PVANET: Deep but Lightweight Neural Networks for Real-time Object Detection

ABSTRACT

Convolutional neural networks (CNNs) have made impressive improvements in object detection for several years. Recent object detection systems have met acceptable accuracies for commercialization in a broad range of markets like automotive and surveillance. In terms of detection speed, however, even the best algorithms are still suffering from heavy computational cost. Although recent work on network compression and quantization shows promising result, it is important to reduce the computational cost in the network design stage. This work presents lightweight feature extraction network architecture for object detection, named PVANET, which achieves real-time object detection performance without losing accuracy compared to the other state-of-the-art systems.

OBJECTIVE

This work presents how we can achieve the state-of-the-art accuracy in multi-category object detection task while minimizing the computational cost by adapting and combining recent technical innovations.

METHODOLOGY

Following the common pipeline of "CNN feature extraction + region proposal + RoI classification", we mainly redesign the feature extraction part, since region proposal part is not computationally expensive and classification part can be efficiently compressed with common techniques like truncated SVD. Our design principle is "less channels with more layers" and adoption of some building blocks including concatenated ReLU, Inception, and HyperNet. The designed network is deep and thin and trained with the help of batch normalization, residual connections, and learning rate scheduling based on plateau detection.

DATA DESCRIPTION

- PVANET was pre-trained with ILSVRC2012 training images for 1000-class image classification. The validation and test data for this competition will consist of 150,000 photographs, collected from flickr and other search engines, hand labeled with the presence or absence of 1000 object categories. The 1000 object categories contain both internal nodes and leaf nodes of ImageNet, but do not overlap with each other.
- VOC2007
- VOC2012

IMPLEMENTATION PLATFORM/TOOLS

Python based machine learning approach.

OUTCOMES

Mean average precision should be calculated and compared with state of art.

TIMELINE

August – September 15 – Dataset collection and Literature Survey

October – Implementation