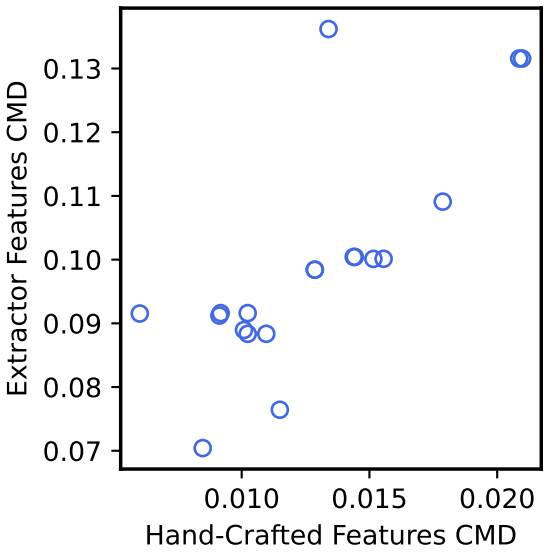
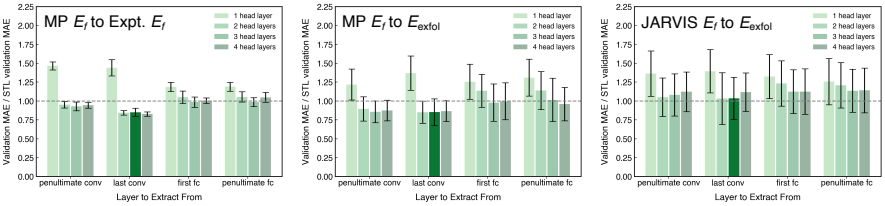
Supplementary Information

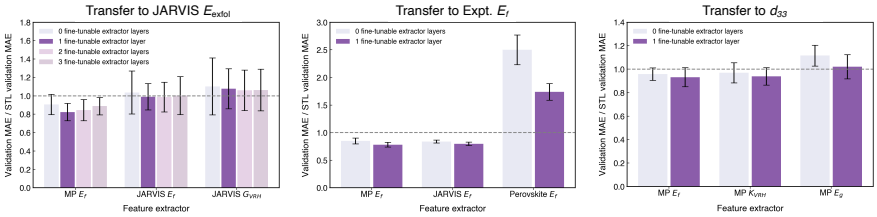
Supplementary Figures



Supplementary Figure 1 Comparison of CMD in feature space between MP E_{f} and downstream tasks using a learned featurization versus a procedurally generated featurization. The procedurally generated featurization uses `matminer.featurizers.site.CrystalNNFingerprint` and `matminer.featurizers.structure.SiteStatsFingerprint` [1]. 500 structures were subsampled from each dataset during procedural generation for efficiency.



Supplementary Figure 2 Bar plots of pairwise transfer learning validation error (normalized by STL validation error on the same random split) with various source and downstream tasks when extracting from different frozen, pre-trained layers and when applying different head sizes. 'Penultimate conv', 'last conv', 'first fc', and 'penultimate fc' denote the second to last convolutional layer, the last convolutional layer, the first head layer, and the second to last head layer, respectively. The darkest color bar indicates the chosen hyperparameter setting. 3 head layers were preferred to match the original CGCNN architecture.



Supplementary Figure 3 Bar plots of pairwise transfer learning validation error (normalized by STL validation error on the same random split) when fine-tuning different numbers of pre-trained extractor layers with various source and downstream tasks. The darkest color bar indicates the chosen hyperparameter setting.

Supplementary Tables

Supplementary Table 1
Hyperparameter settings

Model Hyperparameters	
orig_atom_fea_len	92
n_conv	4
nbr_fea_len	41
atom_fea_len	64
n_h	1
h_fea_len	32
Data Hyperparameters	
max_num_nbr	12
radius	8
dmin	0
step	0.2

Supplementary Table 2 Source tasks and corresponding pre-trained model test mean absolute errors (MAEs) and their mean absolute deviation (MAD)-normalized values

Source task	Dataset size	Test MAE	Test MAE/MAD
MP E_f[1]	132,752	0.0352	0.0350
MP E_g[2]	106,113	0.243	0.183
MP G_{VRH}[3]	10,987	0.0931	0.317
MP K_{VRH}[4]	10,987	0.0749	0.259
n-type σ_e[5]	37,390	1.05	0.106
p-type σ_e[6]	37,390	0.989	0.0605
n-type κ_e[7]	37,390	0.545	0.129
p-type κ_e[8]	37,390	0.611	0.125
p-type S[9]	37,390	105.	0.325
n-type S[10]	37,390	99.0	0.389
n-type $\overline{m}_e^*$[11]	21,037	0.552	0.610
p-type $\overline{m}_h^*$[12]	20,270	0.706	0.711
Perovskite E_f[13]	18,928	0.0488	0.0863
JARVIS E_f[14]	25,923	0.0563	0.0650
JARVIS dielectric constant (Opt)[15]	19,027	0.285	0.245
JARVIS E_g[16]	23,455	0.277	0.224
JARVIS G_{VRH}[17]	10,855	0.238	0.354
JARVIS K_{VRH}[18]	11,028	0.105	0.160

All datasets were acquired through `Matminer` [1]. Examples with NaNs were discarded.

[1]Formation energies from Materials Project [eV/at] [2]

[2]PBE band gaps from Materials Project [eV] [2]

[3]VRH-average shear modulus from Materials Project [$\log_{10}$(GPa)] [2]

[4]VRH-average bulk modulus from Materials Project [$\log_{10}$(GPa)] [2]

[5]Average conductivity eigenvalue with 10^{-18} electrons/cm^{-3} at 300K [log(1+(1/Ω/m/s))] [3]

[6]Average conductivity eigenvalue with 10^{-18} holes/cm^{-3} at 300K [log(1+(1/Ω/m/s))] [3]

[7]Average eigenvalue of electrical thermal conductivity with 10^{-18} carriers/cm^{-3} (n-type) at 300K [log(1+(W/K/m/s))] [3]

[8]Average eigenvalue of electrical thermal conductivity with 10^{-18} carriers/cm^{-3} (p-type) at 300K [log(1+(W/K/m/s))] [3]

[9]Average eigenvalue of the Seebeck coefficient with 10^{-18} holes/cm^{-3} [μ V/K] [3]

[10]Average eigenvalue of the Seebeck coefficient with 10^{-18} electrons/cm^{-3} [μ V/K] [3]

[11]Log of one plus the average eigenvalue of conductivity effective mass with 10^{-18} electrons/cm^{-3} (n-type) at 300K. Values greater than 100 were discarded. [3]

[12]Log of one plus the average eigenvalue of conductivity effective mass with 10^{-18} holes/cm^{-3} at 300K. Values greater than 100 were discarded. [3]

[13]Heat of formation. The reference state for oxygen was computed from oxygen's chemical potential in water vapor [eV/at] [4]

[14]Formation energy per atom from the JARVIS database [eV/at] [5]

[15]Log of one plus the average of static dielectric functions in x, y, and z directions calculated with the OptB88vDW functional [5]

[16]Band gaps from the JARVIS database [eV] [5]

[17]VRH-average shear moduli from the JARVIS database [log(1+GPa)] [5]

[18]VRH-average bulk moduli from the JARVIS database [log(1+GPa)] [5]

Supplementary References

[1] Ward, L., et al.: Matminer: An open source toolkit for materials data mining. Comput. Mater. Sci **152**, 60–69 (2018)

[2] Jain, A., et al.: Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. APL Mater **1**, 11002 (2013)

[3] Ricci, F., et al.: An ab initio electronic transport database for inorganic materials. Sci. Data **4**, 1–13 (2017)

[4] Castelli, I.E., et al.: New cubic perovskites for one- and two-photon water splitting using the computational materials repository. Energy Environ. Sci. **5**, 9034–9043 (2012)

[5] Choudhary, K., et al.: The joint automated repository for various integrated simulations (JARVIS) for data-driven materials design. npj Comput. Mater. **6**, 1–13 (2020)