Check for updates

Facial age estimation using pre-trained CNN and transfer learning

Issam Dagher [1] (ID) · Dany Barbara [1]

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

This paper tackled the problem of human facial age estimation using transfer learning of some pre-trained CNNs, namely VGG, Res-Net, Google-Net, and Alex-Net. Those networks have been fine-tuned with transfer learning and undergone many experiments to get the optimum number of outputs and the optimum age gap. Based on those experiments, a novel hierarchical network that generates high age estimation accuracy was developed. This new network consists of a set of pre-trained 2-classes CNNs (Google-Net) with an optimum age gap which can better organize the face images in the age group they belong to. To show its effectiveness, it was compared with other states of the art techniques on the FGNET and the MORPH databases.

Keywords Facial age estimation · Pretrained CNN · Transfer learning

1 Introduction

Age estimation is defined as the process of determining a person's age based on many biometric attributes [26]. Among the multiple age-related traits, the human face conveys a lot of valuable biological information that allows one to guess the age of others immediately just by looking at their faces. The facial appearance of humans changes remarkably with time due to various extrinsic and intrinsic factors like gender, genes, environment, lifestyle, etc. As a result, that leads to diversity in human aging progress at different age stages. Moreover, aging growth varies among different people. Hence, the human estimation of age cannot always be accurate, especially the fact that people's behaviors and preferences vary at different ages. Therefore, the need for advanced algorithms that can exceed human ability [39] in the field of computer vision has become a more interesting yet difficult problem to handle in recent years [9].

✉ Issam Dagher
dagheri@balamand.edu.lb

[1] Computer Engineering Department, University of Balamand, Tripoli, Lebanon

Some examples of Pre-trained CNN are:

1. Alex-Net: It contains eight layers; the first five were convolutional layers, some of them followed by max-pooling layers, and the last three were fully connected layers. It uses the non-saturating Re-LU activation function, which showed improved training performance over tanh and sigmoid.
2. Google-Net: It consists of a 22-layer deep CNN and implements a novel element that is dubbed an inception module. It reduced the number of parameters from 60 million (Alex-Net) to 4 million.
3. VGG-Net: Its architecture is very uniform making it very attractive and special. It has 16 convolutional layers. It owns 3×3 convolutional layers with several kernels. It has a large number of parameters that is bigger than 130 million weights
4. Res-Net: To avoid going deeper (more training weights), Residual neural networks utilizes *skip connections*, or *shortcuts* to jump over some layers.

Training those networks (millions of parameters) from scratch is very challenging and difficult. A better solution is to use Transfer learning [20]. Transfer learning is a recent benchmark in the deep learning methodology that provides the ability to re-enhance the existed architectures to be suitable for new tasks instead of building new paradigms from scratch. It aims to utilize the knowledge acquired from pre-trained paradigms to solve new-related tasks instead of creating novel models. Thus, the previously learned models can be modified partially by adding new layers or training on new samples. It can improve the performance of deep learning schemes since the existed models learn generic attributes that can be existed in other domains. And it is very useful in overcoming the over-fitting problem, especially if the dataset is not large [20].

When CNN is retrained on new samples for another-related tasks, it is called fine-tuning. In fine-tuning:

1. The last layers will be removed.
2. New layers of the same class of removing ones will be configured.
3. The learnable neural nodes at the last fully connected layer in the classification module are replaced with a new layer from the same class with several neurons equals to the number of classes in corresponding training data.
4. It is applied in cases in which the amount of training samples is small.

Alex-net and Google-net are pre-trained deep convolutional networks that have been trained on over a million images and can classify images into 1000 object categories, such as a keyboard, coffee mug, pencil, and many animals. The networks have learned rich feature representations for a wide range of images which enable them to take an arbitrary image as input and then outputs a label for the object in the image together with the probabilities for each of the object categories.

The objective of this paper is to generate high age estimation accuracy by:

1. Using those powerful networks.
2. Fine-tuning those networks with transfer learning which is usually much faster and easier than training a network from scratch with randomly initialized weights.

To accomplish our objective, we have concentrated on 2 experiments:

1. Getting the optimum number of outputs (classes) of those networks to reach high age estimation results.
2. Finding the optimum age gap: As you increase the age gap between 2 groups, the recognition will increase among the 2 groups at the expense of increasing the age images which are in-between the 2 groups and eventually decreasing the recognition of these in-between ages.

Those experiments led to our novel idea: building a hierarchical network, which will be able to organize the face images in the age group they belong to. This network consists of a set of pre-trained 2-classes CNNs (Google-net) with optimum age gaps.

2 Related work

2.1 Deep learning for age estimation

Y. He. proposed an end-to-end deep embedding network for age estimation [18]. To create a metric embedding, we need to learn a function that can map an input to a feature space where the Euclidean distance in this embedded space directly represent the semantic similarity of the inputs, and in this case, the features were learned by triplet loss using a CNN. Still, on the topic of multi-task classification, [34] uses the age-related ordinal information and proposes a multiple output CNN to perform age ordinal regression. They transform the ordinal regression into a series of binary classification sub-problems. These binary classifiers are responsible for predicting whether the rank of an image is above an rk rank, $k \in K$, where K is the number of ranks, i.e. number of age ranges, which can be discrete or aggregated values. This technique was later improved by [5]. Other similar approaches were developed by [47, 50], where the multi-task classification loss is formulated by adding the gender and gender-specific age estimation losses. Multitask learning can be seen as a form of inductive transfer, which can help improve a model by introducing inductive bias. The inductive bias in the case of multi-task learning is produced by the sheer existence of multiple tasks, which causes the model to prefer the hypothesis that can solve more than one task. Multitask learning usually leads to better generalization [41]. The current state of the art performance is held by [52] on both MORPH2 and FGNET-AD datasets. They extract local facial characteristics of a cropped region estimated by a Long-Short Term Memory (LSTM) unit. They then combine their local extracted features with the global-image level features to perform their final estimation. It should be noted that some of the most popular pre-trained networks are Alex-net [22] and Google-net [45] which have been applied successfully to face datasets like the FGNET [8] and the MORPH [40].

Nowadays Deep learning schemes show impressive results in solving object classification [23]. More recently, multitude deep learning models, especially CNN have achieved a state-of-art performance in the domain of computer vision, especially for face-related assignments like face alignment in [44], facial verification in [46]. A complete assessment on utilizing the deep learning for age estimating is presented in [19] and compared to traditional handcrafted visual features. In practice, the fixed handcrafted fusion attributes are unable to reach the state-of-art performance like other modern deep learning schemes in the age estimation problem. Thus, this conventional model will be replaced with the trendiest techniques. In [51] the authors proposed age estimation using three stages: preliminary abstraction stage for extracting deeper

features, local feature encoding stage to model the relationship between local features and recall stage for the combination of temporary local impressions.

2.2 Facial features

One of the first attempts in the area of age classification of facial images was reported on the work of Kwon and Lobo [25]. Authors in [37] exploit a similar method to design an approach for an age progression in subjects whose ages are under 18. Statistical model of facial appearance was done in [28, 29]. Extracting more detailed information from facial features was achieved by using the Active Appearance Models (AAMs) to project faces into a lower-dimensional space [6]. After that, Geng et al. [11, 12] suggest Aging pattern Subspace (AGES), which relies on training the subspace on a collection of facial images that emerges each person at various ages. Each set of face images is handled as a single sample to project it to a low dimensional space [27]. On the other hand, authors in [48, 49] treat the problem of age estimation as a regression problem. On a different line of work, the authors in [16] propose manifold learning to build common aging patterns in low-dimensional space from multiple facial images for every age. Furthermore, many algorithms have been used successfully to extract appearance features and characterize the facial images such as Local Binary Patterns (LBP) in [15]. In [10] authors utilize Gabor features to learn an age estimator. More advanced technologies have been introduced in [33] in which authors propose a new model based on decreasing the noise of aging features from facial datasets. In [3] the authors convert the age estimation classification problem into smaller binary classification sub-problems based on an ordinal hyper-plane ranking scheme. An enhanced method, called label distribution, is present-ed in [13]. As for age-group research, the authors in [31] organize the dataset in four age groups, namely children, adolescence, moderately aged adults, older adults. The novel scheme utilizes in [43] achieved good performance. Authors in [24] adopted Topological Texture Features (TTF) for facial texture. The research outcomes in [32] show that the TTF method is ineffective to deal with the fast changes in facial textures. A survey of Neural Networks applied to age estimation is done in [36]. A hybrid deep learning CNN–ELM for age and gender classification is presented in [7]. [17] proposes Locally adjusted robust regression applied to age estimation. Age estimation robust to optical and motion blurring by deep residual CNN is done in [21]. The authors in [35] used double-level feature fusion of face and gait images. In [38] Convolutional neural networks for age classification from smart phone based ocular images was applied. Concentration on recognition surgically altered face images was done in [42].

In this paper, we have used a set of two-class CNNs (Google-nets) to tackle the age estimation problem. Different experiments were performed in order to find the optimum number of classes and the optimum age. This led to our new network which is compared with other techniques on both the FGNET and MORPH aging datasets giving smaller Mean Absolute Error (MAE).

3 Experiment setup

3.1 Facial databases

The FG-NET database [8] contains 1002 face images related to 82 individuals with an age range between 0 and 69. It is highly skewed because most of its samples are under the age of

31. A typical order of aging face samples from this dataset is shown in Fig. 1. Fig. 2 shows the age distribution of the FGNET images.

As can be seen from Fig. 2, the FG-NET has a small number of images in older age range. It has a larger number of images in the young and small age ranges. To overcome this limitation and to increase the number of age images, The MORPH dataset is used. The MORPH aging dataset is much larger than FG-NET. It has 55,132 face images from more than 13,000 subjects in this database. The average number of images per subject is 4. The ages of the face images range from 16 to 77 with a median age of 33.

We have used the total number of images in the FG-NET and the MORPH databases and we have randomly selected 80% of images as the training data, including 44,909 images, and the remaining 11,227 images are set as the testing data. Note that there are no duplicate subjects between the training and testing sets.

3.2 Training parameters

The FG-NET + MORPH dataset has been chosen because it covers a wide aging range [0–77]. The typical steps of transfer learning involve replacing the classifier module layers with new ones related to the new studied tasks. In practice, to retrain each of Alex-Net and Google-Net on human facial images, the fully connected layers would be set to have an output equals to the number of examining labels [20]. As for the earlier weights in the first layers, they would not be frozen and therefore new updates would replace the current values during the training phase.

Defining different options for optimal training is selected with a learning factor equals to 0.001, and several batches like the number of input samples. We have used the cross-entropy loss function and the stochastic gradient descent algorithm. Finally, the number of epochs ranges between150 to 300.

3.3 Evaluation methods

To validate the practical implementation of the suggested model, the Mean Absolute Error (MAE) is applied. MAE calculates the average of the absolute errors between the estimated values and the target age, as shown in the following equation:
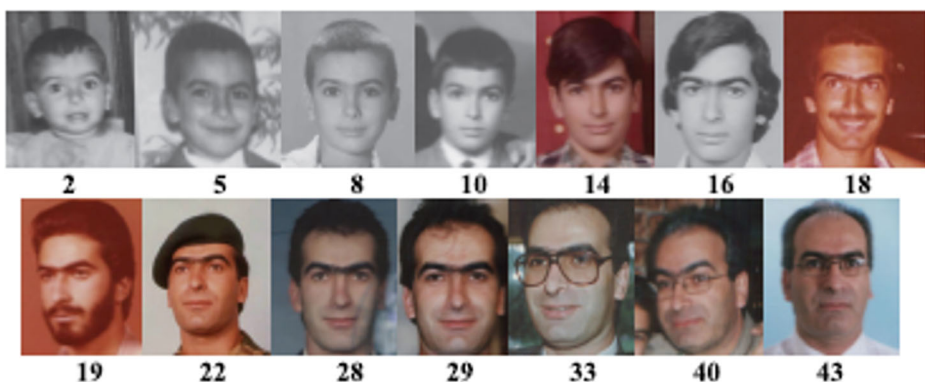

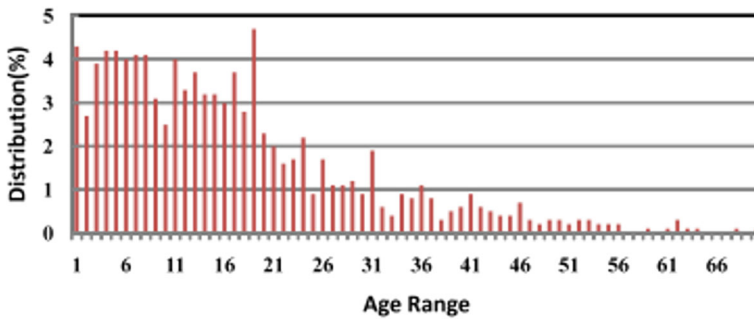**Fig. 1** Aging faces of one person from FG-NET dataset

Fig. 2 Age distribution of FGNET images

$$\text{MAE}_{\text{ABS}} = \frac{1}{N}\left(\sum_{n}^{N}\|k_n - y_n\|\right)$$

Where: K is the real label of the nth image and y is the predicted age based on this proposed model.

4 Experiments and Analysis

We have concentrated on 2 experiments: The effect of the number of classes on the accuracy of the pre-trained CNN and the effect of the age gap on the accuracy of the transfer learning. As a result, we have built a hierarchical network, which will be able to organize the face images in the age group they belong to.

4.1 First experiment: Investigation of the number of classes

In the first test, the relation between the accuracy and number of classes is investigated to measure the impact of the transfer learning on the new task. We have considered different age ranges. Table 1 shows the accuracy results for 6 classes (A, B, C, D, E, and F). We have used an age range of 5 years for classes A, B, C, and 10 years for D, E, and F. This is because "younger aging involves fast cranial changes, older aging involves slow textural changes" [43]. For 5 classes we have used an age gap of 10 years, for 4 classes an age gap of 15 years, for 3 classes an age gap of 20 years and for 2 classes an age gap of 30 years. Tables 1, 2, 3, 4 and 5 show the accuracy outcomes for the different number of classes by using VGG16, ResNet18, Google-Net, and Alex-net.

Table 1 The Accuracy results for 6 classes. A:0–5 B:6–10 C:11–19 D:20–29 E:30–39 F:40–77

Type	Age gap	Accuracy
VGG16	5 years	68%
Alex-net	5 years	69%
Res-Net18	5 years	72%
Google-Net	5 years	74%

Table 2 The accuracy results for 5 classes. A:0–9 B:10–19 C:20–29 D:30–39 E:40–77

Type	Age gap	Accuracy
VGG16	10 years	79%
Alex-net	10 years	81%
Res-Net18	10 years	83%
Google-Net	10 years	85%

Table 3 The accuracy results for 4 classes. A:0–14 B:15–29 C:30–44 D:45–77

Type	Age gap	Accuracy
VGG16	15 years	81%
Alex-Net	15 years	83%
Res-Net18	15 years	86%
Google-Net	15 years	87%

Table 4 The accuracy results for 3 classes. A:0–19 B:20–39 C:40–61

Type	Age gap	Accuracy
VGG16	20 years	86%
Alex-Net	20 years	87%
Res-Net18	20 years	88%
Google-Net	20 years	89%

4.2 Second experiment: Investigation of the age gap

Two-class CNN is chosen to predict which aging groups the testing images belong to. Fig. 3 shows a typical architecture for 2 age ranges. The SoftMax activation function is used. It should be noted that the output of the SoftMax will give the probability for the input image to belong to a certain age range. The comparison between the performance and generalization of the Alex-Net and the Google-Net has been conducted by employing different age ranges for each label in the two-class CNN (Fig. 3). All the train and test sets are from age range A and B. Table 6 shows the recognition results for the VGG16, ResNet18, Google-net, and Alex-net.

Table 6 shows that the best recognition results were attained by using the age ranges A:0–5 and B:10–15. Because all the data are from the ages of 0–15, the errors came from the age

Table 5 The accuracy results for 2 classes. A:0–29 B:30–77

Type	Age gap	Accuracy
VGG16	30 years	90%
Alex-Net	30 years	92%
Res-Net18	30 years	93%
Google-Net	30 years	94%
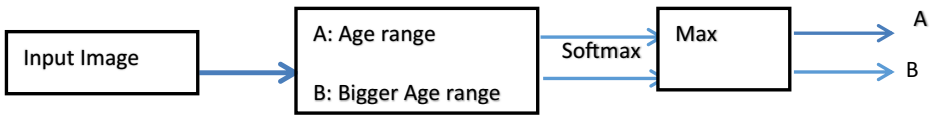
Fig. 3 Two-class CNN

Table 6 Age gap accuracy results

Labels	Google-net	Res-Net18	Alex-net	VGG16
A:0–5 B:6–10	86%	84%	82%	81%
A:0–5 B:10–15	97%	97%	96%	94%
A:0–5 B:15–20	95%	94%	93%	90%
A:0–5 B:20–25	92%	91%	90%	87%
A:0–5 B:25–30	91%	91%	90%	87%

(Train and test sets are only from age range A&B).

range 6–10. Fig. 4 shows 2 two-class CNNs which will give better performance than Fig. 3. The recognition results have increased from 97% to 99% using the Google-net (Table 7). It should be noted that Fig. 4 gave better accuracy because most of the test errors given by the first network in Fig. 4 are in the age range from 10 to 15 which will be covered by the second network.

Figure 3 shows a typical 2-class CNN. Its disadvantage is as you increase the large gap between 2 groups, the recognition will increase among the 2 groups at the expense of increasing the age images which are in-between the 2 groups and eventually decreasing the recognition of these in-between ages. For example: 2classes: Class1:0–5 and Class2:10–15. The in-between age ranges are 6–9. A better network is shown in Fig. 4 (our contribution). Figure 4 gave better accuracy (Table 7) because most of the test errors given by the first network in Fig. 5 are in the age range from 10 to 15 which will be covered by the second network. The in-between age ranges 6–9 will be covered in the first network in Fig. 4.
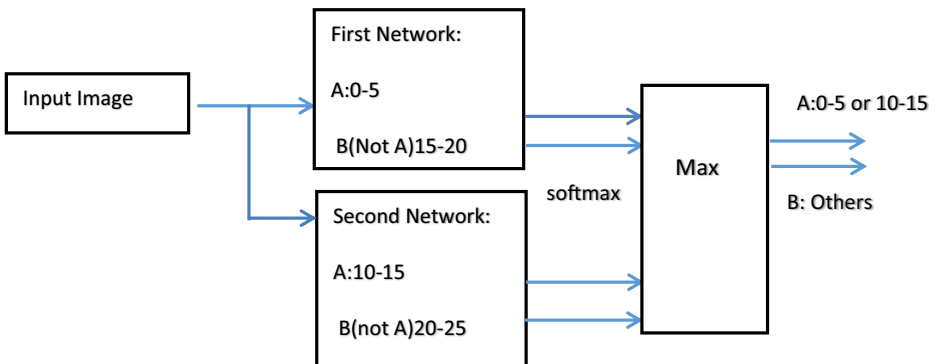

Fig. 4 2 Two-class CNNs

Table 7 Two-class CNN's accuracy Results

Labels	Google-net	Res-Net18	Alex-net	VGG16
A:0–5 B:10–15	99%	98%	97%	94%

(Train and test sets are only from age range A&B).

4.3 Age estimation using age gap

Our Age estimation network is shown in Fig. 5. It consists of a set of two-class CNNs (Google-nets). For example, the first and second CNN's will give an accurate recognition that the input image is in the range 0–5 or 10–15. The second and third CNN's will give an accurate recognition that the input image is in the range 15–20 or 20–25.

It should be mentioned that the above network also applies to targets that represent a small age width. In this case, the prediction process will not only guess the correct aging group of the image but also would give approximately the actual age of this image. Predicting the actual age of images could be achieved by making the age width very small.

4.4 Analysis of the results

We have considered different age ranges corresponding to different classes. The accuracy results for 2 classes gave the best results. And as you increase the number of classes the accuracy results will decrease. This can be interpreted as follows: For the same database, as you increase the number of classes, the number of training data per class will decrease.
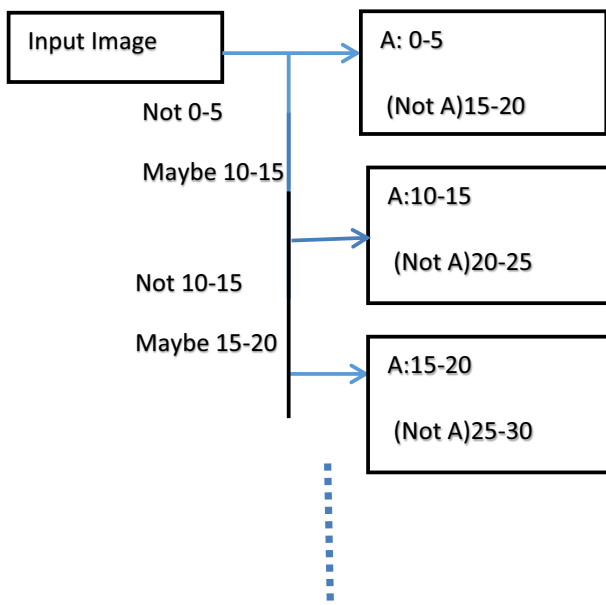

Fig. 5 Age estimation network

We have also considered different age gaps. Given the 2 groups:

A: from age1 to age2.
B: from age3 to age4.

The age gap is the age difference ag3-age2. As you increase the age gap, the number of images which will fall in this age difference range will increase and the accuracy results of these images will decrease.

Our 2-networks (Fig. 4) concentrated on increasing the accuracy results of these images. A fraction of these images will be covered by the first network and the remaining fraction will be covered by the second network.

This approach is generalized to cover all the age ranges (Fig. 5) where each network will cover a certain age range.

4.5 Comparisons with other methods

To evaluate the accuracy of our model the MAE [2] is used. Our network was compared (Table 8) to state-of-the-art models in FGNET and MORPH aging data sets. We have used the total number of images in FG-NET and MORPH. We randomly select 80% of images as the training data, including 44,909 images, and the remaining 11,227 images are set as the testing data. Note that there are no duplicate subjects between the training and testing sets.

5 Conclusion

In this paper, a novel network, which consists of a set of pre-trained 2-classes CNNs with optimum age gaps is presented. This set of two-class CNNs (Google-nets) fits the whole aging estimation task and generates high age estimation accuracy. This new technique was compared with other techniques on both the FGNET and MORPH aging datasets giving smaller Mean Absolute Error (MAE). It should be noted that the limitation of this study was to find a dataset which contains all the age ranges. We have solved this limitation by using the total number of images in the two databases: the FG-NET and the MORPH databases. We have used transfer learning applied to the CNN pre-trained networks; future work can be concentrated on training those networks from scratch.

Table 8 MAE results for different methods including our proposed network

Method	FGNET	MORPH
RED+SVM [2]	6.33	6.57
CS-LBFL [30]	4.36	4.37
CRCNN [1]	4.13	3.74
CPNN [14]	4.76	4.87
OHRANK [4]	6.27	6.28
Proposed Method	2.97	2.94

Our model gave smaller MAE (better accuracy) compared to other techniques.

References

1. Abousaleh FS, Lim T, Cheng W et al (2016) A novel comparative deep learning framework for facial age estimation. J Image Video Proc 2016:47
2. Chang K-Y,Chen C-S,Hung Y-P (2010) A ranking approach for human ages estimation based on face images, Proc. 20th Int. Conf. Pattern Recognit., pp 3396–3399
3. Chang KY, Chen CS, Hung YP (2011) Ordinal hyperplanes ranker with cost sensitivities for age estimation. CVPR 2011:585–592
4. Chang K-Y, Chen C-S, Hung Y-P (2011) Ordinal hyperplanes ranker with cost sensitivities for age estimation in Computer Vision and Pattern Recognition (CVPR), pp. 585–592.
5. Chen S,Zhang C,Dong M, Le J, Rao M (2017) Using ranking-CNN for age estimation. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 742–751
6. Cootes TF, Edwards GJ, Taylor CJ (2001) Active appearance models. IEEE Trans Pattern Anal Machine Intell 6:681–685
7. Duan M, Li K, Yang C, Li K (2018) A hybrid deep learning CNN–ELM for age and gender classification. Neurocomputing 275:448–461
8. FG-Net aging database (2021) http://sting.cycollege.ac.cy/alanitis/fgnetaging
9. Fu Y, Guo G, Huang TS (2010) Age synthesis and estimation via faces: a survey. IEEE Trans Pattern Anal Mach Intell 32(11):1955–1976
10. Gao F, Ai H (2009) Face age classification on consumer images with Gabor feature and fuzzy lda method. In: International conference on biometrics. Springer, Berlin, pp 132–141
11. Geng X, Zhou ZH, Zhang Y, Li G, Dai H (2006) Learning from facial aging patterns for automatic age estimation. In proceedings of the 14th ACM international conference on multimedia, pp 307–316
12. Geng X, Zhou ZH, Smith-Miles K (2007) Automatic age estimation based on facial aging patterns. IEEE Trans Pattern Anal Mach Intell 29(12):2234–2240
13. Geng X, Yin C, Zhou ZH (2013) Facial age estimation by learning from label distributions. IEEE Trans Pattern Anal Mach Intell 35(10):2401–2412
14. Geng X, Yin C, Zhou Z-H (2013) Facial age estimation by learning from label distributions. IEEE Trans Pattern Anal Mach Intell 35(10):2401–2412
15. Gunay A, Nabiyev VV (2008) Automatic age classification with LBP. In 2008 23rd International Symposium on Computer and Information Sciences, pp 1–4.
16. Guo G, Fu Y, Dyer CR, Huang TS (2008) Image-based human age estimation by manifold learning and locally adjusted robust regression. IEEE Trans Image Process 17(7):1178–1188
17. Guo G, Fu Y, Huang TS, Dyer C (2018) Locally adjusted robust regression for human age estimation. In: Proceedings of IEEE workshop on applications of computer vision, pp 19–21
18. He Y,Huang M,Miao Q,Guo H, Wang J (2017) Deep embedding network for robust age estimation. In 2017 IEEE international conference on image processing (ICIP), pages 1092–1096
19. Huerta I, Fernández C, Segura C, Hernando J, Prati A (2015) A deep analysis on age estimation. Pattern Recogn Lett 68:239–249
20. Iorga C, Neagoe V (2019) A deep CNN approach with transfer learning for image recognition, 11th international conference on electronics, vol 2019. Computers and artificial intelligence (ECAI), Pitesti Romania, pp 1–6
21. Kang JS, Kim CS, Lee YW, Cho SW, Park KR (2018) Age estimation robust to optical and motion blurring by deep residual CNN. Symmetry 10(4):108
22. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems pp 1097–1105
23. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems , pp 1097–1105
24. Kumar VV, Kiran JS, Chandana VH (2013) An effective age classification using topological features based on compressed and reduced grey level model of the facial skin. IJ Image Graphics Signal Process 6(1):9–17
25. Kwon YH, da Vitoria Lobo N (1999) Age classification from facial images. Comput Vis Image Underst 74(1):1–21
26. Lanitis A (2010) Facial age estimation. Scholarpedia 5(1):9701
27. Lanitis A (2010) Facial age estimation. Scholarpedia 5(1):9701
28. Lanitis A, Taylor CJ, Cootes TF (2002) Toward automatic simulation of aging effects on face images. IEEE Trans Pattern Anal Mach Intell 24(4):442–455
29. Lanitis A, Draganova C, Christodoulou C (2004) Comparing different classifiers for automatic age estimation. IEEE Trans Syst Man Cybernet Part B (Cybernetics) 34(1):621–628.455
30. Lu J, Liong VE, Zhou J (2015) Cost-sensitive local binary feature learning for facial age estimation. IEEE Trans Image Process 24(12):5356–5368

31. Mohan MC, Vijaya Kumar V, Venkata Krishna V (2010) Novel method of adult age classification using linear wavelet transforms. Int J Comput Sci Network Secur 10(3):61–68
32. Murty GS, Kumar VV, Obulesu A (2013) Age classification based on simple LBP transitions. Int J Comput Sci Eng 5(10):885
33. Ni B, Song Z, Yan S (2009) Web image mining towards universal age estimator. In proceedings of the 17th ACM international conference on multimedia, pp. 85-94
34. Niu Z,Zhou M,Wang L,Gao X, Hua G (2016). Ordinal regression with multiple output CNN for age estimation. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 4920–4928
35. Punyani P, Gupta R, Kumar A (2018) Human age-estimation system based on double-level feature fusion of face and gait images. Int J Image Data Fusion Taylor and Francis 9(3):222–236
36. Punyani P, Gupta R, Kumar (2020) A. Neural networks for facial age estimation: a survey on recent advances. Artif Intell Rev 53:3299–3347
37. Ramanathan N, Chellappa R (2006) Modeling age progression in young faces. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06) Vol. 1, pp. 387–394
38. Rattani A, Reddy N, Derakhshani R (2018) Convolutional neural networks for age classification from smart-phone based ocular images. In: proceedings of IEEE international joint conference on biometrics (IJCB), pp 756–761
39. Rhodes MG (2009) Age estimation of faces: a review. Appl Cognitive Psychol: Off J Soc Appl Res Memory Cognition 23(1):1–12
40. Ricanek K, Tesafaye T (2006) Morph: a longitudinal image database of normal adult age-progression. In 7th international conference on automatic face and gesture recognition (FGR06), pp 341–345
41. Ruder S (2017) An overview of multi-task learning in deep neural networks. CoRR, abs/1706.05098
42. Sabharwal T, Gupta R, Son LH, Kumar R, Jha S (2018) Recognition of surgically altered face images: an empirical analysis on recent advances. Artif Intell Rev
43. Sirovich L, Kirby M (1987) Low-dimensional procedure for the characterization of human faces. Josa a 4(3):519–524
44. Sun Y, Wang X, Tang X (2013) Deep convolutional network cascade for facial point detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3476–3483
45. Szegedy C et al (2015) Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition , pp 1–9
46. Taigman Y, Yang M, Ranzato MA, Wolf L (2014) Deepface: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE conference on computer vision and pattern recognition , pp 1701-1708
47. Tian Q, Chen S, Tan X (2016) A unified gender-aware age estimation. CoRR, abs/1609.03815
48. Yan S, Wang H, Tang X, Huang TS (2007) Learning auto-structured regressor from uncertain nonnegative labels. In 2007 IEEE 11th international conference on computer vision, pp 1–8
49. Yan S, Wang H, Huang TS, Yang Q, Tang X (2007) Ranking with uncertain labels. In 2007 IEEE international conference on multimedia and expo, pp 96–99
50. Yi D, Lei Z, Li SZ (2015) Age Estimation by Multi-scale Convolutional Network. In: Cremers D, Reid I, Saito H, Yang MH (eds) Computer Vision – ACCV 2014. ACCV 2014. Lecture notes in computer science, vol 9005. Springer, Cham. https://doi.org/10.1007/978-3-319-16811-1_10
51. Yu T, Wang J, Wu L, Xu Y (2019) Three-stage network for age estimation. CAAI Trans Intell Technol 4(2):122–126
52. Zhang K, Liu N, Yuan X, Guo X, Gao C, Zhao Z, Ma Z (2020) Fine-grained age estimation in the wild with attention LSTM networks. IEEE Trans Circuits Syst Video Technol 30(9):3140–3152