Project Extension Proposal: Transfer Learning for the Equivariant Prediciton of Tensor Data for Crystal Structures

January 24, 2024

OVERVIEW: Shortly after the advent of the Crystal Graph Convolution Neural Network (CGCNN), a multi-target approach was taken with a similar model [SBY+18]. This multi-target approach simultaneously trained a model on various physical quantities of each input in a supervised manner by simply summing up the loss of each individual target and applying gradient descent to this total loss. Such an approach showed substantial improvement in results, as well as providing a means to generate 'more descriptive' crystal-level features after convolution and pooling. Transfer learning was also applied to CGCNN in other works [LA21], as a way to leverage larger datasets in the prediction of targets with sparser training sets.

Recent works have utilized $SE(3)$ equivariant networks to predict tensorial targets related to physical systems by means of spherical harmonic decompositoins of the target tensors. Such an approach has been used in the prediction of elastic [WHM+23], dielectric [HXG22], and piezoelectric [HXG22] tensors for crystals, as well as NMR spectral shift tensors [VWP23] and other molecular targets [NL22].

While equivariant networks are relatively well-suited for small datasets, their performance still benefits from more expressive and larger datasets. Of the three tensorial targets for crystal systems mentioned above, the elastic tensor has the largest available dataset (10,000 accessible via Materials Project), while data for dielectric (7,700) and especially piezoelectric tensors (X) is more sparse. Some of this discrepancy results from the number of materials that may be modeled with certain relationships. That is, dielectric tensors only describe a subset of materials (insulators) which may have a well-defined elastic tensor (applicable to all homogenous materials), and piezoelectric response is only possible in an even more exclusive set of materials (insulators with no spatial-inversion symmetry). However, the small amount of available data for all three sets is perhaps also a testament to their intensive calculation methods.

We propose a transfer learning scheme, leveraging all three sets of data simultaneously to improve model accuracy in spite of relatively small data sets. Such a model will have two distinct parts: a convolutional representation-updating module, which will be fed into a task-specific MLP. Two approaches may suit us here: either we may (1) train the convolutional layers exclusively on the largest dataset (elastic tensors) and then swap out only the MLPs for different tasks; or (2), we may train the convolutional layers on the largest dataset and then

train, or fine-tune, them also on the other tasks. Note that materials may have only one or two of the different tensors available, making a true simultaneous multi-target model unfeasible, as this would restrict us (at most) to the smallest dataset.

Note further that SUCH AN APPROACH HAS BEEN EXPLICITY MENTIONED in the elastic tensor (MatTen [WHM$^+$23]) paper. This suggests we need to work according to a very short timeline for this project to remain viable!

PROPOSED TIMELINE: 2-3 weeks

WEEK 1: Data gathering/model splitting

- Gather data for dielectric tensors from materials project
- Gather data for piezoelectric tensors from materials project
- Construct and code harmonic decompositon for piezoelectric tensor
- Construct and code harmonic decompositon for dielectric tensor
- Redesign model as two parts: convolution and a task-specific MLP

WEEK 2: Transfer Learning from Larger Datasets

- Train convolutional layers and MLP on elasticity dataset (largest)
- Swap out MLP and train (MLP only) on dielectric dataset
- Swap out MLP and train (MLP only) on piezo dataset

WEEK 2/3: Fine Tuning on Smaller Datasets

- Test fine-tuning (train convolution + MLP) on dielectric dataset
- Test fine-tuning (train convolution + MLP) on piezo dataset

References

[HXG22] Hongyu Yu Hongjun Xiang, Yang Zhong and Xinggao Gong. A general tensor prediction framework based on graph neural networks. 2022.

> This paper proposes a similar SO(3) decomposition of tensors to predict piezoelectric, dielectric, and other tensors.

[LA21] Joohwi Lee and Ryoji Asahi. Transfer learning for materials informatics using crystal graph convolutional neural network. *Computational Materials Science*, 190:110314, 2021.

> Here, transfer learning is utilized for a CGCNN-like model where pretraining is performed on the larger datasets of targets (like formation energy per atom) and then the model is adapted to sparser target sets (like bulk moduli)

[NL22] Vu Ha Anh Nguyen and Alessandro Lunghi. Predicting tensorial molecular properties with equivariant machine learning models. *Phys. Rev. B*, 105:165131, Apr 2022.

> Here, equivariant tensor prediction is applied to various molecular targets.

[SBY+18] Soumya Sanyal, Janakiraman Balachandran, Naganand Yadati, Abhishek Kumar, Padmini Rajagopalan, Suchismita Sanyal, and Partha Talukdar. Mt-cgcnn: Integrating crystal graph convolutional neural network with multitask learning for material property prediction. 2018.

> This was the first attempt to train a multi-target CGCNN for scalar target properties.

[VWP23] Maxwell C Venetos, Mingjian Wen, and Kristin A Persson. Machine learning full nmr chemical shift tensors of silicon oxides with equivariant graph neural networks. *The Journal of Physical Chemistry A*, 127(10):2388–2398, 2023.

> This paper applies symmetry and harmonic decompositions to the NMR spectra shift tensor to predict them equivariantly.

[WHM+23] Mingjian Wen, Matthew K. Horton, Jason M. Munro, Patrick Huck, and Kristin A. Persson. A universal equivariant graph neural network for the elasticity tensors of any crystal system. 2023.

> This paper by the materials project team is our original punchline, giving Backus' spherical harmonic decomposition of the elastic tensor and reading it off of the irreps of the equivariant output.