Face Recognition

One shot learning : learn from 1 example
- learn a similarity function (verify if same person)

Siamese function

run 2 identical CNNs on 2 example, then use the similarity function to compare the distance of the 2 encodings

$$d(x^{(1)}, x^{(2)}) = \| f(x^{(1)}) - f(x^{(2)}) \|_2^2$$

{ small if same person
large if diff person

Triplet loss Anchor Positive Negative

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

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

$$d(A,P) - d(A,N) + \alpha \leq 0$$

hyperparam : margin to prevent d just to output 0 every time

formalize

$$L(A,P,N) = \max(\| f(A) - f(P) \|^2 - \| f(A) - f(N) \|^2 + \alpha , 0)$$

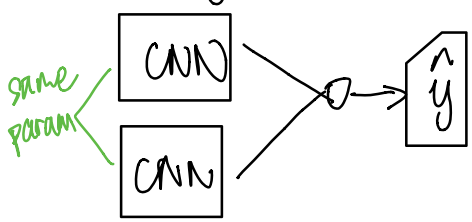
$$J = \sum_{i=1}^{m} L(A^{(i)}, P^{(i)}, N^{(i)})$$

How to choose A, P, N : choose triplets that are hard to train
improve the efficiency of the training algo.

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

↓ ↑

As Binary Classification



$$\hat{y} = \sigma\left(\sum_{k=1}^{128} w_i \left| f(x^{(i)})_k - f(x^{(j)})_k \right| + b\right)$$

Neural Style Transfer

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

content cost $J_{content}(C,G) = \frac{1}{2} \| a^{[l](C)} - a^{[l](G)} \|^2$

style cost

$$G^{[l]}_{kk'} = \sum_{i=1}^{n_H^{[l]}} \sum_{j=1}^{n_W^{[l]}} a^{[l]}_{ijk} \, a^{[l]}_{ijk'}$$

(k: channel, k': channel)

style matrix — "Gram Matrix"

How correlated are activations across different channel ?

$$J^{[l]}_{style}(S,G) = \| G^{[l](S)} - G^{[l](G)} \|_F^2$$

$$\rightarrow \frac{1}{(2n_H n_W)^2} \sum_k \sum_{k'} \left(G^{(S)}_{kk'} - G^{(G)}_{kk'} \right)$$

Conv 3D : (32, 32, 32, 16) → 32 filters (3, 3, 3)

output ⇒ (30, 30, 30, 32)