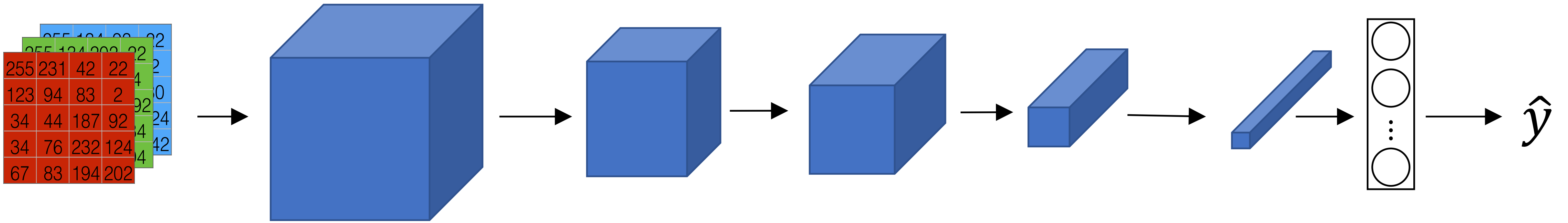
deeplearning.ai

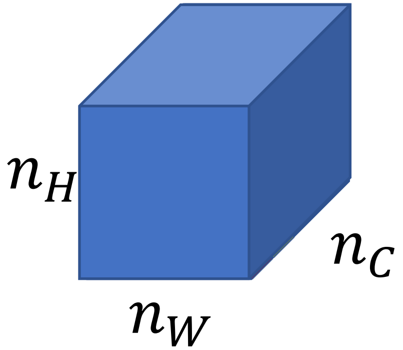
Neural Style Transfer

Style cost function

Meaning of the "style" of an image



Say you are using layer l's activation to measure "style."
Define style as correlation between activations across channels. for a single layer "l"

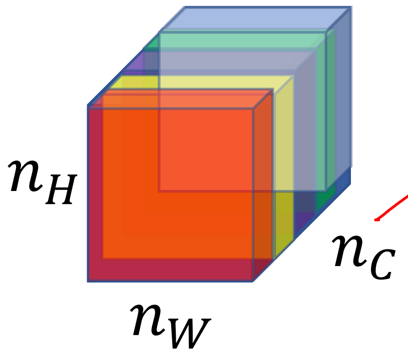


How correlated are the activations across different channels?

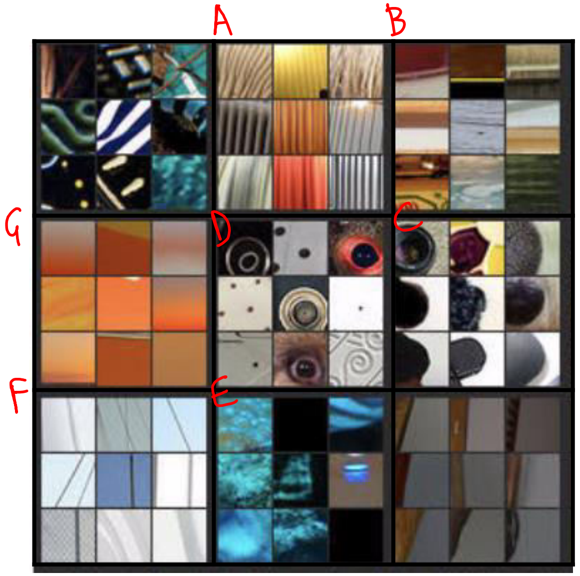
[Gatys et al., 2015. A neural algorithm of artistic style]

Intuition about style of an image

Style image



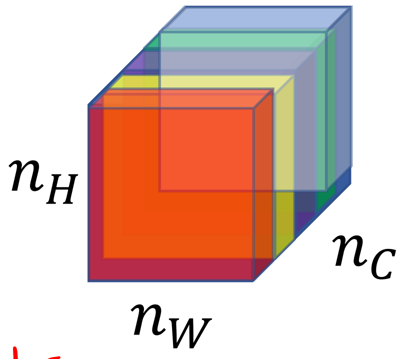
Say $n_C = 5$ then you have 5 blocks of $n_H \times n_W$ pixels



Say you take the Red & yellow channels & take each of their $n_H \times n_W$ cells → Now you consider position wise activations of these cells & try to find correlation b/w them

- Say you find 2 neurons at 2 corresponding positions - one from the Red block, one from the yellow block

Generated Image



- Say Red channel gives you neuron ABCD, yellow gives DEFG

* ABCD → Identifying vertical lines
* DEFG → Identifying orange color

Q How correlated are ABCD & DEFG?

correlated → whenever there is vertical texture, it has an orange-ish color (in part of an Img)

[Gatys et al., 2015. A neural algorithm of artistic style]

Style matrix

index for height
width
channel

Let $a^{[l]}_{i,j,k}$ = activation at (i, j, k). $G^{[l]}$ is $n_c^{[l]} \times n_c^{[l]}$

Correlation

How often do these texture components (vertical lines / color patterns / Rounded edges) occur together

eg In an Img of a water-melon - it is likely that vertical lines neurons & green color neurons correlate well together

- It is also likely that black dot neurons & neurons identifying the color "Red" occur together

$\Rightarrow$ degree of correlation of channels can be used to measure style

- We want to minimize the dis-similarity b/w the original Img & generated Image in terms of their styles, ie, we want the style correlations b/w the channels of the original Img match that of the generated Img

$G^{[l]}_{kk'}$ has the correlation score for the Image b/w Image's channel "k" and Images channel "k'"

$k, k' \in [1, n_c^{[l]}]$ (PTO)

[Gatys et al., 2015. A neural algorithm of artistic style]

Andrew Ng

Style cost function

For layer l

$$J_{style}^{[l]}(S,G) = \frac{1}{\left(2n_H^{[l]}n_W^{[l]}n_C^{[l]}\right)^2}\sum_k\sum_{k'}\underbrace{\left(G_{kk'}^{[l](S)} - G_{kk'}^{[l](G)}\right)}_{\text{element wise}}$$

$$\sum_{i=1}^{n_H^{[l]}}\sum_{j=1}^{n_W^{[l]}} a_{ijk}^{[l](S)} \cdot a_{ijk'}^{[l](S)}$$

$$\sum_{i=1}^{n_H^{[l]}}\sum_{j=1}^{n_W^{[l]}} a_{ijk}^{[l](G)} \cdot a_{ijk'}^{[l](G)}$$

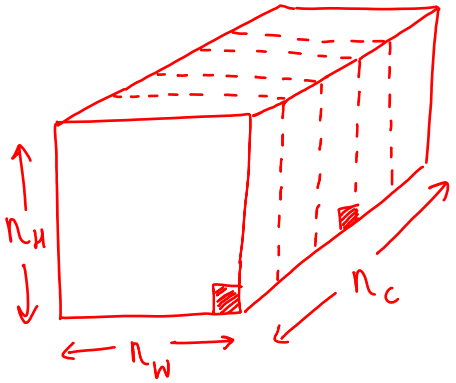
$$G_{kk'}^{[l]} = \sum_{i=1}^{n_H}\sum_{j=1}^{n_W} a_{ijk}^{[l]} \cdot \underbrace{a_{ijk'}^{[l]}}$$

Style matrix

If both a_{ijk} & $a_{ijk'}$ are large, then $G_{kk'}$ is large $\Rightarrow$ highly correlated

for index height i, index width j and channel index k

for the same position of Image i, j, but different channel k'
– one example is the 2 Red boxes



☆ Technically, we calculate cross co-variance & not correlation

We calculate this G matrix (Style matrix) on both style Image "S" & Generated Image "G"
[G matrix also called Gram matrix]

Total style cost

$J_{style}(S,G)$ = sum over all layers

$= \sum_l \lambda^{[l]} \cdot J_{style}^{[l]}(S,G)$

$\lambda^{[l]}$ gives relative Importance to diff layers

Final total cost

$J(G) = \alpha \cdot J_{content} + \beta \cdot J_{style}$

[Gatys et al., 2015. A neural algorithm of artistic style]