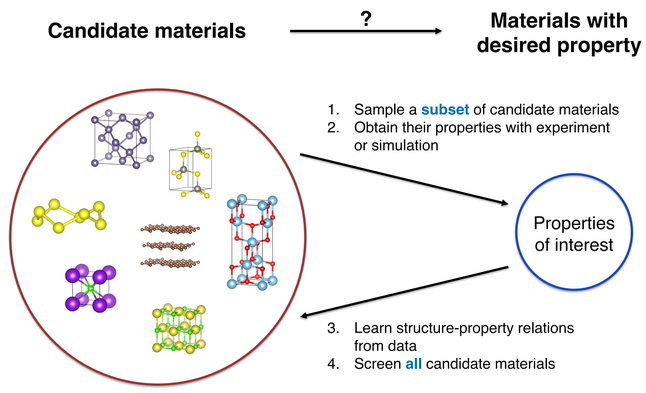
CGCNN: a graph representation of materials for property prediction and materials design

Tian Xie[1], Jeffrey Grossman[1, 2]

[1] Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
[2] Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
Contact: txie@mit.edu; jcg@mit.edu

Accelerate materials design with machine learning

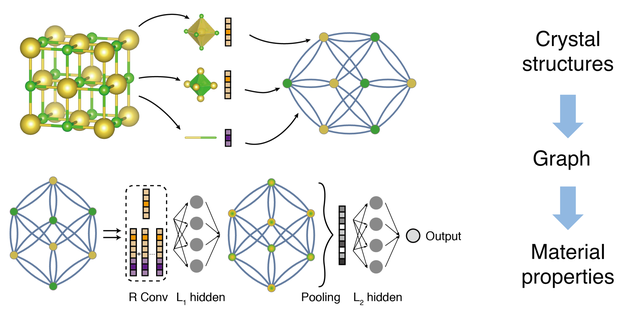


Need machine learning algorithms that:

- Predict material properties with high accuracy at a speed orders of magnitudes faster than quantum simulations/experiments.
- Assist materials design by providing insights of structure-property relations.

A graph-based end-to-end learning framework

Graph-based: consistent with human tuition, easy to understand
End-to-end learning: no prior knowledge, general to arbitrary materials and properties

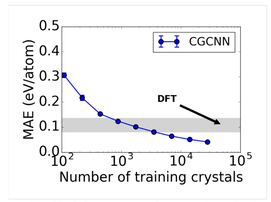


Crystal **G**raph **C**onvolutional **N**eural **N**etworks (**CGCNN**)

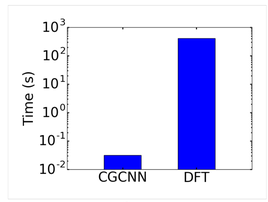
- **Graph representation**: atom → **node**; bond → **edge**
- **Convolution**: incorporate information of surrounding nodes and edges into each node → learns **local** environments
- **Pooling**: summing over each node to obtain an overall representation of the crystal → learns **global** representations

High prediction performance

We obtain highly accurate predictions of materials properties after training with data from the Materials Project.



Reliably reproduce simulation results after ~10^4 training data



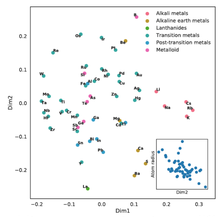
Speed ~10^4 **faster** than density functional theory calculations

- Generalizes to all existing inorganic crystalline materials, covering **various structure types** and **compositions**.
- High prediction performance for **7 different properties**.

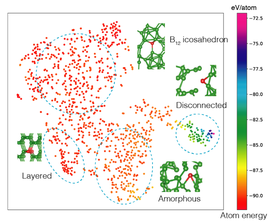
Property	# of train data	Unit	MAE_{model}	MAE_{DFT}
Formation energy	28 046	eV/atom	0.039	0.081–0.136 [28]
Absolute energy	28 046	eV/atom	0.072	···
Band gap	16 458	eV	0.388	0.6 [32]
Fermi energy	28 046	eV	0.363	···
Bulk moduli	2041	log(GPa)	0.054	0.050 [13]
Shear moduli	2041	log(GPa)	0.087	0.069 [13]
Poisson ratio	2041	···	0.030	···

Understand material similarities at different scales

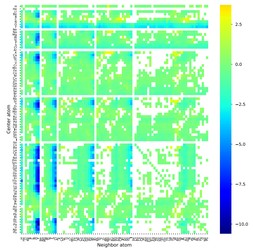
By visualizing different layers of the neural networks, we learn materials similarities at different scales.



Perovskites: same structure, different composition → element similarities



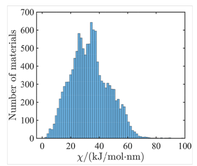
Boron: different structure, same composition → structural similarities



Inorganic crystals: different structure, different composition → a stability map for all possible coordination environments in crystals combining different center and neighbor atoms.

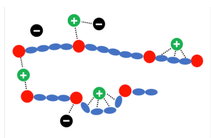
Use cases: screen electrolytes for batteries

We applied CGCNN to screen inorganic materials for solid electrolytes in Li metal anode batteries and polymer materials for solid electrolytes in Li-ion batteries.



Inorganic solid electrolytes for Li metal anode batteries

- Li metal anode batteries have high capacity, but dendritic electrodeposition of lithium metal during cycling causes **short circuit** and **capacity loss**.
- One can estimate the stability of electrodeposition with a stability parameter χ.
- We predicted the stability parameter χ of ~13,000 materials using ~3,400 training data from the Materials Project using CGCNN.



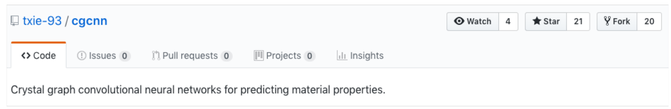
Polymer solid electrolytes for Li-ion batteries

- Polymer electrolytes are safer and less expensive, but they suffer from **low Li-ion conductivity** compared with liquid electrolytes.
- One can control the Li-ion conductivity by modifying the polymer structure.
- We are screening polymer electrolytes combining molecular dynamics and experiments to discover novel polymer electrolytes with high Li-ion conductivity.

Summary

- CGCNN provides a general framework to learn structure-property relations for **arbitrary materials**.
- High prediction performance helps **accelerate** screening new materials by reducing the number of simulations/experiments needed.
- Graph-based architecture makes it easier to **understand** the structure-property relations learned by the model, providing insights for materials design.

CGCNN is an **open-source** software at https://github.com/txie-93/cgcnn.



References:

1. Xie, T. & Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Physical Review Letters* **120,** (2018).
2. Ahmad, Z., Xie, T., Maheshwari, C., Grossman, J. C. & Viswanathan, V. Machine Learning Enabled Computational Screening of Inorganic Solid Electrolytes for Suppression of Dendrite Formation in Lithium Metal Anodes. *ACS Central Science* **4,** 996–1006 (2018).
3. Xie, T. & Grossman, J. C. Hierarchical Visualization of Materials Space with Graph Convolutional Neural Networks. *The Journal of Chemical Physics,* in press.

MIT MATERIALS RESEARCH LABORATORY
G2e
MIT
TOYOTA RESEARCH INSTITUTE
NERSC
XSEDE Extreme Science and Engineering Discovery Environment