

Connections between deep learning and partial differential equations

M. BURGER¹, W. E², L. RUTHOTTO³ and S. J. OSHER⁴

¹*Department Mathematik, Friedrich-Alexander Universität Erlangen-Nürnberg, Cauerstrasse 11,
91058 Erlangen, Germany
email: martin.burger@fau.de*

²*Princeton University, Department of Mathematics, Princeton, NJ 08544-1000, USA
email: weinan@math.princeton.edu*

³*Emory University, Mathematics and Computer Science, 400 Dowman Drive, Atlanta, GA 30322, USA
email: lruthotto@emory.edu*

⁴*UCLA, Department of Mathematics, 520 Portola Plaza, Los Angeles, CA 90095, USA
email: sjo@math.ucla.edu*

(Received 17 March 2021; revised 17 March 2021; accepted 17 March 2021)

The last years have seen a resurgence of interest in machine learning based on (deep) neural networks driven mainly by increasing computational resources and the availability of huge data sets for learning. Aside from their use in main stream data science applications, deep networks (i.e., neural networks with many hidden layers) have led to new techniques for solving partial differential equations (PDEs), particularly in high-dimensional settings. At the same time, the interpretation of some deep neural networks as nonlinear (partial) differential equations has led to a new frontier to gain theoretical insight and design new algorithms for deep learning. Motivated by these trends, this special issue brings together new results at this new interface between applied mathematics and data science. The papers in this special issue provide different insights into the connection between deep learning and PDEs. On the one hand, they discuss the use of deep learning methods for typical tasks in PDEs; on the other hand, they employ PDE techniques to design novel classification schemes or to understand their continuum limits.

Deep learning and numerical solution of PDEs is a field that gained quite some early interest. Training deep neural networks to lower the computational cost of simulating nonlinear PDEs has potential applications in many-query problems such as parameter estimation and uncertainty quantification. In contrast to data science applications, in this context, one can control the training data's accuracy and combine many relatively cheap low-fidelity samples with a few high-fidelity samples. The theoretical arguments and numerical evidence in [5] show that such a multi-level procedure can lower the generalisation error considerably. Becker et al [1] tackle optimal stopping problems for financial derivatives such as American or Bermudan options. A deep neural network predicts the stopping strategy to mitigate the curse of dimensionality when thousands of underlying assets are considered. They show the accuracy of their approach by comparing numerical results to reference values from the literature.

Linearising nonlinear PDEs through coordinate transformations such as the Cole–Hopf transformation or Inverse Scattering Transform is key to simulate, control or estimate nonlinear PDE models. However, finding such a transformation analytically for an arbitrary nonlinear PDE is virtually impossible. To linearise a wide class of PDEs, Gin et al [3] propose an autoencoder architecture. They provide promising results for several examples, including the Kuramoto–Sivashinsky equation. Another relevant topic in PDEs is quantification of uncertainty in presence of random effects. In Khoo et al’s paper [4], the authors use convolutional neural networks to map random parameters of the PDE to physical quantities of interest. Their theoretical motivation links forward propagation to time evolution. In numerical experiments on a diffusion equation and a nonlinear Schrodinger equation, they demonstrate that supervised training of the network can yield accurate surrogates. Chen et al’s paper [2] tackles forward and inverse problems with physics informed neural networks (PINN). In the numerical examples, the PINN approach accurately estimates both random and deterministic parameters of the system from a small number of measurements. They also demonstrate that Bayesian optimisation can automatise the hyper-parameter tuning in PINNs.

As mentioned above, ideas from PDEs increasingly find their way into learning schemes. In their paper, Wang and Osher replace the data-agnostic softmax function with graph-based interpolation to improve deep neural network classifiers’ accuracy and robustness. They show that this choice, in the continuum limit, converges to the Laplace–Beltrami equation on high-dimensional manifolds. Therefore, the work outlines new ways to combine advances in deep neural nets and manifold learning. Savarino and Schnörr [6] propose a novel parameterisation of the assignment flow in a continuous-domain setting. Their method provides a PDE-based approach for solving classification problems on graphs using a sequence of linear elliptic PDEs.

References

- [1] BECKER, S., CHERIDITO P., JENTZEN, A. & WELTI, T. (2021) Solving high-dimensional optimal stopping problems using deep learning. *European Journal of Applied Mathematics*, **32**, 470–514.
- [2] CHEN, X., DUAN J. & KARNIADAKIS, G. E. (2021) Learning and meta-learning of stochastic advection-diffusion-reaction systems from sparse measurements. *European Journal of Applied Mathematics*, **32**, 397–420.
- [3] GIN, C., LUSCH, B., BRUNTON, S. C. & KUTZ, J. N. (2021) Deep learning models for global coordinate transformations that linearize PDEs. *European Journal of Applied Mathematics*, **32**, 515–539.
- [4] KHOO, Y., LU, J. & YING, L. (2021) Solving parametric PDE problems with artificial neural networks. *European Journal of Applied Mathematics*, **32**, 421–435.
- [5] LYE, K. O., MISHRA, S. & MOLINARO, R. (2021) A Multi-level procedure for enhancing accuracy of machine learning algorithms. *European Journal of Applied Mathematics*, **32**, 436–469.
- [6] SAVARINO, F. & SCHNÖRR, C. (2021) Continuous-domain assignment flows. *European Journal of Applied Mathematics*, **32**, 570–597.
- [7] WANG, B. & OSHER, S. J. (2021) Graph interpolating activation improves both natural and robust accuracies in data-efficient deep learning. *European Journal of Applied Mathematics*, **32**, 540–569.