J. Vis. Commun. Image R. 61 (2019) 236–249

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Two-stages based facial demographic attributes combination for age estimation ☆

Mohammed-En-nadhir Zighem [a,*], Abdelkrim Ouafi [a], Athmane Zitouni [a], Yassine Ruichek [b], Abdelmalik Taleb-Ahmed [c]

[a] Laboratory of LESIA, Department of Electrical Engineering, University of Mohamed Khider, P.O. BOX 145 RP, Biskra, Algeria
[b] Le2i EA7508, University Bourgogne Franche-Comté, UTBM, Belfort F-90010, France
[c] IEMN, UMR CNRS 8520, Department OAE, University of Valenciennes and Hainaut-Cambresis, UVHC, Valenciennes, France

ARTICLE INFO

Article history:
Received 1 December 2018
Revised 24 February 2019
Accepted 30 March 2019
Available online 4 April 2019

Keywords:
Age estimation
Demographic classification
Feature extraction
SVR
SVM

ABSTRACT

Automatic age estimation from face images is a topic of growing interest nowadays, because of its great value in various applications. The main challenge in automatic facial age estimation task comes from the large intra-class facial appearance variations due to both gender and race attributes. To this end, in this paper we propose a complete approach for age estimation based on demographic classification. The proposed approach consists of three main parts: (1) Automatic face detection and alignment to extract only the regions of interest. (2) Feature extraction from facial region images using Multi-level face representation. (3) Two-Stages age Estimation (TSE). The main idea of TSE is to classify the input face image into one of demographic classes, then estimate age within the identified demographic class. The experimental results demonstrate that our proposed approach can offer better performance for age estimation when compared to the state-of-the-art methods on MORPH-II, PAL and a subset of LFW databases.

1. Introduction

The widespread topic of automatic demographic classification has been receiving significant attention in recent years. In the facial demographic classification context, gender and race classification has been addressed by several research works. In the recent decade, age estimation has been demanding growing interest, because of its great value in practice such as, business intelligence, access control, human-computer interaction, electronic customer relationship management (*e*CRM). In addition, the estimated age can prevent underages from purchasing cigarettes or alcohol from vending machines [60].

The main goal of age estimation is to predict individuals age based on their face images. In the literature, many methods have been proposed to deal with age estimation. Most of them estimate directly age as a single demographic attribute [10,24]. However, there are some methods which estimate all demographic attributes together [30,11,69]. Despite recent efforts, automatic age estimation still has some difficult problems. Moreover, age estimation has several challenges such as environment settings, temporal variations and especially the large intra-class facial appearance variations due to both gender and race attributes. Fig. 1 shows that race attribute, such as black, white and asian races, could lead to vastly different facial appearances of persons being in the same age. For instance, Guo and Mu in [27], studied the influence of gender and race on age estimation process. They founded that age estimation can be impacted by the gender and race differences considerably.

In this paper, we present a complete framework for age estimation, involving a novel Two-Stages-Estimator (TSE). Our approach consists of three main parts; (1) Automatic face detection and alignment to extract only the regions of interest (facial regions) and to correct the position and the size of faces. (2) Feature extraction from the facial regions image including both global and local texture features. (3) Tow-Stages-Estimator, where the input face is first classified into a specific demographic class using Support Vector Machines (SVM) and then a specific regressor is selected to estimate the exact age using Support Vector Regression (SVR). The performance of our proposed approach is evaluated on MORPH-II, PAL and a subset of LFW databases, and the numerical experiments demonstrate that our TSE obtains better estimation ability in comparison to state of the art methods including Convolutional Neural Networks.

☆ This paper has been recommended for acceptance by Zicheng Liu.

* Corresponding author.
E-mail addresses: zighem25@gmail.com (Mohammed-En-nadhir Zighem), ou_karim@yahoo.fr (A. Ouafi), zitou_a@yahoo.fr (A. Zitouni), yassine.ruichek@utbm.fr (Y. Ruichek), Abdelmalik.Taleb-Ahmed@univ-valenciennes.fr (A. Taleb-Ahmed).

https://doi.org/10.1016/j.jvcir.2019.03.025
1047-3203/

Fig. 1. Examples of different facial appearances of persons at the same age.

Our work has the following main contributions. (i) A complete framework for age estimation based on demographic classification is proposed; (ii) Three different age estimation methods are evaluated; (iii) Performing a hierarchical age estimation method with two different orders and investigating the order that provides better performances. (iv) Demonstrating that the age estimation accuracy is closely depending on the demographic grouping.

The remainder of this paper is organized as follows. In Section 2, we briefly reviewed related previous works. In Section 3, the proposed approach is presented. Experimental setup and results are shown in Section 4. Finally, the conclusions along with future works are given in Section 5.

2. Related works

In the literature, a lot of approaches have been proposed for facial age estimation. Most of them published prior to 2016 are reviewed in a survey paper [13]. In general, the existing approaches for age estimation can be divided into two categories based on the estimation method: direct age estimation [25,26] and age estimation with demographic attributes [6,5,35]. Direct age estimation is an approach that directly predicts the real age or age group as a single facial demographic attribute. Whereas demographic estimation determines the all facial demographic attributes together (i.g., age, gender and race). In this section, we will briefly review some of these works.

2.1. Age estimation

Age estimation is an automatic process of associating the facial image with a classification (year range) or a regression problem (exact age) [26]. The earliest published paper investigating the age classification was the work of Kwon and Lobo [42], where they presented an age classification method based on facial images by computing distance ratios of different facial features (i.e., eyes, nose, mouth, chin,...etc) and detecting the presence of wrinkles. Their method can classify the input faces into one of three age groups (babies, young adults, and senior adults).

Later the exact age estimation was addressed as a regression problem in [43]. Geng et al. [23] proposed an approach for age estimation based on facial aging patterns. The aging pattern was represented by a method called AGing pattErn Subspace (AGES) [22]. The age is indicated through the position of the face in the aging pattern. In [31] Guo et al. developed biologically inspired aging features, through changing the original bio-inspired models by proposing a novel STD operation in creating C_1 features based on Gabor filter responses. They evaluated their approach on YGA and FG-NET databases. Günay and Nabiyev [24] fused local texture features for age estimation. These features are based on Centrally Overlapped Blocks (COB) approach that captures the related information between the blocks of face image. The features are extracted with three texture descriptors: Local Binary Patterns (LBP) [2], Weber Local Descriptor (WLD) [8] and Local Phase Quantization (LPQ) [59]. Then a specific age is estimated by using the Multiple Linear Regression (MLR). Dibeklioglu et al. [14] addressed the presence of facial expressions in estimating age. They combined the dynamic features that are extracted from facial expressions with the appearance features to train the classifiers/regressors model. Their approach showed that smile dynamics can improve the age estimation accuracy.

Additionally, the hierarchical method [73,35,69], which is a combination of classification and regression techniques is also investigated. In [11] Choi et al. designed a hierarchical age estimator consisting of age groups classification followed by age regression based on facial local features (i.e., the wrinkle and skin features), and facial global features (i.e., the appearance and shape of a face), which are combined into a feature vector. Their experiments are conducted on BERC, PAL and FG-Net databases. Also, Luu et al. [53] introduced a hierarchical age estimation method based on a support vector machine (SVM) and support vector regression (SVR) to discriminate between two stages of human development (adulthood and childhood) and estimate a specific age respectively.

Quite recently, deep learning schemes, especially Convolutional Neural Networks (CNNs) [70], have been successfully used for age estimation issue. Hu et al. [37] proposed a CNNs scheme to facial age estimation without age labels by using the age difference information with three kinds of loss functions (i.e., cross entropy loss, entropy loss, and K-L divergence distance). While Rothe et al. [64] offered a solution to real and apparent age estimation by proposing an approach called Deep EXpectation (DEX). They used a deep CNN network pre-trained on the large ImageNet images [65] with VGG-16 architecture [67] followed by a softmax function to expect value formulation for age regression. Their results are reported on FG-NET, MORPH II and CACD datasets for estimating the biological (real) age. Liu et al. [48], exploited the label correlation among face samples in the transformed subspace and proposed a label-sensitive deep metric learning (LSDML) approach for facial age estimation. Then they extended it to a multi-source LSDML (M-LSDML) by using the correlation of multi-source face aging datasets to learn the label-sensitive feature similarity. Their experimental results showed the effectiveness of their approach on MORPH II, AdienceFaces, FG-NET, FACES and ChaLearn databases.

In addition, feature learning methods were incorporated to achieve better facial age estimation performance. Lu et al. [51], learned discriminative local face descriptor directly from raw pixel values into low-dimensional binary codes and encoded them into a real-value histogram feature for face representation by proposing a cost-sensitive local binary feature learning (CS-LBFL). Also, they proposed cost-sensitive local binary multi-feature learning (CS-LBMFL) method to exploit complementary information. Their CS-LBFL and CS-LBMFL are evaluated on FG-NET, MORPH II, LifeSpan, and FACES databases.

2.2. Demographic estimation

Classifying demographic attributes from the human facial images was first introduced by Yang and Ai [72]. The authors introduced a demographic classification method involving age group classification, namely child, youth and old. They considered Local Binary Patterns Histogram(LBPH) feature for ordinary binary classification problems. The step descent method was applied to find an optimal reference template which is used as a measurement of confidence for classification. Their method achieved 6.82%, 7.88% and 12.5% error rates for age classification on SNAPSHOT, FERET and PIE databases respectively.

Guo and Mu [30] proposed a complete framework that can estimate the age, gender, and race traits jointly within a single step. They adopted two methods of linear dimensionality, which are the Partial Least Squares (PLS) model [63] and Canonical Correlation Analysis (CCA) based methods [36]. These methods are also used to reduce the dimensionality of the original feature space. Their framework produced a Mean Absolute Error (MAE) of 3.92 years on the MORPH-II database. Another approach to determine the demographic attributes proposed by Hadid and Pietikäinen [34] considered LBP based spatio-temporal representation as a baseline system for combining spatial (i.e., facial structure information) and temporal (i.e., dynamics information) features for facial demographic classification from video sequences. Moreover, the correlation between the frames through manifold learning has been encoded.

Further in [33] the authors addressed the facial appearance variations which are considered as the more challenging problem of automatic demographic estimation from face images acquired in the wild (real-life face images) by presenting a complete framework that involves face normalization, feature extraction, and age group (or exact age), gender, and race estimation. The Biologically inspired features (BIF) and Support Vector Machines (SVM) were used in extraction and classification steps respectively. The same problem was addressed in [18], by presenting a robust face alignment technique. They provided a unique more challenging benchmark of face images, acquired under real-world conditions, labeled only for age and gender, to develop their proposed method and to evaluate performances. Duana et al. [15] introduced a hybrid architecture to process age and gender classification, which includes: CNNs to extract features from the input face images and Extreme Learning Machine (ELM) in a hierarchical fashion to classify the intermediate results.

In the previous related approaches cited above, there is no relation between the different demographic attributes in the classification step. In other words, each attribute has been classified independently. However, many researchers have been attempted to develop new approaches that can estimate the facial age based on the results of race and gender classification. For instance, Han et al. [35] presented a hierarchical approach consisting of three binary classifiers to build a two-level binary decision to predict gender and race attributes followed by a separate SVM regressor trained within each group to make an accurate age prediction. More recently, the same idea has been explored by Bekhouche et al. [5]. They proposed a hierarchical classifier with three layers to adopt a learning system for facial demographic attributes estimation. According to the highest accuracy of race estimation in their experiments, they chose to estimate the race of the input face as root of their hierarchical estimator. The gender is then estimated based on the race predicted using a corresponding classifier. Finally, based on the predicted race and gender, the age is estimated.

Table 1 summarizes and compares some well known and recent existing age estimation approaches.

3. Proposed approach

In this section, we present our proposed approach for facial age estimation relying on demographic attributes. In our work, we focus on the gender and the race attributes, since they are two of the most frequently facial demographic attributes reported in literature that can influence on age prediction process. Our approach consists of 3 main parts: *(i)* Automatic face detection and alignment, *(ii)* feature extraction, and *(iii)* demographic classification & age estimation. The general scheme of our complete approach is illustrated in Fig. 2. Detailed description of each part is provided in the following subsections.

Table 1
A comparison of some well known and recent age estimation approaches.

Publication	Approach	Dataset	Performance		
			Age (MAE)	Age group (%)	CS (%)
Yang & Ai, 2007 [72]	LBP Real AdaBoost	FERET PIE	N/A	92.1% 87.5%	N/A
Shan, 2010 [66]	LBP+Gabor SVM	Images of Groups	N/A	55.9%	N/A
Hadid & Pietikäinen, 2013 [34]	Manifold learning	Privet	N/A	83.1%	N/A
Guo & Mu, 2014 [30]	BIF KCCA+SVM	MORPH II	3.92	N/A	84.02%
Wang et al., 2015 [70]	Convolutional Neural Network	MORPH II FG-NET	4.77 4.26	N/A	N/A
Han et al., 2015 [35]	BIF SVM	MORPH II PCSO A subset of LFW FG-NET	3.6 4.1 7.8 3.8	N/A	77.80 72.28 43.35 78.91
Bekhouche et al., 2017 [5]	PM-LPQ+PM-BSIF SVM	MORPH II PAL Images of Groups	3.50 5.00 N/A	N/A N/A 68.2%	75.13 57.50 N/A
Günay & Nabiyev, 2017 [24]	WLD+LPQ+LBP SVR	FG-NET PAL MORPH I	4.94 5.75 4.06	N/A	67.97% 73.29% 54.28%
Duan et al., 2018 [15]	CNN-ELM	MORPH II Adience	3.44 N/A	N/A 52.3%	70.11% N/A

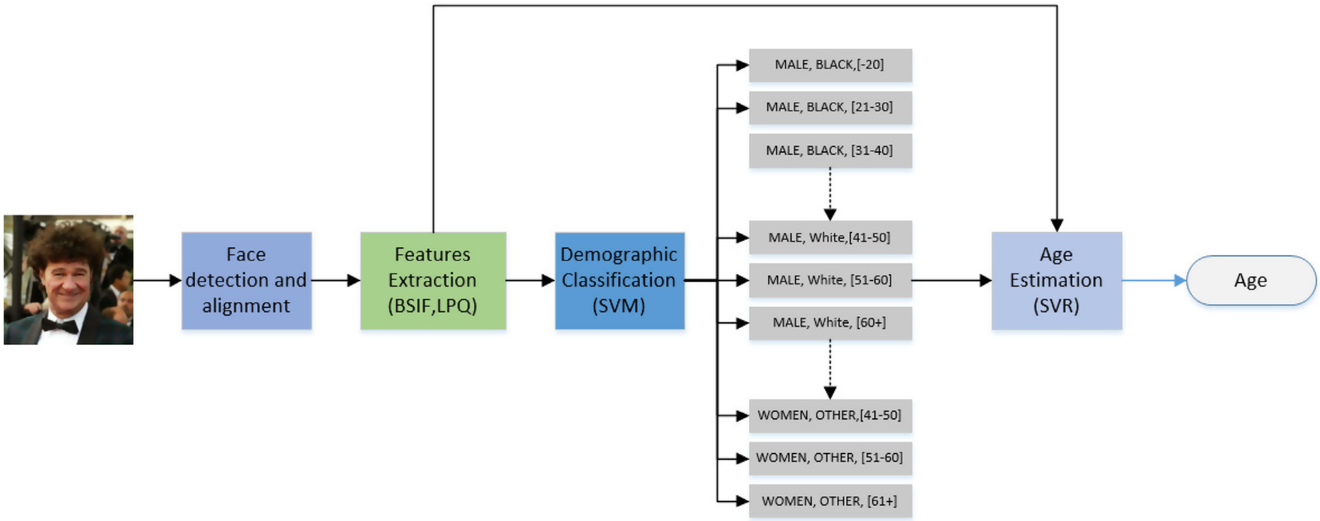


Fig. 2. The general scheme of the complete proposed approach.

3.1. Automatic face detection and alignment

The original face images can be affected by the traditional factors such as illumination conditions, pose variation, etc. These factors led to misalignment between face images. To overcome this issue, in recent years many methods have been proposed for face alignment (i.g., [47,71]). In this work, we adopted alignment process based on eyes center localization, which includes: (i) face and landmarks detection, (ii) eyes localization, (iii) pose correction, and (vi) regions of interest (ROI) cropping (see Fig. 4).

Firstly, to alleviate the influence of inconsistent colors, we transformed the input face color images to gray scale images. Next, we employed the off-the-shelf [41] face and 2D facial landmark detector to obtain the location of the face in each image. It detects 68 facial landmarks located on the mouth, nose, eyes, and eyebrows (Fig. 3). In this paper, we only used 12 facial landmarks which are located around the eyes to set (X_{right}, Y_{right}) and (X_{left}, Y_{left}) coordinates of right and left eye respectively, by applying the center of pressure formula. These two-points coordinates enable us to correct the pose by rotating the face images according to the rotation with an angle β which is defined as follows:

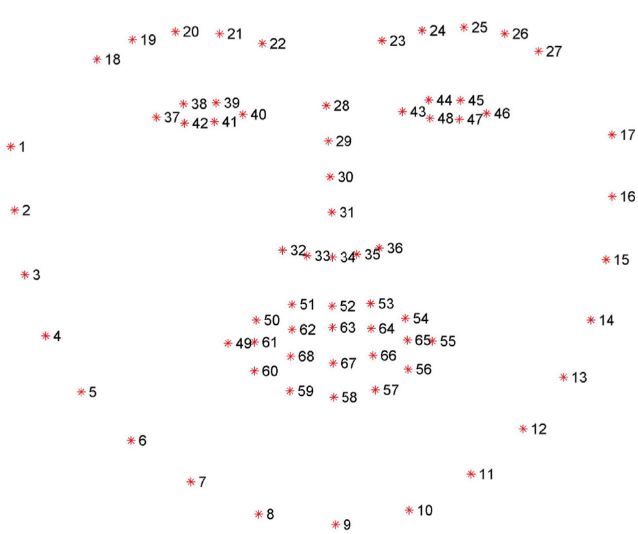


Fig. 3. 2D facial landmarks detected by Kazemi and Sullivan algorithm [41] (figure from google image).

$$\beta = \tan^{-1}\left(\frac{Y_{right} - Y_{left}}{X_{right} - X_{left}}\right) \quad (1)$$

Next, the rotated images are scaled to the same inter-pupillary distance d that is measured from the eyes centers. Finally, the ROIs (aligned faces) are cropped to a standard size of 200×200 pixels. These ROIs are used to train and test a CNN age estimation method [39]. Fig. 4 schematizes the above automatic face detection and alignment procedure.

3.2. Feature extraction

The feature extraction step is performed on each aligned face. In order to extract the features, the Multi-level (ML) face representation [57] is used. The main idea behind applying the ML face representation is to get various global and local texture features from the whole ROI. The ML representation represents the aligned face image at different levels (see Fig. 5). In each level, the ROI is divided into n^2 non-overlapping local blocks, where n is the number of levels (e.g., the first level is the aligned face without dividing). The features extracted from blocks in each level are concatenated to construct the level features vector. The order of concatenating is from left to right and from up to down. The features vectors of all levels are concatenated later to form the global face features vector. Fig. 5 depicts an example of features vector constructing with Multi-Level face representation for 3 levels.

In many researches on age estimation, variety of feature descriptors have been proposed in order to extract aging features from images. They fall into two categories: learning-based and handcrafted-based. Learning-based methods include Discriminant Face Descriptor (DFD) [44], Cost-Sensitive Local Binary Feature Learning (CS-LBFL) [52], Deep Binary Descriptor with Multi-Quantization (DBD-MQ) [17], Discriminative Deep Metric Learning (DDML) [50] and Context-Aware Local Binary Feature Learning (CA-LBFL) [16]. Handcrafted-based methods include Biologically Inspired Features (BIF) [31], AGing pattErn Subspace (AGES) [22], Local Binary Pattern (LBP) [2], Weber Local Descriptor (WLD) [8], and Histograms of Oriented Gradients (HOG) [12]. Some of the latter methods, such as LBP and HOG, consider the face image as a texture pattern. In this study, two texture descriptors are used: *Binarized Statistical Image Feature (BSIF)* and *Local Phase Quantization (LPQ)*. This choice is based on results reported in [5], where demonstrated that BSIF and LPQ descriptors are powerful texture

(a) Input face (b) Landmarks detection (c) Eyes localization (d) Pose correction (e) Regions of interest cropping

Fig. 4. Automatic face detection and alignment.

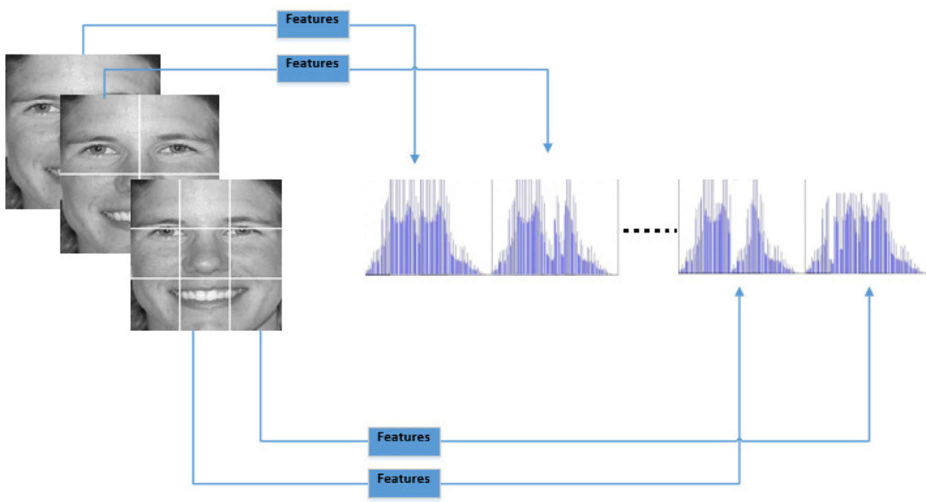


Fig. 5. Exemple of Multi-Level face representation for 3 levels.

descriptors for demographic estimation. These descriptors are explained in the following subsections.

3.2.1. Binarized Statistical Image Feature (BSIF)

A Binarized Statistical Image Feature (BSIF) [40] is a local image texture descriptor that characterizes a pixels surrounding (patch) by a binary code. The bits in the code string are generated by binarizing the convolution results between the patch X of size $M \times M$ pixels with a linear filter H_i $(i = 1 \dots n)$ of the same size. The convolution operation is defined as:

$$S_i = \sum_{v,u} X(v, n) * H_n(v, u) \tag{2}$$

Each bit $b_i(i = 1 \dots n)$ of the binary code is considered by setting a threshold at zero for the response S_i as shown in Eq. (3):

$$b_i = \begin{cases} 1 & \text{if } S_i > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Consequently, the number of filters used determines the length of the binary code. Notice that the filters are learnt from natural image patches by maximizing the statistical independence of the filter responses. In all our experiments, the BSIF descriptor has been used with filters of size 13×13 and 8 bits.

3.2.2. Local Phase Quantization (LPQ)

Ojansivu et al. in [59] proposed a spatial blurring method which used to build LPQ face description. In the frequency domain, the blurred image $G(u)$ is the multiplication of an original image $F(u)$ with the Point Spread Function (PSF) of the blur $H(u)$.

$$G(u) = F(u) \cdot H(u) \tag{4}$$

where u is a vector of coordinates $[\alpha, \beta]^T$. In LPQ, the phase is examined in local neighborhoods N_x at each pixel position x of the image. The short-term Fourier transform (STFT) is applied to compute local Fourier coefficients at four frequency points $u_1 = [a, 0]^T, u_2 = [0, a]^T$, $u_3 = [a, a]^T, u_4 = [a, -a]^T$ where 'a is a sufficiently small scalar to satisfy $H(u) > 0$.

The phase information in these coefficients can be counted by using a simple scalar quantization

$$q_i(x) = \begin{cases} 1 & \text{if } g_i > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

where g_i is the j-th component of the vector $G(u)$. The resulting eight binary coefficients $q_i(x)$ are represented as integer values between 0 and 255 using binary coding:

$$f_{LPQ}(x) = \sum_{j=1}^{8} q_j(x) * 2^{j-1} \tag{6}$$

Finally, the histogram of these integer values is considered as a feature vector.

3.3. Two-stages age estimation

The human age estimation using a generic classifier or regressor is an intrinsically challenging problem and many studies have been proposed to achieve small estimation error. As we mentioned above in Section 2, the age group classification (i.e., age grouping) is one of the famous solutions proposed for this problem. Another solution is introduced in [27] by dividing the prediction age into two phases: The first one classifies the input face images only into gender-race groups. Next, in the second phase, an age prediction model is learned on each classified gender-race group to estimate the age. This approach is conducted on a pre-selected subset of MORPH II database with the same number of face images from two races (Black and White) and two genders (Male and Female).

Unlike these studies, in this paper we present a novel Two-Stages Estimator (TSE), for human age estimation based on three demographic attributes (i.e., gender, race and age ranges). The TSE consists of two stages as shown in Fig. 2: (1) Demographic classification; (2) age estimation within demographic classes; In the first stage, the test subject features vector is automatically classified into one of different demographic classes using support vector machine (SVM), where each class has three demographic information integrated in the same label $Y_{i,j,k}$, where i, j and k denote gender, race, and age range labels respectively. The number of demographic classes is decided based on the number of races and genders in the available training set. For example, if the training set is labeled with two genders, three races and six age ranges, we will have thirty-six demographic classes. After the demographic classification stage is carried out, in the second stage, for each demographic class, an age estimation model is learnt using support vector regression (SVR). Noteworthy that the same features are used in both stages.

The recent CNNs based methods typically implement the extracting features and classifying in just one step with billions of arithmetic operations. They require high-cost hardware to implement high computational complexity of their algorithms [32]. In contrary to these methods, we intentionally adopted a pipeline based on feature extraction and feature classification, which are performed separately, like in [6,4,1]. This pipeline allows making series of stacked classifiers and avoiding disadvantages of the CNN methods.

Table 2
The age range, gender, and race distributions of subjects in the (a) MORPH II, (b) PAL and (c) a subset of LFW databases.

Age range	0–20	21–30	31–40	41–50	51–60	61+	Total
Female	1081	2264	2939	1832	355	17	8488
Male	8372	13236	12690	9427	2678	243	64,646
Black	7571	12077	11848	8626	2263	177	42562
White	1351	2571	3368	2491	740	79	10600
Other	531	852	413	142	30	4	1972
Total	9453	15500	15629	11259	3033	260	55134

(a)

Age range	0–20	21–30	31–40	41–50	51–60	61+	Total
Female	59	161	51	40	39	266	616
Male	57	199	27	32	13	102	430.
Black	43	84	13	22	7	39	208
White	58	189	65	50	45	325	732
Other	15	87	0	0	0	4	106
Total	116	360	78	72	52	368	1046

(b)

Age range	0–20	21–30	31–40	41–50	51–60	61+	Total
Female	48	291	302	246	152	62	1101
Male	52	379	605	693	795	586	3110
White	83	499	727	794	824	574	3501
Other	17	171	180	145	123	74	710
Total	100	670	907	939	947	648	4211

(c)

4. Experiments

In this section, we first introduce the databases and the evaluation protocols of our experiments. After that, we will give description about experimental setups. Finally, we discuss the results obtained on each database.

4.1. Databases

Three databases are used to evaluate the performance of the proposed approach: MORPH-II, PAL and a subset of LFW. The facial images of MORPH-II and PAL databases were collected under controlled and cooperative scenarios: frontal facial images, illumination variation, and facial expression. On the other hand, the subset of LFW database is a very challenging database, since its facial images were captured in uncontrolled environments (from the real life), non-frontal views, occlusions, and low image quality. To the best of our knowledge, there are no other publicly available databases labeled with the three demographic attributes needed for our algorithm.

The MORPH (Album 2) database. [61] is the most frequently used for age estimation in previous works. It contains about 55,000 face images of 13,618 individuals (11,459 males and 2159 females) with the age range from 16 to 77 years old. The MORPH II database is divided into three main races. Black faces about 77%, White faces 19% and the remaining 4% (Hispanic, Asia, India and other). See Table 2a for more details.

The Productive Aging Lab Face (PAL) database. [55] contains totally 1046 frontal face images from 580 subjects with different facial expressions (1.8 images per subject), with the age range from 18 to 93 years old. The PAL database can be divided into three main races: Black subjects (208 images), White subjects (732 images) and other subjects (106 images).

The Labeled Faces in the Wild (LFW) database. is the most widely used database for studying the problem of unconstrained face recognition. Han et al. [35], performed age, gender, and race estimation on a subset of LFW with 4211 subjects (one image per subject), where the face images have relatively small pose variations. Since the label of race and real age are not available, they collected them by using the Amazon Mechanical Turk (MTurk) crowdsourcing service with three workers per task to estimate age, gender and race of each face image. Some of face images from the three databases are shown in Fig. 6.

In our experiments, the entire age of each database is divided into six non-overlapping ranges of 10 years (i.e., decades of life: −20, 21–30, 31–40, 41–50, 51–60 and 61+). We avoided dividing the whole database into more than 6 ranges since the number of training samples in each group becomes too small.

4.2. Experiment setup

The demographic classification is a classification problem. While age estimation is naturally formulated as a regression task. To this end, we used LIBLINEAR library [19], which supports logistic regression and linear support vector machines (SVM). Their optimal parameters c and ϵ were found by a grid search on the training phase for the both stages of our proposed approach (demographic classification and age estimation).

In order to evaluate our proposed approach, we considered 5-fold cross validation and reported the average performance over the 5 folds, which is frequently used for MORPH II and PAL [56,68,5,49]. This protocol randomly considers 1/5 of data samples as test set, and the rest is used as training set. It assures that a sample is not in the both sets simultaneously. For this reason, the 5-fold cross validation protocol is used for all databases.

Fig. 6. Some facial images from the PAL (top), LFW (middle), and MORPH II (bottom) databases.

4.3. Evaluation metrics

To measure the performance of our approach, we adopted two widely used evaluation metrics in literature, the Mean Absolute Error (MAE) in years and the Cumulative Score (CS). The MAE is the average of the absolute error between the ground truth age and the predicted one:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|P_i - G_i| \quad (7)$$

where P_i is the estimated age, G_i is the corresponding ground truth, and N is the total number of samples.

The cumulative score (CS) is used to evaluate the age estimator performance at different absolute error levels, where the age estimation error is lower than a threshold value. The CS is calculated by:

$$CS(k) = \frac{N_{e\leqslant k}}{N} * 100 \quad (8)$$

where k is the threshold (years), $N_{e\leqslant k}$ is the number of test images on which the age estimation makes an absolute error no higher than k and N is the total number of the samples. High CS and low MAE values mean better performance.

4.4. Experimental results

In this study, our purpose is to investigate whether our proposed approach provides more accurate results than other age estimation approach. Towards this goal and to investigate the effectiveness of our approach, we conducted a comparative study for three automatic age estimation methods: (1) direct age regression without relying on results of other attributes (Fig. 7a); (2) Hierarchical age estimation method (Fig. 7b) using two different orders; (3) Our proposed approach TSE to estimate the age based on the result of combining gender, race and age ranges classification. In the next subsections, we present the experimental results of each method.

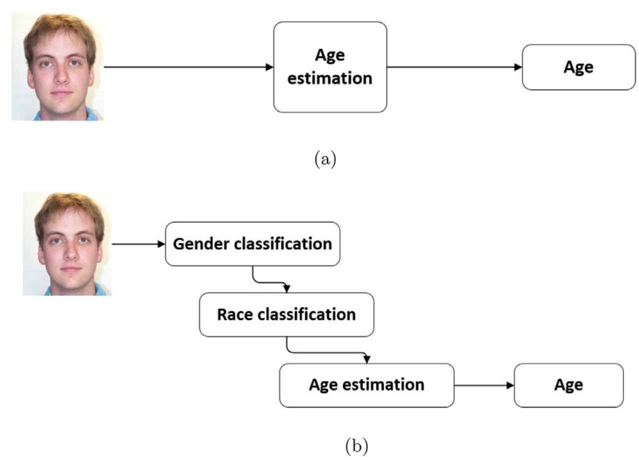


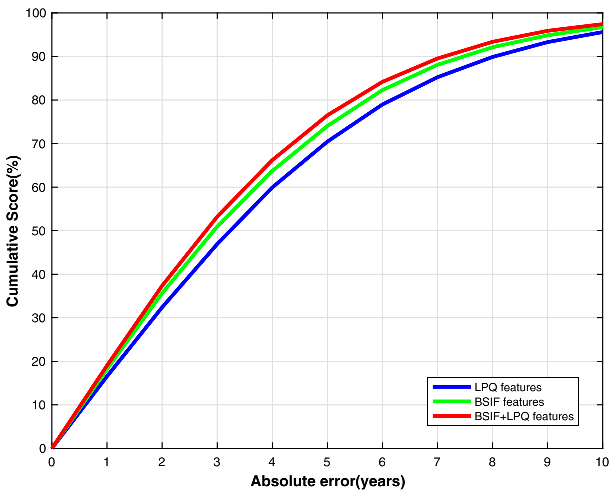
Fig. 7. Two different age estimation methods, (a) direct age estimation (b) Hierarchical age estimation.

4.4.1. Direct age estimation

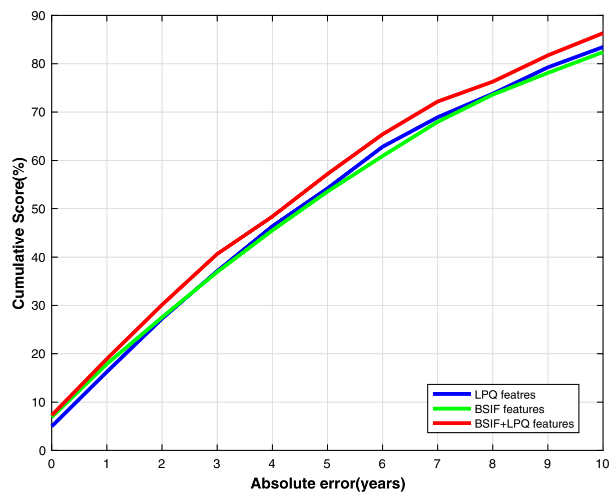
Direct age estimation (Fig. 7a) is the basic pipeline which has been proposed for age estimation [31,4], using a single step age regression. The MAE results for this method on MORPH II, PAL and a subset of LFW databases, with different image descriptors (BSIF, LPQ and BSIF+LPQ) are shown in Table 3. The cumulative score curves are illustrated in Fig. 8. Table 3 lists the results of direct age estimation using different texture features (BSIF, LPQ features and their combination) in terms of MAE. We can observe, when using BSIF features for MORPH II database, the MAE is

Table 3
Mean Absolute Error (MAE) of age estimation on MORPH II, PAL and subset of LFW databases (in years) using direct age estimation.

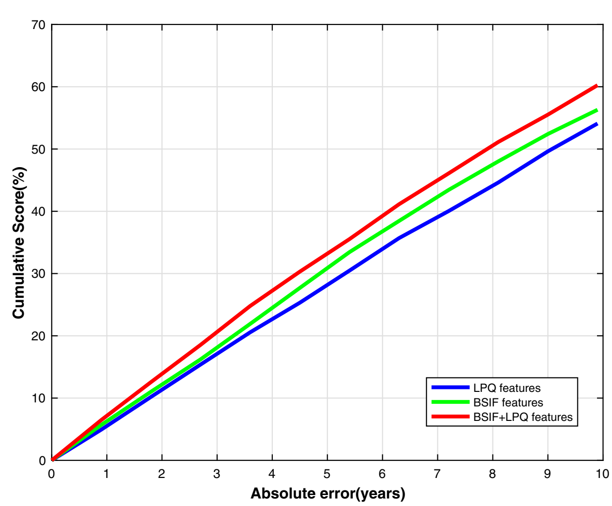
Database	BSIF	LPQ	BSIF+LPQ
MORPH II	3.68	3.88	3.58
PAL	6.01	5.88	5.55
Subset of LFW	9.74	10.18	9.07



(a)



(b)



(c)

Fig. 8. CS curves of direct age estimation at error levels from 0 to 10 years on (a) MORPH II (b) PAL and (c) Subset of LFW databases.

3.68 years, which slightly out-performs LPQ features providing a MAE of 3.88 years. However, the combination of BSIF and LPQ features achieves much lower MAE (3.58 years). In PAL and subset of LFW databases, the combination between BSIF and LPQ features also performs better than BSIF or LPQ features used separately.

Fig. 8, shows the cumulative scores of age estimation on MORPH II, PAL and subset of LFW databases at the error levels from 0 to 10 years. From Table 3 and Fig. 8, the age estimation performance is improved by the combination of BSIF and LPQ features in both MAE and cumulative score.

4.4.2. Hierarchical age estimation

In order to compare our proposed approach with an age estimation method based on demographic attributes, we performed the hierarchical age estimation method for LPQ and BSIF features using two different demographic attributes orders (G→R→AR→Real age) and (R→G→AR→Real age). Where G, R and AR denote gender, race and age range attributes respectively.

The hierarchical method has been studied for age estimation task. It consists of a series of classifiers to classify each attribute separately followed by a regressor to predict age (Fig. 7b).

Table 4 shows the performance of hierarchical age estimation method on MORPH II, PAL, and subset of LFW databases. One can see that the best age estimation results are always obtained with (R→G→AR→real age) order and BSIF+LPQ features on MORPH II and PAL databases. These results can answer the question: Which of the two orders leads to good performances? They make sure that the order adopted in [5], where the race is chosen as the root of the hierarchical estimator, provides better MAE on MORPH II and PAL databases than the order adopted in [35], where the gender is chosen as root. We can see also that the change in MAE value, which depends as said before on the order of attributes, is occurring

Table 4
Age estimation results of hierarchical method for MORPH II, PAL, and subset of LFW databases.

Database	Order of attributes		Features		
			BSIF	LPQ	BSIF+LPQ
MORPH II	G→R→AR	Gender(%)	98.53	98.36	98.91
		Race(%)	98.03	97.83	98.22
		Age ranges(%)	65.42	63.48	67.85
		Age(years)	4.02	3.70	3.70
	R→G→AR	Race(%)	98.12	97.74	98.22
		Gender(%)	98.89	98.50	99.05
		Age ranges(%)	67.25	64.03	68.58
		Age(years)	3.70	4.05	3.52
PAL	G→R→AR	Gender	96.27	94.65	96.75
		Race	97.04	96.65	97.23
		Age ranges(%)	82.79	80.11	82.41
		Age(years)	4.87	5.36	4.84
	R→G→AR	Race(%)	97.32	97.04	97.71
		Gender(%)	96.18	94.93	96.27
		Age ranges(%)	83.08	80.21	83.17
		Age(years)	4.70	5.25	4.69
Subset of LFW	G→R→AR	Gender(%)	93.90	93.56	94.49
		Race(%)	90.48	90.55	91.52
		Age range(%)	35.15	34.08	35.67
		Age(years)	8.54	8.91	8.20
	R→G→AR	Race(%)	90.93	90.81	91.81
		Gender(%)	93.26	92.54	94.04
		Age range(%)	33.84	34.48	35.64
		Age(years)	8.66	8.74	8.27

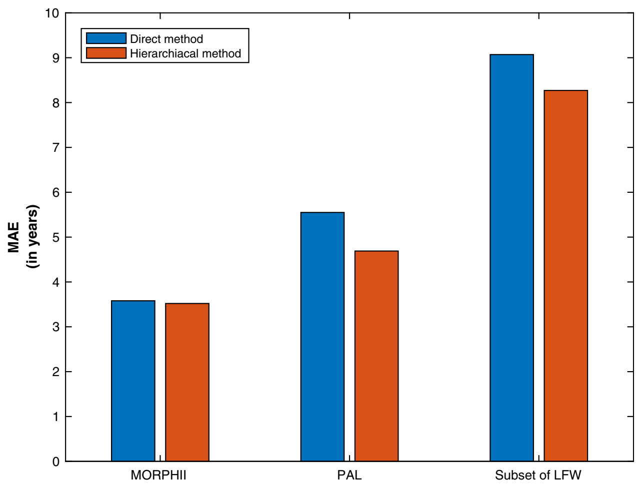


Fig. 9. Comparison of hierarchical and direct age estimation methods.

Table 5
TSE performance results.

Database	Stage	Features		
		BSIF	LPQ	BSIF+LPQ
MORPH II	Demographic Classification(%)	65.49	61.75	66.78
	MAE(years)	3.39	3.68	3.21
PAL	Demographic Classification(%)	79.54	75.81	79.73
	MAE(years)	4.67	4.95	4.49
Subset of LFW	Demographic Classification(%)	29.64	26.29	30.49
	MAE(years)	8.19	8.25	7.93

regardless of their accuracy. However, we note the opposite case on subset of LFW database, where the (G→R→AR→real age) order provides better MAE than the other one. It seems that the better order must start by the attribute providing the higher accuracy (Gender as root provides 94.49% of accuracy whereas Race as root provides only 91.81%).

Fig. 9 shows the comparison of hierarchical and direct age estimation method using BSIF+LPQ features. From the figure, it is clear that the hierarchical method provides more accurate results than the direct age estimation method.

4.4.3. Proposed Two-Stages age Estimator (TSE)

According to the results of the previous subsection, we found that the MAE can be decreased when the age estimation is performed across demographic attributes. In this subsection, the experimental evaluation of our TSE to age estimations is reported. The basic idea of TSE is to classify first the input test face image into one of the demographic classes. Then, based on this classification result, age estimation is performed on each classified group. We believe that estimating age within the same race, gender and age range class will provide smaller estimation errors. We conducted our experiments on MORPH, PAL and subset of LFW databases to evaluate the performance of the TSE. In Table 5, we organized the proposed approach results by grouping them according to the three databases. In each database, 2 results (2 sub-rows) are reported, demographic classification accuracy and MAE.

We can see from the last column in Table 5 that our proposed TSE improves the age estimation accuracy for all databases. MAEs of 3.21 years for MORPH II database, 4.49 years for PAL database and 7.93 years for subset of LFW database are obtained when using the combined (BSIF+LPQ) features. These results are much smaller than direct and hierarchical age estimation results which are shown in Tables 3 and 4 (direct: 3.58/5.55/9.07 years; hierarchical: 3.52/4.69/8.27 years for MORPH II, PAL and subset of LFW databases respectively). The improvements are achieved according to the accuracy of the demographic classification. For example, the demographic classification accuracy of LPQ and BSIF texture features for PAL database are of 75.81% and 79.54% respectively. After the combination of these features, the demographic classification accuracy is increased to 79.73%. The MAEs are reduced from 4.95 to 4.67 then to 4.49 years. The improvement of this demographic classification accuracy is not too much, but it has a great influence on age estimation result. The demographic classification accuracy increases the estimation error decreases.

Considering the performance of the proposed approach on subset of LFW database, demographic classification accuracies did not exceed the threshold of 30.49%. They are not high compared to accuracies on MORPH II and PAL databases, even the MAE is fairly high (7.93 years). Two possible explanations for these poor accuracies: (1) The race and the age which were used as the ground truth on the training phase are not available in LFW database, they are estimated by human [35] using (MTurk) crowdsourcing service [3]; (2) Most of subject images in LFW database are celebrity fig-

Table 6
Overall accuracy of the classication stage of hierarchical age estimation method and proposed TSE method.

Method		Classification accuracy	Database		
			MORPH-II	PAL	Subset of LFW
Hierarchical	G→R→AR Order	Overall (%)	88.32	92.13	73.89
		Correct (%)	65.23	78.75	30.03
		MAE(years)	3.70	4.84	8.20
	R→G→AR Order	Overall (%)	88.61	92.38	73.83
		Correct (%)	65.26	79.42	3.10
		MAE(years)	3.52	4.69	8.27
TSE	Demographic classification	Overall (%)	66.78	79.73	30.49
		MAE(years)	3.21	4.49	7.93.

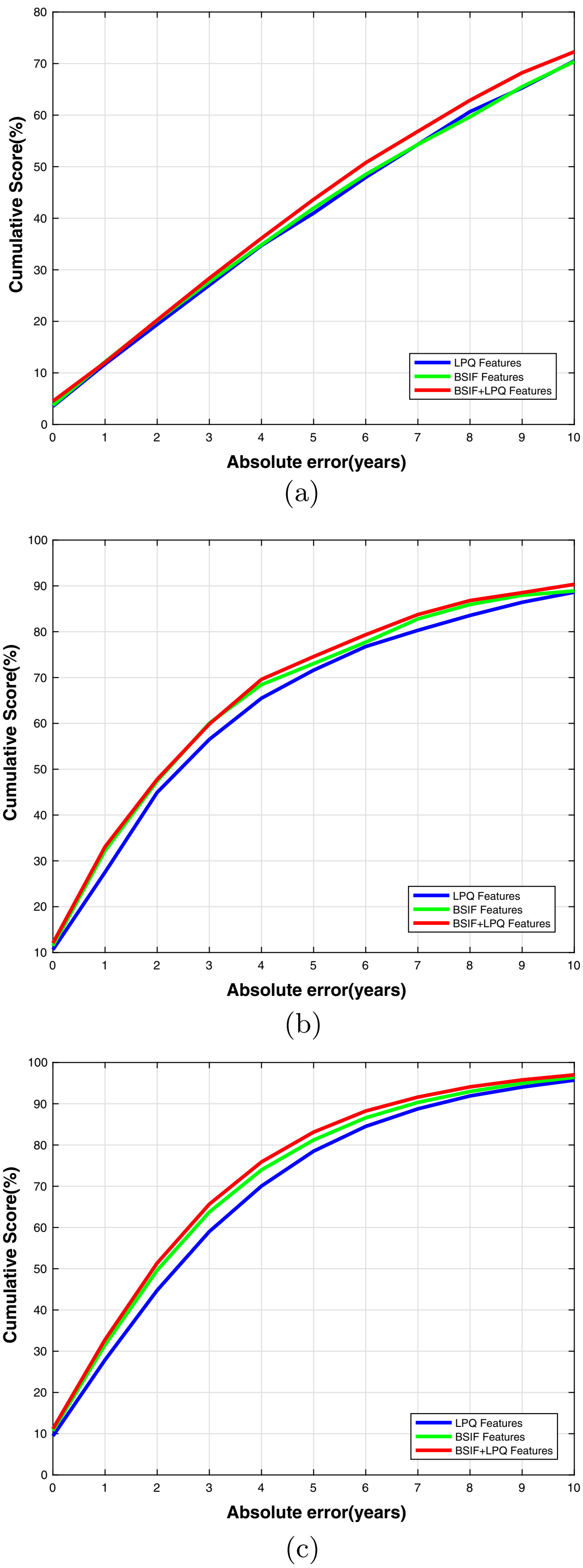


Fig. 10. CS curves of the proposed TSE approach at error levels from 0 to 10 years on (a) Subset of LFW (b) PAL and (c) MORPH II databases.

ures, collected from internet with large variation in pose, illumination, facial makeup and facial dynamics which make it difficult to estimate their real age.

Despite the age is estimated after classifying the integrated demographic attributes in our propose TSE, and after series of demographic attribute classifiers in hierarchical age estimation method, a question arises: why the proposed TSE approach outperforms the hierarchical method? To answer this question, we reported in Table 6 the overall accuracy of the classification stage (race, gender and age range classifiers) of hierarchical age estimation method on MORPH-II, PAL and subset of LFW database. We reported also in Table 6 the classification stage accuracies (extracted from Table 5) of the proposed TSE. Comparing the two results of Table 6, we can see that the overall accuracies of the classification stage of hierarchical method in the two orders are higher. Whereas, we observe the opposite in terms of MAE. This can be explained by misclassification that occurs in one or more of the hierarchical classifiers, which may cause accumulation of errors. For example, on PAL database, the overall accuracies of the classification stage of the hierarchical method are 92.13% and 92.38% with 4.84 years and 4.69 years in MAE for G→R→AR and R→G→AR orders, respectively. While, in our proposed TSE, the demographic classification accuracy is 79.73% with 4.49 years in MAE. If we look at the correct classification rates in race, gender and age range attributes taken simultaneously, the classification stage of the hierarchical method reached only 78.75% and 79.42% in G→R→AR and R→G→AR orders respectively, which are lower than the overall rates obtained by the classification stage of TSE method. This show that the errors accumulation, occurring in the classification stage of the hierarchical method, was discarded in TSE method thanks to integrating race, gender and age range attributes in a single classifier. Therein lies the outperformance of the proposed TSE than hierarchical age estimation method.

Fig. 10 shows the CS curves obtained by the proposed TSE approach, on MORPH, PAL and subset of LFW databases at error levels from 0 to 10 years. The ages of approximately 11.19% of the subjects in the MORPH II database, 12.05% of the subjects in the PAL database and 4.51% of the subjects in the subset of LFW database can be estimated with zero error level. Whereas when the error level increases the estimation accuracy also increases. Our proposed BSIF+LPQ based TSE is able to achieve cumulative scores of 83.13%, 74.75% and 43.64% for an absolute error of 5 years and 97%, 90.34% and 72.29% for an absolute error of 10 years for MORPH, PAL and subset of LFW databases respectively.

Next, in Tables 7 and 8 we compare the performance of our proposed age estimation approach with known and recent state-of-the-art approaches in terms of MAE and CS on MORPH II and PAL databases respectively. Existing methods adopted different experimental settings on MORPH-II. The 5-folds cross validation is introduced in [38,20,7,5]. 10-fold cross validation is another setting which is used in [51,46,21]. To compare performance of our proposed TSE with different state-of-art age estimation methods on MORPH-II database, we performed experimental evaluation with both settings. The upper part of Table 7 shows that the proposed method out performs all the compared handcrafted state-of-the-art methods on MORPH-II database, thanks to adopting demographic classification based age estimation. Further, in the lower part of Table 7, we compared the performance of our proposed TSE with age estimation CNN based methods. The results show that our proposed approach can offer better performance than some CNN-based age estimation methods such as [45,15,70,46,38].

Table 7
Comparison of the proposed approach with known and recent state-of-the-art approaches on MORPH II database.

Publication	Approach	Performance		Protocol
		MAE (years)	CS (%)	
Guo & Mu, 2011 [28]	BIF+KPLS	4.4	–	different split
Chang et al., 2011 [7]	OHRank	6.07	56.4	5-fold cross validation
Fernández et al., 2015 [20]	HOG+SVR	4.83	63.4	5-fold cross validation
Huerta et al., 2015 [38]	rCCA	4.24	71.17	5-fold cross validation
Lu et al., 2015a [51]	CS-LBMFL	4.37	74.10[a]	10-fold cross validation
Guo & Mu, 2013 [29]	BIF	3.98	–	different split
Bekhouche et al., 2017 [5].	MP-BSIF+MP-LPQ+SVM+SVR	3.50	75.13[a]	5-fold cross validation
Our	TSE (ML-BSIF+ML-LPQ+SVM+SVR)	**3.21**	**83.13**	5-fold cross validation
Our	TSE (ML-BSIF+ML-LPQ+SVM+SVR)	**3.17**	**83.72**	10-fold cross validation
Huerta et al., 2015 [38]	CNN	3.88	–	5-fold cross validation
Wang et al., 2015 [70]	CNN	4.77	–	5-fold cross validation
Niu et al., 2016 [58]	OR-CNN	3.27	73.42[a]	5-fold cross validation
Liu et al., 2017a[46]	GA-DFL (MP-CNN)	3.25	80.40[a]	10-fold cross validation
Duan et al., 2018a[15]	CNN+ELM	3.44	70.01	Private
Liao et al., 2018 [45]	DLF+FAM	3.48	71.03[a]	5-fold cross validation
Our	TSE (ML-BSIF+ML-LPQ+SVM+SVR)	**3.21**	**83.13**	5-fold cross validation
Our	TSE (ML-BSIF+ML-LPQ+SVM+SVR)	**3.17**	**83.72**	10-fold cross validation
Chen et al., 2017 [9]	Ranking-CNN	2.96	85.26[a]	5-fold cross validation
Liu et al., 2018 [48]	M-LSDML	2.89	87.03[a]	5-fold cross validation
Gao et al., 2017 [21]	DLDL+VGG-Face	2.42	88.88[a]	10-fold cross validation

The numbers in bold represent the performance of our method.
[a] CS result is obtained from the curves in the original paper.

Table 8
Comparison of the proposed approach with known and recent methods for age estimation on PAL database.

Publication	Approach	Performance		Protocol
		MAE (years)	CS (%)	
Choi et al., 2011 [11]	GHPF+SVR.	8.44	–	5-fold cross validation
Luu et al., 2011 [54]	CAM+SVR	6.00	–	LOPO
Nguyen et al., 2014 [57]	MLBP+GABOR+SVR	6.50	–	2-fold cross-validation
Günay & Nabiyev, 2016 [26]	AAM+GABOR+LBP	5.38	–	3-fold cross validation
Günay & Nabiyev, 2017 [24]	WLD+LBP+LPQ+SVR	5.75	54.28	3-fold cross validation
Choi et al., 2017 [10]	AAM+MLBP	5.50	–	Private
Agrawal et al., 2017 [1]	HOG+GABOR	6.55	–	2-fold cross-validation
Bekhouche et al., 2017 [5]	MP-BSIF+MP-LPQ+SVM+SVR	5.00	57.44[a]	5-fold cross validation
Our	TSE (ML-BSIF+ML-LPQ+SVM+SVR)	**4.49**	**74.57**	5-fold cross validation

[a] CS result is obtained from the curves in the original paper.

Table 9
Comparison of the proposed approach with method [35] for age estimation on subset of LFW database.

Publication	Approach	Performance		Protocol
		MAE (years)	CS (%)	
Han et al., 2015 [35]	DIF+SVM+SVR	**7.8**	43.35[a]	5-fold cross validation
Our	TSE (ML-BSIF+ML-LPQ+SVM+SVR)	7.9	**43.65**	5-fold cross validation

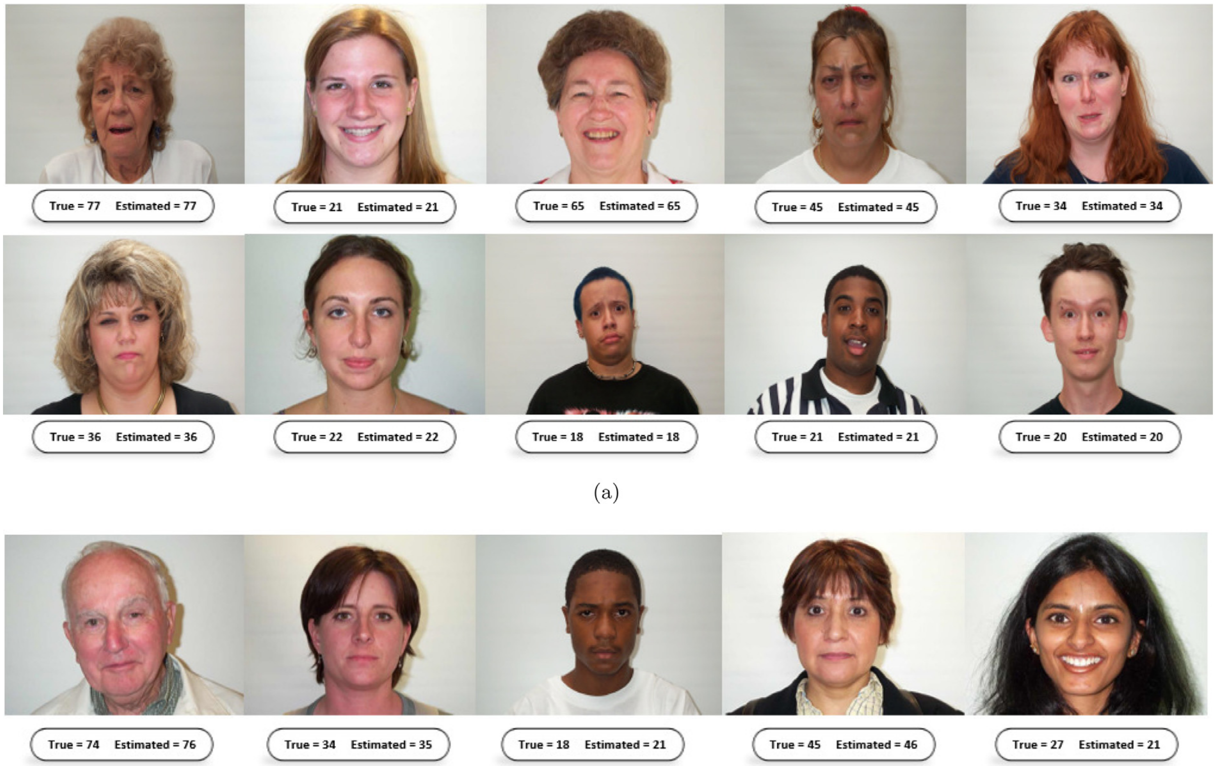
[a] CS result is obtained from the curves in the original paper.

In the literature, some works [62,64] reported a MAE below 3 years on MORPH-II database. Based on deep learning these approaches, reached 2.56 and 2.58 years in MAE, respectively. However, they did not consider the whole database in their evaluation procedure. Indeed, they used a subset of 5475 face images of persons, while the original database contains about 55134 face images. Furthermore, deep learning based methods require important resources in terms of computation and memory.

Table 9, shows the comparison of the proposed approach with [35] which is, to the best of our knowledge, the only work on age estimation including subset of LFW database as part of their experiments. One can see from the results that the TSE method provides competitive performance. Fig. 11 illustrates some examples of correct and incorrect age estimated by the proposed approach on PAL database.

4.4.4. Demographic attributes classification effect

In the proposed two-stages age estimator, the age is estimated after classifying the demographic attributes. The performance of age estimation can deteriorate when poor accuracies are obtained in demographic classification stage. To investigate the effect of misclassification in demographic attributes classification stage on age estimation stage, we have performed age estimation with our proposed TSE by considering the ground-truth of demographic attributes classes as the output of the first stage. In other words,



(b)

Fig. 11. Examples of good and poor age estimation of the proposed method on PAL database using BSIF+LPQ features. (a) Correct age estimates. (b) Incorrect age estimates.

we estimated the age with 100% accuracy of demographic classification stage.

Table 10 illustrates MAEs on MORPH-II, PAL and Subset of LFW databases, obtained by the proposed TSE method with ground-truth demographic attributes classes and with predicted demographic attributes classes. As can be seen, performing our proposed TSE with ground-truth of demographic attributes classes has improved the performance of age estimation (1.17, 2.5 and 2.74 years for MORPH II, PAL and subset of LFW databases respectively) when compared to the results obtained by using predicted demographic attributes classes (3.21, 4.49 and 7.93). The explanation is very intuitive. Estimating age within the correct (race, gender and age range) class provides smaller estimation errors and misclassified images in predicted demographic attributes classes (output of the first stage) can damage the performance of age estimation.

Table 10
MAE of TSE using the ground-truth and the predicted demographic attributes classes in the first stage (in years).

	Databases		
	MORPH-II	PAL	Subset of LFW
TSE with ground-truth classes	1.17	2.50	2.74
TSE with predicted classes	3.21	4.49	7.93

Table 11
MAE of misclassified images in demographic classification stage of the TSE method.

Database	Number of misclassified images	MAE (years)
MORPH II	18278	5.29
PAL	212	10.07
Subset of LFW	2927	9.07

To study the influence of misclassified images in predicted demographic attributes classes on MAE values, we identified them in Table 11, and calculated their MAE for MORPH-II, PAL and subset of LFW databases. We can clearly see, that the MAE of misclassified images in demographic classification stage are high, which adversely affects the performance of age estimation.

5. Conclusion

In this paper, we proposed a novel age estimation approach based on facial demographic attributes. Face detection and alignment to extract only the regions of interest and correct the position and the size of the input face image are first performed. The features are then extracted by Multi-level face representation, using the LPQ and BSIF texture descriptors. Finally, a Two-stages estimator (TSE) for age prediction is proposed. When performing age estimation experiments over three databases using direct and hierarchical (with two different orders) age estimation methods, we observed that TSE approach outperforms significantly the two other methods. A conclusion can be drawn that age estimation can have much small error if age estimation is carried out within the same race, gender and age range. The comparison with known and recent state-of-art age estimation methods, including deep learning based methods, showed that the proposed TSE approach provides the lower MAE on PAL and MORPH II databases and competitive performance on subset of LFW database. The simplicity and efficiency of the proposed approach can provide a helpful guide to facial age estimation on mobile devices with less memory and computation resources.

In future works, a way to further improve the performance of the proposed method is to remove the irrelevant and/or redundant demographic features by using more efficient feature selection methods. This will help also to reduce the training and testing computation times. In addition, since the race and gender are the main challenges associated with apparent age estimation task, we are interested in evaluating the influence of demographic attributes on apparent age estimation. We envision also the possibility to map the proposed hierarchical strategy on CNN based architecture.

References

[1] S.C. Agrawal, R.K. Tripathi, A.S. Jalal, Automatic human age estimation system using fusion of local and global features, in: Multimedia, Signal Processing and Communication Technologies (IMPACT), 2017 International Conference on, IEEE, 2017, pp. 5–9.

[2] T. Ahonen, A. Hadid, M. Pietikainen, Face description with local binary patterns: application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. (2006) 2037–2041.

[3] Amazon Mechanical Turk, The Amazon Mechanical Turk (mturk) Crowdsourcing Service, 2018. <https://www.mturk.com/> (Online; accessed 19-November-2018).

[4] S.E. Bekhouche, A. Ouafi, A. Benlamoudi, A. Taleb-Ahmed, A. Hadid, Facial age estimation and gender classification using multi level local phase quantization, in: 2015 3rd International Conference on Control, Engineering & Information Technology (CEIT), IEEE, 2015, pp. 1–4.

[5] S.E. Bekhouche, A. Ouafi, F. Dornaika, A. Taleb-Ahmed, A. Hadid, Pyramid multi-level features for facial demographic estimation, Expert Syst. Appl. 80 (2017) 297–310.

[6] P. Carcagnì, M. Del Coco, D. Cazzato, M. Leo, C. Distante, A study on different experimental configurations for age, race, and gender estimation problems, EURASIP J. Image Video Process. 2015 (2015) 37.

[7] K.-Y. Chang, C.-S. Chen, Y.-P. Hung, Ordinal hyperplanes ranker with cost sensitivities for age estimation, in: Computer Vision and Pattern Recognition (cvpr), 2011 IEEE Conference on, IEEE, 2011, pp. 585–592.

[8] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, W. Gao, Wld: a robust local image descriptor, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2010) 1705–1720.

[9] S. Chen, C. Zhang, M. Dong, J. Le, M. Rao, Using ranking-cnn for age estimation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5183–5192.

[10] S.E. Choi, J. Jo, S. Lee, H. Choi, I.-J. Kim, J. Kim, Age face simulation using aging functions on global and local features with residual images, Expert Syst. Appl. 80 (2017) 107–125.

[11] S.E. Choi, Y.J. Lee, S.J. Lee, K.R. Park, J. Kim, Age estimation using a hierarchical classifier based on global and local facial features, Pattern Recogn. 44 (2011) 1262–1281.

[12] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1, IEEE, 2005, pp. 886–893.

[13] A. Dantcheva, P. Elia, A. Ross, What else does your biometric data reveal? a survey on soft biometrics, IEEE Trans. Inform. Forensics Secur. 11 (2016) 441–467.

[14] H. Dibeklioğlu, F. Alnajar, A.A. Salah, T. Gevers, Combining facial dynamics with appearance for age estimation, IEEE Trans. Image Process. 24 (2015) 1928–1943.

[15] M. Duan, K. Li, C. Yang, K. Li, A hybrid deep learning cnn-elm for age and gender classification, Neurocomputing 275 (2018) 448–461.

[16] Y. Duan, J. Lu, J. Feng, J. Zhou, Context-aware local binary feature learning for face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 40 (2018) 1139–1153.

[17] Y. Duan, J. Lu, Z. Wang, J. Feng, J. Zhou, Learning deep binary descriptor with multi-quantization, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1183–1192.

[18] E. Eidinger, R. Enbar, T. Hassner, Age and gender estimation of unfiltered faces, IEEE Trans. Inform. Forensics Secur. 9 (2014) 2170–2179.

[19] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, C.-J. Lin, LIBLINEAR: a library for large linear classification, J. Mach. Learn. Res. 9 (2008) 1871–1874.

[20] C. Fernández, I. Huerta, A. Prati, A comparative evaluation of regression learning algorithms for facial age estimation, in: Face and Facial Expression Recognition from Real World Videos, Springer, 2015, pp. 133–144.

[21] B.-B. Gao, C. Xing, C.-W. Xie, J. Wu, X. Geng, Deep label distribution learning with label ambiguity, IEEE Trans. Image Process. 26 (2017) 2825–2838.

[22] X. Geng, Z.-H. Zhou, K. Smith-Miles, Automatic age estimation based on facial aging patterns, IEEE Trans. Pattern Anal. Mach. Intell. 29 (2007) 2234–2240.

[23] X. Geng, Z.-H. Zhou, Y. Zhang, G. Li, H. Dai, Learning from facial aging patterns for automatic age estimation, in: Proceedings of the 14th ACM International Conference on Multimedia, ACM, 2006, pp. 307–316.

[24] A. Günay, V. Nabiyev, A new facial age estimation method using centrally overlapped block based local texture features, Multimedia Tools Appl. (2017) 1–27.

[25] A. Gunay, V.V. Nabiyev, Automatic age classification with lbp, in: Computer and Information Sciences, 2008. ISCIS'08. 23rd International Symposium on, IEEE, 2008, pp. 1–4.

[26] A. Günay, V.V. Nabiyev, Age estimation based on hybrid features of facial images, in: O.H. Abdelrahman, E. Gelenbe, G. Gorbil, R. Lent (Eds.), Information Sciences and Systems 2015, Springer International Publishing, Cham, 2016, pp. 295–304.

[27] G. Guo, G. Mu, Human age estimation: what is the influence across race and gender?, in: Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on, IEEE, 2010, pp 71–78.

[28] G. Guo, G. Mu, Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression, in: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE, 2011, pp. 657–664.

[29] G. Guo, G. Mu, Joint estimation of age, gender and ethnicity: CCA vs. PLS, in: Automatic Face and Gesture Recognition (fg), 2013 10th IEEE International Conference and Workshops on, IEEE, 2013, pp. 1–6.

[30] G. Guo, G. Mu, A framework for joint estimation of age, gender and ethnicity on a large database, Image Vis. Comput. 32 (2014) 761–770.

[31] G. Guo, G. Mu, Y. Fu, T.S. Huang, Human age estimation using bio-inspired features, in: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE, 2009, pp. 112–119.

[32] P. Gysel, M. Motamedi, S. Ghiasi, Hardware-oriented approximation of convolutional neural networks, 2016. arXiv preprint arXiv: 1604.03168.

[33] H. Han, A.K. Jain, Age, gender and race estimation from unconstrained face images. MSU Technical Report, MSU-CSE, 2014, pp. 14-5.

[34] A. Hadid, M. Pietikäinen, Demographic classification from face videos using manifold learning, Neurocomputing 100 (2013) 197–205.

[35] H. Han, C. Otto, X. Liu, A.K. Jain, Demographic estimation from face images: Human vs. machine performance, IEEE Trans. Pattern Anal. Mach. Intell. 37 (2015) 1148–1161.

[36] D.R. Hardoon, S. Szedmak, J. Shawe-Taylor, Canonical correlation analysis: an overview with application to learning methods, Neural Comput. 16 (2004) 2639–2664.

[37] Z. Hu, Y. Wen, J. Wang, M. Wang, R. Hong, S. Yan, Facial age estimation with age difference, IEEE Trans. Image Process. 26 (2017) 3087–3097.

[38] I. Huerta, C. Fernández, C. Segura, J. Hernando, A. Prati, A deep analysis on age estimation, Pattern Recogn. Lett. 68 (2015) 239–249.

[39] J.S. Kang, C.S. Kim, Y.W. Lee, S.W. Cho, K.R. Park, Age estimation robust to optical and motion blurring by deep residual cnn, Symmetry 10 (2018) 108.

[40] J. Kannala, E. Rahtu, Bsif: Binarized statistical image features, in: Pattern Recognition (ICPR), 2012 21st International Conference on, IEEE, 2012, pp. 1363–1366.

[41] V. Kazemi, S. Josephine, One millisecond face alignment with an ensemble of regression trees, in: 27th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, Columbus, United States, 23 June 2014 through 28 June 2014, IEEE Computer Society, 2014, pp. 1867–1874.

[42] Y.H. Kwon, N. da Vitoria Lobo, Age classification from facial images, Comput. Vision Image Understand. 74 (1999) 1–21.

[43] A. Lanitis, C.J. Taylor, T.F. Cootes, Toward automatic simulation of aging effects on face images, IEEE Trans. Pattern Anal. Mach. Intell. 24 (2002) 442–455.

[44] Z. Lei, M. Pietikäinen, S.Z. Li, Learning discriminant face descriptor, IEEE Trans. Pattern Anal. Mach. Intell. 36 (2014) 289–302.

[45] H. Liao, Y. Yan, W. Dai, P. Fan, Age estimation of face images based on cnn and divide and rule strategy, Math. Problems Eng. 2018 (2018) 1–8, https://doi.org/10.1155/2018/1712686.

[46] H. Liu, J. Lu, J. Feng, J. Zhou, Group-aware deep feature learning for facial age estimation, Pattern Recogn. (2017) 66.

[47] H. Liu, J. Lu, J. Feng, J. Zhou, Learning deep sharable and structural detectors for face alignment, IEEE Trans. Image Process. 26 (2017) 1666–1678.

[48] H. Liu, J. Lu, J. Feng, J. Zhou, Label-sensitive deep metric learning for facial age estimation, IEEE Trans. Inform. Forensics Secur. 13 (2018) 292–305.

[49] K.-H. Liu, S. Yan, C.-C.J. Kuo, Age group classification via structured fusion of uncertainty-driven shape features and selected surface features, in: Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on, IEEE, 2014, pp. 445–452.

[50] J. Lu, J. Hu, Y.-P. Tan, Discriminative deep metric learning for face and kinship verification, IEEE Trans. Image Process. (2017) 26.

[51] J. Lu, V.E. Liong, J. Zhou, Cost-sensitive local binary feature learning for facial age estimation, IEEE Trans. Image Process. 24 (2015) 5356–5368.

[52] J. Lu, V.E. Liong, X. Zhou, J. Zhou, Learning compact binary face descriptor for face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 37 (2015) 2041–2056.

[53] K. Luu, K. Ricanek, T.D. Bui, C.Y. Suen, Age estimation using active appearance models and support vector machine regression, in: Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on, IEEE, 2009, pp. 1–5.

[54] K. Luu, K. Seshadri, M. Savvides, T.D. Bui, C.Y. Suen, Contourlet appearance model for facial age estimation, in: Biometrics (IJCB), 2011 International Joint Conference on, IEEE, 2011, pp. 1–8.

[55] M. Minear, D.C. Park, A lifespan database of adult facial stimuli, Behav. Res. Methods Instrum. Comput. 36 (2004) 630–633.

[56] A. Mokadem, M. Charbit, G. Chollet, K. Bailly, Age regression based on local image features, in: Image and Video Technology (PSIVT), 2010 Fourth Pacific-Rim Symposium on, IEEE, 2010, pp. 88–93.

[57] D.T. Nguyen, S.R. Cho, K.Y. Shin, J.W. Bang, K.R. Park, Comparative study of human age estimation with or without preclassification of gender and facial expression, Sci. World J. (2014) 2014.

[58] Z. Niu, M. Zhou, L. Wang, X. Gao, G. Hua, Ordinal regression with multiple output cnn for age estimation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 4920–4928.

[59] V. Ojansivu, J. Heikkilä, Blur insensitive texture classification using local phase quantization, in: International Conference on Image and Signal Processing, Springer, 2008, pp. 236–243.

[60] J.K. Pontes, A.S. Britto Jr, C. Fookes, A.L. Koerich, A flexible hierarchical approach for facial age estimation based on multiple features, Pattern Recogn. 54 (2016) 34–51.
[61] K. Ricanek, T. Tesafaye, Morph: a longitudinal image database of normal adult age-progression, in: Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on, IEEE, 2006, pp. 341–345.
[62] P. Rodríguez, G. Cucurull, J.M. Gonfaus, F.X. Roca, J. Gonzalez, Age and gender recognition in the wild with deep attention, Pattern Recogn. 72 (2017) 563–571.
[63] R. Rosipal, N. Krämer, Overview and recent advances in partial least squares, Lect. Notes Comput. Sci. 3940 (2006) 34.
[64] R. Rothe, R. Timofte, L. Van Gool, Deep expectation of real and apparent age from a single image without facial landmarks, Int. J. Comput. Vision 126 (2018) 144–157, https://doi.org/10.1007/s11263-016-0940-3.
[65] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., Imagenet large scale visual recognition challenge, Int. J. Comput. Vision 115 (2015) 211–252.
[66] C. Shan, Learning local features for age estimation on real-life faces, in: Proceedings of the 1st ACM International Workshop on Multimodal Pervasive Video Analysis MPVA '10, ACM, New York, NY, USA, 2010, pp. 23–28, https://doi.org/10.1145/1878039.1878045. <http://doi.acm.org/10.1145/1878039.1878045> .
[67] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, 2014. arXiv preprint arXiv: 1409.1556.
[68] J. Suo, T. Wu, S. Zhu, S. Shan, X. Chen, W. Gao, Design sparse features for age estimation using hierarchical face model, in: Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on, IEEE, 2008, pp. 1–6.
[69] P. Thukral, K. Mitra, R. Chellappa, A hierarchical approach for human age estimation, in: Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, IEEE, 2012, pp. 1529–1532.
[70] X. Wang, R. Guo, C. Kambhamettu, Deeply-learned feature for age estimation, in: Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on, IEEE, 2015, pp. 534–541.
[71] R. Weng, J. Lu, Y.-P. Tan, J. Zhou, Learning cascaded deep auto-encoder networks for face alignment, IEEE Trans. Multimedia 18 (2016) 2066–2078.
[72] Z. Yang, H. Ai, Demographic classification with local binary patterns, Adv. Biometrics (2007) 464–473.
[73] Y. Zhu, Y. Li, G. Mu, G. Guo, A study on apparent age estimation, in: Computer Vision Workshop (ICCVW), 2015 IEEE International Conference on, IEEE, 2015, pp. 267–273.