This paper focuses primarily on a mean-field analysis of deep neural networks. The first few pages introduce all the necessary information by comparing it with the existing literature to understand the goal of the paper.

The main result, which characterizes the asymptotic behavior of a neural network with two hidden layers when the number of hidden units becomes large, is presented in Section 2. The result can be easily extended to an arbitrary number of hidden layers. Section 3 is devoted to global convergence arguments. In particular, it is shown that under the proper assumption, the limiting problem derived in Section 2 seeks to minimize the limiting objective function, recovering the global minimum. Further, a numerical study is presented to showcase some of the theoretical implications, and, as an example, the limit for a three-layer neural network is discussed in Section 4. The proof of the convergence theorem is given in Section 5. The uniqueness of a solution to the limiting system is established in Section 6. Section 7 has concluding remarks and possible directions for future work. The proof of the limit of the first layer (i.e., the proof of Lemma 2.2) and a few other results are provided in the appendix.

Reviewer: Arzu Ahmadova (Essen)

MSC:
68T07 Artificial neural networks and deep learning
60F05 Central limit and other weak theorems
82C32 Neural nets applied to problems in time-dependent statistical mechanics

Keywords:
weak convergence; deep neural network; mean-field limit; deep learning

Software:
CIFAR; DGM

Full Text: DOI arXiv

References:


