SURF
Summer Undergraduate Research Fellowship

MACHINE LEARNING-BASED SUBSURFACE FLOW SIMULATION

Author: Xinqiao Tong, Binyi Zou Supervisor: Jie Chen
School of Science, Department of Applied Mathematics, Xian'JiaoTong-Liverpool University

ABSTRACT

Apart from oceans and rivers, underground flow is also a significant component of the water system on the earth. It is critical to study the subsurface flow because extractions of underground water, gas and oil largely depend on relevant research. Darcy Equation describes the relationship between the pressure in the field and velocity of the flow, and it has a wide industrial application. In our project, we use two deep learning-based methods, namely Physics Informed Neural Network (PINN) and Deep Ritz Method, to simulate the pressure distribution in the computational domain.

PRINCIPAL OF PHYSICS INFORMED NEURAL NETWORK (PINN)

Physics Informed Neural Network (PINN) was proposed in 2017, which are trained to solve supervised learning tasks while respecting any given law of physics described by general nonlinear partial differential equations. PINN has eight hidden layers, The idea behind PINN is that it regards points which lie on the boundaries (with known boundary condition) as the training set and let the physics law embodied in the area serve as the regularization agent. That is, the network is trained to minimize the mean square error, $MSE = MSE_u + MSE_f$,where MSE_u stands for the error from the boundary and MSE_f represents the error of the collocation points, originating from the constraint from the partial differential equation. Incarnating in formulas, we have $MSE_u = \frac{1}{N_u}\sum_{i=1}^{N_u} |u({x_u}^i, {y_u}^i) - u^i|^2$, $MSE_f = \frac{1}{N_f}\sum_{i=1}^{N_f} |f({x_f}^i, {y_f}^i)|^2$.

We chose to optimize all loss functions using L-BFGS: a quasi-Newton, full-batch gradient-based optimization algorithm.We hoped to achieve a satisfying prediction through a sufficiently expressive neural network architecture, namely training all 3021 parameters of a 9-layer deep neural network using the mean squared error loss. Each hidden layer contained 20 neurons and a hyperbolic tangent activation function (Figure 2). It was claimed to be data-efficient and robust for both linear and nonlinear differential equations.

FIGURE 1

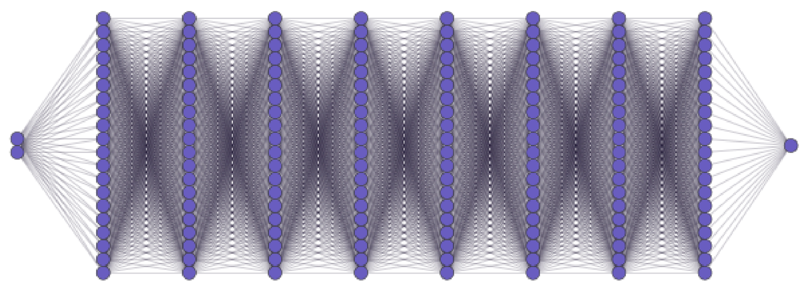
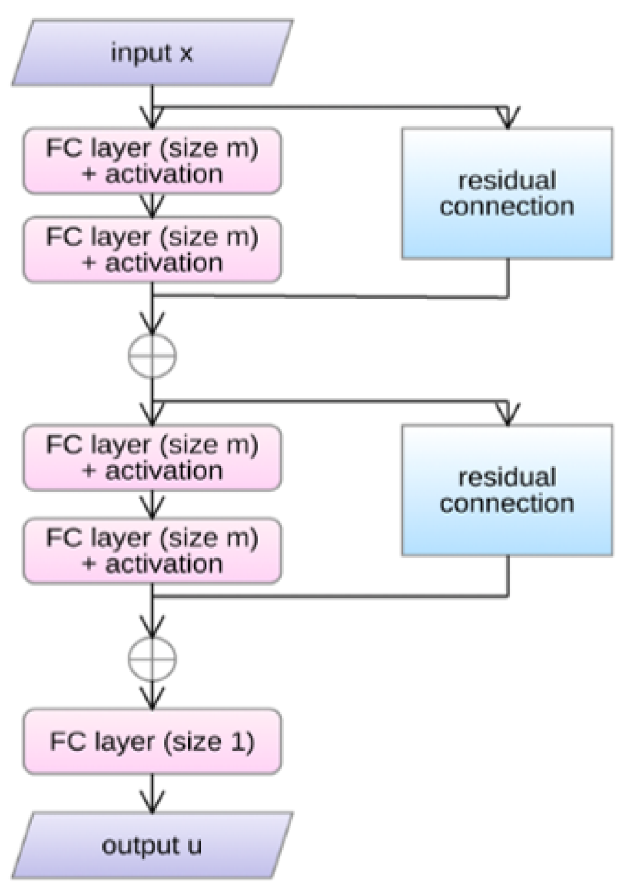


Figure 1: Structure of PINN

PRINCIPAL OF THE DEEP RITZ METHODS

FIGURE 2



Deep Ritz Method Deep Ritz method is a deep learning-based method proposed by Weinan E., and it is designed to solve variational problems numerically. The Deep Ritz Method is based on the following set of ideas:

1. Deep neural network-based approximation of the trial function.
2. A numerical quadrature rule for the functional.
3. An algorithm for solving the final optimization problem.

For example, to solve the PDE problem:

$$-\Delta u(x) = f(x), x \in \Omega$$
$$u(x) = 0, x \in \partial\Omega$$

There are generally two steps to conduct this method.
- Change the PDE problem into an optimization problem. The objection is to minimize the function below, where β is the penalty parameter that we need to set by ourselves.

$$I(u) = \int_\Omega \left(\frac{1}{2}|\nabla_x u(x)|^2 - f(x)u(x)\right) \mathrm{d}x + \beta \int_{\partial\Omega} u(x)^2 \mathrm{d}s.$$

- Use the residual network to solve the optimization problem.

TASKS

In our SURF project, we aims to solve for the numerical solution to the Darcy equation, i.e.,

$$k^{-1} \cdot u + \nabla p = 0,\ in\ \Omega$$
$$\nabla \cdot p = 0,\ in\ \Omega$$

subject to the boundary condition $u|_{x=0} = -\frac{1}{2} \cdot y^2 + \frac{1}{2} \cdot y$, $u|_{x=1} = u|_{y=0} = u|_{y=1} = 0$, where k is the heterogeneous permeability, u is the Darcy velocity, p is the pressure and Ω is the computational domain, a 1×1 *square*.

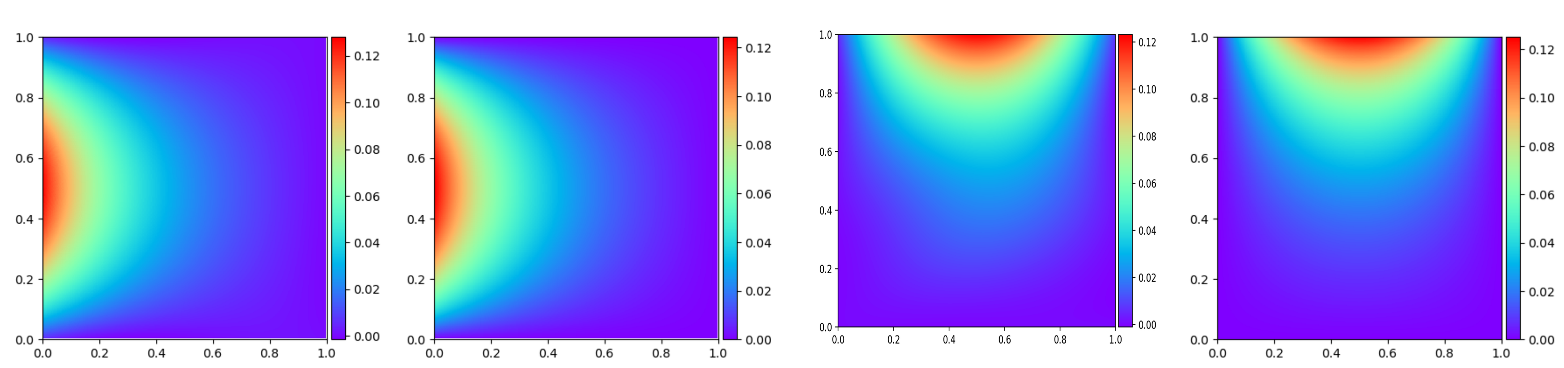
Note that by annihilating the Darcy velocity in the Darcy equation, Darcy equation can be reduced to $k\nabla \cdot u = 0$. There are three scenarios for the permeability k, which are

- $k = 1$ in the whole Ω
- the heterogeneous permeability k fluctuates in fine grids with high-contrast
- there exists fractures in Ω, and the permeability k within the fractures is in high contrast to that outside the fractures.

SCENARIO I: UNIFORM PERMEABILITY

Firstly we focused on the scenario when $k = 1$ in the whole Ω. PINN and Deep Ritz method were reappeared to solve the Darcy equation. To testify the prediction from PINN and Deep Ritz, we generated the conventional numerical solution to the Darcy equation by the PDE Toolbox and compared prediction from PINN and Deep Ritz respectively. Figure 3 shows the comparison. (Post Script: the boundary condition for Deep Ritz was set to be $u|_{y=1} = -\frac{1}{2} \cdot x^2 + \frac{1}{2} \cdot x$, $u|_{x=1} = u|_{y=0} = u|_{y=1} = 0$.Basically the two boundary conditions are the same, except the direction of the frontal surface.)

FIGURE 3



Prediction from PINN | Answer from MATLAB | Prediction from Deep Ritz | Answer from MATLAB

In PINN, we set $N_u = 100$, $N_f = 10000$ and the resulting prediction error was measured at 1.205828×10^{-2} in the relative $L2 - norm$.

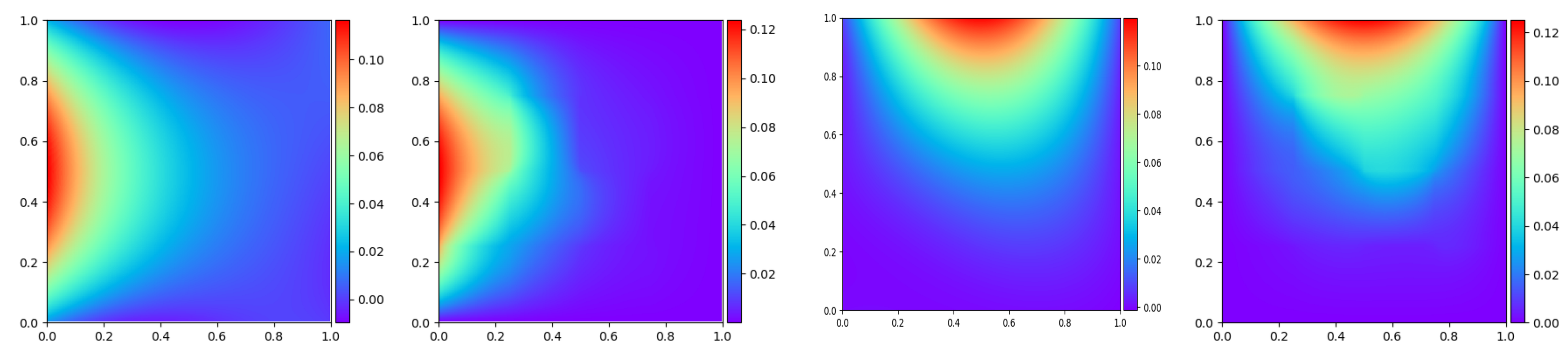
In Deep Ritz, we set $\beta = 50$ and the resulting prediction error was measured at 3.6936×10^{-2}, which is also in the relative $L2 - norm$.

The prediction from both method is satisfying and the error from PINN is nearly one third of that from Deep Ritz. This indicates that PINN is more capable of solving rudimentary problems.

SCENARIO II: HETEROGENEOUS PERMEABILITY IN GRIDS

In order to simulate the real world situation in subsurface flow, the region Ω was divided into 16 areas averagely with different permeability k. We randomly generated the permeability k from [0,50], [0,100] and [0, 250] respectively. Numerical example of k_{50} is shown in figure 4, In terms of PINN, We compare small training set and larger training set performance. With respect to Deep Ritz, we chose the most appropriate penalty parameter β to obtain the feasible solution.

FIGURE 4



PINN ($N_u = 1000, N_f = 80000$) | Answer from MATLAB | Deep Ritz ($\beta = 2500000$) | Answer from MATLAB

relative-error	PINN($N_u = 100, N_f = 10000$)	PINN($N_u = 1000, N_f = 80000$)	Deep Ritz
k_{50}	1.431120e-01	1.073231e-01	**7.14e-02**
k_{100}	2.555351e-01	1.664050e-01	**9.9e-02**
k_{250}	5.150456e-01	3.057628e-01	**1.901e-01**

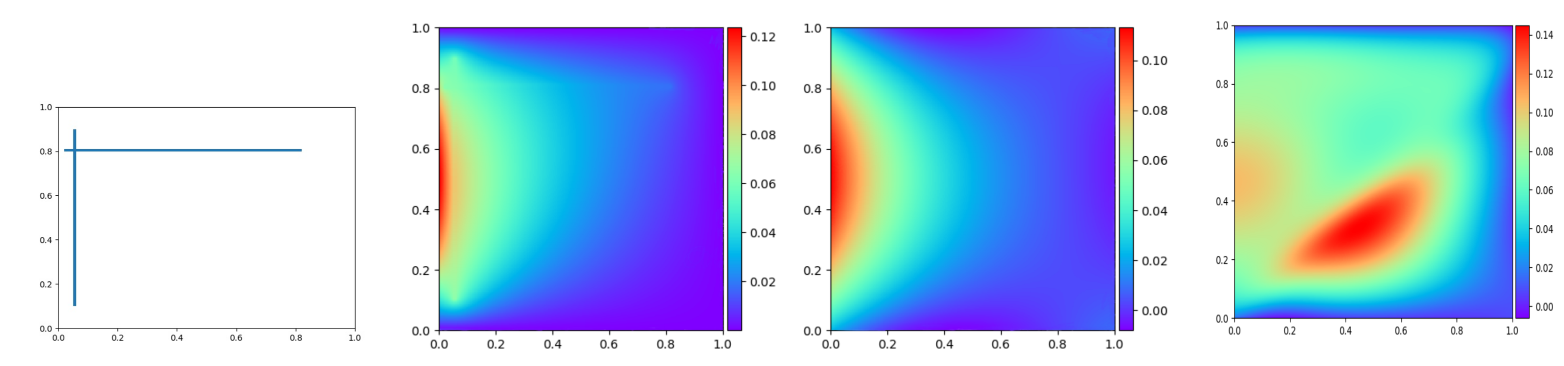
Table: The L_2 relative error for different methods

In view of above analysis, With respect to k_{50} and k_{100}, the performance of Deep Ritz Method is excellent. However,PINN cannot show a reasonable solution. in terms of k_{250}, Deep Ritz Method and PINN both cannot attain a reasonable accuracy. Hence, the result indicates Deep Ritz method is more capable of solving heterogeneous permeability problems.

SCENARIO III: FRACTURES IN Ω

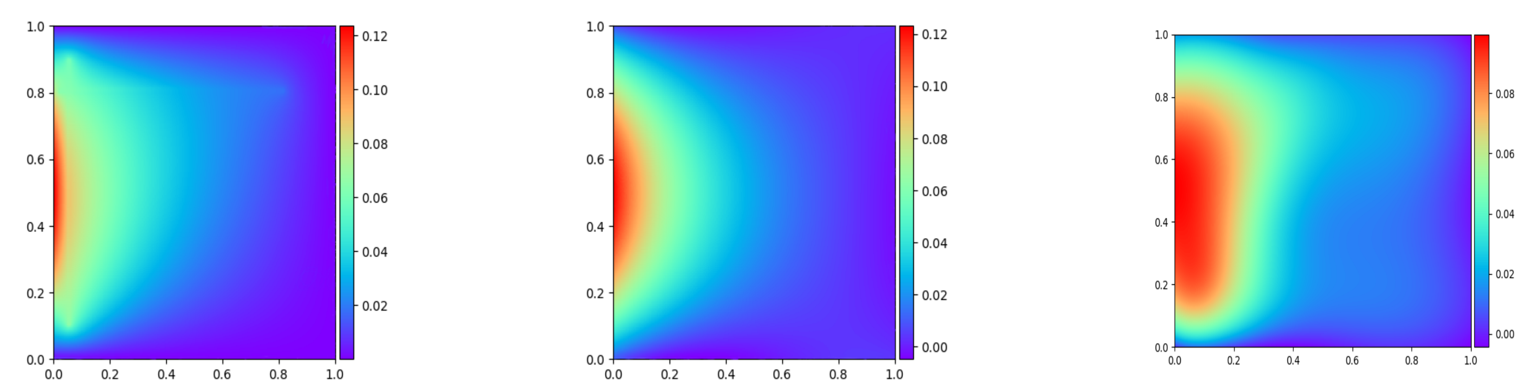
In this final section, we tried to solve the Darcy equation when there are fractures in Ω, i.e.,the permeability inside the fractures is 100 and that outside is 1. The shape of the fractures is shown in Figure 5. Firstly we tried to adapt our previous model to this problem. In terms of PINN, since its performance was not satisfying in scenario ii, we increased $N_u = 1000$ and $N_f = 100000$. However, both methods failed to recognize the existence of the fractures (see in Figure 5). To improve the predictions, we adopted exponentially decaying learning rate in Deep Ritz. In PINN, we attempted to use an adaptive sampling strategy to determine the training points. That is, by iteration, we choose 10000 points by Latin Hyper-cube Sampling and add the top 2000 points with the highest f-prediction value to the final set of collocation points.Finally, we took 1000 training points and 5w adaptive sampled points as collocation points to train the PINN. The results and relative error is shown in Figure 6.

FIGURE 5



Shape of Fractures | Answer from MATLAB | PINN, relative error:0.2631 | Deep Ritz, relative error:0.5347

FIGURE 6



Answer from MATLAB | PINN with Adaptive Sampling error:0.1433 | Deep Ritz with Decaying Learning Rate error:0.1649

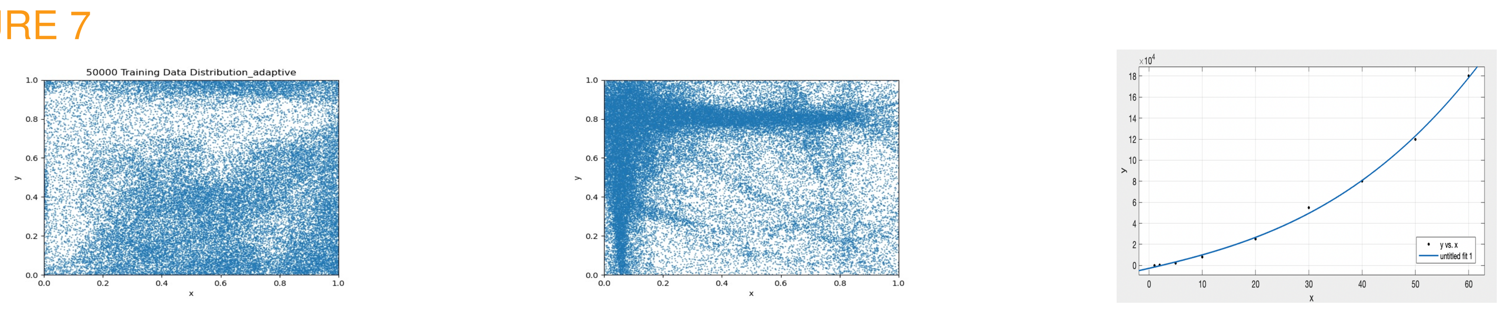
Unfortunately, despite the relative error was reduced significantly after the improvements, the two methods still cannot figure out the location of the fractures. Since we have reached the limit of our computing power, the problem is left for future research.

INTERESTING DISCOVERIES

We now know that even the PINN with adaptive sampling method still cannot figure out the existence of the s. But an interesting phenomenon was discovered in the training process. Since the L-BFGS optimizer is a gradient-based optimization algorithm, it prioritises to minimize part of the *MSE* with the larger weight ($MSE = k \cdot (\frac{\partial^2 u}{\partial^2 x_1} + \frac{\partial^2 u}{\partial^2 y_1}) + (\frac{\partial^2 u}{\partial^2 x_2} + \frac{\partial^2 u}{\partial^2 y_2})$), where x_1, y_1 are collocation points inside the fractures and x_2, y_2 are collocation points outside the fractures. Since L-BFGS optimizer prioritises to minimize the error contributed by the fractures, when we add the top 1000 points with highest f-prediction value, we are not adding the points in the neighbourhood of the fractures as we presumed (in the traditional numerical methods, points near the fractures are considered relatively more difficult to be computed with small error). Actually, we are adding points other than the points near the fractures to the final collocation points.

For Deep Ritz method, the selection of parameter β has a large impact on our results. We found that there is a most suitable β for each case. In general, the selection of β depends on the permeability k and the right side of the equation. In our experiment, we tested the relationship between the permeability k and the best-suitable-β in the case of $f = 0$. We made a fitted curve for it. It can be seen from the figure below that there is a parabolic relation between k and the best-suitable-β. Here x represents the permeability k and y represents the best-suitable-*beta*

FIGURE 7



Points with Largest f-prediction Value | Points with Smallest f-prediction Value | The relation between k and the best-suitable-β

CONCLUSION

In our SURF project, we used two deep-learning based methods, namely Physics Informed Neural Network(PINN) and Deep Ritz method, to simulate the subsurface flow. It can be seen that our two methods both work well when the permeability of the subsurface is uniform or when the range of the permeability of the nonuniform subsurface is within 100. However, our methods cannot be applied well for the case with fractures, or with blocks where the range of the permeability is larger than 100. It provides a challenge for our methods and give us a chance to make them improved in the future.