Identity Verification using Siamese Neural Networks

Group 9
Bharath Challa
Raghunath Siripudi

Problem

The identity verification problem is to verify the identity of a person with a reference image or data.

We need to verify the person's face in real-time at airports, organizations, exam-halls with picture on ID-card as reference.

The appearance of the face changes with time so when verifying the person's identity with an older ID card picture is a difficult task for a human eye.

Problem

If we were to use traditional neural networks, we will have to face two main problems.

First one would be the dataset size. And traditional CNN(Convolutional Neural Networks) won’t be able to learn features with small dataset collection.

Second one, What happens when new persons gets added to our dataset? How can we include those into our facial verification neural network system? Do we need to increase the number of classes and retrain?

Siamese Network

Introduction to Siamese Network

A Siamese Network is a type of network architecture that contains two or more identical subnetworks used to generate feature vectors for each input type and compare them.

Siamese networks are also called as sister networks which uses shared weights with intention to learn similarity and dissimilarity between feature vectors.

“Siamese” Name Origin

Siam is the name of Thailand in ancient times

Like Chinese, Siamese is for "Siam" or "Thai" people.

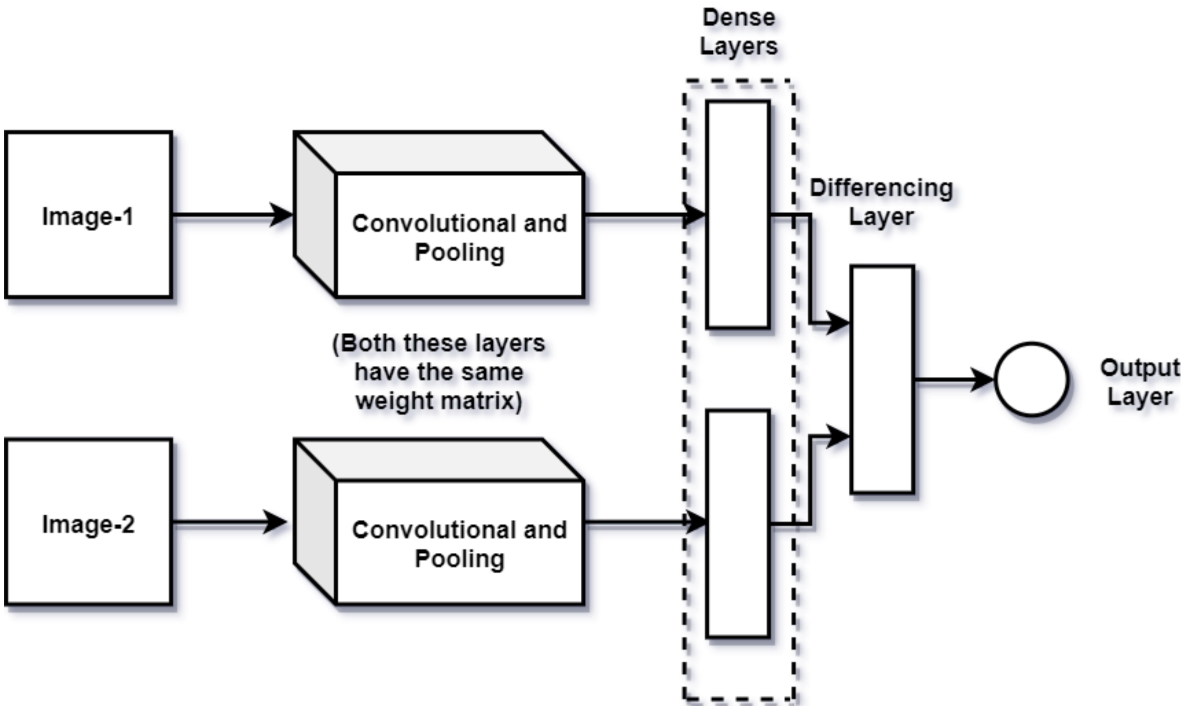
Siamese means "twin" and "conjoined" in English. Why?

A first pair of conjoined twins were born in Thailand in the nineteenth century. They are so popular because medical technology at the time could not separate the two, so the two lived tenaciously.

Applications of Siamese Neural network

- Signature Verification
- Facial Recognition
- Compare Fingerprinting
- Evaluate disease severity based on clinical grading
- Text similarity for a job profile to resume matching
- Text similarity for pairing similar questions

Siamese Architecture



Approach

We use a Siamese Network with three/two identical subnetworks (we are using the same network).

We will provide three images to the model, where two of them will be similar (anchor and positive samples), and the third will be unrelated (a negative example) in case of triplet loss.

Approach

In case of Contrastive loss or Binary Cross Entropy loss we will have a pair of images either positive(similar) or negative(dissimilar).

Our goal is for the model to learn to estimate the similarity between images.

Loss Functions

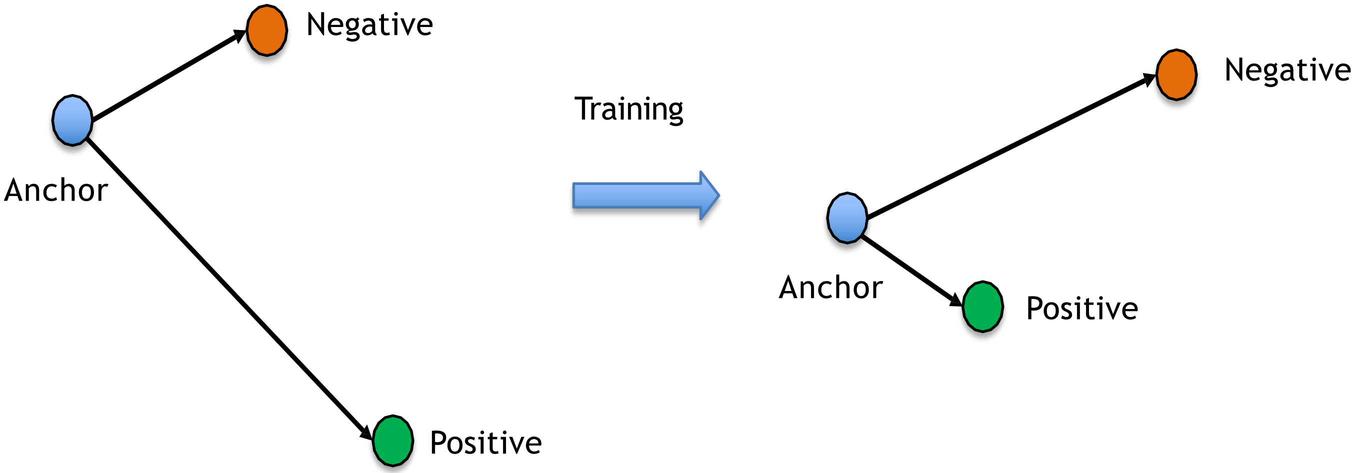
Triplet Loss

The triplet loss is probably the best-known loss function for face recognition and our results also shows the same.

The data is arranged into triplets of images: anchor, positive example, negative example.

The images are passed through a common network and the aim is to reduce the anchor-positive distance while increasing the anchor-negative distance

Triplet Loss



Triplet Loss

Triplet loss allows us to learn a ranking

Anchor (A)

Positive (P)

Negative (N)

$$\text{Goal: } ||f(A) - f(P)||^2 < ||f(A) - f(N)||^2$$

Triplet Loss

$$||f(A) - f(P)||^2 < ||f(A) - f(N)||^2$$

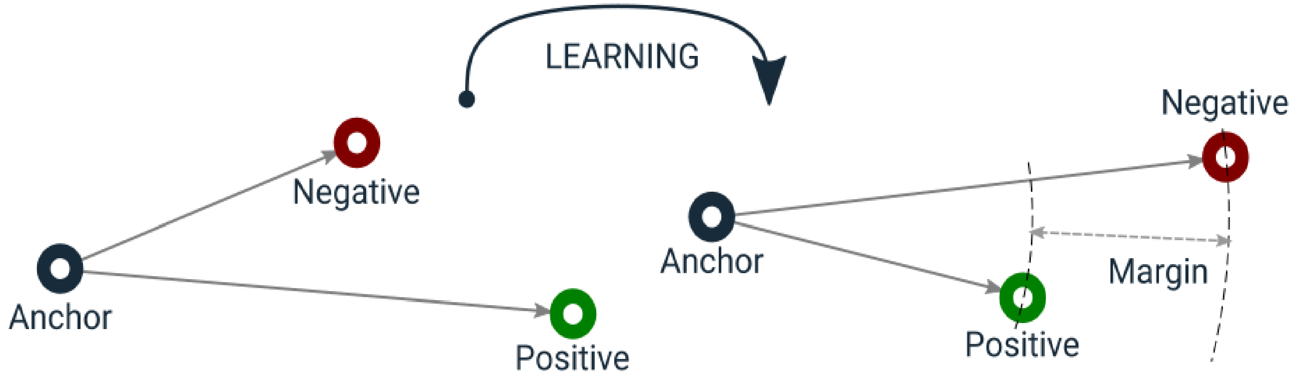
$$||f(A) - f(P)||^2 - ||f(A) - f(N)||^2 < 0$$

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

In the above equation, margin(m) is used to “stretch” the distance differences between similar and dissimilar pairs in the triplet.

Triplet Loss

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



Contrastive Loss

Contrastive Loss is also one of the best-known loss functions for face recognition and results are better than binary cross entropy loss function.

The motivation behind this loss was to develop a model which would learn to represent images in a feature space such that the distance in that space would correspond to semantic distance in the original space.

Contrastive Loss

$$L(A, B) = y\,||f(A) - f(B)||^2$$

$$+$$

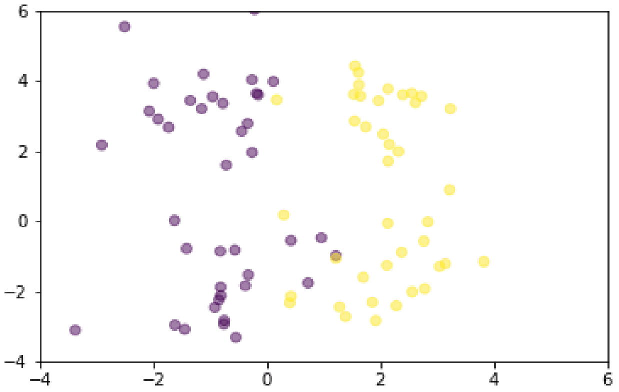
$$(1 - y)\max(0, m^2 - ||f(A) - f(B)||^2)$$

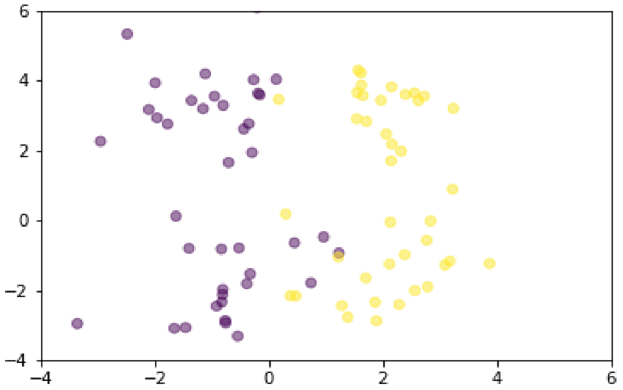
Positive pair, reduce the distance between the elements

Negative pair, brings the elements further apart up to a margin

Contrastive Loss vs Triplet Loss

Contrastive Loss





Triplet Loss

Binary Cross Entropy Loss

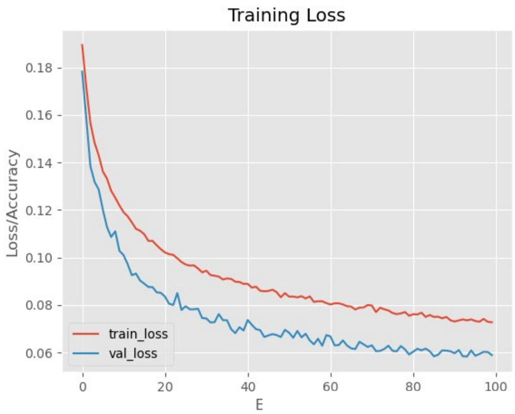
Binary cross entropy checks the each of the predicted probability after the sigmoid activation function to the actual classification output which can be either 0 (dissimilar) or 1 (similar).

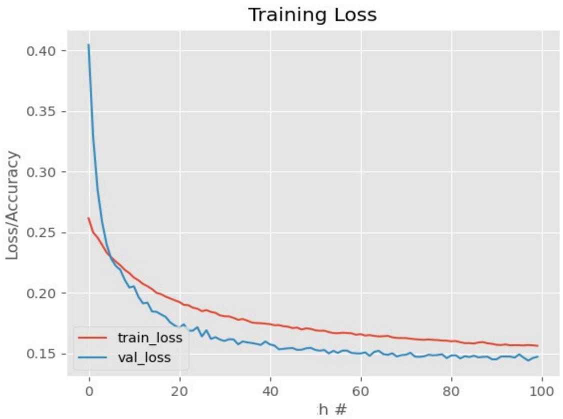
$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

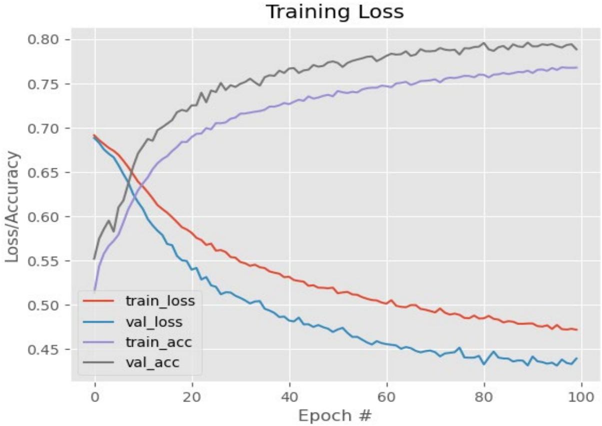
It then calculates the score that penalizes the probabilities based on the distance from the expected value. That means how close or far from the actual value.

Results

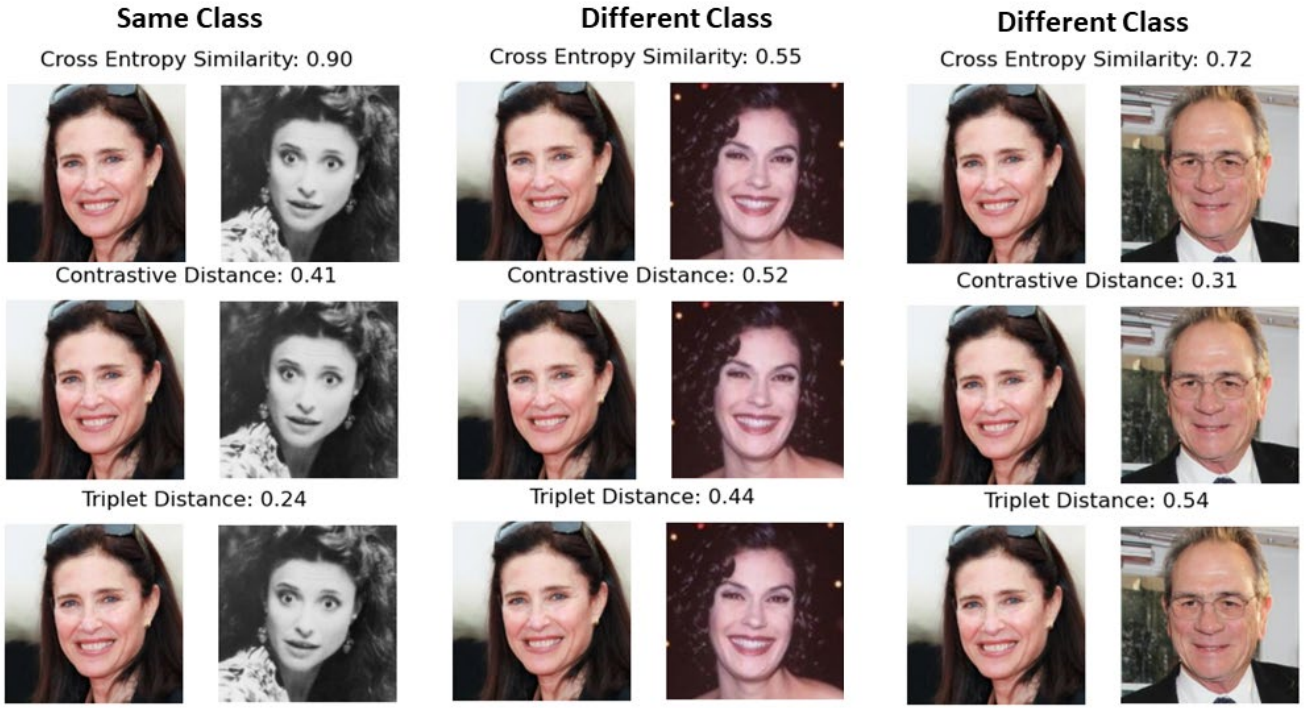
Triplet vs Contrastive vs Binary Cross Entropy







Face Verification



Questions ????