Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2021-0257

Vol. **32**, No. 2, pp. 299-335
August 2022

MOD-Net: A Machine Learning Approach via Model-Operator-Data Network for Solving PDEs

Lulu Zhang[1], Tao Luo[2,3], Yaoyu Zhang[1,3,4], Weinan E[5,6], Zhi-Qin John Xu[1,3,*] and Zheng Ma[2,3,*]

[1] *Institute of Natural Sciences and School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai, 200240, China.*
[2] *School of Mathematical Sciences and Institute of Natural Sciences, Shanghai Jiao Tong University, CMA-Shanghai, Shanghai, 200240, China.*
[3] *MOE-LSC and Qing Yuan Research Institute, Shanghai Jiao Tong University, Shanghai, 200240, China.*
[4] *Shanghai Center for Brain Science and Brain-Inspired Technology, Shanghai, 200031, China.*
[5] *School of Mathematical Sciences, Peking University, Beijing, 100871, China.*
[6] *AI for Science Institute, Beijing, 100080, China.*

Received 28 December 2021; Accepted (in revised version) 10 June 2022

Abstract. In this paper, we propose a machine learning approach via model-operator-data network (MOD-Net) for solving PDEs. A MOD-Net is driven by a model to solve PDEs based on operator representation with regularization from data. For linear PDEs, we use a DNN to parameterize the Green's function and obtain the neural operator to approximate the solution according to the Green's method. To train the DNN, the empirical risk consists of the mean squared loss with the least square formulation or the variational formulation of the governing equation and boundary conditions. For complicated problems, the empirical risk also includes a few labels, which are computed on coarse grid points with cheap computation cost and significantly improves the model accuracy. Intuitively, the labeled dataset works as a regularization in addition to the model constraints. The MOD-Net solves a family of PDEs rather than a specific one and is much more efficient than original neural operator because few expensive labels are required. We numerically show MOD-Net is very efficient in solving Poisson equation and one-dimensional radiative transfer equation. For nonlinear PDEs, the nonlinear MOD-Net can be similarly used as an ansatz for solving nonlinear PDEs, exemplified by solving several nonlinear PDE problems, such as the Burgers equation.

AMS subject classifications: 35C15, 35J05, 35Q20, 35Q49, 45K05

Key words: Deep neural network, radiative transfer equation, Green's method, neural operator.

*Corresponding author. *Email addresses:* xuzhiqin@sjtu.edu.cn (Z.-Q. J. Xu), zhengma@sjtu.edu.cn (Z. Ma)

1 Introduction

Nowadays, using deep neural networks (DNNs) to solve PDEs attracts more and more attention [4–7, 9, 10, 12–20]. Here we review three DNN approaches for solving PDEs.

The first approach is to parameterize the solution by a DNN and use the mean square of the residual of the equation [3, 19] or the variational forms [7, 15] as risk or loss, by minimizing which the DNN output satisfies PDE. A comprehensive overview can be found in [6]. This parameterization approach can solve very high-dimensional PDEs and does not require any labels. However, it only solves a specific PDE during each training trial, that is, if the PDE setup changes, such as the source terms, the boundary conditions or other parameters in the PDE, we have to train a new DNN. An important characteristic of the parameterization approach is slow learning of the high frequency part as indicated by the frequency principle [18,24,25]. To overcome the curse of high frequency, a series of multiscale approaches are proposed [1, 12, 16, 21, 23]. The second approach uses DNN to learn the mapping from the source term to the solution [8]. In this approach, the source function and the solution are sampled at fixed grid points as two vectors. Then the vector of the source function is fed into the DNN to predict the vector of the solution function. The advantage of this mapping approach is that the DNN solves the PDE for any source function, thus it can be very convenient in application. However, the mapping approach can only evaluate the solution at fixed points. DeepOnet [17] is proposed that the source function is still fed into the network on fixed grid points but the output can be evaluated on any points by adding on extra inputs of the points to the network. Such approach requires very large sample points, which is often computational inefficient or intractable, especially when dealing with high-dimensional PDEs or complicated integro-differential equations, such as Boltzmann equation and radiative transfer equation (RTE). The third approach is called neural operator [13, 14], which represents the solution based on the form similar to the idea of the Green's function and the DNN is used to parametrize the Green's function. The neural operator solves a type of PDEs but not a specific PDE and can be evaluated at any time or spatial points. Training of the neural operator is to minimize the difference between the learned solution and the true solution at randomly sampled points. Therefore, the neural operator is a data-driven method and requires a large amount of labels, a similar difficulty to the mapping approach.

In this work, we propose a machine learning approach via model-operator-data network (MOD-Net) for solving PDEs. The MOD-Net has advantages including: (i) obtaining a functional representation of the solution which allows evaluating the solution at any points; (ii) requiring none or few labels numerically computed by a traditional scheme on coarse grid points with cheap computation; (iii) solving a family of PDEs but not a specific PDE. The three key components of MOD-Net are illustrated as follows.

Model driven. MOD-Net is driven by the physical model to avoid using too much expensive labeled data. That is, the empirical risk, i.e., training loss requires the solution satisfying the constraints of the governing PDE or equivalent forms and boundary condi-