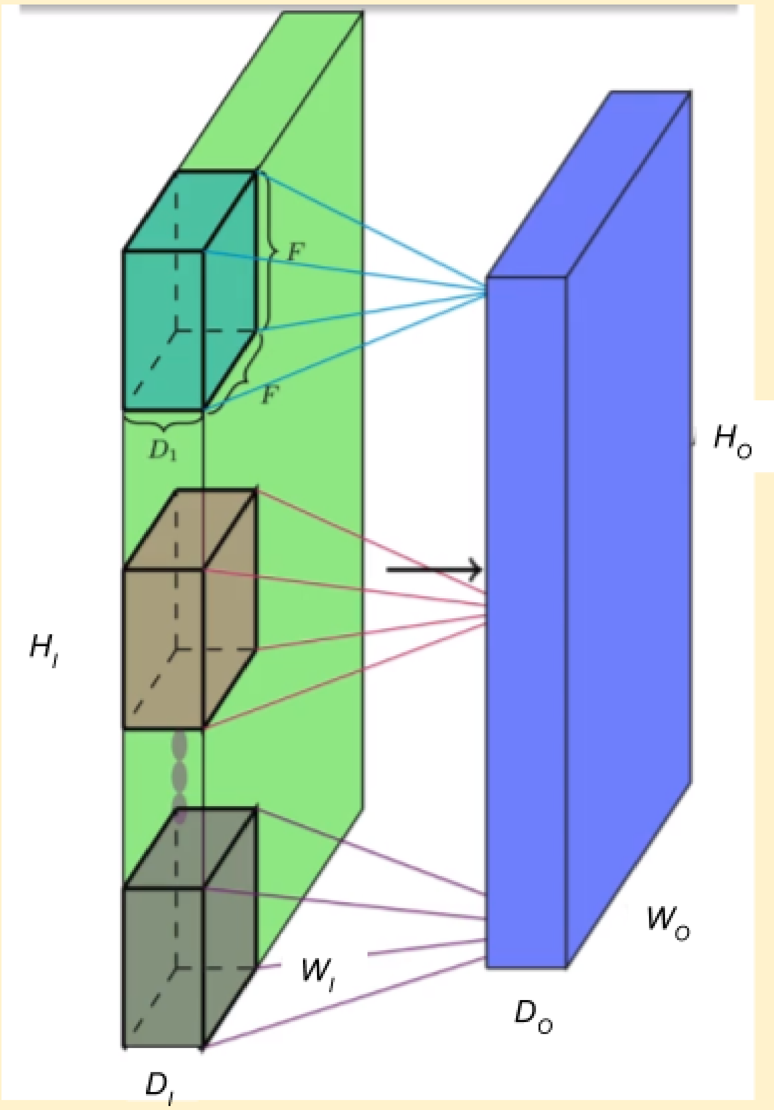
[PadhAI: The Convolution Operation]

One Fourth Labs

Terminology

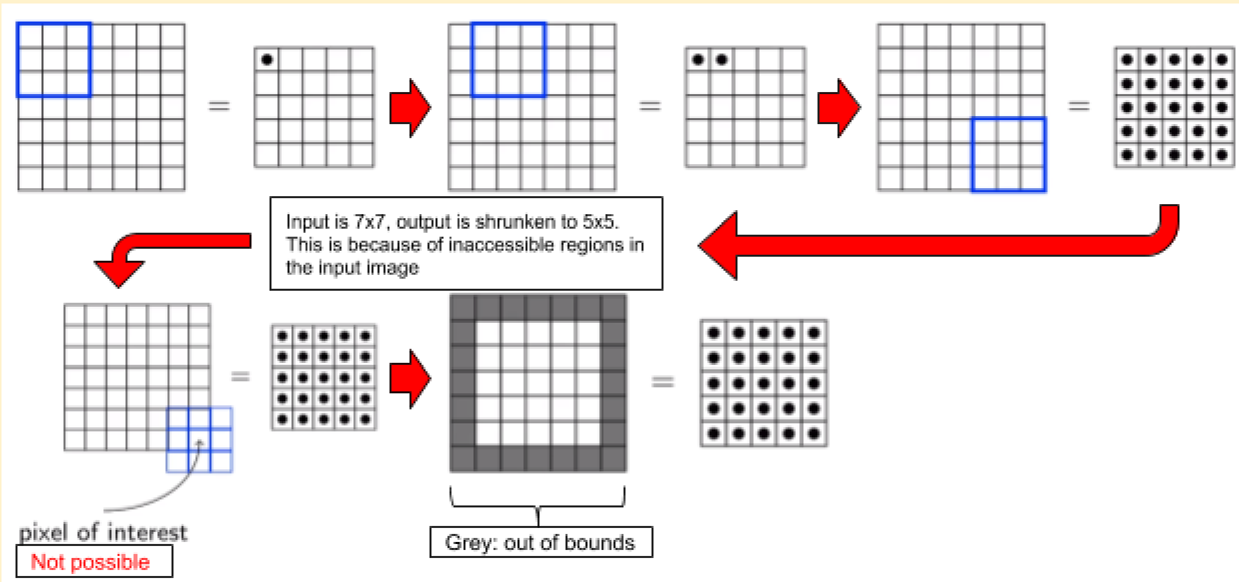
Let's look at some terminology

1. Consider the following 3D convolution operation and look at the terminology associated with it



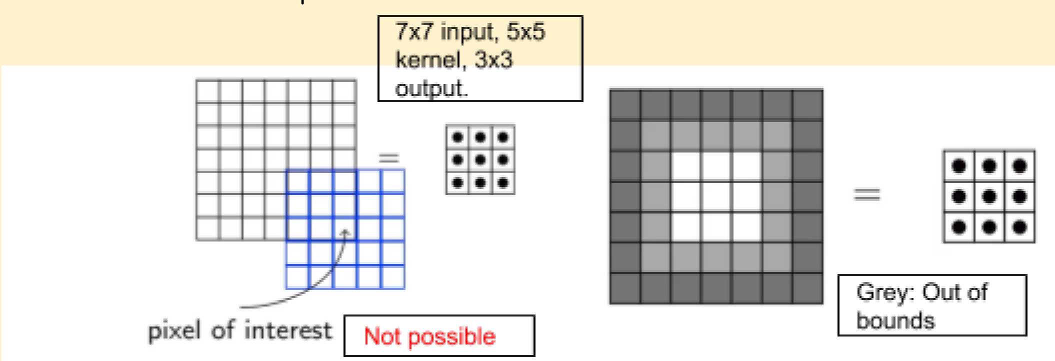
2. **Terminology**:
    a. Input Width (W_I), Height (H_I) and Depth (D_I)
    b. Output Width (W_0), Height (H_0) and Depth (D_0)
    c. The spatial extent of a filter (F), a single number to denote width and height as they are equal
    d. Filter depth is always the same as the Input Depth (D_I)
    e. The number of filters (K)
    f. Padding (P) and Stride (S)
3. **Question**: Given W_I, H_I, D_I, F, K, S and P how do you compute W_0, H_0, and D_0?

4. To answer that, let us look at a sample 3x3 kernel passing over a 7x7 image



   a. Here, we can see that by running the 3x3 kernel over a 7x7 image, we get a smaller 5x5 image.
   b. This is because we can't place the kernel at the corners as it will cross the input boundaries
   c. This is true of all the shaded points.
   d. Hence the size of the output will be smaller than that of the input

5. Let's see another example with a 5x5 kernel



   a. Here, we can see that by running a 5x5 kernel over a 7x7 input, we get a smaller 3x3 image
   b. Here, the out-of-bounds regions are larger.
   c. Thus the output is much smaller.
6. We can see that the reduction in size can be given by the following equations
   a. $W_O = W_I - F + 1$
   b. $H = H_I - F + 1$
7. However in practice, we could still place the kernel on the boundary and take only the valid neighbors. This is roughly what is being done.