Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 21(3): 4485–4500.
DOI: 10.3934/mbe.2024198
Received: 30 December 2023
Revised: 02 February 2024
Accepted: 18 February 2024
Published: 28 February 2024

Research article

Facial age recognition based on deep manifold learning

Huiying Zhang, Jiayan Lin, Lan Zhou, Jiahui Shen and Wenshun Sheng*

Pujiang Institute, Nanjing Tech University, Nanjing 211200, China

* **Correspondence:** Email: 282818344@qq.com.

Abstract: Facial age recognition has been widely used in real-world applications. Most of current facial age recognition methods use deep learning to extract facial features to identify age. However, due to the high dimension features of faces, deep learning methods might extract a lot of redundant features, which is not beneficial for facial age recognition. To improve facial age recognition effectively, this paper proposed the deep manifold learning (DML), a combination of deep learning and manifold learning. In DML, deep learning was used to extract high-dimensional facial features, and manifold learning selected age-related features from these high-dimensional facial features for facial age recognition. Finally, we validated the DML on Multivariate Observations of Reactions and Physical Health (MORPH) and Face and Gesture Recognition Network (FG-NET) datasets. The results indicated that the mean absolute error (MAE) of MORPH is 1.60 and that of FG-NET is 2.48. Moreover, compared with the state of the art facial age recognition methods, the accuracy of DML has been greatly improved.

Keywords: age recognition; manifold learning; deep learning; convolution neural network; feature extraction; mean absolute error

1. Introduction

Age recognition is used to predict a person's age or age range by analyzing the features of the face. Figure 1 shows several images from the publicly available dataset FG-NET [1] with different facial features and ages, such as 2 and 61 years old. Age recognition is widely used in the real world; for example, in e-commerce applications [2], we can analyze the age of people's faces to recommend products they might like; in the human-computer interaction [3], we can prevent minors from buying cigarettes and alcohol; in criminal cases [4], we can determine the age or age range of the victim or suspect.

Many age-related feature extraction algorithms have been proposed to extract facial features from the shape, color, and texture of the face, such as the anthropometric model [5], bio-inspired features

Figure 1. Age recognition based on facial image. Below are their corresponding ground-truth ages.

(BIF), [6] and active appearance models (AAM) [7]. With the development of deep learning, the state of the art facial recognition methods utilize deep convolutional neural networks (DCNNs) to extract face features to identify age, which improves performance of age recognition to human level [8–10]. There is a large amount of redundant information in high-dimensional face features extracted by deep learning, for example, the gender, expression, and posture information, which affects the accuracy of age recognition. It is important to extract age-related information from high-dimensional face features and improve the discriminability of features to better the accuracy of age recognition. Therefore, in order to improve the accuracy of age recognition, it is necessary to eliminate inappropriate age characteristics.

On the other hand, face manifold learning, such as principal component analysis (PCA) [11], local preserving projection (LPP) [12], orthogonal local preserving projection (OLPP), [13] and conformal cembedding analysis (CEA) [14], extracts facial features through original face gray-scale image space. Manifold learning attempts to learn face features through a low-dimensional approach. Manifold learning maps high-dimensional face image data to a low-dimensional manifold space and can extract more discriminative features for better representation and understanding of data. However, it is computationally hard to map high-dimensional features of the original image to low-dimensional manifold space. In the original face space, every pixel is regarded as a feature, resulting in a very high feature dimension, which makes it difficult to deal with for manifold learning. Moreover, correlation and redundancy exist among pixels, resulting in the manifold retaining a large amount of redundant information [15].

In order to extract facial features with high accuracy and low computation cost, this paper combines the advantages of both deep learning and manifold learning, named DML. We first extract high-dimensional aging features from CNN, embed them in the low-dimensional aging subspace based on manifold learning to learn discriminative age features of the face. Finally, we validate the DML in experiments. Experimental results show that DML is superior to the state of the art methods.

The main contributions of this paper are as follows:

1) In the age recognition task, we proposed to combine deep learning and manifold learning for the first time.

2) Combine deep learning with manifold learning to reduce the redundant information of face features and reduce the complexity of face recognition, thus greatly improving the accuracy and efficiency of facial age recognition.

3) We select support vector regression (SVR) for age regression task, and avoid local optimal solution by minimizing structural risk during SVR training, thus improving the stability and generalization ability of the model. SVR has good robustness to noise and outliers.

The structure of this paper is s organized as follows. The related work is reviewed in Section 2. In the third part, the relevant concepts including aging manifold learning, DML, and age regression are introduced. We give the overview description of DML for age estimation. Experiments and analysis of results are introduced in Section 4. Finally, the fifth part summarizes and expounds the future work.

2. Related works

The problem of automatic age recognition is one of the most challenging problems in recent years. Over the last 30 years, a number of methods have been proposed based on facial features extraction. Facial age recognition mainly extracts facial features from shape, skeletal lines, skin color, and texture for age classification or regression analysis. The methods of facial age recognition based on feature extraction are mainly divided into traditional machine learning [16] and deep learning [17].

Early facial age recognition methods were mainly based on traditional machine learning algorithms, such as support vector machine (SVM), stochastic forest, naive bayes classifier [18], so on. Traditional methods of machine learning rely on hand-crafted facial features, such as local binary pattern (LBP) [19], scale invariant feature transform (SIFT), oriented gradient histogram (HOG), BIF, Gabor [14], etc. This traditional machine learning algorithm requires high image quality, is sensitive to illumination change, occlusion and noise, and the computation complexity is very high. On the one hand, manifold learning method is used in facial age recognition by using feature selection methods. LRDOR [20] uses the orthogonal representation idea to obtain uncorrelated features and captures local correlation among features. Karami et al. [21] obtains representative features through spatial distance, that is, the Variance–Covariance distance to handle with the dimensionality reduction and subspace learning. Thanks to the DCNN (AlexNet) proposed by Krizhevsky et al. [22], 1000 objects can be efficiently classified. More and more researchers have devoted themselves to face age recognition in deep learning and achieved good results. Deep learning-based facial feature extraction adopts a CNN to automatically extract facial features. Compared with traditional facial feature extraction, the feature obtained by the deep learning method is more discriminative. Meanwhile, deep learning can also provide convenient end-to-end solutions. Gao et al. [8] used a 16-layer CNN to learn image features and get the age recognition value. Compared with traditional methods, the error is smaller. Zhang et al. [9] proposed that individuals are influenced by internal and external factors to form personal aging patterns, and two deep learning models of CNN and long short term memory (LSTM) are used to identify age. CNN is used to extract facial features, and LSTM is used to learn personal aging pattern in time series images of human faces. Mei et al. [10] combined the output of multiple CNNs into the final result using a framework containing multiple CNNs. Compared with single CNN, the combination of multiple CNNs improves the accuracy of age identification. Researchers also proposed a variety of optimized CNN architectures and conducted a comprehensive experiment on CNN architectures with different convolutional kernels and layers. Recently, researchers introduced transfer learning and attention mechanism into a face age recognition system to extract age-related features [23, 24]. Sonal et al. [25] proposed the relationship between the features of important parts of the face and age recognition, and adopted the rotational local binary pattern (RLBP) to extract and select features of facial images to identify age.

Based on the extracted facial features, facial age recognition is usually treated as a classification problem, a regression problem, or a combination of both. When age is a classification problem, each facial is classified as one of the ages associated with $\{a_1, a_2, \ldots, a_N\}$, where N is the number of classifications. As a regression problem, the purpose is to learn the regression function to recognize the age of the face, map the facial features to the age label, and the age recognized as a value a.

3. Deep manifold learning methods

In this section, we first introduce aging manifold learning, and DML, then we introduce age regression. Finally, we give the overview description of DML for age estimation.

3.1. Preliminary: Aging manifold learning

Although the large amount of high-dimensional data has provided a lot of information for the study of age recognition, it has also greatly increased the workload. More importantly, the high-dimensional data is often correlated in practice, which makes the age recognition more complex and inconvenient. However, the blind reduction of facial features will result in the loss of important information, which will result in the wrong conclusion of age recognition. Manifold learning is a method of mapping high-dimensional data into a low-dimensional space. The purpose of aging manifold learning is to embed high-dimensional features into low-dimensional features and discover geometric features of aging space. Aging manifold extracts latent low-dimensional manifold structure from facial features and finds out aging rule.

Aging manifold learning assumes that facial feature space F is represented as $F = \left\{f_i : f_i \in R^D\right\}_{i=1}^{N}$, where D represents the dimension of feature space, N represents the number of facial images, and the real age label a_i is represented as $y = \{a_i : a_i \in N\}_{i=1}^{N}$ · We want to learn a low-dimensional manifold G embedded in F to obtain a manifold aging feature subspace $X{=}\left\{x_i : x_i \in R^d\right\}_{i=1}^{N}$ with $d << D$. More specifically, our learning objective is to find $D \times d$ and a projection matri $P = [p_1, p_2, \ldots, p_d]$ such that $X = P^T F$, where $F = [f_1, f_2, \ldots, f_N] \in R^{D \times N}$.

3.2. Deep Manifold Learning (DML)

Extracting feature parameters representing age from face images lays a foundation for facial age recognition. Accurate facial features improve the performance of the age recognition system and improve the accuracy of the age recognition algorithm. Compared with facial features extracted by traditional methods, facial features extracted by deep learning improve the accuracy of age recognition.

The deep learning-based CNN includes a series of convolution, pool ing, activation, and full connection layers for obtaining high-dimensional facial discriminative features. In addition to age information, it also contains other information such as gender, ethnicity, expression, and posture. In the case of high correlation, noise, and high dimension, it has been proven to be an effective method to project the observed data into the subspace. In order to improve the accuracy of facial age recognition, it is particularly important to find out age-appropriate features from deep learning features. DML embeds high-dimensional facial features extracted from the deep learning framework into the low-dimensional discriminative aging subspace through manifold learning.

DML assumes that the representation of original facial feature space F is represented as $F = \left\{f_i : f_i \in R^D\right\}_{i=1}^{N}$, where D represents the dimension of the original human facial space, N represents the number of facial images, and the real age label a_i is represented as $y = \{a_i : a_i \in N\}_{i=1}^{N}$. We want

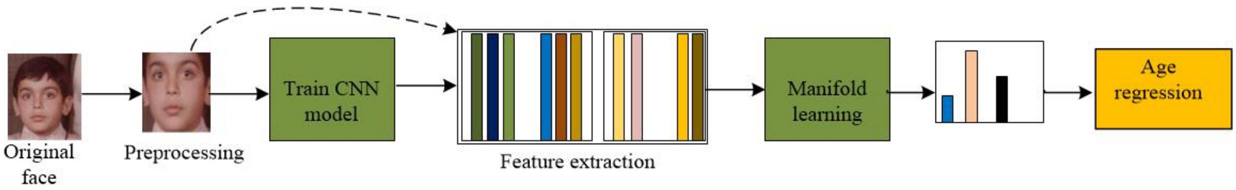

Figure 2. The proposed age recognition framework based on DML.

to extract a high-dimensional deep facial feature space $C = \{c_i : c_i \in R^c\}_{i=1}^N$ through the deep learning model, where C represents the dimension of the deep facial feature space extracted by the CNN model from the original facial feature space. We want to learn a low-dimensional manifold Q embedded in C to obtain a manifold aging feature subspace $X = \left\{x_i : x_i \in R^d\right\}_{i=1}^N$, and $d \ll c \ll D$. More specifically, our learning objective is to find $c \times d$ and projection matrices $P = [p_1, p_2, \ldots, p_c]$ where $Q = [q_1, q_2, \ldots, q_d]$ such that $X = (PQ)^T F$, where $F = [f_1, f_2, \ldots, f_N] \in R^{D \times N}$.

3.3. Age regression

After finding a low-dimensional subspace representation of a facial image via DML learning, we formulate age recognition as a regression problem. Guo and Mu [26] showed that the facial age subspace variable is a complex function of age, and age regression is a curve form. The facial age recognition is a complex nonlinear function, so we use the nonlinear regression method. Compared with logistic regression and neural networks, SVR provides a clearer and more powerful way to learn complex nonlinear equations. It is not only simple in algorithm, but also has better robustness [14]. As the literature [22] concludes: 1) The aging is a complex nonlinear regression problem, especially for a large span of years, e.g., 0–80. The SVR can model properly the complex aging process; 2) The SVR technique is only concerned with errors greater than deviation value. Compared with other loss functions, SVR is less sensitive to outliers and has higher robustness. Therefore, we use nonlinear SVR to perform regression on the subspace of the facial deep manifold.

3.4. The overview framework of DML

The proposed age recognition framework based on DML mainly consists of five modules: facial preprocessing, training deep learning model, feature extra: 1) The facial preprocessing includes facial detection, alignment, and cropping, 2) the deep learning model is trained and the model parameters are saved, 3) the trained model parameters are used to extract high-dimensional facial features, 4) the aging features extracted by CNN are transformed into aging subspace by manifold learning, 5) age regression. The pseudocode for the algorithm is shown in Algorithm 1.

4. Experiments and analysis of results

4.1. Dataset and preprocessing

We verify the effectiveness of our proposed method DML on two publicly available facial datasets MORPH and FG-NET. MORPH contains a total of 55,134 facial images of about 13,000 people of different races, with an average of 4 images per person, ranging from 16 to 77 years old. FG-NET contains 82 subjects labeled with exact ages from 0 to 69 years old, with a total of 1002 high-resolution

Algorithm 1: Deep Manifold Learning

Input: The training set $\boldsymbol{S} = \{(\boldsymbol{s}_i, y_i)\}_{i=1}^{N}$,the facial feature $\boldsymbol{C} = \{\boldsymbol{c}_i : \boldsymbol{c}_i \in R^c\}_{i=1}^{N}$,the facial subspace $\boldsymbol{X} = \left\{\boldsymbol{x}_i : \boldsymbol{x}_i \in R^d\right\}_{i=1}^{N}$,and the convergence criterion ϵ and ε

Output: $y = f(x)$

1. Initialize the model parameter θ
2. $i \leftarrow 0$
3. Repeat
4. Until $T\left(\theta^i\right) - T\left(\theta^{i-1}\right) < \epsilon$
5. Facial feature extraction
6. Manifold learning
7. $f(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b \quad s.t. \begin{cases} \min_{w,b} \frac{1}{2} \parallel W \parallel^2 \\ \mid \mathrm{y_i} - (wx_i + b) \mid \leq \varepsilon \end{cases}$

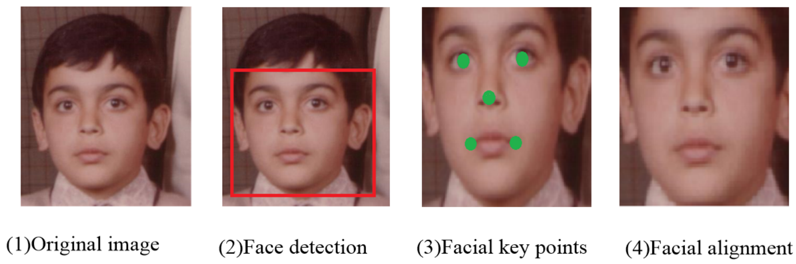

Figure 3. Facial image preprocessing process.

colors or gray facial images.

Facial Image Preprocessing. Facial image preprocessing is the first step of facial age recognition. The result of facial preprocessing directly affects the accuracy of facial feature extraction. The facial preprocessing process is shown in Figure 3.

The deformable parts model (DPM) [9] is used to detect the facial region, then five main key points of face are detected, such as left/right eye center, nose tip, and left/right mouth corner. According to the key points of the face, the facial is aligned with the upright posture and the main facial region is cropped. The cropped facial image is converted into 224 × 224*3 pixels.

4.2. Evaluation criteria

In this paper, mean absolute error (MAE) and dumulative scores (CS) are used to evaluate the performance of the DML. MAE is the mean absolute error between the identified age and the physiological age. CS is the accuracy of age recognition at different levels (years). MAE is calculated by:

$$MAE = \frac{1}{N} \sum_{n=1}^{N} \mid y - y^* \mid, \tag{4.1}$$

where y indicates the ground-truth age, which is labeled in the database, y^* indicates the estimated age from the model, and N is the number of test samples.

Table 1. Parameter setting on MORPH and FG-NET in Inception v4.

parameters	MORPH	FG-NET
dropout rate	0.8	0.8
learning rate	0.001	0.001
epochs	100	100
mini-batchs	80	2
loss function	cross entropy	cross entropy
optimizer	SGD	SGD

The definition of CS is:

$$CS = \frac{N_l}{N} \times 100\%, \tag{4.2}$$

where N_l is the number of test images satisfying $|y - y^*| \leq m$, and m is the set value (years).

4.3. Network initialization and training

Convolutional neural network Inception v4 [27] was proposed by the Google team in 2016, and it improves on the previous Inception architecture. InceptionA, InceptionB, and InceptionC are three important modules in Inception V4. They perform convolution operations and multi-channel operations on different scales (1×1, 3×3, etc.) and splice the obtained feature maps, thus capturing features at different levels and scales. We train the Inception V4 on Windows 10 system using the open source TensorFlow framework, Python 3.7, PyCharm development environment, and NVIDIA Titan XP graphics cards.

We randomly divided the datasets MORPH and FG-NET into 2 parts, in which 80% images were selected for training and the remaining 20% images were used for testing. The stochastic grandient descent (SGD) optimization algorithm was selected. The loss function is the cross entropy function. Table 1 shows the initialization of network parameters.

On MORPH and FG-NET, InceptionV4 was trained with learning rates of 0.1, 0.001 and 0.0001, respectively. The experimental results show that the learning rate of 0.1 would cause the model to not converge easily, and the model convergence would be particularly slow when the learning rate of 0.0001 is selected. Therefore, the learning rate of 0.001 is selected to. The training models of MORPH and FG-NET were saved separately, and the features of each image were extracted. The final layer of Inception v4 is a fully connected layer of 1536 dimensions for facial feature extraction.

4.4. Manifold learning algorithm

There are many kinds of data manifold learning algorithms. Some typical manifold learning algorithms are summarized below. Principal components analysis (PCA) and independent component analysis (ICA) belong to linear dimension reduction algorithms, and isometric feature mapping (ISOMAP) and Locally linear embedding (LLE) belong to nonlinear dimension reduction algorithms.

PCA [11]. PCA is one of the most commonly used linear unsupervised dimension reduction methods. It can identify the main features from the high-dimensional feature space. By rotating the data coordinate axis to the most important direction of the data angle, namely the maximum variance, selecting the principal component to be retained and discarding the secondary component, PCA can

Table 2. MAE of different characteristic dimensions on MORPH and FG-NET datasets by PCA.

Retention Information ratio (%)	MORPH dimension	MORPH MAE	FG-NET dimension	FG-NET MAE
100	1536	2.42	1536	3.74
98	300	1.54	291	3.70
95	111	1.55	171	3.69
90	34	1.56	86	3.67
85	16	1.61	50	3.65
80	10	1.68	35	2.98
75	5	1.72	28	2.54
70	3	1.79	23	2.51
64	2	2.17	20	2.48

use fewer data dimensions and retain more original data features. The original features are simplified without losing information.

ICA [28]. ICA is a linear unsupervised dimension reduction method, which does not require data labels as prior information. In the case of non-Gaussian distribution, the dimension vector obtained by PCA method may not be the optimal solution. In this case, variance cannot be used as the measurement standard, so the orthogonal assumption between dimensions ICA is used.

ISOMAP [15]. ISOMAP is a nonlinear unsupervised dimension reduction algorithm. By preserving the pair geodesic distance between data points, it solves the problem that two points close to each other in the original space are still far apart in the prevalent space in the low-dimensional state. In this technique, the geodesic distance is calculated by introducing the neighborhood graph and PCA is applied to the graph distance matrix to construct the embedded manifold. The limitation of this technique is high computational complexity.

LLE [29]. LLE is a nonlinear unsupervised dimension reduction algorithm, which is a manifold learning method used to preserve local attributes of data. LLE uses the nearest neighbor linear combination to represent the high-dimensional data points and achieves the low-dimensional representation by keeping the reconstructed weights of the linear combinations as identical as possible.

4.5. Experimental results and analysis

We conducted experiments on two datasets, MORPH [30] and FG-NET. Inception v4 extracted the features of the two datasets respectively, and four manifold learning methods PCA, ICA, ISOMAP, and LLE are used to reduce the dimension of facial features. The MAE values obtained by four dimension reduction methods in different dimensions are shown in Tables 2–4.

Test the PCA. As shown in Table 2, the facial features extracted by Inception v4 on the MORPH and FG-NET are 1536 dimensions, and MAE on the two datasets are 2.42 and 3.74, respectively. PCA preserves the same percentage of source information in the two datasets, resulting in different feature dimensions and significant differences in MAE values. **Test the Different Manifold Learning Methods.** As shown in Tables 3 and 4, among PCA, ICA, ISOMAP, and LLE, PCA has the lowest average MAE in different dimensions on the two datasets. PCA performs best. The ICA algorithm is second only to PCA. Tables 3 and 4 show that PCA linear dimension reduction is suitable for learning the subspace of facial features, while two nonlinear methods of ISOMAP and LLE are not ideal for learn-

Table 3. MAE of the four manifold learning algorithms on MORPH under different dimensions.

Dimension \ Method	PCA	ICA	ISOMAP	LLE
2	2.17	2.18	2.11	3.43
3	1.79	1.79	1.81	3.23
5	1.72	1.77	1.77	2.05
10	1.68	1.75	1.73	1.86
20	1.60	1.64	1.72	1.75
30	1.58	1.63	1.70	1.73
40	1.57	1.62	1.70	1.71
50	1.56	1.62	1.71	1.73
Average	1.71	1.75	1.78	2.19

Table 4. MAE of the four manifold learning algorithms on FG-NET under different dimensions.

Dimension \ Method	PCA	ICA	ISOMAP	LLE
2	6.43	6.42	7.06	11.43
3	4.89	4.90	6.76	9.78
5	4.72	4.87	6.51	10.61
10	2.65	2.75	5.70	4.13
20	2.48	2.91	7.46	3.93
30	2.58	2.88	7.39	4.04
40	3.46	3.17	7.93	5.19
50	3.65	3.62	7.91	3.60
Average	3.86	3.94	7.09	6.59

ing the subspace of facial features. From the results in Tables 3 and 4, we can draw conclusions. 1) Compared with FG-NET, MAE on MORPH is better. This is mainly because the two datasets contain very different amounts of data. MORPH is large and the trained Inception v4 is better at learning and extracting facial features. On the contrary, FG-NET is small and the performance of the trained Inception v4 is affected; the model can't extract facial features effectively. 2) Linear unsupervised manifold learning algorithms PCA and ICA can learn facial discriminative aging subspace more effectively.

On MORPH and FG-NET, PCA is superior to other methods for several reasons: 1) PCA searches for principal components through the covariance matrix between the data without relying on the labeling information of the data. In contrast, ICA, ISOMAP, and LLE generally require more prior information to guide the dimensionality reduction process. 2) PCA is a global optimization method that tries to find the projection direction that can maximize the variance of the data. This allows PCA to capture the direction of major changes in the data and to retain the information of the data to the maximum extent possible with fewer features. In contrast, ICA, ISOMAP, and LLE generally focus more on preserving the local structure and nonlinear relationships of the data. 3) PCA projects data into

Table 5. MAE based on different sex/race on MORPH.

Sex/Race	Femal	Male	Race (b)	Race (h)	Race (w)
No.of images	8400	46,600	42,500	1700	10,500
CNN(MAE)	3.14	2.23	3.81	3.94	3.67
CNN(MAE)	2.35	1.47	2.97	3.19	2.89

Table 6. MAE values of 6 methods on MORPH and FG-NET.

Method	MORPH	FG-NET
CS-LBFL	4.52	4.43
CS-LBMFL	4.37	4.36
ODFL	3.12	3.89
MA-SFV2	2.68	3.81
DL-LDL	2.23	3.14
DML	1.60	2.48

a new coordinate system composed of principal components, which can represent the main features of the data. In contrast, ICA, ISOMAP, and LLE are computationally complex and the interpretation of the results is relatively poor.

The FG-NET dataset was too small, so we classified the dataset by gender and race on MORPH, trained and extracted features on Inception v4, and reduced dimensionality according to PCA. The experimental results are shown in Table 5. The MAE results obtained in each group show that the male group has the best MAE value. The results are related to race and gender.

Test the DML Efficiency. The lower dimensions are beneficial for data storage, training regression functions, and predicting age. PCA can use lower dimensions, retain more original data features, and reduce the dimensionality of high-dimensional features when the retention of information is maximum. After comprehensive analysis, both MAE and CS are better when PCA reduces dimension to 20, which is conducive to data storage and improves the computational efficiency of the model. We selected MAE with PCA reduced to 20 dimensions to compare with other methods. Table 6 shows MAE of the state of the art algorithms on MORPH and FG-NET, such as deep learning based on label distribution learning (DL-LDL) [9],cost-sensitive local binary feature learning (CS-LBFL) [31], cost-sensitive local binary multi-feature learning (CS-LBMFL) [31], ordinal deep feature learning (ODFL) [32] and mixed attention-shuffleNetV2 (MA-SFV2) [33]. MAEs of DML are the best on MORPH and FG-NET.

Figures 4 and 5 show the DML test results on CS and compare them with other methods on MORPH and FG-NET respectively. Figures 4 and 5 reveal that DML on MORPH and FG-NET are most accurate. We can simply conclude that the deep learning method is better than the traditional machine learning, and the DML is better than the deep learning to recognize the facial age.

To observe aging manifolds, MORPH and FG-NET visualize the embedded low-dimensional manifolds. On MORPH and FG-NET, we took some faces as samples. Figure 6 shows the visualization of 2-D and 3-D manifolds using PCA, ICA, ISOMAP, and LLE algorithms on MORPH, respectively. Each data point represented an image of a face, and the data points were colored from yellow to black for ages from 16 to 77. Figure 7 shows the 2-D and 3-D manifolds visualization of PCA, ICA, ISOMAP, and LLE algorithms on FG-NET, respectively. Each data point represents an image of a

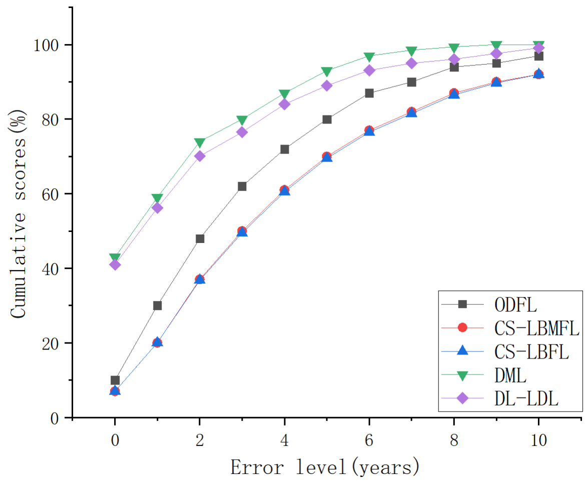

Figure 4. Comparisons of CS on MORPH.

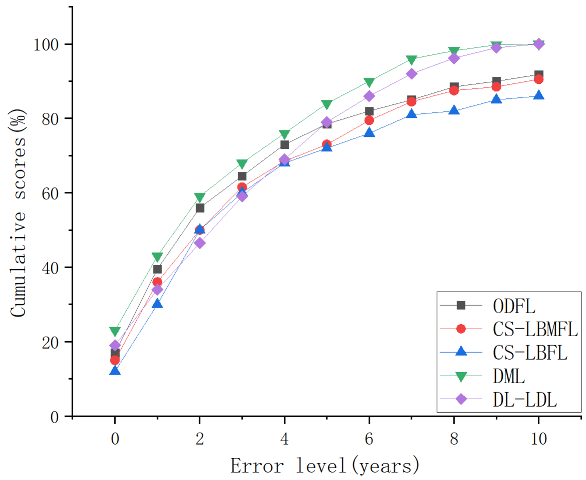

Figure 5. Comparisons of CS on FG-NET.

face. The points of age from 0 to 69 are colored from yellow to black. From Figures 6 and 7, we can see PCA and ICA depict age manifolds better and have obvious discrimination. The manifold formed by ISOMAP is distributed. The manifold formed by LLE is a scattered and the data points are not concentrated, so the discrimination is not obvious. Both ISOMAP and LLE fail to form a good manifold structure. Due to the small FG-NET dataset, the age manifold trend cannot be well reflected. However, the experimental results of FG-NET show that PCA and ICA still have better manifold trend and structure than ISOMAP and LLE methods and show clear clustering and discriminative patterns in visualization images.

In summary, through the analysis of the experimental results, we conclude the main reasons why DML produces excellent results. 1) Inception v4 has good performance. Its small convolutional kernel and deep structure are suitable for complex facial feature extraction. 2) DML based on linear unsupervised method PCA can effectively learn discriminative age feature subspace and improve the accuracy of age recognition. 3) Different dimensions based on DML also directly affect the accuracy of age identification.

DML combines the powerful feature learning ability of deep learning with the data representation

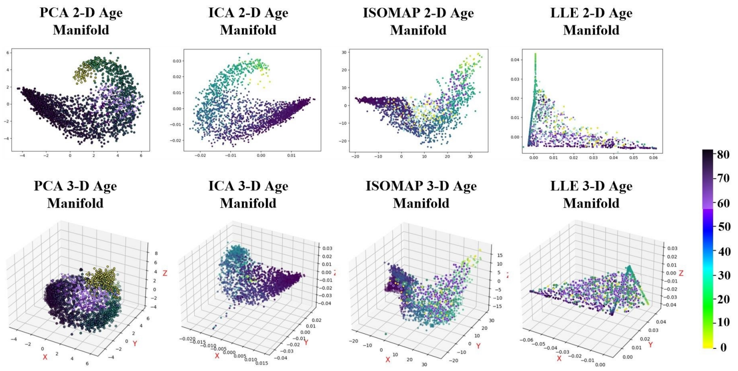

Figure 6. The 2-D and 3-D aging manifold visualization using PCA, ICA, ISOMAP, and LLE algorithms on MORPH, respectively. The color of the dots represent age, with yellow representing the youngest and black representing the oldest.

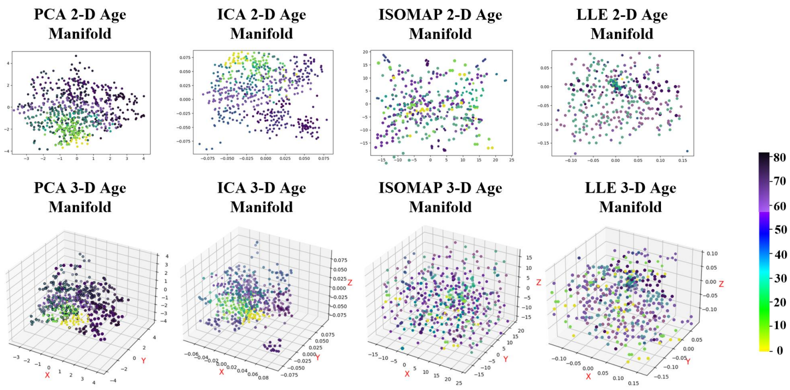

Figure 7. The 2-D and 3-D aging manifold visualization using PCA, ICA, ISOMAP, and LLE algorithms on FG-NET, respectively. The color of the dots represent age, with yellow representing the youngest and black representing the oldest.

ability of manifold learning, which can better mine the complex structure in the data, carry out efficient characterization and modeling of the data, and improve the accuracy of age recognition. Although manifold learning removes redundant information, compared with facial age recognition based on the deep learning framework, the DML method still needs manifold learning to reduce dimension after feature extraction, which increases the complexity of the algorithm.

For DML methods, deep learning cannot correctly extract facial features when the face is occluded or partially occluded or when the face appears in a non-frontal orientation. In the future, consider training models from multi-angle face images, and try to use image repair techniques to recover the blocked parts, thus providing more complete facial information.

5. Conclusions

Deep learning algorithms automatically learn effective facial features and achieve better age recognition accuracy compared with traditional algorithms. However, the high-dimensional facial features extracted by deep learning contain other relevant information, which increases the complexity of the age recognition model. By embedding the high-dimensional features of deep learning into the low-dimensional discriminative subspace through DML, the face discriminative aging subspace is learned. Experiments on two publicly available datasets MORPH and FG-NET show that DML can effectively learn aging subspace, especially PCA and ICA methods, which significantly improve the accuracy of facial age recognition. Moreover, low-dimensional representation of facial features are conducive to data storage and improve the time efficiency of training regression function.

Use of AI tools declaration

The authors declare that they have not used Artiffcial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported in partially by the University Natural Science Research Project of Jiangsu Province, No.19KJD520005; The key Project of Natural Science, Pujiang College, Nanjing University of Technology, No.NJPJ2022-0-06; The Program for outstanding young Talents in universities, No.gxyq2022108.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. G. Panis, A. Lanitis, N. Tsapatsoulis, T. F. Cootes, Overview of research on facial ageing using the FG-NET ageing database, *IET Biom.*, **5** (2016), 37–46. https://doi.org/10.1049/iet-bmt.2014.0053
2. C. Wu, H. J. Lee, Learning age semantic factor to enhance group-based representations for cross-age face recognition, *Neural Comput. Appl.*, **34** (2022), 13063–13074. https://doi.org/10.1007/s00521-022-07176-7
3. Z. Huang, J. Zhang, H. Shan, When age-invariant face recognition meets face age synthesis: A multi-task learning framework, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2021), 7282–7291. https://arxiv.org/abs/2103.01520
4. M. Awais, C. Chen, X. Long, B. Yin, A. Nawaz, S. F. Abbasi, Novel framework: face feature selection algorithm for neonatal facial and related attributes recognition, *IEEE Access*, **8** (2020), 59100–59113. https://doi.org/10.1109/ACCESS.2020.2982865

5. Y. H. Kwon, N. da V. Lobo, Age classification from facial images, *Comput. Vision Image Understanding*, **74** (1999), 1–21. https://doi.org/10.1006/cviu.1997.0549

6. G. Guo, G. Mu, Y. Fu, T. S. Huang, Human age estimation using bio-inspired features, in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, (2009), 112–119. https://doi.org/10.1109/CVPR.2009.5206681

7. K. Y. Chang, C. S. Chen, Y. P. Hung, Ordinal hyperplanes ranker with cost sensitivities for age estimation, in *Conference on Computer Vision and Pattern Recognition (CVPR)*, (2011), 585–592. https://doi.org/10.1109/CVPR.2011.5995437

8. B. B. Gao, C. Xing, C. W. Xie, J. Wu, X. Geng, Deep label distribution learning with label ambiguity, *IEEE Trans. Image Process.*, **26** (2017), 2825–2838. https://doi.org/10.1109/TIP.2017.2689998

9. H. Y. Zhang, W. S. Sheng, Y. Z. Zeng, Face Age recognition algorithm based on label distribution learning, *J. Jiangsu Univ.*, **44** (2023), 180–185.

10. S. Mei, Y. Geng, J. Hou, Q. Du, Learning hyperspectral images from RGB images via a coarse-to-fine CNN, *Sci. China Inf. Sci.*, **65** (2022), 1–14. https://doi.org/10.1007/s11432-020-3102-9

11. M. A. Marjan, M. R. Islam, M. P. Uddin, PCA-based dimensionality reduction for face recognition, *Elecommun. Comput. Electron. Control*, **19** (2021), 1622–1629. http://doi.org/10.12928/telkomnika.v19i5.19566

12. G. Guo, Y. Fu, C. R. Dyer, T. S. Huang, Image-based human age estimation by manifold learning and locally adjusted robust regression, *IEEE Trans. Image Process.*, **17** (2008), 1178–1188. https://doi.org/10.1109/TIP.2008.924280

13. X. He, P. Niyogi, Locality preserving projections, *Adv. Neural Inf. Process. Syst.*, **16** (2003).

14. Y. Fu, T. S. Huang, Human age estimation with regression on discriminative aging manifold, *IEEE Trans. Multimedia*, **10** (2008), 578–584. https://doi.org/10.1109/TMM.2008.921847

15. S. Kshatriya, M. Sawant, K. M. Bhurchandi, Feature selection and feature manifold for age estimation, in *Computer Vision and Image Processing: 5th International Conference*, **1377** (2021), 112–123. https://doi.org/10.1007/978-981-16-1092-9_10

16. R. Zhang, A review of face recognition based on deep learning, in *Institute of Management Science and Industrial Engineering.Proceedings of 2019 3rd International Conference on Computer Engineering,Information Science and Internet Technology (CII 2019)*, 2019. https://doi.org/10.26914/c.cnkihy.2019.037191

17. V. Pouli, S. Kafetzoglou, E. E. Tsiropoulou, A. Dimitriou, S. Papavassiliou, Personalized multimedia content retrieval through relevance feedback techniques for enhanced user experience, in *2015 13th International Conference on Telecommunications (ConTEL)*, 2015. https://doi.org/10.1109/ConTEL.2015.7231205

18. H. Wang, V. Sanchez, C. T. Li, Improving face-based age estimation with attention-based dynamic patch fusion, *IEEE Trans. Image Process.*, **31** (2022), 1084–1096. https://doi.org/10.1109/TIP.2021.3139226

19. N. Sharma, R. Sharma, N. Jindal, Face-based age and gender estimation using improved convolutional neural network approach, *Wireless Pers. Commun.*, **124** (2022), 3035–3054. https://doi.org/10.1007/s11277-022-09501-8

20. M. Samareh-Jahani, F. Saberi-Movahed, M. Eftekhari, G. Aghamollaei, P. Tiwari, Low-redundant unsupervised feature selection based on data structure learning and feature orthogonalization, *Expert Syst. Appl.*, **240** (2024), 122556. https://doi.org/10.1016/j.eswa.2023.122556

21. S. Karami, F. Saberi-Movahed, P. Tiwari, P. Marttinen, S. Vahdati, Unsupervised feature selection based on variance-covariance subspace distance, *Neural Networks*, **166** (2023), 188–203. https://doi.org/10.1016/j.neunet.2023.06.018

22. A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, *Adv. Neural Inf. Process. Syst.*, **25** (2012). https://doi.org/10.1145/3065386

23. O. Sendik, Y. Keller, DeepAge: Deep Learning of face-based age estimation, *Signal Process. Image Commun.*, **78** (2019), 368–375. https://doi.org/10.1016/j.image.2019.08.003

24. S. Lim, Estimation of gender and age using CNN-based face recognition algorithm, *J. Xi'an Univ. Sci. Technol.*, **9** (2020), 203–211. https://doi.org/10.7236/IJASC.2020.9.2.203

25. Sonal, A. Singh, C. Kant, Face and age recognition using three-dimensional discrete wavelet transform and rotational local binary pattern with radial basis function support vector machine method, *Int. J. Electr. Eng. Educ.*, **60** (2023), 389–404. https://doi.org/10.1177/0020720920988489

26. G. Guo, G. Mu, Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression, in *Conference on Computer Vision and Pattern Recognition (CVPR)*, (2021), 657–664. https://doi.org/10.1109/CVPR.2011.5995404

27. C. Szegedy, S. Ioffe, V. Vanhoucke, A. Alemi, Inception-v4, inception-resnet and the impact of residual connections on learning, in *Proceedings of the AAAI Conference on Artificial Intelligence*, **31** (2016). https://doi.org/10.1609/aaai.v31i1.11231

28. L. Y. Song, S. Zhou, H. P. Lu, Direct ICA on data tensor via random matrix modeling, *Signal Process.*, (2022), 508–519. https://doi.org/10.1016/j.sigpro.2022.108508

29. S. T. Roweis, L. K. Saul, Nonlinear dimensionality reduction by locally linear embedding, *Science*, **290** (2000), 2323–2326. https://doi.org/10.1126/science.290.5500.2323

30. K. Ricanek, T. Tesafaye, Image-based human age estimation by manifold learning and locally adjusted robust regression, *IEEE Comput. Soc.*, (2006), 341–345. https://doi.org/10.1109/FGR.2006.78

31. J. Lu, V. Liong, J. Zhou, Cost-sensitive local binary feature learning for facial age estimation, in *IEEE Transactions on Image Processing*, (2015), 144–157. https://doi.org/110.1109/TIP.2015.2481327

32. H. Liu, J. Lu, J. Feng, J. Zhou, Ordinal deep feature learning for facial age estimation, in *12th International Conference on Automatic Face and Gesture Recognition*, (2017), 4465–4470. https://doi.org/10.1109/FG.2017.28

33. X. Liu, Y. Zou, H. Kuang, Face image age estimation based on data augmentation and lightweight convolutional neural network, *Symmetry*, **12** (2020), 146–163. https://doi.org/10.3390/sym12010146

AIMS Press