

# Super Resolution

Qingyun Li

July 6, 2018

## 1. Introduction

The framework can yield specific training algorithms for many kinds of model and optimization algorithm. In the article, the authors explore the special case when the generative model generates samples by passing random noise through a multilayer perceptron, and this discriminative model is also a multilayer perceptron. We refer to this special case as adversarial nets. In this case, we can train both models and using only the highly successful backpropagation and dropout algorithms [3] and sample from the generative model using only forward propagation. No approximate inference or Markov chains are necessary.

## 2. Related work

An alternative to directed graphical models with latent variables are undirected graphical models with latent variables, such as restricted Boltzmann machines (RBMs) [6], deep Boltzmann machines (DBMs) [5] and their numerous variants. The interactions within such models are represented as the product of unnormalized potential functions, normalized by a global summation over all states of the random variables. This quantity and its gradient are intractable for all but the most trivial instances, although they can be estimated by Markov chain Monte Carlo (MCMC) methods<sup>1</sup>. Mixing poses a significant problem for learning algorithms that rely on MCMC [1].

Deep belief networks (DBNs) are hybrid models con-

taining a single undirected layer and several directed layers. While a fast approximate layer-wise training criterion exists, DBNs incur the computational difficulties associated with both undirected and directed models.

Alternative criteria that do not approximate or bound the log-likelihood have also been proposed, such as score matching [4] and noise-contrastive estimation (NCE) [2].

## References

- [1] Y. Bengio, G. Mesnil, Y. Dauphin, and S. Rifai. Better mixing via deep representations. In *ICML*, 2013. 1
- [2] M. Gutmann and A. Hyvärinen. Noise-contrastive estimation: A new estimation principle for unnormalized statistical models. *JMLR*, pages 297–304, 2010. 1
- [3] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. *Computer Science*, pages 212–223, 2012. 1
- [4] A. Hyvärinen. Estimation of Non-Normalized Statistical Models by Score Matching. *JMLR*, pages 695–709, 2005. 1
- [5] R. Salakhutdinov and H. Larochelle. Efficient learning of deep boltzmann machines. In *AISTATS*, 2010. 1
- [6] P. Smolensky. Information processing in dynamical systems: Foundations of harmony theory. Technical report, Computer Science, 1986. 1

---

<sup>1</sup>A Markov chain is a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. Roughly speaking, a process satisfies the Markov property if one can make predictions for the future based only on its present state just as well as one could knowing the process’s full history. In statistics, MCMC methods comprise a class of algorithms for sampling from a probability distribution. By constructing a Markov chain that has the desired distribution as its equilibrium distribution, one can obtain a sample of the desired distribution by observing the chain after a number of steps. The more steps there are, the more closely the distribution of the sample matches the actual desired distribution.