

CNN

Week - 6

Face Verification

New Image



Ground Truth



Is it Dhoni?

YES / NO



Recognition is harder than verification.
Verification may have 1% chance of making mistake, whereas the Recognition has 99%

Face Recognition



Who is this player?

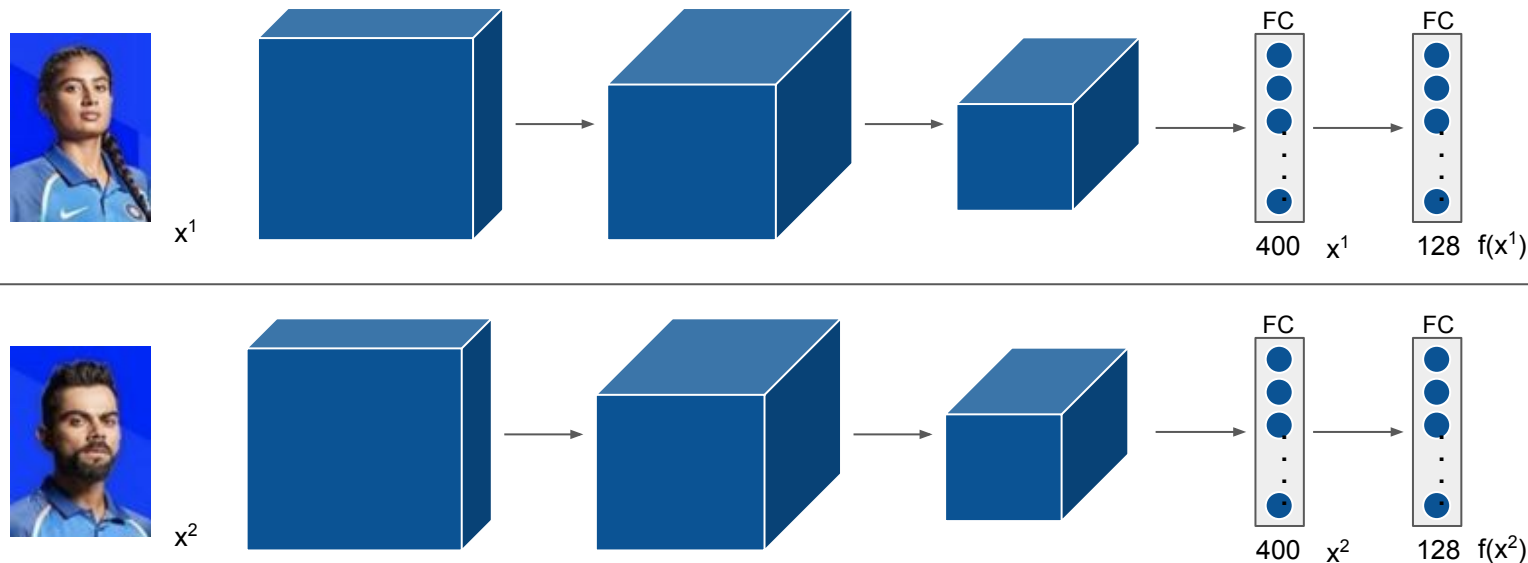


Has a database of K persons.
Output: ID, if the image is any of the K persons, or not recognized



<https://www.linkedin.com/in/arun-kumar-anala-35760523/>

Siamese Network



If x^1 and x^2 are same persons, then $||f(x^1) - f(x^2)||^2$ is small.

If x^1 and x^2 are different persons, then $||f(x^1) - f(x^2)||^2$ is large.

$$d(x^1, x^2) = ||f(x^1) - f(x^2)||^2$$

Siamese Network : What we want?



Anchor Image



Positive Image



Anchor Image



Negative Image

$$d(A, P) = ||f(A) - f(P)||^2$$

$$d(A, N) = ||f(A) - f(N)||^2$$

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

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

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

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

Margin Hyper Parameter pushes the positive match distance and negative match distance further away.

To prevent trivial possibility of network generating the $\vec{0}$ encoding vector for given inputs.

Siamese Network : Loss Function



Anchor Image



Positive Image



Negative Image

$$L(A,P,N) = \text{Max} (||f(A) - f(P)||^2 - ||f(A) - f(N)||^2 + m , 0)$$

$$J = \sum L(A^i, P^i, N^i)$$

Loss function does not care of negative values. It always tries to reduce the positive values to ≤ 0 .

Siamese Network : Choosing Triplets



Ancor Image



Positive Image



Negative Image

VS

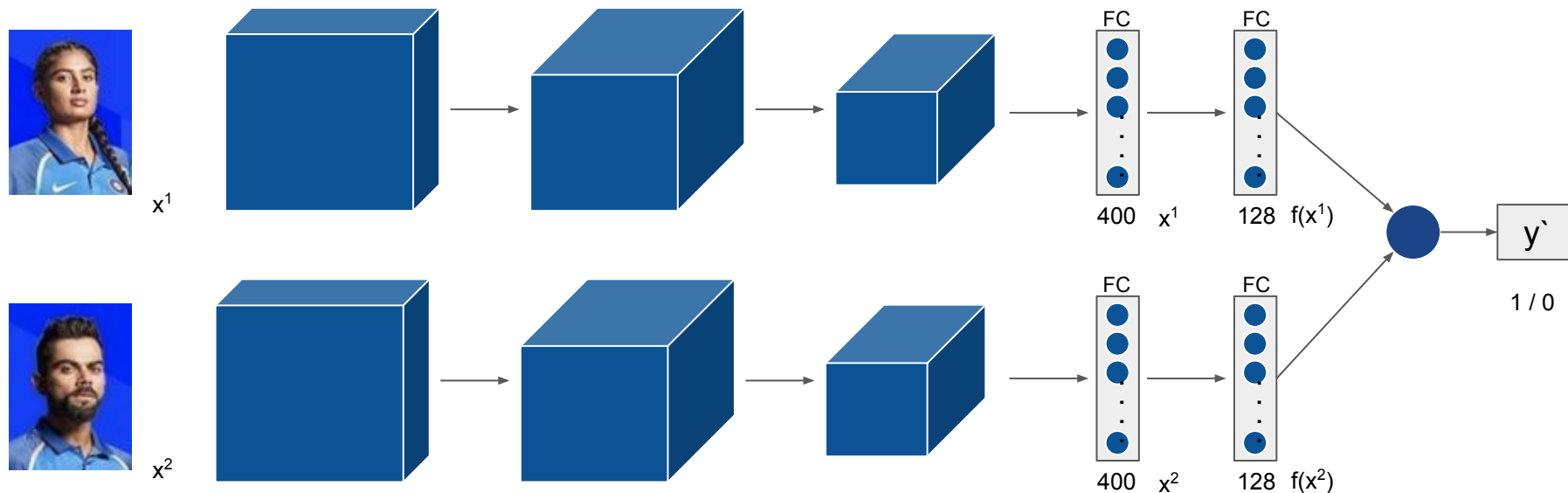


Negative Image

During training, if A, P, N are chosen randomly, then $d(A, P) + m \leq d(A, N)$ is easily learned

Choose the triplets A, P, N that are hard to train on. This will improve the computational efficiency and also model accuracy

Similarity Function



$$Y^{\sim} = \sigma(\sum w^i | f(x^i) - f(x^j) | + b)$$

Store the encoded vector in DB
to save the computation time.