ICIC Express Letters
Volume **18**, Number **9**, September **2024**

ICIC International ©2024 ISSN 1881-803X
pp. 979–986

FACIAL AGE-GROUP CLASSIFICATION WITH ORDINAL RANKING NEURAL NETWORK

Puspaningtyas Sanjoyo Adi[1,3], Agus Harjoko[2] and Wahyono[2,*]

[1]Doctoral Program, Department of Computer Science and Electronics
[2]Department of Computer Science and Electronics
Universitas Gadjah Mada
Bulak Sumur, Yogyakarta 55281, Indonesia
puspaningtyas@mail.ugm.ac.id; aharjoko@ugm.ac.id; *Corresponding author: wahyo@ugm.ac.id

[3]Department of Informatics
Universitas Sanata Dharma
Jl. Affandi, Mrican, Sleman, Yogyakarta 55281, Indonesia
puspa@usd.ac.id

Received September 2023; accepted December 2023

Abstract. *Age-group classification is a technique for classifying a face image into a particular age group. However, due to uncontrollable environments, insufficient and incomplete training data, strong person-specific variations, and large within-age span variations, age-group classification has become a challenging problem. This paper presents a novel neural network with an ordinal ranking approach for automatic age-group classification. After pre-processing, histogram of oriented gradients (HOG) features are extracted. Then, the images are classified into age groups using an ordinal ranking neural network (ORNN) classifier. This classifier consists of a multiclass neural network binary classifier that categorizes the input images into different age groups. We experimented with this approach using four age groups derived from the FG-NET and MORPH-II datasets.*
Keywords: Age-group classification, Ordinal ranking, Neural networks, Facial feature extraction, Histogram of oriented gradients

1. **Introduction.** Age estimation holds crucial importance in various applications, such as vending machines that regulate access to age-restricted items like alcohol and cigarettes. These machines employ age prediction programs to prevent underage purchases, highlighting the significance of accurate age estimation. However, this task is complex due to the stochastic and irreversible nature of aging, resulting in diverse facial shape and texture variations over time. Computer vision systems enable this process through facial recognition technologies, despite the challenges posed by these natural aging processes [1-3].

Age estimation methods rely on two main components: feature representation and estimation, where the system's accuracy hinges on the quality of extracted facial aging features. Techniques like histogram-based descriptors, capturing local intensity variations, have proven crucial in diverse applications. Huerta et al. [4] and Sawant and Bhurchandi [5] utilized HOG in their studies. Additionally, exploring multiple colour spaces, as Boutros et al. [6] discovered, offers varied information for decision-making, with channels like RGB, YUV, and HSV describing high textural detail, significantly improving facial age and gender estimation.

Age estimation models commonly approach the problem as either classification [7,8] or regression [9,10], but there is a growing interest in ordinal ranking methods [11]. For instance, Chang et al. [11] introduced ordinal hyperplanes ranker (OHRank), which uses binary classifiers to order labels based on features from the active appearance model

DOI: 10.24507/icicel.18.09.979

(AAM). Chen et al. [12] and Niu et al. [13] proposed CNN-based approaches treating ages as ranks, while Cao et al. [14] developed the CORAL framework, ensuring rank-monotonicity and consistent confidence scores using CNNs. Shin et al.'s moving window regression (MWR) method [15], also CNN-based, uses global and local regressors across age ranges. Recent CNN-based age estimation research has shown promising improvements [16,17].

Another model of an estimator is the hierarchical approach, which combines classification and regression methods to predict age. This model yields more accurate results and simplifies computational load. In this approach, the data is first classified into age groups, and then the age is predicted based on these pre-defined groups [5,18,19]. Other researchers have also utilized a CNN model that integrates two different methods, namely neural networks and extra learning machines (ELM) [20], to achieve age prediction.

In this paper, we introduce an age-group classification algorithm using local features like HOG along with an ordinal ranking neural network. Improving age-group classification is crucial in hierarchical age estimation. Our method categorizes front-facing images into four well-ordered age groups [5]. Traditional multiclass approaches overlook the ordinal information in age labels, while regression oversimplifies it, despite human aging being nonlinear. People often find it easier to determine if someone is older than a specific age rather than their exact age.

Building upon Cheng et al.'s research [21], where they introduced a neural network for ordinal regression/ranking, we adopted their model as it outperformed standard neural networks used for classification. Additionally, their method achieved similar performance to two state-of-the-art approaches (support vector machines and Gaussian processes) for ordinal regression on the same benchmark. Inspired by Cheng et al. [21], Niu et al. [13] and Chen et al. [12], we employ ordinal ranking neural networks as age group classifiers. A key departure from the approaches of Niu et al. and Chen et al. lies in our utilization of a hand-crafted feature extractor instead of a CNN. Furthermore, unlike Cheng et al., we employ a strategy akin to Niu et al., utilizing multi-binary classifiers to achieve our objectives.

In this paper, we present our proposed age-group classification using ORNN. Our approach consists of three stages: 1) Pre-processing, which involves face alignment and the removal of hair background, 2) Extraction of HOG features, and 3) Age group classification using multiple ORNN classifiers. These classifiers comprise several neural networks that function as binary classifier. They are fine-tuned using supervised training with the ordinal age group labels. The binary outputs are then aggregated to make the final age group prediction. An illustration of our model is shown in Figure 1.

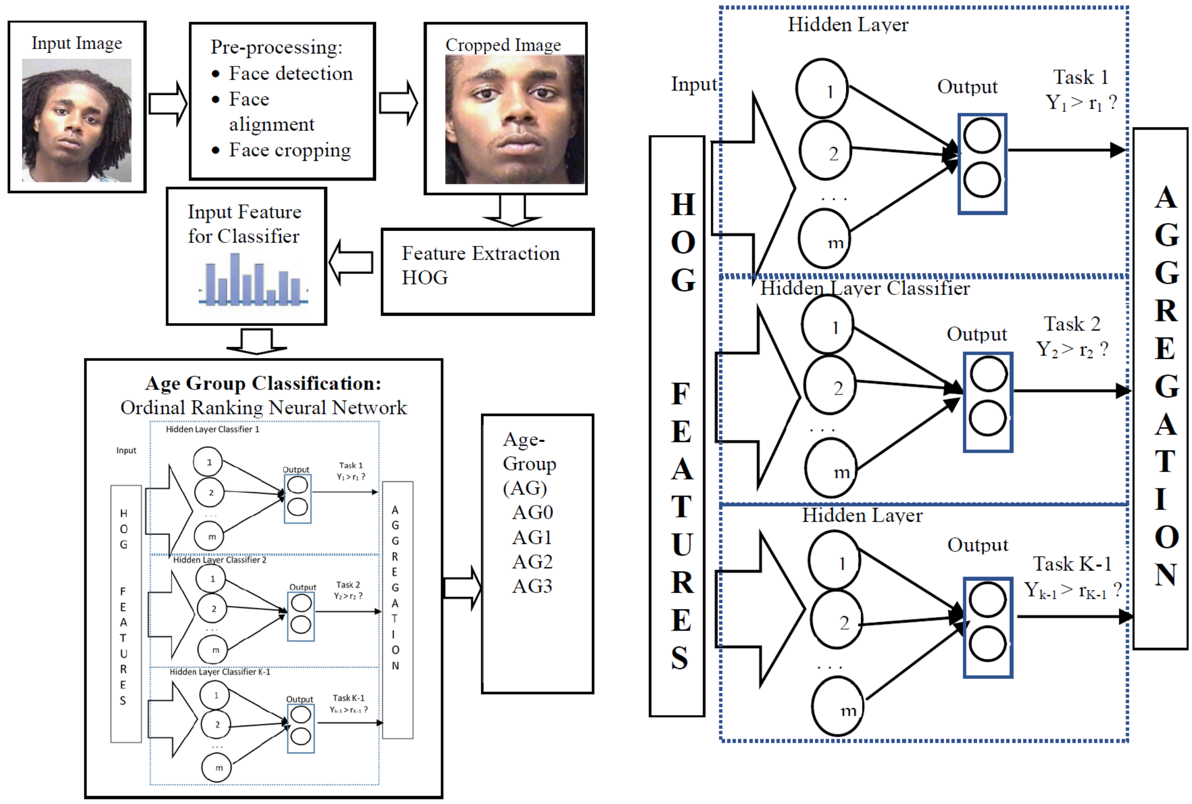
This work has following contributions: 1) we demonstrate effectiveness of simple ordinary ranking neural network in age-group classification; 2) for each age group, our approach tunes specific parameters to learn aging patterns.

The remainder of this paper is organized as follows. Section 2 discusses the material's specifics and proposed methods. Also, Section 3 presents the experimental results, while Section 4 summarizes the conclusion and recommendations.

2. Proposed Method.

The purpose of age group classification is to classify face images into different groups based on their ages. The entire age range is divided into several non-overlapping ranges. Age groups with its dataset can be seen in Table 1.

The proposed method has three key steps: 1) pre-processing: First, images are rotated to align them vertically, second, images are scaled so as to maintain the same distance between eyes for all the images, and third, images are cropped to remove background and hair regions; 2) HOG features extraction: in this step, HOG features are computed and stored in disk/storage; 3) age group classification: an OR-NN classifier is used in this step to classify the facial images into age groups. Figure 1(a) summarizes our proposed approach. Detailed visualization of the proposed method can be seen in Figure 1(b).



(a) (b)

FIGURE 1. (a) Overview of proposed method; (b) architecture of $K-1$ binary classifier ordinal ranking neural network (detailed version of age group classification in (a))

TABLE 1. Age groups partition scheme for FG-NET and MORPH-II dataset

	FG-NET		MORPH-II		
age-group	age-span	#images	age-span	#images	% distribution
0	0-5	233	16-22	6,356	23.6
1	6-12	248	23-35	9,585	35.8
2	13-21	265	36-45	7,265	27.0
3	22-69	256	46-74	3,654	13.6
Total		**1,002**		**26,860**	**100**

The task of ordinal ranking, also known as ordinal regression, involves categorizing data into ordered groups. Therefore, when presented with two labels, $x1$ and $x2$, the information regarding their relationship ($x2 > x1$ or $x1 > x2$) holds greater reliability for age estimation compared to the mere differences between the labels. This 'greater than' relationship offers more consistent data than precise age values.

Suppose we have a dataset of N training samples, $\{(x_i, y_i)|i = 1, \ldots, N\}$, in which x_i represents the ith face image feature vector and y_i represents the corresponding age group label. The age group label y_i is treated as an order of rank, $y_i \in \{0, \ldots, K-1\}$, where K is the total number of age group labels. In particular, an ordinal regression problem with K ranks is transformed into $K-1$ simpler binary classification sub-problems. For each rank $r_k \in \{r_1, r_2, \ldots, r_{K-1}\}$, a binary classifier is constructed to predict whether the rank of a sample y_i is larger than r_k. And then the rank of an unseen sample is predicted based on the classification results of the $K-1$ binary classifiers on this sample.

Given the original training data $D = \{x_i, y_i\}_{i=1}^{N}$, for the kth binary classification sub problem, a specific training data is constructed as $D^k = \left\{x_i, y_i^k, w_i^k\right\}_{i=1}^{N}$, where the $y_i^k \in \{0, 1\}$ is a binary class label indicating whether the rank of the ith sample y_i is larger than r_k as follows:

$$y_i^k = \begin{cases} 1, & \text{if } (y_i > r_k) \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

and w_i^k is the weight for the ith example, kth ranker/classifier, $k = \{1, \ldots, K-1\}$.

The $K-1$ binary classifiers are trained with their corresponding training data. The networks are standard multi-layer neural networks inspired by Cheng et al. [21] where each output layer contains 2 neurons corresponding to a binary classification task. The kth task is to predict whether the age of the ith facial image is larger than the rank r_k. For each binary classifier, we attach a softmax layer to convert the model's linear outputs-logits-to probabilities, which should be easier to interpret. The networks use standard sigmoid function, as shown as Equation (2).

$$\frac{1}{1 + e^{-z_i}} \tag{2}$$

For number of hidden layers, single layer is used because it is sufficient to learn any continuous mapping to an arbitrary accuracy [22]. The numbers of neurons are definite according to empirical equation in the hidden layer [23], as the following:

$$n_H = \sqrt{n_i + n_o} + l \tag{3}$$

where n_H is the number of neurons in hidden layer; n_i is the number of neurons in input layer; n_o is the number of neurons in output layer; l is integer from one to ten.

In our experiments, we use the binary cross entropy cost function because the networks are binary classifier. The binary cross entropy cost function is shown as Equation (4).

$$f_{cost} = -\frac{1}{N} \sum_{i=1}^{N} \left[y_i * \log(o_i) + (1 + y_i) * \log(1 - o_i) \right] \tag{4}$$

where y is target/label vector and o is output vector for the ith example/face image feature and N samples. Target/label vector is changed from standard binary label to one-hot label.

Based on different splitting of D, $K-1$ binary classifier networks can be trained from the base one. Given an unknown input x_i, we first use the basic networks to make a set of binary decisions and then aggregates them to make the final age group prediction $r(x_i)$,

$$r(x_i) = \sum_{k=1}^{K-1} [f_k(x_i) > 0] \tag{5}$$

where $f_k(x_i)$ is output of k binary classifier and [.] is truth-test operator, which is 1 if inner condition is true, and 0 otherwise. The explanation above can be summarized as shown in Figure 1(b).

Algorithm 1 summaries OR-NN for age-group classification.

2.1. **Dataset.** In this paper, we used MORPH album 2 [25] and FG-NET [26] to assess the efficacy of the age classification algorithm. MORPH-II [25] is a benchmark database containing 55,134 images of 13,000 persons ranging in age from 16 to 77 and ethnic from African, European, Asian, Hispanic and others. It was a collection of mugshot images, including information about ethnicity, gender, date of birth and date of acquisition.

Figure 2 displays a few examples from the MORPH album 2 dataset. The FG-NET contains 1,002 images from 82 different subjects collected by mostly scanning photographs of the subjects with ages ranging between new-borns to 69 years old.

Algorithm 1. Algorithm for age-group classification

Training

Input:

training data for each class $D^k = \left\{x_i, y_i^k\right\}_{i=1}^{N}$

x_i represents the ith HOG face image feature vector for training

y_i represents the corresponding age group label

kth binary classifier, $k = \{1, \ldots, K-1\}$

w_i^k is the weight for the ith example and N is training samples

Output: w^k weight of the kth binary classifier

Process: For each kth binary classifier do

- Build the kth multilayer neural network binary classifier, n_H neuron as Equation (3)
- Initialize its weight w^k
- Train the kth binary classifier with loss function Equation (4) with ADAM [24]
- Save weight each kth binary classifier, w^k

End for

Testing

Input:

testing data $\{x_i'\}_{i=1}^{M}$

x_i' represents the ith HOG face image feature vector for testing

w^k weight of the kth binary classifier and M is testing samples

Output: class label $r(x_i')$

Process: For each kth binary classifier do

- Build the kth multilayer neural network binary classifier, n_H neuron as Equation (3)
- Load w^k to the kth binary classifier
- Predict class label for each kth binary classifier, $f_k(x_i)$

End for

Agregate class label $r(x_i')$ with Equation (5)

(a) (b)

Figure 2. Facial image examples of dataset: (a) FG-NET dataset; (b) MORPH-II dataset

3. Experiments.

3.1. **Experimental setup.** FG-NET's limited faces prompt the use of Leave-One-Person-Out (LOPO) testing, where one person's samples are for testing and the rest for training. This prevents the classifier from learning individual traits, reducing data dependency and enhancing result independence.

In the second evaluation approach, over 25,000 images from the MORPH-II dataset [24] are utilized. We randomly selected 26,860 images and conducted a 5-fold cross-validation. Based on Sawant and Bhurchandi's research [5] and life-span development [27], we established four age groups, detailing age spans and image counts per group in Table 1 for balanced representation.

In implementing ORNN, we develop three binary classifiers using TensorFlow and each classifier uses three layers: one input layer, one hidden layer and one output layer. There are three age borders: (6, 13, 22) for FG-NET and (23, 36, 46) for MORPH-II.

3.2. **Result and discussion.** To thoroughly assess the performance of the analyzed features for age classification, we conducted an in-depth analysis of various parameters, including cell sizes, bin sizes, and neuron numbers within the hidden layer, to understand their impact on HOG features. Neuron numbers are based on Equation (3) and we set $l = \{0, 5, 10\}$. Table 2 shows that for FG-NET, 4 × 4 cell, bin = 10 and neuron numbers = 152, gives optimum result. For MORPH-II dataset, 4 × 4 cell, bin = 8 and neuron numbers = 136, gives best result. We observe that the overall classification accuracy increases as the number of features increases. Classification accuracy results vary with changes in the number of neurons.

Table 2. Results for FG-NET and MORPH-II datasets with size image 100 × 100, at varying cell (cell × cell), bin parameters, number of neurons with dropout rate = 0.0

Cell	Bin	Feature	Neuron	FG-NET accuracy (%)	MORPH-II accuracy (%)
12 × 12	8	12,800	113	63.67	64.56
12 × 12	8	12,800	118	62.85	64.92
12 × 12	8	12,800	123	61.71	64.74
12 × 12	9	14,400	120	63.05	65.21
12 × 12	9	14,400	125	62.33	65.01
12 × 12	9	14,400	130	63.26	65.23
12 × 12	10	16,000	126	63.05	65.60
12 × 12	10	16,000	131	63.88	65.23
12 × 12	10	16,000	136	62.95	65.32
8 × 8	8	15,488	124	63.88	66.59
8 × 8	8	15,488	129	63.78	66.88
8 × 8	8	15,488	134	64.81	66.49
8 × 8	9	17,424	132	63.67	67.01
8 × 8	9	17,424	137	64.71	66.72
8 × 8	9	17,424	142	64.40	66.85
8 × 8	10	19,360	139	62.95	66.79
8 × 8	10	19,360	144	63.57	66.88
8 × 8	10	19,360	149	65.02	66.85
4 × 4	8	18,432	136	65.74	**67.28**
4 × 4	8	18,432	141	65.43	66.90
4 × 4	8	18,432	146	64.71	66.88
4 × 4	9	20,736	144	64.50	67.18
4 × 4	9	20,736	149	64.71	67.23
4 × 4	9	20,736	154	65.02	66.99
4 × 4	10	23,040	152	**66.36**	66.79
4 × 4	10	23,040	157	65.63	66.78

We evaluated our method against leading approaches for age group classification, comparing results in Table 3 using the same testing criteria. Our method surpasses [5] and [19] but lags behind [28], highlighting the more effective feature extraction of [28] despite its longer computation time. Notably, our classification excels on smaller databases.

Table 3 shows a summary of comparison of overall performance of this approach with previous methods with MORPH-II dataset. Based on MORPH-II dataset, the performance of our proposed method is still under performance of Liu et al. [28] and Sawant and Bhurchandi [5], but over performance of Pontes et al. [19]. However, when compared to the results of Sawant and Bhurchandi, the research results can still be competitive, with a difference in results of 4%. If compared to the results of Liu et al., our feature

Table 3. Comparison of age-group estimation performance with FG-NET and MORPH-II

Method	FG-NET		MORPH-II	
	Age group	Accuracy (%)	Age group	Accuracy (%)
HOG, Gaussian Process (GP) [5]	0-5, 6-12, 13-21, 22-69	56.07	16-22, 23-35, 36-45, 46-74	71.14
Local Phase Quantization, SVM [19]	0-13, 14-21, 22-39, 40-69	60.00	16-25, 26-35, 36-45, 46-77	50.07
Gradient Oriented Pyramids, SVM [28]	0-5, 6-12, 13-21, 22-69	91.30	15-19, 20-25, 26-35, 36-77	82.80
Proposed: HOG, ORNN	0-5, 6-12, 13-21, 22-69	66.36	16-22, 23-35, 36-45, 46-74	67.28

extraction process is faster than the process of Liu et al. From Table 3, increasing features improves accuracy, but an optimal number is crucial. Textures play a role in age classification.

4. **Conclusions.** This paper introduces ordinal ranking neural networks (ORNN) as an innovative method for age group classification, employing neural networks with an ordinal ranking approach. Compared to methods like GP or SVM, ORNN offers a simpler yet competitive alternative. The experimental results on FG-NET and MORPH-II databases demonstrate the method's effectiveness. Nevertheless, addressing unbalanced training data and reducing feature count remains an area for improvement.

Acknowledgment. The research in this paper is supported by Indonesia Endowment Funds for Education/Lembaga Pengelola Dana Pendidikan (LPDP) and Universitas Gadjah Mada.

REFERENCES

[1] G. Panis, A. Lanitis, N. Tsapatsoulis and T. F. Cootes, Overview of research on facial ageing using the FG-NET ageing database, *IET Biometrics*, vol.5, no.2, pp.37-46, 2016.
[2] O. F. E. Osman and M. H. Yap, Computational intelligence in automatic face age estimation: A survey, *IEEE Trans. Emerg. Top. Comput. Intell.*, vol.3, no.3, pp.271-285, 2019.
[3] Y. Fu, G. Guo and T. S. Huang, Age synthesis and estimation via faces: A survey, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.32, no.11, pp.1955-1976, 2010.
[4] I. Huerta, F. Carles and A. Prati, Facial age estimation through the fusion of texture and local appearance descriptors, *ECCV 2014 Workshops*, pp.667-681, 2015.
[5] M. M. Sawant and K. Bhurchandi, Hierarchical facial age estimation using Gaussian process regression, *IEEE Access*, vol.7, pp.9142-9152, 2019.
[6] F. Boutros, N. Damer, P. Terhorst, F. Kirchbuchner and A. Kuijper, Exploring the channels of multiple color spaces for age and gender estimation from face images, *International Conference on Information Fusion*, 2019.
[7] X. Geng, Z. H. Zhou and K. Smith-Miles, Automatic age estimation based on facial aging patterns, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.29, no.12, pp.2234-2240, 2007.
[8] G. Guo, G. Mu, Y. Fu and T. S. Huang, Human age estimation using bio-inspired features, *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp.112-119, 2009.
[9] Y. Fu, S. Member, T. S. Huang and L. Fellow, Human age estimation with regression on discriminative aging manifold, *IEEE Trans. Multimed.*, vol.10, no.4, pp.578-584, 2008.
[10] G. Guo and G. Mu, Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression, *IEEE Conference on Computer Vision and Pattern Recognition*, 2011.
[11] K. Y. Chang, C. S. Chen and Y. P. Hung, A ranking approach for human age estimation based on face images, *Proc. of Int. Conf. Pattern Recognit.*, pp.3396-3399, 2010.

[12] S. Chen, C. Zhang and M. Dong, Deep age estimation: From classification to ranking, *IEEE Trans. Multimed.*, vol.20, no.8, pp.2209-2222, 2018.

[13] Z. Niu, M. Zhou, L. Wang, X. Gao and G. Hua, Ordinal regression with multiple output CNN for age estimation, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.4920-4928, 2016.

[14] W. Cao, V. Mirjalili and S. Raschka, Rank consistent ordinal regression for neural networks with application to age estimation, *Pattern Recognit. Lett.*, vol.140, pp.325-331, 2020.

[15] N. H. Shin, S. H. Lee and C. S. Kim, Moving window regression: A novel approach to ordinal regression, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.18739-18748, 2022.

[16] X. Feng, C. Shi, S. Zhao and K. Zhang, Multi-task multi-scale attention learning-based facial age estimation, *IET Signal Process.*, vol.17, no.2, 2023.

[17] C. Shi, S. Zhao, K. Zhang and Y. Wang, Face-based age estimation using improved Swin Transformer with attention-based convolution, *Front. Neurosci.*, vol.17, no.4, pp.1-14, 2023.

[18] S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park and J. Kim, Age estimation using a hierarchical classifier based on global and local facial features, *Pattern Recognit.*, vol.44, no.6, pp.1262-1281, 2011.

[19] J. K. Pontes, A. S. Britto, C. Fookes and A. L. Koerich, A flexible hierarchical approach for facial age estimation based on multiple features, *Pattern Recognit.*, vol.54, pp.34-51, 2016.

[20] M. Duan, K. Li, C. Yang and K. Li, A hybrid deep learning CNN-ELM for age and gender classification, *Neurocomputing*, vol.275, pp.448-461, 2018.

[21] J. Cheng, Z. Wang and G. Pollastri, A neural network approach to ordinal regression, *Proc. of Int. Jt. Conf. Neural Networks*, pp.1279-1284, 2008.

[22] L. Fausett, *Fundamental of Neural Network: Architecture, Algorithms, and Applications*, Prentice-Hall, Inc., 1994.

[23] X. W. Chen, Z. H. Liu and Z. K. Zhang, The measurement of planning surface roughness by neural networks based on image, *Proc. of the 6th Int. Conf. Nat. Comput.*, Yantai, China, pp.705-708, 2010.

[24] D. P. Kingma and J. L. Ba, Adam: A method for stochastic optimization, *Proc. of the 3rd Int. Conf. Learn. Represent. (ICLR 2015)*, pp.1-15, 2015.

[25] K. Ricanek Jr. and T. Tesafaye, MORPH: A longitudinal image age-progression, of normal adult, *Proc. of the 7th Int. Conf. Autom. Face Gesture Recognit.*, 2006.

[26] A. Lanitis, C. J. Taylor and T. F. Cootes, Toward automatic simulation of aging effects on face images, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.4, pp.442-455, 2002.

[27] J. W. Santrock, *Life-Span Development*, McGraw-Hill, 2010.

[28] K. H. Liu, S. Yan and C. C. J. Kuo, Age estimation via grouping and decision fusion, *IEEE Trans. Inf. Forensics Secur.*, vol.10, no.11, pp.2408-2423, 2015.