deeplearning.ai

Face recognition

What is face recognition?

Face recognition

[Courtesy of Baidu]

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")

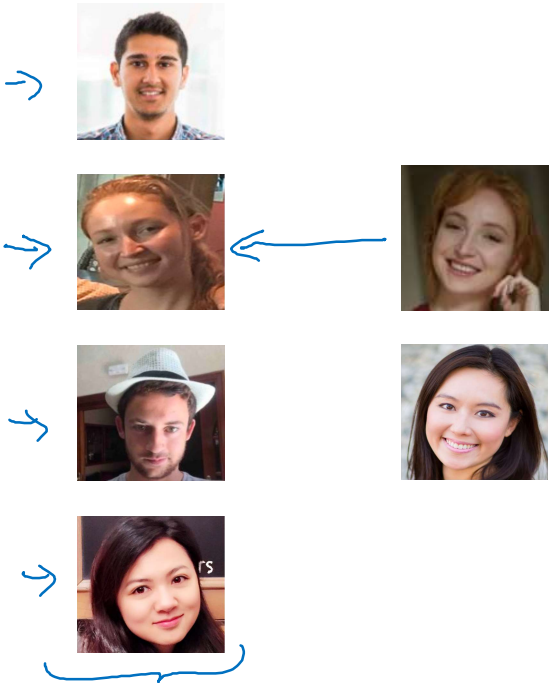
Andrew Ng

deeplearning.ai

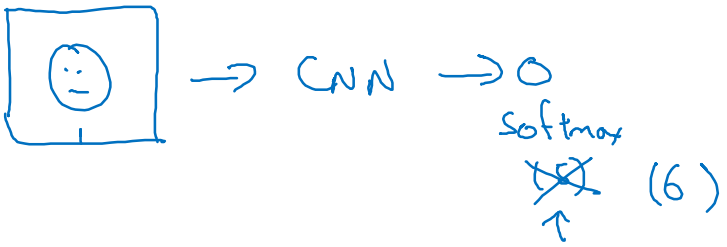
Face recognition

One-shot learning

One-shot learning



Learning from one example to recognize the person again



Andrew Ng

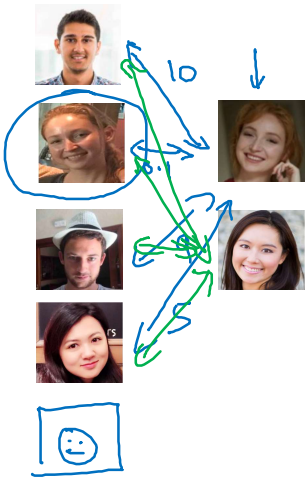
Learning a "similarity" function

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

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

$> \tau$ "different"

} Verification.



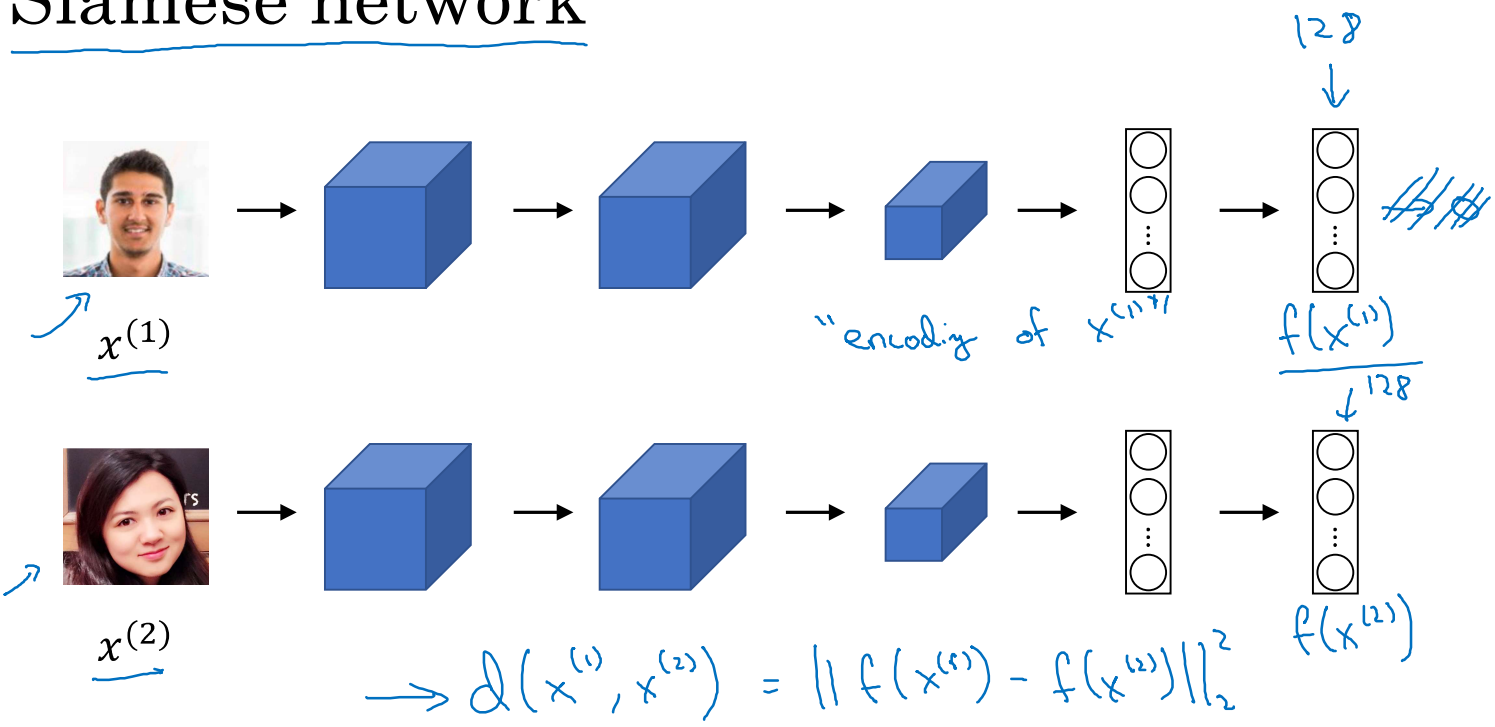
d(img1, img2)

deeplearning.ai

Face recognition

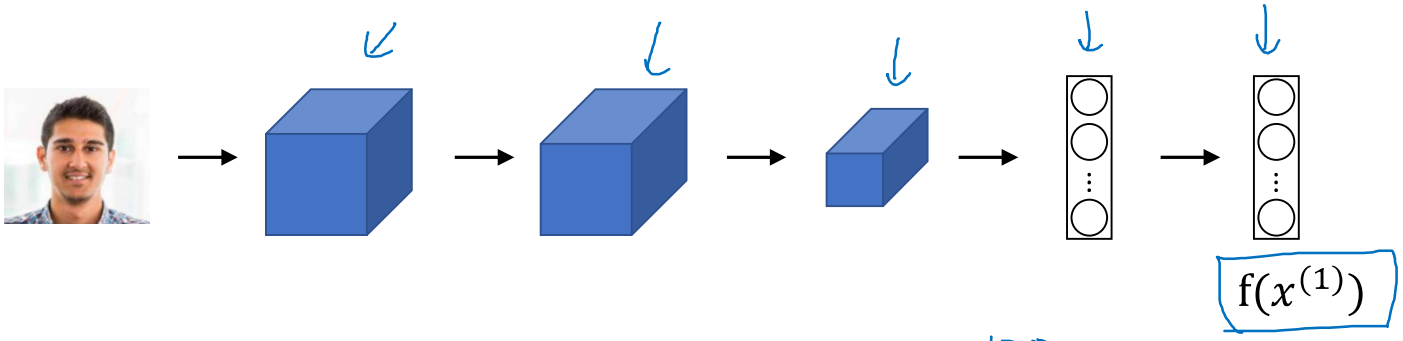
Siamese network

Siamese network



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

Goal of learning



Parameters of NN define an encoding $f(x^{(i)})$ 128

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.

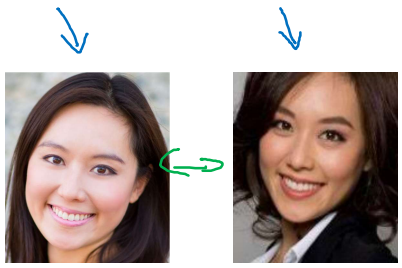
Andrew Ng

deeplearning.ai

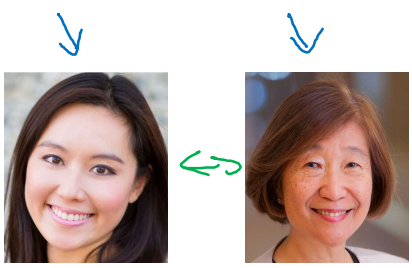
Face recognition

Triplet loss

Learning Objective



Anchor (A) Positive (P)



Anchor (A) Negative (N)

$d(A,P) \leq 0.5$

$d(A,N) \geq 0.7$

Want: $\|f(A) - f(P)\|^2 + \alpha \leq \|f(A) - f(N)\|^2$

$d(A,P)$ $\qquad$ $d(A,N)$

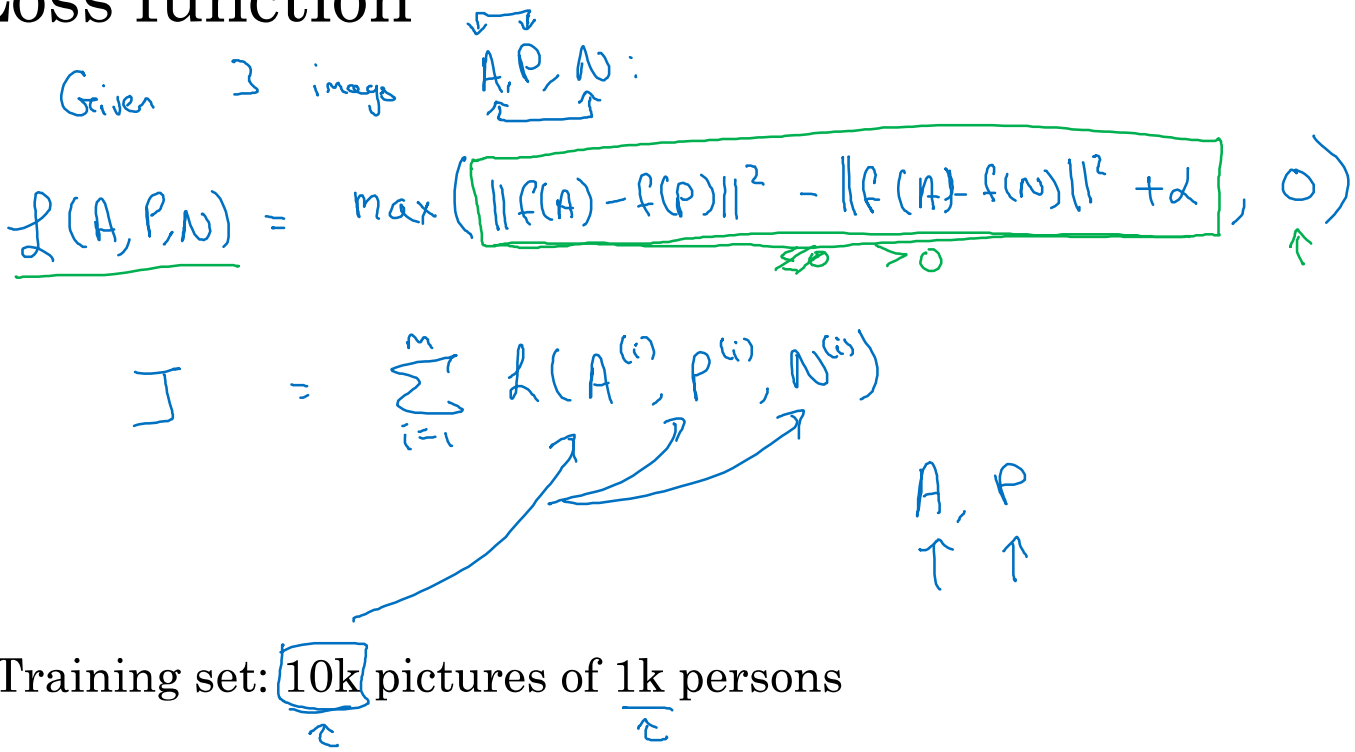
$\|f(A) - f(P)\|^2 - \|f(A) - f(N)\|^2 + \alpha \leq 0$

0 $\qquad$ 0 $\qquad$ margin

$f(img) = \vec{0}$

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

Loss function



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

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.

$\|f(A) - f(P)\|^2 + \alpha \leq \|f(A) - f(N)\|^2$

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

$d(A,P) + \alpha \leq d(A,N)$

$d(A,P) \approx d(A,N)$

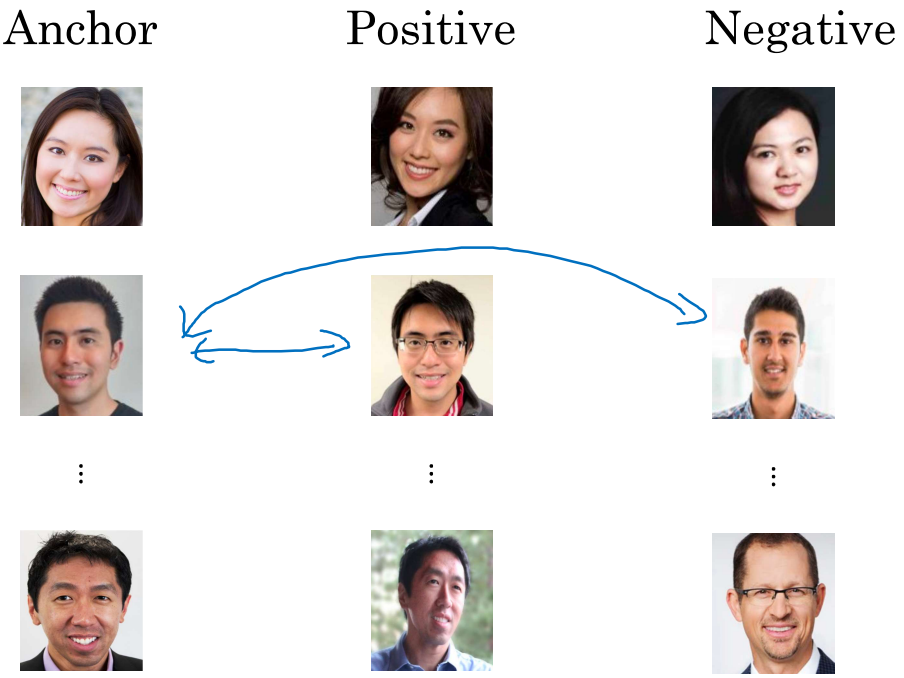
↓ ↑

FaceNet

DeepFace

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

Training set using triplet loss



$$J$$

$$d(x^{(i)}, x^{(j)})$$

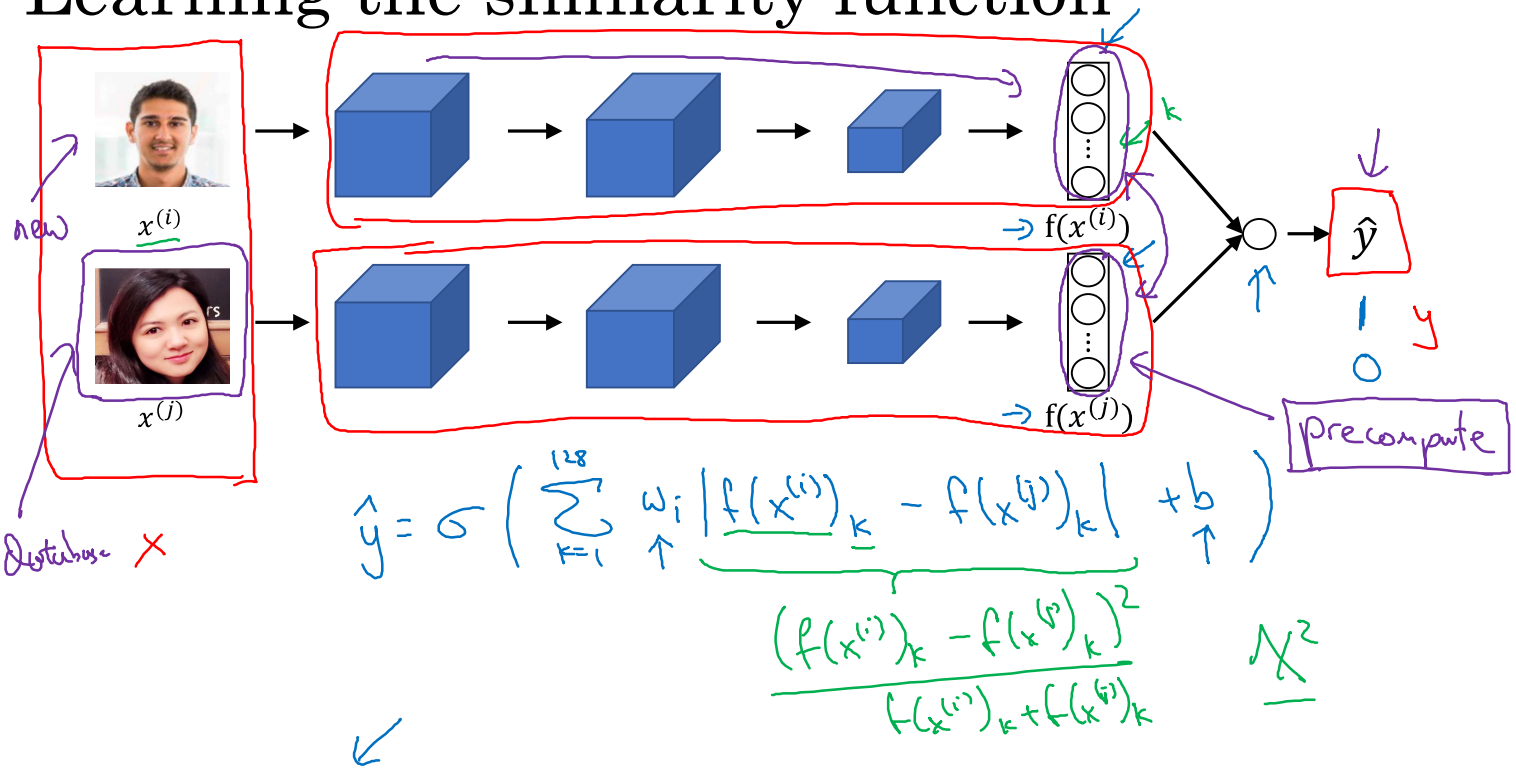
Andrew Ng

deeplearning.ai

Face recognition

Face verification and binary classification

Learning the similarity function



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

Face verification supervised learning

x	y	
	1	"Same"
	0	"Albert"
	0	
	1	

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