

应用深度学习的大姿态人脸对齐*

姜悦卉¹, 张倩¹, 王斌¹, 沈慧中¹, 黄继风¹, 严涛²



¹(上海师范大学 信息与机电工程学院, 上海 201418)

²(莆田学院 信息工程学院, 福建 莆田 351100)

通讯作者: 姜悦卉, E-mail: 1714861314@qq.com; 张倩, E-mail: qianzhang@shnu.edu.cn

摘要: 针对大姿态人脸对齐算法中的精确度低的问题,设计并实现了一种新的分层并行和多尺度 Inception-Resnet 网络来实现大姿态人脸对齐.首先,构建了一个四阶级联沙漏网络模型.该模型通过端到端的方式直接输入图像进行人脸对齐.其次,网络内部使用预先设定的参数进行采样和特征提取.最后,直接输出对应的人脸特征点提取图像以及同等人脸大小的二维坐标点绘制图,并将所提出的方法在 AFLW2000-3D 数据集上进行测试.实验结果表明,对于任意无约束的二维人脸图像,该方法的归一化平均误差为 4.41%.与传统方法相比,该方法输出的正脸姿态图像视觉质量高、保真度更强.

关键词: 人脸对齐;卷积神经网络;大姿态人脸;沙漏网络;端到端

中文引用格式: 姜悦卉,张倩,王斌,沈慧中,黄继风,严涛.应用深度学习的大姿态人脸对齐.软件学报,2019,30(Suppl.(2)):1-8. <http://www.jos.org.cn/1000-9825/19012.htm>

英文引用格式: Jiang YH, Zhang Q, Wang B, Shen HZ, Huang JF, Yan T. Large-pose face alignment based on deep learning. Ruan Jian Xue Bao/Journal of Software, 2019,30(Suppl.(2)):1-8 (in Chinese). <http://www.jos.org.cn/1000-9825/19012.htm>

Large-pose Face Alignment Based on Deep Learning

JIANG Yue-Hui¹, ZHANG Qian¹, WANG Bin¹, SHEN Hui-Zhong¹, HUANG Ji-Feng¹, YAN Tao²

¹(School of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai 201418, China)

²(School of Information Engineering, Putian University, Putian 351100, China)

Abstract: Aiming at the low accuracy of the large-pose face alignment algorithm, this paper designs and implements a new hierarchical parallel and multi-scale Inception-resnet network to achieve large-pose face alignment. Firstly, a four-class Hourglass network model is constructed. The model directly inputs images for face alignment in an end-to-end manner. Secondly, the network internally uses pre-set parameters for sampling and feature extraction. Finally, the corresponding face feature points are directly output. A two-dimensional coordinate point drawing of the image and the equivalent face size is extracted, and the proposed method is tested on the AFLW2000-3D data set. Experimental results show that the normalized average error of this method is 4.41% for any unconstrained two-dimensional face image. Compared with the traditional method, the positive face attitude image outputted in this paper has high visual quality and fidelity.

Key words: face alignment; convolutional neural network; large-pose face; hourglass network; end-to-end

近年来,深度学习在人脸识别领域的成功应用使得人脸识别准确率大幅度提升.然而在自然场景下的人脸识别依然存在许多难题和瓶颈.因此众多研究者开始着手研究影响人脸识别的因素,如人脸对齐、光照以及遮挡等.

在这众多影响人脸识别的因素中,人脸对齐——进行深层人脸分析的必要预处理过程,是阻碍人脸识别技术普及的瓶颈之一,因此对其技术进行突破具有非常重大的意义.人脸对齐是将人脸模型移动,并变形为图像,

* 基金项目: 国家自然科学基金(61741111)

Foundation item: National Natural Science Foundation of China (61741111)

收稿时间: 2019-08-17; 采用时间: 2019-11-04

从而提取面部像素的特征的过程.目前对于人脸对齐的研究可以分为两类:2D 人脸对齐和 3D 人脸对齐.由于人脸识别任务中对大姿态的要求越来越高,传统的 2D 人脸对齐方法无法满足大姿态图像对齐的精确度要求;目前 3D 方法被广泛应用于人脸对齐中,它是利用一个子空间对三维人脸进行建模,通过最小化图像与模型外观的差异来实现拟合,但是仍存在很多的问题有待解决.Kemelmacher-Shizerman 等人^[1]引入了平均三维人脸模型作为约束条件,并取得了不错的效果.然而,其对齐结果往往都接近平均模型,缺少个性化特征;Jourabloo 等人^[2]提出了三维人脸模型拟合方法,其目的是将三维形变模型(3D morphable model,简称 3DMM)拟合到 2D 图像上.该算法在稀疏的 3D 点分布模型的辅助下,模型参数和投影矩阵由级联线性或非线性回归量估计,可以对齐任意姿势的人脸,但是在恢复人脸的细节特征方面,其效果仍然不理想;随后 Jourabloo 等人^[3]又研究出新的方法,即使使用 3D 人脸建模,解决了大姿态下面部特征点定位的问题,将 2D 人脸形状作为 3D 人脸形状通过投影变化的结果,人脸特征点定位问题被转变为预测投影矩阵和 3D 人脸形状模型参数.尽管覆盖任意姿势效果有所改善,但是对齐的准确性仍受到线性参数化 3D 模型的限制.

2014 年前后,由于互联网的发展为海量数据的获取提供了便利,神经网络开始重新回归研究者的视野,香港中文大学的 Sun 等人^[4]首次提出将卷积神经网络(convolutional neural network,简称 CNN)用于原始人脸图像中的地标位置的回归,但该方法只检测 5 个关键点,且单张图片在 CPU 上检测速度过慢,网络结构复杂.Zhang 等人^[5]通过多任务 CNN 进一步将人脸对齐与属性分析相结合,以提升两项任务的执行效率,但此方法参数过多,导致计算量增大,成本大幅度提高.Zhu 等人^[6]通过 CNN 对图像进行密集的三维人脸模型拟合,在人脸侧面视图中合成大规模训练样本,解决了数据标注问题,但同样训练成本过高.Fan 等人^[7]提出两级神经网络级联的方式,实现了由粗到细的地标检测,但是反复迭代参数使得运行时间过长.当前,特征金字塔网络(feature pyramid network,简称 FPN)^[8]也因其丰富的特征而被广泛应用于最先进的人脸检测器中.然而,FPN 只是聚合了高、低输出层之间的层次特征映射,忽略了当前层信息.

综上所述,大多数 CNN 方法描述大姿态人脸的能力有限,因此限制了其在大姿态人脸对齐方面的应用.另外,由于生物特征的多样性和复杂性,尤其是当输入人脸图像单一、具有面部表情、存在特殊姿势、头部姿态角度大、图像清晰度不高、部分受遮挡或者光照强度不同时,同样会降低人脸对齐方法的准确率.因此,针对以上方法成本高、计算量大、人脸检测效果不好、对齐精度不高等不足,本文提出了基于沙漏神经网络的的大姿态人脸对齐,使用新颖的端对端的方式训练模型降低了模型的复杂度,有效地解决了大姿态以及自遮挡的约束问题,提高了传统人脸对齐系统的计算时间以及对齐精度.

1 大姿态人脸对齐方法

人脸对齐是许多人脸分析任务的必要预处理步骤,人脸对齐方法旨在定位诸如“眼角”“鼻尖”“下巴中心”的地标点^[9],基于该人脸地标点完成对人脸图像进行标准化.即使使用的算法不同,所有的人脸对齐方法包括深度学习和非深度学习方法处理流程步骤也是相同的,如图 1 所示.



Fig.1 Face alignment processing flow

图 1 人脸对齐处理流程

人脸对齐是建立在完成检测人脸的基础上,根据输入的人脸图像,经过形状初始化、特征提取、形状预测等步骤自动定位出面部关键特征点,即输入为人脸外观图像,输出为人脸特征点集合..

2 3D 人脸对齐方法

当下,2D 人脸图像及 3D 人脸对齐方法也是多种多样的,为其大致分类如图 2 所示.

传统 3D 人脸对齐方法大多是立足于图像信息,如基于图像亮度、边缘信息、线性透视、颜色、相对高度、视差等 1 种或多种信息建模技术.目前较为广泛应用的是基于模型的 3D 人脸对齐方法,3D 模型主要用三角网格表示.当下流行的模型有通用人脸模型、3DMM 和其变化模型,基于这些模型的 3D 人脸对齐算法,既有传统算法,也有目前流行的深度学习算法.

基于人脸模型的对齐方法存在一些固有的缺陷,首先,常用的基于地标的人脸模型均假定其地标是可见的,因此不适合侧面视图;其次,大姿态的脸部外观从正面视图到侧面视图变化更为明显;再次,必须人工猜测不可见地标,不能达到后续人脸识别的标准.

与人脸模型对齐方法相比,Kevan 等人^[10]规避了上述方法中模型呈现出的缺点,直接利用在大规模数据上训练的级联沙漏模型来预测 3D 面部地标并获得最先进的成果.沙漏网络从高分辨率到低分辨率再从低分辨率到高分辨率的设计结构使得处理数据时容量分布比先前方法更为对称,但是该算法目前还存在计算能力偏低的缺陷,平均每张图片处理时间为 29.98ms.在相同的硬件条件下,本文中提出了基于该级联沙漏神经网络的改进算法,对基本构造模块进行了改进,在保证主观处理精度稳步提高的情况下,平均单张图片处理时间降为 20.42ms,弥补了先前算法计算时间偏高的缺陷.本文算法为端到端的 3D 人脸对齐方法,从输入端的数据直接得到输出端的结果.即直接输入图片数据便可得到最终结果.利用反向传播算法进行误差调整,直到模型收敛或达到预期效果,整个流程均使用神经网络完成,不再分成多个模块处理.端到端最大的优点就是通过缩减人工预处理和后续处理,尽可能地使模型直接从原始输入到最终输出,使模型更自由地根据数据调节,增加模型的整体契合度.本文的人脸对齐算法流程如图 3 所示.

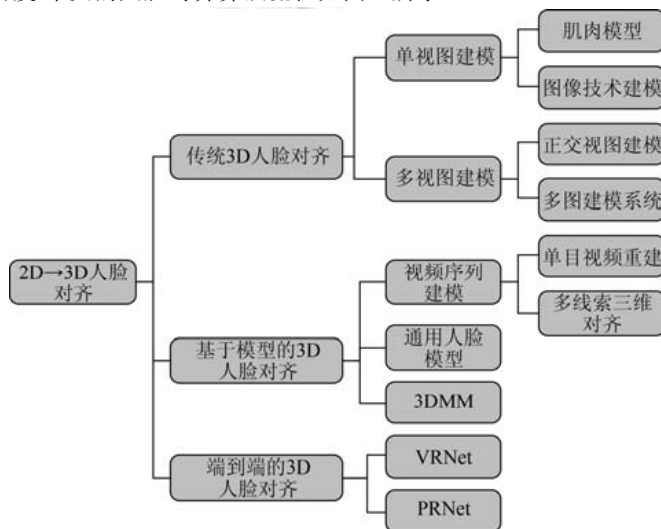


Fig.2 Face alignment methods

图 2 人脸对齐方法

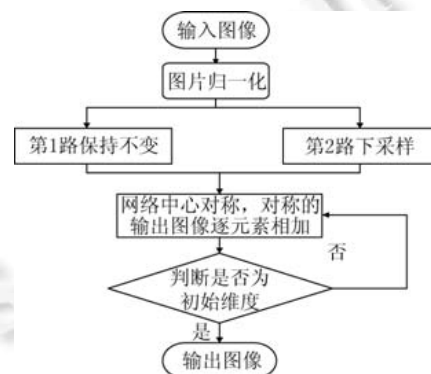


Fig.3 Face alignment algorithm flow

图 3 人脸对齐算法流程

由于初始输入图片大小不一,在进行算法处理前需要对图片进行归一化,维度小的图片使用黑色填充,最终归一化成像素大小为 256×256 的图片.随后归一化后的图片通过池化层进行下采样,使用近邻插值进行上采样.近邻插值算法可以通过缩放后图片中的任意一点求得其对应的原图片中的点.

2.1 基于级联沙漏网络的人脸对齐

2.1.1 Inception-Resnet 模块

沙漏网络模型的初始构造模块为 Resnet 模块,如图 4 所示,左边一列由 3 个核尺度大小不同的卷积层串联而成,第 1 层为大小是 1×1 的卷积核,第 2 层为大小为 3×3 的卷积核,第 3 层为大小为 1×1 的卷积核,其中卷积层之间插有 Batch Normalization 层和 ReLU 激活层.第 2 列为跳级路,只包含 1 个核为 1×1 的卷积层;所有卷积层的步长为 1,填充为 1,不改变数据尺寸,只对数据深度进行变更.为了提高沙漏模型的计算能力,保持沙漏模型的计算复杂度,本文使用并行和多尺度的 Inception-Resnet 块^[11]如图 5 所示,替换原始 Resnet 块如图 4 所示.

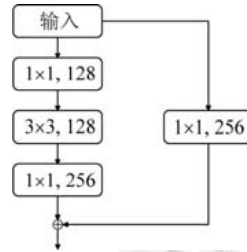


Fig.4 Resnet module

图 4 Resnet 块

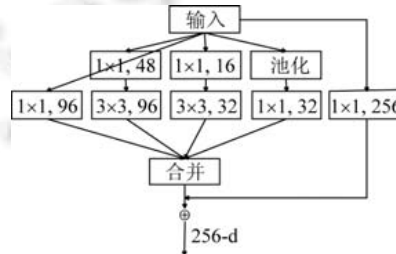


Fig.5 Inception-Resnet module

图 5 Inception-Resnet 块

Inception 结构的主要思想是找出最优的局部稀疏结构并将其覆盖为近似密集组合.如图 5 所示,Inception 模块呈层与层的结构,结构的最后一层进行相关性统计,将高相关性的聚合在一起.这些簇构成下一层的单元,与上一层的单元相连接.为了避免 3×3 的卷积核计算量过大导致特征图厚度过大的情况,本文在 3×3 的卷积核前和池化层后分别加上了 1×1 的卷积核起到了降低特征图厚度的作用.此外,本文直接在输入层接入 Resnet 模块,并与 Inception 模块合并结果堆叠,解决了随着网络深度的增加出现一种退化问题,并最终构成了 Inception-Resnet 模块.

2.1.2 级联沙漏网络模型

沙漏是一个对称的自上而下和自下而上的完全卷积网络.它兼顾了“bottom-up”(从高分辨率到低分辨率)和“top-down”之间的对称分布.在图片归一化之后,每张图片调整为 256×256 像素.Alejandro 等人^[12]用 Resnet 模块构成四阶沙漏网络模型,但计算量过大导致计算成本极高.Deng^[11]等人使用 Inception 模块构成两阶沙漏网络,但是计算精度不高.本文算法将以一个 7×7 、步长为 2 的卷积层开始.而后经过一个最大池化层将图片像素降为 64×64 ,该大小为 64×64 的图片即为算法中沙漏网络模型的实际输入图片.与传统的神经网络方法相比,降低输入图片的维度可以在性能下降不多的前提下,最大限度地减小网络体积.如图 6 所示,不同的大小代表不同维度的图像,立方体越大,代表图像维度越高,中间 3 层为特征提取层.图 6 中, a, b, c, d 每个小块都可以看作是经 Inception-Resnet 模块后所得结果, e, f, g 为特征提取层, a_1, b_1, c_1, d_1 是上采样后的图片,维度大小分别与 a, b, c, d 相同.

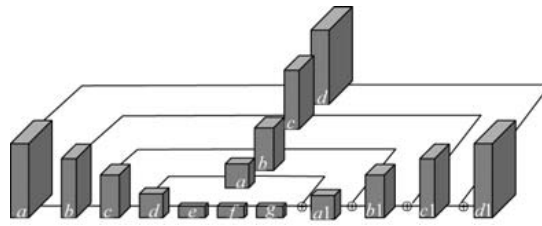


Fig.6 First-order hourglass network structure

图6 一阶沙漏网络结构

每个立方体都是一个 Inception-Resnet 块,沙漏网络设计深度的增加有助于增强前后区域.它包含全局形状推断,并可在局部观测模糊时增强鲁棒性.沙漏是若干个 Inception-Resnet 块构建的,根据阶数不同,有不同的复杂程度.沙漏网络可以公式化表示为

$$X_{n+1} = H(x_n) + F(x_n, W_n) \quad (1)$$

其中, x_n 和 x_{n+1} 是第 n 个单元的输入和输出, $H(x_n)$ 是恒等映射, W_n 是权重, F 是残差函数.沙漏网络输入图像在卷积路中,每经过一次 Inception-Resnet 块,尺度降为原来的一半,再经过特征提取,通过该方法可使网络在不同分辨率下提取出不同特征.

为使计算量尽可能地小并保证算法具有高度的灵活性,本文使用了四阶级联的沙漏网络,即由 4 个一阶级联的沙漏网络联接而成,采用这种结构的优势是:由高维度向低维度提取特征,并反复回归特征点,可获取最接近真实地标的特征点.该沙漏网络模型可以理解为:每次降采样之前,分出上半路保留原尺度信息;每次升采样之后,和上一个尺度的数据相加;两次降采样之间,使用 3 个 Inception-Resnet 模块提取特征;两次相加之间,使用一个 Inception-Resnet 模块提取特征.

3 损失函数与模型训练策略

训练的模型好坏直接影响到人脸对齐的效果,训练模型时须在正则化参数的同时最小化误差.最小化误差是为了让模型拟合训练数据,而正则化参数是防止模型过拟合训练数据.因此需要选择合适的损失函数,尽可能地让训练的神经网络模型拟合测试样本集.由于神经网络参数过多,会导致模型复杂度上升造成过拟合,因此须在保证模型“简单”的基础上最小化训练误差,这样得到的参数才具有良好的泛化性能,而模型“简单”就是通过正则化函数来实现的.本文使用以下目标函数训练模型:

$$\omega^* = \arg \min \sum_i L(y_i, f(x_i; \omega)) + \lambda \Omega(\omega) \quad (2)$$

其中,第 1 项 $L(y_i, f(x_i; \omega))$ 用于衡量模型对第 i 个样本的预测值 $f(x_i; \omega)$ 和真实的标签 y_i 之前的误差.由于模型要拟合训练样本,因此需最小化第 1 项的值,即尽量拟合训练数据.但训练的最终目标是降低模型的测试误差,所以需要使用参数 ω 的正则化函数 $\Omega(\omega)$ 来约束模型.因 L2 范数强大的抗过拟合能力,本文采用 L2 范数作为正则化函数以此提升模型的泛化能力,具体形式如下.

$$\Omega(\omega) = \frac{1}{2} \|\omega\|_2 \quad (3)$$

本文使用 Pytorch 框架进行训练,初始学习率为 10^{-4} ,batch 大小为 8,经过 15 次完整循环迭代后,将学习率降低到 10^{-5} ,再经过 15 次迭代后将学习率降为 10^{-6} ,整个训练共进行 40 次迭代.训练中还应用了随机增强如翻转、增加噪声、随机遮挡等方式来增加训练样本,提高模型的泛化能力.

4 实验

4.1 评价指标

由于传统用于人脸对齐的度量是基于双眼距离归一化的点对点欧基里德距离,然而,对于眼距非常小的轮

廓面,这种误差度量是有偏差的.因此,本文使用归一化平均误差(normalized mean error,简称 NME)^[15]进行人脸对齐精度度量,其定义为

$$NME = \frac{1}{N} \sum_{k=1}^N \frac{\|x_k - y_k\|_2}{d} \quad (4)$$

其中, x_k 表示给定的地面真实地标, y_k 表示相应的预测, d 是地面真实边界框的平方根,计算公式为 $d = \sqrt{w_{bbox}} \times \sqrt{h_{bbox}}$, 其中 w_{bbox}, h_{bbox} 分别是人脸检测边界框的宽和高.

4.2 实验结果与分析

本文使用 Sony Windows10 x64(6GB 内存,500GB 硬盘)电脑搭建实验环境,在 Pycharm 中配置 Pytorch 作为实验平台,依次进行实验及检验全部过程.本文方法用 Python 语言实现,选定一张输入图片输入进网络中,输出结果为人脸检测图像并在三维坐标系中画出地标点.表 1 为现阶段最流行的图像数据集及各自的主要特征.

Table 1 Image data set and main features

表 1 图像数据集及主要特征

数据集	规格	姿态	维度	合成
300-W	4 000	$[-45^\circ, 45^\circ]$	2D	否
300W-LP-2D	61 225	$[-90^\circ, 90^\circ]$	2D	是
300W-LP-3D	61 225	$[-90^\circ, 90^\circ]$	3D	是
AFLW20003D	2 000	$[-90^\circ, 90^\circ]$	3D	否
300-VW	218 595	$[-45^\circ, 45^\circ]$	3D	否

为了验证本文人脸对齐方法在大姿态下的效果,本文给出了基于户外标记人脸数据集(annotated facial landmarks in the wild,简称 AFLW)的实验结果.AFLW 人脸数据库是一个包括了各种姿态、表情、光照、种族等因素影响的图片的数据集,其中大部分的图片都是彩色的,一小部分是黑白的.该数据库适合用于人脸识别、人脸检测、人脸对齐等方面的研究,具有很高的研究价值.

在进行实验之前,先选取 AFLW 中合适的人脸图像作为预处理输入,用各种不同的算法对该图像进行人脸地标检测比对,具体操作见表 2.其中 RCPR 是一种对遮挡有鲁棒性的方法,具备处理自咬合的潜力.ESR、SDM、LBF 和 CFSS 是流行的基于级联回归的方法.

Table 2 Face alignment algorithm results comparison

表 2 人脸对齐算法结果对比

对齐方法	AFLW dataset (68 pts)			Mean
	$[0^\circ, 30^\circ]$	$[30^\circ, 60^\circ]$	$[60^\circ, 90^\circ]$	
LBF ^[13]	6.17	16.48	25.90	16.19
ESR ^[14]	4.38	10.47	20.31	11.72
CFSS ^[15]	3.44	10.90	24.72	13.02
RCPR ^[16]	4.16	9.88	22.58	12.21
SDM ^[17]	3.56	7.08	17.48	9.37
RMFA ^[18]	4.96	8.44	13.93	9.11
3DDFA ^[6]	3.78	4.54	7.93	5.42
CMHM ^[11]	3.43	4.14	6.57	4.71
3D-FAN ^[10]	3.47	4.26	6.78	4.82
Ours	3.33	3.97	5.92	4.41

为了验证本文算法的有效性,整个 AFLW 数据集被分为 3 个角度范围子集: $[0^\circ, 30^\circ]$, $[30^\circ, 60^\circ]$ 和 $[60^\circ, 90^\circ]$.由于目前基于大姿态人脸研究很少,实验选择了一些已开源代码的方法进行验证,并且使用 NME 进行估计度量,验证数据结果见表 2.

通过与上述方法对比可知,对于任意姿势,本文的方法均有所提高,尤其在 $[60^\circ, 90^\circ]$ 效果提升最为明显,归一化均方误差降为 5.92%,整体归一化均方误差降为 4.41%,比现有最优对齐方法降低了 8.9%,进一步说明了本文方法的准确性.最终的输出结果如图 7 所示,由最终输出结果可以看出,本文的算法在各姿态上均有良好的对齐效果.

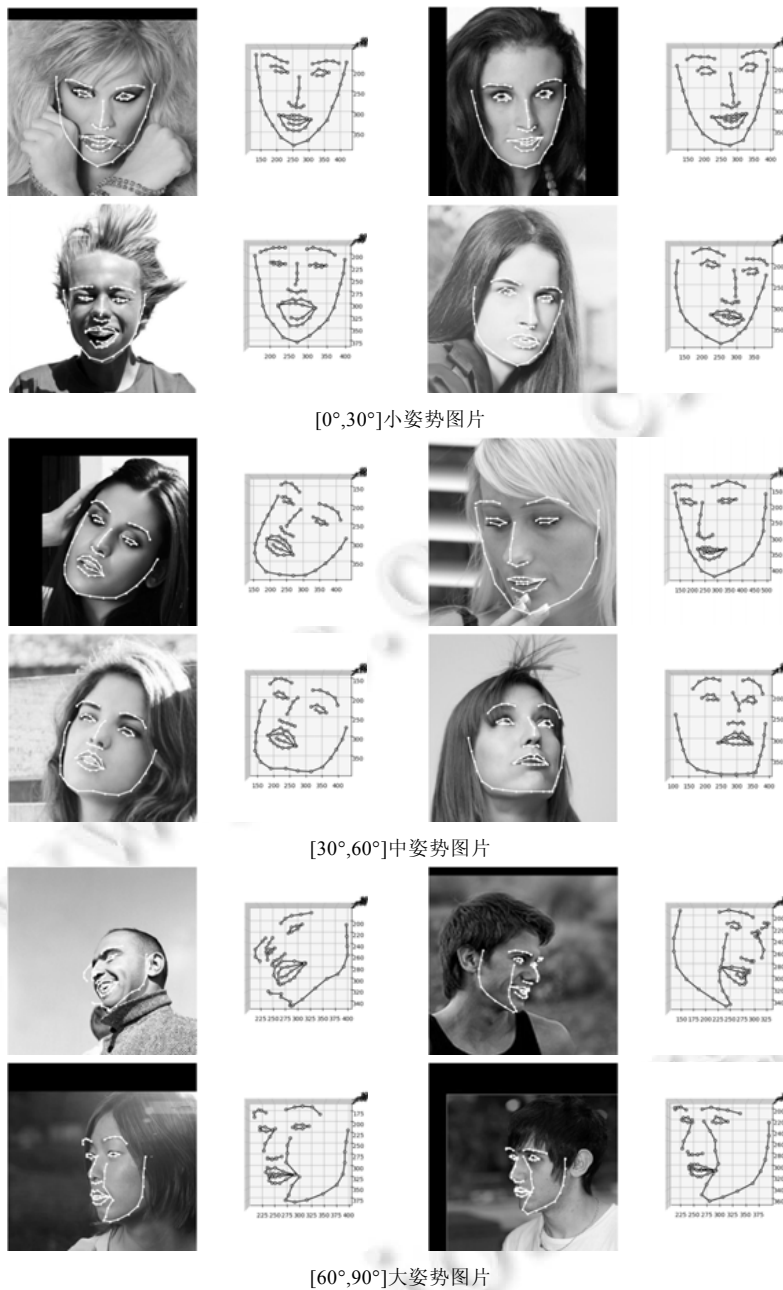


Fig.7 Face alignment results

图7 人脸对齐结果

5 结论

本文针对大姿态环境下的人脸对齐,提出了一种使用沙漏网络模型实现人脸对齐的方法,通过使用深层卷积神经网络直接回归人脸特征点,减小了姿态造成的影响,实现了大姿态环境下的人脸对齐,使用归一化平均误差函数评估对齐精度.实验结果表明,该方法与现有的人脸对齐方法相比,在精度上有明显的优势,但本文提出的方法对于外界事物遮挡下(如手、墨镜、面具等)的人脸难以实现精确对齐.下一步将对该问题进行深入研究.

References:

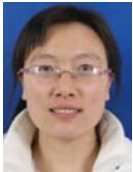
- [1] Kemelmacher-Shlizerman I, Basri R. 3D face reconstruction from a single image using a single reference face shape. *IEEE Trans. on Software Engineering*, 2011,33(2):394–405.
- [2] Jourabloo A, Liu X. Pose-invariant 3D face alignment. In: *Proc. of the 2015 IEEE Int'l Conf. on Computer Vision (ICCV)*. IEEE, 2015. 3694–3702.
- [3] Zhang HW, Li Q, Sun Z, *et al.* Combining data-driven and model-driven methods for robust facial landmark detection. *IEEE Trans. on Information Forensics and Security*, 2018,13(10):2409–2422.
- [4] Sun Y, Wang X, Tang X. Deep convolutional network cascade for facial point detection. In: *Proc. of the 2013 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2013. 3476–3483.
- [5] Zhang Z, Luo P, Loy CC, Tang X. Facial landmark detection by deep multi-task learning. In: *Proc. of the Computer Vision (ECCV) 2014*. Springer-Verlag, 2014. 94–108.
- [6] Zhu X, Lei Z, Liu X, Shi H, Li SZ. Face alignment across large poses: A 3D solution. In: *Proc. of the 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2016. 146–155.
- [7] Fan HJ, Zhou E. Approaching human level facial landmark localization by deep learning. *Image & Vision Computing*, 2016,47:27–35.
- [8] Hong R, Cheng WH, Yamasaki T, *et al.* Small object detection using deep feature pyramid networks. In: *Advances in Multimedia Information Processing—PCM 2018*. 2018. 554–564.
- [9] Jourabloo A, Liu X. Pose-invariant face alignment via CNN-based dense 3D model fitting. *International Journal of Computer Vision*, 2017,124(2):187–203.
- [10] Yuen K, Trivedi MM. An occluded stacked hourglass approach to facial landmark localization and occlusion estimation. *IEEE Trans. on Intelligent Vehicles*, 2017,2(4):321–331.
- [11] Deng JK, Zhou YX, Cheng SY, *et al.* Cascade multi-view hourglass model for robust 3D face alignment. In: *Proc. of the 2018 13th IEEE Int'l Conf. on Automatic Face & Gesture Recognition (FG 2018)*. 2018. 399–403.
- [12] Newell A, Yang KY, Deng J. Stacked hourglass networks for human pose estimation. In: *Proc. of the Computer Vision—ECCV 2016*. Springer-Verlag, 2016. 483–499.
- [13] Lai HJ, Xiao SG, Yan P, *et al.* Deep recurrent regression for facial landmark detection. *IEEE Trans. on Circuits & Systems for Video Technology*, 2018,28(5):1144–1157.
- [14] Peng X, Feris RS, Wang XY, *et al.* RED-net: A recurrent encoder-decoder network for video-based face alignment. *Int'l Journal of Computer Vision*, 2018(2):1–17.
- [15] Liu F, Zeng D, Li J, *et al.* On 3D face reconstruction via cascaded regression in shape space. *Frontiers of Information Technology & Electronic Engineering*, 2017,18(12):1978–1990.
- [16] Barbu A, Lay N, Gramajo G. Face detection with a 3D model. *Academic Press Library in Signal Processing*, 2014,21(4):237–259.
- [17] Ranjan R, Sankaranarayanan S, Bansal A, *et al.* Deep learning for understanding faces: Machines may be just as good, or better, than humans. *IEEE Signal Processing Magazine*, 2018,35(1):66–83.
- [18] Chen FX, Liu F, Zhao QJ. Robust multi-view face alignment based on cascaded 2D/3D face shape regression. In: *Proc. of the Chinese Conf. on Biometric Recognition*. Springer-Verlag, 2016. 40–49.



姜悦奔(1995—),女,江苏东台人,硕士生,主要研究领域为视频编码,模式识别.



沈慧中(1995—),女,硕士生,主要研究领域为模式识别.



张倩(1983—),女,博士,副教授,主要研究领域为视频和图像信息处理.



黄继风(1963—),男,博士,教授,主要研究领域为模式识别,数字图像处理,生物信息学,视频图像识别.



王斌(1987—),女,博士,副教授,主要研究领域为图像处理,大数据.



严涛(1981—),男,博士,主要研究领域为图像处理.