

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# A Deep Learning-Based Aesthetic Surgery Recommendation System

*Phan Chau Phuc Thinh, Bui Thi Xuyen,  
Nguyen Do Trung Chanh, Dao Huu Hung  
and Mimura Daisuke*

## Abstract

We propose in this chapter a deep learning-based recommendation system for aesthetic surgery, composing of a mobile application and a deep learning model. The deep learning model built based on the dataset of before- and after-surgery facial images can estimate the probability of the perfection of some parts of a face. In this study, we focus on the most two popular treatments: rejuvenation treatment and eye double-fold surgery. It is assumed that the outcomes of our history surgeries are perfect. Firstly a convolutional autoencoder is trained by eye images before and after surgery captured from various angles. The trained encoder is utilized to extract learned generic eye features. Secondly, the encoder is further trained by pairs of image samples, captured before and after surgery, to predict the probability of perfection, so-called perfection score. Based on this score, the system would suggest whether some sorts of specific aesthetic surgeries should be performed. We preliminarily achieve 88.9 and 93.1% accuracy on rejuvenation treatment and eye double-fold surgery, respectively.

**Keywords:** aesthetic surgery, rejuvenation treatment, eye double-fold surgery, recommendation system, convolutional neural network, autoencoder

## 1. Introduction

Plastic surgery is a surgical specialty relating to restoration, reconstruction, or alteration of the human body. There are two major categories: (1) reconstructive surgery and (2) aesthetic or cosmetic surgery. The former is intended to correct dysfunctional areas of the body and is reconstructive in nature. Examples of this kind include breast reconstruction, burn repair surgery, congenital defect repair, lower extremity reconstruction, hand surgery, scar revision surgery, etc. The latter focuses on enhancing the appearance of the patient. Improving aesthetic appeal, symmetry, and proportion are among the key goals. The scope of aesthetic surgery procedures includes breast enhancement, facial contouring, facial rejuvenation, body contouring, skin rejuvenation, etc. The scope of this chapter restricts to the latter, aesthetic surgery.

In aesthetic surgery, the treated areas function properly; it is optional based on the willingness of the patient who cares about their beauty. The sense of beauty also varies from geographical areas and sometimes follows either local or global fashion trends. Therefore, the consultation in aesthetic surgery of experienced doctors is extremely important. The severe problem of population aging in developed

countries leads to the shortage of high-skill labors in almost all industrial sectors. Thus this chapter proposes a deep learning-based aesthetic surgery recommendation system, aiming at keeping the valuable know-how of experienced doctors to consult the patient in aesthetic surgery. Moreover, the continuous learning capability of the AI model also facilitates the self-update of the newly fashionable know-how in this field, given a set of rich training data.

Although aesthetic surgery can be performed on all areas of the head, neck, and body, our focused areas in this chapter are the facial area. We take the most popular treatments for facial areas, rejuvenation, and eye double-fold surgery into consideration. In order to build a deep learning system which is capable of predicting the perfection of aesthetic surgery, we collect an in-house training dataset composing of pairs of images capturing the eye area of the same person before and after aesthetic surgery. It is assumed that the beauty of facial areas after surgery is perfect, that is, the know-how of an aesthetic surgeon is embedded into these pairs of images.

In order to keep the know-how of experienced aesthetic surgeon, we propose to train a deep neural network by these pairs of images in our in-house dataset. Among various kinds of neural network architectures, proposed in the literature, convolutional neural networks (CNN) have been demonstrating outstanding performance in image recognition [1]. This was the first time a large and deep CNN—AlexNet model—achieved record-breaking results on highly challenging image recognition dataset with a margin of more than 10% with respect to the second best which makes use of handcrafted features. Even though the performance of AlexNet is still far from the inferotemporal pathway of the human visual system, it created the way of success for successor models such as Inception [2], VGG [3], and Resnet [4]. The convolutional layers learn from data to extract a rich set of features for a variety of purposes such as image classification and recognition [4], visual tracking [5], face recognition [6], object detection [7], person reidentification [8], etc. The power of CNN is enabled by the learning mechanism in which weights of convolutional filters are adjusted toward the adaption to the labels. The generalization of CNN is guaranteed by the availability of a huge dataset to produce an outstanding performance on unseen data.

However, our in-house dataset is not huge enough to guarantee the generalization of CNN for this task. Therefore, we propose to use convolutional autoencoder neural networks to overcome the limitation of our small dataset. The network is firstly trained in a layer-wise mechanism to reconstruct input images in the output layer. This training mechanism is completely unsupervised. After the convolutional autoencoder neural network is trained, the decoder part is truncated. Only the encoder part is kept and is concatenated with fully connected layers. The whole network will be trained by images and their labels, before and after surgery. The weights of the encoder parts are kept intact because the encoder parts have already learned the key features of the training image set. As a result, our proposed model is able to achieve 88.9 and 93.1% accuracy on rejuvenation treatment and eye double-fold surgery, respectively.

The rest of this chapter is organized as follows. Section 2 describes our contribution to the backdrop of related work. The proposed method is presented in detail in Section 3. Finally, we conclude the chapter and delineate future work in Section 4.

## **2. Related work**

The number of aesthetic surgery, particularly for facial areas, has been recently drastically surged. This trend even could spread further in the next few years

due to lowering the average cost of such treatments and the desire of beautification. Numerous research works have been proposed in the literature, especially in the computer vision community to address the challenges posed by aesthetic surgery. These research works are generally broken into three categories, namely, skin quality inspection, face recognition after surgery, and surgery planning and recommendation.

Aesthetic surgery in facial areas to correct facial feature anomaly and to improve the beauty, in general, alters the original facial information. It poses a significant challenge for face recognition algorithms. Majority of proposed methods in the literature focused on the advances of handcraft features. Richa et al. [9] have investigated the effects of aesthetic surgery to face recognition. Amal et al. [10] proposed a face recognition system based on LBP and GIST descriptor to address this problem. Maria et al. [11] combine two methods, face recognition against occlusions and expression variations (FARO) [12] and face analysis for commercial entities (FACE) [13] with split face architecture to deal with the effects of plastic surgery to process each face region as separate biometrics. For a comprehensive survey of face recognition algorithms against variations due to aesthetic surgery, please refer to [14].

Skin quality inspection and assessment is also a potential application of computer vision and deep learning methods. By assessing skin quality, it is able to help the aesthetic surgeon to recommend certain kinds of operation to enhance the beauty of the face. Surface roughness, wrinkle depth, volume, and epidermal thickness of the skin are quantitatively computed by applying deep learning method to images captured by optical coherence tomography [15]. The facial skin is classified into facial skin patches, namely, normal, spots, and wrinkles, by using convolutional neural networks [16]. Batool and Chellappa [17] proposed a method to model wrinkles as texture features or curvilinear objects, so-called aging skin texture, for facial aging analysis. They reviewed commonly used image features to capture the intensity gradients caused by facial wrinkles, such as Laplacian of Gaussian, Hessian filter, steerable filter bank, Gabor filter bank, active appearance model, and local binary patterns.

In the last category of aesthetic surgery planning and recommendation, facial beauty prediction is the first step to assess whether or not a face should perform an aesthetic surgery. Yikui et al. [18] described BeautyNet in which multiscale CNN is employed to obtain deep features, characterizing the facial beauty, and is combined with transfer learning strategy to alleviate overfitting and to achieve robust performance on unconstrained faces. Lu et al. [19] transferred rich deep features from a pretrained VGG16 model on face verification task to Bayesian ridge regression algorithms for predicting facial beauty.

Going beyond the facial beauty prediction, Arakawa and Nomoto [20] removed wrinkles and spots while preserving natural skin roughness by using a bank of nonlinear filters for facial beautification. Eighty-four facial landmark points are represented in a vector of 234 normalized lengths to compare with vectors of beautiful faces to suggest how to warp the triangulation of the original face to the beautiful ones [21]. Bottino et al. [22] presented a quantitative approach to automatically recommend effective patient-specific improvements of facial attractiveness by comparing the face of the patient with a large database of attractive faces. Simulations are performed by applying features of similar attractive faces into the patient faces with a suitable morphing of facial shape.

Our research differs from the above-related works in two senses. Firstly, a convolutional autoencoder is employed to learn rich features and characterize both unattractive and beautiful faces in an unsupervised manner, rather than under supervised learning [18, 19]. The learned features are more discriminative than handcrafted features [9–13, 15–17]. Secondly, the proposed deep learning

framework facilitates a holistic approach to identify what kinds of facial treatment should be performed to enhance the attractiveness, rather than predicting beauty score [20–22].

### 3. Method

The method is divided into two main steps, namely, train feature extractor and train classification model as shown in **Figure 1**. The dataset is first utilized to train the feature extraction model based on convolutional autoencoder. In this step, the convolutional autoencoder is trained in order to learn to encode the input in a set of images and then tries to reconstruct the input from them. Thus, it can learn the feature of the input data by minimizing the reconstruction error. Then we extract the encoder and use it as the feature extractor.

**Figure 2** illustrates the method of training feature extractor. From the first step, the model is trained only to learn filters able to extract feature that can be used to reconstruct the input. These filters in the encoder then are extracted and then utilized as the feature extraction for the classification model which is a fully connected layer.

#### 3.1 Feature extractor

Autoencoders are well-known unsupervised learning algorithm whose original purpose is to find latent lower-dimensional state spaces of datasets, but they are also capable of solving other problems, such as image denoising, enhancement, or colorization. The main idea behind autoencoders is to reduce the input into a latent state space with lower dimensions and then try to reconstruct the input from this representation. Thus the autoencoder uses its input as the reference of the output in the learning phase. The two parts are called encoder and decoder, respectively. By reducing the number of variables which represent the data, we force the model to learn how to keep only meaningful information, from which the input is reconstructable. It can also be viewed as a compression technique as shown in **Figure 2**.

A conventional autoencoder is composed of two layers, corresponding to the encoder  $f_w(\cdot)$  and decoder  $g_v(\cdot)$ . It aims to find a code for each input that minimizes the difference between input,  $x_i$ , and output,  $g_u(f_w(x_i))$ , over all samples [23]:

$$\min_{W,U} \frac{1}{n} \sum_{i=1}^n \|g_u(f_w(x_i)) - x_i\|_2^2, \quad (1)$$

In the fully connected autoencoder,

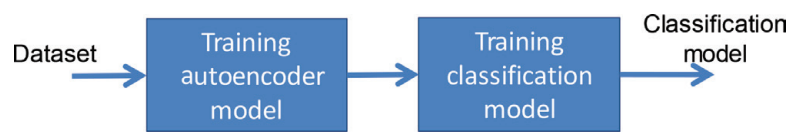
$$\begin{aligned} f_W(x) &= \sigma(Wx) \equiv h \\ g_U(h) &= \sigma(Uh) \end{aligned} \quad (2)$$

where  $x$  and  $h$  are vectors,  $W$  is the learn weights, and  $\sigma$  is the activation function. After learning, the embedded vector  $h$  is a unique representation for input. In our application, the convolutional autoencoders (CAE) are defined as

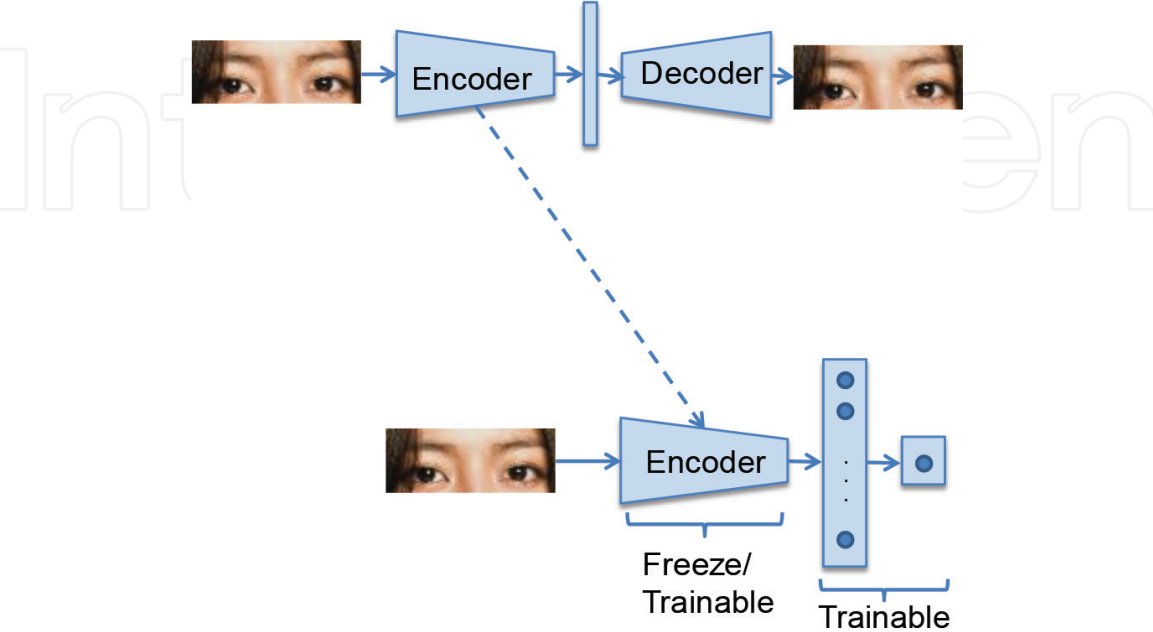
$$\begin{aligned} f_W(x) &= \sigma(x * W) \equiv h \\ g_U(h) &= \sigma(h * U) \end{aligned} \quad (3)$$

where  $x$  and  $h$  are the matrix or tensor and “\*” is the convolutional operator.





**Figure 1.**  
Flowchart of the method.



**Figure 2.**  
Method of training the feature extractor.

We propose the CAE-based feature extraction method that learns the generic feature of the face. The encoder serves as the feature extractor that encodes the image into vector  $h$  that represent the image of the facial part, for example, the eyes.

In **Table 1**, the several model structures that we use in our experiment were shown. We tried four models. The differences between the four models are the number of layers and dimension of the embedded vector. The number of layers in autoencoders vary between 3 and 4 layers, while the dimensions of the embedded vector are 32 and 64.

### 3.2 Perfection prediction system

After training the autoencoder, it will serve as the feature extractor as shown in **Figure 2**. In this step, we apply the transfer learning method to transfer the well-learned filters for facial part feature extraction. We try two different lengths of the embedding vector, say, 32 and 64, as shown in **Table 1**. We also try both to freeze and to retrain the extractor.

This model predicts the probability of perfection; the output range of the model is a one-dimensional vector with the interval of  $[0-1]$  that reflects the probability of perfection. The probability of perfection is defined as follows:

- If the face is perfect (no surgery is needed), the probability is 1.
- If the face is not perfect (surgery is needed to be performed), the probability is 0.

However, in a real situation, we cannot obtain a big dataset of the perfect/non-perfect face. Thus, we assume that the outcome of surgery is perfect (value is 1) and the original face is not perfect (value is 0).

3 layers																							
Autoencoder 1	Encoder	Layer	#kernel	kernel size	stride	input size		output size		Autoencoder 2	Encoder	Layer	#kernel	kernel size	stride	input size		output size					
		conv*	32	5*5	2*2	256	80	3	64			20	32	conv*	16	5*5	2*2	256	80	3	64	20	16
		conv*	32	3*3	1*1	64	20	32	32			10	32	conv*	16	3*3	1*1	64	20	16	32	10	16
	Decoder	conv	64	3*3	1*1	32	10	32	32		10	64	Decoder	conv	32	3*3	1*1	32	10	16	32	10	32
		deconv*	32	3*3	1*1	32	10	64	64		20	32		deconv*	16	3*3	1*1	32	10	32	64	20	16
		deconv*	32	3*3	1*1	64	20	32	128		40	32		deconv*	16	3*3	1*1	64	20	16	128	40	16
Model classifier 1	deconv	3	5*5	2*2	128	40	32	256	80	3	deconv	3		5*5	2*2	128	40	16	256	80	3		
	Layer	output							Model classifier 2	Layer	output												
	Encoder									Encoder													
Dense	64							Dense		32													
Dense	1									Dense	1												
4 layers																							
Autoencoder 3	Encoder	Layer	#kernel	kernel size	stride	input size		output size		Autoencoder 4	Encoder	Layer	#kernel	kernel size	stride	input size		output size					
		conv*	32	5*5	2*2	256	80	3	64			20	32	conv*	16	5*5	2*2	256	80	3	64	20	16
		conv*	32	3*3	1*1	64	20	32	32			10	32	conv*	16	3*3	1*1	64	20	16	32	10	16
	Decoder	conv*	64	3*3	1*1	32	10	32	16		5	64	Decoder	conv*	32	3*3	1*1	32	10	16	16	5	32
		conv	64	3*3	1*1	16	5	64	16		5	64		conv	32	3*3	1*1	16	5	32	16	5	32
		deconv*	64	3*3	1*1	16	5	64	32		10	64		deconv*	32	3*3	1*1	16	5	32	32	10	32
Model classifier 3	deconv*	32	3*3	1*1	32	10	64	64	20	32	deconv*	16		3*3	1*1	32	10	32	64	20	16		
	deconv*	32	3*3	1*1	64	20	32	128	40	32	deconv*	16		3*3	1*1	64	20	16	128	40	16		
	deconv	3	5*5	2*2	128	40	32	256	80	3	deconv	3	5*5	2*2	128	40	16	256	80	3			
Model classifier 3	Layer	output							Model classifier 4	Layer	output												
	Encoder									Encoder													
	Dense	64 <th colspan="2"></th> <th colspan="2"></th> <th colspan="2"></th> <th>Dense</th> <td>32<th colspan="2"></th><th colspan="2"></th><th colspan="2"></th></td>								Dense	32 <th colspan="2"></th> <th colspan="2"></th> <th colspan="2"></th>												
	Dense	1 <th colspan="2"></th> <th colspan="2"></th> <th colspan="2"></th> <th>Dense</th> <td>1<th colspan="2"></th><th colspan="2"></th><th colspan="2"></th></td>								Dense	1 <th colspan="2"></th> <th colspan="2"></th> <th colspan="2"></th>												
conv: convolution. conv*: convolution and max-pooling with filter size (2.2). deconv: transposed convolution. deconv*: (2.2) upsampling with transposed convolution.																							

Table 1. Autoencoder model and classification model.

4. Experiment

4.1 Dataset preparation

In the dataset, the total number of images of rejuvenation treatment and eye double-fold surgery are 5585 and 36,598 pairs of the images, respectively. The data is filtered to select a good image for training. Some errors in the images are not well aligned, only one eye, etc. The reject rates are 82.15 and 59.92%, respectively. The image was crop around the eye area. Then the data is divided into 70, 15, and 15% for training, validation, and testing, respectively. As mentioned earlier, the output of the model is the probability of perfection. Hence, our training data include the image of before and after surgery as shown in Figure 3.

4.2 Prediction accuracy

We chose a thresholding value for the prediction model. If the predicted value is higher than the thresholding value, the face is predicted to be perfect. From that, we obtain accuracy of the different models as shown in Tables 2 and 3. We have two schemes of training in which the feature extractors are freezed (not training together with the classification model) and trainable, resulting in eight models in these two tables. When the feature extractor is freezed, we believe that it has already captured universal features such as edge and curves which is relevant to our tasks. Therefore, we want to keep the weights of the feature extract intact. However, in the second training scheme, we apply different learning rates for the feature extractor and the fully connected layers. The learning rate of the feature extractor is much smaller than that of the fully connected layer because we believe that the weights of the feature extractor is good enough for our tasks and we do not want to distort them too quickly and too much during the training of the classification model. These above training schemes are the best common practices when fine-tuning deep neural networks. We tried both of them, resulting in the following. The best model for double-fold surgery and rejuvenation treatment are Models 1 and 8 (see Tables 2 and 3 for more details), with accuracy on the test dataset of 93.1 and 88.9%, respectively. Model 1 is the model when the encoder



**Figure 3.**  
*Example of original and outcome of surgery.*

Name	Classifier	Training Method	Result on test-data		Result on validation-data		Filter 1 - accuracy	Filter 2 - loss function
			Loss function	Accuracy	Loss function	Accuracy		
Model 1	Classifier 1	Free encoder	0.200	0.931	0.189	0.938	0.935	0.195
Model 2	Classifier 1	Trainable encoder	0.255	0.923	0.199	0.944	0.934	0.227
Model 3	Classifier 2	Free encoder	0.204	0.924	0.179	0.932	0.928	0.192
Model 4	Classifier 2	Trainable encoder	0.368	0.929	0.318	0.940	0.935	0.343
Model 5	Classifier 3	Free encoder	0.194	0.927	0.165	0.939	0.933	0.180
Model 6	Classifier 3	Trainable encoder	0.354	0.939	0.275	0.950	0.945	0.315
Model 7	Classifier 4	Free encoder	0.192	0.925	0.172	0.935	0.930	0.182
Model 8	Classifier 4	Trainable encoder	0.219	0.928	0.173	0.948	0.938	0.196

**Table 2.**  
*Testing accuracy for eye double-fold surgery.*

Name	Classifier	Training Method	Result on test-data		Result on validation-data		Filter 1 - accuracy	Filter 2 - loss function
			Loss function	Accuracy	Loss function	Accuracy		
Model 1	Classifier 1	Free encoder	0.506	0.885	0.859	0.854	0.870	0.683
Model 2	Classifier 1	Trainable encoder	0.938	0.861	0.782	0.902	0.882	0.860
Model 3	Classifier 2	Free encoder	0.370	0.857	0.469	0.844	0.851	0.420
Model 4	Classifier 2	Trainable encoder	0.783	0.895	0.969	0.885	0.890	0.876
Model 5	Classifier 3	Free encoder	0.328	0.854	0.410	0.868	0.861	0.369
Model 6	Classifier 3	Trainable encoder	0.618	0.892	0.594	0.906	0.899	0.606
Model 7	Classifier 4	Free encoder	0.362	0.861	0.542	0.861	0.861	0.452
Model 8	Classifier 4	Trainable encoder	0.662	0.889	0.512	0.923	0.906	0.587

**Table 3.**  
*Testing accuracy for rejuvenation treatment.*

was freezed. However, in Model 8, the encoder was retrained. However, the accuracy differences between the best model and the second best model are less than 1%.

### 5. Conclusion

We have presented in this chapter an interesting application of deep learning in aesthetic surgery recommendation along with its encouraging results. By using our system, the presented deep learning engine will provide a reference decision of taking either rejuvenation treatment or eye double-fold surgery or not to both the surgeon and the patient, just based on the eye photo of the patient. To this end, we trained a deep autoencoder by our in-house dataset, composing of pairs of images captured before and after the surgery. The trained encoder part learned in an unsupervised manner, a rich set of features, characterized both unattractive and beautiful facial features. We concatenate the trained encoder part to a fully connected layer to predict perfection score of an eye photo of a patient, based on which a decision of taking treatment or not will be made.



Even though our preliminary results are promising with 88.9 and 93.1% accuracy on rejuvenation treatment and eye double-fold surgery, respectively, it still has much rooms for improvement. Firstly, we should improve the dimension of our in-house dataset by encouraging more patients to participate in our program, so that we are able to build a deeper network. More and more layers are added; richer and richer learned features are obtained to improve the accuracy of our system. Secondly, we are going to expand the capability of our system to deal with more kinds of treatments to diversify and provide the best services to our clients, rather than focusing on the two above treatment and surgery.

### **Author details**

Phan Chau Phuc Thinh<sup>1</sup>, Bui Thi Xuyen<sup>1</sup>, Nguyen Do Trung Chanh<sup>1</sup>,  
Dao Huu Hung<sup>1\*</sup> and Mimura Daisuke<sup>2</sup>

1 Data Science Laboratory, FPT Software Japan Co. Ltd., Tokyo, Japan

2 Shonan Beauty Clinic, Tokyo, Japan

\*Address all correspondence to: [hungdh3@fsoft.com.vn](mailto:hungdh3@fsoft.com.vn)

### **IntechOpen**

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Alex K, Ilya S, Geoffrey EH. ImageNet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*. 2012
- [2] Christian S, Wei L, Yangqing J, Pierre S, Scott R, Dragomir A, et al. Going deeper with convolutions. In: *Computer Vision and Pattern Recognition*. 2014
- [3] Karen S, Andrew Z. Very deep convolutional networks for large-scale image recognition. In: *Computer Vision and Pattern Recognition*. 2014
- [4] Kaiming H, Xiangyu Z, Shaoqing R, Jian S. Deep residual learning for image recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2015
- [5] Naiyan W, Dit-Yan Y. Learning a deep compact image representation for visual tracking. In: *Advances in Neural Information Processing Systems*. 2013
- [6] Florian S, Dmitry K, James P. FaceNet: A unified embedding for face recognition and clustering. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2015
- [7] Joseph R, Santosh D, Ross G, Ali F. You only look once: Unified, real-time object detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2016
- [8] Bahram L, Mehdi FS, Ihsan U. Survey on deep learning techniques for person re-identification task. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2018
- [9] Richa S, Mayank V, Afzel N. Effects of plastic surgery on face recognition: A preliminary study. In: *IEEE Conference on Computer Vision and Pattern Recognition*; Minami. 2009
- [10] Amal SOA, Vijanth S, Aamir M, Azrina A. Proposed face recognition system after plastic surgery. *IET Computer Vision*. 2016;**10**(5):344-350
- [11] Maria DM, Michele N, Daniel R, Harry W. Robust face recognition after plastic surgery using region-based approaches. *Pattern Recognition*. 2015;**48**(4):1261-1276
- [12] De Marsico M, Nappi M, Riccio D. FARO: Face recognition against occlusions and expression variations. *IEEE Transactions on Systems, Man, and Cybernetics, Part A Systems and Humans*. 2010;**40**(1):121-132
- [13] De Marsico M, Nappi M, Riccio D. FACE: Face analysis for commercial entities. In: *IEEE International Conference on Image Processing*; Hong Kong. 2010
- [14] Michele N, Stefano R, Massimo T. When plastic surgery challenges face recognition. *Image and Vision Computing*. 2016;**54**:71-82
- [15] Sanzhar A, Yujin A, Jiho B, Andrey V, Gil-Jin J, Pilun K, et al. Quantitative classification of OCT skin images with deep learning. In: *Proceedings SPIE 10467, Photonics in Dermatology and Plastic Surgery*. 2018
- [16] Alarifi JS, Goyal M, Davison A, Dancey D, Khan R, Yap M. Facial skin classification using convolutional neural networks. In: *Image Analysis and Recognition ICIAR*. 2017
- [17] Nazre B, Rama C. Modeling of facial wrinkles for applications in computer vision. In: *Advances in Face Detection and Facial Image Analysis*. Springer; 2016. pp. 299-332. <https://www.springerprofessional.de/en/modeling-of-facial-wrinkles-for-applications-in-computer-vision/9975942>

[18] Yikui Z, He C, Wenbo D, Junying G, Vincenzo P, Junying Z. BeautyNet: Joint multiscale CNN and transfer learning method for unconstrained facial beauty prediction. *Computational Intelligence and Neuroscience*. 2019;**2019**:1-14

[19] Lu X, Jinhai X, Xiaohui Y. Transferring rich deep features for facial beauty prediction. 2018. pp. 1-6. arXiv:1803.07253 [cs.CV]

[20] Arakawa K, Nomoto K. A system for beautifying face images using interactive evolutionary computing. In: *International Symposium on Intelligent Signal Processing and Communication Systems*. 2005

[21] Tommer L, Daniel C-O, Gideon D, Dani L. Data-driven enhancement of facial attractiveness. In: *Proceeding ACM Siggraph*. 2008

[22] Bottino A, Laurentini A, Rosano L. A new computer-aided technique for planning the aesthetic outcome of plastic surgery. In: *International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, Plzen—Bory*. 2008

[23] Guo X, Liu X, Zhu E, Yin J. Deep clustering with convolutional autoencoders. In: *International Conference on Neural Information Processing*. 2017