

Multirate Training of Neural Networks

Tiffany Vlaar @ (University of Edinburgh), Benedict Leimkuhler
R 3.07 Tue Z1 17:30-18:00

We propose multirate training of neural networks: partitioning neural network parameters into fast and slow parts which are trained simultaneously using different learning rates. By choosing appropriate partitionings we can obtain large computational speed-ups for transfer learning tasks. We show that for various transfer learning applications in vision and NLP we can fine-tune deep neural networks in almost half the time, without reducing the generalization performance of the resulting model. We also discuss other splitting choices for the neural network parameters which are beneficial in enhancing generalization performance in settings where neural networks are trained from scratch. Finally, we propose an additional multirate technique which can learn different features present in the data by training the full network on different time scales simultaneously. The benefits of using this approach are illustrated for ResNet architectures on image data. Our paper unlocks the potential of using multirate techniques for neural network training and provides many starting points for future work in this area.