

Age prediction based on a small number of facial landmarks and texture features

Mengjie Wang and Weiyang Chen*

College of Computer Science and Technology, Qilu University of Technology (Shandong Academy of Sciences), Jinan, Shandong, China

Abstract.

BACKGROUND: Age is an essential feature of people, so the study of facial aging should have particular significance.

OBJECTIVE: The purpose of this study is to improve the performance of age prediction by combining facial landmarks and texture features.

METHODS: We first measure the distribution of each texture feature. From a geometric point of view, facial feature points will change with age, so it is essential to study facial feature points. We annotate the facial feature points, label the corresponding feature point coordinates, and then use the coordinates of feature points and texture features to predict the age.

RESULTS: We use the Support Vector Machine regression prediction method to predict the age based on the extracted texture features and landmarks. Compared with facial texture features, the prediction results based on facial landmarks are better. This suggests that the facial morphological features contained in facial landmarks can reflect facial age better than facial texture features. Combined with facial landmarks and texture features, the performance of age prediction can be improved.

CONCLUSIONS: According to the experimental results, we can conclude that texture features combined with facial landmarks are useful for age prediction.

Keywords: Facial aging, texture features, age estimation, facial landmarks, regression analysis

1. Introduction

With the rapid development of computer technology and artificial intelligence, human biometric recognition technology has also made rapid development and attracted more and more attention. Age is an essential feature of human beings, which plays a crucial role in many real-world applications. Through modeling, the computer can judge the specific age or the age range of a person according to the visual characteristics of the face image, to realize the non-contact age estimation. It mainly solves how to estimate the exact age of the face according to the facial features. Aging prediction based on facial features has essential applications in age-based human-computer interaction, face image retrieval, filtering, public information collection, public safety, computer vision, and other systems. It has created enormous economic and social benefits. It has positive significance to improve face recognition accuracy and reduce the search range of face image. However, the process of human growth and aging is very complex, affected by many factors, and individual differences are substantial. Therefore, age estimation

*Corresponding author: Weiyang Chen, College of Computer Science and Technology, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, Shandong, China. E-mail: chenweiyang@qlu.edu.cn.

based on facial image is a complicated and essential research work in computer vision, pattern recognition, and artificial intelligence.

Generally speaking, the facial age estimation system consists of two key modules: 1) how to represent a facial image; 2) estimate the age according to the representation. Researchers proposed some methods to predict age [1]. Dornaika et al. explored the application of the depth convolution neural network (CNN) in the image-based age estimation [2]. Angeloni et al. introduced the age estimation method into unconstrained images of facial parts (eyebrows, eyes, nose, and mouth), and clipped landmarks from input images to obtain a compact multi-stream convolutional neural network (CNN) architecture [3]. Liu et al. proposed multi-task learning (MTL) network that combines classification and regression, called a cr-mt network [4]. MTL can improve the generalization performance of the age regression task by sharing information between the two studies. Dong et al. proposed the ordinal relationship between learning differences and age tags from limited training samples. Take advantage of sequential relationships' inherent properties, the learning problem is described as a structured sparse multi-class classification model. Structured sparse regularization encodes the ordinal relationship between different age tags and ensures that samples with similar age tags are close to each other in the feature space [5]. Liu et al. proposed a four-stage fusion framework for facial age estimation. This framework starts with gender recognition, then enters the second stage, gender-specific age grouping, then the third stage, age estimation within the age group, and ends in the fusion stage [6]. Taheri et al. proposed a new multi-stage age estimation system to train the features of convolutional neural network (CNN) from the universal feature extractor and accurately combined these features with the selection of age-related manual features. This is a decision level fusion of two different methods for estimating age, the first method uses feature level fusion of different hand-made local feature descriptors for wrinkles, skin and facial beauty, and the second uses the fractional fusion of different feature layers of CNN to estimate their age [7].

This paper proposes a new facial age estimation method based on texture features and facial feature points. The order of aging faces is preserved so that the compact representation can better describe the aging process from infants to the elderly. Age is an essential feature of people. In life, we can see that each age group's aging will change with time, so the study of facial aging should have a particular significance. The growth of age value is a process of continuous change of an ordered sequence, so age estimation can also be regarded as a regression problem. The regression model uses regression analysis to estimate the age by establishing a function model to represent the law of face age change. The proposed method consists of texture feature extraction, facial feature point extraction, data calibration and age prediction. We also divide the database into different age groups and predict the age group.

2. Methods

This section describes how to obtain facial texture features and facial feature points. We use the gray level co-occurrence matrix (GLCM) to extract the face's texture features, the chehra method is used to locate the feature points, and the age is predicted by regression.

2.1. Gray-level co-occurrence matrix

The basic principle of gray-level co-occurrence matrix is the elements in the GLCM, which represents the joint distribution of the gray levels of two pixels with a particular spatial position relationship. Academic interpretation refers to the process of extracting texture feature parameters by specific image processing technology to obtain the quantitative or qualitative description of the texture [8].

Texture analysis is the extraction and analysis of the spatial distribution pattern of image gray. Texture analysis is widely used in remote sensing image, X-ray photo, cell image interpretation, and processing. The texture is an area property, so it depends on the region's size and shape [9]. The analysis of texture can obtain important information about the object in the image. It is an essential means of image segmentation, feature extraction, and classification recognition.

The texture is a common visual phenomenon. At present, there are different definitions of texture. Definition 1, a repeating pattern formed by arranging elements or primitives according to specific rules. Definition 2, if a set of local attributes of an image function are constant, or slowly varying, or approximately periodic, then the corresponding region in the image has a constant texture [10]. The study of this kind of surface texture is called texture analysis. The texture is related to the local gray level and its spatial organization. It has essential applications in the field of computer vision.

GLCM of an image can reflect the comprehensive information about the direction, adjacent interval, and change the range of the gray image level. It is the basis of analyzing the local patterns of the image and their arrangement rules. GLCM represents the gray level's spatial dependence, representing the pixel gray level's spatial relationship in a texture mode. The gray level co-occurrence matrix is square, and the dimension is equal to the gray level of the image. The value of the elements (i, j) in GLCM represents the number of times when one pixel has a gray value of i , and another pixel has a gray value of j , and the adjacent distance is d direction is A . In practical application, A is generally selected as $0^\circ, 45^\circ, 90^\circ$ and 135° . Generally speaking, the gray level of a gray image is 256. When calculating the texture features derived from the GLCM, the image's gray level is required to be far less than 256 [11,12]. Therefore, before calculating the GLCM, it is necessary to specify the image's histogram to reduce the image's gray level. Generally, the gray level of the specified image is 8 or 16. Many texture features can be derived from GLCM.

2.2. Viola Jones algorithm

In computer vision, face detection or object detection has always been a very concerned field. In face detection, the Viola-Jones face detection algorithm can be said to be a very classic algorithm. Researchers engaged in face detection research will be familiar with this algorithm [13]. Viola-Jones algorithm was proposed on CVPR in 2001 because of its efficient and fast detection. Even now, it is still widely used; OpenCV and MATLAB have written this algorithm into the function library, which can be easily called directly [14].

Haar features are used to describe the common attributes of human faces. Generally speaking, the face will have some basic similarities; for example, the eye area will be much darker than the cheek area, the nose generally belongs to the high light area of the face, the nose will be much brighter than the surrounding cheek, a face image, the relative position of eyes, eyebrows, nose, mouth is regular [15]. Haar features consider the adjacent rectangular regions at a specific location. The pixels of each rectangular area are added and then subtracted. In general, it corresponds to the following situations:

$$[-1, 1], [1, -1], [1, -1, 1], [-1, 1, -1]$$

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

We used to do this for a single pixel, but now we need to focus on the rectangular area, equivalent to turning a single pixel into a rectangular area. Therefore, we need to sum each rectangular area first and then use the above operator to do the operation. This feature is called a Haar-like feature.

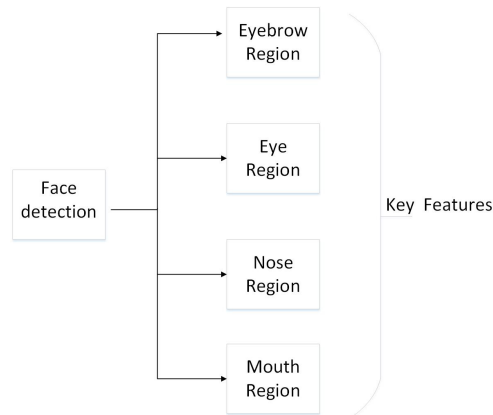


Fig. 1. Process of obtaining facial landmarks.

The rectangle area formed by an image has a large or small size. If each rectangular area is calculated by traversing all the pixels and then summing, the computational burden will be hefty. All VJ face detection algorithms use a brilliant data structure [16], called in the principle of integral image is very simple. The integral image value of any point in the image is equal to the sum of all pixels in the upper left corner of the point.

3. Results

This section introduces the detailed experiments of feature point location and texture feature calculation in a facial image. In the first experiment, the chehra method is used to locate facial feature points automatically. In the second experiment, some facial texture features were extracted. Finally, we use the feature points and texture features to predict the age. We also divide the database into 11 age groups. According to the experimental results, we conclude that combined with facial landmarks and texture features the performance of age prediction can be improved.

3.1. Feature point acquisition

To better study human beings' behavior and emotional state in sociology and clinical medicine, a study by the University of Bradford in the United Kingdom uses a computer software algorithm to detect and track a human face's key points to identify fake smiles automatically [17]. People tend to be friendlier to people who smile, but it is hard to tell whether it is true or not. The researchers first took a video of smiling expressions and then used chehra, a facial tracking software, to monitor vital facial points in real-time. Then they use a lightweight algorithm developed by them to analyze the dynamic changes of eyes, cheeks, mouth, and other parts on a time scale. Using the algorithm's output data, the researchers recorded the movement difference between the genuine smile and the fake smile [18].

As shown in Fig. 1, the computational analysis framework consists of three main components: detection, analysis and output. The detection phase includes face detection, feature point detection. The first step in our framework is to detect faces. To this end, we have used a well-known image processing algorithm called the Viola-Jones algorithm. It is based on Haar feature selection and uses AdaBoost training and cascading to create a complete image classifier. The ability of robust face detection under different

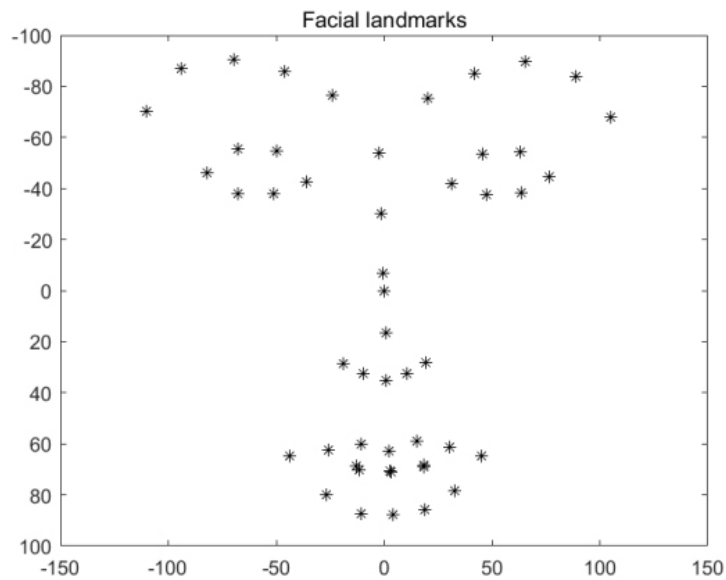


Fig. 2. Facial landmarks labeling.

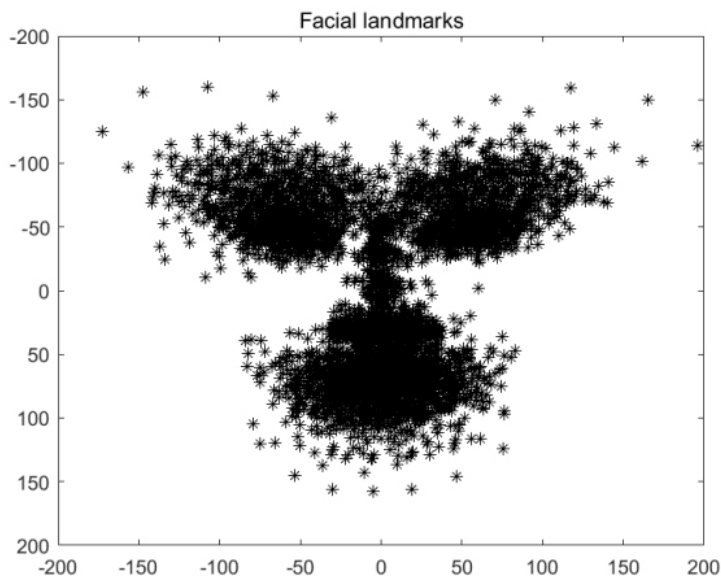


Fig. 3. All the labeled landmarks in the FG-NET database.

illumination conditions has been well established. The second step is feature point detection [19]. Figure 2 shows the example of the detected facial feature points. Figure 3 shows all the labeled feature points in the FG-NET database. We can see that the difference among facial feature points is noticeable.

3.2. Texture feature extraction

The texture feature is an important visual cue in the image. As a fundamental property of object surface,

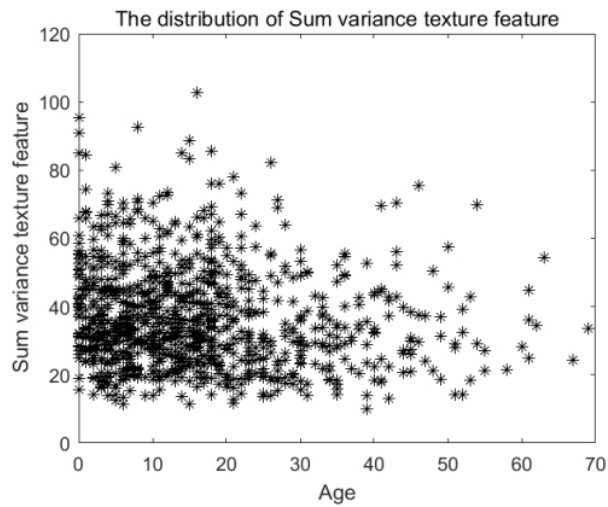


Fig. 4. The distribution of Sum variance texture feature.

texture exists widely in nature, and it is an essential feature to describe and recognize objects. An image’s texture features describe the local patterns and their arrangement rules that repeatedly appear in the image and reflect some gray-level change rules in the macro sense. The image can be regarded as the combination of different texture regions, and the texture is a measure of the relationship between pixels in the local area. Texture features can be used to describe the information in an image quantitatively. Texture analysis technology is an important research field in computer vision, image processing, image analysis and image retrieval. Applying the GLCM method to extract the features of the face image, making it easier to recognize the facial features to obtain a better visual effect than the original image, which proves its effectiveness in improving the classification accuracy [20].

First, the images in the database (FG-NET) were inputted. Then the images were preprocessed and obtained the images’ texture features. Figure 4 shows the example of the computed texture feature.

GLCM is the number of different gray pairs obtained from an image given a specific offset. In terms of a mathematical expression, when the balance is set as $(\Delta a, \Delta b)$, the corresponding GLCM:

$$C(i, j) = \sum_{p=1}^n \sum_{q=1}^s \begin{cases} 1, & \text{if } I(p, q) = i, I(p + \Delta a, q + \Delta b) = j \\ 0 & \text{else} \end{cases}$$

The parameters of the GLCM are selected as follows. For the 8-bit gray-level image, the gray value range of pixels is 0–255, and the offset parameters Δx and Δy of the GLCM are set to 0 or 1. Four angles are selected to calculate GLCM: 0° , 45° , 90° , and 135° , respectively. In this way, four GLCM G1, G2, G3, and G4 are obtained. In principle, the direction can be any value between 0° and 360° , but the texture can be represented by taking four directions of 0° , 45° , 90° , and 135° . And then we calculate some widely used texture features, including Autocorrelation, Contrast, Correlation, Cluster Prominence, Cluster Shade, Dissimilarity, Energy, Entropy, Homogeneity, Maximum probability, Sum of squares, Sum average, Sum variance, Sum entropy, Difference variance, Difference entropy, Information measure of correlation, Inverse difference normalized and Inverse difference moment normalized [21]. Several commonly used texture features are described below.

The contrast describes the relative density contrast between a pixel and a pixel with a relative position

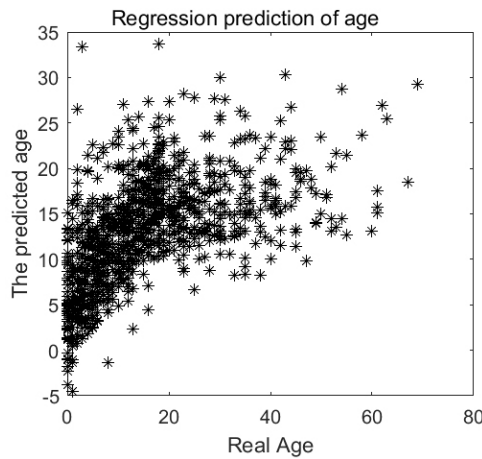


Fig. 5. Regression prediction of age based on facial feature points.

of δ .

$$\text{Contrast} = \sum_{i,j} |i - j|^2 p_{\delta}(i, j)$$

Homogeneity compares the distribution of values on the diagonal of GLCM and the values far away from the diagonal. The diagonal of GLCM represents the frequency of almost all adjacent pixels, which is the critical spatial similarity moment, which describes whether two adjacent pixels in an image are similar.

$$\text{Homogeneity} = \sum_{i,j} \frac{p_{\delta}(i, j)}{1 + |i - j|}$$

Energy is a standard moment, equal to the square of the GLCM moment:

$$\text{Energy} = \sum_{i,j} p_{\delta}(i, j)^2$$

Correlation refers to the correlation index between a pixel in the whole image and the pixels relative to its δ direction and spacing.

$$\text{Correlation} = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j) p_{\delta}(i, j)}{\sigma_i \sigma_j}$$

The formula of the mean value (μ) and variance (σ) is as follows:

$$\mu_i = i \sum_{j=1}^n p_{\delta}(i, j), \mu_j = j \sum_{i=1}^n p_{\delta}(i, j)$$

$$\sigma_i = (i - \mu_i)^2 \sum_{j=1}^n p_{\delta}(i, j), \sigma_j = (j - \mu_j)^2 \sum_{i=1}^n p_{\delta}(i, j)$$

Correlation is a descriptive statistic with a value range of $[-1, 1]$.

3.3. Age prediction

Age is an essential feature of people. Age prediction can not only protect our rights and obligations but also reduce the occurrence of age fraud. Aging degree prediction based on facial features uses computer technology is the basis of studying the law of face changing with age. Through modeling, the computer can judge the specific age or the age range of a person according to the visual characteristics of the face image, to realize the non-contact age estimation. It mainly solves how to estimate the exact age of the face according to the facial features. Aging prediction based on facial features has essential applications in age-based human-computer interaction, information push service, face image retrieval, filtering, public information collection, public safety, computer vision, and other systems. It has created enormous economic and social benefits and has positive significance to improve face recognition accuracy and narrow the search range of face images. However, the process of human growth and aging is very complex, affected by many factors, and individual differences are substantial. Therefore, age estimation based on face image is a challenging and essential research work in computer vision, pattern recognition, and artificial intelligence.

This paper uses SVM regression to predict the age based on facial landmarks and texture features [22]. For the regression model, the goal is to fit each point (x_i, y_i) in training set to a linear model $y_i = w \cdot \phi(x_i) + b$ as far as possible. The loss function of the SVM regression model is different from that of the general regression model. SVM defines a constant $\epsilon > 0$. For a certain point (x_i, y_i) , if $|y_i - w \cdot \phi(x_i) - b| \leq \epsilon$, then there is no loss at all. If $|y_i - w \cdot \phi(x_i) - b| > \epsilon$, then the corresponding loss is $|y_i - w \cdot \phi(x_i) - b| - \epsilon$. The mean square error loss function is different. If it is the mean square error, there will be a loss as long as $y_i - w \cdot \phi(x_i) - b \neq 0$. Generally speaking, the loss function measure of our SVM regression model is:

$$\text{err}(x_i, y_i) = \begin{cases} 0 & |y_i - w \cdot \phi(x_i) - b| \leq \epsilon \\ |y_i - w \cdot \phi(x_i) - b| - \epsilon & |y_i - w \cdot \phi(x_i) - b| > \epsilon \end{cases}$$

The objective function is as follows:

$$\min; \frac{1}{2} \|w\|_2^2; \text{ s.t.}; |y_i - w \cdot \phi(x_i) - b| \leq \epsilon (i = 1, 2, \dots, m)$$

Like the SVM classification model, the regression model can also add relaxation variable for each sample (X_i, Y_i) .

The main advantages of the SVM algorithm are: it is useful to solve the classification and regression problems of high-dimensional features, when the feature dimension is larger than the number of samples, it still has a good effect; only use a part of the support vector to decide hyperplane, without relying on all the data; there are a large number of kernel functions that can be used to solve various nonlinear classification and regression problems; when the sample size is not massive data, the classification accuracy is high, and the generalization ability is strong.

According to the experimental results, we can conclude that texture features combined with facial feature points are very useful for age regression prediction.

From Figs 5 and 6, we can see that facial texture features combined with facial feature points are useful for age regression prediction.

We also divided the samples of the database into 11 segments (age groups) like other studies [23,24], 0–4 years old as segment 1, 5–9 years old as segment 2, 10–14 years old as segment 3, 15–19 years old as segment 4, 20–24 years old as segment 5, 25–29 years old as segment 6, 30–34 years old as segment 7, 35–39 years old as segment 8, 40–44 years old as segment 9, 45–49 years old as segment 10, and more than 50 years old as segment 11. Figure 7 shows our prediction results based on facial texture features and facial landmarks.

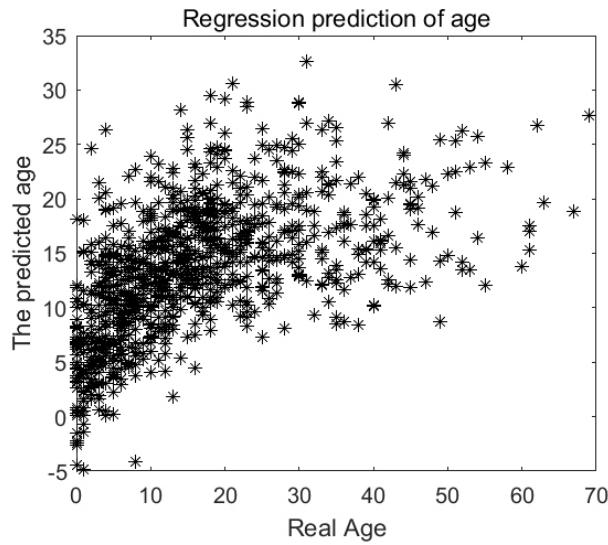


Fig. 6. Regression prediction of age based on facial landmarks and texture features.

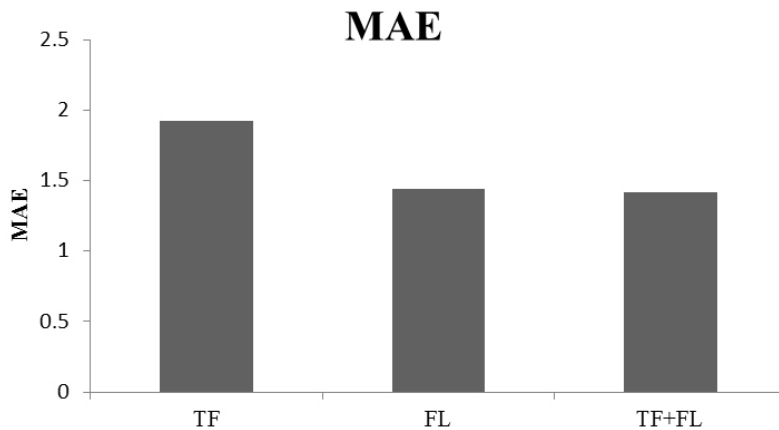


Fig. 7. Regression prediction of age groups based on facial landmarks and texture features. MAE: mean absolute error; TF: the prediction results based on the texture features; FL: the prediction results based on facial landmarks; TF + FL: the prediction results based on the texture features and facial landmarks.

According to Fig. 7, we can see that the facial morphological features contained in facial landmarks can reflect facial age better than facial texture features. Combined with facial landmarks and texture features, the performance of age prediction can be improved.

4. Discussion

In recent years, the analysis of facial age prediction has become a fascinating topic in visual computing. Age is an essential feature of people.

The growth of age is a process of continuous change of an ordered sequence, so age estimation can also be regarded as a regression problem. The regression model uses regression analysis to estimate the

age by establishing a function model to represent the law of facial age change. For the same person, the texture features of each age group will be different. Therefore, we first compute the texture feature of each image in the database. From a geometric point of view, facial feature points will change with age, so it is essential to study facial feature points. We annotate the feature points of each facial image, label the corresponding coordinates of feature point, and then use the coordinates of feature points to predict the age. After that, we use the SVM regression method to analyze the two groups of data, texture features and the coordinates of feature points, and get the predicted age. The experiment results show the prediction effect of the proposed method.

5. Conclusion

We introduce a method of facial age prediction by facial images. The first experiment studies the age prediction based on facial texture features, and the technical scheme of texture feature extraction is first to input the facial image of database (FG-NET). Then the images are preprocessed and the texture features are computed. The processing steps are as follows: firstly, the image is preprocessed, and then the image is grayed, and then the GLCM and texture features are computed. The second experiment studies the age prediction based on facial feature points. The detection phase includes face detection and facial landmark detection. The first step is to detect faces in a given image. We have used a well-known image processing algorithm called the Viola-Jones algorithm. It is based on Haar feature selection and uses AdaBoost training and cascading to create a complete image classifier. The second step is the facial feature point detection. Facial feature point detection is carried out using the chehra model, which is a machine learning algorithm for detecting facial feature points. Finally, we normalize the data set of feature points and texture features and then make a regression prediction on the data set. According to the experimental results, we can conclude that texture features combined with facial feature points are useful for age regression prediction.

Acknowledgments

We thank the FG-NET group for sharing the images. This work was supported by the National Natural Science Foundation of China – Tianyuan Fund for Mathematics (12026210 and 12026209) and the Natural Science Foundation of Shandong Province, China (ZR2017BF041).

Conflict of interest

None to report.

References

- [1] Xie JC, Pun CM. Deep and Ordinal Ensemble Learning for Human Age Estimation from Facial Images. *IEEE Transactions on Information Forensics and Security*. 2020; 15: 2361-2374.
- [2] Dornaika F, Bekhouche SE, Arganda-Carreras I. Robust regression with deep CNNs for facial age estimation: An empirical study. *Expert Systems with Application*. 2020; 141(Mar.): 1129421-112942.7.
- [3] Angeloni MDA, Pereira RDF, Pedrini H. Age Estimation from Facial Parts Using Compact Multi-Stream Convolutional Neural Networks// 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). IEEE, 2020.

- [4] Liu N, Zhang F, Duan F. Facial Age Estimation Using a Multi-Task Network Combining Classification and Regression. *IEEE Access*. 2020; 8: 92441-92451.
- [5] Dong Y, Lang C, Feng S. General structured sparse learning for human facial age estimation. *Multimedia Systems*, 2017.
- [6] Liu KH, Liua TJ. A Structure-Based Human Facial Age Estimation Framework Under a Constrained Condition. *IEEE Transactions on Image Processing*. 2019; PP(99):1-1.
- [7] Taheri S, Toygar O. On the use of DAG-CNN architecture for age estimation with multi-stage features fusion. *Neurocomputing*. 2019; 329(FEB.15): 300-310.
- [8] Kasthuri A, Suruliandi A, Raja SP. Gabor Oriented Local Order Features Based Deep Learning for Face Annotation. *International Journal of Wavelets, Multiresolution and Information Processing*. 2019; 17(5): 51-56.
- [9] Chenqiang G, Xindou LI, Fengshun Z, et al. Face Liveness Detection Based on the Improved CNN with Context and Texture Information. *Chinese Journal of Electronics*, 2019.
- [10] Mohammed AA, Xiamixiding R, Sajjanhar A, et al. Texture features for clustering based multi-label classification of face images// 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). 2017.
- [11] Prasad MVNK. GLCM Based Texture Features for Palmprint Identification System// International Conference on Computational Intelligence in Data Mining. 2014.
- [12] Songpan W. Improved Skin Lesion Image Classification Using Clustering with Local-GLCM Normalization// 2018 2nd European Conference on Electrical Engineering and Computer Science (EECS). IEEE, 2019.
- [13] Peleshko D, Soroka K. Research of Usage of Haar-like Features and AdaBoost Algorithm in Viola-Jones Method of Object Detection// 2013 12th International Conference on the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM). IEEE, 2013.
- [14] Acasandrei L, Barriga A. Accelerating Viola-Jones face detection for embedded and SoC environments// Fifth Acm/ieee International Conference on Distributed Smart Cameras. ACM, 2011.
- [15] Chaudhari MN, Deshmukh M, Ramrakhiani G, et al. Face Detection Using Viola Jones Algorithm and Neural Networks// 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA). 2018.
- [16] Raya IGNMK, Jati AN, Saputra RE. Analysis realization of Viola-Jones method for face detection on CCTV camera based on embedded system// 2017 International Conference on Robotics, Biomimetics, and Intelligent Computational Systems (Robionetics). 2017.
- [17] Asthana, Zafeiriou S, Cheng S, Pantic M. Incremental Face Alignment in the Wild, 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, 2014, pp. 1859-1866, doi: 10.1109/CVPR.2014.240.
- [18] Radlak K, Radlak N, Smolka B. Static Posed Versus Genuine Smile Recognition. 2017.
- [19] Wang JX. Facial feature points detecting based on Gaussian Mixture Models. *Pattern recognition letters*, 2015.
- [20] Ou X, Pan W, Xiao P. In vivo skin capacitive imaging analysis by using grey level co-occurrence matrix (GLCM). *International Journal of Pharmaceutics*. 2014; 460(1-2): 28-32.
- [21] Haralick RM, Shanmugam K, Dinstein I. Dinstein, Textural Features for Image Classification. *IEEE Transactions on Systems Man and Cybernetics*. 1973; 3(6): 610-621.
- [22] Chen F, Zhang Z, Yan D. Image classification with spectral and texture features based on SVM// International Conference on Geoinformatics. IEEE, 2010.
- [23] Li C, Liu Q, Dong W, Zhu X, Liu J, Lu H. Human Age Estimation Based on Locality and Ordinal Information. *IEEE Transactions on Cybernetics*. 2015; 45(11): 2522-2534.
- [24] Jhang K, Cho J. CNN Training for Face Photo based Gender and Age Group Prediction with Camera// International Conference on Artificial Intelligence in Information and Communication. Dept. of Computer Engineering, Chungnam National University, Daejeon, 34134, Korea; Dept. of Computer Engineering, Chungnam National University, Daejeon, 34134, Korea, 2019.