



基于深度学习的脸部年龄预测算法

张亮亮, 张明艳, 程凡永, 周 鹏

(安徽工程大学 安徽省电气传动与控制重点实验室, 安徽 芜湖 241000)

摘要: 为提高脸部年龄预测的准确性, 在深度学习的基础上提出一种可有效预测脸部年龄的算法。通过对人脸图像进行预处理, 获取左眼、右眼、鼻子和嘴巴四个部分的局部图像, 利用迁移 TensorFlow 深度学习库中的 Inception V4 模型, 提取脸部图像四个部分的多尺度局部特征, 并将提取的局部特征使用串联方式相连接以得到融合特征, 再将不同年龄的融合特征输入双向长短期记忆中, 以学习不同年龄融合特征间的相关性, 进而完成年龄预测。在公开数据集 FG-NET 和 MORPH 上的实验结果表明, 该算法通过利用脸部多尺度融合特征和不同年龄融合特征间的相关性, 能够显著提高年龄预测的准确性和鲁棒性。

关键词: 深度学习; 卷积神经网络; 递归神经网络; 脸部图像; 年龄预测

开放科学(资源服务)标志码(OSID):



中文引用格式: 张亮亮, 张明艳, 程凡永, 等. 基于深度学习的脸部年龄预测算法[J]. 计算机工程, 2021, 47(5): 267-272.

英文引用格式: ZHANG Liangliang, ZHANG Mingyan, CHENG Fanyong, et al. Facial age prediction algorithm based on deep learning[J]. Computer Engineering, 2021, 47(5): 267-272.

Facial Age Prediction Algorithm Based on Deep Learning

ZHANG Liangliang, ZHANG Mingyan, CHENG Fanyong, ZHOU Peng

(Anhui Key Laboratory of Electric Drive and Control, Anhui Polytechnic University, Wuhu, Anhui 241000, China)

[Abstract] In order to improve the accuracy of facial age prediction, this paper proposes an algorithm for effective age estimation based on a deep learning. Based on the preprocessed face images, the images of the parts around the left eye, right eye, nose and mouth are obtained. Then by using the Inception V4 model in the deep learning database of TensorFlow, the multiscale local features of these four face parts are extracted and connected in series to obtain the fused features. The fused features of different ages are input into bidirectional Long Short-Time Memory (LSTM) to learn the correlation between the fused features of different ages, completing the age prediction. The experimental results on the open data sets FG NET and MORPH show that the algorithm can significantly improve the accuracy and robustness of age prediction by using the correlation between the fused multiscale facial features and the fused features of different ages.

[Key words] deep learning; Convolutional Neural Network (CNN); Recursive Neural Network (RNN); facial image; age prediction

DOI: 10.19678/j.issn.1000-3428.0057426

0 概述

人脸部年龄的准确预测具有很高的利用价值, 在智能推荐、电子商务广告推送、图像信息检索与图像理解^[1]等人机交互任务中具有广阔的应用前景。目前使用较广泛的年龄预测系统由特征提取和年龄预测两个部分组成。前者涉及到人体测量学模

型^[2]、柔性模型^[3]及流形学习模型^[4]等。人体测量学模型在对人脸器官信息测量的基础上结合数学建模方法, 可有效描述轮廓变化和年龄增长之间的关系。柔性模型关键是将脸部形状和灰度或纹理相联系, 并进行相应的训练获取人脸的特征信息, 进而获得相应的人脸模型。流形学习模型通过使用流形嵌入技术对不同年龄的人脸图像进行训练学习, 进而

基金项目: 国家自然科学基金(61976005); 安徽省教育厅重点项目(KJ2019A0149); 安徽高校协同创新项目(GXXT-2020-070); 安徽工程大学-鸠江区产业协同创新项目(2021cyxta2); 安徽工程大学科研项目(Xjky02201903, 2017YQQ010, 2017YQQ011)。

作者简介: 张亮亮(1995—), 男, 硕士研究生, 主研方向为人工智能、深度学习、图像处理; 张明艳, 讲师、硕士; 程凡永, 讲师、博士; 周 鹏, 副教授、硕士。

收稿日期: 2020-02-19 修回日期: 2020-04-30 E-mail: b12090031@hnu.edu.cn

获取到相应的年龄变化趋势,该算法的优势在于不需要大量的个体样本,甚至在某些特殊情况下单个个体仅需要一个样本即可。年龄预测部分主要包括回归模型^[5]和分类模型^[6]。回归模型依据人类年龄值的连续特点,使用相应的回归函数实现对年龄值的拟合计算,进而能够有效实现对年龄的预测。文献[5]用二次回归函数表示脸部图像的年龄变化规律,该方法在局部年龄段可取得较好的预测结果,但对于该年龄段之外的脸部图像年龄预测效果欠佳。分类模型在对年龄进行预测时,通常是按照提前划分好的年龄段实现对年龄的准确预测,并且最终计算出的预测结果也不是与真实年龄值相似的值,而是一个一定大小的预测范围。文献[6]借助EM算法在低维2DLDA+LDA空间中构建了11种高斯模型,由此能够对各段年龄区间进行明确划分。

上述方法都难以有效提取人脸的复杂特征并挖掘出人脸的年龄变化规律。针对这两个方面的问题,本文结合卷积神经网络(Convolutional Neural Network, CNN)和递归神经网络(Recursive Neural Network, RNN)的优点对脸部图像进行分割,并得到脸部图像的左眼、右眼、鼻子和嘴巴四个部分的局部图像,再通过Inception V4模型^[7]提取人脸面部图像多尺度局部特征,并将这四个部分的局部特征使用

串联的方式得到融合特征。通过将不同年龄的融合特征输入双向长短期记忆(Long Short-Time Memory, LSTM)^[8]中学习不同年龄的特征关联性,以实现对不同年龄的准确预测。

1 年龄特征提取

1.1 基于传统卷积神经网络的年龄特征提取

卷积神经网络^[9]是一种深度学习模型,利用该模型可快速有效地提取图像特征。目前卷积神经网络已广泛应用于图像识别和分析等领域。如图1所示,CNN的层级结构有以下5层构成:第1层是输入层(Input Layer),该层的主要作用是预处理原始图像;第2层是卷积层(Convolution Layer),该层可完成卷积运算;第3层是池化层(Pooling Layer),该层可实现对图像的压缩处理;第4层是全连接层(Fully Connected Layer),该层可实现特征转换处理;第5层是输出层(Output Layer),该层的作用是将分类结果输出。卷积运算的过程如式(1)所示:

$$Y_j^l = \int \left(\sum_{i \in M_j} Y_i^{l-1} \times K_{ij}^l + b_j^l \right) \quad (1)$$

其中, l 为卷积层数, Y_j^l 为第 l 层的第 j 个特征图, M_j 为感受野, K 为卷积核, b 为偏置, \int 为激励函数。

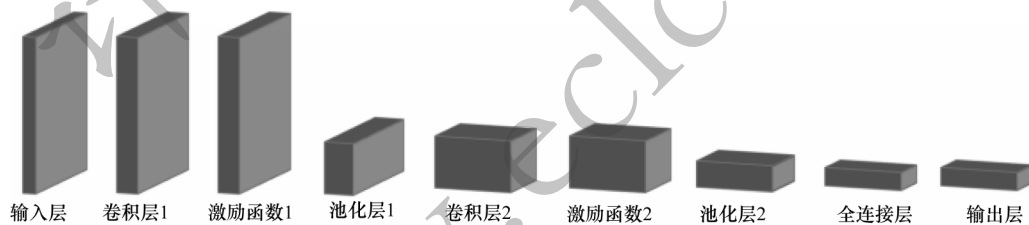


图1 CNN结构

Fig.1 Structure of CNN

当网络结构较深时,传统的CNN模型提取的特征主要为抽象特征,缺少人脸的多尺度信息,而年龄预测不仅需要人脸的抽象特征,而且还需要注重细节特征。因此,直接使用传统的CNN模型对年龄进行预测的效果欠佳。

1.2 基于Inception V4的年龄特征提取

针对传统CNN模型对年龄预测效果较差的问题,本文将Inception V4用于特征提取中。Inception V4是Google在以Inception网络结构构成的前三代分类器的基础上,经过升级得到的第四代卷积神经网络分类器。该网络分类器的结构如图2所示。Inception V4网络对于构成Inception块的每个深度

网络网格大小都进行设计和统一,并通过在每一层中引入多尺度卷积核来实现多尺度特征的提取,且在图像分类中取得较高的分类准确率。

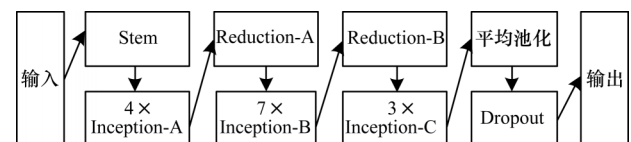


图2 Inception V4网络结构

Fig.2 Network structure of Inception V4

本文用Inception V4提取脸部子区域图像的特征。当脸部图像经过预处理输入Inception网络后,该模块会对同一输入映射上的多个不同变换进行 5×5 和 3×3

的卷积变换以及最大池化,并将相应的结果都传送到输出上,其过程如图3所示。即该模块对每个神经网络层都会进行并行计算,且该模型的下一层会对获取的特征信息进行判断以及如何有效利用。

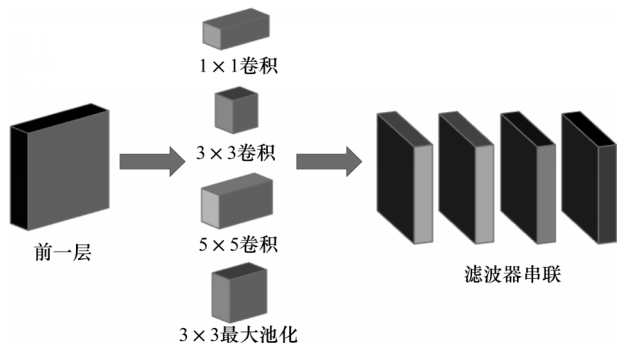


图3 Inception模块工作机理

Fig.3 Working mechanism of Inception module

上述运算在充分保留信息密度的同时,也会引起计算量激增。为了降低计算量,Inception V4借助于 1×1 卷积对输入张量进行特征降维和融合,例如如图4中的Inception-A块,利用 1×1 卷积和多通道的优势,能够有效提取空间信息,并且能够将提取到的空间信息压缩到更低维度。通过以上运算操作,不仅可充分保留脸部图像的信息密度,而且可更有效地提高神经网络的表征能力。由Inception V4得到不同年龄的融合特征后,双向LSTM将对这些不同年龄的多尺度融合特征进行关联性分析,以完成年龄预测。

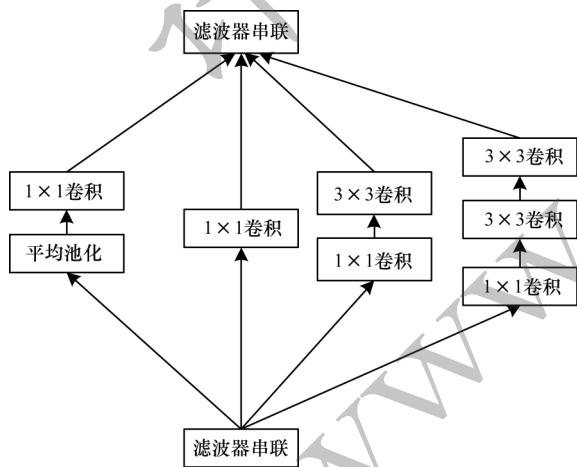


图4 Inception V4中的Inception-A块示意图

Fig.4 Schematic diagram of Inception-A module in Inception V4

2 年龄预测

2.1 整体框架

不同个体的人脸老化过程具有个性化和动态性的双重特点。因此,假如将年龄预测作为一种分类问题,即能够将不同年龄当作彼此相互独立问题来

处理,但是这样的做法忽略了不同年龄之间的相关性,将会导致各个年龄区间的主要信息被遗漏。假设将预测年龄看作是回归分析中的主要问题,并由此展开一系列的具体操作,虽然能够对不同年龄的相关性加以考虑,但是处理过程却将人脸年龄的变化作为一个恒速变化对待,而实际上青年人的人脸变化速度相较于中老年人更快^[10]。

针对以上问题,本文将Inception V4与双向LSTM联合用于年龄预测领域,实现了多尺度特征提取和特征关联性分析的结合,充分挖掘脸部特征变化规律以降低年龄预测误差。本文所提算法先预处理原始脸部图像,进而可以得到人脸局部图像,然后利用Inception V4模型对脸部局部特征进行分类提取,再将不同局部特征相串联进而得到融合特征,继而将不同年龄的融合特征输入至双向LSTM中,以学习不同年龄特征间的相关性,从而完成年龄预测。本文所提年龄预测框架流程如图5所示。

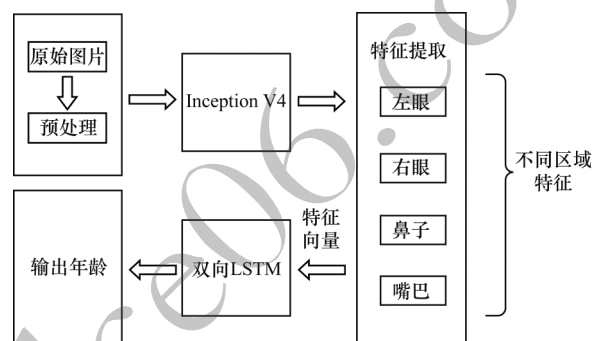


图5 年龄预测流程

Fig.5 Procedure of age prediction

2.2 基于双向LSTM的年龄预测

卷积神经网络在特征提取方面表现较佳且应用广泛。然而,通过卷积神经网络提取的特征通常是静态特征,特征之间缺少相关性,但现实世界中很多元素之间都是相互联系的,人脸年龄特征则是如此。因此,在年龄估计阶段需挖掘出不同年龄脸部特征的相关性以及变化规律。

针对以上问题,本文将双向LSTM应用于算法的年龄预测阶段。LSTM可对时序数据进行合理的处理,并且确保每次输入中的重要信息都可以得到保留,从理论上讲可以将输入的所有核心信息都包括在内,继而通过对所有输入信息进行综合考虑来预测输出。但是这里所包含的所有输入信息仅对预测内容前面的内容进行考虑,即只对上下文中的“上文”信息进行有效考虑,但是却忽略了“下文”信息,则该过程有可能导致重要信息丢失的问题。双向LSTM正是基于这种思想,不仅从前往后保留预测年龄前面的重要信息,而且从后往前保留预测年龄后面的重要信息,即对各个年龄特征进行关联

性分析。年龄预测框架如图6所示。其中,图片1、图片2和图片3表示同一人物不同年龄时刻的图片,将这些人脸图像经过预处理后,按年龄顺序依次输入Inception V4中提取脸部图像多尺度特征,并得到特征向量 x_0 、 x_1 、 x_2 。然后依次将特征向量输

入到双向LSTM中,分别用前向和后向LSTM来学习过去和未来包含的隐藏年龄变化信息,以获得脸部图像的老化模式。在双向LSTM的最后一层中添加一个全连接层,由全连接层和Softmax激励函数输出预测的年龄分布,进而完成年龄预测。

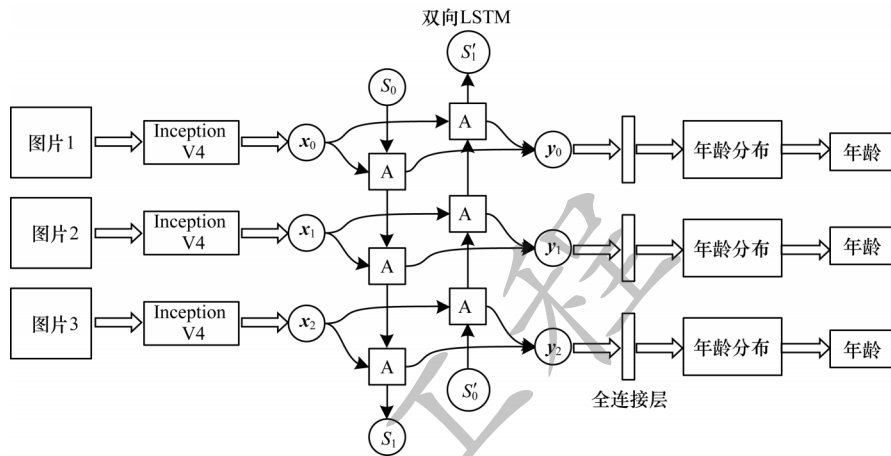


图6 年龄预测模型框架

Fig.6 Framework of age prediction model

3 实验结果与分析

3.1 数据集与实验平台

FG-NET^[11-12]数据集包含固定约束框架下的1 002幅彩色和灰度脸部图像,它是由82个人的人脸图像组成的,平均每人有12幅图像,这些人的年龄为0~69岁。FG-NET中的样本图像如图7所示。

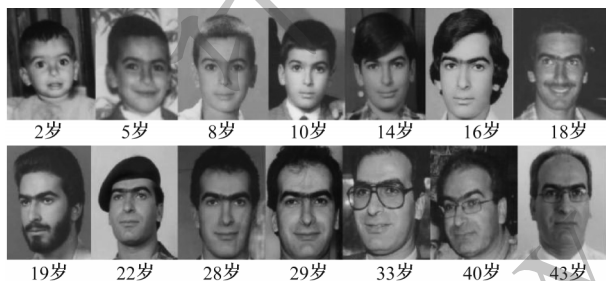


图7 FG-NET数据集示例

Fig.7 Examples of FG-NET dataset

受到数据采集总量的影响,FG-NET借助于留一交叉验证方法来合理划分各类训练集和相应的测试集,该方法的主要操作方式是在所有人中任意选择其中一人,并将其影像作为测试集,而其他所有图像均用作训练集,最终的预测误差取平均值。

MORPH^[13-14]数据集是由北卡罗来纳大学研究人员收集到的真实数据,共包括Album 1和Album 2这2个部分。对论文的具体引用情况进行分析可知,Album2数据集是现阶段应用较为普遍且较为常见的数据集之一,与预测年龄相关的多数理论资料都在该数据集中展开了实际测评。Album2也是一个跨时间的数据集,其收录了同一个人不同年

龄段的图片。该数据集分为商业和学术版本,学术版本包括了13 000个人的55 134张图片,照片收集时间跨度为2003年—2007年,人物年龄为16岁~77岁,且平均年龄为33岁。MORPH数据集中的示例图像如图8所示。



图8 MORPH数据集示例

Fig.8 Examples of MORPH dataset

本文使用TensorFlow^[15]平台进行实验验证,TensorFlow是Google官方对外公开的一款机器学习框架,由于该框架具有较强的灵活性和可移植性,因此利用该框架可使得计算模型轻易地部署到分布式集群、单主机以及移动设备上。

3.2 评价指标

评价算法的优劣不能缺少一个客观公正的指标,针对年龄预测算法优劣的评价存在多种指标。而目前针对年龄预测评价,使用较多的指标之一是平均绝对误差(Mean Absolute Error, MAE)^[16]。

MAE是全部预测值和实际值进行综合比对后

而得出的平均误差,其相应的数学定义可表示为:

$$MAE = \frac{1}{S} \sum_{i=1}^S |x_i - \hat{x}_i| \quad (2)$$

其中, S 表示测试图像的数量, x_i 表示第*i*幅图像的真实值, \hat{x}_i 表示第*i*幅图像的预测值。该指标计算结果越小,则表明算法的性能越佳。

3.3 实验结果与验证

为了从不同角度验证本文算法在年龄预测中的适用性和优越性,本文利用TensorFlow平台在FG-NET和MORPH数据集中完成相关算法的验证。在FG-NET数据集上与文献[17-19]算法进行对比,结果如表1所示。从表1可知,本文算法在FG-NET数据集上对年龄预测的MAE均低于其他3种方法。与本文算法相比,文献[17]提出的LARR算法是一种流形学习方法,该算法主要利用支持向量回归的方法对年龄进行预测,而由对比实验结果可知,使用卷积神经网络提取年龄特征对年龄预测更加有效。文献[18]将年龄预测问题转换为相应的特征优化问题,在确保图像局部流形结构不变化的同时也保留了人脸老化图像的序数信息,而从对比实验结果来看,通过使用双向LSTM学习人脸年龄变化的方式,能够更好地利用不同年龄间的相关性信息,在对年龄进行预测时,则能够获得更佳性能。文献[19]提出一种新的修正对比损失函数,通过模拟单个样本之间的成对序数关系来增强算法的鲁棒性,而实验结果表明使用Inception V4能够更好地提取人脸面部图像的细微特征,从而保留多尺度信息,且结合双向LSTM能够取得更佳的预测准确率。

表1 4种算法在FG-NET数据集上的MAE对比

Table 1 MAE comparison of four algorithms on FG-NET database

算法	MAE
LARR算法	5.68
PLO算法	4.83
VGG+RCL算法	4.26
本文算法	3.64

实验将本文算法和文献[20-22]算法在MORPH数据集上进行对比,结果如表2所示。从表2可以看出,在MORPH数据上,本文算法相比其他3种算法对年龄的预测效果更好。文献[20]将多尺度分析策略从传统方法引入到CNN中,实现了多尺度面部年龄特征分析,该算法对人脸进行区域划分,虽然提取了人脸的多尺度特征,但是不同年龄脸部特征的相关性和变化规律未能充分利用,对比实验结果表明使用双向LSTM可更好地利用不同年龄之间的相关性信息,提高预测准确率。文献[21]将传统的年龄回归问题转化为一组二分类的子问题,但是人脸年龄变化是动态过程,不同年龄之间的相关性信息

对于年龄预测也是十分重要的,由对比实验结果可知,使用双向LSTM学习人脸老化模式的年龄预测方法,在预测准确率上更有优势。文献[22]使用轻量化的模型用于单一图像的年龄估计,可移植性高,但对于特征提取,轻量化模型难以提取人脸多尺度特征,通过对比实验可知,使用Inception V4提取人脸年龄特征,能够更好地保留人脸多尺度信息,并提高年龄预测准确率。

表2 4种算法在MORPH数据集上的MAE对比

Table 2 MAE comparison of four algorithms on MORPH database

算法	MAE
Multi-Scale CNN算法	3.65
OR-CNN算法	3.29
SSR-Net算法	3.21
本文算法	3.18

由以上对比实验可知,在FG-NET和MORPH数据集中,采用本文算法对年龄进行预测的MAE均小于其他3种算法,取得最优年龄预测准确率。这说明与其他算法相比,本文算法综合了卷积神经网络和递归神经网络的优点,不仅实现了多尺度特征提取,而且充分利用了脸部年龄特征变化的相关性,并降低了年龄预测的误差。

4 结束语

针对人脸年龄预测问题,本文提出一种基于深度学习的脸部年龄预测算法。该算法将脸部图像分为四部分人脸局部图像,使用Inception V4分别提取多尺度局部特征,将特征以串联的方式相融合并输入双向LSTM中,以学习年龄变化的相关性,从而完成脸部年龄的准确预测。实验结果表明,与其他算法相比,本文算法在利用多尺度卷积神经网络提取脸部年龄特征后,利用脸部年龄特征变化与时间的相关性,可大幅降低年龄预测误差。下一步将把实验测试扩展到工业细分拣、人的身份识别等更多实际应用中,以更好地验证本文算法的普适性。

参考文献

- [1] RAMANATHANA N, CHELLAPA R, BISWAS S. Computational methods for modeling facial aging: a survey[J]. Journal of Visual Language and Computing, 2009, 20(3): 131-144.
- [2] WANG Xianmei, LIANG Lingyan, WANG Zhiliang, et al. Age estimation by facial image: a survey[J]. Journal of Image and Graphics, 2012, 17(6): 4-19. (in Chinese) 王先梅,梁玲燕,王志良,等.人脸图像的年龄估计技术研究[J].中国图象图形学报,2012,17(6):4-19.
- [3] GUO Jingjing, LIAO Yuming, SUN Ruanhuang. Human face age estimation with adaptive hybrid features[C]// Proceedings of International Conference on System

- Science & Engineering. Washington D. C. , USA : IEEE Press, 2011: 55-58.
- [4] FU Y, HUANG T S. Human age estimation with regression on discriminative aging manifold [J]. IEEE Transactions on Multimedia, 2008, 10(4): 578-584.
- [5] LANITIS A, TAYLOR C J, COOTES T F. Toward automatic simulation of aging effects on face images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 442-455.
- [6] UEKI K, HAYASHIDA T, KOBAYASHI T. Subspace-based age-group classification using facial images under various lighting conditions[C]//Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition. Washington D. C. , USA : IEEE Press, 2006: 43-48.
- [7] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-V4, inception-ResNet and the impact of residual connections on learning[EB/OL]. [2020-01-10]. <https://www.ixueshu.com/document/475e826986754988bafae32faf5e676f.html>.
- [8] ZHANG T, MOUCHERE H, VIARD-GAUDIN C. Tree-based BLSTM for mathematical expression recognition[C]//Proceedings of the 14th IAPR International Conference on Document Analysis and Recognition. Washington D. C. , USA : IEEE Press, 2017: 914-919.
- [9] GU J X, WANG Z H, KUEN J, et al. Recent advances in convolutional neural networks[EB/OL]. [2020-01-10]. https://www.researchgate.net/publication/288060170_Recent_Advances_in_Convolutional_Neural_Networks.
- [10] WANG Qin. Facial age estimation based on adaptive label distribution learning[D]. Nanjing : Southeast University, 2015. (in Chinese)
王芹. 基于自适应标记分布学习的人脸年龄估计[D]. 南京 : 东南大学, 2015.
- [11] PANIS G, LANITIS A, TSAPATSOU LIS N, et al. Overview of research on facial ageing using the FG-NET ageing database[J]. IET Biometrics, 2016, 5(2): 37-46.
- [12] ZHANG Ke, WANG Xinsheng, GUO Yurong, et al. Survey of deep learning methods for face age estimation[J]. Journal of Image and Graphics, 2019, 24(8): 1215-1230. (in Chinese)
张珂, 王新胜, 郭玉荣, 等. 人脸年龄估计的深度学习方法综述[J]. 中国图象图形学报, 2019, 24(8): 1215-1230.
- [13] YANG Guoliang, ZHANG Yu. Facial age estimation method based on convolutional neural network [J]. Journal of Beijing Union University, 2018, 32(1): 59-64. (in Chinese)
杨国亮, 张雨. 基于卷积神经网络的人脸年龄估计方法[J]. 北京联合大学学报, 2018, 32(1): 59-64.
- [14] RICANEK K, TESAFAYE T. Morph: a longitudinal image database of normal adult age-progression[C]//Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition. Washington D. C. , USA : IEEE Press, 2006: 341-345.
- [15] ABADI M, BARHAM P, CHEN J M, et al. TensorFlow: a system for large-scale machine learning[C]//Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation. Washington D. C. , USA : IEEE Press, 2016: 265-283.
- [16] YU Mianshui, ZHU Anqing, XIE Xiaomeng. Age estimation of facial images based on local direction pattern[J]. Science Technology and Engineering, 2014, 14(3): 76-80. (in Chinese)
余棉水, 朱岸青, 解晓萌. 基于局部定向模式的人脸年龄估计[J]. 科学技术与工程, 2014, 14(3): 76-80.
- [17] GUO G D, FU Y, DYER C R, et al. Image-based human age estimation by manifold learning and locally adjusted robust regression [J]. IEEE Transactions on Image Processing, 2008, 17(7): 1178-1188.
- [18] LU Hanqing. Learning ordinal discriminative features for age estimation[C]//Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA : IEEE Press, 2012: 2570-2577.
- [19] PAN Hongyu, HU Han, SHAN Shiguang, et al. Revised contrastive loss for robust age estimation from face[C]//Proceedings of the 24th International Conference on Pattern Recognition. Washington D. C. , USA : IEEE Press, 2018: 3586-3591.
- [20] DONG Y, ZHEN L, LI S Z. Age estimation by multi-scale convolutional network [C]//Proceedings of the 12th Asian Conference on Computer Vision. Berlin, Germany: Springer, 2015: 144-158.
- [21] NIU Zhenxing, ZHOU Mo, WANG Le, et al. Ordinal regression with multiple output CNN for age estimation[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA : IEEE Press, 2016: 4920-4928.
- [22] YANG T Y, HUANG Y H, LIN Y Y, et al. SSR-Net: a compact soft stagewise regression network for age estimation[C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. Washington D. C. , USA : IEEE Press, 2018: 1078-1084.