

基于深度卷积神经网络与中心损失的人脸识别

张延安^{1,2} 王宏玉^{1,3} 徐 方^{1,3}

(中国科学院沈阳自动化研究所机器人国家重点实验室¹,沈阳 110016;中国科学院大学²,北京 100049;
沈阳新松机器人自动化股份有限公司³,沈阳 110168)

摘要 传统人脸识别方法手工设计特征过程复杂、识别率较低,对于开集人脸识别通用深度学习分类模型特征判别能力较弱。针对这两方面的不足,提出了一种以分类损失与中心损失相结合作为模型训练监督信号的深度卷积神经网络。首先,利用构建的应用场景数据集优调从公共数据集获得初始化参数的深度人脸识别模型,解决训练数据过小和数据分布差异问题,同时提高模型训练速度;然后,以传统损失函数和新的中心损失作为迁移学习过程中的监督信号,使得类内聚合、类间分散,提高模型输出人脸特征的判别能力;最后,对人脸特征进行主成分分析,进一步去除冗余特征,降低特征复杂度,提高人脸识别准确率。实验结果表明,与传统人脸识别算法相比该算法可以自动进行特征提取,并且相对于通用深度学习分类模型该算法通过度量学习使特征表示更具判别力。在自建测试集和 LFW、YouTube Faces 标准测试集上都取得了较高的识别率。

关键词 人脸识别 卷积神经网络 深度学习 中心损失 度量学习 主成分分析

中图法分类号 TP391.41; **文献标志码** A

人脸识别作为一种生物特征识别技术,与虹膜识别、指纹扫描、掌形扫描等技术相比,具有高精度、易于使用、稳定性高、难仿冒、性价比高等诸多优势,已经广泛应用于工业领域和我们的日常生活中。近年来,随着人工智能与机器人技术研究的快速发展以及人们对更加智能化生活的追求不断提高,智能人脸识别技术的需求市场进一步扩大。智能化办公考勤,安全金融支付验证,平安城市智能视频监控以及机器人机交互等新的应用场景都对人脸识别算法提出了更高的要求。而人脸识别的主要挑战就是克服由于光照、遮挡、姿态、表情、尺度等复杂场景因素导致的类内差距增大,类间差距减小问题。为了克服这个问题,使人脸识别算法能够很好地区分极其相似的不同人,又能认出场景变化很大的同一个人,进行深入的人脸识别算法研究具有重要意义。针对这个问题主要有两种解决方案,一种是在人脸识别系统的前端通过各种图像预处理来矫正场景变化,另一种是在人脸识别模型的损失函数部分采用度量学习的方法直接进行类内类间距离度量,通过

模型训练的监督信号来优化人脸识别模型。

传统的人脸识别算法主要有通过计算脸部部件的几何特征以及各部件间的相对位置关系的方法,通过库中存储的人脸模板与待识别人脸匹配的方法,将人脸图像映射到特定空间并由空间基向量表示的方法和根据数据训练获得人脸图像的统计特征再进行识别的方法。以上方法所用的特征提取技术在人脸特征的表示中均有不错的效果,但是由于都是手工设计的特征,都介入了许多主观因素,特征提取的好坏依赖于人工选择。传统的人脸识别算法在训练阶段能够利用的数据样本有限,训练出来的算法模型并没有特别好的泛化能力。

由于对人脸的表征十分复杂,而深度卷积神经网络模型具有很强的非线性刻画能力,因此将深度学习方法引入人脸识别领域是非常合适的。深度学习往往含有更深的层次结构,由低层简单特征逐层组合提取形成高层抽象特征,能够充分表达复杂的人脸模型。利用深度学习提取出的人脸特征,相比传统技术,更能表示人脸之间的相关性,辅之有效的分类方法,能够显著提高算法识别率。现在人脸识别率最高的算法基本都是采用了深度卷积神经网络结构,如 DeepID^[1]、FaceNet^[2]、DeepFace^[3]等。

然而,现有人脸识别方法仍然有许多不足之处,例如要求训练样本和测试样本具有相同分布的假设在实际应用中难以满足的问题,以及用主流的 softmax 损失函数作为模型训练的监督信号时,模型输出的特征判别能力较差的问题。针对这两个问题,本文在 VGG-Face^[4] 深度人脸模型的基础上,利用自

2017 年 5 月 14 日收到

国家科技支撑计划基金
(2015BAF13B01)资助

第一作者简介:张延安(1991—),硕士研究生。研究方向:深度学习与计算机视觉。E-mail:zhangyanan@sia.cn。

引用格式:张延安,王宏玉,徐方. 基于深度卷积神经网络与中心损失的人脸识别[J]. 科学技术与工程, 2017, 17(35): 92—97
Zhang Yanan, Wang Hongyu, Xu Fang. Face recognition based on deep convolution neural network and center loss[J]. Science Technology and Engineering, 2017, 17(35): 92—97

已构建的接近应用场景分布状态的训练数据集优调该模型。充分利用原有大数据学习到的特征的同时也通过迁移学习使模型适应了应用环境的数据分布。提出了 softmax 损失函数和新的中心损失函数联合作为模型训练的监督信号,为每一个类设置一个中心点,除了在继续做分类问题外,使所有该类的特征都尽可能的靠近这个中心点。这样通过扩大类间距离的同时,缩小类内距离,显著提高了模型所输出特征的判别能力。最后通过对输出特征进行主成分分析,进一步去除冗余特征,降低特征复杂度,提

高算法识别率。

1 人脸识别系统的组成

如图 1 所示,一个完整的人脸识别系统包括三个主要部分:人脸检测、人脸对齐和人脸识别。人脸检测是指在输入图像中确定所有人脸的位置、大小、位姿的过程。人脸对齐任务即根据输入的人脸图像,自动定位出面部关键特征点,如眼睛、鼻尖、嘴角点、眉毛以对齐到标准正面人脸。

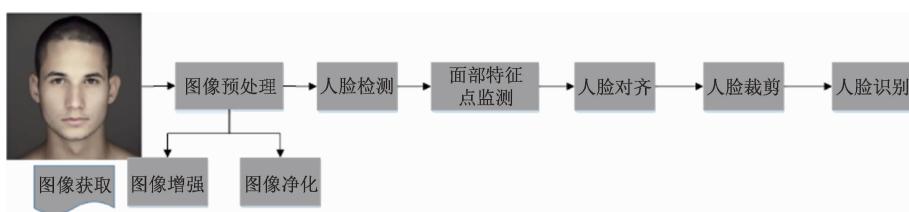


图 1 人脸识别系统

Fig. 1 Face recognition system

一个好的人脸识别系统在图像获取后一般都会进行图像预处理操作,包括常用的图像增强与图像净化部分。下一步就是进行人脸检测框出人脸区域,对人脸区域进行面部特征点检测,然后根据检测到的面部特征点进行仿射变换将人脸对齐到标准的正面人脸。最后对对其的人脸进行裁剪,将其输入到人脸识别网络提取特征向量进行识别。本文利用开源机器学习库 DLIB 中的人脸检测和对齐算法进行识别前的图像处理。人脸检测算法采用了 5 个 HOG 特征检测器级联的形式,人脸特征点检测采用集成回归树算法检测出 68 个面部特征点,然后通过仿射变换进行对其操作。

2 数据集构建

虽然在人脸识别领域已经存在几个数据量比较大的公开人脸数据集,但是这些数据大多是采集于网络搜索引擎,数据的分布状态相差很大。因此对于一个具体应用场景的人脸识别需求,用这些公开数据训练的人脸识别模型效果都不是很好。主要原因就是应用场景采集到的人脸数据与训练数据集的数据分布差异较大。然而要在应用场景中采集十分巨大的人脸数据不仅费时费力,而且成本较高。因此本文提出了通过在实际应用场景中构建的小数据集来优调用公开大数据训练的深度人脸识别模型的策略。这样不仅可以利用公开大数据的优势充分学习人脸的主体表征方式,还可以利用应用场景中的小数据尽可能拉近数据的分布状态。

本文数据集中所有图片都用上面提到的开源机

器学习库中的人脸检测和对齐算法进行了人脸的检测、面部特征点检测、对齐和裁剪。用 68 个面部特征点通过仿射变换进行了人脸对齐。最后,我们将图片裁剪到 128×128 ,并将像素值减去 127.5 后除以 128 进行归一化处理。构建数据集的流程如图 2 所示。



图 2 图像预处理流程示意图

Fig. 2 The flow chart of image preprocessing

通过以上处理方式本文最终构建的实际应用场景人脸训练数据集包括 1 万个人的 10 万张图片。经过数据均衡化处理,每个人有 10 张涵盖光照、姿态、遮挡、年龄变化的图片。部分人脸训练数据样例如图 3 所示。



图 3 训练图片样例

Fig. 3 Example of training pictures

为了最后检验人脸识别算法的效果,我们也构建了一个测试数据集。该数据集数据分布状态同训练数据相同,使用相同的工具进行了人脸检测、对齐、裁剪处理。此测试数据集包含 1 000 对人脸,每两张图片组成一个人脸对,这两张图片可能为同一个人也可能为不同的两个人。其中共包含正例 500 对,负例 500 对,人脸比对结果由待测试的算法提取

的人脸表征向量通过余弦距离度量方式给出。测试集的人脸图片样例如图 4 所示。



图 4 测试图片示例

Fig. 4 Example of test pictures

3 算法描述

3.1 VGG-Face 模型

VGG-Face 模型是用 2 622 个人的 260 万张人脸图片在 VGG^[5]分类模型的基础上训练而来。该模型包含了深层次的卷积层、池化层、RELU 非线性层、全连接层、Dropout 层和 Softmax loss 层。其中卷积层采用了比较小的三层 3×3 卷积核代替了传统的 7×7 大卷积核,既增加了网络的深度又减少了参数量,具有更多的非线性,使得判别函数更加具有判别性。

卷积操作是卷积网络的特征提取操作,该操作相比于传统的全连接层具有部分连接和权值共享的优势,可以极大减少传统全连接网络的参数量,对于符合局部统计特征与全局统计特平衡稳定性的自然图像,具有平移旋转不变性。池化层将局部特征进行聚合,不仅可以通过降维减小模型的参数量,还可以有效地避免模型过拟合,提高模型的泛化能力。常用的池化方法有最大池化、平均池化和随机池化,本文采用了最大池化策略。本文还采用了如式(1)所示的线性修正单元来代替传统的 sigmoid 神经元,可以有效地解决网络模型训练过程中的梯度消失问题,因此网络可以叠加的更深,训练速度也变得更快。Dropout 层通过在模型训练过程中根据设置的比率随机丢弃一部分权值,使网络对部分扰动更鲁棒,防止过拟合,相当于训练了一组弱分类器在测试时进行模型融合提升整体精度。

$$\text{ReLU}(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (1)$$

3.2 迁移学习

机器学习算法的一个大前提是假设训练数据与测试数据服从相同的数据分布。然而在大多数应用场景中,这种同分布的假设通常并不满足。一般而言,一类标记数据只适用于特定应用场景的模型训练,这就意味着在不同的应用场景要使用不同的标记数据训练不同的模型,而进行数据标记是一个非常耗时耗力的工作。迁移学习的目标是将从一个领域中学到的知识用来帮助新领域中的学习任务。换句话说,就是当前只有少量新的标记的数据,但是有

大量旧的已标记的数据,甚至是其他类别的有效数据,这时通过挑选这些旧数据中的有效的数据,加入到当前的训练数据中,训练新的模型。

本文针对人脸识别深度学习方法中要求训练样本和测试样本具有相同分布的假设在实际应用中难以满足的问题,利用模型优调的方法进行迁移学习。

用我们自己构建的实际应用场景数据集优调公开大数据训练的 VGG-Face 模型,将学习到的人脸表征方式迁移到我们的应用环境。在优调阶段我们使用了如式(2)所示的 softmax loss 损失函数作为训练过程中的监督信号,并使用了如图 5 所示的步进策略的学习率方式,每隔固定的迭代次数学习率按衰减因子自动下降。

$$L_S = - \sum_{i=1}^m \lg \frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^n e^{W_j^T x_i + b_j}} \quad (2)$$

式(2)中, $x_i \in R^d$ 是属于第 y_i 类的第 i 个深度特征, d 是特征维数; $W_j \in R^d$ 是 $W \in R^{d \times n}$ 在最后一层全连接层权重的第 j 列; $b \in R^n$ 是偏置项。小批次的尺寸和类别数分别是 m 和 n 。为了分析简便我们省略了偏置项,事实上这对性能几乎没有影响。

3.3 中心损失度量学习

大多数卷积神经网络都是以 softmax loss 损失函数作为深度模型训练过程的监督信号,这种损失函数训练出的模型虽然具有较好的可分性但判别性能力较差。不同于通用物体分类识别等闭集问题,对于人脸识别这个开集识别问题,不仅要求深度模型输出的特征具有可分性,还必须要有较强判别性。因为同一个人可能在人脸的角度、光照、表情、年龄、化妆、遮挡等不同场景变化下,出现类内距离较大的情况。而不同的人在待识别的人数众多时,可能会出现长得极为相似的人,产生类间距离较小的情形。为了有效避免出现类内距离大于类间距离而造成误识别的情况,就必须要求人脸识别模型提取到的特征除了具有可分性以外还要具有较强的判别性。

为了提高深度人脸识别模型学习到特征的判别能力,提出了一种叫做中心损失的新的监督信号。这个中心损失学习每一类深度特征的一个中心,以度量学习的方式对深度特征与它的类中心的距离进行惩罚。我们用 softmax loss 和中心损失共同作为深度模型训练的监督信号进行联合优调,以使得输出的人脸特征类内聚合类间、分散^[6]。

中心损失函数的定义为

$$L_C = \frac{1}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2 \quad (3)$$

式(3)中, $c_{y_i} \in R^d$ 是深度特征的第 y_i 个类中心。理

想的, c_{y_i} 应该随着深度特征变化而更新,但是在每次迭代都把整个训练集考虑在内来求每个类的平均特征是不现实的。为了解决这个问题,我们以每个小批次为单位进行中心值的更新并用标量 α 来控制中心的学习率。 L_c 相对于 x_i 的梯度和 c_{y_i} 的更新计算方式分别如式(4)和式(5)所示。

$$\frac{\delta L_c}{\delta x_i} = x_i - c_{y_i} \quad (4)$$

$$\Delta c_j = \frac{\sum_{i=1}^m \delta(y_i = j) (c_j - x_i)}{1 + \sum_{i=1}^m \delta(y_i = j)} \quad (5)$$

α 的值限制在 $[0, 1]$ 之间。我们所采用的联合监督信号如式(6)所示,其中参数 λ 用来平衡两种损失函数的比重。

$$L = L_s + \lambda L_c = - \sum_{i=1}^m \lg \frac{e^{W_i^T x_i + b_i}}{\sum_{j=1}^n e^{W_j^T x_i + b_j}} + \frac{\lambda}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2 \quad (6)$$

两种损失函数在深度神经网络中的连接方式如图 5 所示,其中 fc7 就是我们最终所要提取的人脸特征。

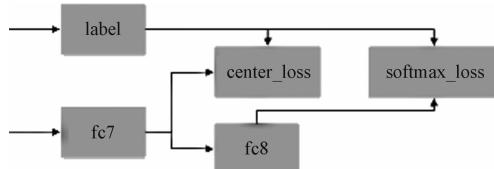


图 5 损失函数连接图

Fig. 5 Loss function connection graph

3.4 主成分分析

在本文中,首先,利用前面所述的算法步骤进行深度卷积神经网络模型的优调训练。然后,对于深度学习方法提取的最终人脸特征,我们用主成分分析的方法进一步去除冗余特征,降低特征复杂度。最后,通过余弦距离度量人脸相似度,画出 ROC 曲线计算 AUC 值。

本文最终模型输出的人脸特征向量为 4 096 维,该向量还存在相关性,也就是说仍然具有冗余特征。而主成分分析是一种可以通过正交变换将一组可能存在相关性的变量转换为只有主成分的多元统计方法。通过人脸特征的主成分分析可以在尽可能多的保留原始人脸模型信息的同时,使特征点之间互不相关,使特征复杂度降低。对于本文中的特征向量降维,经过多次试验发现,将人脸特征维度降到 160 维时,特征的表征能力最强,人脸识别效果最好。

4 实验结果与分析

4.1 实验平台

本实验中硬件配置为 PowerEdge R730 机架式戴尔服务器,英特尔至强处理器 E5—2603 v4, NVIDIA Tesla K80 GPU。软件配置为 Ubuntu 14.04,CUDA7.5 GPU 并行计算库,开源深度学习框架 Caffe^[7]。

4.2 迁移优调实验

首先将前面建立的 10 万训练数据集按照 9:1 的比例划分为训练集和验证集,训练集包含 1 万人的九万张图片,验证集包含 1 万人的 1 万张图片。测试数据集为 1 000 对人脸图片,正样本 500 对,负样本 500 对。同时为了提高实验效果还做了镜像图片反转、图像随机裁剪的数据增广策略,以及图像去均值操作。将数据集中 128×128 像素图片调整尺寸到 256×256 像素,然后做随机裁剪到网络模型所需的 224×224 像素大小。

优调时修改最后一层全连接层的层名,将输出节点数设置为数据集类别数 10 000。由于前面各层使用 VGG-Face 模型的网络参数作为初始化参数,比较接近最优值,基础学习率可以设置为较小的 0.001。最后一层全连接层使用高斯法随机初始化,因此学习率设置因子设置为 10 倍。学习率以 7 000 次迭代为步长,衰减因子 0.1,最大迭代次数 3 万次,损失函数为 softmax loss。

对 1 000 对测试集图片分别用训练好的模型提取人脸特征向量,并用 PCA 进行降维,然后用距离度量的方式计算每对人脸的相似度。常用的距离度量方式有欧式距离和余弦距离^[8],本文采用如式(7)所示的余弦距离进行度量。

$$\cos(x, y) = \frac{(x, y)}{\|x\| \cdot \|y\|} = \frac{\sum_{i=1}^n x_i y_i}{\left(\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2\right)^{\frac{1}{2}}} \quad (7)$$

x 和 y 为一对人脸的特征向量,如果其余弦距离大于某个阈值被认为是同一个人,否则为不同人。不断改变阈值的大小画出如图 6 所示的 ROC 曲线并计算出 AUC 值。

4.3 中心损失函数实验

在上述优调好的模型上添加中心损失函数,以两种损失函数再做联合优调。通过小批次特征向量求平均值为每一个类设置一个聚类中心点,在做分类的同时把图像特征和这个中心点做 L2 loss 约束,让所有该类的特征都尽可能靠近这个中心以提高人脸特征的判别能力。中心学习率(设置为 0.5,中心

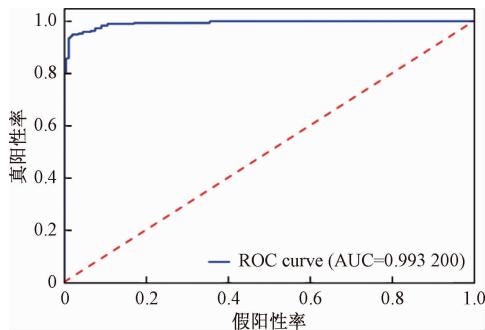


图 6 迁移模型的 ROC 曲线

Fig. 6 The ROC curve of the migration model

损失函数权重设置为 0.003, 以上步优调好的模型参数作为初始化参数, 迭代训练 3 万次。

用训练好的模型对 1 000 对测试集图片分别提取人脸特征向量, 并用主成分分析方法去除冗余特征进行降维, 然后用余弦距离度量的方式计算每对人脸的相似度。不断改变阈值的大小画出如图 7 所示的 ROC 曲线并计算出 AUC 值。

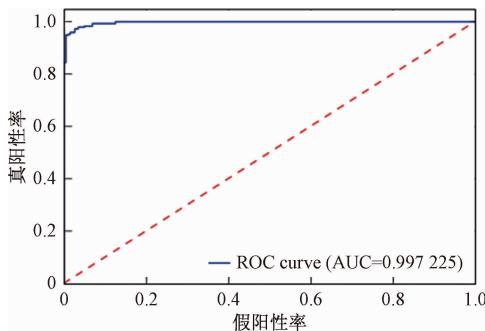


图 7 加入中心损失后模型的 ROC 曲线

Fig. 7 The ROC curve of the model after adding the center loss

4.4 综合实验结果

为了更好的证明算法的有效性以及验证相关参数对模型训练的影响, 本文还做了几个扩展实验, 在自建测试数据及上包括原始 VGG-Face 模型测试实验, 不同批尺寸下的模型优调实验, 不同中心损失函数比重下的联合优调实验, 主成分分析实验。实验结果如表 1 所示。

实验结果表明, 本文提出的通过应用场景数据集进行深度卷积神经网络迁移学习的方法, 在较小训练集的条件下达到了比 VGG-Face 模型更好的人脸识别效果, 且随着训练批尺寸的增加准确率在一定范围内稍有增加。本文提出的新的中心损失函数与 softmax 损失函数联合优调的方法通过提高人脸特征的判别能力进一步提高了准确率, 且在中心损失比重为 0.003 时达到最优。应用 PCA 对最终特征向量进行降维, 不仅降低了特征复杂度, 准确率也有提升。

表 1 各个参数下模型 AUC 值比较

Table 1 Comparison of model AUC values under various parameters

序号	实验方法	AUC 值
1	原始 VGG-Face	0.987 625
2	softmax loss 优调(批尺寸 32)	0.992 500
3	softmax loss 优调(批尺寸 64)	0.993 044
4	softmax loss 优调(批尺寸 128)	0.993 117
5	softmax loss 优调 + PCA(维数 160)	0.993 200
6	softmax loss + center loss 联合优调($\lambda = 0.0005$)	0.995 320
7	softmax loss + center loss 联合优调($\lambda = 0.003$)	0.997 150
8	softmax loss + center loss 联合优调($\lambda = 0.05$)	0.995 574
9	softmax loss + center loss 联合优调 + PCA (维数 160)	0.997 225

通过上述实验证进行了训练参数的选择, 训练批尺寸设为 128, 中心损失的中心学习率设为常数值 0.5, 中心损失函数权重系数设为 0.003, PCA 后的特征维数设为 160, 生成最终的最优模型。然后分别在标准的 LFW^[9]、YouTube Faces^[10] 人脸识别评测集上对各人脸识别算法识别率进行对比, 实验结果如表 2 所示。其中在 LFW 评测集上的测试结果是在非限制协议下获得的。

表 2 各种人脸识别算法准确率比较

Table 2 Comparison of the accuracy of various face recognition algorithms

算法	模型集成数量	LFW	YouTube Faces
LBP	1	0.851 7	0.657
Fisher vector faces	1	0.930 3	0.838
HD-LBP	1	0.931 8	—
DeepFace	3,1	0.973 5	0.914
DeepID	1	0.974 5	—
DeepID2	25	0.991 5	—
DeepID2 +	200	0.994 7	0.932
DeepID3	1	0.995 3	—
FaceNet	1	0.996 3	0.951
VGG-Face	1	0.989 5	0.916
Our algorithm	1	0.982 7	0.911

表 2 中各种人脸识别算法对比结果表明, 我们的算法相对于传统人脸识别算法识别准确率有较大提升, 可见通过深度卷积神经网络自动抽取特征不仅避免了繁杂的手工设计特征过程, 而且所得到的特征对人脸的表征能力更强。相对于其他深度学习算法, 虽然我们的算法在最准确率上没有大幅度提升, 但我们的算法不需要多模型集成就能达到较高的准确率, 同时相对于 DeepID2 和 FaceNet 的度量学习^[11]方式, 我们的中心损失度量计算简单, 避免了模型训练中复杂的训练数据输入对的选取, 提高了模型输出特征的判别能力。另外, 通过迁移学习^[12]中的优调方法, 充分利用相似场景中易于获取的大量数据来学习领域知识作为模型训练的初始化参数, 加速模型优化的收敛过程, 可以很好地解决应

用场景训练数据不足的问题。

5 结束语

本文提出了一种以分类损失与中心损失相结合作为深度卷积神经网络模型训练监督信号来实现人脸识别的方法。利用分类损失来控制深度人脸识别模型的类间区分能力,通过中心损失的度量学习来约束类内变化,提高模型输出人脸特征的判别能力。深度卷积神经网络通过逐层特征抽取组合自动提取最优秀表征人脸的特征向量,并用主成分分析清除噪声和冗余特征,进一步提高人脸特征的表征能力。实验结果表明,该方法在自建测试集和 LFW、YouTube Faces 标准测试集上都取得了较高的人脸识别准确率。然而实际应用场景中往往包含更多诸如光照、姿态、表情、遮挡等干扰因素,因此下一步研究的重点是复杂场景下的鲁棒人脸识别算法。

参 考 文 献

- 1 Sun Y, Wang X, Tang X. Deep learning face representation from predicting 10 000 classes. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 1891—1898
- 2 Schroff F, Kalenichenko D, Philbin J. Facenet: a unified embedding for face recognition and clustering. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 815—823
- 3 Taigman Y, Yang M, Ranzato M A, et al. Deepface: closing the gap to human-level performance in face verification. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 1701—1708
- 4 Parkhi O M, Vedaldi A, Zisserman A. Deep face recognition. Proceedings of the British Machine Vision. 2015; 1(3) : 6
- 5 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv Preprint arXiv:1409.1556 , 2014
- 6 Wen Y, Zhang K, Li Z, et al. A discriminative feature learning approach for deep face recognition. European Conference on Computer Vision. Springer International Publishing, 2016: 499—515
- 7 Jia Y, Shelhamer E, Donahue J, et al. Caffe: convolutional architecture for fast feature embedding. Proceedings of the 22nd ACM International Conference on Multimedia. ACM, 2014: 675—678
- 8 Xu Y, Zhang Z, Lu G, et al. Approximately symmetrical face images for image preprocessing in face recognition and sparse representation based classification. Pattern Recognition, 2016; 54: 68—82
- 9 Huang G B, Ramesh M, Berg T, et al. Labeled faces in the wild: a database for studying face recognition in unconstrained environments; Technical Report 07—49. Amherst: University of Massachusetts, 2007
- 10 Wolf L, Hassner T, Maoz I. Face recognition in unconstrained videos with matched background similarity. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2011: 529—534
- 11 王法强. 基于度量学习的人脸识别方法研究. 哈尔滨: 哈尔滨工业大学, 2013
Wang Faqiang. Research on face recognition method based on metric learning. Harbin: Harbin Institute of Technology, 2013
- 12 张博, 史忠植, 赵晓非, 等. 一种基于跨领域典型相关性分析的迁移学习方法. 计算机学报, 2015; (7) : 1326—1336
Zhang Bo, Shi Zhongzhi, Zhao Xiaofei, et al. A transfer learning method based on cross domain canonical correlation analysis. Journal of Computer Science, 2015; (7) : 1326—1336

Face Recognition Based on Deep Convolution Neural Network and Center Loss

ZHANG Yan-an^{1,2}, WANG Hong-yu^{1,3}, XU Fang^{1,3}

(State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences¹, Shenyang 110016, P. R. China;
University of Chinese Academy of Sciences², Beijing 100049, P. R. China; Shenyang SIASUN Robot &
Automation Co., LTD.³, Shenyang 110168, P. R. China)

[Abstract] For traditional face recognition methods, the process of manual design features is complex and face recognition rate is low. The feature discrimination ability of general deep learning classification model is weak, for the open set face recognition. Aiming at these two problems, a kind of deep convolution neural network is proposed, which combines the classification loss with the central loss as the model training monitoring signal. Firstly, a deep face recognition model based on the initialization parameters obtained from the public dataset is fine tuned using application scene dataset, which can effectively solve the problem of training data is too small and data distribution differences and improve the training speed of the model. Then, the traditional loss function and the new central loss are used as the monitoring signals in the process of transfer learning, which can make the intra-class aggregation and inter-class dispersion and improve the discriminative ability of the model output features. Finally, the principal component analysis is used to remove the redundant face features, reduce the complexity and improve face recognition rate. The experimental results show that our algorithm can automatically extract features compared with the traditional face recognition algorithm and relative to the general deep learning classification model, the algorithm makes the feature representation more discriminative with metric learning. A higher recognition rate has been achieved in the self built test set and the LFW and YouTube Faces Standard test sets.

[Key words] face recognition convolutional neural network deep learning center loss metric learning principal component analysis