Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.OA-2024-0124

Vol. **18**, No. 1, pp. 31-65
February 2025

R-Adaptive DeepONet: Learning Solution Operators for PDEs with Discontinuous Solutions Using an R-Adaptive Strategy

Yameng Zhu[1], Jingrun Chen[2] and Weibing Deng[1,*]

[1] *School of Mathematics, Nanjing University, Nanjing 210093, P.R. China*
[2] *School of Mathematical Sciences and Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou 215127, P.R. China*

Received 24 October 2024; Accepted (in revised version) 7 January 2025

Abstract. When DeepONet approximates solution operators of partial differential equations (PDEs) with discontinuous solutions, it poses a foundational approximation lower bound due to its linear reconstruction property. Inspired by the moving mesh method, we propose an R-adaptive DeepONet method, which consists of: (1) the output data representation is transformed from the physical domain to the computational domain using the equidistribution principle; (2) the maps from input parameters to the solution and the coordinate transformation function over the computational domain are learned using DeepONets separately; (3) the solution over the physical domain is obtained via post-processing methods such as the interpolation method. Additionally, we introduce a solution-dependent weighting strategy in the training process to reduce the error. We establish an upper bound for the reconstruction error based on piecewise linear interpolation and show that the introduced R-adaptive DeepONet can reduce this bound. Moreover, for two prototypical PDEs with sharp gradients or discontinuities, we prove that the approximation error decays at a superlinear rate with respect to the trunk basis size, unlike the linear decay observed in vanilla DeepONets. Numerical experiments on several PDEs with discontinuous solutions are conducted to verify the advantages of the R-adaptive DeepONet over available variants of DeepONet.

AMS subject classifications: 47-08, 47H99, 65D15, 65M50, 68Q32, 68T05, 68T07

Key words: Scientific machine learning, neural operators, DeepONet, R-adaptive method.

---

1. Introduction

Many interesting phenomena in physics and engineering are described by partial differential equations whose solutions contain sharp gradient regions or discontinuities.

*Corresponding author. *Email addresses:* `dg21210022@smail.nju.edu.cn` (Y. Zhu), `jingrunchen@ustc.edu.cn` (J. Chen), `wbdeng@nju.edu.cn` (W. Deng)

The most common types of such PDEs are hyperbolic systems of conservation laws [10], such as Euler equations, inviscid Burgers' equation, etc. It is well-known that solutions of these PDEs develop finite-time discontinuities such as shock waves, even when the initial and boundary data are smooth. Other examples include convection-dominated equations, reaction-diffusion equations, and so on. It is challenging for traditional numerical methods because resolving these discontinuities, such as shock waves and contact discontinuities, requires petite grid sizes. Moreover, characterizing geometric structures, especially in terms of effectively suppressing numerical oscillations near discontinuous interfaces and maintaining the steepness of transition interfaces, is difficult. Specialized numerical methods such as adaptive finite element methods [1] and discontinuous Galerkin finite element methods [12] have been successfully used in this context, but their high computational cost limits their wide use.

At the same time, data-driven approaches are becoming a competitive and viable means for solving these challenging problems. Deep neural networks (DNNs) have shown promising potential for solving both forward and inverse problems associated with PDEs [3]. Numerous researchers have explored methods that utilize DNNs for solving PDEs (see [13,21] and references therein).

Machine learning for PDEs primarily focuses on learning solutions by training a mapping from the computational domain to the solution. This process, known as the solution parameterization, encompasses techniques such as the deep Ritz method [14], deep Galerkin method [34], and physics informed neural networks (PINNs) [5,32]. These methods utilize DNNs to represent the solution and integrate the PDE information into the loss function. The approximate solution is obtained by minimizing the loss function. Since proposed, these methods have been successfully applied to solve both forward and inverse problems for various linear and nonlinear PDEs [5,18,30]. Note that these approaches are tailored to specific instances of PDEs. Consequently, if the coefficients or initial conditions associated with the PDEs change, the model has to be retrained, resulting in poor generalization ability across different PDEs.

Along another line, there is ongoing work on parameterizing the solution map using DNNs, referred to as operator learning [2, 8, 17, 20, 26, 28]. In [8], Chen and Chen introduced a novel learning architecture based on neural networks, termed operator networks, and demonstrated that these operator networks possess an astonishing universal approximation property for infinite-dimensional nonlinear operators. Recently, the authors of [28] replaced the shallow branch and trunk networks in operator networks with DNNs and proposed the deep operator network (DeepONet). Since proposed, it has been successfully applied to a variety of problems with differential equations [6,9,27,39]. In [26], Li *et al.* proposed Fourier neural operators based on a nonlinear generalization of the kernel integral representation for some operators and makes use of the convolutional or Fourier network structure.

Although DeepONets have demonstrated good performance across diverse applications, some studies have pointed out that DeepONets fail to efficiently approximate solution operators of PDEs with sharp gradients or discontinuities [22,24]. In [22], the authors gave a fundamental lower bound on the approximation error of DeepONets and