

Research Article

Age Estimation of Face Images Based on CNN and Divide-and-Rule Strategy

Haibin Liao ¹, Yuchen Yan ², Wenhua Dai,¹ and Ping Fan¹

¹*School of Computer Science and Technology, Hubei University of Science and Technology, Xianning, China*

²*School of Electronic Information, Wuhan University, Wuhan, Hubei, China*

Correspondence should be addressed to Yuchen Yan; yyc@whu.edu.cn

Received 7 March 2018; Revised 28 April 2018; Accepted 7 May 2018; Published 5 June 2018

Academic Editor: Krzysztof Okarma

Copyright © 2018 Haibin Liao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, the research on age estimation based on face images has drawn more and more attention, which includes two processes: feature extraction and estimation function learning. In the aspect of face feature extraction, this paper leverages excellent characteristics of convolution neural network in the field of image application, by using deep learning method to extract face features, and adopts factor analysis model to extract robust features. In terms of age estimation function learning, age-based and sequential study of rank-based age estimation learning methods is utilized and then a divide-and-rule face age estimator is proposed. Experiments in FG-NET, MORPH Album 2, and IMDB-WIKI show that the feature extraction method is more robust than traditional age feature extraction method and the performance of divide-and-rule estimator is superior to classical SVM and SVR.

1. Introduction

In the medical profession, people rely mainly on the analysis of cholesterol, high density cholesterol, albumin, and other blood test indicators to determine a person's "physiological age" and to study the degree of human aging. Unfortunately, this set of indicators is still very imperfect and of great inconvenience to use. If we can use computer and image processing technology to analyze facial images to accurately predict a person's "physiological age" and compare "physiological age" and "actual age", we can know whether you are in "Youth Permanent" or "Premature Aging". Then it will greatly improve research efficiency and reduce research costs. By looking at the "face" to estimate the age, it can not only be used to quantify the aging, but also be applied to smart city and safe city construction. In everyday language conversation, the content and manner of conversation are often affected by other factors such as gender and age. For example, when faced with the elderly, the language of conversation will obviously be formal. More generally, humans quickly estimate each other's gender, age, and identity through the appearance of the other person's face in order to select different social styles.

2. Related Work

Compared with face recognition, age analysis has received much less attention. However, it does not mean that age analysis is not as important as face identification. The use of face age analysis is always involved in the application in face recognition [1, 2] and forensic science [3, 4]. In recent years, face age analysis has attracted more interests and attention from psychology, aesthetics, forensic science, computer graphics, computer vision [5–7], and so on. There is no doubt that the coming applications of automatic age analysis will not just include face recognition, but will be widely used in intelligent ad delivery, bioassays science, biostatistics, electronic customer relationship management, recognition, beauty, law enforcement, security controls demographic census, human-computer interaction, and other fields.

The two main issues in age estimation are "how face images are represented and how ages are estimated?" Namely, they are face representation and estimation function learning. Most of the current estimation functions learning methods are regarded as multiclass classification [8, 9] or regression problems [10, 11] to learn the classifier. The multiclass classification method takes the age value as a separate category

and then learns a classifier for age classification. Therefore, a number of standard classic methods such as K-NN, multilevel perception, AdaBoost, and SVM can be used to perform accurate age prediction and age grouping. The regression method fits the mapping of feature space to age space mainly by regularization method to get the best regression function. Typical nonlinear regression methods are quadratic regression, Gaussian Process, and Support Vector Regression (SVR), which are used for age estimation problems. The classification method only considers the independence of different age categories, but ignores the internal relations between them. In fact, it has a strong internal order between the ages. For example, for a 15-year-old sample, the age is more likely to be 14 or 16 years, rather than 10 or 20 years. Although the regression method considers the connection between different age categories, it views age category as a linearly increasing relationship and does not reflect the diversity of the aging process. For example, in infancy age, aging is mainly reflected in the changes of bones, while in adults it is mainly the changes in texture.

Since the classification method naively treats the age category as an independent label, the regression method simply regards the age category as equilibrium. Therefore, as alternatives to classification and regression methods, recent rank-based (or ordinal regression) methods have been used for age estimation [12, 13]. This approach uses the age-axis strategy for age-class estimation, which is suitable for age estimation because the relative order of age is properly used.

As people of different ages have different age growth, we can divide the training data according to age and use different submodels to predict the age data of different ages. Then, the penalty parameters are weighted to solve the problem of classification bias caused by the imbalance of the sample number, which can make the accuracy of age estimation further improved. In this paper, we will study the rank-based age estimation learning method and propose a divide-and-rule age estimation framework.

Face features extraction is another important problem in image age estimation. In order to extract features of human face, early research is mainly focused on the face geometry ratio feature extraction [14]. In order to extract more detailed features of facial features, active appearance model (AAM) combined with principal component analysis (PCA) gradually replaces the face geometry characteristics of main features for age estimation [15, 16]. Since each test face image has been labelled as a specific age value, this AAM-based age estimation method is more suitable for higher-precision age estimation system. However, the AAM method has some shortcomings because AAM is mainly based on shape and gray scale, in which the average global feature of the training image is extracted, and the characterization of some texture information is not effective for some faces.

All of the above methods require accurate key feature point detection and location technology. However, it is difficult to meet the requirements in practical application. Therefore, with reference to the feature extraction method in face recognition, the current mainstream methods directly extract the face feature vector from human face. For example, Ahonen et al. [17] and Zhou et al. [18] used local binary

model (LBP) and Haar-like wavelet transform to extract face features for age estimation, respectively. In addition, important features such as flow-type learning [19, 20] feature-based dimensionality reduction methods in face recognition are also used to construct face age features. One of the most effective methods for face feature extraction is Bioinspired Features (BIF) [21]. This method imitates the single cell acceptability information domain of the primary visual cortex of vertebrate brain. Firstly, the Gabor filter with different scales and directions is used to convolute the face to extract face features. Then, the extracted feature vector in first step is merged. However, this method tends to generate local transform invariance and reduce texture detail information when combining Gabor convolution coefficients in second step [22].

More recently, CNN-based methods [23–30] have been widely adopted for age estimation due to its superior performance over existing methods. Yi et al. [26] introduced a multitask learning method with a relatively shallow CNN and achieved good results; however, its multiscale network relies on accurate feature point positioning technology, which limits its application. Rothe et al. [25] crawled a largest public dataset (IMDB-WIKI dataset) for age prediction and tackled the estimation of apparent age with CNN using the VGG-16 architecture. In [27], a six-layer CNN was used for gender and age grouping, but it did not address the problem of age estimation. In [24], the six-layer CNN is used to extract the age feature of face, and the manifold learning method is used to estimate the age. Instead of multiclass classification and regression methods, ranking techniques were introduced to the problem of age estimation. In [29], Niu et al. proposed to formulate age estimation as an ordinal regression problem with the use of multiple outputs CNN. In [30], Chen et al. proposed a Convolutional Neural Network- (CNN-) based framework, ranking-CNN, for age estimation. In this paper, we will extract age feature from AgeNet network according to [23] and then use the age-classifying strategy of divide-and-rule to estimate age.

3. Face Age Feature Extraction

3.1. Face Age Description Based on Deep Learning. Since deep CNN has been proved to be rich in scene and target representation ability, it has intelligent and automatic performance compared with the traditional manual feature extraction method. In this paper, the age network (AgeNet) in [23] is used to extract the face age descriptor and then the age estimation is carried out by using the divide-and-rule strategy.

The AgeNet uses an approach based on regression and classification to construct an age-estimated deep CNN. In order to reduce the complexity of the network and training time, we only use the regression-based age estimation deep CNN. Therefore, the following Euclidean loss function can be defined:

$$E(W) = \frac{1}{2N} \sum_{i=1}^N \|\bar{y}_n - y_n\|_2^2 \quad (1)$$

where W is the parameter of deep convolution network, \bar{y}_n is the age prediction value by the deep convolution network,

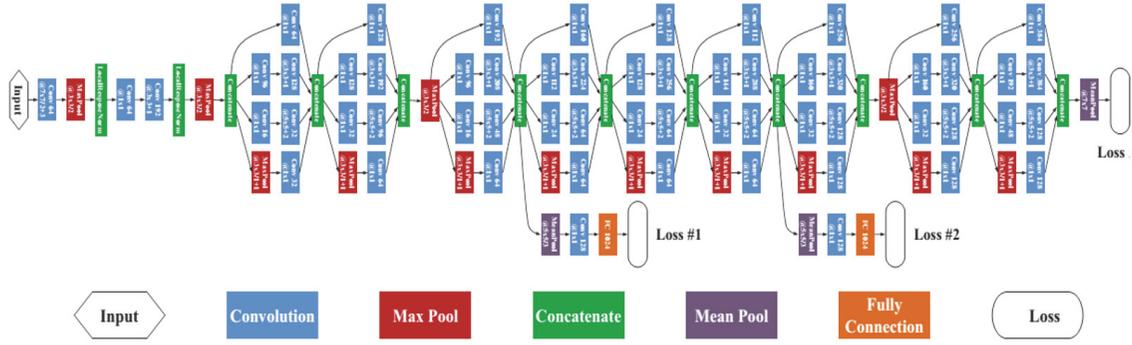


FIGURE 1: GoogLeNet's Deep CNN Hierarchy.

y_n is the actual age value, and N is the batch number. The deep network model is based on GoogLeNet [22], designed by Google et al. [31], which can improve the performance under the same computing cost. It is a very effective network design model, as shown in Figure 1.

In order to further speed up the convergence rate, before the ReLU (activation function) operation, we add the layer of scale normalization (Batch Normalization, BN), while removing all the dropout operation. In order to reduce the risk of overfitting, the training of deep CNN is carried out by means of network migration learning. Thus, the network training consists of two stages: pretrain based on the face identity data set and fine-tune based on the age data set.

Pretrain stage: Although face identity and age are two different concepts, they are related. In addition, there are many public large-scale face identity databases. Therefore, we use the large-scale face database CASIA-WebFace [32] to pretrain the network. Experiments show that the method of network pretraining in face database is better than that of random network initialization.

Fine-tune phase: This stage uses the CACD [33], Morph-II [34], and WebFaceAge [35] face age libraries to fine-tune the network model generated by the pretrain stage to produce a robust age-deep network.

All training and test face images are normalized to 256×256 sizes. In the training stage, the data is augmented. That is, images of 227×227 size are randomly cut out from the images of 256×256 size, and the images are enlarged. In the test, for the input image of 256×256 size, according to the four corners and center point of the image, we cut out five images of 227×227 size; then, all the cropped images are horizontally inverted to generate the corresponding image, a total of 10 images, respectively, into the depth of network learning. Finally, the output of last layer of GoogLeNet is used as the face age descriptor, and features of the 10 images are concatenated to form the final face feature vector.

3.2. Feature Reduction Based on Factor Analysis. Since the facial feature vectors obtained by deep CNN network learning have high dimensionality ($10 * 100 = 1000$ dimensions), both redundancy and noise are inevitable in this process. In order to extract the essence of the discriminative compact feature subset, we regard each age value as a class and use

factor analysis model to deal with original feature dimension reduction.

In the factor analysis model (FAM) [36, 37], it is desirable to find the optimal projection matrix to minimize the difference in style between homogeneous (same age) samples and to maximize the difference in content between different classes (different ages). We consider the features reflecting face of the characteristics of age changes as a content factor, which is defined as follows:

$$F_c = \sum_{i=1}^{C-1} \sum_{j=i+1}^C p_i p_j w(i, j) (\mu_i - \mu_j) (\mu_i - \mu_j)^T \quad (2)$$

where C is the total number of classes, p_i is the prior probability of classes i , μ_i is the average of the original feature vectors of the class i th, and $w(i, j) = 1/(ed)^2$ (ed is Euclidean distance between two classes) are the weights of the i th and the j th class. We consider the change of face illumination and expression as a style factor, which is defined as follows:

$$F_s = \sum_{i=1}^C p_i \sum_{j=1}^{N_i-1} \sum_{k=j+1}^{N_i} w(j, k) (x_j^{(i)} - x_k^{(i)}) (x_j^{(i)} - x_k^{(i)})^T \quad (3)$$

where $x_j^{(i)}, x_k^{(i)}$ are the j -th and k -th feature vectors of class i , respectively, and $w(j, k)$ is the weight of the j -th and k -th vector defined as

$$w(j, k) = \exp\left(-\frac{\|x_j - x_k\|^2}{2t^2}\right) \quad (4)$$

where t is the experience parameter.

Given the weighted F_c, F_s , for FAM, the objective function of the linear factors analysis model is represented as follows:

$$J_F(W) = \arg \max_w \frac{\text{tr}(W^T F_c W)}{\text{tr}(W^T F_s W)} \quad (5)$$

where W is the unresolved linear mapping transformation, which can be obtained by solving the generalized eigenvalue problem:

$$F_c W = \lambda F_s W. \quad (6)$$

Then, we can extract low-dimensional, simplified, and discriminative features \hat{x} from the high-dimensional, redundant, disturbed, and distorted original feature x :

$$\hat{x} = W^T x. \quad (7)$$

By FAM, we finally get the 200 dimensional eigenvectors.

4. Divide-and-Rule Age Estimation Function Learning

After obtaining the feature vector of face image age, the age estimation function can be learned and the corresponding age estimator can be trained. When we compare the two face ages, it is easy for humans to identify which is older, but to accurately guess the age of the face is not so easy. When inferring a person's exact age, we may compare the input face with the faces of many people whose ages are known, resulting in a series of comparisons, and then estimate the person's age by integrating the results. This process involves numerous pairwise preferences, each of which is obtained by comparing the input face to the faces in the dataset. Based on this, we propose a divide-and-rule age estimation function learning scheme.

4.1. Divide. Given the training set of images $I = [I_1, I_2, \dots, I_i, \dots, I_M]$, suppose $x_i \in R^D$ is a face feature vector (obtained by the method of the previous section) and K is the total number of categories (age tags) for which the corresponding age category $y_i \in \{1, 2, \dots, K\}$ is to be treated as order ranks. Suppose, for a given age k , X can be divided into the following two subsets X_k^+ and X_k^- :

$$\begin{aligned} X_k^+ &= \{(x_i, +1) \mid y_i > k\} \\ X_k^- &= \{(x_i, -1) \mid y_i < k\}. \end{aligned} \quad (8)$$

Based on the above two subsets, we can train two-class classifier that only answers the problem of "whether this face is older than k ". Because it only needs to make a yes or no answer in the classification process, the complexity of the problem is reduced. In addition, because $X_k^+ \cup X_k^- = X$, sufficient training samples are ensured for each classification, which overcomes the problem of lack of training samples in classifiers such as parallel-hyperplanes or one-versus-one classifiers (this problem is more prominent in age estimation).

4.2. Rule. Once more than one classifier has been trained, the results of all the two-class classifiers can be aggregated to estimate the age rank as with the multiple hyperplane classification methods. Since there is an inconsistency in all of the two-class classifiers, a classification function $f_k(x)$ with discriminating performance is defined for each of the two-class classifiers. The age rank estimation problem (i.e., the age estimate) becomes

$$r(x) = 1 + \sum_{k=1}^{K-1} \langle f_k(x) > 0 \rangle. \quad (9)$$

In this case, a logical decision is made by $\langle \bullet \rangle$, and if the internal condition is satisfied, it returns 1 and vice versa. Since each classification function uses its own feature space, it is easier to obtain the optimal solution than the method using uniform feature space.

4.3. Categorical Function Learning Based on the Sensitive Cost Function. The estimation of the final age rank in the divide-and-rule age estimation scheme relies on a series of easily classifiable two-class classifiers. Therefore, it is very important to design an appropriate cost function of missed classification for each classifier. In the case of age estimation, the error rate based on 0-1 loss is not of interest. In the 0-1 loss scheme, all missed classifications are given the same price, which is clearly not in line with the reality of the age estimates. For example, when a person's age is k , the 0-1 loss does not identify the miss classification between $k + 1$ and $k + 20$ classes. Obviously, the latter is more serious. Therefore, we study the classification function based on the sensitive cost function.

Given the training set X_k^+ and X_k^- of the k th class, we denote the cost function $\text{cost}_k(l)$ of the missed classification of the age category l th in the subclass of the k th class. And $C(e_i) : R^+ \cup \{0\} \rightarrow R^+ \cup \{0\}$ is recorded as the cost that has occurred, in which $e_i = \|\tilde{l}_i - l_i\|$ indicates the absolute error denoted between the estimated age and the real age. The two most commonly used evaluation criteria in the synthetic age estimation are the mean absolute error (MAE) and the cumulative integral (Cumulative Score, CS) of the absolute error. We define the costs that have already occurred as follows:

$$C(e_i) = \langle e_i > L \rangle \cdot (e_i - L) \quad (10)$$

where L is the acceptable tolerance for the age of the error; you can take 3-5. Therefore, the cost function of the i th sample x_i in the k th subproblem in missed classification is expressed as

$$\text{cost}_k(x_i) = \begin{cases} C(k - y_i + 1), & y_i \leq k \\ C(y_i - k), & y_i > k. \end{cases} \quad (11)$$

With the missed classification cost function, it can be used as an important weight to adjust the data for training of the two-class classifier. In this paper, SVM is used to train the two-class classifier in the subproblems, and the hinge loss is adjusted by using the missed classified cost function as weight item:

$$\begin{aligned} \min_{w_k, b_k, \zeta} \quad & \frac{1}{2} \langle w_k, b_k \rangle + C \left[\sum_i \text{cost}_k(x_k) \zeta_i \right] \\ \text{s.t.} \quad & \dots s_k(i) (w_k^T \phi_k(x_i) + b_k) \geq 1 - \zeta_i, \\ & \zeta_i \geq 0, \forall i \end{aligned} \quad (12)$$

where $s_k(i) = +1$ if $x_i \in X_k^+$ and $s_k(i) = -1$ if $x_i \in X_k^-$, ϕ_k denotes the mapping from feature space to Hilbert space, and $\langle w_k, b_k \rangle$ is hyperplane parameters of the hidden space. For age estimation problems, a single kernel does not apply to

all subproblems. And (12) selects different cores according to the characteristics of the feature space in each subproblem, so that each subproblem can be well projected. The discriminant classification function based on the sensitive cost function eventually becomes

$$f_k(x) = \langle w_k, \phi_k(x) \rangle + b_k. \quad (13)$$

5. Experiments and Analysis

In this section, three common age groups, FG-NET [38], MORPH Album 2 [34], and IMDB-WIKI [39], are used to test the validity of the proposed face feature extraction algorithm and face age estimation function. The FG-NET face database consists of 1002 images of 82 different individuals with different expressions, illumination, and attitude changes. Each one has 6 to 18 images of different ages, ranging from 0 to 69 years old, and FG-NET face age database is one of the most commonly used public databases. The MORPH2 database contains a total of 55,000 pictures of 13,000 volunteers aged 16 to 77 years, 45,000 of which are used for network training, and the remaining 10,000 are used for testing. IMDB-WIKI is the largest publicly available dataset of face images with gender and age labels for training and testing. In total there are 460,723 face images from 20,284 celebrities from IMDb and 62,328 from Wikipedia, being thus 523,051 in total.

The age estimation indexes we used are mean absolute error (MAE) and cumulative index (Cumulative Score, CS), and the expressions are as follows:

$$MAE = \frac{\sum_{k=1}^N |s_k - \tilde{s}_k|}{N} \quad (14)$$

where s_k is the actual age, \tilde{s}_k is the estimated age, and N is the total number of pictures for testing.

$$CS(L) = \frac{N_{e < L}}{N} * 100\% \quad (15)$$

where $N_{e < L}$ denotes the number of test images whose absolute error is not larger than the set value.

5.1. Experiment Based on FG-NET. In this section, we compare the novel face feature extraction method based on deep learning (DLF + FAM) with Gabor + PCA, LBP + PCA, Gabor + LBP + PCA, and BIF + PCA, which are the common face feature extraction methods. The effectiveness of the feature extraction and feature-based multidimensional reduction algorithm is analyzed. The age estimation function we use is the classical SVR. The face database was trained and tested by the Leave-One-Person-Out (LOPO) strategy, which was used to test all age images of the same person as the test set and all the others as training. The test procedure tests each sample and does not overlap with the training set.

Table 1 shows the comparison of MAE values of different feature extraction methods under SVR estimator. Figure 2 shows the CS comparison of different feature extraction methods under SVR estimator. It can be seen that the DLF features adopted in this paper are obviously superior to

TABLE 1: Comparison of MAE values for different feature extraction methods.

Methods	MAE/Age
Gabor+PCA	5.53
LBP+PCA	5.92
Gabor+ LBP+PCA	4.96
BIF+PCA	4.77
DLF+ PCA	4.57
DLF+FAM	4.26

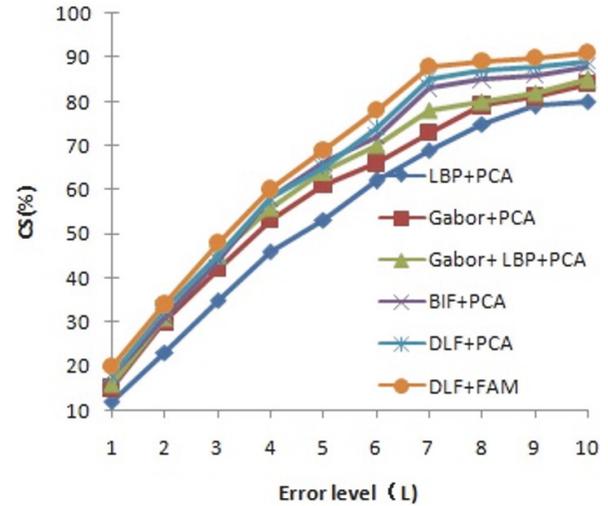


FIGURE 2: CS comparison of different feature extraction methods.

TABLE 2: Comparison of MAE values for different estimators.

Estimators	Feature	MMMAE/Age
SVM	BIF + FAM	5.83
SVR	BIF + FAM	4.58
Our Proposed	BIF + FAM	4.32
SVM	DLF +FAM	4.77
SVR	DLF +FAM	4.26
Our Proposed	DLF +FAM	4.02

the traditional manual features. In order to validate the effectiveness of the method based on factor analysis, this paper gives the comparison of DLF + PCA and DLF + FAM. It can be seen that FAM is better than PCA. This is because FAM makes use of the category information of age tags.

In order to verify the effectiveness of the proposed divide-and-rule age estimation function, BIF and DLF face features with better performance are selected in this experiment, and the divide-and-rule method is compared with classical SVM and SVR methods. The comparison results are shown in Table 2 and Figure 3 (using the DLF feature), and we can see that the age estimator based on divide-and-rule method is about 1 year better than the classic SVM and SVR.

5.2. Experiment Based on MORPH. Due to the small amount of FG-NET face databases, recent studies have shown that

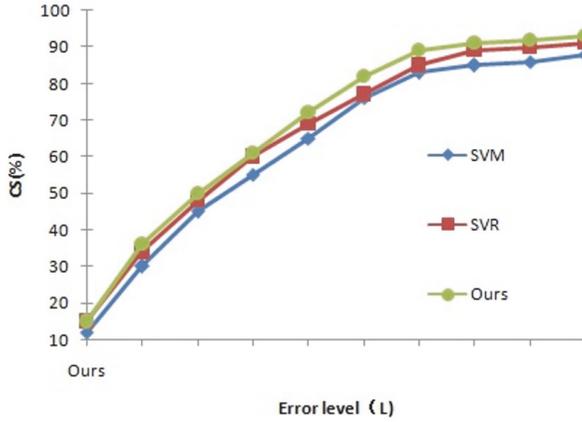


FIGURE 3: CS comparison of different estimators.

TABLE 3: Comparison of MAE values for different estimators.

Estimators	Feature	MMMAE/Age
SVM	BIF + FAM	6.02
SVR	BIF + FAM	4.62
Our Proposed	BIF + FAM	3.77
SVM	DLF+FAM	5.14
SVR	DLF+FAM	4.17
Our Proposed	DLF+FAM	3.48

the corresponding algorithm in this database has very limited space for improvement. In order to further verify the effectiveness of this algorithm, this section will be further experimented on the MORPH Album 2 face database. 10,000 images were further divided, 80% of which were randomly selected for the training of the estimator, and the remaining 20% were used for test. The average of 10 experiments was taken as the final result.

The experimental results of different feature extraction methods are shown in Figure 4 and Table 3. Figure 4 is the age estimation comparison under different features using SVR estimator. Table 3 shows the comparison of the age estimation using the better BIF and DLF features comparison. It can be seen that the feature extraction method proposed in this paper is nearly 3 years better than the Gabor method and nearly 1 year better than the BIF. Using this estimator is nearly 2 years better than SVM and nearly 1 year better than SVR. Figure 5 shows the comparison results of different estimators under the DLF + FAM feature, which further confirms the validity of the proposed estimator.

5.3. Evaluation on Face in the Wild. Experiments on faces in the wild were conducted to demonstrate the performance of different approaches on real-world data. The experiments on IMDB-WIKI [39] were carried out as follows. For the task to be equally discriminative for all ages, we equalize the age distribution; i.e., we randomly ignore some of the images of the most frequent ages. This leaves us with 260,282 face images for our experiment. Further, the 260,000 images are

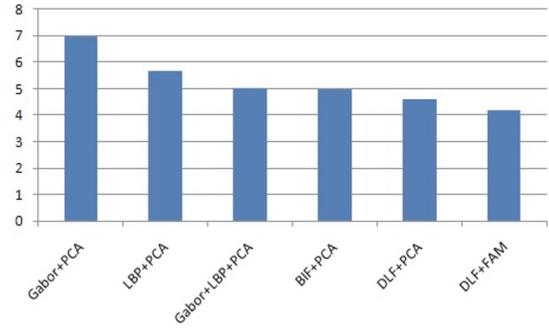


FIGURE 4: Comparison of MAE values of different feature extraction methods.

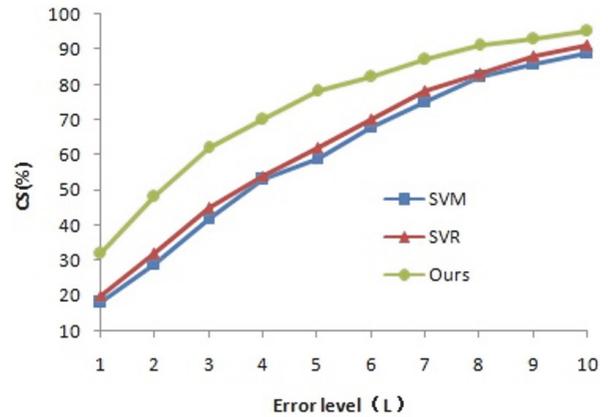


FIGURE 5: CS comparison of different estimators.

split into 200,000 for training, 30,000 images for validation, and 30,282 images for testing.

In order to verify the advancement of this method, we compare the proposed method with the most advanced method based on deep learning [23–25]. In [23], a deep network based on regression and classification is adopted. In [24], a six-layer deep network is used to extract the age feature of the face, and then the manifold learning method is used to estimate the age. In [25], a DEX network using the VGG-16 architecture is proposed, which poses the age regression problem as a deep classification problem followed by a softmax expected value refinement. All the methods use the general-to-specific deep transfer learning scheme for deep network training, which includes two stages, i.e., pretrain with face identities and fine-tune with real age.

Stage 1: we firstly employ the large-scale face identities database CASIA-WebFace [30] to pretrain the deep network, which is much better than random initialization.

Stage 2: we employ IMDB-WIKI training set to fine-tune the deep network from stage 1.

Finally, we employ IMDB-WIKI testing set (30,282 images) for age prediction.

The results of MAE comparisons of different methods on the IMDB-WIKI face database are shown in Table 4. It can be seen from Table 4 that the effect of using 16,22-layer network

TABLE 4: Comparison of different methods based on deep learning.

Methods	MAE/Age
[23]	3.93
[24]	4.22
[25]	3.35
Our Proposed	3.29

is better than that of 6-layer network, which indicates that face age is a complex nonlinear variation problem. Although the deep network in this paper is designed based on [23], the final estimation effect is better than that in [23], which shows the validity of the divide-and-rule strategy proposed in this paper.

6. Conclusion

In this paper, a robust face age feature extraction method is proposed based on the superior image representation ability of depth convolution neural network. In the aspect of feature dimensionality, the feature reduction method based on FAM is used to replace the PCA dimensionality reduction method. In the estimator learning, an ordinal regression problem is transformed into a series of binary classification subproblems, which are collectively solved with the proposed divide-and-rule learning algorithm. By taking the ordinal relation between ages into consideration, it is more likely to get smaller estimation errors, compared with multiclass classification approaches. Experimental results show that the method proposed in this paper is more discriminative and robust than traditional Gabor, LBP, and BIF methods. The performance of the proposed age estimator is superior to SVM and SVR.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported partially by the National Natural Science Fund (No. 61373108), the Hubei Provincial Natural Science Foundation of China (No. 2017CFB168 and No. 2015CFC778), the Hubei Provincial Education Office Science and Technology Research Project, youth talent project (No. Q20172805), the Hubei Provincial Education Science Planning Project (No. 2016GB086), and the Construction of a Special Scientific Research Project for Master's Point (No. 2018-19GZ050).

References

- [1] A. Lanitis and C. J. Taylor, "Towards automatic face identification robust to aging variation," in *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2000*, pp. 391–396, fra, March 2000.
- [2] Z. Li and A. K. Jain, "A discriminative model for age invariant face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1028–1037, 2011.
- [3] N. Dayan, *Skin aging handbook: An Integrated Approach to Biochemistry and Product Development*, Andrew William Press, 2008.
- [4] A. Amadasi, N. Merusi, and C. Cattaneo, "How reliable is apparent age at death on cadavers?" *International Journal of Legal Medicine*, vol. 129, no. 4, pp. 1437–1596, 2015.
- [5] S. B. M. Patzelt, L. K. Schaible, S. Stampf, and R. J. Kohal, "Software-based evaluation of human age: A pilot study," *Journal of Esthetic and Restorative Dentistry*, vol. 27, no. 2, pp. 100–106, 2015.
- [6] W. Chen, W. Qian, G. Wu et al., "Three-dimensional human facial morphologies as robust aging markers," *Cell Research*, vol. 25, no. 5, pp. 574–587, 2015.
- [7] Y. Fu, G. Guo, and T. S. Huang, "Age synthesis and estimation via faces: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 11, pp. 1955–1976, 2010.
- [8] C.-C. Wang, Y.-C. Su, C.-T. Hsu, C.-W. Lin, and H. Y. M. Liao, "Bayesian age estimation on face images," in *Proceedings of the Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '09)*, pp. 282–285, July 2009.
- [9] X. Geng, Z. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2234–2240, 2007.
- [10] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang, "Image-based human age estimation by manifold learning and locally adjusted robust regression," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1178–1188, 2008.
- [11] Y. Zhang and D.-Y. Yeung, "Multi-task warped Gaussian process for personalized age estimation," in *Proceedings of the IEEE Conference on Computer Vision Pattern Recognition*, pp. 2622–2629, 2010.
- [12] J. Lu and Y.-P. Tan, "Ordinary preserving manifold analysis for human age and head pose estimation," *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 2, pp. 249–258, 2013.
- [13] K.-Y. Chang and C.-S. Chen, "A learning framework for age rank estimation based on face images with scattering transform," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 785–798, 2015.
- [14] Y. H. Kwon and N. da Vitoria Lobo, "Age classification from facial images," *Computer Vision and Image Understanding*, vol. 74, no. 1, pp. 1–21, 1999.
- [15] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Toward automatic simulation of aging effects on face images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 442–455, 2002.
- [16] A. Martinez and S. Du, "A model of the perception of facial expressions of emotion by humans: research overview and perspectives," *Journal of Machine Learning Research*, vol. 13, no. 1, pp. 1589–1608, 2012.
- [17] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.
- [18] S. K. Zhou, B. Georgescu, X. S. Zhou, and D. Comaniciu, "Image based regression using boosting method," in *Proceedings of the*

- 10th IEEE International Conference on Computer Vision (ICCV '05), pp. 541–548, October 2005.
- [19] G. Guo, Y. Fu, T. S. Huang, and C. R. Dyer, “Locally adjusted robust regression for human age estimation,” in *Proceedings of the IEEE Workshop on Applications of Computer Vision (WACV '08)*, pp. 1–6, IEEE, Copper Mountain, Colo, USA, January 2008.
- [20] Y. Fu and T. S. Huang, “Human age estimation with regression on discriminative aging manifold,” *IEEE Transactions on Multimedia*, vol. 10, no. 4, pp. 578–584, 2008.
- [21] G. Guo, G. Mu, and Y. Fu, “Human age estimation using bio-inspired features,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW '09)*, pp. 112–119, IEEE, June 2009.
- [22] S. Mallat, “Group invariant scattering,” *Communications on Pure and Applied Mathematics*, vol. 65, no. 10, pp. 1331–1398, 2012.
- [23] X. Liu, S. Li, M. Kan et al., *AgeNet: Deeply Learned Regressor and Classifier for Robust Apparent Age Estimation*, ICCV '16, 2016.
- [24] X. Wang, R. Guo, and C. Kambhamettu, “Deeply-Learned Feature for Age Estimation,” in *Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 534–541, Waikoloa, HI, USA, January 2015.
- [25] R. Rothe, R. Timofte, and L. V. Gool, “DEX: Deep EXpectation of Apparent Age from a Single Image,” in *Proceedings of the 2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pp. 252–257, Santiago, Chile, December 2015.
- [26] D. Yi, Z. Lei, and S. Z. Li, “Age Estimation by Multi-scale Convolutional Network,” in *Computer Vision – ACCV 2014*, vol. 9005 of *Lecture Notes in Computer Science*, pp. 144–158, Springer International Publishing, 2015.
- [27] G. Levi and T. Hassner, “Age and gender classification using convolutional neural networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPRW 2015*, pp. 34–42, usa, June 2015.
- [28] Y. Dong, Y. Liu, and S. Lian, “Automatic age estimation based on deep learning algorithm,” *Neurocomputing*, vol. 187, pp. 4–10, 2016.
- [29] Z. Niu, M. Zhou, L. Wang, X. Gao, and G. Hua, “Ordinal Regression with Multiple Output CNN for Age Estimation,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4920–4928, Las Vegas, NV, USA, June 2016.
- [30] S. Chen, C. Zhang, M. Dong, J. Le, and M. Rao, “Using Ranking-CNN for Age Estimation,” in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 742–751, Honolulu, HI, July 2017.
- [31] C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '15)*, pp. 1–9, Boston, Mass, USA, June 2015.
- [32] D. Yi, Z. Lei, S. Liao, and S. Li, “Learning face representation from scratch,” <https://arxiv.org/abs/1411.7923>.
- [33] B.-C. Chen, C.-S. Chen, and W. H. Hsu, “Cross-age reference coding for age-invariant face recognition and retrieval,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 8694, no. 6, pp. 768–783, 2014.
- [34] K. Ricanek and T. Tesafaye, “MORPH: A Longitudinal Image Database of Normal Adult Age-Progression,” in *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR '06)*, pp. 341–345, 2006.
- [35] B. Ni, Z. Song, and S. Yan, “Web image and video mining towards universal and robust age estimator,” *IEEE Transactions on Multimedia*, vol. 13, no. 6, pp. 1217–1229, 2011.
- [36] G. Zhile, Z. Shaolong, and L. Haibin, “Linear dimensionality reduction method for factor analysis criterion,” *Computer Engineering and Applications*, vol. 52, no. 3, pp. 145–149, 2016.
- [37] N. Hao, J. Yang, H. Liao et al., “A unified factors analysis framework for discriminative feature extraction and object recognition,” *Mathematical Problems in Engineering*, pp. 1–12, 2016.
- [38] “The FGNET Aging Database,” 2013, <http://www.prima.inrialpes.fr/FGnet>.
- [39] R. Rothe, R. Timofte, and L. Van Gool, “Deep expectation of real and apparent age from a single image without facial landmarks,” *International Journal of Computer Vision*, pp. 1–14, 2016.

