Exercises #13 - Convolutional Neural Network

Thales Menezes de Oliveira

Brazilian Center for Physics Research (CBPF)

April 1, 2022

1.What are the advantages of a CNN over a fully connected DNN for image classification?

A:

- Since the layers in a CNN are only partially connected, different than a fully connected DNN, it has fewer parameters than the previous, which makes the training process faster. Due its low number of parameters, the risk of overfitting is also decreased.
- The CNN is capable of learning a particular feature, i.e. an edge, as well as the DNN, but for the first one, it can detect this feature or pattern anywhere in the image, in contrast with the DNN which can only detect the learned feature in a particular location. Since the images typically have very repetitive features, CNNs are able to generalize much better than the DNNs.
- The CNN, due to its architecture, has a prior knowledge of the pixels' organization, in contrast with the DNNs. In a CNN, the lower layers are capable to identify features in small areas of the image, while the higher layers are able to combine these features from the lower layers into larger ones. This procedure works well with most images, which gives the CNN a decisive head start compared to the DNNs.

3. If your GPU runs out of memory while training a CNN, what are the five things you could try to solve the problem?

A: In order to reduce the use of the GPU it is possible to:

1 Reduce the mini-batch size;

2 Reduce dimensionality using a larger stride in one or more layers;

3 Remove one or more layers;

4 Use 16-bit floats instead of 32-bits;

5 Distribute the CNN across multiple devices

4. Why would you want to add a max pooling layer rather than a convolutional layer with the same stride?

A: A convolutional layer carries few parameters, in comparison with the MaxPoolling layer which doesn't carry any parameters at all.

5. Why would you want to add a local response normalization layer?

A:

6. Can you name the main innovations in AlexNet, compared to LeNet-5? What about the main innovations in GoogLeNet and ResNet?

A: The AlexNet was much deeper and larger, compared to LeNet-5, and it stacks convolutional layers directly on top of each other, instead of stack a pooling layer on top of the convolutional one. For the GoogLeNet, the main innovation was the *inception modules*, which makes it possible to have a much deeper CNN architectures than we had before, with fewer parameters. At last, the ResNet introduced the skip connection, allowing to the CNNs to have more than 100 layers.