

基于 HOG-CSLBP 与深度学习的 跨年龄人脸识别算法^①

胡渝苹

重庆水利电力职业技术学院 大数据学院, 重庆 402160

摘要: 针对人脸识别中识别精度低的问题, 提出一种基于深度学习的跨年龄人脸识别算法. 该方法创新性地将方向梯度直方图(Histogram of Oriented Gradient, HOG)和中心对称局部二值模式(Center Symmetric Local Binary Pattern, CSLBPS)组合方法用于人脸图像特征提取, 获得包含结构和强度信息的图像融合特征, 然后使用二叉树对特征信息进行降维, 降维特征作为深度信念网络的可视层输入量, 弥补深度新信念网络无法达到图像局部特征要求的缺陷. 通过训练好的深度网络模型对测试样本进行学习, 在深度信念网络的最顶层对特征进行分类识别. 实验结果表明, 该方法能高精度实现人脸识别, 且与其他方法比较, 该方法性能优于其他方法, 说明该方法具有可行性和有效性.

关键词: 方向梯度直方图; 中心对称局部二值模式; 二叉树; 深度信念网络; 跨年龄人脸识别

中图分类号: TP391.4

文献标志码: A

文章编号: 1000-5471(2020)03-0115-06

人脸识别是一个研究热点, 其面临的挑战之一是实现时间不变性^[1-3], 随着人脸识别算法的发展, 其识别精度也随之增加, 其应用领域也越来越广泛, 如监控、门禁、金融支付和远程身份识别等^[4-5]. 尽管不同外界条件下的人脸识别已经取得了很多研究, 但是跨年龄人脸识别仍然是一个挑战^[6-7]. 老化的人脸识别具有许多有效的实际应用, 例如寻找失踪儿童, 根据照片识别罪犯和多个登记检测问题.

局部二值模式选择是解释数字图像纹理和模型的一种非常强大的方法, 因此是人脸识别系统中特征提取的理想选择. 首先将面部图像分割成小区域, 再提取特征, 然后将其连接成单个特征向量. 该向量形成面部区域的有效表示, 并可用于测量图像之间的相似性^[8]. 从原始面部图像导出面部表示, 是成功的面部表情识别方法的必要步骤.

大多数常规面部识别系统并不适合处理随年龄增长的面部识别, 且许多早期的研究都集中在年龄估计方面, 只有少数研究与衰老的人脸识别直接相关. Zou 等^[9]提出一种跨年龄人脸识别模型, 该模型构造了一种按时间顺序排列的参考集来消除输入图像的年龄差异, 利用图库集的身份信息进行判别分析, 从而可以在参考编码后进一步区分分子空间中的人. Li 等^[10]提出了一种老化人脸识别方法, 其中使用了基于局部模式选择的分层学习模型. 在基于两级学习的分层模型的帮助下解决了该问题. 在第一阶段, 基于局部模式选择从低微结构学习有效特征, 从而减少用户内差异. 在第二阶段, 基于低级结构的输出细化更高级别的信息. Li 等^[11]提出深度卷积网络的跨年龄人脸识别方法, 该模型具有共享相同特征层的年龄歧视网络的辨别网络, 通过交替训练融合网络和组合因子模型, 高级间扩展和类内紧凑性有效地分离跨年龄身份特征. Li 等^[12]等人提出一种新的距离度量优化驱动的深度学习框架, 用于年龄不变的人脸识别, 以端到端的方式同时学习特征表示和相似性度量, 使用新颖的优化方法和精心设计的培训策略培训联合网络.

① 收稿日期: 2018-07-10

基金项目: 重庆市教委科学技术研究项目(KJ1603701).

作者简介: 胡渝苹(1982—), 女, 硕士, 讲师, 主要从事计算机软件应用研究.

针对以上跨年龄人脸识别方法中识别精度低的问题,本文提出一种基于深度信念网络的跨年龄人脸识别方法.该方法将方向梯度直方图(Histogram of Oriented Gridients, HOG)和中心对称局部二值模式(Center Symmetric Local Binary Pattern, CSLBP)方法联合起来,融合提取图像轮廓特征以及纹理特征,然后使用深度信念网络对特征进行训练分类,实验结果表明,对于跨年龄的人脸识别,所提算法能够得到较高的识别精度.

1 基于深度学习的跨年龄人脸识别

1.1 HOG 与 CSLBP

方向梯度直方图 HOG 是图像的密集特征提取方法,在密集特征提取中,算法提取图像中的所有特征位置,通过捕获该区域中的结构形状,获得关于梯度的信息.图像被分成 4×4 个 8×8 像素的单元块.本文采用二阶 HOG 对图像进行特征提取,二阶 HOG 流程如图 1 所示.

对于图像中像素 (x, y) 处的方向梯度直方图梯度幅值和方向的计算如式(1)和式(2)所示.

$$m(x,y)=\sqrt{[I(x,y+1)-I(x,y-1)]^2+[I(x-1,y)-I(x+1,y)]^2}$$

(1)

式(1)表示 (x, y) 处的梯度幅值,其中 $I(x, y)$ 是人脸图像在点 (x, y) 处的像素值.

$$\theta(x,y)=\tan^{-1}\left[\frac{I(x,y+1)-I(x,y-1)}{I(x+1,y)-I(x-1,y)}\right]$$

(2)

式(2)表示 (x, y) 处的梯度方向,其中 $I(x, y)$ 是人脸图像在点 (x, y) 处的像素值.本文采用一阶梯度算子进行水平和垂直方向的梯度计算.

局部二值模式 LBP 产生相当长的直方图,难以在图像描述符中使用.为了解决该问题,使用比较相邻像素的修改方案,即中心对称局部二值模式 CSLBP,旨在使较少数量的 LBP 标签产生更短的直方图,更适合在区域描述符中使用.在 CSLBP 中,像素值不是与中心像素相比,而是相对于中心像素对称地与相对像素相比较.如图 2 所示,该方法比较中心对称的像素对,而不是将每个像素与中心像素进行比较.在 LPS 情况下,8 个邻居产生 256 个不同的二进制模式,而对于 CSLBP,则产生 16 个不同的二进制模式.

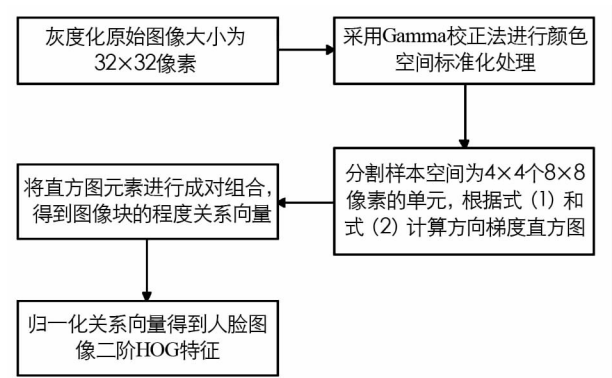


图 1 二阶 HOG 流程图

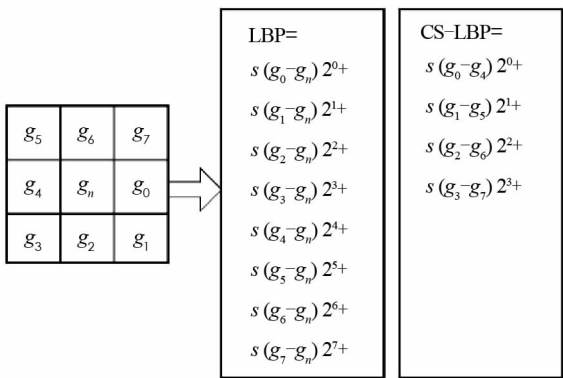


图 2 LPS 与 CSLPS 特征提取比较

1.2 深度信念网络

深度学习中深度信念网络(Deep Belief Networks, DBN)是包含多层隐层的概率生成模型,由多层限制波尔兹曼机(Restricted Boltzmann Machine, RBM)构成,RBM 模型如图 3 所示.与传统训练网络相比,DBN 解决了容易陷入局部最优的问题,DBN 是一个监督学习与非监督学习结合的过程,主要由预训练和微调两个部分组成.

RBM 是由可视层单元和隐藏层单元这两层节点组成的随机神经网络,可以看作一个二分结构无向图,在隐藏层中的两个连续层之间执行每个学习过程.

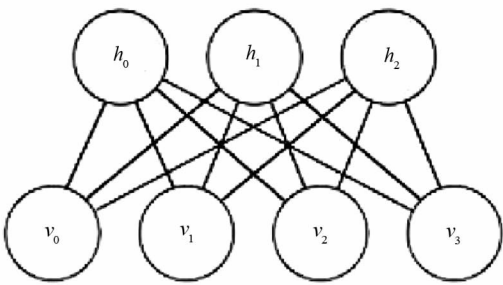


图 3 RBM 模型

RBM 是具有两层结构特定类型的马尔可夫随机场, 其中可见层二进制随机单元 $\mathbf{v} \in \{0, 1\}^D$ 连接到隐藏层二进制随机单元 $\mathbf{h} \in \{0, 1\}^F$, 状态 $\{\mathbf{v}, \mathbf{h}\}$ 的能量可以通过式(3) 计算.

$$E(\mathbf{v}, \mathbf{h}) = - \sum_{i=1}^D \sum_{j=1}^F w_{ij} v_i h_j - \sum_{i=1}^D a_i v_i - \sum_{j=1}^F b_j h_j \quad (3)$$

其中, D 和 F 分别是可见单元和隐藏单元的数量, $\mathbf{v}$ 是可见向量, $\mathbf{h}$ 是隐藏向量, a_i 和 b_j 分别是对应于可见和隐藏单元的偏置. w_{ij} 是可见单元和隐藏单元之间的连接权重. 随机变量 $(\mathbf{v}, \mathbf{h})$ 的联合概率密度可由 Gibbs 分布采样得到.

$$p(\mathbf{v}, \mathbf{h}) = \frac{e^{-E(\mathbf{v}, \mathbf{h})}}{\sum_{\mathbf{v}} \sum_{\mathbf{h}} e^{-E(\mathbf{v}, \mathbf{h})}} \quad (4)$$

Gibbs 采样的每次迭代都使用式(5) 和式(6) 中改了函数更新所有隐藏单元和所有可见单元.

$$p(h_j = 1 | \mathbf{v}) = f\left(\sum_i w_{ij} v_i + b_j\right) \quad (5)$$

$$p(v_i = 1 | \mathbf{h}) = f\left(\sum_j w_{ij} h_j + a_i\right) \quad (6)$$

式(5) 和式(6) 中 $f(\cdot)$ 表示 sigmoid 激活函数.

深度信念网络由多个 RBM 叠加形成, 构成多个随机变量组合的有向无环图. DBN 这个概率生成模型是观察数据和标签之间的联合分布, 其做工原理即使用底层 RBM 来训练下一层 RBM, 最终得到一组 RBM. DBN 结构如图 4 所示. 含有一层隐藏单元的 DBN, 可视单元和隐藏单元之间的联合分布可以表示为

$$p(\mathbf{v}, h^{(1)}, \dots, h^{(l)}) = p(\mathbf{v} | h^{(1)}) \dots p(h^{(l-1)} | h^{(l)}) \quad (7)$$

其中, $\mathbf{v}$ 是深度信念网络可见层, $h^{(k)}$ 表示第 k 层隐藏单元, k 的取值是 $[1, l]$, $\mathbf{v} = h^{(0)}$, 则第 k 层隐藏层与下一隐藏单元之间满足概率分布.

$$p(h^{(k)} | h^{(k+1)}) = \prod_i p(h_i^{(k)} | h^{(k+1)}) \quad (8)$$

$$p(h_i^{(k)} = 1 | h^{(k+1)}) = f\left(\sum_j w_{ij}^{(k)} h_j^{(k+1)} + a_i^{(k)}\right) \quad (9)$$

其中, $a_i^{(k)}$ 表示第 k 层的偏置, $w_{ij}^{(k)}$ 表示第 k 层和第 $k+1$ 层之间的权值.

对于深度信念网络预训练, 此阶段仅需要未标记的数据, 每层的贪婪分层无监督训练原理可以用 RBM 作为每层的构建块. 经过预训练以后, 进行深度信念网络微调, 此阶段仅需要可用的标记数据进行微调, 每次为一个层提供一组合理的权重, 对每一层进行无监督学习, 最后实现对整个网络监督学习的微调.

1.3 本文方法流程

深度学习与 HOG, CSLBP 的人脸识别算法流程如下所示:

- 1) 获取训练集和测试集的二维 HOG 特征, 具体获取步骤如图 1 所示, 得到二维 HOG 特征为 y_1 ;
- 2) 获取训练集和测试集的 CSLBP 特征, 用 CSLBP 算子扫描预处理人脸图像, 得到对应的 CSLBP 编码图像, 用双线性插值来消除编码图像均匀子块图像的边缘效应. 使用直方图统计子块图像的 CSLBP 特征, 顺序级联所有块直方图特征, 形成图像的 CSLBP 特征向量 y_2 ;
- 3) 按照 $\mathbf{Z} = [y_1/\sigma_1 \quad y_2/\sigma_2]^T$ 归一化融合二阶 HOG 特征与 CSLBP 特征, 得到融合特征 $\mathbf{Z}$, 其中 σ_1 和 σ_2 分别为两种特征的标准差;
- 4) 将训练集融合特征作为 DBN 可视层的输入, 对各层进行训练得到所需的最优参数;
- 5) RBM 训练过程是通过使用比散度算法(Contrastive Divergence, CD) 作为梯度近似的 Gibbs 采样来更新权值;

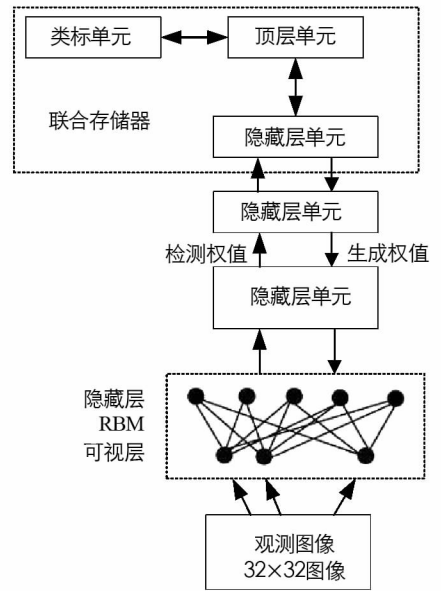


图 4 深度信念网络结构

6) 对每层的权值进行微调, 然后采用反向传播微调整个网络, 使用 Logistic 回归分类器获得测试样本标签, 实现跨年龄人脸识别, 并得到识别率.

本文方法的流程见图 5.

2 实验结果与分析

本文实验数据集从 FGNET 数据集、MORPH 数据集和 CACD 数据集按照条件(首先收集名人名单进行数据集构建, 然后收集不同出生年份的名字)选择图片形成一个综合数据库, 其中包含 16 000 个年龄从 16—62 岁名人的 16 万个图像. 数据集不同的人有不同的年龄, 发现 2 000 个名人具有超过 200 000 个面部图像, 为了移除重复图像, 本文使用重复检测算法删除嘈杂图像后, 最终数据库中有 168 000 张图像. 实验所用的硬件环境为 Intel Core i7, 8 G 内存的笔记本电脑, 使用仿 MATLAB 作为仿真平台. 实验中将本文算法分别与 DBN 算法、LBP+DBN 算法进行比较. 图 6 是实验中用到的部分图像.

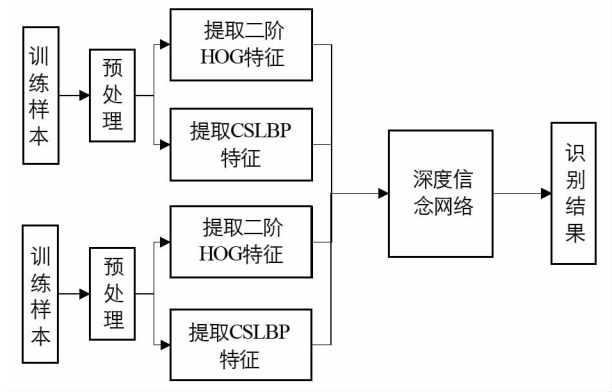


图 5 深度信念网络的跨年龄人脸识别流程图

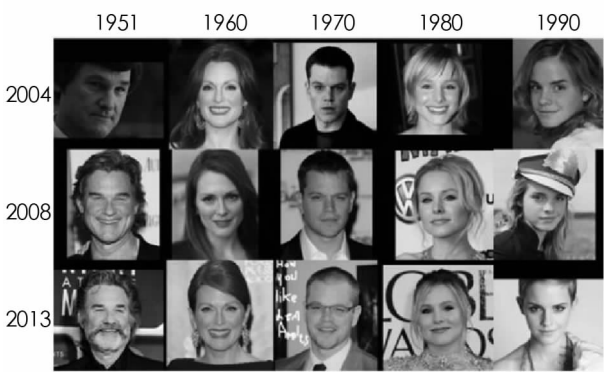


图 6 跨年龄人脸数据库中部分图像

在本文算法中, 深度信念网络隐藏单元数和人脸图像分块数是影响人脸识别效果的两个因素, 因此本文将两个因素对本文算法进行验证. 跨年龄人脸识别率随着深度信念网络中隐藏单元数的变化而发生变化, 这是因为深度学习受到学习有效性的影响. 表 1 给出了在不同的隐藏单元层数条件下, 本文算法与其他算法的识别率以及消耗时间消耗的结果, 图像分块大小为 4×4 个 8×8 像素的单元块.

表 1 不同隐藏层单元数下人脸识别性能

算法	隐藏单元数	识别率/%	训练时间/s	特征匹配时间/s	总时间/s
DBN	1 000	30.85	147.14	8.67	153.5
	2 000	51.17	381.49	13.78	397.3
	3 000	67.47	718.28	28.28	748.62
	4 000	80.59	1162.5	44.37	1 207.89
	5 000	87.32	1704.05	65.45	1 771.53
LBP+DBN	1 000	30.27	469.81	11.02	482.86
	2 000	51.58	848.89	20.83	880.63
	3 000	67.67	2 743.55	76.56	2 822.11
	4 000	82.67	3 150.33	93.84	3 317.82
	5 000	89.34	3 483.55	113.31	3 651.81
本文方法	1 000	31.9	154.55	4.69	161.27
	2 000	53.94	401.24	15.27	418.54
	3 000	78.69	764.58	26.42	793.03
	4 000	88.64	1 171.89	45.27	1 219.19
	5 000	98.5	1 740.97	66.03	1 809.03

由表 1 中数据可以得出, 人脸识别率随着网络中隐藏单元数的增加而提高, 同时人脸识别消耗时间随着增加. 主要原因是深度学习隐藏单元增加导致整个网络计算量增加, 本文深度学习人脸识别算法的识别

率优于 DBN 算法和 DBN+LBP 算法的识别率。

当深度信念网络隐藏层单元一定时，图像的分块大小和 CSLBP 的算子也影响人脸识别率，表 2 给出了不同分块大小条件下，本文方法的识别率对比结果，其中 CLBP4-1²-CLBP16-2² 是 CSLBP 的不同算子。

表 2 不同分块数下人脸识别率 %

算子	1×1	2×2	3×3	4×4	5×5	6×6
CLBP4-1 ²	66.5	84.5	94	96.5	96	95
CLBP8-1 ²	85.5	92.5	96.5	98	96	96
CLBP8-2 ²	87.5	93	96.5	98.5	96.5	97
CLBP16-2 ²	89	96	96	98	94.5	94.5

由表 2 中数据可以得到，在 CLBP8-2² 算子条件下，分块为 4×4 时识别率最好，在后续实验中，将使用 CLBP8-2² 算子和 4×4 分块大小的条件进行实验。

为了说明本文算法的有效性和优越性，将本文算法与文献[4]、文献[5]、文献[9]和文献[11]中算法进行比较，对综合数据库中跨年龄人脸图像进行识别，识别率结果见图 7。

从图 7 中可以看出，本文算法对于跨年龄人脸识别率高于其他 4 种算法，这是因为本文算法中 HOG-CSLBPS 方法能够获得更多包含结构和强度信息的图像融合特征，并使用深度信念网络对降维特征进行预学习与微调，最终提高了跨年龄人脸识别精度。

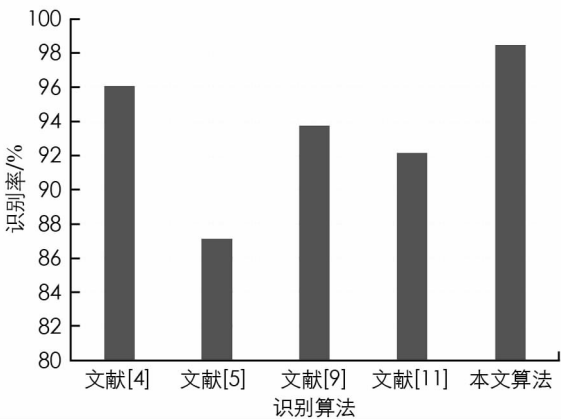


图 7 不同算法的跨年龄人脸识别率

3 结 语

本文提出一种 HOG 和 CSLBP 融合特征的深度学习跨年龄人脸识别算法，该算法使用 HOG 和 CSLBP 组合用于获得图像的结构和颜色信息，然后将融合特征作为深度信念网络输入，对特征图像进行预学习与微调，最后使用 Logistic 回归对人脸图像进行分类。实验结果表明，本文方法能够实现跨年龄人脸的识别，且识别精度优于其他算法，说明本文算法行之有效。未来工作是对本文算法进行深入研究，进一步减少本文算法的时间消耗。

参考文献：

[1] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: a Unified Embedding for Face Recognition and Clustering [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015.

[2] WEN Y D, ZHANG K P, LI Z F, et al. A Discriminative Feature Learning Approach for Deep Face Recognition [M]//Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016.

[3] KARCZMAREK P, KIERSZTYN A, PEDRYCZ W, et al. An Application of Chain Code-based Local Descriptor and Its Extension to Face Recognition [J]. Pattern Recognition, 2017, 65: 26-34.

[4] 嵇新浩. 基于 Gabor 和 HOG 的新型人脸识别算法 [J]. 西南师范大学学报(自然科学版), 2017, 42(12): 96-104.

[5] 潘禹岐, 江铭炎, 张振月. 基于流形的局部加权协从表示人脸识别 [J]. 计算机应用研究, 2016, 33(7): 2206-2209.

[6] DU L, LING H B. Cross-age Face Verification by Coordinating with Cross-face Age Verification [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015.

[7] CHEN B C, CHEN C S, HSU W H. Cross-Age Reference Coding for Age-Invariant Face Recognition and Retrieval [M]//Computer Vision - ECCV 2014. Cham: Springer International Publishing, 2014.

[8] JUEFEI-XU F, SAVVIDES M. Subspace-Based Discrete Transform Encoded Local Binary Patterns Representations for Robust Periocular Matching on NIST’s Face Recognition Grand Challenge [J]. IEEE Transactions on Image Processing, 2014, 23(8): 3490-3505.

[9] ZOU H S, HU H F. Cross-age Face Recognition Using Reference Coding with Kernel Direct Discriminant Analysis [C]//2017 IEEE International Conference on Image Processing (ICIP). Beijing: IEEE, 2017.

[10] LI Z F, GONG D H, LI X L, et al. Aging Face Recognition: a Hierarchical Learning Model Based on Local Patterns Selection [J]. IEEE Transactions on Image Processing, 2016, 25(5): 2146-2154.

[11] LI H X, HU H F, YIP C. Age-Related Factor Guided Joint Task Modeling Convolutional Neural Network for Cross-Age Face Recognition [J]. IEEE Transactions on Information Forensics and Security, 2018, 13(9): 2383-2392.

[12] LI Y, WANG G R, NIE L, et al. Distance Metric Optimization Driven Convolutional Neural Network for Age Invariant Face Recognition [J]. Pattern Recognition, 2018, 75: 51-62.

Cross-Age Face Recognition Algorithm Based on HOG-CSLBP and Deep Learning

HU Yu-ping

Big Data Institute, Chongqing Water Resources and Electric Engineering College, Chongqing 402160, China

Abstract: Aiming at the problem of low recognition accuracy in face recognition, a cross-age face recognition algorithm based on deep learning has been proposed. In this method innovatively, the combination of Histogram of Oriented Gradient (HOG) and Center Symmetric Local Binary Pattern (CSLBPS) has been conducted for face image feature extraction and the image fusion features been obtained including the structure and intensity information, and then the binary tree been used to feature the features. The dimensionality reduction feature has been used as the visual layer input of the deep belief network, which makes up for the defect that the depth of the new belief network cannot meet the local features of the image. The test samples have been learned through the trained deep network model, and the features been classified and identified at the top level of the deep belief network. The experimental results show that the proposed method can realize face recognition with high precision, and compared with other methods, the performance of this method is better than other methods, which shows the feasibility and effectiveness of the proposed method.

Key words: Histogram of Oriented Gradient; Central Symmetric Local Binary Pattern; Binary Tree; Deep Belief Network; Cross-Age Face Recognition

责任编辑 夏 娟