

deeplearning.ai

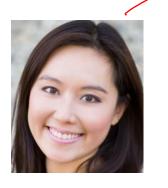
Face recognition

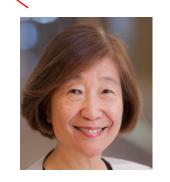
Triplet loss
This is the loss func on which we apply 59D
for facial recognition

Learning Objective

Triplet comes from the fact that you have 3 terms (Anchor, positive, Negative)







Anchor (A)

Positive (?)

Anchor (A)

Negative (N)

- To Apply triplet loss, you need to compare pairs of Images
- We compare 2 pairs (A,P), (A,N)
- where we want the encodings of A,P to be similar so that $\|f(A)-f(P)\|^2$ is small a we want encodings of A,N to be diff so that $\|f(A)-f(N)\|^2$ are large

$$\Rightarrow \| +(A) - +(P) \|^2 \leq \| +(A) - +(N) \|^2 \quad (\text{or } \lambda(A, P) \leq \lambda(A, N))$$

- $d(A,P)-d(A,N) \le 0$ \Rightarrow one way to do this is make f(A)=f(P)=f(N)=constant (but we don't
 - is to make sure that we have some gap b/w d(A,P) & d(A,N) we do $d(A,P)-d(A,N) \leq 0-\lambda$ or $\left[d(A,P)-d(A,N)+\kappa \leq 0\right]$

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

Andrew Ng

always output

a constant)

encoding to

Loss function

Given 3 Images
$$(A,P,N) \rightarrow P = \text{Image of Same person}$$

$$N = \text{Image of Aiff person}$$

$$L(A,P,N) = \max\left(\|f(A) - f(P)\|^2 - \|f(A) - f(N)\|^2 + < < < > 0\right)$$

$$T = \sum_{i=1}^{\infty} L(A^{(i)}, P^{(i)}, N^{(i)})$$

Then generate M rows of 3 Images

Then generate M rows of 3 Images

each (A,P,N) & train NN on each

of these triplets:

Training set: 10k pictures of 1k persons

to have duplicate pictures so you get (A,P)

off et al.,2015, FaceNet: A unified embedding for face ---

Choosing the triplets A,P,N (A,P = Same person) A,N = diff persons

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

If you choose at Random, chances that A,N are very diff is very probable

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

choose triplets where $d(A,P) \approx d(A,N)$

- If you choose Randomly, then
a lot of triplets will be automatically satisfied
a lot of triplets will be automatically satisfied
w/o much change in params => gradient descent wouldn't
be much learning | ", the real learning happens when things are
hard to dicepher"

Training set using triplet loss

