Paper Review

Zhenyu Zhou

目录

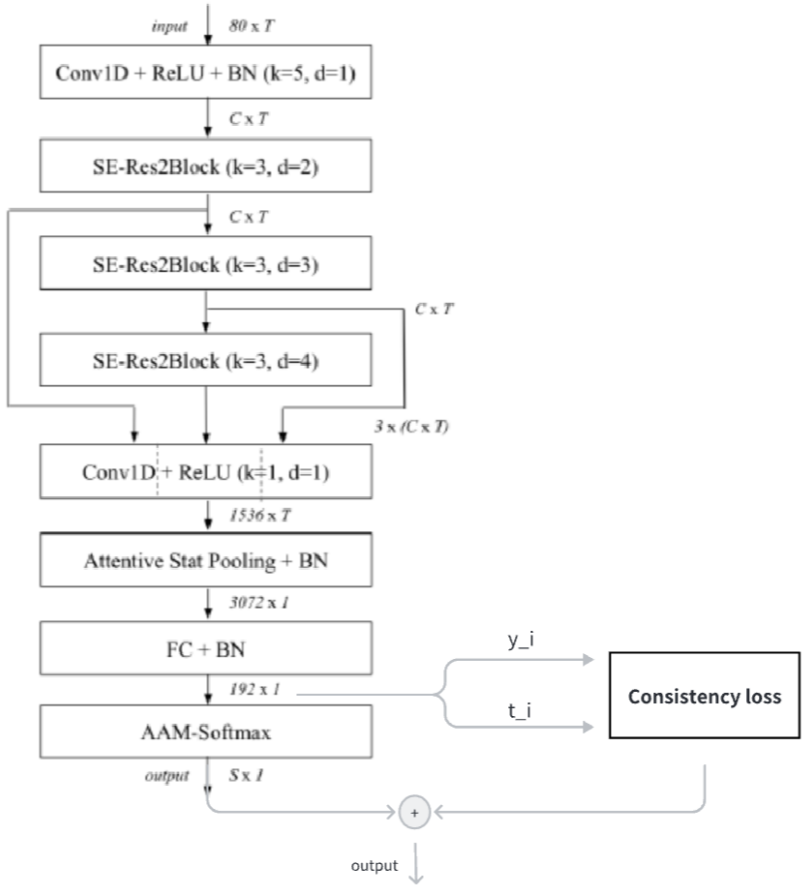
01 Consistency loss

Consistency loss

实现方法：

使用同一个人的两条语音进行配对输入网络中，分别计算两条数据的 AM-softmax loss 和 consistency loss。其中consistency loss 的作用是减小同一个说话人之间的类内方差，使同一类的数据分布更加紧凑。

consistency loss 可以选择MSE loss 或者是余弦相似度，同时consistency loss在网络中放置的位置也是可以变动的。



Consistency loss

Loss Function:

最终的损失函数：

$$L_{MSE} = \frac{1}{2}\sum_{i=1}^{n}(y_i - t_i)^2$$

$$L_{AM\text{-}softmax} = \frac{1}{n}\sum_{i=1}^{n} -\log\left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}\right)$$

$$L = L_{AM\text{-}softmax} + L_{MSE}$$

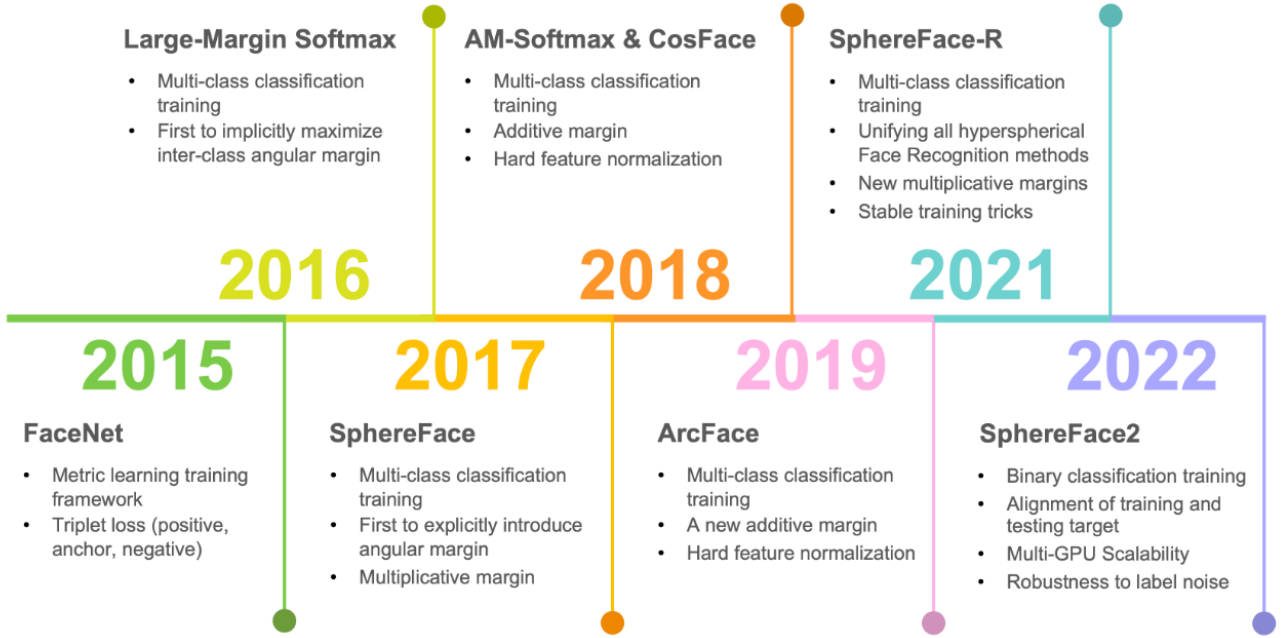
目录

01 SphereFace2

SphereFace2

研究现状：

Brief history of hyperspherical face recognition



	w/o Proxy	w/ Proxy
Triplet	FaceNet [29] VGGFace [26] Triplet [23] MTER [52]	DeepID [32] DeepFace [35] DeepID2* [31] DeepID2+* [33] SphereFace [17] NormFace [38] CosFace [37, 39] ArcFace [4] Ring Loss [51] CurricularFace [10] Circle Loss [34]
Pair	Siamese [3] DeepID2* [31] DeepID2+* [33] Cont. CNN [8]	**SphereFace2**

FaceNet(triplet loss) -> Large-Margin Softmax（L-Softmax) -> Sphereface(A-Softmax) -> Additive Margin Softmax(AM-Softmax)

SphereFace2

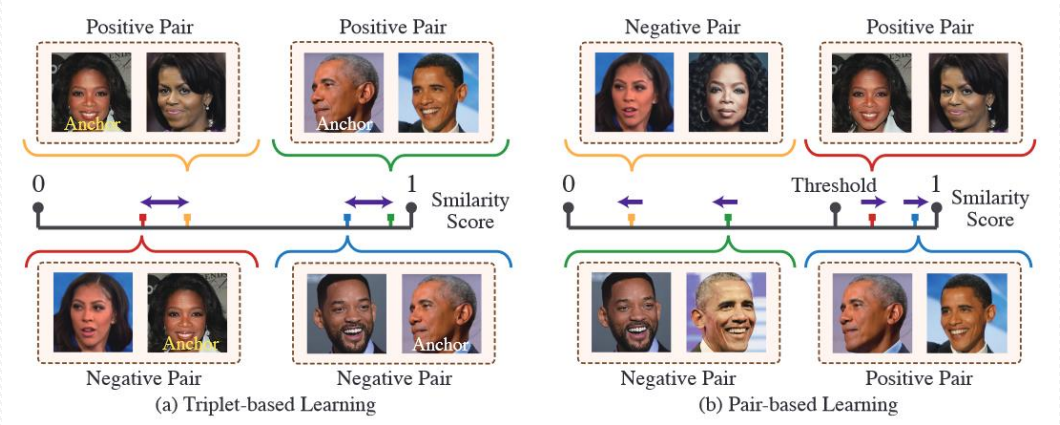
研究现状：

triplet-based learning

- 优化方法是针对每一个 anchor，最大化正对和负对之间的similarity scores距离 。
- 使用正负对和Anchor，没有通用阈值的使用。

pair-based learning

- 优化方法是使用一个通用的阈值，最大化正对的similarity scores，最小化负对的similarity scores。
- 使用正负对和通用阈值。
- 泛化效果更强，更适合开放集FR。

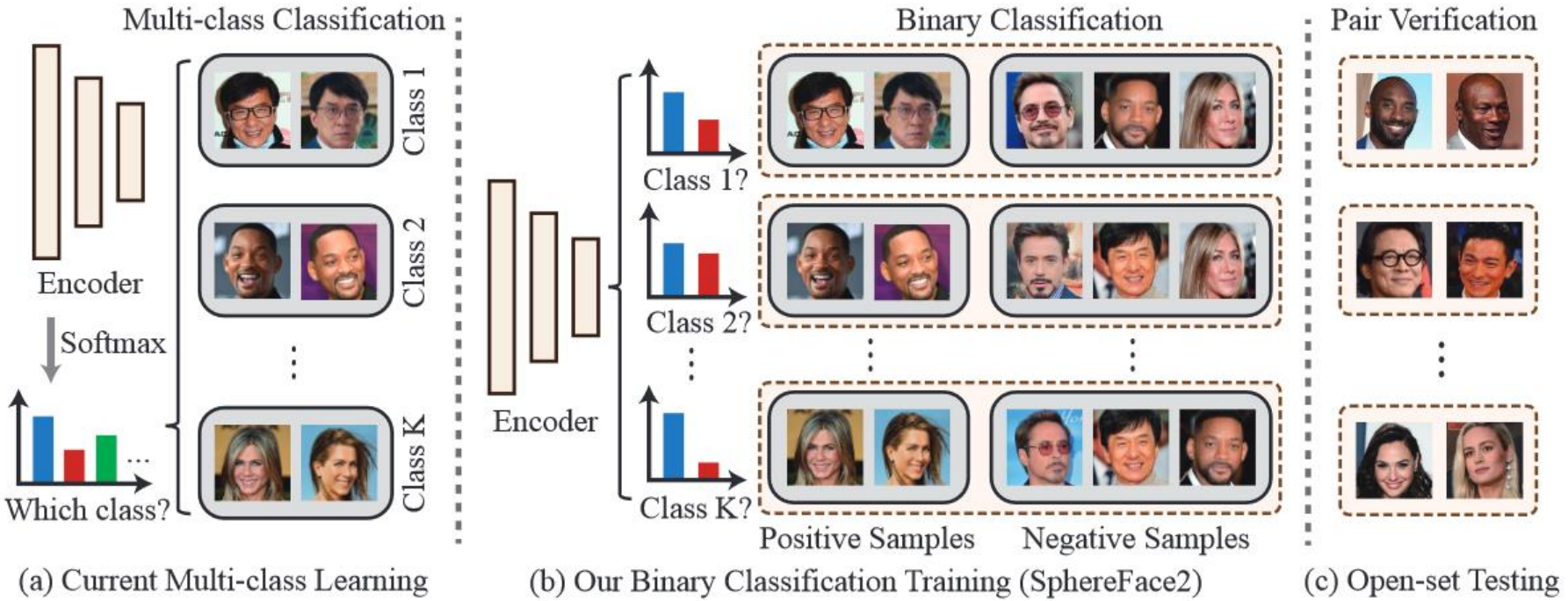


with or without proxies：是否使用proxy代表一组样本。优点：它们大大加快了训练速度并提高了数据效率（尤其是在大规模数据集上）。

SphereFace2：pair-based learning、with proxies

SphereFace2

研究现状：



给定训练集的K类，BCE构建**K个二元分类目标**，其中目标类的数据被视为正样本，其余类的数据被视为负样本。这类似于one-vs-all分类.
多分类训练的softmax cross-entropy loss 类似于拟合一个超平面，该超平面有着K个不同的区域，这导致K个区域相互之间会有竞争，这种竞争往往不利于提高超平面的分割能力，而且仅仅局限于该闭集内的K个类别。
而BCE相当于拟合K个one-vs-all的超平面，每个超平面只需分割两个区域，不仅任务较为简单，而且在每一个二分类任务之间没有任何竞争的情况下提高鉴别能力。

SphereFace2

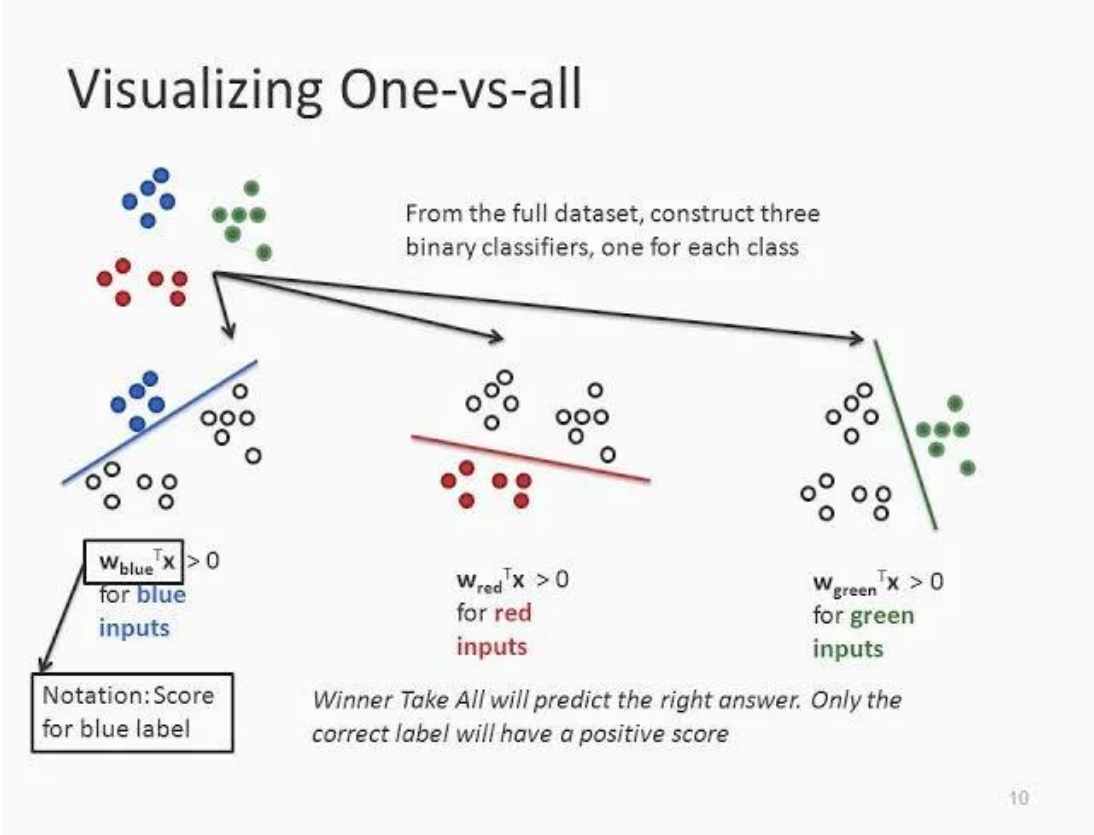
研究现状：

多分类训练(softmax loss及其变种)

- 训练和测试不匹配（闭集训练，开集测试）
- Softmax 损失引入了不同类之间的竞争，增加一个类的置信度必然会降低其他一些类的置信度，这很容易导致over-confident的预测。
- softmax 损失中的梯度更新需要涉及所有分类器、对于标签的要求更高（强监督）

二分类训练(Binary Cross Entropy Loss)

- 训练和测试匹配（开集训练，开集测试）
- SphereFace2 鼓励每个二元分类任务在没有任何竞争的情况下单独提高表示的辨别力。
- SphereFace2每个分类器的梯度仅取决于特征 x 和二元分类器本身。解决了在训练大量人脸身份时的 softmax 损失问题对标签的要求更低（弱监督）,具有更强的噪声鲁棒性



SphereFace2

实现方法：

BCE loss：

构建 K 个二分类器（ K 为类别数），第 i 个分类器的权重是 W_i，正样本为第 i 类图片，其余均为负样本

设 $t_i = W_i^T x + b_i$，sigmoid输出 $p_i = \frac{1}{1+e^{-t}}$，二分类的损失函数为：

$$L = -[y \cdot \log(p) + (1-y) \cdot \log(1-p)]$$

对于第 l 类，$y_l = 1$，$L_l = -\log(p_l) = \log(1+e^{-t}) = \log(1+\exp(-W_y^T x - b_y))$
对其余类，$y_i = 0$，$L_i = -\log(1-p_i) = \log(1+e^{t}) = \log(1+\exp(W_i^T x + b_i))$

所有 K 个Loss相加得到：

$$L_f = \log(1+\exp(-W_y^T x - b_y)) + \sum_{i \neq y} \log(1+\exp(W_i^T x + b_i))$$

将特征和权重归一化到超球面：

$$L_s = \log(1+exp(-cos(\theta_y))) + \sum_{i \neq y} \log(1+\exp(cos(\theta_i)))$$

θ_i 是 W_i 和 x 之间的角度。一般在多分类会移除 b_i，因为close-set训练的到的 b_i 无法再open-set中使用，但是Sphereface2中证明了 b 的有效性。

AM-softmax loss：

使用一个一个K分类的分类器(K为类别数)。

$$L = \frac{1}{N}\sum_i L_i = \frac{1}{N}\sum_i -\log\left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}\right)$$

$$L_{\text{modified}} = \frac{1}{N}\sum_i -\log\left(\frac{e^{\|\boldsymbol{x}_i\|\cos(\theta_{y_i,i})}}{\sum_j e^{\|\boldsymbol{x}_i\|\cos(\theta_{j,i})}}\right)$$

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

SphereFace2

实验结果展示：

PN	EH	AM	SA	LFW	AgeDB-30	CA-LFW	CP-LFW	Combined
✗	✗	✗	✗	93.60	71.67	68.40	74.07	74.49
✓	✗	✗	✗	95.37	72.90	72.93	76.90	78.03
✓	✓	✗	✗	98.60	88.10	89.98	85.23	89.65
✓	✓	✓	✗	**99.62**	92.82	93.07	90.85	93.87
✓	✓	✓	✓	99.50	**93.68**	**93.47**	**91.07**	**94.28**

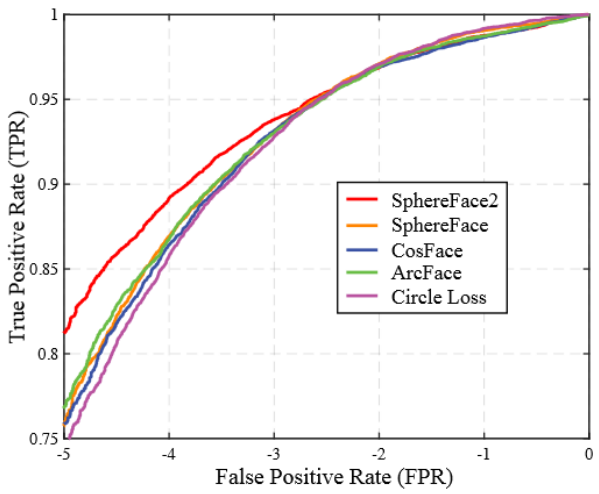
Method	1:1 Veri. TAR @ FAR			1:N Iden. TPIR @ FPIR		
	1e-5	1e-4	1e-3	rank-1	1e-2	1e-1
VGGFace2 (SENet) [2]	74.7	84.0	91.0	91.2	74.6	84.2
MN-vc [47]	-	86.2	92.7	-	-	-
Comparator Nets [46]	-	88.5	94.7	-	-	-
SphereFace	86.07	91.96	95.66	**94.83**	83.15	89.45
CosFace	85.13	90.98	95.47	94.25	80.89	88.17
ArcFace	86.12	91.60	95.72	94.71	82.21	88.94
Circle Loss	84.05	90.83	95.44	93.64	80.07	87.61
SphereFace2	**89.04**	**93.25**	**96.03**	94.70	**85.65**	**91.08**

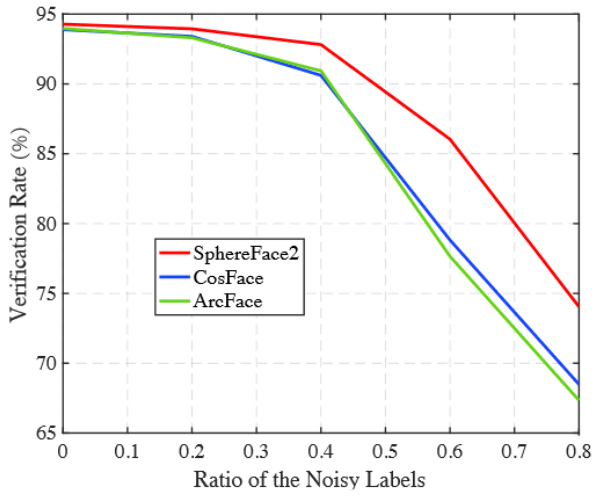
Loss Function	LFW	AgeDB-30	CA-LFW	CP-LFW	Combined
Softmax Loss	98.20	87.23	88.17	84.85	89.05
Coco Loss [20]	99.16	90.23	91.47	89.53	92.4
SphereFace [16, 17]	**99.55**	92.88	92.55	90.90	93.75
CosFace [39]	99.51	92.98	92.83	91.03	93.89
ArcFace [4]	99.47	91.97	92.47	90.85	93.97
Circle Loss [34]	99.48	92.23	92.90	**91.17**	93.78
CurricularFace [10]	99.53	92.47	92.90	90.65	93.70
SphereFace2	99.50	**93.68**	**93.47**	91.07	**94.28**

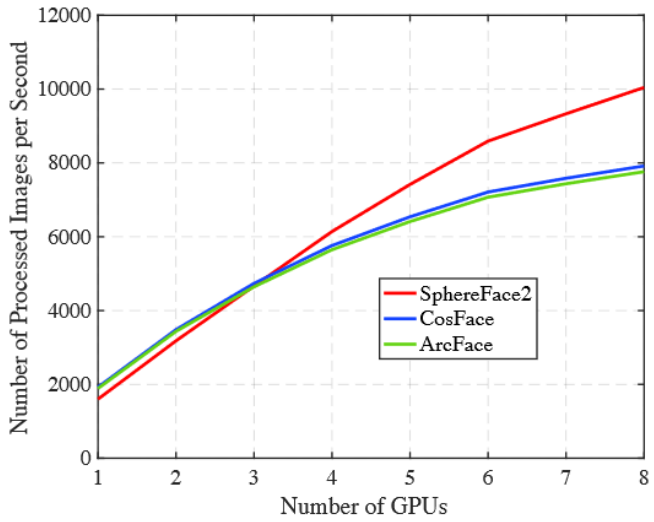
	Iden. (10^6 Distractors)		Veri. (FPR=1e-6)	
label refined	✗	✓	✗	✓
SphereFace	71.53	86.49	85.02	88.48
CosFace	71.65	86.21	85.45	88.35
ArcFace	73.65	87.88	87.77	89.88
Circle Loss	71.32	85.94	84.34	87.96
SphereFace2	**74.38**	**89.84**	**89.19**	**91.94**

SphereFace2

实验结果展示：







SphereFace2具有更加优异的识别性能，更强的噪声鲁棒性，更快的多卡运行速度