Iowa State University

Department of Electrical and Computer Engineering

Deep Machine Learning: Theory and Practice

EE 526X

Midterm Assignment

Author:
Vishal Deep

Instructor:
Dr. Zhengdao Wang

November 2, 2019

Iowa State University

Paper Title: Memory-Optimal Direct Convolutions for Maximizing Classification Accuracy in Embedded Applications

Introduction and Problem Statement

Machine learning on low power and low energy consumption edge devices is growing in popularity these days because of the increased privacy due to the computation done on local devices. But these small edge devices comes with very low memory size (2KB to 16 KB). Therefore, we have to train the network on a server and only store weights, biases, activations and configuration parameters. We know that the classification problems gives best results when Convolution Neural Networks (CNN) are used but the CNNs are memory hungry algorithms. Other works have tried to use alternative algorithms but none of them were able to achieve the accuracy near to CNN. Therefore this paper proposes and provides the method to use CNN or very small memory devices of 2KB to achieve best accuracy for the classification problems.