

Crystal graph convolutional neural networks

Andreas Leitherer

Fritz Haber Institute of the Max Planck Society
Theory department

Coffee talk

July 22, 2019



FRITZ-HABER-INSTITUT
MAX-PLANCK-GESELLSCHAFT

**Crystal Graph Convolutional Neural Networks for an Accurate
and Interpretable Prediction of Material Properties**

Tian Xie and Jeffrey C. Grossman

*Department of Materials Science and Engineering, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA*

 (Received 18 October 2017; revised manuscript received 15 December 2017; published 6 April 2018)

THE JOURNAL OF CHEMICAL PHYSICS **149**, 174111 (2018)

**Hierarchical visualization of materials space with graph
convolutional neural networks**

Tian Xie and Jeffrey C. Grossman

*Department of Materials Science and Engineering, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA*



**Developing an improved Crystal Graph Convolutional Neural Network
framework for accelerated materials discovery**

Cheol Woo Park, Chris Wolverton*

Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois

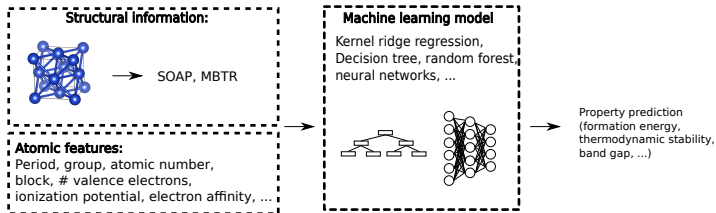
60208, USA

E-mail: c-wolverton@northwestern.edu

arXiv:1906.05267

Goal

- Given chemical (and structural) information, predict material properties (e.g., formation energy) using machine learning model

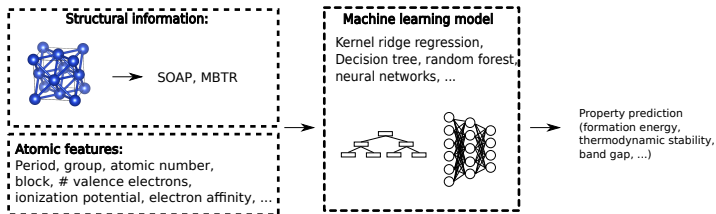


Smooth-overlap-of-atomic-positions (SOAP): A. P. Bartók et al., Phys. Rev. B 87, 184115 (2013)

Many-body tensor representation (MBTR): H. Huo and M. Rupp, arXiv:1704.06439 (2017)

Goal

- Given chemical (and structural) information, predict material properties (e.g., formation energy) using machine learning model
- Close to ab initio accuracy, orders of magnitude faster → accelerate design of new materials

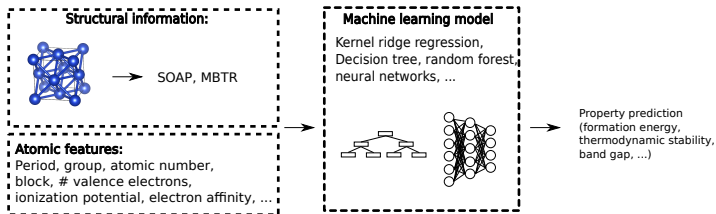


Smooth-overlap-of-atomic-positions (SOAP): A. P. Bartók et al., Phys. Rev. B 87, 184115 (2013)

Many-body tensor representation (MBTR): H. Huo and M. Rupp, arXiv:1704.06439 (2017)

Goal

- Given chemical (and structural) information, predict material properties (e.g., formation energy) using machine learning model
- Close to ab initio accuracy, orders of magnitude faster → accelerate design of new materials
- Large material databases available (Materials Project, AFLOW, OQMD, ICSD, ...)

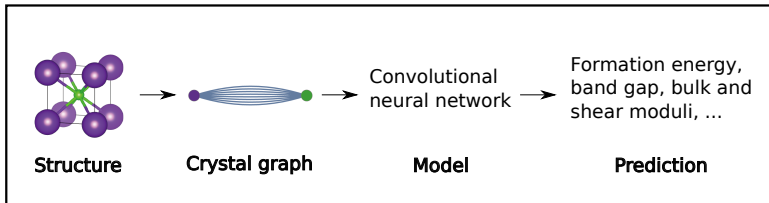


Smooth-overlap-of-atomic-positions (SOAP): A. P. Bartók et al., Phys. Rev. B 87, 184115 (2013)

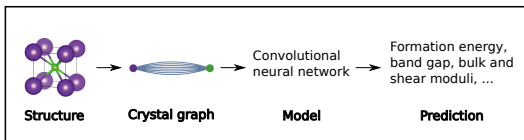
Many-body tensor representation (MBTR): H. Huo and M. Rupp, arXiv:1704.06439 (2017)

This talk: Crystal graph convolutional neural network (CGCNN)

- Connectivity-based approach: input structure (xyz + chemical composition) represented by crystal graph
- Prediction obtained via convolutional neural network

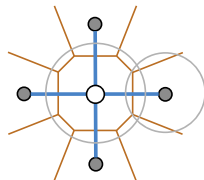


Crystal graph CNN (CGCNN) - Crystal graph representation



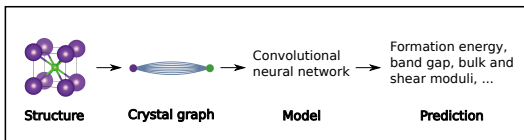
Crystal graph representation

- For each atom, search neighbors within 6.0 \AA ; bond \Leftrightarrow shared Voronoi face and distance $<$ sum of covalent radii



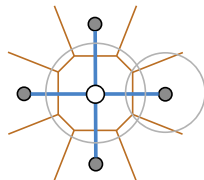
O. Isayev et al. Nat. Commun. 8, 15679 (2017)

Crystal graph CNN (CGCNN) - Crystal graph representation



Crystal graph representation

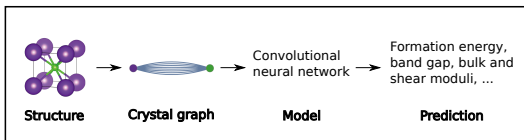
- For each atom, search neighbors within 6.0 \AA ; bond \Leftrightarrow shared Voronoi face and distance $<$ sum of covalent radii
- In practice: take 12 nearest neighbors



O. Isayev et al. Nat. Commun. 8, 15679 (2017)

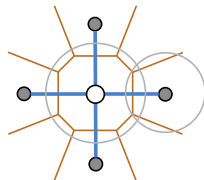
T. Xie and J. C. Grossman, Phys. Rev. Lett. 120,145301 (2018)

Crystal graph CNN (CGCNN) - Crystal graph representation



Crystal graph representation

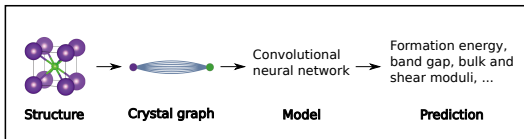
- For each atom, search neighbors within 6.0 \AA ; bond \Leftrightarrow shared Voronoi face and distance $<$ sum of covalent radii
- In practice: take 12 nearest neighbors
- Assign feature vectors \mathbf{v}_i to all vertices (atoms)



O. Isayev et al. Nat. Commun. 8, 15679 (2017)

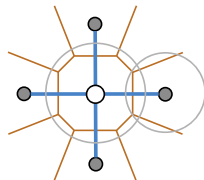
T. Xie and J. C. Grossman, Phys. Rev. Lett. 120,145301 (2018)

Crystal graph CNN (CGCNN) - Crystal graph representation



Crystal graph representation

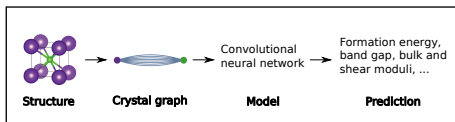
- For each atom, search neighbors within 6.0 \AA ; bond \Leftrightarrow shared Voronoi face and distance $<$ sum of covalent radii
- In practice: take 12 nearest neighbors
- Assign feature vectors \mathbf{v}_i to all vertices (atoms)
- Assign $\mathbf{u}_{ij,k}$ to all k edges (bonds) between atom i and j



O. Isayev et al. Nat. Commun. 8, 15679 (2017)

T. Xie and J. C. Grossman, Phys. Rev. Lett. 120,145301 (2018)

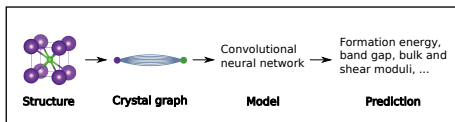
Crystal graph CNN (CGCNN) - Choice of features



T. Xie and J. C. Grossman, Phys. Rev. Lett. 120,145301 (2018)

- For atoms (\mathbf{v}_i): Atomic features (group, period, electronegativity,...)
- For bonds ($\mathbf{u}_{ij,k}$): Distance $d_{ij,k}$

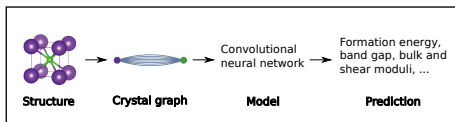
Crystal graph CNN (CGCNN) - Choice of features



T. Xie and J. C. Grossman, Phys. Rev. Lett. 120,145301 (2018)

- For atoms (\mathbf{v}_i): Atomic features (group, period, electronegativity,...)
- For bonds ($\mathbf{u}_{ij,k}$): Distance $d_{ij,k}$
- Discrete features: one-hot encoding

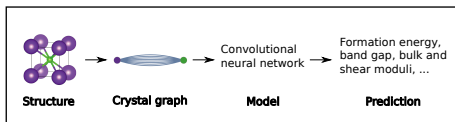
Crystal graph CNN (CGCNN) - Choice of features



T. Xie and J. C. Grossman, Phys. Rev. Lett. 120,145301 (2018)

- For atoms (\mathbf{v}_i): Atomic features (group, period, electronegativity,...)
- For bonds ($\mathbf{u}_{ij,k}$): Distance $d_{ij,k}$
- Discrete features: one-hot encoding
- Continuous features: discretization + one-hot encoding

Crystal graph CNN (CGCNN) - Choice of features



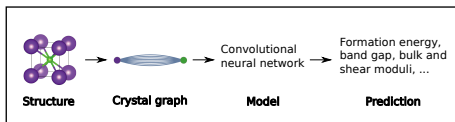
T. Xie and J. C. Grossman, Phys. Rev. Lett. 120,145301 (2018)

- For atoms (\mathbf{v}_i): Atomic features (group, period, electronegativity,...)
- For bonds ($\mathbf{u}_{ij,k}$): Distance $d_{ij,k}$
- Discrete features: one-hot encoding
- Continuous features: discretization + one-hot encoding

T. Xie and J. C. Grossman, J. Chem. Phys. 149, 174111 (2018)

- Random initialization of \mathbf{v}_i (64 dim.), optimized during training

Crystal graph CNN (CGCNN) - Choice of features



T. Xie and J. C. Grossman, Phys. Rev. Lett. 120,145301 (2018)

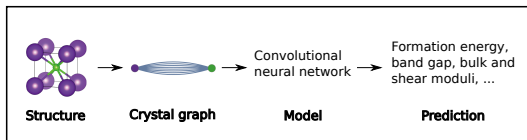
- For atoms (\mathbf{v}_i): Atomic features (group, period, electronegativity,...)
- For bonds ($\mathbf{u}_{ij,k}$): Distance $d_{ij,k}$
- Discrete features: one-hot encoding
- Continuous features: discretization + one-hot encoding

T. Xie and J. C. Grossman, J. Chem. Phys. 149, 174111 (2018)

- Random initialization of \mathbf{v}_i (64 dim.), optimized during training
- Bond feature vector: t -th component

$$[\mathbf{u}_{ij,k}]_t = \exp(-(d_{ij,k} - \mu_t)/\sigma^2), \quad \mu_t = t \cdot 0.2 \text{ \AA}, t = 0, \dots, 40, \sigma = 0.2 \text{ \AA}$$

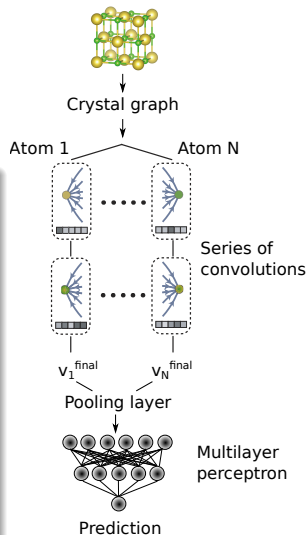
Crystal graph CNN - Convolutional neural network



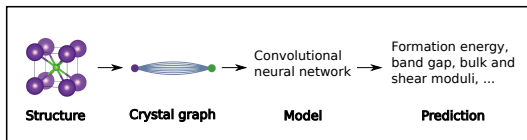
Convolutional on a graph

- Update feature vectors via convolution function - best performing choice:

$$\mathbf{v}_i \rightarrow \text{Conv}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}) = \mathbf{v}_i + \sum_{j,k} \sigma(\mathbf{z}_{ij,k} \mathbf{W}) \odot g(\mathbf{z}_{ij,k} \mathbf{W}')$$



Crystal graph CNN - Convolutional neural network

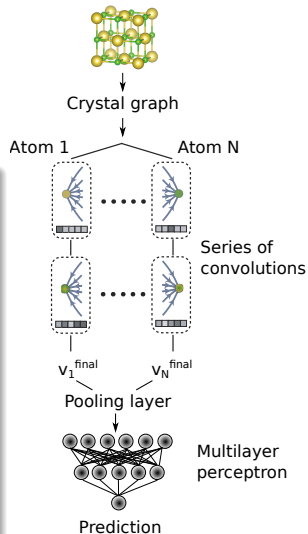


Convolutional on a graph

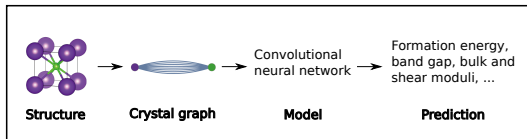
- Update feature vectors via convolution function - best performing choice:

$$\mathbf{v}_i \rightarrow \text{Conv}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}) = \mathbf{v}_i + \sum_{j,k} \sigma(\mathbf{z}_{ij,k} \mathbf{W}) \odot g(\mathbf{z}_{ij,k} \mathbf{W}')$$

- $\mathbf{z}_{ij,k}$: concatenation of $\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}$



Crystal graph CNN - Convolutional neural network

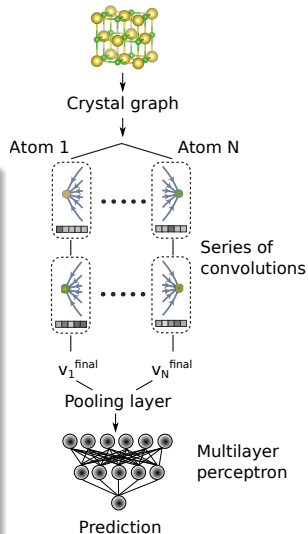


Convolutional on a graph

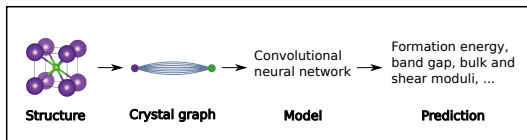
- Update feature vectors via convolution function - best performing choice:

$$\mathbf{v}_i \rightarrow \text{Conv}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}) = \mathbf{v}_i + \sum_{j,k} \sigma(\mathbf{z}_{ij,k} \mathbf{W}) \odot g(\mathbf{z}_{ij,k} \mathbf{W}')$$

- $\mathbf{z}_{ij,k}$: concatenation of $\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}$
- $g(\cdot)$: non-linear activation function



Crystal graph CNN - Convolutional neural network

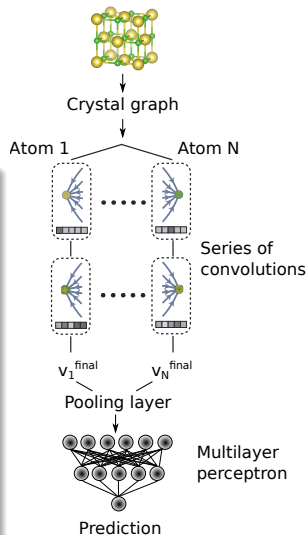


Convolutional on a graph

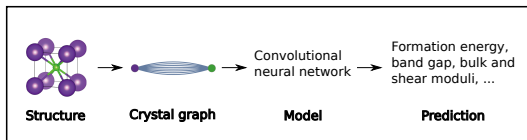
- Update feature vectors via convolution function - best performing choice:

$$\mathbf{v}_i \rightarrow \text{Conv}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}) = \mathbf{v}_i + \sum_{j,k} \sigma(\mathbf{z}_{ij,k} \mathbf{W}) \odot g(\mathbf{z}_{ij,k} \mathbf{W}')$$

- $\mathbf{z}_{ij,k}$: concatenation of $\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}$
- $g(\cdot)$: non-linear activation function
- $\sigma(\cdot) \in [0, 1]$: learned weight to discriminate interaction strength between neighbors



Crystal graph CNN - Convolutional neural network

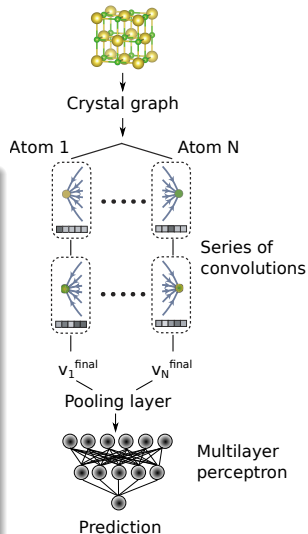


Convolutional on a graph

- Update feature vectors via convolution function - best performing choice:

$$\mathbf{v}_i \rightarrow \text{Conv}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}) = \mathbf{v}_i + \sum_{j,k} \sigma(\mathbf{z}_{ij,k} \mathbf{W}) \odot g(\mathbf{z}_{ij,k} \mathbf{W}')$$

- $\mathbf{z}_{ij,k}$: concatenation of $\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}$
- $g(\cdot)$: non-linear activation function
- $\sigma(\cdot) \in [0, 1]$: learned weight to discriminate interaction strength between neighbors
- Why convolutional? → Shared filter parameters over whole graph



Results - Formation energy

Method	MAE (eV/atom)	Data source	Training	Validation	Test
CGCNN (atomic features)	0.039	MP	28046	9348	9348
CGCNN (random init.)	0.042	MP	28046	9348	9348

Data source

Materials project (MP): extract 46744 materials covering majority of ICSD structures → good representation of known inorganic materials

Table reproduced from T. Xie and J. C. Grossman, J. Chem. Phys. 149, 174111 (2018)

Results - Formation energy

Method	MAE (eV/atom)	Data source	Training	Validation	Test
CGCNN (atomic features)	0.039	MP	28046	9348	9348
CGCNN (random init.)	0.042	MP	28046	9348	9348
SchNet	0.035	MP	60000	4500	5140

Data source

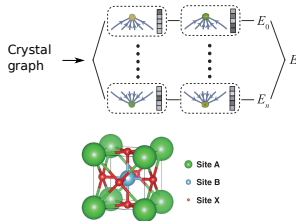
Materials project (MP): extract 46744 materials covering majority of ICSD structures → good representation of known inorganic materials

Table reproduced from T. Xie and J. C. Grossman, *J. Chem. Phys.* 149, 174111 (2018)

SchNet: K. T. Schütt et al., *J. Chem. Phys.* 148, 241722 (2018)

Interpretability I

- Predict formation energies for each Perovskite site (A,B), for all elements; identify stable elements and empirical rules



Database ($\sim 19k$ materials): Available at Computational Materials Repository (“Perovskite water-splitting”)

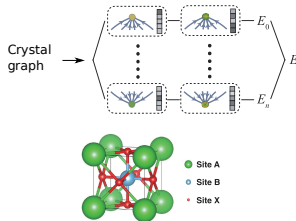
Results - Interpretability

Interpretability I

- Predict formation energies for each Perovskite site (A,B), for all elements; identify stable elements and empirical rules

Interpretability II

- Learn element representations from random initialization; analyze learned insight:



Database (~ 19k materials): Available at Computational Materials Repository (“Perovskite water-splitting”)

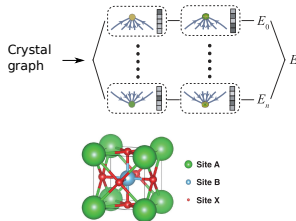
Results - Interpretability

Interpretability I

- Predict formation energies for each Perovskite site (A,B), for all elements; identify stable elements and empirical rules

Interpretability II

- Learn element representations from random initialization; analyze learned insight:
 - Dim. reduction (PCA): Grouping of elements based on stability in Perovskites (e.g., clustering of large elements)



Database (~ 19k materials): Available at Computational Materials Repository (“Perovskite water-splitting”)

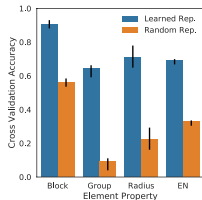
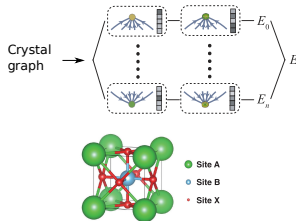
Results - Interpretability

Interpretability I

- Predict formation energies for each Perovskite site (A,B), for all elements; identify stable elements and empirical rules

Interpretability II

- Learn element representations from random initialization; analyze learned insight:
 - Dim. reduction (PCA): Grouping of elements based on stability in Perovskites (e.g., clustering of large elements)
- From learned element representations, predict atomic features of each element
 - Rule out coincidences by comparing to random representations



Linear logistic regression,
3-fold cross validation

Database (~ 19k materials): Available at Computational Materials Repository ("Perovskite water-splitting")

Three recently proposed improvements of CGCNN (iCGCNN)

1. Instead of 12 nearest neighbors, only consider Voronoi neighbors and add solid angle, area, and volume of Voronoi cell as bond features

Three recently proposed improvements of CGCNN (iCGCNN)

1. Instead of 12 nearest neighbors, only consider Voronoi neighbors and add solid angle, area, and volume of Voronoi cell as bond features
2. Add many-body correlations (3-body):

→ Original convolution $\text{Conv}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}) = \mathbf{v}_i + \sum_{j,k} \sigma(\mathbf{z}_{ij,k} \mathbf{W}) \odot g(\mathbf{z}_{ij,k} \mathbf{W}')$

Three recently proposed improvements of CGCNN (iCGCNN)

1. Instead of 12 nearest neighbors, only consider Voronoi neighbors and add solid angle, area, and volume of Voronoi cell as bond features
2. Add many-body correlations (3-body):
 - Original convolution $\text{Conv}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}) = \mathbf{v}_i + \sum_{j,k} \sigma(\mathbf{z}_{ij,k} \mathbf{W}) \odot g(\mathbf{z}_{ij,k} \mathbf{W}')$
 - Add term $\sum_{jl, kk'} \sigma(\mathbf{z}'_{ijl, kk'} \widetilde{\mathbf{W}}) \odot g(\mathbf{z}'_{ijl, kk'} \widetilde{\mathbf{W}}')$
 - $\mathbf{z}'_{ijl, kk'}$ being concatenation of $\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_l, \mathbf{u}_{ij,k}, \mathbf{u}_{il, k'}$

Three recently proposed improvements of CGCNN (iCGCNN)

1. Instead of 12 nearest neighbors, only consider Voronoi neighbors and add solid angle, area, and volume of Voronoi cell as bond features
2. Add many-body correlations (3-body):
 - Original convolution $\text{Conv}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}) = \mathbf{v}_i + \sum_{j,k} \sigma(\mathbf{z}_{ij,k} \mathbf{W}) \odot g(\mathbf{z}_{ij,k} \mathbf{W}')$
 - Add term $\sum_{jl, kk'} \sigma(\mathbf{z}'_{ijl, kk'} \widetilde{\mathbf{W}}) \odot g(\mathbf{z}'_{ijl, kk'} \widetilde{\mathbf{W}}')$
 - $\mathbf{z}'_{ijl, kk'}$ being concatenation of $\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_l, \mathbf{u}_{ij,k}, \mathbf{u}_{il, k'}$
3. Optimize not only atom feature vectors but also bond feature vectors

Three recently proposed improvements of CGCNN (iCGCNN)

1. Instead of 12 nearest neighbors, only consider Voronoi neighbors and add solid angle, area, and volume of Voronoi cell as bond features
2. Add many-body correlations (3-body):
 - Original convolution $\text{Conv}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}) = \mathbf{v}_i + \sum_{j,k} \sigma(\mathbf{z}_{ij,k} \mathbf{W}) \odot g(\mathbf{z}_{ij,k} \mathbf{W}')$
 - Add term $\sum_{jl, kk'} \sigma(\mathbf{z}'_{ijl, kk'} \widetilde{\mathbf{W}}) \odot g(\mathbf{z}'_{ijl, kk'} \widetilde{\mathbf{W}}')$
 - $\mathbf{z}'_{ijl, kk'}$ being concatenation of $\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_l, \mathbf{u}_{ij,k}, \mathbf{u}_{il, k'}$
3. Optimize not only atom feature vectors but also bond feature vectors

Formation energy prediction

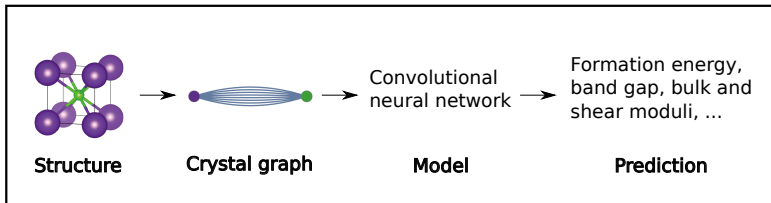
- Dataset (OQMD): train on 180k, validate on 20k, test on 230k

Three recently proposed improvements of CGCNN (iCGCNN)

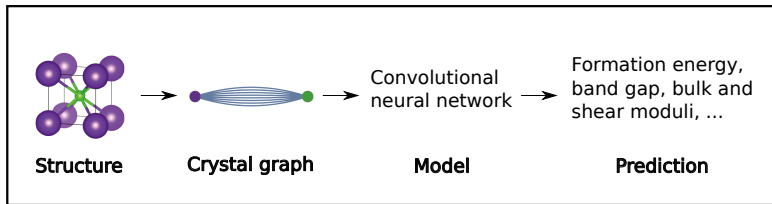
1. Instead of 12 nearest neighbors, only consider Voronoi neighbors and add solid angle, area, and volume of Voronoi cell as bond features
2. Add many-body correlations (3-body):
 - Original convolution $\text{Conv}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_{ij,k}) = \mathbf{v}_i + \sum_{j,k} \sigma(\mathbf{z}_{ij,k} \mathbf{W}) \odot g(\mathbf{z}_{ij,k} \mathbf{W}')$
 - Add term $\sum_{jl, kk'} \sigma(\mathbf{z}'_{ijl, kk'} \widetilde{\mathbf{W}}) \odot g(\mathbf{z}'_{ijl, kk'} \widetilde{\mathbf{W}}')$
 - $\mathbf{z}'_{ijl, kk'}$ being concatenation of $\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_l, \mathbf{u}_{ij,k}, \mathbf{u}_{il, k'}$
3. Optimize not only atom feature vectors but also bond feature vectors

Formation energy prediction

- Dataset (OQMD): train on 180k, validate on 20k, test on 230k
- CGCNN: $R^2 = 0.991$, MAE = 41.3 meV/at., RMSE = 78.6 meV/at.
- iCGCN: $R^2 = 0.993$, MAE = 30.5 meV/at., RMSE = 66.7 meV/at.



- Code availability: <https://github.com/txie-93/cgcnn>
- Blog: <http://txie.me/>



- Code availability: <https://github.com/txie-93/cgcnn>
- Blog: <http://txie.me/>

Thank you for your attention!

Appendix - Park and Wolverton, Voronoi neighbors

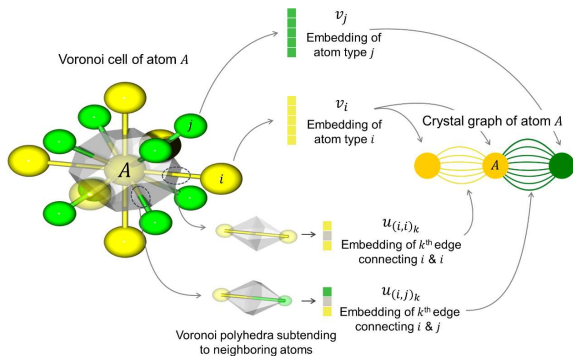


Figure from C. W. Park and C. Wolverton, arXiv:1906.05267 (2019)