Deep Learning Emulators for Pore-Scale CO_2-Water Interaction

ECO-AI project team

Heriot-Watt University

Objective

The goal of this hackathon is to build an emulator that can predict the spatiotemporal evolution of CO_2 invasion in porous media from the benchmark dataset of pore-scale CO_2-water interaction. Successful emulators will serve as fast surrogate models that can accurately approximate complex and computationally expensive multiphase flow simulations performed using the GeoChemFoam solver.

Data

Participants will be provided with results from CO_2-water flow simulations in synthetic porous media. Each simulation represents CO_2 injection into a water-filled porous medium representing heterogeneous porous media.

The dataset consists of:

- Domain arrays: Binary images of size 512×128 pixels representing pore space (1) and solid grains (0), with a resolution of 35μm per pixel
- Invasion arrays: Time series data showing CO_2 displacement of water over 25 discrete time steps
- Training samples with a **single pattern** of heterogeneity (with a random stochastic variations in spacing)

The training dataset captures the complex non-linear dynamics of CO_2 displacing water in porous media at low capillary numbers ($\sim 5 \times 10^{-6}$), where both capillary and viscous forces influence flow patterns.

Challenge

Teams will compete to build the most accurate emulators that can predict the spatiotemporal evolution of the CO_2 invasion field based only on the initial domain geometry. Successful emulators should:

1. Accurately capture how CO_2 progressively invades the pore space over time
2. Maintain physical consistency in the predicted invasion patterns
3. Generalize well to unseen domain geometries
4. Run significantly faster than the original GeoChemFoam simulations

Key Considerations

- Emulators should accurately capture the complex non-linear dynamics of multiphase flow, including preferential flow paths, breakthrough phenomena, and capillary fingering
- Predictions should respect the physical domain constraints (CO_2 invasion can only occur in pore spaces, not in solid grains)
- Uncertainty quantification of the emulator predictions is encouraged
- Memory efficiency and inference speed are important evaluation factors

Evaluation

Emulator performance will be evaluated using several metrics:

1. **Spatial Accuracy**: Mean Squared Error (MSE) and Mean Absolute Error (MAE) between predicted and actual CO_2 invasion fields
2. **Temporal Consistency**: Error growth with time (RMSE vs. time step)
3. **Physical Validity**: Adherence to domain constraints and flow physics
4. **Computational Efficiency**: Training time, memory requirements, and inference speed

Suggested Methods

Participants are encouraged to explore machine learning approaches suited for spatiotemporal modeling of physical processes, such as:

- U-Net architectures with time encoding for direct prediction at specified time steps
- Physics-informed neural networks that incorporate flow constraints
- Temporal models based on ConvLSTM, GRU, or Transformer architectures
- Autoregressive models that predict the next time step from previous ones
- Fourier Neural Operators or other mesh-independent techniques
- Hybrid models that combine data-driven approaches with physical constraints

Starter Code

Participants will be provided with PyTorch starter code that includes:

- Data loading and preprocessing utilities
- Dataset and DataLoader implementations for handling spatiotemporal data
- A simple U-Net architecture with time encoding
- Training loop implementation with custom loss functions
- Evaluation metrics and visualization tools

References

1. Maes, J., Menke, H. P. GeoChemFoam: Direct modelling of flow and heat transfer in micro-CT images of porous media. *Heat and Mass Transfer*, 58(11), 1937-1947, 2022.
2. Abdellatif, A., Menke, H. P., Maes, J., Elsheikh, A. H., Doster, F. Benchmark Dataset for Pore-Scale CO_2-Water Interaction.
3. Ronneberger, O., Fischer, P., Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015.
4. Wen, G., Li, Z., Azizzadenesheli, K., Anandkumar, A., Benson, S. M. U-FNO—An enhanced Fourier neural operator-based deep-learning model for multiphase flow. *Advances in Water Resources*, 163, 104180, 2022.
5. Poels, Y., Minartz, K., Bansal, H., and Menkovski, V. Accelerating Simulation of Two-Phase Flows with Neural PDE Surrogates. *arXiv preprint arXiv:2405.17260*, 2024.

This hackathon aims to bridge the gap between computational fluid dynamics and machine learning by developing efficient emulators for multiphase flow in porous media. The winning solutions will help enable rapid analysis and guide future GeoChemFoam simulations for applications such as carbon capture and storage and related geoenergy applications.