Supplementary Materials

Data-driven strategy for bandgap database construction of perovskites and the potential segregation study

Bobin Wu1,2,3, Xinyu Zhang1,2, Zixuan Wang1,2, Zijian Chen1,2, Shaohui Liu1,2,3, Jie Liu4, Zhenming Xu5,*, Qingde Sun6,7,*, Haitao Zhao1,2,*

1Center for Intelligent and Biomimetic Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong, China.
2Digital Intelligent Manufacturing Research Center, Wenzhou Institute of Technology of the Chinese Academy of Sciences, Wenzhou, 325000, Zhejiang, China.
3Nano Science and Technology Institute, University of Science and Technology of China, Suzhou 215000, Jiangsu, China.
4Department of Chemistry, The University of Hong Kong, Hong Kong 999077, China.
5Jiangsu Key Laboratory of Electrochemical Energy Storage Technologies, College of Materials Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, Jiangsu, China.
6School of Physics and Electronic Science, Changsha University of Science and Technology, Changsha 410114, Hunan, China.
7School of Materials Science & Engineering, Nanyang Technological University, Singapore 639798, Singapore.

*Correspondence to: Dr. Haitao Zhao, Center for Intelligent and Biomimetic Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, 1068 Xueyuan Avenue, Shenzhen University Town, Shenzhen 518055, Guangdong, China. E-mail: haitaozhao@outlook.com; Dr. Zhenming Xu, Jiangsu Key Laboratory of Electrochemical Energy Storage Technologies, College of Materials Science and Technology, Nanjing University of Aeronautics and Astronautics, No. 29 Yudao Street, Nanjing 210016, Jiangsu, China. E-mail: xuzhenming@nuaa.edu.cn; Dr. Qingde Sun, School of Materials Science & Engineering, Nanyang Technological University, 50 Nanyang Ave, Singapore 639798, Singapore. E-mail: qingde.sun@ntu.edu.sg
**Supplementary Table 1. Formulas for the bandgap in eV as a function of relative Br concentration x for the five compounds**

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Bandgap</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cubic FAPb(I_{1-x}Br_x)(_3)</td>
<td>1.48(1 - x) + 2.23x - 0.15x(1 - x)</td>
<td>[1]</td>
</tr>
<tr>
<td>Cubic MAPb(I_{1-x}Br_x)(_3)</td>
<td>1.57(1 - x) + 2.29x - 0.33x(1 - x)</td>
<td>[2]</td>
</tr>
<tr>
<td>Cubic CsPb(I_{1-x}Br_x)(_3)</td>
<td>1.77(1 - x) + 2.38x - 0.35x(1 - x)</td>
<td>[3]</td>
</tr>
<tr>
<td>Tetragonal MAPb(I_{1-x}Br_x)(_3)</td>
<td>1.57(1 - x) + 2.25x - 0.3x(1 - x)</td>
<td>[4]</td>
</tr>
<tr>
<td>Orthorhombic CsPb(I_{1-x}Br_x)(_3)</td>
<td>1.73(1 - x) + 2.34x - 0.22x(1 - x)</td>
<td>[5]</td>
</tr>
</tbody>
</table>
**Supplementary Figure 1.** DFT calculated electronic band structures for six representative mixed halide perovskites in cubic structure. (A) $\text{FA}_{0.75}\text{Cs}_{0.25}\text{PbI}_3$; (B) $\text{FA}_{0.5}\text{MA}_{0.5}\text{PbBr}_3$; (C) $\text{FA}_{0.5}\text{MA}_{0.5}\text{Pb(I0.33Br0.67)}_3$; (D) $\text{FA}_{0.5}\text{MA}_{0.5}\text{Pb(I0.5Br0.5)}_3$; (E) $\text{FA}_{0.5}\text{MA}_{0.5}\text{Pb(I0.67Br0.33)}_3$; (F) $\text{FA}_{0.5}\text{MA}_{0.5}\text{Pb(I0.83Br0.17)}_3$.

**Supplementary Figure 2.** Performances of fine-tuned MEGNet in different datasets. (A) DFT bandgap dataset; (B) The experimental bandgap dataset; (C) The gap between experimental and DFT bandgap dataset.
**Supplementary Figure 3.** Total energy prediction results of different graph neural networks. (A) M3GNet; (B) CHGNet.

**Supplementary Figure 4.** Prediction results of different methods in DFT bandgap dataset. (A) Pretrained MatGL; (B) Pretrained MEGNet; (C) DFT method.

**Supplementary Figure 5.** The probability density distribution of perovskite bandgap database.
**Supplementary Figure 6.** (A) Decomposition energy; (B) Mixing energy; (C) Formation energy; (D) Energy above hull of FAPb(I_{1-x}Br_x)₃.

**Supplementary Figure 7.** $\Delta E_g$ for the different compounds (APb(I_{1-x}Br_x)₃). A = (A) FA_{0.75}Cs_{0.25}, FA_{0.5}Cs_{0.5} and FA_{0.25}Cs_{0.75}; (B) MA_{0.75}Cs_{0.25}, MA_{0.5}Cs_{0.5} and MA_{0.25}Cs_{0.75}; (C) FA_{0.75}MA_{0.25}, FA_{0.5}MA_{0.5} and FA_{0.25}MA_{0.75}.
REFERENCES


