

1126-68-350

Ralph W Crosby* (ralph.crosby@navy.mil), 1 Innovation Drive, N. Charleston, SC 29410.

Training Deep Neural Networks.

Current methods for machine learning focus on the use of artificial neural networks (ANN) as a core technology for a number of processes associated with the classification of inputs into some set of outputs based on opaque characteristics of the data. These inputs potentially include visual data and network data. The outputs potentially include everything from simple classification into a known set (e.g. house, car) to translation from one natural language to another (e.g. English to French).

It has been shown that for any continuous function, it is possible for a single layer neural network of finite width to approximate the function (the Universality approximation theorem). But, even though the width is demonstrably finite, it can become extremely large for even modestly complex functions. The solution to this issue is to generate multi-layer neural networks substituting network depth for width.

The problem that occurs with increases in network depth is that classic methods of “training” the network become progressively less effective as the depth of the network increases. In this talk we present a discussion of the problem of training a deep neural network and further discuss the intuitions that guided the development of modern methods of training deep neural networks. (Received January 17, 2017)