Empirical Evaluation of Gradient-Based Training Algorithms for Ordinary Differential Equation Networks

Authors : Martin K. Steiger, Lukas Heisler, Hans-Georg Brachtendorf

Abstract : Deep neural networks and their variants form the backbone of many AI applications. Based on the so-called residual networks, a continuous formulation of such models as ordinary differential equations (ODEs) has proven advantageous since different techniques may be applied that significantly increase the learning speed and enable controlled trade-offs with the resulting error at the same time. For the evaluation of such models, high-performance numerical differential equation solvers are used, which also provide the gradients required for training. However, whether classical gradient-based methods are even applicable or which one yields the best results has not been discussed yet. This paper aims to redeem this situation by providing empirical results for different applications.