
Crystal graph convolutional neural networks for per-site property prediction

Jessica Karaguesian *

Center for Computational Science and Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139
jkarag@mit.edu

Jaclyn R. Lunger *

Department of Materials Science and Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139
lungerja@mit.edu

Yang Shao-Horn

Department of Mechanical Engineering
Research Laboratory of Electronics
Department of Materials Science and Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139
shaohorn@mit.edu

Rafael Gomez-Bombarelli

Department of Materials Science and Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139
rafagb@mit.edu

Abstract

Graph convolutional neural networks have been shown to accurately predict materials properties by featurizing local atomic environments. However, such models have not yet been employed to predict atom-level properties such as Bader charge, magnetic moment, or site-projected band centers. In this work, we present a per-site crystal graph convolutional neural network that predicts a wide array of such properties. This model captures the chemical environment around each atom and uses it to assign unique prediction labels to each site in a crystal. Using magnetic moments as a case study, we explore an example of underlying physics the per-site model is able to learn.

1 Introduction

Crystal graph convolutional neural networks (CGCNNs) have recently been shown to accurately predict a wide range of crystal properties [Xie and Grossman, 2018, Chen et al., 2019]. Unlike previous efforts based on manual construction of feature vectors, CGCNNs have the distinct advantage of automatically featurizing the neighborhood around each atom. It has thus been claimed that

*authors with equal contribution

CGCNNs are able to leverage the contributions of local environments to make global property predictions. For example, the interpretability of CGCNN’s for crystal property prediction has been shown by extracting the contribution of local chemical environments to the total energy [Xie and Grossman, 2018]. Despite these efforts, CGCNNs have not yet been applied to learning per-site properties and attempts to learn local information have been limited to predicting charge density on grid-points [Gong et al., 2019]. Per-site properties have simply been inferred from the atom feature scalars before the final pooling layer, but have never been rigorously verified.

In this work we develop a CGCNN for learning per-site properties, and show that such a model is able to capture detailed local information such as Bader charge, magnetic moment, atomic vibration frequency, and site-projected O2p-band center and d-band center. These per-site properties were chosen for their diversity of applications, ranging from electronics (Bader charge), spintronics (magnetic moments), and thermodynamic properties (atomic vibration) to catalysis (band centers). We show that this model is able to capture information about the local environment around each site, outperforming an MLP and per-element average baseline.

We additionally explore per-site magnetic moment prediction to explicitly show how the per-site model captures complex physics. Even when magnetism is not the key property of interest, studying magnetic materials with density functional theory requires spin-polarized calculations to find the ground state structure and energy. However, computational exploration of magnetic materials is challenging due to the complexity of their electronic and magnetic states. In particular, calculating the ground state of a magnetic material requires careful initialization because spin-polarized DFT calculations are prone to getting trapped in local minima. Previous efforts to resolve this challenge have included high-throughput workflows wherein many possible magnetic orderings are enumerated and calculated [Frey et al., 2020, Horton et al., 2019]. However, this workflow is computationally intensive and requires many spin-polarized DFT optimizations to correctly identify the magnetic ground state of a single material. To this end, we apply our per-site CGCNN for predicting per-site magnetic moment magnitudes given crystal structure. This model is trained on $\sim 32,000$ magnetic transition metal oxides retrieved from the Materials Project and in-house calculations, and is able to estimate per-site magnetic moment within an MAE of ~ 0.185 Bohr magnetons.

2 Model

The per-site model architecture presented here is an adaptation of the CGCNN developed by Xie and Grossman [2018]. The model takes three-dimensional crystal structures as input (Fig. 1a) and represents them as undirected multigraphs. The graph representation encodes atoms as nodes and connectivity as edges. Each site in the crystal is designated by a feature vector encoding its atomic properties and a neighbor list of sites within a cutoff radius. The neighbor connections, graph edges, are featurized with the distance between atoms. Distances between sites are computed in a periodicity-aware manner and, as such, the crystal graph can have multiple edges between a given pair of nodes. The crystal graph is passed through convolutional layers, which iteratively update the nodal feature vectors with information from their associated neighbors and edge connectivity, thus automatically learning crystal site representations informed by their unique chemical environments (Fig. 1b). The model then passes the convolved per-site representations through several fully connected hidden layers, and finally to an output layer to yield a scalar target property prediction at each site (Fig. 1c). The model is trained by minimizing the difference, defined by a loss function $L(\mathbf{p}, \tilde{\mathbf{p}})$, between target (DFT-calculated) and predicted per-site property values, \mathbf{p} and $\tilde{\mathbf{p}}$, respectively. The vector of per-site predictions is computed as $\tilde{\mathbf{p}} = f(\mathcal{C}; \mathbf{W})$, where f is the CGCNN network, \mathcal{C} is the input crystal, and \mathbf{W} is the set of weights that parametrizes the model layers [Xie and Grossman, 2018]. During training, model weights \mathbf{W} are optimized to minimize $L(\mathbf{p}, \tilde{\mathbf{p}} = f(\mathcal{C}; \mathbf{W}))$ through iterative updates calculated by backpropagation and stochastic gradient descent.

All results presented here are from models trained on 60% of the data, validated on 20%, and tested on 20%. Random train-validation-test splits were used in all cases. Model weights were chosen based on optimal validation set performance. All reported results reflect performance on the test set. Error ranges reflect statistics from initializing each model with three different random seeds. Hyperparameter tuning was performed using SigOpt [Clark and Hayes, 2019] on the magnetic moment dataset, and these same parameters were used to predict each of the other per-site properties. Because all datasets used here consist of DFT-relaxed structures it is not yet clear how the model will generalize to the prediction of per-site properties from unrelaxed structures.

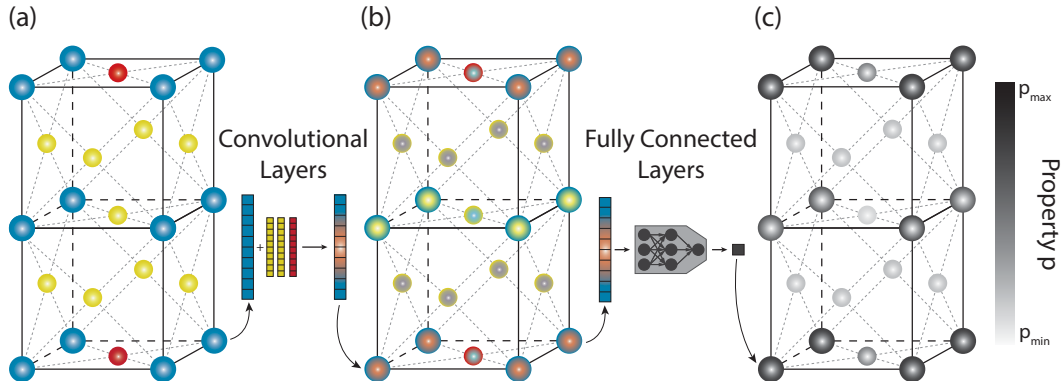


Figure 1: Overview of per-site CGCNN model architecture. (a) Three-dimensional crystal structure is converted to a graph and convolved with information from neighboring sites to give (b) nodal feature vectors encoding local chemical environment. The graphical representation is passed through several fully-connected layers and mapped to a property vector (c) whose entries correspond to each site in the crystal.

3 Experiments and Results

3.1 Datasets and Evaluation

Datasets for several per-site properties were compiled from the Materials Project (MP) [Jain et al., 2013] and in-house calculations of $\sim 9,000$ perovskites. Bader charge, magnetic moment, atomic vibration frequency (site-projected phonon band center), and per-site O2p-band centers were considered both because these are the per-site properties available through MP and because of their wide breadth of potential applications.

All $\sim 120,000$ structures in MP with available charges were used for the Bader charge dataset, where Bader charge in this case is the total number of electrons on a given atom. The Bader charge dataset is not limited to a given materials class, and rather spans all of MP. Similarly, all $\sim 10,000$ structures in MP with available phonon calculations were used to generate the atomic vibration frequency dataset, where atomic vibration frequency refers to the band-center of the site-projected phonon density of states.

The magnetic moment dataset is comprised of magnetic oxides, where we have removed those with unphysical magnetic moments (i.e. greater than 5 for d-band valence). Crystals that included any atoms with f-band valence were also left out of this study, given the scarcity of data and limited accuracy of pseudopotentials. The absolute value of magnetic moments is taken, such that the per-site model is relevant to ferromagnetic structures. We leave a model that handles collinear ordering to future work. In total there are $\sim 35,000$ crystals in this dataset with a wide range of stoichiometries.

The O2p-band center dataset is comprised of $\sim 10,000$ in-house DFT calculations on a wide variety of perovskites. The dataset is extended from existing databases [Castelli et al., 2012, Emery et al., 2016, Jacobs et al., 2018, Jain et al., 2013] by recalculating the structures using DFT calculations and adding novel alloyed systems.

The results from the per-site CGCNN model are summarized in Table 1, showing the MAE of the test set. Bader charge is predicted especially well with an MAE of $0.068 q_e$, likely due to the breadth and depth of the available data (all materials in MP have Bader charges calculated). All models were run on Nvidia GeForce RTX 2080 Ti GPUs.

3.2 Baselines

For each per-site property, we compared our CGCNN model against a multi-layer perceptron (MLP) baseline. This generous baseline is given all the same structural data as the CGCNN, but does not leverage graph structure nor convolutions. In the MLP model, each atom in the crystal is featured as its atomic encoding, along with the atomic encoding of and radial distance to its nearest neighbours,

Table 1: Comparisons between per-site CGCNN, MLP baseline, and per-element average in terms of mean average error (MAE) for several per-site properties. All reported results reflect performance on test datasets. The best results are shown in bold. The number of training examples (# Train crystals) for each property is also tabulated.

Property	# Train Crystals	Units	MAE		
			Per-site CGCNN	MLP	Per-element average
Bader charge	71,787	q_e	0.068 ± 0.001	0.147 ± 0.001	0.578
Magnetic moment	21,113	$ \mu_B $	0.185 ± 0.002	0.377 ± 0.002	0.553
Atomic vibration frequency	5,899	THz	0.817 ± 0.003	1.025 ± 0.011	1.571
O 2p-band center	6,024	eV	0.303 ± 0.004	0.579 ± 0.003	1.227
Metal d-band center	6,024	eV	0.581 ± 0.002	1.266 ± 0.017	1.281

up to the same cutoff used in the CGCNN. This model formulation is memory-inefficient as compared to the per-site CGCNN as it does not link atom featurizations together in a graph but rather repeats each in neighbour lists. Across the properties explored in this work, the MAE attained using the MLP baseline is on average \sim two-fold higher than that achieved by the per-site CGCNN, highlighting the ability of the graphical formulation to capture crystalline environments.

We also compare the performance of the CGCNN and MLP models to the per-element average of each property. This metric is obtained by averaging the per-site value for a given element across the entire dataset. Note that the values reported in Table 1 are the averages across the test set. That this average achieves reasonable results for each of the properties is to be expected, as each property is primarily a function of the atom-type and secondarily a function of the local environment around the site. For example we expect heavier atoms to have a lower atomic vibration frequency. The significantly decreased MAE achieved by the per-site MLP and CGCNN, as compared to this average, indicates the models’ ability to capture the effects of local chemical environment.

4 Prediction of per-site magnetic moments: a case study

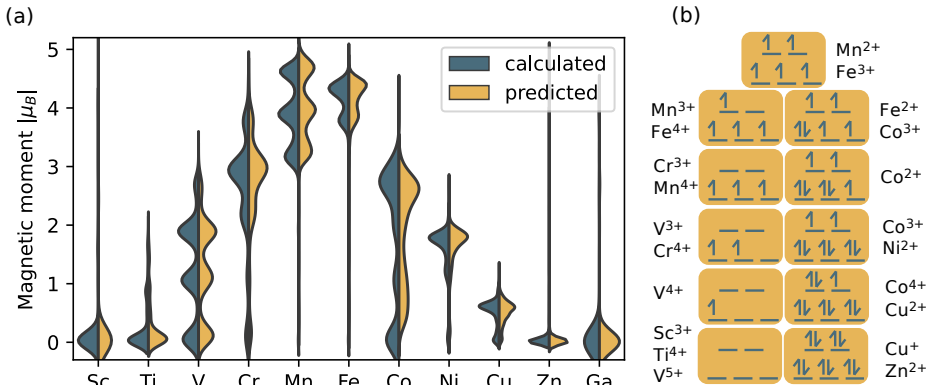


Figure 2: (a) Calculated and predicted magnetic moment distributions per-element across the 3d row. (b) Schematic of how d-band filling is related to magnetic moments.

Using magnetic moments as a case study, we analyze the physical insights gained from training a per-site model. We find that the per-site model is able to predict magnetic moments significantly more accurately than a per-element average baseline (Table 1) because of the wide distribution of magnetic moment for each element as a function of the environment (Fig. 2a). Considering the atoms across the 3d row, the magnetic moment distributions roughly follow electron counting high-spin rules. Namely, as the d-band is filled and electrons are able to exist in separate orbitals the magnetic

moment increases going from Sc to Mn (Fig. 2b). From Fe to Ga, there are more than 5 electrons so electrons must share orbitals and spins are cancelled out as the d-band is further filled. Complexity in the distributions arises from elements that are able to have multiple oxidation states depending on the surrounding atoms. The per-site model is able to estimate these physical effects and match the predicted magnetic moment distributions, achieving an MAE of $0.185 \mu_B$ (Table 1).

5 Conclusion

We present a graph convolutional neural network for predicting per-site materials properties, and show that it is able to accurately learn a diverse set of properties including Bader charge, magnetic moment, atomic vibration frequency, and site-projected electronic structure properties such as O2p-band and d-band center. This model is able to outperform an MLP baseline, showing that convolutions over a graphical representation of crystal structure has successfully incorporated information about the local environments around each atom. We present per-site magnetic moment prediction as a case study, where the model is able to capture the physics of d-band filling as well as the complexity of oxidation states in transition metals.

6 Broader Impact

The presented model has been shown to accurately predict a wide range of per-site crystal properties, and is therefore applicable to atomic-scale design of materials for a variety of applications including but not limited to spintronics, electronics, and catalysis. We are not aware of any potential negative societal impacts of this work.

References

- I. E. Castelli, T. Olsen, S. Dutta, D. D. Landis, S. Dahl, K. S. Thygesen, and K. W. Jacobsen. Computational screening of perovskite metal oxides for optimal solar light capture. *Energy & Environmental Science*, 5:5814–5819, 2012.
- C. Chen, W. Ye, Y. Zuo, C. Zheng, and S. P. Ong. Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, 31:3564–3572, 2019.
- Scott Clark and Patrick Hayes. SigOpt Web page. <https://sigopt.com>, 2019. URL <https://sigopt.com>.
- A. A. Emery, J. E. Saal, S. Kirklin, V. I. Hegde, and C. Wolverton. High-throughput computational screening of perovskites for thermochemical water splitting applications. *Chemistry of Materials*, 28(16):5621–5634, 2016.
- N. C. Frey, M. K. Horton, J. M. Munro, S. M. Griffin, K. A. Persson, and V. B. Shenoy. High-throughput search for magnetic and topological order in transition metal oxides. *Science Advances*, 6(50):1–10, 2020.
- S. Gong, T. Xie, T. Zhu, E. R. Fadel, Y. Li, and J. C. Grossman. Predicting charge density distribution of materials using a local-environment-based graph convolutional network. *Physical Review B*, 100:184103, 2019.
- M. K. Horton, J. H. Montoya, M. Liu, and K. A. Persson. High-throughput prediction of ground-state collinear magnetic order of inorganic materials using density functional theory. *npj Computational Materials*, 5(1):1–11, 2019.
- R. Jacobs, T. Mayeshiba, J. Booske, and D. Morgan. Material discovery and design principles for stable, high activity perovskite cathodes for solid oxide fuel cells. *Advanced Energy Materials*, 8(11):1702708, 2018.
- A. Jain, S. P. Ong, G. Hauttier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1:011002, 2013.
- T. Xie and J. C. Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical Review Letters*, 120(14):145301, 2018.

Checklist

1. For all authors...
 - (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
 - (b) Did you describe the limitations of your work? [Yes] In our datasets section we describe how we are limited to a given list of per-site properties because of data availability issues. Additionally, our datasets are all relaxed structures and it is unclear how the trained per-site model will perform on unrelaxed structures before a DFT calculation has been performed.
 - (c) Did you discuss any potential negative societal impacts of your work? [Yes] We do not believe this work has potential negative societal impacts. On the contrary, our work has applicability to materials design.
 - (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
 - (a) Did you state the full set of assumptions of all theoretical results? [N/A] No theoretical results
 - (b) Did you include complete proofs of all theoretical results? [N/A] No theoretical results
3. If you ran experiments...
 - (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [No] We have not yet included the code/data/instructions, but we plan to release these.
 - (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] We explain the train/test datasplit is random and that hyperparameter tuning was performed on the magnetic moments dataset. The actual hyperparameters will be made available with the code/data release.
 - (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] We ran each model with three different random seeds initializations and reported the resulting error.
 - (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
 - (a) If your work uses existing assets, did you cite the creators? [Yes]
 - (b) Did you mention the license of the assets? [Yes] MP and CGCNN are open source assets
 - (c) Did you include any new assets either in the supplemental material or as a URL? [No] We plan to release our code as a fork of the CGCNN GitHub repository.
 - (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
 - (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
 - (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
 - (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
 - (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]