Noise-enhanced convolutional neural networks

Kartik Audhkhasi, Osonde Osoba, Bart Kosko

PII: S0893-6080(15)00189-6
DOI: http://dx.doi.org/10.1016/j.neunet.2015.09.014
Reference: NN 3538

To appear in: *Neural Networks*

Please cite this article as: Audhkhasi, K., Osoba, O., & Kosko, B. Noise-enhanced convolutional neural networks. *Neural Networks* (2015), http://dx.doi.org/10.1016/j.neunet.2015.09.014

NOISE-ENHANCED CONVOLUTIONAL NEURAL NETWORKS

Kartik Audhkhasi, Osonde Osoba, Bart Kosko

Signal and Information Processing Institute
Electrical Engineering Department
University of Southern California, Los Angeles, CA
Email: kosko@sipi.usc.edu

ABSTRACT

Injecting carefully chosen noise can speed convergence in the backpropagation training of a convolutional neural network (CNN). The Noisy CNN algorithm speeds training on average because the backpropagation algorithm is a special case of the expectation-maximization (EM) algorithm and because such carefully chosen noise always speeds up the EM algorithm on average. The CNN framework gives a practical way to learn and recognize images because backpropagation scales with training data. It has only linear time complexity in the number of training samples. The Noisy CNN algorithm finds a special separating hyperplane in the network's noise space. The hyperplane arises from the likelihood-based positivity condition that noise-boosts the EM algorithm. The hyperplane cuts through a uniform-noise hypercube or Gaussian ball in the noise space depending on the type of noise used. Noise chosen from above the hyperplane speeds training on average. Noise chosen from below slows it on average. The algorithm can inject noise anywhere in the multilayered network. Adding noise to the output neurons reduced the average per-iteration training-set cross entropy by 39% on a standard MNIST image test set of handwritten digits. It also reduced the average per-iteration training-set classification error by 47%. Adding noise to the hidden layers can also reduce these performance measures. The noise benefit is most pronounced for smaller data sets because the largest EM hill-climbing gains tend to occur in the first few iterations. This noise effect can assist random sampling from large data sets because it allows a smaller random sample to give the same or better performance than a noiseless sample gives.

Index Terms— backpropagation, noise benefit, noise injection, convolutional neural network, expectation-maximization algorithm, stochastic resonance, big data, sampling from big data sets, deep learning

1. NOISE-BOOSTED CONVOLUTIONAL NEURAL NETWORKS

This paper presents the Noisy Convolutional Neural Network (NCNN) algorithm for speeding up the backpropagation (BP) training of convolutional neural networks (CNNs). Figure 1 shows the architecture of a CNN with a single hidden convolutional layer and 3 convolutional masks or retinal-like receptive fields. CNNs are standard feedfoward neural networks for large-scale image recognition [1, 2, 3, 4, 5, 6, 7, 8]. Some deep neural nets use on the order of 20 hidden layers of neurons [9, 10]. The NCNN algorithm can noise-boost a CNN with any number of hidden layers so long as the injected noise lies above the NCNN hyperplane in noise space.

The NCNN algorithm exploits two recent theoretical results–a reduction and a noise boost.