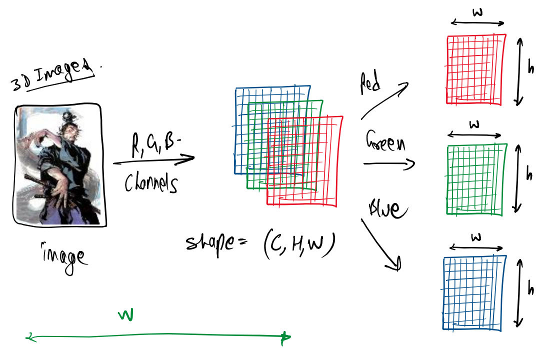
Understanding Images

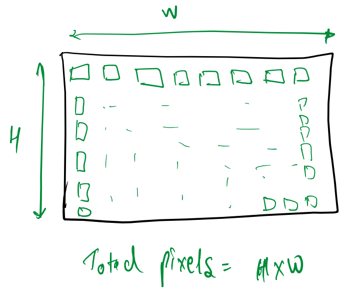
26 December 2025 04:27 PM



Images

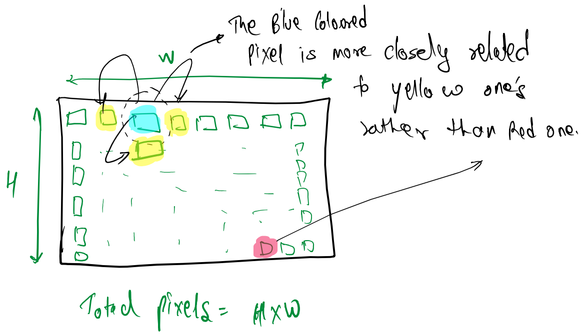
↳ A Discrete Sampling from a continous radiance field.

Physical Properties of Images



a) locality

↳ nearby pixels matters more than distant ones



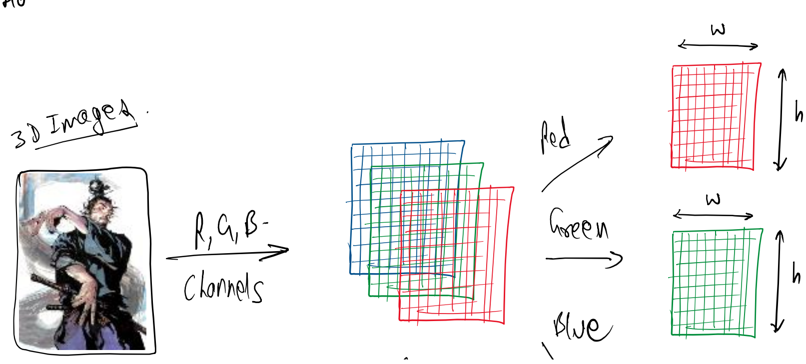
b) spatial feature.

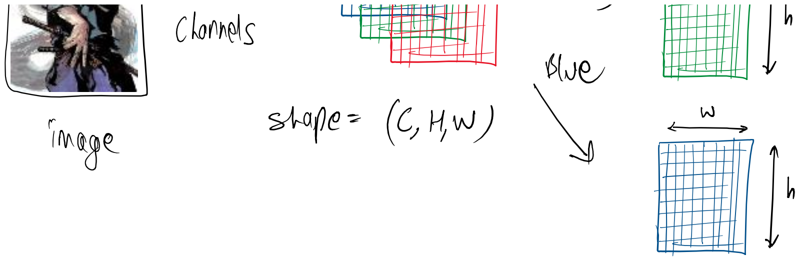
↳ feature formed by neighbourhood pixels

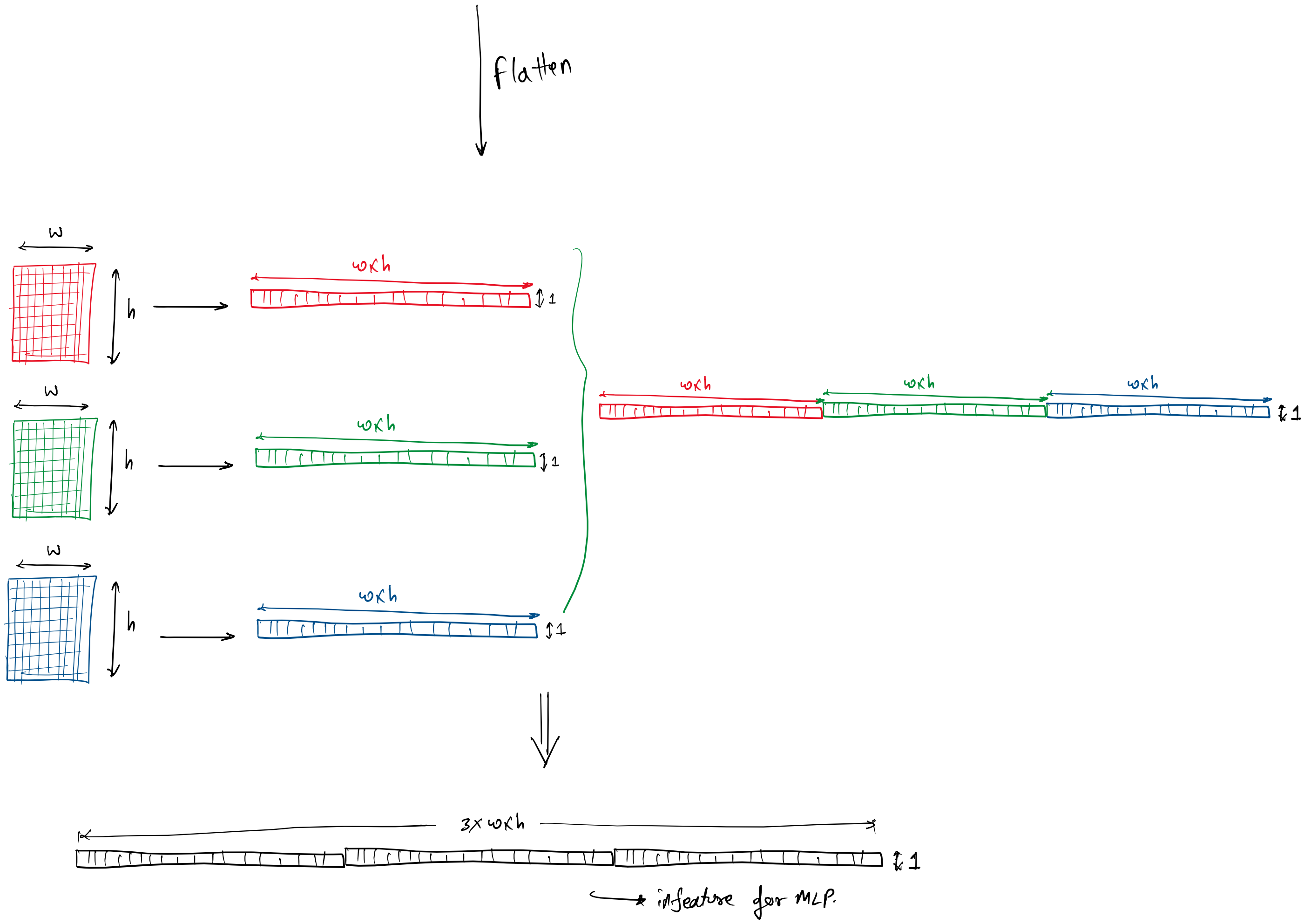
Conclusion:- for learning the Pattern in Images we need to learn the nearby pixels rather than distant ones.

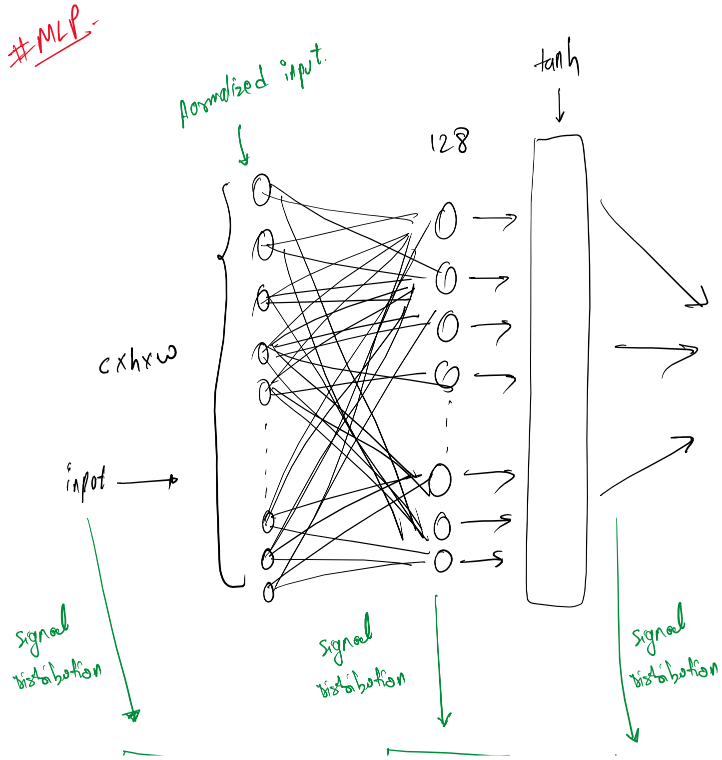
Learning Patterns in Images.

Before learning Pattern in Images let see How Data/Signal flows in MLP (Multi Layer Perceptron.)



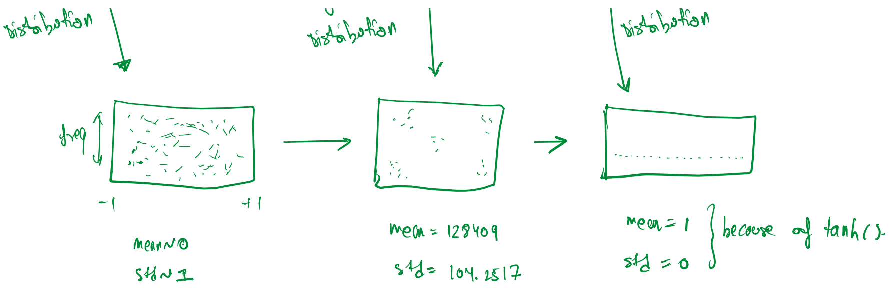






MLP

- Linear Layer :- $y_i = \sum_{i=0}^{N} x_i w_i + b_i$
- Layer Norm :- $\hat{x} \leftarrow \dfrac{\hat{x} - x_{mean}}{\sqrt{x_{var} + \epsilon}}$
- Activation $\longrightarrow$ $f(x) = \tanh(x) = \dfrac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$



Problem:-

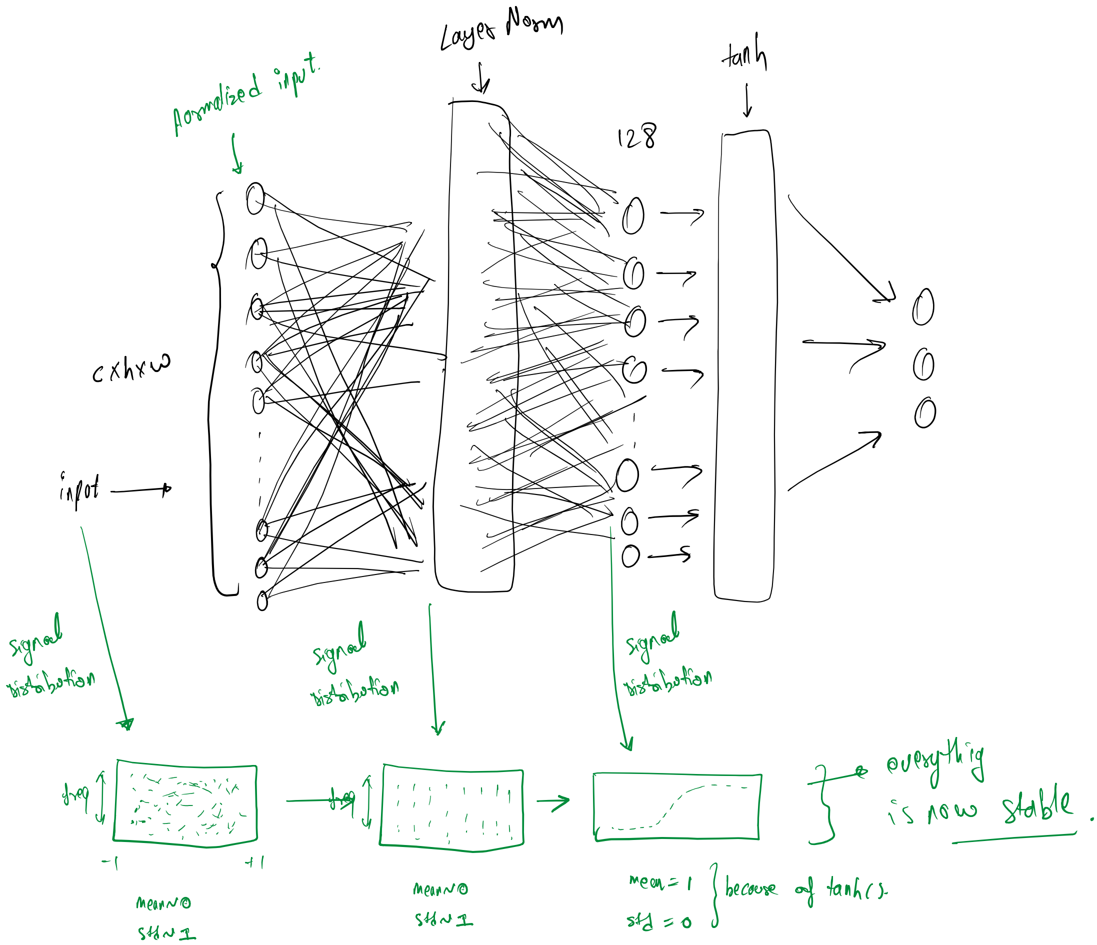
→ Data/Signal's Distribution is always changing and this is Bad for Model.

Solution:-

→ LayerNorm
→ Xavier Initialization
→ He Initialization

} weight Initialization methods.

After LayerNorm and Proper Weight Initialization.



Problem with MLP in Vision Modelling.

Problem with MLP in Vision Modelling.

MLP Assumes :-

Every input Dimension is independent from each other, and equally related to every other Dimension.

example. $x = [x_1, x_2, x_3 \dots x_n]$

for MLP, → x_1 and x_3 are not dependent
→ all 'x' are just a number and co-ordinate in a vector space.

but for images.

$x = [P_1, P_2, P_3 \dots P_n]$

Pixel-1 is closely related than Pixel-2

Images assumes
→ Pixels are related (nearby) } → MLP dsn't care about this. } All relationships must be learned via Data distribution. and Transformation in MLP.

Conclusion :-

→ MLP ignores Geometry
→ Images are Geometry } Mismatch. ⇒ we can't use MLP for understanding patterns in Images.