Making L-BFGS Work with Industrial-Strength Nets

Abhay Yadav
jaiabhay@cs.umd.edu
Tom Goldstein
tomg@cs.umd.edu
David Jacobs
djacobs@cs.umd.edu

University of Maryland
College Park,
Maryland, USA

Supplementary Material

.1 Experiments

.1.1 STL10 [1]

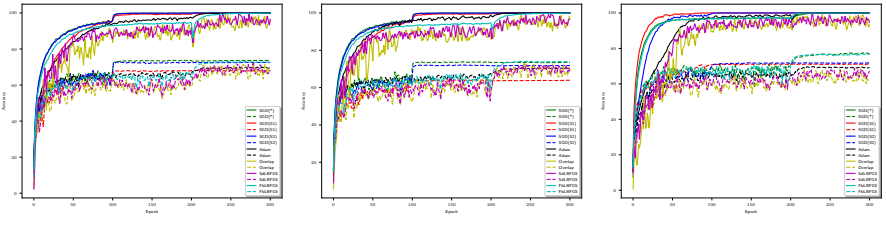

Figure 1: Overview of the performance of STL-10 on ResNet, DenseNet, and Wide ResNet respectively. The solid lines represent train accuracy and dashed lines represent the test accuracy, respectively.

.1.2 CIFAR-10 [3]

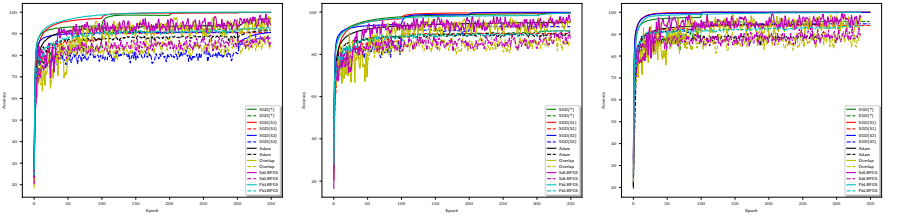

Figure 2: Overview of the performance of CIFAR-10 on ResNet, DenseNet, and Wide ResNet respectively. The solid lines represent train accuracy and dashed lines represent the test accuracy, respectively.

.1.3 CIFAR-100 [2]

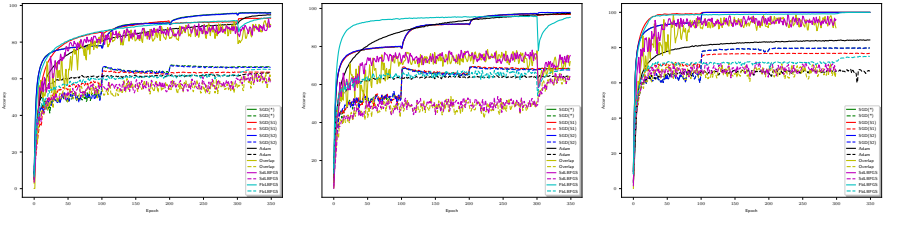

Figure 3: Overview of the performance of CIFAR-100 on ResNet, DenseNet, and Wide ResNet respectively. The solid lines represent train accuracy and dashed lines represent the test accuracy, respectively.

References

[1] Adam Coates, Andrew Ng, and Honglak Lee. An analysis of single-layer networks in unsupervised feature learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 215–223, 2011.

[2] Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton. Cifar-10 and cifar-100 datasets. *URl: https://www. cs. toronto. edu/kriz/cifar. html*, 6, 2009.

[3] Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton. The cifar-10 dataset. *online: http://www. cs. toronto. edu/kriz/cifar. html*, 55, 2014.