Department of Computer Science

St. Francis Xavier University

Presents

Wasserstein Based Feature-Map Knowledge Transfer to Improve the Performance of Small Deep Neural Networks

by

Deepesh Ramegowda
St. Francis Xavier University
M.Sc. Thesis Proposal Presentation
July 6, 2021 at 10am AST via Zoom (Link to follow)

Deep Neural Networks (DNNs) have been successful across various fields such as image recognition, natural language processing, and robotics. With the increased use of DNN applications, training time and memory utilization of the deployed model have become essential factors for good prediction accuracy in DNNs. However, amidst the growth of DNNs, one aspect that has been neglected is the problem of deploying them on devices that can support the computational and memory requirements of DNNs. Many computer vision applications like face detection, object classification, and pedestrian detection require real-time execution, with devices mounted on cameras. These devices are low-powered and do not have the computational resources to run the data through DNNs and get instantaneous results. The DNN architecture consists of several hidden layers. It is important to understand the information in the hidden layers of DNNs. This hidden layer information, which is referred to as feature maps, contains rich spatial and semantic information about the input data. A few techniques like Model Compression, Knowledge Distillation, Deep Mutual Learning, and Adversarial Deep Mutual Learning have been introduced to transfer the knowledge of a large network to a smaller network for improving its prediction accuracy. The existing methods do not effectively use the feature maps as a source of information that can be exploited to improve the knowledge transfer paradigm. This is because of Kullback-Leibler Divergence when used for measuring different feature map distributions, which often leads to the problem of Mode collapse and Non-convergence. The use of Wasserstein metric (a.k.a. Earthmover's distance) when measuring different feature map distributions can alleviate the problem of Mode collapse and Non-convergence. This is because the Wasserstein metric does not require feature map distributions to be on the same probability space. We propose an Adversarial learning-based approach for transferring feature map information from a large pre-trained network to a smaller network using the Wasserstein metric. Our novelty lies in the fact that we eliminate the problem of Mode collapse and Non-convergence in the knowledge transfer paradigm. As a result, we effectively utilize feature map information to improve the predictive performance of small DNNs. We focus on computer vision tasks for our experiments, and we use benchmarks datasets, including CIFAR-10 and CIFAR-100, that provide large training data for different tasks like classification, segmentation, and object localization.