

基于前后端交互的人脸识别系统^①



侯景严, 宋焕生, 梁浩翔, 贾金明, 戴 喆

(长安大学 信息工程学院, 西安 710064)

通讯作者: 宋焕生, E-mail: 838772436@qq.com

摘 要: 针对现有的人脸识别系统计算效率低和鲁棒性较差等问题, 本文提出了一种基于前后端交互的人脸识别系统, 系统包含客户端、数据库以及服务端. 首先, 在客户端提出了基于 GrabCut 的人脸兴趣区域 (ROI) 提取算法. 其次, 将提取到的 ROI 传输到服务端, 并在服务端使用 ResNet 网络根据 ROI 提取人脸特征点. 最后, 将服务端中提取到的人脸特征点返回给客户端, 在客户端将该信息与数据库中预存的特征点进行欧式距离匹配, 得到人脸识别结果. 实验在 CeleA 数据集与随机视频上进行测试, 结果表明提出的 ROI 提取算法明显提升了人脸识别的精度和鲁棒性, 并且系统的前后端交互结构相较于传统的非交互结构极大地提升了人脸识别的计算效率.

关键词: 人脸识别; 前后端交互; 残差神经网络; 特征提取; GrabCut

引用格式: 侯景严, 宋焕生, 梁浩翔, 贾金明, 戴喆. 基于前后端交互的人脸识别系统. 计算机系统应用, 2020, 29(10): 89-96. <http://www.c-s-a.org.cn/1003-3254/7672.html>

Face Recognition System Based on Front and Back Interaction

HOU Jing-Yan, SONG Huan-Sheng, LIANG Hao-Xiang, JIA Jin-Ming, DAI Zhe

(School of Information Engineering, Chang'an University, Xi'an 710064, China)

Abstract: Aiming at the problems of low calculation efficiency and poor robustness of the existing face recognition system, this study proposed a face recognition system based on front and back interaction, including client, database, and server. First, a GrabCut-based facial Region Of Interest (ROI) extraction algorithm was proposed for the client end. Second, the extracted ROI is transmitted to the server, and the ResNet network is used on the server to extract facial feature points according to the ROI. Finally, the facial feature points extracted from the server were returned to the client, and the client performs Euclidean distance matching between this information and the feature points that were pre-stored in the database to obtain the face recognition result. The experiments were performed on the CeleA database and random videos, and the results show that the proposed ROI extraction algorithm significantly improves the accuracy and robustness of face recognition. Moreover, compared with the traditional non-interactive structure, the front and back interactive structure of the system greatly improves the computational efficiency of face recognition.

Key words: face recognition; front and back interaction; residual neural network; feature extraction; GrabCut

人工智能的发展离不开计算机视觉的技术支撑, 计算机视觉^[1] 是使用计算机及其相关设备对生物视觉

的一种模拟, 其主要侧重在图像底层特征提取、视频分析等方面. 其中人脸识别^[2] 就是基于计算机视觉的

① 基金项目: 国家自然科学基金 (61572083); 教育部联合基金 (6141A02022610); 陕西省重点研发计划重点项目 (2018ZDXM-GY-047); 中央高校团队培育项目 (300102248402)

Foundation item: National Natural Science Foundation of China (61572083); Joint Fund of Ministry of Education (6141A02022610); Major Project of Key Research and Development Program of Shaanxi Province (2018ZDXM-GY-047); Team Incubation Project of the Central Universities of China (300102248402)

收稿时间: 2020-02-23; 修改时间: 2020-03-27, 2020-04-29; 采用时间: 2020-05-13; csa 在线出版时间: 2020-09-30

有关生物特征识别的研究技术之一。近年来人们广泛关注有关特征的构建和识别,例如 MSER^[3,4]、HOG^[5]、SIFT^[6]、LBP^[7,8]等。利用普通摄像头就能采集人脸数据模型,通过后台相应的算法就可对个人身份进行识别和确认,因而研究人脸识别本身就有着重要的意义,并且其核心技术的实现展现了弱人工智能向强人工智能的转化。

随着深度学习的发展,卷积神经网络(CNN)^[9]也为人脸识别提供了新的思路,其中经典的模型包含了 AlexNet^[10]、GoogleNet^[11]、VGG^[12],同时也包含了 ConvNet-RBM^[13]、DeepFace^[14]等深度网络人脸识别方法。其中 ConvNet 主要是直接从具有混合深度网络的面部对的原始像素中学习表明身份相似性的关系视觉特征的一种模型。DeepFace 则是使用显式 3D 人脸建模来重新审视对齐步骤和表示步骤,以便应用分段仿射变换,并从 9 层深度神经网络中得出人脸表示。这两种方法都是使用大量的人脸图像样本对神经网络模型进行训练,再手工选取中间层特征作为人脸的特征向量,最后对特征向量进行相似性度量得到人脸识别结果。

ConvNet 和 DeepFace 虽然提高了精度,但是仍存在一定缺陷,首先均需要人为选择特征向量,这将会导致算法具有不确定性,并且在大型数据库上训练神经网络模型时,中间层神经元个数会为了保持特征的完整性而不断增加,导致最后提取的特征向量维度变高,增加了计算成本。其次使用神经网络进行提取全局特征增加了计算机及网络的荷载,降低了人脸识别的速度。

针对以上问题,本文提出基于前后端交互的人脸识别系统,不仅提升了人脸识别的精度和鲁棒性,也有效缓解了计算机及网络的荷载和压力,同时系统的计算效率明显提升。本系统拥有 3 个创新性:

1) 使用了残差神经网络模型(ResNet)^[1],不用手工选取度量方法,而是将高维度的人脸图像映射为 128 维特征向量,再使用欧式距离进行人脸的相似性度量。

2) 采用了前后端交互式系统,提取特征值和特征值距离匹配分别运行于服务端和客户端,同时通信间采用码流传输格式,加快系统传输速率。

3) 提出了基于 GrabCut 的人脸兴趣区域(Region Of Interest, ROI)提取算法,先从系统输入帧图片中提取出人脸前景,排除了带有冗余和干扰信息的背景,提升了系统的计算效率和人脸识别精度。

1 基于前后端交互的人脸识别系统结构

本文提出了一个基于前后端交互的人脸识别系统,系统总体模块设计如图 1 所示,主要包含客户端模块、数据库以及服务端模块。客户端模块的功能是显示输入的视频流画面、收集待识别人脸图像、及对待识别人脸图像进行 GrabCut 图像预处理操作提取人脸前景,并且将从服务端传递来的人脸特征与已知数据库中的人脸特征进行匹配,显示人脸识别结果。本系统实现了客户端与服务端的通信,传递压缩后的人脸图像码流与人脸特征向量码流。服务端模块的功能是通过残差神经网络(ResNet)提取待识别人脸的特征。

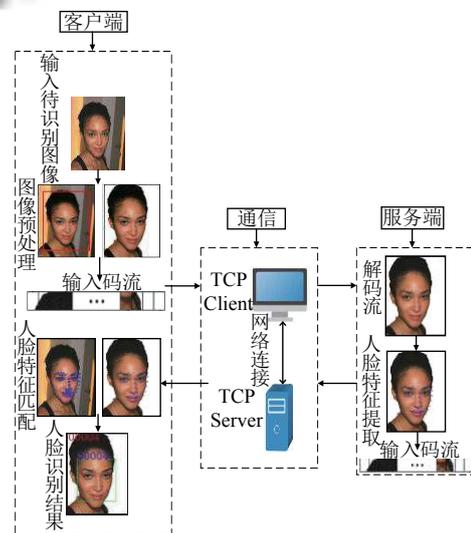


图 1 基于前后端交互的人脸识别系统框架图

本系统可以实现快速高效准确的人脸识别功能,首先在输入端对待识别人脸利用 GrabCut 算法进行图像预处理,获得精准的 ROI 区域,减少外界信息的干扰。其次将通用的人脸识别算法拆分到系统的不同模块中,其中将利用残差神经网络(ResNet)提取人脸特征的复杂操作单独布置在服务端,而简单的人脸特征距离匹配,即在 128 维空间中进行人脸相似性的计算布置在客户端。整个系统通信过程中传输压缩码流,通过多视频并行处理,可以得出人脸识别算法精度提高,并且系统的处理速度明显提升。最后本系统采用 SQL Server 对需要传输的人脸图像进行序列化数据存储,在通信过程中可以实现不同计算机间的数据准确接收,从而避免信息在传输过程中的丢失。

2 算法介绍

2.1 基于 GrabCut 的 ROI 提取算法

首先利用 GrabCut 算法对输入图像进行预处理操作, 排除了系统输入图片中的干扰信息, 得到准确的待识别人脸前景图像. 如图 2 所示, 在输入视频帧中设置人脸框, 图 2(a) 为固定的人脸框, 图 2(b) 为输入待识别人脸图像后的示意, 其中输入的待识别人脸图像可当作一组像素集合, 该像素集合可以分为区域项和边界项, 如图 2(a) 所示, 而区域项和边界项都需被划分为背景模型和前景模型, 最终图像预处理的目标就是获取前景, 即待识别人脸. 该算法结合人为设定的人脸框和 GrabCut 核心算法原理, 主要解决的就是判断某一个像素点属于前景模型还是背景模型, 一旦指定的人脸框内输入框选目标, 即待识别人脸, 则将人脸框外的像素集当作背景区域, 通过对人脸图像的区域像素点和边界像素点进行区分, 从而准确地提取人脸前景.

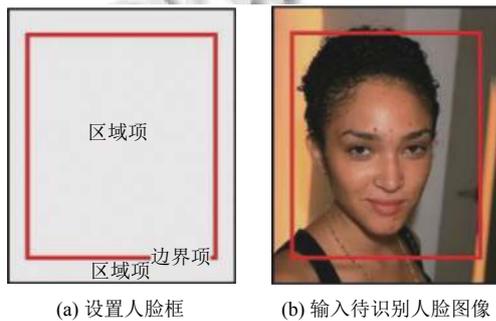


图 2 图像预处理示意图

如流程 图 3 所示, 该算法主要包含初始区和迭代区. 首先初始区得到描述前景的全协方差混合高斯模型 (GMM) 和描述背景的背景 GMM 的 3 个参数, 其次迭代区分配其余像素.

首先使用 k 个高斯分量 (一般取 $k=5$) 的 GMM 来对背景和前景进行建模, 其中背景包含正确的背景像素和可能是背景的像素, 前景包含正确的前景像素和可能是前景的像素, 由此得到背景模型 (BGM) 和前景模型 (FGM). 其次采用聚类算法将各自模型中的像素分到不同的模型样本中, 利用各自的样本来估计高斯模型内的参数, 如式 (1). 其中 π_i 表示每一个高斯分量的权重, μ_i 表示每个高斯分量的均值向量, Σ_i 表示协方差矩阵, x 表示 RGB 三通道向量, 这 3 个参数即为初始区中计算所得的初始化值.

$$D(x) = \sum_{i=1}^k \pi_i g_i(x; \mu_i, \Sigma_i), \sum_{i=1}^k \pi_i = 1 \text{ 且 } 0 \leq \pi_i \leq 1 \quad (1)$$

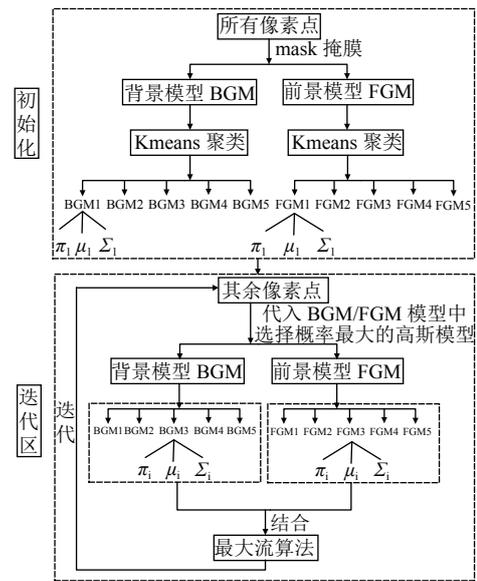


图 3 图像预处理算法流程图

图 3 中 BGM 和 FGM 中 5 个单高斯模型分量组合成了混合高斯模型, 一旦确定了各自 GMM 的 3 个参数, 其余未确定的像素中知道任一像素的 RGB 颜色值后, 代入前景或者背景 GMM, 从而得到该像素是前景或背景的概率, 重新分配未知的像素点, 这样每个模型的像素集合就已更新, 接着对估计的高斯模型重新更新参数, 最后将构建好的高斯模型中的区域项和边界项进行结合, 其中式 (1) 和式 (2) 是区域项像素的计算方法, 式 (3) 是边界项像素的计算方法, 式 (3) 中计算相邻两个相邻像素 m 和 n 之间 RGB 的欧式距离, 参数 β 由图像的对比度确定, 常数 γ 为 50. 最后采用最大流算法, 对未知的像素点重新分配并进行学习, 从而准确区分出每一个像素的归属, 得出其属于 FGM 还是 BGM.

$$g(x; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp \left[-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right] \quad (2)$$

$$V(\alpha, z) = \gamma \sum_{(m,n) \in C} [\alpha_n \neq \alpha_m] \exp -\beta \|z_m - z_n\|^2 \quad (3)$$

2.2 通信过程中的压缩与解压

传统的图像传输大多是基于像素访问进行传输, 传输的大小是图像的分辨率和图像的通道数的乘积. 就一般普通摄像头拍摄到的单通道灰度图像而言, 其分辨率是 640×480 时, 一次要传输的数据量大小是 $640 \times 480 \times 1 = 307\,200$ 个字节, 若是彩色三通道图像分辨率则为 $640 \times 480 \times 3 = 921\,600$ 个字节, 可见采用基于像

素访问的方式传输图像会对网络资耗费巨大, 从而造成视频的卡顿, 所以本文采用二进制码流的传输方式, 对图像进行压缩解压, 同时利用 TcpSocket 通信方式来传送图片二进制数据, 如图 4 所示, 通过压缩解压, 内存里的二进制图片数据和 OpenCV 中的图片格式相互转换, 本文的图片压缩了 20%, 在不影响人脸识别精度的基础上, 一定程度使得系统的运行速度提升, 也缓解了网络的压力。

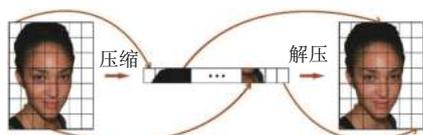


图 4 图像压缩与解压

2.3 基于 ResNet 的人脸识别算法

人脸图像可以用维度很大的矩阵来表征, 人脸识别就是对比对的人脸进行矩阵提取的操作, 本系统先检测到人脸, 之后通过 ResNet 将人脸生成一个 128 维的向量, 该节算法着重介绍人脸特征提取和人脸特征距离匹配。

本系统所使用的实时人脸识别方法是深度残差网络 ResNet-34, 并且利用了人脸识别模型和人脸关键点检测器。首先提前制作了已知人脸图像的特征值库, 其次根据人脸识别模型和人脸关键点检测器, 得到待识别人脸图像的特征值, 最后计算待识别人脸图像的特征值和特征值库的欧式距离, 确定识别结果, 如果两个特征向量的欧式距离小于给定的人脸特征匹配阈值 (本文阈值选定为 0.5), 则判断两幅人脸图像为相同对象, 反之则为不同对象。算法流程如图 5 所示。

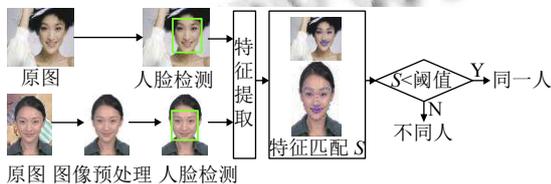


图 5 人脸识别算法流程

2.3.1 深度残差网络

2015 年由 Kaiming He 等 4 名华人共同提出的深度残差网络 ResNet (Residual neural Network)^[15], 通过增加 shortcut connection (Identity Map) 来直接连接深浅层网络, 使得梯度能够很好地传递到浅层, 其结构

可以加速神经网络的训练, 也大大提高了模型的准确度。

图 6 是一个 ResNet 的残差学习单元 (Residual Unit), 从残差学习单元可以看出 ResNet 是将学习目标进行改变, 将学习一个完整的输出 $H(x)$, 改变为学习输出和输入的差值 $H(x)-x$, 即残差。图 6 中假定某段神经网络的输入是 x , 期望输出是 $H(x)$, 如果直接把输入 x 传到输出作为初始结果, 那么此时需要学习的目标为 $F(x)=H(x)-x$ 。

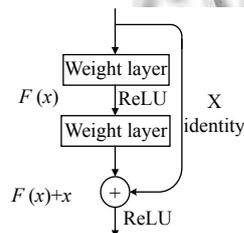


图 6 残差学习单元

2.3.2 人脸特征提取

本文采用 ResNet-34 模型提取人脸图像中具有代表性的特征向量, 在网络训练初始阶段, 每个特征向量表示为一个特征点, 所有的特征点以高斯分布的形式存在于特征空间中, ResNet-34 模型可以使网络在训练过程中, 属于同一人的特征点能够逐渐聚合, 而不属于同一人的特征点能够逐渐分离, 从而对人脸特征点进行识别。

本系统在服务端首先使用预测算子获取得到的人脸区域中的五官的几何点区域, 加载 68 个特征点的人脸标准特征模型; 其次加载 ResNet 模型参数, 获得 ResNet-34 训练模型。在系统输入图片或视频时, 如果存在人脸就使用 ResNet-34 的接口获取人脸特征向量。提取人脸特征点中将 ResNet-34 网络的 layer4 作为人脸的特征层, 此时的特征才是人脸的特征。提取的人脸的 68 个特征点 (例如眼、嘴等位置) 如图 7 所示。

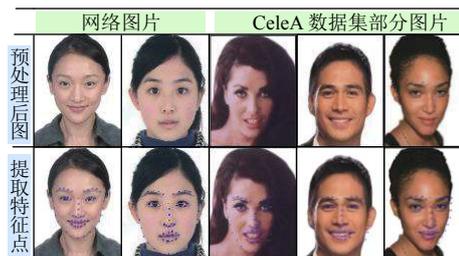


图 7 人脸 68 个特征点示意

2.3.3 人脸特征匹配

人脸特征匹配即距离匹配,是在获取特征向量之后使用欧式距离和本地已知人脸特征数据集的特征向量进行匹配,通过计算得到128维空间中两个点之间的真实距离.

如图8所示,二维坐标图中点 $A(x_1, y_1)$ 和 $B(x_2, y_2)$ 两者相似度为距离 $dist(A, B)$,可通过欧式距离公式计算,如式(4)所示.三维空间,点 $A'(x_1, y_1, z_1)$ 和 $B'(x_2, y_2, z_2)$ 两者相似度为距离 $dist(A', B')$,欧式距离计算如式(5)所示.128维向量空间上两点 $A''(x_{11}, x_{12}, x_{13}, \dots, x_{1(128)})$ 和 $B''(x_{21}, x_{22}, x_{23}, \dots, x_{2(128)})$ 的值,如式(6)所示.

$$dist(A, B) = \sqrt{\sum_{i=1}^2 (x_i - y_i)^2} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (4)$$

$$dist(A', B') = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (5)$$

$$dist(A'', B'') = \sqrt{\sum_{i=1}^{128} (x_{1i} - x_{2i})^2} = \sqrt{(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2 + \dots + (x_{1(128)} - x_{2(128)})^2} \quad (6)$$

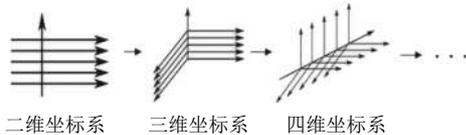


图8 空间欧氏距离示意图

本文采用的人脸识别技术是使用残差神经网络模型将人脸图像映射为128维特征向量,计算特征向量间的欧式距离,进而比较人脸图像的相似度.经过实际验证,本文设置人脸特征匹配阈值为0.5,如果两个特征向量的欧式距离小于给定的阈值0.5,则判断两幅人脸图像为相同对象,反之则为不同对象.与上述深度学习人脸识别方法相比,该方法的网络结构不受数据集影响,也不需要人为地统计选择适合地中间层和特征度量方法.

3 实验结果分析

3.1 实验条件

本实验的硬件支持是具有Intel(R) Core(TM) i7-8700型号的CPU和NVIDIA GTX2080Ti显卡支持的计算机,用来处理人脸识别算法中的提取人脸特征点

操作,摄像头使用TL-IPC42A-4,分辨率为1080P,采集频率15帧/秒.本实验使用了Qt 5.10.1软件,辅助环境为CUDA v10.0, OpenCV 3.4.0.

本文使用CeleA (CelebFaces Attribute,名人人脸属性数据集)数据集进行实验,该数据集包含了已做好特征标记的202599张图片.部分数据集图片如图9所示.

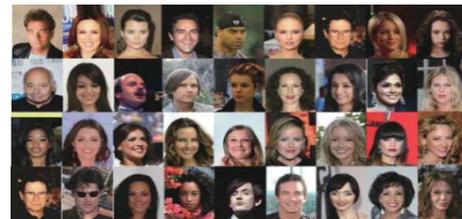


图9 CeleA部分人脸图像

3.2 算法测试

3.2.1 基于GrabCut的ROI提取实验

输入待识别人脸图像,经过图像预处理算法GrabCut处理后如图10所示.精准提取人脸前景区域,为提升人脸识别的算法精度作了铺垫,系统采用基于GrabCut的ROI提取算法有明显的优势,优势体现在两方面,人脸识别精度和计算效率.实验结果如表1和表2所示,在表2中,输入视频的帧率15帧/s,每隔30帧识别一次.



(a) 输入待识别人脸图像 (b) 图像预处理后

图10 图像预处理实验结果图

表1 人脸特征距离匹配值

j	使用ROI算法		精度提高(%)
	i=000004	未使用ROI算法 i=000004	
000000	0.876 967	0.836 678	4.8153
000001	0.901 978	0.858 739	5.0352
000002	0.966 262	0.919 000	5.1428
000003	0.604 038	0.575 066	5.0380
000004	0.101 967	0.097 076	5.0380
...
202597	1.034 830	0.988 836	4.6513
202598	0.884 029	0.838 815	5.3902
202599	0.657 645	0.623 595	5.4602
平均	/	/	5.1067

表 2 针对 ROI 算法系统耗时对比

参数	客户端			服务端			
	视频名称	帧数	$t(\text{min})$	$T(\text{min})$	$t/T(\%)$	$t(\text{s})$	$T(\text{s})$
视频1	240	0.2008	1.5743	12.75	0.4	1.57	74.52
视频2	360	0.3121	2.3614	13.22	0.6	2.36	74.58
视频3	150	0.1325	0.9839	13.47	0.25	0.91	72.53
视频4	270	0.2059	1.7711	11.63	0.45	1.74	74.13
视频5	570	0.4569	3.7389	12.22	0.95	3.65	73.97

表 1 中分别给出了在使用 ROI 提取算法和未使用 ROI 提取算法时姓名为 000004 的人与数据库中的人进行人脸特征距离匹配的结果,即人脸识别结果,最后给出了使用该提取算法对图像做预处理比未使用该算法时的人脸识别精度提高百分比值,当 000004 号人与数据库中 202 599 进行人脸特征距离匹配时,精度提高了 5.4602%,结果表明基于 GrabCut 的 ROI 提取算法在一定程度上提高了人脸识别的精度。

表 2 是分别计算了使用 ROI 算法在客户端和服务端的影响,在客户端中主要测试了 ROI 提取算法的耗时,与客户端总耗时进行了对比,实验结果如表 2 中 t/T 所示,其中 t 表示 ROI 提取算法耗时, T 表示客户端总耗时,当视频 1 输入系统时,ROI 提取算法耗时仅占客户端总耗时的 12.75%,结果表明采用基于 GrabCut 的 ROI 提取算法在客户端耗时比例不大,并不是影响系统计算效率的主要因素;在服务端对提取人脸特征点的计算效率进行了测试,在系统中,由于采用基于 GrabCut 的 ROI 提取算法主要是服务于服务端的提取特征点操作,所以实验对系统在没有使用 ROI 提取算法时服务端提取特征点的耗时,和系统使用 ROI 提取算法进行了耗时对比,计算出了使用 ROI 提取算法对提取人脸特征点的计算效率提高的百分比,计算结果如表 2 中 $(T-t)/T$ 所示,其中 t 表示系统采用 ROI 提取算法服务端耗时, T 表示系统未采用 ROI 提取算法时服务端的耗时,当视频 1 输入系统进行测试时,系统使用 ROI 提取算法时提取人脸特征点的计算效率提高了 74.52%,结果表明该 ROI 提取算法确实有效提升了系统中提取人脸特征点的计算效率。

3.2.2 基于 ResNet 的人脸识别实验

已知人脸数据库如图 11 所示,其中姓名为左上角数字,人脸图像以及人脸特征值如图所示,并且在数据库中采用通用唯一识别码 UUID (Universally Unique

Identifier) 来对样本进行编号,人脸图像及人脸特征值均用二进制数据表示以码流的形式在前后端交互式系统中传递信息。

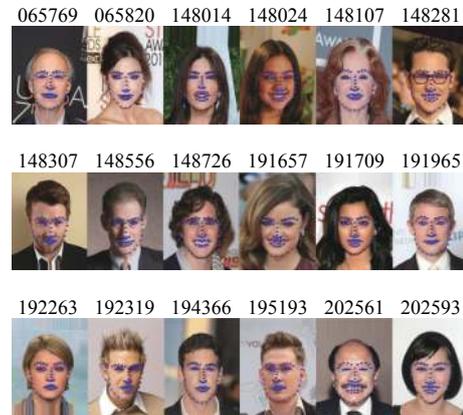


图 11 已知人脸数据库样本示意

本实验选取了 CeleA 中的 202 599 张图片,进行大样本测试的结果部分示意图如图 12 所示,图内小框即为人脸的 68 个特征点区域,下方字体所表示的是识别结果,上方字体则代表了已知人脸数据库的本人身份。若两个表示相同,即代表识别正确,反之则代表识别错误。就姓名为 000004 的人而言,与已知人脸数据库特征值比对结果如表 1 所示,其中 j 表示已知人脸数据库的人脸姓名, $i=000004$ 表示待识别人脸正确识别结果的姓名应为 000004,判断的条件为在该列中找最小数值,即为识别本人,数值计算公式如式 (6)。

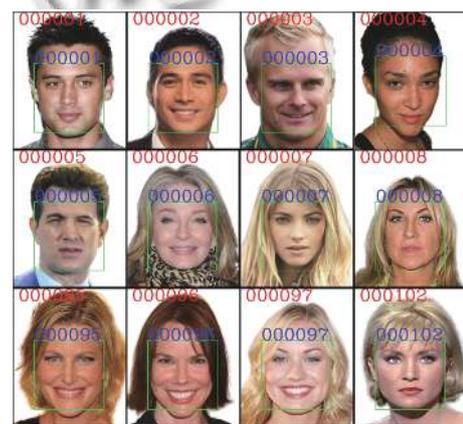


图 12 大样本人脸识别测试结果

实验中判断系统人脸识别算法的优劣,不仅需要精准率,即人脸识别正确的结果占识别结果的比例,同时需要考虑召回率,即人脸识别正确的情况占本身系统

输入中正确情况的比例,表3给出了精准率和召回率评判的4个参数 TP (True Positive), FP (False Positive), FN (False Negative) 和 TN (True Negative), 精准率 P (Precision) 和召回率 R (Recall) 的计算公式如式(7)和式(8)所示.

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

表3 评判参数

	人脸识别结果	人脸识别结果
输入帧中存在人脸	TP (正确人脸)	FP (错误人脸)
输入帧中不存在人脸	FN (识别出人脸)	TN (未识别出人脸)

表4是针对人脸识别算法的评判标准,评判指标包括,已知人脸数据库的图像总数,即样本总数,人脸识别测试结果精准率,召回率以及单帧图像识别平均耗时.由表4可知,在已知的具备人脸不同表情、不同背景、不同光照条件下的人脸数据库中,通过实验结合式(7)和式(8)计算可得,在20万余个样本的测试中,识别精准率高达99.84%,召回率达94.13%,该结果说明了本文的人脸识别算法具有较高的识别精度,同时极大地减少了误判、漏判等情况.其中由于存在部分严重遮挡的图像,导致仍有部分未能正确识别,同时单帧图像识别平均耗时达到0.067 s,实时性很高.

表4 人脸识别算法评判结果

样本总数	识别正确数	精准率	召回率	单帧平均耗时
202 599	202 284	99.84%	94.13%	0.067 s

3.3 系统测试

基于前后端交互的人脸识别系统通过分别将人脸识别核心算法布置在服务端和客户端中,其中服务端进行提取特征点,客户端进行特征点距离匹配,如表5所示,分别对交互系统和非交互系统分别进行测试,同时选用了5个视频进行测试来验证系统的各项处理耗时.

在表5中,输入视频的帧率15帧/s,每隔30帧识别一次,作为该连续30帧的人脸识别结果,前后端交互式系统是将人脸识别算法中的提取特征点和特征点距离匹配拆分在两端并发进行,而非交互式系统整个系统流程并未拆分.从表中交互式系统和非交互式系统的系统总耗时和平均耗时2.0147帧/s和2.5409帧/s

来看,明显交互式系统的要高速.交互式系统中的传输是在局域网内通过二进制形式传输,耗时可忽略不计,证明传输的码流以及图像预处理操作起到了作用.系统提高了网络的传输速率,首先,系统采用的GrabCut算法提取人脸ROI区域,减少了系统输入图像中的冗余和干扰信息;其次,将人脸识别算法中的人脸特征提取和人脸特征匹配分布于服务端和客户端,均衡了服务端和客户端两端的网络以及计算压力;最后,在本系统客户端、数据库、服务端进行数据传输时,图片均经过压缩,使得数据能够快速地通过网络传输.可见本系统有好的人脸识别效果和较快的系统处理速度.

表5 系统测试结果

参数	非交互式		前后端交互系统		
	视频名称	帧数	总耗时(min)	客户端(min)	服务端(s)
视频1	240	1.9853	1.5743	0.4	1.5743
视频2	360	2.978	2.3614	0.6	2.3614
视频3	150	1.2408	0.9839	0.25	0.9839
视频4	270	2.2335	1.7711	0.45	1.7711
视频5	570	4.7152	3.7389	0.95	3.7389
平均	/	2.0147	/	/	2.5409

4 结论

本文提出了一种基于前后端交互的人脸识别系统,该系统提出了一种基于GrabCut的ROI提取算法对视频帧进行预处理,实验结果显示可以精准地提取到ROI区域,对输入的待识别图像帧进行提取人脸特征点和特征点距离匹配操作,不仅具有准确的人脸识别效果,而且减轻了网络的压力和计算机的荷载,传输过程采用码流形式,加快了系统的传输速率.整个系统既提升了人脸识别的精度和鲁棒性,又提升了人脸识别系统的快速性.但目前发展的对于人脸识别特征提取以及相关的表情识别^[15,16]、图像预处理^[17,18]、网络的传输^[19]等技术对整个系统的优化也有帮助,因此在保证较高的识别效果和较快的系统处理速度的条件下又要增加新的实时性功能是本文需要进一步研究的问题.

参考文献

- 1 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 770-778.
- 2 张剑, 何骅, 詹小四, 等. 结合特征适配与拉普拉斯形变的

- 3 3D人脸重建. 中国图象图形学报, 2014, 19(9): 1349–1359. [doi: [10.11834/jig.20140912](https://doi.org/10.11834/jig.20140912)]
- 3 Neumann L, Matas J. Real-time scene text localization and recognition. Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA. 2012. 3538–3545.
- 4 Akbani O, Gokrani A, Quresh M, *et al.* Character recognition in natural scene images. Proceedings of 2015 International Conference on Information and Communication Technologies. Karachi, Pakistan. 2015. 1–6.
- 5 慕春雷. 基于HOG特征的人脸识别系统研究 [硕士学位论文]. 成都: 电子科技大学, 2013.
- 6 Lowe DG. Distinctive image features from scale-invariant key-points. International Journal of Computer Vision, 2004, 60(2): 91–110. [doi: [10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94)]
- 7 Chen D, Cao XD, Wen F, *et al.* Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification. Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA. 2013. 3025–3032.
- 8 Juefei-Xu F, Boddeti VN, Savvides M. Local binary convolutional neural networks. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 4284–4293.
- 9 Hasani B, Mahoor MH. Spatio-temporal facial expression recognition using convolutional neural networks and conditional random fields. Proceedings of the 12th IEEE International Conference on Automatic Face & Gesture Recognition. Washington, DC, USA. 2017. 790–795.
- 10 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, NV, USA. 2012. 1097–1105.
- 11 Szegedy C, Liu W, Jia YQ, *et al.* Going deeper with convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. 2015. 1–9.
- 12 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Proceedings of 2015 International Conference on Learning Representations. San Diego, CA, USA. 2015. 1482–1496.
- 13 Sun Y, Wang XG, Tang XO. Hybrid deep learning for face verification. Proceedings of 2013 IEEE International Conference on Computer Vision. Sydney, NSW, Australia. 2013. 1489–1496.
- 14 Taigman Y, Yang M, Ranzato M, *et al.* DeepFace: Closing the gap to human-level performance in face verification. Proceedings of 2014 IEEE International Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. 2014. 1701–1708.
- 15 袁姮, 王志宏, 姜文涛. 刚性区域特征点的3D人脸识别. 中国图象图形学报, 2017, 22(1): 49–57. [doi: [10.11834/jig.20170106](https://doi.org/10.11834/jig.20170106)]
- 16 梁华刚, 易生, 茹锋. 结合像素模式和特征点模式的实时表情识别. 中国图象图形学报, 2017, 22(12): 1737–1749. [doi: [10.11834/jig.170251](https://doi.org/10.11834/jig.170251)]
- 17 Yu WS, Hou ZQ, Wang P, *et al.* Weakly supervised foreground segmentation based on superpixel grouping. IEEE Access, 2018, 6: 12269–12279. [doi: [10.1109/ACCESS.2018.2810210](https://doi.org/10.1109/ACCESS.2018.2810210)]
- 18 Tang M, Gorelick L, Veksler O, *et al.* GrabCut in one cut. Proceedings of 2013 IEEE International Conference on Computer Vision. Sydney, NSW, Australia. 2013. 1769–1776.
- 19 冯建文, 董剑. 改进的TCP应用层协议在远程实验系统中的应用. 计算机应用与软件, 2018, 35(12): 154–158.