

SFSR-Age: 一种基于人脸强语义的年龄识别算法

孙旭菲^{1,2}, 缪新颖^{1,2}, 毕甜甜^{1,2}, 王水涛^{1,2}, 喻芳宇¹

(1. 大连海洋大学 信息工程学院, 辽宁 大连 116023;

2. 辽宁省海洋信息技术重点实验室, 辽宁 大连 116023)

摘要: 针对因光照、拍摄角度及图片质量等因素导致的经典深度学习算法难以有效提取人脸特征、人物身份识别准确率难以达到理想精度的问题, 提出一种基于人脸强语义的年龄识别算法. 首先, 通过注意力矩阵增强人脸区域的特征权重, 达到提取特征区域的目的; 其次, 使用级联双向长短期记忆(Bi-LSTM)网络学习时序帧之间的特征依赖关系, 弥补部分特征缺失对识别精度的影响. 在人脸数据集 IMDB-WIKI 和数据集 Adience 上进行测试, 该算法的年龄识别准确率分别达到 78.34% 和 77.89%. 实验结果表明, 相比于其他基于深度学习算法的方法, 该算法在基于图片数据集的人物年龄识别任务上具有更高的准确率.

关键词: 年龄识别; 人脸识别; 深度学习算法; 注意力矩阵; 级联 Bi-LSTM

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1671-5489(2024)02-0347-10

SFSR-Age: An Age Recognition Algorithm Based on Strong Facial Semantics

SUN Xufei^{1,2}, MIAO Xinying^{1,2}, BI Tiantian^{1,2}, WANG Shuitao^{1,2}, YU Fangyu¹

(1. College of Information Engineering, Dalian Ocean University, Dalian 116023, Liaoning Province, China;

2. Key Laboratory of Liaoning Provincial Marine Information Technology, Dalian 116023, Liaoning Province, China)

Abstract: Aiming at the problems that the classical deep learning algorithm was difficult to extract facial features effectively and the accuracy of character identification was difficult to reach the ideal accuracy due to factors such as illumination, shooting angle and image quality, we proposed an age recognition algorithm based on strong facial semantics. Firstly, the feature weights of facial regions were enhanced by the attention matrix to achieve the purpose of extracting feature regions. Secondly, a cascaded bi-directional long short-term memory (Bi-LSTM) network was used to learn the feature dependency relationships between temporal frames and compensate for the influence of missing features on recognition accuracy. When tested on IMDB-WIKI facial dataset and Adience dataset, the age recognition accuracy of the algorithm reached 78.34% and 77.89%, respectively. Experimental results show that compared with other methods based on deep learning algorithms, the proposed algorithm has higher accuracy in the task of person age recognition based on image datasets.

Keywords: age recognition; facial recognition; deep learning algorithm; attention matrix; cascaded Bi-LSTM

收稿日期: 2023-05-26.

第一作者简介: 孙旭菲(1999—), 男, 汉族, 硕士研究生, 从事数据挖掘和人脸识别的研究, E-mail: 3522945497@qq.com. 通信

作者简介: 缪新颖(1977—), 女, 满族, 博士, 教授, 从事智能控制和数据挖掘的研究, E-mail: miaoxinying@dlou.edu.cn.

基金项目: 辽宁省揭榜挂帅重点项目(批准号: 2022081)和设施渔业教育部重点实验室开放课题项目(批准号: 202306).

基于人脸的人物年龄识别算法在未成年人上网防沉迷、网络身份识别等领域应用广泛,目前,基于人脸年龄识别方法主要分为两类:基于图像分析的方法和基于回归分析的方法.陈文兵等^[1]提出了基于多模型融合的年龄识别模型 CNN-SE-ELM,该模型首先使用卷积神经网络(CNN)初步提取人脸特征,然后使用 SENet(squeeze-and-excitation networks)进一步提取人脸深层特征,最后使用 EM-ELM(error minimized extreme learning machine)算法实现对面脸年龄的分类及性别分类.实验结果表明,该算法能提取更有效的面部特征,且分类器能快速收敛.陈济楠等^[2]提出了改进的深度神经网络模型,使用大卷积核以及跨连卷积实现对面脸特征的提取,同时使用平均池化替代传统的全连接层,为充分训练网络参数,设置相对较大的学习率以及较小的 Dropout 值,实验结果表明,改进的深度神经网络能在公开人脸数据集上达到较高的年龄识别准确率.周玉阳等^[3]提出了基于改进 VGG(visual geometry group)^[4]的年龄和性别分类算法,在数据集 Adience 上的年龄识别准确率达到 90%,性别识别准确率达到 93%.钱丽萍等^[5]提出了一种基于改进集成学习算法的人脸年龄分类算法,首先使用集成算法对图像进行分类,然后使用基于加权阈值的投票方法将集成学习的弱分类器转换成强分类器,最终得到年龄的分类结果,实验结果表明,该算法在人脸数据集上能得到较高的年龄分类准确率.房国志等^[6]提出了基于多尺度的 YOLO 算法^[7]的人脸年龄估计算法,首先使用多尺度卷积网络以及加权通道方法减缓卷积过程中目标特征丢失问题,基于决策树回归模型实现对年龄的估计,该算法在多个数据集上均达到了较好的效果,且鲁棒性较强.赵丙娟^[8]提出了基于有监督的偏最小二乘降维的人脸年龄估计算法,该算法首先使用 AAM(active appearance models)算法^[9]提取图像特征,然后使用偏最小二乘算法^[10]对特征图降维,最后基于支持向量回归(SVR)算法^[11]实现对年龄估计,实验结果表明,将最小二乘算法与 SVR 相结合的年龄估计算法具有更好的鲁棒性和年龄估计准确率.田会娟等^[12]提出了多任务卷积网络在光线变化条件下的人脸年龄估计算法,利用 YCbCr 色度以及 Retinex 算法增强对面脸特征提取效果,在不同距离以及光照强度条件下的对比实验表明,特征提取增强模块能更好地提取人脸特征,年龄估计绝对平均误差最大下降 2.99%,人脸识别准确率最大提升 25%.田青等^[13]提出了基于 AE(auto encoder)算法的关系自学习年龄估计模型 CRSAE,该方法首先基于人脸协方差矩阵投影获取初始人脸特征,将输入特征与输出编码之间建立关联正则项,然后使用一个额外的结构矩阵学习输入特征与输出特征之间的关联关系,最后求解 CRSAE 最优解,实验结果表明,CRSAE 算法在多个数据集上均能达到较高的鲁棒性.张会影等^[14]提出了改进的标记分布学习用于学习真实年龄及相邻年龄之间的模糊性和多义性,以丰富年龄信息,提高识别精度,在数据集 MORPH 和 FG-NET 上的平均绝对误差均有降低.

上述算法均实现了基于人脸的年龄识别,这些算法大多数改进了现有的深度学习算法实现鲁棒的人脸特征提取,从而达到更高的年龄识别效果,但均是从全局提取人脸特征,缺乏将目标区域与非目标区域进行有效划分,导致缺少对目标区域的高效学习策略.本文提出的 SFSR-age(strong facial semantics recognition for age)算法实现了基于注意力矩阵的目标区域学习机制,提高了算法对目标区域的建模鲁棒性,在多个数据集上的实验结果表明,本文算法对基于人脸的年龄识别结果相对于对比算法有明显优势.

1 预备知识

1.1 卷积神经网络基本组成

卷积神经网络^[15]的基本结构包括卷积层、采样层、全连接层和输出层.卷积层、采样层通常交替衔接.全连接层的作用是将特征图转换成向量,激活后作为最终的输出.

1.1.1 卷积层

基于上层输出,卷积层通过卷积核在特征图上进行卷积操作,得到新的特征图.卷积可分两种形式:1) SAME,即卷积核在卷积过程中可部分移出被卷积特征图,特征图不够卷积的区域使用 0 填充, SAME 类型的特征图卷积过程如图 1 所示;2) VALID,即卷积核在卷积过程中始终不能被移出卷积特征图, VALID 类型的特征图卷积过程中,卷积核在特征图内滑动,且特征图外围不填充.从卷积类

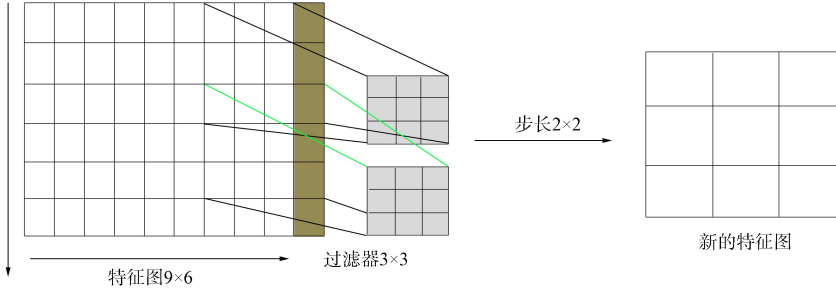


图 1 特征图卷积过程实例

Fig. 1 Example of feature graph convolution process

型看, 卷积得到的新特征图规模满足如下公式:

$$\begin{cases} W = \lceil w // s_w \rceil, \\ H = \lceil h // s_h \rceil, \\ s_w \geq 1, s_h \geq 1, \text{ SMAE}, \end{cases} \quad (1)$$

$$\begin{cases} W = 1 + (w - f_w) // s_w, \\ H = 1 + (h - f_h) // s_h, \\ s_w \geq 1, s_h \geq 1, \text{ VALID}; \end{cases} \quad (2)$$

卷积结果计算过程可表示为

$$\begin{cases} f_{i/s_w, j/s_h} = \text{sum}(f \odot \mathbf{M}_{i \sim i+f_w, j \sim j+f_h}), \\ i/s_w \leq W, j/s_h \leq H, \\ s_w \geq 1, s_h \geq 1, \end{cases} \quad (3)$$

其中 w 和 h 分别表示被卷积特征图的宽度和高度, W 和 H 分别表示卷积运算后特征图的宽度和高度, f 表示卷积核, f_w, f_h 分别表示卷积核的宽度和高度, i 和 j 分别表示水平和垂直方向卷积核移动步长, \odot 表示求内积, s 表示卷积形式, \mathbf{M} 表示被卷积特征图.

由式(1)和式(2)可见, 当卷积形式是 SAME 时, 新特征图的大小只与卷积核滑动步长 Stride 相关, 当卷积形式是 VALID 时, 新生成的特征图大小与步长 Stride 和卷积核大小线性相关. 对于一张确定的特征图以及卷积形式, 均可通过公式计算出新特征图的维度.

1.1.2 采样层

与卷积层类似, 采样层也是对特征图的进一步操作, 采样层也称池化层, 是值全为 1 的特殊卷积核, 采样过程与卷积过程完全相同, 均遵循滑动窗口从左至右、从上到下的移动策略. 采样过程分为两种形式: 最大采样和平均采样. 两种采样过程计算公式如下:

$$\begin{cases} f_{i/s_w, j/s_h} = \max(f \odot \mathbf{M}_{i \sim i+f_w, j \sim j+f_h}), \\ f_{i/s_w, j/s_h} = \text{Avg}(f \odot \mathbf{M}_{i \sim i+f_w, j \sim j+f_h}), \\ s_w \geq 1, s_h \geq 1, \end{cases} \quad (4)$$

由式(4)可见, 采样过程中最大值采样和平均采样后得到的特征图大小完全相同, 只是每个采样值的计算方式不同, 且采样过程中滑动窗口 f 为全 1 的矩阵.

1.1.3 全连接层

全连接层是直接将高层特征进一步通过矩阵的行列式乘积变成若干向量, 然后基于此开展下游任务, 假设全连接任务是单标签预测和单值回归, 全连接过程计算公式如下:

$$\begin{cases} \mathbf{f}_c = \mathbf{M} \otimes \mathbf{F}, \\ \mathbf{F}_w = \mathbf{M}_H, \end{cases} \quad (5)$$

其中 \otimes 表示特征图 \mathbf{M} 与矩阵 \mathbf{F} 之间相乘, 得到目标向量 \mathbf{f}_c , 特征图的最后一个维度与全连接矩阵的第一个维度相同.

1.2 循环神经网络

卷积神经网络以及浅层的全连接网络考虑的是对单个特征图的特征抽取, 缺乏对特征图之间的关联学习, 循环神经网络^[16]解决的是特征图之间的语义关联以及上下文的语义学习. 循环神经网络常被用于知识推理、语言生成和语义理解等领域. 常见的循环神经网络主要包括: GRU(gate recurrent unit)和 Bi-LSTM(bi-directional long short-term memory)等. Bi-LSTM 是常见的网络架构, 如图 2 所示. 由图 2 可见, Bi-LSTM 的基本架构是两层平行的 LSTM(long short-term memory): 前向和后向. 两者基于特征图的不同输入顺序从两个方向学习特征之间的相互依赖关系, 从而得到更鲁棒的特征. 对 Bi-LSTM 的两个维度的输出进行拼接便可得到融合模型, 一种典型的融合如图 3 所示.

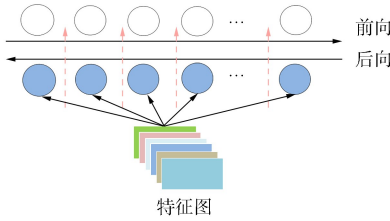


图 2 Bi-LSTM 网络架构

Fig. 2 Bi-LSTM network architecture

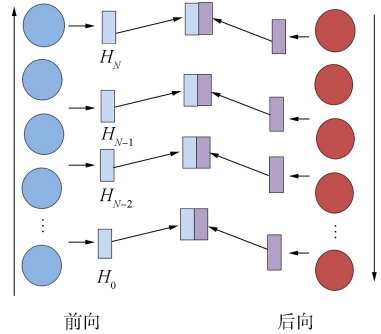


图 3 Bi-LSTM 网络特征融合

Fig. 3 Bi-LSTM network feature fusion

2 算法设计

SFSR-age 算法首先使用残差网络提取人脸图像特征图; 然后基于注意力矩阵动态学习人脸目标区域语义特征, 从语义上区分人脸和背景; 最后基于多层梯形 Bi-LSTM 网络逐层提取人脸区域高层特征, 最终得到年龄分类结果.

SFSR-age 算法主要包括两部分.

1) 基于注意力矩阵的目标人脸特征提取模块. 由于光线、用户距离计算机摄像头远近等的影响, 用户面部图像常出现模糊甚至被遮挡的情况, 如何快速锁定用户面部轮廓、准确提取面部特征对于基于面部识别用户身份至关重要. 本文使用目标人脸注意力矩阵训练目标图片, 提取目标图片人脸特征, 最大程度地学习有用特征. 因此提出基于注意力矩阵的人脸区域快速定位和特征抽取算法.

2) 基于分层级联的 Bi-LSTM 人脸特征抽取算法. 不同时刻摄像头采集的人脸图像清晰度不同, 有时人脸被遮挡, 甚至部分人脸没有出现在画面中, 基于上下文信息借助衣着、面部表情、动作等识别出用户身份, 能有效解决用户面部短暂脱离摄像头产生的身份识别不准问题. 因此, 本文提出基于梯形结构的分层 Bi-LSTM 抽取用户面部特征.

本文将 1) 和 2) 相结合共同提取用户的面部特征, 用于实现对基于人脸年龄的识别. SFSR-age 算法基本框架如图 4 所示.

2.1 SFSR-age 算法注意力矩阵

通过对大量人脸图像进行训练, 优化全局动态注意力矩阵(attention matrix, AM)提取图像中人脸特征, 同时弱化非人脸区域对人脸区域特征边界的干扰, 达到准确提取人脸特征的目的. 由于目前大多数人脸年龄数据集缺少对人脸的标注信息, 为满足本文研究的需要, 需要对 IMDB-WIKI 和 Adience 等人脸年龄数据集进行预处理得到人脸标定信息. 本文使用基于 YOLO5 预训练模型实现对所使用的数据集人脸进行初步标注, 得到每张人脸的 Bounding-Box, Bounding-Box 区域即为目标区域, 区域外即为背景. 只有 Bounding-Box 区域包含人脸, 对年龄识别起关键作用. 将目标与背景进行精确分割是准确识别人脸轮廓的关键, 基于此本文设置全局注意力矩阵识别目标轮廓, 过程如下.

1) 将数据集中的每张图像进行如下处理:

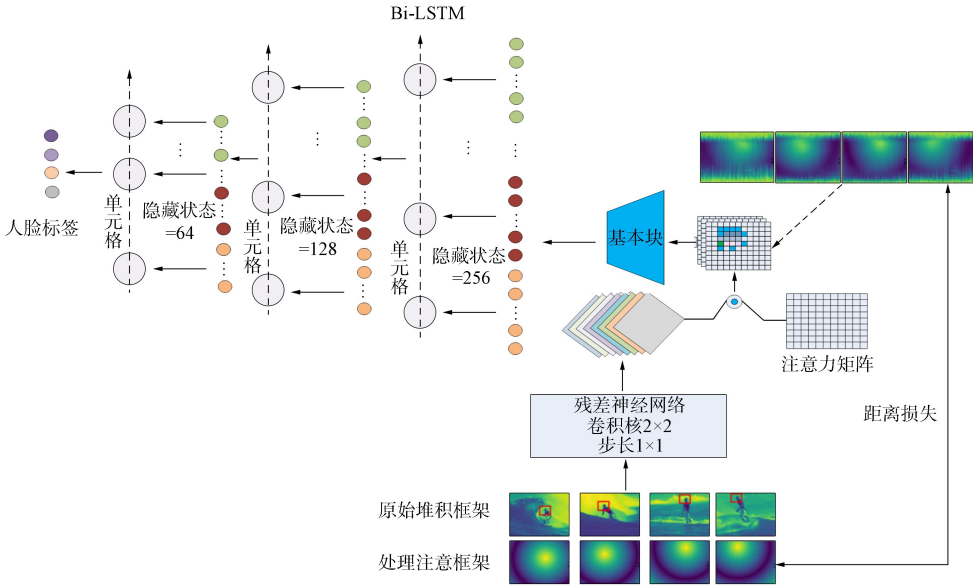


图 4 SFSR-age 算法基本架构

Fig. 4 Basic architecture of SFSR-age algorithm

$$\text{AttentionMatrix} = \begin{cases} M_n(\sqrt{(A_{i,j,h} - H/2)^2 + (A_{i,j,w} - W/2)^2}, 255.0), \\ \mathbf{A} = \begin{pmatrix} 0,0 & 0,1 & \dots & 0,M \\ 1,0 & 1,1 & \dots & 1,M \\ \vdots & \vdots & & \vdots \\ N,0 & N,1 & \dots & N,M \end{pmatrix}, \end{cases} \quad (6)$$

得到注意力矩阵. 其中 H 和 W 分别表示目标人脸区域特征图的高度和宽度, \mathbf{A} 表示特征图像素位置矩阵, $A_{i,j,h}$ 和 $A_{i,j,w}$ 表示特征图位置, (i,j) 对应相对原点的高度和宽度, 坐标 $(W/2, H/2)$ 表示人脸目标中心, 由式(1)可见, 越远离目标中心的像素点与目标的关联越小, 映射后的值越大. 其效果就是对于 AttentionMatrix 矩阵, 越靠近目标中心的像素点映射后值越大, 颜色越浅, 越远离目标中心的像素点映射后值越小, 颜色越深. 矩阵 \mathbf{A} 的大小与原始图像大小完全相同.

2) 原始图像特征提取. 使用残差网络提取图像静态特征, 实现对面脸轮廓及人脸区域的建模. 残差网络输出的特征图即为人脸特征, 使用目标轮廓与预测轮廓进行距离求解. 本文基于余弦距离和欧氏距离求解预测区域特征与真实目标区域之间的差异, 为计算方便, 首先将目标注意力矩阵与预测注意力矩阵进行灰度化处理, 得到灰度矩阵, 计算两者之间的距离损失, 即

$$\text{dis} \leftarrow \sum_{q=0}^B \sum_{k=0}^A \sum_{j=0}^Z 1 - \cos(p_{k,j}, t_{k,j}) + \| p_{k,j}, t_{k,j} \|, \quad (7)$$

其中 B, A, Z 分别是批大小、图像宽和高, k 表示 x 轴坐标, j 表示 y 轴坐标, q 表示 z 轴坐标. 假设坐标原点为 o , 图像任意一点均位于二维空间中, 如图 5 所示.

对面脸目标进行识别时, 预测向量与目标向量越重叠, 余弦值越大, 优化目标是使 $1 - \cos(p_{k,j}, t_{k,j})$ 最小, 即 $\cos(p_{k,j}, t_{k,j})$ 越大越好, 余弦距离只能保证向量平行, 平行包含重叠, 但对于目标定位, 满足平行不足以定位, 欧氏距离 $\| p_{k,j}, t_{k,j} \|$ 是求解预测轮廓与实际轮廓之间的绝对距离, 能保证预测轮廓绝对靠近目标轮廓. 因

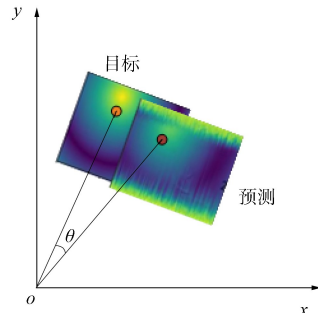


图 5 目标注意力矩阵和预测注意力矩阵的空间表示

Fig. 5 Spatial representation of target attention matrix and predicted attention matrix

此, 基于上述两个损失即可保证预测轮廓的准确性.

数据集 IMDB-WIKI 和 Adience 部分原始样本和标注后的样本分别如图 6 和图 7 所示. 部分图像的注意力矩阵如图 8 所示. 在图 6 和图 7 中, 下半部分被红色矩形围起来的即为借助 YOLO5 算法框定的人脸区域, 人脸区域是年龄识别的主要学习目标. 图 8 是 YOLO5 算法识别出的人脸区域基于式 (6) 计算得到的注意力矩阵, 矩阵中心颜色最深, 向外颜色逐渐变浅, 注意力逐渐减弱.



图 6 数据集 IMDB-WIKI 部分人脸原图及标注结果
Fig. 6 Partial facial original pictures and annotation results of IMDB-WIKI dataset

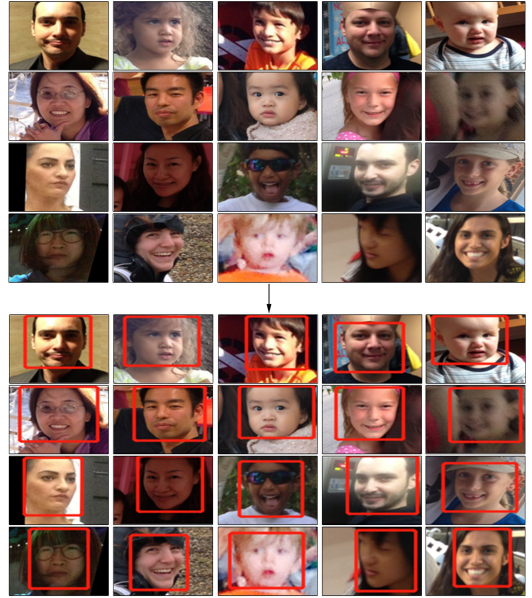


图 7 数据集 Adience 部分人脸原图及标注结果
Fig. 7 Partial facial original pictures and annotation results of Adience dataset



图 8 图像注意力矩阵(部分)

Fig. 8 Image attention matrix (partial)

2.2 基于分层 Bi-LSTM 人脸特征抽取

基于残差网络提取的特征图使用若干个层叠 Bi-LSTM 实现对目标人脸的特征抽取, 得到分类特征, 完成对年龄阶段及年龄的预测. 人脸特征的形式化表示如下:

$$fea \leftarrow \begin{cases} f(f(x_n) \parallel \theta, \lambda_0, \dots, \lambda_n), \\ f(x_0) = f(X \mid \lambda_0, B), \end{cases} \quad (8)$$

其中 B 表示 BasicCNN 的网络参数, λ_n 为第 n 层 Bi-LSTM 的隐藏层节点数, X 为第一层 Bi-LSTM 的输入.

本文为进一步分层学习目标人脸区域的深层次特征, 层叠 Bi-LSTM 之间均设置门控参数, 实现对重要特征的蒸馏, 每层 Bi-LSTM 之间依次基于上一层的建模输出乘以门控因子作为本层的输入. 图 4 中已经给出了基于分层 Bi-LSTM 过程用来提取人脸特征. 基于层叠 Bi-LSTM 特征分层提取的网络结构如图 9 所示. 由图 4 和图 9 可见, 本文基于人脸的年龄分类算法的误差由两部分组成: 人脸全域建模与特征提取和全域真实注意力矩阵之间的误差, 误差计算公式由式(7)给出.

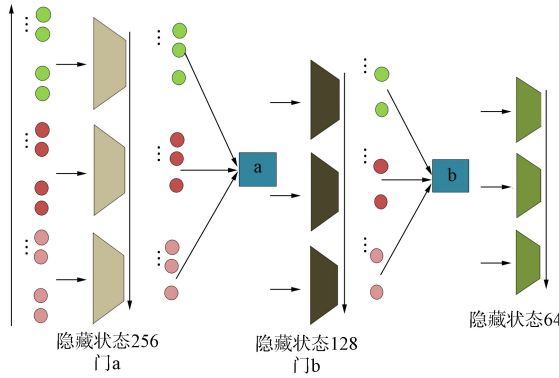


图 9 分层 Bi-LSTM 网络特征提取

Fig. 9 Feature extraction of hierarchical Bi-LSTM network

基于门控分层 Bi-LSTM 对年龄预测的损失, 由于本文使用的数据集均为带年龄标签的数据集, 将对年龄的预测转换成对年龄的分类问题, 误差由交叉熵给出, 总的损失为

$$\text{Loss} \leftarrow \text{dis} - \sum_{q=0}^B \sum_{k=0}^C P_{q,k} \log(p_{q,k}), \tag{9}$$

其中 C 表示年龄类别, $P_{q,k}$ 表示样本 q 被分类为类别 k 的概率.

3 实验结果与分析

3.1 实验环境与参数配置

本文使用的数据集为 IMDB-WIKI 和 Adience 人脸年龄数据集. 本文提出的 SFSR-age 算法可在 GPU 快速部署运行, 与其他算法在 GPU 上进行对比实验, 对比实验硬件配置: GPU 为 RTX 2080Ti, 操作系统为 Ubuntu18.04, 内存为 16 GB, 硬盘大小为 1 000 GB, CPU 核个数为 4, GPU 核个数为 1~6. 实验超参数设置: 批次大小设为 64, 学习率设为 0.001, 截断损失为 100, 优化函数采用 AdamOptimizer.

3.2 数据集实验结果

将 SFSR-age 算法与对比算法在多个数据集上进行对比实验, 本文使用的数据集基本信息列于表 1. 其中数据集 IMDB-WIKI 和 Adience 的 8 个类别样本分布列于表 2. 每个类别均给出了相应的数据规模, 为使训练过程稳定, 本文在选取样本时尽量使男女比例为 1:1, 并且在整个数据集的所有年龄类别进行采样, 其中, 数据集 IMDB-WIKI 的每个年龄段选取 1 000 名男性和 1 000 名女性; 数据集 Adience 中, 由于数据分布特征, 每个类别的样本均按最大限度采样, 而又不使不同类别间样本失衡, 最终选取出 16 080 个样本数据. 基于以上两个数据集的采样, 得到对应的子集合并成新的数据集, 然后进行对比实验, 获得实验结果.

表 1 实验数据集的基本信息

Table 1 Basic information of experimental datasets

数据集	规模	类别	训练	测试
IMDB-WIKI	16 000	8	12 800	3 200
Adience	16 080	8	12 864	3 216

注: 本文使用的数据集 IMDB-WIKI 和 Adience 均来源于原始数据集的子集, 男:女=1:1.

表 2 实验数据集的样本分布

Table 2 Sample distribution of experimental datasets

数据集	年龄/岁							
	0~2	4~6	8~13	15~20	25~32	38~43	48~53	60~100
IMDB-WIKI	2 000	2 000	2 000	2 000	2 000	2 000	2 000	2 000
Adience	1 400	2 100	2 200	1 600	4 800	2 300	820	860

注: 数据集 IMDB-WIKI 和 Adience 样本覆盖所有的光线强度和年龄类别.

本文实验的对比算法为 CNN^[17], VGG-Face-SVM^[18], CNN-SVM^[19], VGG16-DEX-IMDB^[20], CNN-SE-ELM, VGG-Face-FC. 评估指标分别为: 准确率(Accuracy)、精确率(Precision)、召回率(Recall)、 F_1 值(F_1 -Score), 直接调用 Sklearn.metrics 封装好的计算公式计算这 4 个指标值. 4 种评估指标计算公式分别为

$$\begin{aligned}
 \text{Precision}(p) &= \frac{TP}{TP + FP}, \\
 \text{Recall}(R) &= \frac{TP}{TP + FN}, \\
 F_1 &= \frac{2 \times P \times R}{P + R}, \\
 \text{Accuracy} &= \frac{TP + TN}{TP + FP + TN + FN},
 \end{aligned}
 \tag{10}$$

其中 TP 为真阳例, FP 为假阳例, FN 为假阴例, TN 为真阴例.

3.3 数据集实验分析

本文算法在两个数据集上的实验结果列于表 3. 由表 3 可见, 本文算法具有较好的鲁棒性, 4 种评估指标下的实验结果均较高. 为验证本文提出的 SFSR-age 算法对人脸图像特征提取的有效性以及年龄分类的准确率, 本文算法与目前已有的算法实验结果对比列于表 4. 由表 4 可见, 本文 SFSR-age 算法在数据集 IMDB-WIKI 和 Adience 上的实验结果优于其他算法在原始数据集上的实验结果.

表 3 本文算法在两个数据集上的实验结果

Table 3 Experimental results of proposed algorithm on 2 datasets

数据集	准确率	F_1 值	精确率	召回率
IMDB-WIKI	78.34	76.68	77.12	77.12
Adience	77.89	76.91	77.90	77.88

表 4 不同算法的准确率对比实验结果

Table 4 Comparison of experimental results on accuracy of different algorithms

算法	准确率		算法	准确率	
	IMDB-WIKI	Adience		IMDB-WIKI	Adience
CNN	43.0	52.9	CNN-SE-ELM	67.0	65.0
VGG-Face-SVM	58.4	58.0	VGG-Face-FC	54.3	53.9
CNN-SVM	65.3	64.0	SFSR-age	78.34	77.89
VGG16-DEX-IMDB	64.9	64.0			

图 10 为本文算法在数据集训练过程中预测的注意力矩阵逐步优化过程, 其中第 2 行是目标注意力矩阵, 第 3 行后的所有图片均是算法训练过程中对注意力矩阵优化过程中产生的. 由图 10 可见, 算法训练早期, 预测的注意力矩阵是随机的, 随着训练的深入, 预测的注意力矩阵逐渐向目标注意力矩阵靠拢直至重合.

为进一步量化 SFSR-age 算法相比其他对比算法在权威公开人脸数据集上年龄分类的准确率提升情况, 不同算法在两个数据集上评估指标提升结果列于表 5.

表 5 不同算法在两个数据集上 F_1 值的对比提升结果

Table 5 Comparison and improvement results of F_1 values of different algorithms on 2 datasets

算法	F_1 值		算法	F_1 值	
	IMDB-WIKI	Adience		IMDB-WIKI	Adience
CNN	82.186	47.221	CNN-SE-ELM	16.925	19.815
VGG-Face-SVM	34.144	34.276	VGG16-DEX-IMDB	20.708	21.688
CNN-SVM	19.969	21.688	VGG-Face-FC	44.273	44.490

由表 4 和表 5 可见, 以 F_1 为评估指标条件下, 本文算法在两个数据集的 F_1 值最大提升 82.186%, 最小提升 16.925%; 本文算法实验准确率完全超过 CNN 以及 VGG-Face 算法, 与目前最

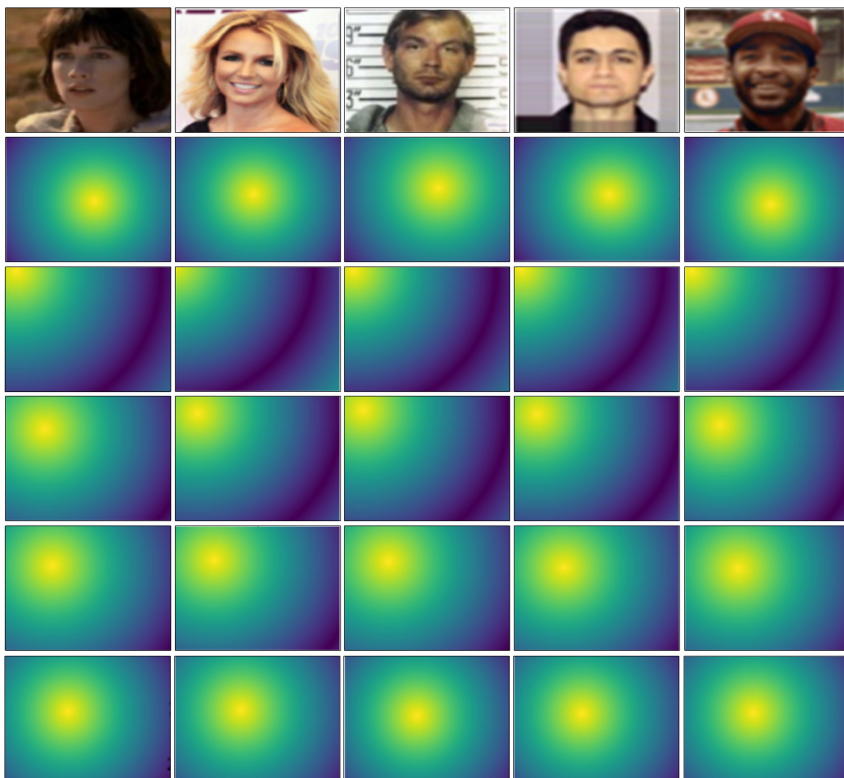


图 10 本文算法注意力矩阵优化过程(部分)

Fig. 10 Attention matrix optimization process of proposed algorithm (partial)

新的人脸年龄识别算法 CNN-SE-ELM 以及 VGG16-DEX-IMDB 算法相比, 本文算法在 F_1 值上仍分别提升了 19.815% 和 21.688%。上述对比结果表明, 本文算法具有一定的优势。

综上所述, 针对因光照、拍摄角度及图片质量等因素导致的经典深度学习算法难以有效提取人脸特征、人物身份识别准确率难以达到理想精度的问题, 本文提出了一种基于人脸强语义的年龄识别算法。首先采用残差网络提取人脸图像特征图, 基于注意力矩阵动态学习人脸目标区域的语义特征, 从语义上区分人脸和背景; 其次, 利用多层梯形 Bi-LSTM 网络逐层提取人脸区域的高层特征, 最终得到年龄分类结果。在实验中, 分别对数据集 IMDB-WIKI 和 Adience 进行了预处理, 将每个样本打上 Bounding-Box 标签, 使用多个损失共同优化算法在人脸数据集上的年龄分类误差。实验结果表明, 本文算法在数据集上的实验具有更高的准确率, 且优于其他对比算法。

参 考 文 献

- [1] 陈文兵, 李育霖, 陈允杰. 一种基于 CNN-SE-ELM 的年龄和性别识别模型 [J]. 计算机工程与科学, 2021, 43(5): 872-882. (CHEN W B, LI Y L, CHEN Y J. An Age and Gender Recognition Model Based on CNN-SE-ELM [J]. Computer Engineering and Science, 2021, 43(5): 872-882.)
- [2] 陈济楠, 李少波, 高宗, 等. 基于改进 CNN 的年龄和性别识别 [J]. 计算机工程与应用, 2018, 54(16): 135-139. (CHEN J N, LI S B, GAO Z, et al. Age and Gender Recognition Based on an Improved CNN [J]. Journal of Computer Engineering and Applications, 2018, 54(16): 135-139.)
- [3] 周玉阳, 秦科. 改进的 VGG 网络可提升年龄与性别预测准确率 [J]. 计算机工程与应用, 2019, 55(18): 173-179. (ZHOU Y Y, QIN K. Improved VGG Network Can Enhance the Accuracy of Age and Gender Prediction [J]. Journal of Computer Engineering and Applications, 2019, 55(18): 173-179.)
- [4] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition [EB/OL]. (2014-09-14)[2023-05-01]. <https://arxiv.org/abs/1409.1556>.
- [5] 钱丽萍, 俞宁宁, 黄玉蕻, 等. 一种基于改进集成学习策略的人脸图像年龄识别方法: CN201910027193.2 [P]. 2021-06-18. (QIAN L P, YU N N, HUANG Y P, et al. An Age Recognition Method for Face Images Based on

- an Improved Ensemble Learning Strategy: CN201910027193.2 [P]. 2021-06-18.)
- [6] 房国志, 孙康瞳. 多尺度 YOLO 人脸年龄估计方法研究 [J]. 计算机工程与应用, 2019, 55(21): 135-141. (FANG G Z, SUN K T. Research on a Multi-scale YOLO Method for Face Age Estimation [J]. Journal of Computer Engineering and Applications, 2019, 55(21): 135-141.)
- [7] 于秀萍, 吕淑平, 陈志韬. 基于 YOLO 算法的多类目标识别 [J]. 实验室研究与探索, 2019(3): 34-36. (YU X P, LÜ S P, CHEN Z T. Multi-class Object Identification Based on YOLO Algorithm [J]. Laboratory Research and Exploration, 2019(3): 34-36.)
- [8] 赵丙娟. 基于偏最小二乘降维的人脸年龄估计研究 [D]. 南昌: 南昌大学, 2017. (ZHAO B J. Research on Face Age Estimation Based on Partial Least Square Dimension Reduction [D]. Nanchang: Nanchang University, 2017.)
- [9] COOTES T F, EDWARDS G J, TAYLOR C J. Active Appearance Models [C]//European Conference on Computer Vision. Berlin: Springer, 1998: 484-498.
- [10] WOLD H. Path Models with Latent Variables: The NIPALS Approach [J]. Quantitative Sociology, 1975, 11: 307-357.
- [11] GUO G D, MU G W, FU Y, et al. Human Age Estimation Using Bioinspired Features [C]//The 2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2009: 112-119.
- [12] 田会娟, 乔明天, 蔡敏鹏. 基于变化光照下的人脸识别与年龄估计 [J]. 激光与光电子学进展, 2022, 59(2): 252-262. (TIAN H J, QIAO M T, CAI M P. Face Recognition and Age Estimation under Changing Illumination Conditions [J]. Laser & Optoelectronics Progress, 2022, 59(2): 252-261.)
- [13] 田青, 毛军翔, 曹猛. 耦合关系自学习的人脸年龄估计研究 [J]. 智能系统学报, 2022, 17(2): 257-265. (TIAN Q, MAO J X, CAO M. Self-learning Coupled Relationships for Face Age Estimation [J]. Journal of Intelligent Systems, 2022, 17(2): 257-265.)
- [14] 张会影, 圣文顺, 曾耀徽. 基于标记分布学习的人脸年龄识别算法 [J]. 江苏大学学报(自然科学版), 2023, 44(2): 180-185. (ZHANG H Y, SHENG W S, ZENG Y Z. Face Age Recognition Algorithm Based on Label Distribution Learning [J]. Journal of Jiangsu University (Natural Science Edition), 2023, 44(2): 180-185.)
- [15] 窦慧, 张凌茗, 韩峰, 等. 卷积神经网络的可解释性研究综述 [J/OL]. 软件学报, (2023-02-23)[2023-04-26]. <https://kns.cnki.net/kcms/detail//11.2560.TP.20230222.1625.013.html>. (DOU H, ZHANG L M, HAN F, et al. A Review of the Interpretability of Convolutional Neural Networks [J/OL]. Journal of Software, (2023-02-23)[2023-04-26]. <https://kns.cnki.net/kcms/detail//11.2560.TP.20230222.1625.013.html>.)
- [16] 胡中源, 薛羽, 查加杰. 演化循环神经网络研究综述 [J]. 计算机科学, 2023, 50(3): 254-265. (HU Z Y, XUE Y, ZHA J J. A Review of Evolutionary Recurrent Neural Networks [J]. Computer Science, 2023, 50(3): 254-265.)
- [17] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C]//The 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 7132-7141.
- [18] OZBULAK G, AYTAZ Y, EKENEL H K. How Transferable Are CNN-Based Features for Age and Gender Classification? [C]//2016 International Conference of the Biometrics Special Interest Group (BIOSIG). Piscataway, NJ: IEEE, 2016: 1-6.
- [19] WOLFSHAAR J, KARAABA M F, WIERING M A. Deep Convolutional Neural Networks and Support Vector Machines for Gender Recognition [C]//IEEE Symposium Series on Computational Intelligence. Piscataway, NJ: IEEE, 2015: 188-195.
- [20] ROTHE R, TIMOFTE R, VAN GOOL L. Deep Expectation of Real and Apparent Age from a Single Image without Facial Land-Marks [J]. International Journal of Computer Vision, 2018, 126(2): 144-157.

(责任编辑: 韩 啸)