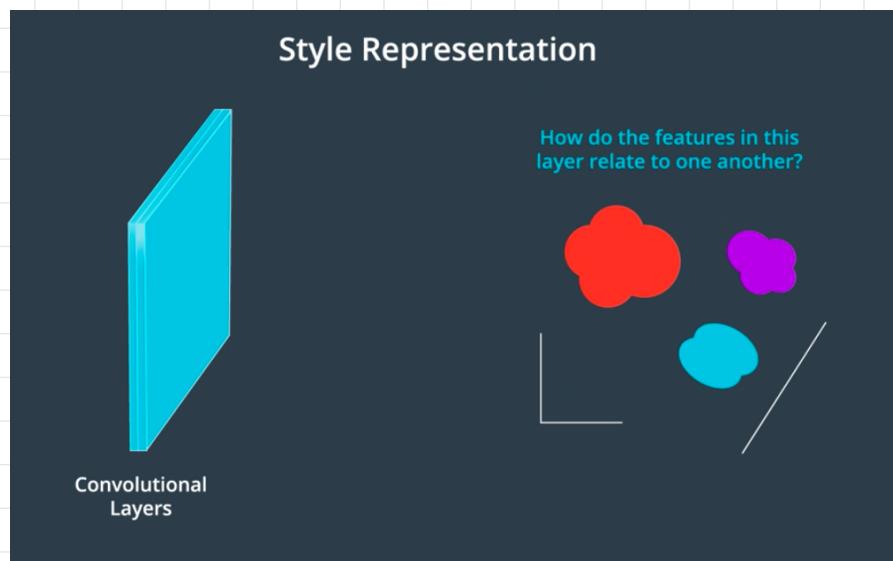
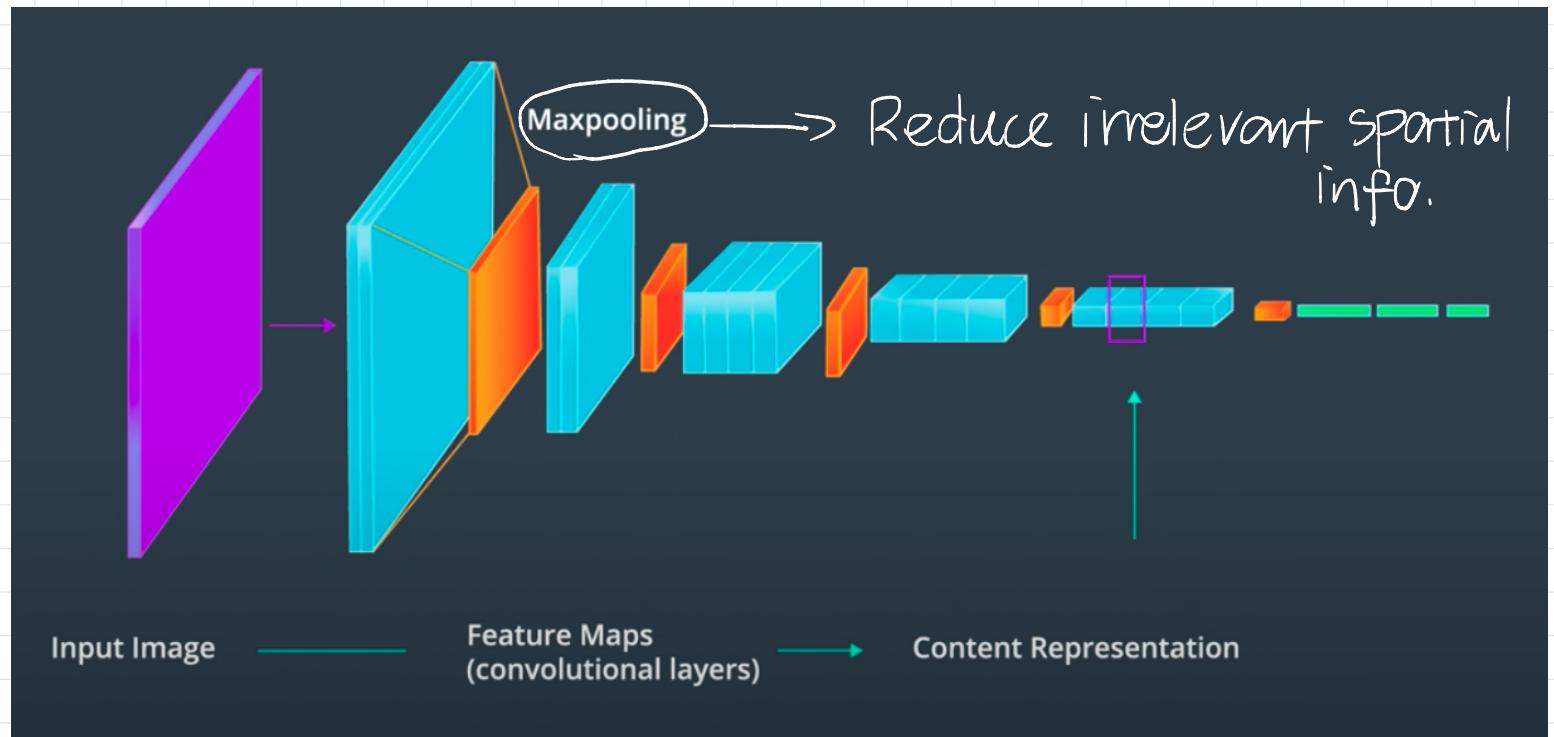


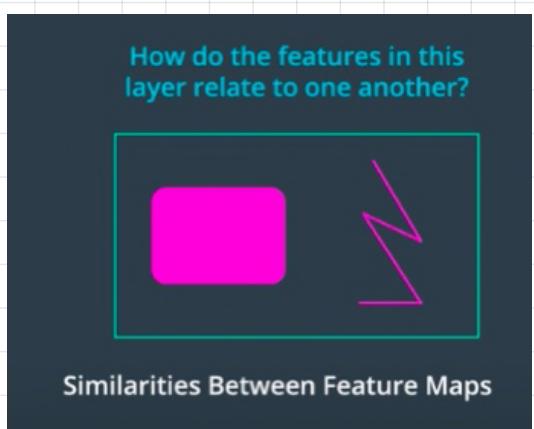
Style Transfer



Separating Style & Content

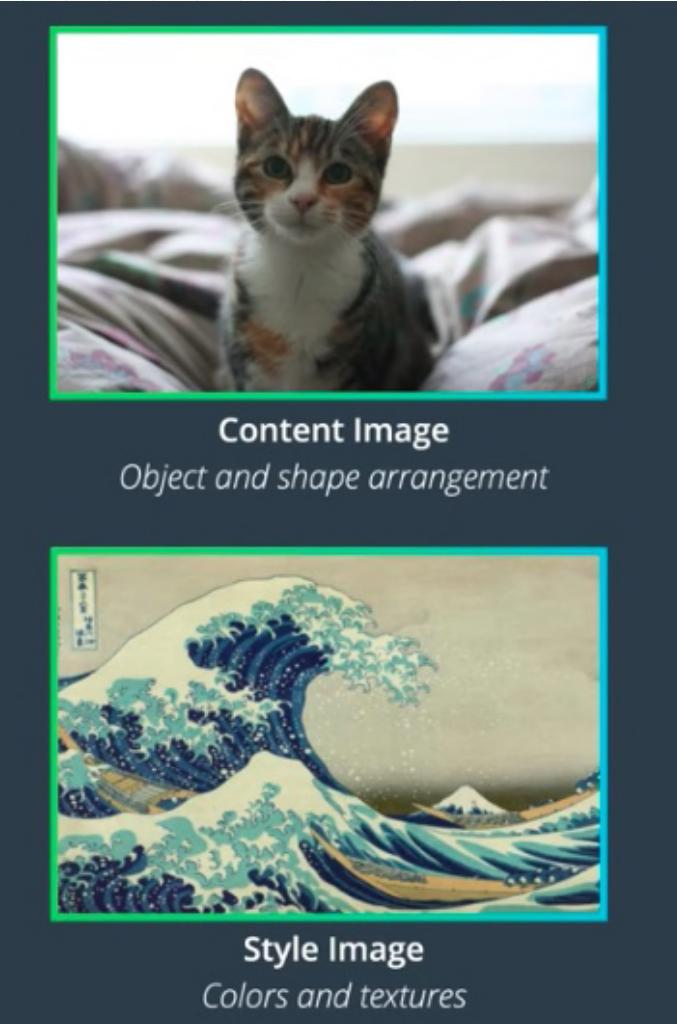


For each feature map, measure how strongly its detected features relate to the other features.
or w/ other feature map.



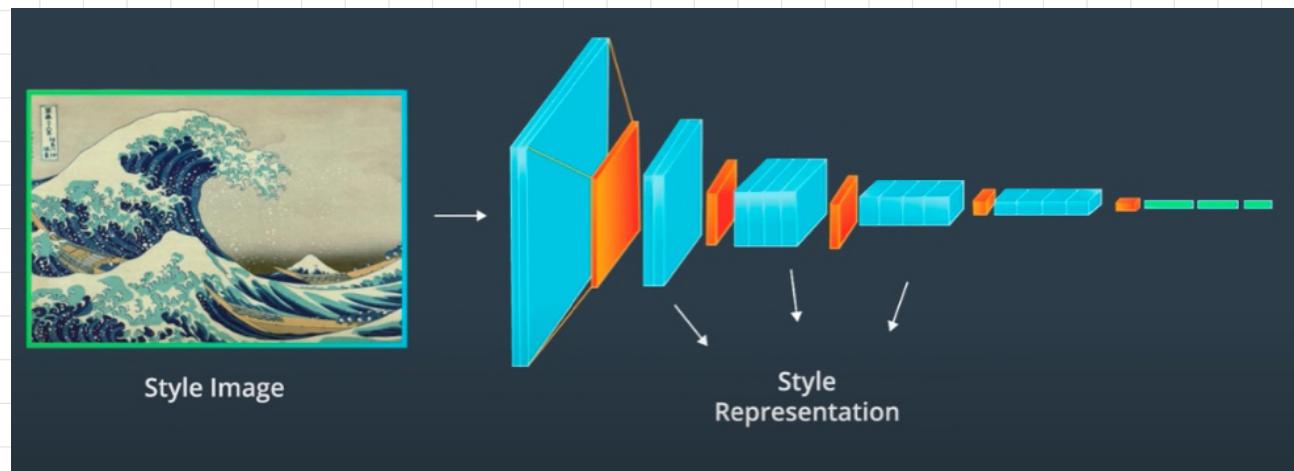
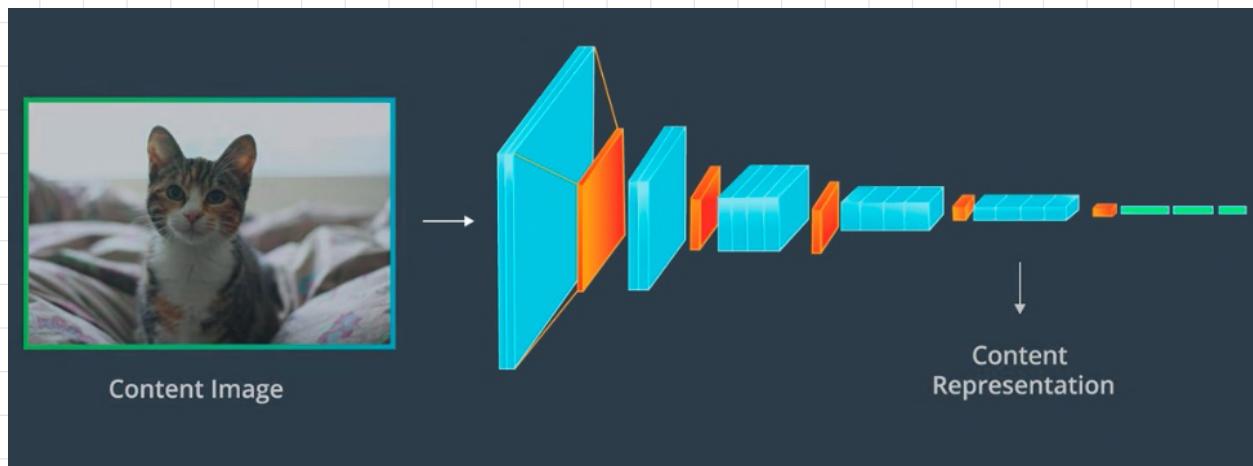
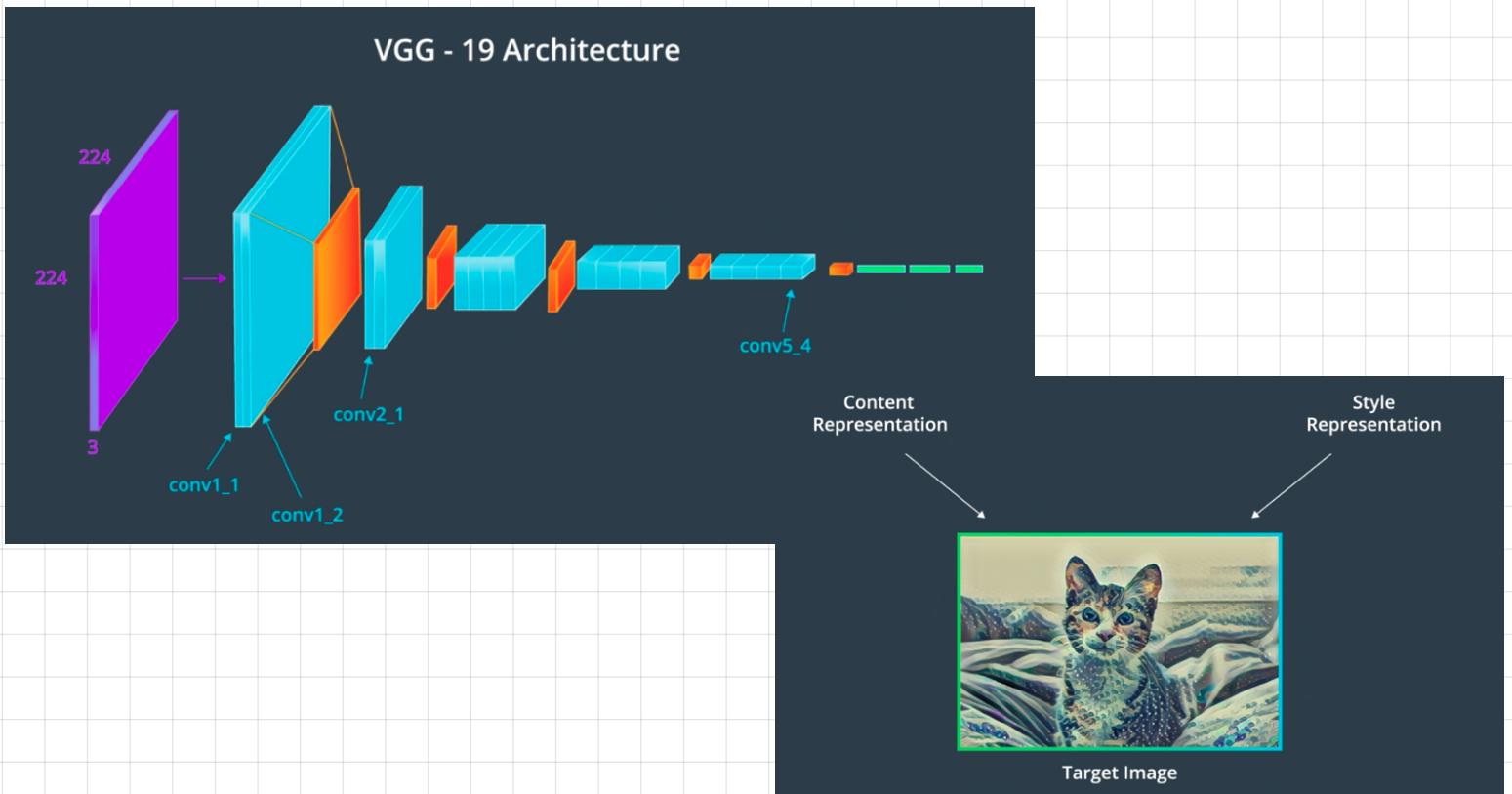
For instance, diff. feature maps show pink color. So this may be the style of the img.

Should not tell us anything about the identity or placement of objects in an img.

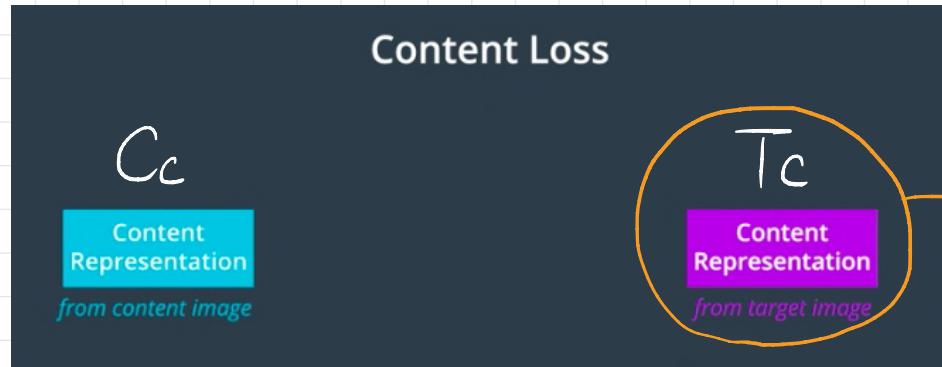


Merge style of one img. w/ content of another

VGG-19 & Content Loss



Content Loss



$$L_{content} = \frac{1}{2} \sum (T_c - C_c)^2 \Rightarrow \text{To minimize}$$

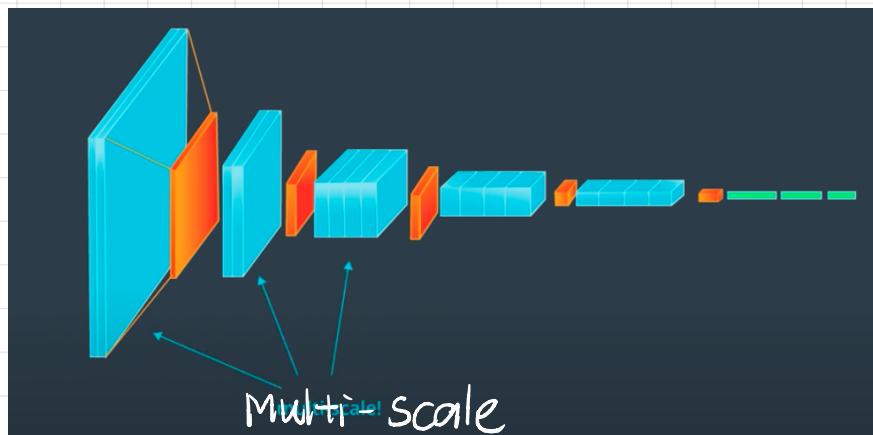
However, we aren't training the CNN.

- * Recall from the VGG19 representation in previous slide, use VGG19 as a feature extractor since we slice off the classifier portion!
- * Still using backprop!
- * Same steps when creating style representation.

Gram Matrix

- Recall content loss.
- Style representation looks at correlations bt. features in individual layers of VGG-19.
 - How similar the features in a single layer are

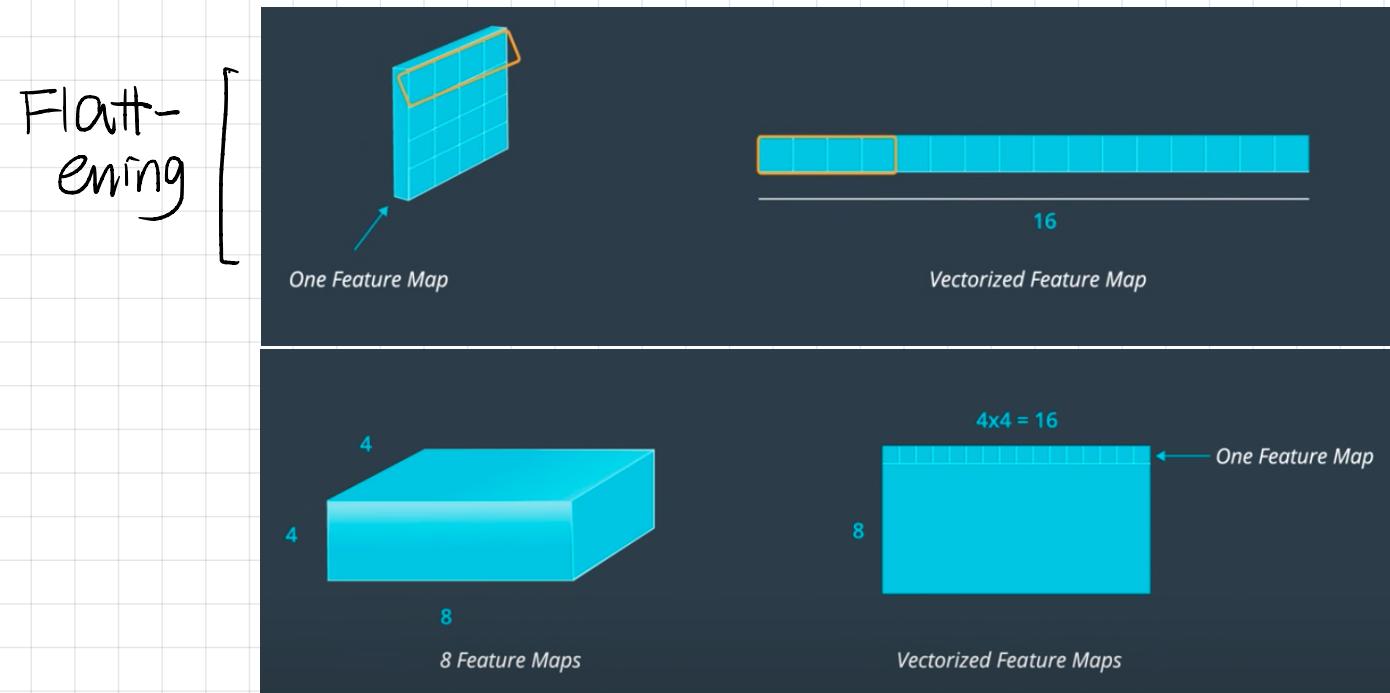
Color Texture.



Multiscale!

Evaluating correlation of features in layers of diff. sizes.

The Math





- Each value in feature map = unrelated to space & other values → non-localized info.
- G_{42} = similarity bt. 4th & 2nd feature map of that layer.
- Dim of matrix \propto # of feature maps, not \propto dim of input images.

Style Loss

Mean Squared Distance bt.

Style img Gram Matrix
Target img GM

CONV1_1 ~ conv5_1

5 pairs of GMs.

Style Loss

S_s T_s

Account for # of values in each layer

$$\mathcal{L}_{style} = a \sum_i w_i (T_{s,i} - S_{s,i})^2$$

Only thing that changes.

style weights

Given at each layer Changing how much each layer style representation will have on our final img.

Total Loss

$$\mathcal{L}_{content} = \frac{1}{2} \sum (T_c - C_c)^2 + \mathcal{L}_{style} = a \sum_i w_i (T_{s,i} - S_{s,i})^2$$

- Apply backprob & optimize, train on iterations to decrease this loss.

Loss Weights

- $L_{content} \propto L_{style}$ will be pretty different
- We need the targeting to take in both consideration equally

