

Transfer Learning Adaptive Facial Attractiveness Assessment

I Lebedeva, Y Guo and F Ying

School of Information Science and Engineering,
East China University of Science and Technology,
Shanghai, China

irina@mail.ecust.edu.cn

Abstract. Recent advances in Computer Vision and Artificial Intelligence have brought the opportunity to automate facial attractiveness evaluation. A range of studies have been addressed to the task and have achieved reasonable prediction accuracy. However, most of these methods work well only on photos with restrictions on expression, posture, illumination, but not on real-world face photos. This work is aimed to improve the attractiveness assessment state-of-the-art in both cases. To this end, an approach that employs transfer learning methodology as well as shallow machine learning was proposed for highly accurate facial attractiveness prediction. Specifically, a Convolutional Neural Network (CNN), Facenet, originally designed and pre-trained for the face recognition task is utilized. High-level facial features were extracted by using the network and then fed into Support Vector Regression in order to predict facial attractiveness. Extensive experiments conducted on widely used facial beauty datasets Gray and SCUT-FBP5500 demonstrated that the proposed method outperformed other attractiveness prediction approaches. The experimental results also confirmed the effectiveness of the method in both constrained and unconstrained environment.

1. Introduction

A primary task in Artificial Intelligence and Computer Vision is to reproduce the human ability to recognize various visual concepts. Facial beauty is one of these concepts [10], and to let a machine predict attractiveness is a challenging task. Automatic beauty evaluation has application potential in face beautification [14], facial makeup recommendation [16], content-based image retrieval [18] and, as a result, attracts considerable interest among both the research community and industry.

A range of approaches was proposed for facial beauty prediction. One category of studies applies shallow machine learning predictors trained by hand-crafted features, e.g. geometric ratios, landmark distances [11], [6] and SIFT, Gabor features [1]. Another category of works is advanced by deep learning and thus have achieved relatively high accuracy on attractiveness prediction [15]. However, there are still many issues that should be investigated. First of all, many existing beauty evaluation methods, including of deep learning based methods train a neural network from scratch [15], a large-sized dataset and long training process are required for prediction and make them time- and power-consuming. Thirdly, the prediction accuracy is still can be improved even on face images in constrained environment.

In order to solve the problems above and explore the best automatic beauty evaluation method, the state-of-the-art of other face-related tasks, such as gender and age detection [19] have been studied. The conclusion about main points was drawn, which make that approaches highly effective and accurate, especially on face photos in unconstrained environment. First of all, the input of that face-



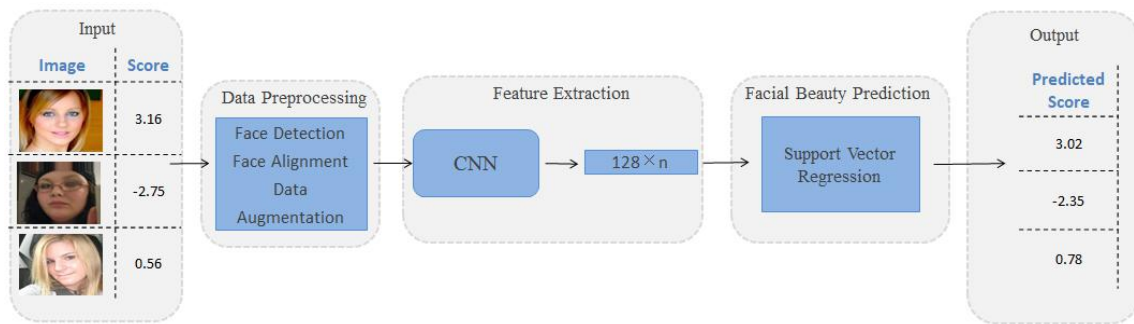


Figure 1. Pipeline. The pipeline includes the images preprocessing, CNN feature extractor, SVR predictor.

related methods is not pixels, but high-level features, obtained by well proven pre-trained networks, e.g. VGG16 [21], AlexNet [13]. Secondly, these high-level features are used to train shallow machine learning algorithms in order to predict beauty scores. In other words, up-to-date age and gender assessment methods utilize transfer learning methodology as well as shallow predictors to achieve high prediction accuracy. At the same time, involve much less computation power and data needed for training.

Moreover, the face recognition state-of-the-art and the possibility of applying its advances to the attractiveness assessment issue have been investigated. Facenet [20] is a neural network designed specifically for face recognition and pre-trained on a large-sized facial photos dataset VGGFace2 [3]. All this makes its application for facial beauty prediction potentially effective. In this work, transfer learning methodology was utilized and the advantages of Facenet facial classification model were taken to obtain high-level and robust features of face images (Facenet embeddings) and, as a result, make attractiveness evaluation more accurate and less time-consuming. The pipeline is presented in figure 1.

Since data preprocessing improves the accuracy of deep learning methods, various image preprocessing approaches also have been studied and the results of their performance on photos in all type of environment are presented in the work. Firstly, face detection and alignment were applied, and then data augmentation was conducted. Extensive experiments were conducted on two widely used facial attractiveness datasets with photos with and without variations in pose, expression, illumination - on Gray dataset [10] and SCUT-FBP 5500 [15], respectively. The results demonstrate the effectiveness of the proposed method in both cases. The main contributions of the study are the following:

- A novel automatic attractiveness evaluation method that employs transfer learning and support vector regression is proposed. Specifically, a network designed and pre-trained for the face recognition issue, namely Facenet, is used for high-level features extraction.
- The results of data preprocessing methods performed in various combinations were demonstrated and compared in order to find the most effective approaches for face images in both constrained and unconstrained environment.
- Extensive experiments conducted on facial attractiveness datasets with and without restrictions on pose, expression, illumination, and blurriness proved the effectiveness of the proposed method in both cases.

2. Related Works

Studies addressed to the facial attractiveness prediction task can be divided into 2 categories. The first category uses the combination of hand-crafted features and shallow machine learning; the second one is advanced by neural networks, especially the state-of-the-art deep learning. In turn, hand-crafted features can be divided into geometrical and textual features. Facial attractiveness evaluation methods based on deep learning technologies can be divided into deep CNN based and transfer learning approaches.

The earliest attempts on facial attractiveness prediction utilized hand-crafted features and shallow machine learning techniques [22], [12]. Geometric ratio and landmark distance were widely employed as extracted features for shallow predictors [7]. However, these approaches have low prediction accuracy and require a huge amount of time to manually specify facial landmarks for each image in a dataset. Eienthal *et al.* applied an ensemble of features that included ratios and landmark distances, an indicator of facial symmetry, hair color, skin smoothness and the coefficients of an eigenface decomposition [6]. Kagian *et al.* later enhanced their approach using an improved feature selection [11]. The examples of the textual features used for the beauty assessment issue are SIFT and Gabor [30], [1].

Significant results in automatic attractiveness evaluation have been achieved due to the advancement in deep learning [26], [8]. Gray *et al.* [10] made the first attempt to create a fully automatic beauty prediction approach that does not require manual annotation of facial features and used a hierarchical feed-forward model for attractiveness prediction. The authors in [25], [15] utilized CNN advantages to improve the automatic attractiveness assessment state-of-the-art [25].

Transfer learning has been widely applied in recent facial attractiveness methods. Donahue [4] showed the efficiency of features extracted by CNN pre-trained on ImageNet. Rothe *et al.* employed pre-trained VGG16 for automatic age, gender and generic attractiveness prediction [19]. Xu *et al.* proposed an approach which extracts rich deep facial features through knowledge adaptation, and then trains Bayesian ridge regression algorithm for facial beauty prediction [28]. The method in [17] is using deep CNNs to extract and then select facial attributes with significant contributions to attractiveness verified by statistical tests. A multi-task network with fully convolutional architecture that predicts facial beauty was presented in [27].

Various neural network structures have also improved the facial beauty prediction state-of-the-art. Zhai *et al.* proposed an effective CNN using a novel Softmax-MSE loss function and a double activation layer [29]. A three-level residual-in-residual structure for facial beauty prediction was presented in [2]. Gan *et al.* introduced a multi-input multi-task 2M BeautyNet to jointly learn facial beauty prediction and gender recognition. The proposed network transfers pre-trained parameters when inputting from different databases simultaneously [9]. The use of graph-based semi-supervised learning for face beauty scoring was introduced in [5]. The authors also adapted a linear Flexible Manifold Embedding scheme to the case of real scores propagation.

3. Methodology

A range of works has demonstrated that the last layer of a CNN is a powerful feature extractor. Following the same strategy, in the proposed method, A well-proven neural network pre-trained on a large-sized dataset was applied in order to make the method more accurate and less time-consuming. Since we worked with face images, a neural network, specifically designed for a face-related task and pre-trained on a facial dataset, was exploited. Data preprocessing, e.g. face detection, face alignment, data augmentation, was also exploited to make attractiveness assessment more effective. Support vector regression (SVR) trained on CNN-based features was chosen for the facial beauty prediction for the following reasons. Firstly, SVR is more robust than other shallow machine learning techniques, such as Naive Bayes, Decision Tree, K-Nearest Neighbors, due to optimal margin gap between separating hyperplanes. It is computationally more efficient by reason of using Kernel Trick. Thus, it is faster in training, better in accuracy with stability.

3.1. FaceNet

Since the Facenet network structure [20] is designed for face recognition and pre-trained on a large sized facial dataset VGGFace2 [3], applying FaceNet for the facial attractiveness evaluation issue potentially results in higher prediction accuracy. The FaceNet model generates 128 embeddings for the face recognition purpose. As a result, Facenet was used to turn a color face image into 128 embeddings or a vector of 128 floating points numbers. This vector can be used as features for the facial attractiveness task.

3.2. Feature Extraction

When data preprocessing is done, an image is fed into FaceNet network to extract the facial features as embeddings. These embeddings from the last layer of FaceNet can be thought of as the unique facial features that describe an individual's face. Then the FaceNet embeddings are used to train out prediction model. The whole dataset for the training and prediction can be described as

$$M = [x_1, x_2, \dots, x_n] \quad (1)$$

where M - dataset, x - face image - number of images. Then every face image can be presented as a vector

$$x_n = [z_1, z_2, \dots, z_{128}] \quad (2)$$

where v, k, z are embedding, generated by Facenet model. 128 embeddings were used for image because this number of embeddings is well-proven for the face recognition task. As we see, the input features are computationally light and can be efficiently trained by a shallow machine learning approach. It makes the proposed method less time- and power-consuming and, as a result, more applicable for real-life.

3.3. Prediction Model

In this work, facial attractiveness is regressed using Support Vector Regression (SVR). We use 60% of images for training and 40% for testing. Training data has n samples. The goal is to find a $f(x)$ that has at most ϵ from the obtained targets y_i for the training data. The function must be as flat as possible

$$f(x) = \langle w, x \rangle + b \quad (3)$$

Flatness in the case of (7) means that one seeks a small w . In order to ensure this, we minimize the norm, i.e. $|w|^2 = \langle w, w \rangle$. We can write this as a convex optimization problem:

$$\text{Minimize } \frac{1}{2} \|w\|^2 \quad (4)$$

$$\text{Subject to } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon \\ \langle w, x_i \rangle + b - y_i \leq \epsilon \end{cases} \quad (5)$$

However, there are slack variables, their concept is simple - for any value that falls outside of ϵ , we can denote its deviation from the margin as ξ_i, ξ_i^* .

$$\text{Minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (6)$$

$$\text{Subject to } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon \\ \langle w, x_i \rangle + b - y_i \leq \epsilon \\ \xi_i, \xi_i^* \leq \epsilon \end{cases} \quad (7)$$

The constant $C > 0$ means the trade-off between the flatness of f and the amount up to which deviations bigger than ϵ are tolerated. This corresponds to dealing with ϵ -insensitive loss function ξ_i that is outlined by

$$\text{Subject to } \begin{cases} 0 & \text{if } |\xi| \leq \epsilon \\ |\xi| - \epsilon & \text{otherwise} \end{cases} \quad (8)$$

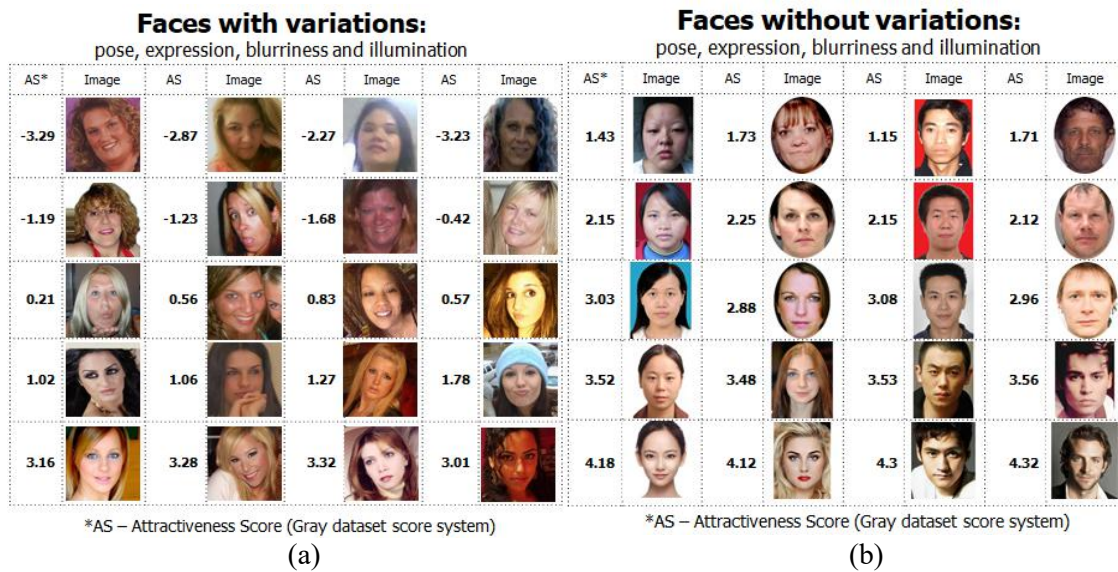


Figure 2. Face images with and without variations, (a) Gray dataset samples - female Caucasian face images in unconstrained environment: various postures, blurriness, illumination; (b) SCUT-FBP 5500 dataset samples - female/male Asian/Caucasian frontal face images in constrained environment.

4. Experimental Results

Footnotes should be avoided whenever possible. If required they should be used only for brief notes that do not fit conveniently into the text.

4.1. Datasets

The proposed method was conducted on different datasets for 3 reasons. Firstly, there is still no benchmark dataset for attractiveness prediction. Secondly, these datasets have different consistency, for instance, Gray has only female faces in-the-wild (figure 3 (a)), SCUT-FBP 5500 has only front faces with lighting restrictions (figure 3 (b)). Thirdly, score systems used in datasets are different. For examples, in order to evaluate attractiveness, some authors use scores from 1 to 5, where 5 is the highest score, other authors employ scores from -5 to 5. In order to evaluate the proposed methods and compare them with previous works 2 widely used attractiveness datasets were exploited.

Table 1. The Method Results on SCUT-FBP 5500 Dataset.

	Original Images	Face Cropping	Original Photos + Face Alignment	Face Cropping + Face Alignment
PC	0.8834	0.8728	0.8941	0.8807
MAE	0.2415	0.2565	0.2365	0.2455
RMSE	0.3221	0.2319	0.3101	0.3252

Table 2. The Method Results on Gray Dataset.

	Original Images	Face Cropping	Original + Alignment	FaCropping+ Face Alignment
PC	0.4638	0.4704	0.4665	0.4732
MAE	0.1417	1.1337	1.1378	1.1315
RMSE	0.9467	0.9290	0.9371	0.9254

SCUT-FBP 5500. The dataset presented in [15] contains 5500 frontal faces with diverse properties (male/female, Asian/Caucasian, ages). Although the dataset has been used in a range of works,

restrictions in terms of illumination, blurriness and pose may limit the performance of the attractiveness prediction models.

Gray Dataset. The dataset proposed in [10] contains 2056 Caucasian female faces and there are no restrictions in terms of pose, background, lighting, expression. Most images were cropped from low-quality photos taken by phone cameras and, the implication of the dataset into experiments makes the work closer to real life scenario.

4.2. Performance Evaluation

The performance was reported in terms of Pearson's Correlation (PC), Mean Absolute Error (MAE) and Root Mean Square Deviation (RMSE).

$$PC = \frac{\sum_{i=0}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (9)$$

$$MAE = \frac{1}{m} \sum_{i=1}^m |f(x_i) - y_i| \quad (10)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (f(x_i) - y_i)^2} \quad (11)$$

where n denotes the number of image samples, x_i - the input feature vector of image i , $f(\bullet)$ - the learning algorithm, y_i - the ground truth beauty score of image i , \bar{x} denotes the sample mean, and analogously for \bar{y} . PC is a measure of the linear correlation between $h(x_i)$ and y_i . It has a value between 1 and -1, where 1 means total positive linear correlation, 0 - no linear correlation, and -1 means total negative linear correlation. MAE and $RMSE$ measure of the quality of a machine learning model, the performance is better if MAE and $RMSE$ are closer to zero.

4.3. Results

Table 1 and table 2 summarize the results of extensive experiments on SCUT-FBP 5500 and Gray dataset respectively. The tables contain the prediction accuracy obtained on original photos of the datasets, the photos after face cropping and face alignment. Since Gray dataset contains unaligned faces photos with variant postures, face cropping makes the prediction accuracy of the proposed method higher, while the same preprocessing makes the result even worse on SCUT-FBP 5500 dataset that has only frontal faces. Moreover, a SCUT-FBP 5500 photo contains not only a face but also some backgrounds. Thus, we can conclude, that background, added to the cropped faces, provides better results for automatic attractiveness prediction.

Most faces in SCUT-FBP 5500 dataset are centralized and, as result, the positions of eyes on the photos are almost the same. Thus, as we see in table 1, face alignment increases the prediction accuracy slightly. However, the same preprocessing significantly improves the results on the dataset in unconstrained environment (table 2).

Table 3. The Proposed Method with and without Data Augmentation.

	Original Data			Augmented Data		
	PC	MAE	RMSE	PC	MAE	RMSE
SCUT	0.8941	0.2365	0.3101	0.8964	0.2341	0.3064
Gray	0.4732	1.1315	0.9254	0.4914	1.1165	0.9187

Since data augmentation is critical for deep learning techniques and can improve the prediction accuracy for another face-related tasks, such as age, gender estimation, face recognition, we decided to run the proposed method on both datasets Gray and SCUT-FBP 5500 with data augmentation in order

to increase the accuracy of the automatic attractiveness evaluation. As we can see in table 3, data augmentation slightly improve the prediction results on SCUT-FBP 5500 dataset but makes much better work on Gray dataset.

4.3.1. Comparison with the State-of-the-Art Methods. In order to demonstrate the effectiveness and robustness of the proposed approach, the achieved results are compared with other state-of-the-art approaches on datasets with photos in constrained (SCUT-FBP 5500) and unconstrained environment (Gray dataset). Since these datasets have different consistency and use different score systems, we make the comparison separately. A summary of the attractiveness prediction methods conducted on SCUT-FBP 5500 as well as their accuracy in terms Pearson's correlation (PC), Mean Absolute Error (MAE) and Root Mean Square Deviation (RMSE) are presented in table 4. The comparison on Gray dataset is presented only in term of PC, because the authors who performed experiments on the dataset reported their results only in term of PC (table 5).

The proposed method was compared with facial beauty prediction approaches that are based on hand-crafted features and shallow predictors: Geometric Features and Linear Regression (GF + Linear Reg), Gabor and Support Vector Regression (Gabor + SVR) on SCUT-FBP 5500 dataset and Eigenface [6] on Gray dataset. As you see on table 4 and table 5, the methods provided the lowest accuracy on both datasets.

Table 4. Comparison with other state-of-the-art approaches on SCUT-FBP 5500 dataset.

Method	PC	MAE	RMSE
GF + Linear Reg.[15]	0.6738	0.3914	0.5085
Gabor SVR [15]	0.8065	0.3976	0.5126
AlexNet [13]	0.8298	0.2938	0.3819
ResNeXt50 [15]	0.8777	0.2518	0.3325
HMTNet [27]	0.9783	0.2501	0.3263

Table 5. Performance comparison with the related state-of-the-art works on Gray dataset.

Method	PC
Eigenface.[6]	0.180
Auto Encoder [24]	0.437
Multiscale Model [10]	0.458
Feature Transferring [28]	0.468
VGG16+SVR [19]	0.478
Proposed Method	0.491

The CNN-based methods showed much better performance [13],[27], [28]. For SCUT-FBP 5500 dataset, we made the comparison with study in [15], where the neural network ResNeXt-50 was employed for the attractiveness prediction task. The network structure was originally designed not for a face-related task, but was fully trained on 5500 face images from SCUT-FBP 5500. As a result, the prediction accuracy of the approach is lower than the proposed method based on specific face-related neural network. For Gray dataset, we also made a comparison with the original Gray approach [10]. The authors made the first attempt to apply deep learning for automatic attractiveness prediction and, as we see, significantly improved its state-of-the-art.

A comparison with the approach presented in [19] was also provided. The basic idea is applying the pre-trained network VGG-16 for facial feature extraction and SVR for prediction. This method is similar to the proposed method, but since VGG-16 was created not for a face-related task, the accuracy is lower. Experimental results on SCUT-FBP 5500 demonstrated that the proposed approach achieved the highest Pearson's correlation of 0.8964, minimum MAE of 0.2361 and RMSE of 0.304 among all the methods in table 4. The experiments on Gray dataset indicated that the method showed a superb performance as well and overcame the state-of-the-art. This work took advantages of transfer learning

methodology as well as a network designed and pre-trained specially for face recognition. Moreover, relevant data preprocessing approaches improved the prediction results, especially on a dataset with photos in-the-wild.

5. Conclusion and Future Works

In this work, a method for automatic attractiveness assessment that employs transfer learning methodology as well as Facenet face classification model is proposed. The network designed and pre-trained specially for face recognition was exploited to extract high-level and robust features of face images. Then these features are used to train Support Vector Regression algorithm in order to predict facial attractiveness.

Extensive experiments were conducted on datasets with and without restrictions on posture, expression, blurriness, illumination, in other words, with photos in constrained (SCUT-FBP 5500) and unconstrained environment (Gray dataset). The most effective data preprocessing approach and pipeline were found for both types of face images. The experimental results proved that the proposed method achieved the improved or comparable performance and demonstrated its effectiveness in both constrained and unconstrained environment.

6. References

- [1] Altwaijry H and Belongie S 2013 Relative ranking of facial attractiveness *In Applications of Computer Vision (WACV) 2013 IEEE Workshop on. IEEE* **7**
- [2] Cao K, Choi K, Jung H and Duan L 2020 Deep learning for facial beauty prediction *Information* **11** 8
- [3] Cao Q, Shen L, Xie W, Parkhi O and Zisserman A 2017 VGGFace2: a dataset for recognizing faces across pose and age **10**
- [4] Donahue J, Jia Y, Vinyals O, Hoffman J, Zhang N, Tzeng E and Darrell T 2014. Decaf: a deep convolutional activation feature for generic visual recognition. *In International conference on machine learning* 647–655.
- [5] Dornaika F, Wang K, Arganda-Carreras I, Elorza A and Moujahid A 2020 Toward graph-based semi-supervised face beauty prediction. *Expert Systems with Applications* **142**, 112990.
- [6] Eysenck H, Dror G and Ruppin E 2006 Facial attractiveness: beauty and the machine *Neural Computation* **18**, 1 (2006), 119–142.
- [7] Fan J, Chau K, Wan X, Zhai L and Lau E 2012 Prediction of facial attractiveness from facial proportions *Pattern Recognition* **45**, 6 (2012), 2326–2334.
- [8] Gan J, Li L, Zhai Y and Liu Y 2014 Deep self-taught learning for facial beauty prediction *Neurocomputing* **144** (2014), 295–303.
- [9] Gan J, Xiang L, Zhai Y, Mai C, He G, Zeng J, Bai Z, Donida Labati R, Piuri V and Scotti F 2020 2M BeautyNet: Facial Beauty Prediction Based on Multi-Task Transfer Learning *IEEE Access* **8**, 20245–20256.
- [10] Gray D, Yu K, Xu W and Gong Y 2010 Predicting facial beauty without landmarks *In European Conference on Computer Vision. Springer*, 434–447.
- [11] Kagan A, Dror G, Leyvand T, Cohen-Or D and Eytan Ruppin 2007 A humanlike predictor of facial attractiveness. *In Advances in Neural Information Processing Systems*. 649–656.
- [12] Kalayci S, Ekenel H and Gunes H 2014 Automatic analysis of facial attractiveness from video. *In Image Processing (ICIP), 2014 IEEE International Conference on. IEEE*, 4191–4195.
- [13] Krizhevsky A, Sutskever I and Hinton G 2012 Imagenet classification with deep convolutional neural networks. *In Advances in neural information processing systems*.

1097–1105.

- [14] Li J, Xiong Ch, Liu L, Shu X and Yan Sh 2015 Deep face beautification. In *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 793–794.
- [15] Liang L, Lin L, Jin L, Xie D and Li M 2018 SCUT-FBP5500: A diverse benchmark dataset for multi-paradigm facial beauty prediction. In *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, 1598–1603.
- [16] Liu L, Xing J, Liu S, Xu H, Zhou X and Yan 2014 Wow! you are so beautiful today! *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 11, 1s (2014), 20.
- [17] Liu X, Li T, Peng H, Chuoying Ouyang I, Kim T and Wang R 2019 Understanding beauty via deep facial features. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 0–0.
- [18] Murray N, Marchesotti L and Perronnin F 2012 AVA: A large-scale database for aesthetic visual analysis In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2408–2415.
- [19] Rothe R, Timofte R and Van Gool L 2016 Some like it hot-visual guidance for preference prediction. In *Proceedings CVPR 2016*. 1–9.
- [20] Schroff F, Kalenichenko D and Philbin J 2015 Facenet: A unified embedding for face recognition and clustering In *Proceedings of the IEEE conference on computer vision and pattern recognition* 815–823.
- [21] Simonyan K and Zisserman A 2014 Very deep convolutional networks for large-scale image recognition (*Preprint gr-qc/1409.1556*).
- [22] Sutić D, Brešković I, Huić R and Jukić I 2010 Automatic evaluation of facial attractiveness In *MIPRO, 2010 Proceedings of the 33rd International Convention IEEE*, 1339–1342.
- [23] Viola P and Jones M 2001 Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, Vol. 1. IEEE, I–I.
- [24] Wang S, Shao M and Yun Fu 2014 Attractive or not?: Beauty prediction with attractiveness-aware encoders and robust late fusion. In *Proceedings of the 22nd ACM international conference on Multimedia*. ACM, 805–808.
- [25] Xie D, Liang L, Jin L, Xu J and Li M 2015 SCUT-FBP: A benchmark dataset for facial beauty perception. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*. IEEE, 1821–1826.
- [26] Xu J, Jin L, Liang L, Feng Z and Xie D 2015 A new humanlike facial attractiveness predictor with cascaded fine-tuning deep learning model (*Preprint gr-qc/1511.02465*).
- [27] Xu L, Fan H and Xiang J 2019. Hierarchical Multi-Task Network For Race, Gender and Facial Attractiveness Recognition. In *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, 3861–3865.
- [28] Xu L, Xiang J and Yuan X 2018. Transferring Rich Deep Features for Facial Beauty Prediction (*Preprint gr-qc/1803.07253*).
- [29] Zhai Y, Huang Y, Xu Y, Gan J, Cao H, Deng W, Donida Labati R, Piuri V and Scotti F. 2020 Asian Female Facial Beauty Prediction Using Deep Neural Networks via Transfer Learning and Multi-Channel Feature Fusion. *IEEE Access* 8 (2020), 56892–56907.

Acknowledgements

This research is financially supported by The National Key Research and Development Program of China (grant number 2018YFC0807105) National Natural Science Foundation of China (grant number 61462073) and Science and Technology Committee of Shanghai Municipality (STCSM) (under grant numbers 17DZ1101003 , 18511106602 and 18DZ2252300).

Partially Supported by Open Funding Project of the State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai, China.

Reproduced with permission of copyright owner. Further reproduction
prohibited without permission.