Solving partial differential equations by a supervised learning technique, applied for the reaction–diffusion equation

Behzad Zakeri¹ · Morteza Khashehchi² · Sanaz Samsam³ · Atoosa Tayebi⁴ · Atefeh Rezaei⁵

Received: 17 September 2019 / Accepted: 2 November 2019 / Published online: 9 November 2019
© Springer Nature Switzerland AG 2019

Abstract

Deep learning is a crucial point of valuable intelligence resources to deal with complicated mathematical problems. The effectiveness of deep learning in solving differential equations has been considered over the past few years. Supervision of the learning process requires significant information to be marked in order to train the network. Nevertheless, this approach could not be a helpful strategy in case of unknown differential equations that we have no identified data. In order to address this problem, a new method for solving differential equations will be introduced in this paper using only the boundary and initial conditions. As an efficient method, inadequate monitoring can provide an ideal bed to fix boundary and initial value issues. For verification of the proposed method, a reaction–diffusion equation was performed. This equation has a variety of applications in engineering and science.

Keywords Deep learning · Weakly supervised learning · Reaction–diffusion equation · Danckwert’s transform

1 Introduction

Reaction–diffusion equation (RDE) is one of the well-known partial differential equations (PDEs) in engineering sciences as well as chemistry and finance [12]. RDE, by considering the simultaneous diffusion and dissipation of a property in the system gives an accurate model to predict the value of that property in the space-time domain. Coefficients, variables, boundary conditions type and dimensions of the RDE are highly dependent on the case which is studying, and based on the way that they are chosen, proper technique must be taken. All techniques aim to represent a precise solution ($C(X, t)$) of the RDE which describe diffusive property concentration in each time and position [21].

Although there are numerous numerical and analytical techniques to solve PDEs with different conditions, there

is no global method which can handle all kinds of PDEs. In recent years, by the explosive increment in available data, several heuristic approaches which all rely on the Machine Learning (ML) have been introduced for solving linear and non-linear PDEs [1, 2]. In all of them, models, by the utility of available data set tries to discovery the governing rule between initial/boundary conditions and the solution. Nevertheless, data-driven models can adequately learn the physics of the PDEs, but it requires the solution of the PDE before the solving process, which is not available in most cases.

In this work, weakly supervised learning was utilized to solve the one-dimensional RDE with constant coefficients and Dirichlet boundary condition. The benefit of this technique is that the network is trained whit only the boundary condition without providing any solution before the solving process. For the validation purpose, the analytical

✉ Morteza Khashehchi, M.khashehchi@ut.ac.ir; Behzad Zakeri, Behzad.Zakeri@ut.ac.ir; Sanaz Samsam, sanazsamsam@cmail.carleton.ca | ¹Department of Food Technology, College of Aburaihan, University of Tehran, Tehran, Iran. ²Department of Agro-Technology, College of Aburaihan, University of Tehran, Tehran, Iran. ³Department of Mechanical and Aerospace Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada. ⁴Department of Chemical Engineering, Amirkabir University of Technology, Tehran, Iran. ⁵Department of Chemical Engineering, Islamic Azad University, South Tehran Branch, Tehran, Iran.

solution of the corresponding problem has been provided by taking advantage of Danckwert's transform. Also, the dimensional analysis method was used to assess the function of the neural network in solving pure reaction and pure diffusion without any training process.

2 Related work

The present study is a connection between artificial intelligence and dynamical systems, which each of them has been conducted in a variety of research studies. The former subject is researched by data scientists and AI developers. The main objective of weakly supervised learning is to provide a general platform technique for learning algorithms to be able to learn at the training stage with minimal initial marked data. The latter topic is aiming for a reliable method for solving PDEs. The use of various mathematical techniques for the estimation of the solution of PDEs is one of the essential aspects of this work.

Changing traditional numerical methods to alternative meshless approaches, such as machine learning, has become increasingly popular in recent years. Particularly in the event of problems with complex mathematical formulation, machine learning schemes are replaced by classical models [11]. Oquab et al. [13] have used a weakly supervised convolutional neural network to identify objects in image processing to reduce the number of input labelled images. This method was a general concept and has been used in various applications, such as automated identification, medical image analysis and differential equation resolution [3, 8, 18]. Sharma et al. [17] trained an encoder-decoder U-Net architecture, which was a completely convoluted neural network to solve a steady-state two-dimensional heat equation on a regular square. For this reason, weakly supervised learning techniques have been used to describe the proper convolutional kernel and loss function to train the network only by using the boundary conditions of the PDE, rather than having a large number of marked data sets [17]. Han et al. [6] have introduced a new approach to use deep learning to solve high-dimensional PDEs. In the form of backward stochastic differential equations (BSDEs), they reformulate the PDEs and then use deep learning to approximate the gradient of the solution. Though their method is effective in dealing with high-dimensional situations, this method's limitations justify looking for the comprehensive approach to resolving linear and low-dimensional PDEs [6, 19].

Then again, customary numerical strategies, for example, FDM and FVM, have been broadly created to handle various sorts of scientific issues which portrayal of a precise answer for them is not available [12]. For example, for the above case of the utilization of the response dispersion condition

in sulfate assault, Zuo et al. [22] have utilized the limited contrast technique to discover the fixation dissemination of sulfate particles in concrete. Additionally, expanded looks into have led on the use of AI in various designing fields, for example, handling with violent streams and control theory [5, 10, 20].

3 Physics

3.1 Reaction-diffusion equation

For explaining 1-D RDE, a primary line has been considered as a domain with Dirichlet boundary condition at the parts of the bargains. By doling out the self-assertive consistent to the dissemination coefficients, we are able to control the part of the fabric on the transport phenomena. Too, the response coefficient indicated the impact of the interaction between the diffusive substance and medium. In this recreation, a high concentration connected to the boundaries, and the point is modelling the engendering of that substance among the space.

The general form of the RDE in one-dimensional space is shown in Eq. 1. Also, the initial/boundary conditions are given in Eq. 2, and we want to determine $C(x, t)$, the concentration field in arbitrary time.

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} - RC \quad (1)$$

$$\begin{cases} C(0 < x < L, 0) = 0 \\ C(0, t) = C(L, t) = C_0 \end{cases} \quad (2)$$

where $D, R > 0$ are the diffusion coefficient and reaction rate between specified material and domain respectively.

The analytical procedures are not accessible in most cases. Be that as it may, by expecting the straight line as space and Dirichlet boundary condition, the expository arrangement of the RDE is appeared in taking after. Since convergence analysis within the neural networks is an impossible task [14], this arrangement plays a critical part within the approval of profound learning comes about.

3.2 Analytical solution

For the utility of Danckwert's method to represent the analytical solution and by considering Eq. 1 with constant coefficients (D and R), firstly, the equation must be solved without any source term (reaction term). Consequently, Eq. 1 reform to the Eq. 3 as follows:

$$\frac{\partial C_1}{\partial t} = D \frac{\partial^2 C_1}{\partial x^2} \quad (3)$$

where C_1 represents the solution of the RDE without any reactive term.

Let us attempt to find a nontrivial solution of (3) satisfying the boundary conditions (2) using separation of variables:

$$C_1(x, t) = X(x)T(t) \quad (4)$$

Substituting $C_1(x, t)$ back into Eq. 3 one obtains:

$$\frac{1}{D} \frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)} = -\lambda^2 \quad (5)$$

where λ is an arbitrary positive coefficient. The solution of the corresponding ODEs of Eq. 5 are:

$$T(t) = e^{-\lambda^2 Dt} \quad (6)$$

$$X(x) = A \sin(\lambda x) + B \cos(\lambda x) \quad (7)$$

leading to a of Eq. 3 of the form:

$$C_1(x, t) = (A \sin \lambda x + B \cos \lambda x) \exp(-\lambda^2 Dt) \quad (8)$$

where A and B are constants of integration. Since Eq. 3 is a linear equation, the most general solution is obtained by summing solutions of type Eq. 8, so that we have:

$$C_1(x, t) = \sum_{\alpha=1}^{\infty} (A_{\alpha} \sin \lambda_{\alpha} x + B_{\alpha} \cos \lambda_{\alpha} x) \exp(-\lambda_{\alpha}^2 Dt) \quad (9)$$

where A_{α} , B_{α} , and λ_{α} are determined by the initial and boundary conditions for any particular problem. The boundary conditions Eq. 2 demand that:

$$A_{\alpha} = 0, \quad \lambda_{\alpha} = \frac{\alpha\pi}{L} \quad (10)$$

By utility of Fourier series, the final solution of the Eq. 3 have been extracted as the following form:

$$C_1(x, t) = \frac{4C_0}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \exp\left(\frac{-D(2n+1)^2 \pi^2 t}{L^2}\right) \cos\left(\frac{(2n+1)\pi x}{L}\right) \quad (11)$$

Based on the Danckwert's transform [9], the solution of the Eq. 1 can be calculated by the following integral transform:

$$C(x, t) = k \int_0^t C_1 e^{-kt'} dt' + C_1 e^{-kt} \quad (12)$$

Finally, after the integration, the final solution of the 1-D RDE (Eq. 1) can be shown as follow:

$$C(x, t) = -\frac{4C_0}{\pi} \sum_{n=0}^{\infty} \left(a_n \cos(\omega_n x) \left(k \Psi_n \left(\exp\left(\frac{t}{\Psi_n}\right) - 1 \right) + \exp\left(\frac{t}{\Psi_n}\right) \right) k \right) + C_0 \quad (13)$$

where a_n , ω_n and Ψ_n are represented as follows:

$$a_n = \frac{(-1)^n}{(2n+1)}, \quad \omega_n = \frac{(2n+1)\pi}{2L}, \quad \Psi_n = \frac{4L^2}{(-D_e(2n+1)^2 \pi^2) - (4kL^2)} \quad (14)$$

3.3 Finite difference method

The finite-difference is considered as a powerful numerical strategy which is utilized to compute the precise solution of the PDEs in arbitrary domains. In this strategy, governing physical equations and space both discretized, and all equations solve iteratively on the discrete space. Considering the discretization of the domain 16 and time 17, the discretized shape of RDE for position m and time n would be:

$$\frac{C_m^{n+1} - C_m^n}{\Delta t} = D \frac{C_{m-1}^n - 2C_m^n + C_{m+1}^n}{\Delta x^2} - RC_m^n \quad (15)$$

$$\frac{x_L - x_0}{m} = \Delta x \quad (16)$$

$$\frac{t_{\infty} - t_0}{n} = \Delta t \quad (17)$$

where C_m^n is the concentration in time n and position m , also indices 0, L and ∞ represent the boundaries in time and positions, ends of the domain and last time step respectively.

With solving the Eq. 15 iteratively, the value of the C in each time and position converge to the correct value.

4 Deep learning solver

The point of this work is utilizing the deep neural network to solve the RDE with as it was utilizing boundary and initial conditions, without knowing the numerical or analytical solution or indeed having any labelled information. For this reason, the differential form of the equation has been decoded into a *physical-informed loss function*. This method makes a difference for us to discover the solution of the PDE without utilizing supervision within the frame of information.

To define the initial and boundary conditions of the problem into the deep neural network, we utilized a $n \times m$ matrix which its columns and rows represent the positions and time steps respectively. All of the matrix components for the input matrix are zero except the primary and final columns which their values represent to the boundary condition values (which in this case is C_0). Moreover, in this

matrix each row demonstrates the concentration distribution in a specified time-frame.

4.1 Deep learning architecture

A fully convolutional encoder–decoder organize within the shape of U-Net engineering has been utilized in this consider, as Ronneberger et al. [15] have utilized this design for biomedical picture segmentation. The most reason for choosing a fully convolutional design among other structures is the adaptability of this structure to fathom issues at different scales. The arrange contains a few encoding convolutional and translating pooling layers which spare the input matrix estimate amid the learning handle. At last, the output matrix gives the solution of the PDE within the discretized space-time space. The schematic structure of the network has appeared in Fig. 1.

As it appears in Fig. 1, each encoding layer has been connected to the corresponding decoder layer utilizing Fusion link. The reason for the utility of fusion link is to pass the boundary values of the input to the output layers, and by this procedure, the network is not constrained to memorize the structure of the input in its bottleneck layers. The number of layers in our design is self-assertive, and it is conceivable to include layers into the network as much as essential.

4.2 Kernel

To make an intelligent network that can solve the equation in any time and position, it is necessary to define the governing rule in that equation in a simple way for the neural network. It is similar to the method that FDM use for solving the discretized equation. In fact, by

discretization of a continues equation and transferring that equation into the algebraic form we can observe the governing rule for every point in space and time.

By reforming the Eq. 15, we can find the state of an arbitrary point in the space-time domain based on its neighbours as shown in Eq. 18:

$$C_m^{n+1} = C_m^n + B(C_{m-1}^n - 2C_m^n + C_{m+1}^n) - RC_m^n \Delta t \quad (18)$$

And B defined as follow:

$$B = \frac{D\Delta t}{\Delta x^2} \quad (19)$$

For transferring the relation among variables into the neural network, Eq. 18 have been decoded into the 3×3 convolutional kernel as follow:

$$\begin{pmatrix} -a & -b & -c \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (20)$$

where

$$a \rightarrow Bc \rightarrow Bb \rightarrow (1 - 2B - R\Delta t) \quad (21)$$

Discussed Kernel has been convolved into the across the input matrix, and the output matrix after normalization was used to calculate the Loss function:

$$\sum_{i,j} (\text{Conv2D}(\text{Kernel}, \text{Output})_{i,j})^2 \quad (22)$$

By minimizing the Eq. 22, the deep neural network tries to make its' solution closer to the real values which can be found in Eq. 18 and changing in boundary and initial conditions train the network for solving any type of problems governed by reaction–diffusion physics.

Fig. 1 Schematic diagram of the network architecture

5 Results

In this section, the solutions of the RDE which have been conducted by analytical and Deep Learning methods analyze and compare with previously validated solutions. The solution of the proposed equation has been represented by taking advantage of the numerical method (FDM) for determining the concentration of the sulfate ion in concrete. To have better judgment, boundary and initial condition of the demonstrated solution in this section have been chosen precisely similar to the Zuo et al. [22].

As it was raised in the deep learning section, the output of the U-Net network for specific input is a matrix which its columns and rows represent the position and time, and the value of each element demonstrates the concentration in i th time and j th position. On the other hand, the proposed analytical solution in Sect. 3.2 gives the continues field of solution. Figures 2 and 3 illustrate the contour of the analytical solution of the proposed equation.

Looking at Fig. 2 in more detail, the concentration diffusion along the time axis is apparent. Both the left and right edges of the demonstrated illustration represent the boundary conditions of the problem within this case are the same and equal to (C_0) . By moving through the y-axis (Time-axis) of the contour, the propagation of the concentration advance, although this phenomenon is less visible in this image and more evident in Fig. 3, because of the way the coefficients have chosen.

Figure 3 shows the contours of the RDE solution with different coefficients compares to Fig. 2. Reaction and Diffusion coefficients have chosen in a way that after some

Fig. 2 The contour of the concentration conducted by the analytical solution. ($D = \mathcal{O}(10^{-8})$, $R = \mathcal{O}(10^{-4})$)

Fig. 3 The contour of the concentration conducted by the analytical solution. ($D = \mathcal{O}(10^{-8})$, $R = \mathcal{O}(10^{-6})$)

time, the diffusion process fills all the domain rather than Fig. 2. In both Figs. 2 and 3, the reaction and diffusion process plays a significant role, and none of them can not be neglected.

One of the demanding features of Fig. 3 is the bell shape of the iso-contours of the concentration. The reason for this shape is the nature of the propagation in such cases which behave such as waves. Although the wave base solution of such systems is much developed for semi-infinite domains with pure diffusion, the nature of the wave behaviour of the RDE is evident in Fig. 3. Several

Fig. 4 2D comparison of deep learning with FDM solver in different times

(a) Deep Learning, $D_a = 1$ (b) Analytical, $D_a = 1$ (c) Deep Learning, $D_a = 10^4$ (d) Analytical, $D_a = 10^4$ (e) Deep Learning, $D_a = 10^{-4}$ (f) Analytical, $D_a = 10^{-4}$

Fig. 5 3D comparison of deep learning with analytical solution

Table 1 Accuracy analyze based on changing coefficients

<i>R</i> coefficient	<i>MSE</i> value
(a) $D = 2.7 \times 10^{-9}$	
2.25×10^{-2}	0.587
2.25×10^{-5}	0.495
2.25×10^{-7}	0.623
2.25×10^{-10}	0.341
<i>D</i> coefficient	<i>MSE</i> value
(b) $R = 2.25 \times 10^{-7}$	
2.7×10^{-2}	0.305
2.7×10^{-7}	0.576
2.7×10^{-8}	0.588
2.7×10^{-10}	0.534

cases of wave solution of such systems can be found in the mathematical literature of diffusion phenomena [4, 7, 16].

Figure 4 compares the deep learning solution with finite difference method comes about in terms of concentration propagation along with space at various times. With looking more accurately to Fig. 4, we watch that deep learning comes about to have high consistency with FDM comes about. However, it seems that deep learning can not accurately forecast the concentrations values in primary time-steps such as 1s. Many reasons can be named for the lack of deep learning in predicting the correct values in primary time-steps, but the most critical and compelling reason is the high gradient in this space-time.

In Fig. 5, 3D results of the reaction–diffusion solutions conducted by deep learning and analytical solution have been demonstrated. To have a better perspective of the physics of the reaction–diffusion process, the proportion of the reaction rate and diffusion coefficient have chosen meticulously. The proportion of reaction and diffusion coefficient can determine the physics of the system among pure diffusion, pure reaction, and simultaneous reaction–diffusion. In fact, by choosing the correct range of coefficient, the role of the arbitrary term in the equation (reaction or diffusion) can outweigh the other term. For this reason, a dimensionless coefficient has been characterized which offer assistance us to calculate the right extent of response and dissemination coefficients to have all three state of the arrangement in our computation.

Damköhler number is an important dimensionless parameter in chemical engineering which clarifies the role of diffusion, reaction or simultaneous reaction–diffusion phenomena in transport phenomena and define as follow:

$$D_a = \frac{\text{Rate of reaction}}{\text{Diffusion rate}} \quad (23)$$

In our model Eq. 1, Damköhler number is defined as:

$$D_a = \frac{RL^2}{D} \quad (24)$$

This number represents the states of reaction–diffusion in different states where $D_a \cong 1$, $D_a \gg 1$, and $D_a \ll 1$ mean the physics of *Reaction–Diffusion*, *pure Reaction*, and *pure Diffusion* respectively.

The Mean Square Error (MSE) index has been utilized to calculate the deep learning faults in predicting the correct values of the equation. It has been observed that the final value of the concentration in space-time is dependent on the reaction and diffusion coefficients. However, this dependency is not as much as affect the accuracy of the final results in a way that they become unreliable.

To have a quantitative assessment of deep learning solution, we assumed one of the coefficients constant, and by changing the other coefficient MSE value has been computed, and the result of this analysis is reported in Table 1.

6 Conclusion

In this paper, the capability of weakly supervised learning in comprehending the transient one-dimensional reaction–diffusion equation has been studied. Also, an analytical solution for the RDE based on the separation of variable technique and the utility of Danchwert's transform has proposed.

It appeared that the results conducted by deep learning method have grate consistency with analytical and numerical results. Moreover, it was observed that the values of the reaction and diffusion coefficients could cause the miss estimation by deep learning. Although these noises

were not in such a level that manipulates our results in this case, it could be the source of destructive faults in other problems like BSDEs.

Finally, it is worth emphasizing that weakly supervised learning, cloud successfully tackle the lack of sufficient labelled data to learn the physics of the governing equations. Furthermore, this method can be considered for complex problems with limited labelled data and complex governing equation.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Berg J, Nyström K (2018) A unified deep artificial neural network approach to partial differential equations in complex geometries. *Neurocomputing* 317:28–41
- Berg J, Nyström K (2019) Data-driven discovery of pdes in complex datasets. *J Comput Phys* 384:239–252
- Ciampi F, de Hoop B, van Riel SJ, Chung K, Scholten ET, Oudkerk M, de Jong PA, Prokop M, van Ginneken B (2015) Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2d views and a convolutional neural network out-of-the-box. *Med Image Anal* 26(1):195–202
- Crank J et al (1979) *The mathematics of diffusion*. Oxford University Press, Oxford
- Guo X, Yan W, Cui R (2019) Integral reinforcement learning-based adaptive NN control for continuous-time nonlinear MIMO systems with unknown control directions. *IEEE Trans Syst Man Cybern Syst*. <https://doi.org/10.1109/TSMC.2019.2897221>
- Han J, Jentzen A, Weinan E (2018) Solving high-dimensional partial differential equations using deep learning. *Proc Natl Acad Sci* 115(34):8505–8510
- Kuttler C (2011) Reaction–diffusion equations with applications. In: Internet seminar
- Liu P, Gan J, Chakrabarty RK (2018) Variational autoencoding the Lagrangian trajectories of particles in a combustion system. *arXiv preprint* [arXiv:1811.11896](https://arxiv.org/abs/1811.11896)
- McNabb A (1993) A generalized Danckwerts transformation. *Eur J Appl Math* 4(2):189–204
- Mohan AT, Gaitonde DV (2018) A deep learning based approach to reduced order modeling for turbulent flow control using LSTM neural networks. *arXiv preprint* [arXiv:1804.09269](https://arxiv.org/abs/1804.09269)
- Monsefi AK, Zakeri B, Samsam S, Khashehchi M (2019) Performing software test oracle based on deep neural network with fuzzy inference system. In: Grandinetti L, Mirtaheri SL, Shahbazian R (eds) *High-performance computing and big data analysis*. Springer, Cham, pp 406–417
- Morton KW, Mayers DF (2005) *Numerical solution of partial differential equations: an introduction*. Cambridge University Press, Cambridge
- Oquab M, Bottou L, Laptev I, Sivic J (2015) Is object localization for free?-weakly-supervised learning with convolutional neural networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 685–694
- Raissi M, Perdikaris P, Karniadakis GE (2019) Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys* 378:686–707
- Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: *International conference on medical image computing and computer-assisted intervention*. Springer, Berlin, pp 234–241
- Schell KG, Fett T, Bucharsky EC (2019) Diffusion equation under swelling stresses. *SN Appl Sci* 1(10):1300. <https://doi.org/10.1007/s42452-019-1343-1>
- Sharma R, Farimani AB, Gomes J, Eastman P, Pande V (2018) Weakly-supervised deep learning of heat transport via physics informed loss. *arXiv preprint* [arXiv:1807.11374](https://arxiv.org/abs/1807.11374)
- Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J (2016) Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans Med Imaging* 35(5):1299–1312
- Weinan E, Han J, Jentzen A (2017) Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Commun Math Stat* 5(4):349–380
- Wu JL, Xiao H, Paterson E (2018) Physics-informed machine learning approach for augmenting turbulence models: a comprehensive framework. *Phys Rev Fluids* 3(7):074602
- Zakeri B, Monsefi AK, Samsam S, Monsefi BK (2019) Weakly supervised learning technique for solving partial differential equations; case study of 1-d reaction–diffusion equation. In: Grandinetti L, Mirtaheri SL, Shahbazian R (eds) *High-performance computing and big data analysis*. Springer, Cham, pp 367–377
- Zuo XB, Sun W, Yu C (2012) Numerical investigation on expansive volume strain in concrete subjected to sulfate attack. *Constr Build Mater* 36:404–410

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.