❒   1833

Optimization of deep learning features for age-invariant face recognition

Amal A. Moustafa, Ahmed Elnakib, Nihal F. F. Areed
Electronics and Communications Engineering Department, Faculty of Engineering,
Mansoura University, Mansoura, Egypt

Article Info	ABSTRACT
	This paper presents a methodology for Age-Invariant Face Recognition (AIFR), based on the optimization of deep learning features. The proposed method extracts deep learning features using transfer deep learning, extracted from the unprocessed face images. To optimize the extracted features, a Genetic Algorithm (GA) procedure is designed in order to select the most relevant features to the problem of identifying a person based on his/her facial images over different ages. For classification, K-Nearest Neighbor (KNN) classifiers with different distance metrics are investigated, i.e., Correlation, Euclidian, Cosine, and Manhattan distance metrics. Experimental results using a Manhattan distance KNN classifier achieve the best Rank-1 recognition rates of 86.2% and 96% on the standard FGNET and MORPH datasets, respectively. Compared to the state-of-the-art methods, our proposed method needs no preprocessing stages. In addition, the experiments show its privilege over other related methods.

Corresponding Author:

Ahmed Elnakib,
Electronics and Communications Engineering Department,
Faculty of Engineering, Mansoura University,
El-Gomhoria ST, Mansoura 35516, Dakahlia, Egypt.
Email: nakib@mans.edu.eg

1. INTRODUCTION

Identification of a person based on his face images over different ages is mandatory for security and forensic applications, e.g., identification of criminals and missing persons. However, this problem is very challenging due to the rapid change in face images with age, especially when images from different age-stages are considered, e.g., newborn, toddler, teenage, and adult face images. In addition, the facial changes are very specific for each person based on his genes, lifestyle. For these reasons, developing a reliable, accurate method for AIFR is necessary

In the literature, many research groups have presented different methods to identify person based on facial age images. These methods can be categorized into three categories: (i) generative methods, (ii) discriminative methods, and (iii) deep learning methods. Generative methods construct a personal face model based on the collected facial-age images. For example, shape and intensity features are used by Lanitis et al. [1] to build a 3D age model, which achieves 68.5% recognition accuracy on a private aging data. Pose correction is added, by Park at. al [2], to a 3D shape- and texture-based face model to achieve an accuracy of 37.4% and 79.8% on the two popular online standard databases for AIFR, i.e., FGNET [3] and MORPH-II [4] databases, respectively. The main drawback of these methods is the need for unrealistic parametric assumptions [5].

On the other hand, discriminative methods do not rely on face modeling. They extract direct facial features for classification, such as Gradient Orientation Pyramid (GOP), Scale-Invariant Feature Transform (SIFT) [6], and Multi-Scale Local Binary Pattern (MLBP) [7, 8]. For example, Ling et al. [9] used GOP

features to describe faces. Then, the extracted features were input to a Support Vector Machine (SVM) for personal identification. Li et al. [10] used SIFT and MLBP features as local descriptors for face recognition achieving recognition rates of 47.5% and 83.9% for FGNET and MORPH-II datasets, respectively. Gong et al [11] 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. For matching faces, they further applied an Identity Factor Analysis (IFA) to achieve recognition rates of 76.2% for FGNET dataset and 92.3% for MORPH-II dataset. Li et al. [12] applied a Modified Hidden Factor Analysis (MHFA) to a set of discriminative features, achieving recognition rates of 72.8% and 87.9% on FGNET and MORPH-II datasets, respectively. Although these methods achieved success in face recognition, their generalization to recognize age-invariant faces is still under investigation [5].

To investigate more general features for AIFR, the third category of face recognition methods use deep learning. These methods usually use different Convolutional Neural Network (CNN) architectures to extract compact face features. CNN architectures, composed of convolutional layers, pooling layers, and fully connected layers, had achieved a remarkable success in AIFR. Y. Li et al. [13] used a CNN architecture composed of seven layers that can learn face features, distance between faces, and discriminating thresholds simultaneously to achieve an accuracy of 93.6% on MORPH-II dataset. Parkhi et al. [14], a group in Oxford University, namely, Visual Geometry Group (VGG), have built a deep learning architecture for face recognition named as VGG-Face CNN model. The model has been trained using more than 2.6 million face images achieving high face recognition rate. Later on, this model has been repeatedly invoked by other research teams, using transfer learning, for age estimation [15, 16] and for AIFR [17]. ElKhiyari and Wechsler [17] used this model with an Ensemble of Subspace Discriminant (ESD) classifiers to achieve an accuracy of 80.6% for FGNET dataset and 92.2% for MORPH-II dataset. In this paper, we employed the VGG-face CNN model, with two major contributions:

- The proposed method involves no preprocessing steps, i.e., images are only resized to the standard size of VGG-model input without any form of registration. Registration steps, usually used in the literature [13, 17-19], may not only involve alignment errors [20, 21], but also make the method less robust, time consuming, and sensitive to the success of the registration step.
- The extracted features using the transfer learning VGG-model are further optimized using a sophisticated GA in order to select the most relevant features to AIFR. In addition, GA performs feature selection, which significantly reduces the feature space and improves the recognition speed.

This paper is organized as follows: Section 2 illustrates the details of the proposed methodology for face recognition. Section 3 outlines the results with detailed discussions, and finally, section 4 concludes the paper and the future work.

2.　METHODS

The proposed method for AIFR consists of two main steps, as shown in Figure 1. First, the VGG-face model is used to extract compact face features from of the input raw images. The activations of the fully connected layer FC6, of length $N=4096$, the first fully connected layer, is used as the face features for the proposed system. Second, feature selection is applied by the aid of an offline smart Genetic Algorithm (GA) optimization. Using a training database, the GA iteratively evolves to the best chromosome that represents the most revenant FC6 features of length $M<N$ for the problem of AIFR. Finally, a classification step is applied to perform the face identification. The details of each of these steps are illustrated in this section.

2.1.　Feature extraction

The proposed method applies transfer learning [22, 23] using the VGG-Face model [14] to extract compact face features for face recognition. The model is composed of 12 convolutional layers and three fully connected layers trained using 2.6 million face images collected from 2,622 persons [14]. Since the utilized age invariant face databases, i.e., FGNET and MORPH-II, are of much smaller sizes, transfer learning is preferred to avoid over-fitting [22, 23]. Applying transfer learning, the trained convolutional layers' parameters are transferred to the proposed model, whereas the fully connected layers are trained using the utilized age invariant face databases, i.e., FGNET and MORPH datasets. In addition, transfer learning involves replacing the classification layer (the last fully connected layer) with the proposed system classifier [24]; in this manuscript, a KNN classifier is used. In the proposed system, the activations of the first fully connected layer, FC6, of length $N=4096$ are used as the compact face features.
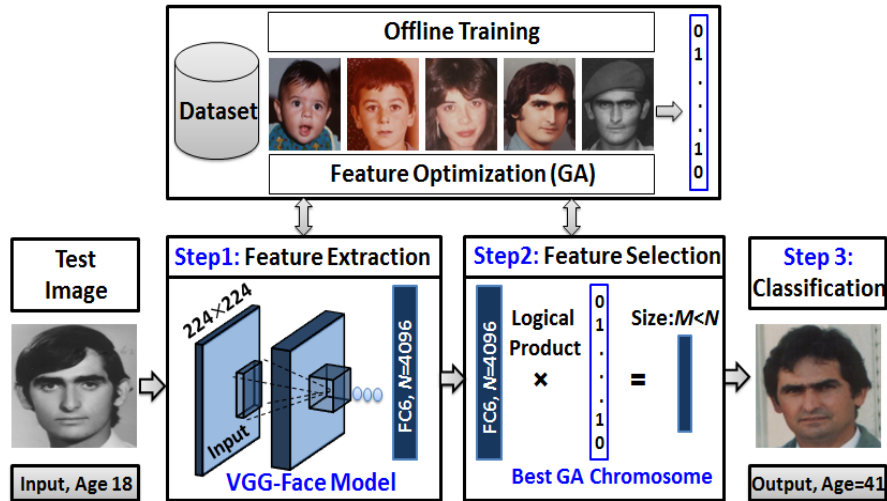
Figure 1. Diagram of the proposed method for AIFR. Input raw image is processed by the VGG-Face model. The activations of the fully connected layer, FC6, of length N=4096, is used as a feature descriptor. Feature selection is applied offline using a GA, which selects the most relevant features for AIFR of length $M < N$. Finally, classification is performed for face identification

2.2. Genetic algorithm

To select the most relevant features to the problem of AIFR, a smart GA is trained offline, using the activations of the fully connected layer FC6 of size N=4096. The proposed GA block diagram is illustrated in Figure 2. Initially, a random population of size S_P is generated from L logical vectors of length N=4096, where L is the total number of images (total data size). A logic one indicates that the feature is selected, whereas a logic zero indicates that the feature is not selected. The cost function (CF) of each chromosome in the population is estimated as the inverse of the recognition rate (RR) as follows:

$$CF = \frac{1}{RR} = 1 \div \frac{\text{Number of Recognized Faces}}{\text{Total Data Size}}$$
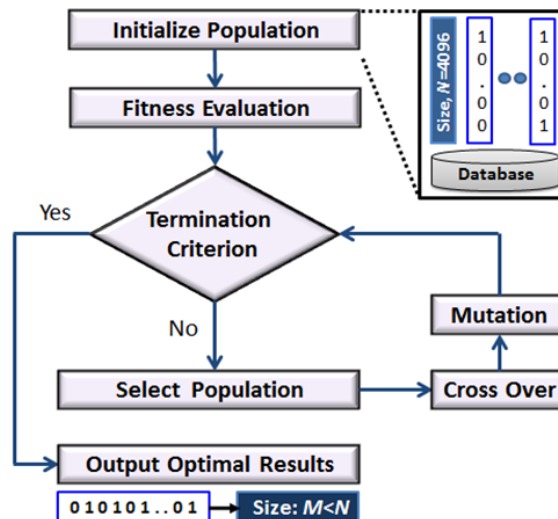

Figure 2. Diagram of the proposed smart GA optimization performed offline at the training phase of the proposed method. The initial population is randomly generated from logic vectors of length N=4096, where a logic one indicates that the feature is selected and a logic zero indicates that the feature is notthe best chromosome of the selected most relevant features to AIFR of length $M < N$

The GA runs through a number of generations N_G to minimize the cost function, i.e., maximize the RR through performing GA operations (selection, cross over and mutation). The chromosome of the minimum CF (maximum RR) represents the selected features of size $M < N$ as shown in Figure 2. The proposed GA uses the following setting: tournament selection [25], with a tournament size of two, blending crossover [26], and adaptive feasible mutation [27]. The best population size S_P and best number of generations N_G for the proposed GA are investigated throughout experimentations.

2.3. Classification

The proposed method investigates using different distance metrics (Euclidian, Manhattan, Correlation, and Cosine) for the K-Nearest Neighbor (KNN) [28] classifier to perform face recognition using the selected GA features of length $M < N$. The Rank-1 KNN classifier is selected in the proposed system for three reasons: (*i*) its simplicity and efficient time performance and computations, (*ii*) to investigate the quality of the selected GA features by applying a simple classifier, and (*iii*) KNN classifier does not need training; since it performs recognition by simply selecting the database image of feature vector y that has the minimum distance to the test image feature vector x. Let M be length of the GA feature vector. Let i be the index of a feature point in a vector; $i \in \{0, 1 ... M\}$. The following distance metrics are investigated for the proposed system:

Euclidian distance:

$$d_E(x, y) = \sqrt{\sum_i^M (x_i - y_i)^2}$$

(1)

Manhattan (Cityblock) distance:

$$d_M(x, y) = \sum_i^M |x_{i-} y_i|$$

(2)

Correlation distance:

$$d_C(x, y) = \frac{\sum_i^M (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^M (x_i - \bar{x})^2 \sum_i^M (y_i - \bar{y})^2}}$$

(3)

Cosine distance:

$$d_{cosine}(x, y) = \frac{\sum_i^M (x_i y_i)}{\sqrt{\sum_i^M (x_i)^2 \sum_i^M (y_i)^2}}$$

(4)

3. EXPERIMENTAL RESULTS AND DISCUSSIONS

This section describes in details the utilized aging databases, experimental setup, experimental results, as well as the comparative results to other state-of-the-art methods.

3.1. Datasets

Two well-known aging databases are used to test the performance of the proposed method, namely, FGNET [3] and MORPH-II (available online for academic research) [4]. Table 1 describes the details of these data. As shown in the Table, MORPH contains more than 55 thousand images from 13 thousand subjects, ranging from 16 to 77 years old. On the other hand, FGNET contains 1002 images from 82 persons, ranging from 0 to 69 years old. Samples for a span of one subject aging images for each dataset are exemplified in Figure 3 and Figure 4. The biometric recognition of faces in these databases is very challenging due to the limited number of images per subject across the different ages, making the training process training harder. In addition, the images are of different poses, blurring, illumination, and camera distances as shown in Figure 3 and Figure 4. Moreover, at lower ages, i.e., between 0 and 2 as in FGNET, the personal face attributes are still developing and it is difficult to relate the image to the corresponding subject, even visually, see the first row in Figure 3.

Table 1. FGNET and MORPH (album-II) database description

Data	# Subjects	# Images	# Images/Subj.	Age Range	Age Gap
FGNET	82	1002	5-13	0-69	0-45
MORPH	13,617	55,134	2-5	16-77	0-5

Figure 3. A sample of an image span of one subject over age for FGNET database. The subject age is provided below each image

Figure 4. A sample of an image span of one subject over age for MORPH database. The subject age is provided below each image

3.2. Experimental setup

The raw images, as shown in Figure 3 and Figure 4, are input to the proposed method without any preprocessing steps. The proposed system processes the raw images using the VGG-Face model for features extraction, offline GA for feature selection, and KNN for classification as shown Figure 1. To select the best parameters of the proposed GA, i.e., different values for the number of generations N_G and the population size S_P are investigated, i.e., N_G=60, 80, and 100 and S_P=30 and 40, respectively. In order to train the GA, data is divided into train and test datasets using Leave-One-Person-Out (LOPO) scheme. For a database of size L, LOPO runs a number of L experiments, on each uses $L-1$ images for training and only one image for testing. The LOPO scheme, used for AIFR throughout literature (e.g., in [10, 19, 29]), is usually used in forensic applications and missed person identification, where one image of the person at a certain age is available and it is required to retrieve the correct person identity from the aging database. KNN classifier [28] is used for matching with $K = 1$.

3.3. Results of VGG-face model

The activations of the fully connected layer, FC6, of the VGG-Face model are used as the features for recognition, achieving the best recognition rates RR of 80.7% and 92.8% on FGNET and MORPH databases, respectively, as shown in Table 2. The visual results, represented in Figure 5, and the quantitative results, reported in Table 2, indicate that FC6 can represent a good initial feature vector for AIFR.

Table 2. Comparison between using all VGG-model's FC6 features (*N*=4096) and the optimized GA feature vector of length *M* in terms of the Rank=1 *RR* for FGNET and MORPH-II datasets with different KNN distance metrics (i.e., Manhattan, Euclidean, Correlation, and Cosine) using population size of S_P =30 and number of generations of N_G=60

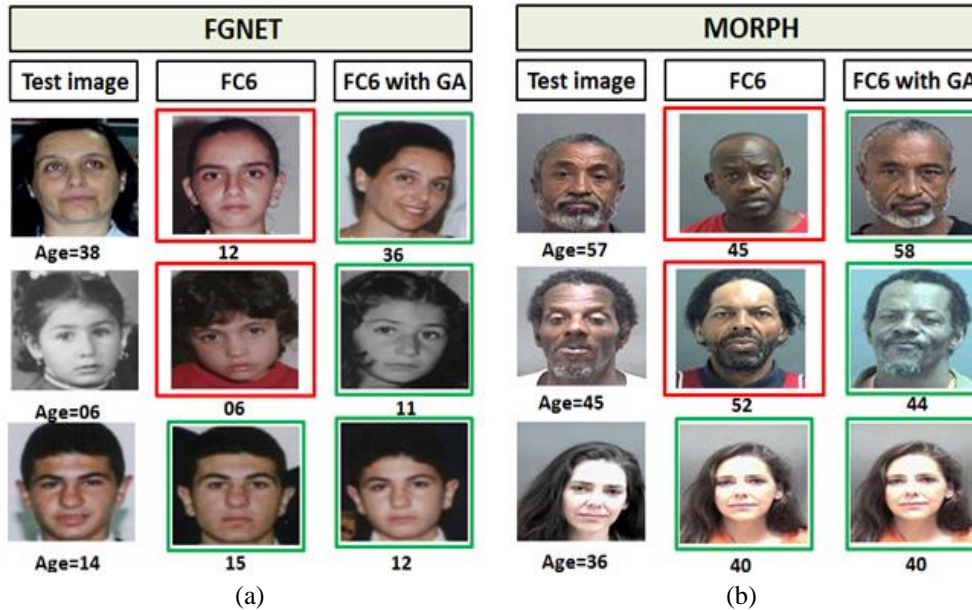| | FGNET | | | | MORPH | | | |
| | VGG, FC6 | | GA | | VGG, FC6 | | GA | |
Distance	N	RR%	M	RR%	N	RR	M	RR%
Euclidean	4096	80.7	2022	85.7	4096	92.8	2010	95.7
Correlation	4096	80.7	2022	84.0	4096	92.8	2085	94.0
Cosine	4096	80.7	2022	85.5	4096	92.8	2049	95.4
Manhattan	4096	80.7	2021	85.8	4096	92.8	2005	96.0

Figure 5. Visual system recognition for (a) three samples from FGNET and (b) three samples from MORPH (album-II) datasets using LOPO scheme and KNN classifier with Manhattan distance. For each dataset, the first column is the test image. Columns 2 and 3 represent visual results of using individual FC6 layer and of using FC6 layer after being optimized by the GA (with S_P =30 and N_G=60), respectively. A large red rectangular frame indicates failed recognition, whereas a small green rectangular frame indicates correct recognition

3.4. Results of GA optimization of VGG-face model

To improve the accuracy of the VGG-Face model, the proposed GA is used offline for feature selection. To select the best distance for the KNN classifier, recognition rates are calculated using each distance, keeping the GA parameters fixed (S_P =30 and N_G=60). As reported in Table 2, the GA optimization, using a Manhattan distance KNN classifier, is able to improve the *RR* for FGNET dataset from *RR*=80.7% using the whole 4096 FC6 features to *RR*=85.8% using the optimal feature set of length *M* = 2021. For MORPH-II dataset, GA optimization successfully improves the *RR* from *RR*=92.2% to *RR*=96.0% using an optimal feature set of length *M* =2005. In additional to the quantitative results, visual results, represented in Figure 5, support the efficiency of the proposed GA optimization to improve the performance. To further investigate the best parameter setting of the GA, different numbers of population sizes of S_P =30 and 40 and different number of generations of N_G=60, 80, and 100 are investigated using the best KNN distance metric, i.e., the Manhattan distance. As shown in Figure 6, *RR* values are slightly varying for the FGNET dataset using different parameters, between 85% and 86.2%, with the best setting N_G=60 and S_P =40, achieving *RR*=86.2%, with a feature vector of length *M*=2014. In contrast, changing the parameters for MORPH do not change the *RR* and is kept fixed at 96.0%, using a fixed feature vector of length *M* = 2005. These results support the fact that the FGNET dataset is more challenging than the MORPH dataset, and therefore requires more effort to reach the optimal performance. Figure 7 summarizes the improvement of the proposed GA optimization on the achieved *RR* values.
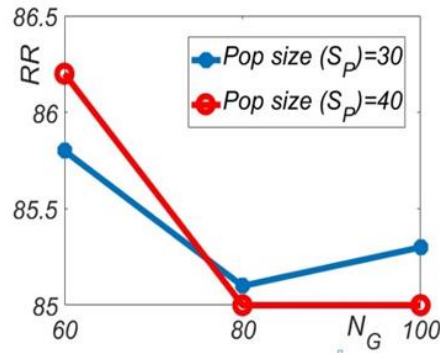
Figure 6. Comparison between RR for FGNET dataset using different GA parameter setting
(N_G=60, 80, or 100 and S_P =30 or 40), and a Manhattan distance KNN classifier
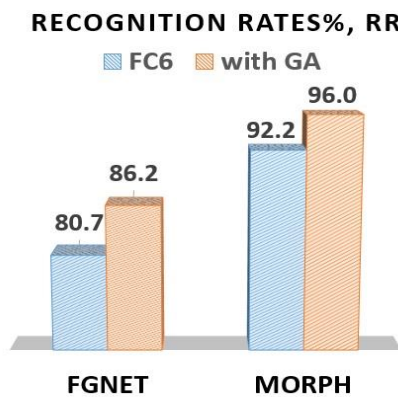

Figure 7. For FGNET, the Rank-1 RR improves from 80.7% (using all FC6 features, N=4096) to 86% after
GA feature selection, achieved using a Manhattan distance KNN classifier. For MORPH, the Rank-1 RR
improves from 92.8% (using all FC6 features, N=4096) to 96% after GA feature selection, achieved using
a Manhattan distance KNN classifier

3.5. Comparing with state-of-the-art algorithms

To highlight the contribution of the proposed method, Table 3 compares the performance of the proposed method with the state-of-the-art techniques, in terms of the RR on two standard biometric datasets, i.e. FGNET and MORPH. As reported in the Table, the proposed GA optimization shows an ability to achieve a superior performance, evidenced by the highest RR. Note that some of the state-of-the-art algorithms can't be compared to the proposed method since they either use a subset of the data (not the whole dataset, as the proposed method do with the FGNET dataset) and/or using different training scheme (e.g., 5-fold cross-validation instead of leave-one-out as the proposed method does). The elapsed times for the proposed method for feature extraction and classification are 0.792 sec and 0.0139 sec, respectively. Therefore, the overall time of the complete proposed framework is 0.806 sec.

Table 3. Comparative results in terms of the Rank-1 RR between the proposed method and the state-of-
the-art methods for AIFR on FGNET and MPROP-II datasets

Method	preprocessing	FGNET	MORPH
Generative Model [28]	YES	37.4%	79.8%
Discriminative Model [9]	YES	47.5%	83.9%
Hidden Factor Analysis [5]	YES	69%	86.8%
Maximum Entropy model [11]	YES	76.2%	92.3%
Deep Joint Learning approach [13]	YES	NA	93.6%
FR across time lapse using CNN [17]	YES	80.6%	92.2%
Modified Hidden Factor Analysis [12]	YES	76.8%	88%
AIFR-CNN model [19]	YES	76.5%	92.5%
Proposed System	NO	86.2%	96.0%

3.6. Limitations

As expected, experiments on MORPH database achieve significantly better results than FGNET. This is due to the fact that the FGNET is being more challenging since it includes higher age gaps than MORPH (maximum age gap between images of MORPH subjects is 4 year). Figure 8 shows some samples of failed recognized images using the proposed method. Figure 8 (b) indicates that image properties like illumination and distance from the camera may degrade the result of MORPH dataset. Experimental result on FGNET dataset shows a very high similarity among children in age gap (0-3 years old), so most incorrect matching occurs in this age gap, as shown in Figure 8 (a). However, in the teenage gap the face features appear to be unstable in this stage and this is clearly obvious in the mismatching images. More stability in features occurs in adult stage that leads to improving the recognition rates.
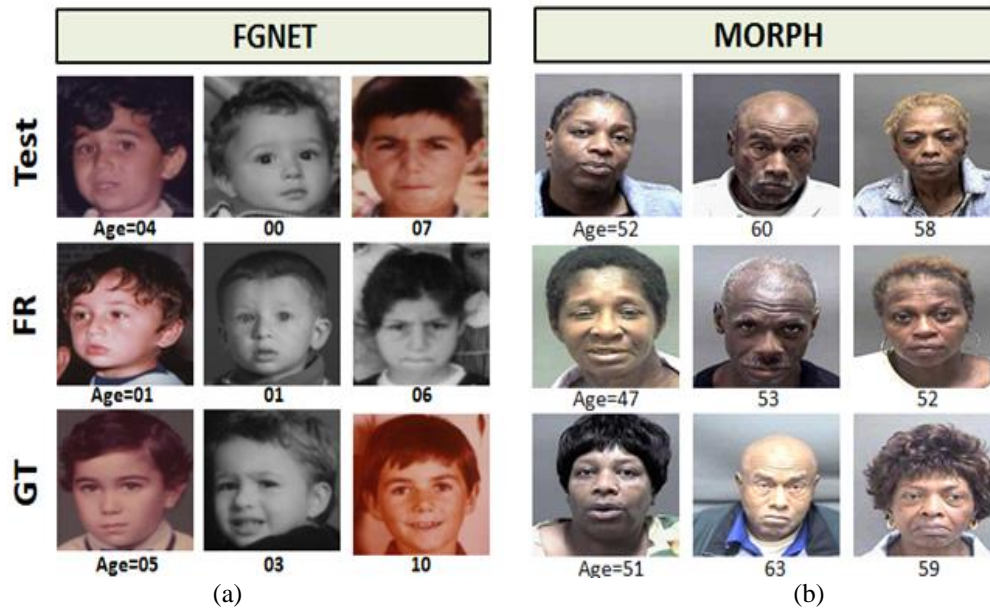

Figure 8. Visual system of failed image retrieval (a) three samples of FGNET and (b) three samples from MORPH (album-II) datasets. The first row is the test image. The second row shows Failed Retrieved (FR) images using our system. The last row is the ground truth (GT), determined as the recognized person image of the nearest age

4. CONCLUSION

A deep learning method for AIFR is proposed. The proposed method involves no registration steps, where raw face data are input directly to a VGG-Face deep learning model to extract face features, avoiding time-consuming, unreliable preprocessing steps. The extracted face features are further optimized using a smart GA that performs feature selection and reduction, resulting on the most relevant features to the problem of AIFR. In addition, the proposed GA optimization shows an ability to improve the recognition accuracy from 80.7% and 92.8% to 86.2% and 96% on two publicly available databases, i.e., FGNET and MORPH, respectively. Comparing to the state-of-the-art method, the proposed method has shown promising results.

REFERENCES

[1]   A. Lanitis, C. J. Taylor, and T. F. Cootes, "Toward automatic simulation of aging effects on face images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 442–455, 2002.
[2]   U. Park and A. K. Jain, "Face matching and retrieval using soft biometrics," *IEEE Trans. Inf. Forensics Secur.*, vol. 5, no. 3, pp. 406–415, 2010.
[3]   The Face and Gesture Recognition Network (FG-NET) Aging Database, https://www-prima.inrialpes.fr/FGnet/html/benchmarks.html [accessed December 2019].
[4]   K. Ricanek and T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," in *Automatic Face and Gesture Recognition, 7th International Conference on FGR*, pp. 341–345, 2006.

[5]   D. Gong, Z. Li, D. Lin, J. Liu, and X. Tang, "Hidden factor analysis for age invariant face recognition," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2872–2879, 2013.

[6]   D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[7]   T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.

[8]   S. R. Konda, V. Kumar, and V. Krishna, "Face recognition using multi region prominent lbp representation," *Int. J. Electr. Comput. Eng.*, vol. 6, no. 6, p. 2781, 2016.

[9]   H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs, "Face verification across age progression using discriminative methods," *IEEE Trans. Inf. Forensics Secur.*, vol. 5, no. 1, pp. 82–91, 2010.

[10] Z. Li, U. Park, and A. K. Jain, "A discriminative model for age invariant face recognition," *IEEE Trans. Inf. forensics Secur.*, vol. 6, no. 3, pp. 1028–1037, 2011.

[11] D. Gong, Z. Li, D. Tao, J. Liu, and X. Li, "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.

[12] H. Li, H. Zou, and H. Hu, "Modified hidden factor analysis for cross-age face recognition," *IEEE Signal process. Lett.*, vol. 24, no. 4, pp. 465–469, 2017.

[13] Y. Li, G. Wang, L. Lin, and H. Chang, "A deep joint learning approach for age invariant face verification," in *Chinese Conference on Computer Vision CCF*, pp. 296–305, 2015.

[14] O. M. Parkhi, A. Vedaldi, A. Zisserman, and others, "Deep face recognition.," in *Proceedings of the British Machine Vision Conference (BMVC),Swansea*, vol. 1, no. 3, p. 6.

[15] Z. Qawaqneh, A. A. Mallouh, and B. D. Barkana, "Deep Convolutional Neural Network for Age Estimation based on VGG-Face Model," *arXiv Prepr. arXiv1709.01664*, 2017.

[16] A. Anand, R. Donida Labati, A. Genovese, E. Munoz Ballester, V. Piuri, and F. Scotti, "Age estimation based on face images and pre-trained convolutional neural networks," in *CISDA*, 2017.

[17] H. El Khiyari and H. Wechsler, "Face recognition across time lapse using convolutional neural networks," *J. Inf. Secur.*, vol. 7, no. 03, pp. 1028–1037, 2016.

[18] C. Xu, Q. Liu, and M. Ye, "Age invariant face recognition and retrieval by coupled auto-encoder networks," *Neurocomputing*, vol. 222, pp. 62–71, 2017.

[19] M. Nimbarte and K. Bhoyar, "Age Invariant Face Recognition Using Convolutional Neural Network," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 4, 2018.

[20] A. Dutta, M. Günther, L. El Shafey, S. Marcel, R. Veldhuis, and L. Spreeuwers, "Impact of eye detection error on face recognition performance," *IET biometrics*, vol. 4, no. 3, pp. 137–150, 2015.

[21] H. Mohammadzade, A. Sayyafan, and B. Ghojogh, "Pixel-level alignment of facial images for high accuracy recognition using ensemble of patches," *JOSA A*, vol. 35, no. 7, pp. 1149–1159, 2018.

[22] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: an astounding baseline for recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 806–813, 2014.

[23] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker, "Feature Transfer Learning for Deep Face Recognition with Long-Tail Data," arXiv Prepr. arXiv1803.09014, 2018.

[24] C.-L. Zhang, H. Zhang, X.-S. Wei, and J. Wu, "Deep bimodal regression for apparent personality analysis," in *European Conference on Computer Vision*, pp. 311–324, 2016.

[25] B. L. Miller, D. E. Goldberg, and others, "Genetic algorithms, tournament selection, and the effects of noise," *Complex Syst.*, vol. 9, no. 3, pp. 193–212, 1995.

[26] A. E. Eiben and C. A. Schippers, "On evolutionary exploration and exploitation," *Fundam. Informaticae*, vol. 35, no. 1–4, pp. 35–50, 1998.

[27] D. Thierens, "Adaptive mutation rate control schemes in genetic algorithms," in *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, vol. 1, pp. 980–985, 2002.

[28] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. theory*, vol. 13, no. 1, pp. 21–27, 1967.

[29] U. Park, Y. Tong, and A. K. Jain, "Age-invariant face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 947–954, 2010.