# Chapter 6
# Morph Creation and Vulnerability of Face Recognition Systems to Morphing


**Matteo Ferrara and Annalisa Franco**

**Abstract** Face recognition in controlled environments is nowadays considered rather reliable, and very good accuracy levels can be achieved by state-of-the-art systems in controlled scenarios. However, even under these desirable conditions, digital image alterations can severely affect the recognition performance. In particular, several studies show that automatic face recognition systems are very sensitive to the so-called face morphing attack, where face images of two individuals are mixed to produce a new face image containing facial features of both subjects. Face morphing represents nowadays a big security threat particularly in the context of electronic identity documents because it can be successfully exploited for criminal intents, for instance to fool Automated Border Control (ABC) systems thus overcoming security controls at the borders. This chapter will describe the face morphing process, in an overview ranging from the traditional techniques based on geometry warping and texture blending to the most recent and innovative approaches based on deep neural networks. Moreover, the sensitivity of state-of-the-art face recognition algorithms to the face morphing attack will be assessed using morphed images of different quality generated using various morphing methods to identify possible factors influencing the probability of success of the attack.

## 6.1 Introduction

Face morphing is generally described as a seamless transition transforming a facial image into another. Morphing was initially proposed as an image generation technique for computer graphics applications [1] or psychological studies [2, 3]. However, only in recent years it has emerged as a potential and severe security thread for Face Recognition Systems (FRS). The main risk deriving from face morphing is especially related to the adoption of automatic face-based identity verification in various

M. Ferrara (✉) · A. Franco
Department of Computer Science and Engineering, University of Bologna, via dell'Università, 50, Cesena, Italy
e-mail: matteo.ferrara@unibo.it



C. Rathgeb et al. (eds.), *Handbook of Digital Face Manipulation and Detection*,
Advances in Computer Vision and Pattern Recognition,
https://doi.org/10.1007/978-3-030-87664-7_6

applications like civilian identity management, Machine Readable Travel Documents (eMRTD), or visa management. A possible attack in relation to the use of MRTD in Automated Border Control (ABC) gates has been firstly identified in [4] and later confirmed by several research works. Identity verification at an ABC relies on the comparison of a live captured probe face image with a digital face image stored in an eMRTD such as an e-passport. If a morphed image, which is similar enough to the face of the two parent subjects, can be included in an eMRTD, then two persons can share the document. In this scenario, a criminal could exploit the passport of an accomplice with no criminal records to overcome the security controls. In more details, the subject with no criminal records (i.e., the accomplice) could apply for an eMRTD by presenting the morphed face photo; if the image is not noticeably different from his/her face, the police officer accepts the photo and releases the document (see Fig. 6.1).

The attack will be successful if the morphed image contemporarily meets two conditions.

- It is able to fool the human expert, i.e., the morphed face must be very similar to the accomplice who applies for the document and no elements (e.g., morphing artifacts) of the image should raise suspicions;
- the image fools at the same time the FRS used for automatic identity verification, meaning that the morphed face can be successfully matched with both subjects (criminal and accomplice).

Some studies confirm that morphed faces can be very realistic and able to fool human experts [5–7]. It is well known, in fact, that unfamiliar face recognition is a

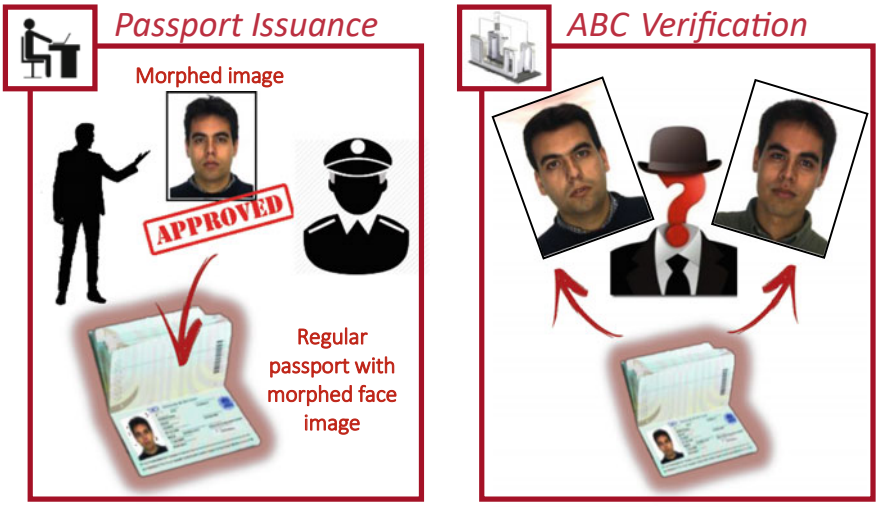


**Fig. 6.1** The face morphing attack in the eMRTD scenario. The morphed ID photo delivered to the officer is very similar to the applicant, but also contains facial features of a different subject

hard task for humans and it becomes even harder when it has to be accomplished based on a small-size id photo such as the one used by the citizens to apply for an identity or travel document. This photo is generally obtained by printing a high-quality digital image on photographic paper (typical size is 3.5 cm × 4.5 cm) and is then scanned to be included into the document. This printing and scanning process (P&S) hides many small details of the image (e.g., artifacts introduced by the morphing process) thus making it more difficult for human examiners to spot the attack attempt.

Figure 6.2 shows two examples of morphing. In the first case (top row), the morphed image (b) is obtained with an almost equal contribution of the two subjects (a) and (c); the result is quite similar to subject (a) but a human expert could notice some differences. In the second example, the morphed image (e) has been generated from (d) and (f), but with a stronger contribution of subject (d). Visually the morphed image is almost indistinguishable from the accomplice (d) and is very unlikely that it would raise some suspicion by the officer. Both these morphed images, (b) and (d), contain enough information of the "criminal" subject to fool commercial FRSs.

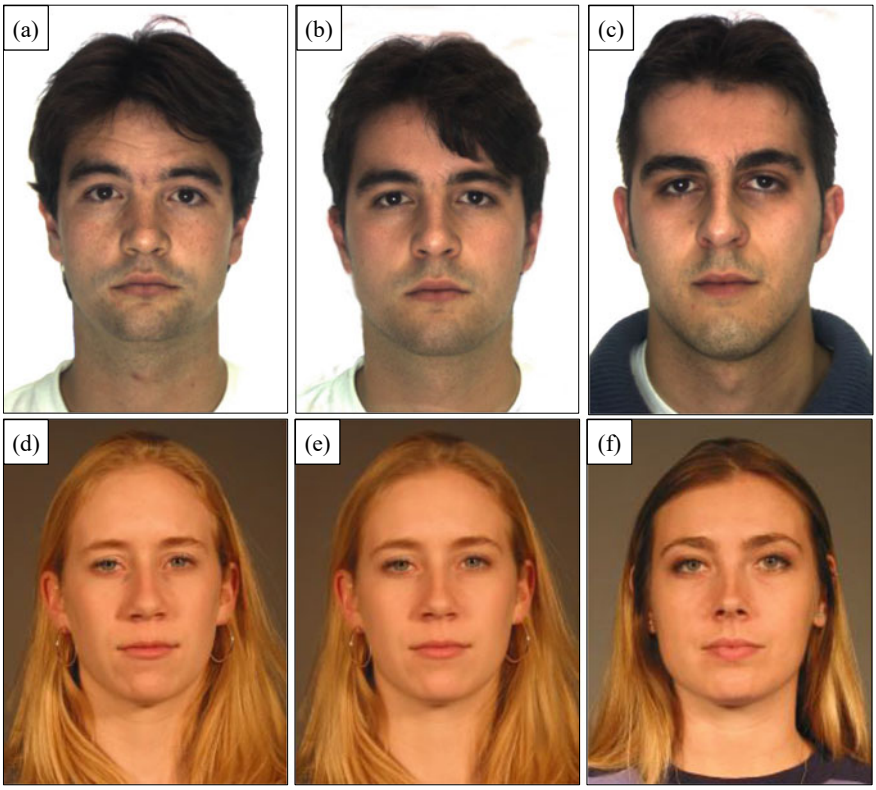


**Fig. 6.2** Two example of morphed images: **b** obtained from the subjects (**a**) and (**c**); **e** obtained from the subjects (**d**) and (**f**)

It is worth noting that in case of successful attack, the document issued is perfectly regular; the attack does not consist of altering the document content but in deceiving the officer while issuing the document. The document released will thus pass all the integrity checks (optical and electronic) performed at the gates.

This attack is made possible in practice by the procedure adopted in several countries where there is no live enrolment for facial images and citizens apply for the document by providing an ID photo printed on photographic paper. The trust chain is thus broken since citizens could intentionally alter the image content by different possible digital image manipulations [5], even with criminal intents. Switching to live enrolment would certainly be the most effective solution, but its adoption by all the involved countries is very unlikely; moreover, we have to consider the huge number of documents already issued since the introduction of eMRTDs, which still represent a potential risk. In fact, governmental agencies already reported a few real morphing attack attempts and recent news confirm that the criticalities related to the morphing attack have reached a wide public audience [8–10]. Estimating the real extent of this phenomenon is hard, due to the practical impossibility of spotting the cases of successful attack. Unfortunately, the analysis of the vulnerability of FRSs to morphing attack, discussed later in this chapter, is not encouraging and confirm once again that designing effective countermeasures is quite urgent.

This chapter is organized as follows. Section 6.2 describes the face morphing generation algorithms, presenting both traditional landmark-based approaches, as well as innovative solutions based on deep learning. Section 6.3 analyzes and discusses the vulnerability of commercial FRSs to morphing attack; finally, Sect. 6.4 draws some concluding remarks.

## 6.2 Face Morphing Generation

Nowadays, the generation of a morphed image has become quite an easy and inexpensive task. Open-source solutions are publicly available, such as for instance general image processing software with specific plugins (e.g., the GAP plugin for GIMP [11]). Moreover, a number of free or commercial tools (e.g., FaceMorpher [12] or FantaMorph [13]), as well as applications for mobile devices or online services are available. Interested readers can refer to [14] for a comprehensive review of publicly available morphing tools. It is however worth noting that the images obtained with these fully automated systems are usually affected by the presence (more or less accentuate) of clearly visible artifacts that would probably cause a rejection of the image by the human officer during the document issuing process. As discussed later in this chapter, the creation of a high quality and credible morphed image usually requires an accurate manual intervention aimed at removing the most relevant defects and make the image undistinguishable from a bona fide one.

### 6.2.1 Landmark Based Morphing

Landmark-based approaches for face morphing allow synthesizing a fluid and gradual transformation from one image to another by exploiting facial landmark points in the involved images. Reference points usually correspond to prominent facial components such as mouth, nose, eyes or eyebrows, and approximately outline their shape. Such reference points can be either manually annotated or automatically determined using facial landmark detection algorithms such as Dlib [15], which is the most widely used for this purpose. Of course, the effort needed in the two cases is different, and manual annotation is a boring and time-consuming task; on the other hand, if properly executed, manual landmark labeling usually provides more precise landmark locations and achieves a better image coverage. Automatic landmark detection algorithms, in fact, usually adopt standard facial models that consider the central part of the face and the chin but ignore for instance the forehead region. As we will discuss later, the accuracy of landmark detection has a direct impact on the quality and effectiveness of the generated morphed images.

Starting from the facial landmarks, the morphing process can be generally described as follows. Let $I_0$ and $I_1$ be the two parent images to morph and let $P_0$ and $P_1$ be the two sets of correspondence points in $I_0$ and $I_1$, respectively. For most of the landmark-based approaches, the transformation between the two images is ruled by the so-called morphing factor, a parameter $\alpha$ representing a weighting factor for the two images. The morphing process is therefore generating a set of intermediate frames $\mathbb{M} = \{I_\alpha, \alpha \in \mathbb{R}, 0 < \alpha < 1\}$ representing the transformation of the first image ($I_0$) into the second one ($I_1$) as shown in Fig. 6.3. Note that, to obtain realistic results, the two images need to be aligned in advance (e.g., by overlaying the eye centers).

In general, each frame is a weighted linear combination of $I_0$ and $I_1$ (based on $\alpha$ value), obtained by combining (i) *geometric warping* [16] of the two images based on correspondence points and (ii) *texture blending*.

Formally:

$$I_\alpha(\mathbf{p}) = (1-\alpha) \cdot I_0\big(w_{P_\alpha \to P_0}(\mathbf{p})\big) + \alpha \cdot I_1\big(w_{P_\alpha \to P_1}(\mathbf{p})\big), \tag{6.1}$$

where

- $\mathbf{p}$ is a generic pixel position;
- $\alpha$ is the weight factor, representing the contribution of image $I_1$ to the morphing ($\alpha = 0.3$ indicates that the morphed image will be obtained for the 30% from $I_1$ and 70% from $I_0$);
- $P_\alpha$ is the set of correspondence points aligned according to the weight factor $\alpha$;
- $w_{P_B \to P_A}(\mathbf{p})$ is a warping function.

Several warping techniques have been proposed in the literature [17]. A common approach consists in representing the two sets of points ($P_A$ and $P_B$) by means of topologically equivalent (i.e., no folding or discontinuities are permitted) triangular

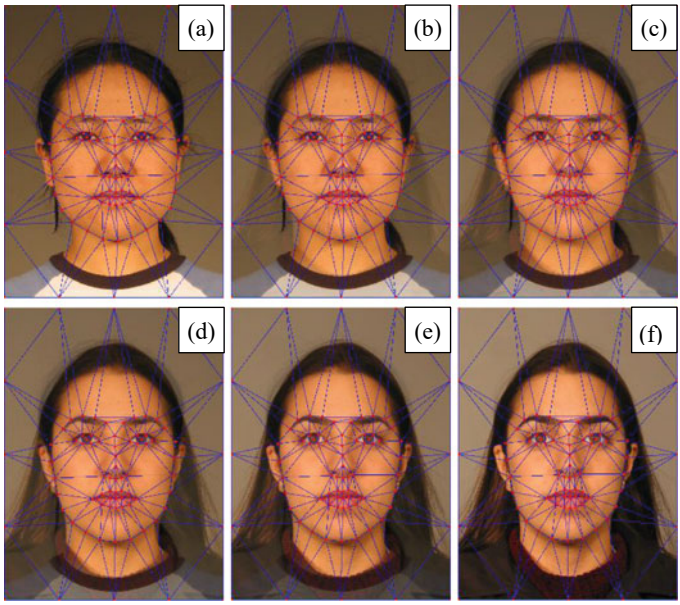


**Fig. 6.3** Morphing of image $I_0$ (**a**) to $I_1$ (**f**). **b**, **c**, **d** and **e**) are intermediate frames, obtained by the morphing procedure, gradually moving from $I_0$ to $I_1$. The correspondence points and the triangular meshes are highlighted in red and blue, respectively.

meshes (see Fig. 6.3) and computing local spatial transformations that map each warped triangle to the corresponding original one [18]. Note that the meshes are constrained to cover the whole images and not to cause self-intersection (i.e., each pixel position is contained in exactly one mesh). A triangular mesh can be derived from a set of points via Delaunay triangulation [19]. Given a generic pixel position **p** in the warped image, the transformation used to map **p** onto the original image $I$ is the local transformation corresponding to the warped triangle that contains **p** (see Fig. 6.4).

The set of aligned correspondence points $P_\alpha$ in Eq. (6.1) is computed as follows (see Fig. 6.5):

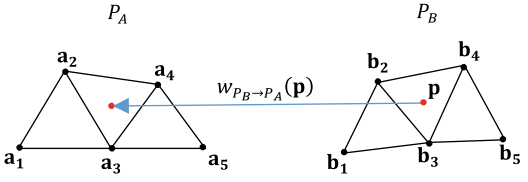


**Fig. 6.4** Example of image warping using triangular meshes. The point **p** in the warped image is mapped into the original image using the inverse mapping of triangle $\Delta\, \mathbf{b_2 b_3 b_4}$ into $\Delta\, \mathbf{a_2 a_3 a_4}$

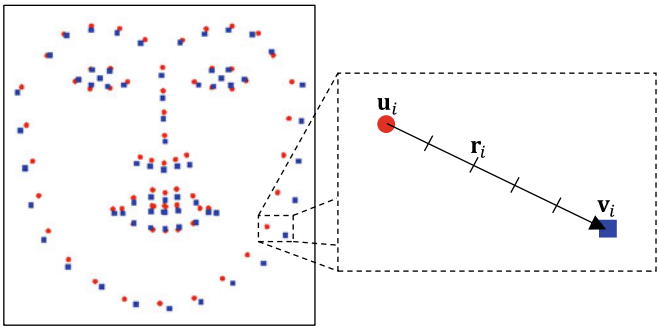


**Fig. 6.5** On the left, $P_0$ (red circles) and $P_1$ (blue squares) are the corresponding points of images in Fig. 6.3a and f, respectively. On the right, the region containing points $\mathbf{u}_i$ and $\mathbf{v}_i$ is zoomed to show point $\mathbf{r}_i$ corresponding to morphed frame $I_{0.4}$ (see Eq. (6.2) and Fig. 6.3c)

$$P_\alpha = \{\mathbf{r}_i | \mathbf{r}_i = (1-\alpha)\cdot\mathbf{u}_i + \alpha\cdot\mathbf{v}_i, \mathbf{u}_i \in P_0, \mathbf{v}_i \in P_1\}. \tag{6.2}$$

A more general formulation of the morphing process has been proposed in [20]; here geometric warping and image blending are ruled by two different factors. Equation (6.1) can be generalized as follows:

$$I_{\alpha_B,\alpha_W}(\mathbf{p}) = (1-\alpha_B)\cdot I_0\big(w_{P_{\alpha_W}\to P_0}(\mathbf{p})\big) + \alpha_B \cdot I_1\big(w_{P_{\alpha_W}\to P_1}(\mathbf{p})\big), \tag{6.3}$$

where $\alpha_B$ and $\alpha_W$ are the blending and warping factors, respectively.

The effects of blending and warping are shown in Fig. 6.7 where two very different subjects have been selected (see Fig. 6.6) to highlight the influence of $\alpha_B$ and $\alpha_W$. From a visual point of view, the result from different combinations is overall quite similar, but the effects produced on the probability of success of the attack by the possibility of acting separately on geometry warping and image blending have to be carefully considered. Several studies in fact show that, in the context of face recognition, humans are more sensitive to texture than to geometry [21]; the study [20] reveals that the same holds for FRSs, as confirmed by the experimental results reported in Sect. 3.2. Assigning different weighting factors to texture blending and geometry warping during the face morphing process significantly increases the chances of success, especially in the presence of look-alike subjects.

The automatic generation of morphed images can produce some visible artifacts that might be easily spotted by a human observer, thus drastically reducing the probability of success of the face morphing attack. The adoption of automatically detected facial landmarks, further increase the probability of artifacts in case of inaccurate point identification. The following visible artifacts are generally detectable:


**Fig. 6.6** Images $I_0$ and $I_1$ used to generate the morphed images in Fig. 6.7

- Macroscopic ghost artifacts in the face surrounding area (see Fig. 6.8a). Facial landmarks are usually exclusively located in the facial region, and no reference points are considered for hairs, ears, and ecc. No accurate warping is therefore carried out for these regions, and the blending process produces therefore visible artifacts due to different characteristics (e.g., hair style or background) of the two contributing images.
- Minor artifacts close to the facial reference points (eyes, eye brows, mouth, nose, chin, and nostrils) mainly due to insufficient or inaccurate landmarks. Typical patterns are double edges or double reflections on irises (see Fig. 6.9a).

A widely used solution to remove the macroscopic artifacts in the face surrounding area is background substitution; the background region is typically replaced by the corresponding region of one of the parent images (the one with the highest blending factor), after a proper alignment (see Fig. 6.8b). An additional step is recommended in this case, aimed at equalizing the skin color before background substitution. In fact, due to different illumination conditions or skin color between the two face images, the retouching result could be unsatisfactory, in particular when the facial landmarks do not include the forehead region, thus causing a strong edge with the central face region. To overcome this issue, the histogram matching method described in [22] could be applied.

The second category of artifacts is more difficult to address, and no effective automatic solutions have been identified so far. At present, only a very careful manual post-processing is able to remove them, with a combination of low-level image processing operations such small region cloning from the contributing images, direct painting or edge smoothing (see Fig. 6.9b). Of course, this manual intervention is not trivial and requires some practice to achieve a good result. However, manual post-processing is a key element for the success of the morphing attack, in particular to fool human experts, which could quite easily spot morphing artifacts if not carefully removed.

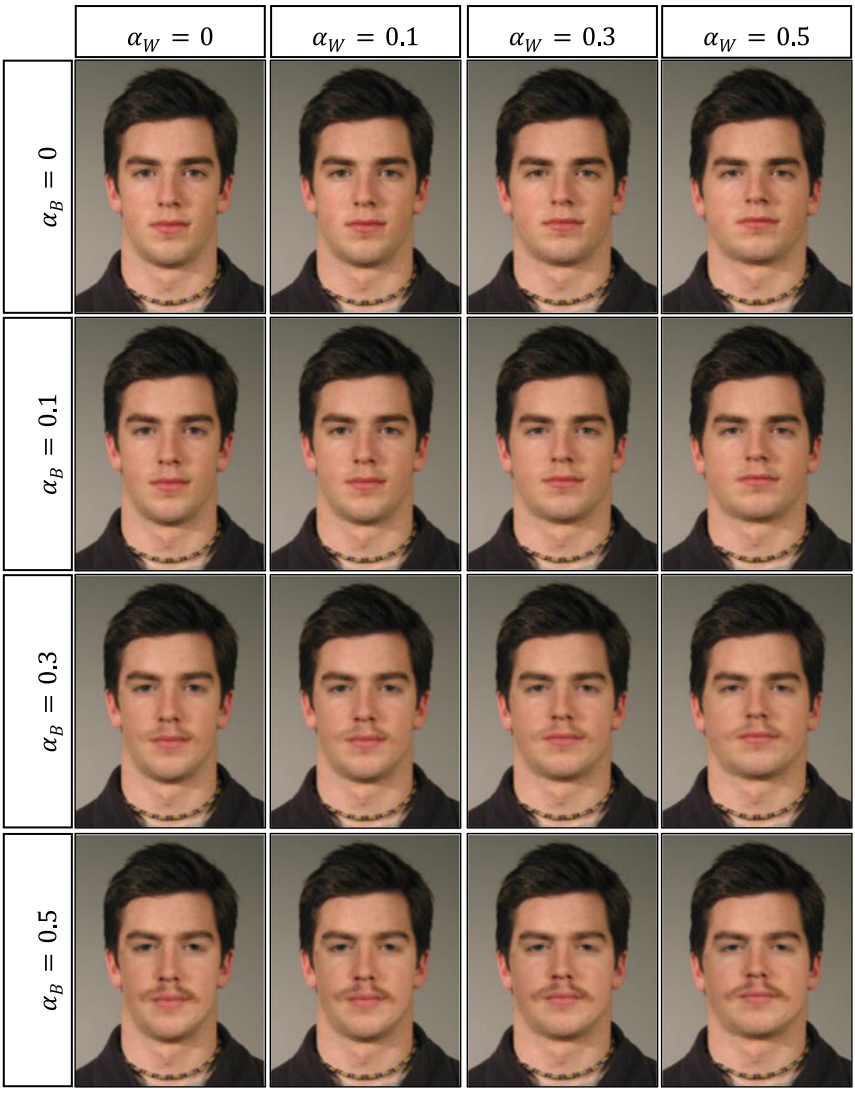


**Fig. 6.7** Morphed images obtained with different blending and warping factors by combining Fig. 6.6a ($I_0$) and Fig. 6.6b ($I_1$)

### *6.2.2 Deep Learning-Based Face Morph Generation*

The face morphing approaches presented in the previous section provide a precise control on the morphing process in relation for instance to the contribution of the two subjects in the resulting image. On the other hand, since the process relies on facial landmarks, an inaccurate detection of such reference points, as well as the lack


**Fig. 6.8** Morphed image obtained from the two subjects in Fig. 6.3 with macroscopic artifacts in the region around face; **b** morphed image in (**a**) after automatic background substitution


**Fig. 6.9** **a** Small artifacts in the eye region, with double edge effect and multiple light reflections in the iris; **b** eye region after manual post-processing for artifact removal

of reference points in specific face regions, determine in most cases the presence of some ghost artifacts in the morphed image, which a human expert observing the image could spot quite easily. As mentioned above, the realization of an "ideal" morphed image requires a difficult and time-consuming manual post-processing aimed at removing all visible artifacts. To overcome this limitation, some innovative solutions for face morphing generation have been recently proposed, with the aim of fully automating the generation process. In particular, a few recent works in the literature exploit the potential of Generative Adversarial Networks (GAN) to synthesize

morphed images by sampling the two contributing facial images in the latent space, without requiring preliminary landmark extraction and alignment.

GANs are based on the combined action of two different agents, a generator and discriminator. The first one, the generator $G$, produces samples from a distribution which should be ideally indistinguishable from the training distribution. The discriminator $D$ is trained to determine if the incoming samples are drawn from the real set of training images or are fake samples generated by $G$. The training process gradually improves the samples produced by the generator $G$, which learns the most effective way to fool the discriminator.

The first approach for GAN-based face morphing generation, called MorGAN, was proposed in [23]. The network architecture is inspired by the work [24] where the Bidirectional Generative Adversarial Network (BiGAN) is introduced. In addition to the generator $G$ from the standard GAN framework BiGAN includes an encoder $E$ which maps data $\mathbf{x}$ to latent representations $\mathbf{z}$. The BiGAN discriminator $D$ discriminates not only in data space ($\mathbf{x}$ versus $G(\mathbf{z})$), but jointly in data and latent space (tuples $(\mathbf{x}, E(\mathbf{x}))$ versus $(G(\mathbf{z}), \mathbf{z})$, where the latent component is either an encoder output $E(\mathbf{x})$ or a generator input $\mathbf{z}$. The idea is that the BiGAN encoder $E$ should learn to invert the generator $G$, even if the two modules cannot directly "communicate". This architecture is adapted by the authors of [23] to the problem of face morph generation. The generator is split into two components, complementary inverse to each other, and the discriminator is trained to distinguish between joint pairs (samples from the encoder and samples from the decoder). The main limitation of the MorGAN approach is the limited size of the generated morphed images, 64 $\times$ 64 pixels, which is quite far from the resolution needed to fulfill the ISO/ICAO quality standards (minimum inter-eye distance of 90 pixels) and to successfully fool commercial FRSs. This last aspect is confirmed in [25] where the authors evaluate the vulnerability of state-of-the-art face recognition systems to MorGAN morphed images.

The same work [25] focuses on the generation of high-quality morphed images, with the aim of overcoming the key limitation of the MorGAN approach. In particular, the authors propose the adoption of StyleGAN [26] for morphing generation. Given the latent code $L_1$ of the face, StyleGAN maps the inputs to an intermediate latent space through the mapping network. The mapping layer consists of 8 fully connected layers serially connected. The approach synthesizes a data-subject-specific morphed face by forcing a strategy to embed the face image into the latent space. The subject-specific embedded latent space passes through the synthesis network consisting of 18 layers, thus obtaining a representation in 18 latent spaces (dimension 512) which is further concatenated. The loss function driving the embedding measures the similarity between the input image and the reconstructed image. The images of the two contributing subjects are both processed according to the procedure described above and a weighted average (to recall the idea of morphing factor) of the corresponding latent codes is computed to obtain the morphed image latent code, which is finally passed through the synthesis network to generate the high-resolution morphed image (1024 $\times$ 1024).

The morphing approach based on StyleGAN has been successively improved by the same authors in [27] where the MIPGAN (Morphing through Identity Prior driven GAN) approach is presented. The introduction of a loss function aimed at preserving the identity of the generated morphed image, through enforced identity priors represents the main element of novelty. Given the images of the two contributing subjects, the corresponding latent vectors are first computed using a latent prediction network. The morphed image latent vector is again obtained by a weighted average of the two input vectors and is finally passed through the synthesis network to obtain a morphed image of size $1024 \times 1024$. The last step consists of a final optimization stage based on the identity preserving loss function. The authors propose two different versions of MIPGAN, obtained using two versions of StyleGAN, [26] and [28], respectively. The MIPGAN approach achieves interesting results in terms of efficacy of the attack, as shown by the results reported in the next section.

Besides image resolution, another important aspect to consider is the similarity of the morphed image to the two contributing subjects. From this point of view, the landmark-based approaches certainly allow to better preserve the identity of the two contributing subjects and to control quite easily (via the morphing factor) the similarity of the resulting morphed images to one of the two individuals. GAN-based approaches seem to have less control on this aspect, even when an identity preserving loss function is adopted. Even if the morphed images generated using GAN-based approaches can fool automatic FRSs, we believe that further work is needed to make the generated images able to fool the human expert.

## 6.3 Vulnerability of Face Recognition Systems to Face Morphing

In this section, we describe the experiments carried out using three commercial face recognition SDKs (referred to as $SDK_1$, $SDK_2$, and $SDK_3$) which provided top performance in the "Face Recognition Vendor Test (FRVT)—1:1 Verification" [29, 30]; the names of the SDKs cannot be disclosed, and the results will be therefore presented in anonymous form.

In order to simulate a realistic attack to an ABC system, the operational threshold of the face recognition software have been fixed according to the Frontex guidelines [31]. In particular, for ABC systems operating in verification mode, the face recognition algorithm has to ensure a *False Acceptance Rate* (FAR) equal to 0.1% and a *False Rejection Rate* (FRR) lower than 5%. During the experimentation, for each SDK, the security threshold indicated in the corresponding documentation to achieve FAR $= 0.1\%$ has been used. Since we focus on morphing attacks, the performance is evaluated in terms of Mated Morph Presentation Match Rate (MMPMR) [32] with the aim to quantify the percentage of morphing attacks able to fool the SDKs. To this purpose the MMPMR for all SDKs have been measured by comparing morphed

face images against probe images of both subjects involved in the generation of the morphed image.

### *6.3.1 Data Sets*

The SDKs have been evaluated on five data sets:

- BIOLAB-1.0 [5]: it contains 80 morphed images generated using the GIMP software [11, 33] after a manual labeling of the facial reference points and a first manual alignment based on eyes superimposition; a final manual retouch was carried out to remove visible artifacts. For each morphed image, it contains two probe images, one for each parent subject.
- MorphDB [34]: the aim of this dataset is to reproduce the typical scenario where the ID photo is provided by the citizens printed on photographic paper and then scanned by the officer during the issuing process. It contains 100 morphed images generated using the Sqirlz Morph 2.1 software [35] with facial landmarks automatically detected and a morphing factor in the range [0.3;0.4]. After the generation, the morphed images have been manually retouched to remove visible artifacts introduced by the morphing procedure. The P&S images have been created by printing the digital version on high quality photographic paper by a professional photographer and scanned at 300 DPI. For each morphed image, it includes a variable number of probe images of the two parent subjects.
- SOTAMD [36]: it contains 5748 high quality images for benchmarking under realistic conditions. The dataset consists of facial images from subjects of various ethnicities, age-groups, and both genders. After a careful subject pre-selection, the morphed images have been created using seven different morphing algorithms and applying manual post-processing to remove visible artifacts. Moreover, the images have been also printed and scanned. For each morphed image, it includes 10 probe images, for each contributing subject, captured under a simulated ABC gate operational scenario presenting more variations with respect to other datasets.
- AMSL [37]: a dataset containing images from the Face Research Lab London Set [38]. 2175 morphed face images were generated using the morphing approach described in [39]. All images were modified in the way to comply with the requirements of the ICAO portrait quality standard for eMRTD [40] and to fit on a chip of an eMRTD including cropping, down-scaling, and JPEG compression. For each morphed image, it contains two probe images, one for each subject.
- B&W [20]: a dataset containing morphed images automatically generated by separately varying the blending and the warping factors $\alpha_B$ and $\alpha_W$ to evaluate their importance in fooling face recognition systems. It contains 560 morphed images for each combination of $\alpha_B$ and $\alpha_W$ and for each of them, a probe image for each contributing subject.

BIOLAB-1.0, MorphDB, and SOTAMD datasets are available for testing on the Bologna Online Evaluation Platform (BOEP) [41] hosted in the FVC-onGoing framework [42, 43].

### 6.3.2 Results

Table 6.1 reports the single MMPMR of the three SDKs and their average on all datasets (except for B&W data set whose results are reported below).

For all SDKs, the most difficult datasets seem to be both BIOLAB-1.0 and AMSL with an average MMPMR of 95.0 and 92.7%, respectively. This is probably due to a combination of different elements:

- *morphingfactor*—both BIOLAB-1.0 and AMSL datasets contain symmetric morphed images (i.e., morphing factor equal to 0.5) while MorphDB dataset contains asymmetric morphed images generated with a morphing factor in the range [0.3;0.4] and SOTAMD dataset contains morphed images generated with two different morphing factors (0.3 and 0.5);
- *facial landmarks manually labeled*—to generate BIOLAB-1.0 morphed images, the facial landmarks have been manually selected, while automatically detected facial landmarks have been used to generate MorphDB and SOTAMD morphed images;
- *forehead landmarks*—BIOLAB-1.0 morphed images have been generated using also landmarks manually labeled on the hairline (see Fig. 11 in [5]) which have not been used to generate the other databases;
- *facial outer region substitution*—as shown in Fig. 6.3, the intermediate morphed frames could present double exposure effects outside the facial region (e.g., background, hair, shoulders, and body). To make morphed images more realistic and therefore more difficult to be detected, usually a retouching is applied. MorphDB and SOTAMD morphed images have been automatically retouched by replacing

**Table 6.1** MMPMR of the three SDKs on different data sets

| Database | Format | Morphed images | Probe images per parent subject | $SDK_1$ (%) | $SDK_2$ (%) | $SDK_3$ (%) | AVG (%) |
|---|---|---|---|---|---|---|---|
| BIOLAB-1.0 | Digital | 80 | 1 | 98.75 | 96.25 | 90.00 | 95.00 |
| MorphDB | Digital | 100 | Variable | 78.00 | 60.00 | 50.00 | 62.67 |
| | P&S | | | 74.00 | 59.00 | 50.00 | 61.00 |
| SOTAMD | Digital | 2045 | 10 | 69.10 | 50.81 | 46.41 | 55.44 |
| | P&S | 3703 | | 69.89 | 42.07 | 44.64 | 52.20 |
| AMSL | Digital | 2175 | 1 | 99.08 | 94.25 | 84.78 | 92.70 |

the pixels outside the face region with those of the accomplice image, while BIOLAB-1.0 morphed images have been manually retouched.

- *probe images*—to simulate an ABC gate operational scenario, in the SOTAMD database, the morphed images are compared against face images acquired using ABC gates. Such images present different lighting conditions, and some of them have been acquired as grayscale images. Such differences could decrease the chance to fool the SDKs.

As the SOTAMD dataset [36] presents meta-data regarding the characteristics of the parent subjects used for morphing (e.g., gender) and of the morphing generation pipeline (e.g., morphing approach), the MMPMR of the three SDKs and their average on different subsets are reported in Tables 6.2 and 6.3 (digital and P&S versions, respectively).

Some interesting results can be observed, in relation to the main attributes characterizing the database images:

- *gender*—the chance of fooling SDKs for female subjects looks on average higher than for male subjects (about 10% better on both digital and P&S versions).
- *post-processing*—as expected manual retouching increases the probability of fooling the SDKs with respect to automatic post-processing, even if the difference is not so evident (about 5% better on both digital and P&S versions).
- *morphingalgorithm*—SDKs exhibit different behaviors as the morphing algorithm changes; algorithms C02 and C01 present a higher change to fool SDKs

**Table 6.2** MMPMR of the three SDKs on digital version of SOTAMD subsets

| Attribute | Subset | # Morphed images | $SDK_1$ (%) | $SDK_2$ (%) | $SDK_3$ (%) | AVG (%) |
|---|---|---|---|---|---|---|
| Gender | Female | 876 | 71.69 | 58.33 | 53.42 | 61.15 |
| | Male | 1169 | 67.15 | 45.17 | 41.15 | 51.15 |
| Post processing | Automatic | 1575 | 67.87 | 49.46 | 45.78 | 54.37 |
| | Manual | 470 | 73.19 | 55.32 | 48.51 | 59.01 |
| Morphing algorithm | C01 | 325 | 79.08 | 64.92 | 55.08 | 66.36 |
| | C02 | 200 | 91.00 | 82.00 | 62.00 | 78.33 |
| | C03 | 400 | 65.00 | 43.75 | 40.75 | 49.83 |
| | C05 | 420 | 67.38 | 46.90 | 45.24 | 53.17 |
| | C06 | 400 | 61.75 | 40.00 | 40.50 | 47.42 |
| | C07 | 300 | 61.33 | 44.00 | 43.67 | 49.67 |
| Morphing factor | 0.3 | 1035 | 47.54 | 25.89 | 22.32 | 31.92 |
| | 0.5 | 1010 | 91.19 | 76.34 | 71.09 | 79.54 |
| Morph quality | High | 1059 | 89.99 | 76.11 | 66.19 | 77.43 |
| | Low | 986 | 46.65 | 23.63 | 25.15 | 31.81 |

**Table 6.3** MMPMR of the three SDKs on P&S version of SOTAMD subsets

| Attribute | Subset | # Morphed images | $SDK_1$ (%) | $SDK_2$ (%) | $SDK_3$ (%) | AVG (%) |
|---|---|---|---|---|---|---|
| Gender | Female | 1661 | 71.76 | 49.91 | 50.87 | 57.52 |
| | Male | 2042 | 68.36 | 35.70 | 39.57 | 47.88 |
| Post processing | Automatic | 1453 | 66.83 | 37.72 | 45.42 | 49.99 |
| | Manual | 2250 | 71.87 | 44.89 | 44.13 | 53.63 |
| Morphing algorithm | C01 | 500 | 79.80 | 53.00 | 57.60 | 63.47 |
| | C02 | 500 | 95.00 | 79.00 | 57.40 | 77.13 |
| | C03 | 1264 | 60.21 | 28.64 | 34.97 | 41.27 |
| | C05 | 939 | 68.26 | 38.45 | 43.66 | 50.12 |
| | C06 | 500 | 62.40 | 35.00 | 45.20 | 47.53 |
| Morphing factor | 0.3 | 1853 | 49.00 | 23.48 | 22.23 | 31.57 |
| | 0.5 | 1850 | 90.81 | 60.70 | 67.08 | 72.86 |
| Morph quality | High | 1920 | 90.73 | 64.90 | 63.39 | 73.00 |
| | Low | 1783 | 47.45 | 17.50 | 24.45 | 29.80 |
| Image compression | Uncompressed | 380 | 82.37 | 67.89 | 51.32 | 67.19 |
| | Compressed | 3323 | 68.46 | 39.12 | 43.88 | 50.49 |

**Table 6.4** MMPMR of $SDK_1$ on B&W data set for each combination of $\alpha_B$ and $\alpha_W$. Different values are represented by different blue levels (the darker, the greater)

| $\alpha_B$ \ $\alpha_W$ | **0** | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** |
|---|---|---|---|---|---|---|
| **0** | 1.4% | 1.6% | 2.1% | 2.7% | 4.3% | 4.5% |
| **0.1** | 5.4% | 7.7% | 8.4% | 9.8% | 11.1% | 11.6% |
| **0.2** | 18.0% | 20.2% | 22.3% | 25.0% | 27.1% | 29.8% |
| **0.3** | 40.5% | 46.4% | 49.6% | 55.0% | 58.6% | 61.8% |
| **0.4** | 73.0% | 79.3% | 82.7% | 86.4% | 88.9% | 90.4% |
| **0.5** | 93.0% | 95.2% | 96.6% | 97.5% | 97.7% | 97.9% |

with respect to algorithms C06, C07, and C03. Please refer to [36] for a detailed description of the different morphing algorithms.

- *morphingfactor*—as expected symmetric morphing (morphing factor equals to 0.5) fools the SDKs more easily (more than 40% better on both digital and P&S versions) than asymmetric morphing (morphing factor equals to 0.3).
- *morph quality*—as expected high quality morphs are more difficult to detect than low quality morphs (about 45% better on both digital and P&S versions).
- *image compression*—the uncompressed images present a higher probability to fool SDKs with respect to the compressed version (about 15% better).

Tables 6.4, 6.5, 6.6, and 6.7 report the MMPMR of the three SDKs and their average on B&W data set. For all SDKs blending and warping present a very different impact on the probability of success of the attack, while geometric modifications obtained increasing the warping factor $\alpha_W$ do not heavily affect recognition accuracy (see ranges $\alpha_B \in [0; 0.1]$, $\alpha_W \in [0.4; 0.5]$), an opposite behavior is observed for the blending factor $\alpha_B$ ($\alpha_B \in [0.4; 0.5]$, $\alpha_W \in [0; 0.1]$). Hence, for a criminal it would be much more convenient to create a morphed image with $\alpha_B = 0.5$ and $\alpha_W \in [0; 0.2]$ instead of using a balanced morphing factor in the range [0.2; 0.3] as

**Table 6.5** MMPMR of $SDK_2$ on B&W data set for each combination of $\alpha_B$ and $\alpha_W$. Different values are represented by different blue levels (the darker, the greater)

| $\alpha_B$ \ $\alpha_W$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|
| **0** | 1.4% | 1.8% | 2.1% | 2.1% | 3.0% | 2.7% |
| **0.1** | 3.6% | 4.5% | 5.7% | 6.1% | 7.0% | 8.9% |
| **0.2** | 9.3% | 11.6% | 15.7% | 18.9% | 23.8% | 26.6% |
| **0.3** | 27.1% | 32.0% | 38.4% | 43.2% | 47.7% | 54.1% |
| **0.4** | 50.9% | 59.5% | 66.4% | 71.6% | 76.3% | 79.6% |
| **0.5** | 72.0% | 78.9% | 85.0% | 88.0% | 91.1% | 93.2% |

**Table 6.6** MMPMR of $SDK_3$ on B&W data set for each combination of $\alpha_B$ and $\alpha_W$. Different values are represented by different blue levels (the darker, the greater)

| $\alpha_B$ \ $\alpha_W$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|
| **0** | 0.5% | 0.9% | 1.3% | 1.1% | 1.6% | 2.1% |
| **0.1** | 2.5% | 3.0% | 3.4% | 4.6% | 5.5% | 7.0% |
| **0.2** | 7.5% | 9.8% | 11.1% | 13.2% | 15.5% | 18.6% |
| **0.3** | 21.4% | 23.9% | 28.4% | 33.0% | 38.0% | 42.5% |
| **0.4** | 44.6% | 51.3% | 56.4% | 61.1% | 66.1% | 69.3% |
| **0.5** | 70.4% | 75.5% | 81.4% | 85.9% | 89.3% | 91.6% |

**Table 6.7** Average MMPMR of the three SDKs on B&W data set for each combination of $\alpha_B$ and $\alpha_W$. Different values are represented by different blue levels (the darker, the greater). The green region represents the most promising combinations of blending and warping factors to successfully perpetrate the attack

| $\alpha_B$ \ $\alpha_W$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|
| **0** | 1.1% | 1.4% | 1.9% | 2.0% | 3.0% | 3.1% |
| **0.1** | 3.8% | 5.1% | 5.8% | 6.9% | 7.9% | 9.2% |
| **0.2** | 11.6% | 13.9% | 16.4% | 19.1% | 22.1% | 25.0% |
| **0.3** | 29.7% | 34.1% | 38.8% | 43.8% | 48.1% | 52.8% |
| **0.4** | 56.2% | 63.3% | 68.5% | 73.0% | 77.1% | 79.8% |
| **0.5** | 78.5% | 83.2% | 87.7% | 90.5% | 92.7% | 94.2% |


**Fig. 6.10** Example images from the database used for the experiment. The morphed images are obtained combining the two images $I_0$ and $I_1$ with different blending ($\alpha_B$) and warping ($\alpha_W$) factors

stated in [34, 44]. This choice would increase the chances of successful attack at the border (from about 16–44 to 78–88%, on the average) keeping unaltered the chances of fooling the human officer during the document issuing process. In fact, a visual inspection of several generated morphs reveals that the difference between the two images is imperceptible, in particular when look-alike subjects are involved (see the example of Fig. 6.10). Moreover, we should always consider that human recognition capabilities are surprisingly error-prone in front of unfamiliar faces [45] and small appearance variations would probably be neglected. Finally, it is important to note that the MMPMR values could be even higher because, in a real scenario, a criminal would try to produce high quality morphed images, discarding the morphs with a low probability of success and applying manual retouching to remove unrealistic artifacts.

### *6.3.3 Deep Learning-Based Morphing Results*

Currently no databases of morphed images generated by GANs are publicly available; therefore, the vulnerability assessment we did only focus on images generated by landmark-based approaches. However, as a reference, we think it is worth reporting the preliminary results reported by the authors of the GAN-based approaches in their paper [27].

**Table 6.8** MMPMR of a face recognition system on morphed images generated by GANs as reported in [27]

| Format | Morph generation type | | | |
|---|---|---|---|---|
| | Facial landmark [46] (%) | StyleGAN | MIPGAN-I (%) | MIPGAN-II (%) |
| Digital | 100 | 64.68 | 94.36 | 92.93 |
| P&S | 97.64 | 61.72 | 92.97 | 80.56 |
| P&S with compression | 97.84 | 58.92 | 92.29 | 90.24 |

Table 6.8 compares the MMPMR of a state-of-the-art FRS on morphed images generated by (i) GANs and (ii) a landmark-based morphing method [46]. While StyleGAN generates morphed images with a low chance to fool the FRS (about 60%), the MIPGAN approach achieves interesting results in terms of efficacy of the attack (about 90%) even if lower than the facial landmark method (about 98%).

On the other hand, even if MIPGAN seems able to fool a FRS, some further efforts are necessary to improve the similarity with the contributing subjects thus increasing the effectiveness of the attack against human experts.

## 6.4 Conclusions

The general trust on automatic face recognition systems has recently been undermined by several possible kinds of attack, among which the face morphing is one of the most insidious and difficult to address. Dealing with face morphing is particularly complex in the context of ePassports; FRS are requested to work at fixed operational thresholds that guarantee a good trade-off between security and convenience in the use of ABC gates. Unfortunately, at these thresholds, it is very hard for FRSs to reject morphed images, thus making them quite vulnerable to the face morphing attack. This is particularly true when the morphed facial image is accurately prepared, with a manual intervention for facial landmark selection and artifact removal. Studies in the literature show that humans are easily fooled by accurate morphed images. Moreover, the high success rate measured in this chapter for landmark-based morphing techniques and the preliminary results reported in research papers for the GAN-based approaches confirm that face morphing is a real security threat. Recently, several research groups working on face recognition devoted significant efforts in designing face morphing attack detection techniques but, as discussed in a later chapter, further improvements are still needed to achieve good generalization capabilities.

## References


1. Beier T (1992) Feature-based image metamorphosis. Comput Graph 26:35–42
2. Steyvers M (1999) Morphing techniques for manipulating face images. Behav Res Meth Instrum Comput 359–369
3. Jäger T, Seiler KH, Mecklinger A (2005) Picture database of morphed faces (MoFa): technical report. Experimental neuropsychology unit. Department of Psychology, Saarland University, Saarbrücken, Germany
4. Ferrara M, Franco A, Maltoni D (2014) The magic passport. In: International joint conference on biometrics, clearwater (FL), pp 1–7
5. Ferrara M, Franco A, Maltoni D (2016) On the effects of image alterations on face recognition accuracy. In: Face recognition across the imaging spectrum, pp 195–222
6. Robertson DJ et al. (2018) Detecting morphed passport photos: a training and individual differences approach. Cogn Res Princ Implic 3(27)
7. Robertson DJ (2020) Morphed passport photo detection by human observers. In: International conference on biometrics for borders. Warsaw
8. Spiegel (2021) Aktivisten schmuggeln Fotomontage in Reisepass. https://www.spiegel.de/netzwelt/netzpolitik/biometrie-im-reisepass-peng-kollektiv-schmuggelt-fotomontage-in-ausweis-a-1229418.html
9. Monroy M (2021) Laws against morphing. https://digit.site36.net/2020/01/10/laws-against-morphing/
10. The Peng! Collective (2021) Mask.ID Part II—We send our passports to Libya. https://pen.gg/campaign/mask-id-2/
11. GIMP (2021) GIMP animation package. https://www.gimp.org/news/2009/06/05/gimp-animation-package-260-released/
12. Luxand (2021) FaceMorpher. http://www.facemorpher.com/
13. Abrosoft (2021) FantaMorph. https://www.fantamorph.com/
14. Scherhag U, Rathgeb C, Merkle J, Breithaupt R, Busch C (2019) Face recognition systems under morphing attacks: a survey. IEEE Access, pp 23012–23026
15. (2021) Dlib C++ Library. http://dlib.net/
16. Wikipedia (2021) Image warping. http://en.wikipedia.org/wiki/Image_warping
17. Wolberg G (1994) Digital image warping, 1st edn. IEEE Computer Society Press, Los Alamitos, CA, USA
18. Rogers DF, Adams JA (1989) Mathematical elements for computer graphics, 2nd ed. McGraw-Hill Higher Education
19. Delaunay BN (1934) Sur la sphère vide. Bulletin de l'Académie des sciences de l'URSS, Classe des sciences mathématiques et naturelles 6:793–800
20. Ferrara M, Franco A, Maltoni D (2019) Decoupling texture blending and shape warping in face morphing. In: International conference of the biometrics special interest group (BIOSIG), Darmstadt, pp 1–5
21. Lai M, Oruç I, Barton JS (2013) The role of skin texture and facial shape in representations of age and identity. Cortex, pp 252–265
22. Gonzalez RC, Woods RE (2017) Digital image processing, 4th ed. Pearson
23. Damer N, Saladi AM, Braun A, Kuijper A (2018) Morgan: recognition vulnerability and attack detectability of face morphing attacks created by generative adversarial network. In: Internationa conference on biometrics theory, applications and systems, pp 1–10
24. Donahue J, Krähenbühl P, Darrell T (2017) Adversarial feature learning. https://arxiv.org/abs/1605.09782
25. Venkatesh S et al. (2020) Can gan generated morphs threaten face recognition systems equally as landmark based morphs?—vulnerability and detection. In: 8th International workshop on biometrics and forensics (IWBF). Porto Portugal, pp 1–6
26. Karras T, Laine S, Aila T (2019) A style-based generator architecture for generative adversarial networks. In: IEEE conference on computer vision and pattern recognition, pp 4401–4410

27. Zhang H et al. (2020) MIPGAN—generating robust and high quality morph attacks using GAN. https://arxiv.org/abs/2009.01729
28. Karras T et al. (2020) Analysing and improving the image quality of StileGAN. In: IEEE/CVF conference on computer vision and pattern recognition, pp 8110–8119
29. NIST (2021) Face recognition vendor test (FRVT) 1:1 verification. https://pages.nist.gov/frvt/html/frvt11.html
30. Grother P, Ngan M, Hanaoka K (2021) Ongoing face recognition vendor test (FRVT)—Part 1: verification. NIST, Gaithersburg, MD
31. FRONTEX—R&D Unit (2015) Best practice technical guidelines for automated border control (ABC) systems. FRONTEX, Warsaw, Poland, ISBN: 978–92–95205–50–5. https://doi.org/10.2819/39041
32. Scherhag U et al. (2017) Biometric systems under morphing attacks: assessment of morphing techniques and vulnerability reporting. In: International conference of the biometrics special interest group (BIOSIG). Darmstadt, Germany
33. GIMP (2021) GNU image manipulation program web site. http://www.gimp.org/
34. Ferrara M, Franco A, Maltoni D (2018) Face demorphing. IEEE Trans Inf Forensics Secur 13(4):1008–1017
35. Xiberpix (2021) Sqirlz Morph 2.1 web site. http://www.xiberpix.net/SqirlzMorph.html
36. Raja K et al. (2020) Morphing attack detection - database, evaluation platform and benchmarking. In: IEEE transactions on information forensics and security (TIFS)
37. (2021) AMSL face morph image data set. https://omen.cs.uni-magdeburg.de/disclaimer/index.php
38. DeBruine L, Jones B (2021) Face research lab London set. https://doi.org/10.6084/m9.figshare.5047666.v3
39. Neubert T, Makrushin A, Hildebrandt M, Kraetzer C, Dittmann J (2018) Extended stirtrace benchmarking of biometric and forensic qualities of morphed face images. IET Biometrics 7(4):325–332
40. Wolf A (2016) ICAO: portrait quality (reference facial images for MRTD), version 0.7
41. BioLab (2021) Bologna online evaluation platform web site. https://biolab.csr.unibo.it/fvcongoing/UI/Form/BOEP.aspx
42. Dorizzi B et al. (2009) Fingerprint and online signature verification competitions at ICB 2009. In: Proceedings 3rd IAPR/IEEE international conference on biometrics (ICB09). Alghero
43. BioLab (2021) FVC-ongoing web site. http://biolab.csr.unibo.it/fvcongoing
44. Robertson DJ, Kramer RSS, Burton AM (2017) Fraudulent ID using face morphs: experiments on human and automatic recognition. PLoS ONE 12(3)
45. Young AW, Burton AM (2017) Recognizing faces. Curr Dir Psychol Sci 26(3):212–217
46. Raghavendra R, Raja KB, Venkatesh S, Busch C (2017) Face morphing versus face averaging: vulnerability and detection. In: IEEE international joint conference on biometrics (IJCB). Denver, CO, USA