Computational aspects of deep neural networks

List of authors:Eldad Haber and Lars Ruthotto

E. Haber [1]

Partial differential equations (PDEs) are indispensable for modeling many physical phenomena and also commonly used for solving image processing tasks. In the latter area, PDE-based approaches interpret image data as discretizations of multivariate functions and the output of image processing algorithms as solutions to certain PDEs. Posing image processing problems in the infinite dimensional setting provides powerful tools for their analysis and solution. Over the last three decades, the reinterpretation of classical image processing tasks through the PDE lens has been creating multiple celebrated approaches that benefit a vast area of tasks including image segmentation, denoising, registration, and reconstruction. In this work, we establish a new PDE-interpretation of deep convolution neural networks (CNN) that are commonly used for learning tasks involving speech, image, and video data. Our interpretation includes convolution residual neural networks (ResNet), which are among the most promising approaches for tasks such as image classification having improved the state-of-the-art performance in prestigious benchmark challenges. Despite their recent successes, deep ResNets still face some critical challenges associated with their design, immense computational costs and memory requirements, and lack of understanding of their reasoning.

Guided by well-established PDE theory, we derive three new ResNet architectures that fall two new classes: parabolic and hyperbolic CNNs. We demonstrate how PDE theory can provide new insights and algorithms for deep learning and demonstrate the competitiveness of three new CNN architectures using numerical experiments.

[1] Department of Earth and Ocean Science