Deep Learning for Computer Vision - A crash course

Shubham Arora

1 Standard tasks in Computer Vision

- Recognition

  - Image classification
  - Image captioning
  - Object localization
  - Segmentation
  - Object Detection

- Motion Analysis

  - Tracking
  - Optical flow

- Other

  - Image recolorization
  - Super-Resolution

2 Important Ideas of Deep Learning-Vision. A historical perpective

2.1 Convolutional Neural Networks (CNN)

LeNet   1998 - LeNet-5 was one of the first NN that utilized backpropagation using Supervised Learning. CNN's are by far the most used Neural layer architecture in any machine vision tasks. Yann LeCun was a co-recipient of the 2018 Turing Award for his work in Ai & vision, owning a lot to the success of CNN in practical applications.

How does it work

- ConvNet architecture is particularly optimized for images - vastly reduces the number of parameters neede to train the network.

- essentially a combination of matrix dot products and $max()$ operations

- Function from raw pixels $->$ n numbers (classs scores in the case of image classification)

2.2 Deep Convolutional Neural Networks (CNN)

AlexNet   2012 - Image classifier on the ImageNet database. It Built upon the 1998 work on CNN, it was scaled massively due to availability of exponentially more data and parallel compute using GPU's

Major Improvements

- Reduced code complexity as it utilized homogeneous architectures

- No need to do complex step of feature extraction

- Made it easier to do Transfer Learning

Industry applications

- Face Recognition

- Self driving cars

- Image captioning?

- Building block in Reinforcement Learning

2.3   Residual Networks (ResNet)

ResNet   2015 - Microsoft Research - Much better performance than plain deep neural networks. Utilized the idea of skip connections across non sequential layers. Won the ImageNet challenge in 2015.

3   Why vision is important?

- Vision and NLP have been the core of Machine Learning innovation

- Vision is a very important ability in robotics for perception, localization, mapping and motion planning tasks.

- Automate critical tasks that rely on human vision - fault detection, autonomous vehicles

4   Important Tools

- OpenCV

- Keras

- Tensorflow

5   Refrences

- Andrej Karpathy

- Wikipedia

- Stanford - cs231n