Contents

List of Figures

Chapter 1

Results & Discussion

THE PROPOSED FRAMEWORK is tested on the University of Ottawa (UO) database, for predicting CO_2 uptake in metal-organic frameworks (MOFs), the gas that mainly "triggered" the development of energy-based descriptors. In order to evaluate the transferability of the approach, a different host-guest system is also examined. We apply the suggested approach in the database created by Mercado et al. (2018), for predicting CH_4 uptake in covalent organic frameworks (COFs). In both cases, the resulting machine learning (ML) models are compared with conventional ones, built upon geometric descriptors. In the rest of this chapter, results from these comparisons are presented, followed by discussion for improvements of the proposed framework. Before delving into the results, we first take a look at RetNet, the 3D convolutional neural network (CNN) under the hood, that takes as input a voxelized potential energy surface (PES) and outputs a prediction for a gas adsorption property, hereon gas uptake.

1.1 Visualizing RetNet

Figure 1.1 illustrates the processing a voxelized PES undergoes, as it is passing through RetNet. For the purpose of this visualization, we use the model trained on the MOFs dataset with the largest training set size (see Section **??**). Moreover, for the ease of visualization, only some feature maps of RetNet are visualized. Please note, that each feature map of a given layer, combines all the feature maps of the precedent layer. The only exception are the pooling layers, which just downsample the feature maps from the previous layers.

For example, each feature map of the Conv2 layer takes into account all the twelve feature maps of Conv1 layer. In contrast, the feature maps of the MaxPool1 layer, are just downsampled versions of the corresponding feature maps in Conv2 layer. Although feature maps of CNNs are not meant to be interpreted by humans—especially the ones found deeper in the network—it is worth noticing that early Conv layers (i.e. Conv1 and Conv2) emphasize the texture of the structure. For instance, the third feature map of Conv1 layer delineates the skeleton of the framework.

Moving towards the output layer, the alternation of MaxPool and Conv layers continues until the Flatten layer, which just flattens out and concatenates[1] all feature maps from Conv2 layer into a single vector of size 3240. This vector is then processed by a fully connected neural network (FCNN)—i.e. the stack of Dense and Output layers—to give the final prediction. Since the Output layer is really nothing more than a linear layer, all that RetNet does is the following:

[1] Given m feature maps of size $n \times n \times n$, a Flatten layer converts them into a vector of size mn^3.

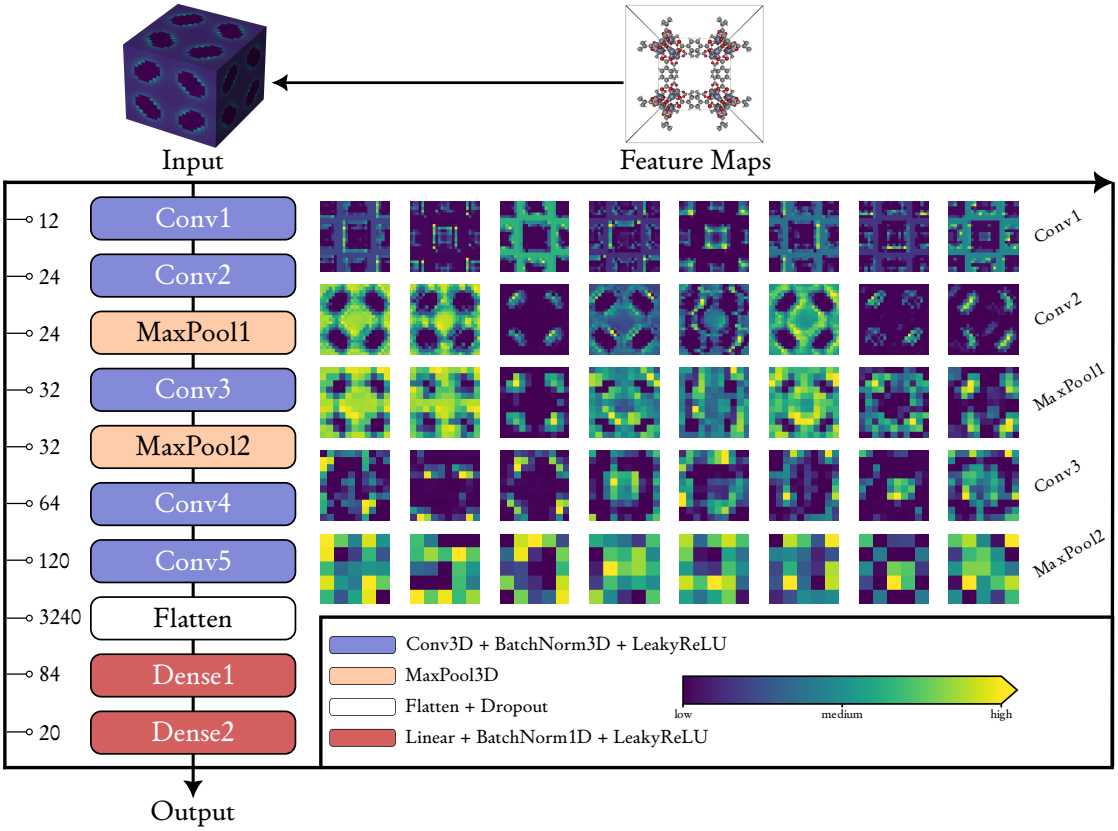
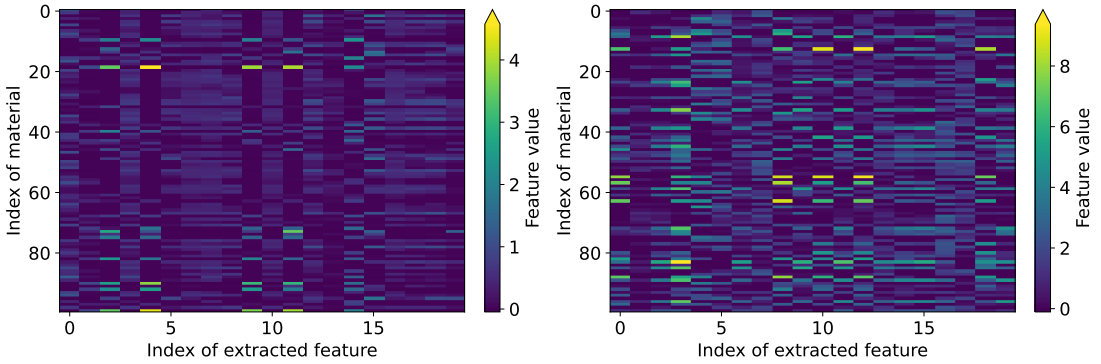


Figure 1.1: Forward pass of IRMOF-1 through RetNet. For the sake of visualization, only slices (feature maps are 3D matrices) of eight feature maps from the first five layers are visualized. For Conv1 layer, the fifth slice is presented, while for the remaining layers, the first slice is presented. The IRMOF-1 structure was visualized with the iRASPA software (Dubbeldam et al. 2018).

$$\underbrace{\boldsymbol{x}}_{\substack{\text{PES}\\ \text{input}}} \longrightarrow \underbrace{\phi(\boldsymbol{x};\boldsymbol{\theta})}_{\substack{\text{fingerprint}\\ \text{feature extraction}}} \longrightarrow \overbrace{\boldsymbol{\beta}^\top \phi(\boldsymbol{x};\boldsymbol{\theta}) + \beta_0}^{\substack{\text{gas uptake}\\ \text{output}}} \tag{1.1}$$

Equation 1.1 says that RetNet, *starting from the PES, extracts a fingerprint—that is, a high level representation of the PES—and then predicts the gas uptake by using a linear model on top of this fingerprint*. All intermediate layers between Input and Output layer participate in this feature extraction step, with the Dense2 layer determining the size of the fingerprint, which is a vector of size 20, i.e. $\phi(\boldsymbol{x}) \in \mathbb{R}^{20}$ (see Figure 1.2). The fact that *this fingerprint extraction step is learnable*—the parameters $\boldsymbol{\theta}$ of ϕ are learned during the training phase—*is what fundamentally distinguishes the proposed approach from methods that use hand-crafted fingerprints* (see Section **??**). *In these methods the fingerprint or extraction step is fixed*, and based on some heuristic, such as energy histograms (Bucior et al. 2019) or average interactions (Fanourgakis et al. 2020). *Hereon, feature extraction from the PES is no longer fixed, but is*



(A) Fingerprints extracted from the MOFs dataset.

(B) Fingerprints extracted from the COFs dataset.

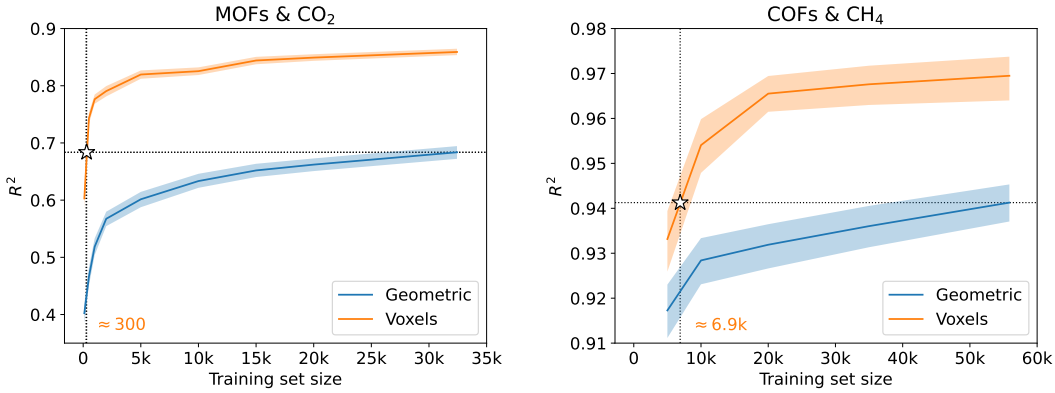
Figure 1.2: Output of the last LeakyReLU layer of RetNet trained on MOFs (left) and COFs (right) datasets, with the corresponding maximum training set size. The fingerprints of the first 100 materials in the training set are depicted.

an essential part of the training phase.

1.2 Learning Curves

For the performance boost, please see Section **??**

1.3 Discussion



(A) Learning curves for MOFs & CO_2.

(B) Learning curves for COFs & CH_4.

Figure 1.3: Performance (R^2 score) on test set as function of the training set size for conventional and CNN models. Shaded areas correspond to the 95 % confidence interval (CI). The x-coordinate of the white star denotes the trainig set size where the CNN model reaches the performance of the conventional one, the y-coordinate. "Geometric" stands for geometric descriptors, while "Voxels" stands for energy voxels.

Index

Acronyms

ANG adsorbed natural gas.

CNG compressed natural gas.

COF covalent organic framework.

CSD Cambridge Structural Database.

DAC direct air capture.

LNG liquefied natural gas.

MAE mean absolute error.

ML machine learning.

MOF metal-organic framework.