Article reviewed: Deep Residual Learning for Image Recognition – Meetu Malhotra

Reference: https://arxiv.org/abs/1512.03385

Summary: Image recognition pertains to the field of computer vision and involves the automatic identification and interpretation of various objects within pictures. Image recognition involves use of Convolutional network, which are hard to train. As deeper networks begin to converge, a degradation issue emerges and as the depth of the network increases, the accuracy becomes stagnant. At this point, if we further add more layers to the network, accuracy falls steadily. To overcome the issue, by introducing the shortcut connections which neither add extra parameter nor computation complexity. The implementation of residual network is tested on various datasets such as ImageNet, CIFAR-10 etc and the analysis showed promising results, with the error percentage of 3.75% on ImageNet test dataset.

Five C's:

Category: As per arxiv, this paper falls under the category of Computer vision and Pattern Recognition specifically deep Convolutional networks, focusing on RESNETs

Context: The paper focuses on residual learning framework to ease the training of networks such that that these residual networks are easier to optimize, and can gain accuracy from considerably increased depth

Correctness: The paper presents accurate information on developing residual learning framework as shown in the experiments section, covering comparison on plain networks, residual networks, and deeper bottleneck architectures.

Contributions: This paper significantly changed the domain of deep learning, especially in computer vision, by training a lot deeper neural networks and addressed the degradation problem by introducing a deep residual learning framework. The paper focuses on residual concept, where the layers learn the residual or difference between the input and the desired output. The architecture is trained up to 152 layers, resulting in 3.57% error on the ImageNet test set.

Clarity: The paper is written in a clear and understandable manner. It starts with the introduction of Residual network and shows 3.57% of error on imageNet test set, eliminating the degradation problem. It shows clear architectural design through ResNets in section 3 of deep residual learning. It provides comparisons on different layers and provides clear and concise results in tables on ImageNet dataset. Additionally, it shows experiments conducted on CIFAR 10 dataset. Further, it covers analysis on COCO object detection dataset and ILSVRC dataset. These experiments show validity of the claims made by author and the practical implementation on the concept on different datasets.

Outline: The paper follows a logical structure, starting with an introduction to Resnet, the issue of degradation that it resolved and the results. It then talks about its architecture, concept of residual learning, Identity Mapping by Shortcut, explaining with architectural diagram in figure 3. The paper also discusses the analysis of implementation on different image datasets and on different layers, starting from 18 layers to aggressively on 1000 layers. Furthermore, it emphasizes good generalization of performance on object recognition tasks also, including ImageNet Detection and ImageNet Localization. It further covers the object detection improvements also.

Discussion:

Innovations: It brought the innovative idea of residual deep learning networks, which revolutionized the world of deep learning. It addressed the key issue of degradation in deep neural networks. As we add more layers to the network, the accuracy saturates and eventually degrades rapidly after a certain point. This results in high training error. The paper introduced the concept of residuals in layers that overcome this issue.

Assumptions: The paper assumes that the readers have fundamental understanding of neural network concept such as shallow and deep neural networks, identity mapping, identity shortcuts, the degradation issue with deep learning neural networks.

Faults: the paper is a bit complicated; the concept is difficult to understand, not explained in depth.

Terminology: Deep neural networks, Deep Residual Learning, degradation, shortcut connections, identity mapping, ImageNet classification, Object Detection