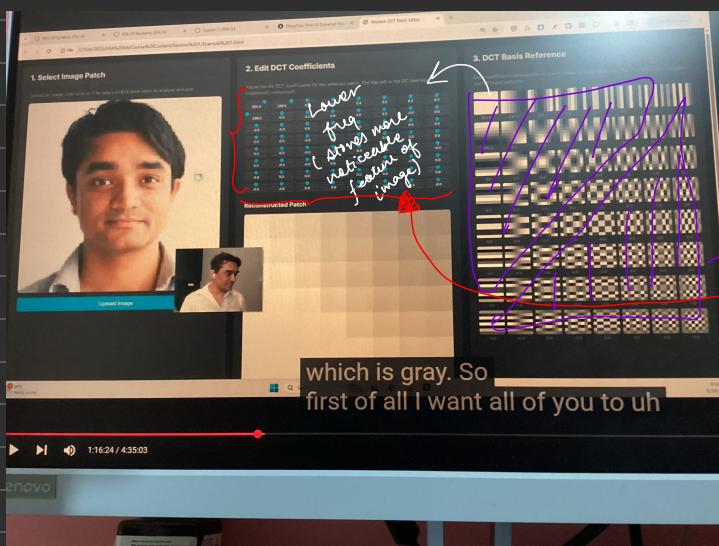


Discrete Cosine Transform:- Any image of size 8×8 can be made using DCT (Discrete cosine transform).



• Main idea behind DCT: Any image can be broken down into these 8×8 boxes or to put it generally, any data can be broken down into smaller set of rules.

→ Higher freq. (stores minute detail)

Image can be broken down into smaller 8×8 boxes

Similarly we can break any data into smaller constituent.

SVD (Rank Approximation):-



Let's say we have different pixels with its importance.

$K = 5 \Rightarrow$ 5 most important pixels.

$K = 10 \Rightarrow$ 10 most important pixels

$K = 20 \Rightarrow$ 20 most important pixels

⋮

$K = 100 \Rightarrow$ Taking complete image

NOTE:- We are going to use this technique to compress our model.

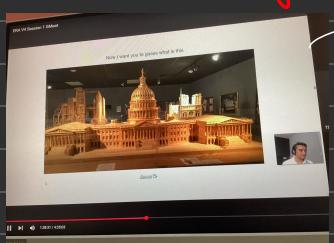
CHANNELS OR BREAKING IT DOWN!:- Any color can be made through a combination of different kinds of "primary color" combinations.

Eg:- RGB, CMYK, CMK, YIQ etc.

RGB is made up of three channels $\begin{cases} R: - \text{stores how much Red is present in each pixel.} \\ G: - \text{stores how much Green is present in each pixel.} \\ B: - \text{stores how much Blue is present in each pixel.} \end{cases}$

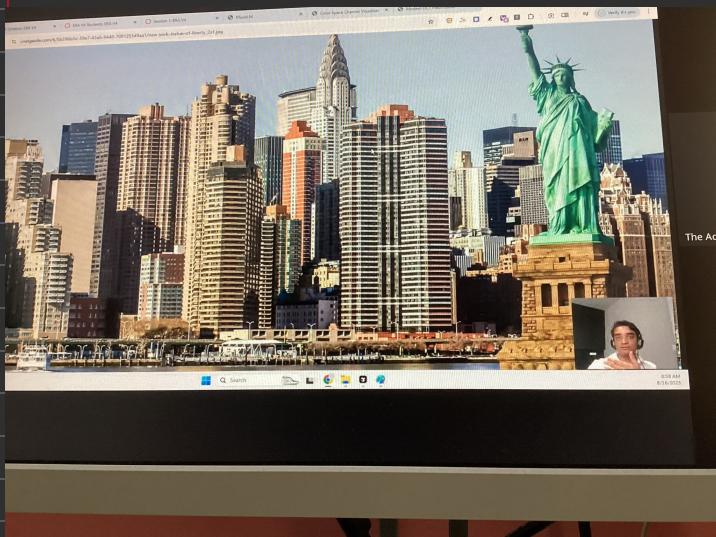
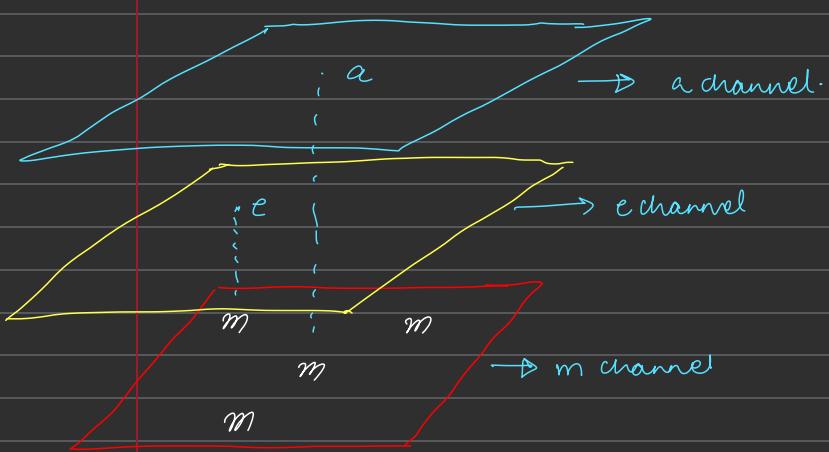
Each channel is a grayscale image where brighter values mean high intensity of that channel.

Made Up By:-



It's made up by tooth picks.

Channel of Effect: 2

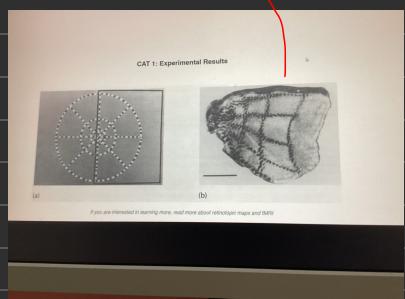
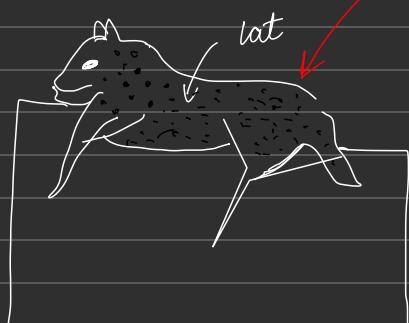
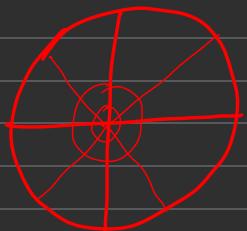


Extra Gyan:- In neural network, the initial layers is to break things down and task of later layer is to combine these things and make something out of it.

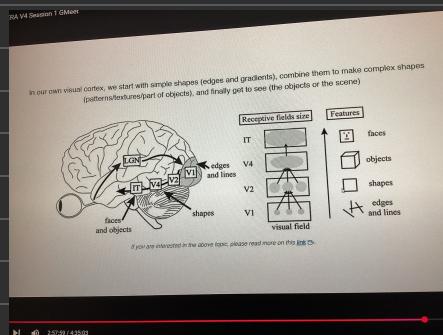


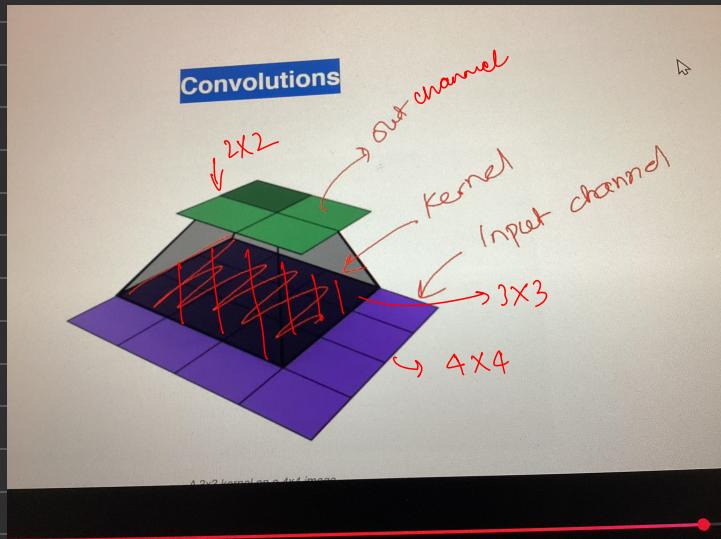
Origin or The story of two missing

Brain of cat!



Same things happens to us as well, when we see a object, like in artificial neural network the initial layers extract edge and line and slowly the more we go deeper the layer combine those to get something meaningful.



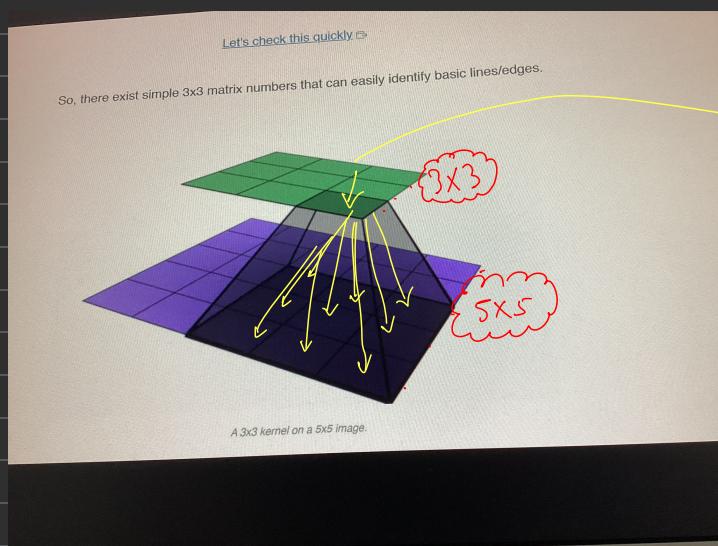


Here,

Input channel: 4×4

Kernel: 3×3

Output channel: 2×2



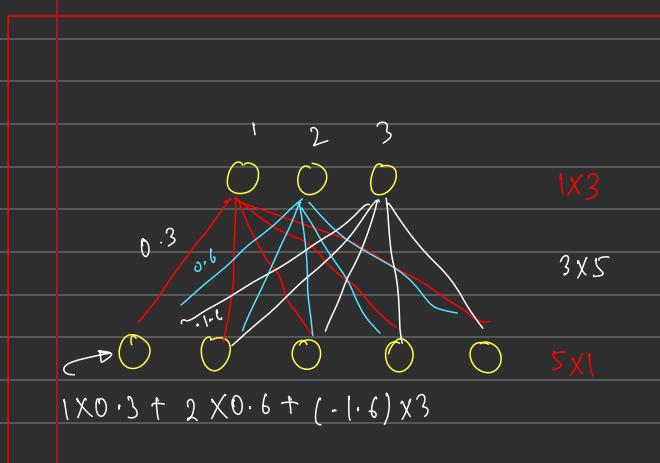
Receptive field of this is 3×3

How computing happens in GPU?

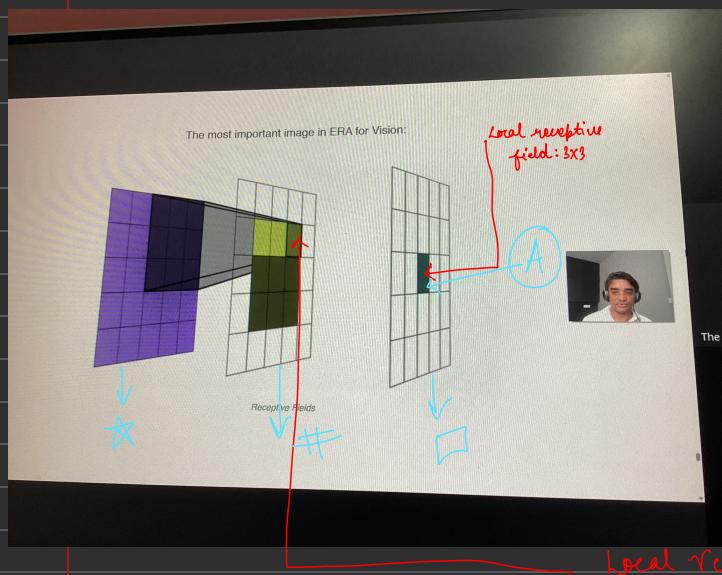
In CPU, it happens sequentially

In GPU, it happens parallelly

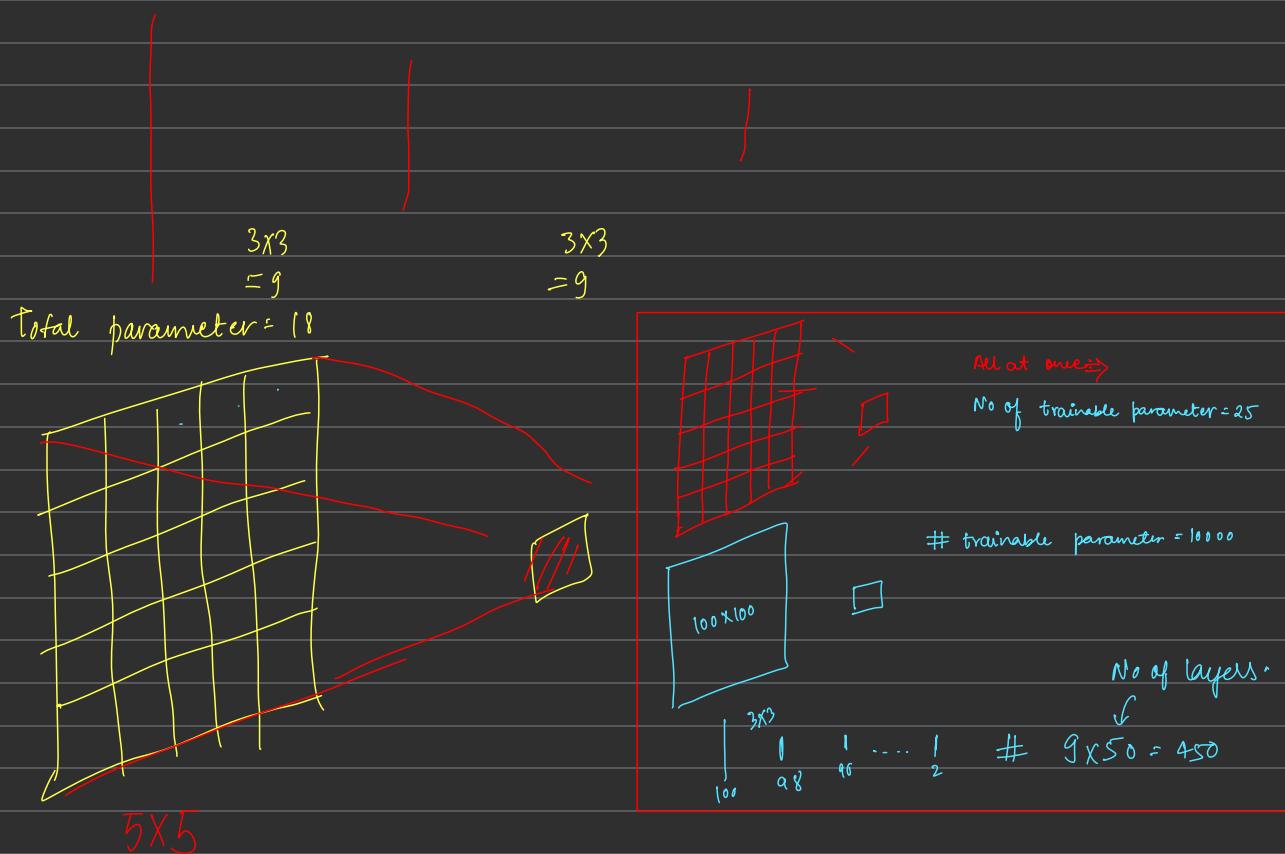
In CPU	In GPU
$4 \times 3 = 12$	$\frac{3}{1}$
$16 \times 3 = 48$	$\frac{4}{2} \rightarrow 12$
$5 \times 3 = 15$	$\frac{16}{4} \rightarrow 48$ 3 steps $\frac{16}{8} \rightarrow 15$ 3 steps



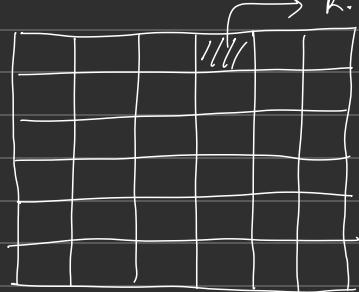
Receptive field & Attention: ➔

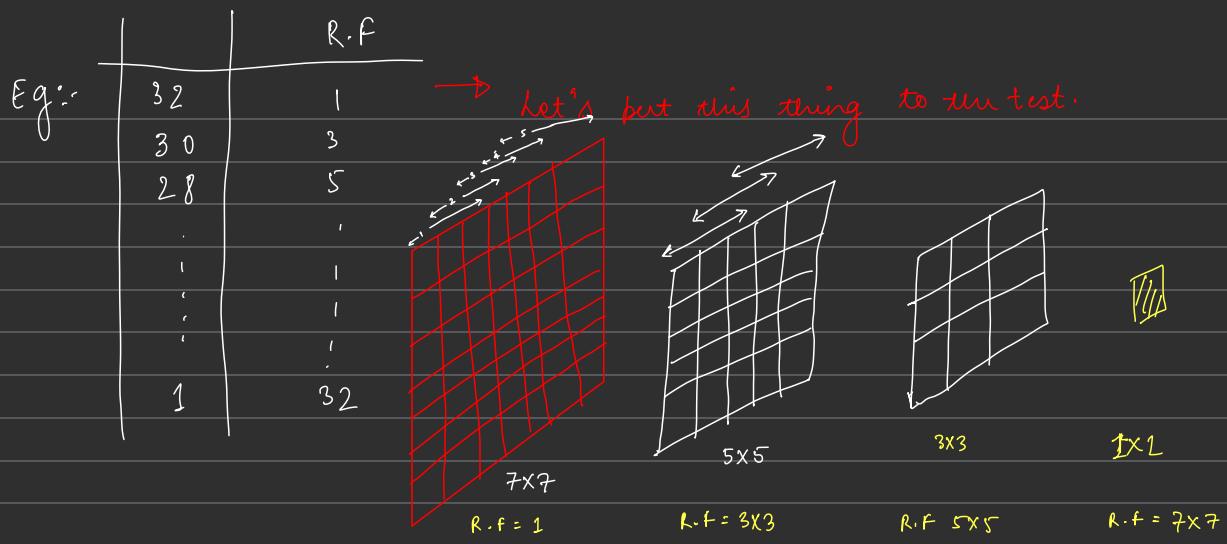


Actual receptive field of (A) is 5×5
because all of the data of \star is getting contained in this # and all # data is getting projected into □



Q:- $R.F. = 1$ (it's getting data from itself)





what is VAT?

