halide Perovskite by Classical Molecular Dynamics: The Role of Temperature. J. Phys. Chem. C 2015, 119, 17421–17428.

(24) Gutierrez-Sevillano, J. J.; Ahmad, S.; Calero, S.; Anta, J. A. Molecular Dynamics Simulations of Organohalide Perovskite Precursors: Solvent Effects in the Formation of Perovskite Solar Cells. Phys. Chem. Chem. Phys. 2015, 17, 22770–22777.

(25) Lahmster, J.; Kresse, G.; Heinen, J.; Bokdam, M. Finite-Temperature Structure of the Mph3i Perovskite: Comparing Density Functional Approximations and Force Fields to Experiment. Phys. Rev. Mater. 2018, 2, 073604.

(26) Takaba, H.; Kimura, S.; Alam, M. K. Crystal and Electronic Structures of Substituted Halide Perovskites Based on Density Functional Calculation and Molecular Dynamics. Chem. Phys. 2017, 485–486, 22–28.

(27) Zhou, N.; Shen, Y.; Zhang, Y.; Xu, Z.; Zheng, G.; Li, L.; Chen, Q.; Zhou, H. C. Pre-Intercalation in the Inorganic Framework for Efficient and Stable Fa1–xCxapid(3i)(3) Perovskite Solar Cells. Small 2017, 13, 1700484.

(28) Kubicki, D. J.; Prochowicz, D.; Hofstetter, A.; Zakeeruddin, S. M.; Grätzel, M.; Emley, L. Phase Segregation in Cs–Rb- and K-Doped Mixed-Cation (Ma)X(Fa)1–Xpbi3 Hybrid Perovskites from Solid-State Nmr. J. Am. Chem. Soc. 2017, 139, 14173–14180.

(29) Grätzel, M. The Rise of Highly Efficient and Stable Perovskite Solar Cells. Accounts Chem. Res. 2017, 50, 487–491.
(30) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. Phys. Rev. Lett. 2007, 98, 146401.

(31) Chmiela, S.; Tkatchenko, A.; Saucesa, H. E.; Poltavsky, I.; Schütt, K. T.; Müller, K.-R. Machine Learning of Accurate Energy-Conserving Molecular Force Fields. Sci. Adv. 2017, 3, No. e1603015.

(32) Behler, J. Perspective: Machine Learning Potentials for Atomistic Simulations. J. Chem. Phys. 2016, 145, 170901.

(33) Jain, A.; et al. Commentary: A Materials Genome Approach to Accelerating Materials Innovation. APL Mater. 2013, 1, 011002.

(34) Mansouri Tehrani, A.; Olynyk, A. O.; Parry, M.; Rizvi, Z.; Couper, S.; Lin, F.; Miyagi, L.; Sparks, T. D.; Brogch, J. Machine Learning Directed Search for Ultrafastcompressible, Superhard Materials. J. Am. Chem. Soc. 2018, 140, 9844–9853.

(35) Zimmermann, N. E. R.; Hannah, D. C.; Rong, Z.; Liu, M.; Ceder, G.; Haranczyk, M.; Persson, K. A. Electrostatic Estimation of Intercalant Jump-Diffusion Barriers Using Finite-Size Ion Models. J. Phys. Chem. Lett. 2018, 9, 628–634.

(36) Artrith, N.; Kolpak, A. M. Grand Canonical Molecular Dynamics Simulations of Cu–Au Nanoalloys in Thermal Equilibrium Using Reactive Ann Potentials. Comput. Mater. Sci. 2015, 110, 20–28.

(37) Artrith, N.; Urban, A. An Implementation of Artificial Neural-Network Potentials for Atomistic Simulations: Performance for TiO2. Comput. Mater. Sci. 2016, 114, 135–150.

(38) Kobayashi, R.; Girofè, D.; Junge, T.; Ceriotti, M.; Curtin, W. A. Neural Network Potential for Al-Mg-Si Alloys. Phys. Rev. Mater. 2017, 1, 053604.

(39) McCulloch, W. S.; Pitts, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. Bull. Math. Biol. 1943, 52, 99–115.

(40) Schütt, K. T.; Sauceda, H. E.; Kindermans, P. J.; Tkatchenko, A.; Müller, K.-R. Schnet – a Deep Learning Architecture for Molecules and Materials. J. Chem. Phys. 2018, 148, 241722.

(41) Lubbers, N.; Smith, J. S.; Barros, K. Hierarchical Modeling of Molecular Energies Using a Deep Neural Network. J. Chem. Phys. 2018, 148, 241715.

(42) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. J. Chem. Phys. 2011, 134, 074106.

(43) Khoshshir, A.; Peterson, A. A. Amp: A Modular Approach to Machine Learning in Atomistic Simulations. Comput. Phys. Commun. 2016, 207, 310–324.

(44) Mosconi, E.; Quart, C.; Ivanovska, T.; Ruani, G.; De Angelis, F. Structural and Electronic Properties of Organo-Halide Lead Perovskites: A Combined Ir-Spectroscopy and Ab Initio Molecular Dynamics Investigation. Phys. Chem. Chem. Phys. 2014, 16, 16137–16144.

(45) Eperon, G. E.; Stranks, S. D.; Menelaou, C.; Johnston, M. B.; Herz, L. M.; Snaith, H. J. Formamidinium Lead Trihalide: A Broadly Tunable Perovskite for Efficient Planar Heterojunction Solar Cells. Energy Environ. Sci. 2014, 7, 982–988.

(46) Baike, T.; Fang, Y.; Kadro, J. M.; Schreyer, M.; Wei, F.; Mhysairak, S. G.; Graetzel, M.; White, T. J. Synthesis and Crystal Chemistry of the Hybrid Perovskite (CH3NH3)PbI3 for Solid-State Sensitised Solar Cell Applications. J. Mater. Chem. A 2013, 1, 5628–5641.

(47) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. J. Comput. Phys. 1995, 117, 1–19.

(48) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. Phys. Rev. B: Condens. Matter Mater. Phys. 1993, 47, 558–561.

(49) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Open-Shell Transition Metals. Phys. Rev. B: Condens. Matter Mater. Phys. 1993, 48, 13115–13118.

(50) Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metal-Amorphous-Semiconductor Transition in Germanium. Phys. Rev. B: Condens. Matter Mater. Phys. 1994, 49, 14251–14269.

(51) Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Ziolhais, C. Atoms, Molecules, Solids, and Surfaces: Applications of the Generalized Gradient Approximation for Exchange and Correlation. Phys. Rev. B: Condens. Matter Mater. Phys. 1992, 46, 6671–6687.

(52) Perdew, J. P.; Wang, Y. Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. Phys. Rev. B: Condens. Matter Mater. Phys. 1992, 45, 13244–13249.

(53) Blöchl, P. E. Projector Augmented-Wave Method. Phys. Rev. B: Condens. Matter Mater. Phys. 1994, 50, 17953–17979.

(54) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. Phys. Rev. B 1999, 59, 1758–1775.

(55) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (Dft-D) for the 94 Elements H-Pu. J. Chem. Phys. 2010, 132, 154104.

(56) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. J. Comput. Chem. 2011, 32, 1456–1465.

(57) Kresse, G.; Furthmüller, J.; Hafner, J. Ab Initio Force Constant Approach to Phonon Dispersion Relations of Diamond and Graphite. Europhys. Lett. 1995, 32, 729–734.

(58) Xu, J.; Xu, Q.; Han, G.; Wu, B.; Huan, C. H. A.; Leek, M. L.; Sum, T. C. Hot Carrier Cooling Mechanisms in Halide Perovskites. Nat. Commun. 2017, 8, 1300.

(59) Quarti, C.; Grancini, G.; Mosconi, E.; Bruno, P.; Ball, J. M.; Lee, M. M.; Snaith, H. J.; Petrozza, A.; De Angelis, F. The Raman Spectrum of the CH3NH3PbI3 Hybrid Perovskite: Interplay of Theory and Experiment. J. Phys. Chem. Lett. 2014, 5, 279–284.