Decoupled Neural Interfaces Using Synthetic Gradients

Qilei Zhang

Jun 12 2018

Abstract

Training directed neural networks typically requires forward-propagating data through a computation graph, followed by backpropagating error signal, to produce weight updates.

1. Introduction

Each layer (or module) in a directed neural network can be considered a computation step, that transforms its incoming data. These modules are connected via directed edges, creating a forward processing graph which defines the flow of data from the network inputs, through each module, producing network outputs [1].

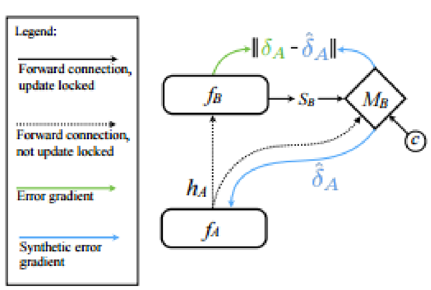

Figure 1. Transfer ratios on the Amazon benchmark. Both SDA-based systems outperforms the rest, and SDAsh (unsupervised training on all domains) is best.Reproduced from Glorot et al. (2011b).

$$x'(t) = -V'(x) + A_0 cos(wt + o) + u(t) \tag{1}$$

2. Decoupled Neural Interfaces

We begin by describing the high-level communication protocol that is used to allow asynchronously learning agents to communicate [2].

3. Synthetic Gradient for Recurrent Networks

We begin by describing how our method of using synthetic gradients applies in the case of recurrent networks; in some ways this is simpler to reason about than feed-forward networks or more general graphs [3].

4. Arbitrary Network Graphs

Although we have explicitly described the application of DNIs for communication between layers in feed-forward networks [4], and between recurrent cores in recurrent networks, there is nothing to restrict the use of DNIs for arbitrary network graphs [5].

References

[1] C. Barat and C. Ducottet. String representations and distances in deep convolutional neural networks for image classification. *Pattern Recognition*, 54(1):104–115, 2016. 1

[2] C. D. Eiber, N. H. Lovell, and G. J. Suaning. Attaining higher resolution visual prosthetics: a review of the factors and limitations. *Journal of Neural Engineering*, 10(1):1102, 2013. 1

[3] J. X. Liu, L. I. Ying-Hong, Y. G. Chen, and X. K. Wei. Aero-engine separate decoupling control based on recursive neural networks. *Journal of Aerospace Power*, 20(2):287–292, 2005. 1

[4] A. Mahendran and A. Vedaldi. Visualizing deep convolutional neural networks using natural pre-images. *International Journal of Computer Vision*, 120(4):1–23, 2015. 1

[5] L. I. Yi and J. H. Gong. Effective global image delivery and dispatch strategies in net environment. *Science of Surveying*, 35(4):52–53, 2010. 1