World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:12, No:10, 2018

Understanding and Improving Neural Network Weight Initialization

Authors : Diego Aguirre, Olac Fuentes

Abstract : In this paper, we present a taxonomy of weight initialization schemes used in deep learning. We survey the most representative techniques in each class and compare them in terms of overhead cost, convergence rate, and applicability. We also introduce a new weight initialization scheme. In this technique, we perform an initial feedforward pass through the network using an initialization mini-batch. Using statistics obtained from this pass, we initialize the weights of the network, so the following properties are met: 1) weight matrices are orthogonal; 2) ReLU layers produce a predetermined number of non-zero activations; 3) the output produced by each internal layer has a unit variance; 4) weights in the last layer are chosen to minimize the error in the initial mini-batch. We evaluate our method on three popular architectures, and a faster converge rates are achieved on the MNIST, CIFAR-10/100, and ImageNet datasets when compared to state-of-the-art initialization techniques.

Keywords : deep learning, image classification, supervised learning, weight initialization
Conference Title : ICPRML 2018 : International Conference on Pattern Recognition and Machine Learning
Conference Location : Paris, France
Conference Dates : October 29-30, 2018