# On the Influence of Ageing on Face Morph Attacks: Vulnerability and Detection

Sushma Venkatesh Kiran Raja Raghavendra Ramachandra Christoph Busch
Norwegian University of Science and Technology (NTNU), Norway
E-mail: {sushma.venkatesh;kiran.raja;raghavendra.ramachandra;christoph.busch} @ntnu.no

## Abstract

*Face morphing attacks have raised critical concerns as they demonstrate a new vulnerability of Face Recognition Systems (FRS), which are widely deployed in border control applications. The face morphing process uses the images from multiple data subjects and performs an image blending operation to generate a morphed image of high quality. The generated morphed image exhibits similar visual characteristics corresponding to the biometric characteristics of the data subjects that contributed to the composite image and thus making it difficult for both humans and FRS, to detect such attacks. In this paper, we report a systematic investigation on the vulnerability of the Commercial-Off-The-Shelf (COTS) FRS when morphed images under the influence of ageing are presented. To this extent, we have introduced a new morphed face dataset with ageing derived from the publicly available MORPH II face dataset, which we refer to as MorphAge dataset. The dataset has two bins based on age intervals, the first bin - MorphAge-I dataset has 1002 unique data subjects with the age variation of 1 year to 2 years while the MorphAge-II dataset consists of 516 data subjects whose age intervals are from 2 years to 5 years. To effectively evaluate the vulnerability for morphing attacks, we also introduce a new evaluation metric, namely the Fully Mated Morphed Presentation Match Rate (FMMPMR), to quantify the vulnerability effectively in a realistic scenario. Extensive experiments are carried out using two different COTS FRS (COTS I Cognitec FaceVACS-SDK Version 9.4.2 and COTS II - Neurotechnology version 10.0) to quantify the vulnerability with ageing. Further, we also evaluate five different Morph Attack Detection (MAD) techniques to benchmark their detection performance with respect to ageing.*

## 1. Introduction

Facial characteristics have been well explored for identifying and verifying individuals and numerous biometric systems have been deployed in operational applications for many years [21, 23]. The preference towards face based biometric systems is founded on multiple factors such as ease of capture of facial characteristic without invasive imaging, capturing at a stand-off distance both in semi-cooperative (voluntary identification/verification) and uncooperative scenarios (surveillance) [27, 5, 26]. While many of the breakthrough articles detailing iris and vein recognition systems have shown impeccable accuracy with very low false accepts and false rejects, those systems suffer from highly constrained image capturing processes. In order to reach the performance of such iris and vein recognition systems, face biometrics has seen benefits from recent algorithmic advancements, which was focused on features that have been engineered in a robust manner [31, 30, 6], and pre-processing that has been improvised [35] by including end-to-end learning using Deep Neural Networks (DNN) even in large scale applications [26, 39].

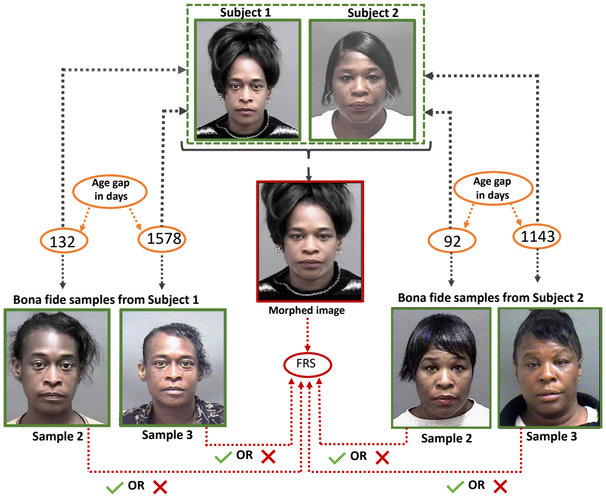


Figure 1: Illustration of the influence of ageing on face morphing

Such attractive and inherent advantages of the face modality have led to a wider deployment of Face Recognition Systems (FRS) in passport issuance processes, visa

978-1-7281-9186-7/20/$31.00 ©2020 IEEE

management, identity management and Automated Border Control (ABC). Despite the high accuracy and convenience of face biometrics, FRS systems are impeded by various factors such as ageing [34, 22], partial face availability [22] and also imperilled by various attacks that include presentation attacks (spoofing-attacks) with print, display or silicon mask attack instruments [32], make-up attacks [48, 2], coverted mask attacks [8], morphing attacks [11], database level attacks [25, 3] and comparison level attacks [18, 24]. While many of the attacks have been addressed through mitigation measures over the period of time, we focus on recently surfaced face morphing attacks [11, 36] in this work. Despite some of the recent works proposing measures to mitigate these attacks through various approaches [42, 14, 13, 36] a number of covariates are reported to impact the attack detection performance. A list of covariates impacting the performance of morphing attack detection include the techniques used to generate the morphed image [11], the configuration of the print-scan pipeline [37], factors of age and ethnicity [40] among many other unknown factors. With a clear introspection of the existing works, we observe that both the FRS vulnerability and also the Morphing Attack Detection (MAD) performance under variation of age is not studied in the context of morphing attacks, despite the fact that the issue was pointed out already in the early works in this domain [40, 13].

Starting with this observation, we focus in this work on establishing the impact of ageing on morphing attacks by carefully studying the vulnerability of FRS and MAD performance of currently reported MAD algorithms under the influence of ageing. The key motivation stems from earlier works who have disentangled the impact of ageing on face recognition systems with respect to recognition performance [34, 22, 1] and a number of works that have proposed approaches to handle the associated performance limitations [34, 29, 19, 28, 1]. We therefore provide a brief overview of impact of ageing in the subsequent section and thereafter illustrate the impact of ageing specifically for morphing. Further, we focus our work on investigating the impact using the digital images alone due to two primary factors: (i) many countries across the world allow to upload digital images via web-portal for passport renewal and visa issuance, and (ii) to align our works with recent studies focusing on digital MAD [13].

### 1.1. Facial Ageing

Facial ageing is a commonly observed phenotype of human ageing, which is visibly seen. Despite the complexity of understanding the characteristic changes associated with the facial ageing, a number of works have reported the role of skin and soft tissues and their impact on visible changes of facial appearance [10]. Complementary works have demonstrated the role of loss of facial bone volume to contribute to facial appearance under ageing progress [45]. As it can be deduced, facial ageing being a complex process involving soft tissues and skeletal structure changes, it is influenced by many factors, such as exposure to sunlight and body weight among others. As an additional factor, large variations in facial ageing across individuals and ethnic populations can further be observed [33]. While in face recognition, the main differences in exterior facial structure making individuals distinguishable from each other allows recognition analysis to achieve high identification accuracy, a longitudinal study of the same face over a period of time has shown to challenge the accuracy [1].

### 1.2. Facial Ageing and Morphing Attacks

Under the observation of complex changes of facial appearance, which bring down the recognition accuracy of FRS unless proper measures are taken, our assertion is that the effect and impact on morphing attacks may change. For electronic Machine Readable Travel Documents (MRTDs) a typical life-cycle of 10 years is recommend [9] meaning that the drastic changes in facial appearance must be tolerated as intra-class variance during that life-cycle, while up to now the impact of morphing and its correlation with the progressing of the potentially morphed reference image in this life-cycle, has neither been considered nor investigated. Initial studies on morphing attacks have demonstrated the ability to fool a human expert (i.e. trained border guards) with morphed facial images. The changes of facial appearance, which are caused by ageing, are illustrated in Figure 1. Our assertion is to validate the impact of ageing and thus we formulate three specific research questions:

- How vulnerable are COTS FRS when a composite morph image is enrolled and is after a period of ageing probed against a live image from one of the contributing subjects?
- Do current Morphing Attack Detection (MAD) algorithms scale-up to detect such attacks under the influence of ageing?
- What is the impact of different alpha (or blending, morphing) factors used to generate the morphed image under the constraint of ageing, specifically with respect to MAD?

We address each of these questions in a systematic manner through our contributions. We focus in this work to first establish the impact on FRS through an extensive empirical evaluation. While a detailed study of appearance change is more of a cognitive study, it is beyond the scope of the current work.

### 1.3. Contributions of Our Work

While the hypothesis is well justified, we also note that there exists no database with morphing and ageing according to the current literature. With such a caveat, we focus on first creating a database to facilitate and validate our assertion.

- The first key contribution is the creation of a (moderately) large-scale database of morphed faces with ageing covariate by employing the MORPH II non-commercial face dataset [4], which is hereafter referred as MorphAge Database.
- We investigate the vulnerability of FRS to such attacks by employing two widely used Commercial-Off-The-Shelf (COTS) FRS systems. This contribution not only helps in verifying our assertion but also validates the usefulness of the newly created database. Further, we also investigate the role of alpha (or blending, morphing) factor (with $\alpha =$ 0.3. 0.5 and 0.7) while analysing the vulnerability under ageing.
- As a third contribution, we employ a set of recently reported morphing attack detection algorithms to benchmark detection performance and thereby identify the impediments if any.

In the rest of the paper, we first provide details on the newly constructed database in Section 2 and in Section 3 we investigate the vulnerability of FRS using two COTS FRS. Further, the benchmarking of morphing attack detection systems is detailed in Section 4 while the key observations and conclusions are reported in Section 6.

## 2. MorphAge Database Construction

To effectively study the influence of ageing on face morphing vulnerability and morph detection, we introduce a new dataset, which is derived from the MORPH II non-commercial dataset [4] that is publicly available. The MORPH II dataset consists of a total of 55000 unique samples captured from 13000 data subjects. The images are captured over the time span from 2003 to 2007. The age of the subjects varies from 16 to 77 years. The dataset consists of male and female subjects with different ethnicity (African, European, Asian, Hispanic). In this work, we choose the MORPH II dataset motivated by the large number of subjects, the quality of the captured data and the variation in age for one and the same subject across different capture sessions.

The newly constructed MorphAge dataset is binned in two age groups from MORPH II dataset. The first bin - Age Group (MorphAge-I) consists of 1002 unique data subjects with a gender distribution of 143 female and 859 male subjects. For each data subject, three different samples are chosen such that the first session corresponds to the high quality data capture (younger age), second session corresponds to the aged capture of 1-8 months from first session and third session corresponds to the aged capture of same subject between 1-2 years from first session. The second bin - Age Group (MorphAge-II) is comprised of 516 unique data subjects sub-sampled from the MORPH II dataset with 62 female and 454 male data subjects. Each data subject was captured in three different sessions. The first session corresponds to the high quality data capture (younger age), the second session corresponds again to a time lapse of 1-8 months from the first session and the third session corresponds to an aged capture of 2 years up to 5 years after the first session.

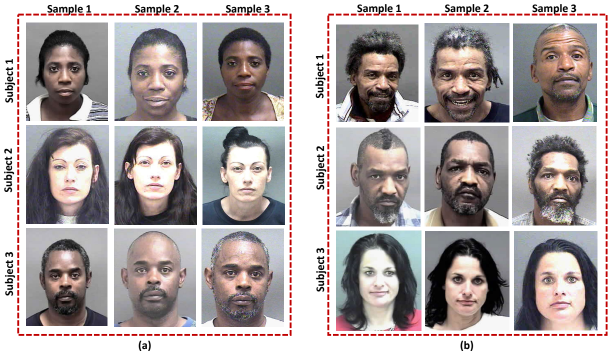


Figure 2: Illustration of sample images from newly constructed MorphAge dataset (a) MorphAge-I (1 year to 2 years) (b) MorphAge-II (2 years to 5 years)

Table 1: Statistics of bona fide and morphed images in MorphAge Database

| Session | Dev | Training | Testing | Total |
|---|---|---|---|---|
| MorphAge-I Subset | | | | |
| Session 1<br>(used for morphing) | 251 | 500 | 251 | 1002 |
| Session 2<br>(used for vulnerability) | 251 | 500 | 251 | 1002 |
| Session 3<br>(with age difference) | 251 | 500 | 251 | 1002 |
| Morphed Images | 1980 | 6614 | 1944 | 10538 |
| MorphAge-II Subset | | | | |
| Session 1<br>(used for morphing) | 130 | 257 | 129 | 516 |
| Session 2<br>(used for vulnerability) | 130 | 257 | 129 | 516 |
| Session 3<br>(with age difference) | 130 | 257 | 129 | 516 |
| Morphed Images<br>(with different morphing factors) | 648 | 2310 | 809 | 3767 |

### 2.1. MorphAge-I and MorphAge-II - Bonafide Set

In both bins, i.e. MorphAge-I and MorphAge-II, we select for each data subject three samples (one of each session) such that the *first session sample is used only to generate the morphing image, the sample from the second session is used as bona fide sample in the morph attack detection experiments and the third session to analyze the vulnerability of the commercial FRS.* As seen from the Figure 2, the facial appearance changes significantly with the increasing age which cannot be modelled geometrically or morphologically for any particular ethnicity or age group for both the bins (MorphAge-I and MorphAge-II).

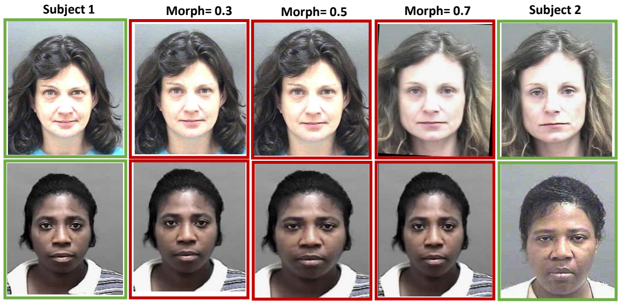


Figure 3: Example of generated morphed images

### 2.2. MorphAge-I and MorphAge-II - Morphed Image Set

To generate the morphed image datasets for the subjects represented in our newly constructed dataset, we have used the face morph generation tool from Ferrara et al. [15] [12], which is based on facial landmarks based warping and weighted linear blending to generate a high quality morphed image. We particularly, choose this technique for morphing generation over other type of generators based on GAN [7] by considering: (1) high quality of the generated morphed images, in order to establish a significant threat to the tested commercial FRS [15] (2) high quality of generated morphed image, such that the submitted images are considered compliant with the requirements in the ICAO standards and (3) feasibility to create the morphed images with various blending and warping factors.

In this work, the morphing process is carried out between only two data subjects by considering its use-case in a real-life scenario where typically one criminal morphs his/her face image with the image of an accomplice. To carefully select the pair of images for the morphing process, we use the COTS-I FRS, which is widely used in Automated Border Control installations. Through the FRS, a set of similarity scores is obtained between the probe image of a selected data subject against the reference images of all data subjects. We then choose the pair of images that are successfully verified at FMR = $0.1\%$ with high scores to retain a high degree of similarity between two constituting subjects for the morphed image. Additional care is exercised not to combine data subjects with different genders and also to separate the data subject into three independent groups such as non-overlapping training, testing and development sets [40, 43]. For a selected image pair, we generate three morphed images at three different morphing (or blending) factors $\alpha = 0.3, 0.5, 0.7$ to obtain insights with regard to the impact of ageing at different blending factors. Figure 3 shows the example of morphed face images with three different blending factor within our MorphAge dataset.

Table 1 presents the statistics of the generated dataset corresponding to the two bins - MorphAge-I and MorphAge-II. Further, in order to evaluate the MAD performance, we have divided the whole datasets into three independent and non-overlapping subsets for training, development and testing. The training subset is used purely to train the MAD techniques, the development subset is used to optimize and adjust the operating threshold for the MAD techniques and finally the testing subset is solely used to analyze the detection performance obtained at the optimal threshold.

## 3. Vulnerability Analysis

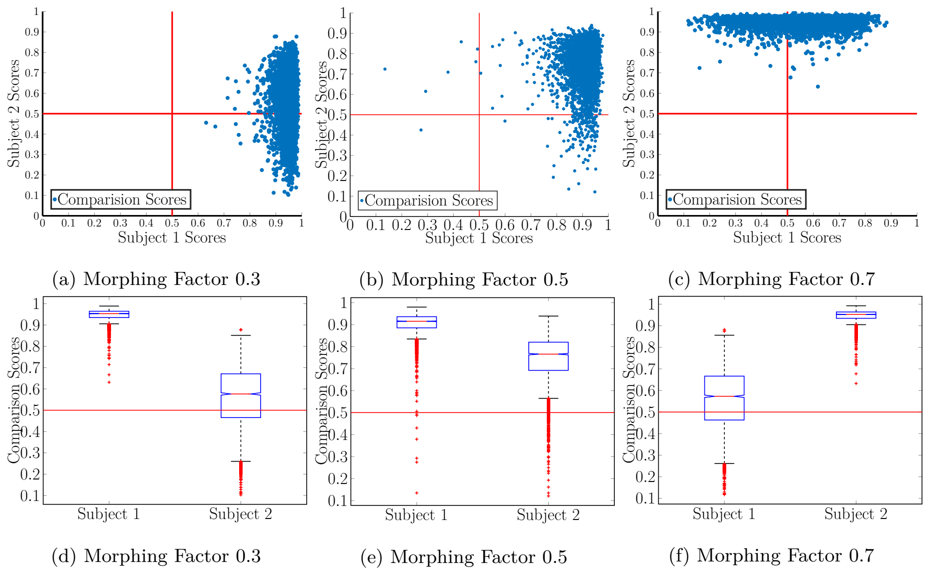
In this section, we present the vulnerability analysis of the FRS, when confronted with the morphed images under variation of age. To this extent, we employ two different COTS Face Recognition Systems (FRS) namely, COTS-I Cognitec FaceVACS-SDK Version 9.4.2 and COTS-II Neurotechnology Version 10.0. To effectively measure the vulnerability of the FRS against morphed face samples, we set a realistic constraint that all contributing data subjects (in our case two) must exceed the verification threshold of the FRS. Further, in this work, we set the operating threshold of both COTS FRS to FAR = $0.1\%$ following the guidelines of FRONTEX [17] for automated border control. Thus, we coin the new realistic constraint using a new vulnerability metric as *Fully Mated Morphed Presentation Match Rate (FMMPMR)* that can be computed as:

$$\begin{aligned} &FMMPMR \\ &= \frac{1}{P}\sum_{M,P}(S1_M^P > \tau)\&\&(S2_M^P > \tau)\ldots\&\&(Sk_M^P > \tau) \end{aligned} \tag{1}$$

Where $P = 1, 2, \ldots, p$ represent the number of attempts made by presenting all the probe images from the contributing subject against $M^{th}$ morphed image, $K = 1, 2, \ldots, k$ represents the number of contributing data subjects to the constitution of the generated morphed image (in our case $K = 2$), $Sk_M^P$ represents the comparison score of the $K^{th}$ contributing subject obtained with $P^{th}$ attempt (in our case the $P^{th}$ probe image from the dataset) corresponding to $M^{th}$ morph image and $\tau$ represents the threshold value corresponding to FAR = $0.1\%$.

We have employed the new metric FMMPMR considering the fact that the existing vulnerability metric MMPMR[40] accounts only for the morphed images getting verified with the contributing subjects without taking into account the number of attempts. However, the new metric FMMPMR overcomes this drawback and considers each and every attempt a morphed image gets verified with

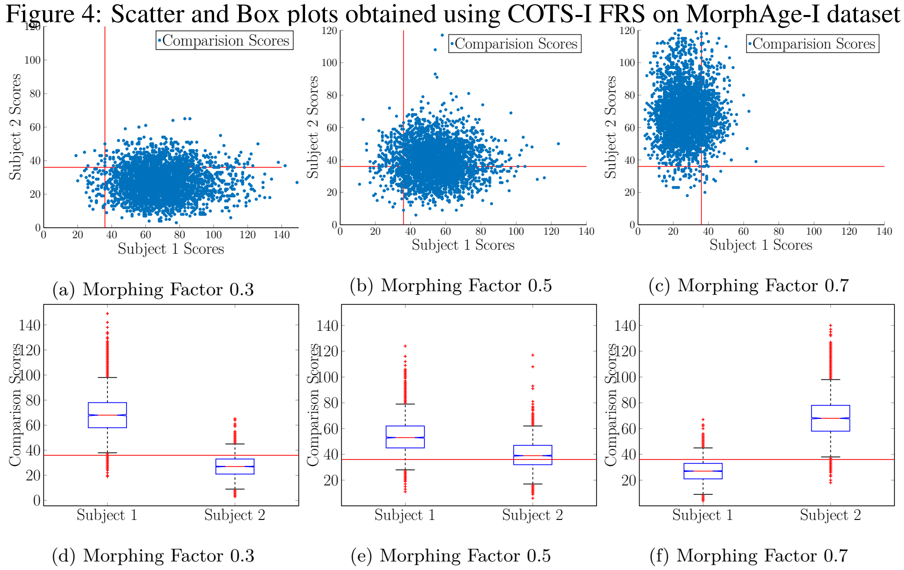
Outcome not necessarily constitutes the best the algorithm can do.



(a) Morphing Factor 0.3 (b) Morphing Factor 0.5 (c) Morphing Factor 0.7

(d) Morphing Factor 0.3 (e) Morphing Factor 0.5 (f) Morphing Factor 0.7

Figure 4: Scatter and Box plots obtained using COTS-I FRS on MorphAge-I dataset



(a) Morphing Factor 0.3 (b) Morphing Factor 0.5 (c) Morphing Factor 0.7

(d) Morphing Factor 0.3 (e) Morphing Factor 0.5 (f) Morphing Factor 0.7

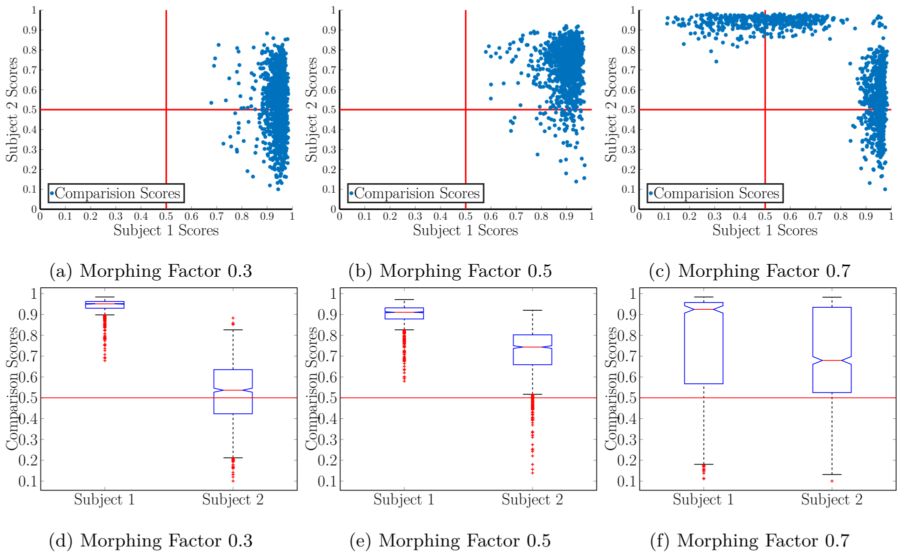
Figure 5: Scatter and Box plots obtained using COTS-II FRS on MorphAge-I dataset

the pair of contributing subjects, i.e., reflecting the actual vulnerability of a FRS.

Table 2: Vulnerability analysis: FMMPMR (%)

| Morphing factor ($\alpha$) | **FMMPMR**(%) | | | |
|---|---|---|---|---|
| | MorphAge-I | | MorphAge-II | |
| | COTS-I | COTS-II | COTS-I | COTS-II |
| 0.3 | 66.24 | 18.42 | 58.47 | 17.29 |
| 0.5 | 95.07 | 56.96 | 93.81 | 51.27 |
| 0.7 | 67.32 | 18.21 | 58.18 | 15.61 |

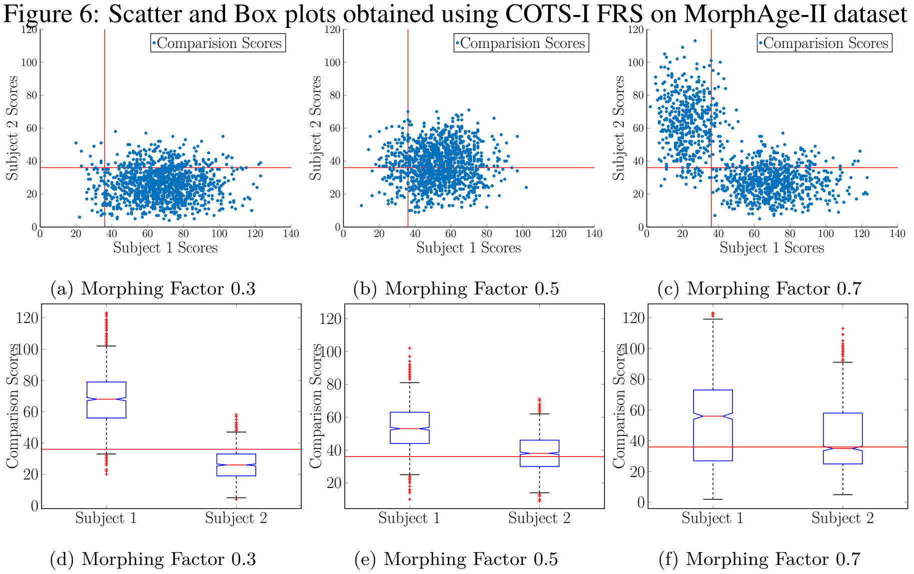
Table 2 indicates the FMMPMR (%) computed using the two COTS FRS on both bins - MorphAge-I and MorphAge-II. Figure 4 and Figure 5 shows the scatter plot and box plot for MorphAge-I dataset from two COTS FRS respectively. Figure 4(a), 4(b), 4(c) and Figure 5(a), 5(b), 5(c) provides the visualization of the comparison scores when the morphed image is enrolled, and both contributing data subjects are probed for both FRS. In the most serve conditions, meaning a high vulnerability of the FRS with regards to morphing attacks, we will obtain comparison scores that are clustered in the top right corner of the figure. The vertical and horizontal lines indicate the threshold that is recommended by the COTS FRS for operational settings in the border control application corresponding to $FMR = 0.1\%$. Figure 4(d), 4(e), 4(f) and 5(d), 5(e), 5(f) shows the box



(a) Morphing Factor 0.3 (b) Morphing Factor 0.5 (c) Morphing Factor 0.7

(d) Morphing Factor 0.3 (e) Morphing Factor 0.5 (f) Morphing Factor 0.7

Figure 6: Scatter and Box plots obtained using COTS-I FRS on MorphAge-II dataset



(a) Morphing Factor 0.3 (b) Morphing Factor 0.5 (c) Morphing Factor 0.7

(d) Morphing Factor 0.3 (e) Morphing Factor 0.5 (f) Morphing Factor 0.7

Figure 7: Scatter and Box plots obtained using COTS-II FRS on MorphAge-II dataset

plot that provides insight on the distributions of comparison scores corresponding to the contributor probe images allowing us to understand which of the two probe images (of the contributing subjects) are more vulnerable for the FRS. In similar lines, Figure 6 and 7 shows the scatter plot and box plot for the MorphAge-II dataset that are computed from the two COTS FRS. Based on the obtained results the following are our main observations:

- *Intra-Age Groups:* As expected the morphed image with the morphing factor of 0.5 indicates the highest vulnerability as reflected by both COTS FRS. However, the morphing factor of 0.3 and 0.7 indicates a reduced vulnerability that can be attributed to the morphing factor weights leaning toward only one of the contributing data subjects. This fact is illustrated in Figure 4, 5, 6 and 7, where we can observe that with a morphing factor of 0.3, the subject 1 is likely to be verified. While with a morphing factor of 0.7, in most cases, subject 2 is likely to be verified rather than subject 1. While not so surprising, the morphing factor of 0.5 indicates (almost) equally both contributing subjects can be verified.
- *Inter-Age groups:* Based on the obtained results, it is also interesting to note the direct influence on the morphing factor on the vulnerability. Thus, with the morphing factor of 0.3 and 0.7, both COTS FRS shows a greater reduction value of FMMPMR on MorphAge-II dataset. This indicates morphing attacks pose lesser threats to FRS under the influence of ageing. However, with the morphed

Table 3: Experiment-I: Quantitative performance of the MAD techniques on MorphAge-I

| Algorithm | Development Set | Testing set | | | |
|---|---|---|---|---|---|
| | EER (%) | EER (%) | BPCER (%) @ APCER (%) = | | |
| | | | 1 | 5 | 10 |
| Morphing factor ($\alpha$) 0.3 | | | | | |
| LBP-SVM [36, 46, 16, 41] | 28.14 | 35.11 | 84.4 | 68.8 | 56.8 |
| BSIF-SVM [36, 41] | 31.82 | 37.59 | 98.8 | 90 | 73.2 |
| HOG-SVM citeScher2017 | 32.09 | 33.51 | 84.4 | 63.6 | 53.6 |
| AlexNet-SVM [16, 38, 44] | 4.38 | 2 | 7.2 | 3.2 | 0.8 |
| Color Denoising [47] | 1.63 | 3.65 | 5.2 | 0.4 | 0.4 |
| Morphing factor ($\alpha$) 0.5 | | | | | |
| LBP-SVM [36, 46, 16, 41] | 27.82 | 33.76 | 75.2 | 59.2 | 57.2 |
| BSIF-SVM [36, 41] | 31.82 | 36.9 | 98.8 | 89.21 | 73.6 |
| HOG-SVM [41] | 30.73 | 34.1 | 81.2 | 63.2 | 56.8 |
| AlexNet-SVM [16, 38, 44] | 3.18 | 2.01 | 4.12 | 0 | 0 |
| Color Denoising [47] | 1.63 | 1.21 | 7.6 | 0.4 | 0 |
| Morphing factor ($\alpha$) 0.7 | | | | | |
| LBP-SVM [36, 46, 16, 41] | 28.86 | 34.92 | 88.4 | 66.8 | 57.2 |
| BSIF-SVM [36, 41] | 31.9 | 37.98 | 98.8 | 88 | 73.2 |
| HOG-SVM [41] | 32.98 | 33.38 | 80 | 62.8 | 57.2 |
| AlexNet-SVM [16, 38, 44] | 5.08 | 2.78 | 5.6 | 2 | 0 |
| Color Denoising [47] | 2.75 | 2.43 | 13.2 | 2 | 0.4 |

Table 4: Experiment-I: Quantitative performance of the MAD techniques on MorphAge-II

| Algorithm | Development Set | Testing set | | | |
|---|---|---|---|---|---|
| | EER (%) | EER (%) | BPCER (%) @ APCER (%) = | | |
| | | | 1 | 5 | 10 |
| Morphing factor ($\alpha$) 0.3 | | | | | |
| LBP-SVM [36, 46, 16, 41] | 30.64 | 29.21 | 61.24 | 48.83 | 44.96 |
| BSIF-SVM [36, 41] | 33.35 | 39.17 | 58.91 | 51.16 | 48.83 |
| HOG-SVM [41] | 32.56 | 32.56 | 66.66 | 51.93 | 45.73 |
| AlexNet-SVM [16, 38, 44] | 4 | 5.49 | 7.75 | 4.65 | 4.65 |
| Color Denoising [47] | 3.15 | 1.7 | 3.1 | 0 | 0 |
| Morphing factor ($\alpha$) 0.5 | | | | | |
| LBP-SVM [36, 46, 16, 41] | 28.71 | 32.39 | 68.99 | 48.06 | 41.08 |
| BSIF-SVM [36, 41] | 32.65 | 39 | 63.56 | 51.16 | 48.83 |
| HOG-SVM [41] | 30.33 | 32.56 | 62.02 | 52.71 | 44.96 |
| AlexNet-SVM [16, 38, 44] | 2.92 | 3.78 | 6.2 | 3.87 | 3.11 |
| Color Denoising [47] | 3.77 | 0.75 | 1.55 | 0.77 | 0.77 |
| Morphing factor ($\alpha$) 0.7 | | | | | |
| LBP-SVM [36, 46, 16, 41] | 29.33 | 27.03 | 58.91 | 51.98 | 45.73 |
| BSIF-SVM [36, 41] | 34.7 | 33.07 | 58.91 | 51.16 | 48.06 |
| HOG-SVM [41] | 31.02 | 29.09 | 73.64 | 58.91 | 44.96 |
| AlexNet-SVM [16, 38, 44] | 3.15 | 5.5 | 11.62 | 5.42 | 4.65 |
| Color Denoising [47] | 3.15 | 0.75 | 3.1 | 0 | 0 |

factor of 0.5, the COTS-II FRS indicates lower values of FMMPMR, while COTS-I indicates a moderate reduction in the vulnerability despite being very significant.

- Observing the box plots for the morphing factor of 0.5 from both MorphAge-I and MorphAge-II, it can be noted that, both the median and whiskers corresponding to the comparison scores from both subjects are reduced in MorphAge-II when compared to MorphAge-I. These observations, together with the quantitative value of FMMPMR, indicate the reduced threats to morphing attacks on FRS under ageing. This fact is consistently observed for both COTS FRS and are statistically significant as observed in the box plots.
- *Role of COTS FRS:* The COTS-I FRS indicates the highest vulnerability on three morphing factors when compared to that of the COTS-II FRS. The morphing factor with 0.5 shows the highest FMMPMR with 95.07% on MorphAge-I and 93.81% on MorphAge-II with COTS-I FRS. The lowest value of FMMPMR is noted with COTS-II FRS with a morphing factor of 0.7 in the MorphAge-II dataset.

## 4. Face Morph Attack Detection Performance

In this section, we benchmark the most recent digital MAD techniques on the newly created MorphAge dataset. The goal of this experiment is to understand the impact of ageing on the detection performance of the MAD techniques. To this extent, we design two different experiments to reflect the variation in the performance of the MAD techniques under the influence of ageing. **Experiment-I:** the evaluation protocol is designed to evaluate the MAD detectors in the same age group. Thus, the MAD detectors are trained and tested with the same group data. **Experiment-II:** is designed to evaluate the performance of MAD detection with the variation in age. Thus, MAD detectors are trained with the MorphAge-I data and tested with only the MorphAge-II dataset. In both experiments, the corresponding development dataset is used to tune the parameters of the algorithm and also to compute the operating threshold at APCER = 1%, 5% and 10%. In this work, we have evaluated five different MAD schemes such as: Local Binary Pattern (LBP) LBP-SVM [36, 46, 16, 41], Binarized Statistical Image Features (BSIF) [36, 41], Histogram of Oriented Gradients (HOG) [41], AlexNet [16, 38, 44] and Color Denoising [47]. We have considered these five MAD techniques as they have indicated good performance on three different large scale digital morphing datasets [47]. The quantitative results are presented according to the ISO/IEC 30107-3 [20] metrics such as Bona fide Presentation Classification Error Rate (BPCER(%)) and Attack Presentation Classification Error Rate (APCER (%)) along with D-EER(%).

Table 3 and Table 4 indicates the quantitative results of the MAD schemes on two different age groups MorphAge-I and MorphAge-II respectively on the Experiment-I protocol. Based on the obtained results, it can be noticed that:

- The traditional MAD methods based on LBP, BSIF, and HOG fail to indicate acceptable detection performance for both MorphAge-I and MorphAge-II dataset.
- Recently introduced MAD techniques based on AlexNet

and Color denoising techniques have shown excellent performance in detecting morphing attacks.

- It is interesting to note that the MAD methods do not show any influence of the different morphing factors on the detection performance. The detection performance with different morphing factor did further not vary irrespective of the age group as well.
- Among the five benchmarked different MAD techniques, the color denoising MAD has indicated the best performance across various morphing factors ($\alpha$) for both MorphAge-I and MorphAge-II.

Table 5: Experiment-II: Quantitative detection performance of MAD techniques on MorphAge-I v/s. MorphAge-II

| **Algorithm** | **Development Set** | **Testing set** | | | |
|---|---|---|---|---|---|
| | EER (%) | EER (%) | BPCER (%) @ APCER (%) = | | |
| | | | 1 | 5 | 10 |
| Morphing factor ($\alpha$) 0.3 | | | | | |
| LBP-SVM [36, 46, 16, 41] | 28.14 | 34.19 | 92.24 | 65.89 | 47.28 |
| BSIF-SVM [36, 41] | 31.82 | 44.13 | 100 | 98.44 | 84.49 |
| HOG-SVM [41] | 32.09 | 41.86 | 91.47 | 70.54 | 62.01 |
| AlexNet-SVM [16, 38, 44] | 4.38 | 3.03 | 8.52 | 3.10 | 2.32 |
| Color Denoising [47] | 1.63 | 2.27 | 1.55 | 0.45 | 0 |
| Morphing factor ($\alpha$) 0.5 | | | | | |
| LBP-SVM [36, 46, 16, 41] | 27.82 | 33.29 | 86.04 | 66.66 | 48.06 |
| BSIF-SVM [36, 41] | 31.82 | 45.42 | 100 | 96.89 | 84.49 |
| HOG-SVM [41] | 30.73 | 37.95 | 85.27 | 67.44 | 57.36 |
| AlexNet-SVM [16, 38, 44] | 3.18 | 0.94 | 3.10 | 0.77 | 0.39 |
| Color Denoising [47] | 1.63 | 1.59 | 0.7 | 0 | 0 |
| Morphing factor ($\alpha$) 0.7 | | | | | |
| LBP-SVM [36, 46, 16, 41] | 28.86 | 32.24 | 93.20 | 65.89 | 49.61 |
| BSIF-SVM [36, 41] | 31.90 | 37.33 | 100 | 96.89 | 84.49 |
| HOG-SVM [41] | 32.98 | 33.54 | 88.37 | 68.99 | 58.13 |
| AlexNet-SVM [16, 38, 44] | 5.08 | 2.27 | 6.97 | 3.87 | 0.77 |
| Color Denoising [47] | 2.75 | 2.46 | 3.10 | 0.40 | 0 |

Table 5 indicates the quantitative detection performance of MAD methods in Experiment-II. Based on the obtained results, it can be noted that the ageing does not influence the performance of the MAD methods. It is worth noting that, in this protocol, MAD methods are trained using only MorphAge-I dataset and are tested on the MorphAge-II dataset with the age difference up to 5 years. Further, the data subjects in MorphAge-I and MorphAge-II do not overlap. Among the five different MAD methods, color denoising based MAD has again indicated the best performance for all three morphing factors ($\alpha$). As it can be deduced, ageing does not influence the detection capabilities of MAD under the performed experimental settings.

## 5. Discussion

Based on the observations made above from the experiments and obtained results, the research questions formulated in Section 1.2 are answered below.

- Q1. How vulnerable are COTS FRS when a composite morph image is enrolled and is after a period of ageing probed against a live image from one of the contributing subjects?
  - Supported by the obtained experimental results reported in Table 2, it is interesting to note that the value of FMMPMR is reduced to certain extent in case of MorphAge-II dataset. The morphed images are not easily verified against the probe images after a certain degree of ageing making FRS less vulnerable.
- Q2. Do current Morphing Attack Detection (MAD) algorithms scale-up to detect such attacks under the influence of ageing?
  - Based on the experimental results reported in Table 4, ageing has negligible impact on the MAD and thereby the existing MAD schemes can detect the attacks even under ageing.
- Q3. What is the impact of different alpha (or blending, morphing) factors used to generate the morphed image under the constraint of ageing, specifically with respect to MAD?
  - Based on the experimental results, it is interesting to note that the morphing factors $alpha = 0.3$ and $0.7$ show greater reduction in the vulnerability in both the COTS FRS with respect to ageing as reported in Table 2. It has to be however noted that COTS-II FRS indicates lower vulnerability when a morphing factor of 0.5 is employed.

## 6. Conclusion

We have presented an empirical study on quantifying the vulnerability of COTS FRS with regards to morphing attacks under the influence of ageing. We have introduced a new dataset with two different age groups derived from the publicly available MORPH II face dataset referred as MorphAge-I and MorphAge-II. Further, we have also introduced a new evaluation metric namely, Fully Mated Morphed Presentation Match Rate(FMMPMR) to quantify the vulnerability effectively. Extensive experiments were carried out using two different COTS FRS and three different morphing factors(with $\alpha$ = 0.3, 0.5 and 0.7). Based on the obtained results, it is observed that impact of ageing reduces the vulnerability from morphing attacks on COTS FRS. The reduction in the vulnerability is more prominent when the morphing factor is $\alpha$ = 0.3 and 0.7. However with a morphing factor of $\alpha$ = 0.5, the vulnerability does not change significantly with the COTS-I, while COTS-II FRS still indicates a significant reduction in the vulnerability. Extensive experiments were performed to quantify the performance variation of the MAD methods under the influence of ageing. To this extent, three different evaluation protocols are presented that show no influence of ageing on morph attack detection performance. It is also interesting to note that robust MAD methods are not sensitive to variations of the morphing factor even under the influence of ageing.

# References

[1] L. Best-Rowden and A. K. Jain. Longitudinal study of automatic face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 40(1):148–162, 2017.

[2] A. Bharati, R. Singh, M. Vatsa, and K. W. Bowyer. Detecting facial retouching using supervised deep learning. *IEEE Transactions on Information Forensics and Security*, 11(9):1903–1913, 2016.

[3] B. Biggio, L. Didaci, G. Fumera, and F. Roli. Poisoning attacks to compromise face templates. In *2013 International Conference on Biometrics (ICB)*, pages 1–7. IEEE, 2013.

[4] G. Bingham, K. Kempfert, B. Yip, J. Fabish, M. Ferguson, C. Nansalo, K. Park, R. Towner, T. Kling, Y. Wang, et al. Preliminary studies on a large face database morph-ii.

[5] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman. Vggface2: A dataset for recognising faces across pose and age. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pages 67–74. IEEE, 2018.

[6] J. Chen, V. M. Patel, L. Liu, V. Kellokumpu, G. Zhao, M. Pietikäinen, and R. Chellappa. Robust local features for remote face recognition. *Image and Vision Computing*, 64:34–46, 2017.

[7] N. Damer, A. M. Saladié, A. Braun, and A. Kuijper. Morgan: Recognition vulnerability and attack detectability of face morphing attacks created by generative adversarial network. In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–10, Oct 2018.

[8] N. Erdogmus and S. Marcel. Spoofing 2d face recognition systems with 3d masks. In *2013 International Conference of the BIOSIG Special Interest Group (BIOSIG)*, pages 1–8. IEEE, 2013.

[9] M. Erickson. Passport canada's citizen engagement process and the new 10-year epassport. 2013.

[10] J. P. Farkas, J. E. Pessa, B. Hubbard, and R. J. Rohrich. The science and theory behind facial aging. *Plastic and Reconstructive Surgery Global Open*, 1(1), 2013.

[11] M. Ferrara, A. Franco, and D. Maltoni. The magic passport. In *IEEE International Joint Conference on Biometrics*, pages 1–7. IEEE, 2014.

[12] M. Ferrara, A. Franco, and D. Maltoni. *Face Recognition Across the Imaging Spectrum*, chapter On the Effects of Image Alterations on Face Recognition Accuracy, pages 195–222. Springer International Publishing, 2016.

[13] M. Ferrara, A. Franco, and D. Maltoni. Face demorphing. *IEEE Transactions on Information Forensics and Security*, 13(4):1008–1017, 2017.

[14] M. Ferrara, A. Franco, and D. Maltoni. Face demorphing in the presence of facial appearance variations. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 2365–2369. IEEE, 2018.

[15] M. Ferrara, A. Franco, and D. Maltoni. Decoupling texture blending and shape warping in face morphing. In *2019 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–5. IEEE, 2019.

[16] M. Ferrara, A. Franco, and D. Maltoni. Face morphing detection in the presence of printing/scanning and heterogeneous image sources. *CoRR*, abs/1901.08811, 2019.

[17] FRONTEX. Best practice technical guidelines for automated border control ABC systems, 2015.

[18] J. Galbally, C. McCool, J. Fierrez, S. Marcel, and J. Ortega-Garcia. On the vulnerability of face verification systems to hill-climbing attacks. *Pattern Recognition*, 43(3):1027–1038, 2010.

[19] D. Gong, Z. Li, D. Lin, J. Liu, and X. Tang. Hidden factor analysis for age invariant face recognition. In *Proceedings of the ieee international conference on computer vision*, pages 2872–2879, 2013.

[20] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 30107-3. Information Technology - Biometric presentation attack detection - Part 3: Testing and Reporting*. International Organization for Standardization, 2017.

[21] A. K. Jain, P. Flynn, and A. A. Ross. *Handbook of biometrics*. Springer Science & Business Media, 2007.

[22] A. K. Jain, B. Klare, and U. Park. Face recognition: Some challenges in forensics. In *Face and Gesture 2011*, pages 726–733. IEEE, 2011.

[23] A. K. Jain and S. Z. Li. *Handbook of face recognition*. Springer, 2011.

[24] A. K. Jain, K. Nandakumar, and A. Ross. 50 years of biometric research: Accomplishments, challenges, and opportunities. *Pattern Recognition Letters*, 79:80–105, 2016.

[25] A. K. Jain, A. Ross, and U. Uludag. Biometric template security: Challenges and solutions. In *2005 13th European signal processing conference*, pages 1–4. IEEE, 2005.

[26] I. Kemelmacher-Shlizerman, S. M. Seitz, D. Miller, and E. Brossard. The megaface benchmark: 1 million faces for recognition at scale. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4873–4882, 2016.

[27] B. F. Klare, B. Klein, E. Taborsky, A. Blanton, J. Cheney, K. Allen, P. Grother, A. Mah, and A. K. Jain. Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1931–1939, 2015.

[28] Y. Li, G. Wang, L. Nie, Q. Wang, and W. Tan. Distance metric optimization driven convolutional neural network for age invariant face recognition. *Pattern Recognition*, 75:51–62, 2018.

[29] Z. Li, U. Park, and A. K. Jain. A discriminative model for age invariant face recognition. *IEEE transactions on information forensics and security*, 6(3):1028–1037, 2011.

[30] J. Lu, V. Erin Liong, and J. Zhou. Simultaneous local binary feature learning and encoding for face recognition. In *Proceedings of the IEEE international conference on computer vision*, pages 3721–3729, 2015.

[31] J. Lu, V. E. Liong, X. Zhou, and J. Zhou. Learning compact binary face descriptor for face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(10):2041–2056, 2015.

[32] S. Marcel, M. S. Nixon, and S. Z. Li. *Handbook of biometric anti-spoofing*, volume 1. Springer, 2014.

[33] G. Panis, A. Lanitis, N. Tsapatsoulis, and T. F. Cootes. Overview of research on facial ageing using the fg-net ageing database. *Iet Biometrics*, 5(2):37–46, 2016.

[34] U. Park, Y. Tong, and A. K. Jain. Age-invariant face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 32(5):947–954, 2010.

[35] A. Punnappurath, A. N. Rajagopalan, S. Taheri, R. Chellappa, and G. Seetharaman. Face recognition across non-uniform motion blur, illumination, and pose. *IEEE Transactions on image processing*, 24(7):2067–2082, 2015.

[36] R. Raghavendra, K. B. Raja, and C. Busch. Detecting Morphed Face Images. In *8th IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS)*, pages 1–8, 2016.

[37] R. Raghavendra, K. B. Raja, S. Venkatesh, and C. Busch. Transferable deep-cnn features for detecting digital and print-scanned morphed face images. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1822–1830. IEEE, 2017.

[38] R. Raghavendra, S. Venkatesh, K. Raja, and C. Busch. Detecting face morphing attacks with collaborative representation of steerable features. In *IAPR International Conference on Computer Vision & Image Processing (CVIP-2018)*, pages 1–7, 2018.

[39] R. Ranjan, V. M. Patel, and R. Chellappa. Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(1):121–135, 2017.

[40] U. Scherhag, A. Nautsch, C. Rathgeb, M. Gomez-Barrero, R. N. Veldhuis, L. Spreeuwers, M. Schils, D. Maltoni, P. Grother, S. Marcel, et al. Biometric systems under morphing attacks: Assessment of morphing techniques and vulnerability reporting. In *2017 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–7. IEEE, 2017.

[41] U. Scherhag, R. Raghavendra, K. Raja, M. Gomez-Barrero, C. Rathgeb, and C. Busch. On the vulnerability of face recognition systems towards morphed face attack. In *International Workshop on Biometrics and Forensics (IWBF 2017)*, pages 1–6, 2017.

[42] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch. Face recognition systems under morphing attacks: A survey. *IEEE Access*, 7:23012–23026, 2019.

[43] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch. Face recognition systems under morphing attacks: A survey. *IEEE Access*, 7:23012–23026, 2019.

[44] C. Seibold, W. Samek, A. Hilsmann, and P. Eisert. Detection of face morphing attacks by deep learning. In C. Kraetzer, Y.-Q. Shi, J. Dittmann, and H. J. Kim, editors, *Digital Forensics and Watermarking*, pages 107–120. Springer International Publishing, 2017.

[45] R. B. Shaw Jr, E. B. Katzel, P. F. Koltz, M. J. Yaremchuk, J. A. Girotto, D. M. Kahn, and H. N. Langstein. Aging of the facial skeleton: aesthetic implications and rejuvenation strategies. *Plastic and reconstructive surgery*, 127(1):374–383, 2011.

[46] L. Spreeuwers, M. Schils, and R. Veldhuis. Towards robust evaluation of face morphing detection. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 1027–1031, Sep. 2018.

[47] S. Venkatesh, R. Raghavendra, K. Raja, L. Spreeuwers, R. Veldhuis, and C. Busch. Morphed face detection based on deep color residual noise. In *ninth International Conference on Image Processing Theory, Tools and Applications (IPTA 2019)*, pages 1–5. IEEE, 2019.

[48] T. Y. Wang and A. Kumar. Recognizing human faces under disguise and makeup. In *2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*, pages 1–7. IEEE, 2016.