Face recognition

Triplet loss

deeplearning.ai

Learning Objective
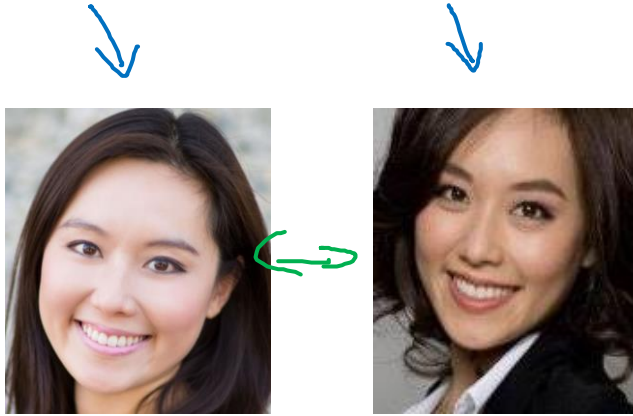
Anchor   Positive   Anchor   Negative

$$A \quad d(A,P) = \underline{0.5} \qquad \nearrow 0.2 \qquad A \quad d(A,N) = \overset{N}{\cancel{0.51}} \; \boxed{0.7}$$

$$\text{Want:} \quad \underbrace{\| f(A) - f(P) \|^2}_{d(A,P)} + \alpha \; \leq \; \underbrace{\| f(A) - f(N) \|^2}_{d(A,N)}$$

$$\boxed{\underbrace{\| f(A) - f(P) \|^2}_{0} - \underbrace{\| f(A) - f(N) \|^2}_{0} + \underbrace{\alpha}_{margin} \; \leq \; \underline{0}} \quad \text{4/0} \qquad f(img) = \vec{0}$$

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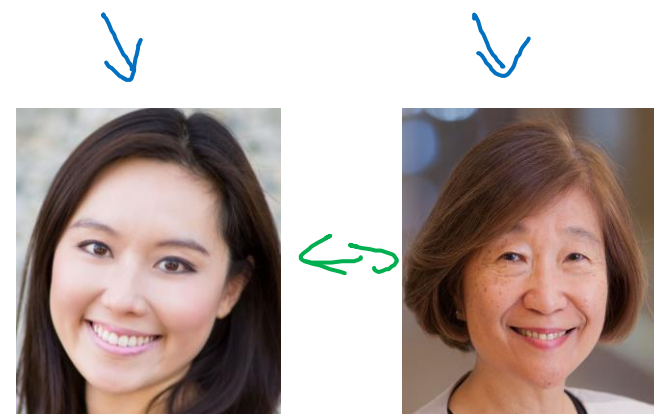
Loss function

Given 3 images A, P, N:

$$\mathcal{L}(A, P, N) = \max\left(\|f(A) - f(P)\|^2 - \|f(A) - f(N)\|^2 + \alpha, \; 0\right)$$

$\leq 0 \qquad > 0$

$$J = \sum_{i=1}^{m} \mathcal{L}\left(A^{(i)}, P^{(i)}, N^{(i)}\right)$$

A, P

Training set: 10k pictures of 1k persons

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

Andrew Ng

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.

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

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

Face Net

Deep Face

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

Andrew Ng

Training set using triplet loss

Anchor Positive Negative

$$J$$

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