Selective Multilayer Fusion Convolutional Neural Network for Facial Age Estimation*

LI Chao[1*], *TONG Lin*[2], *LIU Yonghui*[3], *ZHU Daomeng*[1], *MA Yanchen*[1]

(1.*Guizhou Vocational Technology College of Electronics & Information, Department of Mechanical and Electrical Engineering, Kaili Guizhou* 556099, *China*;
2.*School of Physics and Electrical Engneering Liupanshui Normal University, Liupanshui Guizhou* 553004, *China*;
3.*School of Information Engineering, Zhejiang University of Technology, Hangzhou Zhejiang* 310014, *China*)

Abstract: Facial age estimation using traditional neural network algorithm can bring gradient dispersion or gradient explosion, which leads to the risk of network degradation. Hence, a Selective multilayer fusion convolutional neural network structure is proposed. Firstly, the SMLF-Net with SE-like (Squeeze-and-Excitation Networks) structure is designed to enhance the model's nonlinearity and feature recalibration. Secondly, in order to improve the classification accuracy of deep network model, a structure similar to RE (ResNets) is utilized to solve the network degradation problem. Finally, the network structure is selected dynamically according to the moved step of convolution kernel to avoid excessive network parameters. The proposed method builds RE and SE structures to extract high-dimensional, middle-dimensional and low-dimensional features of human faces. Meanwhile, it uses Batch Normalization to reduce network training difficulty. Imdb-wiki data set after Asian data expansion is used to train the proposed methed and tested on Raspberry Pi. The results show that the average absolute error (MAE) of SMLF-NET is 3.09. This result is better than Google InceptionNet with the average absolute error 3.32 and ShuffleNet with the average absolute error 3.54. The results demonstrate adequately the advance and effectiveness of the proposed method.

Key words: deep learning; convolutional neural network; face detection; activation function; age estimation
EEACC: 6150P; 7230 **doi**: 10.3969/j.issn.1004-1699.2020.09.014

面向人脸年龄估计的选择性多层融合神经网络*

李 超[1*], 童 林[2], 刘永辉[3], 朱道萌[1], 马延臣[1]

(1.贵州电子信息职业技术学院机电工程系, 贵州 凯里 556099; 2.六盘水师范学院物理与电气工程学院, 贵州 六盘水 553004;
3.浙江工业大学信息工程学院, 浙江 杭州 310014)

摘 要: 针对使用传统神经网络算法进行人脸年龄估计可能会带来梯度弥散或梯度爆炸引起网络退化的风险, 提出了选择性多层融合的卷积神经网络结构(SMLF-Net, Selective multilayer fusion convolutional neural network)。首先在 SMLF-Net 设计中融合了类似 SE(Squeeze-and-Excitation Networks)结构增强了模型的非线性和特征重标定, 其次为了提高深层网络模型的分类精度, 引入了类似 RE(ResNets)结构解决网络退化问题, 最后根据卷积核移动步长动态选择网络结构避免网络参数过多。该方法通过构建 RE 与 SE 结构提取人脸的高维、中维、低维特征, 并使用 BN(Batch Normalization)操作降低网络训练难度。结果在扩充亚洲数据后的 IMDB-WIKI 数据集上对本文方法进行训练, 并在树莓派上进行测试, 结果显示 SMLF-Net 方法能够取得平均绝对误差(MAE)为 3.09 的估计精度, 优于 Google InceptionNet 方法(MAE)为 3.32 和 ShuffleNet 方法(MAE)为 3.54, 充分证实了本文所提方法的先进性和有效性。

关键词: 深度学习; 卷积神经网络; 人脸检测; 激活函数; 年龄估计
中图分类号: TN929 **文献标识码**: A **文章编号**: 1004-1699(2020)09-1313-07

项目来源: 贵州省科技厅联合基金项目(黔科合 LH 字[2014]7456 号); 教育部高等学校大学物理课程教学指导委员会高等学校教学研究项目(DJZW201934xn)
收稿日期: 2020-05-05 **修改日期**: 2020-07-23

近年来,随着人工智能技术的发展,人脸年龄属性存在很多实际应用价值,能够与人脸检测,行人再识别等技术结合,应用在人工智能机器人、智能安防、人物行为预测、智能客流量统计系统[1]。虽然绝大部分人的面部特征随着年龄的变化趋势相似,但在实际生活中人的面部年龄表现受生活方式、区域、种族、遗传基因等外在和内在的因素影响[2],导致同一年龄的人面部外观上相差很大。而且摄像头采集到的图像容易受到光线、角度及清晰度的影响,这使得研究人员不断加深网络以提取深层特征。但庞大的网络需要进行大量的计算,因此对多张人脸实时预测年龄的硬件要求十分高。提高人脸年龄预测的模型效率及速度仍是计算机视觉领域的一个极具挑战的任务。

卷积神经网络最早由 Yann LeCun 教授提出[3],主要是解决目标分类问题,其中的局部连接,权值共享,空间和时间上的采样使得在目标物体检测与识别上具有扭转不变性,并且在物体的分类上具有很好的鲁棒性。在二十一世纪后,随着深度学习理论的提出和计算机计算速度的提高,卷积神经网络取得了飞速发展,并被大量应用在计算机视觉、自然语言处理等领域。尤其是在图像处理领域[4]。深度学习是一种特征学习,能够把原始数据由简单的非线性处理转变为更高层次的、更加抽象的表达。和传统神经网络相比,深度学习能够通过无监督方式逐层提取图像的高层次特征[5],从而得到更能准确表示图像的本质特征。受到深度学习在人脸识别中良好的表现,引发了许多研究者的兴趣。Wang 等[6]人设计一个具有 6 层卷积神经网络,其中第 2 到 5 层特征图谱包含 2 个卷积层和两个池化层,将其首尾相连构成一个向量后应用流形学习的方法做降维处理以此作为人脸图像表征送入到 SVM、SVR、PLS(partial least squares)与 CCA(canonical correlation analysis)进行最终分类或做回归预测。Levi 等人[7]构造了一个 5 层的类似于 Alexnet 的卷积神经网络,在 Adience 数据集上进行年龄分类。Yi 等人[8]先将人脸图像划分为不同的尺寸,接着在其内部划分出多个局部对齐块(local aligned patches),最后将得到的块输送到有 4 层卷积神经网络中。这样一方面在增加了样本空间的同时也增加了网络的拟合(fitting)能力。Niu 等人[9]把年龄标签当作连续的数值,让一个具有 5 层卷积神经网络的最后一层神经元个数为 1,这样就得到年龄的具体数值。随后,构造类似 OH-ranker 的排序方法,唯一的区别是将 SVM 替换成一个 5 层的神经网络,这样构造结果也是最好的年龄估计方法之一。

综上所述虽然卷积神经网络的参数相比传统网络减少了极多。但为了提取高维特征,需要不断增加网络深度,因此卷积神经网络一般都有很多层[10],这带来的结果是网络参数十分庞大。庞大的网络参数势必会带来大量的计算,容易造成网络过拟合,给网络训练带来很大的困难。而且随着网络层数的加深还会出现梯度弥散或梯度爆炸[11]而造成网络退化。为了解决此问题,文中提出了基于深度学习的选择性多层融合的卷积神经网络人脸年龄识别算法。首先对网络训练的数据集进行扩充及数据增强处理。其次构建所提出的选择性多层融合的卷积神经网络结构并使用 LeakyRelu 作为激活函数,网络最后的特征输出通过 softmax 进行分类并加权求和得到预测年龄。

1 基于 SMLFnet 网络的人脸年龄估计方法

本文提出的基于 SMLFnet 网络的人脸年龄估计方法流程如图 1 所示,主要分为两个部分包括图像预处理和基于 SMLFnet 网络的年龄估计。下面对这两部分详细介绍。

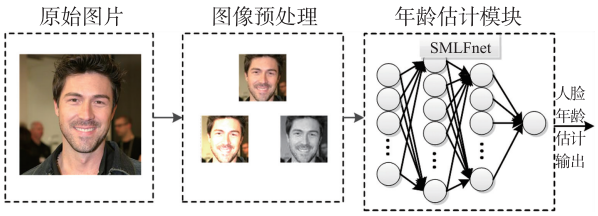


图 1 基于 SMLFnet 网络的人脸年龄估计算法

1.1 图像预处理

人脸年龄估计一般需要和人脸检测配合使用。待处理的图片先经过人脸检测模块对人脸进行定位,将定位后的人脸进行过滤(模糊过滤,大小过滤),之后对人脸进行对齐,裁剪,最后通过年龄估计模块对人脸进行特征输出、softmax 分类输出人脸年龄。

在进行网络训练之前通过对数据集进行数据增强处理来丰富图像数据,增强网络的泛化性。为了避免训练图片中人脸大小不统一造成网络学习困难,需要对数据进行规范化处理,主要将图片中的人脸区域扩充为 1.2 倍并进行图片裁剪。在图 2(a)为原始图片,经过裁剪后得到图 2(b)。对裁剪后的照片进行随机亮度增强,亮度值取自上下极限值之间,增强模型对人脸图片的明暗光线的适应性,图 2(b)经过区间随机增加亮度后得到图 2(c)。对图片进行灰度化处理可以增强模型对灰度图片处理能力同时也能增强模型对图片本质特征的提取。对图 2(b)进行灰度

处理后得到图 2(d)。此外,文中也随机对数据集随机进行旋转、左右镜像、高斯模糊等处理。

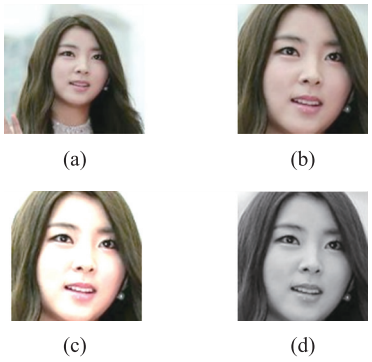


图 2　图片预处理

1.2　SMLFnet 网络算法

1.2.1　网络结构

参考 Google 团队提出的高效率深度可分离卷积减少网络参数的方法[12],解决本文所提网络因增加网络深度而带来参数量增加的问题。文中实现卷扩张压缩是先通过卷积核为 1×1 的卷积对网络进行扩张(提升 channel),并对该层输出进行 BN(Batch Normalization)操作,经过 BN 后对结果进行 LeakyRelu 激活函数激活。再通过一个卷积核为 3×3 的卷积层,再次对输出进行 BN 和 LeakyRelu 激活函数操作。最后通过卷积核为 1×1 的卷积层对网络输出通道进行压缩。图 3 为网络结构图,其中 t 为扩充倍数,s 为卷积核移动步长。

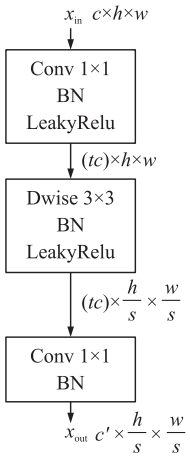


图 3　Bottleneck 结构图

图 4 为 SE 网络结构图,该结构的输入为原始网络的输出,是一个对网络进行优化的结构。当上层网络输出特征图为 $c\times h\times w$ 时,经过 Globalpooling 将特征图压缩为 $c\times 1\times 1$ 的特征图,然后经过一个全链接层将通道数进行压缩,压缩参数为 r,输出特征图的大小为 $(c/r)\times 1\times 1$;该全链接层的输出经过 Relu 激活函数对输出的特征图进行非线性增强;激活层的输出再次经过全链接层将输出通道进行扩充为 c,该层的输出特征图为 $c\times 1\times 1$ 再经过 Sigmoid 函数将特征信息归一化,最后通过 Scale 操作实现对原始网络输出特征进行特征重标定,起到特征增强的作用。

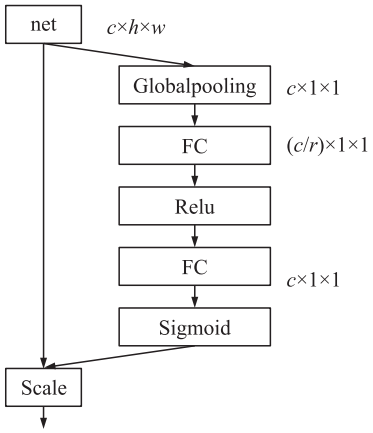


图 4　SE 网络结构图

由于深层网络能够提取图像的特征越抽象越具有语义信息,相应的对图像的分类也会更准确。但更深的网络可能会带来梯度弥散或梯度爆炸带来网络退化的风险。为了避免文中网络出现上述问题,提出了选择性多层融合的卷积神经网络结构(SMLF-Net, Selective multilayer fusion convolutional neural network),在 SMLF-Net 设计中融合了类似 SE(Squeeze-and-Excitation Networks)[13]结构增强了模型的非线性和特征重标定,同时加入类似 RE(ResNets)结构解决网络退化问题[14],提高深层网络模型的分类精度。为保证引入 Residual block 层不会给网络增加太多参数,对于文中的网络模型仅在输入通道和输出通道一致的卷积结构中加入 RE 网络结构,图 5 为 bottleneck 加入 RE 和 SE 后的 SMLF-Net 结构。

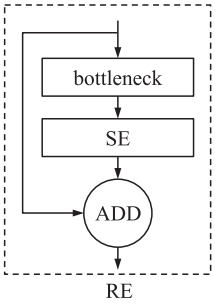


图 5　bottleneck 与 RE、SE 融合的 SMLF-Net 结构

表 1 为 SMLF-Net 网络结构的具体配置。其中,C 表示通道数,t 表示通道扩充倍数,RE 表示使用残差结构,SE 表示使用压缩激励结构来增强网络的特征筛选能力,s 为卷积核移动的步长,R 表示该层网络重复使用的次数。第一层采用卷积核为 3×3,步

长 s=2 的卷积。DW1 结构为一个卷积核 1×1 的卷积和一个卷积核为 3×3 的可分离卷积组成并进行 BN 和 LeakyRelu 激活函数操作。DW2 为一个卷积核 7×7 的可分离卷积和一个卷积核为 1×1 的 conv 组成并进行 BN 和 LeakyRelu 激活函数操作。在 $s=1$ 的 bottleneck 卷积层中使用残差与 SE 结构,在 $s=2$ 的 bottleneck 卷积层中仅使用 SE 结构。全链接层使用的网络输出为 66(年龄为 65 及以上按同一类处理),最后是 softmax 将全链接层输出的 66 个特征转化为相应的概率值再对应的年龄进行加权求和。

表 1 整体 SMLF 网络结构配置

Input	Operator	C	t	RE	SE	s	R
112×112	Conv1	32				2	1
56×56	DW1	32			√	1	1
56×56	bottleneck	64	2		√	2	2
28×28	bottleneck	64	2	√	√	1	3
28×28	bottleneck	128	4		√	2	1
14×14	bottleneck	128	2	√	√	1	4
14×14	bottleneck	256	4		√	2	1
7×7	bottleneck	256	2	√	√	1	2
7×7	DW2	128			√	1	1
1×1	FC	66				1	1
Softmax+加权求和							

1.2.2 激活函数

在神经网络中,使用激活函数能够对上一层结点的输出做非线性转换,对网络进行非线性化,使神经网络学习到复杂的映射关系,因此合适激活函数对提高神经网络的精度具有十分重要的影响[15]。文中使用 LeakyRelu 作为激活函数,当 $X\geqslant 0$ 时,表达式为 $Y=X$ 的正比例函数,导数为 1,当 $X<0$ 时,表达式为 $Y=aX$,导数为 a,具有在 X 的正负半轴上均具有导数输出的优点。LeakyRelu 激活函数表达式为:

$$\mathrm{LeakyRelu}(X)=\begin{cases}X & X\geqslant 0\\ aX & X<0\end{cases} \tag{1}$$

由(1)求导得的导数:

$$\mathrm{LeakyRelu}'(X)=\begin{cases}1 & X\geqslant 0\\ a & X<0\end{cases} \tag{2}$$

1.2.3 softmax 分类

softmax 常用在神经网络最后一层的多分类任务中,它能够将多个神经元的特征输出映射到(0,1)区间内[16],且经过 softmax 后所有的输出和为 1,从而来实现多分类。式(3)为 softmax 函数公式。

$$p(y^{(i)}=j\mid x^{(i)};\theta)=\frac{\mathrm{e}^{\theta_j^{\mathrm{T}}x^{(i)}}}{\sum_{i=1}^{k}\mathrm{e}^{\theta_j^{\mathrm{T}}x^{(i)}}} \tag{3}$$

式中:j 为第 j 个分类,θ 为网络中的权重参数,x 为输入。

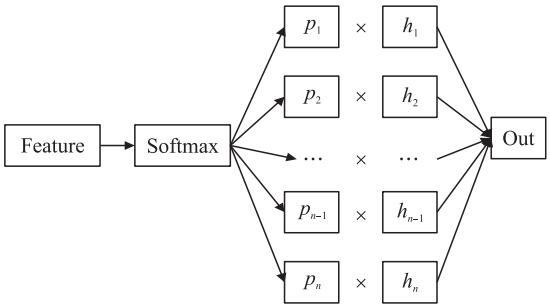


图 6 softmax 输出加权求和

文中年龄估计视为不同年龄人脸的分类问题。一般选取 softmax 概率最大(也就是值对应最大的)节点作为预测目标。但这种方式不能充分应用每一个特征输出而且分类效果也不好(如当 softmax 输出为[0.1,0.3,0.3,0.3]时,因为较大的几个数值一样或相差不大,这时候再选择最大值很容易误分类)。图 6 中将 softmax 的输出结果和对应的年龄类别分别相乘加权求和,这样可以更好的利用每一个输出特征值。

特征值经过 softmax 函数后输出 p_i,然后将 p_i 与相应的年龄类别 h_i 进行相乘,最后再将所有相乘的结果进行求和得到最终的年龄预测结果。

$$h_{\mathrm{out}}=\sum_{i=1}^{n}p_i l_i \tag{4}$$

2 实验结果与分析

2.1 实验与测试环境

文中模型训练采用 Google 的 TensorFlow 深度学习框架并在 Ubuntu 18.04 操作系统下进行网络训练。测试则是在嵌入式开发板树莓派 3b CPU 上进行模型性能测试。表 2 为具体的实验配置信息。

表 2 实验配置信息

软硬件	训练环境	测试环境
操作系统	Ubuntu 18.04	Raspbian
CPU	i7-9700k	BCM2837
GPU	GeForce 1080Ti	—
内存	DDR4	900 MHz 1G
CUDA	10.0	—
TensorFlow	1.13.2-GPU	1.14.0-CPU
numpy	1.16.4	1.16.2

2.2 数据库及评价指标

由于 IMDB-WIKI 数据欧美数据偏多缺乏亚洲数据,而且各个年龄的人数相差比较大。因此针对数据及进行扩充完善,本文工作者增加从生活中采集的亚洲数据并进行年龄标注处理。最后获得亚洲

人脸数据 35 126 张,将处理后的数据集与 IMDB-WIKI 数据集合并,并对各年龄段的数据进行均衡处理使每个年龄段的人脸数控制在[10 000,12 000]之间。图 7 为处理后的数据集各年龄样本数分布直方图,可以看出处理后的数据集分布性更好。

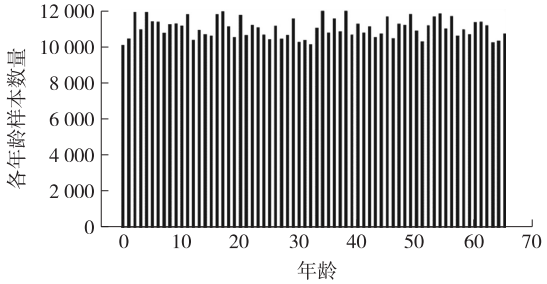


图 7　数据集中各年龄段图像数量的具体分布

人脸年龄预测评价指标通常使用结对平均绝对误差(Mean Absolute Error,MAE)来表示[17],MAE 是一种线性平均值,在所有个体差异在平均值上的权重均相等,是所有单个观测值与算术平均值的偏差的绝对值的平均。使用平均绝对误差可以避免误差相互抵消的问题,因而可以准确反映实际预测误差的大小。对于一组数据平均绝对偏差的 MAE:

$$\mathrm{MAE}(x,h)=\frac{1}{m}\sum_{i=1}^{n}|h(x_i)-y_i| \tag{5}$$

式中:x 为输入变量,h 为变量 x 的函数,$h(x_i)$ 为变量 x_i 的函数值。

另外,人脸年龄估计性能评价中,人们关注更多的是所估计出的年龄值的绝对误差范围是否在人们能接受的范围内,因此累积指数被用于年龄估计的性能评价中,累积指数(Cumulative Score,CS)的定义如下:

$$\mathrm{CS}=\frac{N_{\theta\leqslant j}}{N}\times 100\% \tag{6}$$

式中:$N_{\theta\leqslant j}$表示测试图像中估计年龄与真实年龄的绝对误差不超过 j 年的测试图像数,分母 N 为所有测试图像的总数。

2.3　实验结果分析

为了体现提出方法的优势,通过在 MORPH 2 人脸年龄数据集上对模型性能进行评价,在表 3~表 5 和图 8 中呈现了与其他方法的对比。

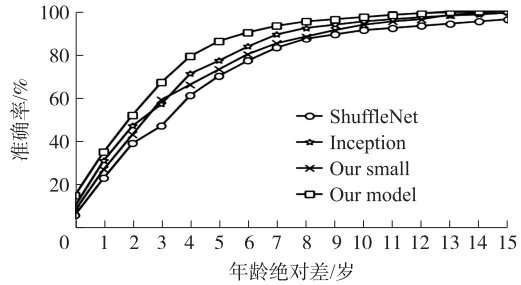


图 8　测试模型 CS 值(准确率)对比

在图 8 中可以看到,本文方法在 0~15 岁年龄差条件下,准确率明显优于文献[18]和文献[19]分别提出了 Inception 和 ShuffleNet 网络,为了验证本文所提网络效果,采取对比试验,其中在年龄岁数差为 2~8 岁时,本文方法显著优于 Inception Net 和 ShuffleNet 方法。本文方法在接近五岁年龄差时,准确率能够达到 90%,而 Inception 方法的准确率需要在 7 岁左右的年龄差条件下能够达到 90%,ShuffleNet 方法需要在 10 岁年龄差的条件下准确率能够达到 90%。另外,相同年龄差条件下,本文中的 Small(当 SMLF-Net 输入的分辨率降为 64×64,通道降低一半时称为 Our small)模型准确率优于 ShuffleNet 方法。

表 3 为本文方法与 Inception 和 ShuffleNet 的 MAE 在 MORPH 2 数据集上测试结果对比,本文方法在全尺寸模型时测试 MAE 值为 3.09 均优于其他两种方法,本文的 Our small 模型测试 MAE 值为 3.41 其表现也优于 ShuffleNet,但在表 4 可以看到 small 模型测试时间远远优于 ShuffleNet 方法和 Inception 方法,牺牲了一定的准确性来换取预测的快速性。

表 3　测试模型 MAE 比对

Model	MAE
Our model	3.09
Our small	3.41
Inception[18]	3.32
ShuffleNet[19]	3.54

表 4 本文模型与 Inception 和 ShuffleNet 模型在树莓派 3b 上推理一张人脸图像所用时间对比,通过对比可以看出本文方法在全尺寸模型时用时 169 ms,识别时间优于 Inception 和 ShuffleNet 模型,而本文的 Our small 模型的识别时间 31 ms,其识别时间表现最佳。

表 4　测试模型预测时间比对

Model	Time
Our model	169 ms
Our small	31 ms
Inception[18]	2.1 s
ShuffleNet[19]	337 ms

将训练好的模型参数进行量化,实现模型训练参数的固定,方便模型移植应用,而模型所占内存大小是该方法能否应用在硬件上的重要条件之一。在表 5 中对比本文模型与 Inception 和 ShuffleNet 模型内存及参数,通过对比可以看出本文方法在全尺寸模型时所占内存为 5.5 M,参数值为 1.37 M 明显优于 Inception 和 ShuffleNet 模型,相比 Inception 方法模型

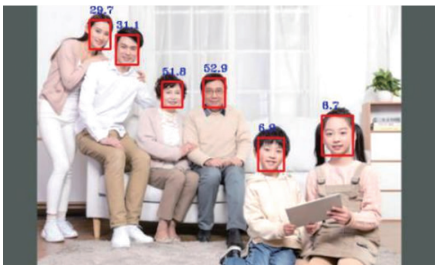
减小了 14 倍以上。本文的 Our small 模型内存仅为 1.6 M,参数仅为 0.35 M 是四个模型中最好的一个。

表 5 测试模型所占内存及参数比对 单位:M

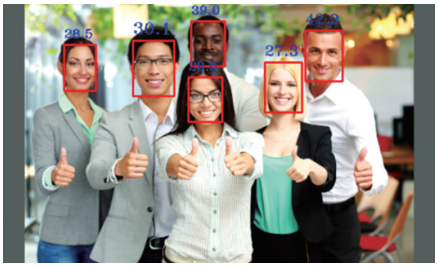
Model	FLOPs	Params
Our model	5.5	1.37
Our small	1.6	0.38
Inception[18]	87	22.8
ShuffleNet[19]	10.7	2.47

通过以上的对比分析,本文所提基于 SMLFnet 的人脸年龄估计方法在精确度、快速性和内存空间消耗上均优于 ShuffleNet 方法和 Inception 方法。另外,考虑到部分场景下对内存和时间性能要求较高,提出的 Small 方法更加适用于快速预测人脸年龄估计,能够以较小的空间和较高的速度准确通过人脸识别预测年龄。

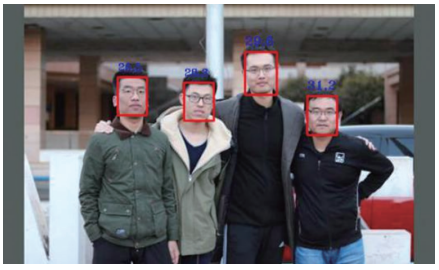
最后,为验证实际场景下的人脸年龄估计,在实际场景中对文中所提算法进行测试,其中人脸检测采用的是 PyramidBox 算法对人脸进行定位。从图 9(a)~(c)中可以看出文中所提算法对各年龄阶段的人脸年龄估计均达到与视觉年龄相近的结果,具有很好的泛化性。



(a)



(b)



(c)

图 9 实验测试图片

3 结论

针对人脸年龄具有个体差异性大、干扰因素多、标准卷积神经网络计算参数量大导致人脸年龄难以实时准确识别的问题,提出一种融合了 RE 及 SE 网络结构的可融合深度可分离卷积神经网络结构,此网络加入了 BN 层提高了网络收敛速度,降低了网络训练难度。采用 RE 算法与 SE 算法结构提取人脸的高维特征、中维特征和低维特征。同时增加了网络的非线性以及较好的避免网络梯度消失,退化。在网络中采用的深度可分离网络降低了网络的参数量,使该网络在增加网络深度的同时不会导致参数量巨增。通过实验与其他模型进行对比,该网络结构具有更少的参数,更快的预测速度,更好的性能。下一步研究方向是结合性别和种族等人脸属性因素对人脸年龄估计的影响,以及进一步压缩网络参数使模型推理速度更快。

参考文献:

[1] Geng X, Yin C, Zhou Z H. Facial Age Estimation by Learning from label distributions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(10): 2401-2412.

[2] Yi D L, Lei Z, Li S Z. Age Estimation by Multi-Scale Convolutional Network[C]//Proceedings of Asian Conference on Computer Vision. Springer, Cham. 2014: 144-158.

[3] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition [C]//Computer Vision and Pattern Recognition. IEEE, 2016: 770-778.

[4] Szegedy C, Liu W, Jia Y, et al. Going Deeper with Convolutions [C]//Computer Vision and Pattern Recognition. IEEE, 2015: 1-9.

[5] Hu Z, Wen Y, Wang J, et al. Facial Age Estimation with Age Difference[J]. IEEE Transactions on Image Processing a Publication of the IEEE Signal Processing Society, 2017, 26(7): 3087-3097.

[6] Wang Xiao Long, Guo Rui, Kambhamettu Chandra. Deeply-Learned Feature for Age Estimation[C]//2015 IEEE Winter Conference on Applications of Computer Vision. IEEE, 2015.

[7] Levi Gil, Hassncer Tal. Age and Gender Classification Using Convolutional Neural Networks[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE Computer Society, 2015: 34-42.

[8] Yi Dong, Lei Zhen, Li Stan. Age Estimation by Multi-Scale Convolutional Network [C]//Asian Conference on Computer Vision. 2015.

[9] Niu Zhen Xing, Zhou Mo, Wang Le, et al. Ordinal Regression with Multiple Output CNN for Age Estimation[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2016: 4920-4928.

[10] Choi S E, Lee Y J, Lee S J, et al. Age Estimation Using a Hierar-

chical Classifier Based on Global and Local Facial Features[J]. Pattern Recognition,2011,44(6):1262-1281.

[11] Chen K, Gong S, Xiang T, et al. Cumulative Attribute Space for Age and Crowd Density Estimation[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society,2013:2467-2474.

[12] Feng W, Jian C, Liu W, et al. Additive Margin Softmax for Face Verification[J]. IEEE Signal Processing Letters, 2018, 25(7): 926-930.

[13] Eidinger E, Enbar R, Hassner T. Age and Gender Estimation of Unfiltered Faces[J]. IEEE Transactions on Information Forensics and Security,2014,9(12):2170-2179.

[14] Jie H, Li S, Albanies, et al. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019,42(8):2011-2023.

[15] 王红霞,周家奇,辜承昊,等. 用于图像分类的卷积神经网络中激活函数的设计[J]. 浙江大学学报,2019,53(7):1363-1373.

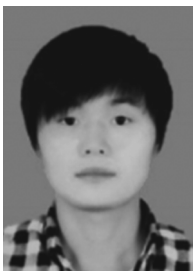
[16] Bin Y U, Kumbier K. Artificial Intelligence and Statistics[J]. Frontiers of Information Technology & Electronic Engineering, 2018,19(1):9-12.

[17] Guo G, Fu Y, Dyer C R, et al. Image-Based Human ge Estimation by Manifold Learning and Locally Djusted Robust Regression[J]. IEEE Transactions on mage Processing,2008,17(7):1178-1188.

[18] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2016:2818-2826.

[19] Zhang X, Zhou X, Lin M, et al. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices[C]//CVPR. 2017.

李　超(1990—),男,山东省临沂人,贵州电子信息职业技术学院讲师,主要研究方向是机器学习,2275769994@qq.com;



刘永辉(1994—),男,河南省商丘人,浙江工业大学硕士研究生,主要研究方向是计算机视觉,嵌入式系统,lyhgalaxy@aliyun.com。

童　林(1995—),男,黑龙江省鸡西人,六盘水师范学院讲师,主要研究方向为机器学习,tonglin0123@foxmail.com;