ForceNet: A Graph Neural Network for Large-Scale Quantum Calculations

Carnegie Mellon University

Weihua Hu, Muhammed Shuaibi, Abhishek Das, Siddharth Goyal, Anuroop Sriram, Jure Leskovec, Devi Parikh, Larry Zitnick

FB

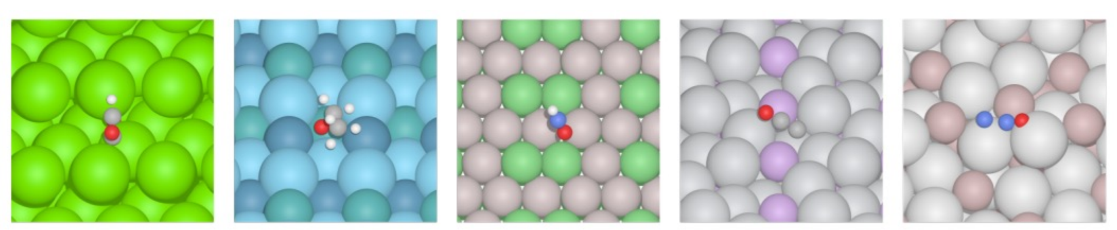
Overview

We consider a new approach for GNNs to predict atomic force prediction. Unlike existing GNN models, our approach is **both** scalable and expressive. Physical rules are encouraged through **data augmentation.**

	Physical rules	Expressiveness	Scalability
SchNet	✓	X	✓
DimeNet	✓	✓	X
Our approach	△ Use data augmentation	✓	✓

DFT-based Atomic Force Calculations

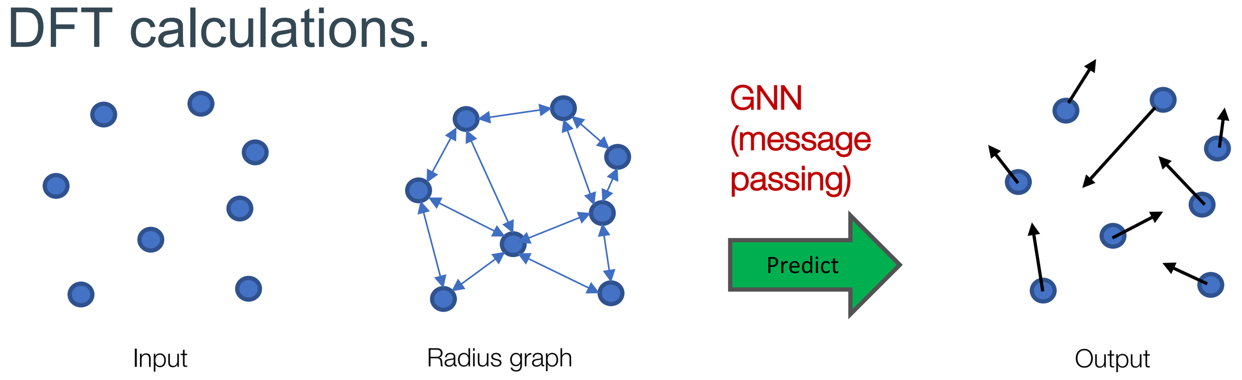
Given a molecular system (a set of atoms in 3D space), we want to calculate **per-atom forces**---fundamental quantities for molecular simulation.



Density Functional Theory (DFT) is an accurate quantum mechanical method to calculate atomic forces.
DFT is widely used in drug and material discovery.
However, **DFT is computationally expensive.**

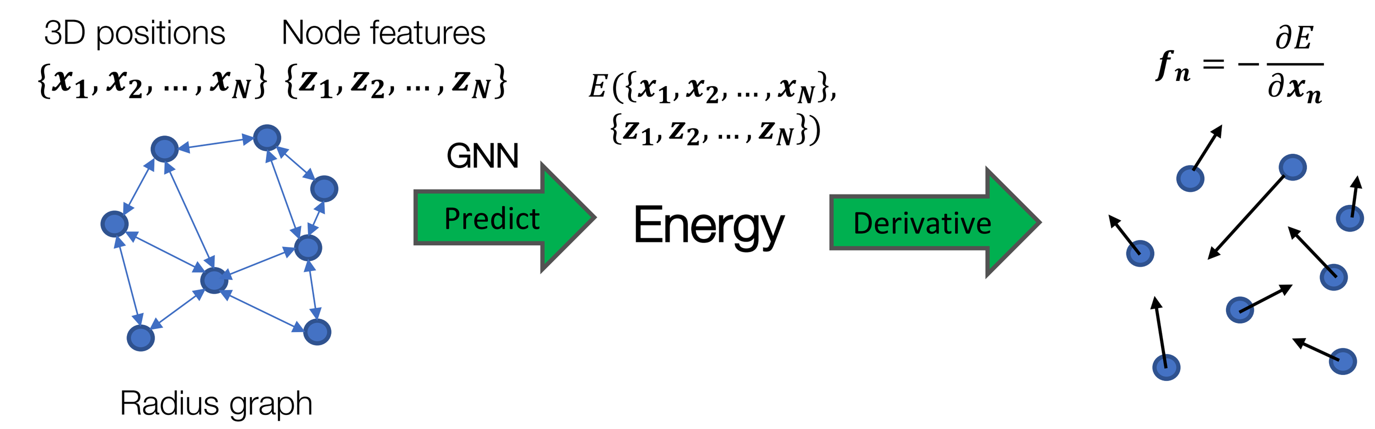
ML-Based Force Prediction

Use ML models (especially **Graph Neural Networks**) to approximate DFT-calculated atomic forces.
- Inference in ML models is **very fast.**
- With ever-increasing scientific compute, a huge number of DFT training data has been generated.
E.g., The recent OC20 dataset [Chanussot et al. 2020] includes 130 million DFT calculations.



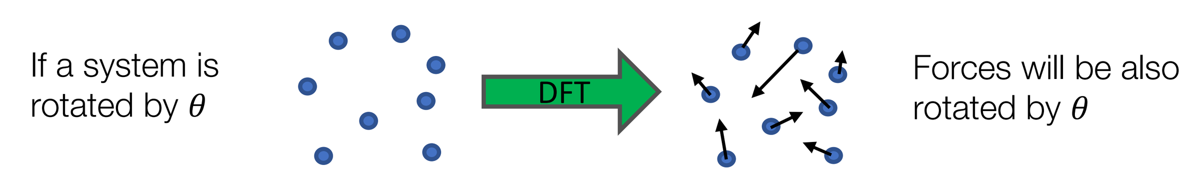
SoTA Approach: Energy-centric

- First, predict the energy of the system in a **rotation-invariant** way.
- Then, predict per-atom forces by taking derivatives



Advantage of Energy-centric approach

Predicted atomic forces obey the basic physical rules
- Rotation-covariance



- **Energy conservation:** Forces are derivatives of the underlying energy

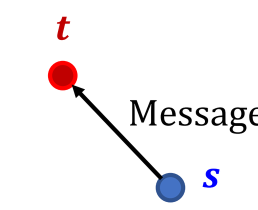
Issues of Energy-centric approach

Model architecture needs to be carefully constrained to ensure the rotation invariance. SoTA GNN models suffer from the trade-off between model expressiveness and computational efficiency.

- Schnet

Only uses the **distance** to model its message.

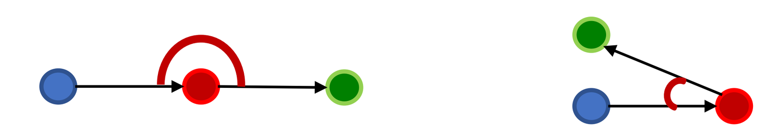
$$h_t^{(k+1)} = \sum_{t \in N(s)} \text{Message}(h_s^{(k)}, \|x_s - x_t\|)$$



But… SchNet does not explicitly capture angular information.
Limited expressiveness leads to suboptimal performance.

- DimeNet

Explicitly incorporates angular information in its message.
- Messages are defined over triplets of nodes



- DimeNet is expressive and provides better performance than SchNet.

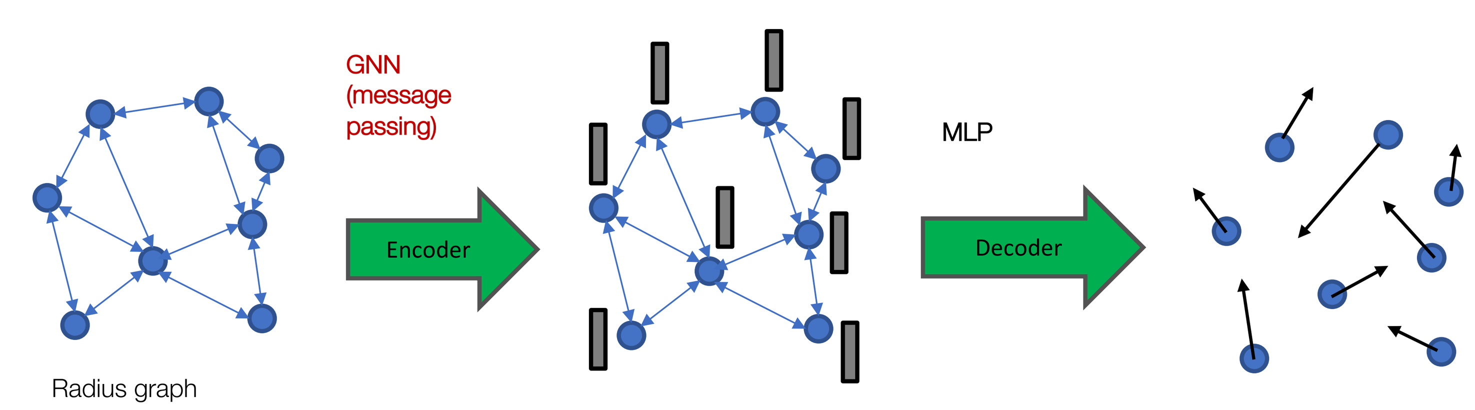
But… DimeNet is computationally very expensive.
E.g., #Messages is ~40x more than SchNet in OC20. 1600 GPU days to train.

Our Approach

Not explicitly enforcing physics rules (e.g., rotation invariance/covariance) in model architecture.
We can flexibly design scalable and expressive model.
Physical rules are encouraged through **data augmentation**

Our Force-centric Model: ForceNet

To instantiate the approach, we design a GNN model, called ForceNet.



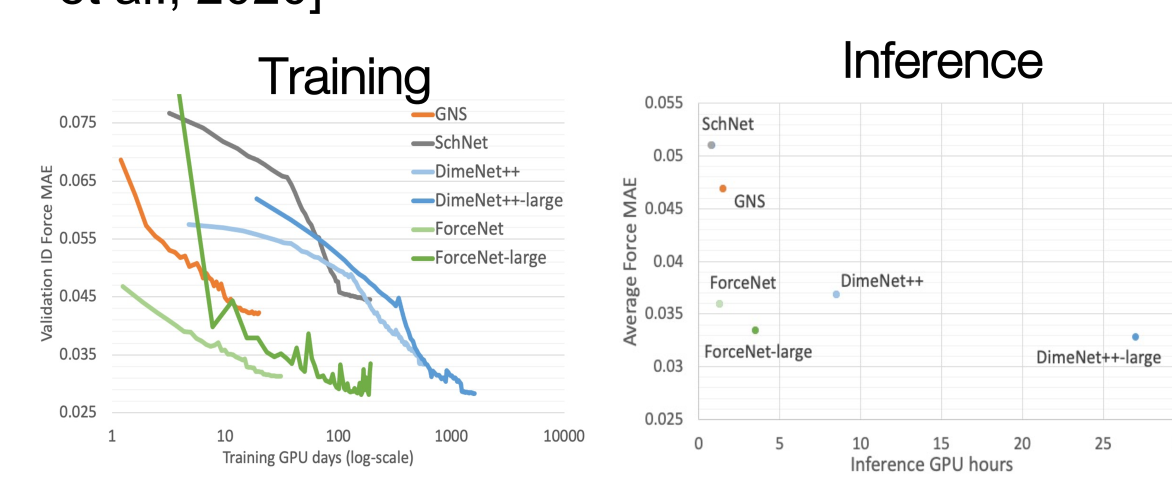
Three key components in message modeling
- Expressive message architecture
- Appropriate basis and non-linear activation functions
- Scaling up model sizes.

Rotation Augmentation

The prediction of ForceNet is not necessarily rotation-covariant.
We use **rotation data augmentation** to encourage rotation covariance.

Experiments

We evaluate ForceNet on the OC20 dataset [Chanussot et al. 2020].
Evaluation metric : Force MAE.
Baseline models: SchNet, DimeNet++, and GNS [Sanchez-Gonzalez et al., 2020]



Faster and more accurate

Dataset	Rotation aug.	Average instability of per-atom force pred.	Val Force MAE ID	Val Force MAE Average
All (130M)	✓	**0.0037**	**0.0313**	**0.0360**
All (130M)		0.0069	0.0314	0.0366
2M	✓	0.0041	0.0332	0.0382
2M		0.0093	0.0346	0.0400

Close to rotation-covariant