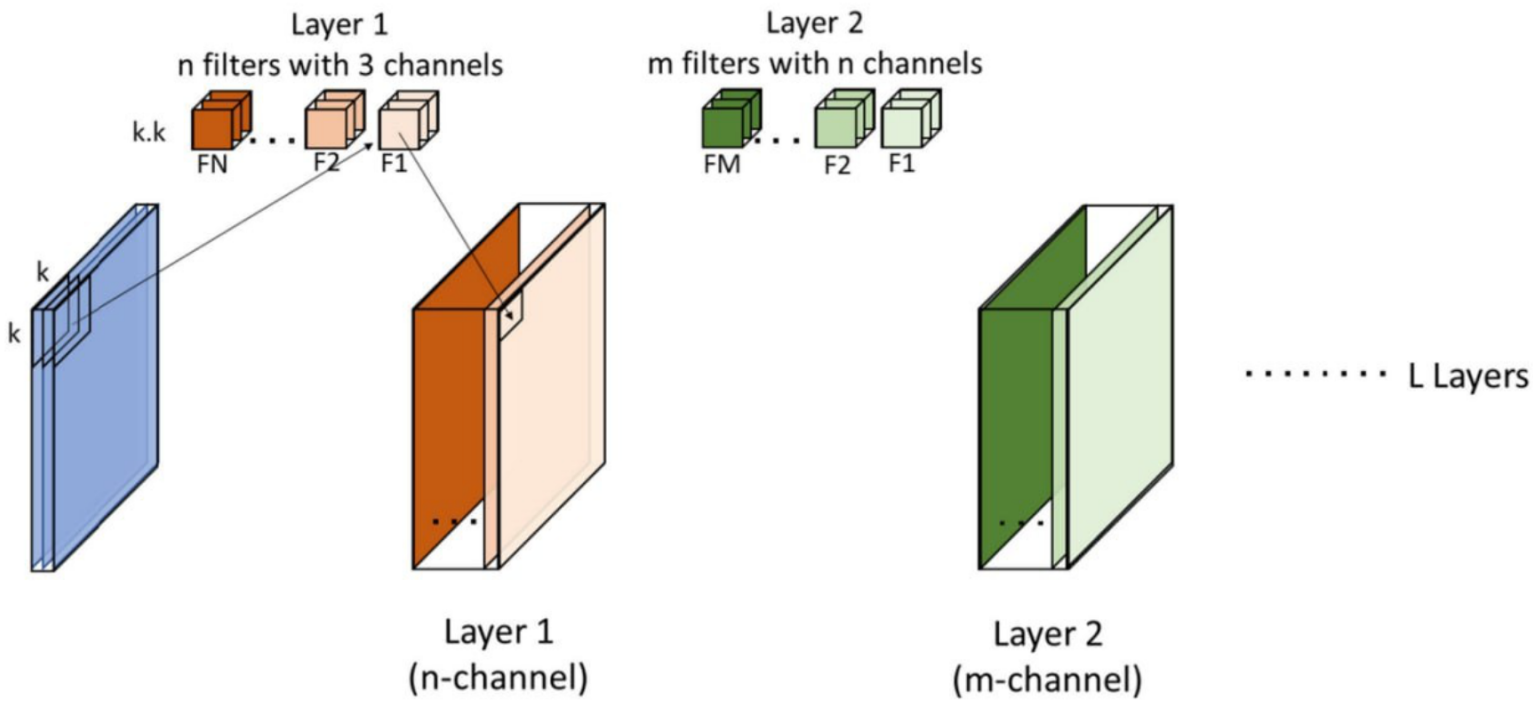
Approach for Convolution

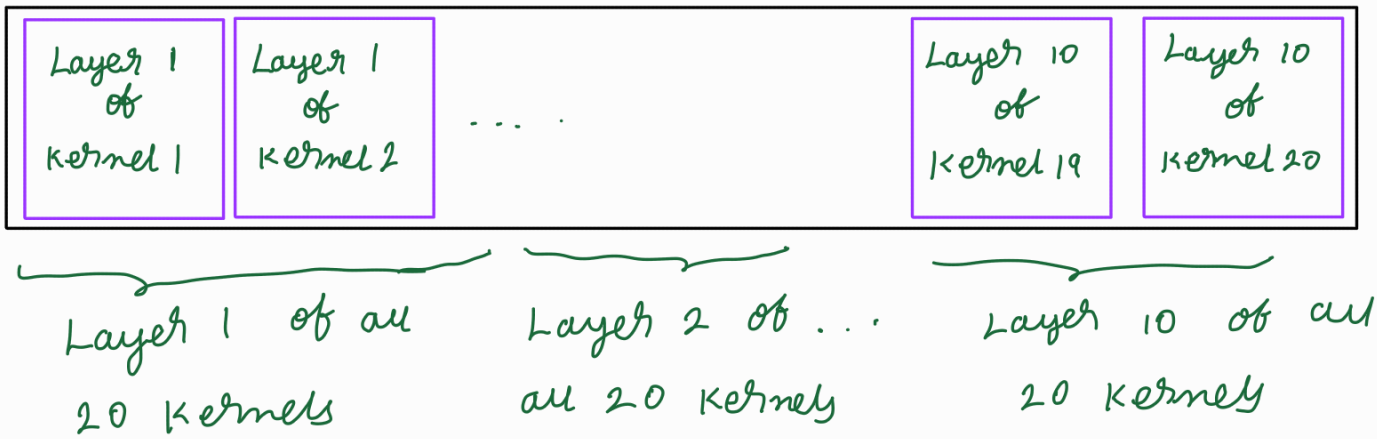
Let input size be $28 \times 28 \times 10$

kernel size be $3 \times 3 \times 10$

↳ Let us use 20 such kernels for 20 output channels



1) Kernels partition

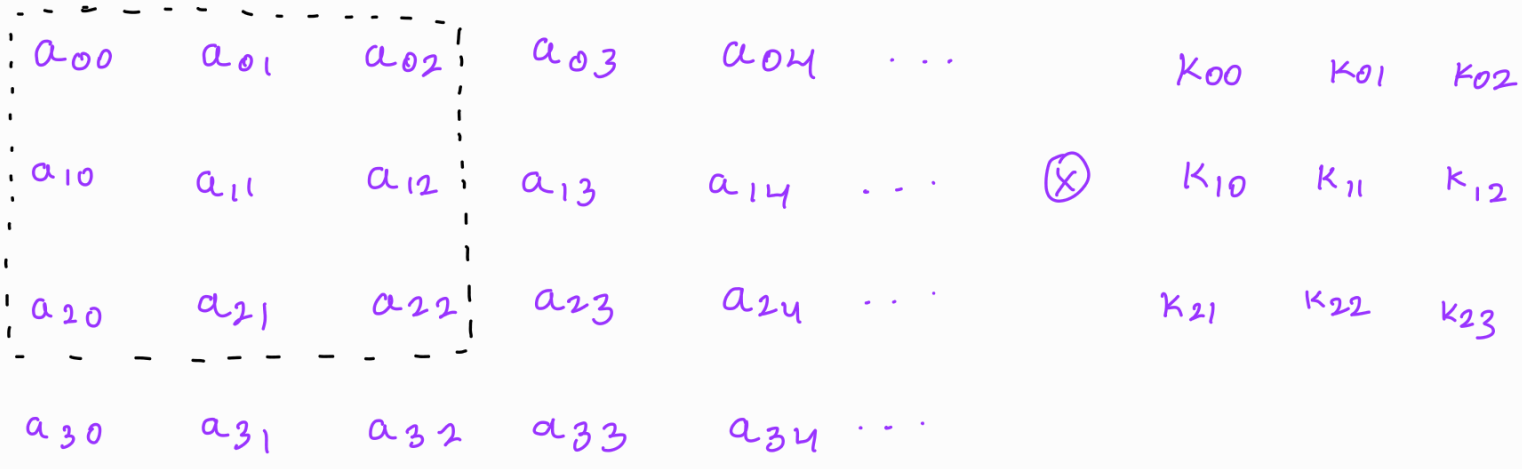


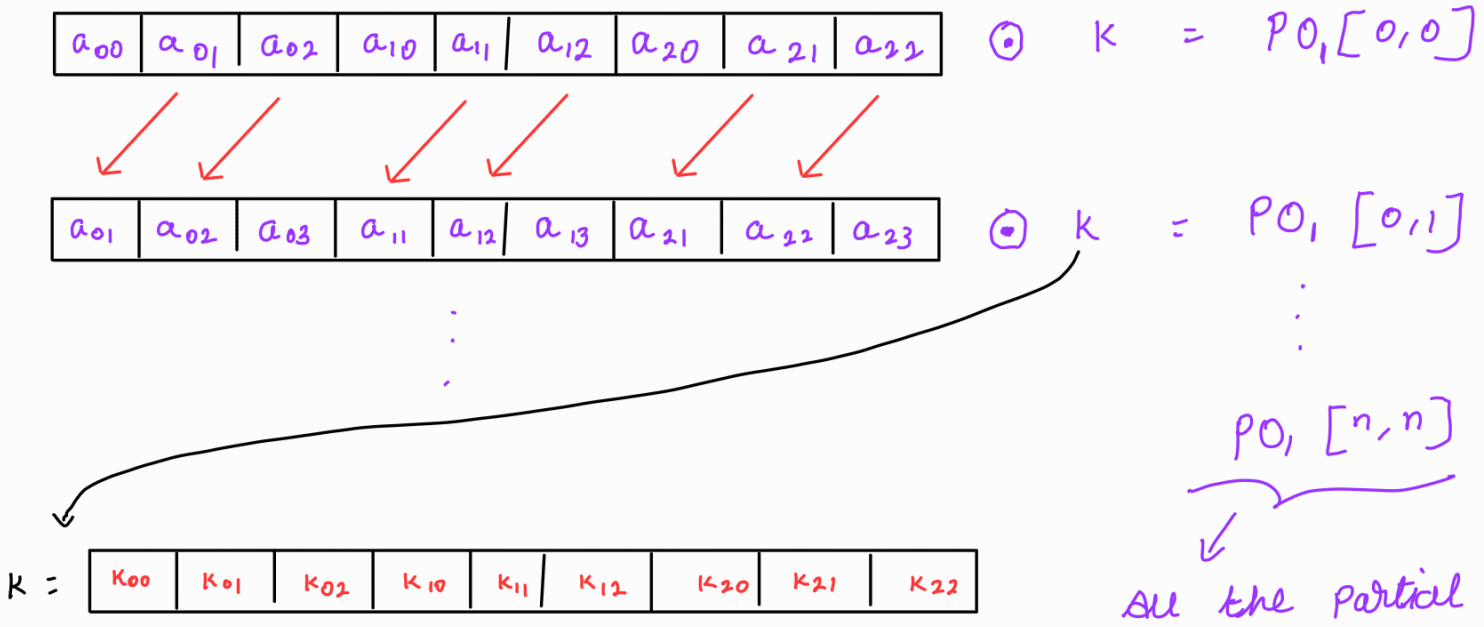
2) Input

→ Fetch only one layer of input

→ Store and reuse it until convolution with layer 1 of all 20 kernels is completed

3) 2D - Convolution





All the partial outputs of 1st channel are calculated

→ Now continue the above process with 1st layer of Kernel 2 → This will give PO_2

→ Once all 20 kernels are used, now we have 20 channels of partial outputs

→ Next use 2nd input layer, and 2nd layer of 20 kernels → This will again give 20 channels of partial outputs, which we need to add with previous partial outputs