NUS National University of Singapore | Department of Mathematics Faculty of Science

Research Highlight: Deep Learning from A Dynamical Viewpoint

Work of Assistant Professor Li Qianxiao

NUS mathematicians have developed a new theoretical framework based on dynamical systems to understand when and how a deep neural network can learn arbitrary relationships.

Despite achieving widespread practical success, understanding the theoretical principles of deep learning remains a challenging task. One of the most fundamental questions is: can deep neural networks learn arbitrary input-output relationships (in mathematics, these will be called functions), and does the way they achieve this differ from traditional methodologies?

To understand this question, it is necessary to think about what exactly is new in deep neural networks compared to traditional function approximation paradigms. For example, classical Fourier series approximates complicated functions as a weighted sum of simpler functions, such as sines and cosines. Deep neural networks operate quite differently. Instead of weighted sums, they build complex functions out of repeated stacking of simple functions (layers). This is also known as function composition in mathematics. The key question is on how complicated functions can be built out of simple ones by stacking them together. It turns out that this is quite a new problem in the branch of mathematics known as approximation theory.

In this study, Assistant Professor Qianxiao LI from the Department of Mathematics, National University of Singapore and his collaborators developed a new theory of approximation capabilities of function composition. An interesting observation is that while function composition is challenging to analyse in practice due to its discrete and nonlinear structure, this is not the first time that there are such problems. In the study of the motion of solids and fluids, they are often idealized as a continuum of particles, satisfying some continuous equations (ordinary or partial differential equations). This allows the difficulty of modelling such systems at the discrete, atomic level to be side-stepped. Instead, continuum equations that model their behaviour at the macroscopic level were derived. The key idea in the study is that this concept can be extended to deep neural networks, by idealizing the layered structure as a continuous dynamical system (See Figure 1). This connects deep learning with the branch of mathematics known as dynamical systems. Such connections allow for the development of new tools to understand the mathematics of deep learning, including a general characterisation of when it can approximate arbitrary relationships.

Prof Li said, “The dynamical systems viewpoint of deep learning offers a promising mathematical framework that highlights the distinguishing aspects of deep neural networks compared with traditional paradigms. This brings exciting new mathematical problems on the interface of dynamical systems, approximation theory and machine learning.”

“A promising area of future development is to extend this framework to study other aspects of deep neural networks, such as how to train them effectively and how to ensure they work better on unseen datasets,” added Prof Li.

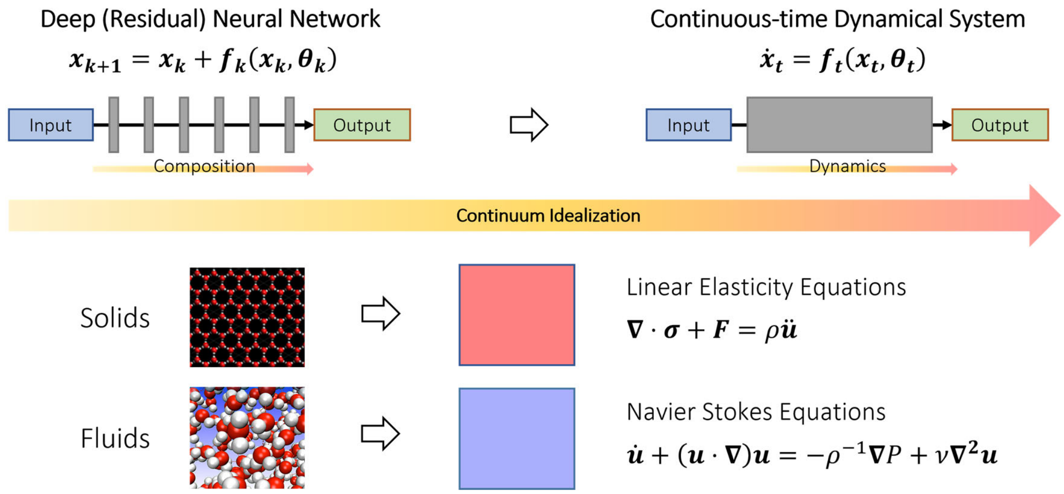


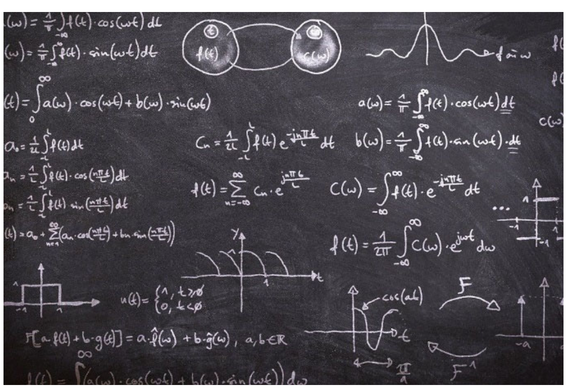
Figure 1: The continuous viewpoint of deep learning. Similar to how scientists study solid and fluid mechanics in the continuum limit, one can also idealize deep neural networks as a discretization of a continuous dynamical system. In the schematic, the fictitious time parameter represents a continuous analogue of layers and the dynamics model layer stacking.

Reference

Li Q*; Lin T*; Shen Z*, "Deep learning via dynamical systems: An approximation perspective", JOURNAL OF THE EUROPEAN MATHEMATICAL SOCIETY, DOI: 10.4171/JEMS/1221, Published: 2022.

[Internal reference]

Cover Image



[obtained from pixabay: https://pixabay.com/photos/board-mathematics-school-to-learn-4855963/]