Paper Link

Ring loss: Convex Feature Normalization for Face Recognition

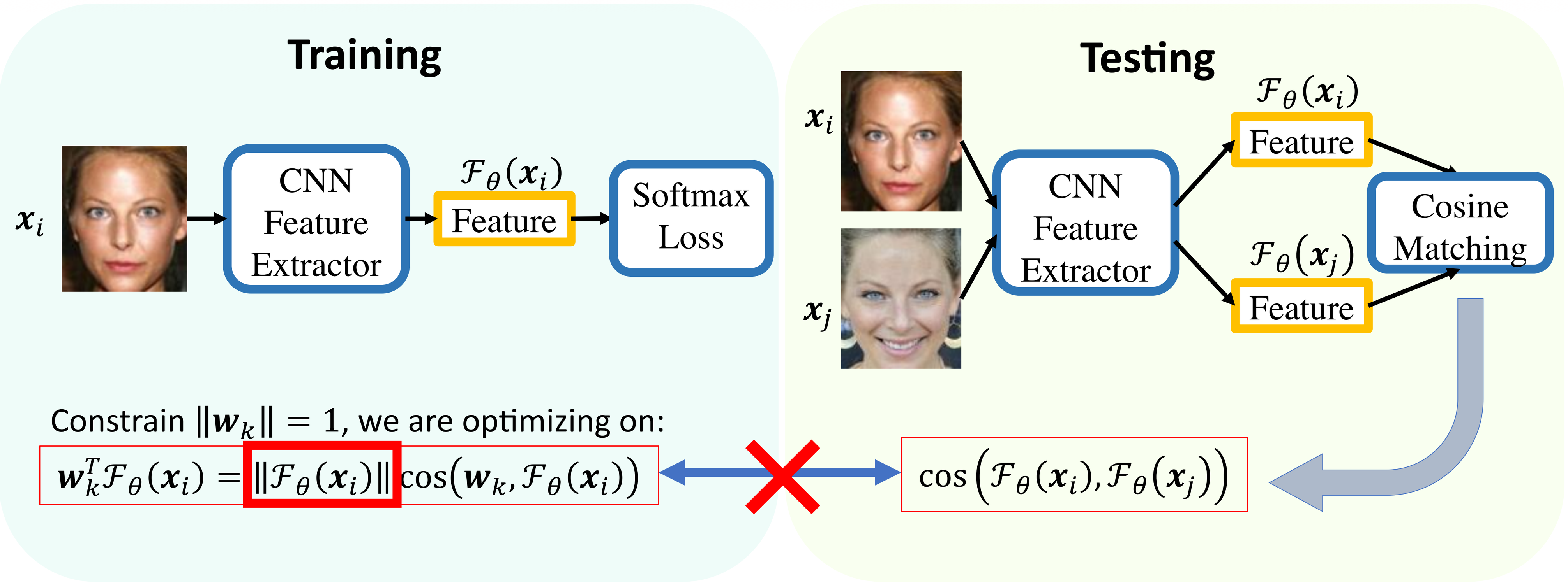
Yutong Zheng, Dipan K. Pal, and Marios Savvides

yutongzh@andrew.cmu.edu, dipanp@andrew.cmu.edu, msavvid@ri.cmu.edu

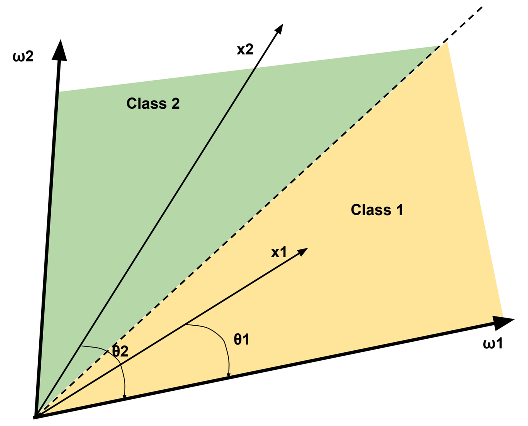
2018 CVPR SALT LAKE CITY • JUNE 18-22

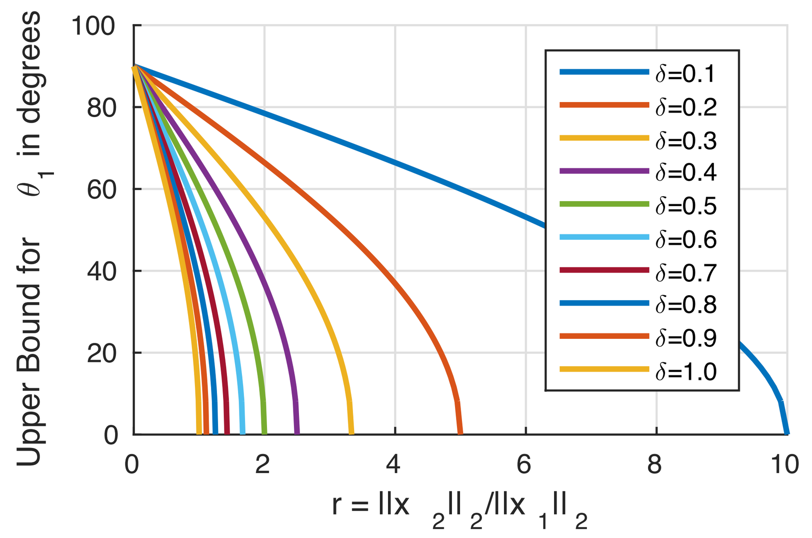
The Problem

Problem 1: Inconsistency between Training and Testing



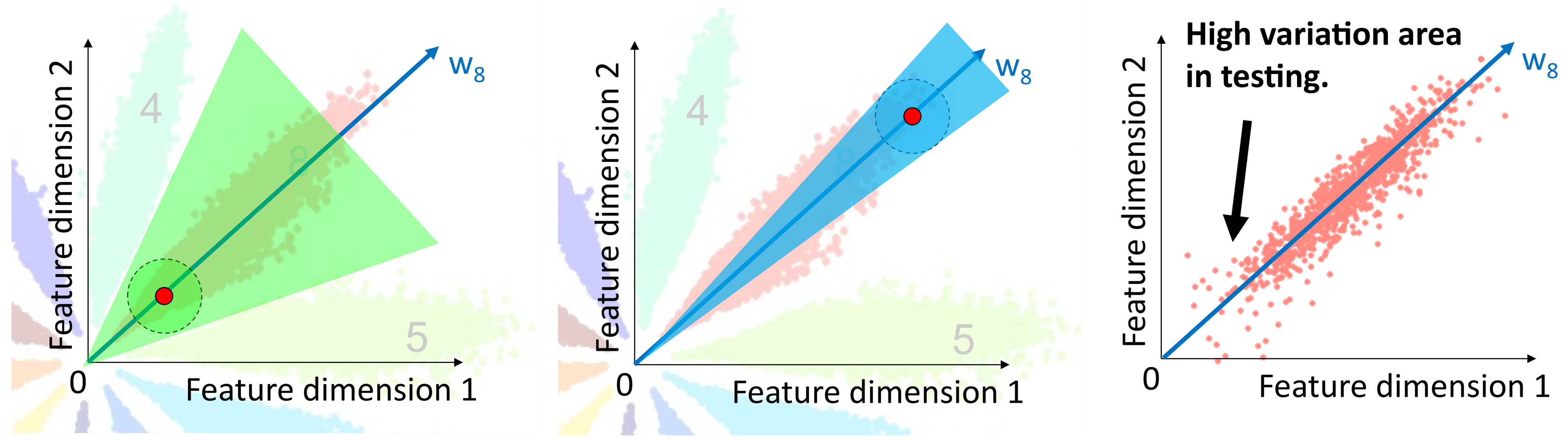
Effect on Classification Angular Margin:





Difference in feature norm leads to **imbalance in class angular margin**

Problem 2: High Angular Variation for Low-norm Features

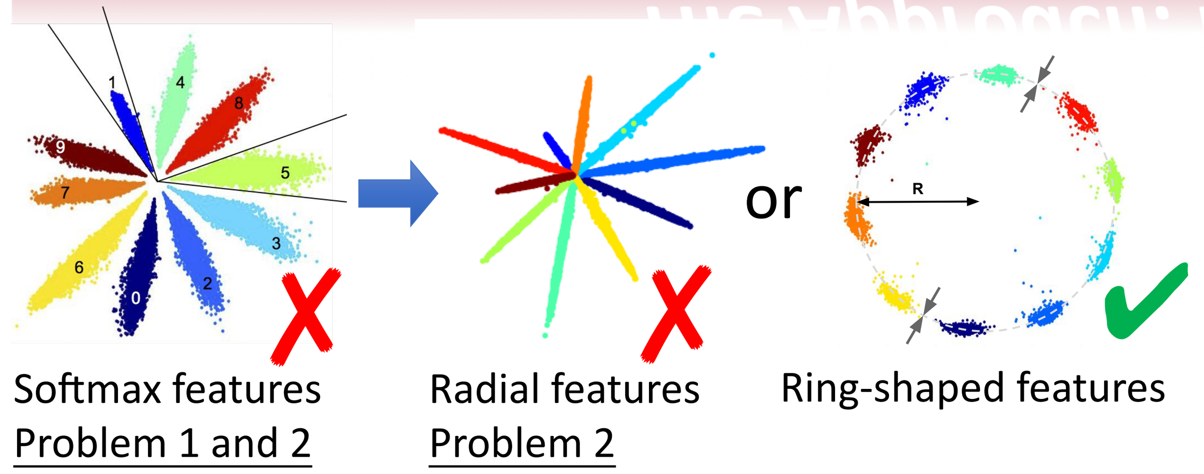


Angular variation for **low norm features** is usually higher than **high norm features** due to variation in feature samples (e.g. ***during testing***).

The Problem

For faces, **low norm features** are usually *low-resolution, blurry and off-angled* **hard examples**, which **we also want to handle**. *Test samples follow similar behavior.*

The Approach: Ring Loss



$$\min_{\theta,w} L_{SM}$$

constrain $\|\mathcal{F}_\theta(x_i)\|$

$$\min_{\theta,w,R} L_{SM} \quad s.t.\ \|\mathcal{F}_\theta(x_i)\| = R, \forall i$$

reshape

$$\min_{\theta,w,R} L_{SM} + L_R$$

$$L_R = \frac{\lambda}{2N}\sum_{i}^{N}(\|\mathcal{F}_\theta(x_i)\| - R)^2$$

- λ: loss weight $\in \mathbf{R}^+$
- R: target radius for all feature (trainable) $\in \mathbf{R}^+$

Objective: Minimize the difference between *feature norm* and a *shared radius R*.

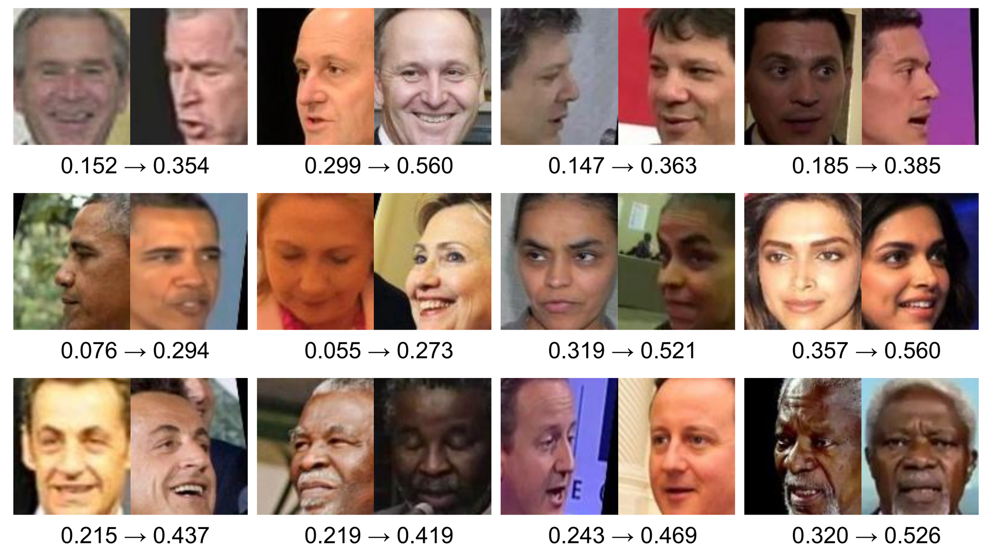
Result: All features share roughly the same norm R.

- ✓ Classification loss only depend on cosine.
- ✓ Reduction in **low norm features** during testing.

The Experiments

1. Face Matching Score Improvement:

Softmax Loss → Softmax Loss + Ring Loss



L2-Softmax Loss → Softmax Loss + Ring Loss

The Experiments

2. Face Verification Results:

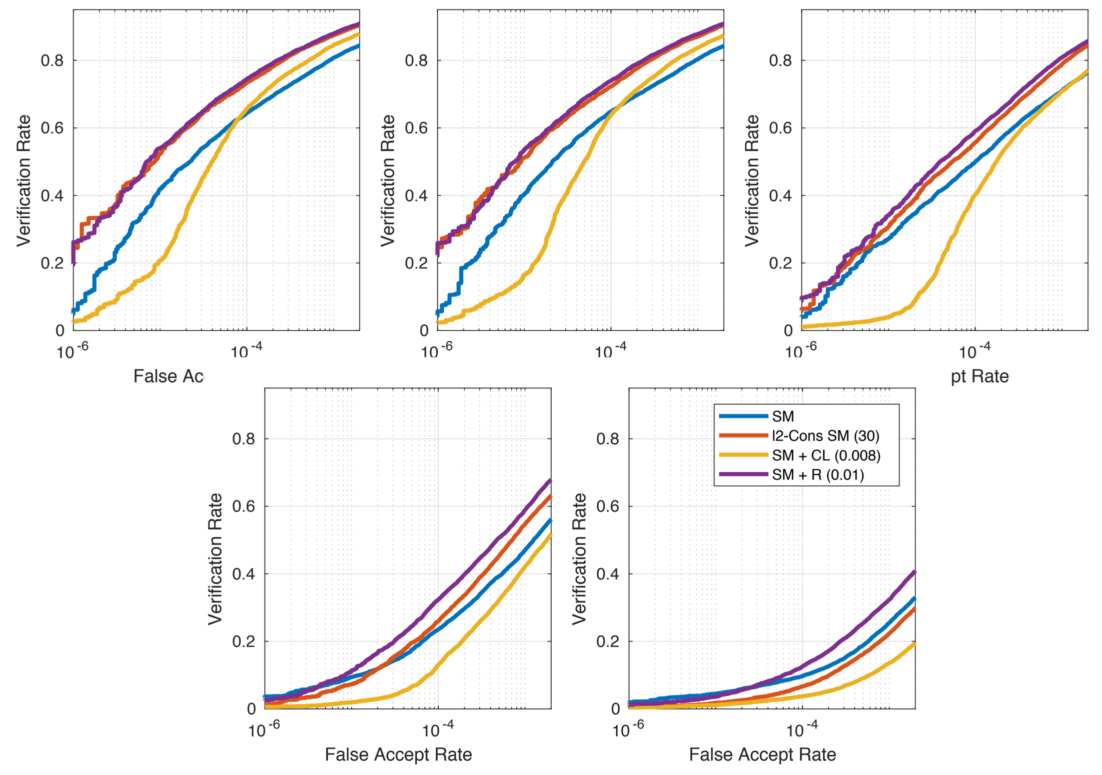
Method	Acc % (MegaFace)	10^{-3} (CFP)
SM	56.36	55.86
$l2$-Cons SM (30) [17]	72.22	82.14
$l2$-Cons SM (20) [17]	70.29	83.69
$l2$-Cons SM (10) [17]	66.20	76.77
SM + CL [26]	67.24	78.94
SF [14]	74.95	89.94
SF + CL [26, 14]	71.15	82.97
SM + R (0.01)	71.10	87.43
SM + R (0.001)	71.67	81.29
SM + R (0.0001)	69.41	76.30
SF + R (0.03)	73.05	86.23
SF + R (0.01)	74.93	**90.94**
SF + R (0.001)	**75.22**	87.69
SF + R (0.0001)	74.45	88.17

Table 2: Identification rates on MegaFace with 1 million distractors (Accuracy %) and Verification rates at 10^{-3} FAR for the CFP Frontal vs. Profile protocol.

Method	10^{-6}	10^{-5}	10^{-4}	10^{-3}
Bodla *et. al.* Final1 [3]	-	-	69.81	82.89
Bodla *et. al.* Final2 [3]	-	-	68.45	82.97
Lin *et. al.* [12]	-	-	72.52	83.55
SM	6.16	42.03	64.52	80.86
$l2$-Cons SM (30) [17]	24.47	52.32	73.36	87.46
$l2$-Cons SM (20) [17]	21.14	48.82	68.84	85.34
$l2$-Cons SM (10) [17]	13.28	36.08	57.80	78.36
SM + CL [26]	2.88	20.87	65.71	84.55
SF [14]	28.51	63.92	82.29	90.58
SF + CL [26, 14]	28.99	53.36	72.91	86.14
SM + R (0.01)	25.17	52.60	73.56	87.50
SM + R (0.001)	26.62	54.13	74.56	87.93
SM + R (0.0001)	17.35	50.65	71.06	85.48
SF + R (0.03)	27.27	56.84	76.97	88.75
SF + R (0.01)	**35.18**	**65.02**	**82.74**	**90.99**
SF + R (0.001)	32.19	63.13	81.62	90.17
SF + R (0.0001)	32.01	63.12	81.57	90.24

Table 4: Verification % on the Janus CS3 1:1 verification protocol.

3. Performance on Low-resolution images:



Method	10^{-5}	10^{-4}	10^{-3}
$l2$-Cons SM* (101) [17]	-	87.9	93.7
$l2$-Cons SM* (101x) [17]	-	88.3	93.8
SM	60.52	69.69	83.10
$l2$-Cons SM (30) [17]	73.29	80.65	90.72
$l2$-Cons SM (20) [17]	67.63	76.88	89.89
$l2$-Cons SM (10) [17]	53.74	68.58	83.42
SM + CL [26]	46.01	74.10	88.32
SF [14]	78.52	88.0	**93.24**
SF + CL [26, 14]	72.35	81.11	89.26
SM + R (0.01)	72.53	79.1	90.8
SM + R (0.001)	78.41	85.0	91.5
SM + R (0.0001)	69.23	82.30	89.20
SF + R (0.03)	79.54	85.37	91.64
SF + R (0.01)	**82.41**	**88.5**	**93.22**
SF + R (0.001)	79.74	87.71	92.62
SF + R (0.0001)	80.13	86.34	92.57

Table 3: Verification % on the IJB-A Janus 1:1 verification protocol. $l2$-Cons SM* indicates the result reported in [17] which uses a 101 layer ResNet/ResNext architecture.

As we manually decrease the resolution of the image, performances of all face recognition methods drop. However, since Ring loss is able to efficiently learn hard examples while **regularizing the norm**, it *consistently outperforms other methods*.

Paper Link

Carnegie Mellon University
CyLab Biometrics Center