Imperial College London

# Learning Deep Convolutional Embeddings for Face Representation Using Joint Sample- and Set-based Supervision

Baris Gecer, Vassileios Balntas, Tae-Kyun Kim
{b.gecer, v.balntas15, tk.kim}@imperial.ac.uk

ICCV 17
AMFG 2017
IEEE Workshop on Analysis and Modeling of Faces and Gestures

## Abstract


- We propose a novel loss function called Max-Margin Loss that benefits from set-based information by drawing inter-class margins. It maximizes the maximum possible inter-class margin that is calculated by SVMs. It implicitly pushes all the samples towards correct side of the margin with a vector perpendicular to the hyperplane and a strength inversely proportional to the distance to the hyperplane.
- We review existing set-based DML approaches and evaluate them and their combinations together with Max-Margin Loss and Softmax Loss
- We build a framework where such functions can operate jointly with sample-based ones and investigate the strategies to maintain set information during training


## Related Work

1. Sample-based DML methods (e.g., Softmax, Contrastive, Triplet, LSE, Quadruplet Losses)
2. Set-based DML methods
   (a) Center Loss [4]: Biased centroid estimation and Do not tolerate intra-class variation
   (b) Magnet Loss [3] : Not combinatorial with other functions
3. SVM based Deep Learning [5] : Penalizes only slack variables

## References


[1] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. Deepface: Closing the gap to human-level performance in face verification. In *CVPR*, 2014.
[2] Omkar M Parkhi, Andrea Vedaldi, and Andrew Zisserman. Deep face recognition. In *BMVC*, 2015.
[3] Oren Rippel, Manohar Paluri, Piotr Dollar, and Lubomir Bourdev. Metric learning with adaptive density discrimination. *ICLR*, 2016.
[4] Yandong Wen, Kaipeng Zhang, Zhifeng Li, and Yu Qiao. A discriminative feature learning approach for deep face recognition. In *ECCV*, 2016.
[5] Yichuan Tang. Deep learning using linear support vector machines. In *In ICML workshop*, 2013.


## Overview of Set-based Learning Framework

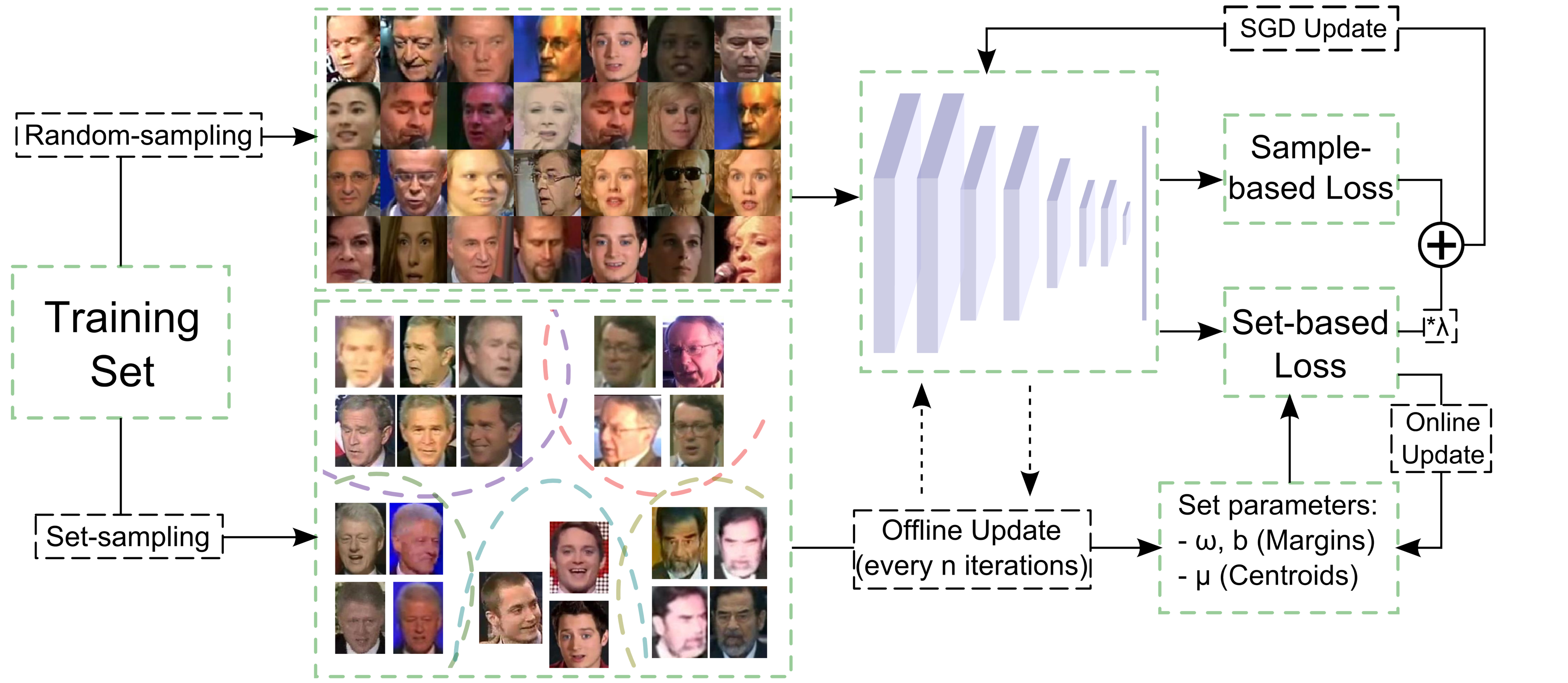


**Figure 1:** Random batch of face images are sampled from the training set to train a CNN with sample-based loss function (i.e., Softmax).
**Offline Update:** In every n iterations, set of face images are sampled that consist significant amount of images from each identities and fed into the network while training is paused. Resulting feature vectors are used to calculate set parameters whose way is specific to set-based loss used.
**Online Update:** While the training is going, set parameters are updated with a small weight by the parameters calculated with current random batches.

## Max-Margin Loss

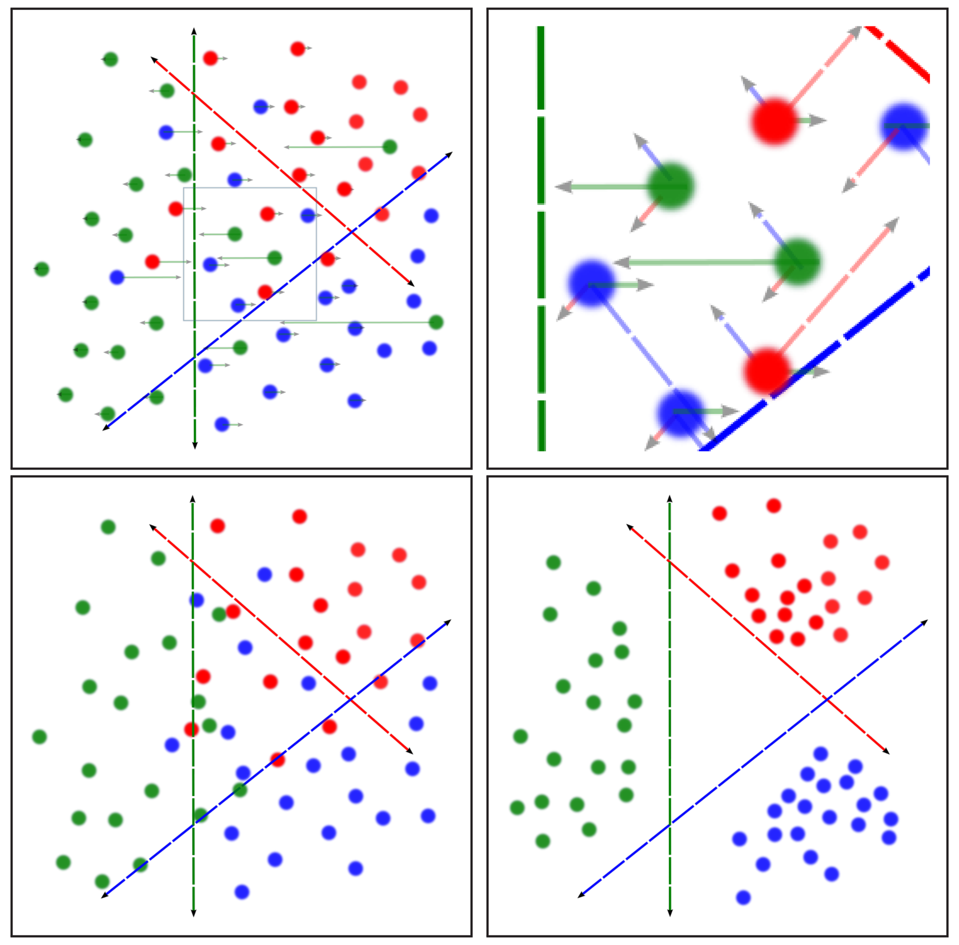

**Figure 2:** (top-left) First, good embedding is used to calculate separating hyperplanes. The loss applies to all samples by green plane is indicated with arrows. (bottom-left) After one update for only green plane. (bottom-right) After many iteration, almost convergence.

The loss is computed by the following formula:

$$\mathcal{L}_M=\lambda_M\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1-\bar{\delta}(y_i=j)}{m-1}e^{-\frac{\bar{\delta}(y_i=j)(w_j^T x_i+b)}{||w_j||_2}}$$

where separating hyperplane for $j$th class is defined as $w_j^T x+b=0$ and $\bar{\delta}(condition)$ equals to 1 if the condition is satisfied and $-1$ otherwise.

## Comparison to State-of-the-art

| Method | #Training Images | #Ids | Input Size | Network (#Params.) | FT on YTF or LFW | Accuracy on YTF (%) | Accuracy on LFW (%) |
|---|---|---|---|---|---|---|---|
| DeepFace [1] | 4.4M | 4,030 | 152×152 | AlexNet(120M) | No | 91.4 | 97.35 |
| VGG Face [2] | 2.62M | 2,622 | 224×224 | VGG(138M) | Yes | 97.3 | 98.95 |
| VGG Face [2] | 2.62M | 2,622 | 224×224 | VGG(138M) | No | 91.6 | - |
| $\mathcal{L}_S+\mathcal{L}_M$ | 0.83M | 2,558 | 96×96 | NNS1(26M) | No | **92.44** | **96.03** |

## Embedding of 3 Identities

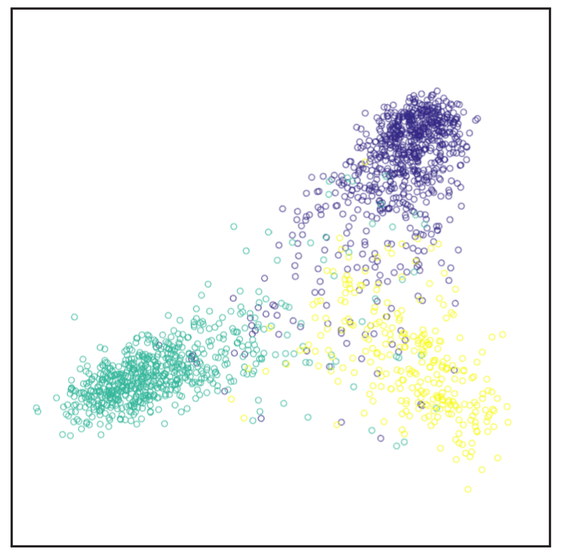

**Figure 3:** Softmax Loss Alone

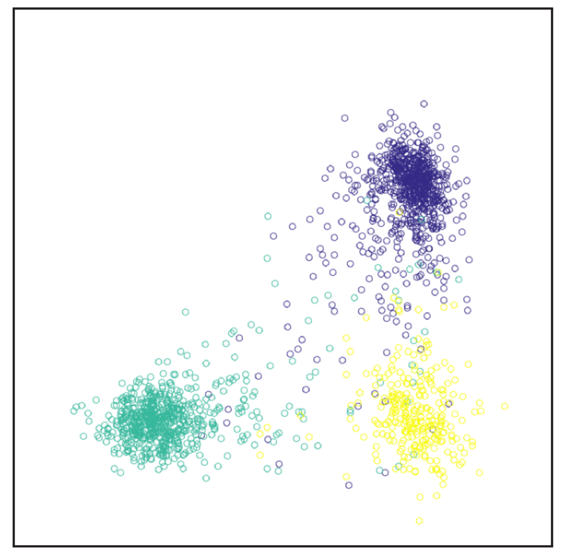

**Figure 4:** Softmax & Center L.

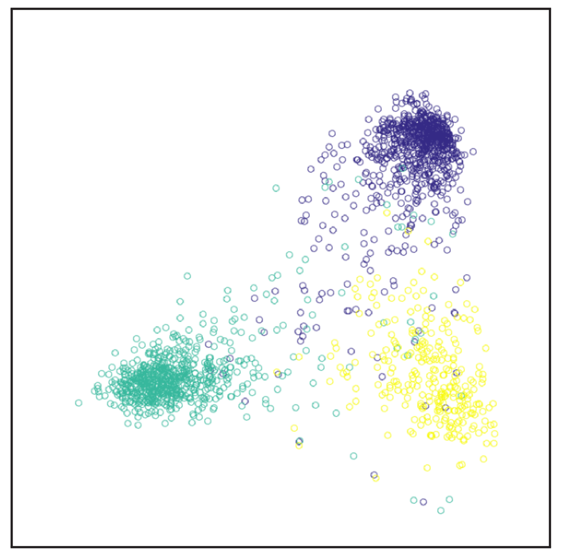

**Figure 5:** Softmax & Pushing L.

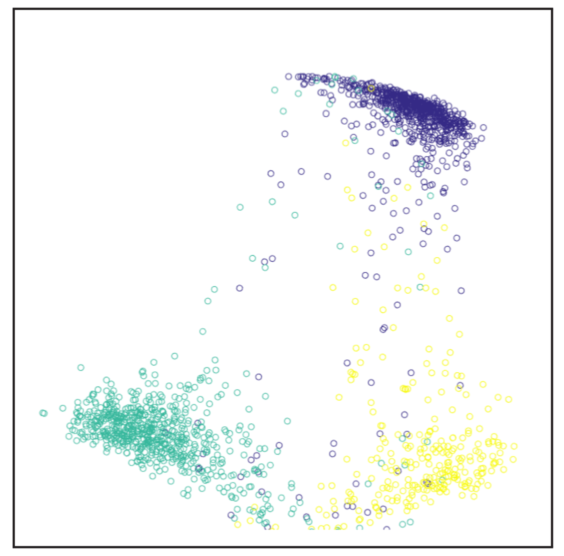

**Figure 6:** Soft. & Max-Margin L.

## Comparison of Set-based Losses

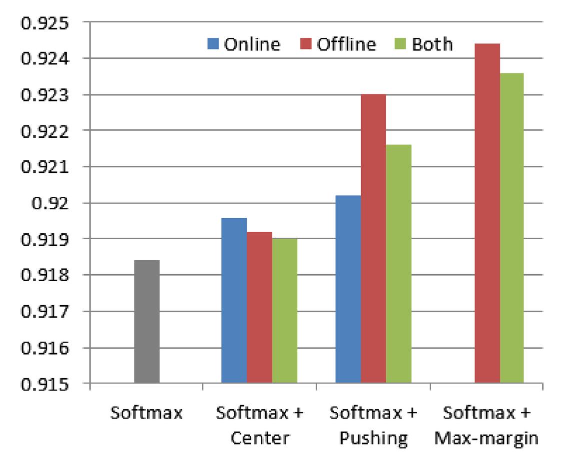


**Figure 7:** Accuracies on YTF

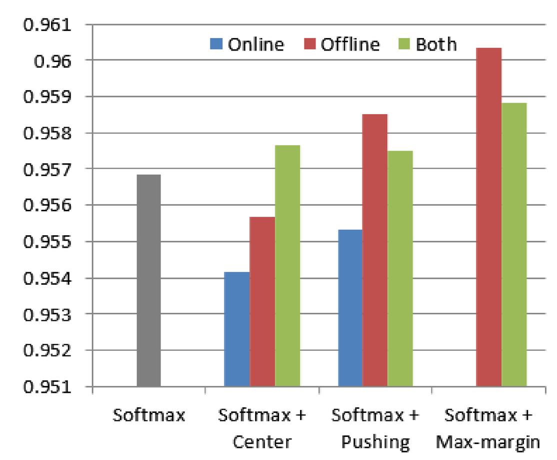


**Figure 8:** Accuracies on LFW

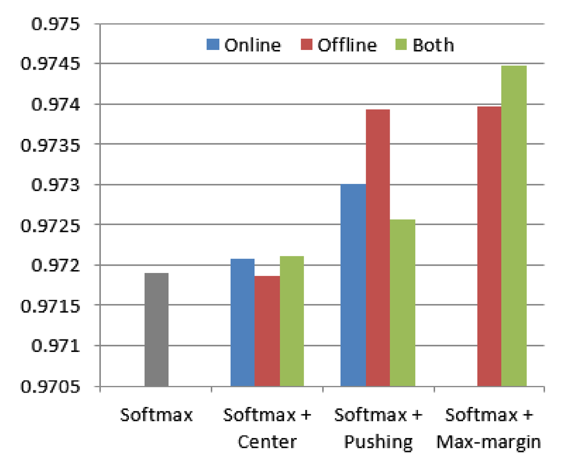


**Figure 9:** AUC on YTF

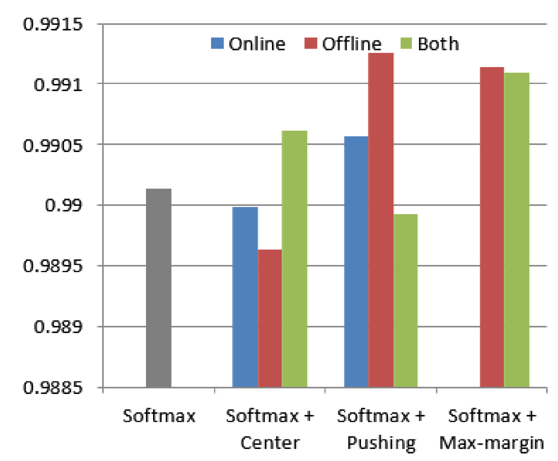


**Figure 10:** AUC on LFW

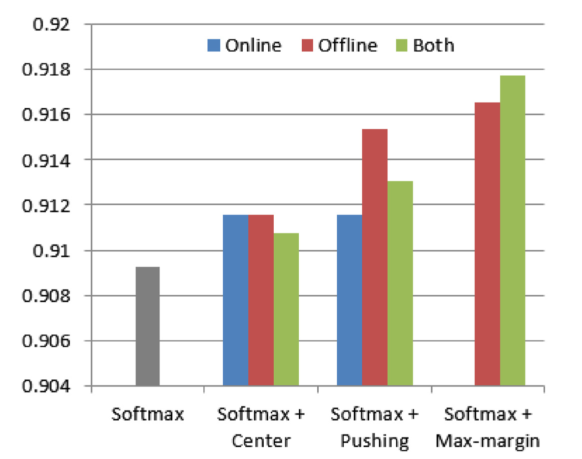


**Figure 11:** 100%- EER on YTF

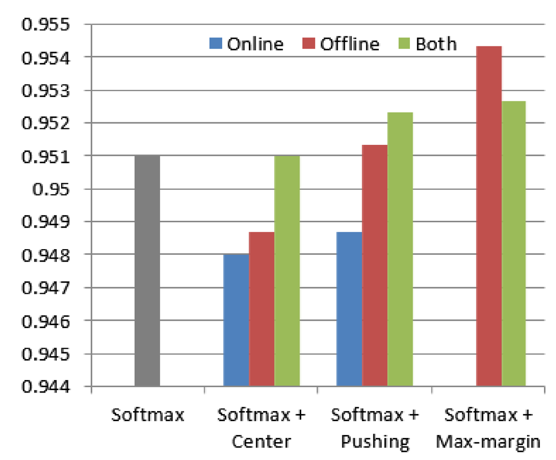


**Figure 12:** 100%- EER on LFW