deeplearning.ai

Face recognition

What is face recognition?

1

Face recognition

[Courtesy of Baidu]

Andrew Ng

2

Face verification vs. face recognition

Verification
- Input image, name/ID
- Output whether the input image is that of the claimed person

Recognition
- Has a database of K persons
- Get an input image
- Output ID if the image is any of the K persons (or "not recognized")

Andrew Ng

3

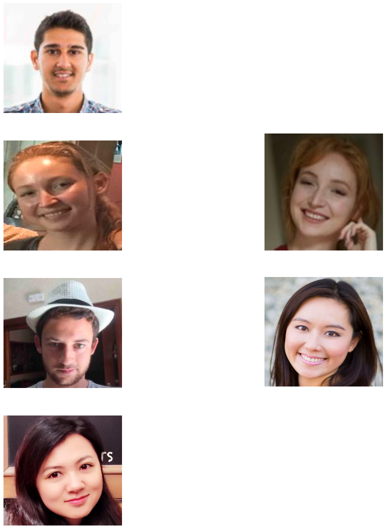
deeplearning.ai

Face recognition

One-shot learning

4

One-shot learning



Learning from one example to recognize the person again

Andrew Ng

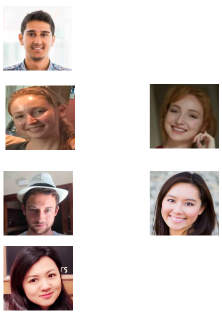
5

Learning a "similarity" function

d(img1,img2) = degree of difference between images

If d(img1,img2) $\leq \tau$

$> \tau$



Andrew Ng

6

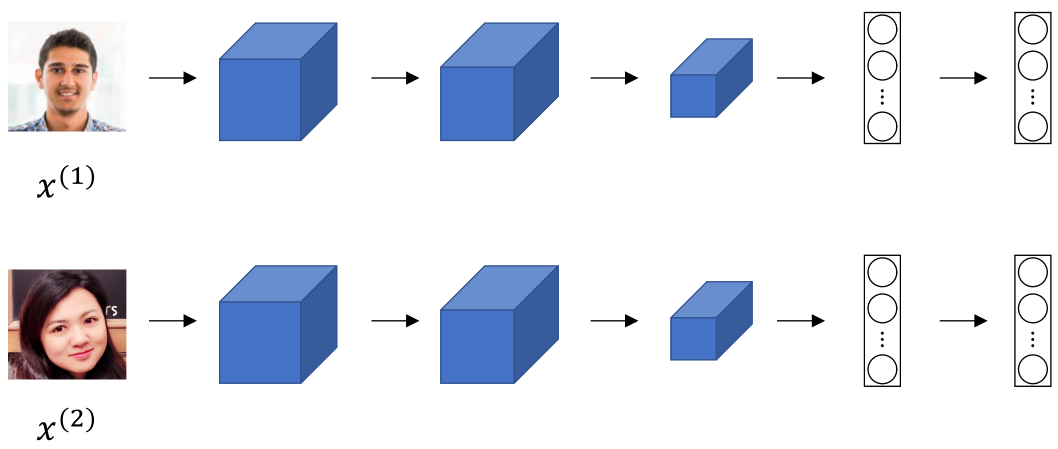
deeplearning.ai

Face recognition

Siamese network

7

Siamese network

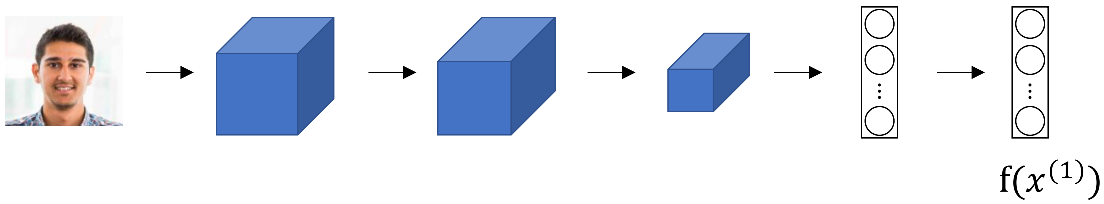


[Taigman et. al., 2014. DeepFace closing the gap to human level performance]

Andrew Ng

8

Goal of learning



Parameters of NN define an encoding $f(x^{(i)})$

Learn parameters so that:

If $x^{(i)}, x^{(j)}$ are the same person, $\|f(x^{(i)}) - f(x^{(j)})\|^2$ is small.

If $x^{(i)}, x^{(j)}$ are different persons, $\|f(x^{(i)}) - f(x^{(j)})\|^2$ is large.

Andrew Ng

9

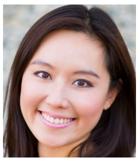
deeplearning.ai

Face recognition

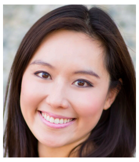
Triplet loss

10

Learning Objective


Anchor


Positive


Anchor


Negative

[Schroff et al.,2015, FaceNet: A unified embedding for face recognition and clustering]

Andrew Ng

11

Loss function

Training set: 10k pictures of 1k persons

[Schroff et al.,2015, FaceNet: A unified embedding for face recognition and clustering]

Andrew Ng

12

Choosing the triplets A,P,N

During training, if A,P,N are chosen randomly,
$d(A,P) + \alpha \leq d(A,N)$ is easily satisfied.

Choose triplets that're "hard" to train on.

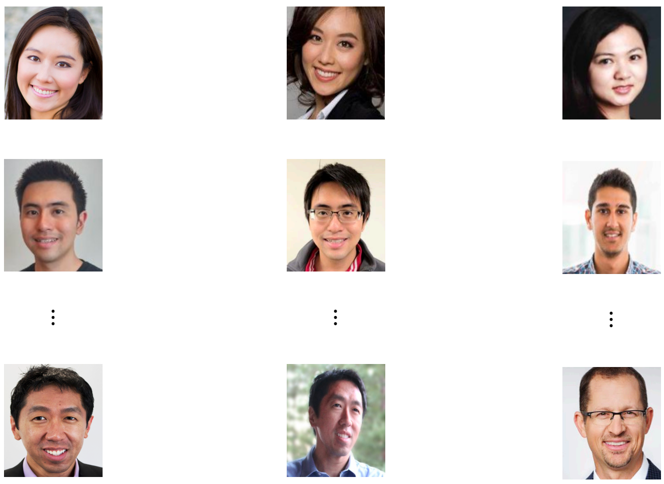
[Schroff et al.,2015, FaceNet: A unified embedding for face recognition and clustering]

Andrew Ng

13

Training set using triplet loss

Anchor | Positive | Negative



Andrew Ng

14

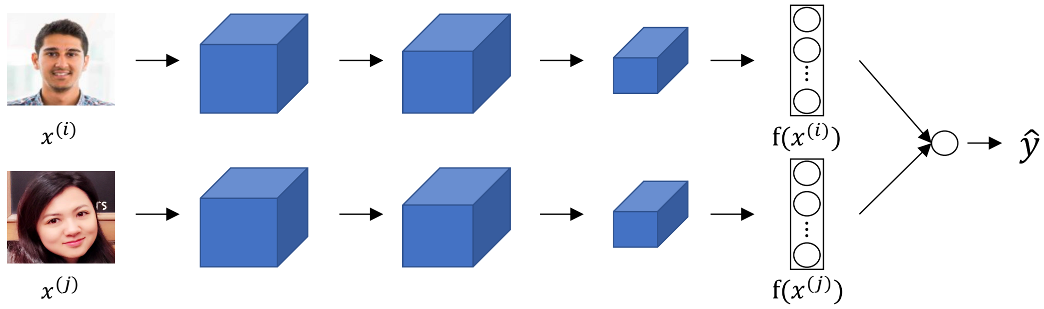
deeplearning.ai

Face recognition

Face verification and binary classification

15

Learning the similarity function



[Taigman et. al., 2014. DeepFace closing the gap to human level performance]

Andrew Ng

16

Face verification supervised learning

x y

1

0

0

1

[Taigman et. al., 2014. DeepFace closing the gap to human level performance]

Andrew Ng

17

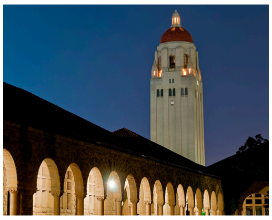
deeplearning.ai

Neural Style Transfer

What is neural style transfer?

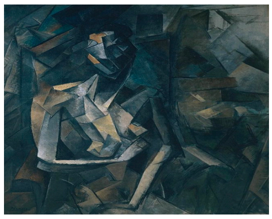
18

Neural style transfer




Content Style

Generated image

Content Style

Generated image

[Images generated by Justin Johnson]

Andrew Ng

19

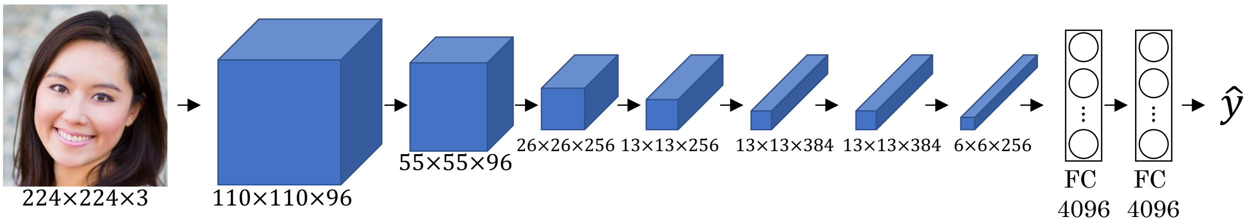
deeplearning.ai

Neural Style Transfer

What are deep ConvNets learning?

20

Visualizing what a deep network is learning



Pick a unit in layer 1. Find the nine image patches that maximize the unit's activation.

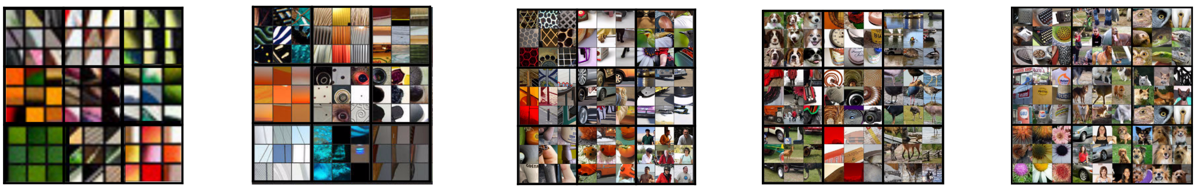
Repeat for other units.

[Zeiler and Fergus., 2013, Visualizing and understanding convolutional networks]

Andrew Ng

21

Visualizing deep layers



Layer 1 Layer 2 Layer 3 Layer 4 Layer 5

Andrew Ng

22

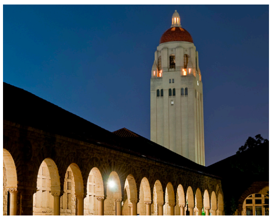
deeplearning.ai

Neural Style Transfer

Cost function

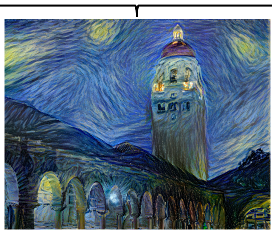
23

Neural style transfer cost function



Content C

Style S



Generated image G

[Gatys et al., 2015. A neural algorithm of artistic style. Images on slide generated by Justin Johnson] Andrew Ng

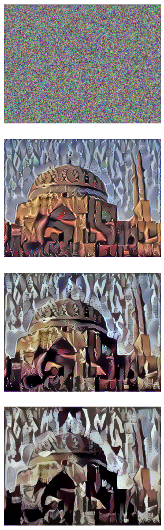
24

Find the generated image G

1. Initiate G randomly

   G: 100×100×3

2. Use gradient descent to minimize $J(G)$



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

Andrew Ng

25

deeplearning.ai

Neural Style Transfer

Content cost function

26

Content cost function

$$J(G) = \alpha\, J_{content}(C, G) + \beta\, J_{style}\,(S, G)$$

- Say you use hidden layer l to compute content cost.
- Use pre-trained ConvNet. (E.g., VGG network)
- Let $a^{[l](C)}$ and $a^{[l](G)}$ be the activation of layer l on the images
- If $a^{[l](C)}$ and $a^{[l](G)}$ are similar, both images have similar content

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

Andrew Ng

27

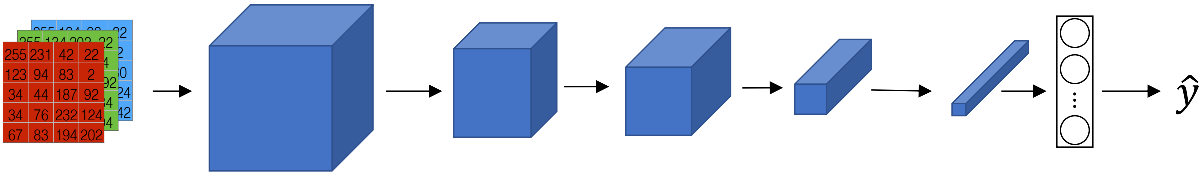
deeplearning.ai

Neural Style Transfer

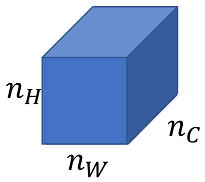
Style cost function

28

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.



How correlated are the activations across different channels?

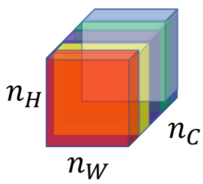
[Gatys et al., 2015. A neural algorithm of artistic style]

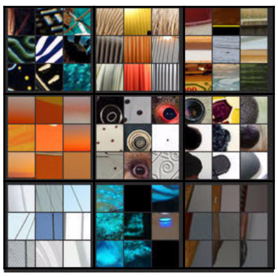
Andrew Ng

29

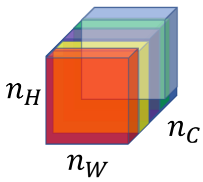
Intuition about style of an image

Style image





Generated Image



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

Andrew Ng

30

Style matrix

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

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

Andrew Ng

31

Style cost function

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

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

Andrew Ng

32

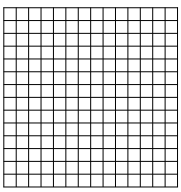
deeplearning.ai

Convolutional Networks in 1D or 3D

1D and 3D generalizations of models

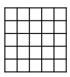
33

Convolutions in 2D and 1D

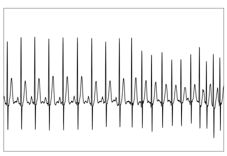


2D input image
14×14

*



2D filter
5×5



1	20	15	3	18	12	4	17

*

1	3	10	3	1

Andrew Ng

34

3D convolution