

# Arcface 參數定義及原 理推論

---

# 甚麼是Arcface?

---

Arcface 即 Additive Angular Margin Loss，是一種用於人臉識別任務的損失函數。

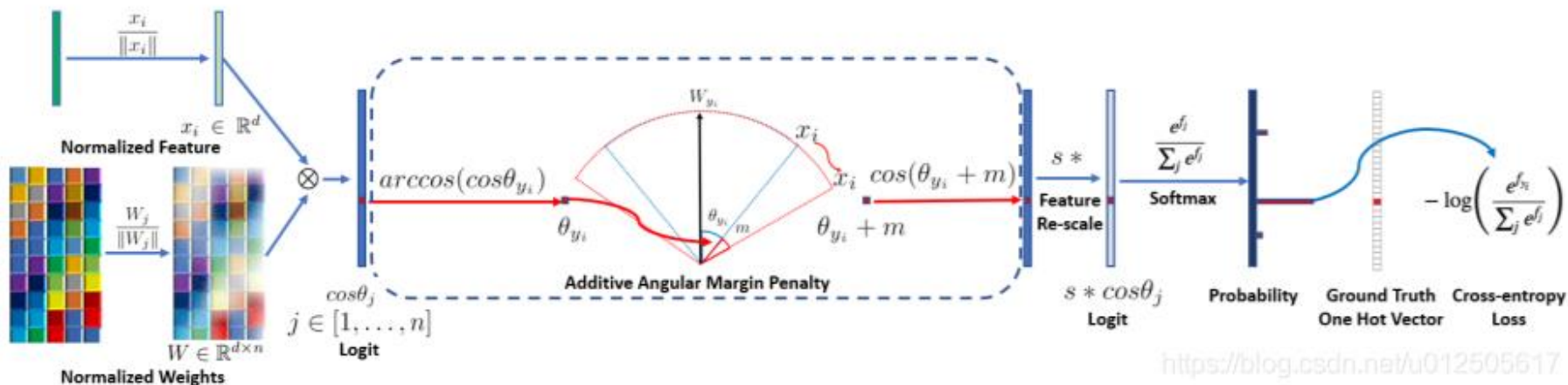
Arcface 是基於SphereFace基礎上改進了對特徵向量歸一化和加性角度間隔，提高了類間可分性同時加強類內緊度和類間差異。

優點：

- 性能高，易於編程實現，複雜性低，訓練效率高
- ArcFace直接優化geodesic distance margin(弧度)，因為歸一化超球體中的角和弧度的對應。 $L_S$

# ArcFace 處理流程

先將特徵向量L2歸一化，權重L2歸一化，他倆的夾角為 $\theta$ ，計算 $\cos(\theta_j)$ ，求反餘弦 $\arccos(\cos\theta_{y_i})$ 得到特徵 $x_i$ 與真實權值 $W_{y_i}$ 之間的夾角 $\theta_{y_i}$ ，添加角度間隔 $m$ ，再求餘弦 $\cos(\theta_{y_i}+m)$ ，將所有的log乘以特徵尺度 $s$ ，然後將log送到softmax函數得到各類別概率。再用Ground Truth和One Hot Vector一起算出交叉熵損失。即DCNN特徵和最後一個完全連接層之間的點積等於特徵和權重歸一化後的餘弦距離。我們利用arc-cosine函數來計算當前特徵和目標權重之間的角度。然後，在目標角上加上一個附加的角度間隔，用餘弦函數重新計算邏輯回歸的反向傳播過程。



# ArcFace loss介紹

---

ArcFace loss: Additive Angular Margin Loss (加性角度間隔損失函數)

- 對特徵向量和權重歸一化，對 $\theta$ 加上角度間隔 $m$ ，角度間隔比餘弦間隔在對角度的影響更加直接。幾何上有恆定的線性角度margin。
- 是直接在角度空間 $\theta$ 中最大化分類界限

Softmax loss公式↓

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

ArcFace loss

在 $x_i$ 和 $W_{j_i}$ 之間的 $\theta$ 上加上角度間隔 $m$ （注意是加在了角 $\theta$ 上），以加法的方式修改深度特徵與其相應權重之間的角度，從而同時增強了類內緊度和類間差異。

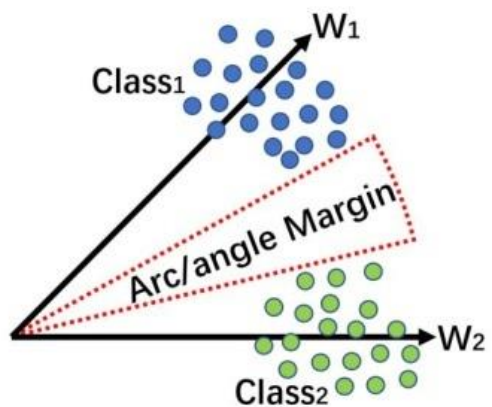
降低 $\theta$ 角度的意思就是：訓練時加上 $m$ 就會使 $\theta$ 降低

L2歸一化來修正單個權重 $\|W_j\|=1$ ，還通過L2歸一化來固定嵌入特徵 $\|x_i\|$ ，並將其重新縮放成 $s$ 。特徵和權重的歸一化步驟使預測僅取決於特徵和權重之間的角度。因此，所學的嵌入特徵分佈在半徑為 $s$ 的超球體上。

$$L_3 = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s(\cos(\theta_{y_i} + m))}}{e^{s(\cos(\theta_{y_i} + m))} + \sum_{j=1, j \neq y_i}^n e^{s \cos \theta_j}} \quad (3)$$

# 弧矩

ArcFace是在角度空間分割，ArcFace的angular margin對應著弧距(arc margin)（也叫geodesic測地距離），就是在超球面曲面上的最小距離。而不是兩特徵點之間直接相連的距離。



(a) ArcFace

# 參考資料

---

[ArcFace: Additive Angular Margin Loss for Deep Face Recognition](#)