**Assignment 1**

Q1.What are channels and kernels(according to EVA)?

Ans:

To start with, in basic terms we can understand:

**Channel as feature container and**

**Kernel as a feature extractor.**

**Channel:**

Moving forward with a channel we can understand it as anything basic or primary, combining which we can obtain anything useful.

Channels and nothing but primary features. There can be as number of channels as required but after combining all the channels we should be able to get something of our choice.

**Channels are primary in nature and it should be combinable.**

Eg1:

We can consider the colours red, green and blue called as RGB as 3 channel.

Now combining these three colours or basically combining these three channels we can obtain any colour of our choice. Therefore, these three channels are nothing but three basic primary colours or the basic features of any colour image.

Eg2:

We can consider the digits from 0 to 9 as 9 different channels.

Now combining these nine channels accordingly we can get any positive number in the universe of our choice.

**Kernel:**

The dimension of the kernel can also be of any choice and what a kernel does is basically **convolves on the top of a channel and extracts its features and creates a new intermediate channe**l or a final output channel.

The kernel of any given size has to go around all the possible ways to cover a channel and extract all of its features and store them into a cell of a new intermediate of final channel.

Q2. Why should we (nearly) always use 3x3 kernels?

Ans:

What basically a kernel is doing is studying the parameters or the features over the channel it is convolving on.

When we are choosing a kernel of size 3 x 3, we are basically forcing a kernel to study the channel in the size of 3 x 3. The advantage this is providing is is that it goes in depth of the extraction of features since it **focuses on a small area at a time**.

Also, when we are choosing a kernel of bigger size it’s also possible that it may study unrelated features which we do not want.

What advantage a 3 x 3 kernel is providing is that it is covering all possible scenarios and hence making the output channel more feature-rich after each convolution.

Adding to this it is also proven to be GPU efficient.

Q3. How many times do we need to perform 3x3 convolutions operations to reach close to 1x1 from 199x199 (type each layer output like 199x199 > 197x197...)

Ans: 100

199 x 199, 197 x 197, 195 x 195, 193 x 193, 191 x 191, 189 x 189, 187 x 187, 185 x 185, 183 x 183, 181 x 181, 179 x 179, 177 x 177, 175 x 175, 173 x 173, 171 x 171, 169 x 169, 167 x 167, 165 x 165, 163 x 163, 161 x 161, 159 x 159, 157 x 157, 155 x 155, 153 x 153, 151 x 151, 149 x 149, 147 x 147, 145 x 145, 143 x 143, 141 x 141, 139 x 139, 137 x 137, 135 x 135, 133 x 133, 131 x 131, 129 x 129, 127 x 127, 125 x 125, 123 x 123, 121 x 121, 119 x 119, 117 x 117, 115 x 115, 113 x 113, 111 x 111, 109 x 109, 107 x 107, 105 x 105, 103 x 103, 101 x 101, 99 x 99, 97 x 97, 95 x 95, 93 x 93, 91 x 91, 89 x 89, 87 x 87, 85 x 85, 83 x 83, 81 x 81, 79 x 79, 77 x 77, 75 x 75, 73 x 73, 71 x 71, 69 x 69, 67 x 67, 65 x 65, 63 x 63, 61 x 61, 59 x 59, 57 x 57, 55 x 55, 53 x 53, 51 x 51, 49 x 49, 47 x 47, 45 x 45, 43 x 43, 41 x 41, 39 x 39, 37 x 37, 35 x 35, 33 x 33, 31 x 31, 29 x 29, 27 x 27, 25 x 25, 23 x 23, 21 x 21, 19 x 19, 17 x 17, 15 x 15, 13 x 13, 11 x 11, 9 x 9, 7 x 7, 5 x 5, 3 x 3, 1 x 1.

Q4. How are kernels initialized?

Ans:

The initialisation of kernel with **random values** has proven to be very efficient and useful.

(Why?)

To understand this how let’s consider this. When a particular image/object has to be processed completely by a mapping process that has predefined method then we have to ensure that the processing has to be done in a manner such that it is able to process the entire object or image.

Hence the initialisation for the starting point of that mapping algorithm has to be in a random manner and not from a fixed point or with fixed values.

The randomness of the random values is what will differentiate the same mapping algorithm to get an overview or the idea of the entire object or image **from all the possible perspective**.

Therefore, when initialised with random values it covers the entire area **(entire features in all possible direction covering each and every scenario)** even if it is being processed by the same mapping algorithm and hence the training is **very information rich.**

Each time we initiate the training each of the time we will be having different random values on which the mapping algorithm will be working on hence it helps in **enhancing the training data**.

Q5. What happens during the training of a DNN?

Ans:

The basic step what’s happening in a deep neural network is that we have a channel of a given dimension, we have a kernel convolving around it which is also of a given dimension and output layer or output channel which is being obtained after the convolution happens.

The cell of the kernel has some random values, which after processing the cells of the given input channel by a particular defined algorithm, stores the information in the cell of the output channel.

According to the area which kernel is processing at a time the entire information of those cells of a particular dimension is being stored in a single cell of the output channel.

For example, say suppose we have an input channel of a dimensional 5 x 5, and a kernel of dimension 3 x 3, so the kernel of this particular size is processing 3 x 3 area of the given input channel according to its algorithm and storing the information in a particular cell of the output channel. So, in the output channel initially we have one cell which contains the information of 9 cells of the input channel.

Similarly, this process is carried on by the kernel by moving one cell upwards and one cell downwards (or horizontal or vertical) and covering the entire region of the input channel. Therefore, by carrying out the single operation on a 5 x 5 channel by 3 cross 3 kernel we get 2 x 2 output channel.

This was just a primary step of what’s happening in the entire system, we may have an input channel of a very big size and we have a given kernel of a particular size. This process of convolution will begin with the input channel and will be carried on until required. So the size of the output or intermediate channels will keep on decreasing and the cells in it will be storing information of a greater view of the previous input channel.