Supplementary Materials

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

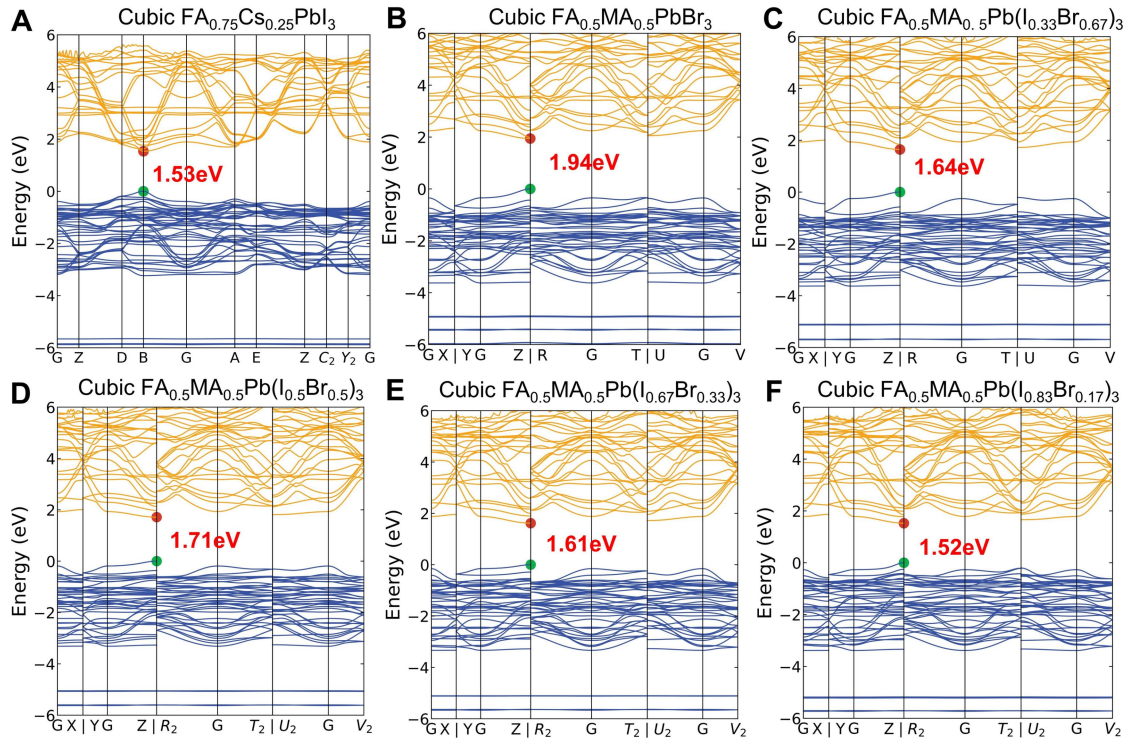
Bobin Wu[1,2,3], Xinyu Zhang[1,2], Zixuan Wang[1,2], Zijian Chen[1,2], Shaohui Liu[1,2,3], Jie Liu[4], Zhenming Xu[5,*], Qingde Sun[6,7,*], Haitao Zhao[1,2,*]

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

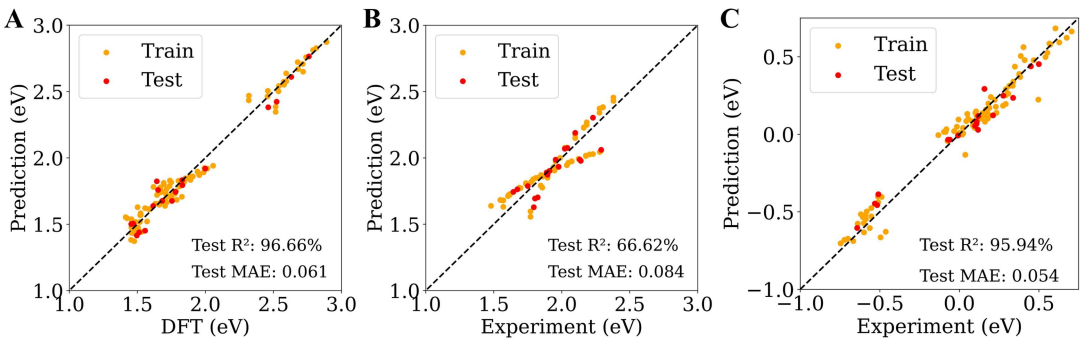
[*]**Correspondence to:** Dr. Haitao Zhao, Center for Intelligent and Biomimetic Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, 1068 Xueyuan Avenue, Shenzhen University Town, Shenzhen 518055, Guangdong, China. E-mail: haitaozhao@outlook.com; Dr. Zhenming Xu, Jiangsu 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

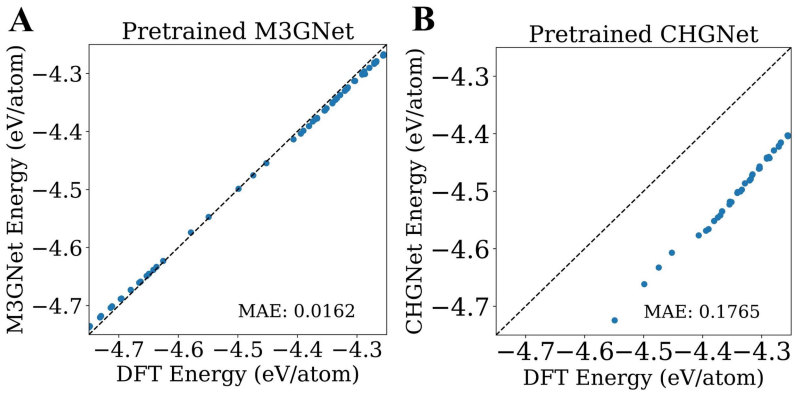
Compounds	Bandgap	Ref
Cubic $FAPb(I_{1-x}Br_x)_3$	$1.48(1 - x) + 2.23x - 0.15x(1 - x)$	[1]
Cubic $MAPb(I_{1-x}Br_x)_3$	$1.57(1 - x) + 2.29x - 0.33x(1 - x)$	[2]
Cubic $CsPb(I_{1-x}Br_x)_3$	$1.77(1 - x) + 2.38x - 0.35x(1 - x)$	[3]
Tetragonal $MAPb(I_{1-x}Br_x)_3$	$1.57(1 - x) + 2.25x - 0.3x(1 - x)$	[4]
Orthorhombic $CsPb(I_{1-x}Br_x)_3$	$1.73(1 - x) + 2.34x - 0.22x(1 - x)$	[5]



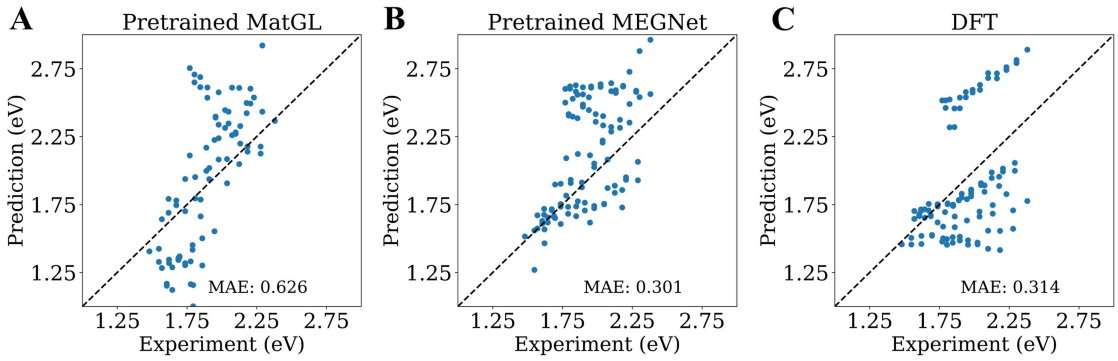
Supplementary Figure 1. DFT calculated electronic band structures for six representative mixed halide perovskites in cubic structure. (A) $FA_{0.75}Cs_{0.25}PbI_3$; (B) $FA_{0.5}MA_{0.5}PbBr_3$; (C) $FA_{0.5}MA_{0.5}Pb(I_{0.33}Br_{0.67})_3$; (D) $FA_{0.5}MA_{0.5}Pb(I_{0.5}Br_{0.5})_3$; (E) $FA_{0.5}MA_{0.5}Pb(I_{0.67}Br_{0.33})_3$; (F) $FA_{0.5}MA_{0.5}Pb(I_{0.83}Br_{0.17})_3$.



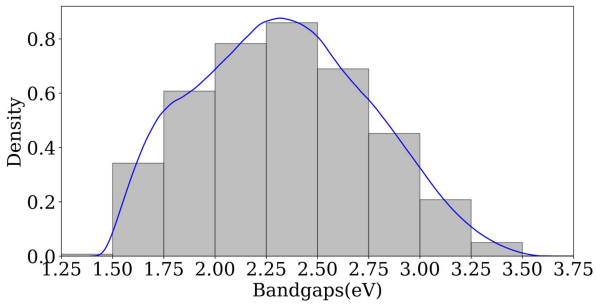
Supplementary Figure 2. Performances of fine-tuned MEGNet in different datasets. (A) DFT bandgap dataset; (B) The experimental bandgap dataset; (C) The gap between experimental and DFT bandgap dataset.



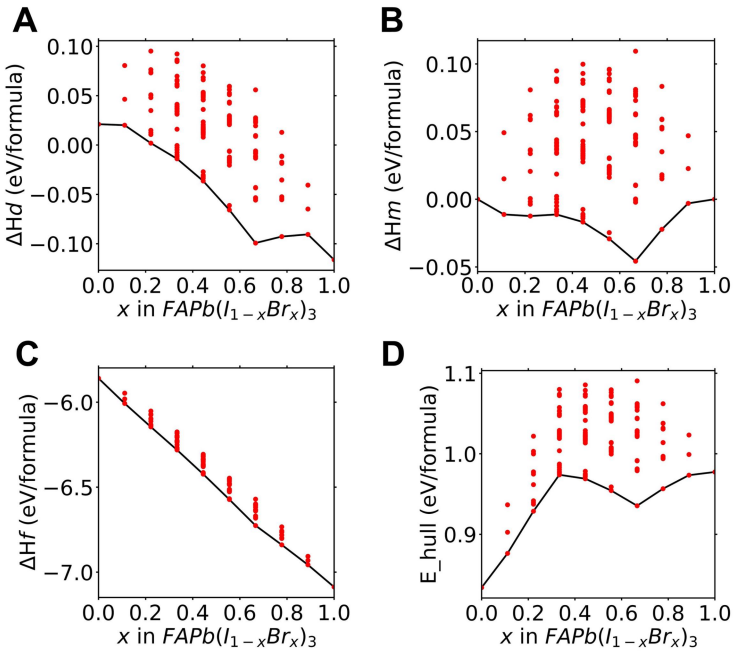
Supplementary Figure 3. Total energy prediction results of different graph neural networks. (A) M3GNet; (B) CHGNet.



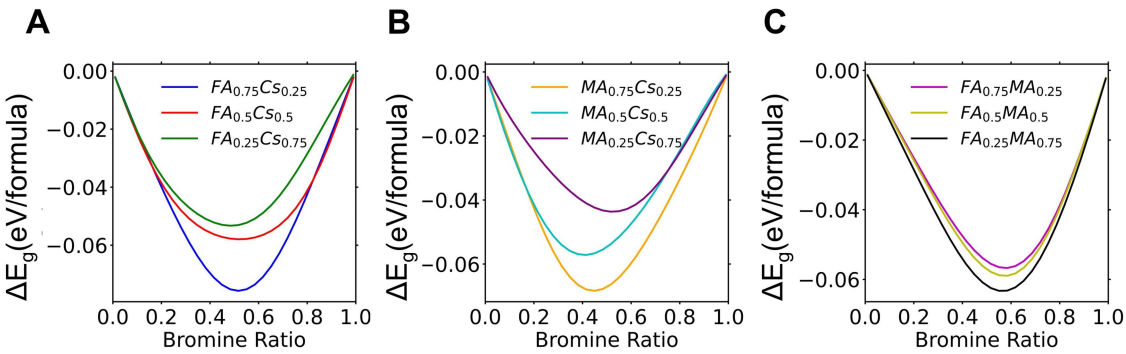
Supplementary Figure 4. Prediction results of different methods in DFT bandgap dataset. (A) Pretrained MatGL; (B) Pretrained MEGNet; (C) DFT method.



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. ΔE_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

1. Eperon GE, Stranks SD, Menelaou C, Johnston MB, Herz LM, et al. Formamidinium lead trihalide: a broadly tunable perovskite for efficient planar heterojunction solar cells. *Energy & Environmental Science* 2014;7:982. https://doi.org/10.1039/c3ee43822h

2. Noh JH, Im SH, Heo JH, Mandal TN, Seok SI. Chemical Management for Colorful, Efficient, and Stable Inorganic–Organic Hybrid Nanostructured Solar Cells. *Nano Letters* 2013;13:1764-9. https://doi.org/10.1021/nl400349b

3. Beal RE, Slotcavage DJ, Leijtens T, Bowring AR, Belisle RA, et al. Cesium Lead Halide Perovskites with Improved Stability for Tandem Solar Cells. *The Journal of Physical Chemistry Letters* 2016;7:746-51. https://doi.org/10.1021/acs.jpclett.6b00002

4. Atourki L, Vega E, Marí B, Mollar M, Ahsaine HA, et al. Role of the chemical substitution on the structural and luminescence properties of the mixed halide perovskite thin MAPbI3−xBrx ($0 \leq x \leq 1$) films. *Applied Surface Science* 2016;371:112-7. https://doi.org/10.1016/j.apsusc.2016.02.207

5. Jung J, Yun Y, Yang SW, Oh HG, Jeon A-Y, et al. Ternary diagrams of phase, stability, and optical properties of cesium lead mixed-halide perovskites. *Acta Materialia* 2023;246:118661. https://doi.org/10.1016/j.actamat.2022.118661