



Age-invariant face recognition based on deep features analysis

Amal A. Moustafa¹ · Ahmed Elnakib¹ · Nihal F. F. Areed¹

Received: 23 July 2019 / Revised: 14 November 2019 / Accepted: 5 January 2020 / Published online: 13 January 2020
© Springer-Verlag London Ltd., part of Springer Nature 2020

Abstract

Age-invariant face recognition is one of the most crucial computer vision problems, e.g., in passport verification, surveillance systems, and missing individuals identification. The extraction of robust face features is a challenge since the facial characteristics change over age progression. In this paper, an age-invariant face recognition system is proposed, which includes four stages: preprocessing, feature extraction, feature fusion, and classification. Preprocessing stage detects faces using Viola–Jones algorithm and frontal face alignment. Feature extraction is achieved using a CNN architecture using VGG-Face model to extract compact face features. Extracted features are fused using the real-time feature-level multi-discriminant correlation analysis, which significantly reduces feature dimensions and results in the most relevant features to age-invariant face recognition. Finally, K -nearest neighbor and support vector machine are investigated for classification. Our experiments are performed on two standard face-aging datasets, namely FGNET and MORPH. Rank-1 recognition accuracy of the proposed system is 81.5% on FGNET and 96.5% on MORPH. Experimental results outperform the current state-of-the-art techniques on same data. These preliminary results show the promise of the proposed system for personal identification despite aging process.

Keywords Face recognition · Aging · Convolutional neural networks (CNN) · Transfer learning · Feature fusion · Multi-discriminant correlation analysis (MDCA)

1 Introduction

Age-invariant face recognition (AIFR) plays a vital role in several applications, such as biometric security systems [5], forensic applications [10], and identification of missing individuals [11]. Facial characteristics largely change over time with different age pattern per person. This refers personal genes, individual health, lifestyle, etc. Extracting robust features describing aging facial details is challenging research problem, especially when large age gaps between face images are considered [24]. In this work, two well-known standard databases are investigated for the job of AIFR, namely

MORPH [26] and FGNET [6] databases. The challenges in these data can be summarized as follows:

- Each subject has a variable low number of images of different ages. This makes the learning process very hard due to the limited number of training images per person and the fact that each person has his/her own age pattern based on genes, health, etc.
- Pose, illumination, blurring, and distance from the camera are different among images.
- The part of the data which is between ages 0–2 years, where strong personalized face features were still not established yet. At this age, it is very hard to associate the face of this age (0–2) with the correct same person's elder faces even with human eyes.
- Existence of large age gaps, especially for FGNET (ages are between 0 and 69 years).

✉ Ahmed Elnakib
nakib@mans.edu.eg

Amal A. Moustafa
amal.soltan55@gmail.com

Nihal F. F. Areed
nfayez@zewailcity.edu.eg

¹ Electronics and Communication Engineering Department, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt

Recently, convolutional neural networks (CNN) architectures have shown a remarkable success to address the challenge of recognizing faces in spite of their age [1,5,15,25,39]. In this paper, analysis of CNN extracted features is investi-

gated to optimize a set of features that precisely describe the faces invariantly of age. To achieve this goal, the proposed recognition system used the pre-trained VGG-Face deep-learning CNN architecture to extract face features. To improve the recognition rate (RR), the proposed system performs the following contributions.

- Analyzing the potential of the extracted features from VGG-Face model layers that were unexplored in the literature (i.e., flatten layer) for AIFR.
- Discriminating related features to AIFR using MDCA, which significantly reduces the feature space dimensions and improves the recognition rate.
- Performing space reduction of the extracted features, which in turn reduces the overall system time.

The proposed MDCA fusion system has achieved a RR of 81.5% on FGNET and 96.5% on MORPH. The rest of this paper is organized as follows. Section 2 summarizes the related work for AIFR. Section 3 describes the proposed system. Section 4 presents the results and discussions. Finally, Sect. 5 concludes the paper.

2 Related work

AIFR techniques are classified into generative, discriminative, and deep-learning approaches, as follows:

2.1 Generative approaches

Generative approaches attempt to simulate age process by constructing a synthesized face image (a pseudo-photograph) from individual's aging images that are acquired before performing face recognition. For example, Lanitis et al. [13] built a 3D simulated aging model based on shape and intensity features, achieving a RR of 68.5% on a privately collected database. Park et al. [24] added a pose correction stage to a simulated 3D aging model of shape and texture, achieving RR of 37.4% and 79.8% on FGNET and MORPH databases, respectively. Although generative approaches could get simulated-age models, these models are limited by strong unrealistic parametric assumptions [7].

2.2 Discriminative approaches

Discriminative approaches extract discriminative facial features that mainly describe the aging process. For example, Ling et al. [19] used a gradient orientation pyramid (GOP) to describe the aging process and a support vector machine (SVM) was used for classification. Li et al. [17] used scale-invariant feature transform (SIFT) [18] and multi-multiple local binary pattern (MLBP) [23] as discriminative features

to achieve age-invariant recognition. Gong et al [8] applied a maximum entropy feature descriptor (MEFD) that encoded the microstructure of facial images into a set of discrete codes in terms of maximum entropy. Li et al. [14] analyzed discriminative features using a modified hidden factor analysis (MHFA); they considered the correlation between age and identity, instead of assuming them to be mutually independent. Recently, Zhou et al. [41] used an identity inference model for AIFR based on a probabilistic linear discriminant analysis and expectation–maximization (EM) algorithm.

2.3 Convolutional neural networks (CNN)

To take into account the local correlations in natural images, CNN includes both fully connected hidden layers and locally connected convolutional layers, which use parameter sharing, pooling, and dropout to greatly reduce the number of parameters (“features”) learned by the CNN [5]. Recently, researches have shown that the output of CNN layers produces highly discriminative descriptors for AIFR. In this regard, Yan et al. [39] used a CNN architecture to extract facial features and a SVM classifier is used to select the belonging age range. Additionally, Li et al. [15] employed a deep CNN model that performed both feature extraction and classification. Xu et al. [37] used a coupled auto-encoders network to extract identity features from face image for AIFR. Li et al. [16] applied an optimization model that learn features and a distance metric simultaneously, for AIFR. Shakeel et al [27] used a pre-trained CNN model to extract face features. The extracted features were further encoded using a learned codebook, followed by linear regression mapping for face matching. Recently, the VGG-Face CNN pre-trained model [25] was widely used in face recognition applications, e.g., in [1,5,25]. This model is also adopted in our work to extract discriminate features of face images.

3 Proposed system

This section details the proposed system for age-invariant face recognition. The system is composed of four modules, namely preprocessing, feature extraction, fusion, and classification, as shown in Fig. 1. The details of each of these stages are illustrated below.

3.1 Image preprocessing

The input of the preprocessing step is the raw facial image (as acquired from its database), and the output is an aligned frontal face image. The objective of this step is to register all images, i.e., on the basis of eye coordinates, such that they are aligned on the same standard size, i.e., each point on any given face is aligned to the same point on all of the images. To

Fig. 1 Proposed framework for age-invariant face recognition composed of four stages: preprocessing to align faces frontally, VGG-deep-learning feature extraction, MDCA fusion, and KNN classification

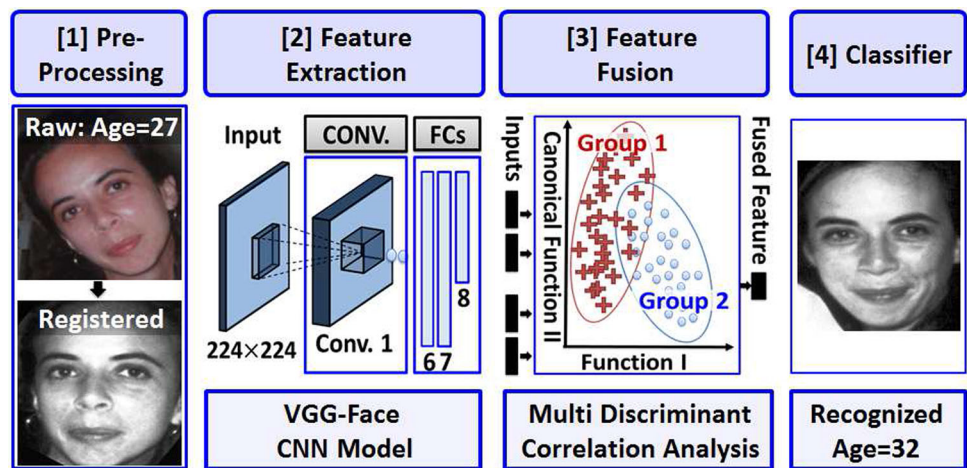


Fig. 2 Block diagram of the steps of the proposed image preprocessing

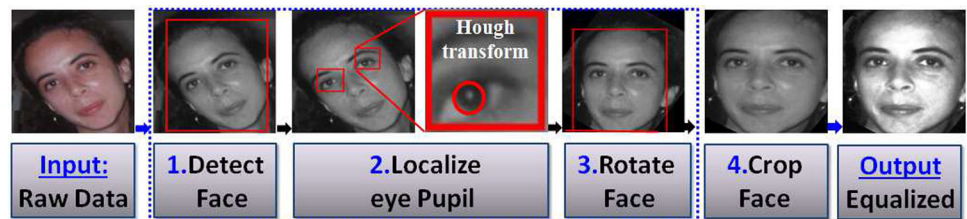


Fig. 3 The top row shows three raw data samples from FGNET and three raw data samples from MORPH datasets, respectively. The bottom row shows the same samples after the proposed preprocessing steps



do this job, face is detected for each raw image using Viola–Jones method [36]. Hough transform [2] is then applied to detect the center of each pupil as clearly discussed in the work of [12,29]. Afterward, a face image is aligned frontally by rotating the face such that the detected centers are on the same horizontal line. Finally, the image is contrast-enhanced, auto-cropped, and resized to the standard 224×224 , which is the input image size for of VGG-Face model. The details of the preprocessing step is shown in Algorithm 1. Figure 2 illustrates the results of each step of the algorithm. Figure 3 shows sample results of the preprocessing step for different samples from FGNET and MORPH raw data.

3.2 Feature extraction

The second stage of the proposed system uses VGG-Face CNN model [25] to extract the face features. This model is an application of the very deep convolutional network VGG-

Algorithm 1 Frontal Face Image Alignment

Input: Raw face image (from the selected album)

Output: Frontally aligned, standard size face image

1. Detect face from input image via Viola–Jones method
2. Detect pupil centers by circular Hough transform
3. Rotate face frontally on the basis of centers’ locations
4. Auto-crop rotated face image, as shown in Fig. 2, based on the distance between the pupils to a standard size 224×224 (size of input image of VGG-Face)
5. Apply histogram equalization to enhance image contrast

16 architecture that is trained specifically on a very large-scale face database (about 2.6 million face images of 2622 different people). The architecture of VGG-Face model is composed of 15 layers (12 convolutional layers and three fully connected layers, named FC6 and FC7). The vector of activities of the fully connected VGG layers is repeatedly used in the literature as a powerful generic image descriptor that can be applicable to other face databases [1,28,33]). In the proposed system, transfer learning is applied to train the

VGG-Face model on new datasets (FGNET or MORPH) for age-invariant face recognition. This involves removing the last output layer of a trained VGG-Face model and using the activations of the last fully connected layers (FC6 and FC7) as the extracted features. Additionally, the proposed system investigates the power of extracting features from the last convolutional layer (regularly called “flatten layer”). Thus, these extracted features (of length 4096 for both FC6 and FC7 and of length 25,088 for flatten layer) are used in the proposed system as the input features.

3.3 Feature fusion

Fusion can be performed either in the base of feature-level or decision-level (multi-classifier outputs). The proposed system uses feature-level fusion since it is expected to provide better recognition results than decision-level as this fusion tries to enrich the information extracted from input image fusion other than the matching score or output decision of a classifier [21,38]. Traditional feature fusion techniques, such as parallel, which is based on creating a complex vector [40], and serial fusion [20], which is based on concatenating two different features directly. However, these methods of fusion may cancel out the discriminative power of the individual input vectors [41]. Recently, for feature fusion canonical correlation analysis (CCA) is repeatedly used in the literature for different pattern recognition problems [30–32], based on maximizing the correlation of corresponding features across the two input feature sets. Discriminant correlation analysis (DCA) [9] is an extension of CCA that takes into account the class labels. The idea behind DCA is to maximize the correlation of corresponding features across the two input feature sets (as done in CCA-methods) in addition to performing a correlation for features that belong to different classes within each feature set. In addition to feature-level fusion, DCA performs significant dimension reduction, which implies low computational complexity. Therefore, the extracted fused features can be employed in real-time applications.

In the proposed system, a cascade version of DCA, called multimodal discriminant correlation analysis (MDCA) [9] is used for fusion process. While DCA works on two sets of input variables at a time, MDCA can fuse any number of input pairs. Figure 4 illustrates how MDCA is employed in the proposed system to combine the extracted features from three VGG layers with different feature length vectors as shown in Fig. 4. Using a cascaded two-level DCA, level one is used to fuse the three VGG layers through two blocks of DCA and level two is used to produce the final fused output vector of reduced length. Note that FC6 is used as a common input for the two DCA blocks in level one since the results in this paper show that it is the most discriminative layer among the three investigated layers.

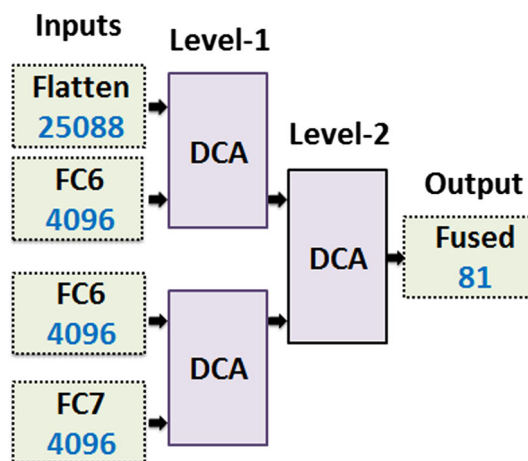


Fig. 4 MDCA fusion using a cascaded two-level DCA : level one is used to fuse the different investigated layers of VGG model illustrating the extracted feature length from each layer and level two is used to produce the final fused output vector and it’s significant reduced length (81 in case of FGNET). A number in inputs and output blocks indicate the number of features per block

3.4 Classification methods

To recognize faces across ages, a K -nearest neighbor (KNN) [4] classifier, based on the Euclidian distance, is applied on the extracted deep features. In our experiments, $K = 1$ is used to select the image of the nearest feature vector to the feature vector of the input test image. KNN classifier is selected to offer a very simple design and significant less computations. In addition to the KNN classifier, a support vector machine (SVM) [34] classifier is investigated to check its ability to improve system accuracy. Binary linear kernels are selected to support multiclass SVM classification.

4 Experimental results

4.1 Datasets

To test the proposed system, two well-known databases for face aging are for the job of age-invariant face recognition, including MORPH [26] FGNET [6] databases (see Table 1). Table 2 describes these databases. As shown in Table 2, FGNET [6] is a relatively small database (1002 images of 82 subjects of different ages that vary between zero and 69 with a large age gap per subject (0–45), making it more challenging. For the MORPH album-II [26], which is available for academic research purposes, contains 55,134 images of 13,000 subjects collected over 4 years and includes age between 16 and 77 years.

Table 1 FGNET and MORPH (album-II) Rank-1 recognition rates (RRs) for different VGG-Face CNN layers (Flatten, FC6, and FC7), their DCA fusion, their serial fusion, and the proposed MDCA fusion using LOPO scheme and KNN classifier

	Flatten	FC6	FC7	FC6 \diamond FC7	Flatten \diamond FC6	Flatten \diamond FC7	Serial fusion	Proposed fusion
FGNET RR	68.2%	74.9%	73.0%	79.3%	79.8%	79.0%	74.8%	80.4%
MORPH RR	89.8%	93.6%	91.2%	95.6%	96.2%	96.4%	92.3%	96.5%

\diamond Denotes DCA fusion

Table 2 Description of FGNET and MORPH databases

Data	# Subjects	# of Images	Age range	Age gap	#Images subject
FGNET	82	1002	0–69	0–45	5–13
MORPH	13,617	55,134	16–77	0–5	2–5

4.2 Experimental setup

Rank-1 recognition rate (RR) is used to evaluate the performance of the proposed system. A scheme of leave one image of a person out (LOPO) is used for testing. LOPO is widely used to test recognition accuracy on FGNET and MORPH datasets as in [22], as LOPO scheme identifies a person face from his previous face images. This is repeatedly applied in security applications (i.e., forensic history to identify wanted criminals and missing individuals, based on their previous images at different ages).

The reason behind using LOPO is that it provides robust evaluation especially in case of small data size [3], which is the case for FGNET dataset as it contains 1002 images of 82 different subjects which are all used for our experiments. In MORPH (album-II) dataset, each person has limited face images, a maximum of five per person, a random subset of 1100 images from MORPH corresponding to 372 subjects are used in our experiments. (This number of samples is chosen randomly to nearly equal FGNET samples in consideration to the computation.)

4.3 Experiments on FGNET dataset

The proposed system in Fig. 1 is applied on FGNET data. In the first stage of the proposed system, preprocessing of the raw images is performed to produce frontally aligned images with the standard VGG model size. In the second stage of the system, features acquired of the VGG-Face model from different layers (Flatten, FC6, and FC7) are extracted. To evaluate the efficiency of the VGG-Face model, Rank-1 RRs of using each individual layer features (Flatten, FC6, and FC7) are evaluated using LOPO scheme and KNN classifier, with $K = 1$, on the whole FGNET dataset. As shown in Table 1, recognition rates achieve scores of 68.2%, 74.9% and 74.75% using each of VGG model layer features: Flatten, FC6, and FC7, respectively. These results confirm the capability of these features to achieve good initial results; however, further improvement is required. To enhance the

Table 3 Rank-1 recognition results of proposed system using KNN and SVM classifiers

Classifier	FGNET RR%	Morph RR%
KNN classifier	80.4	96.5
SVM classifier	81.5	94.3

RR, the third stage of the proposed system attempts to fuse between the Flatten, FC6, and FC7 features. As shown in Table 1, applying a naive serial fusion, by concatenating the features, is not able to improve the RR, because it cancels out their discriminating power.

The proposed system uses a smart fusion between the Flatten, FC6, and FC7 features using the MDCA feature fusion. MDCA significantly reduces the feature dimension space to output the most powerful features for age-invariant face recognition, i.e., only 81 discriminant features in FGNET case. This in turn reduces the computational cost of the proposed system. The significant reduction is around 99.76% of the input space (the input space is of size 25088 (Flatten) + 4096 (FC6) + 4096 (FC7)), see Table 4. In addition, the visual results in Fig. 5 show that the proposed MDCA is able to achieve the correct matching in many cases, even for cases where none of the individual layers achieves the correct matching (as in the last row sample data in Fig. 5). In additional, quantitative results confirmed the privileges of the proposed MDCA fusion, which achieves an improved recognition rate of 80.4% using KNN.

To investigate the potential of more sophisticated classifiers, SVM is applied on the fused data; the results in Table 3 show that SVM successfully improves RR for FGNET dataset from 80.4% using KNN to 81.5%

4.4 Experiments on MORPH dataset

Same experiments on FGNET are repeated using a random subset of 1100 images from MORPH corresponding to 372 subjects with same experimental setting. Figure 5 and Table 1

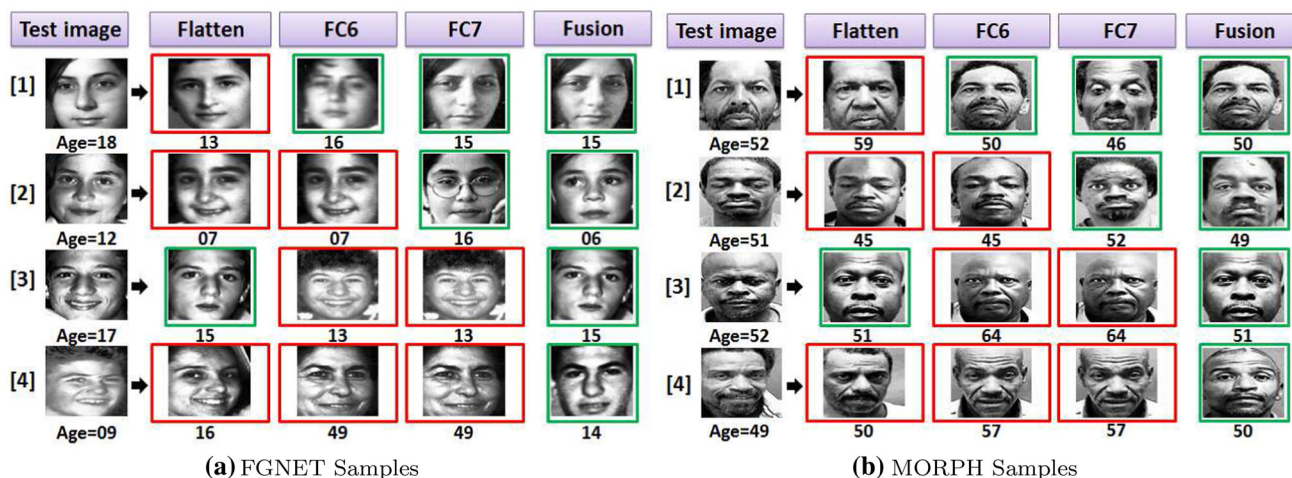


Fig. 5 Visual system recognition for **a** four samples from FGNET and **b** four samples from MORPH (album-II) datasets using LOPO scheme and KNN classifier. For each dataset, the first column is the test image. Columns 2,3,4 represent visual results of using individual Flatten, FC6,

and FC7 layers, respectively. The last column represents the results of the proposed MDCA fusion system. A large red rectangular frame indicates failed recognition, whereas a green small rectangular frame indicates correct recognition

summarize the visual and quantitative results of the experiments. As in FGNET, the results show the impact of the proposed MDCA fusion in achieving better RR. In addition, since the data are collected over only 5 years (a maximum of 5 years age gap between the images) with ages ranging from 16 to 77 years, the recognition rate has reached 96.5% using the proposed system, in comparison with FGNET, where age gap between images can reach 45 years, with age ranging from 0–69 years, making it more challenging.

4.5 Time performance

All experiments in the paper are implemented using MATLAB 2017a (64-bit) with 2.5 GHz Intel CORE i5 2450M CPU and 8 GB of RAM. The library MatConvNet [35] is used for GPU parallelization. To evaluate time performance, the processing time has been evaluated for a number of 30 different trials and the mean and standard deviation are calculated, and noted as mean ± SD.

The proposed MDCA fusion successfully does the following :

- reducing the feature space size to only 81 FGNET features, as illustrated in Table 4

- enhancing the estimated speed of the proposed system up to 7.7× over using all the features of the individual FC6 or FC7 layers and 52× over using all concatenated serial fusion features

The elapsed times for the proposed system stages (preprocessing, feature extraction, feature fusion, and classification) are 0.784 s, 0.792 s, 0.005 s, and 0.004 s, respectively. Therefore, the overall time of the complete proposed framework is 1.58 sec.

4.6 Comparing with state-of-the-art algorithms

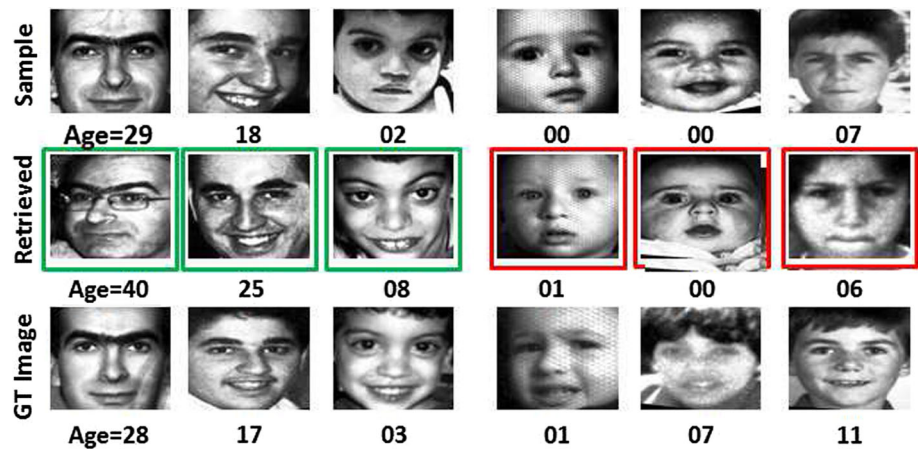
To verify the quality of the proposed system for age-invariant face recognition, a comparison with other age-invariant face recognition algorithms is summarized in Table 5. Note that some of these methods cannot be compared, since they use different number of samples per databases or different experimental setting. As shown in Table 5, the proposed MDCA fusion step in the system has improved the recognition accuracy over other compared systems.

Table 4 Time performance evaluation of the KNN classification step elapsed for an FGNET test sample

Method	Proposed system	Flatten	FC6 or FC7	Serial fusion
Input feature space size	81	25,088	4096	250,688 + 2 × 4096
Mean ± SD	0.003 ± 0.001 s	0.139 ± 0.0035 s	0.0235 ± 0.0017 s	0.174 ± 0.0027 s

Table 5 Comparative Rank-1 recognition results of state-of-the-arts for age-invariant face recognition

Methods	FGNET RR%	Morph RR%
3D aging model (2010) [24]	37.4	79.8
Discriminative aging model (2011) [17]	47.5	83.9
Hidden factor analysis model (2013) [7]	69	86.8
Maximum entropy model (2015) [8]	76.2	92.3
Face recognition across time lapse using CNN (2016) [5]	80.6	92.2
Modified hidden factor analysis model (2017) [14]	76.8	88.0
AIFR-CNN (2018) [22]	76.5	92.5
Distance metric optimization driven CNN (2018) [16]	NA	93.6
Proposed system	81.5	96.5

Fig. 6 Proposed system retrieved images for six samples from FGNET dataset using our proposed framework. The first row is the test image. The second row shows retrieved images using our system. The last row is the ground truth (GT), determined as the recognized person image of the nearest age. A green rectangular frame indicates correct recognition, and a red rectangular frame indicates failed recognition

4.7 Discussion and limitations

Table 1 indicates that FC6 features give the highest Rank-1 RR for the all investigated layers' features (Flatten, FC6, and FC7). Moreover, combining the features from FC6 with each of Flatten and FC7 using MDCA fusion has shown a significant improvement in the Rank-1 RR even with a simple KNN classifier, with $K = 1$, in addition to the significant dimension reduction of the input feature space. Note that FC6 features are used as a common input for each input pair of the MDCA fusion since it achieves the highest RR.

The difference in the Rank-1 RR between FGNET dataset (large age gap) and MORPH is due to that most of the FGNET images appear in the childhood and teenage stages, as face features appear to be unstable as shown in the failed retrievals examples in Fig. 6. But on the contrary, most of the MORPH images (the maximum age gap between images of each person is 4 years) lie in the elder age groups as face features tend to be more stable; this reflects the higher RR that MORPH dataset achieves. In the future, more features will be investigated to improve the accuracy of the system, specially if large age gap is considered.

5 Conclusion

In this paper, a system is proposed for age-invariant face recognition. The proposed system applies an image preprocessing step, followed by features extraction using the utility of a pre-trained VGG-Face CNN to get highly discriminative descriptors (rather than hand-crafted). Then, a dimension reduction using an efficient MDCA fusion has significantly reduced the input feature space and improved the recognition rate of the system. The proposed system achieves 81.5% accuracy for challenging FGNET (with large age gap (0–45) between face images) dataset and 96.5% for MORPH (album-II) dataset. In addition, it outperforms the state-of-the-art methods on the same data. These results show the promise of the proposed system in age-invariant face recognition.

References

1. Anand, A., Donida Labati, R., Genovese, A., Munoz Ballester, E., Piuri, V., Scotti, F.: Age estimation based on face images and pre-trained convolutional neural networks. In: CISDA. IEEE (2017)
2. Ballard, D.H.: Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognit.* **13**(2), 111–122 (1981)

3. Cawley, G.C.: Leave-one-out cross-validation based model selection criteria for weighted ls-svms. In: International Joint Conference on Neural Networks, 2006. IJCNN'06, pp. 1661–1668. IEEE (2006)
4. Cover, T., Hart, P.: Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **13**(1), 21–27 (1967)
5. El Khyari, H., Wechsler, H.: Face recognition across time lapse using convolutional neural networks. *J. Inf. Secur.* **7**(03), 1028–1037 (2016)
6. Face, group, G.R.W., et al.: Fg-Net Aging Database (2000)
7. Gong, D., Li, Z., Lin, D., Liu, J., Tang, X.: Hidden factor analysis for age invariant face recognition. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2872–2879 (2013)
8. Gong, D., Li, Z., Tao, D., Liu, J., Li, X.: A maximum entropy feature descriptor for age invariant face recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5289–5297 (2015)
9. Haghighat, M., Abdel-Mottaleb, M., Alhalabi, W.: Discriminant correlation analysis: real-time feature level fusion for multimodal biometric recognition. *IEEE Trans. Inf. Forensics Secur.* **11**(9), 1984–1996 (2016)
10. Jain, A.K., Klare, B., Park, U.: Face matching and retrieval in forensics applications. *IEEE Multimed.* **19**(1), 20 (2012)
11. Jain, A.K., Li, S.Z.: Handbook of Face Recognition. Springer, New York (2011)
12. Jan, F., Usman, I., Agha, S.: Reliable iris localization using hough transform, histogram-bisection, and eccentricity. *Signal Process.* **93**(1), 230–241 (2013)
13. Lanitis, A., Taylor, C.J., Cootes, T.F.: Toward automatic simulation of aging effects on face images. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(4), 442–455 (2002)
14. Li, H., Zou, H., Hu, H.: Modified hidden factor analysis for cross-age face recognition. *IEEE Signal Process. Lett.* **24**(4), 465–469 (2017)
15. Li, Y., Wang, G., Lin, L., Chang, H.: A deep joint learning approach for age invariant face verification. In: CCF Chinese Conference on Computer Vision, pp. 296–305. Springer, New York, (2015)
16. Li, Y., Wang, G., Nie, L., Wang, Q., Tan, W.: Distance metric optimization driven convolutional neural network for age invariant face recognition. *Pattern Recognit.* **75**, 51–62 (2018)
17. Li, Z., Park, U., Jain, A.K.: A discriminative model for age invariant face recognition. *IEEE Trans. Inf. Forensics Secur.* **6**(3), 1028–1037 (2011)
18. Lindeberg, T.: Scale Invariant Feature Transform (2012)
19. Ling, H., Soatto, S., Ramanathan, N., Jacobs, D.W.: Face verification across age progression using discriminative methods. *IEEE Trans. Inf. Forensics Secur.* **5**(1), 82–91 (2010)
20. Liu, C., Wechsler, H.: A shape-and texture-based enhanced fisher classifier for face recognition. *IEEE Trans. Image Process.* **10**(4), 598–608 (2001)
21. Monwar, M.M., Gavrilova, M.L.: Multimodal biometric system using rank-level fusion approach. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **39**(4), 867–878 (2009)
22. Nimbarte, M., Bhojar, K.: Age invariant face recognition using convolutional neural network. *Int. J. Electr. Comput. Eng. (IJECE)* **8**(4), 2126–2138 (2018)
23. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
24. Park, U., Tong, Y., Jain, A.K.: Age-invariant face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(5), 947–954 (2010)
25. Parkhi, O.M., Vedaldi, A., Zisserman, A., et al.: Deep face recognition. In: BMVC , vol. 1, p. 6 (2015)
26. Ricanek, K., Tesafaye, T.: Morph: A longitudinal image database of normal adult age-progression. In: 7th International Conference on Automatic Face and Gesture Recognition, FGR 2006, pp. 341–345. IEEE (2006)
27. Shakeel, M.S., Lam, K.M.: Deep-feature encoding-based discriminative model for age-invariant face recognition. *Pattern Recognit.* **93**, 442–457 (2019)
28. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. Published as a Conference Paper at ICLR 2015 (2014)
29. Song, F., Tan, X., Chen, S., Zhou, Z.H.: A literature survey on robust and efficient eye localization in real-life scenarios. *Pattern Recognit.* **46**(12), 3157–3173 (2013)
30. Sun, Q.S., Heng, P.A., Jin, Z., Xia, D.S.: Face recognition based on generalized canonical correlation analysis. In: International Conference on Intelligent Computing, pp. 958–967. Springer, New York (2005)
31. Sun, Q.S., Liu, Z., Heng, P.A., Xia, D.S.: A theorem on the generalized canonical projective vectors. *Pattern Recognit.* **38**(3), 449–452 (2005)
32. Sun, Q.S., Zeng, S.G., Liu, Y., Heng, P.A., Xia, D.S.: A new method of feature fusion and its application in image recognition. *Pattern Recognit.* **38**(12), 2437–2448 (2005)
33. Sun, Y., Liang, D., Wang, X., Tang, X.: Deepid3: Face Recognition with Very Deep Neural Networks. arXiv preprint [arXiv:1502.00873](https://arxiv.org/abs/1502.00873) (2015)
34. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, New York (2013)
35. Vedaldi, A., Lenc, K.: Matconvnet: Convolutional neural networks for matlab. In: Proceedings of the 23rd ACM International Conference on Multimedia, pp. 689–692. (2015)
36. Viola, P., Jones, M.J.: Robust real-time face detection. *Int. J. Comput. Vis.* **57**(2), 137–154 (2004)
37. Xu, C., Liu, Q., Ye, M.: Age invariant face recognition and retrieval by coupled auto-encoder networks. *Neurocomputing* **222**, 62–71 (2017)
38. Xu, X., Mu, Z.: Feature fusion method based on KCCA for ear and profile face based multimodal recognition. In: 2007 IEEE International Conference on Automation and Logistics, pp. 620–623. IEEE (2007)
39. Yan, C., Lang, C., Wang, T., Du, X., Zhang, C.: Age estimation based on convolutional neural network. In: Pacific Rim Conference on Multimedia, pp. 211–220 (2014)
40. Yang, J., Yang, Jy, Zhang, D., Lu, Jf: Feature fusion: parallel strategy vs. serial strategy. *Pattern Recognit.* **36**(6), 1369–1381 (2003)
41. Zhou, H., Lam, K.M.: Age-invariant face recognition based on identity inference from appearance age. *Pattern Recognit.* **76**, 191–202 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.