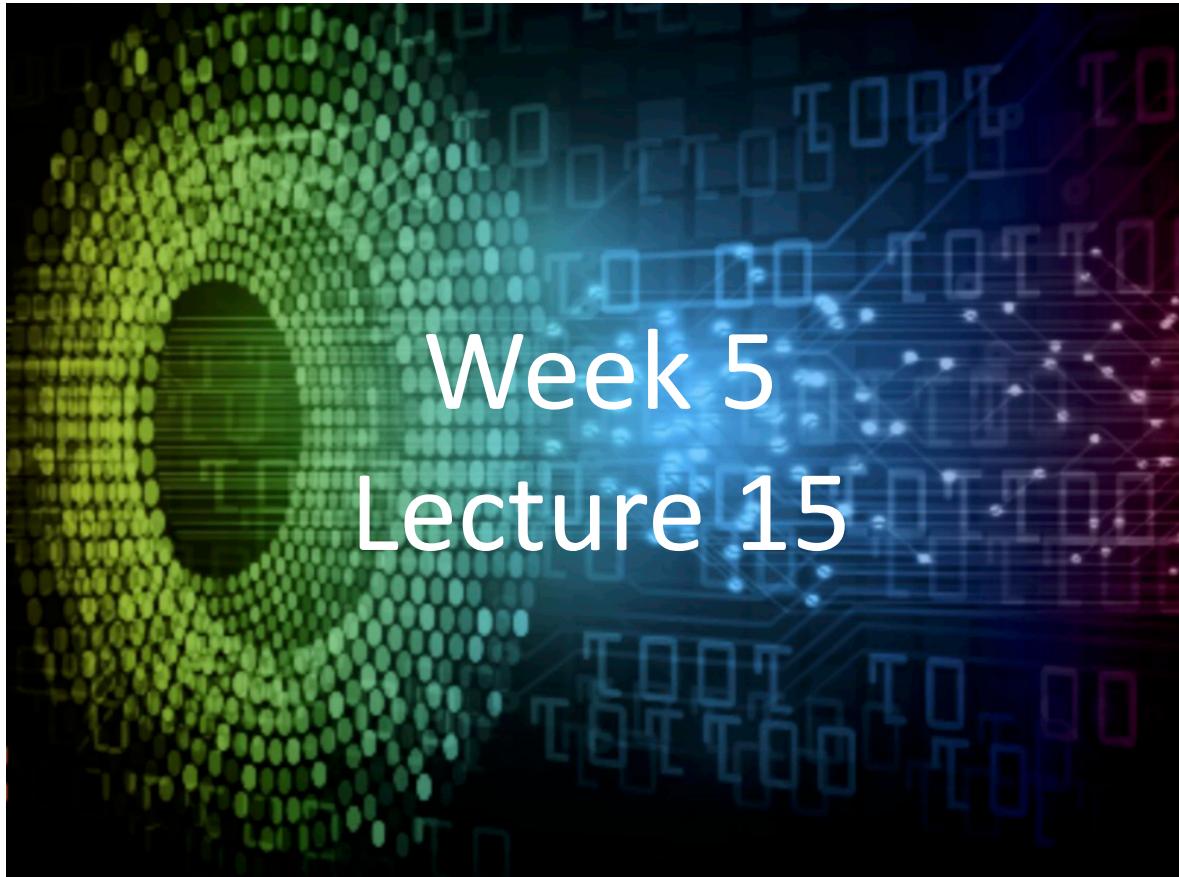


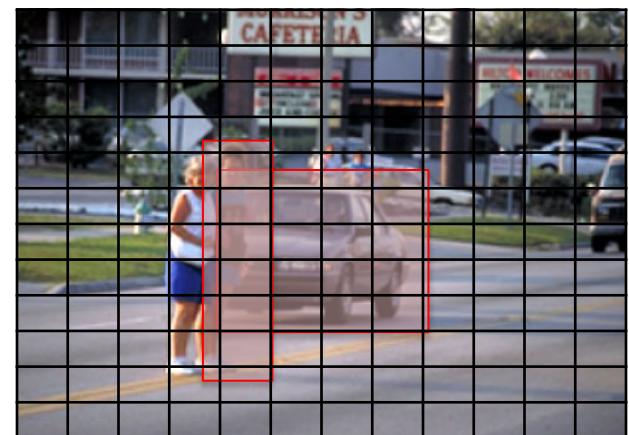
# Introduction to Deep Learning Applications and Theory



ECE 596 / AMATH 563

# Previous Lecture: Advanced CNNs

- Generalization of CNN
  - Complex Output
  - Extension of output through FCONV
- Complex Problems
  - Localization
  - Detection



Anchor boxes

# This Lecture: Interpretation of CNNs

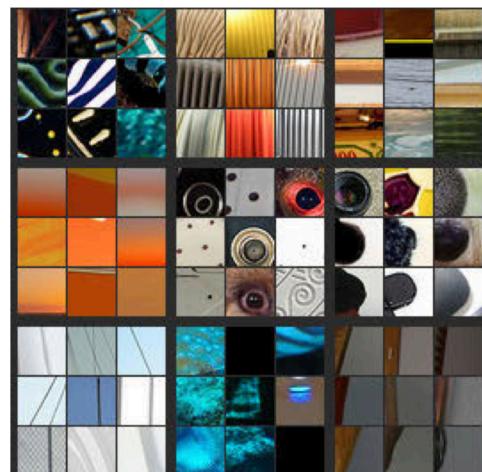
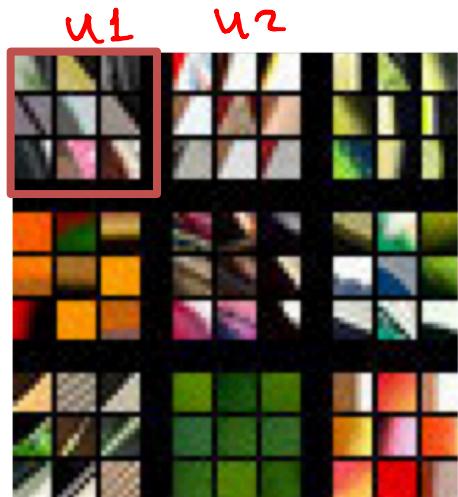
## Pre-Trained CNNs

- Latent (Hidden) Features
  - Linking hidden layers to input
  - Visualization
- Neural Transfer
  - Content Loss
  - Style Loss

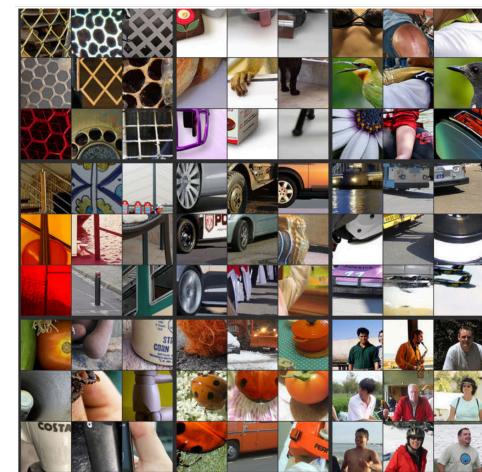
# What is encoded by CNN?

Pick AlexNet like network

For a unit  $k$  in layer  $l$  plot the patches in the training set image that maximizes its activation



Layer1, 9 units



Layer2, 9 units

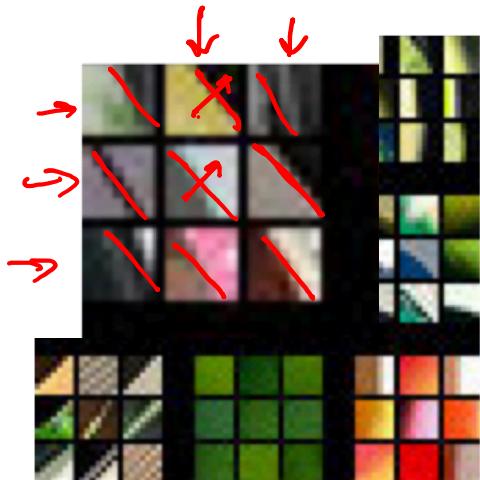
Layer3, 9 units

M.Zeiler, R. Fergus, Visualizing and Understanding Convolutional Networks, CVPR 2013

# What is encoded by CNN?

Pick AlexNet like network

For a unit  $k$  in layer  $l$  plot the patches in the training set image that maximizes its activation

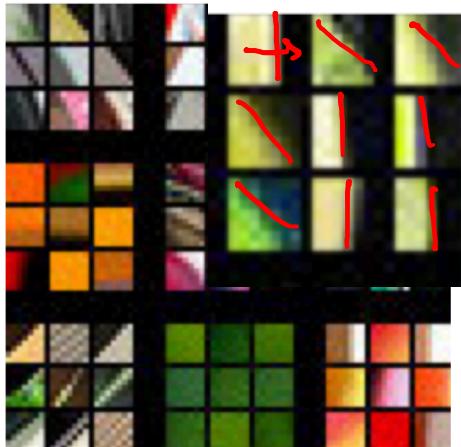


Layer1, 9 units

# What is encoded by CNN?

Pick AlexNet like network

For a unit  $k$  in layer  $l$  plot the patches in the training set image that maximizes its activation



Layer1, 9 units

# What is encoded by CNN?

Pick AlexNet like network

For a unit  $k$  in layer  $l$  plot the patches in the training set image that maximizes its activation

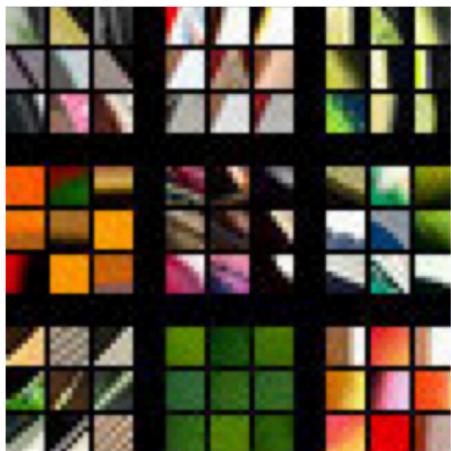


Layer1, 9 units

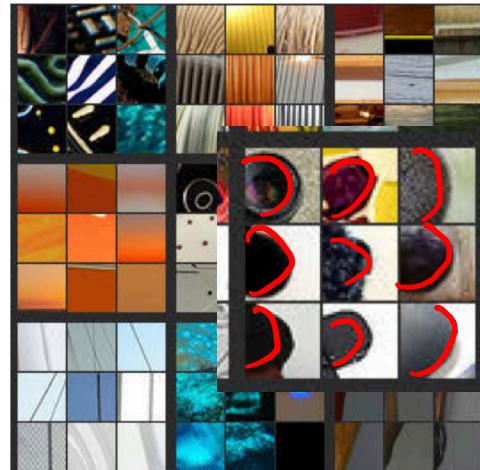
# What is encoded by CNN?

Pick AlexNet like network

For a unit  $k$  in layer 1 plot the patches in the training set image that maximizes its activation



Layer1, 9 units



Layer2, 9 units

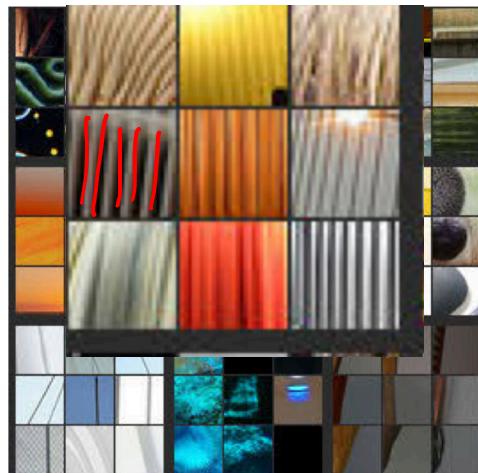
# What is encoded by CNN?

Pick AlexNet like network

For a unit  $k$  in layer  $l$  plot the patches in the training set image that maximizes its activation



Layer1, 9 units

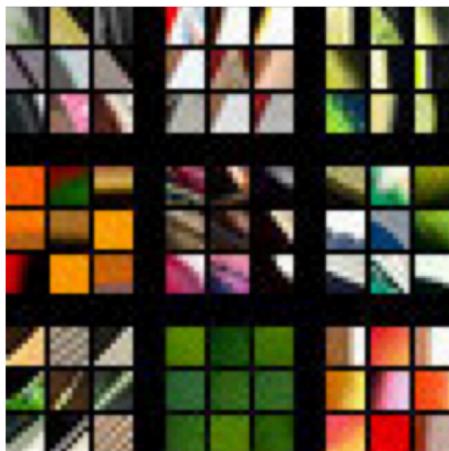


Layer2, 9 units

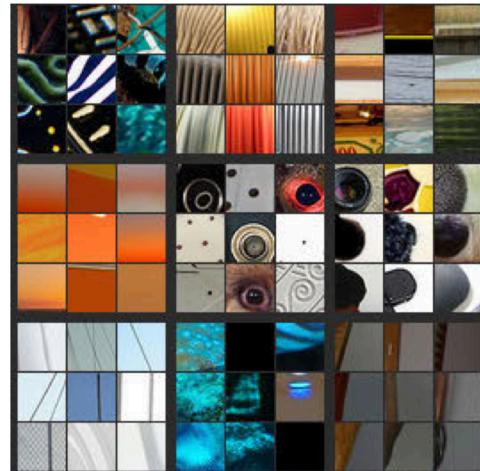
# What is encoded by CNN?

Pick AlexNet like network

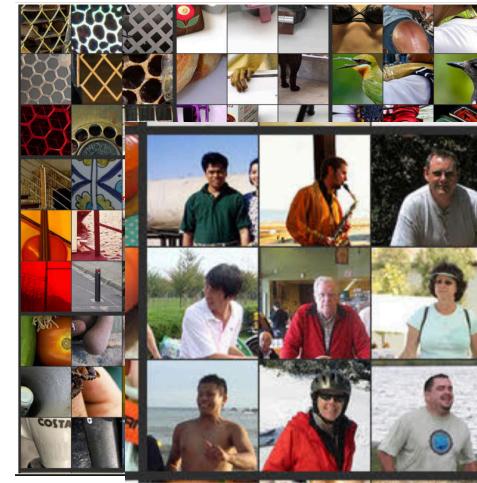
For a unit  $k$  in layer  $l$  plot the patches in the training set image that maximizes its activation



□  
Layer1, 9 units



□  
Layer2, 9 units

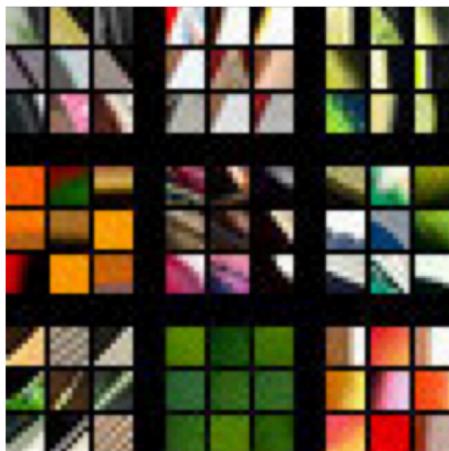


□  
Layer3, 9 units

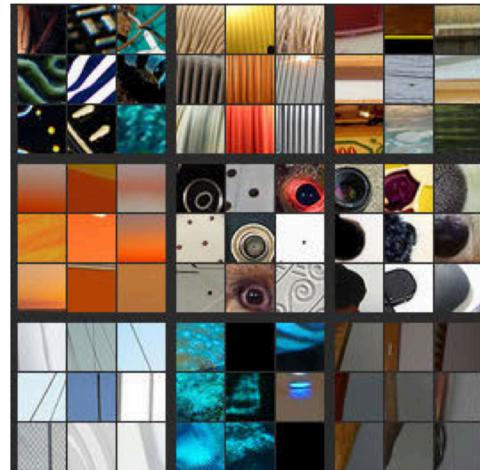
# What is encoded by CNN?

Pick AlexNet like network

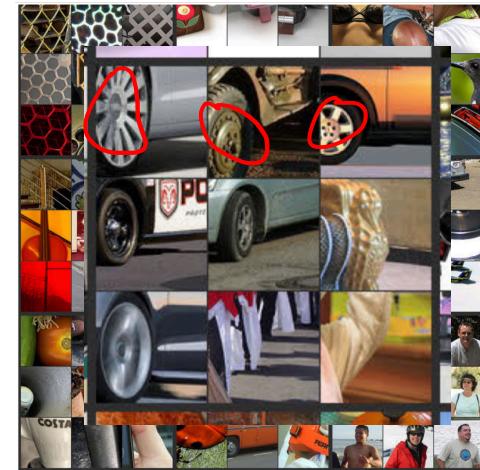
For a unit  $k$  in layer  $l$  plot the patches in the training set image that maximizes its activation



Layer1, 9 units



Layer2, 9 units



Layer3, 9 units

# Neural Style Transfer

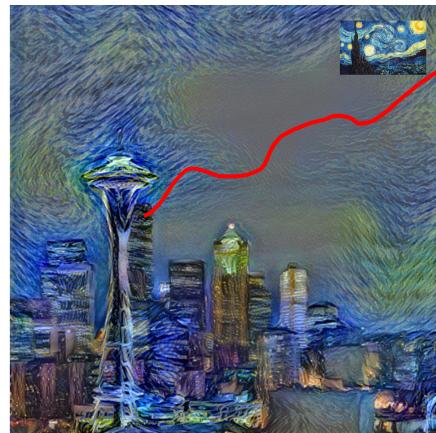
---



Content (C)



Style (S)



Generated (G)

*Generative*

(c) Steve Murch

# Neural Style Transfer

- Define a cost function with two components

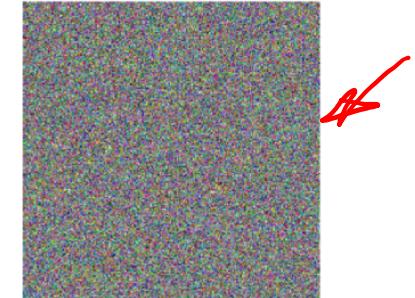
$$\underline{J(G)} = \alpha \underline{J_c(C, G)} + \beta \underline{J_s(S, G)}$$

= content + style losses

# Neural Style Transfer

- Algorithm:

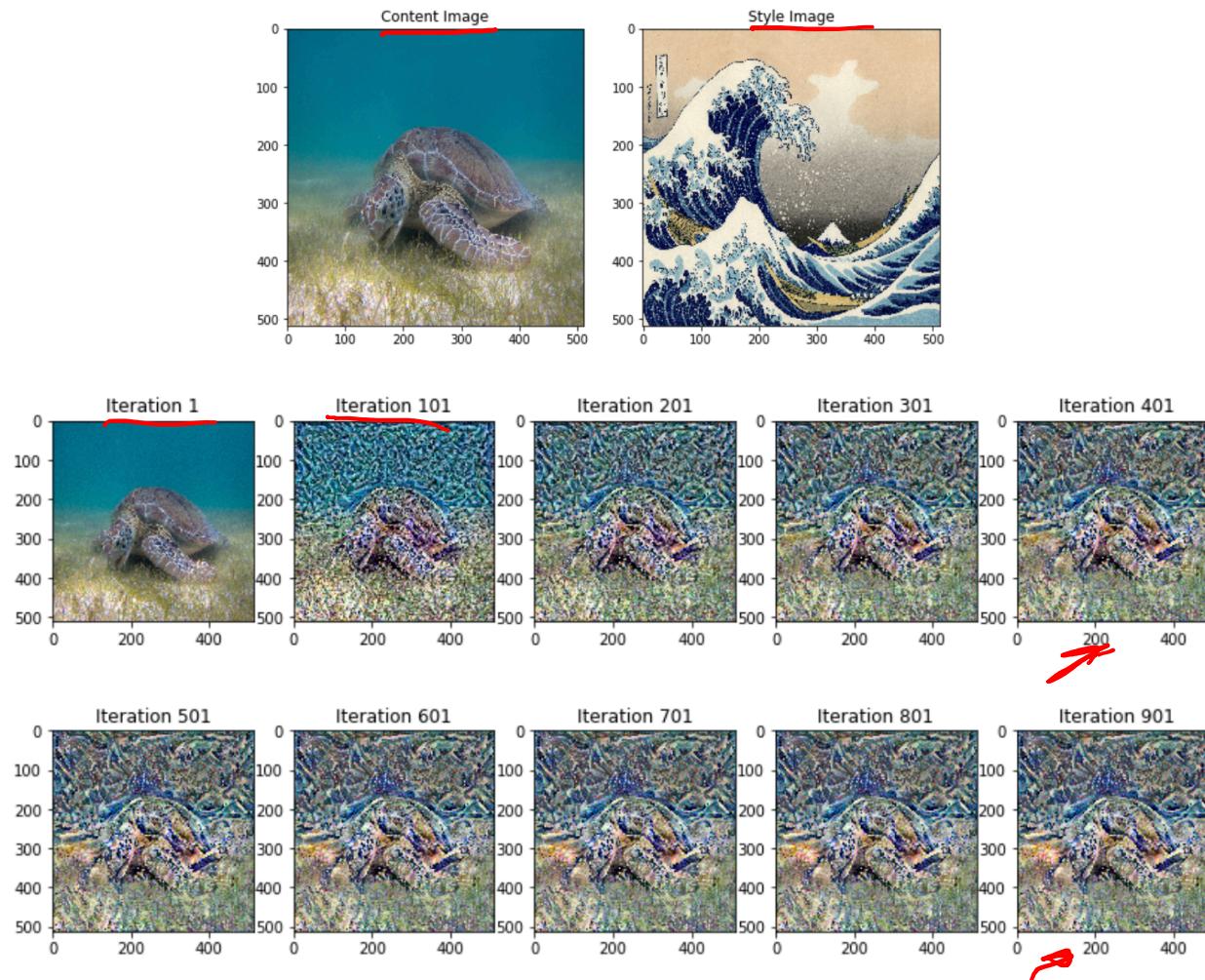
1. Initiate  $G$  randomly =  $(\underbrace{100, 100, 3})$



2. Use GD to minimize  $J(G)$

$$G_{k+1} = G_k - \nabla_G J(G)$$

# Neural Style Transfer



# Neural Style Transfer

- Content Loss:

- Use pretrained object classification/recognition Network (VGG)
- How similar the activations of C and G when they propagate through layer l?

$$\underline{J_c(C, G)} = \frac{1}{2} \|\underline{a^{[l][C]} - a^{[l][G]}}\|_2$$

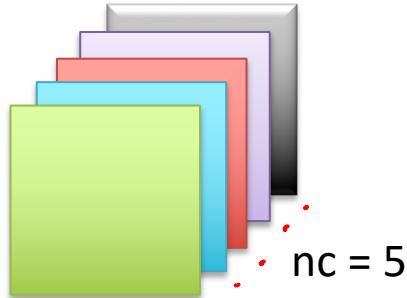
*layer*

↑    ↓

1

# Neural Style Transfer

- Style Loss:
  - “Style”: how consistent are activations across channels?
  - Define pairwise correlation between activations: how often texture components occur together



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

# Neural Style Transfer

$$\begin{aligned} J_s^{[l]}(S, G) &= \frac{1}{(n_H^{[l]} n_W^{[l]} n_c^{[l]})^2} \sum_{\underline{k}} \sum_{\underline{k'}} (\underline{\underline{G}}_{kk'}^{[l](S)} - \underline{\underline{G}}_{kk'}^{[l](G)})^2 \\ &= \frac{1}{(\underline{n}_H^{[l]} \underline{n}_W^{[l]} \underline{n}_c^{[l]})^2} \|\underline{\underline{G}}_{kk'}^{[l](S)} - \underline{\underline{G}}_{kk'}^{[l](G)}\|_F \end{aligned}$$

$$J_s(S, G) = \sum_l \lambda_{\underline{l}}^{[l]} J_s^{[l]}(S, G)$$

# Neural Style Transfer

