MagFace: A Universal Representation for Face Recognition and Quality Assessment

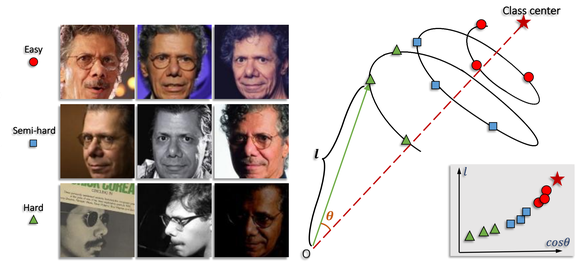
Qiang Meng, Shichao Zhao, Zhida Huang, Feng Zhou
Algorithm Research, Aibee Inc.
{qmeng,sczhao,zdhuang,fzhou}@aibee.com

Introduction

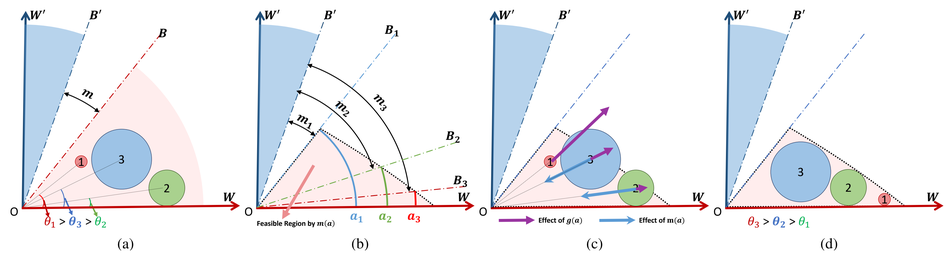
MagFace is a category of losses which

- only requires recognition labels
- pulls the easier samples closer to the class center
- pushs the easier samples away from the origin o

feature direction $\rightarrow$ *recognition*
feature magnitude $\rightarrow$ *quality*



Methods



(a) (b) (c) (d)

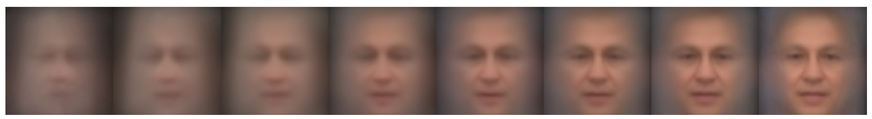
$$L_{Mag} = \frac{1}{N}\sum_{i=1}^{N} L_i, \quad \text{where } L_i = -\log \frac{e^{s\cos(\theta_{y_i}+m(a_i))}}{e^{s\cos(\theta_{y_i}+m(a_i))}+\sum_{j\neq y_i} e^{s\cos\theta_j}} + \lambda_g g(a_i).$$

a_i – feature manigutde for sample i
$m(a_i)$ –the magnitude-aware angular margin
$g(a_i)$ – the regularizer

Requirements for MagFace

- $m(a_i)$ is an increasing convex function in $[l_a, u_a]$ and $m'(a_i) \in (0, k]$, where l_a, u_a, k is a pre-defined and arbitary.
- $g(a_i)$ is a strictly convex function with $g'(u_a) = 0$;
- $\lambda_g \geq \frac{sk}{-g'(l_a)}$.

Visualization of 100k Faces from IJB-C



(a) mean: 22.84 range: (-∞, 24) # of faces: 3692
(b) mean: 25.13 range: [24, 26) # of faces: 9955
(c) mean: 27.03 range: [26, 28) # of faces: 15459
(d) mean: 29.03 range: [28, 30) # of faces: 17565
(e) mean: 31.01 range: [30, 32) # of faces: 20627
(f) mean: 32.99 range: [32, 34) # of faces: 19743
(g) mean: 34.80 range: [34, 36) # of faces: 11238
(h) mean: 36.55 range: [36, ∞) # of faces: 1721

Properties of MagFace

Property of Convergence. *For $a_i \in [l_a, u_a]$, L_i is a strictly convex function which has a unique optimal solution a_i^*.*

Lemma 1. *Assume that f_i is top-k correctly classified and $m(a_i) \in [0, \pi/2]$. If the number of identities n is much larger than k (i.e., $n \gg k$), the probability of $\theta_{y_i} + m(a_i) \in [0, \pi/2]$ approaches 1.*

Lemma 2. *A unique optimal solution a_i^* exists in $[l_a, u_a]$.*

Lemma 3. *For $a_i \in [l_a, u_a]$, L_i is a strictly convex function of a_i.*

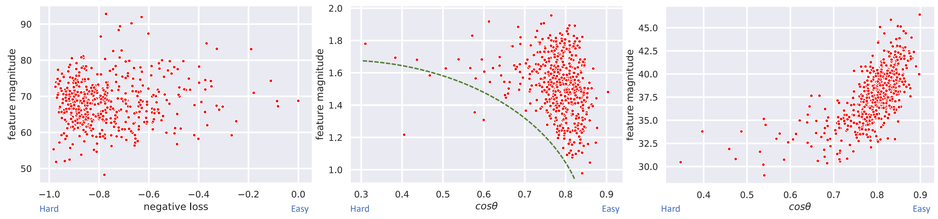
Property of Monotonicity. *The optimal a_i^* is monotonically increasing as the cosine-distance to its class center decreases and the cos-distances to other classes increase.*

Lemma 4. *With fixed f_i and $W_j, j \in \{1, \cdots, n\}, j \neq y_i$, the optimal feature magnitude a_i^* is monotonically decreasing if the cosine-distance to its class center W_{y_i} increases.*

Lemma 5. *With other things fixed, the optimal feature magnitude a_i^* is monotonically decreasing with a decreasing inter-class distance B.*

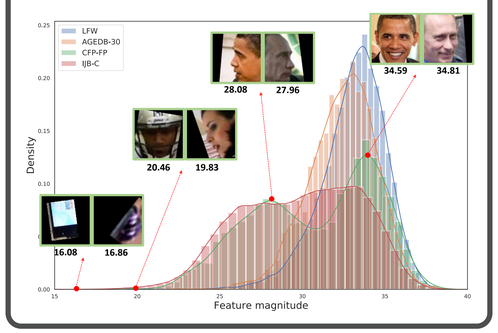
Analysis on Feature Magnitude

1. In MagFace, feature magnitudes and difficulties for recognition are positively related.
2. A lower bound exists in ArcFace.



(a) Softmax (b) ArcFace (c) MagFace

Magnitude Distributions



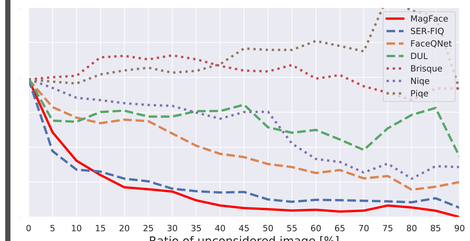
Experiment - Recognition

Method	IJB-C (TAR@FAR)		
	1e-6	1e-5	1e-4
VGGFace2*	-	74.70	84.00
CenterFace*	-	78.10	85.30
CircleLoss*	-	89.60	93.95
ArcFace*	-	-	95.60
Softmax	64.07	83.68	92.40
SV-AM-Softmax	63.45	80.30	88.34
SphereFace	68.86	83.33	91.77
CosFace	87.96	92.68	95.56
ArcFace	85.65	92.69	95.74
MagFace	89.26	93.67	95.81
MagFace+	**90.24**	**94.08**	**95.97**

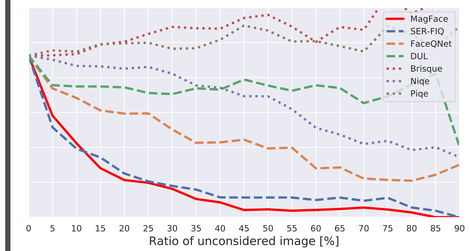
Experiment - Quality

FNMR on CFP-FP.

Evaluated by MagFace:



Evaluated by ArcFace released by its authors:



Experiment - Clustering

F-score (%) and NMI (%) on IJB-B.

Method	Net	IJB-B-1845	
		F	NMI
K-means	ArcFace	66.70	88.83
	MagFace	**66.75**	**88.86**
AHC	ArcFace	69.72	89.61
	MagFace	**70.24**	**89.99**
DBSCAN	ArcFace	72.72	90.42
	MagFace	**73.13**	**90.61**
L-GCN	ArcFace	84.92	93.72
	MagFace	**85.27**	**93.83**

Visualization of magnitudes and confidences of being class centers.