



facenet皮尔森判别网络的人脸识别方法

谷凤伟, 陆军, 夏桂华

引用本文:

谷凤伟, 陆军, 夏桂华. facenet皮尔森判别网络的人脸识别方法[J]. *智能系统学报*, 2022, 17(1): 107–115.

GU Fengwei, LU Jun, XIA Guihua. Face recognition method based on facenet Pearson discrimination network[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(1): 107–115.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202104008>

您可能感兴趣的其他文章

一种多层特征融合的人脸检测方法

Face detection method fusing multi-layer features

智能系统学报. 2018, 13(1): 138–146 <https://dx.doi.org/10.11992/tis.201707018>

基于卷积特征和贝叶斯分类器的人脸识别

Face recognition based on convolution feature and Bayes classifier

智能系统学报. 2018, 13(5): 769–775 <https://dx.doi.org/10.11992/tis.201706052>

人脸识别背后的数据清理问题研究

Research on the data cleansing problem for face recognition technology

智能系统学报. 2017, 12(5): 616–623 <https://dx.doi.org/10.11992/tis.201706025>

一种新融合算法的维吾尔族人脸识别

A new fusion algorithm for uyghur face recognition

智能系统学报. 2018, 13(3): 431–436 <https://dx.doi.org/10.11992/tis.201710014>

基于嘴巴状态约束的人脸特征点定位算法

A facial feature point locating algorithm based on mouth-state constraints

智能系统学报. 2016, 11(5): 578–585 <https://dx.doi.org/10.11992/tis.201602006>

一种基于P稳定局部敏感哈希算法的相似人脸检索系统设计

A similar-face-image-retrieval system design based on a P-stable locality-sensitive Hashing algorithm

智能系统学报. 2017, 12(3): 392–396 <https://dx.doi.org/10.11992/tis.201607005>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.202104008

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.tp.20211214.1634.005.html>

facenet 皮尔森判别网络的人脸识别方法

谷凤伟^{1,2}, 陆军^{1,2}, 夏桂华^{1,2}

(1. 哈尔滨工程大学 智能科学与工程 黑龙江 哈尔滨 150001; 2. 哈尔滨工程大学 船海装备智能化技术与应用教育部重点实验室 黑龙江 哈尔滨 150001)

摘要: 非限制场景下存在光照、遮挡和姿态变化等问题, 这严重影响了人脸识别模型的性能和准确度。针对该问题, 本文对 facenet 进行改进, 提出了一种基于 facenet 皮尔森判别网络的人脸识别方法 facenetPDN。首先, 构建 facenetPDN 深度卷积神经网络, 在 facenet 前端融合多任务级联卷积神经网络进行人脸检测提取目标人脸。然后, 通过深度神经网络提取人脸深度特征信息, 采用皮尔森相关系数判别模块替换 facenet 中的欧氏距离判别模块实现人脸深度特征判别。最后, 使用 CASIA-WebFace 和 CASIA-FaceV5 人脸数据集训练网络。为了证明本文方法的有效性, 训练后的模型在 LFW 和 celeA 人脸数据集进行测试和评估, 并进行对比分析。实验结果表明, 改进后的 facenetPDN 方法的准确度比原来整体提高了 1.34%, 在融合训练集上提高了 0.78%, 该算法鲁棒性和泛化能力优良, 可实现多人种的人脸识别, 对非限制场景下人脸目标具有良好的识别效果。

关键词: 非限制场景; 人脸识别; facenet; 多任务级联卷积神经网络; 人脸检测; 皮尔森相关系数; 欧氏距离; 人脸数据集

中图分类号: TP391.41 文献标志码: A 文章编号: 1673-4785(2022)01-0107-09

中文引用格式: 谷凤伟, 陆军, 夏桂华. facenet 皮尔森判别网络的人脸识别方法 [J]. 智能系统学报, 2022, 17(1): 107-115.

英文引用格式: GU Fengwei, LU Jun, XIA Guihua. Face recognition method based on facenet Pearson discrimination network[J]. CAAI transactions on intelligent systems, 2022, 17(1): 107-115.

Face recognition method based on facenet Pearson discrimination network

GU Fengwei^{1,2}, LU Jun^{1,2}, XIA Guihua^{1,2}

(1. College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China; 2. Key Laboratory of Intelligent Technology and Application of Marine Equipment, Harbin Engineering University, Harbin, 150001, China)

Abstract: In unrestricted scenes, there are problems such as illumination, occlusion, and pose changes, which seriously affect the performance and accuracy of the face recognition model. This paper improves facenet to solve this problem, and proposes a facenetPDN method based on facenet Pearson discriminant network. Firstly, the deep convolutional neural network in facenetPDN is constructed, and the multi-task cascaded convolutional neural network is fused on the front-end of facenet to detect and extract the target face. Then, the facial depth feature information is extracted through the deep convolutional neural network, and the Pearson correlation coefficient discrimination module is used to replace the Euclidean distance discrimination module in the facenet algorithm to realize the facial depth feature discrimination. Finally, CASIA-WebFace and CASIA-FaceV5 face datasets are used to train the network. The trained model is tested and evaluated on the LFW and celeA face datasets to prove effectiveness of the method in this paper, and a comparative analysis is performed. The experimental results show that the accuracy of the improved facenetPDN method is 1.34% higher than that of the original method as a whole, and the accuracy of the model after training in the fusion training dataset is improved by 0.78%. The algorithm has excellent robustness and generalization ability, which can realize multi-ethnic face recognition, and has a good recognition effect on face targets in unrestricted scenes.

Keywords: unrestricted scene; face recognition; facenet; multi-task cascaded convolutional neural network; face detection; Pearson correlation coefficient; euclidean distance; face dataset

人脸识别是计算机视觉领域中的一个颇具综合性的研究方向, 涉及图像处理、模式识别和计算机图形学等多种技术。人脸识别与虹膜识别、

指纹识别等技术相比, 具有数据采集方便快捷、成本低、无接触等优点, 容易使用户接受^[1], 广泛应用于身份校验、金融安全和智能监控^[2]等领域, 获得了良好的应用效益。但是, 在非限制场景下人脸识别技术仍然面临着背景复杂、光照、遮挡、人脸姿态多变等诸多干扰因素^[3], 因此设计

收稿日期: 2021-04-06. 网络出版日期: 2021-12-15.

基金项目: 黑龙江省自然科学基金项目(F201123).

通信作者: 陆军. E-mail: lujun0260@sina.com.

出一种鲁棒性和泛化能力良好的人脸识别算法实现对人脸的准确识别仍然是一个严峻的挑战。

传统的人脸识别算法主要以人工特征和机器学习算法为基础,如基于稀疏表示的方法、基于子空间的方法等。主要技术有方向梯度直方图法^[4](histogram of oriented gradient, HOG)、主成分分析法^[5](principal component analysis, PCA)、支持向量机^[6](support vector machine, SVM)和线性判别分析法^[7](linear discriminative analysis, LDA)等。其中, HOG 通过计算和统计目标图像局部区域的梯度方向直方图^[8]来获得特征,可与 SVM 分类器一同应用在人脸图像识别任务,获得了不错的效果。但是由于梯度计算的原因,算法容易受到噪声的干扰。PCA 利用坐标系变换去除人脸图像的冗余信息,可以大幅度降低人脸图像的维度^[9],从而提取主要的人脸特征用于人脸识别。该法应用简单高效,但是要想取得优良的识别效果,需要训练集和测试集高度相关。SVM 是一种浅层的机器学习模型,对人脸数据虽然具有一定的学习能力^[10],但是无法提取人脸深层次的特征信息。LDA 是一种线性分类^[11]的特征提取方法。在人脸识别过程中,该方法可利用较大的类间距离区分不同的人脸,较小的类内距离将同类别人脸归为一类,但是对于非线性因素,该法效果不佳。

虽然传统的人脸识别方法在一定程度上取得了不错的识别效果,但是仍然无法得到人脸图像的深度特征。随着计算机硬件条件的不断提升,深度学习技术^[12]被逐渐应用到目标检测、自然语言处理和图像生成等领域,因而人脸识别技术也得到了飞速的发展。基于深度学习的人脸识别方法能够获得人脸图像的深度特征,深度模型表现优良,已成为主流方法。DeepFace 模型^[13]首次在人脸识别中利用深度学习技术,该模型涉及了 2D 和 3D 人脸对齐操作,在非限制条件下利用卷积神经网络获得的准确度和人工接近。DeepID^[14]通过将人脸图像分割成图像块的方式训练深度神经网络,能够更好地获得人脸图像的深度特征,该方法在非限制场景下具有优秀的性能表现。VGGNet 模型经过 VGGFace 人脸数据集^[15]的训练也达到了一个不错的准确度。Wang 等^[16]提出

了一种新的损失函数,解决了传统的深度卷积神经网络中 softmax 损失缺少辨别能力的问题。Scherhag 等^[17]通过利用人脸数据库的子集创建来测试和评估真实数据,提高了算法的检测性能。Prasad 等^[18]利用卷积神经网络改善了模型的鲁棒性。上述基于深度特征的人脸识别方法在特定条件下取得了良好准确度,但是在非限制场景下存在光照、遮挡等一系列的干扰因素,人脸识别的难度加大。为了保证人脸识别的准确度,提高系统的鲁棒性和泛化能力,本文提出了一种基于 facenet 皮尔森判别网络的人脸识别方法 facenetPDM。该方法首先建立了 facenetPDM 深度神经网络结构,在 facenet 网络前端通过多任务级联卷积神经网络 (multi-task cascaded convolutional neural network, MTCNN)^[19]完成人脸检测,获得目标人脸图像,然后确定 facenet^[20]中的深度卷积神经网络 Inception-ResNet-v2^[21],利用该结构得到目标人脸的深度特征信息,经皮尔森相关系数判别模块 (the discriminant module of pearson correlation coefficient, PDM)^[22]判定人脸深度信息的相关性,并给出量化指标,从而实现人脸识别。为验证本文方法有效性,使用 CASIA-WebFace 和 CASIA-FaceV5^[23]训练算法模型,并在 LFW 和 celeA^[24]人脸数据集测试和评估模型性能,将 facenetPDM 与 facenet 以及几种常见的人脸识别方法进行对比分析。

1 人脸识别算法

人脸识别算法是人脸识别系统的核心所在,本文将利用所提出的人脸识别算法构建一套人脸识别系统,以满足在非限制场景下的应用需求。人脸识别算法 facenetPDM 的网络框架如图 1 所示,人脸识别系统的图像处理流程如图 2 所示。人脸识别的整体流程是:首先,在非限制场景下的任意人脸图像对进入人脸识别系统,经过 MTCNN 提取目标人脸区域图像块;然后,利用 facenetPDM 提取对应图像块中的人脸深度特征向量,并经 PDM 模块判定特征向量的相关程度;最后,人脸识别系统给出皮尔森相关系数和判定结果。若该流程中的一张人脸图像来自于人脸数据库,这就完成了人脸身份的查找和确认。

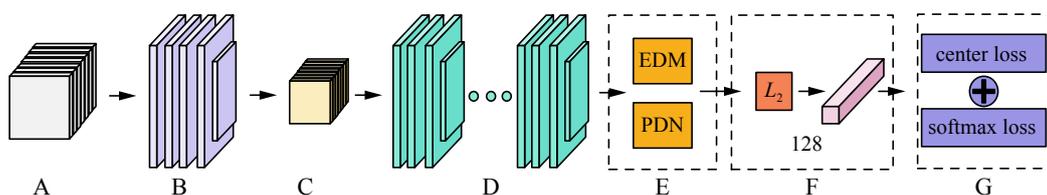


图 1 facenetPDM 的网络框架

Fig. 1 Network framework of facenetPDM

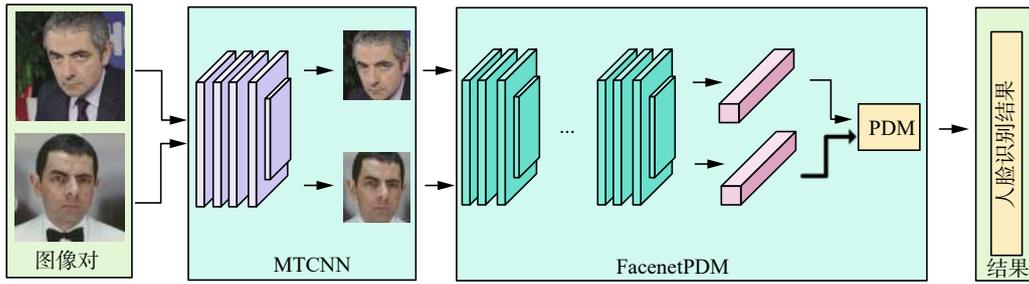


图 2 人脸识别系统的图像处理流程

Fig. 2 Image processing flow of face recognition system

1.1 算法的网络框架

本文以 facenet 网络为人脸识别方法的基础框架, 在 facenet 网络中插入 MTCNN 网络, 将 facenet 中原有的欧氏距离模块 (euclidean distance module, EDM) 去除, 在相应的位置处插入皮尔森相关系数判别模块 PDM, 得到 facenetPDM 人脸识别算法。在图 1 所示的 facenetPDM 的网络框架中, A 是人脸图像, B 是多任务级联卷积神经网络, C 是人脸数据, D 是深度卷积神经网络, E 是度量模块 (EDM 或 PDM), F 是归一化过程, G 是损失函数。当度量模块为 EDM 时, facenetPDM 退化为 facenet 网络结构, facenet 利用 EDM 得到的特征距离, 距离越大表明人脸图像来自于同一人的可能性越小, 但是 facenetPDM 经 PDM 得到的相似度值越大, 表明人脸图像来自于同一人的可能性就越大, 两者识别结果呈负相关趋势。

facenetPDM 把 MTCNN 的人脸检测输出作为网络输入, 利用深度卷积神经网络将人脸图像映射到高维特征空间, 提取人脸深度特征向量信息, 经 L_2 归一化后得到 128 维特征向量, 再利用 center 损失和 softmax 损失优化网络模型, 训练后的模型通过 PDM 进行人脸深度特征向量的皮尔森相关系数判定, 这就完成了人脸识别任务。

1.2 深度卷积神经网络的结构

在非限制场景下, 人脸特征的提取过程存在光照、遮挡等多种环境因素的干扰, 这就要求深度卷积神经网络具有足够的特征提取能力。在 facenetPDM 的网络框架中, 深度卷积神经网络是其核心结构, 主要作用是人脸图像映射到高维特征空间, 得到人脸图像的深度特征, 利用该高维度的深度特征实现人脸识别任务。因而, 本文采用 Inception-ResNet-v2 深度卷积神经网络作为 facenetPDM 的主干网络。Inception-ResNet-v2 的网络结构如图 3 所示。主要包括 stem 模块、Inception-ResNet- $M(i)$ 模块 (模块类型 $M=A, B, C$, 对应的模块个数 $i=5, 10, 5$)、Reduction- N 模块 (模块类型 $N=A, B$) 等。其中, stem 部分为 Inception-ResNet-v2 网络的数据输入结构, 用于对输入数据的初步处理, 其输入和输出数据的形状分别

为 $299 \times 299 \times 3$ 和 $35 \times 35 \times 384$, 显然其输出特征已经达到了一定的深度。Inception-ResNet- X ($X=A, B, C$) 模块在增加了网络深度的同时保证了网络宽度, 可以提取人脸的深度特征, 并且更有利于加快网络的收敛速度。Reduction- Y ($Y=A, B$) 模块的主要作用是补偿目标人脸深度特征的维度, 可以利用滤波器提高特征的维度。实际上, Inception-ResNet-v2 将 inception 网络和残差思想结合, 在提高网络深度后可以有效提取人脸深度特征, 并且保证了深度网络模型训练的可行性。

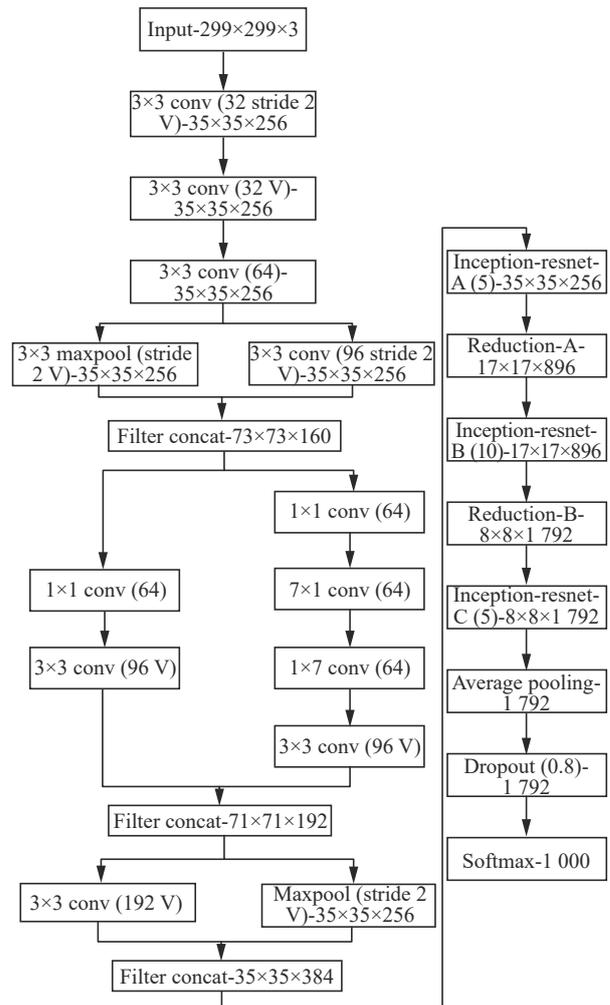


图 3 Inception-ResNet-v2 的网络结构

Fig. 3 Network structure of Inception-ResNet-v2

1.3 facenetPDM 的损失函数

损失函数是算法模型训练的关键因素,能够反映出模型输出值和真实值之间的误差距离。在本文 facenetPDM 网络模型的训练过程中,采用 softmax 损失^[25]和 center 损失^[26]相结合的方式优化网络模型。其中,softmax 损失强调类别之间的可分性,但是对类别之内的分布效果不佳;center 损失可以将类内的差异最小化,在功能上可以和 softmax 损失互相补充。因此,采用组合损失函数的方式既可以解决单一损失训练模型准确率不佳的问题,又可以改善整体模型的训练效果。facenetPDM 的损失函数为

$$L_{\text{facenetPDM}} = \lambda L_{\text{center}} + L_{\text{softmax}} \quad (1)$$

式中: λ 为比例因子; L_{center} 为 center 损失; L_{softmax} 为 softmax 损失。

center 损失和 softmax 损失分别为

$$L_{\text{center}} = \frac{1}{2} \sum_{i=1}^N \|g(x_i) - c_{j_i}\|_2^2 \quad (2)$$

$$L_{\text{softmax}} = - \sum_{i=1}^n \log \frac{e^{w_{ji}^T g(x_i) + b_{ji}}}{\sum_{k=1}^n e^{w_k^T g(x_i) + b_k}} \quad (3)$$

式中: N 为训练数据中的人脸样本数; x_i 为训练数据中的第 i 个人脸样本; $g(x_i)$ 为 x_i 人脸样本对应的人脸特征; j_i 为对应的人脸样本类别; c_{j_i} 为人脸类别对应的特征中心; n 为人脸样本的类别数; w_{j_i} 为第 i 个人脸类别样本的权重; b_{j_i} 为人脸类别样本的权重对应的偏置值; w_k 为第 k 个人脸类别权重; b_k 为第 k 个人脸类别权重对应的偏置值。

1.4 皮尔森相关系数判别模块

人脸图像数据在经过 facenetPDM 时会得到对应的人脸深度特征向量,此时需要根据特征向量给出判定结果。本文采用皮尔森相关系数判别模块 PDM 完成人脸深度特征向量的识别,需要先将 facenet 中的 EDM 模块替换为 PDM 模块,从而给出向量间的相关系数值以及最终的判定结果。现假设存在两个 n 维的人脸深度特征向量 X 、 Y , 其中 $X = [X_1 X_2 \cdots X_m]^T$, $Y = [Y_1 Y_2 \cdots Y_m]^T$, 则 X 、 Y 的皮尔森相关系数如式(4)所示:

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (4)$$

式中: X_i 、 Y_i 分别为人脸深度特征向量 X 、 Y 的元素; \bar{X} 、 \bar{Y} 分别为人脸深度特征向量 X 、 Y 的所有元

素的均值。

皮尔森相关系数 ρ_{XY} 的取值范围为 $[-1, 1]$ 。当 ρ_{XY} 值在 $[-1, 0)$ 之间时,表示 X 、 Y 之间是负相关的,可认为人脸特征不是来自同一人;当 $\rho_{XY} = 0$ 时,表示 X 、 Y 之间没有线性相关关系,也可认为人脸特征不是来自同一人;当 ρ_{XY} 值在 $(0, 1]$ 之间时,表示 X 、 Y 之间是正相关的,且越接近于 1 人脸特征来自于同一人的可能性越大。当然,可根据需求将 ρ_{XY} 的取值范围映射到 $[0, 1]$ 的范围内,若 G 为映射关系,则此时的人脸深度特征向量的皮尔森相关系数如式(5)所示:

$$\text{SIM}_p = G(\rho_{XY}) \quad (5)$$

式中: SIM_p 为映射结果;取值范围 $\text{SIM}_p \in [0, 1]$ 。

EDM 模块计算人脸深度特征向量间欧氏距离的过程中,对特征向量的每一个维度赋予了相同的权重大小,一旦出现不同维度间的取值范围差别很大的情况,就会很容易导致判定结果被某些维度决定,处理的人脸深度特征向量都是高维向量,很容易发生这种情况。PDM 模块可以避免这种情况的发生,这要得益于皮尔森相关系数获取过程中存在数据中心化的操作。数据中心化的主要过程是对每个人脸深度特征向量的每个维度的元素值都减去向量所含元素的均值,虽然这需要提前计算均值,但是均值计算很容易实现。

2 实验

2.1 实验环境和数据集

本文实验采用 tensorflow2.0 框架搭建的网络结构,在具有 Nvidia GPU 的 Windows10 系统中完成网络模型的训练与测试评估, GPU 的具体参数为 RTX 2080Ti 11GB。在模型训练时,学习率初始值可以设定为 0.05,并且随着训练轮数的增加学习率可根据实际情况动态调整,应用 Adam 优化器不断更新和优化网络模型参数以减小损失值。采用 CASIA-WebFace 和 CASIA-FaceV5 人脸数据集融合的方式一起训练模型,在 CASIA-WebFace 中融入 CASIA-FaceV5 可在一定程度上达到提高模型的鲁棒性和泛化能力的目的。在模型测试评估阶段,采用 LFW 和 celeA 人脸数据集。为验证方法的有效性,分别在 LFW 和 celeA 中选出一系列的人脸图像组成图像对,并且对应分组,利用本文提出的 facenetPDM 以及 facenet 方法分别对每组图像进行测试评估,并以 LFW 为基础绘制 ROC 曲线^[27]来对比分析几种常见人脸识别方法的性能。

2.2 实验结果和分析

为了验证本文提出的人脸识别方法的有效性,在实验中从 LFW 和 celeA 人脸数据集中各挑

选出 10 对处于非限制场景下的人脸图像进行测试分析, 对应的人脸图像为 A 和 B 两组, 并把每个组别中的图像对依次编号为 N01~N10, 同时以标签的形式给出人脸识别的度量值和判定结果, A 和 B 两组人脸识别的结果分别如图 4、5 所示。如

图 4 中的第一对图像对在第一张人脸图像的左上角给出图像对编号 N01, 在两张人脸图像之间位置处标出对应方法下的度量值标签, 此处为 0.828 5, “Y”表示人脸识别的判定结果“正确”, 若标注“N”则表示判定结果为“错误”。



图 4 A 组人脸识别对比

Fig. 4 Comparison of face recognition in group A

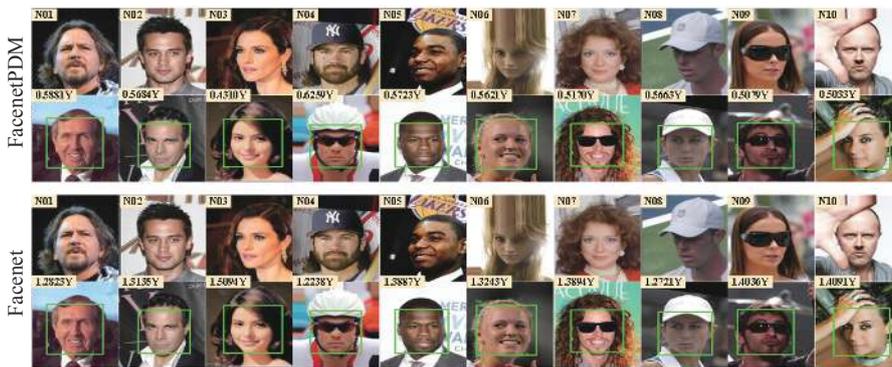


图 5 B 组人脸识别对比

Fig. 5 Comparison of face recognition in group B

从图 4、5 中 A、B 两组人脸整体的识别结果上看, 所选图像对在 facenetPDM 方法下的人脸识别结果均正确, 在 facenet 方法下 N09、N10 识别错误, facenetPDM 方法的识别准确度较高。如图 4 所示, A 组人脸图像是同一人在不同非限制场景下的选取结果, 主要存在人脸姿态变化、光照、遮挡、表情夸张和背景变化剧烈等问题, 且某些图像对受多种因素干扰。图像对 N01、N09、N10 主要存在遮挡的问题, 图像中人脸被眼镜、话筒和手臂部分遮挡, 很可能发生人脸特征提取不完整的情况, 但在实验中利用 facenetPDM 方法时识别结果均正确, 而利用 facenet 方法时 N09、N10 识别结果错误。N03、N06、N07 中人脸姿态变化较大, 同时 N03、N06 较 N07 表情变化略大, 两种方法下的人脸识别结果均正确。N04、N05 背景变化较为剧烈, 且存在轻微的人脸姿态变化。N01、N02、N03、N08、N10 主要存在光照影响, 致

使人脸图像肤色发生严重改变, 其中 N01、N03 肤色变化相对较轻, 其余 3 个图像对肤色变化严重, 这在很大程度上会干扰人脸特征向量信息, 导致识别失败, 虽然 N02、N08、N10 在利用 facenetPDM 方法下均取得了正确的识别结果, 但是 N10 的度量值较低, 仅为 0.748 8, 可能是眼镜遮挡以及表情变化等因素引起的这种现象。然而, 利用 facenet 方法时对 N10 识别错误。上述实验结果分析表明 facenetPDM 具有较好的鲁棒性和泛化能力。

B 组人脸图像是不同人在不同非限制场景下的选取结果, 人脸识别对比结果如图 5 所示。B 组人脸图像主要存在光照、遮挡、图像模糊等问题, 而且部分图像对受到多种干扰因素的严重影响, 例如 N05 背景反差、人脸姿态变化, 以及背景与肤色的区分。但是 B 组人脸图像在 facenetPDM、facenet 方法下的识别结果均正确, 这说明一

一般情况下不同人的脸特征间的差异较大,但是这仍然需要人脸识别方法要足够的稳健。N01、N04、N07、N09 存在胡须遮挡干扰;N02、N03、N05 存在人脸姿态和肤色对比;N06 和 N08 存在人脸姿态变化、图像模糊、光线干扰;N04、N07、N08、N09、N10 存在帽子、眼镜、手掌的遮挡。这些干扰因素会对人脸特征的提取产生一定的不良影响,在这种条件下获得的人脸特征与人脸原有特征相比会有较大偏差。但从实验结果来看,这种偏差对人脸识别的影响多数是在可控范围内的,并且 facenetPDM、facenet 方法表现良好且人脸识别结果均正确,没有严重地影响到人脸识别最终的判定结果,可见该方法较为稳健。

A、B 两组人脸图像中含有多类别人种的人脸,如 A 组人脸图像中的 N03 和 B 组人脸图像中的 N02、N05 等为黑色人种的人脸识别,如 A 组人脸图像中的 N05、N06、N07 和 B 组人脸图像中的 N04、N07 等为白色人种的人脸识别。从图 4、5 人脸识别的结果可知,本文的 facenetPDM 可有效获得多人种对应的人脸深度信息完成人脸识别,且获得了优良的识别效果。

A、B 两组图像对的实验数据进行量化统计,结果如表 1 所示,其中的 Th1、Th2 代表对应方法下的阈值。从统计结果可以看出, facenet 的整体准确率为 90%, facenetPDM 的整体准确率为 100%, facenetPDM 和 facenet 方法均取得了良好的人脸识别效果,但 facenetPDM 方法的准确度整体上要优于 facenet 方法。在本实验中, N01--N08 图像对在两种方法下的度量值与对应阈值的偏差较大,这正是所希望的结果,一旦出现噪声等干扰因素也不会对识别的结果产生太大的影响,即可以保证识别准确度,当然干扰是在一定范围内的。但是对于 A 组 N09、N10 图像对具有多种干扰因素,导致 facenet 的识别结果错误,而 facenetPDM 的识别结果正确,这也表明了 facenetPDM 方法具有良好的鲁棒性和泛化能力。

接下来,本文在 LFW 人脸数据集上对人脸识别方法 facenetPDM 和 facenet 进行整体测试和评估。采用 CASIA-WebFace 和 CASIA-FaceV5 人脸数据集一起训练本文方法的算法模型, CASIA-FaceV5 中具有不同的人脸姿态、光照等条件下的人脸图像,可以进一步丰富训练集人脸信息,在一定程度上也可以改善网络模型的鲁棒性和泛化能力。现在将 CASIA-WebFace 人脸数据集定义为 FD1,将 CASIA-WebFace 和 CASIA-FaceV5 的融合数据集定义为 FD2,训练后的网络模型在

LFW 上的测试准确度如表 2 所示。从整体测试结果来看, facenetPDM 的识别准确度要高于 facenet;从细节上看, facenetPDM 的识别准确度整体上较 facenet 提高了 1.34%, facenetPDM 的识别准确度在 FD1 和 FD2 上较 facenet 分别提高了 0.56% 和 0.22%, facenetPDM 在 FD2 上的识别准确度较 FD1 提高了 0.78%, facenet 在 FD2 上的识别准确度较 FD1 提高了 1.12%。因此,分析可知适当地增加数据集可以在一定程度上提高模型准确度,并且在数据集不变的条件下适当地改变模型结构也可以提高模型的准确度,本文的 facenetPDM 方法在识别准确度上取得了不错的效果。

表 1 A、B 两组图像对的实验数据量化统计表
Table 1 Quantitative statistics of experimental data in the two image pairs A and B

方法	facenet (Th ₁ =0.88)		facenetPDM (Th ₂ =0.72)	
	A	B	A	B
N01	0.823 1Y	1.282 3Y	0.828 5Y	0.588 1YY
N02	0.786 3Y	1.313 5Y	0.844 4Y	0.568 4Y
N03	0.736 7Y	1.509 4Y	0.789 9Y	0.431 0Y
N04	0.796 3Y	1.223 8Y	0.840 3Y	0.625 9Y
N05	0.814 4Y	1.388 7Y	0.836 6Y	0.572 3Y
N06	0.838 6Y	1.324 3Y	0.824 1Y	0.562 1Y
N07	0.624 1Y	1.389 4Y	0.9023Y	0.517 0Y
N08	0.655 8Y	1.272 1Y	0.897 2Y	0.566 3Y
N09	1.014 7N	1.403 6Y	0.7433Y	0.507 9Y
N010	1.004 3N	1.409 1Y	0.748 8Y	0.503 3Y
Y/N	8/2	10/0	10/0	10/0
单组准确率/%	80.00	100	100	100
整体准确率/%	90.00		100	

表 2 模型的测试准确度
Table 2 Test accuracy of model

方法	数据集	facenet/%	facenetPDM/%
LFW	FD1	97.16	97.72
	FD2	98.28	98.50

为了评估本文模型的整体性能水平,本文以 LFW 为基础绘制了几种常见人脸识别方法 (DeepFace^[28]、TL Joint Bayesian^[29]、High-dim LBP^[30]、Deep FR^[31]、WebFace^[32]) 的 ROC 曲线进行对比分析。如图 6 所示给出了各个人脸识别方法的 ROC 曲线,可知 ROC 曲线的纵轴表示真阳性率,横轴表示假阳

性率。事实上,真阳性率的值越大越好,而假阳性率的值越小越好,但是在真阳性率的值增大的同时假阳性率的值也在随着不断地增大,因此需要找到一个平衡点作为模型性能的判定依据。由图 6 中各个方法 ROC 曲线的变化趋势分析可知,ROC 曲线与直线 $y=-x+1$ 的交点可以使真阳性率和假阳性率的值达到一个平衡状态,在此情形下如果交点越接近于点 $(0,1)$ 就代表真阳性率的值越大,假阳性率的值越小,这也就表示 ROC 曲线下的面积越大,ROC 曲线对应的模型的性能越好,这正是所期望达到的效果。如图 6 可知, facenetPDM 方法对应的 ROC 曲线下的面积 $AUC_{\text{facenetPDM}}=0.9081$, facenet 方法对应的 ROC 曲线下的面积 $AUC_{\text{facenet}}=0.9037$,显然 $AUC_{\text{facenetPDM}}$ 的值更大一些,因此改进后的 facenetPDM 比 facenet 的模型性能更好。再将本文方法对应的 ROC 曲线与其他方法对比可知, facenetPDM 方法模型的 ROC 曲线最靠近点 $(0,1)$,且曲线下的面积 AUC 值最大为 0.9081,因而本文的 facenetPDM 方法可以有效提高模型性能。

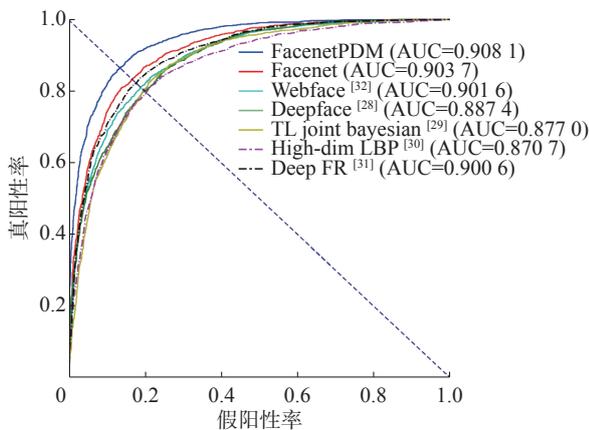


图 6 ROC 曲线对比

Fig. 6 Comparison of ROC curves

3 结束语

针对非限制场景下人脸图像存在光照、遮挡等干扰,本文提出了一种基于 facenet 皮尔森判别网络的人脸识别方法 facenetPDN,目的在于提高人脸识别模型的准确度和性能。facenetPDN 方法利用皮尔森相关系数判别模块替换 facenet 中的欧氏距离判别模块完成人脸深度特征判别,利用 CASIA-WebFace 和 CASIA-FaceV5 人脸数据集结合的方式一起训练网络模型,并在 LFW 和 celeA 人脸数据集上进行测试与评估。实验结果表明,在非限制场景下 facenetPDN 方法可以很好地提高目标人脸识别的准确度,实现多人种的人

脸识别任务,并具备良好的鲁棒性和泛化能力,可将其应用于线下的人脸识别系统,避免线上系统的网络限制,实用性更佳。

本文未来的研究内容可从以下几个方面进行:1)丰富数据集的多样性,使数据更加全面;2)研究模型结构,进一步提高目标特征提取的能力;3)在保证模型性能和准确度的前提下,进行模型优化。

参考文献:

- [1] GALBALLY J, MARCEL S, FIERREZ J. Image quality assessment for fake biometric detection: application to iris, fingerprint, and face recognition[J]. IEEE transactions on image processing, 2014, 23(2): 710–724.
- [2] HOANG V D, DANG V D, NGUYEN T T, et al. A solution based on combination of RFID tags and facial recognition for monitoring systems[C]//2018 5th NAFOSTED Conference on Information and Computer Science (NICS). Ho Chi Minh City, Vietnam: IEEE, 2018: 384–387.
- [3] CAI Weidong, MA Bo, ZHANG Liu, et al. A pointer meter recognition method based on virtual sample generation technology[J]. Measurement, 2020, 163: 107962.
- [4] NIGAM S, SINGH R, MISRA A K. Efficient facial expression recognition using histogram of oriented gradients in wavelet domain[J]. Multimedia tools and applications, 2018, 77(21): 28725–28747.
- [5] VYANZA V E, SETIANINGSIH C, IRAWAN B. Design of smart door system for live face recognition based on image processing using principal component analysis and template matching correlation methods[C]//2017 IEEE Asia Pacific Conference on Wireless and Mobile (AP-WiMob). Bandung, Indonesia: IEEE, 2017: 23–29.
- [6] ZHU Yue, GAO Wanrong, GUO Zhenyan, et al. Liver tissue classification of en face images by fractal dimension-based support vector machine[J]. Journal of biophotonics, 2020, 13(4): e201960154.
- [7] OUYANG Aijia, LIU Yanmin, PEI Shengyu, et al. A hybrid improved kernel LDA and PNN algorithm for efficient face recognition[J]. Neurocomputing, 2020, 393: 214–222.
- [8] 张毅, 廖巧珍, 罗元. 融合二阶 HOG 与 CS-LBP 的头部姿态估计 [J]. 智能系统学报, 2015, 10(5): 741–746.
ZHANG Yi, LIAO Qiaozhen, LUO Yuan. Head pose estimation fusing the second order HOG and CS-LBP[J].

- CAAI Transactions on Intelligent Systems, 2015, 10(5): 741-746.
- [9] 杨恢先, 刘建, 张孟娟, 等. 双差值局部方向模式的人脸识别[J]. 智能系统学报, 2018, 13(5): 751-759.
YANG Huixian, LIU Jian, ZHANG Mengjuan, et al. Face recognition with double difference local directional pattern[J]. CAAI transactions on intelligent systems, 2018, 13(5): 751-759.
- [10] 刘峤, 秦志光, 陈伟, 等. 基于零范数特征选择的支持向量机模型[J]. 自动化学报, 2011, 37(2): 252-256.
LIU Qiao, QIN Zhiguang, CHEN Wei, et al. Zero-norm penalized feature selection support vector machine[J]. Acta automatica sinica, 2011, 37(2): 252-256.
- [11] 朱换荣, 郑智超, 孙怀江. 面向局部线性回归分类器的判别分析方法[J]. 智能系统学报, 2019, 14(5): 959-965.
ZHU Huanrong, ZHENG Zhichao, SUN Huaijiang. Locality-regularized linear regression classification-based discriminant analysis[J]. CAAI transactions on intelligent systems, 2019, 14(5): 959-965.
- [12] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [13] TAIGMAN Y, YANG Ming, RANZATO M A, et al. DeepFace: closing the gap to human-level performance in face verification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014: 1701-1708.
- [14] SUN Yi, WANG Xiaogang, TANG Xiaoou. Deep learning face representation from predicting 10,000 classes[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014: 1891-1898.
- [15] BanBANSAL A, NANDURI A, CASTILLO C D, et al. UMDFaces: an annotated face dataset for training deep networks[C]//2017 IEEE international joint conference on biometrics. Denver, USA: IEEE, 2017: 464-473.
- [16] WANG Hao, WANG Yitong, ZHOU Zheng, et al. CosFace: large margin cosine loss for deep face recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 5265-5274.
- [17] SCHERHAG U, RATHGEB C, MERKLE J, et al. Deep face representations for differential morphing attack detection[J]. IEEE transactions on information forensics and security, 2020, 15: 3625-3639.
- [18] PRASAD P S, PATHAK R, GUNJAN V K, et al. Deep learning based representation for face recognition [M]. ICCCE 2019. Singapore: Springer, 2020: 419-424.
- [19] ZHANG Kaipeng, ZHANG Zhanpeng, LI Zhifeng, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE signal processing letters, 2016, 23(10): 1499-1503.
- [20] 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. Boston, USA: IEEE, 2015: 815-823.
- [21] THOMAS A, HARIKRISHNAN P M, PALANISAMY P, et al. Moving vehicle candidate recognition and classification using inception-ResNet-v2[C]//2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC). Madrid, Spain: IEEE, 2020: 467-472.
- [22] ZHAO Yuanfang, ZHEN Zonglei, LIU Xiqin, et al. The neural network for face recognition: insights from an fMRI study on developmental prosopagnosia[J]. NeuroImage, 2018, 169: 151-161.
- [23] GAO Wen, GAO Bo, SHAN Shiguang, et al. The CASPEAL large-scale Chinese face database and baseline evaluations[J]. IEEE transactions on systems, man, and cybernetics-part A: systems and humans, 2008, 38(1): 149-161.
- [24] LIU Ziwei, LUO Ping, WANG Xiaogang, et al. Deep learning face attributes in the wild[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 3730-3738.
- [25] LI Xiaobin, WANG Weiqiang. Learning discriminative features via weights-biased softmax loss[J]. Pattern recognition, 2020, 107: 107405.
- [26] FARZANEH A H, QI Xiaojun. Discriminant distribution-agnostic loss for facial expression recognition in the wild[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle, USA: IEEE, 2020: 1631-1639.
- [27] ALBIERO V, BOWYER K W, VANGARA K, et al. Does face recognition accuracy get better with age? deep face matchers say no[C]//Proceedings of the IEEE Winter Conference on Applications of Computer Vision. Snowmass, USA: IEEE, 2020: 250-258.
- [28] TAIGMAN Y, YANG M, RANZATO M, et al. Deep-

Face: closing the gap to human-level performance in face verification[C]// Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2014.

- [29] CAO Xudong, WIPF D, WEN Fang, et al. A practical transfer learning algorithm for face verification[C]//2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE, 2013: 3208–3215.
- [30] CHEN Dong, CAO Xudong, WEN Fang, et al. Blessing of dimensionality: high-dimensional feature and its efficient compression for face verification[C]//2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013: 3025–3032.
- [31] PARKHI O M, VEDALDI A, ZISSERMAN A. Deep face recognition[C]//Proceedings of the British Machine Vision Conference. Swansea, UK: BMVA Press, 2015: 41.1–41.12.
- [32] YI Dong, LEI Zhen, LIAO Shengcai, et al. Learning face representation from scratch[J]. arXiv preprint arXiv: 1411.7923, 2014.

作者简介:



谷凤伟, 博士研究生, 主要研究方向为人脸识别、机器视觉、图像处理、目标跟踪。



陆军, 教授, 博士生导师, 主要研究方向为计算机视觉、智能控制技术、图像处理、高性能船舶控制和机械臂控制。科技部科技型中小企业创新基金项目评审专家, 国家自然科学基金同行评议专家。获省部级以上奖励多项。发表学术论文 70 余篇。



夏桂华, 教授、博士生导师, 主要研究船舶运动建模与仿真、智能船舶技术、船舶驾控模拟系统、图像处理和智能机器人控制。享受国务院特殊津贴专家, 获省部级以上奖励多项。发表学术论文 70 余篇。

首期人工智能创新技术研讨会

由中国人工智能学会主办, 浙江杭州未来科技城管理委员会、中国人工智能学会会士之家(杭州站)承办的“元宇宙与人工智能”首期人工智能创新技术研讨会活动将在于 2022 年 1 月 15 日杭州未来科技城国际人才园举办。

当下, 人工智能已成为新一轮科技及产业革命的重要驱动力。“元宇宙”作为近期的热门话题, 受到社会各界的广泛关注与讨论。在“元宇宙”的世界里, 人工智能也将担任重要角色, 为元宇宙赋予智能的“大脑”。

报名链接: <https://www.wenjuan.com/s/Mfmeay6/#>

联系方式:

中国人工智能学会会士之家(杭州站) 电话: 0571-88697861

邮箱: fellow_hz@caai.cn