Abstract

Institut für Neuroinformatik
International Graduate School of Neuroscience
Philosophiae Doctoris (PhD) in Neuroscience

Learning Natural Image Statistics with Variants of Restricted Boltzmann Machines

by Nan Wang

Motivated by modeling the visual system in the brain, this thesis focuses on modeling natural image statistics with a class of causal models that are variants of Restricted Boltzmann Machines (RBMs). Concretely, we consider Gaussian RBMs, covariance RBMs, mean-covariance RBMs, Gaussian deep belief networks, and Gaussian deep Boltzmann machines.

From the perspective of causal models, $p(observations, causes)$ is the probability distribution defined in the models that describes the causal relationships between the observations and the underlying causes. In the first part of this work, we examine the heuristics in each model by decomposing $p(observations, causes)$ as a product between $p(observations|causes)$ and $p(causes)$. In this way, we show that the variants of RBMs construct $p(observations)$ with a weighted sum of Gaussian distributions. This perspective enables us to further demonstrate the constraints of these models’ heuristics in modeling distributions of observations.

Second, we train the models on natural image patches to examine their performances in modeling natural image statistics. After training, on the one hand, we analyze the first-layer weights as linear filters and examine the second-layer units’ weight connections to the filters. On the other hand, we look into the locations and covariance matrices of the models’ components from the perspective of causal models. The results suggest that all the models learn meaningful parameters from natural image patches.
Furthermore, we synthesize image patches and compare the statistics of the synthesized patches with natural image statistics. In reduced space and pixel space, the statistics of the mean-covariance RBM match natural image statistics best. The Gaussian deep Boltzmann machine takes the second place. In Haar wavelet space, although all the models are able to capture some important statistical features, none of them fully demonstrates the properties of the natural image patches. We suggest the models’ failures of fully capturing nature image statistics are due to the inappropriate constraints of the heuristics.

Finally, we relate the samples generated from the models’ prior probability to the spontaneous activity in the primary visual cortex; we examine whether the hidden units in these samples show the correlations between different neurons in the spontaneous activity, which are reported by Kenet et al. [2003]. In our computational experiment, only the Gaussian deep Boltzmann machine trained on the natural image patches demonstrates similar properties as the spontaneous activity. Our results suggest Gaussian deep Boltzmann machine is a meaningful model for basic receptive field properties and the emergence of spontaneous activity in early visual cortex.