



PCCP

**Accurate prediction of grain boundary structures and energetics in CdTe: A machine-learning potential approach**

Journal:	<i>Physical Chemistry Chemical Physics</i>
Manuscript ID	CP-ART-09-2021-004329.R1
Article Type:	Paper
Date Submitted by the Author:	14-Dec-2021
Complete List of Authors:	Yokoi, Tatsuya; Nagoya University, Department of Materials Physics Adachi, Kosuke; Nagoya University, Department of Materials Physics Iwase, Sayuri; Nagoya University, Department of Materials Physics Matsunaga, Katsuyuki; Nagoya University, Materials Physics

SCHOLARONE™  
Manuscripts

## Accurate prediction of grain boundary structures and energetics in CdTe: A machine-learning potential approach

Tatsuya Yokoi,<sup>a\*</sup> Kosuke Adachi<sup>a</sup>, Sayuri Iwase<sup>a</sup>, Katsuyuki Matsunaga<sup>a,b</sup>

<sup>a</sup>*Department of Materials Physics, Nagoya University, Nagoya 464-8603, Japan*

<sup>b</sup>*Nanostructures Research Laboratory, Japan Fine Ceramics Center, Nagoya, 456-8587, Japan*

\* Corresponding author.

E-mail address: [yokoi@mp.pse.nagoya-u.ac.jp](mailto:yokoi@mp.pse.nagoya-u.ac.jp)

Phone number: +81-52-789-3365

### Abstract

To accurately predict grain boundary (GB) atomic structures and their energetics in CdTe, the present study constructs an artificial-neural-network (ANN) interatomic potential. To cover a wide range of atomic environments, large amounts of density-functional-theory (DFT) data are used as a training dataset including point defects, surfaces and GBs. Structural relaxation combined with the trained ANN potential is applied to symmetric tilt and twist GBs, many of which are not included in the training dataset. The relative stability of the relaxed structures and their GB energies are then evaluated with the DFT level. The ANN potential is found to accurately predict low-energy structures and their energetics with reasonable accuracy with respect to DFT results, while conventional empirical potentials critically fail to find low-energy structures. The present study also provides a way to further improve the transferability of the ANN potential to more complicated GBs, using only low- $\Sigma$  GBs as training datasets. Such improvement will offer a way to accurately predict atomic structures of general GBs within practical computational cost.

### 1. Introduction

Polycrystalline semiconductors are actively investigated as clean-energy and low-cost materials in the photovoltaic and thermoelectric research areas. For CdTe, with the zincblende structure, its polycrystal has already drawn much attention as a p-type absorber material in thin-film solar cells, due to its high absorption coefficient and low processing cost. However, it exhibits lower efficiency than the single-crystal counterpart, mainly because of the presence of grain boundaries (GBs). They are considered to act as nonradiative recombination centers by creating deep gap states inside the band gap, as indicated by electron beam induced current (EBIC) mapping and cathodoluminescence (CL) measurements.<sup>1,2</sup> Recent CL measurements combined with electron backscattered diffraction (EBSD) mapping indicated that carrier recombination velocity does not simply depend on grain misorientations but varies with GB planes.<sup>3</sup> With the goal of uncovering

the exact role of GBs in carrier recombination, a better understanding of what type of GB creates detrimental gap states is a critical step in the research field of polycrystalline CdTe.

To reveal the relationship between GB structures and electronic properties, theoretical calculations have been widely used,<sup>4-8</sup> sometimes combined with electron microscopy observations.<sup>4,5</sup> Previous density-functional-theory (DFT) calculations indicated that specific symmetric tilt GBs (STGBs) have close Te-Te and Cd-Cd pairs whose antibonding and bonding states create deep gap states, respectively.<sup>4,5,7</sup> On other hand, DFT calculations on other STGBs indicated that those GBs do not create deep gap states, although there are shallow gap states arising from GB atoms with dangling bonds.<sup>5,7</sup> Park recently indicated that even for crystallographically the same GB, different atomic structures behave differently with respect to electronic structure: the lowest-energy structure with non-mirror symmetry does not have deep gap states, while a commonly adopted structure with mirror symmetry exhibits a significant gap state in the band gap.<sup>8</sup> These studies indicate that whether a GB acts as a recombination center strongly depends on its atomic structure. The accurate determination of stable atomic structures for GBs is thus essential, in order to reveal their exact electronic structures as well as other physical properties.

In most cases, however, stable atomic structures of GBs are difficult to determine using DFT calculations alone due to their huge computational cost. To explore low-energy GB structures,  $\gamma$ -surface methods are most commonly used.<sup>9-11</sup> In those approaches, many structures are initially generated by incrementally shifting one grain relative to the other, in order that they have different relative positions of two grains. Even for low- $\Sigma$  STGBs, dozens of structures are often needed to sufficiently sample the  $\gamma$  surface at GBs. All initial structures must then be fully relaxed. More efficient methods have been also used to search for a wider range of candidate structures, including simulated annealing<sup>12,13</sup> and evolutionary algorithm.<sup>14,15</sup> Those methods need long-time molecular dynamics (MD) simulations and larger numbers of total-energy calculations. Thus those search methods are typically combined with empirical interatomic potentials rather than DFT calculations. However, the accuracy of empirical potentials significantly depends on substances and also GB structures.<sup>16,17</sup> For CdTe GBs, the predictive ability of commonly used empirical potentials is severely impaired in prediction of relative GB energies between different atomic structures, as is discussed later.

To construct interatomic potentials with DFT accuracy, machine-learning potentials have become a mainstream.<sup>18,19</sup> In this approach, machine-learning models are fitted to large amounts of DFT data (e.g., total energy, atomic force and stress tensor), so that they can approximate the DFT potential energy surface without numerically solving the Kahn-Sham equation. They have been focused on not only perfect crystal but also lattice defects, including point defects,<sup>20,22</sup> surfaces,<sup>23,24</sup> dislocations,<sup>25,26</sup> and have demonstrated excellent transferability to

such complicated systems. Among such studies, their application to GBs is still very limited to simple systems such as Si<sup>27,28</sup> and fcc metals.<sup>29</sup> Although Mińkowski et al. recently reported artificial-neural-network (ANN) potentials for CdTe, they were aimed at predicting diffusion properties in the bulk.<sup>30</sup> So far it remains unclear whether machine-learning potentials can be used for accurately predicting GB structures in CdTe. In addition, there is still no report on an efficient way of improving their transferability to general GBs, which obviously cannot be contained in the training dataset due to the need for very large computational cells.

In this work, an ANN potential is constructed with the goal of accurately predicting stable atomic structures of GBs and their physical properties in CdTe. The ANN potential is trained with large numbers of total energies and atomic forces obtained from DFT calculations. The training dataset contains point defects, surfaces and tilt and twist GBs, in order to sample a wide range of atomic environments of lattice defects including GBs. The trained ANN potential is combined with a  $\gamma$ -surface method with structural relaxation and then applied to ten different GBs. Their GB energies are again calculated with the DFT level, in order to evaluate whether the ANN potential can accurately distinguish between low- and high-energy structures and also to predict their GB energies. Furthermore, the present study provides an efficient way to generate training datasets, in order to improve the transferability of the ANN potential to crystallographically complex GBs that are not included in the training dataset.

## 2. Computational method

### 2.1 DFT calculation

To generate training and test datasets, DFT calculations based on the projector augmented wave (PAW) method<sup>31,32</sup> were performed using the Vienna *Ab initio* Simulation Package (VASP).<sup>33,34</sup> A plane-wave basis set was used to expand wave functions, and its energy cutoff was chosen to be 500 eV. The electrons in the 4*d* and 5*s* orbitals for Cd and in the 5*s* and 5*p* orbitals for Te were explicitly treated as valence electrons. The revised generalized gradient approximation parameterized by Perdew, Burke and Ernzerhof (GGA-PBESol)<sup>35</sup> was employed to calculate the exchange-correlation energy. The convergence criterion of total energy was set to 10<sup>-6</sup> eV. Training datasets based on DFT-MD simulations were also obtained using the Parrinello-Rahman dynamics with the Langevin thermostat.<sup>36,37</sup> Both the NVT and NPT conditions were used with a time step of 2 fs.

### 2.2 Construction of ANN potential

To construct an ANN potential, a feed-forward network with two hidden layers was implemented in our in-house code, which is based on that reported by Behler.<sup>38</sup> Two- and three-body symmetry

functions,<sup>38</sup> denoted by  $G_i^2$  and  $G_i^3$  for atom  $i$ , were employed. Their functional forms are defined as

$$G_i^2 = \sum_j^{N_i^c} e^{-\eta_2(R_{ij}-R_s)^2} \cdot f_c(R_{ij}) \text{ and (1)}$$

$$G_i^3 = 2^{1-\zeta} \sum_{j \neq ik \neq ij}^{N_i^c} \sum_{k \neq ij}^{N_i^c} \left[ \{1 + \cos(\theta_{ijk} - \theta)\}^\zeta \cdot e^{-\eta_3(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}) \cdot f_c(R_{jk}) \right] e^{-\eta_2(R_{ij}-R_s)^2} \text{ , (2)}$$

where  $\eta_2$ ,  $R_s$ ,  $\zeta$ ,  $\theta$ , and  $\eta_3$  are the hyperparameters, and  $N_i^c$  is the number of atoms in the cutoff radius for atom  $i$ .  $R_{ij}$ ,  $R_{ik}$  and  $R_{jk}$  represent the distance between atoms  $ij$ ,  $ik$ , and  $jk$ , respectively.  $\theta_{ijk}$  represents the angle between bonds  $ij$  and  $ik$ . Their hyperparameters are listed in Tables S1 and S2 in the Supporting Information. A cutoff function  $f_c(R_{ij})$  is defined by

$$f_c(R_{ij}) = \begin{cases} \frac{1}{2} \left[ \cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1 \right] & (R_{ij} \leq R_c) \\ 0 & (R_{ij} > R_c) \end{cases} \text{ . (3)}$$

The cutoff distance  $R_c$  was chosen to be 7 Å. This value was found to give the minimum residual among 6, 7, 8, 9 and 10 Å. The number of symmetry functions was chosen to be 66 (18 for  $G_i^2$  and 48 for  $G_i^3$ ), corresponding to the number of nodes in the input layer. The test calculations for determining the ANN hyperparameters are described Tables S3 and S4 in the Supporting Information.

Table 1 summarizes the training dataset generated. To account for various atomic environments, not only perfect-crystal cells but also simulation cells containing vacancies, anti-site defects and surfaces were also contained. Additionally, symmetric tilt, twist and mixed GBs were considered. For STGBs, the  $\Sigma 5(210)/[001]$ ,  $\Sigma 5(310)/[001]$  and  $\Sigma 3(112)/[\bar{1}10]$  GBs were constructed. For the  $\Sigma 3(112)/[\bar{1}10]$  STGB, crystallographically two different structures can exist, so-called the Cd- and Te-core, depending on the Cd-Te bond direction with respect to the normal to the GB plane. Thus both structures were constructed. For a twist GB, the  $\Sigma 3(112)/[112]$  GB with a twist angle of 180° was constructed. To further consider various GB atomic environments for the training dataset, mixed GBs were also constructed by rotating one grain of the  $\Sigma 3(112)/[\bar{1}10]$  STGB by 180° about the  $[112]$  axis. As a result, this mixed GB has the (112) GB plane and the misorientation angles of 109.47° and 180° for tilt and twist components, respectively. The mixed GB also has the Cd- and Te-core, and both structures were thus included in the training dataset. To obtain reference structures of each GB, initial structures

were created by shifting one grain relative to the other along the GB plane and then fully relaxed. The relaxed structures were then used to generate the training dataset as described below.

To obtain various atomic environments from one reference structure, two approaches were used. One is to randomly displace atoms and to randomly change the cell shape. The maximum atomic displacement was set to 0.4 Å along each direction in the Cartesian coordinate. The computational cell was also subjected to contraction, expansion and shear deformation, by varying axis lengths and angles between axes with the maximum of  $\pm 20\%$ , respectively. Total energies and atomic forces were then calculated by performing DFT single-point calculation and structural relaxation for a few iterations, and they were added to the training dataset. The other is to perform DFT-MD simulations for several hundreds of steps in the temperature range between 200 and 1600 K. For single-crystal and GB cells, the NVT and NPT conditions were used, whereas only the NVT condition was used for surface cells. Snapshots extracted from those DFT-MD simulations were then included in the training dataset. Totally 118050 total energies and 6214545 atomic forces were used to train the ANN potential. The ANN potential was trained using the extended Kalman filter with weighted least squares.<sup>39</sup>

### 2.3 Structural optimization with ANN potential

The ANN potential trained was used to predict the GB-structure-energy relationship. For this purpose, it was combined with the quasi-Newton method with the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm to perform structural relaxation, for which the atomic positions are updated using the atomic forces based on the ANN potential. In this calculation, a simulation cell with one GB plane and two surfaces was initially constructed as illustrated in Fig. 1(a). Here, the  $x$ -axis is aligned perpendicular to the GB plane, and the  $y$ - and  $z$ -axes are parallel to the GB plane. For all GBs examined, the  $x$ -axis length of a grain was chosen to be more than 40 Å to ensure that there is no interaction between the surfaces and GB. The width of a vacuum layer along the  $x$ -axis was set to 10 Å, so that two surface do not interact. To obtain various pairs of GB structures and energies, one grain was stepwise shifted along the  $y$ - and  $z$ -axes with an increment of 0.5 Å, so that initial structures with different atomic positions at the GB were generated. All structures were then fully relaxed using the ANN potential until the atomic forces were less than  $10^{-3}$  eV/Å. In structural relaxation, all atoms were allowed to move whereas the cell dimensions were fixed. From a relaxed structure (Fig. 1(b)), a simulation cell with two identical GB structures was reconstructed with three-dimensional (3D) periodic boundary conditions so as to eliminate the surface region, as illustrated in Fig. 1(c). The reconstructed cell was finally used to calculate the GB energies for the ANN and DFT calculation. Similar calculations were also performed using conventional empirical potentials implemented in the LAMMPS code,<sup>40</sup> in order to examine their

transferability to GBs. For this purpose, Stillinger-Weber (SW)<sup>41</sup> and bond-order (BO) potentials<sup>42</sup> were selected.

The abovementioned procedure was applied to two classes of GB structures. One is GBs contained in the training dataset, the  $\Sigma 5(210)/[001]$ ,  $\Sigma 5(310)/[001]$  STGBs and  $\Sigma 3(112)/[\bar{1}10]$  twist GB. The other is GBs not included in the training dataset. The latter is more important in the sense that ultimately the ANN potential must sustain its transferability even to general GBs, which are naturally difficult to contain in the training dataset due to the need for very large computational cells. Here, the latter class contained STGBs ( $\Sigma 13(320)/[001]$  and  $\Sigma 13(510)/[001]$ ) and five twist GBs with the twist angle of  $180^\circ$  ( $\Sigma 3(111)/[111]$ ,  $\Sigma 9(114)/[114]$ ,  $\Sigma 9(221)/[221]$ ,  $\Sigma 11(113)/[113]$  and  $\Sigma 11(332)/[332]$ ).

### 3. Results and discussion

#### 3.1 Errors of ANN potential for training and test datasets

Figure 2 shows the correlation between the ANN and DFT values. For the training dataset (red point), including perfect-crystal structures, point defects, surfaces and GBs, all data points are distributed near the diagonal line, without large deviation. This indicates that the present ANN potential have sufficiently learned all types of structures without significant errors. The mean absolute error (MAE) values over all data points are calculated to be 3.21 meV/atom and 40.1 meV/Å for the total energies and atomic forces, respectively. These values are the same level as those of previous studies on machine-learning potentials,<sup>28,29,43</sup> although a quantitative comparison is often difficult since MAE values strongly depend on training datasets. Table 1 summarizes the MAE value for each training dataset. Although the MAE values vary depending on the individual structures, they are comparable to each other. It is thus expected that even for GBs, standard molecular simulations with the ANN potential are possible with acceptable accuracy.

It is noted that the energy MAE of the single-crystal-based structures is evaluated to be 5.90 meV/atom, slightly larger than the other datasets. This is because this dataset includes simulation cells that are substantially contracted or compressed by the maximum of  $\pm 20\%$  from the original cell dimension. We empirically find that it is difficult for ANN potentials to learn such structures with small errors when other various structures are also included in the training dataset at the same time. Nevertheless, the ANN potential well reproduces the volume dependence of the potential energy obtained from DFT calculations, as shown in Fig. 3(a). For the zincblende structure of CdTe, the ANN potential predicts the lattice constant to be 6.51 Å, which is in quantitative agreement with the DFT value (6.50 Å). Figure 3(b) also shows the phonon dispersion curve and density of states (DOS) predicted by the ANN potential. They reasonably agree with the DFT results, although there is a small difference in the low-frequency mode. These

results indicate that the present ANN potential reproduces basic properties of bulk CdTe with sufficient accuracy.

For the test dataset constructed from the  $\Sigma 3(112)$ ,  $\Sigma 5(210)$ ,  $\Sigma 5(310)$  GBs, the blue points are also distributed along the diagonal line. The MAE value of each GB is comparable to those for the training dataset, as listed in Table 2. This indicates that the trained ANN potential can sustain the reasonable accuracy for atomic environments of GBs included in the training dataset. It is thus expected that if a training dataset captures a wide range of atomic environments at GBs, ANN potentials trained with the dataset can be applied to GBs with reasonable accuracy.

### 3.2 GB structures and energetics predicted by ANN potential

This section shows predictive accuracy of the trained ANN potential to GBs, comparing the SW and BO potentials. For this purpose, structural relaxations with the ANN and empirical potentials are performed to obtain the GB-structure-energy relationship of each potential. Figure 4 shows GB energies of the  $\Sigma 9(114)/[114]$ ,  $\Sigma 13(320)/[001]$  and  $\Sigma 13(510)/[001]$  GBs, all of which are not included in the training dataset. On the left side of each panel, the individual data points correspond to the GB energies of different GB structures obtained from structural relaxation. These structures are then used to calculate GB energies by performing DFT single-point calculations. The DFT values correspond to the data points on the right side of each panel. Two data points connected with the line are obtained from the same GB structure. The dashed lines indicate the lowest DFT values obtained from the ANN GB structures.

It is found that the lowest-energy structures predicted by the ANN potential are also the energetically lowest ones for DFT calculations. For these structures, the energy differences between the ANN and DFT values are less than  $0.09 \text{ J/m}^2$ , indicating that their energetics is also accurately predicted by the ANN potential. Furthermore, the ANN potential correctly distinguishes between low- and high-energy states as their order in GB energy is maintained for DFT calculations in most cases. Similar trends are also observed for the other GBs, as shown in Fig. S1 in the Supporting Information. It is noted that for the  $\Sigma 3(112)/[112]$ ,  $\Sigma 5(310)/[001]$  and  $\Sigma 11(332)/[332]$  GBs, the lowest energy structures are very close in GB energy to their metastable structures. In these cases, the ANN potential fails to distinguish the lowest-energy and metastable structures. Nevertheless, it successfully distinguishes between low- and high-energy states for all GBs examined (see Fig. S1). These results demonstrate that even for GBs absent in the training dataset, the present ANN potential enables screening candidate structures without DFT calculations, which will greatly reduce computational cost to explore low-energy GB structures.

Comparing the DFT GB energies of the lowest-energy structures obtained from the ANN, SW and BO potentials, the ANN values are found to be the lowest for the three GBs, as

indicated by the dashed lines in Fig. 4. The lowest-energy structures predicted by the ANN potential have lower GB energies than those within the empirical potentials. Importantly, the lowest-energy structures predicted by the empirical potentials are not the lowest-energy ones for the DFT level. Rather, they have clearly high GB energies for the  $\Sigma 13(320)/[001]$  GB. Table 3 summarizes the DFT values for all GB examined. These values are obtained from the lowest-energy structures that are predicted by each potential. In all cases except for the  $\Sigma 3(111)/[111]$  GB, the GB energies calculated from the ANN GB structures are lower than those for the SW potential. This clearly demonstrates that at atomic environments at GBs, the PES of the SW potential greatly deviates from that of DFT and cannot be used to explore low-energy GB structures with reliable accuracy.

It should be noted that structural relaxation algorithms also potentially affect the results in Fig. 4. To examine this possibility and to confirm that the ANN potential is exactly superior to the SW and BO potentials, GB energies based on these empirical potentials were calculated for the lowest-energy structures obtained from the ANN potential. The calculated values for the three GBs in Fig. 4 are listed in Table. 4. The values of single-point calculations with the empirical potentials were entirely higher than those of the energy-minimum structures obtained from the empirical potentials. In addition, structural relaxation with the empirical potentials was performed by using the ANN structures, and GB energies of the relaxed structures were calculated. The obtained GB energies were found to be lower than those of single-point calculations, because the ANN structures were not local minima for the empirical potentials, and consequently, they changed to other local minimum structures during the structural relaxation. As a result, the ANN structures changed to the same structures (having the same GB energies) with the SW and BO lowest-energy ones for the  $\Sigma 9(114)$  GB. Nevertheless, the GB energies were still higher than those of the energy-minimum structures obtained from the empirical potentials for the  $\Sigma 13(320)$  and  $\Sigma 13(510)$  GBs. These results demonstrate that the SW and BO potentials incorrectly predict the low-energy ANN structures.

Figure 5 compares the lowest-energy structures obtained from the ANN and empirical potentials. There are some critical differences between them: for the  $\Sigma 9(114)/[114]$  GBs (Fig. 5(a)), the ANN potential exhibits asymmetric atomic arrangements in terms of the GB planes, whereas the SW and BO potentials predict the symmetric ones. For  $\Sigma 13(320)/[001]$  (Fig. 5(b)), the ANN potential predicts an asymmetric structure while the empirical potentials predict different symmetric structures. For  $\Sigma 13(510)/[001]$  (Fig. 5(c)), the three potentials predict asymmetric structures, but the structural units of the empirical potentials are less distorted. These trends of the empirical potentials probably reflect the fact that they are typically fitted to reproduce reference bulk structures, which usually have higher symmetry, simpler atomic arrangements than

GBs. Such a fitting procedure is most likely to result in their limited transferability to GBs in CdTe.

Figures 6(a)-6(c) show the correlation in GB energy of DFT single-point calculations to the ANN, SW and BO potentials, respectively, for all GBs examined. For low GB energies, the ANN and DFT values are well correlated with small deviations from the diagonal line. Although the errors of ANN predictions become large for high GB energies, most of the data points are still distributed along the diagonal line. This clearly demonstrates that the ANN potential can entirely distinguish between low- and high-energy structures, even for GBs that are not included in the training dataset. The average MAE value over all GBs is calculated to be 0.07 J/m<sup>2</sup>, suggesting that reasonable estimation of GB energies is possible by means of the ANN potential. It is thus expected that further improvement in predictive ability of the ANN potential enables us to predict more complicated GB structures and their GB energies, which typically need large-scale simulation cells, with practical computational cost. An improvement strategy is discussed in the next section.

Unlike the ANN potential, Fig. 6(b) shows that most of the GB energies predicted by the SW potential greatly deviate from the DFT values, and their order is often inconsistent with the DFT result. This trend is observed for all GBs examined except for the  $\Sigma 3(111)/[111]$  GB, which has a low GB energy of 0.01 J/m<sup>2</sup>. Although the BO potential shows a smaller error than the SW potential, the data points still significantly deviate from the diagonal line, with a MAE value of 0.19 J/m<sup>2</sup>. Thus the SW and BO potentials are difficult to use for determination of low-energy GB structures with sufficient accuracy.

### 3.3 Improvement in Predictive Ability of ANN Potential

The previous section demonstrates that the ANN potential outperforms the SW and BO potentials in the prediction of GB energies and low-energy structures. However, it is found that there exist two problems of applying the ANN potential to general GBs, which are most likely to have a wider range of GB-energy values than low- $\Sigma$  GBs. One is that the predictive power of the ANN potential tends to become low for high-energy structures, as indicated by Fig. 6(a). The other is that if a GB structure predicted by the ANN potential has a large error in GB energy from the corresponding DFT single-point calculation, its relaxed structure with the DFT level is often significantly different from that predicted by the ANN potential. For example, Fig. 7(a) shows that the GBs pointed by the black arrows differ in GB energy by 0.15 J/m<sup>2</sup>. As displayed in Fig. 8, their GB structures also have different atomic arrangements, with different relative positions of two grains along the GB plane. In this case, the GB structure relaxed by DFT calculation is slightly lower in GB energy than the lowest-energy one predicted by the ANN potential, as

indicated by the black dashed line. This problem would be critical when one uses the ANN potential prior to DFT calculations, in order to distinguish between low- and high-energy GBs.

Whether the above problems occur depends on individual GBs and it is difficult to know beforehand what types of GB structures involve such large errors. For example, the  $\Sigma 9(221)/[221]$  twist GB are not included in the training dataset, but the error in GB energy is found to be relatively small as shown in Fig. 7(b). Only with the information of GB type, it is thus difficult to find GB atomic environments that are not covered by the current training dataset.

In order to efficiently explore GB atomic environments that are not sufficiently covered by the training dataset, the ANN potential trained once was utilized in the follow manner. From now this ANN potential is referred to as ANN-1. If a GB structure obtained from ANN-1 has an error in GB energy larger than  $0.05 \text{ J/m}^2$  from the DFT single-point calculation, it was used as a reference structure to generate an additional training dataset. For the  $\Sigma 5(310)/[001]$  and  $\Sigma 9(221)/[221]$  GBs (Fig. 7), the corresponding structures are indicated by the red arrows. The GBs other than the  $\Sigma 3(111)/[111]$  GB were used to obtain their reference structures in the same way. The  $\Sigma 5(310)/[001]$ ,  $\Sigma 5(210)/[001]$  and  $\Sigma 3(112)/[112]$  GBs were used for additional training datasets, whereas the remaining six GBs were used for test datasets.

The training and test datasets were generated by randomly displacing atoms and randomly changing the cell shape of the reference structures, with the same condition described in the section of Computational method. Single-point calculation and structural relaxation for a few iterations were then performed using the generated structures. Finally, 3290 energies and 177105 atomic forces were added to the first training dataset, and ANN-1 was trained again. Here this second ANN potential is referred to as ANN-2.

Figure 9 shows the MAE values of ANN-1 and ANN-2 for the additional training and test datasets. For the training dataset, the ANN-2 values are lower than the ANN-1 values for the three GBs, particularly for  $\Sigma 5(210)/[001]$  GB. This trend is reasonable since ANN-2 has learned these structures. More importantly, ANN-2 also exhibits lower MAEs than ANN-1 for the GBs in the test dataset, although these GBs were not added to the additional training dataset.

ANN-2 was also used to perform the  $\gamma$ -surface method with structural relaxation to predict various GB-structure-energy relationships and to compare them with the corresponding DFT single-point calculations, in the same manner as the case of ANN-1 (Fig. 6(a)). Table 5 compares the MAE values of the two ANN potentials. The ANN-2 values are found to be smaller than the ANN-1 values, not only for the three GBs in the additional training dataset but also the other GBs that are not included in the training dataset. This may indicate that inclusion of high- $\Sigma$  GBs into the training dataset is not necessarily need to improve predictive ability of ANN structural relaxation. Rather, searching for low- $\Sigma$  GB structures with large errors by means of an

ANN potential trained once is probably an efficient way to reduce computational cost for ANN training.

Figure 10 shows the GB-structure-energy relationship predicted by ANN-2. The two points indicated by the black arrows correspond to the same GB structure and thereby the same GB energy, which is the lowest-energy relaxed structure with the DFT level. Recall that ANN-1 fails to find the lowest-energy relaxed structure with the DFT level, as shown in Fig. 7(a). This means that ANN-1 does not work well in screening of candidate relaxed structures for the  $\Sigma 5(310)/[001]$  STGB. This problem is remedied by ANN-2: Fig. 10 indicates that its second lowest-energy structure corresponds to the lowest-energy one obtained from DFT structural relaxation. This allows us to apply DFT calculation to only low-energy structures predicted by ANN-2, in order to find the exact lowest-energy structure with the DFT level.

It should be noted that similar to ANN-1, ANN-2 still has the same problem that some high-energy structures predicted by ANN-2 transform to lower-energy structures during DFT structural relaxation, as indicated by the red arrows in Fig. 10. This problem may be resolved by repeatedly applying the data generation scheme proposed above to the ANN potential. With this approach, the predictive power is expected to improve systematically and efficiently. Our future work will improve the ANN potential in this way and apply it to more complicated GBs such as asymmetric GBs and high- $\Sigma$  GBs.

#### 4. Conclusions

An ANN interatomic potential is constructed in order to accurately predict GB atomic structures and their energetics in CdTe, with practical computational cost. To construct the ANN potential transferable to various atomic environments including GBs, a training dataset is generated by considering not only perfect-crystal structures but also point defects, surfaces and GBs. Using the trained ANN potential, a  $\gamma$ -surface method with structural relaxation is applied to ten different GBs, in order to obtain GB atomic structures with various energy states. The relaxed structures obtained from the ANN potential are then used to calculate their GB energies with the DFT level. As a result, the ANN potential is found to correctly distinguish between low- and high-energy structures with respect to DFT results, even for GBs that are not included in the training dataset. In contrast, the SW and BO potentials fail to predict the energetics of GBs and low-energy structures, indicating that conventional empirical potentials are difficult to use to predict physically reasonable GB structures for CdTe. In order to improve the predictive ability of the ANN potential to GBs, this work also presents an approach in which the ANN potential trained once is utilized to explore GB structures that are not sufficiently covered by a given training dataset. It is indicated that the ANN potential constructed in this way exhibits a systematic improvement in the prediction of low-energy structures and GB energies. This approach will offer

a way to construct machine-learning potentials transferable to general GBs for not only CdTe but also other materials.

### Conflicts of interest

There are no conflicts to declare.

### Acknowledgements

This work was supported by JSPS KAKENHI (Grant Number JP19H05786 and JP21K14405) and JST-CREST (Grant Number JPMJCR17J1).

### References

1. J. D. Poplawsky, N. R. Paudel, C. Li, C. M. Parish, D. Leonard, Y. Yan, S. J. Pennycook, Direct imaging of Cl- and Cu-induced short-circuit efficiency changes in CdTe solar cells, *Adv. Energy Mater.*, 2014, **4**, 1400454.
2. J. Moseley, W. K. Metzger, H. R. Moutinho, N. Paudel, H. L. Guthrey, Y. Yan, R. K. Ahrenkiel, M. M. Al-Jassim, Recombination by grain-boundary type in CdTe, *J. Appl. Phys.*, 2015, **118**, 025702.
3. G. Stechmann, S. Zaefferer, T. Schwarz, P. Konijnenberg, D. Raabe, C. Gretener, L. Kranz, J. Perrenoud, S. Buecheler, A. N. A. Tiwari, Correlative investigation of grain boundary crystallography and electronic properties in CdTe thin film solar cells, *Sol. Energy Mater. Sol.*, 2017, **166**, 108-120.
4. Y. Yan, M. M. Al-Jassim, K. M. Jones, Structure and effects of double-positioning twin boundaries in CdTe, *J. Appl. Phys.*, 2003, **94**, 2976-2979.
5. C. Li, Y. Wu, J. Poplawsky, T. J. Pennycook, N. Paudel, W. Yin, S. J. Haigh, M. P. Oxley, A. R. Lupini, M. Al-Jassim, S. J. Pennycook, Y. Yan, Grain-boundary-enhanced carrier collection in CdTe solar cells, *Phys. Rev. Lett.*, 2014, **112**, 156103.
6. J. S. Park, J. Kang, J. H. Yang, W. Metzger, S. H. Wei, Stability and electronic structure of the low- $\Sigma$  grain boundaries in CdTe: a density functional study, *New J. Phys.*, 2015, **17**, 013027.
7. C. J. Tong, K. P. McKenna, Passivating Grain Boundaries in Polycrystalline CdTe, *J. Phys. Chem. C*, 2019, **123**, 23882-23889.
8. J. S. Park, Stabilization and self-passivation of symmetrical grain boundaries by mirror symmetry breaking, *Phys. Rev. Mater.*, 2019, **3**, 014602.
9. H. Nishimura, K. Matsunaga, T. Saito, T. Yamamoto, Y. Ikuhara, Atomic structures and energies of  $\Sigma 7$  symmetrical tilt grain boundaries in alumina bicrystals, *J. Am. Ceram. Soc.*, 2003, **86**, 574-580.

10. E. N. Hahn, S. J. Fensin, T. C. Germann, M. A. Meyers, Symmetric tilt boundaries in body-centered tantalum, *Scripta Mater.*, 2016, **116**, 108-111.
11. L. Wang, W. Yu, S. Shen, Revisiting the structures and energies of silicon <110> symmetric tilt grain boundaries, *J. Mater. Res.*, 2019, **34**, 1021-1033.
12. H. Guhl, H. S. Lee, P. Tangney, W. M. C. Foulkes, A. H. Heuer, T. Nakagawa, Y. Ikuhara, M. W. Finnis, Structural and electronic properties of  $\Sigma 7$  grain boundaries in  $\alpha$ -Al<sub>2</sub>O<sub>3</sub>, *Acta Mater.*, 2015, **99**, 16-28.
13. T. Yokoi, Y. Kondo, K. Ikawa, A. Nakamura, K. Matsuanga, Stable and metastable structures and their energetics of asymmetric tilt grain boundaries in MgO: a simulated annealing approach, *J. Mater. Sci.*, 2021, **56**, 3183-3196.
14. T. Frolov, W. Setyawan, R.J. Kurtz, J. Marian, A.R. Oganov, R.E. Rudd, Q. Zhu, Grain boundary phases in bcc metals, *Nanoscale*, 2018, **10**, 8253-8268.
15. Q. Zhu, A. Samanta, B. Li, R.E. Rudd, T. Frolov, Predicting phase behavior of grain boundaries with evolutionary search and machine learning, *Nat. Commun.*, 2018, **9**, 467.
16. N. A. Benedek, A. L. S. Chua, C. Elsässer, A. P. Sutton, M. W. Finnis, Interatomic potential for strontium titanate: An assessment of their transferability and comparison with density functional theory, *Phys. Rev. B*, 2008, **78**, 064110.
17. D. Scheiber, R. Pippan, P. Puschnig, L. Romaner, Ab initio calculations of grain boundaries in bcc metals, *Modelling Simul. Mater. Sci. Eng.*, 2016, **24**, 035013.
18. J. Behler, M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surface, *Phys. Rev. Lett.*, 2007, **98**, 146401.
19. A. P. Bartók, M. C. Payne, R. Kondor, G. Csányi, Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons, *Phys. Rev. Lett.*, 2010, **104**, 136403.
20. H. Babaei, R. Guo, A. Hashemi, S. Lee, Machine-learning-based interatomic potential for phonon transport in perfect crystalline Si and crystalline Si with vacancies, *Phys. Rev. Mater.*, 2019, **3**, 074603.
21. A. S. Bochkarev, A. van Roekeghem, S. Mossa, N. Mingo, Anharmonic thermodynamics of vacancies using a neural network potential, *Phys. Rev. Mater.*, 2019, **3**, 093803.
22. J. Byggmästar, A. Hamedani, K. Nordlund, F. Djurabekova, Machine-learning interatomic potential for radiation damage and defects in tungsten, *Phys. Rev. B*, 2019, **100**, 144105.
23. S. K. Natarajan, J. Behler, Neural network molecular dynamics simulations of solid-liquid interfaces: water at low-index copper surfaces, *Phys. Chem. Chem. Phys.*, 2016, **18**, 28704.
24. M. C. Groenenboom, T. P. Moffat, K. A. Schwarz, Halide-induced step faceting and dissolution energetics from atomistic machine learned potentials on Cu(100), *J. Phys. Chem. C*, 2020, **124**, 12359-12369.

25. H. Mori, T. Ozaki, Neural network atomic potential to investigate the dislocation dynamics in bcc iron, *Phys. Rev. Mater.*, 2020, **4**, 040601.
26. D. Dragoni, T. D. Daff, G. Csányi, N. Marzari, Achieving DFT accuracy with a machine-learning interatomic potential: Thermomechanics and defects in bcc ferromagnetic iron, *Phys. Rev. B*, 2018, **2**, 013808.
27. A. P. Bartók, J. Kermode, N. Bernstein, G. Csányi, Machine learning a general-purpose interatomic potential for silicon, *Phys. Rev. X*, 2018, **8**, 041048.
28. T. Yokoi, Y. Noda, A. Nakamura, K. Matsunaga, Neural-network interatomic potential for grain boundary structures and their energetics in silicon, *Phys. Rev. Mater.*, 2020, **4**, 014605.
29. T. Nishiyama, A. Seko, I. Tanaka, Application of machine learning potentials to predict grain boundary properties in fcc elemental metals, *Phys. Rev. Mater.*, 2020, **4**, 123607.
30. M. Mińkowski, K. Hummer, C. Dellago, Cation interstitial diffusion in lead telluride and cadmium telluride studied by means of neural network potential based molecular dynamics simulations, *J. Phys.: Condens. Matter.*, 2021, **33**, 015901.
31. P. E. Elöchl, Projector augmented-wave method, *Phys. Rev. B*, 1994, **50**, 17953-17979.
32. G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B*, 1999, **19**, 1758-1775.
33. G. Kresse, J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci.*, 1996, **6**, 15-50.
34. G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B*, 1996, **54**, 11169-11186.
35. J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, K. Burke, Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces, *Phys. Rev. Lett.*, 2008, **100**, 136406.
36. M. Parrinello, A. Rahman, Crystal Structure and Pair Potentials: A Molecular-Dynamics Study, *Phys. Rev. Lett.*, 1980, **45**, 1196-1199.
37. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.*, 1981, **52**, 7182-7190.
38. J. Behler, Constructing high-dimensional neural network potentials: A tutorial review, *Int. J. Quantum Chem.*, 2015, **115**, 1032-1050.
39. T. B. Blank, S. D. Brown, Adaptive, global, extended Kalman filters for training feedforward neural networks, *J. Chemom.*, 1994, **8**, 391-407.
40. S. Plimpton, Fast Parallel Algorithms for Short-Range Molecular Dynamics, *J. Comp. Phys.*, 1995, **117**, 1-19.
41. Z. Q. Wang, D. Stroud, A. J. Markworth, Monte Carlo study of the liquid CdTe surface, *Phys. Rev. B*, 1989, **40**, 3129-3132.

42. D. K. Ward, X. Zhou, B. M. Wong, F. P. A. Doty, Refined parameterization of the analytical Cd–Zn–Te bond-order potential, *J. Mol. Model.*, 2013, **19**, 5469–5477.
43. N. Artrith, J. Behler, High-dimensional neural network potentials for metal surfaces: A prototype study for copper, *Phys. Rev. B*, 2012, **85**, 045439.
44. A. Togo, I. Tanaka, First principles phonon calculations in materials science, *Scripta Mater.*, 2015, **108**, 1-5.

Table 1 Training datasets of single-crystal, surface and GB structures.  $N_{\text{str}}$  and  $N_{\text{force}}$  are the number of total energies and atomic forces for training the ANN potential, respectively.  $N_{\text{atom}}$  is the number of atoms contained in a simulation cell.

Dataset	$N_{\text{str}}$	$N_{\text{atom}}$	$N_{\text{force}}$	MAE		
				Energy	Force	
				[meV/atom]	[meV/Å]	
Single crystal	33650	7-64	1790100	5.90	40.41	
Surface	(100)	10680	16-64	544320	1.61	40.77
	(110)	5700	12	102600	1.51	21.67
	(111)	6700	12	120600	2.61	21.40
	(112)	5690	24	204840	1.37	23.15
	(310)	6700	20	201000	1.59	28.68
	(510)	1200	52	93600	1.46	33.13
GB	$\Sigma 3(112)/[112]$ twist	10010	46-48	720720	2.17	42.32
	$\Sigma 3(112)/[\bar{1}10]$ STGB Cd-core	3830	46-48	273924	2.16	36.01
	$\Sigma 3(112)/[\bar{1}10]$ STGB Te-core	3850	46-48	275448	2.24	37.70
	$\Sigma 3(112)$ mix Cd-core	3780	46-48	270456	2.51	35.18
	$\Sigma 3(112)$ mix Te-core	3800	46-48	271884	2.67	35.63
	$\Sigma 5(210)/[001]$ STGB	10990	38-40	658149	2.02	42.61
	$\Sigma 5(310)/[001]$ STGB	11470	38-40	686904	3.23	55.76

Table 2 Test dataset consisting the three GB structures. Note that these GBs are the same as those contained in the training dataset, but the simulation cells in this test dataset are not included in the training dataset.

Structure	$N_{\text{str}}$	$N_{\text{atom}}$	$N_{\text{force}}$	MAE	
				Energy	Force
				[meV/atom]	[meV/atom]
$\Sigma 3(112)$ twist	600	48	86400	3.12	41.32
$\Sigma 5(210)$ tilt	600	40	72000	2.50	39.53
$\Sigma 5(310)$ tilt	600	40	72000	4.02	55.79

Table 3 DFT GB energies of the lowest-energy structures obtained from the ANN, SW and BO potentials.

GB	GB energy [J/m <sup>2</sup> ]		
	ANN	SW	BO
$\Sigma 3(111)$	0.01	0.01	0.00
$\Sigma 3(112)$	0.45	1.01	0.63
$\Sigma 5(210)$	0.49	0.80	0.77
$\Sigma 5(310)$	0.60	0.66	0.65
$\Sigma 9(114)$	0.40	0.48	0.44
$\Sigma 9(221)$	0.50	0.64	0.63
$\Sigma 11(113)$	0.25	0.33	0.28
$\Sigma 11(332)$	0.49	0.53	0.52
$\Sigma 13(320)$	0.53	1.00	0.88
$\Sigma 13(510)$	0.51	0.76	0.70

Table 4 GB energies calculated from the SW and BO potentials by using the lowest-energy structures obtained from the ANN potential for the  $\Sigma 9(114)$ ,  $\Sigma 13(320)$  and  $\Sigma 13(510)$  GBs.

GB	GB energy [J/m <sup>2</sup> ]					
	SW potential			BO potential		
	SW lowest-energy structure	ANN lowest-energy structure		BO lowest-energy structure	ANN lowest-energy structure	
		Single-point calculation	Structural relaxation		Single-point calculation	Structural relaxation
	$\Sigma 9(114)$	0.58	0.86	0.57	0.58	0.70
$\Sigma 13(320)$	0.43	1.37	0.81	0.70	1.29	0.92
$\Sigma 13(510)$	0.43	0.91	0.61	0.58	0.98	0.75

Table 5 MAE with respect to DFT single-point calculations. To obtain these values, various GB structures and their GB energies are obtained by performing the  $\gamma$ -surface method with ANN-1 and ANN-2, and then their GB energies with the DFT level are calculated by performing DFT single-point calculations.

GB type	MAE [J/m <sup>2</sup> ]	
	ANN-1	ANN-2
$\Sigma 3(112)/[112]$	0.07	0.05
$\Sigma 5(210)/[001]$	0.04	0.01
$\Sigma 5(310)/[001]$	0.09	0.07
$\Sigma 9(114)/[114]$	0.12	0.10
$\Sigma 9(221)/[221]$	0.05	0.04
$\Sigma 11(113)/[113]$	0.06	0.05
$\Sigma 11(332)/[332]$	0.05	0.05
$\Sigma 13(320)/[001]$	0.06	0.04
$\Sigma 13(510)/[001]$	0.09	0.07
Average	0.075	0.060

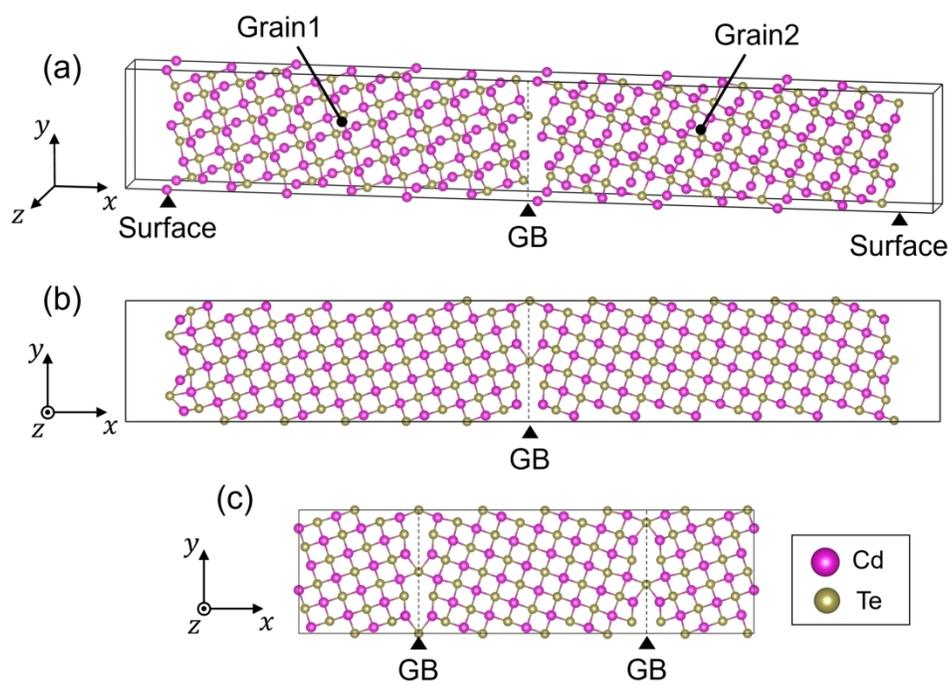


Fig. 1. Schematic simulation cell used in the  $\gamma$ -surface method with the ANN potential: (a) An initial cell containing one GB and two surfaces, (b) the corresponding relaxed structure, and (c) the final structure reconstructed from structure (b) with 3D periodic boundary conditions, where two same GB structures are present whereas the surface region is absent.

195x142mm (300 x 300 DPI)

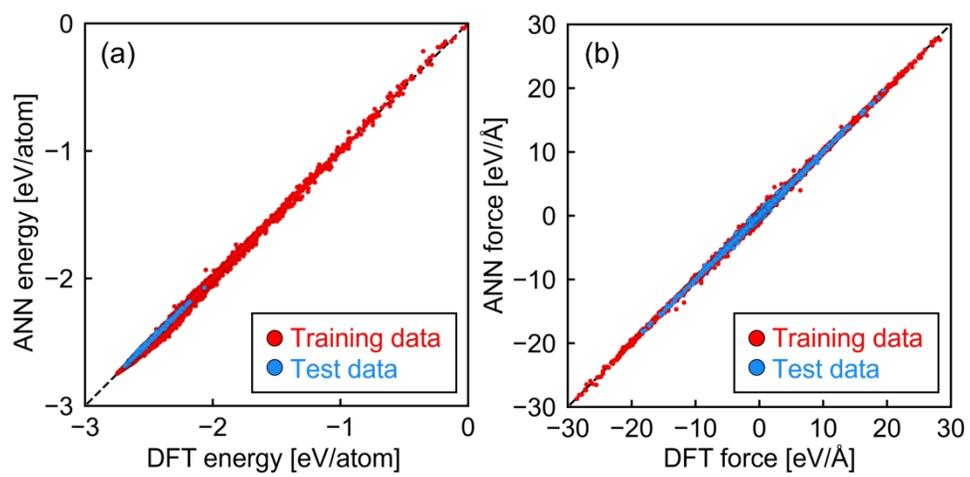


Fig. 2. Correlation between values obtained from the ANN potential and DFT calculations for (a) the total energy and (b) the atomic force in the training (red point) and test (blue point) datasets.

243x124mm (300 x 300 DPI)

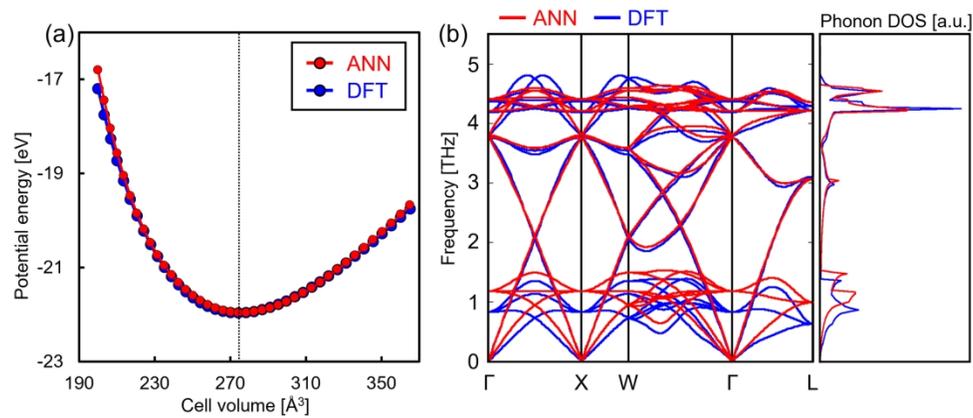


Fig. 3. (a) Potential-energy increase as a function of cell volume and (b) the phonon dispersion curve and density of states (DOS). The red and blue lines correspond to the values obtained from the ANN potential and DFT calculations, respectively. The lattice vibration properties were obtained by performing lattice dynamics calculations implemented in the PHONOPY code<sup>44</sup>. In this calculation, a  $3 \times 3 \times 3$  supercell of the zincblende structure were used.

231x101mm (300 x 300 DPI)

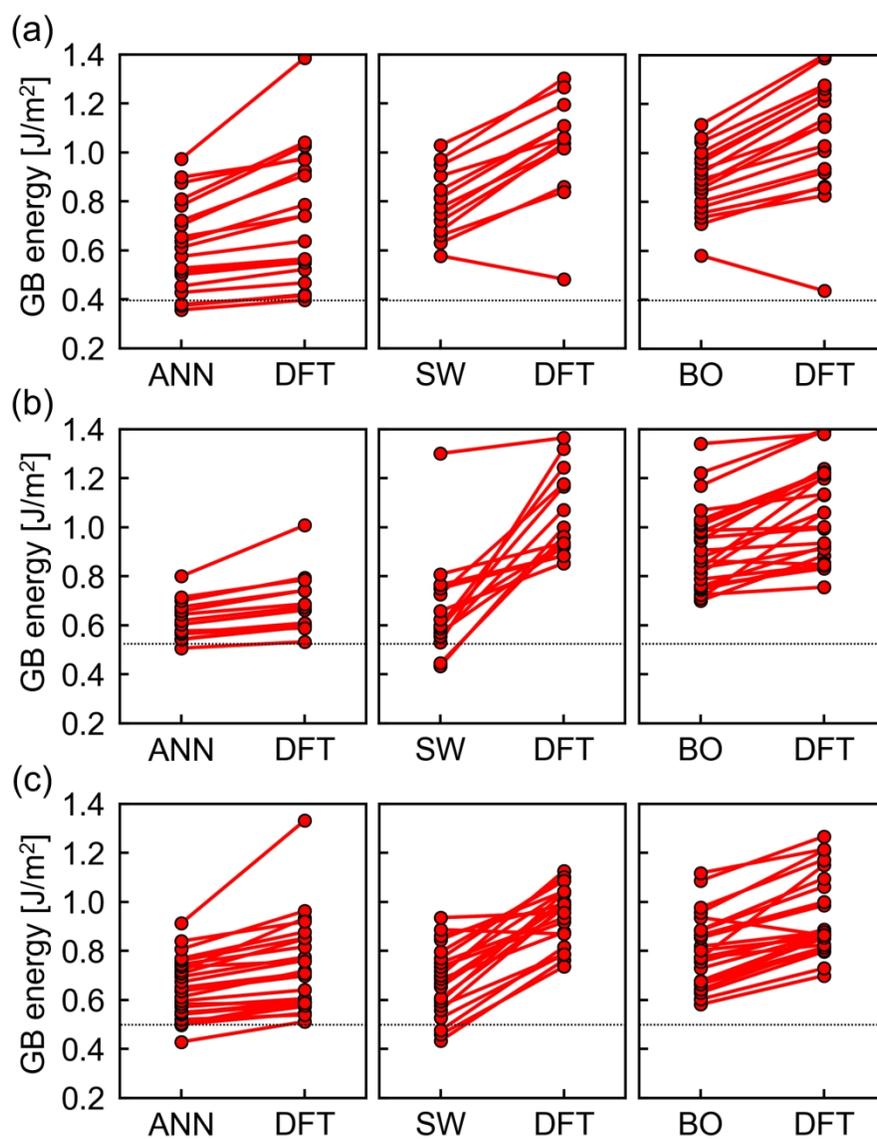


Fig. 4. GB energies of the ANN, SW and BO potentials for (a) the  $\Sigma 9(114)/[114]$  twist GB, (b) the  $\Sigma 13(320)/[001]$  STGB and (c) the  $\Sigma 13(510)/[001]$  STGB. Using relaxed structures obtained from each of the potentials, GB energies are calculated by performing DFT single-point calculations.

187x238mm (300 x 300 DPI)

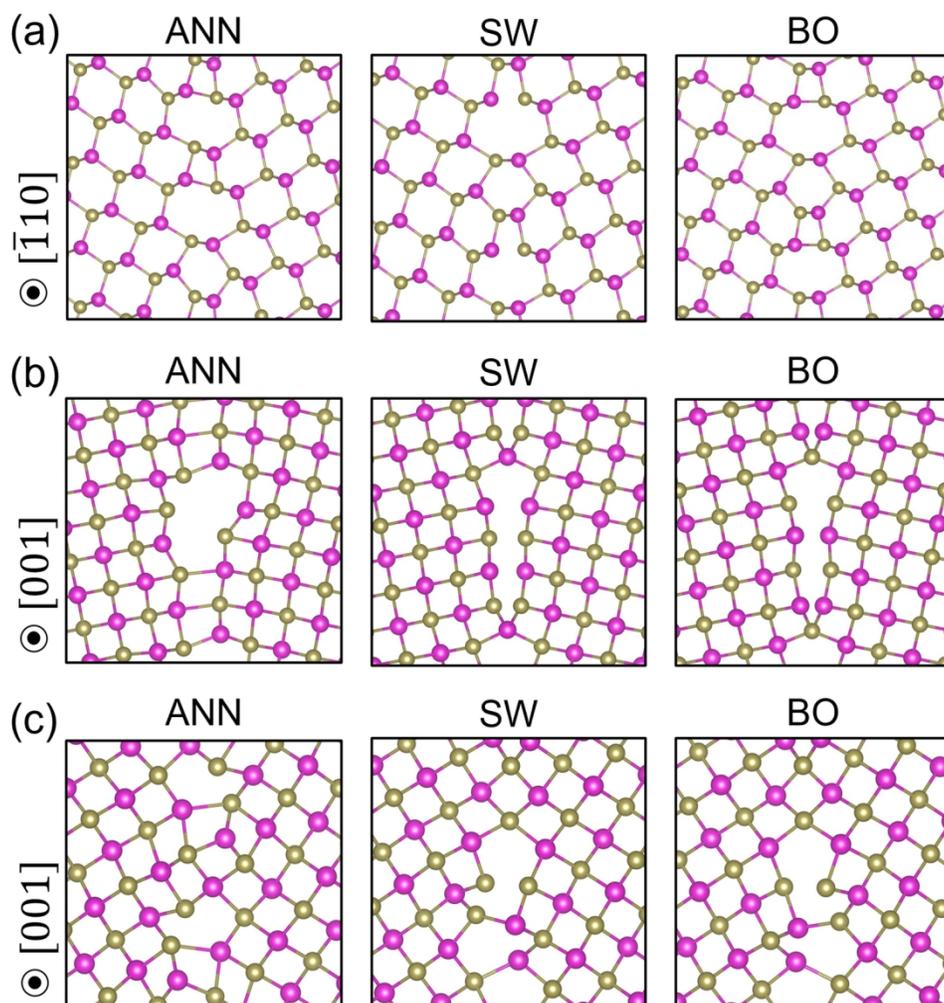


Fig. 5. Atomic arrangements of the lowest-energy GB structures obtained from the ANN, SW and BO potentials for (a) the  $\Sigma 9(114)/[114]$  twist GB, (b) the  $\Sigma 13(320)/[001]$  STGB, and (c) the  $\Sigma 13(510)/[001]$  STGB.

152x159mm (300 x 300 DPI)

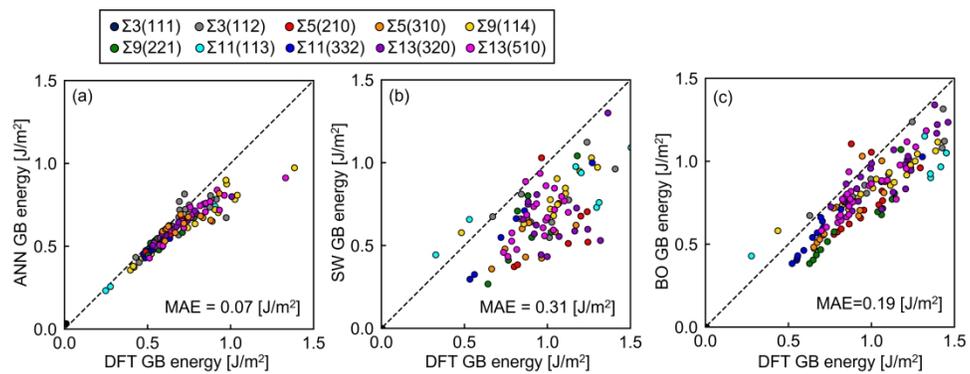


Fig. 6. Correlation in GB energy of DFT single-point calculations to (a) the ANN potential and (b) the SW potential.

375x146mm (300 x 300 DPI)

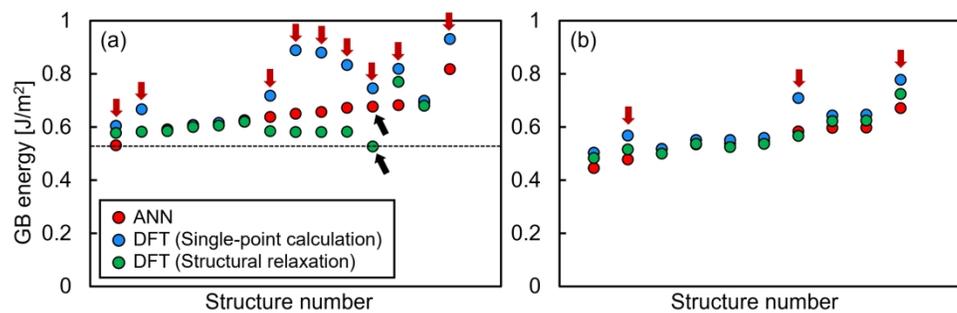


Fig. 7. GB-structure-energy relationship calculated from the ANN potential, DFT single-point calculation and DFT structural relaxation, corresponding to the red, blue and green points, respectively: (a) the  $\Sigma 5(310)/[001]$  STGB and (b) the  $\Sigma 9(221)/[221]$  twist GB.

259x88mm (300 x 300 DPI)

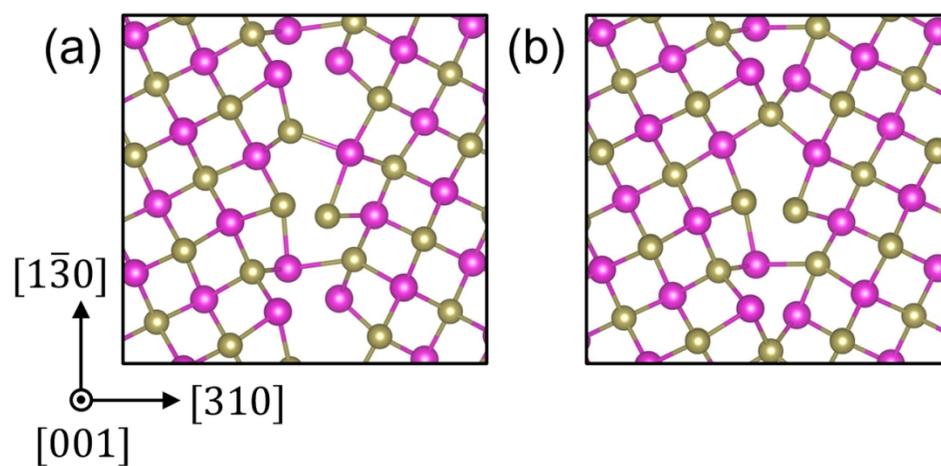


Fig. 8 GB structures predicted by (a) the ANN potential and (b) obtained from DFT structural relaxation. Their GB energies are pointed by the black arrows in Fig. 7(a).

116x59mm (300 x 300 DPI)

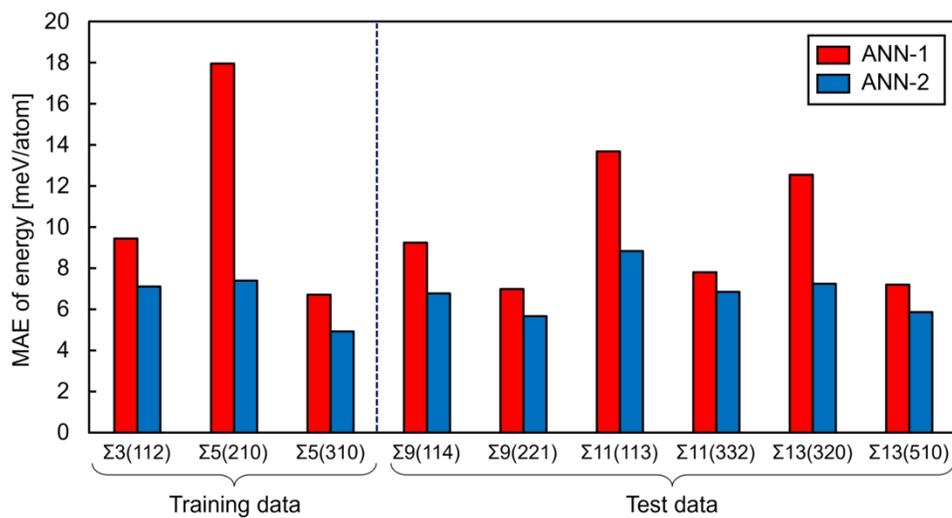


Fig. 9. Energy MAE of the training and test datasets. The red and blue bars corresponds to the results of ANN-1 and ANN-2, respectively.

212x114mm (300 x 300 DPI)

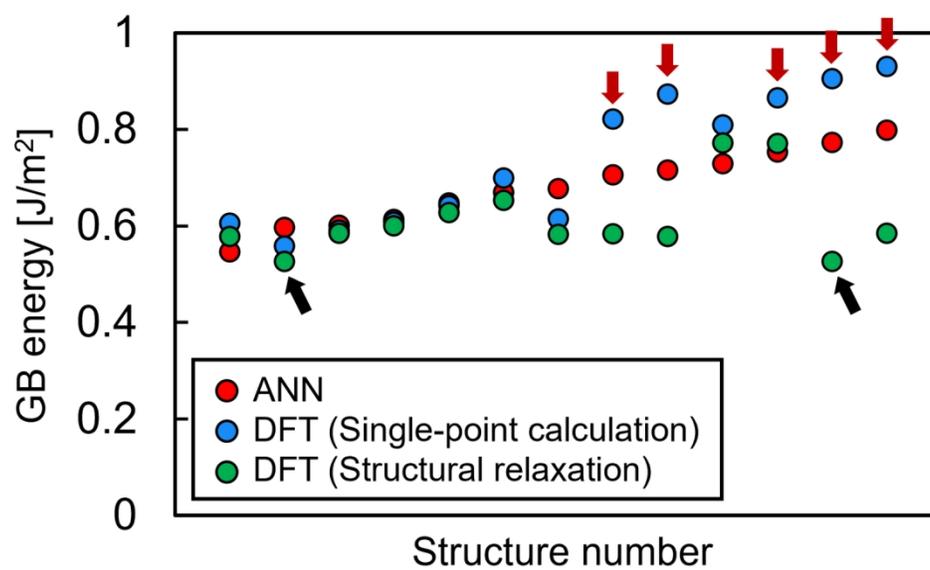


Fig. 10. GB-structure-energy relationship obtained from ANN-2, DFT single-point calculation and DFT structural relaxation for the  $\Sigma 5(310)/[001]$  STGB.

144x89mm (300 x 300 DPI)