Summary of "ImageNet Classification with Deep Convolutional Neural Networks"

Isaac Akintaro

18/11/2023

1 Introduction

This paper, also known as AlexNet, was groundbreaking in reviving interest in Deep Learning. It paved the way for significant advancements in image recognition, object detection and other areas.

The paper used Convolutional neural networks (CNNs) which makes use of convolutional layers which capture patterns in the features, whilst the fully-connected layers turned learned features into abstract representations. The network is trained on the ImageNet dataset of over 1 million labeled high-resolution images in 1000 categories. The architecture consists of 5 convolutional layers for feature extraction, followed by 3 fully-connected layers for classification.

2 Novel Elements

Several novel elements are introduced:

- Using rectified linear units (ReLUs) as the nonlinear activation function. ReLUs do not saturate (output becomes constant at the extremes) like tanh units, allowing several times faster training.
- Training the network across 2 GPUs using a parallelization scheme to distribute computation. Half the neurons/kernels are assigned to each GPU.
- Applying local response normalization in some layers, which performs lateral inhibition between kernel maps to introduce competition.
- Using overlapping pooling for robustness.

3 Reducing Overfitting

From my understanding although the dataset was 1.2 million labelled images, there were 60 million parameters needed to be learnt, so data augmentation was used to increase training set by a factor of 2048. To reduce overfitting:

- Augmenting the data by generating transformed versions of images (translations, reflections, altering color intensities).
- Using dropout, where random neurons are dropped during training to prevent co-adaptation of features.

4 Results

The network achieved top-1 and top-5 error rates of 37.5