

Neural Style



Gatys, Ecker, Bethge. "A Neural Algorithm of Artistic Style"
(<http://arxiv.org/abs/1508.06576>)

CSC321: Intro to Machine Learning and Neural Networks, Winter 2016

Michael Guerzhoy

Neural Style

- Task: given an input photo I and a painting P , produce a photo with the same contents as I , but with the style P
- Idea: use gradient descent again to change the input with the weights of the ConvNet constant, with a cost function that keeps x close to I (i.e., preserve the content) and makes the style of x close to the style of P
- Use a ConvNet that works for image classification
 - It knows how to represent images well

Cost function: content

- To keep the content close to the input I , make sure that $|x - I|^2$ stays small
 - Not the best idea! The pixels might be a completely different colour in the x , but x and I can still be similar in content
- Even better: make sure that all the activations in all the different layers for the original image and for x stay the same:
- Make sure that $L_{content}(x, I) = \sum_{l,i,j} (F(x)_{i,j}^l - F(I)_{i,j}^l)^2$ is small
 - $F(y)_{(i,j)}^l$ is the activation at layer l at location j in feature map i , for input y
- Will make sure that the high-level features in the image stay the same, too

Style

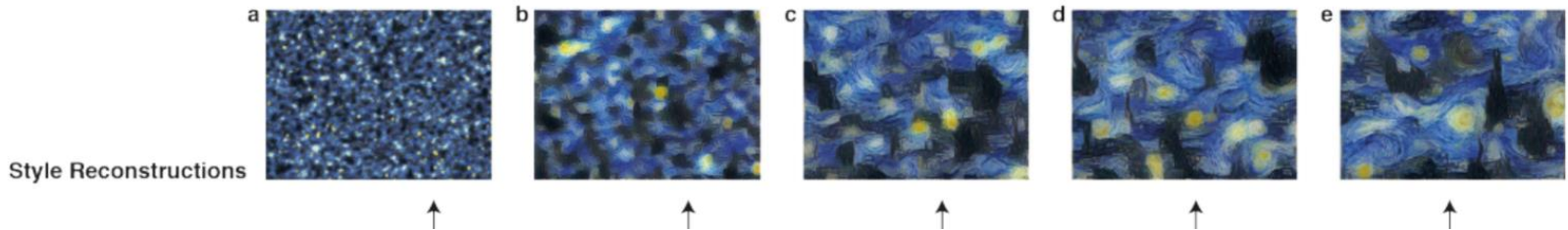
- Define the Gram matrix at layer l as

$$G_{ij}^l(\mathbf{y}) = \sum_k F(\mathbf{y})^l_{ik} F(\mathbf{y})^l_{jk}$$

- Discovery: the Gram matrix represents the style
 - $G_{ij}^l(\mathbf{y})$ is large if at layer l , the i -th feature and the j -th feature tend to be discovered together a lot
 - (i.e., a lot of k 's for which the product of both activations is large)

Style Reconstructions

(explanation in a couple of slides)



- “Blue strokes at 40 degrees co-occur with blue strokes at 45 degrees?”

Cost function: style

Divide by the size of the layer to make sure every layer influences things equally for now

- $E_l(x, P) = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l(x) - G_{ij}^l(P))^2$
- $L_{style}(x, P) = \sum_l w_l E_l(x, P)$

Cost function: overall

- $cost(x, I, P) = \alpha L_{style}(x, P) + \beta L_{content}(x, I)$