



# DocFace: Matching ID Document Photos to Selfies

---

Yichun Shi and Anil K. Jain

Michigan State University

# Why ID – selfie ?



Border Control



Access Control



Online Service

# Datasets



## MS-Celeb-1M:

- 5M Images after cleaning
- 98K identities
- User for pre-training



## Private ID-selfie Dataset:

- 116,914 Images
  - 53,591 Identities
  - 53,054 classes have only a pair of ID photo and selfie
  - 5-fold for cross-validation
- Heterogeneous images
- Shallow and wide dataset

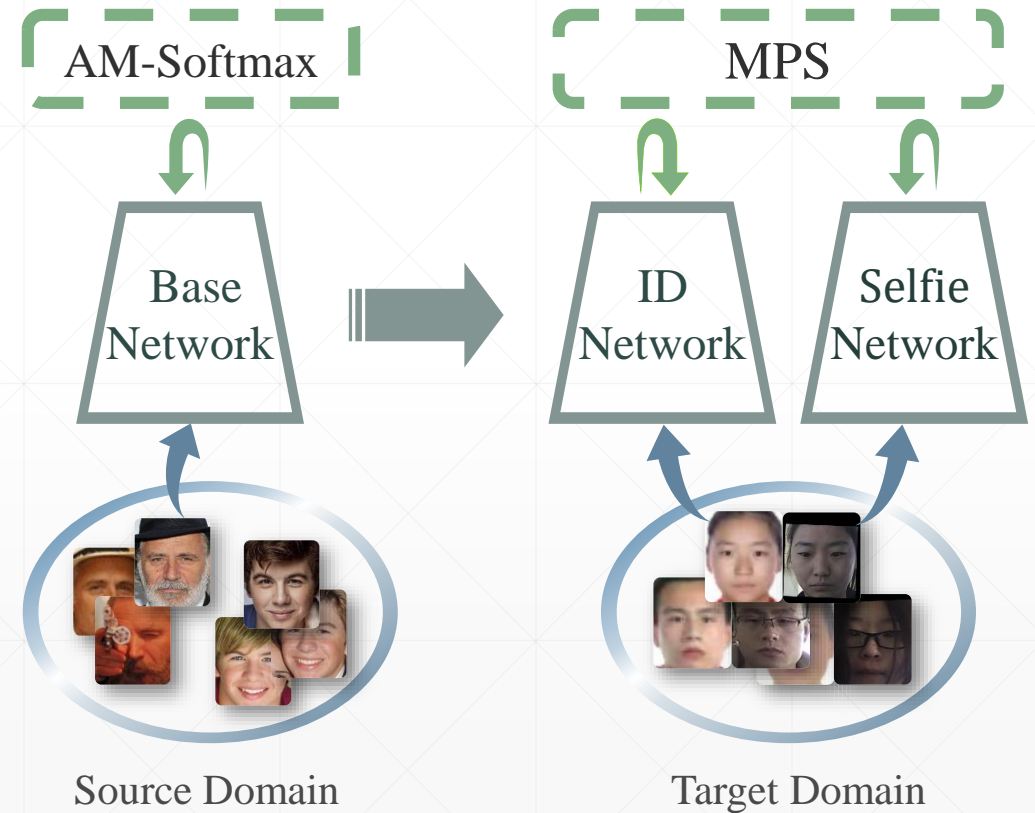
# Methodology (BTAS paper)

- Pre-training a model on MS-Celeb-1M
- Parameters transferred to a pair of networks.
- Maximum Pairwise Score loss:

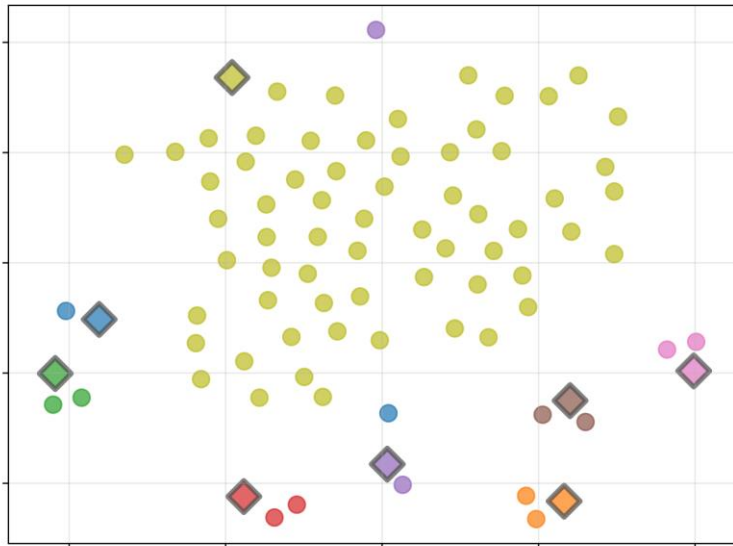
$$\mathcal{L}_t = [\max_{j \neq i} (\max(\cos \theta_{j,i}, \cos \theta_{i,j})) - \cos \theta_{i,i} + m']_+$$

where

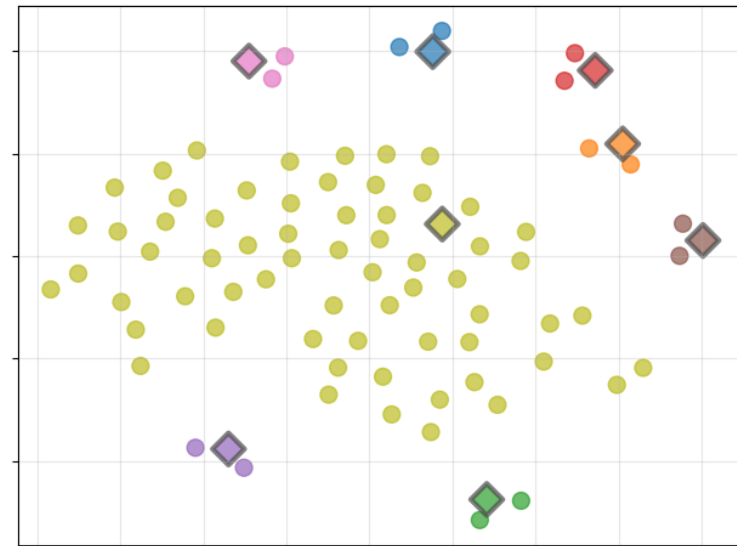
$$\begin{aligned} \cos \theta_{i,j} &= g_i^T h_j \\ g_i &= \frac{\mathcal{G}(x_{i1}^t)}{\|\mathcal{G}(x_{i1}^t)\|_2^2} \\ h_i &= \frac{\mathcal{H}(x_{i2}^t)}{\|\mathcal{H}(x_{i2}^t)\|_2^2} \end{aligned}$$



# Problem with classification loss (ongoing work)



Original



DWI

Underfitting of the classifier weights

# Dynamic Weight Imprinting

- The reason for the worse performance classification-based loss is the underfitting of the classifier weights.

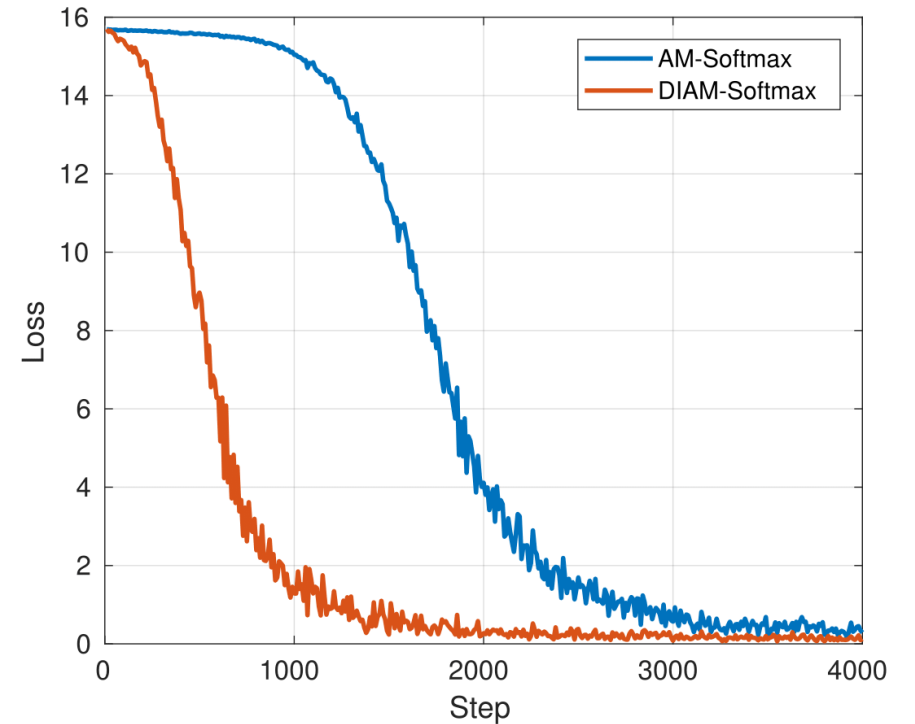
$$\mathcal{L} = -\log \frac{\exp(a_j^{(i)})}{\sum_k \exp(a_k^{(i)})}$$

$$a_j^{(i)} = \begin{cases} s \mathbf{w}_j^T \mathbf{f}_i - m, & \text{if } j = y_i \\ s \mathbf{w}_j^T \mathbf{f}_i, & \text{otherwise} \end{cases}$$

- Dynamic Weight Imprinting (DWI):

$$\mathbf{w}_j = \frac{\mathbf{w}_j^*}{\|\mathbf{w}_j^*\|_2},$$

$$\mathbf{w}_j^* = (1 - \alpha) \mathbf{w}_j + \alpha \mathbf{w}_j^{\text{batch}}$$



Accelerated Convergence



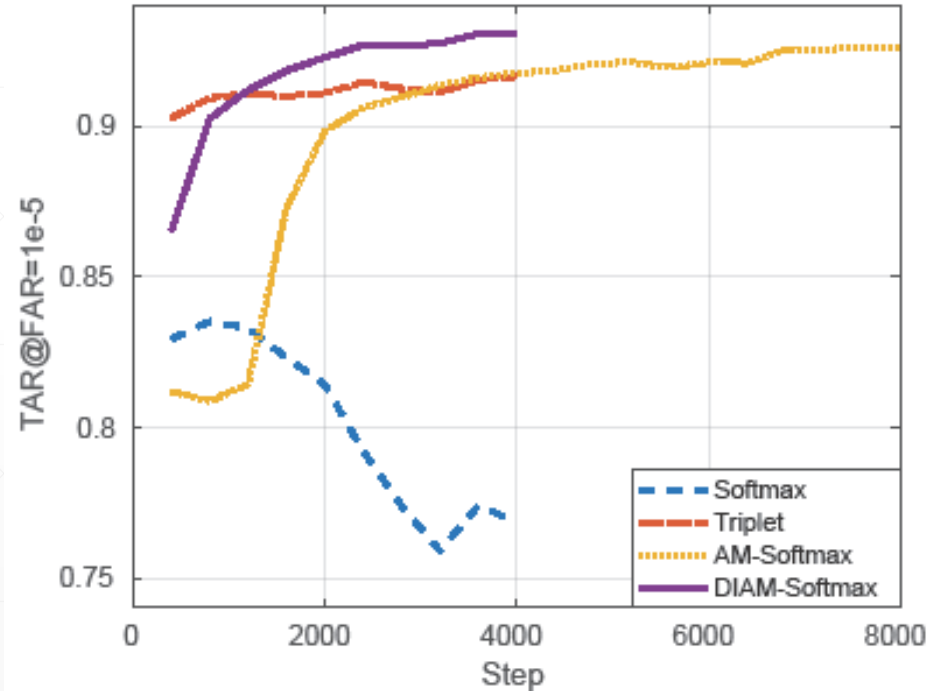
# Experiments

## Comparison of difference loss functions

Method	True Accept Rate (%)		
	FAR=0.001%	FAR=0.01%	FAR=0.1%
Base Model	77.69 ± 2.02	85.95 ± 1.67	92.43 ± 0.98
Softmax	83.51 ± 1.76	90.14 ± 1.44	94.53 ± 0.81
A-Softmax [20]	N/C	N/C	N/C
AM-Softmax [22]	92.53 ± 1.09	95.57 ± 0.57	97.23 ± 0.42
Contrastive [39]	91.13 ± 1.65	95.05 ± 0.77	97.18 ± 0.47
Triplet [19]	91.68 ± 1.21	95.42 ± 0.70	97.26 ± 0.45
MPS [39]	91.79 ± 1.16	95.43 ± 0.65	97.27 ± 0.44
<i>DIAM-Softmax</i>	<b>93.16 ± 0.85</b>	<b>95.95 ± 0.54</b>	<b>97.51 ± 0.40</b>

## Comparison with static weight imprinting

Weight Update	True Accept Rate (%)		
	FAR=0.001%	FAR=0.01%	FAR=0.1%
Static - fixed	90.97 ± 1.01	94.76 ± 0.64	96.91 ± 0.46
Static - periodical	92.95 ± 0.85	95.88 ± 0.50	97.43 ± 0.39
<i>DWI</i>	<b>93.16 ± 0.85</b>	<b>95.95 ± 0.54</b>	<b>97.51 ± 0.40</b>



Comparison with three representative loss functions

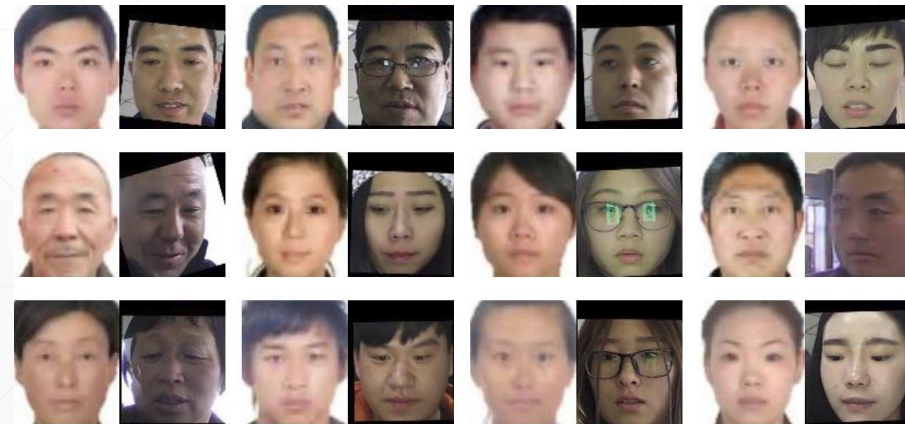
# Experiments

Comparison on Private ID-Selfie Dataset

Method	True Accept Rate (%)		
	FAR=0.001%	FAR=0.01%	FAR=0.1%
COTS-1	58.62 ± 2.30	68.03 ± 2.32	78.48 ± 1.99
COTS-2	91.53 ± 1.96	94.41 ± 1.84	96.50 ± 1.78
CenterFace [49]	27.37 ± 1.27	41.38 ± 1.43	59.29 ± 1.42
SphereFace [20]	7.96 ± 0.68	21.15 ± 1.63	50.76 ± 1.55
<i>Base model</i>	77.69 ± 2.02	85.95 ± 1.67	92.43 ± 0.98
<i>DocFace+</i>	<b>93.16 ± 0.85</b>	<b>95.95 ± 0.54</b>	<b>97.51 ± 0.40</b>



Example False Accept Pairs



Example False Accept Pairs





**Thank you!**