

基于属性驱动损失函数的人脸识别算法

李燊, 苏寒松, 刘高华*, 吴慧华, 王萌

天津大学电气自动化与信息工程学院, 天津 300072

摘要 为使通过卷积神经网络学习到的人脸识别特征更容易判别,在角度距离损失函数 A-Softmax 的基础上进行改进,将人脸属性融入到训练过程中,如性别、年龄和种族。使用属性驱动损失函数,利用属性邻近性对特征映射进行正则化,实验结果表明本方法学习到更多与属性相关的鉴别特征。本文改进算法在人脸验证数据集(包括 LFW,CFP, AgeDB 和 MegaFace)上均取得不错的效果,验证了该改进算法的有效性。

关键词 机器视觉; 人脸识别; 卷积神经网络; 人脸属性

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP56.241505

Face Recognition Algorithm Based on Attribute-Driven Loss Function

Li Shen, Su Hansong, Liu Gaohua*, Wu Huihua, Wang Meng

School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

Abstract To make the face recognition features learned from the convolutional neural network easier to identify, this paper improves the angular distance loss function A-Softmax by incorporating the facial attributes, such as gender, age, and race, into the training process. By using an attribute-driven loss function and regularizing the feature mapping with attribute proximity, the experimental result shows that more attribute-related discriminating features are learned by the proposed method. The improved algorithm has achieved good results in the face verification datasets, such as LFW, CFP, AgeDB, and MegaFace, verifying the effectiveness of the improved algorithm.

Key words machine vision; face recognition; convolutional neural network; facial attributes

OCIS codes 150.1135; 100.2000; 100.3008

1 引言

近些年来,人脸识别取得巨大的发展,这得益于深度神经网络的飞速发展,尤其是卷积神经网络的发展^[1]。卷积神经网络(CNN)在人脸分析任务中发挥着重要的作用,如关键点检测、人脸识别及三维人脸重建等。目前,一些主流方法在某些基准数据集上的表现已经超过了人类,如 LFW(Labeled Faces in the Wild)数据集,这些方法通常将人脸映射到高维欧氏空间的判别特征向量上,以确定一对人脸是否属于同一类别。例如,深度度量学习(如对比损失^[2]或三元组损失^[3])通过比较成对或三个一组的面部图像来训练卷积神经网络,从而使得网络

可以从面部图像中提取出有区分力的特征。之后,不同 Softmax 损失函数的变体在 CNN 中作为监督信号,来提取有区分力的特征,在小训练集的协议下实现卓越的性能^[4]。这些方法利用 CNN 学习强区分力的特征,通过使用损失函数来增大类间距离,减小类内距离。

虽然上述两类方法都取得了显著的效果,但它们也有各自的局限性。因为构造了大量的图像对或三个一组的图像组,对比损失和三元组损失在训练过程中收敛缓慢。相比之下,Softmax 损失函数及其变体对训练数据没有这样的要求,收敛速度更快。中心损失^[5]是第一个在 Softmax 损失函数的深层特征上添加软约束以最小化类内距离的方法,显著提高了

收稿日期: 2019-04-25; 修回日期: 2019-05-27; 录用日期: 2019-06-24

基金项目: 广州市科技计划(201802020008)

* E-mail: suppig@126.com

Softmax 损失的性能。随后,角度 Softmax 损失^[6]对超球面施加判别约束,进一步提高了 Softmax 损失函数的性能。然而,通过在训练数据之间强制类内聚合和类间分离,即使训练数据可能不是均匀采样的,现有的 Softmax 损失函数需要训练数据的特征向量均匀分布。因此,对于两个测试数据,学习到的特征向量之间的接近度可能并不能正确地表示底层人脸之间的接近度,这可能会影响基于特征接近度的人脸识别算法的准确性。为解决这个问题,在原本的乘性角度距离损失函数上进行改进,使用其他属性(如性别、种族和年龄)来规范所学习的特征映射。改进后的损失函数在相邻训练数据的特征差与属性差之间建立全局线性关系,使得具有相似属性的人脸数据的特征向量聚拢在一起。此外,由于这些属性与人脸几何形状、外观相关,改进后的损失函数也隐含地将特征邻近性相对于人脸邻近性进行正则化,这有助于解释训练集中潜在的采样偏差。

改进算法在 CASIA Webface 数据集上训练,该训练集包括 10^4 个身份, 4.6×10^5 张照片,人工给该数据集加上属性标签,包括性别、种族和年龄。改进算法在不同人脸识别验证集上都表现出很好的效果,包括 LFW 数据集、CFP 数据集、AgeDB 数据集和 MegaFace 数据集。为确保网络中各层之间信息传递的最大化,直接让所有层相互连接,并引入残差学习模块,这减少了传统 CNN 模型的参数,提高了识别精度。虽然网络有较深的深度,但并不会出现梯度消失和过拟合的现象。

2 基本原理

2.1 人脸识别度量学习

经典识别网络有一个问题:必须提前设定好类别数。这也意味着,每增加一个新种类,就要重新定义网络模型,并从头训练一遍,比如需要做一个门禁系统,每增加或减少一个员工(相当于一个新类别),就要修改识别网络并重新训练。很明显,这种做法在某些实际运用中很不科学。因此,度量学习作为经典识别网络的替代方案,可以很好地适应某些特定的图像识别场景,一种较好的做法是丢弃经典神经网络最后的 Softmax 层,改成直接输出特征向量,按照度量学习,在特征库里寻找最近邻的类别作为匹配项。

人脸识别存在训练集不包含测试集的问题,使得较新的人脸识别方法的核心在于度量学习。度量学习尝试以监督的方式优化距离参数,在目标空间上使相似的对象在附近,不相似的对象相距很远。

在文献[7]中,当给定一些相似的样本对时,通过寻找带矩阵参数的马氏距离来完成学习。为处理更有挑战性的问题,在度量学习中引入卷积核技巧^[8]来提取非线性嵌入。近年来,利用先进的网络结构可以学习到更多有区分力的特征,这些网络结构使一些基于欧氏距离的损失函数最小化,如对比损失函数和三元组损失函数。此外,这些损失函数可以通过允许多个负样本之间的联合比较或最小化总体分类错误来改进^[9]。

2.2 人脸识别损失函数

在图像分类任务中,最常用的分类损失函数是 Softmax 损失函数,Softmax 损失函数可表示为

$$L_1 = -\frac{1}{N} \sum_{i=1}^N \log_2 \frac{\exp(\mathbf{W}_{y_i}^T \mathbf{x}_i + b_{y_i})}{\sum_{j=1}^n \exp(\mathbf{W}_j^T \mathbf{x}_i + b_j)}, \quad (1)$$

式中: \mathbf{x}_i 表示第 i 个样本的深度特征,属于第 j 类; y_i 表示样本的标签; \mathbf{W}_j 表示全连接层的权重矩阵 \mathbf{W} 的第 j 列; b_j 和 b_{y_i} 表示偏置项; N 表示批量处理的图片数; n 表示类别数。Softmax 损失函数将图像映射到深度特征,然后再映射到预测标签。经过 Softmax 损失函数训练的 CNN 可以产生判别特征向量。

然而,Softmax 损失函数主要鼓励类间分散,因此不能产生强的鉴别特征。为提高深度特征的区分力,中心损失函数被提出并用来加强类内聚集和类间分散,表示为

$$L_c = \frac{1}{2} \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{c}_{y_i}\|_2^2, \quad (2)$$

式中: L_c 表示中心损失; \mathbf{c}_{y_i} 表示第 y_i 个类别的深度特征的中心。中心损失函数常与 Softmax 损失函数一起使用,用参数 λ 控制两者的比重。联合损失 L 表示为

$$L = L_1 + \lambda L_c = -\frac{1}{N} \sum_{i=1}^N \log_2 \frac{\exp(\mathbf{W}_{y_i}^T \mathbf{x}_i + b_{y_i})}{\sum_{j=1}^n \exp(\mathbf{W}_j^T \mathbf{x}_i + b_j)} + \frac{\lambda}{2} \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{c}_{y_i}\|_2^2. \quad (3)$$

与此同时,观察到 Softmax 损失函数偏向于样本分布^[10],能很好地适应高质量的面孔,但忽略了低质量的面孔。在使用 Softmax 损失函数学习到的特征上添加 L2 约束,可以使得到的特征与那些使用中心损失函数训练得到的特征一样具有区分力。随后,有研究人员进一步改进了这一特性,将一

个角度距离代替欧氏距离加入到 Softmax 损失函数中,提出 sphereface 算法。该算法采用角度间隔损失函数,即 A-Softmax 损失函数,可表示为

$$L_{\text{ang}} = -\frac{1}{N} \sum_i \log_2 \cdot \left\{ \frac{\exp[\|\mathbf{x}_i\| \varphi(\theta_{y_i,i})]}{\exp[\|\mathbf{x}_i\| \varphi(\theta_{y_i,i})] + \sum_{j \neq y_i} \exp[\|\mathbf{x}_i\| \cos(\theta_{j,i})]} \right\}, \quad (4)$$

式中: $\varphi(\theta_{y_i,i}) = (-1)^k \cos(m\theta_{y_i,i}) - 2k$, $\theta_{y_i,i}$ 表示权重向量 \mathbf{W}_{y_i} 和特征向量 \mathbf{x}_i 之间的角度, \mathbf{W}_{y_i} 表示权重矩阵的 y_i 列,并且 $\theta_{y_i,i} \in \left[\frac{k\pi}{m}, \frac{(k+1)\pi}{m} \right]$, $k \in [0, m-1]$, m 是一个大于 1 的整数,控制着角度间隔的大小; $\theta_{j,i}$ 表示权重向量 \mathbf{W}_j 和特征向量 \mathbf{x}_i 之间的角度。

就多分类任务中最简单的情形二分类而言, Softmax 损失函数两种类别的决策边界为 $(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{x} + b_1 - b_2 = 0$, 其中 $\mathbf{W}_1, \mathbf{W}_2$ 分别表示权重矩阵的第 1 列和第 2 列, \mathbf{x} 表示特征向量, b_1, b_2 分别表示偏置矩阵的第 1 列和第 2 列。定义规范化处理后的 Softmax 损失函数为改进的损失函数,规范化处理包括权重归一化处理和偏置项置零处理。改进的损失函数的分类边界为 $\|\mathbf{x}\| \cos \theta_1 - \|\mathbf{x}\| \cdot \cos \theta_2 = 0$, 其中 θ_1, θ_2 分别表示特征向量与类别 1 和类别 2 的分类向量的角度,当 $\|\mathbf{x}\| \neq 0$ 时,分类边界变成 $\cos \theta_1 - \cos \theta_2 = 0$, 即分类边界完全取决于 θ_1 和 θ_2 。A-Softmax 的基本思想就是通过控制分类边界来产生类间的角度距离。在 A-Softmax 损失函数中,同样有归一化和偏置置零两个限制条件,特征向量 \mathbf{x} 被分类到类别 1 时的分类边界为 $\|\mathbf{x}\| [\cos(m\theta_1) - \cos \theta_2] = 0$, 特征向量 \mathbf{x} 被分类到类别 2 的分类边界为 $\|\mathbf{x}\| \cdot [\cos \theta_1 - \cos(m\theta_2)] = 0$ 。A-Softmax 损失函数和改进的损失函数的不同之处在于两个不同类的决策平面不是同一个, A-Softmax 损失函数有两个分隔的决策平面且决策平面分隔的大小还与 m 的大小成正相关。表 1 为二分类任务中 Softmax 损失函数、改进的 Softmax 损失函数和 A-Softmax 损失函数的分类边界。

2.3 基于角度间隔和人脸属性的损失函数

为获得较高的人脸识别精度,需要不同类别的人脸特征尽可能远离,相同类别的人脸特征尽可能靠近。理想情况下,两个人脸越不相似,对应的特

表 1 二分类任务下不同损失函数的决策边界的比较

Table 1 Comparison of decision boundaries of different loss functions in binary case

Loss function	Decision boundary
Softmax loss	$(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{x} + b_1 - b_2 = 0$
Modified Softmax loss	$\ \mathbf{x}\ \cos \theta_1 - \ \mathbf{x}\ \cos \theta_2 = 0$
A-Softmax	class 1: $\ \mathbf{x}\ [\cos(m\theta_1) - \cos \theta_2] = 0$ class 2: $\ \mathbf{x}\ [\cos \theta_1 - \cos(m\theta_2)] = 0$

征簇之间的距离就应该越远。然而, Softmax 损失函数的变体并不能很好地保证这一点。学习后的特征映射使得类内变化最小化,最大限度地增大了训练数据的类间距离,使得训练集人脸的特征向量分布趋于均匀;另一方面,不能保证训练集中的人脸图像在整个人脸空间中均匀分布。因此,当训练数据集中存在较大的采样密度变化时,所学习的特征映射可能不能正确地表示底层人脸的邻近性。为解决这个问题,可以引入一个损失函数项,它可对特征邻近性相对于人脸邻近性进行正则化。然而,这是一项具有挑战性的任务,因为人脸图像只能从一定的视角方向揭示潜在的脸型,并且会受到光照条件、传感器噪声等多种因素的影响。因此,仅通过比较两张人脸的扫描图像很难可靠地计算出它们之间的距离。

除了脸型的相似性,还希望学习到的特征映射与其他属性(如性别、种族和年龄)之间的相似性有关。例如,如果将一个图像与人脸图像数据库进行比较,通过特征邻近性来识别 K 个最可能匹配的图像,那么最好所有返回的图像都来自具有相同或相似属性的人,而上述 Softmax 损失函数的变体也不能保证这一特性,因为它们只考虑训练过程中的面部图像。

基于这些观察结果,本文将基于角度距离的损失函数进行改进,改进为人脸属性驱动。除标签信息,训练数据集中还给出人脸图像的性别、民族、年龄等其他属性。人脸属性损失部分可表述为

$$L_{\text{att}} = \frac{1}{2} \sum_{\substack{y_i < y_j \\ d(\mathbf{p}_i, \mathbf{p}_j) < \eta}} \|(\mathbf{f}_i - \mathbf{f}_j) - \mathbf{G}(\mathbf{p}_i - \mathbf{p}_j)\|_2^2, \quad (5)$$

式中: \mathbf{G} 表示一个需要通过训练来学习的矩阵参数; \mathbf{p}_i 和 \mathbf{p}_j 表示属性向量; $d(\mathbf{p}_i, \mathbf{p}_j) = \|\mathbf{p}_i - \mathbf{p}_j\|_2$ 表示两个属性向量之间的欧氏距离; η 表示自定义的阈值; \mathbf{f}_i 和 \mathbf{f}_j 分别表示 CNN 学习到的特征映射; L_{att} 表示人脸属性损失。

直观地说,这个损失项可以通过将特征差异

与属性差异关联起来的全局线性映射 \mathbf{G} , 将具有相似属性的特征集群聚拢在一起, 使得具有相似属性的人脸特征更加接近。如图 1 所示, 图中左侧代表输入的图片, 示例中两位男性具有相同的属性特征, 右侧代表使用 A-Softmax 损失函数和属性驱动损失函数学习的特征, 方块形点代表女性图片学习的特征, 圆点和三角形代表两位男性图

片学习到的特征。仅使用 A-Softmax 损失函数的人脸识别的特征分布趋于均匀, 当训练数据在人脸空间中采样不均匀时, 效果会变差。使用人脸属性损失项对损失函数进行扩展, 得到的特征具有与属性相似的分布, 更适用于识别, 增加人脸属性的正则化项后, 具有相似属性的特征簇变得更加接近。

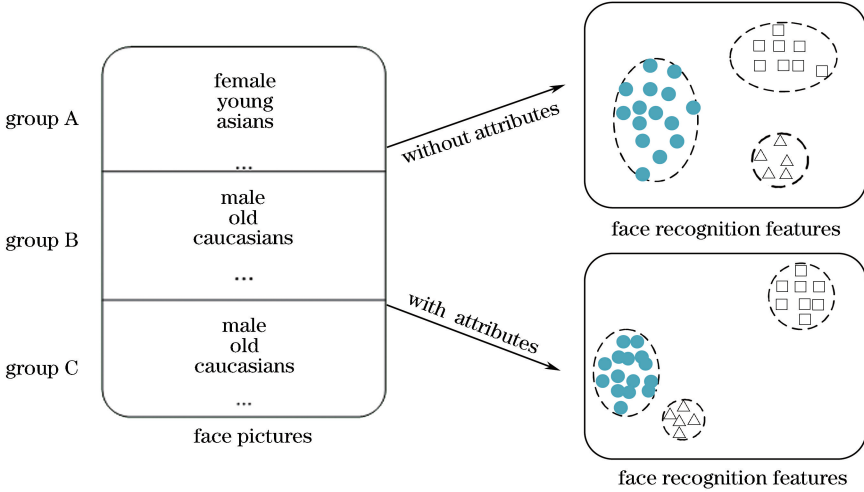


图 1 使用 A-Softmax 损失函数训练的特征分布和使用属性驱动损失函数训练的特征分布比较

Fig. 1 Comparison of feature distributions trained by A-Softmax loss function and attribute-driven loss function

从另一个角度看, 使用额外属性的正则化可以帮助网络获取其他有用的面部识别线索, 因为性别、种族和年龄等属性与面部形状和外观高度相关, 例如, 两个不同性别的人的面部表情可能会有显著的差异。因此, 属性驱动损失函数可以隐式地利用与这些属性

相关的外观变化来改善所学习的特征映射。

上述的属性损失函数可以作为一个人脸识别训练过程中的辅助监督信号, 与基于角度距离的损失函数相结合, 两种损失函数联合监督, 以提高人脸识别的准确率。联合监督损失函数表示为

$$L = L_{\text{ang}} + \lambda L_{\text{att}} = -\frac{1}{N} \sum_i \log_2 \left\{ \frac{\exp[\|\mathbf{x}_i\| \varphi(\theta_{y_i, i})]}{\exp[\|\mathbf{x}_i\| \varphi(\theta_{y_i, i})] + \sum_{j \neq y_i} \exp[\|\mathbf{x}_i\| \cos(\theta_{j, i})]} \right\} + \frac{\lambda}{2} \sum_{\substack{y_i < y_j \\ d(\mathbf{p}_i, \mathbf{p}_j) < \eta}} \|(\mathbf{f}_i - \mathbf{f}_j) - \mathbf{G}(\mathbf{p}_i - \mathbf{p}_j)\|_2^2, \quad (6)$$

式中, λ 为控制两个损失函数项比重的自定义参数。

3 分析与讨论

3.1 实验设置

3.1.1 训练数据

使用公开的人脸数据集 CASIA Webface 作为训练集, 该训练集包含 10575 个不同身份的 494414 张图片, 是相对较小的数据集, 并且该数据集不包含人脸属性。使用 MOON 损失函数^[11], CeleBA 作为

训练集, Inception-ResNet v2 网络作为网络结构, 训练人脸属性识别模型, 将该模型在 CASIA Webface 上的属性识别结果作为属性标注。

3.1.2 测试数据

测试数据包括 LFW 数据集和 MegaFace 数据集。LFW 数据集包含 13233 张图片, 来自 5749 个不同身份, 在姿态、表情和光照方面有很大的变化。按照官方给出的测试协议^[12], 对 LFW 数据集的 6000 张人脸对进行人脸验证。MegaFace 数据

集^[13]是一个公开的人脸测试数据集,引入了百万数量级的干扰项,用来评估人脸识别算法的性能。如果训练集的规模小于 5×10^5 ,那么该数据集可视为一个小数据集^[14]。本文是在小训练集数据量协议下对提出的算法进行评价。

3.1.3 训练细节

为比较损失函数的效果,分别使用 A-Softmax 损失函数和结合属性损失的 A-Softmax 损失函数进行实验。在实验中,采用随机梯度下降法训练网络参数,属性损失的比重 λ 设置为 0.0001,参数 η 设置为 0.01。为保证实验的公平性,实验中均采用 MTCNN 算法进行人脸检测和对齐^[15],提取特征的网络结构均使用 Inception-ResNet v2 网络^[16]。在训练过程中,所有实验的环境使用的是 PyTorch,编程语言为 Python3,使用两块 NVIDIA 1080Ti 显卡进行训练,批处理大小设为 256,学习率最初设置为 0.1,总迭代次数为 48 万次,分别在 24 万次、36 万次和 42 万次时将学习率除以 10。在训练时,权重衰减设置为 0.0005,权重动量设置为 0.9,对训练数据仅采用随机镜像对称作为数据增强。

3.2 超参数 η 和 λ 的影响

在基于角度距离和属性驱动的损失函数中,参数 λ 控制属性损失和角度距离损失的比重,决定 L_{att} 在总损失里的重要性,而参数 η 决定在属性损失中,一对属性向量是否足够接近。因为这两个参数对损失函数都极为重要,本文用 2 个实验来说明 λ 和 η 对人脸识别性能的影响。以 CASIA Webface

为训练集, LFW、CFP^[17]和 AgeDB^[18]为测试集验证算法的准确性。在训练集中,年龄和性别分布比较均衡,选择性别、种族和年龄作为训练模型的属性。对于性别属性,使用 1 表示男性,使用 -1 表示女性。对于种族属性,使用 1 表示亚洲人,而 -1 表示白种人。对于年龄,首先将年龄值截断为 100,然后从范围 $[0, 100]$ 线性映射到 $[-1, 1]$ 。将属性表示为三维向量 $\mathbf{p}_i = (p_i^{(g)}, p_i^{(e)}, p_i^{(a)})$,其中上标 g、e 和 a 分别表示性别、种族和年龄。

在第 1 个实验中,参数 $\eta = 0.02$, λ 的取值范围从 0 到 0.003。人脸识别的经典评价标准是闭集识别任务的性能。图 2(a)显示的是不同 λ 取值下得到的模型识别率,可以看到属性损失可以大大提高人脸识别的性能,尤其在 λ 的范围为 $[10^{-5}, 10^{-3}]$ 时。当 $\lambda < 3 \times 10^{-5}$ 时,识别准确率随着 λ 的增大而增大,刚开始增长得很快,后面增长得越来越慢;在 $\lambda = 3 \times 10^{-5}$ 时,达到识别准确率的最大值;当 $\lambda > 3 \times 10^{-5}$ 时,识别准确率随着 λ 增大而减小。在第 2 个实验中, $\lambda = 0.001$,参数 η 取值范围从 0 到 0.04。测试集对应的识别率如图 2(b)所示,可以观察到,当 $\eta < 0.01$ 时,识别准确率随着 η 的增大而增大;在 $\eta = 0.01$ 时,取得最大的准确率;当 $\eta > 0.01$ 时,识别准确率先下降而后趋于稳定。当 η 的取值在 $[0.01, 0.04]$ 范围时,在一个批处理中有 150 到 630 对具有不同身份的相似属性向量。在实践中,为减小计算量,通常会取较小的参数 η 。在后续的实验中,参数 λ 取值为 3×10^{-5} ,参数 η 取值为 0.01。

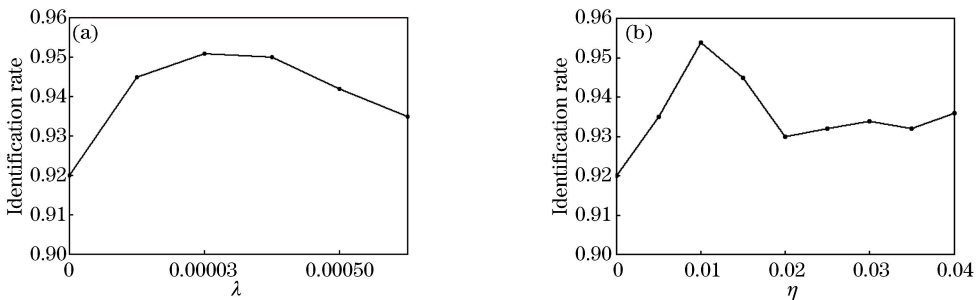


图 2 超参数 λ 和 η 对验证准确率的影响。(a) η 相同 λ 不同的验证准确率;
(b) η 不同 λ 相同的验证准确率

Fig. 2 Influences of super parameters λ and η on verification accuracy. (a) Verification accuracy with same η and different λ ; (b) verification accuracy with different η and same λ

3.3 实验结果

3.3.1 在不同网络结构下的实验结果

基于 CNN 模型结构设计的最新进展,在实验中探索了不同的网络结构对人脸识别验证准确率的影响,其中包括 ResNet、MobileNet、Inception-ResNet

v2、DenseNet (densely connected convolutional networks)、SENet (squeeze and excitation networks)。表 2 为不同的网络结构的验证准确率的比较,验证数据集包括 LFW、CFP-FP 和 AgeDB-30。由表 2 可见, Inception-ResNet v2 结构在所有网络结构中具有最高

的验证准确率,基于这个实验结果,在比较不同的损失函数的验证准确率时,采用 Inception-ResNet v2 作为实验的网络结构。

表 2 不同网络结构的验证准确率比较

Table 2 Comparison of verification accuracy of different network structures

Network structure	LFW /%	CFP-FP /%	AgeDB-30 /%
ResNet50	99.27	91.34	94.31
ResNet101	99.35	92.15	95.82
MobileNet	99.13	90.10	93.88
Inception-ResNet v2	99.67	93.00	97.42
DenseNet	99.54	92.39	96.40
SE-ResNet101	99.48	92.78	96.67

3.3.2 LFW 数据集的实验结果

网络结构采用相同的 Inception-ResNet v2, 训练数据集采用相同的 CASIA Webface 数据集, 得到不同损失函数在 LFW、CFP-FP、AgeDB-30 数据集的验证准确率, 其中包括 Softmax 损失函数、三元组损失函数 (Triplet loss)、中心损失函数 (Center loss)、L-Softmax 损失函数、A-Softmax 损失函数, 实验结果如表 3 所示。通过获取一个额外的属性监督信号, 基于角度距离和属性的改进损失函数在 LFW、CFP-FP、AgeDB-30 验证数据集上均取得优于其他损失函数的准确率。在姿态差异较大的 CFP-FP 数据集和年龄变化大的 AgeDB-30 数据集上, 改进损失函数的优越性更加明显。可见, 改进损失函数不仅能提升人脸识别的准确率, 而且能增强对年龄变化和姿态变化的鲁棒性。

表 3 不同损失函数的验证准确率

Table 3 Verification accuracy of different loss functions

Loss function type	LFW /%	CFP-FP /%	AgeDB-30 /%
Softmax	97.78	89.64	93.04
Triplet loss	98.65	90.22	95.88
Center loss	99.02	91.10	96.12
L-Softmax loss	99.15	91.90	96.20
A-Softmax loss	99.42	92.80	96.83
Modified A-Softmax loss	99.67	93.00	97.42

3.3.3 在 MegaFace 数据集的实验结果

MegaFace 数据集在评估人脸识别表现方面极具挑战。MegaFace 有多个测试场景, 包括在两个协议 (大训练集或小训练集) 下的识别、验证和姿态不变性。实验中采用两种场景, 人脸鉴别和人脸验证。实验结果如表 4 所示, 两种准确率分别为人脸鉴别任务下的准确率和人脸验证任务下的准确率, 从表 4 中可以看出, 与使用相同的 CNN 架构和数据集训

练的损失函数相比, 改进损失函数也显示出良好的性能。并且, 采用小规模训练集和属性驱动损失函数得到的识别准确率, 面对采用大规模训练集的模型仍然具有很大的竞争力。仅人脸鉴别准确率, 改进损失函数比 Vocord-DeepVo1 低 0.596 个百分点, 人脸验证准确率比 Vocord-DeepVo1 高 19.816 个百分点, 准确率更是高于大规模协议下的 Google-FaceNet V8。这些结果证明基于角度距离和属性的损失函数是一个良好的损失函数设计和监督信号, 属性驱动的特征提取可以明显地提升人脸识别的性能。

表 4 MegaFace 数据集下不同损失函数的准确率

Table 4 Accuracy of different loss functions in MegaFace dataset

Method	Protocol	Identification accuracy /%	Verification accuracy /%
Vocord-DeepVo1	Large	75.127	67.318
Google-FaceNet V8	Large	70.496	86.493
Softmax loss	Small	54.628	65.732
Triplet loss	Small	64.698	78.030
Center loss	Small	65.334	80.106
L-Softmax	Small	67.035	80.185
A-Softmax	Small	72.729	85.561
Modified A-Softmax loss	Small	74.531	87.134

4 结 论

提出一种基于角度距离和属性驱动的人脸识别损失函数, 该函数将学习到的识别特征相对于附加属性的分布进行正则化。该方法能够有效地解决训练数据集采样不均匀的问题, 提高了人脸识别的精度。采用提出的人脸识别算法在 CASIA Webface 数据集上进行训练, 然后在 LFW 和 MegaFace 人脸识别数据集上进行测试。实验结果证明了该算法的有效性, 训练好的人脸识别模型具有良好的泛化能力。在这项工作中, 虽然本文只使用性别、年龄和种族属性进行正则化, 但是实验结果仍然得到很好的改进, 后续还可以添加其他属性以进行进一步研究。

参 考 文 献

- [1] Li J N, Zhang B H. Face recognition by feature matching fusion combined with improved convolutional neural network [J]. Laser & Optoelectronics Progress, 2018, 55(10): 101504.
李佳妮, 张宝华. 特征匹配融合结合改进卷积神经网络

- 络的人脸识别[J]. 激光与光电子学进展, 2018, 55(10): 101504.
- [2] Hadsell R, Chopra S, LeCun Y. Dimensionality reduction by learning an invariant mapping[C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR'06), June 17-22, 2006, New York, NY, USA. New York: IEEE, 2006.
- [3] 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), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 815-823.
- [4] Ranjan R, Castillo C D, Chellappa R. L2-constrained softmax loss for discriminative face verification [J/OL]. (2017-06-07) [2019-04-24]. <https://arxiv.org/abs/1703.09507>.
- [5] Wen Y D, Zhang K P, Li Z F, *et al.* A discriminative feature learning approach for deep face recognition[M]//Leibe B, Matas J, Sebe N, *et al.* Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9911: 499-515.
- [6] 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), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6738-6746.
- [7] Xing E P, Ng A Y, Jordan M I, *et al.* Distance metric learning with application to clustering with side-information [C] // Advances in Neural Information Processing Systems, December 8-13, 2003, Vancouver and Whistler, British Columbia, Canada. Canada: NIPS, 2003: 521-528.
- [8] Hernandez M, Choi J, Medioni G. Near laser-scan quality 3-D face reconstruction from a low-quality depth stream [J]. Image and Vision Computing, 2015, 36: 61-69.
- [9] Sohn K. Improved deep metric learning with multi-class N-pair loss objective[C]//Advances in Neural Information Processing Systems, December 5-10, 2016, Barcelona, Spain. Canada: NIPS, 2016: 1857-1865.
- [10] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//Advances in Neural Information Processing Systems, December 3-6, 2012, Lake Tahoe, Nevada, United States. Canada: NIPS, 2012: 1097-1105.
- [11] Rudd E M, Günther M, Boulton T E. MOON: a mixed objective optimization network for the recognition of facial attributes[M]//Leibe B, Matas J, Sebe N, *et al.* Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9909: 19-35.
- [12] Huang G B, Learned-Miller E. Labeled faces in the wild: updates and new reporting procedures: University of Massachusetts, Amherst Technical Report UM-CS-2014-003[R]. Boston: University of Massachusetts, 2014.
- [13] Kemelmacher-Shlizerman I, Seitz S M, Miller D, *et al.* The MegaFace benchmark: 1 million faces for recognition at scale[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 4873-4882.
- [14] Yang H X, Liu J, Zhang M J, *et al.* Face recognition algorithm based on orthogonal gradient difference local directional pattern [J]. Laser & Optoelectronics Progress, 2018, 55(4): 041008. 杨恢先, 刘建, 张孟娟, 等. 基于正交梯度差局部分向模式的人脸识别算法[J]. 激光与光电子学进展, 2018, 55(4): 041008.
- [15] Zhang K P, Zhang Z P, Li Z F, *et al.* Joint face detection and alignment using multitask cascaded convolutional networks [J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.
- [16] Szegedy C, Ioffe S, Vanhoucke V, *et al.* Inception-v4, inception-ResNet and the impact of residual connections on learning [C] // Thirty-First AAAI Conference on Artificial Intelligence, February 4-10, 2017, Hilton, San Francisco Union Square. USA: AAAI, 2017: 4278-4284.
- [17] Sengupta S, Chen J C, Castillo C, *et al.* Frontal to profile face verification in the wild[C]//2016 IEEE Winter Conference on Applications of Computer Vision (WACV), March 7-10, 2016, Lake Placid, NY, USA. New York: IEEE, 2016: 16051830.
- [18] Moschoglou S, Papaioannou A, Sagonas C, *et al.* AgeDB: the first manually collected, in-the-wild age database[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 1997-2005.