ANN
COURSE 4 - WEAK 1
CNN SUMMARY

By:
Mohamed Ahmed Nassar

- Computer Vision is one of the widely used technology in many areas like self driving cars.
- Before Deep learning, There were many classical technique were used in image processing
- Edge, corner and interest point detection are the fundamental blocks of any image processing algorithm
- Any complex image feature are composed of simpler features (Edges, corner and interset points)
- in Deep-learning, convolutional networks outperform all other techniques in image processing field
- convolution is a linear process where a filter or (Kernel) is slide on image and multiplied element-wise then summed
- Convolution networks uses the divide-and-conquer approach to handle any problem
- Earlier layers of CNN are learn simple features Like edges and corner
- intermediate layers learn combination from those simple features maybe simple shapes (circles, triangles and squires)
- Deeper layers learn another combination from the simple shapes, now it can fire for a specific object like face, nose, ear or any other object depending on problem
- Convolution layers have many advantages
  - Shared weights, where a one small filter are slided the whole image with the same weights
  - CNN are translation invariant
  - small number of weights due to shared weights
- Convolution output size is smaller than it's input image size
- Size shrinking on convolution depend on stride value
- Stride is the filter shifting steps used after each convolution step
- The more the stride value the less the output size
- Padding is used to overcome the problem of size shrinking
- Padding is nothing than add zero rows and columns around the image to impute the size shrinking
- Convolution are done on volumes in case of multi-channel images
- Depth of convolution filter is always equal to the number of input channels
- Number of output channels is a hyper-parameter
- The more the output channels to more the features detected
- Pooling layers is widely used after convolution layers
- Pooling in a simple term is a down-sampling
- Two types of pooling are introduced "Max" and "Average" pooling
- Max pooling select the largest feature value within window
- Average pooing take the window average
- Pooling Layers has many advantages
  - Translation invariance, it doesn't matter where the feature is
  - pooling decreases the volume size
  - in max pooling it keeps only the important features
- Non-linear activation function is also introduced after max pooling layer
- One or more fully Connected layers are used after CNN
- the last layer of the network is a Fully Connected Layer to do the regression or classification
- Cost function, Optimiser, regularization and batch normalization are still used as before with the same effect
- The Rule of thumb pipeline of CNN is as follow
  - Conv → pooling → ReLu → [Repeated many times] → [one or more] Fully connected