

An Optimal Time Variable Learning Framework for Deep Neural Networks

Evelyn Herberg¹ Harbir Antil² Hugo Diaz³

Feature propagation in Deep Neural Networks (DNNs) can be associated to nonlinear discrete dynamical systems. The novelty, in this paper, lies in letting the discretization parameter (time step-size) vary from layer to layer, which needs to be learned, in an optimization framework. The proposed framework can be applied to any of the existing networks such as ResNet, DenseNet or Fractional-DNN. This framework is shown to help overcome the vanishing and exploding gradient issues. Stability of some of the existing continuous DNNs such as Fractional-DNN is also studied. The proposed approach is applied to an ill-posed 3D-Maxwell's equation.

References:

[1] <https://arxiv.org/abs/2204.08528>

¹George Mason University, Mathematical Sciences
evelyn@ib-herberg.de

²George Mason University, Mathematical Sciences
hantil@gmu.edu

³University of Delaware, Department of Mathematical Sciences
hugodiaz@udel.edu