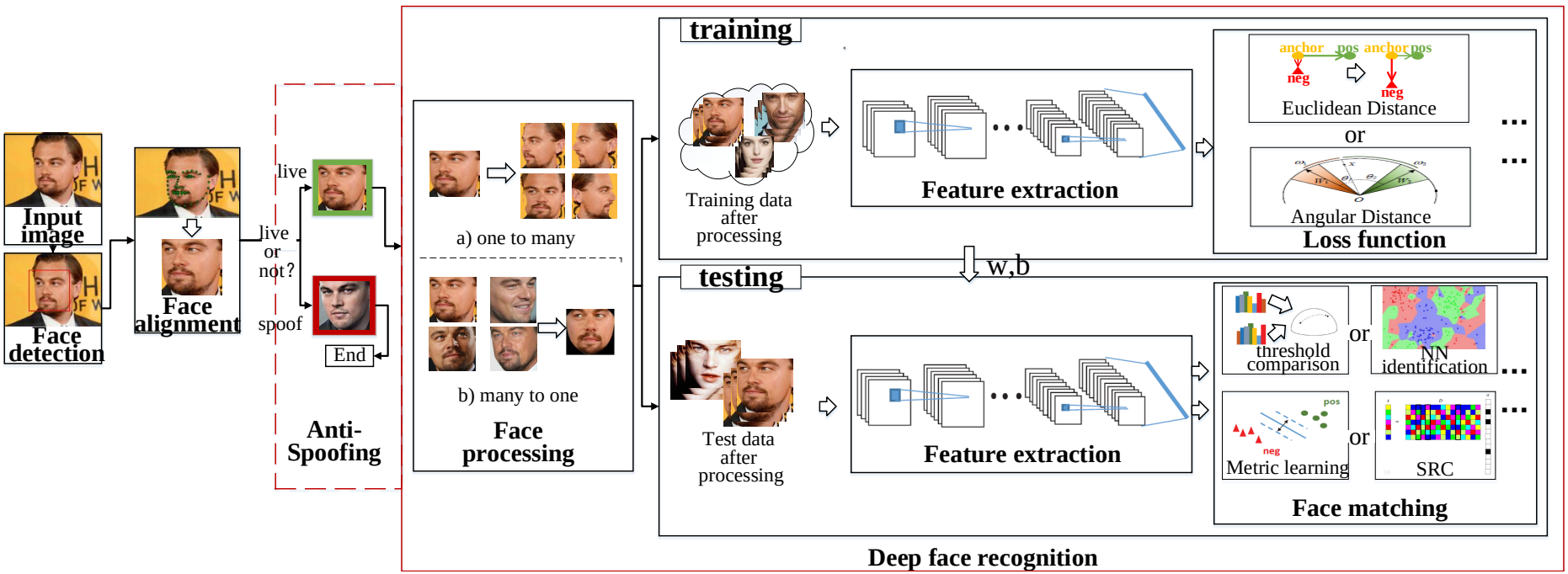
ArcFace: Additive Angular Margin Loss for Deep Face Recognition

Jiankang Deng, Jia Guo, Niannan Xue, Stefanos Zafeiriou

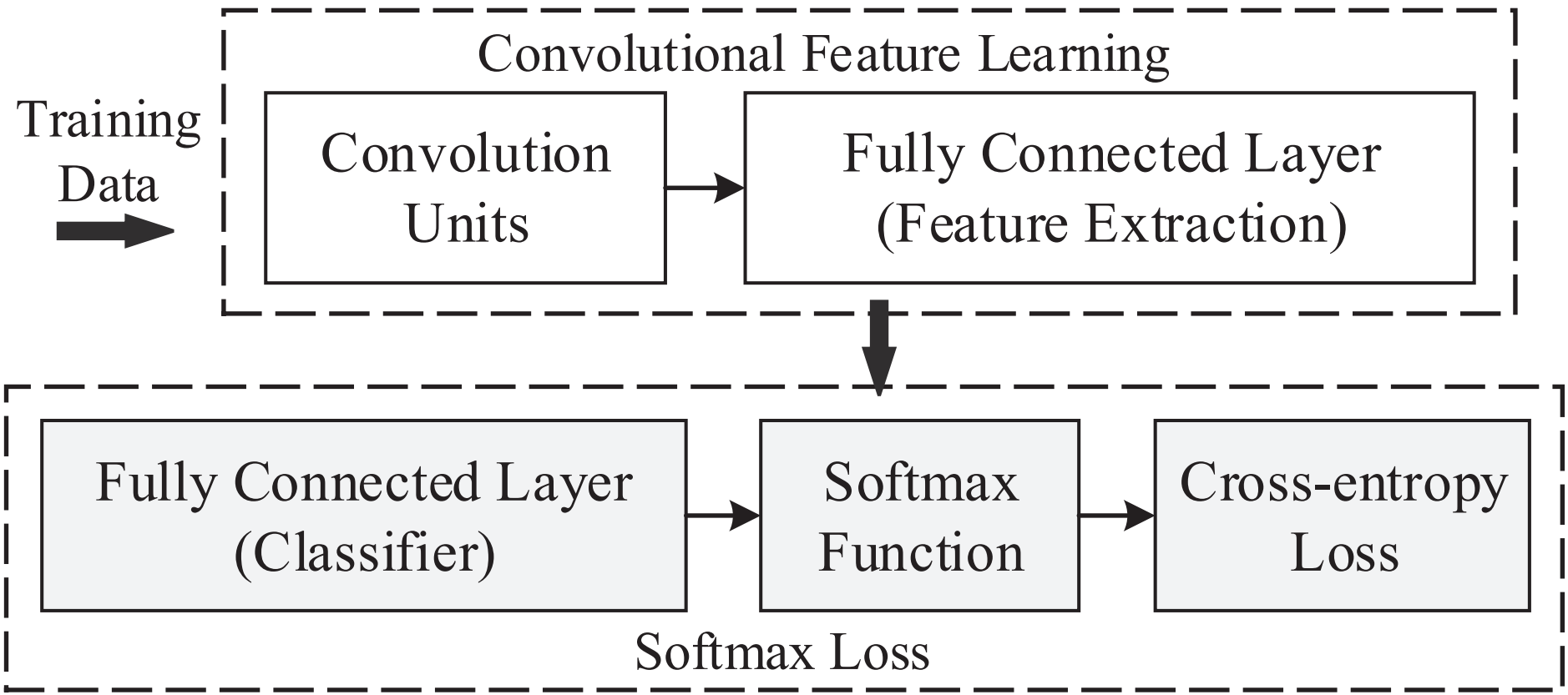
Imperial College London, InsightFace

February 9, 2019

Пайплайн системы распознавания лиц



Softmax Loss



Softmax Loss

- Softmax:

  $$Softmax(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)}$$

- Cross Entropy:

  $$CE = -\sum_i^C t_i \log\left(f_i(s)\right)$$

- Softmax Loss:

  $$L = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{\exp(x_i)}{\sum_j \exp(x_j)}$$

Softmax Loss

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

Softmax Loss

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

Softmax Loss

$$L = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{\left\|W_j\right\|\|x_i\|\cos\theta_j}}{\sum_{j=1}^{n}e^{\left\|W_j\right\|\|x_i\|\cos\theta_j}}$$

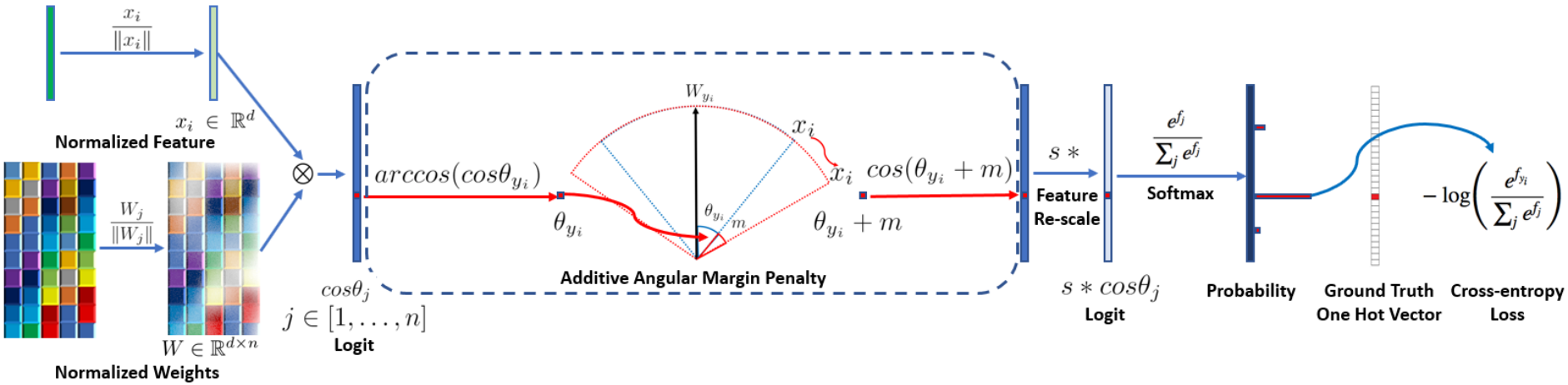
Softmax Loss

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

ArcFace Loss

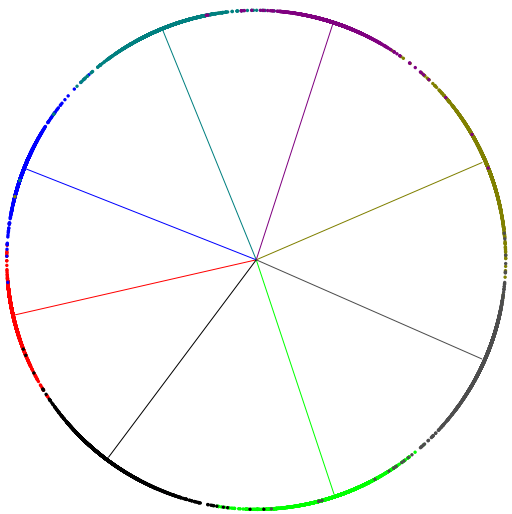
$$L = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{e^{s\left(\cos\left(\theta_{y_i}+m\right)\right)}}{e^{s\left(\cos\left(\theta_{y_i}+m\right)\right)} + \sum_{j=1, j\neq y_i}^{n} e^{s\cos\theta_j}}$$

ArcFace

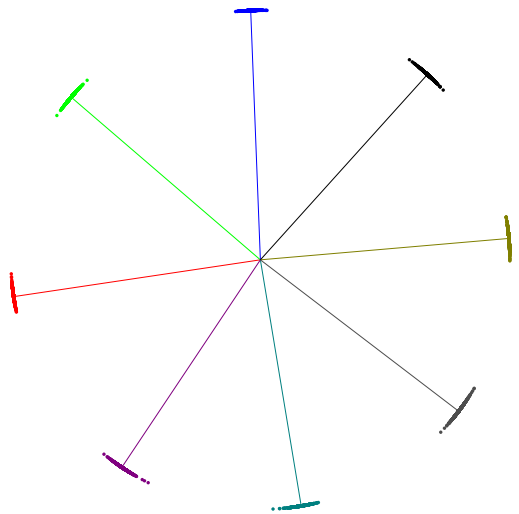


- Обнулить смещения: $b_j = 0$
- Преобразовать произведение: $W_j^T x_i = \|W_j\| \, \|x_i\| \cos\theta_j$
- Нормализовать W_j: $\|W_j\| = 1$
- Нормализовать x_i и отмасштабировать s: $\|x_i\| = s$

Softmax vs ArcFace

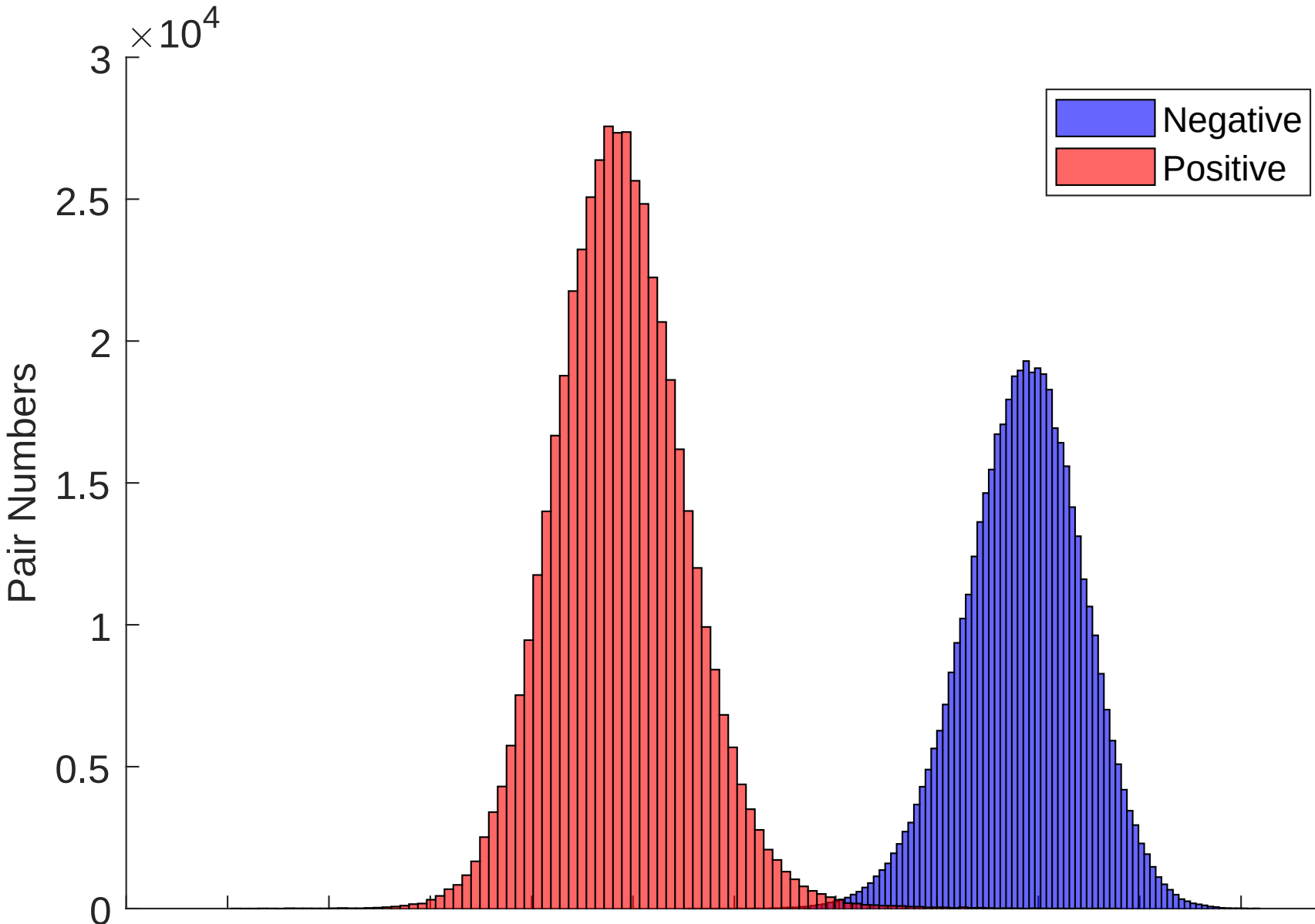


Softmax embeddings



ArcFace embeddings

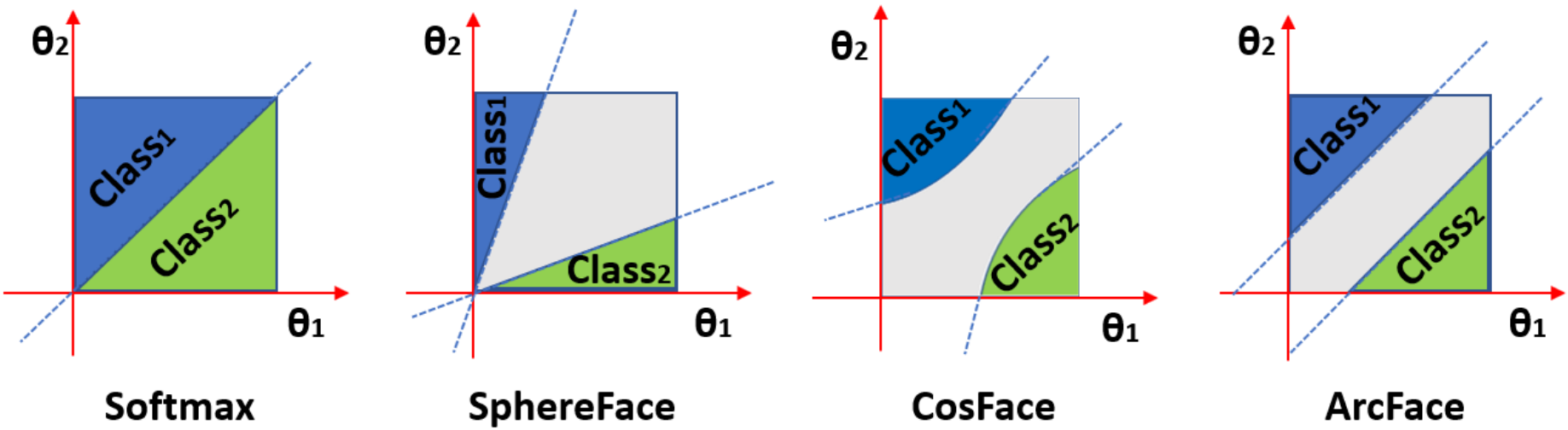
Распределение значений углов между сэмплами



ArcFace, SphereFace и CosFace вместе

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

- SphereFace: $\cos\left(m_1\theta_{y_i}\right)$
- ArcFace: $\cos\left(\theta_{y_i}+m_2\right)$
- CosFace: $\cos\left(\theta_{y_i}\right)-m_3$



Бенчмарки [MS1MV2 Dataset, ResNet-100]

Method	Images	LWF	YTF
DeepID	0.2M	99.47	93.20
Deep Face	4.4M	97.35	91.4
VGG Face	2.6M	98.95	97.30
FaceNet	200M	99.63	95.10
Center Loss	0.7M	99.28	94.9
Range Loss	5M	99.52	93.70
SphereFace	0.5M	99.42	95.00
CosFace	5M	99.73	97.60
ArcFace	5.8M	**99.83**	**98.02**

Плюсы и минусы

Pros

- State-of-the-art качество
- Простота имплементации
- Эффективность

Cons

- Не сходится в обучении с нуля модели с небольшой размерностью эмбеддинга (128d)

AirFace [Li-ArcFace Loss]

$$L = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s\left(\pi-2\left(\theta_{y_i}+m\right)\right)/\pi}}{e^{s\left(\pi-2\left(\theta_{y_i}+m\right)\right)/\pi}+\sum_{j=1,j\neq y_i}^{n}e^{s\left(\pi-2\theta_j\right)/\pi}}$$

- Монотонно уменьшается при угле от 0 до $\pi + m$, что положительно влияет на сходимость, особенно при малых размерах эмбеддинга

Ссылки

- ArcFace: https://arxiv.org/abs/1801.07698
- Deep Face Recognition: A Survey: https://arxiv.org/abs/1804.06655
- AirFace: https://arxiv.org/abs/1907.12256