WEEK_5

Deep Learning for Image Search

Computational Graph:

The entire purpose of TensorFlow is to have a so-called computational graph that can be executed much more efficiently than if the same calculations were to be performed directly in Python.

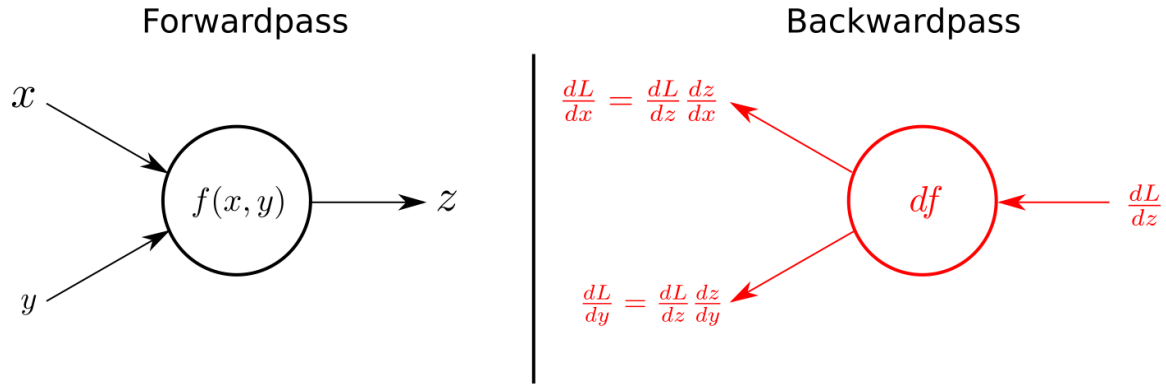
TensorFlow can also automatically calculate the gradients that are needed to optimize the variables of the graph so as to make the model perform better.

Computational Graph consists of:

- Placeholder variables used to change the input to the graph.
- Model variables that are going to be optimized so as to make the model perform better.
- The model which is essentially just a mathematical function that calculates some output given the input in the placeholder variables and the model variables.
- A cost measure that can be used to guide the optimization of the variables.
- An optimization method which updates the variables of the model.

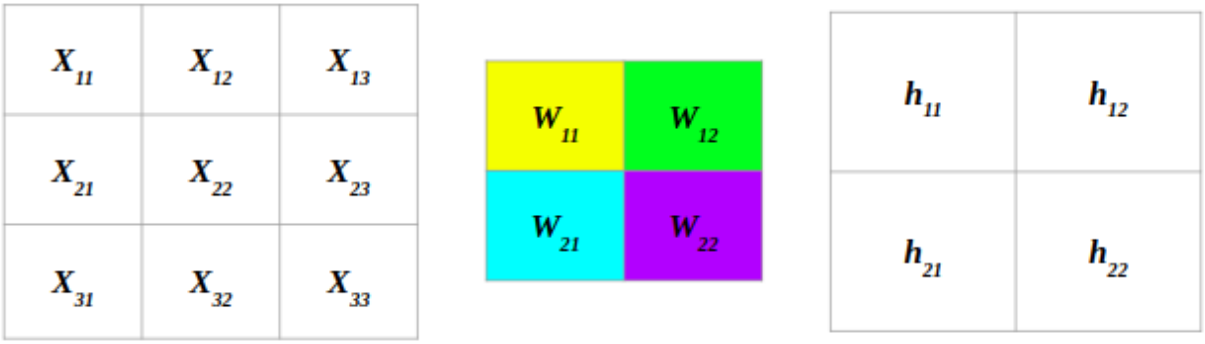
The Chain Rule

The following figure summarises the use of chain rule for the backward pass in computational graphs.



Integrating with CNN:

Suppose a convolution operation takes an input X of size 3x3 using a single filter W of size 2x2 without any padding and stride = 1 generating an output H of size 2x2.



$$h_{11} = W_{11}X_{11} + W_{12}X_{12} + W_{21}X_{21} + W_{22}X_{22}$$

$$h_{12} = W_{11}X_{12} + W_{12}X_{13} + W_{21}X_{22} + W_{22}X_{23}$$

$$h_{21} = W_{11}X_{21} + W_{12}X_{22} + W_{21}X_{31} + W_{22}X_{32}$$

$$h_{22} = W_{11}X_{22} + W_{12}X_{23} + W_{21}X_{32} + W_{22}X_{33}$$

Now using above formula and feeding into Computational Graph we get the desired output because every node of Graph i.e. tensor can be vector, function, matrix, etc. so, it makes easy and more efficient.

Also, Computational Graph makes use of chain rule to calculate derivative which helps in reducing the time. Now, for implementing the back-propagation step for the current layer, we can assume that we get ∂h as input and our aim is to calculate ∂w and ∂x. It is important to understand that ∂x would be the input for the backward pass of the previous layer. This is the core principle behind the success of back propagation.