

# 基于中心偏差估计和自适应间隔的人脸识别算法

何志浩<sup>1</sup>, 王 浩<sup>1,2,3\*</sup>, 曹文明<sup>1,2,3</sup>, 何志权<sup>1,2,3</sup>

(1. 深圳大学广东省多媒体信息服务工程技术研究中心, 广东深圳 518060; 2. 广东省智能信息处理重点实验室, 广东深圳 518060;  
3. 深圳大学射频异质异构集成全国重点实验室, 广东深圳 518060)

**摘 要:** 损失函数的设计在深度人脸识别中至关重要. 常见做法是给所有类别添加固定的间隔项, 以修改类别间的决策边界, 压缩类内特征间距, 提高模型分离不同类别特征的能力. 然而, 为所有类别添加相同的间隔项可能会忽略人脸识别数据集内类别间的不一致性. 为进一步提升模型效果, 模型应依据类别的学习难易程度, 对不同类别样本特征给予不同程度的关注. 文中设计了基于类均值中心与类权重中心之间的偏差挖掘难类的方法, 称之为中心偏差估计. 本文提出的方法会根据中心偏差估计的程度, 为不同类别自适应分配不同大小的间隔项. 同时, 为解决训练前期中心偏差计算不稳定问题, 提出了动态变化的收敛参数, 调整中心偏差估计的可信度, 开展相关实验验证收敛参数的有效性. 在人脸验证基准数据集中, 本文提出的方法比基准方法的平均准确率提高了 0.26%, 达到 96.62%. 在 2 个大型人脸验证测试数据集上, 在 FPR 等于 0.01% 时, 提出方法的 TPR 分数分别提高了 0.58% 和 0.22%, 获得 88.47% 和 92.29% 的实验结果, 且多组实验结果表明提出的方法优于一般现有算法. 实现代码参见 <https://github.com/TCCof-WANG/FR-Centers-Bias>.

**关键词:** 深度人脸识别; 困难类别挖掘; 类别不平衡; 中心偏差估计; 自适应间隔

**基金项目:** 国家自然科学基金 (No.62206178); 深圳市稳定支持 A 类项目 (No.20200826104014001)

**中图分类号:** TP391.4; TN911.73 **文献标识码:** A **文章编号:** 0372-2112(2024)08-2866-12

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20230533

## Face Recognition Based on Center Bias Estimation and Adaptive Margin

HE Zhi-hao<sup>1</sup>, WANG Hao<sup>1,2,3\*</sup>, CAO Wen-ming<sup>1,2,3</sup>, HE Zhi-quan<sup>1,2,3</sup>

(1. Guangdong Multimedia Information Service Engineering Technology Research Center, Shenzhen University, Shenzhen, Guangdong 518060, China;

2. Guangdong Key Laboratory of Intelligent Information Processing, Shenzhen, Guangdong 518060, China;

3. State Key Laboratory of Radio Frequency Heterogeneous Integration, Shenzhen University, Shenzhen, Guangdong 518060, China)

**Abstract:** The design of the loss function is crucial in deep face recognition. A common practice is to add a fixed margin term to all classes to modify the decision boundary between classes, compress the distance between intra-class features, and improve the ability of the model to separate features of different classes. However, adding the same margin term for all classes may ignore the inconsistency between classes in the face recognition dataset. In order to further improve the effectiveness of the model, we argue that the model should pay different attention to the samples of different classes according to the learning difficulty of the class. In this paper, we introduce a method for hard class mining based on the bias between the center of the class mean and the center of the class weight, called center bias estimation. The method proposed in this paper adaptively assigns margin terms of different sizes to different classes according to the value of center bias estimation. At the same time, to solve the problem of unstable calculation of center bias estimation in the early stage of training, we propose an adaptively changing convergence parameter to adjust the credibility of center bias estimation and design relevant experiments to prove the effectiveness of the convergence parameters. In the face verification baseline dataset, the proposed method in this paper is improved by 0.26% on average accuracy compared with the baseline method, reaching 96.62%. In two large face verification test datasets, when FPR is equal to 0.01%, the TPR scores of our method is improved

by 0.58% and 0.22%, respectively, and the experimental results of 88.47% and 92.29% are obtained, and multiple experimental results show that our method is better than the general existing algorithms. The implementation code is published on <https://github.com/TCCofWANG/FR-Centers-Bias>.

**Key words:** deep face recognition; hard class mining; class imbalance; center bias estimation; adaptive margin

**Foundation Item(s):** National Natural Science Foundation of China (No. 62206178); Shenzhen Stability Support General Project (Type A) (No.20200826104014001)

## 1 引言

近年来,计算机视觉领域的人脸识别技术取得显著进展.在损失函数的设计和优化方面,相关研究大致可划分为以下3类:第1类是采用成对优化方式的损失函数设计.Sun等人<sup>[1]</sup>提出基于特征学习的损失函数,称为对比损失(Contrastive Loss).该损失函数的目标是在特征空间中尽量减小来自同类别的样本特征间距离,同时尽量增大不同类别的样本特征间距.Schroff等人<sup>[2]</sup>提出了三元组损失(Triplet Loss),该损失函数通过考虑2个正样本和1个负样本之间的特征距离增强模型的判别能力.第2类是将类间和类内特征距离相结合的损失函数设计.Wen等人<sup>[3]</sup>提出了中心损失(Center Loss)模型,该模型提出参数化的类中心设计,利用Softmax损失函数确保类间可分性,通过压缩类内样本与类中心间距来提升类内紧凑性.第3类是基于角度或余弦相似度的损失函数设计.Liu等人<sup>[4]</sup>提出名为L-Softmax(Large margin Softmax)损失函数,该损失函数通过对乘法角度间隔进行约束,间接提高了模型的判别能力.Wang等人<sup>[5]</sup>提出基于余弦相似度的损失函数,称为CosFace.它在Softmax损失函数的基础上增加了余弦边界项.Deng等人<sup>[6]</sup>提出名为ArcFace的损失函数,通过角度边界约束进一步提升类间特征的可分性和类内特征的紧凑性.

上述方法多是在类别一致性假设前提下进行优化的<sup>[7,8]</sup>.在人脸识别中,类别一致性体现为不同身份(类别)收集到的人脸图像应具有相似的样本数量、样本分类难易程度等.然而,在实际情况下,这些条件通常很难满足.因人脸图像数据的隐私保护等问题,收集数据往往难确保每个类别拥有相同数量的采样图像.此外,由于每个人的长相或拍摄环境的差异,不同类别的人脸图像之间也难以保证有相同的分类难易程度.因此,不同类别的分类难易程度也难保持一致.在分类难度不一致的情况下,一般损失函数仍然赋予不同类别间相同的间隔,这可能引起模型分类准确率下降.为解决上述问题,从空间几何角度上看,给学习难度更高的类别分配更大类间间隔,使决策边界偏离难类而靠近易类,借此提高模型的泛化能力.

文中提出了中心偏差估计来挖掘困难的类别,依据中心偏差估计的程度,为不同学习难度的类别自适应分配不同大小的间隔项.具体来说,设定基准间隔项

为 $m_{\text{base}}$ ,通过中心偏差估计值计算类别困难度,依据类别困难度的大小为不同类别配置不同的附加间隔项 $m_{\text{add}}$ ,模型在训练过程中会动态调整附加间隔项,缓解因人脸识别数据集类别一致性假设不成立带来的负面影响.

除此之外,近期工作表明在人脸识别损失函数中纳入图像质量的影响能使类内分布更稳定,提高模型性能<sup>[9-11]</sup>.因此,在中心偏差估计中,同样利用图像质量来辅助优化模型.另外,在模型训练初期,中心偏差估计可能不准确,导致模型对难易类关系估计错误、决策边界反向偏移.为防止这种情况发生,依据模型的收敛程度,需要动态调整中心偏差估计的可信度.因此,在依据中心偏差估计计算附加间隔项的同时,引入依据模型收敛程度设定的收敛参数 $t$ .从训练开始到模型收敛,收敛参数 $t$ 会从0逐渐增大到稳定值.在每个训练步骤中,根据当前收敛参数 $t$ 控制附加间隔项的变化,将动态间隔应用到损失函数.

本文设计目标是解决人脸识别数据固有的类别间数据不一致问题,论文的主要贡献如下:

(1)本文提出了中心偏差估计,在训练过程中动态评估不同类别的学习难度,通过实验验证该方法的可行性.

(2)本文设计了依据类别学习困难度的动态类间间隔分配方法,自适应调整难类的类内空间,获得更优决策边界.

(3)为提升模型的性能,本文提出根据模型收敛情况的自适应收敛参数,缓解训练前期因模型训练不完全导致的中心偏差估计错误问题.

(4)实验表明,本文提出的方法优于一般现有基线算法.

## 2 背景

### 2.1 困难样本挖掘

难样本挖掘(Hard Example Mining, HEM)通过在训练数据中挑选出难以区分的样本,即难样本,增加它们在训练中受关注程度,提升模型的性能.而在线难样本挖掘(Online Hard Example Mining, OHEM)<sup>[12]</sup>专注于在1个batch中挖掘高价值样本,通过训练过程中获得的经验策略得到难样本,给困难样本调整更大权重,逐渐降低简单样本的影响.在人脸识别领域也有部分相关工作基于在线难样本挖掘算法展开,例如MV-Softmax(Mis-

classified Vector guided Softmax)、CurricularFace 等<sup>[13-16]</sup>. 在 MV-Softmax 方法中, Softmax 函数首先识别出误分类样本, 将这些样本视作困难样本, 再将相应的特征用于指导模型更新, 提升模型对这些难样本的关注度; 在 CurricularFace<sup>[14]</sup>方法中, 将课程学习的方式引入到人脸识别框架中, 同样也将分类错误的样本定义为难样本, 随着训练次数增加, 模型会逐渐提升自身对困难样本的关注程度. 即使如此, 在人脸识别中, 仅依据单样本的难易程度改善训练过程是不够的, 训练过程中不同类别学习难易程度的差异对分类的准确率带来负面影响.

## 2.2 困难类别挖掘

在许多深度学习任务中, 很难保证训练数据中不同类别之间是完全均衡一致的. 例如, 类别不均衡是类别间不一致的原因之一<sup>[17-19]</sup>. 当数据集出现类别不均衡问题时, 数量较少的类别往往更难被正确分类, 这是因为模型在训练过程中接触到这些类别的样本较少, 导致模型在这些类别上的识别准确率较差. 此外, 类间样本分类难易程度的不同也会导致类别间的不一致性. 某些类别的样本可能具有明显的特征或内在结构, 使模型能够较容易地区分这些类别, 而某些类别的样本可能具有相似的特征、模糊的边界或高度重叠的属性, 导致模型难以准确区分这些类别.

困难类别挖掘(Hard Class Mining, HCM)是通过克服不同类别学习难度不一致性来提高模型性能的策略<sup>[20]</sup>, 常用于降低类别间不一致性对模型训练的负面影响. 在人脸识别中, 通过调节不同类别的类间隔达到困难类别挖掘的效果. 例如在 Fair Loss<sup>[21]</sup>中采用了 Margin-Aware 损失函数, 通过引入强化学习来优化间隔, 动态调整类别间隔. 在 AdaptiveFace<sup>[22]</sup>中, 提出用于人脸识别的自适应方法, 将间隔设定为可训练参数, 通过调整分类边界的间隔和采样策略提高识别性能. 此外, 在 PFC<sup>[23]</sup>(Partial FC)中提出了新的全连接层变体, 该方法在每次迭代中仅选择部分负类中心和正类中心来计算损失, 有效克服类别间不一致性对模型训练造成的影响.

## 3 分析与实现

### 3.1 损失函数分析

原始 Softmax 损失函数通过最大化正类的概率分离不同类别特征. 在人脸识别中,  $\mathbf{x}_i \in \mathbb{R}^D$  表示模型从人脸图像中抽取得到的低维特征,  $i$  表示  $\mathbf{x}_i$  属于第  $i$  类样本,  $\mathbf{W}_i \in \mathbb{R}^D$  表示属于类别  $i$  的权重向量,  $\mathbf{b}_i \in \mathbb{R}$  表示对应偏置量, 一般基于 Softmax 的人脸识别损失函数定义为

$$L = -\log p_i = -\log \frac{e^{\mathbf{W}_i^T \mathbf{x}_i + \mathbf{b}_i}}{\sum_{j=0}^N e^{\mathbf{W}_j^T \mathbf{x}_i + \mathbf{b}_j}} \quad (1)$$

在人脸识别的测试过程中, 通常会基于 2 张人脸图

像对应特征向量的余弦相似度判断 2 者是否属于同一身份. 为了使训练与测试时的度量标准保持一致, 相关工作中将原始的 Softmax 损失函数中的内积计算转变为余弦相似度计算. 具体做法是设置偏置量为 0, 且利用  $L_2$  范数对权重向量  $\mathbf{W}_i$  进行规范化<sup>[24]</sup>, 约束权重向量模长即令  $\|\mathbf{W}_i\| = 1$ , 同时固定  $\|\mathbf{x}_i\| = s$  来消除特征在径向方向上的变化<sup>[25]</sup>, 让模型在超球的表面空间学习可分离特征. 依据上述过程, Softmax 的 logit 可修改为

$$\begin{aligned} \mathbf{W}_i^T \mathbf{x}_i + \mathbf{b}_i &= \|\mathbf{W}_i\| \|\mathbf{x}_i\| \cos(\mathbf{W}_i, \mathbf{x}_i) \\ &= s \cdot \cos(\mathbf{W}_i, \mathbf{x}_i) = s \cdot \cos\theta_i \end{aligned} \quad (2)$$

其中,  $\cos\theta_i$  表示  $\mathbf{W}_i$  与  $\mathbf{x}_i$  之间的余弦相似度. 这样相应的损失函数可以修改为

$$L = -\log \frac{e^{s \cdot \cos\theta_i}}{e^{s \cdot \cos\theta_i} + \sum_{j=0}^{N-1} e^{s \cdot \cos\theta_j}} \quad (3)$$

这里  $L_2$  正则化的  $\mathbf{W}_i$  与重新缩放大小的  $\mathbf{x}_i$ , 使优化过程有了一定几何意义, 最小化损失相当于最小化特征  $\mathbf{x}_i$  到权重向量  $\mathbf{W}_i$  之间的夹角. 在超球空间中,  $\mathbf{W}_i$  被称为类别  $i$  的类权重中心. 增加类别间的距离与减少类别内的距离对类权重中心而言, 等价于增大类权重中心  $\mathbf{W}_i$  与其他类权重中心之间的角度并尽可能减少类内所有特征向量到  $\mathbf{W}_i$  的夹角.

经过修改的 Softmax 损失函数会将类内特征在角度域上聚集, 但无法确保类间特征的可分性. 在损失函数中加入额外的间隔项能扩张类间特征的决策边界, 使模型得到的类间特征更加分离. 类间特征的分离反过来又会压缩特征的类内距离, 使类内特征间的余弦相似度增加. 加入间隔项  $m$  的人脸识别损失函数如下:

$$L = -\log \frac{e^{s \cdot f(\cos\theta_i, m)}}{e^{s \cdot f(\cos\theta_i, m)} + \sum_{j=0, y_j \neq y_i}^{N-1} e^{s \cdot \cos\theta_j}} \quad (4)$$

在不同方法中, 间隔  $f(\cos\theta_i, m)$  的具体形式有所差异. 例如, 在 CosFace<sup>[5]</sup>中用输出项的余弦值减去间隔项来引入间隔  $f(\cos\theta_i, m) = \cos\theta_i - m$ , 相应的超球空间决策边界为  $\theta_i = \arccos(m + \cos\theta_j)$ ; 在 SphereFace<sup>[4]</sup>中, 对角度项赋予倍数间隔, 通过角度的倍数关系, 引入重叠的间隔  $f(\cos\theta_i, m) = \cos(m \cdot \theta_i)$ , 相应的超球空间决策边界为  $\theta_i = \frac{\theta_j}{m}$ ; 在 ArcFace<sup>[6]</sup>中, 对角度赋予加性间隔项, 使邻近类之间存在角度间隔  $f(\cos\theta_i, m) = \cos(\theta_i + m)$ , 相应的超球空间决策边界为  $\theta_i = \theta_j - m$ .

但这些损失函数给所有类分配的间隔是相同的, 这就要求所有类别的特征向量应具有一致性, 即每个类中的样本数量和各个样本的分类难易程度应尽可能

能相似,否则会影响算法的分类效果.对实际应用而言,上述假设过于严格.首先,现有的人脸识别数据集普遍存在着类别不平衡问题;其次,很难保证训练集中各类别是在相同的环境条件下进行的拍摄,不同样本的光照分布、人脸角度偏移等影响因素很难完全一致,因此,不同样本的分类难易程度普遍存在差异.在这样的情况下,对不同类别设定相同的间隔,会影响模型的最终效果,需要寻找1种规则,在人脸识别的训练过程中,根据分类的难易程度对不同类别赋予不同的间隔项.

### 3.2 难类挖掘策略分析

接下来,将从不同类别的样本分类难易程度和类别样本数不均衡这2个角度分析困难类别的特征,设计算法挖掘难类,依据分类难易程度给不同类别分配不同间隔,提高模型的泛化能力.

#### 3.2.1 类样本质量与难易度分布

在人脸识别中,不同类别分类的识别难易程度往往是指模型从人脸图像中抽取的特征与对应的类中心之间余弦相似度的大小.直观而言,如果原始的人脸图像样本存在人脸遮挡、大幅度偏转,导致模型难以从抽取的特征中识别出对应身份,此时特征向量与类权重向量间的余弦相似度较小,这种样本被认定为难样本,余弦相似度较大的样本被认定为简单样本.根据近几年的研究发现<sup>[9-11,26,27]</sup>,在人脸识别中,人脸图像质量也是影响识别效果的重要因素.图像分类难易程度与图像质量之间并不等同.图像质量是1个衡量图像视觉效果指标,它涉及到许多因素,包括分辨率、噪声、色彩是否偏移等.在人脸识别领域,先前的工作说明人脸图像质量与特征模长有相关性<sup>[11]</sup>,并将特征模长作为1个衡量指标来修改损失函数,有助于提升模型性能.以往的工作多专注于根据样本的特征模长给单独的样本计算不同的间隔项,希望将模长更大即质量更高的特征移动到类权重中心<sup>[9]</sup>,显然这些是针对单个样本的优化过程,但可能会破坏类内特征的模长分布规律.

本文认为不合理的难类分布如图1(a)所示,其中, $W_i$ 表示正同类权重中心, $W_j$ 表示其他类的类权重中心.浅红色区域代表模长较高的特征分布集中区域,米色区域代表模长较低的特征分布区域.红色虚线所代表的向量表示均值中心计算位置与模长的范围示意.由于带间隔项的Softmax损失函数假设了模型在超球面上优化,其优化目的是使类内所有特征与类权重中心的余弦相似度尽可能大,这个过程忽略了模长信息,因此,可能会出现类似图1(a)的情况,类别中大量模长较高的特征分布在离类权重中心较远距离,由于特征模长的大小与人脸图像质量的高低有相关性,意味着形成高质量

样本比低质量样本更难以分类的类内模长分布,这种分布是不合理的.为了解决这种不合理的分布,需要分析间隔与样本难易度之间的关系.对于在式(4)中 $f(\cos\theta_i, m)$ 是固定间隔的损失函数形式,如CosFace、ArcFace、SphereFace,从梯度的角度在这些损失函数上分析间隔项对难易样本的优化作用,梯度的变化过程如图2所示.其中,2条垂直红色虚线分别表示在角度等于 $10^\circ$ 和角度等于 $50^\circ$ 时,大间隔和小间隔之间的梯度差距.

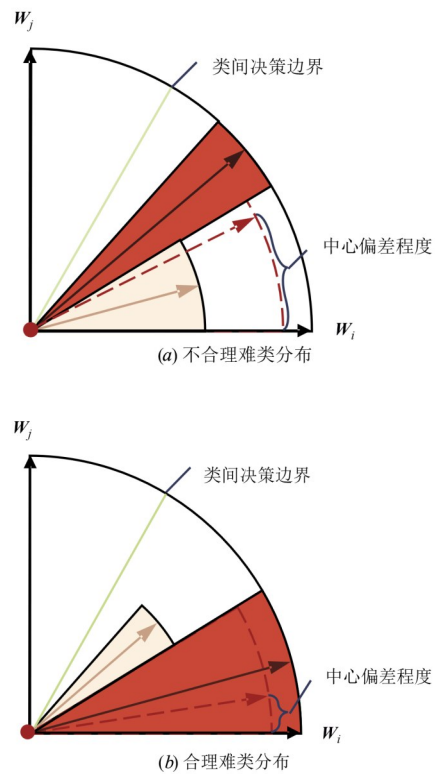


图1 2种特征模长与分布位置图

从图2可以看出,3种不同损失函数的图例中,右侧红色虚线的增长幅度大于左侧红色虚线,表明增大间隔项,难样本的梯度增加程度会高于简单样本,即在优化过程中,若梯度方向正确指向类权重中心,难样本会比简单样本更容易靠近类权重中心.如图3所示,其中,蓝色虚线表示模型在较小的间隔( $m_{base}$ )时,抽取简单样本的特征位置,橙色虚线表示困难样本,绿色虚线表示对应的难类与易类间的原始决策边界.当间隔增加时,用 $g_{hard}$ 和 $g_{easy}$ 表示计算难样本和简单样本的梯度对相应样本的缩放作用,用 $\Delta m$ 表示决策边界的偏移,绿色实线表示新的决策边界位置.

当提高某一类的间隔,模型会更加专注于优化难样本.对某一类别,如果类别中较多的图像质量高的特征分

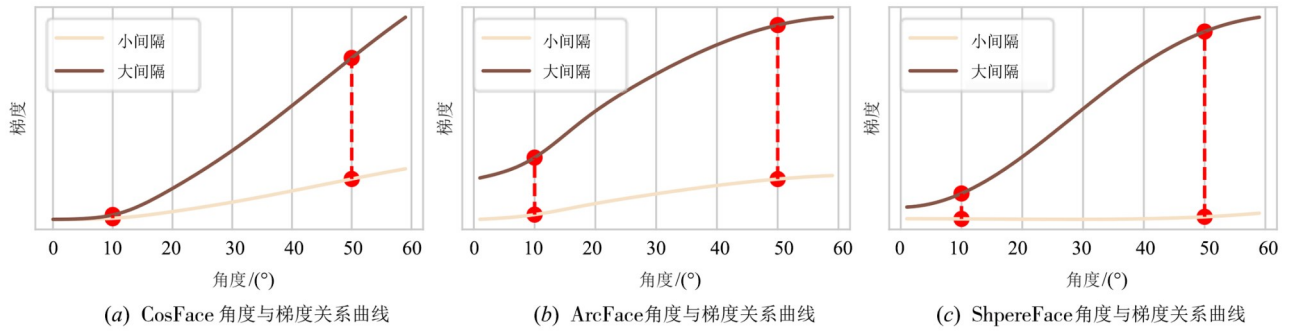


图2 CosFace、ArcFace、ShpereFace 样本特征到某类权重中心的角度与梯度模长的关系

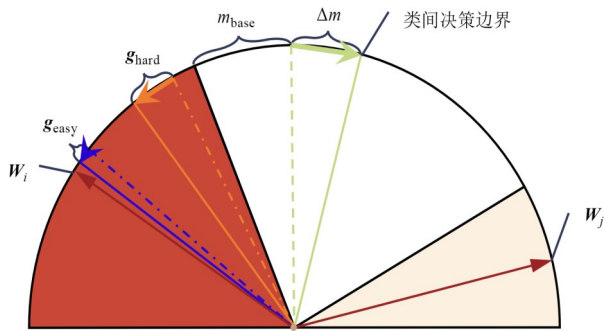


图3 类间决策边界与类内特征位置变化图

布在离类权重中心余弦相似度较小位置,应该在训练过程中增大间隔,挤压类内空间,在不改变类内模长分布规

律的同时,使模长相对较长的特征更接近类权重中心;但对于存在较多图像质量低且难分类特征的类别,应避免挤压类内空间,防止权重中心向错误的方向偏移.

### 3.2.2 类别不平衡

现有的大部分人脸训练普遍存在类别不平衡的问题,如图4所示,常用的数据集 MS-Celeb-1M、Web-Face1M、Glint 大多服从长尾分布,多数类别拥有样本数量较少.对于样本数量较少的类别,模型需要更多的迭代次数来学习样本特征<sup>[17]</sup>.此外,每次迭代中,样本数量较多的类别可能会主导损失函数的计算,使训练过程关注样本的更多类别.因此,类中样本数量的多少是衡量某类学习是否困难的标准之一.

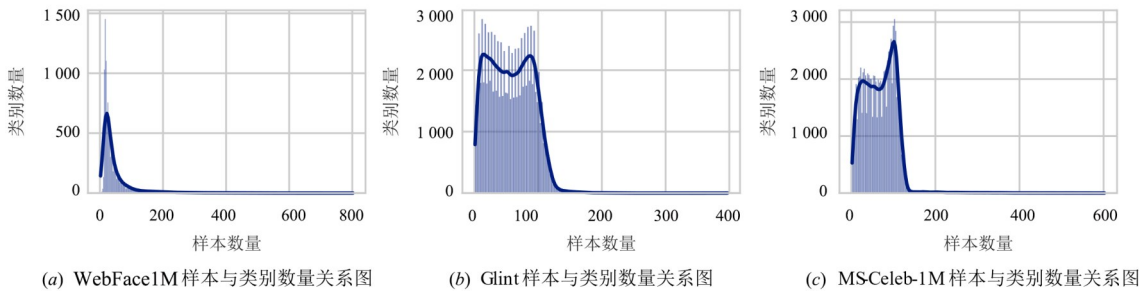


图4 不同样本数量对应常用数据集 WebFace1M、Glint 和 MS-Celeb-1M 的类别数

首先,分析样本数量较少的类对训练模型带来的影响.假设  $G_i$  表示类别  $i$  在  $D$  维的超球空间中真实的类中心.对于第  $i$  类,基于间隔的 Softmax 损失函数使用余弦相似度的度量方式会忽略特征的模长,可以将数据集中属于第  $i$  类的归一化特征  $\frac{\mathbf{x}_i}{\|\mathbf{x}_i\|}$  视作随机变量,相应的类均值中心为  $C_i = E \left[ \frac{\mathbf{x}_i}{\|\mathbf{x}_i\|} \right]$ ,依据大数定理,当某类中的样本数接近无限大时,类均值中心  $C_i$  会趋近于  $G_i$ .但在实际优化过程中,由于类内人脸图像质量的不一致,对于相同类别,模型难以将所有特征同时约束在超

球的表面空间.为了同时考虑特征的模长对类均值中心的影响,使用从模型直接输出的特征向量  $\mathbf{x}_i$  作为随机变量,类均值中心则改为  $C_i = E[\mathbf{x}_i]$ ,从而在优化过程中体现模长的作用.

在人脸识别中,Softmax 损失函数通过最大化后验概率使模型获得分离不同类别特征的能力,且通常引入权重矩阵  $W$  来辅助计算后验概率.权重矩阵的基本假设是,给定类别  $i$  的条件下,输入样本  $\mathbf{x}_i$  的概率分布是以对应类中心为期望的多元正态分布.根据式(1)中权重矩阵与特征向量间的计算可知,将权重矩阵  $W$  的任意第  $i$  行转置得到的列向量  $W_i$  作为对应的类权重中心.在最大化后验概率过程中,通过优化权重矩阵  $W$

来调整每一类的类权重中心. 在理想情况下, 当模型完全收敛时, 学到的类权重中心将使给定输入样本的后验概率最大化, 类权重中心  $\mathbf{W}_i$  会趋近  $\mathbf{G}_i$ , 类权重中心与类均值中心在  $D$  维超球空间中应收敛到相同位置.

但由于类中的样本缺失, 类均值中心会与真实类中心之间存在位置偏差, 另外由于模型的容量不足<sup>[28]</sup>, 特征不能完全地张成类别  $i$  的特征子空间, 可能导致其与错误的类权重中心进行余弦相似度计算. 因此, 实际类权重中心与类均值中心往往无法收敛到相同位置, 2者之间的位置偏差, 称为中心偏差.

图5统计了中心偏差大小与对应样本数量间的变化关系. 用预训练模型计算类均值中心, 类权重中心则由预训练模型的权重矩阵  $\mathbf{W}$  得到, 使用2者计算得到中心偏差程度, 统计不同偏差程度下, 对应类别的样本数. 从图5看出, 随着中心偏差程度的增加, 对应类别的样本数呈递减趋势. 因此, 利用类均值中心与类权重中心的中心偏差程度估计样本的缺少程度是有意义的.

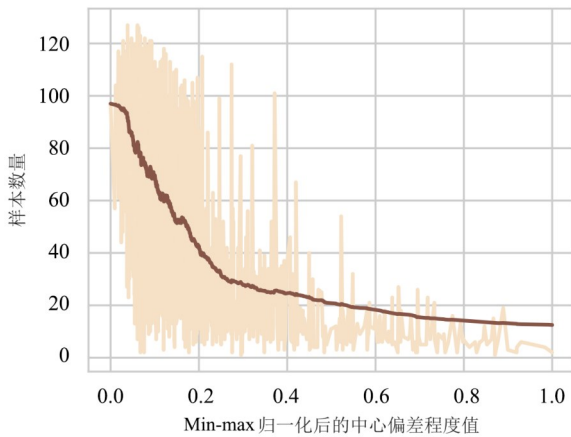


图5 中心偏差与样本数量的变化趋势

### 3.3 自适应角间隔增长损失

基于上述分析, 本文提出直接利用类均值中心  $\mathbf{C}_i$  和类权重中心  $\mathbf{W}_i$  之间的中心偏差程度来挖掘难类的策略. 类权重中心由损失函数优化得到, 分布在靠近类内特征簇位置, 而类均值中心受特征模长影响, 会相对偏向模长较高的特征簇位置. 研究利用角度衡量中心偏差的程度, 当类内分布如图1(a)所示时, 类内样本特征的分类难易度与模长分布不合理, 即高质量样本的分类难度更大, 这种不合理的分布使类均值中心与类权重中心的偏角增加, 即中心偏差程度更大. 类内分布如图1(b)时, 高质量样本更被模型关注, 因此, 类均值中心更接近类权重中心, 形成稳定的类内分布, 此时的中心偏差程度会相应减小. 依据中心偏差的大小, 定义不同类别的学习困难程度, 用  $h_i$  表示类别  $i$  的困难度, 具体如式(5):

$$h_i = 1 - \cos(\mathbf{C}_i, \mathbf{W}_i) \quad (5)$$

其中,  $\cos(\mathbf{C}_i, \mathbf{W}_i)$  用于衡量2种中心之间的偏移程度. 当计算得到的余弦相似度越低, 对应的中心偏差程度越大, 相应的类困难度就越高. 考虑到实际优化时, 类均值中心可能不会与类权重中心完全垂直, 即所有类别的最大类困难度不为1, 为易于优化, 需要对  $h_i$  进行 Min-Max 归一化,  $h_{\max}$  表示所有类困难度中的最大值,  $h_{\min}$  表示所有类困难度中的最小值, 这样归一化后的类困难度为

$$\hat{h}_i = \frac{h_i - h_{\min}}{h_{\max} - h_{\min}} \quad (6)$$

由于使用了中心偏差程度定义了不同类别的学习难易程度, 类  $f(\cos\theta_i, m)$  中  $m$  对所有类而言不再是固定值, 根据困难度  $h_i$  的计算结果对不同类别自适应分配不同大小的间隔  $m_i$ , 即表示为  $f(\cos\theta_i, m_i)$ . 具体做法是, 为所有类设定了相同的基础间隔  $m_{\text{base}}$ , 同时设定可调节的间隔增量  $m_{\text{add}}$ ,  $m_{\text{base}} + m_{\text{add}}$  决定了最大可扩张的间隔大小. 利用归一化后的类困难度  $\hat{h}_i$ 、基础间隔  $m_{\text{base}}$  和间隔增量  $m_{\text{add}}$  构成新的增长间隔  $m_i$ :

$$m_i = m_{\text{base}} + \hat{h}_i \cdot m_{\text{add}} \quad (7)$$

考虑到实际训练过程中, 中心偏差估计的准确性会根据模型的收敛程度发生改变, 如在训练前期, 由于模型分离特征的能力不足, 利用特征计算得到的类均值中心不完全准确, 类权重中心也未移动到相应位置, 导致出现中心偏差估计不稳定. 因此, 能够在不同训练阶段赋予中心偏差估计相同的信任程度. 为稳定中心偏差估计的准确性, 引入收敛参数  $t$  表示模型的收敛程度, 在训练不同阶段赋予归一化类困难度  $\hat{h}_i$  不同大小的权重, 解决估计不稳定问题. 收敛参数  $t$  具体计算由正样本到对应类别的余弦相似度计算得到, 实际模型训练过程, 采用基于 mini-batch 的形式进行学习, 使用基于指数加权移动平均的方式来动态更新收敛参数  $t$ :

$$t^k = \frac{1}{N} (1 - \alpha) \sum_{i=0}^N \cos\theta_i + \alpha \cdot t^{k-1} \quad (8)$$

其中,  $k$  表示训练中第  $k$  次迭代,  $N$  表示在第  $k$  步更新的 mini-batch 中的样本数量,  $\alpha$  表示过去输入的权重. 在训练开始时, 设置  $t=0$ , 使用指数加权移动平均来更新类均值中心  $\mathbf{C}_i$ :

$$\mathbf{C}_i^k = \frac{1}{N_i} (1 - \alpha) \sum_{j=0}^{N_i} \mathbf{x}_j + \alpha \mathbf{C}_i^{k-1} \quad (9)$$

其中,  $N_i$  表示在第  $k$  步更新的 mini-batch 中属于第  $i$  类的样本数量. 在实验中令  $\alpha = 0.99$ , 引入收敛参数  $t$  后, 式(7)也变为自适应增长的间隔计算:

$$m_i = m_{\text{base}} + t \cdot \hat{h}_i \cdot m_{\text{add}} \quad (10)$$

自适应增长间隔能够在许多固定间隔方法中使用,如CosFace、ShpereFace等,本文使用在ArcFace中添加间隔的方法进行改进,依据不同类困难度大小,为不同类别分配不一致的角间隔,完整的损失函数定义如式(11):

$$L = -\log \frac{e^{s \cdot \cos(\theta_i + m_i)}}{e^{s \cdot \cos(\theta_i + m_i)} + \sum_{j=0, j \neq y_i}^{N-1} e^{s \cdot \cos \theta_j}} \quad (11)$$

## 4 实验与结果分析

### 4.1 数据集

本实验中使用MS-Celeb-1M改进后的MS1MV2, MS1MV3和Glint的子集GlintMini作为训练集,它们分别包含约580万、510万、520万张人脸图像。其中,MS1MV2与MS-Celeb-1M数据集相比,MS1MV2数据集对图像进行了清理和增强,减少噪声和不相关的图像,提高数据质量。而MS1MV3是在MS1MV2基础上利用更优秀的特征点检测模型标记关键点精炼后的数据集。GlintMini是Glint数据集的1个子集,一定程度上代表Glint数据集的数据特征。

本文使用了6个常用的基准数据集作为验证集来进行无约束人脸验证,分别为AgeDB(Age DataBase)、LFW(Labeled Faces in the Wild)、CA-LFW(Cross-Age Labeled Faces in the Wild)、CFP-FF(Celebrities in Frontal-Profile Frontal to Frontal)、CFP-FP(Celebrities in Frontal-Profile Frontal to Profile)、CP-LFW(Cross-Pose LFW)。这些数据集提供了不同场景下的人脸图像,能够评估人脸识别算法的验证能力。此外,使用了更大的评估数据集IJB-B(IARPA Janus Benchmark-B)和IJB-C(IARPA Janus Benchmark-C)作为测试集,进一步检验模型性能。IJB-B数据集包含了来自2029个个体的118446张人脸图像,这些图像包含了各种姿态、表情、光照和年龄变化。IJB-C数据集包含了来自3531个个体的140695张人脸图像,与IJB-B数据集相比,IJB-C数据集的图像具有更广泛变化,主要用于评估算法在更具挑战性的人脸识别任务中的性能,能更好地检验人脸识别算法的鲁棒性。遵循ArcFace<sup>[6]</sup>中的测试协议,在IJB-B和IJB-C测试集上分别进行了匹配精度验证。所有数据集均按照insightface方式进行预处理,即使用五锚点的特征点检测模型对齐后裁切,得到112×112像素大小的RGB图像作为模型输入。

### 4.2 训练设置

对于输入图像,在训练时,按照文献[6]的方式进行数据增强。对112×112的输入图像的数值从0~255范围缩放到-0.5~0.5范围,在训练时按照0.5的概率随机水平翻转作为数据增强方式。

在实验中使用了2类模型,分别是ResNet<sup>[29]</sup>(Residual Networks)和MobileFaceNet<sup>[30]</sup>,ResNet使用了在ArcFace论文<sup>[6]</sup>中修改过后结构,MobileFaceNet的结构与原论文一致。所有模型的输出特征维度都是512维。模型使用SGD优化器进行训练,初始学习率设置为0.1, batch大小设置为1024,一共对训练集遍历25次,即epoch=25,在epoch为8、16、22时,学习率缩减为当前值的0.1倍。所有实验在2张RTX3090上使用半精度训练。如果没有特殊说明,以下所有实验的模型训练都采用以上的训练设置。

### 4.3 超参选择实验

中心偏差估计中的 $s$ 参照以往的经验<sup>[5,6]</sup>固定为64,设计了相关实验对超参数基础间隔 $m_{\text{base}}$ 和附加间隔 $m_{\text{add}}$ 进行超参选择。此次实验的模型结构为ResNet-18,随机选择MS1MV3数据集原有类别数的1/5所对应的样本作为训练集。在AgeDB、CA-LFW、CFP-FF、CFP-FP、CP-LFW和LFW上依照1:1的方式计算验证准确率,后将结果平均得到AVG项。对于IJB-B和IJB-C,使用原定的协议<sup>[6]</sup>进行测试,记录TPR@FPR=1%时的结果,验证结果见表1。

根据式(7)可知,对于 $\hat{h}_i = 0$ 的类别,其间隔为 $m_{\text{base}}$ ,表示最易学习的类; $\hat{h}_i = 1$ 的类别,是最难学习的类,能够被分配到最大的间隔,间隔值为 $m_{\text{base}} + m_{\text{add}}$ ,当 $m_{\text{base}} + m_{\text{add}}$ 的最大间隔值过高时可能导致模型难以收敛。表1中,设置 $m_{\text{base}}$ 分别为0.4、0.45、0.5, $m_{\text{add}}$ 为0.1或0.05,依据ArcFace<sup>[6]</sup>中 $m$ 的建议范围值,将 $m_{\text{base}} + m_{\text{add}}$ 设置在0.45~0.6区间内,通过对比得知在 $m_{\text{base}}$ 为0.4时,模型取得比其他2个参数更高的性能。本文探索了 $m_{\text{add}}$ 的可能性,将 $m_{\text{base}} + m_{\text{add}}$ 仍然限制在相同区间内,设置 $m_{\text{add}}$ 为0.2或0.15,在 $m_{\text{base}}$ 为0.4, $m_{\text{add}}$ 为0.15时,模型能够取得在组合参数中的最好性能。

表1 不同参数在相同模型、训练集下的消融实验结果 单位:%

基础间隔 $m_{\text{base}}$	附加间隔 $m_{\text{add}}$	AVG	IJB-B	IJB-C
0.4	0.1	94.15	96.01	96.91
0.4	0.05	93.69	94.58	95.47
0.45	0.05	93.43	94.05	94.86
0.45	0.1	93.67	94.63	95.27
0.5	0.05	93.47	91.94	93.33
0.5	0.1	93.30	93.95	94.67
0.4	0.2	92.92	81.44	79.86
<b>0.4</b>	<b>0.15</b>	<b>94.51</b>	<b>96.46</b>	<b>97.17</b>

注:加粗字体表示效果最好的1组参数。

### 4.4 收敛参数 $t$ 的作用

在中心偏差估计中,使用收敛参数 $t$ 来稳定前期的训练,分析了随着训练迭代过程的进行,收敛参数 $t$ 的

收敛情况,如图6所示.相同模型的情况下,模型在不同数据集上有不同的收敛过程,同样,相同数据集根据模型的特征学习能力不同,收敛参数最终的值也不一致.在容量偏小的模型中,如 ResNet-18 和 MobileFaceNet,收敛参数  $t$  最终会稳定在 0.7~0.73 区间内.

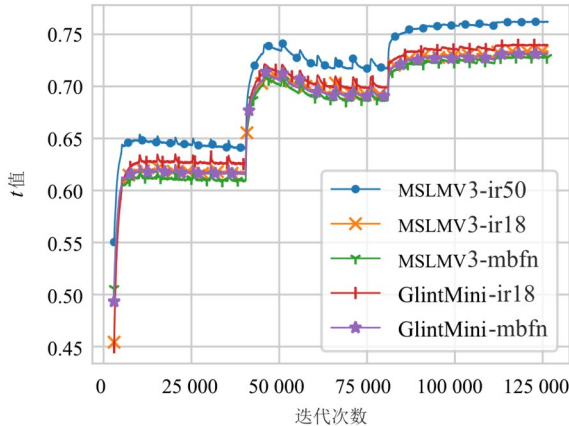


图6 不同数据集和不同模型收敛参数  $t$  的变化趋势

为了证明收敛参数  $t$  的有效性,设计了实验对比设置收敛参数  $t$  与固定  $t=1$  之间的区别,利用 2 种不同的模型,分别是 ResNet-18 和 MoBileFaceNet (MBFN),在 MS1MV3、MS1MV2、GlintMini 上进行实验,在 IJB-B 和 IJB-C 按照原定的协议<sup>[6]</sup>进行测试,记录 TPR@FPR=0.1% 时的结果,陈列在表 2 中.

表 2 显示,在 MS1MV3、MS1MV2、GlintMini 数据集中,使用自适应参数  $t$  辅助训练的模型比固定参数  $t$  为 1 的模型在 IJB-B 和 IJB-C 测试集上表现更好.且对比 ResNet-18 与 MobileFaceNet 的结果可发现,自适应参数  $t$  在 MobileFaceNet 上性能提升的幅度比在 ResNet 上更多,由于设计的轻量化,MobileFaceNet 的模型容量相对 ResNet 更低,导致训练前期,中心偏差估计的错误可能性会比容量更大的模型结构有所增加,因此,使用自适

表 2 不同训练集和不同模型验证收敛参数  $t$  的有效性 单位:%

训练集	模型	$t$ 是否固定为 1	IJB-B	IJB-C
MS1MV3	ResNet	变化	92.96	94.70
MS1MV3	ResNet	固定	92.90	94.67
MS1MV3	MBFN	变化	92.20	94.09
MS1MV3	MBFN	固定	91.99	93.94
MS1MV2	ResNet	变化	92.40	94.24
MS1MV2	ResNet	固定	92.36	94.15
MS1MV2	MBFN	变化	91.63	93.55
MS1MV2	MBFN	固定	91.32	93.47
GlintMini	ResNet	变化	92.97	94.77
GlintMini	ResNet	固定	92.94	94.75
GlintMini	MBFN	变化	92.05	94.04
GlintMini	MBFN	固定	92.04	94.02

应参数  $t$  能稳定前期训练过程,更好地提升性能.

#### 4.5 对比实验

将使用固定间隔的损失函数<sup>[5,6]</sup>、有难样本挖掘纳入的损失函数<sup>[13-15]</sup>、同样有困难类别挖掘的损失函数<sup>[22]</sup>和引入样本质量估计的损失函数<sup>[9]</sup>与本文的方法用 ResNet-18 在 MS1MV3 数据集上进行对比.为了与原本 ArcFace 公平地比较,在 MS1MV2 上也进行了模型训练对比.本文记录了在 6 个验证集 AgeDB、CA-LFW、CFP-FF、CFP-FP、CP-LFW、LFW 和 2 个测试集 IJB-B、IJB-C 上的实验结果.在 6 个验证集上,依照 1:1 的方式计算验证准确率,记录每个验证集的准确率结果,将结果平均得到 AVG 项,此外,计算了平均的  $F_1$  分数,模型测试的结果陈列在表 3. 测试集 IJB-B 和 IJB-C 则按照测试协议<sup>[6]</sup>记录 AUC 和 TPR@FPR=0.01% 时的结果,模型测试结果陈列在表 4.

表 3 中,本文的方法在多数基准人脸验证数据集上优于现有基线方法,且在 6 个基准数据集上的 AVG 和  $F_1$  得分最高,表明与其他方法相比,本文方法有更高的分类准确率,更充分解决类别间不一致性的问题.在实

表 3 验证集上不同损失函数的实验结果

单位:%

损失函数类型	训练集	AgeDB	CA-LFW	CFP-FF	CFP-FP	CP-LFW	LFW	AVG	$F_1$ 分数
ArcFace <sup>[6]</sup>	MS1MV3	97.27	95.75	99.70	95.50	90.23	99.68	96.36	96.26
CosFace <sup>[5]</sup>	MS1MV3	97.35	95.77	99.76	95.84	90.45	99.65	96.47	96.39
AdaptiveFace <sup>[22]</sup>	MS1MV3	97.10	95.63	<b>99.83</b>	95.70	90.55	99.70	96.42	96.31
MV-Softmax <sup>[13]</sup>	MS1MV3	97.47	95.67	99.81	<b>96.57</b>	90.40	<b>99.72</b>	96.61	96.40
MagFace <sup>[9]</sup>	MS1MV3	97.35	95.78	99.73	95.91	89.77	99.68	96.37	96.34
BoundaryFace <sup>[15]</sup>	MS1MV3	97.30	95.67	99.71	95.87	90.40	99.52	96.41	96.31
CurricularFace <sup>[14]</sup>	MS1MV3	97.25	95.57	99.80	94.74	90.52	99.57	96.24	96.12
本文方法	MS1MV3	<b>97.50</b>	<b>95.82</b>	99.79	96.24	<b>90.72</b>	99.63	<b>96.62</b>	<b>96.48</b>
ArcFace <sup>[6]</sup>	MS1MV2	97.40	<b>95.80</b>	<b>99.77</b>	93.69	89.75	<b>99.67</b>	96.01	95.82
本文方法	MS1MV2	<b>97.57</b>	95.70	99.74	<b>94.69</b>	<b>90.00</b>	99.57	<b>96.21</b>	<b>96.06</b>

注:加粗字体表示相同数据集下表现最好的损失函数方法.

表4 测试集上不同损失函数的实验结果

单位:%

损失函数类型	训练集	IJB-B	IJB-B	IJB-C	IJB-C
		TPR(FPR=0.01%)	AUC	TPR(FPR=0.01%)	AUC
ArcFace <sup>[6]</sup>	MS1MV3	87.89	99.49	92.07	99.58
CosFace <sup>[5]</sup>	MS1MV3	87.15	99.45	92.21	99.56
AdaptiveFace <sup>[22]</sup>	MS1MV3	87.84	99.37	91.48	99.51
MV-Softmax <sup>[13]</sup>	MS1MV3	87.52	99.49	91.39	99.59
MagFace <sup>[9]</sup>	MS1MV3	86.69	99.48	90.69	99.57
BoundaryFace <sup>[15]</sup>	MS1MV3	87.17	99.47	91.02	99.59
CurricularFace <sup>[14]</sup>	MS1MV3	88.02	99.38	92.07	99.55
本文方法	MS1MV3	<b>88.47</b>	<b>99.52</b>	<b>92.29</b>	<b>99.63</b>
ArcFace <sup>[6]</sup>	MS1MV2	85.30	<b>99.42</b>	90.82	99.51
本文方法	MS1MV2	<b>85.74</b>	<b>99.42</b>	<b>91.21</b>	<b>99.54</b>

注:加粗字体表示相同数据集下表现最好的损失函数方法。

际应用中,人脸识别应用通常对错误预测有一定容忍度,在衡量不同方法之间的实际差异时,按照 IJB-B 和 IJB-C 测试集的测试协议<sup>[6]</sup>,研究人员通常会将固定 FPR 值下的 TPR 分数作为评估指标.观察表4可知,在 TPR 分数上,实验结果表明本文方法在2个测试集上的测试性能大幅优于现有方法.在 MS1MV3 数据集上,提出的方法与 ArcFace 进行对比,能够在 IJB-B 和 IJB-C 测试集上获得 0.58% 和 0.22% 的提升,且在 ArcFace 原论文中的 MS1MV2 数据集上优于 ArcFace 方法,得到 0.44% 和 0.39% 的提升,提升幅度较为显著.此外,基于测试协议<sup>[6]</sup>进一步计算了 AUC 分数,AUC 分数能整体评估模型的预测性能.在 AUC 分数上,实验结果表明提出的方法在2个测试集上的测试性能也优于现有方法.

#### 4.6 探索性实验

对于难以学习类的挖掘,普遍认为使用某类的收敛程度会更直观,具体来说,类别的收敛程度是指类内所有特征到对应类权重中心的余弦相似度的平均.类别不均衡与类内样本难易程度分布不一致,同样体现在类别的特征与类权重中心的余弦相似度平均上,容易学习的类会拥有更大的余弦相似度均值,难类的均值则较低,显然,这样也能够被作为一种定义难类的方式,这种困难类别估计方式称为收敛余弦估计.为了区别中心偏差估计与直观的收敛余弦估计,为后续工作提供研究思路,本文设计了1组对比实验,将式(5)中的类困难度  $h_i$  替换成了以下形式:

$$h_i = 1 - r_i \quad (12)$$

其中,  $r_i$  表示类别  $i$  的余弦相似度,为了符合深度学习利用 mini-batch 的训练方式,  $r_i$  使用指数加权移动平均进行数值更新:

$$r_i^k = \frac{1}{N_i} (1 - \alpha) \sum_{i=0}^{N_i} \cos \theta_i + \alpha \cdot r_i^{k-1} \quad (13)$$

在数据集 MS1MV3、MS1MV2、GlintMini 上利用

ResNet-18 和 MoBileFaceNet (MBFN) 进行实验,模型性能使用 IJB-B 和 IJB-C 测试集并按照相应协议<sup>[6]</sup>测试,记录 TPR@FPR=0.1% 时的结果显示在表5中.

表5 实验验证中心偏差估计在不同数据集和不同模型上的有效性  
单位:%

训练集	模型	$t$ 是否替换	IJB-B	IJB-C
MS1MV3	ResNet	原始	92.96	94.70
MS1MV3	ResNet	替换	92.81	94.57
MS1MV3	MBFN	原始	92.20	94.09
MS1MV3	MBFN	替换	91.92	94.05
<b>MS1MV2</b>	<b>ResNet</b>	原始	92.40	94.24
<b>MS1MV2</b>	<b>ResNet</b>	替换	92.38	94.24
MS1MV2	MBFN	原始	91.63	93.55
MS1MV2	MBFN	替换	91.48	93.40
GlintMini	ResNet	原始	92.97	94.77
GlintMini	ResNet	替换	92.94	94.68
<b>GlintMini</b>	<b>MBFN</b>	原始	92.05	94.04
<b>GlintMini</b>	<b>MBFN</b>	替换	92.41	94.28

注:加粗字体表示值得注意的对比数据组。

分析表5中的数据,发现在 MS1MV3、MS1MV2 上, ResNet 和 MobileFaceNet 的中心偏差挖掘难类方式都能比收敛余弦估计的方式取得更好的性能表现.上节讨论了样本分类难易度与模长分布的意义,相比收敛余弦估计,中心偏差估计考虑了特征的模长,证实了引入模长评估难类是有意义的.同时,分析使用中心偏差估计得到的类困难度分布,如图7所示.类困难度的分布与数据的长尾分布规律相似,说明中心偏差估计方式有选择地对难易类进行筛选.其中,米色虚线表示类困难度的均值,棕色虚线表示类困难度的中值.值得注意的是,在 GlintMini 数据集上,MBFN 的表现与前2个数据集中呈现的规律不符合,由于 GlintMini 数据集是 Glint 数据集的子集,所抽取的图像质量普遍较高,选取

的大部分类别可能都拥有较多的样本数量,减弱了长尾分布效应,因此,使用中心偏差挖掘出难类对模型优化带来的提升不明显,特别是对于小容量模型而言. 相较于在 MS1MV3 数据集上中心偏差估计与收敛余弦估计提升的幅度,模型在 MS1MV2 上提升并不显著,研究认为存在的问题是中心偏差估计难以处理带有标签噪声的数据集,标签噪声会降低中心偏差估计的准确性,导致错误的间隔分配,这给后续基于中心偏差估计的工作提供一些研究思路.

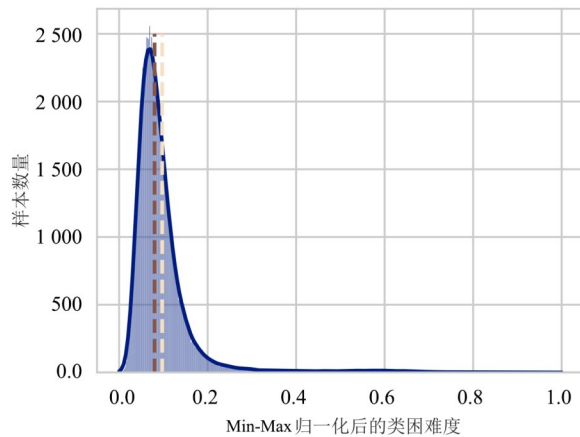


图7 类困难度的分布图

## 5 结论

文中提出基于中心偏差估计和自适应间隔的人脸识别算法. 通过设定基础间隔  $m_{base}$ , 依据类别难度估计值对不同类别自适应分配不同的附加间隔  $m_{add}$ , 实现类间间隔的动态分配. 同时,为了解决训练前期中心偏差计算不稳定的问题,本文提出动态变化的收敛参数,调整中心偏差估计的可信度. 基于多个模型在大型人脸识别数据集上的实验结果验证了本文方法的有效性. 实验结果表明本文的算法优于现有方法,解决了类别间不一致性的学习问题,但该方法无法对错误标记的样本进行正确处理. 因此,样本可能会导致错误的间隔分配. 未来工作将进一步优化算法,能够同时处理类别不一致和标签噪声问题. 此外,如何在训练过程中利用丰富的信息对间隔进行调整,以及如何利用类别中心设计更加有效的类别困难程度评估指标也是值得探索的方向.

### 参考文献

- [1] SUN Y, CHEN Y H, WANG X G, et al. Deep learning face representation by joint identification-verification[C]// Proceedings of the 27th International Conference on Neural Information Processing Systems. New York: ACM, 2014: 1988-1996.
- [2] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: A unified embedding for face recognition and clustering[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2015: 815-823.
- [3] WEN Y, ZHANG K, LI Z, et al. A discriminative feature learning approach for deep face recognition[C]//2016 14th Computer Vision - ECCV. Amsterdam: Springer, 2016, 499-515.
- [4] LIU W Y, WEN Y D, YU Z D, et al. SphereFace: Deep hypersphere embedding for face recognition[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 6738-6746.
- [5] WANG H, WANG Y T, ZHOU Z, et al. CosFace: Large margin cosine loss for deep face recognition[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 5265-5274.
- [6] DENG J K, GUO J, XUE N N, et al. ArcFace: Additive angular margin loss for deep face recognition[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 4685-4694.
- [7] RIPPEL O, PALURI M, DOLLAR P, et al. Metric learning with adaptive density discrimination[J]. (2015-11-18) [2023-06-08]. <http://arXiv.org/pdf/1511.05939>.
- [8] BOUTROS F, DAMER N, KIRCHBUCHNER F, et al. ElasticFace: Elastic margin loss for deep face recognition[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2022: 1577-1586.
- [9] MENG Q, ZHAO S C, HUANG Z D, et al. MagFace: A universal representation for face recognition and quality assessment[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 14220-14229.
- [10] TERHÖRST P, IHLEFELD M, HUBER M, et al. QMagFace: Simple and accurate quality-aware face recognition[C]//2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE, 2023: 3473-3483.
- [11] KIM M, JAIN A K, LIU X M. AdaFace: Quality adaptive

- margin for face recognition[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 18729-18738.
- [12] SHRIVASTAVA A, GUPTA A, GIRSHICK R. Training region-based object detectors with online hard example mining[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 761-769.
- [13] WANG X B, ZHANG S F, WANG S, et al. Mis-classified vector guided softmax loss for face recognition[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12241-12248.
- [14] HUANG Y G, WANG Y H, TAI Y, et al. CurricularFace: Adaptive curriculum learning loss for deep face recognition[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 5900-5909.
- [15] WU S J, GONG X. BoundaryFace: A mining framework with noise label self-correction for face recognition[C]//European Conference on Computer Vision. Cham: Springer, 2022: 91-106.
- [16] HUANG Y G, WU J X, XU X K, et al. Evaluation-oriented knowledge distillation for deep face recognition[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 18719-18728.
- [17] CHAWLA N V, BOWYER K W, HALL L O, et al. SMOTE: Synthetic minority over-sampling technique[J]. Journal of Artificial Intelligence Research, 2002, 16: 321-357.
- [18] 吴洪兴, 彭宇, 彭喜元. 适用于不平衡样本数据处理的支持向量机方法[J]. 电子学报, 2006, 34(S1): 2395-2398. WU H X, PENG Y, PENG X Y. Support vector machine method for unbalanced sample data processing[J]. Acta Electronica Sinica, 2006, 34(S1): 2395-2398. (in Chinese)
- [19] KIM B, KO Y, SEO J. Novel regularization method for the class imbalance problem[J]. Expert Systems with Applications, 2022, 188: 115974.
- [20] SUH Y, HAN B, KIM W, et al. Stochastic class-based hard example mining for deep metric learning[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 7244-7252.
- [21] LIU B Y, DENG W H, ZHONG Y Y, et al. Fair loss: Margin-aware reinforcement learning for deep face recognition[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2019: 10051-10060.
- [22] LIU H, ZHU X Y, LEI Z, et al. AdaptiveFace: Adaptive margin and sampling for face recognition[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 11939-11948.
- [23] AN X, DENG J K, GUO J, et al. Killing two birds with one stone: Efficient and robust training of face recognition CNNs by partial FC[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 4032-4041.
- [24] WANG F, XIANG X, CHENG J, et al. NormFace:  $L_2$  hypersphere embedding for face verification[C]//Proceedings of the 25th ACM international conference on Multimedia. New York: ACM, 2017: 1041-1049.
- [25] RANJAN R, CASTILLO C D, CHELLAPPA R.  $L_2$ -constrained softmax loss for discriminative face verification[J]. (2017-03-28)[2023-06-08]. <http://arXiv.org/pdf/1703.09507>.
- [26] 汪宇玲, 陈立, 黎明, 等. 基于迹变换和旋转增量调制特征的模糊人脸识别[J]. 电子学报, 2021, 49(12): 2437-2448. WANG Y L, CHEN L, LI M, et al. Rotational delta modulation feature and its application in blurry face recognition based on trace transform[J]. Acta Sinica Electronica, 2021, 49(12): 2437-2448. (in Chinese)
- [27] EBRAHIMI SAADABADI M S, RAHIMI MALAKSHAN S, ZAFARI A, et al. A quality aware sample-to-sample comparison for face recognition[C]//2023 IEEE/VF Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE, 2023: 6118-6127.
- [28] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. (2014-09-04)[2023-06-08]. <http://arXiv.org/pdf/1409.1556>.
- [29] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Confer-

ence on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.

- [30] CHEN S, LIU Y, GAO X, et al. MobileFaceNets: Efficient CNNs for accurate real-time face verification on mobile devices[C]//Chinese Conference on Biometric Recognition. Cham: Springer, 2018: 428-438.

#### 作者简介



何志浩 男, 2002年1月出生, 广东佛山人. 深圳大学电子与信息工程学院本科生, 主要研究方向为人工智能算法.  
E-mail: hezhihao2020@email.szu.edu.cn



王浩 男, 1990年1月出生, 河南鹤壁人. 深圳大学电子与信息工程学院助理教授. 主要研究方向为机器学习、深度学习、大数据挖掘.  
E-mail: haowang@szu.edu.cn



曹文明 男, 1965年11月出生, 江苏南京人. 深圳大学广东省多媒体信息服务工程中心主任, 教授, 博士生导师, 俄罗斯自然科学院外籍院士. 主要研究方向为几何代数在传感器网络、多光谱图像、目标检测与评估.  
E-mail: wmcao@szu.edu.cn



何志权 男, 1978年2月出生, 湖南邵阳人. 深圳大学电子与信息工程学院副教授. 主要研究方向为机器学习、深度学习、图像处理.  
E-mail: zhiquan@szu.edu.cn