Comparison of Loss Functions on Face Embedding Learning

Yu Sun (ys3225) Hetian Bai (hb1500) Jieyu Wang (jw4937)

Center for Data Science, New York University

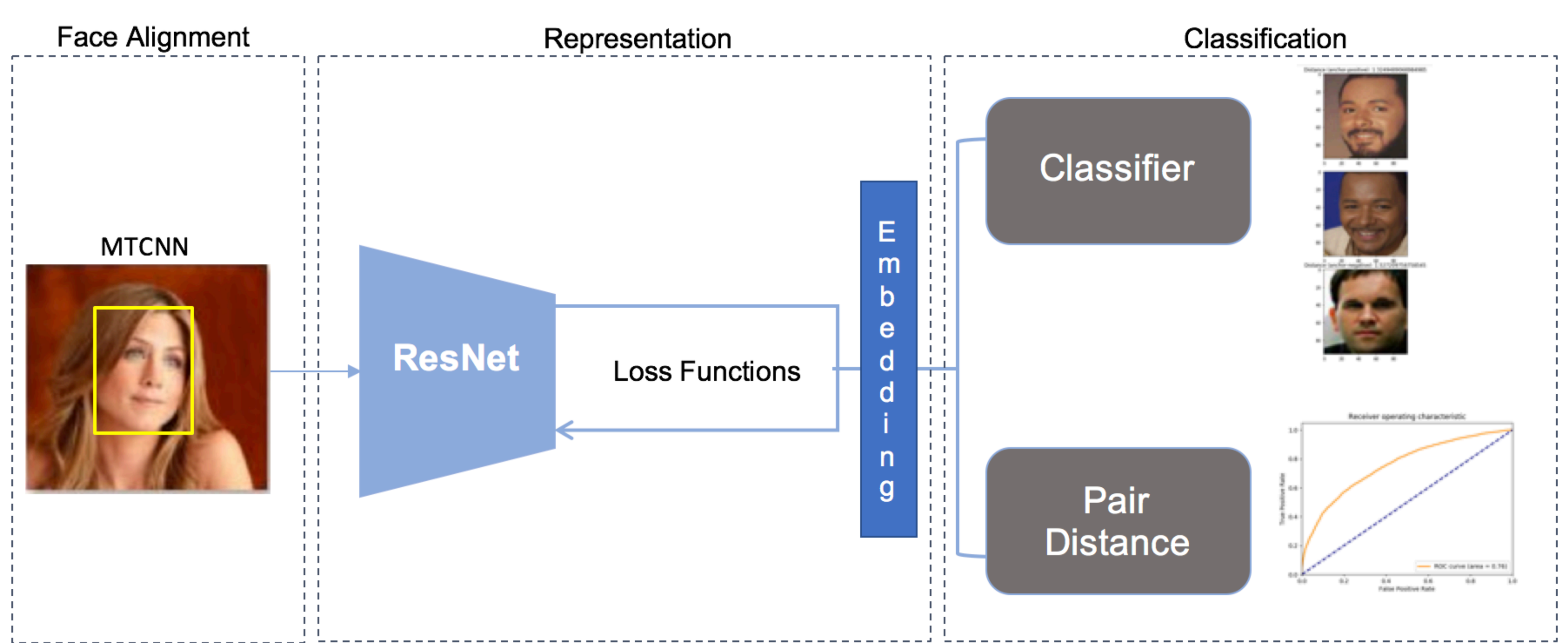
Introduction

The goal of this project is to compare performances of models with different loss functions on face verification and recognition tasks. We conduct five experiments:

1. Center Loss (baseline model)
2. Softmax Loss (cross-entropy loss)
3. Triplet Loss: l^2-norm triplet loss and random-sampling to generate triplets
4. Margin-based Loss: distance-weighted sampling
5. Hard Margin-based Loss: distance-weighted sampling, large penalty on mislabeled hard-pairs

We focus more on comparing loss functions on **face recognition** task than **verification**, because there are insufficient research papers that are evaluating model's recognition performances, so there is a lot for us to explore.

Framework



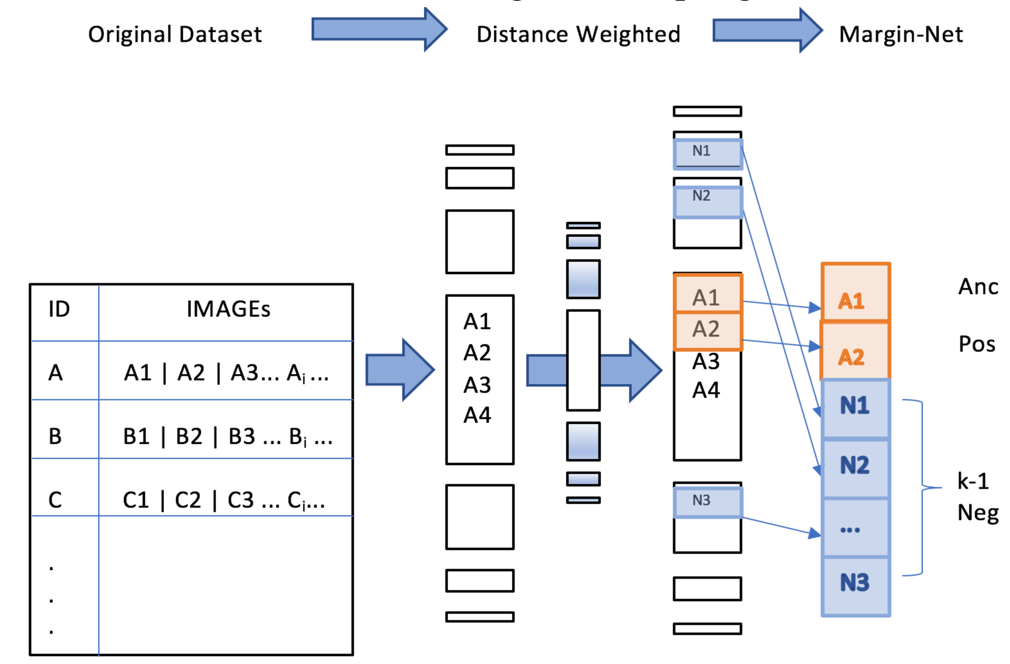
- **Alignment Layer**
  - Detect and crop face by applying the pre-trained CNN model – MTCNN
- **Representation Layer**
  - Images transformation: resize into 96*96; horizontal-flipped to increase data robustness; turn into tensor; and normalization
  - DNN architecture: Resnet18 (ave-pooling layer is dropped) + two fully-connected layers
  - Implementation of 5 different loss functions
- **Classification Layer**
  - Face recognition: linear classifier + Softmax => multi-class predictions
  - Face verification: calculate D_{ij} (l^2 distance between pairs in LFW)

  IF D_{ij} < threshold => same identity

  IF D_{ij} > threshold => different identities

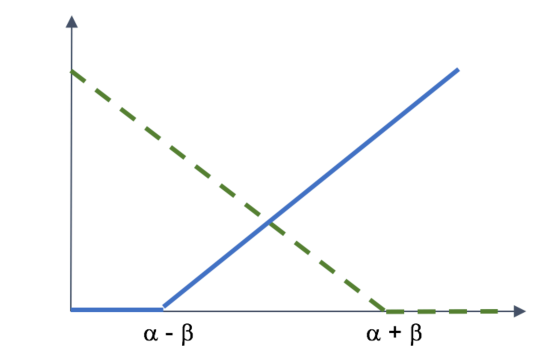
Innovation

Distance-Weighted-Sampling



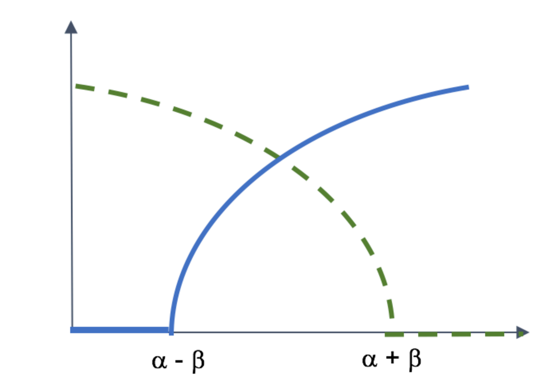
Distance-Weighted Sampling uniformly samples data based on distance

Margin-based Loss



$$L^{margin}_{i,j} := [\alpha + y_{ij}(D_{ij} - \beta)]_+$$

Hard Margin-based Loss



$$L^{hard\_margin}_{i,j} := log([\alpha + y_{ij}(D_{ij} - \beta)]_+ + 1)$$

Based on margin-based loss, we propose a new margin-based loss which gives different weights of penalty by distance. It gives higher penalty to misidentified closed-pairs and lower penalty to pairs that are far away than margin, which may contribute to separate face images better.

Experiments & Results

Datasets

- Training set: 68 images for 6400 identities from MS-Celeb-1M
- Verification testing set: 6000 image-pairs from LFW
- Recognition testing set: 2 images for 6400 identities from MS-Celeb-1M

Evaluation

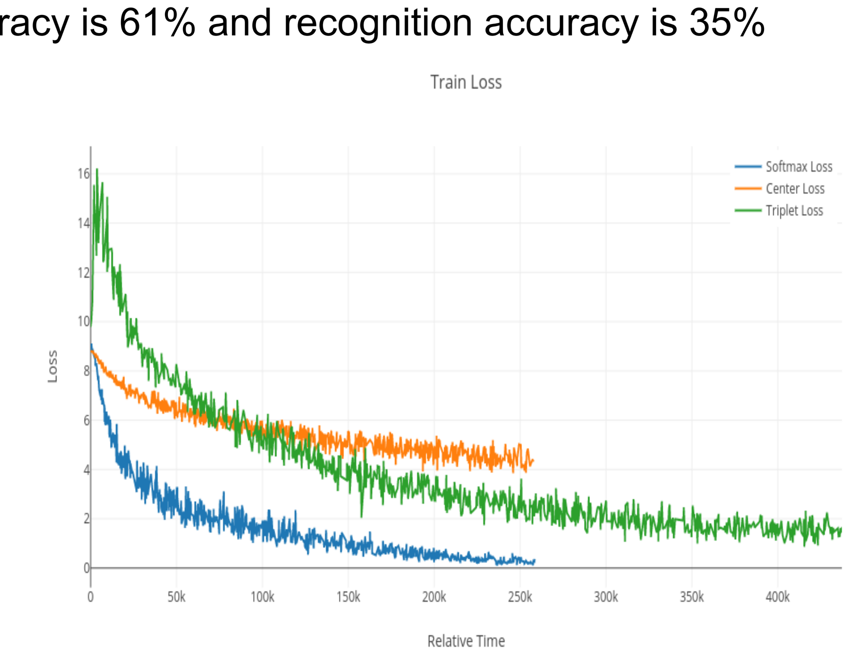
- Face recognition task: Accuracy, compare prediction with true label
- Face verification task: AUC, best accuracy over all thresholds

Results and Analysis

❖ Center Loss Baseline: The verification accuracy is 61% and recognition accuracy is 35%

❖ Compare Converge Speed:

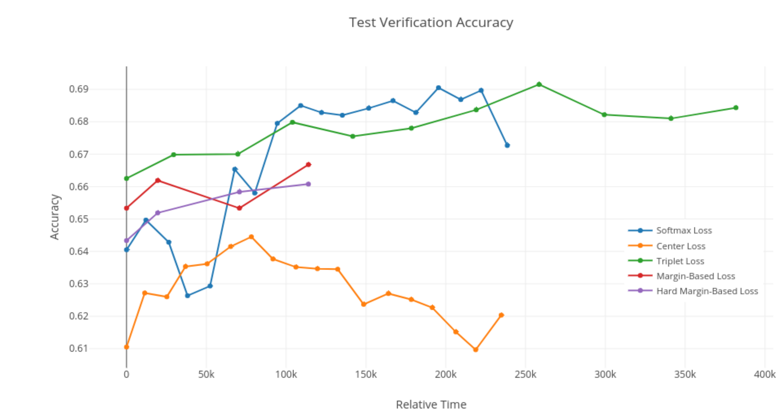
As showed in figure, we find softmax loss converges very fast and gets the lowest loss. Although triplet loss converges faster than the baseline, it fluctuates wildly because we use random sampling and the induced loss depends on triplets in each batch.



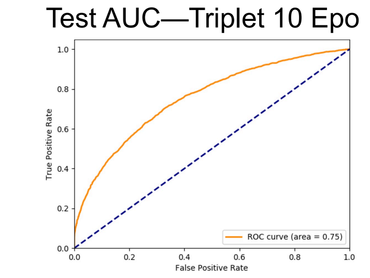
❖ Experiments on Verification Task

Triplet loss has better performance than others. It is more stable and has higher AUC than softmax loss, because it ensures negative-pairs has larger embedding distance than positive-pairs, which fits verification tasks well.

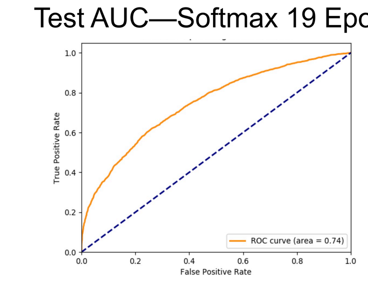
Without fully trained, both margin-based loss and hard-margin-based loss have the trends to outperform baseline, we may expect them to have good performances after parameter-tuning.



Test AUC—Triplet 10 Epo



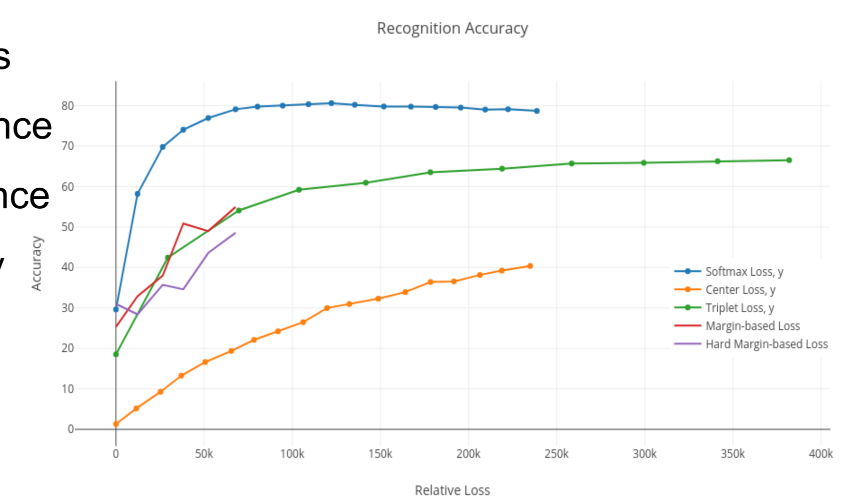
Test AUC—Softmax 19 Epo



❖ Experiments on Recognition Task

Softmax loss has the best performance over all in terms of accuracy and speed. Triplet loss yields good result on verification but a poor performance on recognition task because it only takes distance into account, so it contributes slightly to identify one image from 6400 classes.

Margin-based loss and new margin-based loss yield close performances with triplet loss.



Conclusions

1. Softmax loss yields good results on both face verification and recognition tasks.
2. Even though triplet loss takes more time to converge, it gives good performance on verification task.
3. Constrained by time, we have not complete training and tuning two margin-based function models, however, based on their initial performances and trends, we expect it could outperform other loss functions models after parameter-tuning.

Future Work

1. Continue to train both margin-based and hard-margin-based model, and observe their speed of convergence, verification and recognition accuracy;
2. In order to obtain more rigid results, each model should be carefully turned to achieve the best performance, so hyper-parameter tuning needs to be done;
3. Conduct error analysis on images to figure our why faces were incorrect labeled by model and to compare analysis results between models to see what role loss functions involved.