TRIANGULAR DROPOUT: VARIABLE NETWORK WIDTH WITHOUT RETRAINING

Anonymous authors
Paper under double-blind review

ABSTRACT

One of the most fundamental design choices in neural networks is layer width: it affects the capacity of what a network can learn and determines the complexity of the solution. This latter property is often exploited when introducing information bottlenecks, forcing a network to learn compressed representations. However, such an architecture decision is typically immutable once training begins; switching to a more compressed architecture requires retraining. In this paper we present a new layer design, called Triangular Dropout, which does not have this limitation. After training, the layer can be arbitrarily reduced in width to exchange performance for narrowness. We demonstrate the construction and potential use cases of such a mechanism in three areas. Firstly, we describe the formulation of Triangular Dropout in autoencoders, creating an MNIST autoencoder with selectable compression after training. Secondly, we add Triangular Dropout to VGG19 on ImageNet, creating a powerful network which, without retraining, can be significantly reduced in parameters with only small changes to classification accuracy. Lastly, we explore the application of Triangular Dropout to reinforcement learning (RL) policies on selected control problems, showing that it can be used to characterize the complexity of RL tasks, a critical measurement in multitask learning and meta-learning domains.