Boosting Cross-Age Face Verification via Generative Age Normalization

Grigory Antipov[⋆†] Moez Baccouche[⋆] Jean-Luc Dugelay[†]
[⋆] Orange Labs, 4 rue Clos Courtel, 35512 Cesson-Sévigné, France
[†] Eurecom, 450 route des Chappes, 06410 Biot, France
{grigory.antipov,moez.baccouche}@orange.com jean-luc.dugelay@eurecom.fr

Abstract

Despite the tremendous progress in face verification performance as a result of Deep Learning, the sensitivity to human age variations remains an Achilles' heel of the majority of the contemporary face verification software. A promising solution to this problem consists in synthetic aging/rejuvenation of the input face images to some predefined age categories prior to face verification. We recently proposed [3] Age-cGAN aging/rejuvenation method based on generative adversarial neural networks allowing to synthesize more plausible and realistic faces than alternative non-generative methods. However, in this work, we show that Age-cGAN cannot be directly used for improving face verification due to its slightly imperfect preservation of the original identities in aged/rejuvenated faces. We therefore propose Local Manifold Adaptation (LMA) approach which resolves the stated issue of Age-cGAN resulting in the novel Age-cGAN+LMA aging/rejuvenation method. Based on Age-cGAN+LMA, we design an age normalization algorithm which boosts the accuracy of an off-the-shelf face verification software in the cross-age evaluation scenario.

1. Introduction

Being the cornerstone of face recognition and face retrieval, face verification is a fundamental biometrics problem. Its objective is to compare a pair of human photos and to tell whether they belong to the same person or not. In recent years, face verification has been revolutionized with the surge of deep learning (in particular, Convolutional Neural Networks (CNN) [18]) setting the state-of-the-art scores [30, 27] which surpass human performances on popular face verification benchmarks [17].

However, face verification remains a challenging and unsolved problem in the cross-age evaluation scenario. Indeed, a number of studies [20, 19, 6, 12, 33] have shown that the performances of off-the-shelf face verification software significantly degrades in the presence of age variation. This is a serious issue for biometrics applications such as passport renewal, border control or research of wanted criminals, where automatic face verification systems are often required to deal with big age gaps between the tested and the reference photos [22, 8].

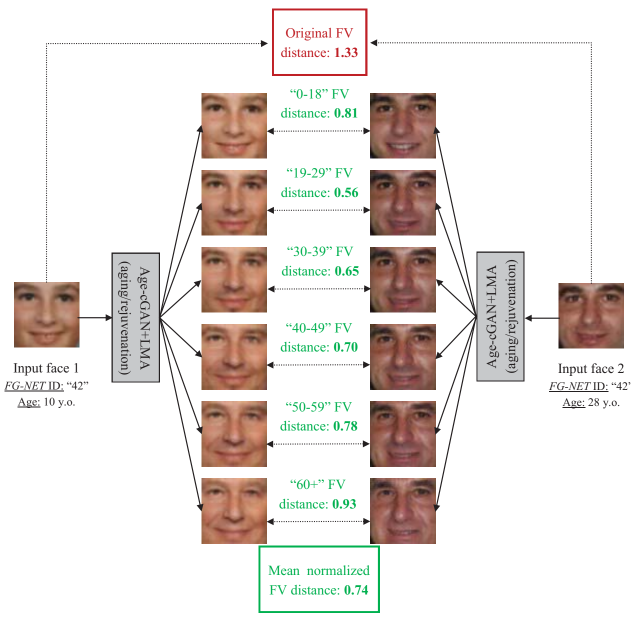


Figure 1. (Better viewed in color). Our fully-synthetic age normalization algorithm uses the proposed Age-cGAN+LMA aging/rejuvenation method to improve cross-age Face Verification (FV). A pair of faces from the *FG-NET* dataset [1] belonging to the same person at different ages are compared with the OpenFace FV software [2]. Without age normalization, the software incorrectly classifies the faces as a negative pair: the estimated FV distance of 1.33 (in red rectangle) is well above the software rejection threshold of 0.99. After aging/rejuvenation both faces to 6 age categories ("0-18", "19-29", "30-39", "40-49", "50-59" and "60+"), the mean estimated FV distance between 6 synthetic pairs by the same software is of 0.74 (in green rectangle). This allows to correctly classify the initial pair as positive.

Therefore, the problem of cross-age face verification has attracted a large body of research studies, which can be roughly split into two categories: (1) the ones improving face verification software in order to make it invariant to age variations, and (2) the ones which synthetically modify the input face images to normalize ages between them prior to face verification. The approaches belonging to the first category [14, 10, 34, 7, 33] aim to design face descriptors which can effectively separate identity and age information in face images. At the same time, the vast majority of existing face verification software remains highly sensitive to age gaps between the tested faces making the approaches of the second category particularly important.

The core of age normalization is the underlying method for synthetic face aging/rejuvenation (also known as age progression/regression [28]). Its goal is to aesthetically age or rejuvenate an input face preserving the original identity. Thus, in our previous work [3], we proposed an aging/rejuvenation method called Age-cGAN which is based on Generative Adversarial Networks (GANs) [11] and produces plausible, naturally looking faces. However, as shown in this work, the level of the identity preservation of Age-cGAN is insufficient for cross-age face verification.

The present study builds upon Age-cGAN proposing a novel approach to further enhance its identity preservation. By doing so, we allow application of Age-cGAN for age normalization prior to face verification (cf. Figure 1). More precisely, our contributions are two-fold:

1. We extend our previous Age-cGAN aging/rejuvenation method [3] with a novel Local Manifold Adaptation (LMA) approach which allows to quasi-perfectly preserve the original person's identity.

2. Based on Age-cGAN+LMA, we design a fully-synthetic age normalization algorithm. As a result, we improve the accuracy of a popular deep learning face verification software in cross-age evaluation scenario.

The rest of the paper is organized as following: in Section 2, we discuss the most notable existing works on synthetic aging/rejuvenation; in Section 3, we briefly present GANs in general and Age-cGAN, in particular; in Section 4 we detail our LMA approach, the resulting Age-cGAN+LMA aging/rejuvenation method and our fully-synthetic algorithm of age normalization; in Section 5, we find optimal hyperparameters for LMA, qualitatively compare our method with alternatives and quantitatively show the improvements of the proposed age normalization on cross-age face verification; and finally, Section 6 concludes the study and gives directions for the future work.

2. Related Work

Existing works on synthetic face aging/rejuvenation mainly fall into one of the following 2 groups: modelling-based and prototype-based.

Modelling-based methods employ parametric models to simulate the physical aging mechanism of muscles, skin and skull of an individual. Different modelling-based methods were proposed including Active Appearance Models (AAMs) [15], craniofacial growth model [25], and-or graph model [29] and statistical model [23]. The common limitation of modelling-based methods is their requirement for long-term aging sequences of the same person in order to model the subtle face aging mechanism. Obviously, it is very costly to collect big enough training datasets with such sequences. Modelling-based methods are also known to be computationally expensive.

Prototype-based methods [5, 31, 13] define average faces calculated on training images of certain age categories as prototypes for these categories. Differences between the prototype faces constitute the aging patterns which are further used to transform an input face image into the target age category. The prototype approaches are often fast, but due to the fact that they discard personalized information, these methods are prone to lose the identity while aging. Therefore, contemporary prototype-based methods use explicit mechanisms to preserve the original identity.

Thus, the two most recent studies [28, 32] which applied face aging/rejuvenation to improve cross-age face verification are prototype-based. Shu et al. [28] proposed Coupled Dictionary Learning (CDL). The authors learn a dictionary per each age category, and the aging pattern is encoded by the dictionary bases. Pairs of neighbouring dictionaries are learned jointly, and the identity information is regarded as the reconstruction error which is added into the aged face directly. Despite convincing aging results, CDL suffers from the ghosting artefacts. The other Recurrent Face Aging (RFA) aging/rejuvenation method proposed by Wang et al. [32] uses Recurrent Neural Networks (RNNs) for age pattern transition between the coefficients of eigenfaces calculated in different age categories. Contrary to CDL, RFA smooths the ghosting artefacts but poorly preserves the original person's identity.

Despite aging and rejuvenation being equally important, once trained, the vast majority of previous models can change ages only in one direction (i.e. either to age or to rejuvenate, but not both). This is a serious limitation which has been addressed by the most recent generative neural models, namely: Adversarial AutoEncoder (AAE) [36] and our Age-cGAN [3]. Both AAE and Age-cGAN learn a synthetic manifold which models human faces of all ages. When an input face is projected on the designed manifold, its age can be effortlessly changed to any required category. The two generative aging/rejuvenation methods have the opposite advantages and downsides: the synthetic faces produced by Age-cGAN have fewer visible artefacts than the ones produced by AAE. At the same time, AAE better preserves the original person identity than Age-cGAN.

3. GANs for Face Aging/Rejuvenation

3.1. Generative Adversarial Networks

General Idea Introduced in [11], Generative Adversarial Network (GAN) is a pair of neural networks: the generator G and the discriminator D. G generates synthetic images x from random latent vectors z which are sampled from a standard normal distribution $p_z \sim N(0, I)$. In other words, G defines a mapping from the latent space N^z to the image space N^x. In this work, N^x is the space of human face images. The goal of the generator is to imitate the distribution p_{data} of natural faces. At the same time, the discriminator tries to distinguish natural face images coming from the image distribution p_{data} and the synthetic images produced by the generator. Given an input face x (either synthetic or natural), D outputs the probability that x is a natural human face rather than a synthetic one. Thus, the two networks have opposite objectives, and they are iteratively optimized against each other in a minimax game (hence the name "adversarial"). More formally, GAN training can be expressed as an optimization of the function $v(\theta_G, \theta_D)$, where θ_G and θ_D are parameters of G and D, respectively:

$$\begin{aligned}\min_{\theta_G} \max_{\theta_D} v(\theta_G, \theta_D) = \mathbf{E}_{x \sim p_{data}}[\log D(x)] \\ + \mathbf{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))]\end{aligned} \tag{1}$$

Conditional GAN Conditional GAN (cGAN) [21, 9] extends the GAN model allowing the generation of images with certain attributes ("conditions"). In practice, conditions $y \in N^y$ can be any information related to the target face image: level of illumination, facial pose or facial attribute. The optimization objective for a cGAN is very similar to the one for a standard GAN presented in Equation 1, but in the case of cGAN, additional conditional information is provided both to G and D:

$$\begin{aligned}\min_{\theta_G} \max_{\theta_D} v(\theta_G, \theta_D) = \mathbf{E}_{x,y \sim p_{data}}[\log D(x, y)] \\ + \mathbf{E}_{z \sim p_z(z), \widetilde{y} \sim p_y}[\log(1 - D(G(z, \widetilde{y}), \widetilde{y}))]\end{aligned} \tag{2}$$

Synthetic Face Manifold Once cGAN training is finished, G can generate plausible synthetic faces following a distribution similar to the one of natural faces. Varying the latent vectors z and the conditions y at the input of G results in different synthetic faces $\bar{x}$ at its output. Importantly, the mapping learned by the generator is continuous, meaning that small variations in latent vectors z and conditions y result in small variations in the generated faces $\bar{x}$. Thus, an ensemble of all possible synthetic faces produced by G with various $z \in N^z$ and $y \in N^y$ taken together form a *synthetic manifold* $\bar{N}^x$ (the term was coined in [37]).

Face Editing with cGAN In order to edit an input face image x with a cGAN, it is firstly necessary to reconstruct (i.e. to approximate) it with a synthetic image $\bar{x}$ belonging to the manifold $\bar{N}^x$. In other words, one need to find a latent vector z and a condition y which, when given at the input of G, would produce a plausible reconstruction $\bar{x} = G(z, y)$ of the input x. Here and below, we refer to such reconstruction as *projection of an input face x onto the synthetic manifold $\bar{N}^x$*. Contrary to autoencoders, GANs do not have an explicit mechanism for manifold projection. However, this problem is often [37, 24] circumvented by training a separate neural network, called encoder E, which approximates an inverse mapping for the generator: $E \approx G^{-1}$. Having trained the encoder, the reconstruction can be simply done as $\bar{x} = G(E(x))$.

Once a sufficiently good reconstruction $\bar{x}$ of an input face x is obtained, the face editing is trivial with cGANs. Indeed, the generator G of a cGAN disentangles the information encoded by latent vectors z and by conditions y making the two independent [24, 16]. It means that if we keep z fixed and change the condition y, then only the conditional information encoded by y will change in the generated synthetic face.

For example, if the conditions y encode age information (as in this work), the latent vectors would encode everything except for the age (i.e. person's identity, facial pose, presence of glasses etc.) Therefore, if we manage to approximate an input natural face x^0 at age y^0 with a synthetic face $\bar{x^0} = G(z^*, y^0)$ (where z^* is an optimal latent vector), then the aged/rejuvenated version of the initial face x^1 at age y^1 can be simply obtained by swapping the age condition at the input of the generator $\bar{x^1} = G(z^*, y^1)$.

3.2. Age-cGAN Aging/Rejuvenation Method

In our previous work [3], we proposed an Age-conditioned cGAN (Age-cGAN) which is able to generate realistic synthetic faces at 6 age categories: "0-18", "19-29", "30-39", "40-49", "50-59" and "60+". Moreover, in order to use Age-cGAN for aging/rejuvenation, we also designed [3] an algorithm to find an optimal latent vector z^* for manifold projection preserving the original identity of the person in an input image x of age y (the age of the person is assumed given). z^* is selected as the solution of the optimization problem which minimizes the Euclidean distance between the face recognition embeddings of the original x and the reconstructed faces $\bar{x}$:

$$\begin{aligned}z^* &= \underset{z}{\operatorname{argmin}} \, ||FR(x) - FR(\bar{x})||_{L_2} \\ &= \underset{z}{\operatorname{argmin}} \, ||FR(x) - FR(G(z, y))||_{L_2}\end{aligned} \tag{3}$$

where $FR(\cdot)$ are face recognition embeddings of face images calculated with a separately trained CNN FR.

In this paper, we propose an extension which allows to improve the presented algorithm.

In particular, as reported in [3], Age-cGAN obtains a decent level of about 80% of identity preservation when measured with an open-source face verification software OpenFace [2]. However, the fact that in almost 20% of cases, Age-cGAN loses the original person's identity even before aging/rejuvenation makes impossible its practical application for improving cross-age face verification. Indeed, the negative effect of the poor identity reconstruction is so big that it cannot be compensated by age normalization. This is experimentally illustrated in Figure 5 which is discussed in Subsection 5.4. Therefore, in order to apply Age-cGAN for age normalization in the cross-age face verification scenario, the original method from [3] must be improved to even better preserve the identities of input faces in the synthetic reconstructions.

Finally, the ultimate objective of this study is boosting cross-age face verification, so unlike the original paper [3], in this work, we have trained Age-cGAN uniquely on face crops discarding all unnecessary background information (cf. Figures 1, 2, 3, 4).

4. Proposed Approach

4.1. Local Manifold Adaptation

GANs are arguably the most powerful generative models today. Nevertheless, their variability and expressiveness are obviously limited. In other words, the generator G which is trained on (no matter how big but) *finite* number of faces cannot exactly reproduce the details of all real-life face images with their *infinite* possibilities of minor facial details, accessories etc. Therefore, even the optimal projection $\bar{x}^*$ of a natural input face x on the manifold $\bar{N}^x$ is still quite different from the original (cf. Figure 2-(a)) in terms of subtle facial details (i.e. unique form of the mouth, of the eyes, of the nose, skin texture, etc.). Taken together, these subtle facial details result in rather significant identity differences between the original and the reconstructed faces.

A natural way to remedy this problem is to modify the designed synthetic face manifold $\bar{N}^x$ given an input face x in order to bring closer x and its projection $\bar{x}^*$. In particular, we propose changing only the *local* area around the projection $\bar{x}^*$ in the manifold $\bar{N}^x$. Hence, we refer to our approach as Local Manifold Adaptation (LMA) (cf. Figure 2-(b)).

Since the synthetic manifold $\bar{N}^x$ is completely defined by the generator, LMA is performed by a slight automatic modification of the generator G with respect to an input image x. More precisely, the key idea of LMA is the following: instead of performing aging/rejuvenating with a general generator G (issued from the Age-cGAN training), we firstly customize the general generator G to better fit an input image x_0 of (a known) age y_0 obtaining a new generator G_{x_0}. After LMA, G_{x_0} can produce a quasi-perfect reconstruction $\widehat{x_0} = G_{x_0}(z^*, y_0)$ of the input face x_0. Our intuition suggests that if the LMA reconstruction $\widehat{x_0}$ is closer to the original face x_0 than the Age-cGAN reconstruction $\bar{x}_0^*$, then the aged/rejuvenated face $\widehat{x_1} = G_{x_0}(z^*, y_1)$ will also better preserve the original identity than the one obtained via the general generator G. This intuition is confirmed by the experiments in Subsection 5.4.

The same optimization objective as in Equation 3 (i.e. the distance between natural and reconstructed face recognition embeddings: $||FR(x) - FR(G(z^*, y))||_{L_2}$) is used to customize G for an input face x. However, in case of LMA, we "freeze" the previously found z^* and optimize G instead. Given the fact that inputs of the generator are fixed (z^* and y), we employ a classical backpropagation algorithm to optimize G. In order to make the changes of $\bar{N}^x$ local and to preserve the continuity of the synthetic manifold which is learned during the adversarial training, the number of backpropagation iterations N and the used learning rate μ should be limited. In Subsection 5.2, we experimentally find the optimal N and μ for LMA and demonstrate both quantitatively and qualitatively the positive effect of our approach on preserving the original identities in the input face reconstructions.

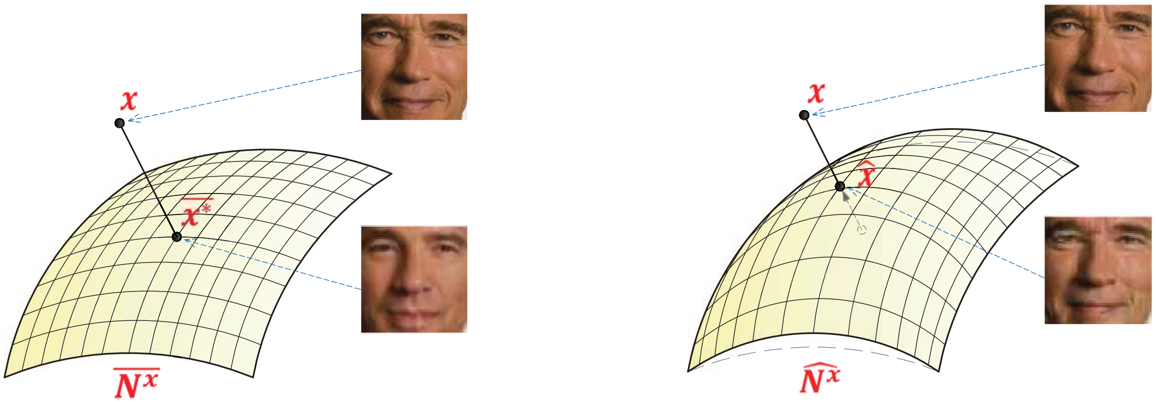
In order to avoid confusion, in this work, we refer to face aging/rejuvenation with Age-cGAN empowered by the proposed LMA approach as "Age-cGAN+LMA", while the basic algorithm from [3] is referred as "Age-cGAN".

4.2. Age Normalization

The objective of age normalization is to remove age related differences from pairs of face images and in so doing improve the performance of general-purpose face verification software. Usually, the software outputs the relative distance between a pair of input faces (the lower is the distance, the more similar are the identities in the faces according to the software). After age normalization this distance should be invariant to age differences in input faces.

Obviously, age normalization can be done differently. For example, in recent works [28, 32], the authors synthetically aged the younger face of a pair to the age category of the other older one. Aging was preferred over rejuvenation because the respective aging methods could perform the age change only in one direction (make faces older). Unlike these methods, our Age-cGAN+LMA is able to equally easy generate synthetic faces in 6 age categories while the optimization procedure is done *only once* per input image.

Thus, the first age normalization algorithm tested in this work is the combination of both aging and rejuvenation. In particular, given a pair of faces x_1 of age y_1 and x_2 of age y_2, we generate 2 pairs for face verification: $(\widehat{x_1}^{(y_2)}, x_2)$ and $(x_1, \widehat{x_2}^{(y_1)})$. Both pairs are evaluated by a face verification software which outputs 2 respective face distances,



(a): before LMA (b): after LMA

Figure 2. (Better viewed in color). Local Manifold Adaptation (LMA) approach to improve the identity preservation in the synthetically reconstructed face. (a) Input face x is reconstructed by projecting it on the synthetic manifold $\bar{N^x}$ (as proposed in [3]). (b) LMA locally modifies the synthetic manifold $\bar{N^x}$ transforming it to the new manifold $\widehat{N^x}$. As a result, the initial face x and its projection $\widehat{x}$ on the new manifold are brought closer than they were before LMA.

and the mean of these distances is taken as the final result of face verification. We refer to this algorithm of age normalization as *Half-Synthetic* (HS) because each tested pair of faces is composed of a natural face and a synthetic one.

However, another algorithm of age normalization has been found more effective with our Age-cGAN+LMA aging/rejuvenation method (cf. Subsection 5.4 for the quantitative comparison). The idea is the following: for each pair of face images x_1 and x_2, we generate 6 face verification pairs $\widehat{x_1}^{(i)}$ and $\widehat{x_2}^{(i)}$, $i \in \{1, 2, 3, 4, 5, 6\}$ belonging to each age category. The mean of the 6 corresponding face distances is taken as the final normalized result. We refer to this algorithm as *Fully-Synthetic* (FS) as all tested pairs are composed only of synthetic faces. FS age normalization is illustrated in the introductory Figure 1, and unless said otherwise, this algorithm is used for age normalization below.

5. Experiments

5.1. Datasets and Face Verification Solution

Datasets We have used *IMDB-Wiki* dataset [26] composed of photos of celebrities for training Age-cGAN. 2 other face datasets have been employed for evaluation in this work: *LFW* and *FG-NET*.

The Labelled Faces in the Wild (*LFW*) dataset [17] containing about 13K photos is the standard benchmark for face verification systems. In this work, we employ it as a validation dataset to find the optimal hyper-parameters for our LMA approach and to evaluate its impact on the identity preservation in the reconstructed face images. The standard *LFW* face verification pairs are composed of face images of (almost) same ages, so we do not use this dataset for cross-age face verification evaluation.

Finally, all cross-age face verification experiments have been performed on the *FG-NET* [1] dataset containing 975 face images of 82 persons with age annotations. *FG-NET* and *CACD* [6] are the two most used benchmarks for cross-age face verification. However, *CACD* is composed of celebrities and have a lot of intersections with *IMDB-Wiki* which is utilized for training of Age-cGAN. Moreover, OpenFace [2] face verification software was trained using the images of the same celebrities as in *CACD*. Therefore, in order to avoid biased results, we have not used the *CACD* dataset for cross-age face verification in this work.

Faces in all used datasets have been detected and aligned with a commercial solution which is based on [35, 4]. Contrary to many previous works on face aging/rejuvenation [28, 32], we do not employ any sort of face warping to normalize face expressions or poses. The designed generative Age-cGAN+LMA aging/rejuvenation method absorbs these variations implicitly.

OpenFace Face Verification Software We employ the recently developed deep learning-based solution OpenFace [2] for face verification experiments in this work. OpenFace has been chosen because (1) it is one of the most popular and easy-to-use open-source projects for face verification which ensures the reproducibility of our results, and (2) it has not been explicitly trained to be robust to age variations. The software has been used as a black box: a pair of natural or synthetically normalized images is given to its input, and OpenFace outputs a single value which is the relative distance between the provided pair of faces. The software requires a specific alignment of input faces, so before being processed by it, both natural and synthetic faces are aligned accordingly.

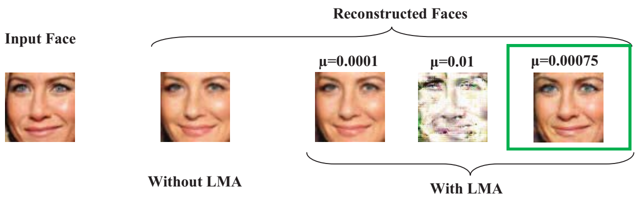


Figure 3. (Better viewed in color). Face reconstruction with and without Local Manifold Adaptation (LMA). For LMA-enhanced reconstructions, the impact of the learning rate μ is illustrated.

Tested Pairs	FV score on *LFW*
Original	**89.4%**
Reconstructed (Age-cGAN) [3]	82.0%
Reconstructed (Age-cGAN+LMA)	**88.7%**

Table 1. "OpenFace" Face Verification (FV) scores calculated on the *LFW* dataset.

5.2. Face Reconstruction with LMA

Optimal Hyperparameters for LMA As explained in Subsection 4.1, LMA requires 2 hyper-parameters to be selected: the number of backpropagation iterations N per face image and the respective learning rate μ. In order to find optimal values for these hyperparamters, we perform a grid search varying N from 10 to 50 with a step of 10 and trying 7 different values for μ between 0.0001 and 0.01 (the extreme values for the tested parameters have been selected empirically). Thus, 35 pairs of N and μ have been tested in total. For each pair of hyperparameters N and μ, we measure how well LMA manages to preserve the identity of the original faces x by the synthetic ones $\widehat{x}$.

More precisely, both faces $x_i, i \in \{1, 2\}$ of a face verification pair are firstly projected on the synthetic manifold $N^{\bar{x}}$, and then the resulting synthetic reconstructions $\bar{x}_i^*$ are improved by LMA with the tested hyperparameters N and μ to obtain $\widehat{x_i}$. Finally, the obtained pair of synthetic images: $\widehat{x_1}$ and $\widehat{x_2}$ is fed to OpenFace and the software estimates the distance between the corresponding identities.

The face verification accuracies have been calculated following the standard *LFW* evaluation protocol[1]. Depending on the tested hyperparameters, the face verification scores have varied between 73.0% and 88.7%. The best face verification score of 88.7% has been obtained with $N = 50$ and $\mu = 0.00075$. In general, we have identified that our LMA approach is much more sensitive to the learning rate μ than to the number of backpropagation iterations N. Thus, in Figure 3, we provide an example of the reconstructed faces with the too small, the too big and the optimal learning rates of 0.0001, 0.01 and 0.00075, respectively. When the learning rate is too small, the reconstructed face identity is far from the original one (in this case, LMA does not bring anything with respect to the initial reconstruction), while the excessively big learning rate destroys the synthetic manifold and the resulting reconstruction is completely degenerated.

Improving the Identity Preservation In Figure 3, it is qualitatively visible that LMA significantly improves the identity preservation in the reconstructed face with respect to the basic Age-cGAN. In order to quantitatively measure the impact of LMA, we also evaluate the OpenFace software on the original *LFW* images x and on the reconstructions $\bar{x}^*$ obtained by the basic Age-cGAN. The face verification scores compared in Table 1 demonstrate the high effectiveness of LMA. Indeed, the difference between the scores on the original image pairs and on the ones reconstructed without LMA is more than 7 points, while our approach reduces this gap to only 0.7 points.

Finally, it is important to highlight that the proposed LMA approach does not extend a lot the execution time of the aging/rejuvenation process with respect to the basic Age-cGAN. Thus, 50 backpropagation iterations of LMA take about 0.4 second on Tesla K40c GPU for a single image, while the process of finding an optimal z^* with L-BFGS-B takes about 1.5 seconds.

5.3. Qualitative Comparison with Alternative Face Aging/Rejuvenation Methods

In this Subsection, we compare the proposed Age-cGAN+LMA aging/rejuvenation method with the most recent alternatives [28, 32, 36, 3] mentioned in Section 2. Despite the authors of [28, 32] reported improving cross-age face verification accuracy with their respective methods, it is unfortunately impossible to quantitatively compare their improvements with our results from Table 2 (presented in the next Subsection 5.4), because of the differences in experimental protocols. In particular, unlike this work, the authors of [28, 32] employed private face verification software, and did not provide explicit details on how the face verification pairs were composed from the *FG-NET* images.

Therefore, we propose to qualitatively compare the faces aged by our method and by the listed alternatives on some examples from *FG-NET* in Figure 4 (for [28, 32, 36], we use illustrations from the respective articles which have been automatically aligned to correspond to the used face crop).

The images produced by AAE [36] visually better preserve the original person's identity than the ones generated by CDL [28] and RFA [32], but the images of AAE suffer from a number of artefacts (most visible in lines 1 and 3) which make them unrealistic. At the same time, the faces aged by the basic Age-cGAN are plausible, but do not preserve the original identity as well as the other methods. Finally, thanks to our LMA approach, the proposed

[1] The LFW evaluation protocol for face verification is explained here: http://vis-www.cs.umass.edu/lfw/.

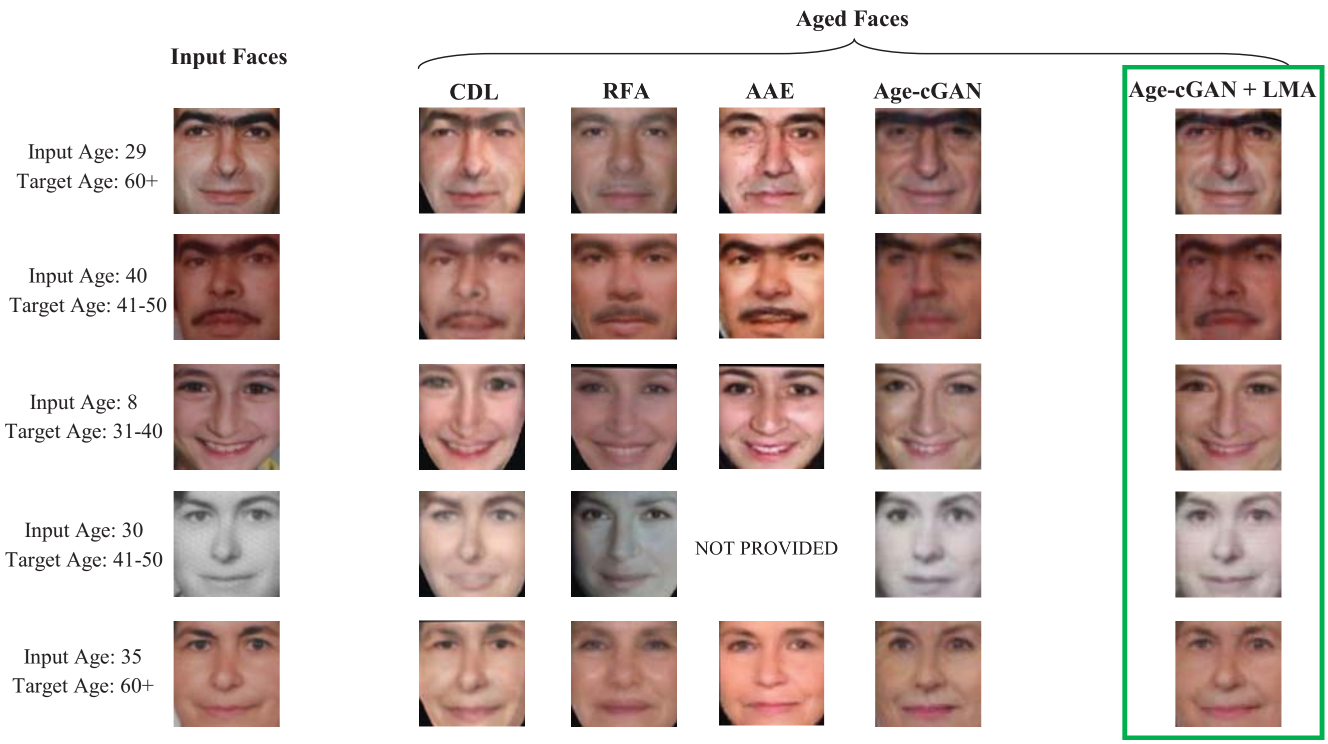


Figure 4. (Better viewed in color). Comparison of face aging by our Age-cGAN+LMA and by alternative methods from previous works. Each line corresponds to aging of a face image from the *FG-NET* dataset: the initial and the target ages are provided at the beginning of the line. 5 methods are compared: Coupled Dictionary Learning (CDL) [28], Recurrent Face Aging (RFA) [32], Adversarial AutoEncoder (AAE) [36], Age-cGAN [3], and Age-cGAN+LMA proposed in this work (highlighted by a green rectangle).

Age-cGAN+LMA aging/rejuvenation method remedies the problem of Age-cGAN producing aged faces preserving the original identity and of high visual fidelity. In addition, contrary to alternative non-generative methods (like CDL [28] and RFA [32]), our Age-cGAN+LMA can be easily adapted to also model other face attributes (like the presence of glasses or the presence of beard) just by adding the corresponding conditions to the trained cGAN. It can further improve the robustness of face verification to the variation of the respective attributes.

5.4. Boosting Cross-Age Face Verification with Age Normalization

In order to evaluate the impact of age normalization by Age-cGAN+LMA on cross-age face verification, we have selected all pairs from *FG-NET* where both face images of a pair (1) belong to the same person, (2) belong to different age categories (according to the age categories defined in Subsection 3.2), and (3) are of at least 10 years old (there are almost no children younger than 10 years old in the training *IMDB-Wiki* dataset, so we do not evaluate our algorithm on images of children). In the *FG-NET* dataset, there are 1519 face pairs fulfilling the three conditions, and we select all of them as positive pairs for face verification. We also randomly select the same number of negative pairs (composed of photos of different persons) following the same conditions (2) and (3) as for positive pairs. Among the selected pairs (both positive and negative), 61.4% have an age gap of 10-20 years, 24.0% of 20-30 years, 11.1% of 30-40 years, and 3.5% of 40+ years.

In Figure 5, we compare the “False Acceptance Rate (FAR) vs. False Rejection rate (FRR)” curves calculated on the original pairs and on the ones after age normalization. In order to highlight the necessity of the proposed LMA approach, we also plot the curve for the age-normalized pairs for which the corresponding aging/rejuvenation is performed without LMA (i.e. following the basic Age-cGAN [3]). Age normalization by Age-cGAN+LMA (the red curve) significantly increases the accuracy of face verification with respect to the original pairs (the green dashed curve). Obviously, the biggest improvements are due to falsely rejected original positive pairs (the upper left corner of Figure 5), i.e. pairs of images of the same person at different ages which are initially rejected by OpenFace, but which are accepted after removing the age difference. At the same time, the blue curve of Figure 5 perfectly demonstrates that LMA is indispensable for age normalization with Age-cGAN, because without it, the positive effect of

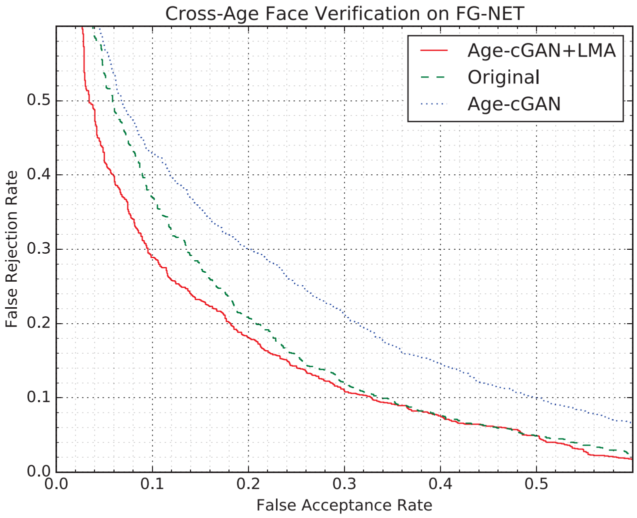


Figure 5. "FAR vs. FRR" curves of cross-age face verification on the *FG-NET* dataset. The curves have been calculated calculated (1) on age-normalized pairs generated by Age-cGAN+LMA, (2) on original pairs, and (3) on age-normalized pairs generated by basic Age-cGAN [3].

Tested Cross-Age Pairs		AUC	FRR@10	EER
All	Normalized (FS)	**89.5%**	**29.3%**	**18.9%**
	Normalized (HS)	89.2%	31.2%	19.3%
	Original	87.6%	37.1%	20.5%
≥ 40	Normalized (FS)	**84.5%**	**37.7%**	**23.6%**
	Original	80.4%	49.1%	26.4%

Table 2. Impact of age normalization with Age-cGAN+LMA on cross-age Face Verification (FV) on the *FG-NET* dataset, and comparison of 2 age normalization methods presented in Subsection 4.2: (1) Fully-Synthetic (FS) and (2) Half-Synthetic (HS). Evaluation on all cross-age positive/negative pairs and also on the pairs with a particularly huge age gap (at least 40 years of difference). Results are provided for 3 metrics: Area Under ROC Curve (AUC), False Rejection Rate (FRR) when False Acceptance Rate (FAR) is of 10% (FRR@10), and Equal Error Rate (EER).

age normalization is not enough to compensate the degradation of performances due to imperfect face reconstructions.

Table 2 summarizes the comparison of face verification on original and age-normalized image pairs according to 3 popular metrics, namely: Area Under ROC Curve (AUC), False Rejection Rate (FRR) when False Acceptance Rate (FAR) is of 10% (FRR@10) and Equal Error Rate (EER). Age normalization significantly boosts face verification with respect to all 3 metrics, and in particular, for FRR@10, it reaches almost 8 points of improvement. Moreover, the bigger is the initial age gap between the tested faces, the more important is the impact of age normalization. This is illustrated in the lower part of Table 2, where we present the scores for face verification on the test subset containing pairs with at least 40 years of age gap. In this case, FRR@10 is boosted by almost 12 points.

Finally, Table 2 (in its upper part) also compares 2 age normalization algorithms which are presented in Subsection 4.2. Despite, the fact that HS age normalization also improves the face verification accuracy with respect to original pairs, FS normalization demonstrates slightly better performances according to all 3 metrics. Moreover, the execution time difference is negligible between HS and FS algorithms because face synthesis and face verification takes less than 0.1 seconds for a pair of faces (i.e. only about 3% of the time required for reconstruction of the two original faces: cf. Subsection 5.2). Therefore, the fully-synthetic algorithm is selected for age normalization with Age-cGAN+LMA.

6. Conclusion and Future Work

In this work, we have proposed a novel generative Age-cGAN+LMA aging/rejuvenation method as well as associated fully-synthetic age normalization algorithm allowing to boost the accuracy of an off-the-shelf face verification software which is sensitive to age variations. More precisely, the following results have been obtained:

1. The proposed LMA approach significantly improves the input face's identity preservation of the basic Age-cGAN aging/rejuvenation method [3] by 7 points in terms of face verification accuracy and extending its running time by only 0.4 second per image.

2. Our fully-synthetic age normalization which is based on Age-cGAN+LMA boosts FRR@10 of OpenFace face verification software by about 8 points in average and by about 12 points in case of the biggest age gaps.

There is an important difference between Age-cGAN+LMA and alternative non-generative methods of synthetic aging/rejuvenation. Indeed, our method models not only face aging process but also human faces in general via the designed synthetic face manifold. Therefore, Age-cGAN+LMA can be adapted to also compensate other recurrent variations in human faces which can confound face verification software (e.g. facial pose, presence of beard/moustache/eyeglasses, facial expression, irregular illumination etc.) This can be done by adding the corresponding conditions to the trained cGAN (and it obviously requires a training dataset with the respective annotations). We are planning to explore this path in our future work.

References

[1] Fg-net aging dataset. `http://fipa.cs.kit.edu/433_451.php`.

[2] B. Amos, B. Ludwiczuk, and M. Satyanarayanan. Openface: A general-purpose face recognition library with mobile applications. Technical report, CMU-CS-16-118, CMU School of Computer Science, 2016.

[3] G. Antipov, M. Baccouche, and J.-L. Dugelay. Face aging with conditional generative adversarial networks. In *ICIP*, Beijing, China, 2017.

[4] P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Kumar. Localizing parts of faces using a consensus of exemplars. *IEEE TPAMI*, 35(12):2930–2940, 2013.

[5] D. M. Burt and D. I. Perrett. Perception of age in adult caucasian male faces: Computer graphic manipulation of shape and colour information. *the Royal Society of London B: Biological Sciences*, 259(1355):137–143, 1995.

[6] B.-C. Chen, C.-S. Chen, and W. H. Hsu. Cross-age reference coding for age-invariant face recognition and retrieval. In *ECCV*, Zurich, Switzerland, 2014.

[7] L. Du and H. Ling. Cross-age face verification by coordinating with cross-face age verification. In *CVPR*, Boston, USA, 2015.

[8] Y. Fu, G. Guo, and T. S. Huang. Age synthesis and estimation via faces: A survey. *IEEE TPAMI*, 32(11):1955–1976, 2010.

[9] J. Gauthier. Conditional generative adversarial nets for convolutional face generation. *Class Project for Stanford CS231N: Convolutional Neural Networks for Visual Recognition*, 2014.

[10] D. Gong, Z. Li, D. Lin, J. Liu, and X. Tang. Hidden factor analysis for age invariant face recognition. In *ICCV*, Sydney, Australia, 2013.

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *NIPS*, Montreal, Canada, 2015.

[12] I. Kemelmacher-Shlizerman, S. M. Seitz, D. Miller, and E. Brossard. The megaface benchmark: 1 million faces for recognition at scale. In *CVPR*, Las Vegas, USA, 2016.

[13] I. Kemelmacher-Shlizerman, S. Suwajanakorn, and S. M. Seitz. Illumination-aware age progression. In *CVPR*, Columbus, USA, 2014.

[14] B. Klare and A. K. Jain. Face recognition across time lapse: On learning feature subspaces. In *IJCB*, Washington DC, USA, 2011.

[15] A. Lanitis, C. J. Taylor, and T. F. Cootes. Toward automatic simulation of aging effects on face images. *IEEE TPAMI*, 24(4):442–455, 2002.

[16] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther. Autoencoding beyond pixels using a learned similarity metric. In *ICML*, New York, USA, 2016.

[17] E. Learned-Miller, G. B. Huang, A. RoyChowdhury, H. Li, and G. Hua. Labeled faces in the wild: A survey. In *Advances in Face Detection and Facial Image Analysis*, pages 189–248. Springer, 2016.

[18] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *IEEE*, 86(11):2278–2324, 1998.

[19] Z. Li, U. Park, and A. K. Jain. A discriminative model for age invariant face recognition. *IEEE TIFS*, 6(3):1028–1037, 2011.

[20] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs. A study of face recognition as people age. In *ICCV*, Rio de Janeiro, Brazil, 2007.

[21] M. Mirza and S. Osindero. Conditional generative adversarial nets. In *NIPS*, Montreal, Canada, 2014.

[22] U. Park, Y. Tong, and A. K. Jain. Age-invariant face recognition. *IEEE TPAMI*, 32(5):947–954, 2010.

[23] P. Paysan. *Statistical modeling of facial aging based on 3D scans*. PhD thesis, University of Basel, 2010.

[24] G. Perarnau, J. van de Weijer, B. Raducanu, and J. M. Álvarez. Invertible conditional gans for image editing. In *NIPS Workshops*, Barcelona, Spain, 2016.

[25] N. Ramanathan and R. Chellappa. Face verification across age progression. *IEEE TIP*, 15(11):3349–3361, 2006.

[26] R. Rothe, R. Timofte, and L. V. Gool. Dex: Deep expectation of apparent age from a single image. In *ICCV Workshops*, Santiago, Chile, 2015.

[27] F. Schroff, D. Kalenichenko, and J. Philbin. Facenet: A unified embedding for face recognition and clustering. In *CVPR*, Boston, USA, 2015.

[28] X. Shu, J. Tang, H. Lai, L. Liu, and S. Yan. Personalized age progression with aging dictionary. In *ICCV*, Santiago, Chile, 2015.

[29] J. Suo, S.-C. Zhu, S. Shan, and X. Chen. A compositional and dynamic model for face aging. *IEEE TPAMI*, 32(3):385–401, 2010.

[30] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. Deepface: Closing the gap to human-level performance in face verification. In *CVPR*, Columbus, USA, 2014.

[31] B. Tiddeman, M. Burt, and D. Perrett. Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, 21(5):42–50, 2001.

[32] W. Wang, Z. Cui, Y. Yan, J. Feng, S. Yan, X. Shu, and N. Sebe. Recurrent face aging. In *CVPR*, Las Vegas, USA, 2016.

[33] Y. Wen, Z. Li, and Y. Qiao. Latent factor guided convolutional neural networks for age-invariant face recognition. In *CVPR*, Las Vegas, USA, 2016.

[34] H. Yang, D. Huang, and Y. Wang. Age invariant face recognition based on texture embedded discriminative graph model. In *IJCB*, Florida, USA, 2014.

[35] L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Li. Face detection based on multi-block lbp representation. *Advances in Biometrics*, pages 11–18, 2007.

[36] Z. Zhang, Y. Song, and H. Qi. Age progression/regression by conditional adversarial autoencoder. In *CVPR*, Honolulu, USA, 2017.

[37] J.-Y. Zhu, P. Krähenbühl, E. Shechtman, and A. A. Efros. Generative visual manipulation on the natural image manifold. In *ECCV*, Amsterdam, Netherlands, 2016.