Convexified Convolutional Neural Networks

Kyle Daruwalla and Akhil Sundararajan

ECE 901 Fall 2016

November 4, 2016

Overview

Paper Overview

1. Start generic two-layer CNN
2. Convex relaxation
   - 2.1 Linear activation – optimize for a low-rank matrix A instead of filter weights and coefficients
   - 2.2 Non-linear activation – frame problem in terms of RKHS
3. Introduce a kernel-based algorithm for CCNNs
4. Provide theoretical guarantees on the generalization error
5. Explain extensions like pooling and multi-layer CNNs
6. Provide experimental results on MNIST and CIFAR-10

Convolutional Neural Networks

Setting up CNN problem in context of paper. (two slides)

Optimizing For Low-Rank Matrix (Linear Activation)

Explain the process (show figure) and emphasize optimization over nuclear norm ball

Reproducing-Kernel Hilbert Space (RKHS)

Background on RKHS and representer theorem

Framing Problem Using RKHS

Framing problem using RKHS

CCNN Algorithm

Introduce algorithm

Solving ERM in CCNN Algorithm

Details on solving minimization problem embedded in algorithm

Bounds on Generalization Error

Theorem 1

Proof of Theorem 1

Proof of theorem 1

Experimental Results

State results

References

Y. Zhang, et al. (2016, Sept. 4). *Convexified Convolutional Neural Networks* (v1) [Online]. Available: `https://arxiv.org/abs/1609.01000`

The End