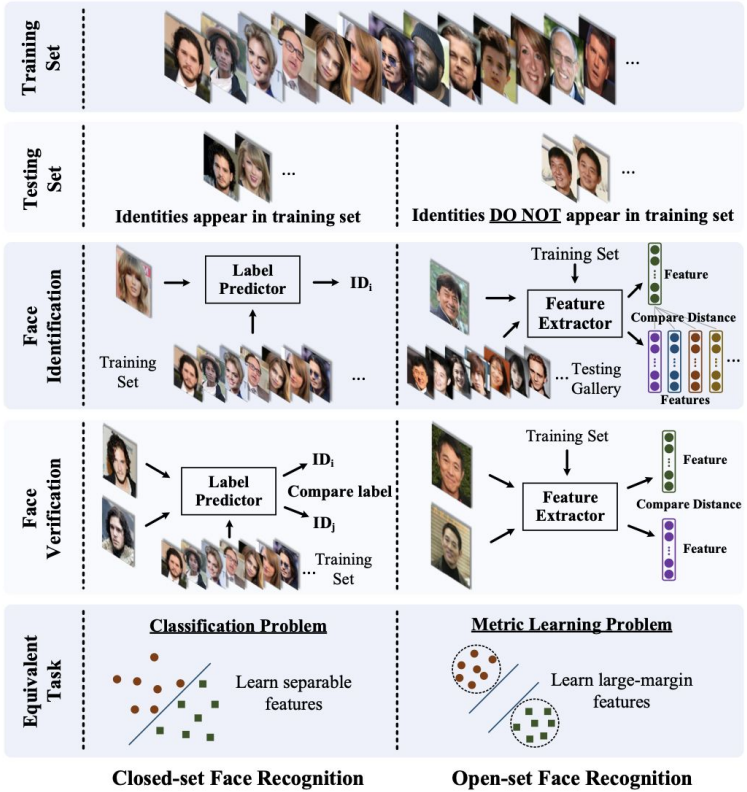
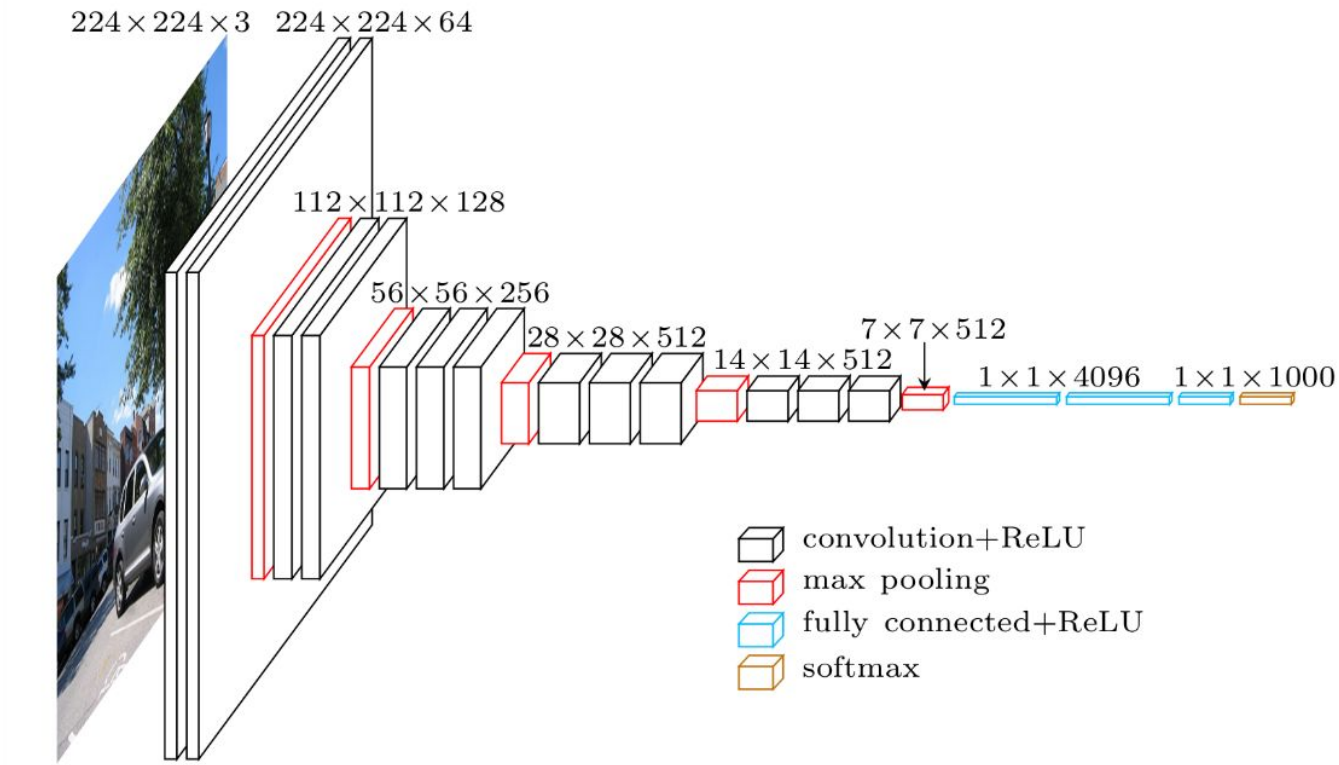
Deep Learning
Metric Learning

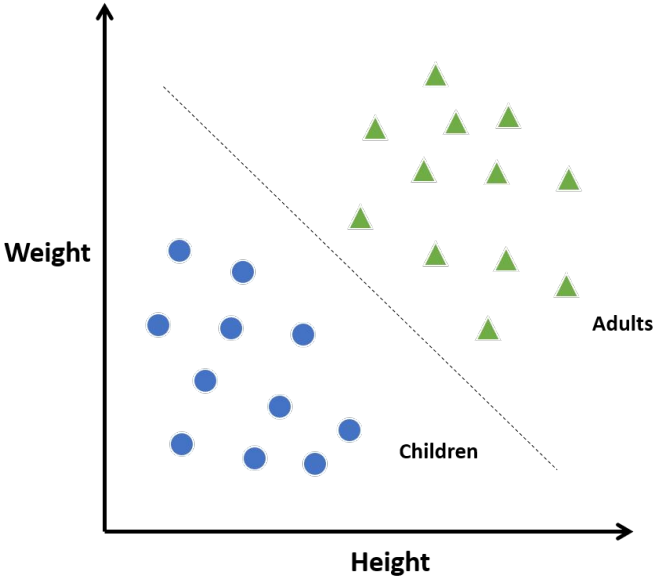
Person Identification

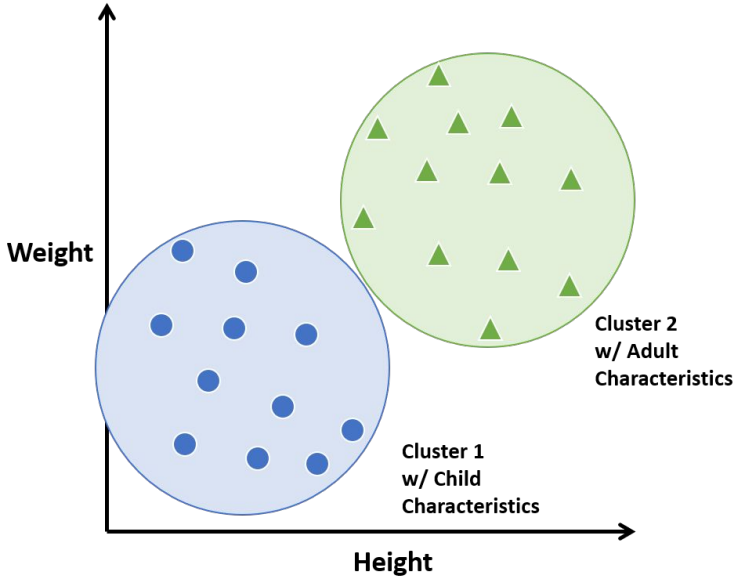


Feature Extraction

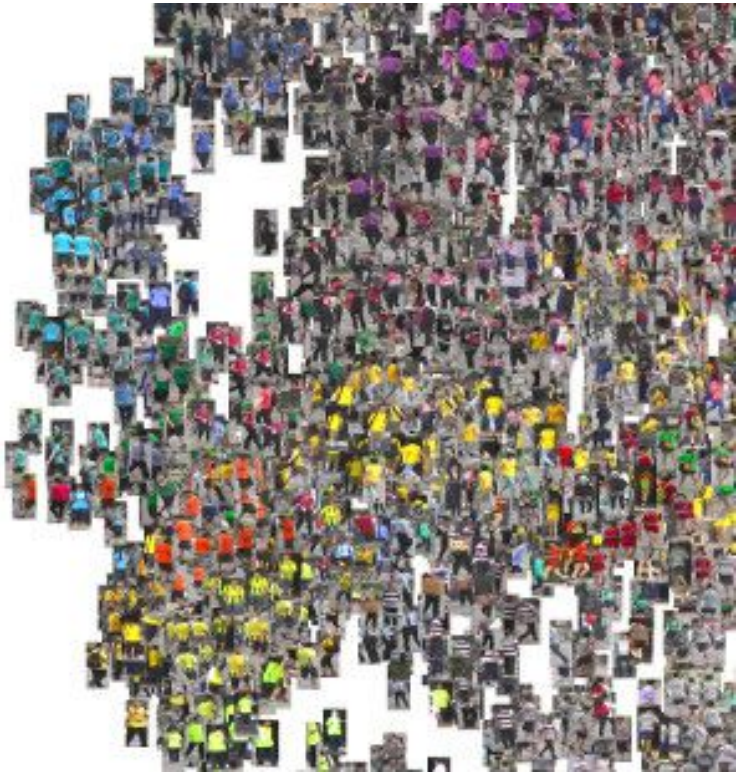


Classification vs Clustering





Classification vs Clustering



Class is important

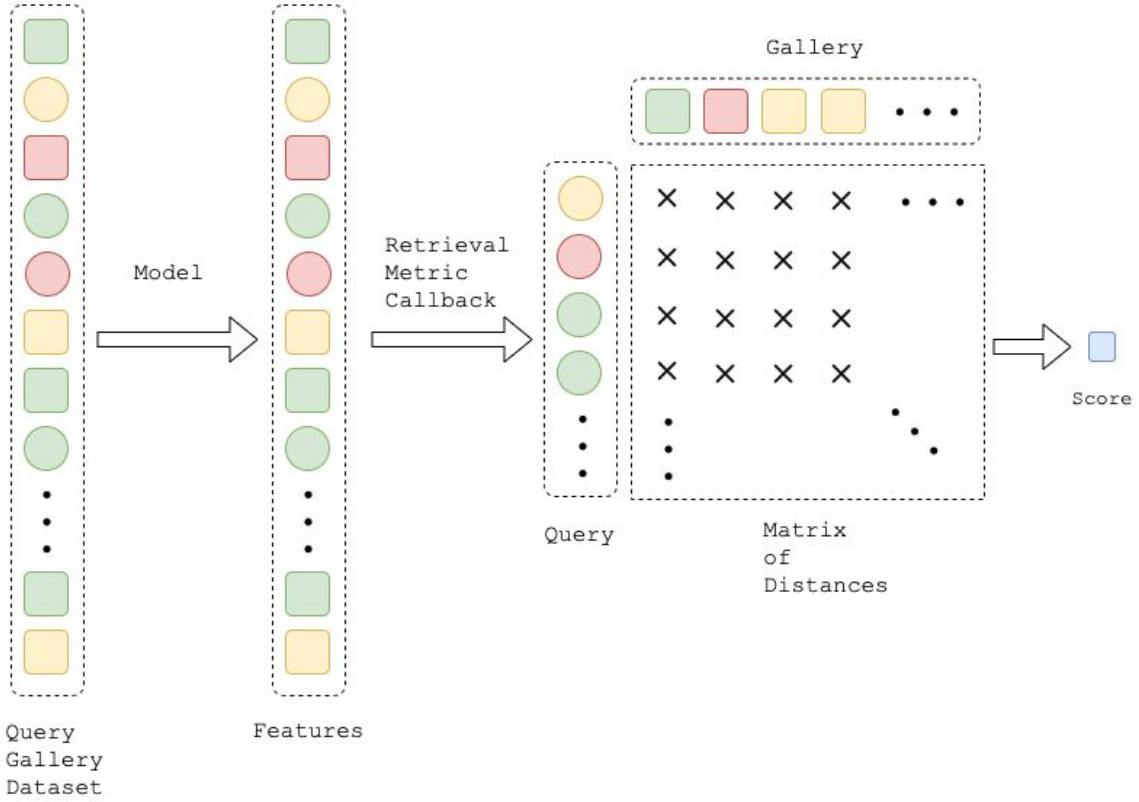
Datasets

Many classes, few objects pre class

- Omniglot
- Person ReID
- LFW

Datasets. ReID

Metric. Query Gallery

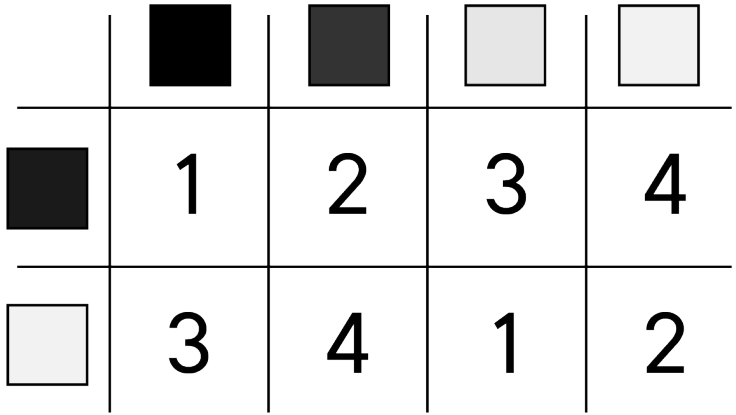


Metrics. mAP

$$\text{Recall}_k = \frac{\text{\# predicted labels that are relevant}}{\text{\# all labels}}$$

$$\text{Precision}_k = \frac{\text{\# predicted labels that are relevant}}{\text{\# all predicted labels}}$$

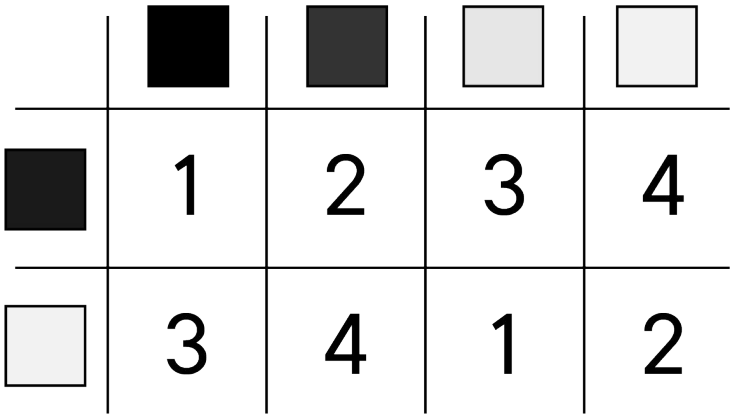
$$\text{AP}_K = \sum_{k=1}^{K} (\text{Recall}_k - \text{Recall}_{k-1}) \text{Precision}_k$$



Metrics. Cumulative Matching Characteristics

$$Acc_k = \begin{cases} 1 & \text{if top-}k\text{ ranked gallery samples contain the query identity} \\ 0 & \text{otherwise} \end{cases}$$

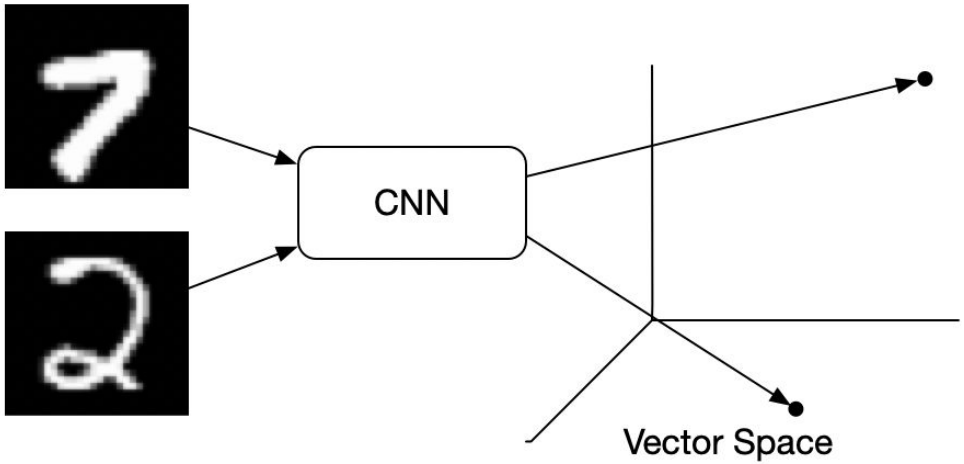
$$\text{CMC} = \frac{1}{C}\sum_{C} Acc_k$$



Metric Learning

- Contrastive Loss
- Triplet Loss
- Cross Entropy
- Siamese Network

Contrastive Loss



Contrastive Loss

$$y_{\text{pred}} = \text{dist}(x_i, x_j)$$

$$y_{\text{target}} = \begin{cases} 1, \text{ if } c_i == c_j, \\ 0, \text{otherwise} \end{cases}$$

Contrastive Loss

$$y_{\text{pred}} = \text{dist}(x_i, x_j)$$

$$y_{\text{target}} = \begin{cases} 1, \text{ if } c_i == c_j, \\ 0, \text{otherwise} \end{cases}$$

$$L = y_{\text{target}} y_{\text{pred}} - (1 - y_{\text{target}}) y_{\text{pred}}$$

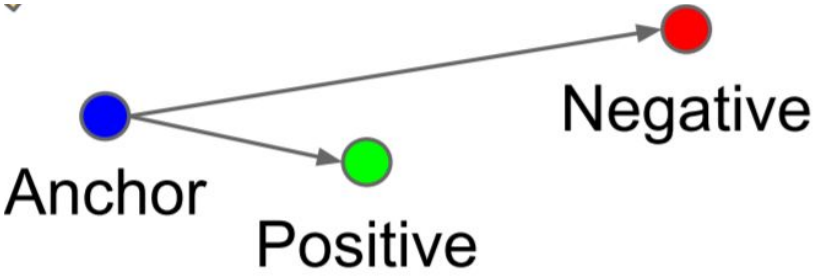
Contrastive Loss

$$y_{\text{pred}} = \text{dist}(x_i, x_j)$$

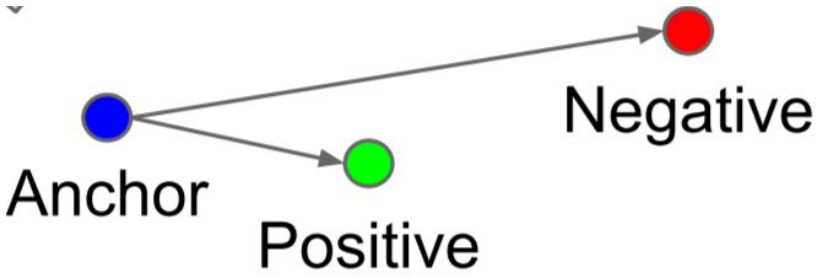
$$y_{\text{target}} = \begin{cases} 1, \text{ if } c_i == c_j, \\ 0, \text{otherwise} \end{cases}$$

$$L = y_{\text{target}} y_{\text{pred}} + (1 - y_{\text{target}}) \max(0, m - y_{\text{pred}})$$

Triplet Loss

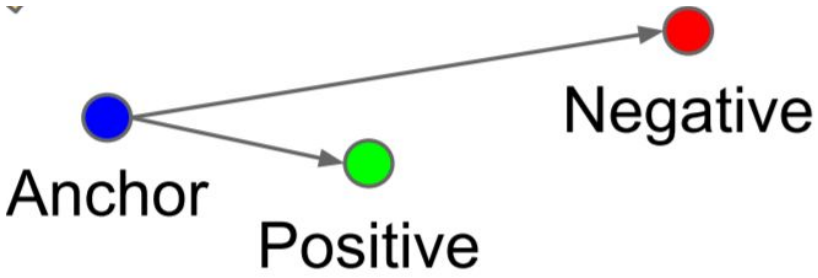


Triplet Loss



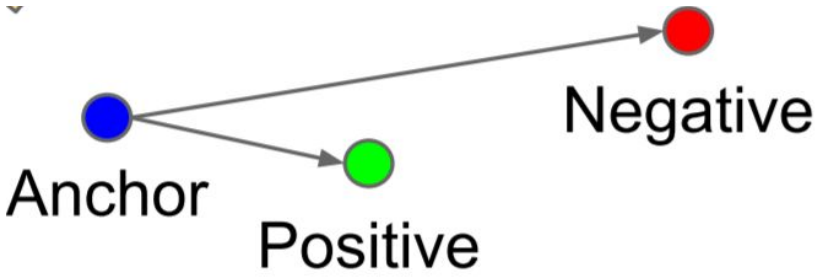
$$L = \text{dist}(x_a, x_p) - \text{dist}(x_a, x_n)$$

Triplet Loss



$$L = \text{dist}(x_a, x_p) + \max(0, m - \text{dist}(x_a, x_n))$$

Triplet Loss + Classification



$$L = L_{\text{triplet}} + L_{\text{CE}}$$

Sampling

- All Triplet
- Hard Triplet
- Hard Cluster Triplet
- Hierarchical Triplet

All Sampling

All Sampling

$$\mathcal{L}_{\mathrm{BA}}(\theta; X) = \overbrace{\sum_{i=1}^{P}\sum_{a=1}^{K}}^{\text{all anchors}} \overbrace{\sum_{\substack{p=1 \\ p\neq a}}^{K}}^{\text{all pos.}} \overbrace{\sum_{\substack{j=1 \\ j\neq i}}^{P}\sum_{n=1}^{K}}^{\text{all negatives}} \left[m + d^{i,a,p}_{j,a,n}\right]_+, \quad (6)$$

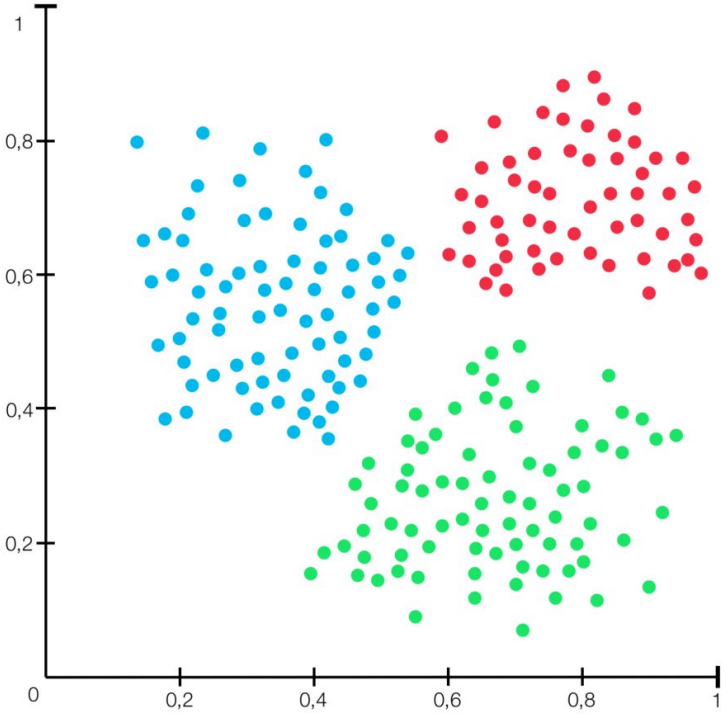
$$d^{i,a,p}_{j,a,n} = D\left(f_\theta(x^i_a), f_\theta(x^i_p)\right) - D\left(f_\theta(x^i_a), f_\theta(x^j_n)\right).$$

Hard Sampling

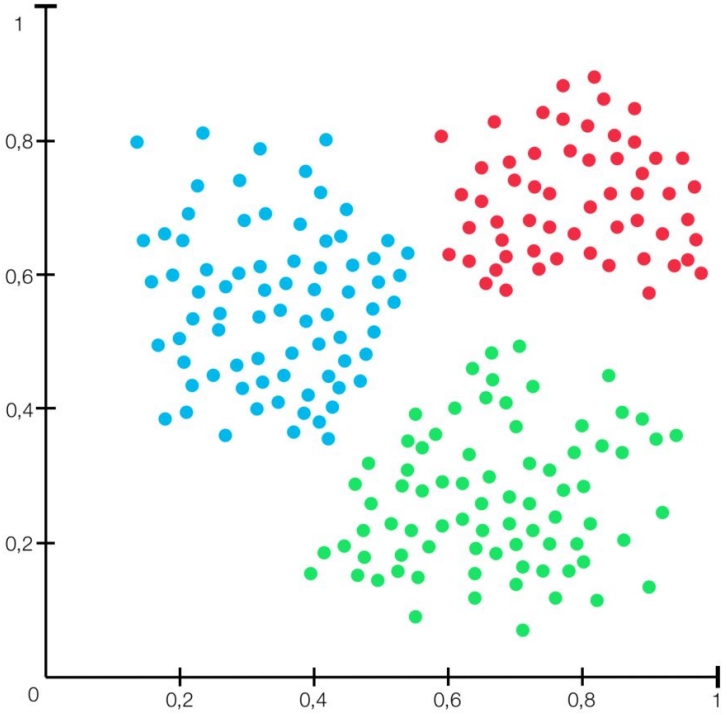
Hard Sampling

$$\mathcal{L}_{\text{BH}}(\theta; X) = \overbrace{\sum_{i=1}^{P} \sum_{a=1}^{K}}^{\text{all anchors}} \Big[m + \overbrace{\max_{p=1\ldots K} D\left(f_\theta(x_a^i), f_\theta(x_p^i)\right)}^{\text{hardest positive}} - \underbrace{\min_{\substack{j=1\ldots P \\ n=1\ldots K \\ j \neq i}} D\left(f_\theta(x_a^i), f_\theta(x_n^j)\right)}_{\text{hardest negative}} \Big]_+, \quad (5)$$

Hard Cluster Sampling



Hard Cluster Sampling

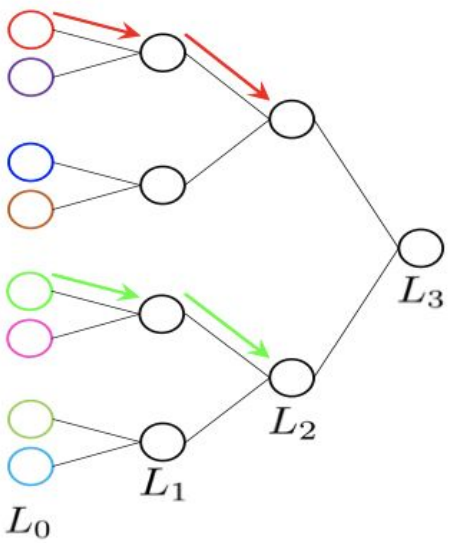


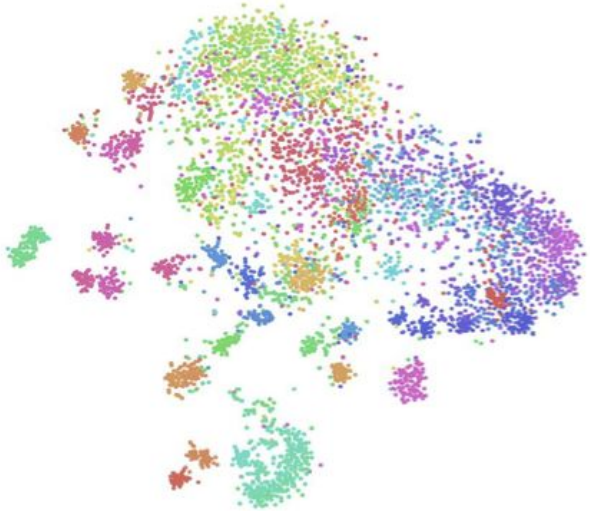
$$d_i^{inter} = \min_{\forall i_d \in P, i_d \neq i} ||f_i^m - f_{i_d}^m||_2^2$$

$$d_i^{intra} = \max_K ||f(x) - f_i^m||_2^2$$

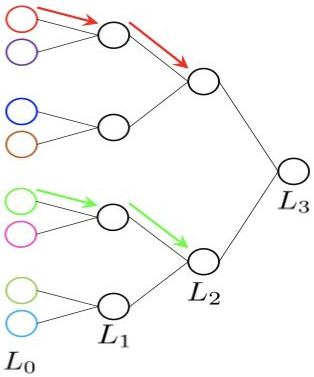
$$Lb_c = \sum_i^P \max((d_i^{intra} - d_i^{inter} + \alpha), 0)$$

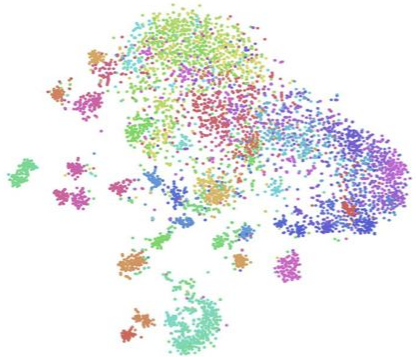
Hierarchical Sampling



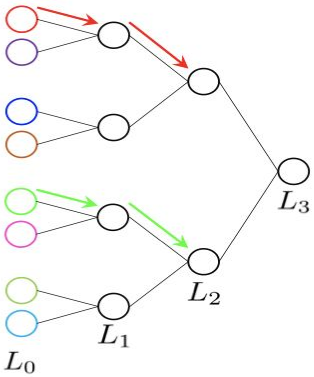


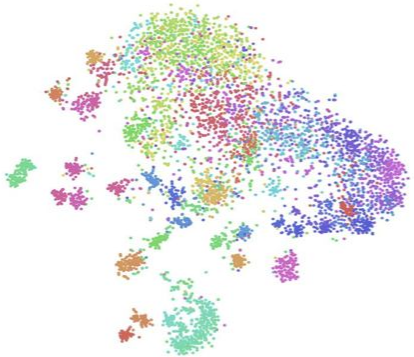
Hierarchical Sampling





Hierarchical Sampling

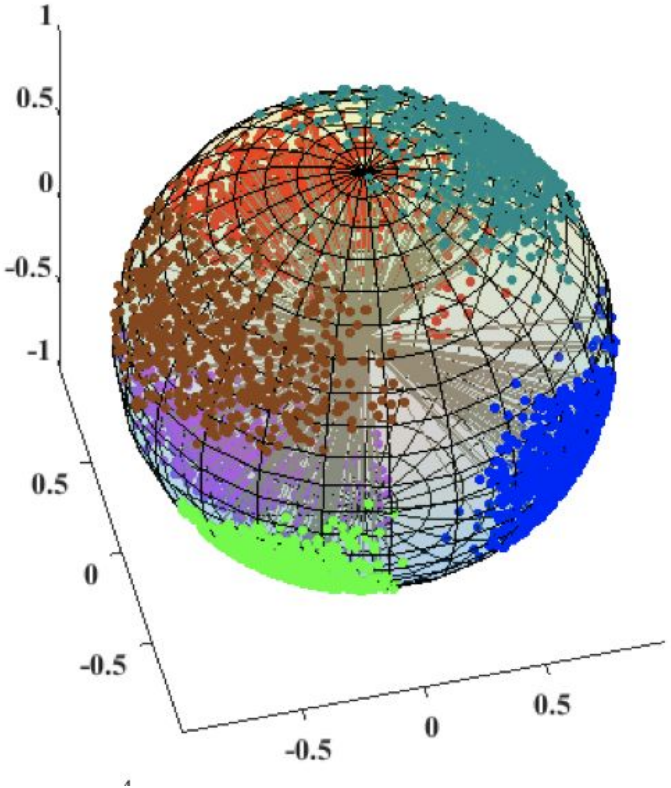




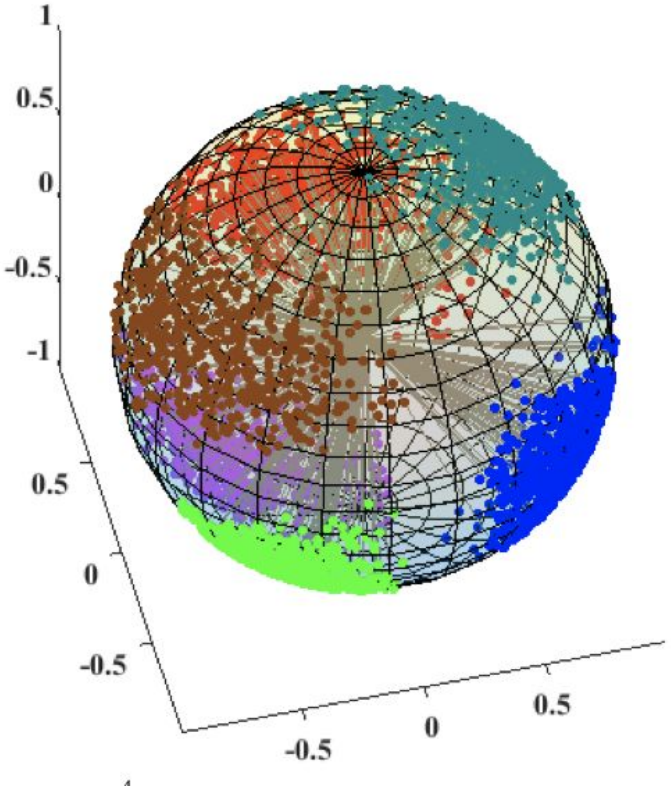
$$\mathcal{L}_{\mathcal{M}} = \frac{1}{2Z_{\mathcal{M}}} \sum_{\mathcal{T}^z \in \mathcal{T}^{\mathcal{M}}} \left[\left\| \boldsymbol{x}_a^z - \boldsymbol{x}_p^z \right\| - \left\| \boldsymbol{x}_a^z - \boldsymbol{x}_n^z \right\| + \alpha_z \right]_+ .$$

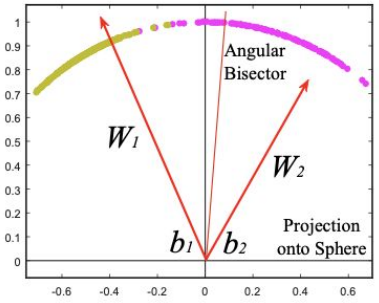
$$\alpha_z = \beta + d_{\mathcal{H}(y_a, y_n)} - s_{y_a}$$

SphereFace

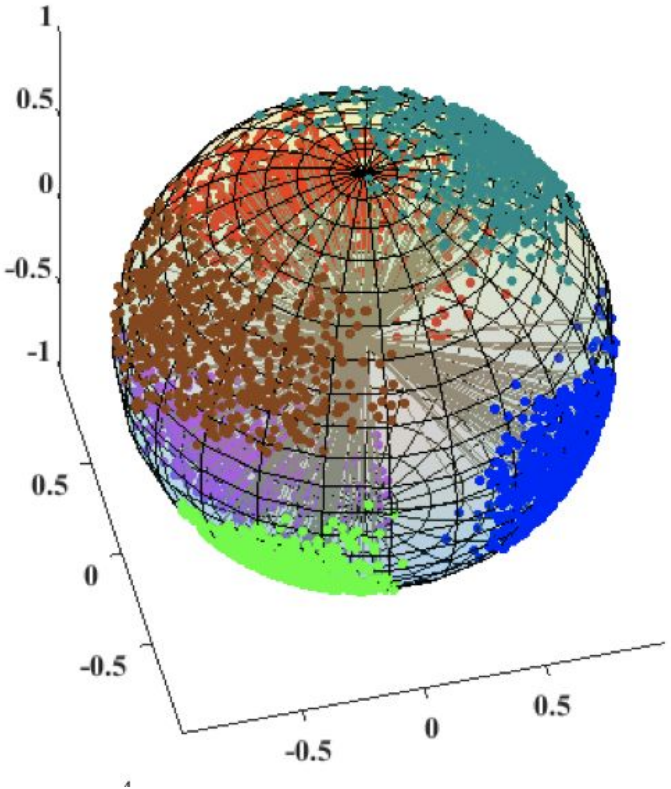


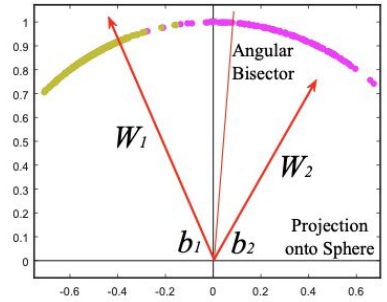
SphereFace


1
0.5
0
-0.5
-1
0.5
0
-0.5
-0.5
0
0.5


Angular
Bisector
W_1
W_2
b_1
b_2
Projection
onto Sphere

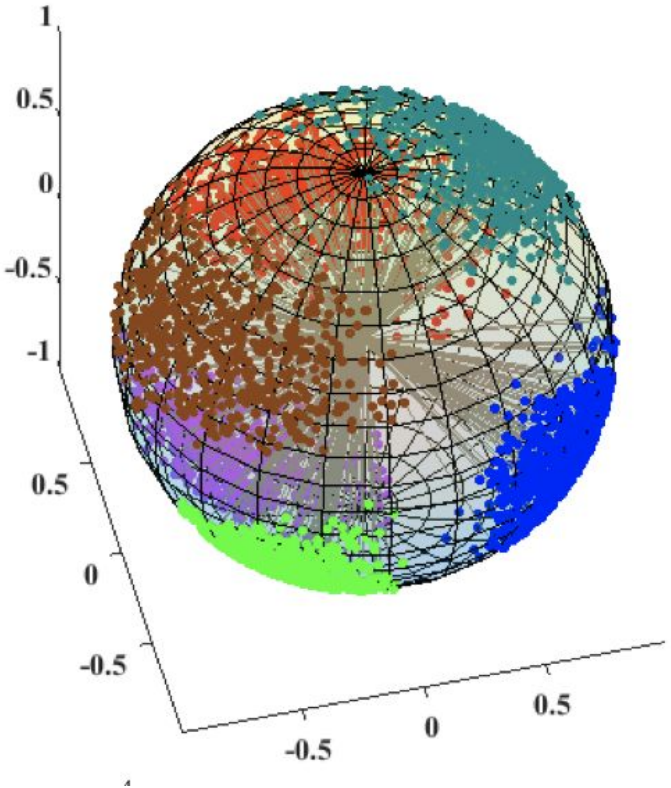
SphereFace

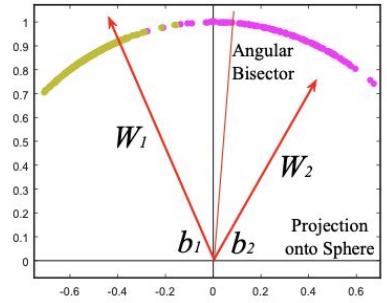




$$f_1 = W_1 x$$
$$f_2 = W_2 x$$
$$\text{dist}(f_1, f_2) = \cos(f_1, f_2)$$

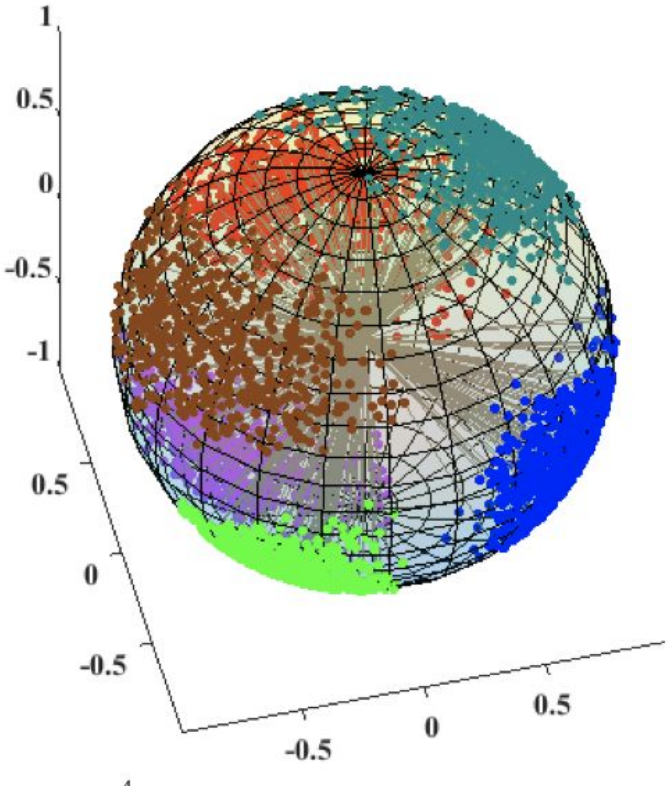
SphereFace

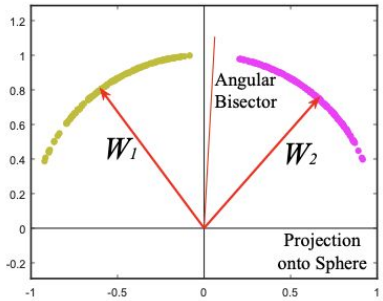




$$f_1 = W_1 x = \cos(\theta_1)||W_1||||x||$$
$$f_2 = W_2 x = \cos(\theta_2)||W_2||||x||$$
$$\text{dist}(f_1, f_2) = \cos(f_1, f_2)$$

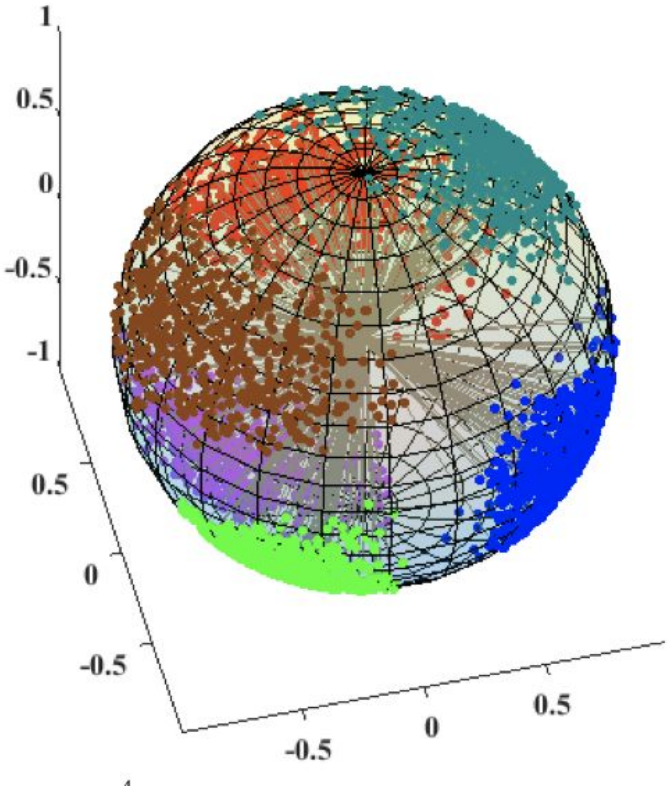
SphereFace

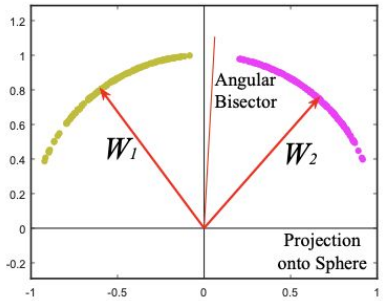




$$f_1 = W_1 x = \cos(\theta_1)||W_1||||x||$$
$$f_2 = W_2 x = \cos(\theta_2)||W_2||||x||$$
$$\text{dist}(f_1, f_2) = ||x||(\cos(\theta_1) - \cos(\theta_2))$$

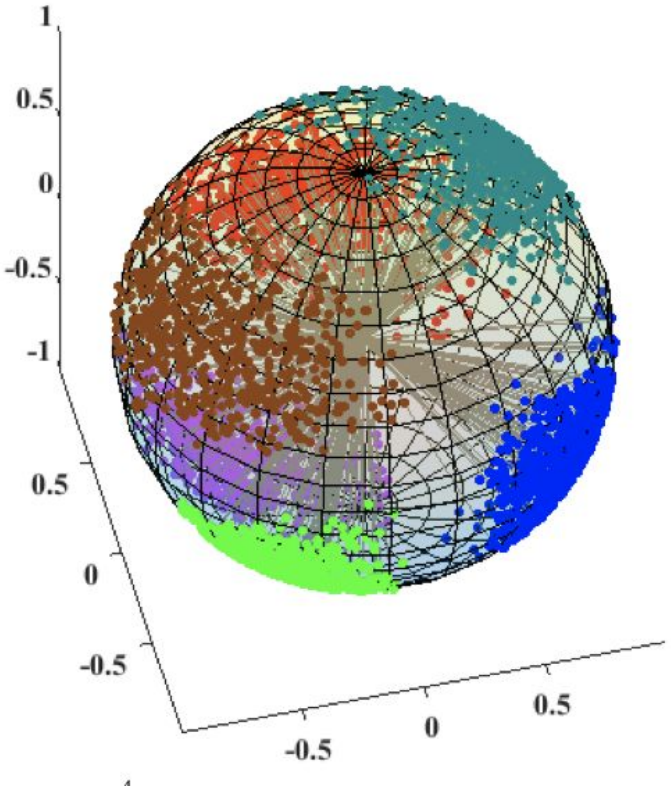
SphereFace

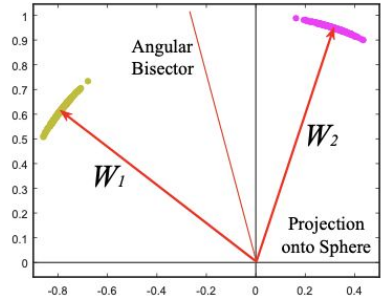




$$f_1 = W_1 x = \cos(\theta_1)||W_1||||x||$$
$$f_2 = W_2 x = \cos(\theta_2)||W_2||||x||$$

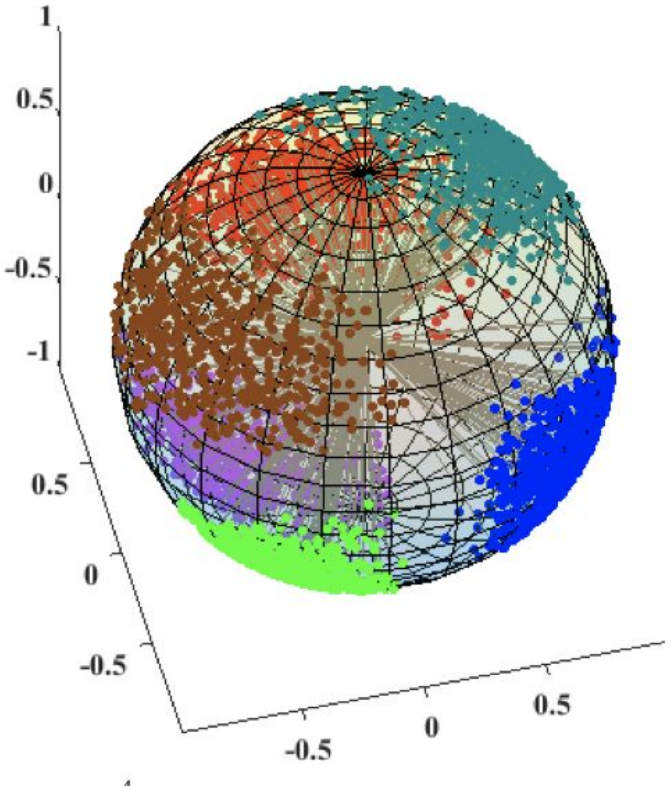
SphereFace

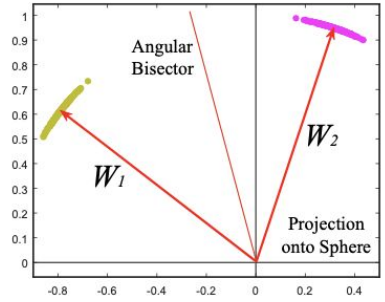




$$f_1 = W_1 x = \cos(\theta_1)||W_1||||x||$$
$$f_2 = W_2 x = \cos(\theta_2)||W_2||||x||$$

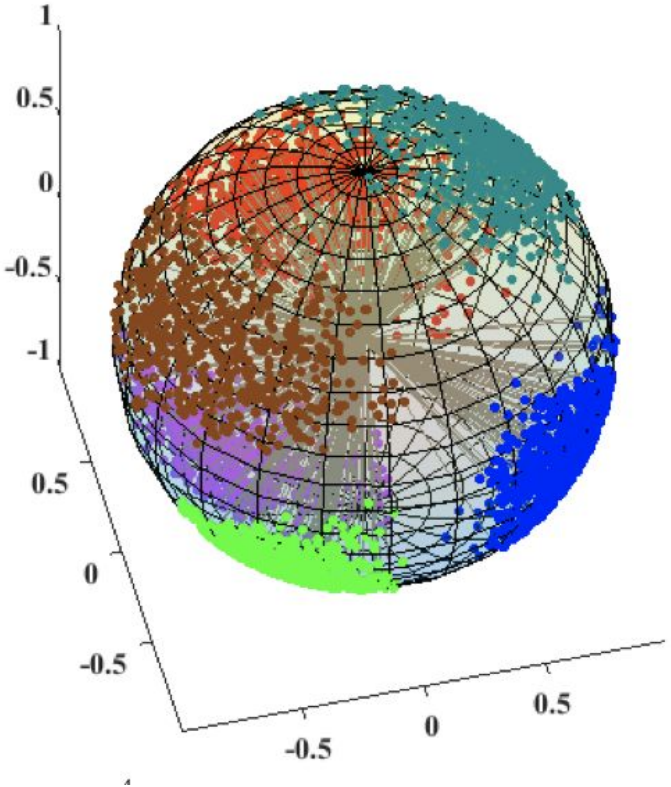
SphereFace





$$f_1 = W_1 x = \cos(\theta_1)||W_1||||x||$$
$$f_2 = W_2 x = \cos(\theta_2)||W_2||||x||$$
$$\text{dist}(f_1, f_2) = ||x||(m\cos(\theta_1) - \cos(\theta_2))$$
$$\text{dist}(f_2, f_1) = ||x||(\cos(\theta_1) - m\cos(\theta_2))$$

SphereFace



Method	Models	Data	LFW	YTF
Softmax Loss	1	WebFace	97.88	93.1
Softmax+Contrastive [26]	1	WebFace	98.78	93.5
Triplet Loss [22]	1	WebFace	98.70	93.4
L-Softmax Loss [16]	1	WebFace	99.10	94.0
Softmax+Center Loss [34]	1	WebFace	99.05	94.4
SphereFace	1	WebFace	**99.42**	**95.0**

CosFace

CosFace

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

CosFace

Method	LFW	YTF
Softmax Loss [23]	97.88	93.1
Softmax+Contrastive [30]	98.78	93.5
Triplet Loss [29]	98.70	93.4
L-Softmax Loss [24]	99.10	94.0
Softmax+Center Loss [42]	99.05	94.4
A-Softmax [23]	**99.42**	95.0
A-Softmax-NormFea	99.32	95.4
LMCL	99.33	**96.1**

ArcFace

ArcFace

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

ArcFace

Loss Functions	LFW	CFP-FP
ArcFace (0.4)	99.53	95.41
ArcFace (0.45)	99.46	95.47
ArcFace (0.5)	**99.53**	**95.56**
ArcFace (0.55)	99.41	95.32
SphereFace [18]	99.42	-
SphereFace (1.35)	99.11	94.38
CosFace [37]	99.33	-
CosFace (0.35)	99.51	95.44

Siamese Network