Deep Learning Final Project Proposal

SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy (CVPR 2023)

Member: 111703040 游宗諺, 111703004 林子齊, 110703066 王冠智, 111703009 嚴聲遠

A. Introduction and Problem Description

Convolutional Neural Networks (CNNs) have achieved impressive results in computer vision tasks such as image classification and object detection. However, as the paper states, traditional convolutional layers often extract redundant spatial and channel features, which leads to unnecessary computations and memory usage.

There are some clever methods that were introduced by the professor during classes to alleviate this problem. For example, ResNet utilizes an efficient shortcut connection to improve the network topology and diminish the redundant parameters. At the same time, depth-wise convolution (DWC) reduces the spatial redundancy, and point-wise convolution (PWC) reduces the channel redundancy. However, these methods don't directly look into the feature maps to check and reduce feature redundancy. They simply apply computationally cheaper operations and hope that feature redundancy can be reduced.

This paper presents a new "plug-and-play" convolutional module to solve this issue. The module is composed of two specialized units:

1. **Spatial Reconstruction Unit (SRU):** Uses a "separate-and-reconstruct" method to **suppress** spatial redundancy and **highlight** more informative features, instead of straightly using DWC to compress feature maps.
2. **Channel Reconstruction Unit (CRU):** Employs a "split-transform-and-fuse" strategy to adaptively merge the high-level features and the detailed-supplementary features via a soft-attention mechanism, instead of straightly using PWC to decrease the parameter count.

B. Project Goal

The objective of this project is to reproduce and evaluate the SCConv concept on a smaller scale using the CIFAR dataset and ResNet50 architecture in PyTorch. Specifically, the project will go through the following steps:

1. Train a baseline ResNet50 on CIFAR-10 (or CIFAR-100) using a Google Colab T4 GPU.
2. Implement the Spatial Reconstruction Unit (SRU) and Channel Reconstruction Unit (CRU) modules as defined in the paper.
3. Integrate SRU + CRU into the ResNet50 architecture by replacing the standard bottleneck convolutional blocks.
4. Train the modified SCConv-ResNet50 model under the same training setup as the baseline.
5. Compare the two models in terms of:
    a. Classification accuracy
    b. Training time and computational efficiency
    c. Number of learnable parameters

Expected outcomes:

Demonstrate reduced model redundancy and parameter count while maintaining or improving classification accuracy.

C. Prerequisite Knowledge

To successfully complete this project, the following knowledge areas are required:

1. All deep learning techniques that were taught in class, including:
    a. Group Normalization + their respective trainable parameters γ and β
    b. Applying activation functions (Sigmoid, SoftMax) on CNN
    c. PWC, DWC and GWC
    d. Global pooling
2. PyTorch Programming
    a. Building and training models using PyTorch modules
    b. Implementing custom layers and integrating them into existing architectures
3. Google Colab and Working Environment
    a. Efficient usage of T4 GPUs on Colab for training deep networks
    b. Being able to use Git with Google Colab