

Machine Learning Application for Permeability Estimation of Three-Dimensional Rock Images

Hongkyu Yoon¹, Darryl Melander², and Stephen J. Verzi²

¹Geomechanics Department, Sandia National Laboratories, Albuquerque, NM 87123

²Complex System for National Security, Sandia National Laboratories, Albuquerque, NM 87123
hyoon@sandia.gov

Abstract

Estimation of permeability in porous media is fundamental to understanding coupled multi-physics processes critical to various geoscience and environmental applications. Recent emerging machine learning methods with physics-based constraints and/or physical properties can provide a new means to improve computational efficiency while improving machine learning-based prediction by accounting for physical information during training. Here we first used three-dimensional (3D) real rock images to estimate permeability of fractured and porous media using 3D convolutional neural networks (CNNs) coupled with physics-informed pore topology characteristics (e.g., porosity, surface area, connectivity) during the training stage. Training data of permeability were generated using lattice Boltzmann simulations of segmented real rock 3D images. Our preliminary results show that neural network architecture and usage of physical properties strongly impact the accuracy of permeability predictions. In the future we can adjust our methodology to other rock types by choosing the appropriate architecture and proper physical properties, and optimizing the hyperparameters.

Introduction

Recent advances in multiscale imaging techniques for the analysis of complex pore structures and compositions have revolutionized our ability to characterize various porous media systems (Yoon and Dewers, 2013). Applications of imaging for porous media systems have been expanded for multi-interdisciplinary areas including fractured and porous natural media, biofilm, human bones/bodies, and various materials among many others. Flow and transport properties in porous media are very important to control and impact a variety of Earth science applications. Imaging methods have been tremendously advanced to produce 2D/3D structures and compositions of porous media over a range of scales, and numerical methods also have been advanced to fully

understand multiphysics behaviors in complex porous media (e.g., Yoon et al., 2013 and 2015). Although it is now largely possible to understand how pore topology, structure, and composition impact multiple processes affecting flow patterns, transport process, and evolution of porous media by combining a suite of imaging techniques and advanced numerical methods, integration of these techniques requires tremendous computational powers and expenses.

Recent advances in machine learning provide a great opportunity to enhance image-based property estimation and modeling capabilities (e.g., Raissi et al., 2018; Wu et al., 2018). In addition, combination of image data with other numeric and categorical data has improved the prediction of various quantities such as house prices (e.g., Rosebrock, 2019) and image classification (Aimone and Severa, 2017).

In this work, we explore how machine learning can be used to predict the permeability of porous media with physical properties. An emerging challenge for machine learning/deep learning in engineering and scientific research is the ability to incorporate physics into machine learning process. We used convolutional neural networks (CNNs) to estimate permeability by training a set of segmented data of carbonate chalk and physical properties such as porosity and surface area of fractured and porous media. We evaluated the effect of hyperparameters on permeability prediction.

Related work

Convolutional neural networks (CNNs) have been very successfully utilized for image classification and segmentation and have also been adopted for various scientific and engineering problems including permeability estimation in network systems (Wu et al., 2018), physics-informed reduced

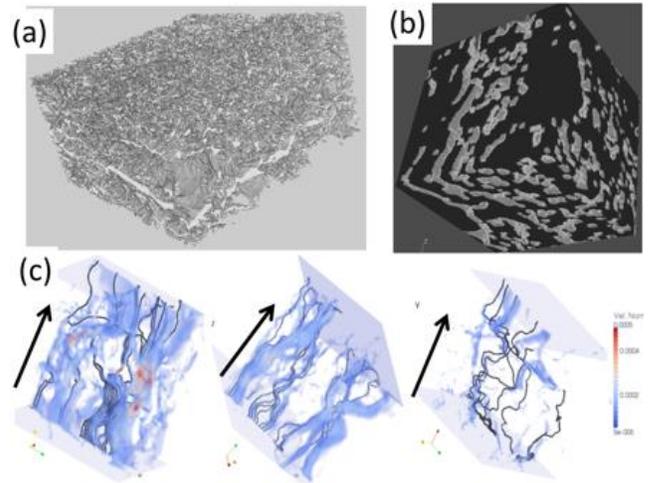
order modeling combined with high fidelity turbulence simulations (Ling et al., 2016), and extraction of flow features (Ströfer et al., 2018). In particular, recent works (Ling et al., 2016, Raissi et al., 2018) demonstrated that deep neural network architectures have an ability to account for underlying physics behind the data.

Dataset and Physical Properties

First, 256 images representative of a three-dimensional fractured and porous media system with binary phase were extracted from a larger, raw image of 3D segmented chalk, as shown in Figure 1. The raw image was obtained using focused-ion beam-scanning electron microscopy (FIB-SEM) and a corresponding segmented image (Figure 1a) constructed in Yoon and Dewers (2013) was used in this study. A total of 256 image sets were constructed from the original data of $930 \times 520 \times 962$ by sampling a subset of 256^3 images. Subsampling was performed with a regular moving volume in all three directions to represent the physical continuity of fracture features. For computational purpose, a size of 256^3 image was averaged to make a size of 128^3 as shown in Figure 1b. Each voxel resolution of the training data is 31 nanometers. Each directional permeability was calculated using lattice Boltzmann simulations (Latt et al., 2020, Yoon et al., 2015) where simulated pressure drop over each dimension and flowrate are used to compute the directional Darcy’s permeability. In addition, porosity was computed by counting the number of void voxels and the surface area was computed using Minkowsky function (Legland and Grothausmann, 2020). All physical data (permeability, porosity, and surface area) were normalized from 0 to 1. Since the logarithmic scale of the permeability is more correlated with porosity, we use a logarithmic permeability in this work. Figure 2 shows the relationship between permeability and porosity-surface area. As seen, the permeability has positive and negative correlations with porosity and surface area, respectively, however, due to the presence of fractures (Figure 1) the surface area does not have a good correlation with permeability.

Model Architecture and Training

Additional physical information can provide physical constraints for training the model. The combination of image and numerical data allows us to build and train a hybrid physics-constrained machine learning model. To handle processing of the porous media images, we have developed convolutional neural networks (CNNs) whose input consists of binary phase images. We started with a general notion of a desired learning model architecture that takes 3D image data and pore topology metrics as input, but wanted to explore potential permutations within that notional



architecture. We developed a general framework to perform optimization of hyperparameters for convolutional neural network architectures. A concrete learning model can be generated by

Figure 1. (a) original segmented image (white: fractures and pores; dark gray: solid matrix), (b) an example of the subset with a size of $128 \times 128 \times 128$ voxels, (c) directional velocity profiles from lattice Boltzmann simulations.

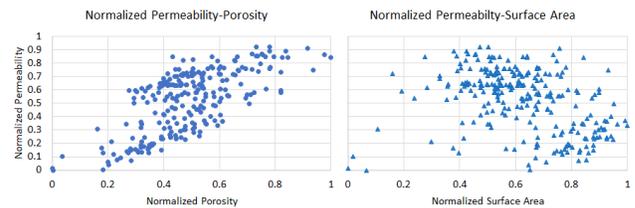


Figure 2. Normalized permeability in x-direction ($\ln k_x$) vs. normalized porosity (left) and surface area (right) for all 256 data.

choosing specific values for architectural options and hyperparameter variables as shown in Figure 3, including:

- The number of convolutional layers
- The size, stride, and number of filters in the convolutional layers
- Whether to use batch normalization, max pooling, or both
- Which activation function to use
- The size and number of fully-connected dense layers at certain locations in the model.

In this work, as a basis we use a 3D CNN (conv3D) layer for 3D image data and a multilayer perception for physical numeric data. 256 3D datasets and corresponding porosity and surface area data split into training, validation, and testing data consisting of 164, 41, and 51 samples, respectively. The training set was also increased through data augmentation by rotating each image along the x-axis by 90° , 180° , and 270° . When rotating along the x-axis, the permeability

in the x-direction is the same for all rotations. The resulting size of the training set is 656 images. Mean squared error (MSE) was used as loss function and early stopping with 35 epochs as patience was used.

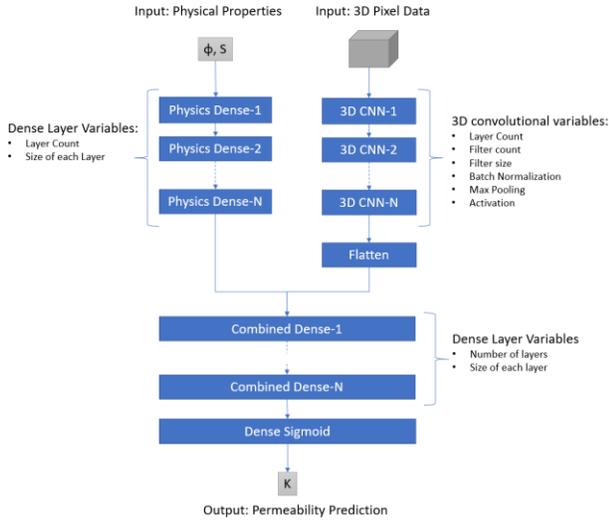


Figure 3. Schematic of convolutional neural network structure and additional information stream to optimize network architecture and hyperparameters.

Results and Discussion

In our preliminary results, the learning model that gave the best results had the following architecture:

- 5 convolutional layers
- 16 3x3 filters per convolutional layer
- 2 dense layers dedicated to physical properties
- 2 dense layers that operate on both image and physical quantities together

The mean squared error (MSE) and mean absolute error (MAE) values for models with and without physical data (porosity and surface area) are reported in Table 1. Results with testing data sets are shown in Figure 4 with a linear

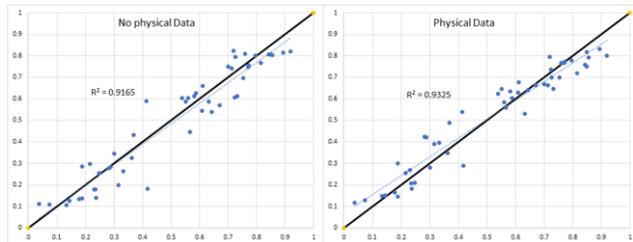


Figure 4. The permeability prediction against test set without (left) and with (right) physical data. The linear regression fitting is also shown. The black line represents the 1 to 1 line.

regression fitting (dotted blue line) and an associated R^2 value. As a reference, the single perfect black line is also shown. The slope shows the overall performance of each model with a better performance closer to one, while the R^2 value shows the proximity of predicted data along the linear regression line.

Table 1 shows that addition of physical quantities (porosity and surface area) during training reduced the MAE by $\sim 10\%$ and MSE by $\sim 20\%$, compared to the model without physical data. As shown in Figure 2, the porosity and surface area are correlated with the permeability, so both sources of information would provide additional physical constraints that are combined with features extracted from image data. Although there is a need to study what features are extracted from image and how two input data can be used to learn the underlying feature to the permeability, Figure 4 shows that the model architecture with both types of physical data tend to predict the permeability with a smaller scattered pattern. More interestingly, the regression lines of both results are well aligned with the 1-to-1 line, indicating that both trained models tend to converge to the perfect predictive model, but this needs to be evaluated further with more training data.

Another important aspect is that the improvement with physical data in this work was lower, compared to our previous study with synthetic 2D bead packing cases (Yoon et al., 2019) where training and testing results were improved with two physical data by more than 300 % and 30%, respectively. This may imply that the physical constraints from the numeric data would be not greatly informative or the amount of training data with the 3D CNN architectures may be not big enough to extract 3D features. In fact, the impact of including physical quantities varied significantly from run to run, and between model architectures. This observation may be related to the relation between the surface area and permeability (Figure 2) which is strongly influenced by the features of original chalk data. As seen in Figure 1, the chalk has strong microfracture networks, hence instead of surface area other physical quantities such as multipoint statistics including percolation length, connectivity length, and Euler number may represent the permeability property better. For 3D CNN architectures the amount and types of 3D data used in this study may not be enough as

Table 1. Summary of results with and without physical data for model training.

	MAE	MSE
No physical data	0.0622	0.00576
Physical data (porosity and SA)	0.00547	0.00461

MSE with normalized permeability values.

seen in image segmentation of 3D data. In the future it will be pursued with comprehensive 2D architectures to explore hyperparameters more computationally efficiently.

Conclusions

We evaluated how additional physical information can enhance the permeability prediction with CNN models. As it is now well accepted in the community that a physics-informed and/or physics-constrained machine learning model can overcome overfitting to the training data and improve the features underlying the physical processes, there is a strong need to improve how the physical constraints and/or additional information (e.g., equations and theory) can enhance the learning process in machine learning. Our results clearly show that the optimal neural network architecture and implementation of physics-informed constraints are important to properly improve the model prediction of permeability. The analysis of the features learned by each layer and the output data from the MLP will reveal a better mechanistic understanding of the machine learning processes. A comprehensive analysis of hyperparameters with different CNN architectures and the data implementation scheme of the physical properties will be performed to optimize our learning system for various porous media system.

Acknowledgments

This work was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

Aimone, J.B.; Severa, W.M. 2017. Context-modulation of hippocampal dynamics and deep convolutional networks: Using parallel pathways to limit network size. In *Proceedings of 2017 NIPS Cognitive Influenced Artificial Intelligence Workshop*. arXiv preprint arXiv:1711.09876.

Bultreys, T.; De Boever, W.; Cnudde, V. 2016. Imaging and image-based fluid transport modeling at the pore scale in geological materials: A practical introduction to the current state-of-the-art. *Earth-Science Reviews* 155: 93-128.

Gostick, J.; Khan, Z.A.; Tranter, T.G.; Kok, M.D.R.; Agnaou, M.; Sadeghi, M.A.; Jervis, R. 2019. PoreSpy: A Python Toolkit for

Quantitative Analysis of Porous Media Images. *Journal of Open Source Software*. doi:10.5281/zenodo.2633284.

Latt, J., Malaspinas, O., Kontaxakis, D., Parmigiani, A., Lagrava, D., Brogi, F., Belgacem, M.B., Thorimbert, Y., Leclaire, S., Li, S. and Marson, F., 2020. Palabos: Parallel Lattice Boltzmann solver. *Computers & Mathematics with Applications*.

Ling, J.; Kurzawski, A.; Templeton, J. 2016. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics* 807: 155-166.

Legland, D.; Grothausmann, R. 2020. Image Processing library for Matlab, MatImage v1.1, <https://github.com/mattools/matImage>

Raissi, M.; Karniadakis, G.E. 2018. Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics* 357: 125-141.

Rosebrock, A., Keras: Multiple Inputs and Mixed Data, <https://www.pyimagesearch.com/2019/02/04/keras-multiple-inputs-and-mixed-data/>, accessed on July 15, 2019.

Ströfer, C.M.; Wu, J.; Xiao, H.; Paterson, E. 2018. Data-driven, physics-based feature extraction from fluid flow fields. arXiv preprint arXiv:1802.00775.

Wu, J.; Yin, X.; Xiao, H. 2018. Seeing permeability from images: fast prediction with convolutional neural networks. *Science bulletin* 63(18): 1215-1222.

Yoon, H.; Dewers, T.A. 2013. Nanopore structures, statistically representative elementary volumes, and transport properties of chalk. *Geophysical Research Letters* 40(16): 4294-4298.

Yoon, H.; Kang, Q.; Valocchi, A.J. 2015. Lattice Boltzmann-based approaches for pore-scale reactive transport. *Rev. Mineral. Geochem* 80: 393-431. <http://dx.doi.org/10.2138/rmg.2015.80.12>.

Yoon, H., Melander, D.J. and Verzi, S.J., 2019. Permeability Prediction of Porous Media using Convolutional Neural Networks with Physical Properties. AAAI 2020 Spring Symposium on Combining Artificial Intelligence and Machine Learning with Physics Sciences. http://ceur-ws.org/Vol-2587/article_16.pdf.