University of Fribourg / Faculty of Science and Medicine / Department of Computer Science

Novel Deep Neural Networks Initialisations Techniques

Learning from Data to Better Learn from Data

Michele Alberti

The widespread success of Deep Learning (DL) applications is generally accompanied with an important resource investment. The aim of this study is to investigate the use of novel Deep Neural Networks (DNNs) initialisation techniques to reduce these costs. Recent studies show that a careful initialisation is vital for success in training a DNN, however, the landscape of their initialisation is not sufficiently understood and thus offers inaccurate and limited guidance to practitioners. This work examines novel initialisation methods in two different directions. First, it presents layer-wise data-driven approaches which leverage deterministic and optimal algorithms, such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Second, it proposes adapting the Artificial Neural Network (NN) architecture to support the computation of matrix trans-forms, such as spectral ones. To demonstrate the impact of our approach the performances of DNN initialised with our approaches have been measured and com-pared against the traditional random initialisation on a variety of datasets and with a thorough hyper-parameters search setup. It was observed that networks initialised with data-driven approaches are characterised by a faster convergence, are less subject to stochastic effects, and denoted a higher accuracy for the task of image classification. Moreover, introducing modules into the networks architecture which can support signal-processing operations (such as spectral trans-form) resulted in competitive performances overall and significantly higher performances on selected datasets, both in terms of accuracy and training speed. Besides these scientific contributions, we realized three publicly available technical frameworks (N-light-N, DeepDIVA, and GALE) for the research community. These frameworks have been adopted by both, researchers and industry alike and contribute towards the reproducibility issue and establishing good programming practices in the field. This study provides compelling evidence that there are better alter-natives to random initialisation; and contributes to close the knowledge gap on DNN initialisation techniques. Therefore, in good agreement with recent advances in the literature, this study demonstrates that through a careful initialisation, it is possible to mitigate the high costs associated with the development of DL applications.

Jury:

Prof. Rolf Ingold (thesis supervisor)
Prof. Marcus Liwicki (thesis supervisor)
Prof. Seiichi Uchida (external co-examiner)
Dr. Ashok Popat (external co-examiner)
Prof. Philippe Cudre-Mauroux (president of the jury)