


**PAPER****ENGINEERING SCIENCES**

Muhammad Sajid,¹ Ph.D.; Imtiaz Ahmad Taj,¹ Ph.D.; Usama Ijaz Bajwa ,² Ph.D.; and Naeem Iqbal Ratyal,¹ Ph.D.

Facial Asymmetry-Based Age Group Estimation: Role in Recognizing Age-Separated Face Images

ABSTRACT: Face recognition aims to establish the identity of a person based on facial characteristics. On the other hand, age group estimation is the automatic calculation of an individual's age range based on facial features. Recognizing age-separated face images is still a challenging research problem due to complex aging processes involving different types of facial tissues, skin, fat, muscles, and bones. Certain holistic and local facial features are used to recognize age-separated face images. However, most of the existing methods recognize face images without incorporating the knowledge learned from age group estimation. In this paper, we propose an age-assisted face recognition approach to handle aging variations. Inspired by the observation that facial asymmetry is an age-dependent intrinsic facial feature, we first use asymmetric facial dimensions to estimate the age group of a given face image. Deeply learned asymmetric facial features are then extracted for face recognition using a deep convolutional neural network (dCNN). Finally, we integrate the knowledge learned from the age group estimation into the face recognition algorithm using the same dCNN. This integration results in a significant improvement in the overall performance compared to using the face recognition algorithm alone. The experimental results on two large facial aging datasets, the MORPH and FERET sets, show that the proposed age group estimation based on the face recognition approach yields superior performance compared to some existing state-of-the-art methods.

KEYWORDS: forensic science, facial asymmetry, temporal variations, age group estimation, age-separated face images, support vector machine, identification, verification

Accurate age group estimation has many applications in fields such as homeland security, forensic science, and passport services; it can also be used for locating missing persons, determining the age of asylum seekers with missing legal documents, controlling pedophilia, conducting statistical analysis (e.g., class-wise age distribution), limiting access to the purchase of certain commodities (e.g., alcohol and tobacco), and controlling some human-computer interactions (HCI) (e.g., limiting Internet access to certain age groups). As an active research area, both human perception-based and machine-based age group estimation algorithms have been reported in the literature. Despite significant advances and ample work in related research areas, see, for example, Fu and Huang (1), Yang and Ai (2), Ylioinas et al. (3), Lu and Tan (4), Hu and Jain (5), Han et al. (6), and Guo and Mu (7), existing methods have certain limitations, which include the following: (i) the use of handcrafted features that have poor discriminative power and that are vulnerable to certain extrinsic variations, such as facial makeup (8); and (ii) the lack of related studies that integrate the knowledge learned from age group estimation into face recognition algorithms. In this work, we have used deeply learned asymmetric features. A face

recognition algorithm is developed that seeks the knowledge learned from age group estimation to recognize face images. The current research focuses on age group estimation instead of the exact age, for the purpose of using certain characteristics of a given age group to recognize age-separated face images.

Related Work

This study consists of two components: (i) age group estimation and (ii) recognition of age-separated face images. Hence, the related work of the two components is discussed separately.

Age Group Estimation

The motivation for age estimation systems has grown over the past few decades, given the rise of the digital age and the increase in human-computer interaction. Recently, the Face Recognition Vendor Test (FRVT) evaluated the performance of facial age estimation as a new area of study with Still Images Track (9). The main objectives of this evaluation include the assessment of current age estimation technology and the investigation of the estimation accuracy on large-scale datasets across demographic variations.

A literature survey related to age estimation can be broadly classified into two categories: (i) human perception-based age estimation and (ii) machine-based age estimation. With regard to human perception-based age estimation, a detailed survey presented by Rhodes (10) shows that humans can estimate the age

¹Vision and Pattern Recognition Systems Research Group, Capital University of Science and Technology, Expressway, Zone V, Islamabad, Pakistan.

²Department of Computer Science, COMSATS Institute of Information Technology, Off Raiwind Road, Lahore, Pakistan.

Received 12 June 2017; and in revised form 31 Oct. 2017; accepted 8 Mar. 2018.

of previously unseen face images quite accurately. The age estimation accuracy evaluated by two groups of human participants is discussed in Burt and Perrett (11), with a reported deviation of 2.39 years. The effect of feature manipulation and the influence of certain local regions in age estimation are presented in George and Hole (12) and Jones and Smith (13).

With regard to facial age estimation by machines, a number of age estimation methods have been reviewed in Fu et al. (14). In a pioneer work presented by Kwon and Lobo (15), facial age group classification is performed using facial features. In (1), aging manifolds have been used toward accurate age estimation. Local facial features have been used in (2) for binary age group classification. Variants of local binary patterns (LBP) with support vector machines (SVM) have been employed in (3) toward accurate age group classification. Manifolds of raw intensity are used in (4) for age and head pose estimation. In (5), biologically inspired features (BIF) have been used for age, gender, and ethnical group classification. A recent study (6) gives a comprehensive account of human versus machine performance for the joint estimation of age, gender, and ethnical group. In most of the related studies, it is observed that the age estimation accuracy decreases with the age progression due to many factors, such as the gender, ethnical group, stress levels, eating, and sleeping habits, which altogether result in an individual's age progression function.

Recognition of Age-separated Face Images

The second aspect of this study aims at recognizing face images across aging variations. Recently, a number of related studies have been reported in the literature, such as Ling et al. (16), Park et al. (17), Li et al. (18), Yadav et al. (19), and Sungatullina et al. (20). Overall, the existing methods can be classified as generative or discriminative. The generative methods focus on facial age synthesis. A representative study in this category is reported in (17). Despite their strong discriminative power, such methods depend on complex age synthesis. Discriminative methods, on the other hand, use multiple face features to achieve robust recognition performance. Some of the representative works in this category include (18–20). In (18), facial overlapping patches are used to extract local features, and subsequently, discriminant analysis is performed for recognition. In (19), a bacteria foraging fusion (BFF) scheme achieves superior performance by optimizing the region-specific weights. A local feature-based multiview discriminative learning (MDL) approach is presented in (20). More recently, the effects of facial aging on recognition performance across a large population have been studied in (21). Despite the state-of-the-art performance, these approaches exhibit two main limitations: (i) lack of facial feature evaluations with temporal variations and (ii) a lack of discriminative information because most of these methods utilize a single set of facial features. For example, appearance-based facial features show poor performance in recognizing face images across temporal variations, owing to their sheer dependence on raw pixel intensities (22). Local features, although used in many age invariant face recognition methods, are vulnerable to external face variations, for example, facial makeup (8). Table 1 summarizes some recent studies that cover age group estimation and face recognition. The results reported in the previous related studies suggest that there is a vast scope to improve the accuracy of age group estimation and age-separated face recognition algorithms. The aim of the current study is to develop an age-assisted face recognition algorithm based on facial asymmetry.

Some of the existing methods report facial asymmetry to recognize face images, such as Liu et al. (23) and Gutta and Wechsler (24). In (23), facial asymmetry was used for expression-invariant human identification. In (24), left- or right-sided asymmetry was used to recognize face images, especially in situations in which a full face is not available. To understand the possibility of matching the left and right ears, an analysis of the symmetry of human ears is presented in (25). Keeping in view the existing methods, the role of facial asymmetry in age group estimation and age-assisted face recognition is yet to be explored. The proposed method differs significantly from previous approaches such that existing approaches are discriminative or generative in nature, and our approach belongs to the discriminative category with integration of the knowledge learned from age group estimation into a face recognition algorithm.

The main contributions of this research are as follows:

- Facial asymmetry-based age group estimation.
- Study of the role of different asymmetric facial regions in age group estimation and age-assisted face recognition.
- Design of a face recognition algorithm inspired from age group estimation.

The Facial Aging Databases

The current study focused on two publicly available face databases, MORPH II (26) and FERET (27), as described below.

- MORPH II is a large longitudinal face database. It contains 55,134 face images of over 13,000 subjects with aging variations. MORPH II has been extensively used in research related to age estimation and face recognition, such as (18,20).
- FERET is another large face database, which contains 3540 frontal face images of 1196 subjects. The database contains two subsets, dup I and dup II, which contain facial images that have small and large temporal variations in addition to a gallery set called the fa set.

Unlike some other large face databases (e.g., the LFW database [28]), both MORPH II and FERET provide necessary demographic information on the subjects, including the age, gender, and ethnical group, which is useful for conducting research related to age group estimation and face recognition across a time lapse. Example face images from the MORPH II and FERET datasets are shown in Fig. 1, while key statistics are given in Table 2.

Etiology, Measurement, and Assessment of Facial Asymmetry

Etiology of Facial Asymmetry

Facial asymmetry refers to noncorrespondence in shape, size, and arrangement of bilateral facial landmarks. Facial asymmetry is an intrinsic facial characteristic, even in young healthy subjects (Ercan et al. [29]). Facial asymmetry can be classified as follows: (i) acquired, (ii) congenital, and (iii) developmental (Cheong and Lo [30]). Acquired facial asymmetry results from pathological or traumatic reasons. The congenital type is seen at birth, and its origin is in genetics, while the developmental type is commonly seen in the general population and appears with the progression of age. The developmental facial asymmetry finds its origin in one-sided habitual chewing and one-sided sleep in addition to other reasons. In this study, we present an assessment of facial asymmetry across aging variations and its

TABLE 1—Summary of some recent approaches on (a) age estimation and (b) recognition for age-separated face images.

(a)					
Representative Work (year)	Features Employed	Face Database	Performance Measure	Age Estimation	Face Recognition
Ylioinas et al. (2012) (3)	Variants of local binary patterns (LBP)	Images of groups	Rank 1 classification accuracy	Studied	Not studied
Lu and Tan (2013) (4)	Manifolds of raw intensity	MORPH II	Mean absolute error (MAE)	Studied	Not studied
Hu and Jain (2014) (5)	Biologically inspired features (BIF)	Images of groups, LFW+	Age group classification accuracy	Studied	Not studied
Han et al. (2015) (6)	Biologically inspired features (BIF)	FG-NET, MORPH II, PCSO, LFW	Mean absolute error (MAE)	Studied	Not studied
(b)					
Representative Work (year)	Features Employed	Face Database	Performance Measure and Accuracy	Age Estimation	Face Recognition
Park et al. (2010) (17)	Aging patterns based on PCA coefficients	FG-NET, MORPH II	Rank 1 recognition accuracy	Not studied	Studied
Li et al. (2011) (18)	Scale invariant feature transform (SIFT) and multiscale local binary patterns (MLBP)	FG-NET, MORPH II	Rank 1 recognition accuracy	Not studied	Studied
Yadav et al. (2013) (19)	Local binary patterns (LBP)	FG-NET, IIIT Delhi	Rank 1 recognition accuracy	Not studied	Studied
Sungatullina et al. (2013) (20)	Local binary patterns, scale invariant feature transform, and gradient orientation pyramid (GOP)	FG-NET, MORPH II	Rank 1 recognition accuracy	Not Studied	Studied
Proposed work	Deeply learned aggregated asymmetric facial features (da-AFF)	MORPH II, FERET	Age group classification accuracy/Rank 1 recognition accuracy	Studied	Studied



FIG. 1—Sample image pairs showing intrasubject temporal variations from MORPH II (top row) and FERET (bottom row) datasets.

role in age group estimation and age-assisted face recognition. For this purpose, the face images are first preprocessed, and facial landmarks are detected as described below.

Preprocessing and Facial Landmark Detection

To mitigate the effects of extrinsic variations, the face images were preprocessed as follows:

- All of the face images were rotated and aligned based on eye coordinates such that they are vertically upright.
- All of the face images were then cropped into a 128×128 pixel size with an equal interpupillary distance (IPD).
- Color face images were converted to grayscale images to alleviate unwanted color effects. For this purpose, a luminance model adopted by NTSC and JPEG was used, as shown in Eq. 1.

$$Y = 0.299R + 0.587G + 0.114B, \quad (1)$$

where R , G , and B represent the red, green, and blue color channels and Y is the resulting grayscale image.

- To remove unwanted illumination variations from the resulting grayscale images, histogram equalization was used.

TABLE 2—Statistics of publicly available MORPH II and FERET databases.

	MORPH II	FERET		
		Fa set	Dup I set	Dup II set
# Subjects	13,000+	1196	243	75
# Images	55,134	1196	722	234
# Average images/subject	4	1	2.97	3.12
Temporal variations	Minimum: 1 day Maximum: 1681 days	N/A	Images were acquired later in a time with maximum temporal variation of 1031 days as compared to corresponding gallery images	Images were acquired at least 1.5 years later than corresponding gallery images
Age range (years)	16–77	10–60+		

After the preprocessing stage, 64 facial landmarks were detected (Fig. 2) using the following two methods:

(i) Manual annotation as per the definitions of the corresponding landmarks given in Table 3 and (ii) a state-of-the-art face recognition and analysis tool, Face++ (31). To check the validity of the chosen methods, the average error was calculated, which resulted in 1 pixel per image for all of the selected landmarks. The error results show a negligible error between the landmark points that were extracted automatically and their actual physical correspondence. Among the detected facial landmarks, 36 landmarks, previously used by Ercan et al. (29), were selected (shown by the numbers in Fig. 2 and illustrated in Table 3) for the measurement and assessment of facial asymmetry across temporal variations, as described below.

Measurement and Assessment of Facial Asymmetry across Temporal Variations

This section presents an account of the facial asymmetry measurement and assessment across temporal variations, as described below.

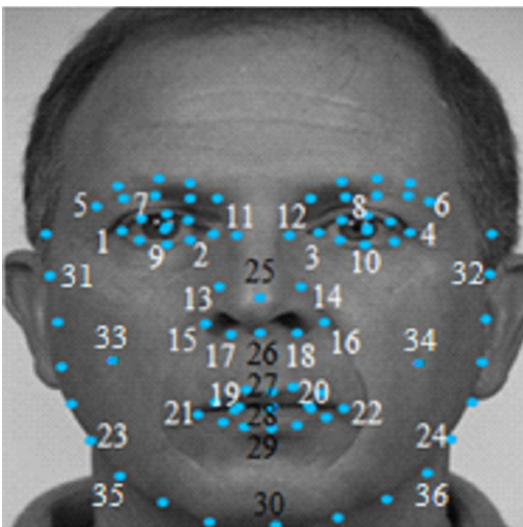


FIG. 2—Automatic facial landmark detection using Face++ (31) (face image taken from FERET dataset).

Measurement of Facial Asymmetry—To measure facial asymmetry, the horizontal facial asymmetry (HFA), vertical facial asymmetry (VFA), and angular facial asymmetry (AFA) were used. For HFA, a vertical axis ($P3P4$) was determined by fitting a least square regression line (termed as the facial midline in the remainder of the text) through six facial landmarks (#25–30), while for VFA, a horizontal axis ($P1P2$) was determined by joining the eye coordinates of a subject looking straight, as shown in Fig. 3a–c. As the distance between the eye coordinates remains the same for the subjects not looking straight, the same procedure was used to determine the horizontal axis for such subjects, as shown in Fig. 3d–f.

Finally, fifteen horizontal and eleven vertical facial asymmetric distances of selected pairs of bilateral facial landmarks were measured, referenced as $P3P4$ and $P1P2$, using Eqs (2) and (3), respectively.

$$\text{HFA} = d_{lh} - d_{rh}, \quad (2)$$

$$\text{VFA} = d_{lv} - d_{rv}, \quad (3)$$

where d_{lh} , d_{rh} , d_{lv} , and d_{rv} represent the left- and right-sided horizontal and vertical interpixel distances of the selected bilateral facial landmarks, respectively.

In addition to the horizontal and vertical linear asymmetric measurements, eight angular facial asymmetric distances were measured to define the angular facial asymmetry (AFA), as shown in Eq. 4.

$$\text{AFA} = g_l - g_r, \quad (4)$$

where g_l and g_r represent the left- and right-sided angular measurements in degrees.

The selected horizontal, vertical, and angular asymmetric distances (called the horizontal, vertical, and angular dimensions in the remainder of the text) are given in Tables 4–6, respectively.

The selected dimensions were used to (i) evaluate how facial asymmetry varies across time, (ii) assess the correlation between

TABLE 3—Illustration of facial landmarks (29).

Type of Facial Landmarks	Landmark #	Location	Name of Facial Landmark	Abbreviation
Pairs of bilateral facial landmarks	1 and 4	Eyes	Exocanthion	Exo
	2 and 3	Eyes	Endocanthion	End
	5 and 6	Eyebrows	Frontotemporale	Fronto
	7 and 8	Eyes	Palpebrale superius	Palp-Sup
	9 and 10	Eyes	Palpebrale inferius	Palp-Inf
	11 and 12	Nose bridge	Maxillofrontale	Maxil
	13 and 14	Nose	Lateral cartilage	Lat-Cart
	15 and 16	Nose	Alare	Ala
	17 and 18	Nose	Subalare	Sub-Ala
	19 and 20	Lips	Crista philter	Crist
Single facial landmarks	21 and 22	Lips	Cheilion	Cheil
	23 and 24	Face	Gonion	Goni
	31 and 32	Face	Zygion	Zyg
	33 and 34	Face	Cheek	Chk
	35 and 36	Face	Inferior border	Inf-Bord
	25	Nose	Pronasale	Prona
	26	Nose	Subnasale	Sub-Nas
	27	Lips	Labiale superius	Lab-Sup
	28	Lips	Stomion	Stomio
	29	Lips	Labiale inferius	Lab-Inf
30	Chin	Gnathion	Ganth	

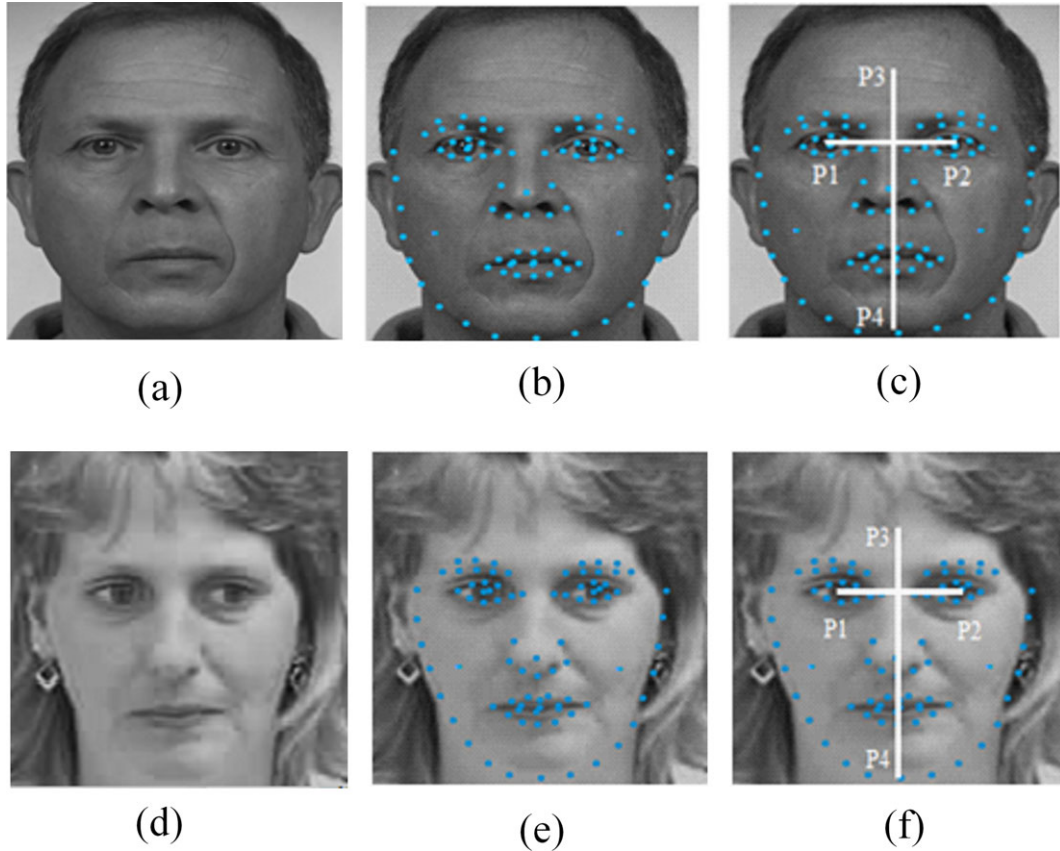


FIG. 3—Top row: (a) original face image of a subject looking straight, (b) detected facial landmarks, (c) facial midline (P3P4) and line joining eye coordinates (P1P2). Bottom row: (d) original face image of a subject not looking straight, (e) detected landmarks, (f) facial midline (P3P4) and line joining eye coordinates P1P2.

the horizontal, vertical, and angular dimensions, and (iii) develop a hierarchical approach to perform age group estimation, as described below.

Facial Asymmetry Assessment across Temporal Variations—

To assess the facial asymmetry, we select subsets of face images from the MORPH II and FERET databases, such that the selected images represent small and large temporal variations, as described below.

- In the case of the MORPH II database, a subset that consists of 30,000 images of 10,000 subjects (three images per subject) was selected representing the youngest, older, and oldest face images of each selected subject. Here, the older and oldest images represent small and large temporal variations, respectively.
- In the case of the FERET database, a subset that consists of 75 images from fa set, 75 images from dup I set, and 75 images from dup II set was selected such that the images of the subjects present in fa set were also present in the dup I and dup II sets.

To assess the facial asymmetry across temporal variations, the mean and standard deviation (SD) of 15 horizontal, 11 vertical, and 8 angular dimensions were computed and normalized in the range of (0–1) for the face images in the selected subsets. The results were reported in terms of bar graphs, as shown in Fig. 4, with the following two important conclusions:

- Horizontal, vertical, and angular facial asymmetry increases from the upper to lower parts of the face. For example, a facial region that contains the cheilion and gonion landmarks

is more asymmetric compared to a facial region that contains the frontotemporale and exocanthion landmarks. More pronounced facial asymmetry for the lower parts of the face can be attributed to the response of functional adaptation to asymmetrical masticatory activity in these parts, as suggested by Vig and Hewitt in (32). The control of the facial musculature is complex, with different patterns of neural innervations present for the upper versus the lower face, depending on the nature of the neurological control of the bilateral facial parts by the two cerebral hemispheres (32).

- Horizontal and vertical facial asymmetry increases with age progression. A small increase in facial asymmetry is observed when the mean and SD of the image sets with small aging variations are compared with those of the youngest face image set from MORPH II and from fa set in the FERET dataset, respectively. On the other hand, such variations are more pronounced when the mean and SD of the image sets with large aging variations are compared with the youngest face image set from MORPH II and the dup II set from the FERET database, respectively. Hence, it is concluded that the facial asymmetry is a strong indicator of the facial age, which increases with age progression.

To analyze the effects of rotation and resizing of the face images on facial asymmetry, we measured and evaluated the facial asymmetry on raw face images (i.e., without preprocessing). The mean and SD of the horizontal, vertical, and angular dimensions were also normalized in the range of (0–1), as shown in Fig. 5. The comparison of the mean and SD of the

TABLE 4—Horizontal dimensions.

Pairs of Bilateral Facial Landmarks	1,4	2,3	5,6	7,8	9,10	11,12	13,14	15,16	17,18	19,20	21,22	23, 24	31,32	33,34	35,36
Horizontal dimensions	HA1	HA2	HA3	HA4	HA5	HA6	HA7	HA8	HA9	HA10	HA11	HA12	HA13	HA14	HA15

TABLE 5—Vertical dimensions.

Pairs of Bilateral Facial Landmarks	9,10	11,12	13,14	15,16	17,18	19,20	21,22	23,24	31,32	33,34	35,36
Vertical dimensions	VA1	VA2	VA3	VA4	VA5	VA6	VA7	VA8	VA9	VA10	VA11

TABLE 6—Angular dimensions.

Triads of Bilateral Facial Landmarks	(19,21,30) and (20,22,30)	(1,7,9) and (4,8,10)	(13,15,17) and (14,16,18)	(1,7,11) and (4,8,12)	(11,13,17) and (12,14,18)	(21,25,30) and (22,25,30)	(23,27,30) and (24,27,30)	(25,27,30) and (25,28,30)
Angular dimensions	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8

preprocessed and raw face images reveals that facial asymmetry is not affected by the alignment process. This alignment invariance can be attributed to the structured nature of anthropometric-based asymmetric facial features compared to the unstructured microlevel features, as suggested in (22).

We also analyzed facial asymmetry variations for an individual subject across a time lapse. Figure 6 shows the variations in the horizontal, vertical, and angular dimensions of an individual subject at three different ages, 38, 40, and 42 years of age. The dimension variations show that facial asymmetry is an age-dependent facial characteristic that increases with age.

Once facial asymmetry is assessed across temporal variations, the next goal is to evaluate any correlation between the horizontal, vertical, and angular dimensions.

To analyze the mutual independence of the horizontal, vertical, and angular dimensions, we performed the Pearson's correlation test between 15 horizontal, 11 vertical, and 8 angular dimensions. It was observed that 82.35% of the horizontal, vertical, and angular dimensions represent weak correlations (<0.5) and, hence, are mutually independent. Figure 7 shows correlation matrices in the form of correlograms that highlight the degree of correlation between the horizontal, vertical, and angular dimensions. Motivated by this fact, the horizontal, vertical, and angular dimensions were used as independent measures of facial asymmetry in subsequent analysis for the age group estimation task, as described below.

Age Group Estimation

This section presents a facial asymmetry-based hierarchical approach for age group estimation. For this purpose, face images were normalized against different types of appearance variations, as explained in Section 2. To perform facial asymmetry-based age group estimation, it was desirable to represent face images in terms of horizontal, vertical, and angular dimensions and select a subset of these dimensions such that the selected subset could discriminate between images that belong to two different age groups. The F -test (Snedecor and Cochran [33]) was used for dimension selection, while

canonical correlation analysis (CCA) (Hotelling [34,35]) was used to find a linear combination of a selected subset of dimensions and hence calculate a discriminating score to classify the face images that belong to two different age groups. Finally, the discriminating scores were applied as an input to a support vector machine (SVM)-based binary tree classifier to achieve the age group estimation. An overview of the proposed approach is shown in Fig. 8.

Dimension Selection

A qualitative analysis of the projection of the original 15 horizontal, 11 vertical, and 8 angular dimensions on the first four principal components showed that it was not possible to discriminate face images that belong to two given age groups based on these 34 dimensions, altogether. This outcome revealed the fact that the original 34-dimensional space was unable to discriminate between two age groups, based on their simple linear combinations. Hence, a statistical procedure was used to determine whether a subset of the original dimensions could be discriminative between two age groups for a given set of face images, as illustrated below.

Given a training set G with S samples, such that,

$$G = \{(x_i, y_i) : x_i \in R^v, y_i \in N, i \in [1, s]\}, \quad (5)$$

where x_i represents v -dimensional feature vectors, which represent the original dimensions from the i th training face image and y_i represents the corresponding labels. The aim is to select a subset H with dimension v' of the original dimensions such that,

$$H = \{x'_i : x'_i \in R^{v'}, x'_i \subset x_i, i \in [1, s]\}, \quad (6)$$

where $v' < v$, which can retain discriminative dimensions to classify face images that belong to two given age groups. For the age group classification task, the labels $y_i \in \{0, 1\}$ represent the respective age groups, that is, {Agegroup1, Agegroup2}. To accomplish this goal, we divided a set of face images into two groups, with a set of 34 original dimensions extracted for each face image. The univariate F -test (33) was performed for each of

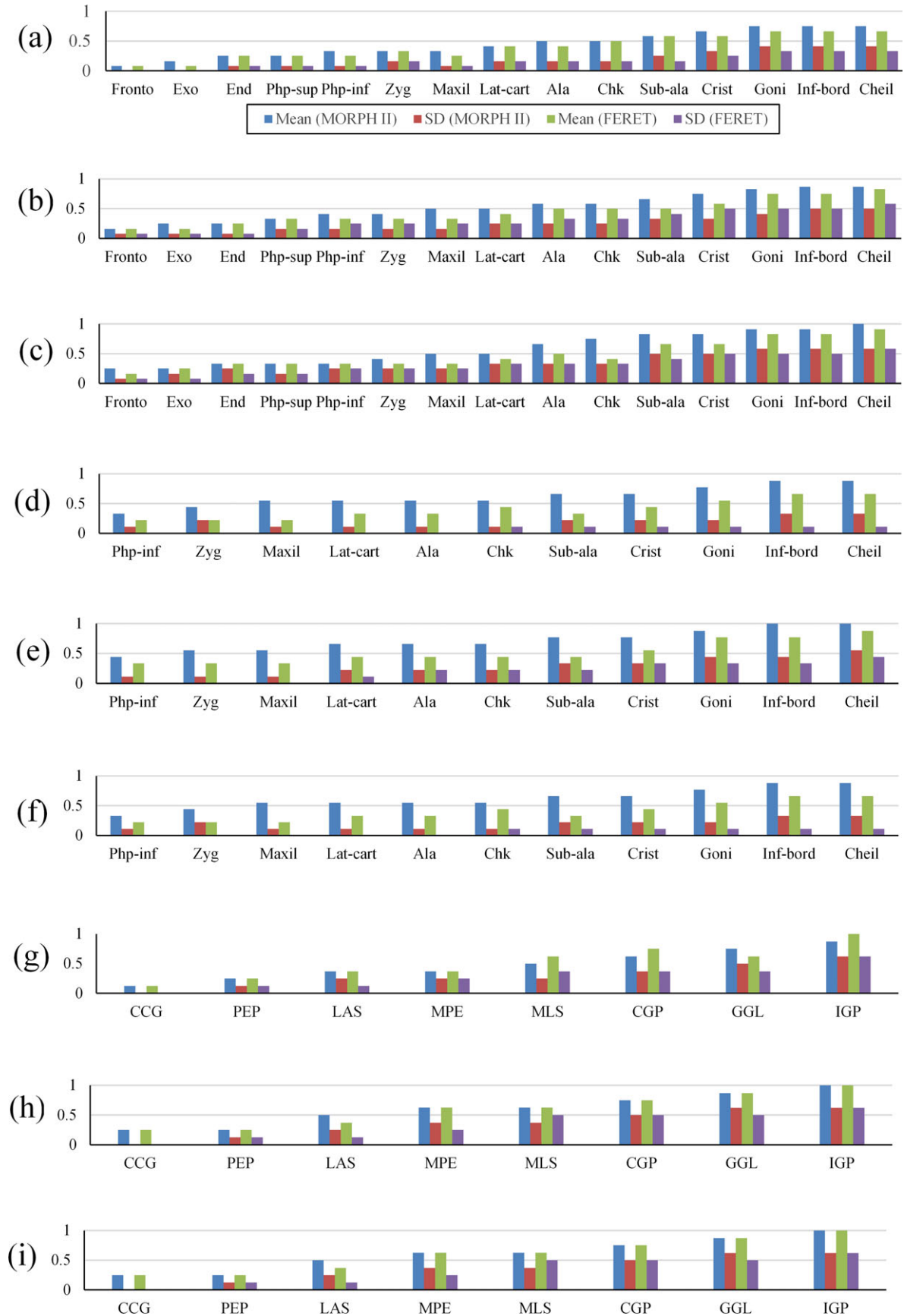


FIG. 4—Comparison of mean and standard deviation: (a-c) horizontal dimensions, (d-f) vertical dimensions, and (g-i) angular dimensions for preprocessed face images from FERET and MORPH datasets.

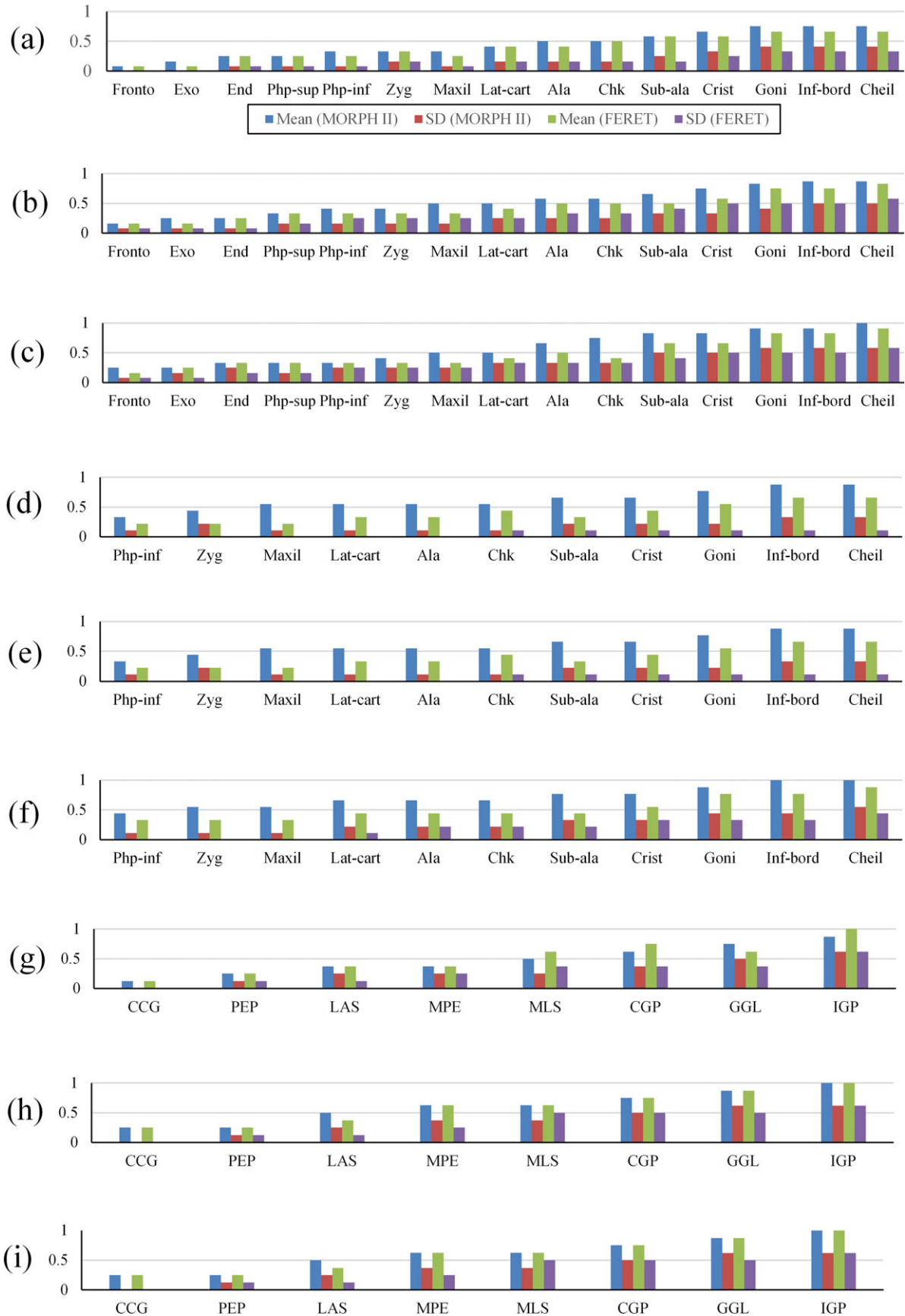
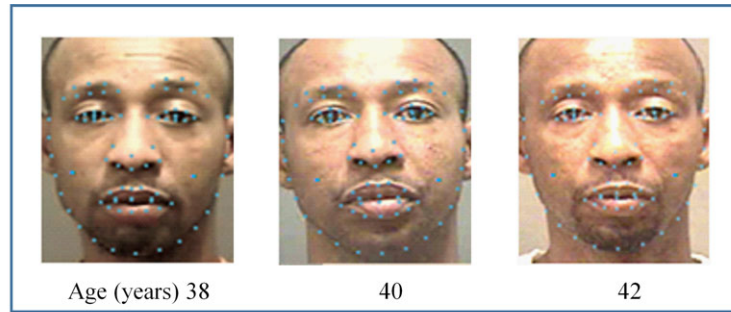
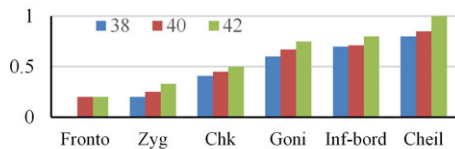


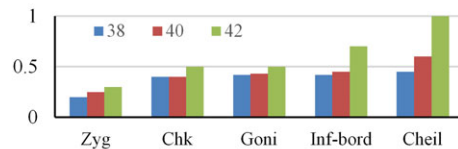
FIG. 5—Comparison of mean and standard deviation: (a-c) horizontal dimensions, (d-f) vertical dimensions, and (g-i) angular dimensions for raw face images from FERET and MORPH datasets.



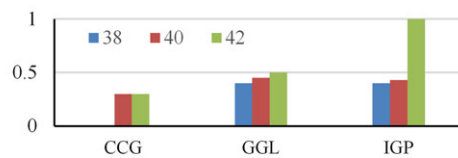
(a)



(b)



(c)



(d)

FIG. 6—Facial asymmetry variation across time lapse for a single subject: (a) face images with temporal variations, (b) horizontal, (c) vertical, and (d) angular asymmetric variations for selected landmarks of subject shown in part (a).

the 34 original dimensions to measure the significant difference between the means of the dimensions of two age groups. An F -test for a particular dimension can be defined as in Eq. 7.

$$F = \frac{\text{Total variance}}{\text{Average within group variance}}. \quad (7)$$

A larger value of F for a given dimension shows that it is more discriminating between two age groups, compared to other dimensions. The procedure was repeated for all dimensions, and we selected only those dimensions whose F values were greater than a predetermined threshold, as calculated empirically. All of the other dimensions were discarded, and further analysis was performed using the selected dimensions only. An F -test was performed on the MORPH II and FERET datasets, as described below.

- Ten thousand youngest face images from the MORPH II database (one image per subject) were selected, and 1196 face images of 1196 subjects (one image per subject) from the fa set of the FERET database were selected. An F -test was performed on the selected image sets by dividing the images into two age groups, age group 1 (16–30 for MORPH II and 10–30 for FERET datasets) and age group 2 (31–60+ in the case of both the MORPH II and FERET datasets), based on the ground truth information. Table 7 illustrates the F -test results for the two age groups for the MORPH II and FERET datasets, with the top seven selected dimensions

(HA1, HA2, HA3, VA1, VA2, AA2, and AA4) along with the corresponding p values, which suggests the significance of the selected dimensions. It can be observed that these seven dimensions (called the discriminating dimensions in the remainder of the text) are the most discriminating among the other dimensions for discriminating between the two mentioned age groups. This observation can be attributed to increased facial asymmetry with increasing age, because the facial hemi-sides are functionally asymmetric with increasing age (29), which is not surprising given the morphogenetic link between the brain and the craniofacial appearance.

- To assess the robustness of the proposed method, the procedure was repeated for selected sets from the MORPH II and FERET databases. The rate of recurrence of each measurement was recorded for each of the three image sets, as shown in Fig. 9a,b, respectively. The results show the most frequent occurrence of HA1, HA2, HA3, VA1, VA2, AA2, and AA4, which suggests the robustness of the proposed approach in selecting the most discriminating dimensions for the age group classification.

We also selected two other feature selection methods to check the validity of the F -test described above. These methods include (i) sequential floating forward selection (SFFS) (36) and (ii) correlation-based feature selection (37) (CFS). The chosen SFFS can be considered to be floating up and down during the process of dimension subset selection. After each forward step, SFFS performs backward steps as long as

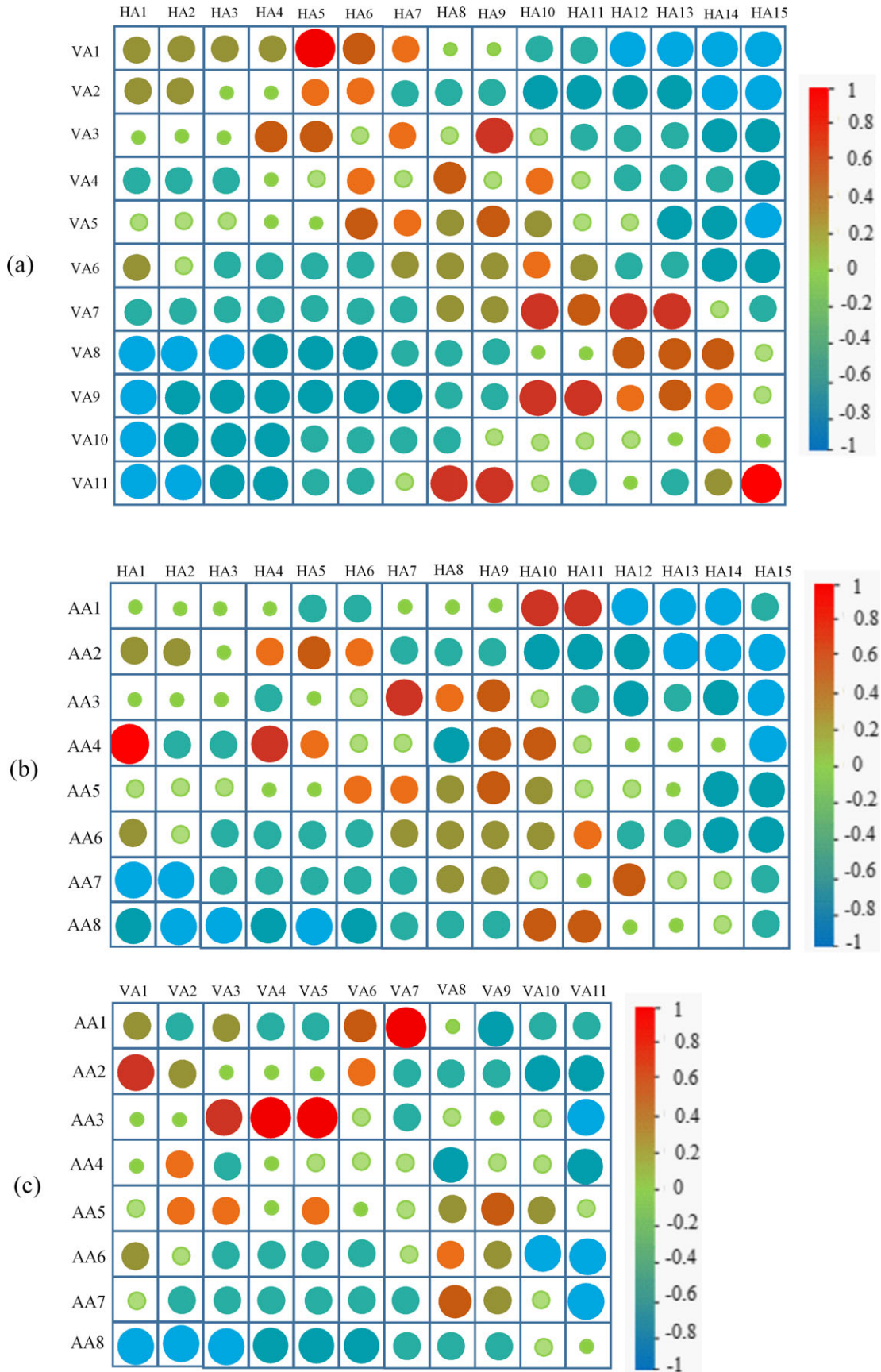


FIG. 7—Correlation matrices: (a) horizontal versus vertical, (b) horizontal versus angular, and (c) vertical versus angular dimensions. (Positive correlations are displayed in red and negative correlations in blue color. Color intensity and the size of the circle are proportional to the correlation coefficients. In the right side of each correlogram, the legend color shows the correlation coefficients in the range of -1 to 1 and the corresponding colors).

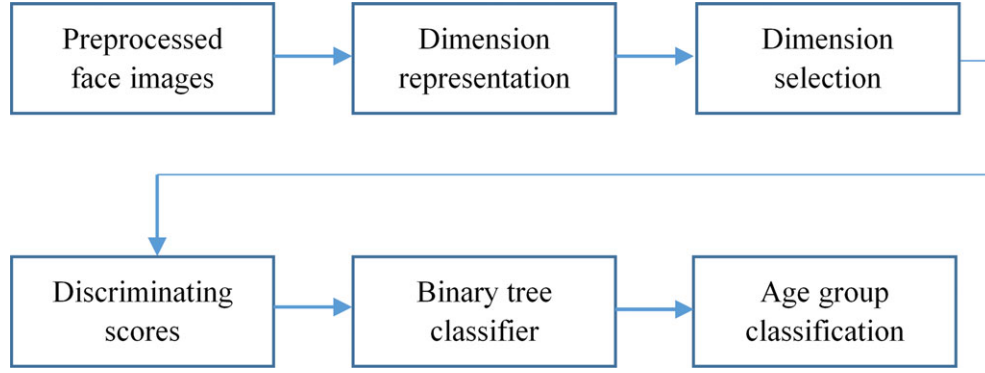


FIG. 8—Overview of the proposed age group estimation approach.

the objective function increases. The CFS method consists of two parts, a dimension correlation measure based on Pearson’s correlation coefficient and a best-first heuristic search algorithm that moves through the search space by greedy hill-climbing supported by a back-tracking facility. In practice, we performed dimension selection using F -test, SFFS, and CFS. The selected dimensions by all three methods are shown in Fig. 9a,b for both the MORPH II and FERET datasets, respectively. It can be observed that the occurrence of selected dimensions was the same for all three methods, which suggests the efficacy of the proposed method for dimension selection.

The same procedure was repeated by splitting the selected image sets from the MORPH II and FERET datasets into four age groups, age group 11, age group 12, age group 21, and age group 22 (16–20, 21–30, 31–45, and 46–60+ for MORPH II, while 10–20, 21–30, 31–45, and 46–60+ for the FERET datasets, respectively). It can be observed that dimensions HA1, HA4, HA7, VA1, VA3, AA1, and AA4 are the most discriminating between age group 11 and age group 12, while the dimensions HA2, HA7, HA13, VA1, VA3, VA6, and AA7 are most discriminating between age group 21 and age group 22, for both the MORPH II and FERET datasets.

Discriminating Scores

After the selection of a subset of the most discriminating dimensions through the F -test, the next stage is the learning of (i) a subspace, based on a linear combination of discriminating dimensions, to calculate the discriminating scores to classify the face images into two age groups and (ii) the relative contribution of each discriminating dimension to classify the two age groups. For this purpose, the well-known canonical correlation analysis (CCA) was used, as described below.

Canonical Correlation Analysis (CCA)—For a given set of objects with two different representations, canonical correlation analysis (CCA) aims at computing a projection such that the correlation between two representations is maximized in a subspace with reduced dimensionality.

Consider two multivariate random variables, $u \in \mathbb{R}^{D_x}$ and $v \in \mathbb{R}^{D_y}$. Let the groups $G_x = \{x_1, x_2, \dots, x_n\}$ and $G_y = \{y_1, y_2, \dots, y_n\}$ be paired. CCA aims at finding new coordinates for x and y by selecting directions $\mathbf{a} \in \mathbb{R}^{D_x}$ and $\mathbf{b} \in \mathbb{R}^{D_y}$ such that the

correlation between the projections of G_x and G_y on \mathbf{a} and \mathbf{b} is maximized, that is,

$$\rho = \max_{\mathbf{a}, \mathbf{b}} \frac{\mathbf{a}' C_{xy} \mathbf{b}}{\sqrt{\mathbf{a}' C_{xx} \mathbf{a}} \sqrt{\mathbf{b}' C_{yy} \mathbf{b}}}, \quad (8)$$

where ρ is the correlation, C_{xy} is the between-group covariance matrix, and C_{xx} and C_{yy} are the within-group covariance matrices. The problem can be reduced to a generalized eigenvalue problem, as shown in Eq. 9.

$$C_{xx}^{-1} C_{xy} C_{yy}^{-1} C_{yx} = \lambda \mathbf{a}, \quad (9)$$

where \mathbf{a} corresponds to the top eigenvector.

In the case of the age group classification task, it has been observed that a projection direction with the largest eigenvalue was significantly higher than the second largest eigenvalue, which verifies the possibility of representing and hence discriminating the data in a single dimension. The relative contribution of each discriminating dimension was also calculated in terms of the magnitude and polarity of the relevant canonical coefficients. The linear combination of such dimensions then resulted in a linear function that represents the face images from the two age groups. For a set of 200 face images from the youngest face image set of the MORPH II dataset, sample canonical coefficients with relative magnitude and polarity are shown in Table 8. It has been observed that the larger the values of HA1 and VA1 are, the more likely it is that the subject belongs to age group 1. Similarly, the larger the values are of HA2, HA3, and VA2, the more likely that the subject will belong to age group 2. A sample scatter plot that shows the discriminating scores for two age groups (16–30 and 31–60+) for the 200 youngest face image set from the MORPH II dataset is shown in Fig. 10.

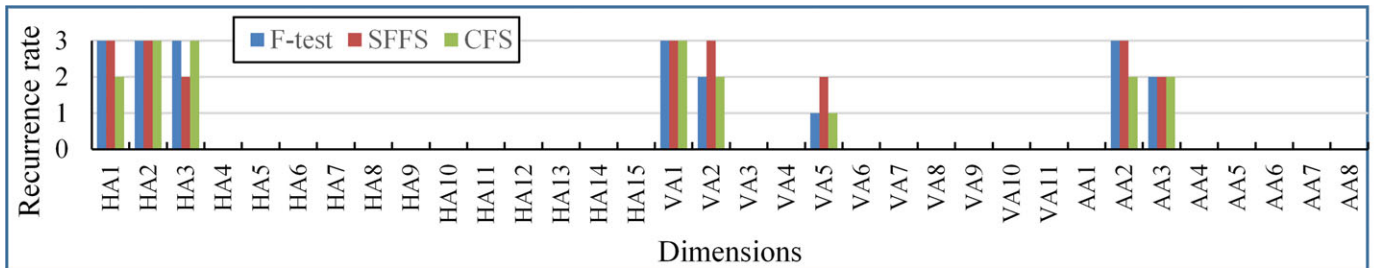
From Fig. 10, it can be observed that the discriminating scores determined through CCA cannot completely distinguish between the two age groups due to overlap. To improve the classification, the discriminating scores were applied to an SVM-based binary tree classifier, as explained in the following subsection.

Binary Tree Classification for Age Group Estimation

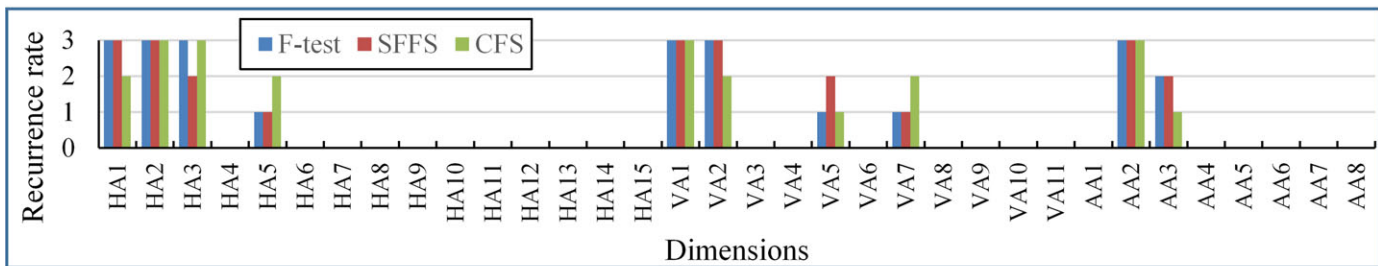
The age group classification problem presented in this study can be defined as a binary classification task. To handle such a classification problem, a two-level hierarchical age group

TABLE 7—F-test results for top seven discriminating dimensions for MORPH II and FERET datasets.

Selected Dimensions	MORPH II							FERET						
	HA1	HA2	AA2	VA2	AA4	HA3	VA1	HA1	HA2	AA2	VA2	AA4	HA3	VA1
F values	7.0	4.1	3.1	2.5	2.0	1.8	1.2	6.6	3.2	3.0	2.0	2.1	1.1	1.0
p values	0.001	0.000	0.002	0.003	0.001	0.002	0.001	0.000	0.000	0.001	0.004	0.002	0.001	0.002



(a)



(b)

FIG. 9—Rate of recurrence of selected dimensions extracted from three different image sets from (a) MORPH II and (b) FERET datasets.

classification method was proposed, which uses discriminating scores, calculated from CCA, as an input to the SVM-based binary tree classifiers, as shown in Fig. 11. Three binary SVM classifiers (C , C_1 , and C_2) were used to learn a two-level binary decision tree (Vapnik [38]). A given test face image can then be classified either into one of the binary age groups (age group 1 or age group 2) at level 1 or into one of four age groups (age group 11, age group 12, age group 21, and age group 22) at level 2 of the binary decision tree.

The support vector machine has been successfully used for age classification problems, as reported in (5). In the current study, a publicly available LIBSVM source (Chang and Lin [39]) was used. The labels (0, 1) were used for binary classification of two age groups when training the SVM-based binary tree classifiers. The RBF kernel was used with SVM with parameters c and γ selected using a 5-fold cross-validation protocol on the training set. The thresholds t , t_1 and t_2 , which were used to partition the age range, were empirically determined, due to the overlapping nature of the discriminating scores learned from the CCA algorithm.

Previously, hierarchical methods have been studied for age group estimation problems, such as in Choi et al. (40) and Thukral et al. (41). The current approach differs from these methods in that instead of dividing the entire age range directly into a certain number of age groups (e.g., four age groups for

the age range of 0–70 years in [41]), the proposed approach uses a coarse-to-fine strategy for the age group classification, which is similar to that proposed in Han et al. (6). However, the proposed approach uses a two-stage age group classification, first by calculating discriminating scores using canonical correlation analysis (CCA) and then using an SVM binary classifier. The second-stage classification mitigates the classification errors produced by the CCA due to the overlap of the discriminating scores.

Evaluation and Experimental Results of the Age Group Estimation Approach

To evaluate the performance of the proposed approach, the 10,000 youngest face images (of 10,000 subjects) from the MORPH II database were partitioned into a training set and a test set with 5000 face images in each set, with an age range of 16–60+ years. Similarly, 1196 face images from the fa set of the FERET database were partitioned into a training set and a test set, with 598 face images in each set and an age range of 10–60+ years. A fivefold cross-validation method was used to evaluate the proposed age group estimation algorithm. The age group estimation accuracy is shown in a trained predicted confusion matrix in Table 9a,b, in which the ages of the subjects belong to one of four age groups: 16–20, 21–30, 31–45, and

46–60+ for the MORPH II dataset, while 10–20, 21–30, 31–45, and 46–60+ for the FERET dataset.

Comparison with Existing Methods—The performance of the proposed age group estimation approach was compared with the following existing state-of-the-art approaches.

- A local binary pattern (LBP)-based approach proposed in Ylioinas et al. (3).
- Face ++ (31), an independently trained, state-of-the-art age group estimation algorithm.

Sensitivity, which represents the true-positive performance (42), was computed as a measure of the overall accuracy for the proposed and existing methods. The LBP-based approach (Ylioinas et al. [3]) gave an overall age group accuracy of 79.41% and 77.99% on the MORPH II and FERET datasets, respectively. The Face ++ (31) gave an overall age group estimation accuracy of 87.26% and 86.23% on the MORPH II and FERET datasets, compared to the proposed method, which resulted in an overall accuracy of 93.86% and 92.98% on the MORPH II and FERET datasets, respectively. Table 10*a,b* summarizes the age group estimation accuracies for the proposed and existing methods.

Effect of Asymmetric Facial Regions on Age Group Estimation

Variations in the facial asymmetry were observed for different facial landmarks while evaluating the facial asymmetry in Section 2. Motivated by this fact, the next goal is to understand the effect of various asymmetric facial regions on the age group estimation performance. To accomplish this goal, each face image was divided into four distinct regions, region 1 (landmarks: 5, 6, 7, and 8), region 2 (landmarks: 1, 2, 3, 4, 9, 10, 11, and 12), region 3 (landmarks: 13, 14, 15, 16, 17, and 18), and region 4 (landmarks: 19, 20, 21, 22, 23, and 24). Each region was used individually to estimate the age groups using a similar experimental setup as employed in the age group estimation task described in Section 3.4. The results for this experiment are shown in Table 11*a,b*, for both the MORPH II and FERET datasets. It can be observed that the age group estimation accuracy increases from the upper to the lower regions of the face, which is attributed to more pronounced facial asymmetry for these regions and hence a more prominent role of the corresponding dimensions in the age group estimation. For example, in the case of age group 16–20 for the MORPH II dataset, region 4 (composed of the lower facial landmarks) gave an estimation accuracy of 46.56% compared to region 1, which had an estimation accuracy of 37.95% (composed of the upper facial landmarks). Thus, we conclude that the more asymmetric a facial region is, the more it contributes to estimating the age group of a given facial image. To further validate the role of different asymmetric facial regions in the age group estimation, we generated a symmetrical face by combining the left half-face and its mirror image, as shown in Fig. 12. A low, moderate, and high level of facial asymmetry is then induced in the symmetrical face image using a deformation technique based on moving least squares (43), such that the line that joins the left-

TABLE 8—Relative contributions of discriminating dimensions determined using CCA.

Selected dimensions	HA1	HA2	HA3	VA1	VA2
Canonical coefficients	−12.05	3.6	7.9	−6.0	2.5
(magnitude and polarity)					

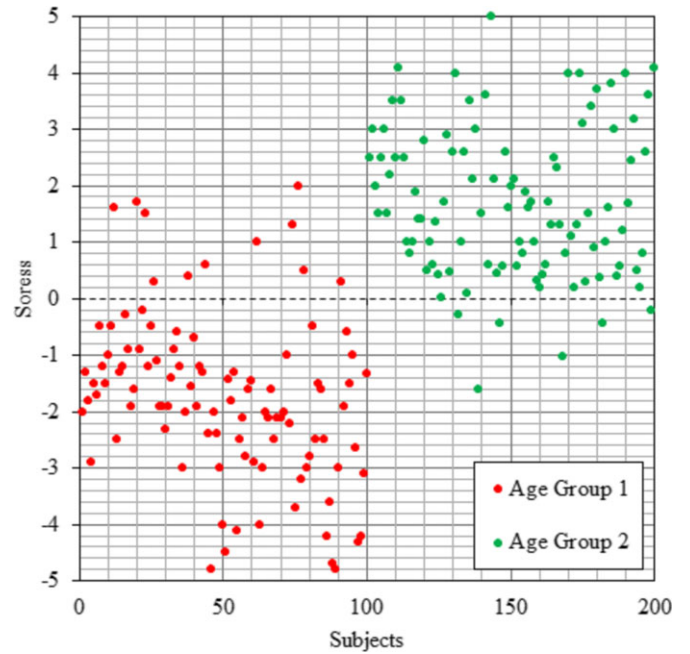


FIG. 10—A sample scatter plot showing discriminating scores for two age groups of 200 subjects from MORPH II dataset.

and right-sided lateral cartilage is rotated 5°, 10°, and 15° for the low, moderate, and high level of facial asymmetry, respectively. The deformed face images with induced facial asymmetry are shown in Fig. 12*c–e*. We repeated the age group estimation experiments using probe face image sets with low, moderate, and high facial asymmetry. The age group estimation results for the three probe sets are reported in Table 12*a,b* for the MORPH II and FERET datasets, respectively. The results show that higher estimation accuracies are achieved with higher levels of facial asymmetry compared to probe sets with low and moderate levels of asymmetry.

Recognition of Age-Separated Face Images

After the age group estimation task, the next goal is to design an age-assisted face recognition algorithm. The proposed algorithm seeks to recognize age-separated face images based on the knowledge learned from age group estimation of a query face image. Before the implementation of such an algorithm, it is desirable to understand the roles of various asymmetric facial regions in recognizing face images across aging variations. For this purpose, a feature extraction scheme that is suitable for extracting asymmetrical facial features is proposed, as follows:

Deeply learned Aggregated Asymmetric Facial Features (da-AFF)

In this work, we aim at extracting deeply learned asymmetric facial features (da-AFF) for face recognition. To accomplish this goal, we first extract a difference half-face image followed by extracting asymmetric facial regions and asymmetric facial features (AFF), as follows:

Extraction of the Difference Half-Face Image—First, the pre-processed frontal face images from the MORPH II and

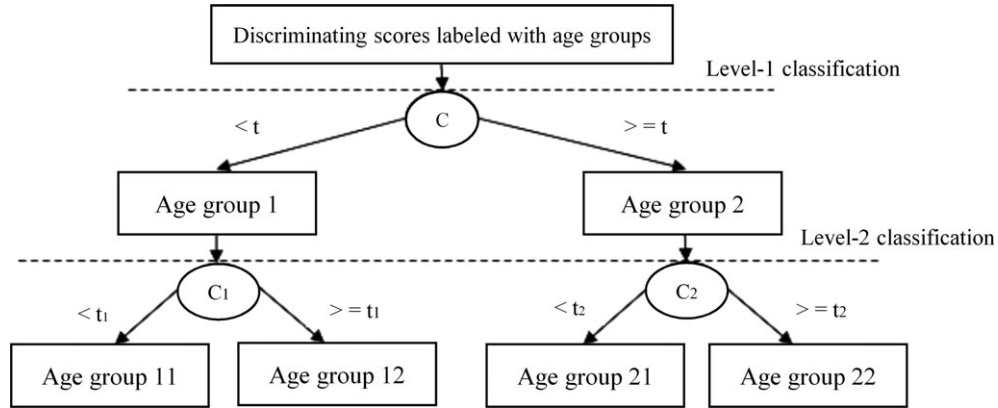


FIG. 11—Two-level hierarchical binary tree classifier for age group classification.

TABLE 9—Confusion matrix for age group classification accuracy (%) for (a) MORPH II (b) FERET datasets.

(a)				
Training Age Groups	Predicted Age Groups			
	16–20	21–30	31–45	46–60+
16–20	748	58	10	6
21–30	20	1349	40	11
31–45	5	30	2048	50
46–60+	11	9	21	584

(b)				
Training Age Groups	Predicted Age Groups			
	10–20	21–30	31–45	46–60+
10–20	211	11	7	5
21–30	2	122	3	3
31–45	0	1	80	3
46–60+	1	3	7	139

The bold values show the classification accuracies of respective age groups

FERET datasets were masked to remove unwanted hair and background regions. A left–right difference half-face image (called the difference half-face in the remainder of the text)

was extracted by subtracting the mirror image of the left half-face (LHF) from the right half-face (RHF). Let $I_L(x, y)$ and $I_R(x, y)$ represent the left and right half-face images, respectively. The difference half-face image is then represented as shown in Eq. 10.

$$\text{Difference half face} = I_R(x, y) - I'_L(x, y), \quad (10)$$

where $I'_L(x, y)$ is the mirror image of $I_L(x, y)$. The whole procedure is illustrated in Fig. 13.

Extraction of Asymmetric Facial Regions—The difference half-face image extracted above contains the left–right asymmetric facial differences. To extract the corresponding asymmetric facial regions from the difference half-face, attribute profiles (APs) have been proposed. Morphological profiles (MPs), defined in Pesaresi and Benidiktsson (44), are based on antigranulometry and granulometry (closing and opening by reconstruction transformations), respectively. The morphological opening profile (Π_γ) for an image f can be defined as an array of n openings, which is calculated on the original image, as shown in Eq. 11.

$$\Pi_\gamma(f) = \{\Pi_{\gamma\lambda}(f) : \Pi_{\gamma\lambda}(f) = \gamma_\lambda^{\lambda}(f)\}; \quad \lambda = 0, 1, 2, \dots, n. \quad (11)$$

Similarly, the morphological closing profile (Π_ϕ) is defined in Eq. 12.

TABLE 10—Age group estimation performance comparison of proposed and existing methods for (a) MORPH II (b) FERET datasets.

(a)					
Method	Age Groups				Sensitivity or Overall Accuracy (%)
	16–20	21–30	31–45	46–60+	
Variants of LBP (Ylioinas et al. [3])	81.99	80.00	76.98	78.67	79.41
Face++ (31)	86.49	86.97	90.62	84.96	87.26
Proposed	90.99	95.00	96.01	93.44	93.86

(b)					
Method	Age Groups				Sensitivity or Overall Accuracy (%)
	10–20	21–30	31–45	46–60+	
Variants of LBP (Ylioinas et al. [3])	80.76	77.69	76.19	77.33	77.99
Face++ (31)	85.47	86.15	89.28	84.00	86.23
Proposed	90.17	93.84	95.23	92.67	92.98

TABLE 11—Age group estimation accuracy of different facial regions for (a) MORPH II and (b) FERET datasets.

Asymmetric Facial Regions' Accuracy (%)				
(a)	Region 1	Region 2	Region 3	Region 4
Age Group				
16–20	37.95	43.66	44.02	46.56
21–30	35.03	39.50	40.03	40.96
31–45	31.99	41.47	38.02	42.08
46–60+	27.61	37.53	38.02	40.48

Asymmetric Facial Regions' Accuracy (%)				
(b)	Region 1	Region 2	Region 3	Region 4
Age Group				
10–20	38.03	43.07	45.23	47.33
21–30	33.33	38.46	39.28	40.00
31–45	30.34	36.15	37.90	39.33
46–60+	26.06	36.15	38.09	41.33

$$\Pi_{\varphi}(f) = \{\Pi_{\varphi\lambda}(f) : \Pi_{\varphi\lambda}(f) = \varphi_{\mathbb{R}}^{\lambda}(f)\}; \lambda = 0, 1, 2, \dots, n. \quad (12)$$

Hence, a morphological profile of size $2n-1$ (at $\lambda = 0$, the opening and closing profiles are equal to the original image, and hence, both are considered only once) can be obtained by combining closing and opening profiles, each of size n . Hence, MPs can be defined as given in Eq. 13.

$$\text{MP}(f) = \left\{ \begin{array}{ll} \Pi_{\varphi\lambda}(f), & \lambda = (n-1+i) \quad i \in [1, n] \\ \Pi_{\gamma\lambda}(f), & \lambda = (i-n-1) \quad i \in [n+1, 2n+1] \end{array} \right\}, \quad (13)$$

In Eq. 13, i is the size of the structuring element. Although the MPs are successful in analyzing the image structures that are under consideration, such an analysis is only partial. To describe a shape more precisely, other features, such as the shape and texture, can also be accounted for. This approach will lead to an increase in the discriminative power of such analysis in classification problems. Second, the structuring elements in the MPs are not able to completely describe the gray-level characteristics of the underlying shape, such as the homogeneity and contrast. Finally, MPs suffer from higher computational complexity. These drawbacks of MPs can be eliminated by defining attribute profiles (APs).

Attribute profiles perform image analysis similar to that of MPs but with greater flexibility in terms of the filtering criterion. The selection of different attributes makes the characterizations of the underlying structures possible. Analogously to Eq. 13, the attribute profile (AP) can be defined as shown in Eq. 14.

$$\text{AP}(f) = \left\{ \Pi_i : \left\{ \begin{array}{ll} \Pi_i = \Pi_{\varphi}u_{\lambda}, & \lambda = (n-1+i), \quad \forall \lambda \in [1, n]; \\ \Pi_i = \Pi_{\gamma}u_{\lambda}, & \lambda = (i-n-1), \\ & \forall \lambda \in [n+1, 2n+1] \end{array} \right. \right\}. \quad (14)$$

In the current study, APs have been proposed to extract asymmetric facial regions on a given set of difference face images. A set of four APs was generated for each difference face image, which includes the following:

- Area attribute to measure the area of the facial asymmetric regions.

- Moment of inertia attribute to describe the geometry of the underlying asymmetric facial shapes.
- Diagonal attribute to measure the length of the diagonal of the bounding box. It is worthwhile to note that the diagonal attribute can detect the facial asymmetries that are caused by facial marks such as moles or scars by encoding each mark with a bounding box.
- Standard deviation to measure the homogeneity of the intensity values of the underlying pixels.

Once generated, the combined use of four APs together can extract the underlying facial asymmetric regions with corresponding temporal variations, from a difference half-face image, more effectively. A sample difference half-face image and corresponding image with asymmetric facial regions extracted using the combined use of four APs are shown in Fig. 14a,b, respectively.

Extraction of Asymmetric Facial Features (AFF)—Once the asymmetric facial regions have been extracted, the next stage is extracting the asymmetric facial features (AFF). For this purpose, each difference half-face image was divided into a set of overlapping patches. To characterize the asymmetric facial regions, seven geometrical Hu moment invariants (45) were extracted for each patch and concatenated together to form a single AFF feature vector. These Hu moments are invariant to rotation, translation, and scale changes. The procedure to extract the AFF feature vector is as follows:

Consider a difference half-face image of size $h \times w$, with a patch size $p \times q$ and overlapping radius r ; then, the number of horizontal (A) and vertical (B) patches obtained is given in Eqs 15 and 16, respectively.

$$A = \frac{h-p}{p-r} + 1, \quad (15)$$

$$B = \frac{w-q}{q-r} + 1. \quad (16)$$

For each patch, a d -dimensional Hu moment invariant was extracted, which resulted in a $d \times A \times B$ -dimensional AFF feature vector. A similar procedure was adopted to extract the AFF features from each facial region of the difference half-face image. To achieve this goal, the difference half-face was first divided into four regions, Region 1, Region 2, Region 3, and Region 4. Each region was then divided into $C \times D$ overlapping patches to extract d' -dimensional Hu moment invariants for each patch, which resulted in a $d' \times C \times D$ -dimensional AFF descriptor for each region.

The feature extraction step explained above generates a set $A = a_1, a_2, \dots, a_t \in \mathbb{R}^d$ of local descriptors. However, these handcrafted descriptors are not optimal for local feature representation. To address this problem, we propose deeply learned asymmetric facial features in the following two steps.

Step 1: First, we extract a compact AFF code based on feature aggregation. Successful techniques for feature aggregation for deriving image representations from local descriptors include (i) bag-of-words (BoW) (46), (ii) Fisher vectors (FV) (47), and (iii) vector of locally aggregated descriptors (VLAD) (48).

The BoW is essentially a histogram in which local features are extracted from each image, and each feature is assigned to the nearest visual word from a visual vocabulary. The result of



FIG. 12—Face image deformation: (a) original face image, (b) symmetrical face, (c) deformed face with low asymmetry, (d) deformed face with moderate asymmetry, and (e) deformed face with high asymmetry.

TABLE 12—Age group estimation accuracies for deformed face images with low, moderate, and high facial asymmetry for (a) MORPH II and (b) FERET datasets.

Accuracy (%)												
(a)	Probe Set 1 (low asymmetry)				Probe Set 2 (moderate asymmetry)				Probe Set 3 (high asymmetry)			
Age Group	Region 1	Region 2	Region 3	Region 4	Region 1	Region 2	Region 3	Region 4	Region 1	Region 2	Region 3	Region 4
16–20	32.05	38.00	40.02	42.80	37.00	44.00	43.00	45.00	42.00	48.66	47.70	51.01
21–30	31.00	35.50	35.67	37.00	36.90	39.00	41.00	41.00	39.65	43.50	44.44	45.00
31–45	29.00	36.30	33.78	37.00	31.00	40.60	39.21	41.99	36.20	43.74	42.00	45.98
46–60+	22.00	32.94	34.00	36.01	28.90	38.82	37.99	40.50	31.00	39.92	42.00	44.00

Accuracy (%)												
(b)	Probe Set 1 (low asymmetry)				Probe Set 2 (moderate asymmetry)				Probe Set 3 (high asymmetry)			
Age Group	Region 1	Region 2	Region 3	Region 4	Region 1	Region 2	Region 3	Region 4	Region 1	Region 2	Region 3	Region 4
16–20	32.21	37.99	40.00	43.10	37.81	45.01	42.50	45.80	43.23	47.99	47.10	52.00
21–30	32.32	35.00	36.10	37.32	37.00	38.77	41.07	40.79	39.50	43.00	44.00	46.00
31–45	29.99	37.50	34.01	38.67	33.00	40.00	39.00	40.39	37.00	44.00	43.00	46.00
46–60+	22.05	33.00	33.99	36.00	29.53	39.23	37.90	41.20	31.93	39.95	42.60	44.56

this process is a fixed length n -dimensional representation of an image. Despite their success, BoW-based methods suffer from computational constraints.

Fisher vector (FV) encoding aggregates local image descriptors based on a Fisher kernel framework. Compared to the BoW, which only considers the sum of local descriptors in each visual word, the FV accounts for the higher order statistics, which results in a more discriminative representation and hence improved performance. However, FV assumes that the local descriptors follow a Gaussian mixture model, which rarely occurs in practice.

To overcome this limitation, a simplified version of FV called VLAD was introduced in (48). In case of VLAD, a codebook $\{c_1, c_2, \dots, c_n\}$ of n cluster centers is learned by applying the K -means algorithm, and each descriptor $a_t \in \mathbb{R}^d$ is hard-assigned to its nearest cluster center $\text{NN}(a_t)$. More precisely, we compute the cluster-level representations $\delta_i \in \mathbb{R}^d$ by aggregating the differences between the descriptors and their corresponding cluster centers, that is,

$$\delta_i = \sum_{a_t: \text{NN}(a_t)=i} a_t - c_i. \quad (17)$$

Finally, D -dimensional VLAD is obtained by concatenating all of the aggregated vectors δ_i for all clusters $1, 2, \dots, n$. The resulting aggregated AFF-VLAD representation is called a-AFF in the remainder of the text.

Step 2: Having obtained a-AFF, we also derive deep convolutional neural networks (dCNN)-based a-AFF features, called da-AFF. Recent research shows that image descriptors that use

dCNNs achieve state-of-the-art performance, such as in (49), (50). To encode dCNN-based a-AFF descriptors, a difference half-face image of size $h \times w$ is passed through a deep network architecture called AlexNet (51). The choice of this architecture was made due to its better performance. This network consists of five convolutional layers (C1, C2, C3, C4, C5), three pooling layers (P1, P2, P3), and 3 fully connected layers (f6, f7, f8). Each convolutional layer is followed by a rectified linear unit (ReLU). Fully connected layers are regularized using dropout. Figure 15 shows the graphical representation of the AlexNet architecture. Facial features are extracted from the second to last fully connected layer followed by an L_2 normalization step. The output of layer r is a set $A^r = \{a_1^r, a_2^r, \dots, a_n^r\}$ of asymmetric facial features. A codebook $\{c_1^r, c_2^r, \dots, c_n^r\}$ of n cluster centers is also learned in training. For each centroid, the difference vectors $a_n^r - c_n^r$ are computed and concatenated to form the da-AFF representation for all of the visual words. Figure 16 shows the pipeline of the da-AFF feature extraction.

Effect of the Asymmetric Facial Regions on the Face Recognition Accuracy

To understand the effect of the different asymmetric facial regions on the recognition accuracy, identification experiments were performed using the MORPH II and FERET training and probe sets. For this purpose, the 10,000 youngest face images from MORPH II were used in a gallery, while the older and

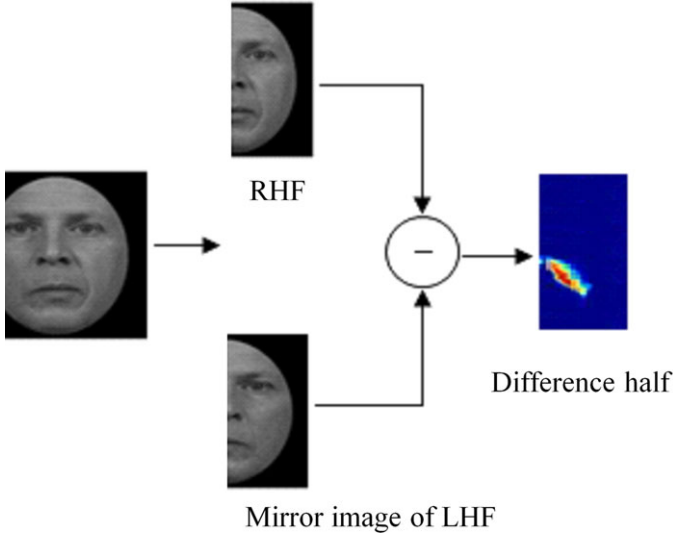


FIG. 13—Extraction of difference half-face.

oldest face image sets with 10,000 images each were used as probe sets. In the case of the FERET dataset, 1196 face images from the fa set were used as the gallery set, while the dup I and dup II sets were used as the probe sets. For each 128×64 difference half-face image, we extracted a 2048-dimensional da-AFF feature vector. Similarly, for each region of size 32×64 , a 512-dimensional da-AFF feature vector was extracted. Face identification experiments were conducted for the experimental setup described above.

The rank 1 identification accuracies for the difference half-face and four regions are reported in Table 13a,b for the MORPH II and FERET datasets, respectively.

On analyzing the accuracies of the various age groups, it can be observed that the recognition accuracy decreases for older age groups, compared to younger age groups. The decrease in the recognition accuracy can be attributed to the significant variations in the facial asymmetry with increasing age. For younger age groups (16–20 and 21–30 for MORPH II and 10–20 and 21–30 for FERET datasets), region 4 gave the highest recognition accuracy of 57.80% and 56.58% for the older image set and 52.10% and 51.50% for the oldest image set on MORPH II. The rank 1

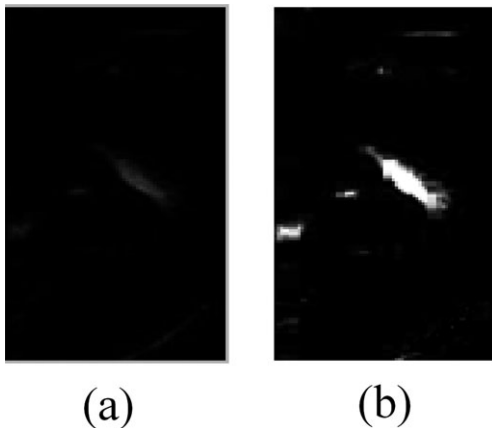


FIG. 14—Extraction of asymmetric facial regions using attribute profiles (a) original difference half-face image and (b) asymmetric facial regions extracted using attribute profiles.

recognition accuracy of 58.17% and 55.26% was achieved for the dup I set and 51.70% and 50.00% for the dup II set for the FERET datasets, respectively. The highest recognition accuracy of region 4 is attributed to having more symmetric information in this region. On the other hand, for the older age groups (31–45 and 46–60+), the highest recognition accuracy was achieved for the difference half-face, for both the MORPH II and FERET datasets, compared to individual regional accuracies. This finding occurs because facial asymmetry increases with the passage of time for all of the facial regions for these age groups. Thus, the difference half-face plays a significant role in recognizing face images across aging variations for older age groups, while for younger age groups, region 4 is the most discriminative.

Age-Separated Face Recognition Algorithm

The objective of this component of study is to build a face recognition algorithm based on the knowledge learned from the roles of various asymmetric facial regions in recognizing age-separated face images. The proposed approach is formulated as follows:

- Age group of probe face image is estimated.
- da-AFF feature vectors are extracted for difference half-face images, for both the gallery and probe images.
- Matching scores of da-AFF feature vectors are calculated for the gallery and probe images.
- To perform recognition of face images with aging variations, weights are assigned to different regions of the difference half-face images, based on their recognition significance as related to different age groups.

Figure 17 gives an overview of the proposed age group estimation-based face recognition approach. The implementation of the proposed face recognition approach is given in the following steps:

- To extract facial features, each face image was preprocessed.
- A 2048-dimensional da-AFF feature vector is extracted for each difference half-face image using features extracted from the f7 layer of AlexNet. Similarly, a 512-dimensional feature vector is extracted for each region of a difference half-face image.
- Matching scores of gallery and probe face images were calculated by comparing the respective L_2 -normalized da-AAF feature vectors. A matching-score matrix M of size $A \times B$ was calculated for the da-AFF-based matching scores. Here, A and B represent the number of probe and gallery images, respectively. The matching-score matrix M has a negative polarity, that is, smaller matching scores represent a higher similarity between the gallery and probe images and vice versa. Five matching-score matrices M_j were calculated, one for the normalized da-AFF features of the difference half-face and four for the da-AFF feature vectors of the four regions.
- Each matching-score matrix was normalized before fusion in the f8 layer of AlexNet. A simple min-max rule (51,52) was used to normalize each row, on a scale of 0–1, as shown in Eq. 18.

$$M'_{jrow} = \frac{M_{jrow} - \min(M_{jrow})}{\max(M_{jrow} - \min(M_{jrow})) - \min(M_{jrow} - \min(M_{jrow}))}, \quad (18)$$

where $\max(M_{jrow})$ and $\min(M_{jrow})$, respectively, represent the highest and lowest row-specific values in the matching-score matrix.

- The five normalized matching-score matrices, for the difference half-face and four regions, were then fused to obtain a combined matching-score matrix M_{row} using the weighted sum rule shown in Eq. 19.

$$M_{\text{row}} = \sum_{i=1}^5 w_i M_{\text{row}i}, \quad (19)$$

In Eq. 18, w_i represent the weights that pertain to the i th difference half-face or region and can be assigned using the recognition rates of the difference half-face and four regions, as given in Eq. 20.

$$w_i^k = \frac{r_i^k}{\sum r_i^k}, \quad (20)$$

where r_i^k represents the recognition rates of the k th age group of the gallery and probe face images and the i th difference half-face or region. For example, if the predicted age group of a probe image was 16–20, then weights could be assigned to the difference half-face and each region as per the recognition accuracies, as calculated for the corresponding age group given in Table 13. The matching-score matrix M_{row} was once again normalized using the min-max rule (53,54) to obtain M'_{row} , as shown in Eq. 21.

$$M'_{\text{row}} = \frac{M_{\text{row}} - \min(M_{\text{row}})}{\max(M_{\text{row}} - \min(M_{\text{row}})) - \min(M_{\text{row}} - \min(M_{\text{row}}))}. \quad (21)$$

- Finally, the matching scores obtained from M'_{row} were used to calculate the recognition accuracies in the SoftMax layer of the AlexNet.

Evaluation and Experimental Results of the Proposed Face Recognition Algorithm—To evaluate the performance of the proposed face recognition algorithm, face identification and face verification experiments were performed on the MORPH II and FERET aging datasets, as described below.

- In the case of the MORPH II database, the 10,000 youngest face images of 10,000 subjects were used as the gallery, while 10,000 face images each from the older and oldest face image sets were used as aging probe sets.
- In the case of the FERET database, standard gallery and probe image sets were used to evaluate the performance of the proposed face recognition algorithm. A total of 1196 face images from the fa set were used as the gallery set, while the dup I and dup II sets were used as the probe sets.

As shown in Fig. 16, the age group of the probe image was first predicted using the hierarchical age group estimation approach proposed in Section 3. Based on the predicted age group, weights were assigned to the appropriate difference half-face or regions depending on their significance in recognizing the age-separated face images of different age groups, as evaluated in Table 13. The rank 1 recognition accuracies and verification accuracies at 0.001 false accept rate (FAR) are reported in Table 14 for all of the face recognition experiments.

Comparison with Existing Methods—The performance of the proposed face recognition approach was compared with the following existing state-of-the-art approaches.

- A commercial off-the-shelf (COTS) system Verilook (52) is used as a baseline algorithm.
- Discriminative model (DM) (18). In DM, the MLBP and SIFT feature vectors were extracted for overlapping patches of face images. A discriminant analysis was then performed to construct a random subspace-based fusion model for face recognition.
- Bacteria foraging fusion (BFF) (19). In BFF, a local binary pattern (LBP) was used to calculate the matching scores for the face regions, including the mouth, binocular, and periocular. Region-specific weights were then optimized to achieve better recognition accuracy.
- Multiview discriminative learning (MDL) (20). In MDL, three locals (LBP, SIFT, and GOP) were extracted from overlapping patches of face images. Subsequently, the interclass separation was maximized and intraclass separation minimized by solving an optimization problem to recognize face images effectively.
- We also performed face identification and face verification experiments using a-AFF feature vectors to compare their performance with da-AFF feature vectors and the proposed approach.

The performance of the proposed face recognition method is quantified through two measures: (i) cumulative match characteristic (CMC) curves for face identification experiments and (ii) receiver operating characteristic (ROC) curves for face verification experiments, as shown in Figs 18 and 19, respectively. An analysis of the error introduced by the age group estimation of the probe images compared to the actual age groups, in recognizing face images, is also presented in row (vii) of Table 14.

Computational Complexity Analysis—It is worthwhile to underline the two-phased nature of the proposed method presented in this study. The first phase contains the procedures for age group estimation. Here, the computational complexity of the dimension selection through the F -test is $O(n^2)$ in terms of Big-O notation. The computational complexity of calculating the discriminating scores using CCA is $O(nd^2) + O(d^3)$, where $d = \max(D_x, D_y)$. For the classification stage, as we only need to train $N-1$ binary SVM classifiers, the order of complexity is $O(\log N)$.

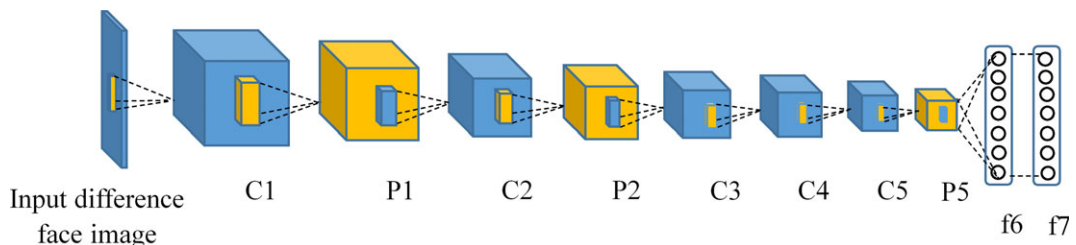
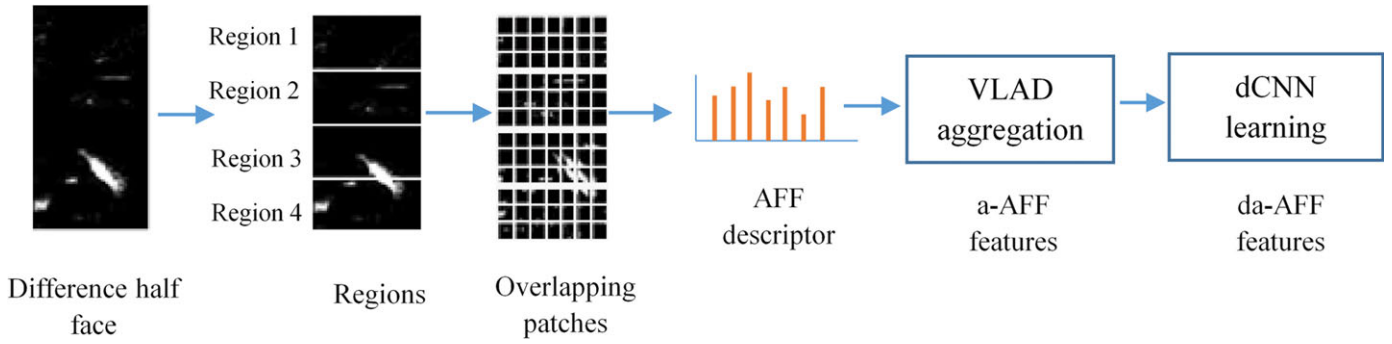


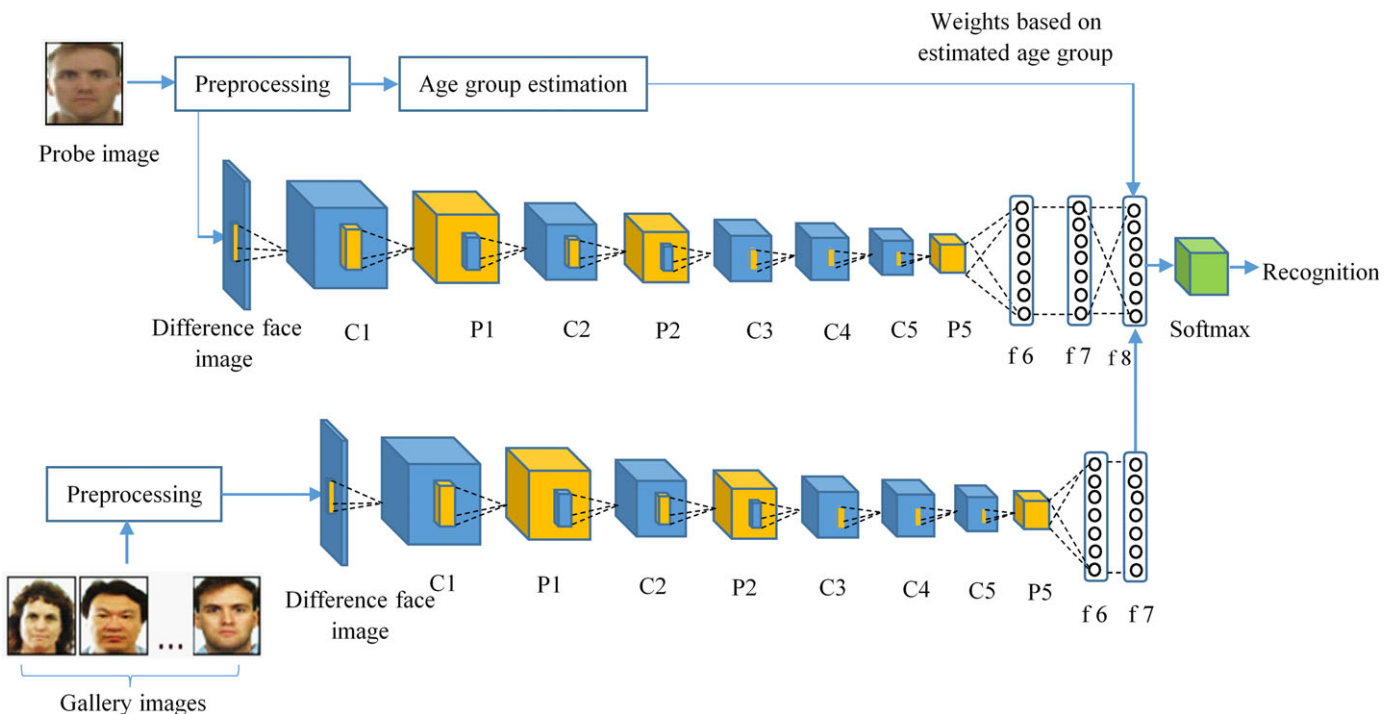
FIG. 15—Graphical representation of dCNN-based feature extraction using AlexNet.

FIG. 16—*da-AFF extraction pipeline.*TABLE 13—*Rank 1 recognition accuracies of different facial regions for (a) MORPH II and (b) FERET datasets.*

(a)	Rank 1 Recognition Accuracy for Difference Half-face and Four Regions for Older Image Set					Rank 1 Recognition Accuracy for Difference Half-face and Four regions for Oldest Image Set				
	Age Groups	Difference Half-Face	Region 1	Region 2	Region 3	Region 4	Difference Half-Face	Region 1	Region 2	Region 3
16–20	68.00	57.50	61.60	61.00	69.00	59.00	53.98	61.20	52.00	63.00
21–30	61.00	56.60	63.50	64.00	66.66	49.15	50.00	58.00	59.50	62.70
31–45	62.99	56.00	62.00	57.21	57.84	63.00	49.10	58.00	56.20	55.66
46–60+	63.41	50.00	60.00	58.50	56.00	55.31	43.00	40.00	51.92	54.00

(b)	Rank 1 Recognition Accuracy for Difference Half-face and Four Regions for Dup I Set					Rank 1 Recognition Accuracy for Difference Half-face and Four Regions for Dup II Set				
	Age Groups	Difference Half-face	Region 1	Region 2	Region 3	Region 4	Difference Half-face	Region 1	Region 2	Region 3
10–20	63.98	54.98	60.94	62.04	67.03	59.82	52.13	59.82	50.42	60.25
21–30	60.94	53.46	62.04	63.98	66.06	58.11	51.28	58.11	53.41	60.25
31–45	60.94	53.04	59.97	57.61	54.01	52.13	50.42	58.11	53.41	52.13
46–60+	63.98	51.93	59.00	57.06	54.98	54.70	40.17	34.18	52.13	51.28

The bold values show the highest accuracies.

FIG. 17—*Overview of the proposed method to recognize age-separated face images.*

The second phase consists of face recognition. The extraction of AFF feature vectors results in a linear complexity of $O(GN_{dp} + 4LN_{dp})$ for the APs, while it is $O(N_{dp}^2)$ for the Hu moment invariants, where G , L , and N_{dp} represent the number of gray levels, number of transformation levels in APs, and number of pixels in the difference half-face image, respectively. The total time complexity of AlexNet is computed in terms of all of its convolutional layers, as $O(\sum_{i=1}^n y_{i-1}x_i^2y_iz_i^2)$. Here, n is the number of convolutional layers, y_{i-1} is the number of input channels of the i th layer, y_i is the number of filters in the i th layer, x_i is the spatial size of the filters, and z_i is the size of the output feature map.

The above analysis shows that the calculation of discriminating scores using CCA and feature extraction using AlexNet are computationally the most demanding stages in the age group estimation and face recognition components of the proposed approach, respectively.

Results-Related Discussion

The research conducted in this paper covers two important facial aging aspects: age group estimation and face recognition. The performance of these two aspects is discussed separately.

Performance of the Age Group Estimation Algorithm

In the case of the age group estimation task,

- The proposed method achieved an overall age group estimation accuracy of 93.86% and 92.98% on the MORPH II and FERET datasets, compared to 79.41% and 77.99% given by LBP variants proposed in Ylioinas et al. (3), while Face++ gave an overall accuracy of 87.26% and 86.23% on the MORPH II and FERET datasets, respectively. Hence, the proposed method outperformed both of the existing methods considerably.
- The superior age group estimation performance shows that facial asymmetry is a strong indicator of age, as depicted in facial asymmetry measurement and evaluation.

Performance of the Face Recognition Algorithm

The accuracy of the proposed face recognition approach was compared with existing state-of-the-art methods. The effect of

errors introduced by the age group estimation of the probe face images into the face recognition algorithms was also analyzed.

- In the case of the MORPH II database, the rank 1 recognition accuracies of 90.02% and 80.00% were achieved for the older and oldest image sets, respectively. For the FERET database, the rank 1 recognition accuracies of 81.99% and 75.21% were achieved for the dup I and dup II sets, respectively. For face verification experiments, accuracies of 91.13% and 82.50% at 0.001 FAR were obtained for the selected subsets of the MORPH database. Similarly, for the FERET database, we obtained verification accuracies of 83.10% and 77.77% for the selected subsets.
- The identification and verification results reveal that the proposed approach outperformed the existing approaches, including COTS (52), sum rule fusion, DM (18), BFF (19), and MDL (20), on both the MORPH II and FERET datasets, which shows the effectiveness of the proposed approach in recognizing age-separated face images.
- The proposed automatically learned da-AFF descriptors yield superior recognition performance compared to the handcrafted features that are employed in the existing methods, such as DM (18), BFF (19), and MDL (20).
- To analyze the errors introduced by the age group estimation of the probe images, the rank 1 accuracies for face, difference half-face, and four regions were calculated for both scenarios: (i) with actual age groups and (ii) with estimated age groups. A decrease in accuracy of 2.68% and 1.46% was observed for the older and oldest image sets for the MORPH II dataset. For the FERET dataset, the accuracy decreases by 1.06% and 1.11% for the dup I and dup II sets, respectively. For the face verification experiments, these errors are 1.72% and 0.88% for the selected subsets of the MORPH II database, while they are 0.98% and 1.61% for the selected subsets of the FERET database. The minimal decrease in accuracy for both the MORPH II and FERET aging datasets shows the superior performance of the proposed age group estimation approach.
- The face recognition results across aging variations suggest that integration of the knowledge learned from age group estimation into the face recognition algorithm boosts the recognition accuracy of the face images considerably.

TABLE 14—Comparison of face recognition accuracies of proposed approach with existing algorithms for MORPH II and FERET databases.

Approach	Facial Region(s)	Rank 1 Recognition Accuracy (%)				Verification Accuracy @ 0.001 FAR			
		MORPH II PROBE SET		FERET Probe Set		MORPH II		FERET	
		Older Images	Oldest Images	Dup I	Dup II	Older Images	Oldest Images	Dup I	Dup II
(i) a-AFF	Difference half-face	70.00	66.00	61.35	60.68	72.02	64.10	66.00	64.10
(ii) da-AFF	Difference half-face	87.11	78.50	72.02	70.08	87.95	81.00	76.00	71.79
(iii) COTS (Verilook) (52)	Face	84.07	76.00	69.25	67.09	86.01	80.00	73.00	70.08
(iv) DM (18)	Overlapping patches	75.50	65.60	69.94	64.10	76.10	67.52	70.08	64.95
(v) BFF (19)	Face, mouth, binocular, periocular	71.00	64.00	68.42	61.11	72.01	65.39	68.97	62.39
(vi) MDL (20)	Overlapping patches	67.00	59.90	67.03	60.68	67.41	60.00	68.05	61.96
(vii) Proposed approach with weights assigned to difference half-face and regions (actual age groups)	Face, difference half-face, region 1, region 2, region 3, region 4	92.50	81.19	83.10	76.06	92.73	83.24	83.93	79.05
(viii) Proposed approach with weights assigned to difference half-face and regions (estimated age groups)	Face, difference half-face, region 1, region 2, region 3, region 4	90.02	80.00	81.99	75.21	91.13	82.50	83.10	77.77

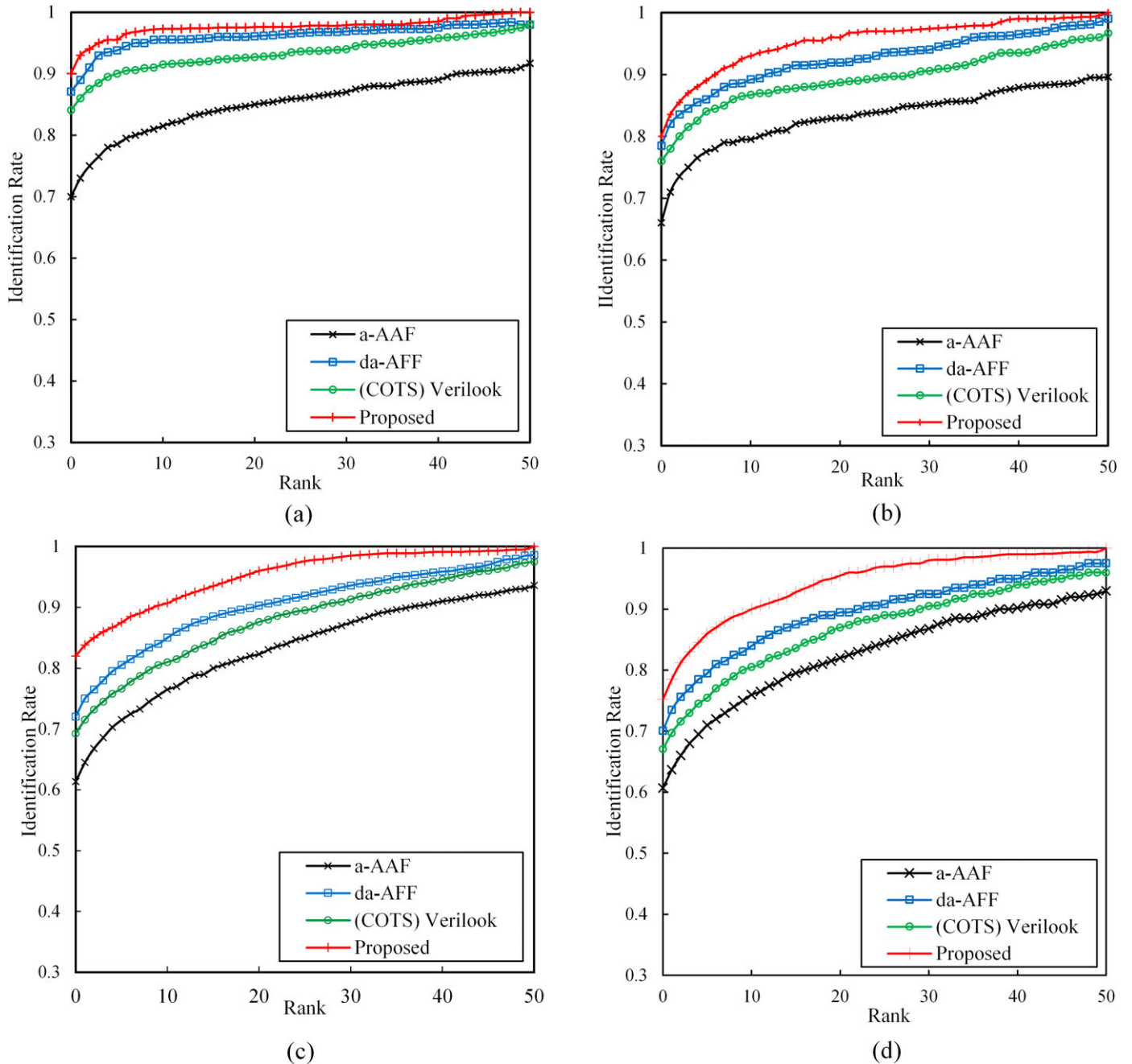


FIG. 18—CMC curves showing face identification performance for (a) MORPH II small, (b) MORPH II large, (c) FERET small, and (d) FERET large temporal variations.

Conclusions

In this paper, we presented two important facial aspects: age group estimation and age-assisted face recognition. We also analyzed facial asymmetry as an intrinsic facial feature to estimate the age group of a given face image. The role of different asymmetric facial regions in estimating age groups has been explored. Deeply learned asymmetric facial features were used to recognize face images while incorporating the knowledge learned from age group estimation. In conclusion, first, it was observed that facial asymmetry is an age-dependent feature. Second, it is deduced that asymmetric facial dimensions can be effectively used for age group estimation. Third, we observe that deeply

learned asymmetric facial features have more discriminative power compared to the handcrafted features employed in some existing methods. Finally, the experimental results on two large facial aging datasets, MORPH and FERET, show that inserting the knowledge learned from age group estimation into the face recognition algorithm can enhance the face recognition accuracies considerably. Age group estimation and face recognition accuracies clearly show the efficacy of the proposed approach compared to existing state-of-the-art methods.

In the future, we plan to extend this work in the following directions:

- Facial asymmetry-based demographic estimation, including gender and ethnic information.

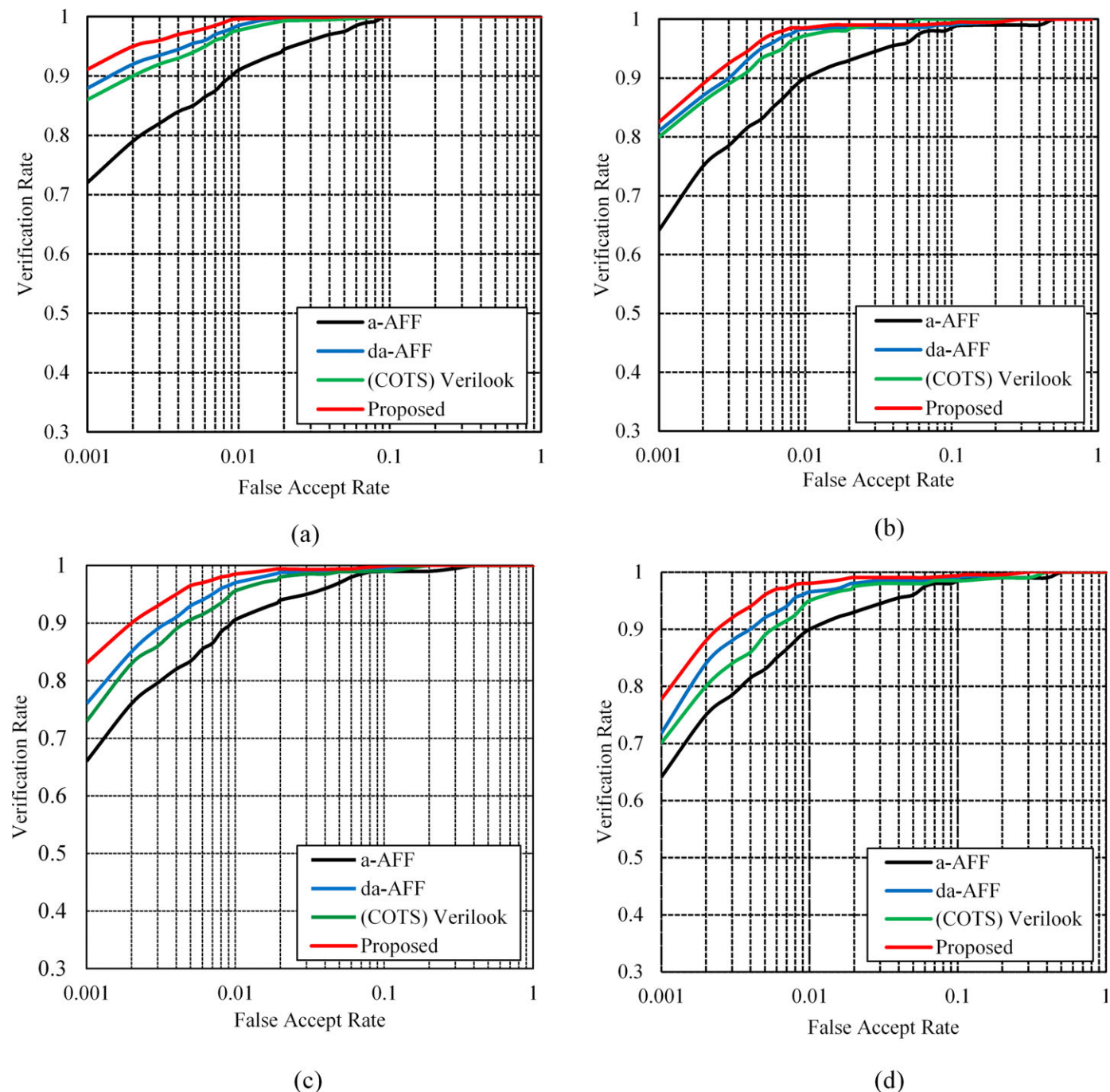


FIG. 19—ROC curves showing face verification performance for (a) MORPH II small, (b) MORPH II large, (c) FERET small, and (d) FERET large temporal variations.

- Integration of demographic information into face recognition algorithms to boost the performance.
- Building facial aging models that incorporate facial asymmetry.

Acknowledgments

Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office. The authors would also like to acknowledge the University of North Carolina Wilmington (UNCW) for providing MORPH database used in current research.

References

1. Fu Y, Huang TS. Human age estimation with regression on discriminative aging manifold. *IEEE Trans Multimedia* 2006;10(4):578–84.
2. Yang Z, Ai H. Demographic classification with local binary patterns. *Proceedings of the International Conference on Biometrics: Advances in Biometrics*; 2007 Aug 27–29; Seoul, South Korea. Berlin, Heidelberg: Springer;2007, 4642:464–73.
3. Ylioinas J, Hadid A, Pietikainen M. Age classification in unconstrained conditions using LBP variants. *Proceedings of the 21st International Conference on Pattern Recognition*; 2012 Nov 11–15; Tsukuba, Japan. Piscataway, NJ: IEEE, 2012.
4. Lu J, Tan Y. Ordinary preserving manifold analysis for human age and head pose estimation. *IEEE Trans Hum Mach Syst* 2013;43(2): 249–58.

5. Hu H, Jain AK. Age, gender and race estimation from unconstrained face images. East Lansing, MI: Michigan State University, 2014; Report No.: MSU-CSE-14-5.
6. Han H, Otto C, Liu X, Jain AK. Demographic estimation from face images: human vs. machine performance. *IEEE Trans Pattern Anal Mach Intell* 2015;37(6):1148–61.
7. Guo G, Mu G. Joint estimation of age, gender and ethnicity: CCA vs. PLS. *Proceedings of the 10th International Conference and Workshops on Automatic Face and Gesture Recognition*; 2013 Apr 22–26; Shanghai, China. Piscataway, NJ: IEEE, 2013:1–6.
8. Chen C, Dantcheva A, Ross A. Impact of facial cosmetics on automatic gender and age estimation algorithms. *Proceedings of the 9th International Conference on Computer Vision Theory and Applications*; 2014 Jan 5–8; Lisbon, Portugal. Piscataway, NJ: IEEE, 2014.
9. Ngan ML, Grother PJ. Face Recognition Vendor Test (FRT) – performance of automated age estimation algorithms, 2014; <https://www.nist.gov/publications/face-recognition-vendor-test-frvt-performance-automated-age-estimation-algorithms> (accessed March 8, 2018).
10. Rhodes MG. Age estimation of faces: a review. *Appl Cogn Psychol* 2009;23(1):1–12.
11. Burt D, Perrett D. Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information. *Proc Biol Sci* 1995;259(1355):137–43.
12. George PA, Hole GJ. The influence of feature-based information in the age processing of unfamiliar faces. *Perception* 1998;27(3):295–312.
13. Jones G, Smith PK. The eyes have it: young children's discrimination of age in masked and unmasked facial photographs. *J Exp Child Psychol* 1984;38(2):328–37.
14. Fu Y, Guo G, Huang TS. Age synthesis and estimation via faces: a survey. *IEEE Trans Pattern Anal Mach Intell* 2010;32(11):1955–76.
15. Kwon YH, Lobo NV. Age classification from facial images. *Comput Vis Image Underst* 1999;71(1):1–21.
16. Ling H, Soatto S, Ramanathan N, Jacobs D. Face verification across age progression using discriminative methods. *IEEE Trans Inf Forensics Sec* 2010;5(1):82–91.
17. Park U, Tong Y, Jain AK. Age-invariant face recognition. *IEEE Trans Pattern Anal Mach Intell* 2010;32(5):947–54.
18. Li Z, Park U, Jain AK. A discriminative model for age-invariant face recognition. *IEEE Trans Inf Forensics Sec* 2011;6(3):1028–37.
19. Yadav D, Vatsa M, Singh R, Tistarelli M. Bacteria foraging fusion for face recognition across age progression. *Proceedings of the International Conference on Computer Vision and Pattern Recognition Workshops*; 2013 June 23–28; Portland, OR. Piscataway, NJ: IEEE, 2013.
20. Sungatullina D, Lu J, Wang G, Moulin P. Multiview discriminative learning for age-invariant face recognition. *Proceedings of the 10th International Conference and Workshops on Automatic Face and Gesture Recognition*; 2013 Apr 22–26; Shanghai, China. Piscataway, NJ: IEEE, 2013:1–6.
21. Best-Rowden L, Jain AK. Longitudinal study of automatic face recognition. *IEEE Trans Pattern Anal Mach Intell* 2017;40(1):148–62.
22. Klare B, Jain AK. On a taxonomy of facial features. *Proceedings of the Fourth International Conference on Biometrics: Theory Applications and Systems*; 2010 Sept 27–29; Washington, DC. Piscataway, NJ: IEEE, 2010.
23. Liu Y, Schmidt K, Cohn J, Mitra S. Facial asymmetry quantification for expression-invariant human identification. *Comput Vis Image Underst* 2003;91(1–2):138–59.
24. Gutta S, Wechsler H. Face recognition using asymmetric faces. *Proceedings of the First International Conference on Biometric Authentication*; 2004 July 15–17; Hong Kong, China. Berlin, Heidelberg: Springer, 2004;3072:162–8.
25. Abaza A, Ross A. Towards understanding the symmetry of human ears: a biometric perspective. *Proceedings of the Fourth International Conference on Biometrics: Theory Applications and Systems*; 2010 Sept 27–29; Washington, DC. Piscataway, NJ: IEEE, 2010.
26. Ricanek KJr, Tesafaye T. MORPH: a longitudinal image database of normal adult age-progression. *Proceedings of the Seventh International Conference on Automatic Face and Gesture Recognition*; 2006 Apr 10–12; Southampton, U.K. Piscataway, NJ: IEEE, 2006;341–5.
27. Color FERET Database; <https://www.nist.gov/itl/iad/image-group/color-feret-database> (accessed March 8, 2018).
28. Huang GB, Ramesh M, Berg T, Learned-Miller E. Labeled faces in the wild: a database for studying face recognition in unconstrained environments. University of Massachusetts, 2007; Report; 07-49.
29. Ercan I, Ozdemir ST, Etoz A, Sigirli D, Tubbs RS, Loukas M, et al. Facial asymmetry in young healthy subjects evaluated by statistical shape analysis. *J Anat* 2008;213:663–9.
30. Cheong YW, Lo LJ. Facial asymmetry: etiology, evaluation and management. *Chang Gung Med J* 2011;34(4):341–51.
31. Face++; <https://www.faceplusplus.com> (accessed March 8, 2018).
32. Vig PS, Hewitt AB. Asymmetry of the human facial skeleton. *Angle Orthod* 1975;45:125–9.
33. Snedecor GW, Cochran WG. *Statistical methods*, 8th edn. Ames, IA: Iowa State University Press, 1989.
34. Hotelling H. Relations between two sets of variates. *Biometrika* 1936;28:321–77.
35. Hotelling H. The most predictable criterion. *J Educ Psychol* 1935;26:139–42.
36. Pudil P, Novovicova J, Kittler J. Floating search methods in feature selection. *Pattern Recognit Lett* 1994;15(11):1119–25.
37. Hall MA. Correlation-based feature subset selection for machine learning. [dissertation]. Hamilton, New Zealand: University of Waikato, 1999.
38. Vapnik VN. *Statistical learning theory*. New York, NY: John Wiley & Sons Inc, 1998.
39. Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2011;2:1–27.
40. Choi SE, Lee YJ, Lee SJ, Park KR, Kim J. Age estimation using a hierarchical classifier based on global and local facial features. *Pattern Recognit* 2011;44(6):1262–81.
41. Thukral P, Mitra K, Chellappa R. A hierarchical approach for human age estimation. *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*; 2012 Mar 25–30; Kyoto, Japan. Piscataway, NJ: IEEE, 2012;1529–32.
42. Altman DG, Bland JM. Diagnostic tests. 1: sensitivity and specificity. *BMJ* 1994;308(6943):1552.
43. Schaefer S, McPhail T, Warren J. Image deformation using moving least squares. *ACM Trans Graph* 2006;25:533–40.
44. Pesaresi M, Benidiktsson JA. A new approach for the morphological segmentation of high resolution satellite imagery. *IEEE Trans Geosci Remote Sens* 2001;39(2):309–20.
45. Hu MK. Visual pattern verification by moment invariants. *IEEE Trans Inf Theory* 1962;8:179–87.
46. Sivic J, Zisserman A. Video Google: efficient visual search of videos. In: Ponce J, Hebert M, Schmid C, Zisserman A, editors. *Toward category-level object recognition: lecture notes in computer science*. Berlin, Heidelberg: Springer, 2006;4170:127–44.
47. Perronnin F, Liu Y, Sanchez J, Poirier H. Large-scale image retrieval with compressed Fisher vectors. *Proceedings of the International Conference on Computer Vision and Pattern Recognition*; 2010 June 13–18; San Francisco, CA. Piscataway, NJ: IEEE, 2010;3384–91.
48. Jegou H, Perronnin F, Douze M, Sanchez J, Perez P, Schmid C. Aggregating local image descriptors into compact codes. *IEEE Trans Pattern Anal Mach Intell* 2012;34(9):1704–16.
49. Babenko A, Lempitsky VS. *Aggregating deep convolutional features for image retrieval*. Ithaca, NY: Cornell University Library, 2015.
50. Razavian AS, Sullivan J, Maki A, Carlsson S. *Visual instance retrieval with deep convolutional networks*. Ithaca, NY: Cornell University Library, 2014.
51. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems*; 2012 Dec 03–06; Lake Tahoe, NV. New York, NY: ACM, 2012;1097–105.
52. VeriLook SDK: face identification for stand-alone or web applications; <http://www.neurotechnology.com/verilook.html> (accessed March 8, 2018).
53. Ross A, Jain AK. Information fusion in biometrics. *Pattern Recognit Lett* 2003;24(13):2115–25.
54. Kittler J. Combining classifiers: a theoretical framework. *Pattern Anal Appl* 1998;1(1):18–27.

Additional information and reprint requests:
 Usama Ijaz Bajwa, Ph.D.
 Department of Computer Science
 COMSATS Institute of Information Technology
 Off Raiwind Road
 Lahore
 Pakistan
 E-mail: usamabajwa@ciitlahore.edu.pk; usama@usamaijaz.com