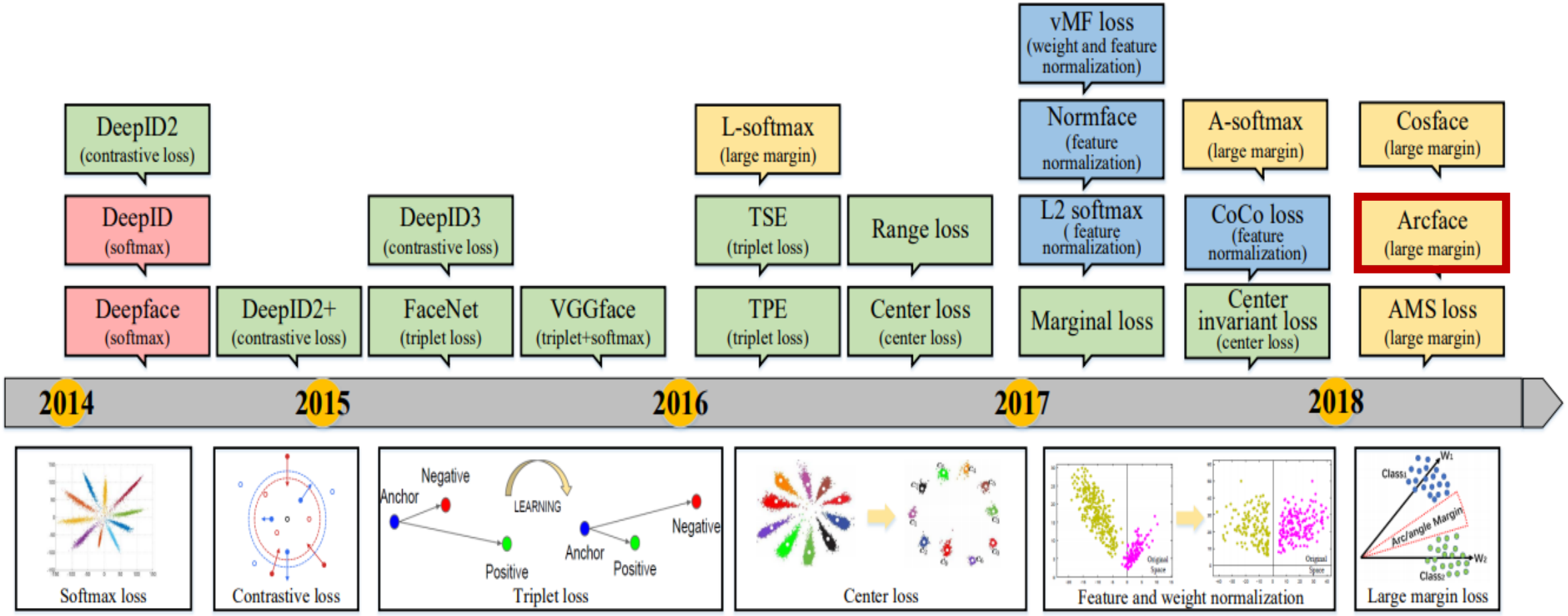
2020.07 하계방학 Part.3

ArcFace

Jiankang Deng Jia Guo Niannan Xue Stefanos Zafeiriou
Imperial College London InsightFace FaceSof

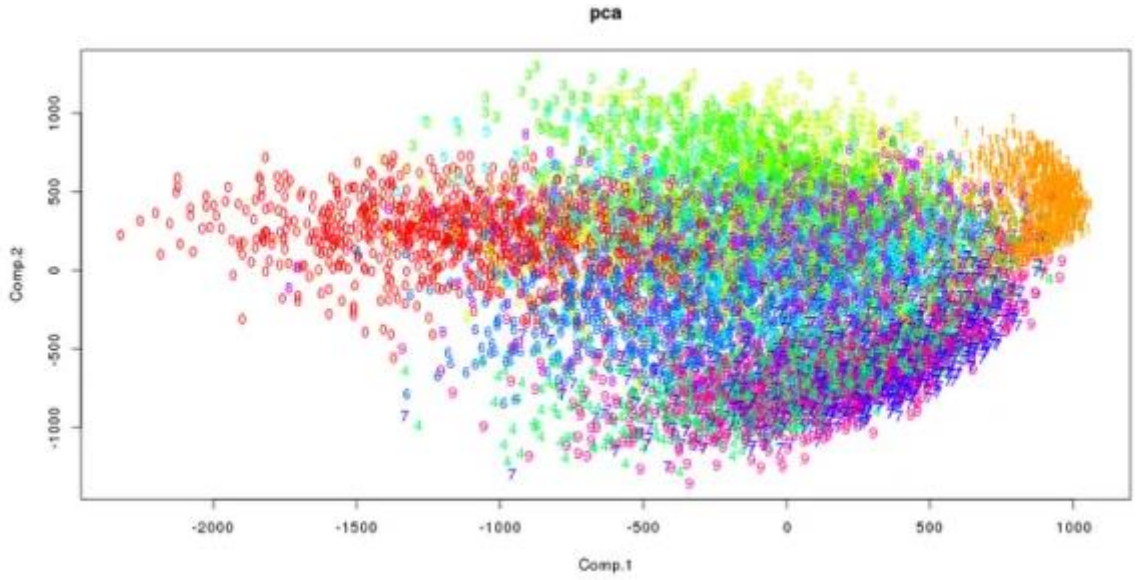
20151739 공대현

- # Face Recognition(FR) Loss의 발전

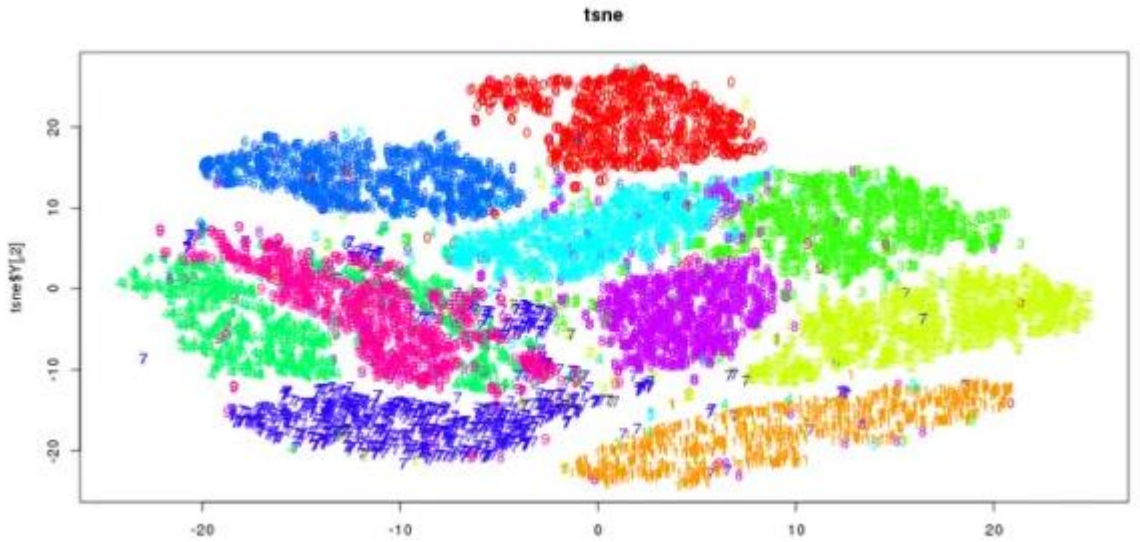


- # Metric learning Method

Before



After



Previous Skills

- Triplet Loss: Euclidean-distance-based loss

$$\mathcal{L} = y_{ij} max\left(0, \|f(x_i) - f(x_j)\|_2 - \epsilon^+\right) + (1 - y_{ij}) max\left(0, \epsilon^- - \|f(x_i) - f(x_j)\|_2\right)$$

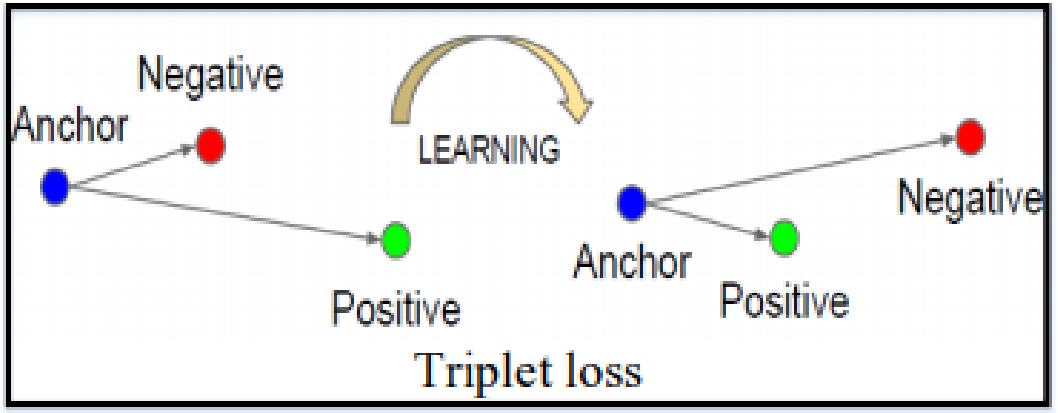
If $x_i = x_j$, y_{ij} = 1 → verification
If $x_i \neq x_j$, y_{ij} = -1 → identification

FaceNet : Triplet Loss

$$\|f(x_i^a) - f(x_i^p)\|_2^2 + \alpha < -\|f(x_i^a) - f(x_i^n)\|_2^2$$

Minimizes the distance between an anchor and a positive sample of the same identity and maximizes the distance between the anchor and a negative sample of a different identity

단점: Training instability due to
the selection of effective training samples



❖ Softmax-BCE Loss

input: x의 class가 2일 때...

$$P_1 = \frac{e^{W_1 x}}{e^{W_1 x} + e^{W_2 x} + e^{W_3 x}} \quad 0.05 \qquad P_2 = \frac{e^{W_2 x}}{e^{W_1 x} + e^{W_2 x} + e^{W_3 x}} \quad 0.8$$

$$P_3 = \frac{e^{W_3 x}}{e^{W_1 x} + e^{W_2 x} + e^{W_3 x}} \quad 0.15$$

$\log P_1 = \log 0.05 = -1.3$, $\log P_2 = \log 0.8 = -0.0969$, $\log P_3 = -0.8239$

$-\log P_1 = 1.3$, $-\log P_2 = 0.0969$, $-\log P_3 = 0.8239$

⇒ cross entropy : $Loss = -\sum_{c=1}^{C} L_c \log P_c$, $L_c = [0, 1, 0]$

$\therefore Loss = 0 \times 1.3 + 1 \times 0.0969 + 0 \times 0.8239 = 0.0969$ ← 줄어드는 방향으로 W update

예측값이 잘못됐을 때 : $Loss = 1 \times 1.3 + 0 \times 0.0969 + 0 \times 0.8239 = 1.3$

- **Previous Skills**
  - Angular loss: Softmax → Revised Softmax → SphereFace → Cosface → ArcFace → Combined

Original Softmax Loss:

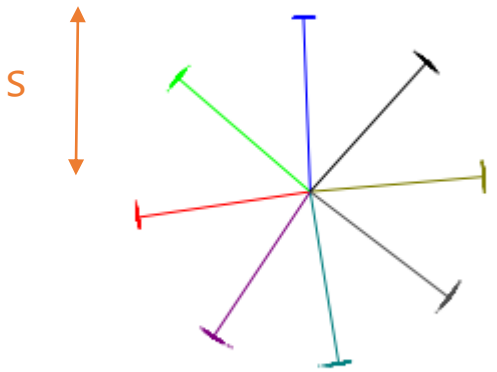
The most widely used classification loss function, softmax loss, is presented as follows:

$$L_1 = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^{n} e^{W_j^T x_i + b_j}}, \quad (1)$$

b =0 ,
x is normalized and rescaled to s ,
W x = ||W|| ||x|| cosθ ,

Revised Softmax Loss:

$$L_2 = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s\cos\theta_{y_i}}}{e^{s\cos\theta_{y_i}} + \sum_{j=1, j\neq y_i}^{n} e^{s\cos\theta_j}}.$$



(b) ArcFace

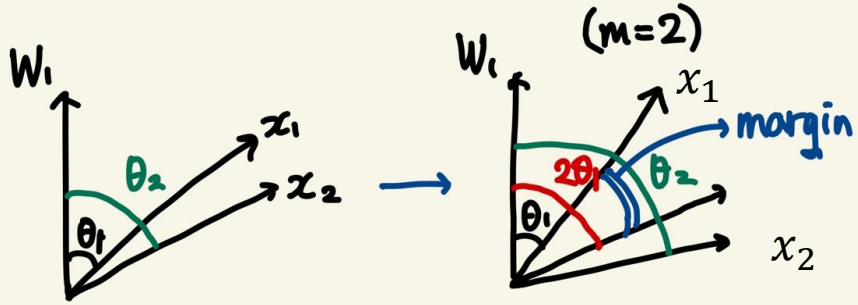
Previous Skills

- Angular loss: Softmax → Revised Softmax→ SphereFace → Cosface → ArcFace → Combined

Revised Softmax Loss: $$L_2 = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s\cos\theta_{y_i}}}{e^{s\cos\theta_{y_i}}+\sum_{j=1,j\neq y_i}^{n}e^{s\cos\theta_j}}.$$

Decision boundary
$\|\boldsymbol{x}\|(\cos\theta_1-\cos\theta_2)=0$

$cos\theta \rightarrow \cos(m\theta)$



A-Softmax (SphereFace): $$L_{ang} = \frac{1}{N}\sum_{i}-\log\left(\frac{e^{\|\boldsymbol{x}_i\|\cos(m\theta_{y_i,i})}}{e^{\|\boldsymbol{x}_i\|\cos(m\theta_{y_i,i})}+\sum_{j\neq y_i}e^{\|\boldsymbol{x}_i\|\cos(\theta_{j,i})}}\right)$$

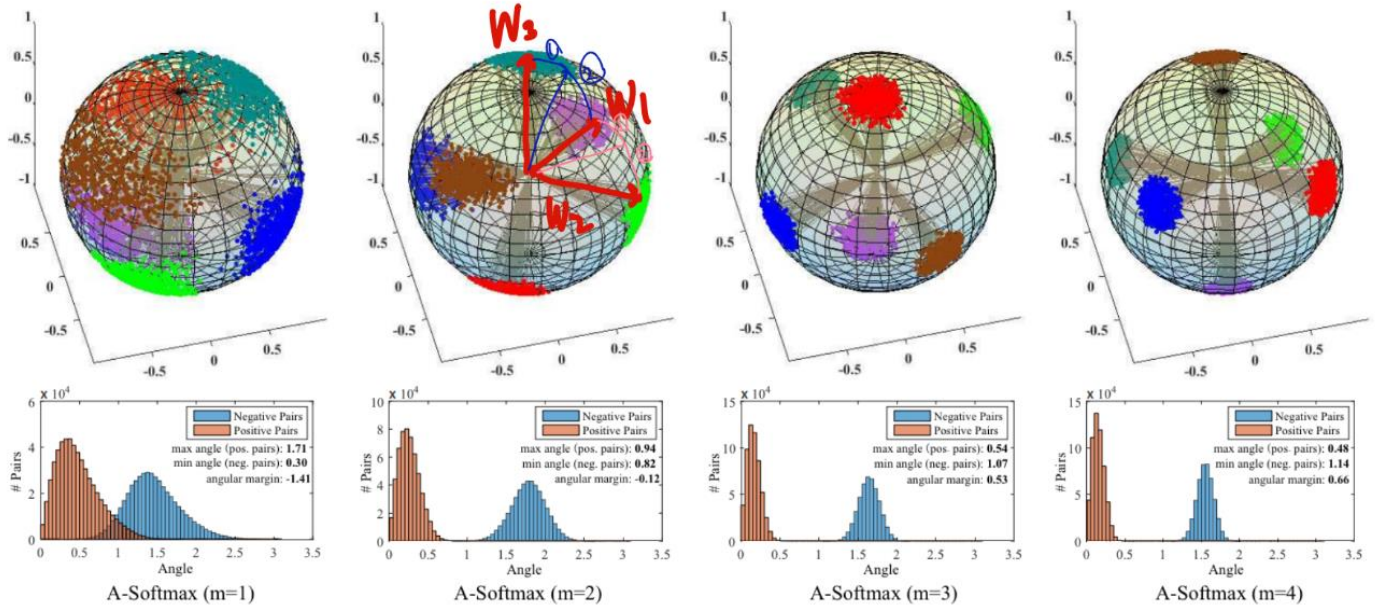
$\|\boldsymbol{x}\|(\cos m\theta_1-\cos\theta_2)=0$ for class 1 $\|\boldsymbol{x}\|(\cos\theta_1-\cos m\theta_2)=0$ for class 2

Dataset	Original	m=1	m=2	m=3	m=4
LFW	97.88	97.90	98.40	99.25	**99.42**
YTF	93.1	93.2	93.8	94.4	**95.0**

Previous Skills

- Sphere Face 의 Angular margin effect

Dataset	Original	m=1	m=2	m=3	m=4
LFW	97.88	97.90	98.40	99.25	**99.42**
YTF	93.1	93.2	93.8	94.4	**95.0**



- **Previous Skills**
  - Angular loss: Softmax → Revised Softmax→ SphereFace → Cosface → ArcFace → Combined

A-Softmax (SphereFace) : $L_{ang} = \frac{1}{N}\sum_i -\log\left(\frac{e^{\|\boldsymbol{x}_i\|\cos(m\theta_{y_i,i})}}{e^{\|\boldsymbol{x}_i\|\cos(m\theta_{y_i,i})} + \sum_{j\neq y_i} e^{\|\boldsymbol{x}_i\|\cos(\theta_{j,i})}}\right)$

$\cos(m\theta) \rightarrow cos\theta - m$

LMCL(CosFace) : $L_{lmc} = \frac{1}{N}\sum_i -\log\frac{e^{s(\cos(\theta_{y_i,i})-m)}}{e^{s(\cos(\theta_{y_i,i})-m)} + \sum_{j\neq y_i} e^{s\cos(\theta_{j,i})}}$

Decision boundary
$\|\boldsymbol{x}\|(\cos m\theta_1 - \cos\theta_2) = 0$ for class 1 $\|\boldsymbol{x}\|(\cos\theta_1 - \cos m\theta_2) = 0$ for class 2
$C_1 : \cos(\theta_1) \geq \cos(\theta_2) + m,$ $C_2 : \cos(\theta_2) \geq \cos(\theta_1) + m.$

Previous Skills

- Angular loss: Softmax → Revised Softmax→ SphereFace → Cosface → ArcFace → Combined

LMCL(CosFace) : $L_{lmc} = \frac{1}{N}\sum_{i} -\log \frac{e^{s(\cos(\theta_{y_i,i})-m)}}{e^{s(\cos(\theta_{y_i,i})-m)} + \sum_{j\neq y_i} e^{s\cos(\theta_{j,i})}}$

$\cos\theta - m \rightarrow \cos(\theta + m)$

Additive Angular Margin Loss(ArcFace) : $L_3 = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{e^{s(\cos(\theta_{y_i}+m))}}{e^{s(\cos(\theta_{y_i}+m))} + \sum_{j=1,j\neq y_i}^{n} e^{s\cos\theta_j}}.$

Decision boundary

$C_1 : \cos(\theta_1) \geq \cos(\theta_2) + m,$
$C_2 : \cos(\theta_2) \geq \cos(\theta_1) + m.$

$\hat{x}\left(cos\left(\theta_1 + m\right) - cos\theta_2\right) = 0$

- **Previous Skills**
  - Angular loss: Softmax → Revised Softmax→ SphereFace → Cosface → ArcFace → Combined

Decision boundary

Additive Angular Margin Loss(ArcFace) :

$$L_3 = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s(\cos(\theta_{y_i}+m))}}{e^{s(\cos(\theta_{y_i}+m))}+\sum_{j=1,j\neq y_i}^{n}e^{s\cos\theta_j}}.$$

$$\hat{x}\left(cos\left(\theta_1+m\right)-cos\theta_2\right)=0$$

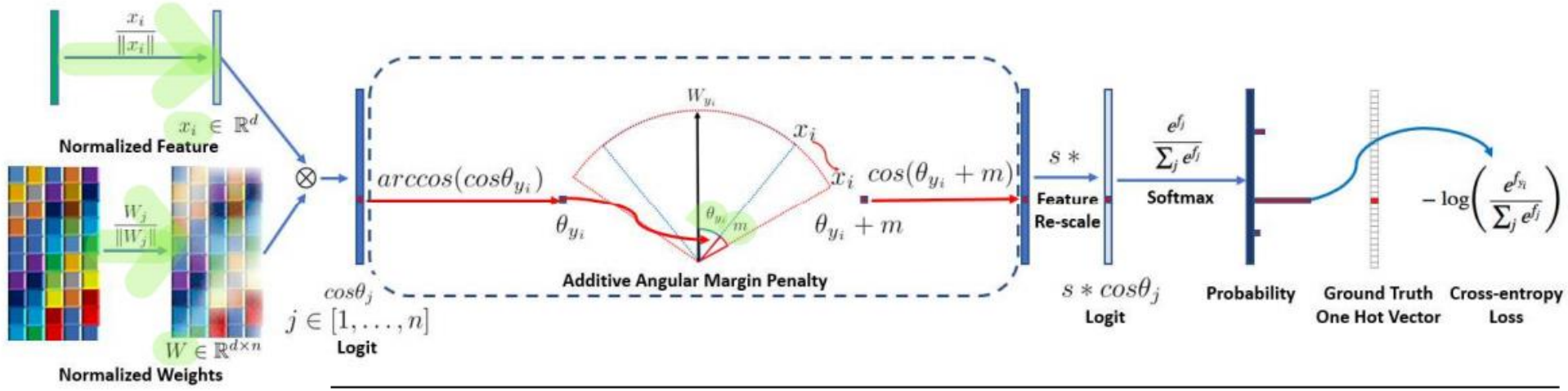
$\cos(\theta + m) \rightarrow \cos(m_1\theta + m_2) - m_3$

Combined margin :

$$L_4 = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s(\cos(m_1\theta_{y_i}+m_2)-m_3)}}{e^{s(\cos(m_1\theta_{y_i}+m_2)-m_3)}+\sum_{j=1,j\neq y_i}^{n}e^{s\cos\theta_j}}$$

ArcFace

- $$L_3 = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s(\cos(\theta_{y_i}+m))}}{e^{s(\cos(\theta_{y_i}+m))}+\sum_{j=1,j\neq y_i}^{n}e^{s\cos\theta_j}}.$$



Algorithm 1 The Pseudo-code of ArcFace on MxNet

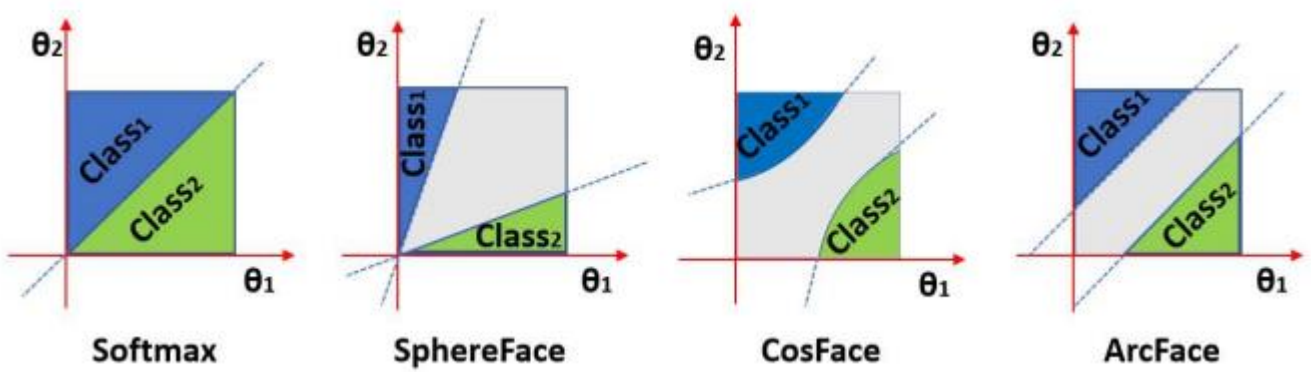
Input: Feature Scale s, Margin Parameter m in Eq. 3, Class Number n, Ground-Truth ID gt.

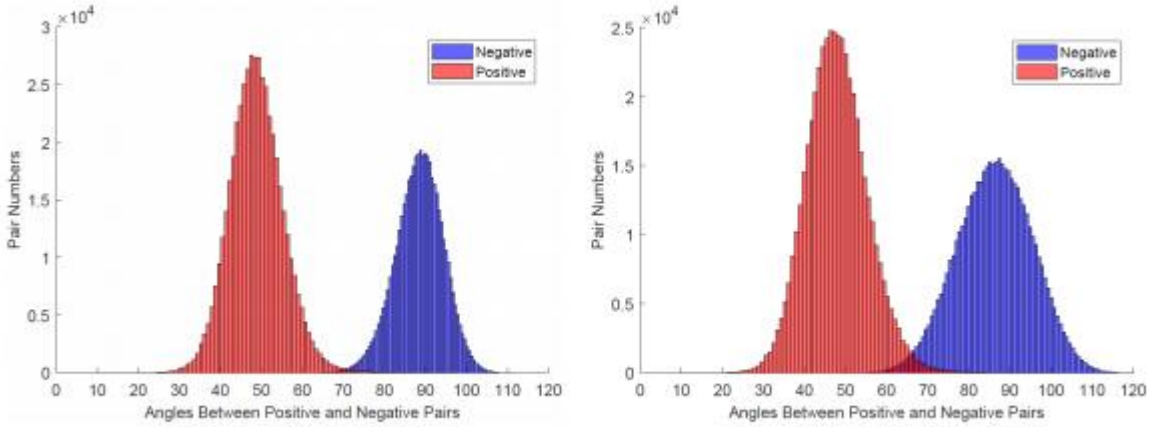
1. x = mx.symbol.L2Normalization (x, mode = 'instance')
2. W = mx.symbol.L2Normalization (W, mode = 'instance')
3. fc7 = mx.sym.FullyConnected (data = x, weight = W, no_bias = True, num_hidden = n)
4. original_target_logit = mx.sym.pick (fc7, gt, axis = 1)
5. theta = mx.sym.arccos (original_target_logit)
6. marginal_target_logit = mx.sym.cos (theta + m)
7. one_hot = mx.sym.one_hot (gt, depth = n, on_value = 1.0, off_value = 0.0)
8. fc7 = fc7 + mx.sym.broadcast_mul (one_hot, mx.sym.expand_dims (marginal_target_logit - original_target_logit, 1))
9. fc7 = fc7 * s

Output: Class-wise affinity score $fc7$.

ArcFace

Result





(a) ArcFace (b) Triplet-Loss

Loss Functions	LFW	CFP-FP	AgeDB-30
ArcFace (0.4)	99.53	95.41	94.98
ArcFace (0.45)	99.46	95.47	94.93
ArcFace (0.5)	**99.53**	**95.56**	**95.15**
ArcFace (0.55)	99.41	95.32	95.05
SphereFace [15]	99.42	-	-
SphereFace (1.35)	99.11	94.38	91.70
CosFace [35]	99.33	-	-
CosFace (0.35)	99.51	95.44	94.56
CM1 (1, 0.3, 0.2)	99.48	95.12	94.38
CM2 (0.9, 0.4, 0.15)	99.50	95.24	94.86
Softmax	99.08	94.39	92.33
Norm-Softmax (NS)	98.56	89.79	88.72
NS+Intra	98.75	93.81	90.92
NS+Inter	98.68	90.67	89.50
NS+Intra+Inter	98.73	94.00	91.41
Triplet (0.35)	98.98	91.90	89.98
ArcFace+Intra	99.45	95.37	94.73
ArcFace+Inter	99.43	95.25	94.55
ArcFace+Intra+Inter	99.43	95.42	95.10
ArcFace+Triplet	99.50	95.51	94.40