Chapter 16
Practical Evaluation of Face Morphing Attack Detection Methods

Luuk Spreeuwers, Maikel Schils, Raymond Veldhuis, and Una Kelly

Abstract Face morphing is a technique to combine facial images of two (or more) subjects such that the result resembles both subjects. In a morphing attack, this is exploited by, e.g., applying for a passport with the morphed image. Both subjects who contributed to the morphed image can then travel using this passport. Many state-of-the-art face recognition systems are vulnerable to morphing attacks. Morphing attack detection (MAD) methods are developed to mitigate this threat. MAD methods published in literature are often trained on a limited number of or even a single dataset where all morphed faces are created using the same procedure. The resulting MAD methods work well for these specific datasets, with reported detection rates of over 99%, but their performance collapses for face morphs created using other procedures. Often even simple image manipulations, like adding noise or smoothing cause a serious degradation in performance of the MAD methods. In addition, more advanced tools exist to manipulate the face morphs, like manual retouching or morphing artifacts can be concealed by printing and scanning a photograph (as used in the passport application process in many countries). Furthermore, datasets for training and testing MAD methods are often created by morphing images from arbitrary subjects including even male-female morphs and morphs between subjects with different skin color. Although this may result in a large number of morphed faces, the created morphs are often not convincing and certainly don't represent a best effort attack by a criminal. A far more realistic attack would include careful selection of subjects that look alike and create high quality morphs from images of these subjects using careful (manual) post-processing. In this chapter we therefore argue that for robust evaluation of MAD methods, we require datasets with morphed images created using a large number of different morphing methods, including

L. Spreeuwers (✉) · M. Schils · R. . Veldhuis · U. Kelly
University of Twente,Enschede, Netherlands
e-mail: l.j.spreeuwers@utwente.nl

R. . Veldhuis
e-mail: r.n.j.veldhuis@utwente.nl

U. Kelly
e-mail: u.m.kelly@utwente.nl

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_16

various ways to conceal the morphing artifacts by, e.g., adding noise, smoothing, printing and scanning, various ways of pre- and post-processing, careful selection of the subjects and multiple facial datasets. We also show the sensitivity of various MAD methods to the mentioned variations and the effect of training MAD methods on multiple datasets.

16.1 Introduction

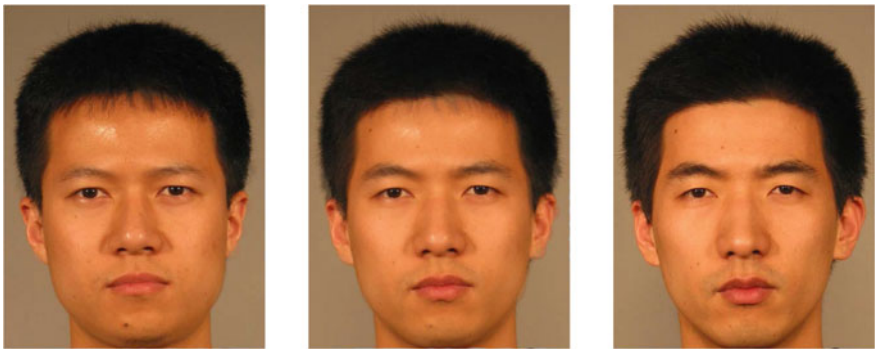
A morphed face image is a combination of two or more face images, created in a way that all contributing subjects are verified successfully against the morphed image. Suppose A' and B' are images of two distinct subjects A and B, shown in Fig. 16.1a and b. With face morphing, the two images are combined to create attack sample M, see Fig. 16.1c. If we perform identification tasks with state-of-the-art facial recognition software, a good morph will generate high comparison scores between morph M and templates of subjects A and B. It is obvious that face morphing poses a severe threat to all processes where face recognition is used to establish the identity of subjects, as first reported in [4]. Also human face recognition is vulnerable, as reported by Robertson et al. [15].

Automated morphing attack detection can be the solution to this problem. The morphing process leaves certain traces in the morphed image because the image is locally stretched or compressed and the images are combined. In high quality morphs, these textures differences are not visible to humans. Automated morphing attack detection scenarios can be subdivided into two types; morphing attack detection with or without a sample as reference. The scenario with reference sample means that apart from the morphed image, also an image of one of the original contributing subjects is available, which in principle makes morphing attack detection simpler. In this research we primarily address automated morphing attack detection without reference sample.

Many of the published methods for face morphing attack detection are developed and tested using a single dataset with morphed and bona fide samples and often good detection results are reported. However, the use of a single dataset and therefore a single, specific way to generate morphed images, may result in a morphing attack detection method that works well only for this specific type of face morphing. An example is morphing attack detection based on so-called double JPEG compression detection—detection of artifacts that occur because the morphed images are created from JPEG compressed images and compressed again when they are stored. Such a method will fail to detect morphed images if they are stored uncompressed.

The aim of this chapter is to demonstrate evaluation of morphing attack detection methods using single datasets and cross dataset testing and sensitivity to several simple morphing disguise techniques. It is based on research at the University of Twente, Netherlands, published in [18, 19].

In the remainder of this chapter, first a brief overview of some related work on face morphing attack detection is presented. Next, the creation of 4 datasets with morphed



(a) Face image A' (b) Morph M_m (c) Face image B'

Fig. 16.1 Bona fide face samples (left and right) and manual face morph (center). Images from FRGC [11]

face images is described that are used to train and test morphing attack detection methods. Multiple datasets are required to investigate cross dataset performance of morphing attack detection. Subsequently, a morphing attack detection method based on Local Binary Patterns (LBP) and a Support Vector Machine (SVM) is presented which will be used as a representation of morphing attack detection methods that are trained using a dataset with morphed and bona fide images. Next, two approaches to disguise morphing: adding nose and scaling images are presented for which we will investigate morphing attack detection robustness. Then, experiments and results are presented concerning within and cross dataset performance of morphing attack detection and robustness against morphing disguise and the effect of selection of faces that look alike. Finally, conclusions are presented.

16.2 Related Work

In order to evaluate the performance of morphing attack detection methods, the following metrics were introduced in ISO/IEC 30107-3 [2]:
Attack Presentation Classification Error Rate (APCER) Proportion of attack presentations incorrectly classified as bona fide presentations.
Bona Fide Presentation Classification Error Rate (BPCER) Proportion of bona fide presentations incorrectly classified as presentation attacks.

A bona fide sample refers to a non-morph and an attack sample refers to a morph. The trade-off between APCER and BPCER can be represented in a Detection Error Trade-Off (DET)-curve and also Equal Error Rates (EER) can be reported.

Currently, much published work on face morphing attack detection is based on textural feature classifiers, e.g., LBP features or features obtained using Convolutional Neural Networks, followed by an SVM classifier or other, see, e.g., [13, 20]. Tested

on single datasets of morphed face images good results are reported in literature. Creation of good datasets with morphed face images is one of the most important steps in the development of reliable face morphing attack detection methods. In [13] 450 morphed faces are created manually from a dataset comprised of 110 subjects. The face region is detected with Viola Jones detection. Various features like LBP, LBQ, 2DFFT (Fourier Transform) and BSIF filters are extracted. The combination of BSIF [6] with 7×7 and 12bit and SVM yields an Attack Presentation Classification Error Rate (APCER) of 1.73%. The dataset of 450 morphs was split into three subsets; training, testing, and validation. A problem with the dataset however is that these sets are not split according to the original 115 subjects. This means a morph in the training set may share a contributing subject with a morph in the test or validation set. In [17] the experiments from [13] are repeated, but instead the morphing attack detection process at a passport control is simulated by printing and scanning the face images. Morphing attack detection performance was analyzed before and after printing and scanning. It is found that printing and scanning images add noise and granularity, causing a loss in morphing attack detection performance. The dataset was split into training and testing sets without overlapping subjects. The reported performances are in the order of 40% BPCER at 10% APCER.

Apart from the various ways to split data in training and test sets, there are also various methods to create morphed images. The most popular method is based on the detection of landmarks in faces, triangularization, and warping of the triangles. More details are provided in Sect. 16.3. But there are various ways to define the landmarks and triangulation and each of them leads to small differences in the created morphs. It is also possible to create morphs manually using graphical software or to manually or automatically post-process the created morphs. Again this leads to variations in the types of morphs. Finally, also deep-learning methods for creation of face morphs are being developed, again leading to different types of morphs, see, e.g., [3].

In the next sections, it will be demonstrated that using only a single dataset for training and testing, even though it may be split into disjunct sets for training and testing, may lead to far too optimistic performance results. If the morphing attack detection methods are evaluated using datasets with morphed faces that were created using a different procedure or the images are manipulated by, e.g., adding some noise, the performance tends to be much worse.

16.3 Creation of Morphing Datasets

For experiments with morphing attack detection a large number of face morph images is required. We use automated morphing algorithms to quickly generate morphs. The dataset is split in a part for training and a part for testing with no overlap in subjects.

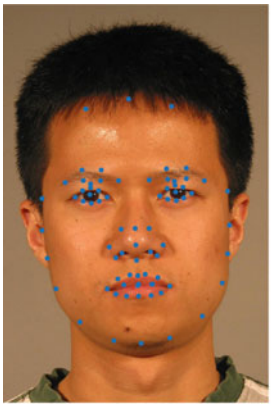
16.3.1 Creating Morphs

Various ways exist to create morphed face images. Nowadays, much research concentrates on the use of Generative Adversary Networks (GANs) for this purpose. However, the simpler landmark-based approaches still result in higher quality morphs. Therefore in this chapter, we chose this method to create morphs.

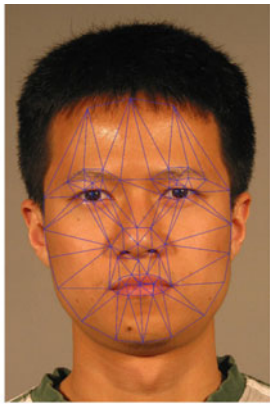
To create a face morph, the first step is to extract landmarks from both face images. For manual morphing the landmarks can be selected by hand, for automated morphing we use an existing landmark localisation algorithm. For morphing it is critical to know which parts in the image of one contributing subject correspond to the parts of another. Therefore it is vital that landmarks are accurately extracted, if they are placed incorrectly, it can lead to extremely poor morphs. There are several landmark localisation algorithms available. We found that STASM [10] and DLIB [7] result in high quality morphs. Figure 16.2a shows STASM landmarks on a face sample A'. A triangular mesh is defined over the landmarks using Delaunay Triangulation [8] (Fig. 16.2b). Now each triangle can be related to its corresponding triangle from the other contributing image. The triangles are morphed toward average triangles located in the final morph M_a using an affine transformation.

A blending value α defines the weight of contribution of the involved subjects. There are various ways of selecting α: we can set $\alpha = 0.5$ so that both subjects contribute equally to the morph or face recognition software can be used to set α so that the morph generates approximately the same comparison score for both contributing subjects. If the morph should resemble one of the subjects more than the other (the passport application is considered more critical than the use of the passport for automated border control), α can be set to a value of, e.g., 0.3 or 0.7.

The automatically generated morphs normally suffer from artifacts near the boundaries of the face and around the eyes, nose and mouth, because of the lim-



(a) STASM Landmarks



(b) Delaunay Triangulation

Fig. 16.2 Initial steps of the morphing process (images from FRGC [11])

Table 16.1 Characteristics of the datasets, resolution is given in pixels Inter Eye Distance (IED)

Dataset	Resolution IED (pix)	Morph train images	Bona fide train images	Morph test images	Bona fide test images
FRGC	129	500	150	500	150
ARF	177	500	150	500	100
Feret Color	177	750	250	750	250
Feret Gray	60	500	200	500	200

ited number of landmarks. In our research on morphing attack detection, we only used the inner part of the face.

When creating morphed face images, it is vital to save them in a lossless format like ".png" to ensure the morphing attack detection methods do not detect compression artifacts.

16.3.2 Datasets

We created four datasets with images of different quality and properties, originating from different facial datasets: FRGC [11], ARF [9], Feret color and Feret gray [12].

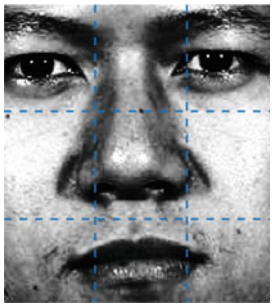
An overview of the created datasets with information on resolution (Inter Eye Distance, IED), number of training and testing images is given in Table 16.1.

Note that the resolution of the Feret Gray dataset is much lower than the resolution of the other datasets. This may impact morphing attack detection performance. Care was taken to use different subjects for each of the subsets: Morph Train, Non-Morph Train, Morph Test and Non-Morph Test. For all morphs, we used $\alpha = 0.5$ for the blending factor.

16.4 Texture-Based Face Morphing Attack Detection

To demonstrate the effects of within and cross dataset testing and concealing morphing artifacts, we chose a simple example of a trained texture-based morphing attack detection method. Even though BSIF filters perform better in literature, we chose to use LBP to extract features as it is not trained and shows results close to that of BSIF. With the use of landmarks the face region as shown in Fig. 16.3a is extracted and resized to a fixed size. The face region is cut off at the top of the eyebrows and somewhat below the mouth. With this region we ensure that the sides of the face which often contain obvious morphing artifacts are not present in the face image. We convert the image to gray scale and apply histogram equalization, enhancing image contrast (Fig. 16.3b). Using the FRGC dataset we performed a parameter sweep for

(a) Original cropped face



(b) Gray, histogram equalised cropped face

Fig. 16.3 Region of interest for LBP operator, the dashed lines show the areas for which local LBP histograms are obtained

LBP parameters: uniform/non-uniform LBP, number of neighbors n and radius r. We find that uniform LBP features with "standard" parameters, ($n = 8, r = 1$) and a 3×3 histogram result in a good performance. Increasing the number of histograms; e.g., 4×4 or 5×5 layout, only slightly increases the performance but also the dimensionality of the feature space increases. We therefore decided to use the "standard" parameters. For uniform LBP, a single histogram contains 59 feature values, which means for a 3×3 layout the feature space has 531 dimensions. The SVM classifiers are trained on between 650 and 1,000 samples.

16.5 Morphing Disguising

As pointed out earlier, often morphing attack detection methods are trained on a single dataset with morphed images. This may result in a morphing attack detection method that only detects a certain property of the morphing creation process. If the morphing creation process is slightly disturbed, these methods will fail.

Here, we investigate two simple ways to disguise the morphing process: adding Gaussian noise to the image and rescaling. In the first approach, a small amount of Gaussian noise is added to the image, masking certain noise characteristics of the morphing process that a morphing attack detection method may have learnt. The noise is kept small, such that to the human eye it is barely noticeable, see Fig. 16.4.

In the second approach, the image is down-scaled using a scaling factor s and then up-scaled again to its original resolution. In this way, some of the higher spatial frequencies are lost also masking the typical noise characteristics of morphed images. Examples of down-up scaled images are shown in Fig. 16.5. Again the manipulation is barely noticeable to the human eye.

Another way to hide the artifacts of face morphing is to print the photograph on paper and next scan it to obtain a digital photograph again. This is still common

(a) Example of morph with $\sigma = 0.01$

(b) Example of morph with $\sigma = 0.025$

Fig. 16.4 Morphs with added Gaussian noise. The gray level range of the image is 0.1

(a) Example of a morph with $s = 0.8$

(b) Example of a morph with $s = 0.5$

Fig. 16.5 Down-up scaled morphs to disguise morphing

practice for passport application in many countries, where the photographer prints the photograph and the subject brings the printed photograph to the municipality to apply for a new passport. The printed photograph is scanned in order to obtain a digital representation that is stored in the chip of the passport and is printed on the passport data page. The effect of printing and scanning has been thoroughly investigated in [5], where a significant decrease in morphing attack detection performance is reported. If the morphing attack detection methods are also trained on printed and scanned photographs, the performance improves again but is still significantly lower that on digital-only images. The effect is very comparable to the effects of adding noise and scaling we demonstrate in Sect. 16.6.

16.6 Experiments and Results

In order to demonstrate the impact of a number of the described factors on the performance of the LBP/SVM morphing attack detector, we present the following experiments:

1. Within dataset performance
2. Cross dataset performance
3. Mixed dataset performance
4. Robustness against additive Gaussian noise
5. Robustness against down-up scaling
6. Selection of similar subjects

16.6.1 Within Dataset Performance

With this experiment we investigate if the morphing attack detection method we used performs in line with the results reported in literature. Furthermore, we use the performance as a baseline to compare the results of the other experiments with.

For each of the datasets listed in Table 16.1 the SVM of the morphing attack detector was trained on features extracted from the training set and the morphing attack detection was determined using the test set.

The results are shown in the form of a DET-curve in Fig. 16.6. We can observe that the performance for 3 of the 4 datasets is similar (EER 2.5–5%), while for the low resolution Feret Gray set the results are poorer (EER = 17%). The reason for the poorer results is likely that the image quality (resolution) of the Feret Gray dataset is significantly lower.

The EER for the various datasets is shown in the top of Table 16.2. The MAD methods trained on the different datasets are called LBP-SVM1-LBP-SVM4.

The performance on the other datasets is in line with results reported in literature (EER = 1.7% in [13]).

16.6.2 Cross Dataset Performance

Next the cross dataset morphing attack detection performance is determined. In this experiment the SVMs are trained using the binary pattern features of the one dataset and tested using the test set of another dataset. The experiments were only conducted for the FRGC and ARF datasets and the results are shown in the middle part of Table 16.2.

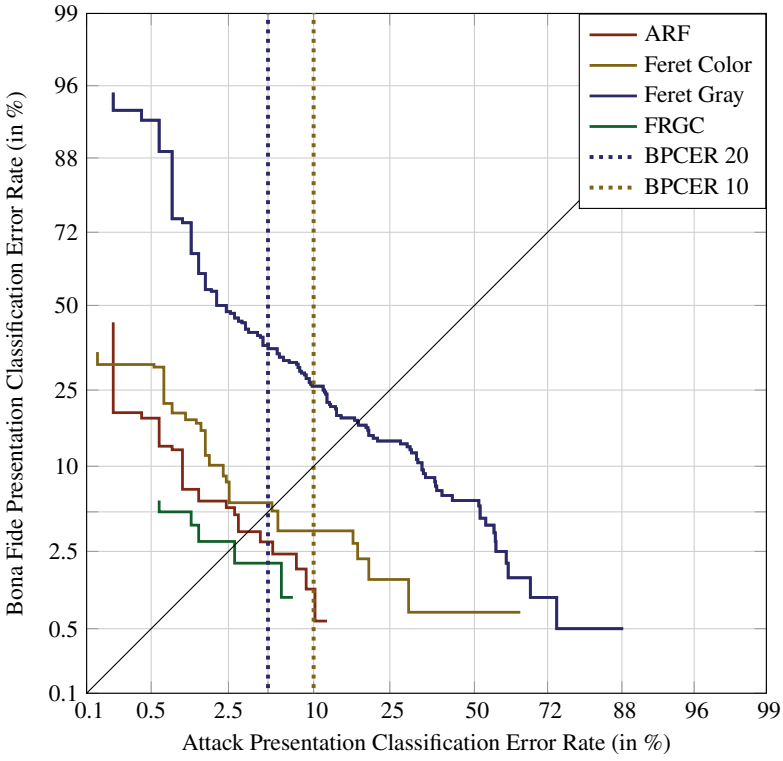


Fig. 16.6 DET-curve of LBP experiments on all datasets

Table 16.2 MAD performance reported as EER for within, cross and mixed dataset testing for various datasets

MAD method	Training set	Test set	Test proc.	EER (%)
LBP-SVM1	FRGC	FRGC	Within	2.5
LBP-SVM2	ARF	ARF	Within	3
LBP-SVM3	Feret Color	Feret Color	Within	5
LBP-SVM4	Feret Gray	Feret Gray	Within	20
LBP-SVM1	FRGC	ARF	Cross	80
LBP-SVM2	ARF	FRGC	Cross	79
LBP-SVM5	FRGC+ARF	FRGC+ARF	Mixed	35

The cross dataset performances were much worse than the within dataset performances, suggesting that indeed the morphing attack detector learnt features very specific for the dataset it was trained on: the EER of the LBP-SVM1 and LBP-SVM2 methods increases to 80% resp. 79%.

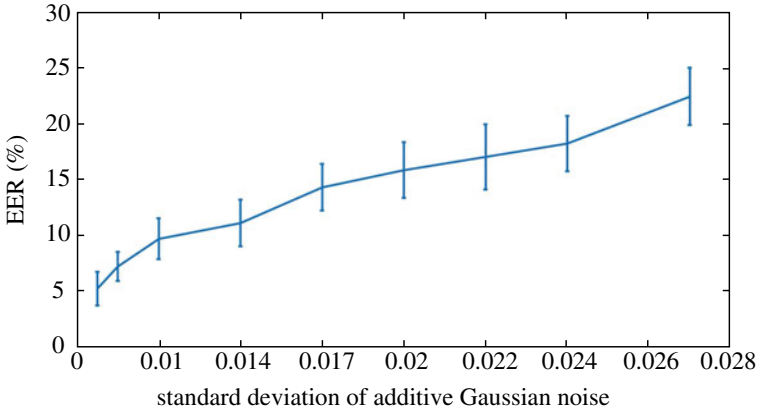


Fig. 16.7 Morphing attack detection performance for added Gaussian noise

16.6.3 Mixed Dataset Performance

In this experiment the SVMs are trained using 50% of both of the datasets FRGC and ARF and tested using the test set of both datasets. The results are given at the bottom of Table 16.2. The EER for this mixed test set is equal to 35%.

The mixed dataset performance is better than the cross dataset performances, suggesting that if multiple datasets are used for training, the morphing attack detector becomes more robust. The performance is still much worse than the within dataset performance, though.

16.6.4 Robustness Against Additive Gaussian Noise

In this experiment, we add Gaussian noise to the morphed images in order to disguise artifacts generated by the morphing process. The standard deviation of the noise was varied from 0.004 to 0.027, where the gray level range was normalized to 0.1. Only within dataset performance is reported.

The results are depicted in Fig. 16.7. We can observe that for small σ of the noise, the EER of the morphing attack detection is still around 5%, close to the baseline experiment. When the noise increases, the EER increases to above 20% for $\sigma = 0.027$. Note that even this noise will not be observed by human inspection, so it seems morphing artifacts can quite successfully be disguised by adding a bit of noise to the morphed images.

The experiments were done several times for different divisions of the data in training and test sets. The error bars show the minimum and maximum EER values obtained.

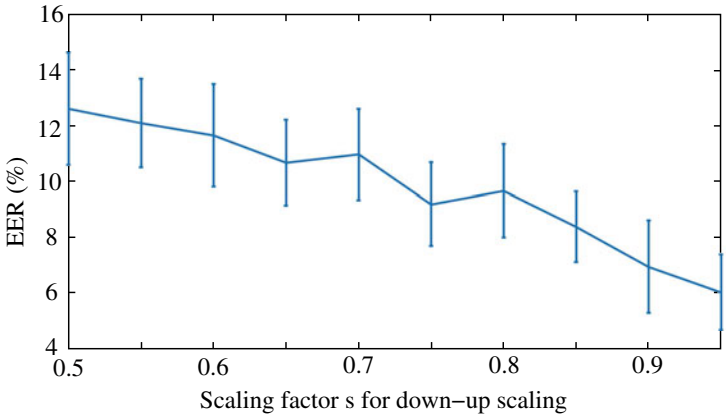


Fig. 16.8 Morphing attack detection performance for down and up scaling with scaling factor s.

16.6.5 Robustness Against Scaling

In this experiment, the original face images are first down-scaled with a factor s and then up-scaled again to their original resolution. In this way, some fine detail, i.e., high spatial frequency information is lost. Since morphing also influences (high) frequency contents of the face images, it is likely that traces caused by morphing can be obscured by this down-up scaling of the image. We investigated the impact on the morphing attack detection performance for a scaling range of $s = 0.5.0.95$. Only within dataset performance is reported.

The results are depicted in Fig. 16.8. We can observe that for $s = 0.95$, i.e., hardly any high frequency information is lost, the EER of the morphing attack detection is still around 5%, close to the baseline experiment. When the down scaling factor is lower, the EER increases to above 12% for $s = 0.5$. Note that even for this scaling factor, the difference to the original image will not be observed by human inspection, so it seems morphing artifacts can successfully be disguised by down-up scaling as well.

The experiments were done several times for different divisions of the data in training and test sets. The error bars show the minimum and maximum EER values obtained.

16.6.6 Selection of Similar Subjects

For this experiment, we created two sets of morphed faces. For the first set, arbitrary images were used to create morphs without paying attention to the similarity between the subjects. Indeed, even morphs between male and female subjects occur in this

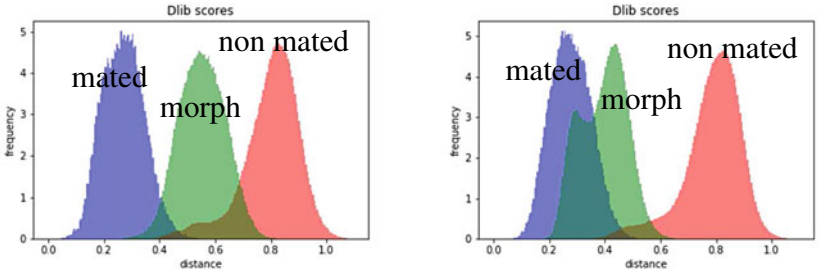


Fig. 16.9 Distance score distributions of morphs of arbitrary subjects (left) and of subjects selected on their resemblance (right). The distance scores of the latter are much closer to those of bona fide images

dataset. For the second set, the subjects used to create morphs were selected in such a way that gender matched and according to the DLIB face recognition system [1] they are reasonably similar. In Fig. 16.9 the distance scores of mated comparisons (2 images of the same subject), non-mated comparisons (two different subjects) and morph comparisons (a morph with an image of one of the contributing subjects) are depicted. The DLIB face recognition system decides that two images originate from the same subject if the distance score is below 0.6. In Fig. 16.9 on the left it can be seen that for morphs from arbitrary subjects about 70% of the morphs are accepted as genuine images, while for the morphs created from subjects selected on their resemblance, nearly all morphs are accepted (Fig. 16.9 right). Of course, criminals will attempt to create as good morphed face images as they can, thus the 2nd scenario is much more likely in practice. Therefore, it is important that morphing attack detection systems should not only be evaluated using morphs created using various different morphing methods, but also with morphs created from carefully selected similar subjects representing a criminals best effort to create high quality face morphs, see, e.g., [16].

16.7 The SOTAMD Benchmark

A very good attempt at creating a versatile benchmark for morphing attack detection methods was developed in the framework of the European SOTAMD (State Of The Art of Morphing attack Detection) project [14]. It includes morphed images created using 7 different morphing algorithms with various post-processing methods including manual post-processing for part of the dataset. In addition it includes printed and scanned bona fide and morphed images using several print and scan protocols. The subjects used to create morphs were selected based on various criteria including facial recognition scores and human observation. In [14] several morphing attack detection algorithms are tested. On the hardest tests, all these algorithms fail to provide acceptable results, which demonstrates the great challenge of reliable morphing attack detection.

16.8 Conclusion

Face morphing, the combination of two face images of distinct subjects into one image that resembles both subjects, poses a serious threat to face recognition. In several publications it is claimed that reliable morphing attack detection is possible. We noticed that often morphing attack detection methods are developed and tested using a single dataset with morphed face images. In this chapter we show that this results in morphing attack detection that only works well for a single type of morph or dataset. Using a LBP/SVM based morphing attack detection method that performs well on a single dataset (around 2% EER), we show that for cross dataset testing, the performance collapses resulting in an EER as high as 80%. Experiments with mixed datasets suggest that morphing attack detection can be made more robust if trained on multiple datasets. In addition, we show that the morphing artifacts that are used as features for detection can be obscured by simple image manipulations like adding Gaussian noise or down-up scaling the morphed images. The EER for within dataset detection increased from below 5% to above 20% for adding noise and above 12% for down-up scaling. In both cases the manipulation was almost invisible to the human observer.

We therefore argue that morphing attack detection methods should be tested extensively on multiple datasets obtained from different sources and morphing methods and a range of image manipulations. Furthermore, they should be tested on morphed face images that were created from similar subjects rather than arbitrary subjects and carefully post-processed in order to mimic a criminal's best effort at creating high quality facial morphs.

References

1. http://dlib.net/face_recognition.py.html
2. Information technology-biometric presentation attack detection-part 3: Testing and reporting, jtc 1/sc 37. ISO/IEC FDIS 30107-3:2017 (2017)
3. Damer N, Saladié AM, Braun A, Kuijper A (2018) 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), pp 1–10. https://doi.org/10.1109/BTAS.2018.8698563
4. Ferrara M, Franco A, Maltoni D (2014) The magic passport. In: IEEE international joint conference on biometrics, pp 1–7. https://doi.org/10.1109/BTAS.2014.6996240
5. Ferrara M, Franco A, Maltoni D (2019) Face morphing detection in the presence of printing/scanning and heterogeneous image sources. CoRR arXiv:1901.08811
6. Kannala J, Rahtu E (2012) Bsif: binarized statistical image features. In: Proceedings of the 21st international conference on pattern recognition (ICPR2012), pp 1363–1366
7. King DE (2009) Dlib-ml: a machine learning toolkit. J Mach Learn Res 10:1755–1758
8. Lee DT, Schachter BJ (1980) Two algorithms for constructing a delaunay triangulation. Int J Comput Inf Sci 9(3):219–242. https://doi.org/10.1007/BF00977785.
9. Martinez AM, Benavente R (1998) The AR Face Database. Tech. rep, CVC
10. Milborrow S, Nicolls F (2014) Active shape models with SIFT descriptors and MARS. VISAPP

11. Phillips PJ, Flynn PJ, Scruggs T, Bowyer KW, Chang J, Hoffman K, Marques J, Min J, Worek W (2005) Overview of the face recognition grand challenge. In: Proceedings of the 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol 1. IEEE Computer Society, Washington, DC, USA (2005), pp 947–954. https://doi.org/10.1109/CVPR.2005.268
12. Phillips PJ, Wechsler H, Huang J, Rauss PJ (1998) The feret database and evaluation procedure for face-recognition algorithms. Image Vis Comput 16(5):295–306
13. Raghavendra R, Raja KB, Busch C (2016) Detecting morphed face images. In: 2016 IEEE 8th international conference on biometrics theory, applications and systems (BTAS), pp 1–7. https://doi.org/10.1109/BTAS.2016.7791169
14. Raja K, Ferrara M, Franco A, Spreeuwers L, Batskos I, De Wit F, Gomez-Barrero M, Scherhag U, Fischer D, Venkatesh S, Singh JM, Li G, Bergeron L, Isadskiy S, Ramachandra R, Rathgeb C, Frings D, Seidel U, Knopjes F, Veldhuis R, Maltoni D, Busch C (2020) Morphing attack detection-database, evaluation platform and benchmarking. IEEE Trans Inf Forensics Secur 1–1. https://doi.org/10.1109/TIFS.2020.3035252
15. Robertson D, Kramer R, Burton A (2017) Fraudulent id using face morphs: experiments on human and automatic recognition. PLoS One 12(3). https://doi.org/10.1371/journal.pone.0173319
16. Röttcher A, Scherhag U, Busch C (2020) Finding the suitable doppelgänger for a face morphing attack. In: 2020 IEEE international joint conference on biometrics (IJCB), pp 1–7. https://doi.org/10.1109/IJCB48548.2020.9304878
17. Scherhag U, Raghavendra R, Raja KB, Gomez-Barrero M, Rathgeb C, Busch C (2017) On the vulnerability of face recognition systems towards morphed face attacks. In: 2017 5th international workshop on biometrics and forensics (IWBF), pp 1–6. https://doi.org/10.1109/IWBF.2017.7935088
18. Schils M (2017) Towards a structured approach for face morphing detection. University of Twente, Master EE Biometrics and Computer Vision
19. Spreeuwers L, Veldhuis R, Schils M (2018) Towards robust evaluation of face morphing detection. In: 2018 26th European signal processing conference, EUSIPCO 2018, European signal processing conference. IEEE, United States, pp 1027–1031. https://doi.org/10.23919/EUSIPCO.2018.8553018. http://www.eusipco2018.org/. (26th European Signal Processing Conference, EUSIPCO 2018, EUSIPCO ; Conference date: 03-09-2018 Through 07-09-2018)
20. Wandzik L, Kaeding G, Vicente-Garcia R (2018) Morphing detection using a general- purpose face recognition system. In: 26th European signal processing conference, EUSIPCO 2018, Roma, Italy, 3–7 Sept 2018. IEEE, pp 1012–1016. https://doi.org/10.23919/EUSIPCO.2018.8553375.