

A hierarchically trained generative network for robust facial symmetrization

Shu Zhang^b, Ting Wang^{a,*}, Yanjun Peng^a and Junyu Dong^b
^a*Shandong University of Science and Technology, Qingdao, Shandong, China*
^b*Ocean University of China, Qingdao, Shandong, China*

Abstract. Face symmetrization has extensive applications in both medical and academic fields, such as facial disorder diagnosis. Human face possesses an important characteristic, which is as known as symmetry. However, in many scenarios, the perfect symmetry doesn't exist in human faces, which yields a large number of studies around this topic. For example, facial palsy evaluation, facial beauty evaluation based on facial symmetry analysis, and many among others. Currently, there are still very limited researches dedicated for automatic facial symmetrization. Most of the existing studies only utilized their own implementations for facial symmetrization to assist their interdisciplinary academic researches. Limitations thus can be noticed in their methods, such as the requirements for manual interventions. Furthermore, most existing methods utilize facial landmark detection algorithms for automatic facial symmetrization. Though accuracies of the landmark detection algorithms are promising, the uncontrolled conditions in the facial images can still negatively impact the performance of the symmetrical face production. To this end, this paper presents a joint-loss enhanced deep generative network model for automatic facial symmetrization, which is achieved by a full facial image analysis. The joint-loss consists of a pair of adversarial losses and an identity loss. The adversarial losses try to make the generated symmetrical face as realistic as possible, while the identity loss helps to constrain the output to have the same identity of the person in the original input as much as possible. Rather than an end-to-end learning strategy, the proposed model is constructed by a multi-stage training process, which avoids the demand for a large size of the symmetrical face as training data. Experiments are conducted with comparisons with several existing methods based on some of the most popular facial landmark detection algorithms. Competitive results of the proposed method are demonstrated.

Keywords: Facial symmetry, facial palsy image, symmetrization, multi-stage training, Generative Adversarial Networks

1. Introduction

Human face possesses an important character, which is as known as facial symmetry. However, in practical scenarios, many conditions can compromise this symmetry. Asymmetry can always be spot between left and right half faces with or without facial expressions. It is most obvious especially for patients with facial disorders like facial palsy or facial nerve paralysis. Due to the malfunctions of the controls for the facial muscles, facial disordered patients have difficulties to perform normal facial expressions, which results in certain facial asymmetries to various extents. Most existing diagnose approaches for such facial disorders are based on the evaluations of the facial symmetry degrees. The evaluations highly relies on the predictions of the facial symmetries. A facial symmetrization process is therefore commonly utilized in most of the existing researches, which provides a perfectly symmetrical face based on a given face. The performance of the facial symmetrization is thus critical to those facial

*Corresponding author: Ting Wang, Shandong University of Science and Technology, Qingdao, Shandong, China. E-mail: wangting@sdust.edu.cn.

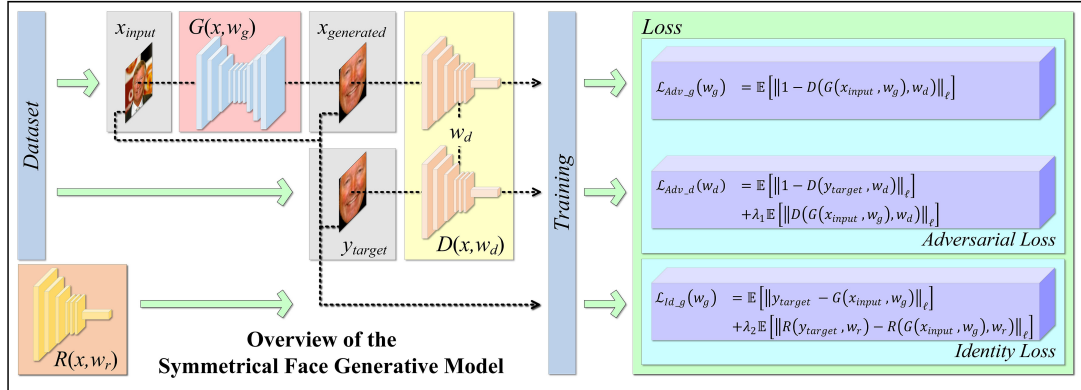


Fig. 1. The overview of the symmetrical face generative model, which consists of a generator (red block), a discriminator (yellow block), and a facial identifier (orange block). The generator tries to produce a symmetrical face that can fool the discriminator block-by-block. The facial identifier helps to provide the metrics for identity loss of the model. The identifier network is trained independently with respect to the facial symmetrization model.

disorder diagnose applications, as well as many other studies involving facial symmetry based facial analysis.

Most of the existing approaches for face symmetrization are achieved by the facial regional geometric calculations based on the facial landmark extractions. However, the accurate of those extracted landmarks can drop dramatically under certain circumstances, for example, the variations in the head poses, illuminations, occlusions and many other aspects. Especially, the traditional facial landmark detection can fail for most of the facial data from facial disordered patients. To avoid such error accumulation during the process, this paper presents a novel method for facial symmetrization by full facial image analysis. A joint-loss enhanced generative adversarial network model is proposed as shown in Fig. 1, which takes a facial image as input and tries to generate a symmetrical face that has the same identity of the person in the original input data. A supervised multi-stage training process is utilized to achieve the model learning.

The contributions of this paper are as follows:

A novel joint-loss enhanced model based on the Deep Convolutional Generative Adversarial Networks (DCGAN) is proposed for facial image synthesis. It utilizes a Variational Autoencoder (VAE) as the generator, a Convolutional Neural Networks (CNN) as the discriminator, and a CNN as the facial identifier network, as Fig. 1 shows. The joint-loss consists of a pair of adversarial losses and an identity loss.

A multi-stage training process is presented, which can learn the model parameters in a coarse-to-fine manner, as shown in Fig. 2. It avoids the demand for large amount of the symmetrical face training data.

A competitive performance is demonstrated compared with several existing methods based on a range of most popular facial landmark detection algorithms.

2. Related work

The generating of the symmetrical face has a great number of applications in many fields such as computer vision assisted bio-medical cares. For example, most computer vision based facial palsy studies are deeply coupled with facial symmetry analysis to visually and metrically diagnose the facial regional movement functions [1–4]. Apart from this, the facial symmetry is also an important factor to evaluate

the facial beauty or facial attractiveness reported in many studies [5,6]. Besides those medical studies, the analysis of the facial symmetry can also contribute to many computer vision based bio-information applications, such as facial frontalisation [7–9], facial feature detection [10], facial recognition [11] and many other facial manipulations [12–14].

Despite such large applicational and academic demands in medical and bio-information fields, there are still very limited researches dedicated for symmetrical face generating. Most of the studies only utilize their own implementations of symmetrical face generating for their studies in other academic fields. For example, Chang et al. [15] presented a study on a pediatric facial palsy quantification based on Sunnybrook grading system, which utilized facial symmetry for the grading process. They employed digital image correlation technique in the help of facial symmetry analysis. Their method however still involved manual interventions during the process. Song et al. [16] introduced a research into the facial nerve paralysis assessment based on the facial asymmetry. They utilized a facial landmark detection method implemented in dlib to find the facial symmetry axis for facial symmetrization. Katsumi et al. [17] explored the quantitative analysis of facial palsy by a 3D facial measurement system. Their method utilized the facial symmetry obtained by a self-defined 3D facial landmark system, which consisted of 9 landmarks in total for a face. Hassner et al. [18] explored the face frontalisation problem with the consideration of the facial symmetry. Their facial symmetry model, called soft-symmetry, was achieved with the facial features extracted by Supervised Descent Model (SDM) [19], which is one of the most popular methods for facial landmark localization. Xu et al. [11] discussed the facial recognition based on the facial symmetry property. Their method was based on the generated virtual facial images that were axis-symmetrical. Those virtual facial images were synthesized by flipping the original face that were manually centered in the images beforehand. Their method required manual interventions during the symmetrical face preparation step. Harguess and Aggarwal [20] proposed a study to explore the relationship between the facial symmetry and the face recognition. They included the discussions to utilize an average-half-face [21] to obtain a symmetrical face. To achieve the average-half-face, the facial image was firstly centred about the nose, then divided into two symmetric halves, and finally averaged altogether. However, in their paper, they didn't mention about how to localize the nose automatically. In practice, this can be done by either manual annotations or facial landmark detection algorithms, such as AAM [22], SDM [19], CLM [23] or CLNF [24].

Up to now, almost all the existing researches related to face symmetrization utilize a facial landmark extraction algorithm to find a vertical axis, which can symmetrize the face. Despite that the performance of the facial landmark extraction algorithms is promising, the facial landmark based face symmetry axis estimation can still be less accurate since there are so many uncontrolled conditions in the facial images, such as different facial expressions, head poses and face occlusions. The severe asymmetry in the face of facial disordered patient can also negatively impact the symmetrization process based on the facial landmark extractions.

The rest of the paper is organized as follows. The facial symmetrization problem is formally formulated in Section 3. The proposed model is introduced in Section 4 in terms of the network composition and the joint-loss functions for parameter optimization. Section 5 discusses the model implementation in detail, and the experiments are demonstrated in Section 6. The paper is finally concluded in Section 7.

3. Face symmetrization problem formulation

Consider N facial data $X = \{x^{(i)}\}_{i=1}^N$, where $x^{(i)} \in \mathbb{R}^{(N_h \times N_w \times 3)}$ is a three-channel RGB image with its height and width being N_h and N_w respectively. The face symmetrization is trying to find the

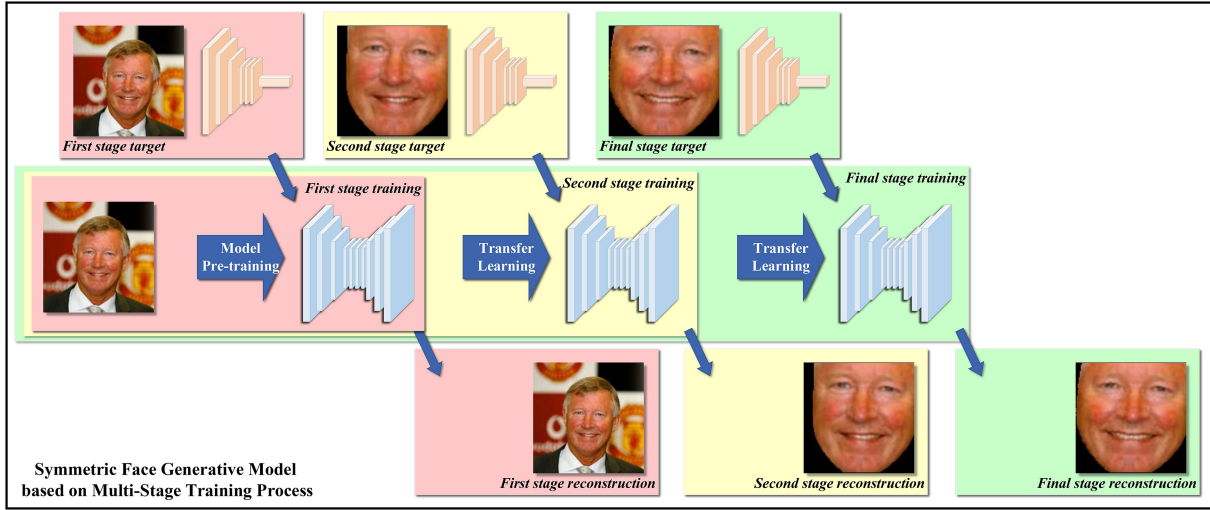


Fig. 2. The hierarchically structured generative network training process for facial symmetrization.

corresponding N data points $Y = \{y^{(i)}\}_{i=1}^N$, where $y^{(i)} \in \mathbb{R}^{(N_h \times N_w \times 3)}$ is also a three-channel RGB image that contains a symmetrical face having the same identity as in $x^{(i)}$. The solution to this face symmetrization is to learn a mapping M to translate each $x^{(i)}$ into an image $y^{(i)}$ that has a face with a same identity but symmetrized. The mapping $M: (X) \rightarrow Y$ can be described using a deep convolutional generative model, which takes an image as input and generates a target image.

The modelling of this mapping M is a supervised learning process, which means the training data should contain both input X and target Y . However, in practice, compared to normal facial images, the symmetrical faces can be less common and difficult to obtain for training data preparation. Due to the insufficient training data, the end-to-end learning process can lead to overfitting and potential drops of the generalization for the mapping M . To this end, we propose to train this mapping M in a coarse-to-fine manner. As shown in Fig. 2, a hierarchically structured supervised training process with three stages is proposed in this paper to tackle the problem of insufficient training data.

4. Methodology

4.1. Symmetrical face generative model composition

As shown in Fig. 1, the proposed facial symmetrization model is inspired by the deep generative adversarial network, which contains a generator network and a discriminator network. The generator $G(x, w_g)$ aims at producing a $y_{generated}^{(i)}$ given the input data $x_{input}^{(i)}$, where $y_{generated}^{(i)}$ approximates the target $y_{target}^{(i)}$ as possible as it can. The discriminator $D(x, w_d)$ provides the distance penalties on how far the generated data point $y_{generated}^{(i)}$ is deviated from the desired image $y_{target}^{(i)}$, for example, being a realistic-looking face. In this paper, the proposed facial symmetrization model utilizes a variational autoencoder (VAE) as the generator and a convolutional neural network (CNN) as the discriminator. Additionally, the proposed model also includes a facial recognition network as facial identifier $R(x, w_r)$ to produce the cost for $y_{generated}^{(i)}$ and $y_{target}^{(i)}$ being different facial identities.

VAE based generator. The goal of the generator $G: X \rightarrow \hat{X}$ is to map a high-resolution input pixel grid (facial images $X = \{x_{input}^{(i)}\}$) to a high-resolution output pixel grid (symmetrical facial images $\hat{Y} = \{y_{generated}^{(i)}\}$). Specifically, the problem targeted in this paper considers the output data that has a latent relationship with the input data, such as the facial identity. Moreover, all the information in the output data (symmetrized faces $Y = \{y_{target}^{(i)}\}$) is actually encoded in the input data (original faces X). Thus, an encoder-decoder network is used in the generator. More precisely, we use variational autoencoder to generate the desired output data assuming that the symmetrized face $y_{target}^{(i)}$ can be a sample of the input data $x_{input}^{(i)}$.

The VAE network commonly consists of two parts: an encoder network and a decoder network. The encoder compresses the input data (a facial image) $x^{(i)} \in \mathbb{R}^{(N_h \times N_w \times 3)}$ into a vector z of latent variables following a probabilistic distribution $P(z|X)$, where X is the datapoints in the dataset. The vector z is a compressed representation of the input data. It contains the necessary information for a image reconstruction. The decoder translates a certain sampled feature from this latent vector z back into the desired image, where the sampling process follows a certain distribution $P(z)$. Different sampling distribution $P(z)$ can lead to different image reconstruction. Thus, a supervised training process is required to properly model this generator.

CNN based discriminator. It is well known that the traditional autoencoder network produces blurry results due to ℓ_2 - norm or ℓ_1 - norm loss directly computed between the generated data \hat{X} and the data Y from the training set. Those losses can encouraging the accuracies in low frequencies of the images, however, ignore the high frequency information. This will potentially lead to a lower resolution output compared to the input. This motivates utilizing a discriminator D in the proposed model to produce additional penalties for the generation errors. It acts as a classifier to discriminate between $\{y_{generated}^{(i)}\}$ and $\{y_{target}^{(i)}\}$, in other words, between the faked and the real facial images. Inspired by [25], we construct a discriminator $D: Y \rightarrow I$ to enforce the realism in the generated data in a block-wised way. I is a matrix containing the block-wised discrimination results of output Y . Each block has a value between 0.0 and 1.0.

CNN based facial identifier. The proposed model includes a facial identifier $R(x, w_r)$ to enforce identity similarities between $y_{generated}^{(i)}$ and $y_{target}^{(i)}$. The facial identifier is actually a facial recognition network pre-trained independently with respect to the facial symmetrization model. It can produce compressed features for the facial identity representations, which will be contributed to the distance measurement between $y_{generated}^{(i)}$ and $y_{target}^{(i)}$. The facial identifier model parameters w_r are fixed during the training of the proposed facial symmetrization model.

4.2. Loss functions

The traditional generative adversarial network takes a norm-distributed random feature z_0 as input and generates a realistic-looking output image $y_{generated}^{(i)}$ in a certain domain. Its adversarial training process only try to make the generated image look real, which may lead to an arbitrary surface appearance in the output. Differing from the previous approaches, the facial symmetrization model is trained in a supervised manner, which tries to generate a deterministic output given a certain input. Thus, the loss functions in the proposed model are twofold: an adversarial loss and an identity loss. The adversarial loss tries to constrain the output to be as real as possible as a symmetrical facial image. The identity loss aims at making the generated symmetrical face to be as the same identity of the input face as possible.

Adversarial loss. The proposed facial symmetrization model contains a generator G and a discriminator D . The goal of the training process is to make G be able to produce a realistic-looking symmetrical facial image, while enabling D to accurately distinguish between the real images and the faked ones. Therefore, the loss function in this part has two components:

$$\mathcal{L}_{Adv_g}(w_g) = E_{x_{input} \sim P_{data}(X)} [||1 - D(G(x_{input}, w_g), w_d)||_l] \quad (1)$$

$$\begin{aligned} \mathcal{L}_{Adv_d}(w_d) = & E_{y_{target} \sim P_{data}(Y)} [||1 - D(y_{target}, w_d)||_l] \\ & + \lambda_1 E_{x_{input} \sim P_{data}(X)} [||D(G(x_{input}, w_g), w_d)||_l] \end{aligned} \quad (2)$$

where $\mathcal{L}_{Adv_g}(w_g)$ indicates a distance of how far the generated image $y_{generated}^{(i)} = G(x_{input}^{(i)}, w_g)$ is from a real one. $\mathcal{L}_{Adv_d}(w_d)$ implies that from the discriminator's perspective, the target image $y_{target}^{(i)}$ should be classified as a real symmetrical face image and the generated image $G(x_{input}^{(i)}, w_g)$ should be identified as a faked one. This will ultimately match the distribution of the generated data to the distribution of the data in the target domain.

Identity loss. The adversarial loss can help to learn a mapping G that produces the outputs following a data distribution from the target domain. However, this generated symmetrical face may still belong to an arbitrary person rather than the one of the same identity in the input facial image. To tackle this, we include an identity loss in the facial symmetrization model, which will guide the gradient descent during the learning process towards a solution to reconstruct the symmetrical face for the same person.

$$\begin{aligned} \mathcal{L}_{id_d}(w_g) = & E_{x_{input} \sim P_{data}(X), y_{target} \sim P_{data}(Y)} [||y_{target} - G(x_{input}, w_g)||_l] \\ & + \lambda_2 E_{x_{input} \sim P_{data}(X), y_{target} \sim P_{data}(Y)} [||R(y_{target}, w_r) - R(G(x_{input}, w_g), w_r)||_l] \end{aligned} \quad (3)$$

where $R(x, w_r)$ is a CNN based facial identifier, which produces a facial identity representation given an input facial data x . The parameter w_r is learnt independently with respect to the facial symmetrization model.

5. Model implementation details

Network architecture. The generator network G is implemented according to a VAE architecture, which can provide an image-to-image mapping. It utilizes multiple Convolution-BatchNorm-LRelu sequences as modules to compose the encoder network. There are fully-connected layers before z is reached. Inspired by the work from Isola et al. [25], we adopt skips in the network connections to construct the generator G , which give G the option of bypassing the encoding or decoding block if it doesn't have a use for it. For discriminator network D , we utilize similar concept from PatchGAN [26,27] to build the network structure, which provides block-by-block discriminations. It has fewer parameters compared to a full-image D , and can work with arbitrarily sized image data.

Parameter optimization. The generator and the discriminator are trained alternately. The loss function for G is a weighted sum of $\mathcal{L}_{Adv_g}(w_g)$ and $\mathcal{L}_{Id_g}(w_g)$ with a hyperparameter λ_3 as a weight to adjust the proportion between them. The loss function for D is composed using $\mathcal{L}_{Adv_d}(w_d)$. The discriminator's parameters w_d are fixed when training the generator G , and vice versa. There are another two hyperparameters λ_1 and λ_2 as the weights adjusting the proportions of the components in $\mathcal{L}_{Adv_d}(w_d)$ and $\mathcal{L}_{Id_g}(w_g)$ respectively.

It is a fact that the symmetrical facial images are less common among all available facial image databases. Therefore, the end-to-end model training will fail due to the insufficient training data. To tackle this problem, the proposed facial symmetrization model learns its parameters in a hierarchical manner. We propose a multi-stage training process to achieve a coarse-to-fine parameter optimization. The first stage tries to reconstruct the original facial image, which can be trained using the face training data that are sufficient enough in publicly available databases. The second stage tries to re-train the model to extract facial regions in the facial images. The third stage tries to transfer the model to achieve the symmetric face generating. The training data in the second and third stages can be substantially less than what are needed in the first stage.

6. Experiments and evaluations

6.1. Dataset for model training and testin

For the demonstrative purpose, the proposed facial symmetrization model is trained using the facial database of *Labelled Faces in the Wild* (LFW) [28], which contains more than 13 K wild facial images collected from the web. Those data include the facial images with a combination of the different head poses, different ambient illuminations, and different face sizes. Some of them are even presented with different facial occlusions, for example, by shades or by hands. The facial images are taken from the real photographs unconstrainedly, which covers different skin colors, genders and ages. The only constraint of those facial images is that they are detected by the Viola-Jones face detector [29]. This database provides facial labels, yet only the facial images are utilized to train the proposed model. We use 11 K images as the training data for the first training stage, and 2 K images for the evaluation of the generating performance of the proposed model. We also manually crop facial regions for 5 K of the facial images as the training data for the second training stage. The facial middle lines on those face-only images are manually marked as well to generate 5 K symmetric faces for the third training stage.

To further evaluate the performance of the proposed model, we also conduct experiments on 2.5 K 2D facial images from the database of *Binghamton University 3D Facial Expression* (BU-3DFE) [30], which includes seven facial expressions for each individual such as anger, disgust, fear, happiness, sadness, surprise and neutral face. Those facial expressions provide facial asymmetries to various extents, which can pose challenges for traditional facial symmetrization methods.

Especially, we evaluate the proposed model using 1.2 K 2D facial palsy images that we have collected from hospital. It can demonstrate the performance of our model when dealing with images of facial paralysis patients. Those facial paralysis images present facial asymmetries to a high degree, which can make the traditional facial auto-symmetrization methods fail easily. The collected facial palsy database contains 50 subjects with five facial performances for each. Those facial performances include raising eyebrows, closing eye, screwing up nose, drumming cheek and opening mouth. We record the facial performances into videos and extract 4 to 5 frames from each of them as 2D facial palsy images. For the testing purpose of the proposed facial symmetrization model, we resize the images into the size of 256 by 256. The equipment used for data collection is demonstrated in Fig. 3.

6.2. Facial symmetrization

The proposed model is evaluated using facial palsy images that can pose much more challenges for facial landmark extraction algorithms. The experiments are also based on the face database of both LFW

Table 1
The performance comparisons in terms of facial recognition errors

	Collected facial palsy images	LFW dataset	BU3D dataset
CLNF [24]	0.335	0.262	0.284
SDM [19]	0.401	0.306	0.216
CLM [23]	0.492	0.342	0.304
Dlib [31]	0.312	0.288	0.274
End-to-end training	0.749	0.701	0.685
Ours	0.186	0.116	0.105



Fig. 3. The equipment used for facial paralysis image acquisition.

and BU3D, which respectively provide the uncontrolled and controlled facial images. Figure 4 demonstrates the visual examples of the experiment results, and the performance comparisons with several existing approaches are illustrated in Table 1. The severe facial disorders in the facial palsy images can obviously lead to a facial landmark tracking failure, which make the traditional facial symmetrization method out of order easily. The uncontrolled facial images in LFW database, which represents the practical scenarios, can also provide obstacles for traditional facial auto-symmetrization methods due to different head poses, occlusions and illuminations.

On the other hand, it can be observed from Fig. 4 that despite the high degree of the asymmetries existing in the input images, the proposed model can robustly produce perfect symmetrical faces. And the reconstruction can sufficiently hold the same identity of the person in the original input image. The proposed method analyses the whole facial images, which makes the proposed model more robust to aforementioned conditions.

Beside those visual demonstrations, we also compare our method with several other traditional approaches that based on the different facial landmark detection algorithms, including some of the most robust and popular ones such as CLNF [24], SDM [19], CLM [23] and Dlib [31]. The generated symmetrical faces should hold the same facial identities of the persons in the original images, based on which we employ facial recognition algorithm to identify whether the generated symmetric face belongs to the same person in the original image. This yields a recognition errors that can reflect the facial symmetrization performance.

The *OpenFace* [32] is utilized to achieve the facial recognition in this evaluation experiment due to its excellent recognition performance on LFW database [33] and its independence with respect to our model. The performance comparisons are demonstrated in Table 1. As it can be seen, the results using the traditional facial landmark based methods can be negatively impacted when high degree facial asymmetry exists in the input images, such as facial palsy images. The uncontrolled conditions can also negatively affect the facial symmetrization using the traditional methods that based on the facial landmark localization algorithms, even those landmark localization algorithms have a good performance on facial landmark detection task isolatedly.

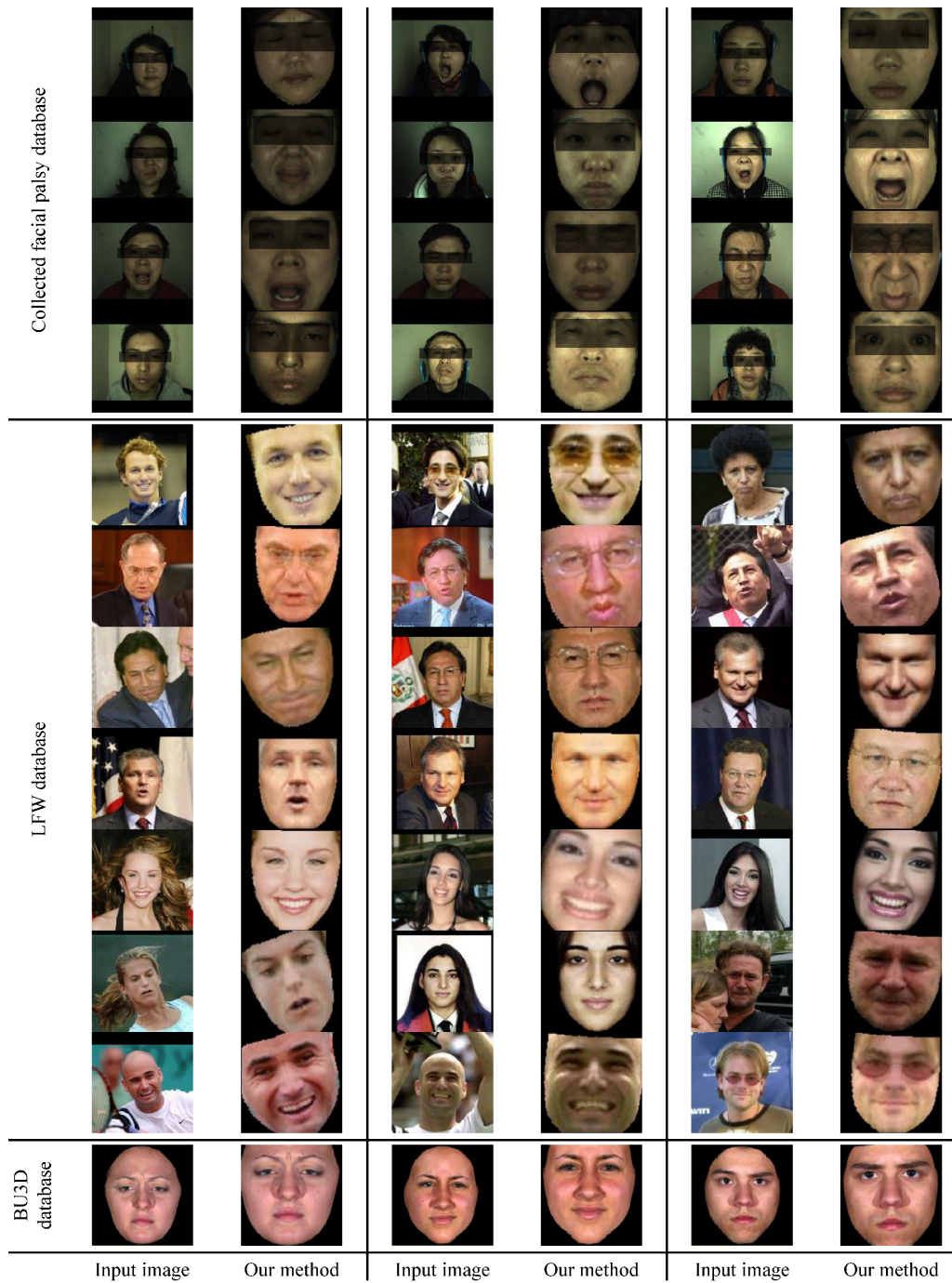


Fig. 4. Samples of the experiment results for facial symmetrization on collected facial palsy images, LFW database and BU3D database. It is noticeable that some test examples with large facial performances, head poses or facial occlusions like shades, can be well handled by the proposed model, whereas those conditions can pose more challenges for traditional facial landmark localization based approaches.

7. Conclusion

This paper presents a facial symmetrization model based on a joint-loss enhanced deep generative network architecture, which consists of a VAE as the generator network, a CNN as the discriminator network, and a CNN as the facial identifier network. The joint loss is composed of a pair of adversarial losses and an identity loss. The adversarial loss pair tries to provide penalties on the distance of how far the generated data is from a symmetrical facial image. The identity loss tries to constrain the output to remain the same identity of the person in the input image as much as possible. Additionally, a hierarchically structured training process is proposed to learn the model parameters on a limited number of symmetrical facial data as training samples. A multi-stage optimization strategy is presented to train the model in a coarse-to-fine manner. In the first stage, the model is trained to reproduce the whole facial images. The training data for this stage is sufficient enough since there are plenty of facial image database publicly available. In the second stage, the model is re-trained to extract the face-only images from the input data. In the final stage, the model is transferred to generate symmetric faces based on the manually crafted training data, which can be substantially less than what required in the first training stage. The performance of the proposed model is demonstrated in the experiments. Along with the visual demonstration, the proposed method also achieves competitive performance compared with several other approaches based on some of the most popular facial landmark detection algorithms in this field.

Acknowledgments

This work was supported by the Project of Shandong Province Higher Educational Science and Technology Program (J17KA076) and the Scientific Research Foundation of Shandong University of Science and Technology for Recruited Talents (2017RCJJ047).

Conflict of interest

None to report.

References

- [1] Wang T, Zhang S, Dong J, et al. Automatic evaluation of the degree of facial nerve paralysis. *Multimedia Tools and Applications*. 2016; 75(19): 11893-11908.
- [2] Wang T, Dong J, Sun X, et al. Automatic recognition of facial movement for paralyzed face. *Bio-Medical Materials And Engineering*. 2014; 24(6): 2751-2760.
- [3] Sajid M, Shafique T, Baig MJA, et al. Automatic Grading of Palsy Using Asymmetrical Facial Features: A Study Complemented by New Solutions. *Symmetry*. 2018; 10(7): 242.
- [4] Fujiwara K, Furuta Y, Yamamoto N, et al. Factors affecting the effect of physical rehabilitation therapy for synkinesis as a sequela to facial nerve palsy. *Auris Nasus Larynx*. 2018; 45(4): 732-739.
- [5] Saegusa C, Intoy J, Shimojo S. Visual attractiveness is leaky: the asymmetrical relationship between face and hair; 2015.
- [6] Zhang G, Wang Y. Asymmetry-based quality assessment of face images. In: *International Symposium on Visual Computing*. Springer; 2009. p. 499-508.
- [7] Wang Y, Yu H, Dong J, et al. Cascade Support Vector Regression-based Facial Expression-Aware Face Frontalization; 2017.
- [8] Yin X, Yu X, Sohn K, et al. Towards Large-Pose Face Frontalization in the Wild. In: *The IEEE International Conference on Computer Vision (ICCV)*; Oct 2017.

- [9] Sagonas C, Panagakis Y, Zafeiriou S, et al. Robust statistical face frontalization. In: Proceedings of the IEEE International Conference on Computer Vision; 2015. pp. 3871-3879.
- [10] Sikander G, Anwar S, Djawad YA. Facial Feature Detection: A Facial Symmetry Approach. In: 2017 5th International Symposium on Computational and Business Intelligence (ISCBI); 2017. pp. 26-31.
- [11] Xu Y, Zhang Z, Lu G, et al. Approximately symmetrical face images for image preprocessing in face recognition and sparse representation based classification. *Pattern Recognition*. 2016; 54: 68-82.
- [12] Yu H, Liu H. Regression-based facial expression optimization. *IEEE Trans Hum Mach Syst*. 2014; 44(3): 386-394.
- [13] Salam H, Séguier R. A survey on face modeling: building a bridge between face analysis and synthesis. *The Visual Computer*. 2018; 34(2): 289-319.
- [14] Zhang S, Dong J, Yu H. Feature Matching for Underwater Image via Superpixel Tracking. In: Automation and Computing (ICAC), 2017 23rd International Conference on; Huddersfield, UK; 2017. p. 1-5.
- [15] Chang BL, Wilson AJ, Samra F, et al. Quantifying Pediatric Facial Palsy: Using Digital Image Correlation to Objectively Characterize Pediatric Facial Symmetry. *The Cleft Palate-Craniofacial Journal*. 2017 Dec; 55(1): 119-126.
- [16] Song A, Xu G, Ding X, et al. Assessment for facial nerve paralysis based on facial asymmetry. *Australasian Physical & Engineering Sciences in Medicine*. 2017; 40(4): 851-860.
- [17] Katsumi S, Esaki S, Hattori K, et al. Quantitative analysis of facial palsy using a three-dimensional facial motion measurement system. *Auris Nasus Larynx*. 2015; 42(4): 275-283.
- [18] Hassner T, Harel S, Paz E, et al. Effective face frontalization in unconstrained images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015. pp. 4295-4304.
- [19] Xiong X, De la Torre F. Supervised descent method and its applications to face alignment. In: Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. IEEE; 2013. pp. 532-539.
- [20] Harguess J, Aggarwal JK. Is there a connection between face symmetry and face recognition? In: CVPR 2011 WORKSHOPS; 2011. pp. 66-73.
- [21] Harguess J, Aggarwal JK. A case for the average-half-face in 2D and 3D for face recognition. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops; 2009. pp. 7-12.
- [22] Matthews I, Baker S. Active appearance models revisited. *International Journal of Computer Vision*. 2004; 60(2): 135-164.
- [23] Cristinacce D, Cootes T. Automatic feature localisation with constrained local models. *Pattern Recognition*. 2008; 41(10): 3054-3067.
- [24] Baltrusaitis T, Robinson P, Morency LP. Constrained Local Neural Fields for Robust Facial Landmark Detection in the Wild. In: The IEEE International Conference on Computer Vision (ICCV) Workshops; Jun 2013.
- [25] Isola P, Zhu JY, Zhou T, et al. Image-to-Image Translation with Conditional Adversarial Networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017. pp. 5967-5976.
- [26] Zhu JY, Park T, Isola P, et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In: Computer Vision (ICCV), 2017 IEEE International Conference on; 2017.
- [27] Li C, Wand M. Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks. In: Leibe B, Matas J, Sebe N, et al., editors. European conference on computer vision; Cham. Springer International Publishing; 2016. pp. 702-716.
- [28] Learned-Miller E, Huang GB, RoyChowdhury A, et al. Labeled Faces in the Wild: A Survey. In: Kawulok M, Celebi ME, Smolka B, editors. Advances in face detection and facial image analysis. Cham: Springer International Publishing; 2016. pp. 189-248.
- [29] Viola P, Jones MJ. Robust Real-Time Face Detection. *International Journal of Computer Vision*. 2004; 57(2): 137-154.
- [30] Yin L, Wei X, Sun Y, et al. A 3D facial expression database for facial behavior research. In: 7th International Conference on Automatic Face and Gesture Recognition (FGR06); Apr 2006. p. 211-216.
- [31] King DE. Dlib-ml: A Machine Learning Toolkit. *Journal of Machine Learning Research*. 2009; 10: 1755-1758.
- [32] Amos B, Ludwiczuk B, Satyanarayanan M. OpenFace: A general-purpose face recognition library with mobile applications. CMU-CS-16-118, CMU School of Computer Science; 2016.
- [33] Schroff F, Kalenichenko D, Philbin J. FaceNet: A Unified Embedding for Face Recognition and Clustering. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 2015.