CNNs under Attack: On the Vulnerability of Deep Neural Networks Based Face Recognition to Image Morphing

Lukasz Wandzik, Raul Vicente Garcia, Gerald Kaeding, and Xi Chen
lukasz.wandzik@ipk.fraunhofer.de, raul.vicente@ipk.fraunhofer.de,
gerald.kaeding@ipk.fraunhofer.de, xi.chen@ipk.fraunhofer.de

Fraunhofer Institute for Production Systems and Design Technology IPK,
Pascalstraße 8 – 9, 10587 Berlin, Germany

Abstract. Facial recognition has become a critical constituent of common automatic border control gates. Despite many advances in recent years, face recognition systems remain susceptible to an ever evolving diversity of spoofing attacks. It has recently been shown that high-quality face morphing or splicing can be employed to deceive facial recognition systems in a border control scenario. Moreover, facial morphs can easily be produced by means of open source software and with minimal technical knowledge. The purpose of this work is to quantify the severeness of the problem using a large dataset of morphed face images. We employ a state-of-the-art face recognition algorithm based on deep convolutional neural networks and measure its performance on a dataset of 7260 high-quality facial morphs with varying blending factor. Using the *Inception-ResNet-v1* architecture we train a deep neural model on 4 million images to obtain a validation rate of 99.96% at 0.04% false acceptance rate (FAR) on the original, unmodified images. The same model fails to repel 1.13% of all morphing attacks, accepting both the impostor and the document owner. Based on these results, we discuss the observed weaknesses and possible remedies.

Keywords: face recognition, biometric spoofing, face morphing, deep learning

1 Introduction

The detection of biometric counterfeits, commonly known as anti-spoofing, is a very active field of research. A diversity of techniques has been proposed in literature for protecting face recognition systems in real authentication scenarios like border control. However, as shown in recent surveys [4, 6], most works have only considered presentation attacks, neglecting the often fragile face biometrics enrollment process. In particular, a simple face morphing trick can be applied so that two different persons can potentially pass through an automated border control gate with the same electronic machine readable travel document (eMRTD). This attack simply consists in submitting for enrollment a facial image

that is obtained from morphing the face of the legitimate document owner (accomplice) and the face of a reasonably similar looking impostor. In a successful attack, the face recognition system positively matches the tampered template stored in the eMRTD to the live image of the impostor. The authors of [3] originally presented the possibility of such an attack. However, in that work only commercial, closed source face recognition systems and a very limited set of test images were employed. In contrast hereto, the main contribution of this work is a quantitative estimate of the severeness of the so-called *morphing attack* (see Fig. 1 for an example illustration) based on extensive experimental results, using a large test dataset and an exemplar face recognition method that is both state-of-the-art and publicly accessible. The employed exemplary methods and datasets are selected in order to model a sufficiently realistic, yet optimistic scenario. Herewith, we aim to estimate the lower bounds of the success rate of morphing attacks against current face recognition methods.

2 Related Work

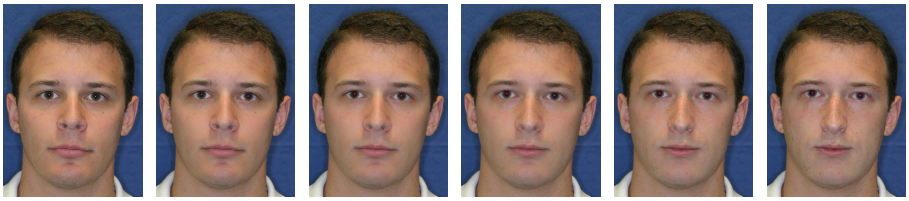
Intentional manipulation of images and its impact on modern generic image recognition systems has been addressed in the case of image classification. The authors of [20] and [12] show how to perform image modifications that are imperceptible to the human eye, and yet drastically change the outputs of the attacked recognition system. Also unintentional effects produced by the image acquisition system, like perspective distortion, can pose a problem for face recognition. This has been addressed in [21] through image de-warping models. The dependency of recognition performance on image quality in a broader sense has been the subject of research in several works [1, 8]. The threat of a morphing attack in a eMRTD scenario using commercial face recognition software was first identified in [3]. More recent publications [11, 15] have addressed this specific problem and propose methods for both the generation of morphs and for the detection of manipulation traces in the image.

3 Methods

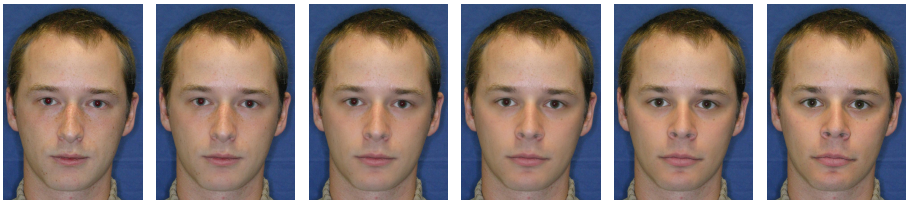
In this section, we introduce the methods employed for generating our face morphing dataset and describe the face recognition pipeline.

3.1 Face Morphing

Face morphing is usually done by projecting and blending the coordinates and texture informations of many corresponding image regions [23] from two source images S_1 and S_2 into a new synthesized destination image D. These regions can be generated by an automatic landmark detection followed by a Delaunay triangulation. One problem with this approach is the occurrence of blending artifacts, especially in non related regions. Since we only have reliable landmarks



(a) 1.0 (b) 0.8 (c) 0.6 (d) 0.4 (e) 0.2 (f) 0.0



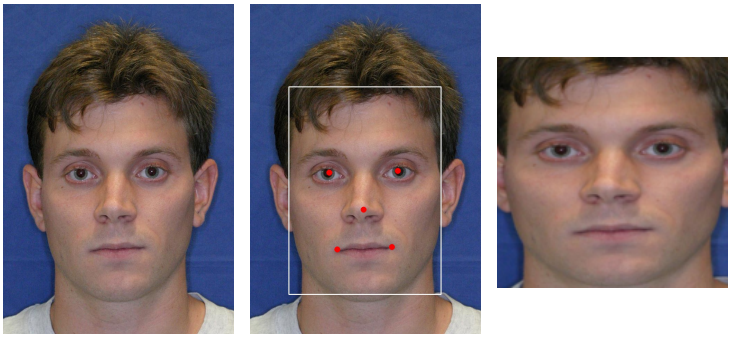
(g) 1.0 (h) 0.8 (i) 0.6 (j) 0.4 (k) 0.2 (l) 0.0

Fig. 1. Visualization of the morphing process with respect to blending factors. The images were generated using two identities 191 (top) and 217 (bottom) from the *MultiPIE* dataset [16]. A blending factor of 1.0 corresponds to the original, unmanipulated image, whereas 0.0 represents the case where the image information of the destination image was entirely replaced by the image data of the source image. All other images show a linear transformation between these two extremes.

for the facial regions, such unrelated assignments will occur outside the convex hull of these landmarks. A way to suppress such artifacts could be the manual retouch of the synthesized image, but this is infeasible for generating a large amount of examples. To address this problem, we only blend the facial regions $\hat{S}_1$ and $\hat{S}_2$ that lie within the convex hulls of our landmarks. To obtain these regions we first detect 68 landmarks [9] using the Dlib library [10] within S_1 and S_2, followed by a Delaunay triangulation. We next determine corresponding triangles to compute the projections. To obtain our morphing image we first set $D = S_2$. We next blend the geometry and texture informations of eyes, nose, and mouth of $\hat{S}_1$ and $\hat{S}_2$. By adopting the outer shape of $\hat{S}_2$ we obtain the destination image $\hat{D}$. At this moment $\hat{D}$ consists only of the morphed facial region. Finally, we blend $\hat{D}$ with D using the Poisson blending of [14]. An exemplary face morphing is depicted in Figure 1. One can see that the outer shape of S_2 will be preserved, while the geometry and texture information of the relevant facial regions of both input images will be blended.

3.2 Face Recognition Methods

Face Registration. In order to improve performance of facial recognition systems a number of preprocessing steps are applied to the input images, including



(a) Input image (b) Bounding box (c) Face crop

Fig. 2. Input image preprocessing using the deep cascaded multi-task framework [24].

face detection and alignment. In our work, we use the recently proposed deep cascaded multi-task framework for face detection and alignment [24]. Its hierarchical architecture with three stages of deep convolutional networks predicts face and landmark location in a coarse-to-fine manner. The resulting face bounding box is padded by a margin of 16 pixels on each side and finally resized to a 160×160 pixel square in order to fit the input of our neural network (Fig. 2). The input image size translates to about 90 pixels interpupillary distance, which complies with the ICAO recommendations[1] for eMRTDs and corresponds to the actual conditions of a real morphing attack.

Feature Extraction. We have tested several network architectures for face verification including the works of Parkhi et al. [13] and Schroff et al. [18]. The best performing network is based on the recently purposed *Inception-ResNet-v1* architecture [19]. Its topology includes inception modules as well as residual connections and achieves a state-of-the-art accuracy of $99.3\% \pm 0.04$ on the challenging LFW benchmark dataset [7]. We compute the face embeddings using a tensorflow implementation [17] of the *FaceNet* architecture introduced by Schroff et al. [18]. Instead of a pure inception model, as described in the original paper, [17] employs the *Inception-ResNet-v1* network architecture. Furthermore, the network is trained as a classifier using the Center Loss approach [22] instead of Triplet Loss. We extract the feature vectors by tapping the last fully connected layer before the softmax. The output of this layer is a 1792-dimensional L2-normalized vector. For decision making, we threshold the Euclidean distance between pairs of feature vectors computed by the neural network.

[1] ICAO, Machine Readable Travel Documents, Seventh Edition 2015, Part 9: Deployment of Biometric Identification and Electronic Storage of Data in eMRTDs.

4 Experiments

4.1 Datasets

Training. In order to train the deep neural network from scratch, we use the MS-Celeb-1M dataset which contains the top 100k subjects from the 1M most popular celebrities list [5]. There are approximately 100 images of each celebrity, resulting in about 10M images in total. Due to the vast number of automatically acquired data some images suffer from mislabeling. Therefore, only a subset of them was used for training. The dataset was cleaned up by computing the distance between a given image embedding and its class center. Keeping only 75% of the images that are closest to its class center reduces the dataset size to about 4M images divided into 51k classes.

Morphing. The face images used for investigating the impact of a morphing attack on our system come from the well known *MultiPIE* [16] dataset. The image data was collected under controlled conditions, similar to those prescribed for eMRTDs. *MultiPIE* contains images of 337 subjects from four different photo sessions, including variations in viewpoint, illumination and expression. In this work, we consider only frontal views and discard coarse illumination changes. To create our morphing dataset we only use images from session one and restrict ourselves to subjects that occur in at least one more session. Furthermore, we only consider subjects not wearing glasses and those belonging to the same gender, in order to avoid unnatural image artifacts. In total, we create 3630 pairs with 22 morphed images each. These images are used for enrollment purposes only, whereas images of session two to four are used for verification.

4.2 Data Preparation

We subdivide the dataset into *positive* and *negative* pairs of identities and use this structure for all experiments throughout the study. The positive pairs consist of the morphed image and the accomplice image whereas the negative pairs contain the impostor and the morphed image. Given subjects from session one of the *MultiPIE* dataset, we select two images I_a and I_b such that

$$\|\phi(I_\mathrm{a}) - \phi(I_\mathrm{b})\|_2 > t \tag{1}$$

where ϕ is an embedding of facial features. By this constraint, we make sure that the recognition algorithm is able to discern the individuals. Next, we merge those two images using the method described in the previous section and obtain a morphed image I_ab

$$\mathrm{morph}(I_\mathrm{a},\, I_\mathrm{b}) \rightarrow I_\mathrm{ab} \tag{2}$$

During the evaluation process, the generated image I_ab is matched against images from session two to four, provided that the respective individual is available in that session

$$\|\phi(I_\mathrm{ab}) - \phi(I_x)\|_2 < t, \;\; \text{where} \;\; x \neq \mathrm{a} \wedge x \neq \mathrm{b} \tag{3}$$

so that the images used for producing the morphs are not being used in the evaluation process. For any given identity pair, we generate two kinds of morphing images I_{ab} and I_{ba} by swapping the source and destination identities. Following the described procedure, we generate 17992 image pairs for both positive and negative class, resulting in a total number of 35984 pairs.

4.3 Evaluation Procedure

The evaluation procedure starts by combining identity pairs that share a common morphed image into a triplet. We generate a total number of 44546 triplets using positive and negative identity pairs. The morphed image acts as reference and the accomplice or impostor image as query. For the purpose of this study we define the morphing attack as follows:

Definition 1. *Let I_{a} be the accomplice image and I_{b} the impostor image. Let t be some optimal threshold and ϕ an embedding of facial features. A morphing attack using a tempered template I_{ab} is successful if and only if*

$$\|\phi(I_{\mathrm{ab}}) - \phi(I_{\mathrm{b}})\|_2 < t \quad \wedge \quad \|\phi(I_{\mathrm{ab}}) - \phi(I_{\mathrm{a}})\|_2 < t \tag{4}$$

We only consider triplets of images consisting of the morphed image I_{ab}, the impostor image I_{a}, and the accomplice image I_{b}. From that, we derive four possible outcome cases, as shown in Figure 3.

Case 1. The first case represents a successful morphing attack where both the accomplice and the impostor succeed and get accepted by the system. We also consider cases where only one of the individuals gets accepted, either the accomplice or the impostor.

Case 2. The second case reflects a correct operation mode of a face recognition system where the accomplice gets accepted and the impostor is rejected.

Case 3. Case three represents the reverse situation where the impostor gets accepted and the accomplice is rejected. This case could also be seen as a successful morphing attack, but in this work we restrict ourselves only to Case 1 as defined in 1.

Case 4. In the last, fourth case, both the accomplice and impostor are rejected. In this situation, a total failure of an morphing attack is simulated.

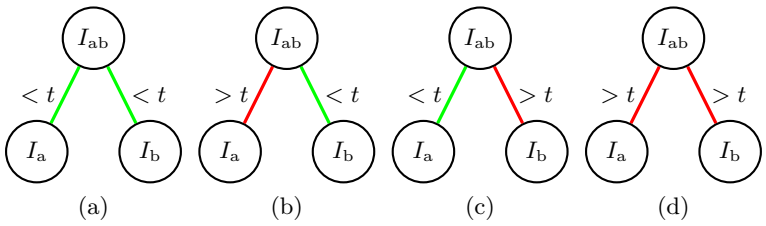


Fig. 3. Four possible outcome cases: a) Successful morphing attack. b) Correct verification. c) Accomplice rejection. d) Accomplice and impostor rejected.

4.4 Analysis of Variance

The success of a morphing attack is closely related to the discriminatory abilities of the employed classifier. By this token, a successful attack could be regarded as a misclassification error related to poor class separability in the feature space. In our case, the discriminant function was learned by the feature extracting neural network which encodes facial features and maps them to a new data distribution. We perform an analysis of variance in order to test for discriminatory power of features computed by our neural network. The goal is to determine at which blending factor a misclassification is most likely to occur. The analysis of variance (ANOVA) within and between the identity groups was conducted for positive as well as negative pairs of images by computing a test statistic. With this criterion, the quality of separability of classes in the feature space is measured. We perform a one-way ANOVA on the original images by using the Euclidean distance as a factor.

5 Results

5.1 Threshold Selection

Using the original images from the *MultiPIE* dataset, we compute the Euclidean distances between mated and unmated pairs of facial features and select a threshold at 0.1% FAR, as commonly prescribed for border control applications [2]. The optimal baseline threshold of 0.78 was selected using a 8-fold cross-validation procedure with a verification rate of 99.96% at 0.04% FAR.

5.2 Evaluation of Cases

Case 1: Successful Morphing Attack. As shown in Figure 4 the *optimal* blending factor from the perspective of a potential impostor is 0.4 for our setup. The plot also shows the maximum percentage of successful attacks on our system which amounts to 1.13% (Tab. 1). The low success rate is due to an *early* rejection of the accomplice while the blending factor increases and not due to rejection of the impostor.

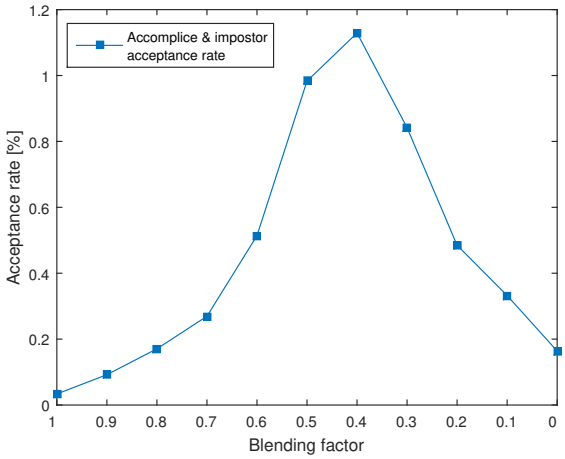


Fig. 4. Case 1: Successful morphing attack. The peak of 1.13% is reached at a blending factor of 0.4.

Table 1. Case 1: Successful attack rates for different blending factors. A factor of 1.0 corresponds to the original image and a factor of 0.0 to a complete replacement of the original face region. In this case both the accomplice and the impostor are accepted by the system.

Blending Factor	**Acceptance Rate [%]**	**Threshold**
1.0	0.03 %	0.78
0.9	0.09 %	0.78
0.8	0.17 %	0.78
0.7	0.27 %	0.78
0.6	0.51 %	0.78
0.5	0.98 %	0.78
0.4	**1.13** %	**0.78**
0.3	0.84 %	0.78
0.2	0.49 %	0.78
0.1	0.33 %	0.78
0.0	0.16 %	0.78

Case 2: Correct Verification. This case reflects the correct operation mode of a face recognition system. The performance of the system drops drastically for blending factors less than 0.7 (Fig. 5). The impostor rejection rate reaches its maximum of 29.86% at a blending factor of 0.0 (Tab. 2). This discrepancy is due to the fact that the accomplice fails to pass the verification process as we already stated in the previous paragraph. This is also the main reason for the low verification rate which, at a blending factor of 0.5, amounts to only 34.66%.

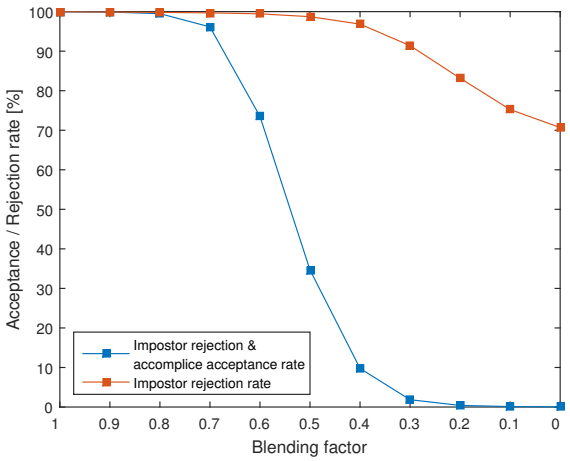


Fig. 5. Case 2: Correct verification rates for different blending factors. A factor of 1.0 corresponds to the original image and a factor of 0.0 to a complete replacement of the original face region. The main (blue) curve represents the correct verification case where the impostor gets rejected and the accomplice gets accepted. The impostor rejection rate was plotted for comparison.

Table 2. Case 2: The table shows accomplice acceptance rates with respect to the blending factor. A factor of 1.0 corresponds to the original image and a factor of 0.0 to a complete replacement of the original face region. Here the accomplice is accepted and the impostor is rejected.

Blending Factor	Acceptance / Rejection Rate [%]	Threshold
1.0	**99.97** %	**0.78**
0.9	99.89 %	0.78
0.8	99.49 %	0.78
0.7	96.14 %	0.78
0.6	73.50 %	0.78
0.5	34.66 %	0.78
0.4	9.76 %	0.78
0.3	1.85 %	0.78
0.2	0.41 %	0.78
0.1	0.12 %	0.78
0.0	0.03 %	0.78

Case 3: Accomplice Rejection. In this case the accomplice gets rejected by the recognition system while the impostor gets accepted. There is almost no rejection of the accomplice down to a blending factor of 0.7 (Fig. 6). However, the system fails to accept most of the accomplices at a blending factor of 0.5. This is due to manipulations introduced by our morphing tool which was to

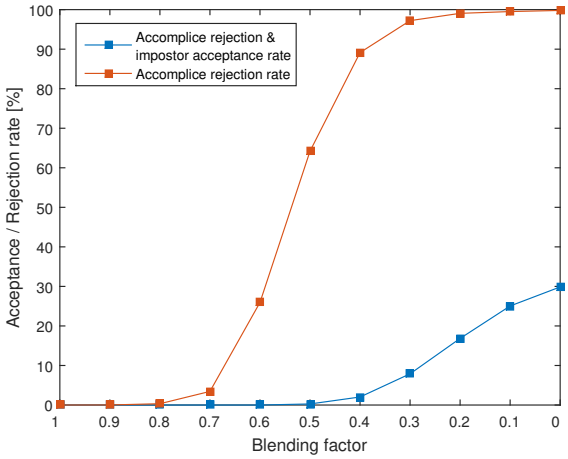


Fig. 6. Case 3: Accomplice rejection rates for different blending factors. A factor of 1.0 corresponds to the original image and a factor of 0.0 to a complete replacement of the original face region. The impostor is accepted and the accomplice is rejected in this case. The main curve corresponding to Case 3 is depicted in blue.

strong for the accomplice to get accepted. Up to this point there were also no impostor images that got accepted by the system.

Table 3. Case 3: The table shows impostor acceptance rates with respect to the blending factor. A factor of 1.0 corresponds to the original image and a factor of 0.0 to a complete replacement of the original face region. In this case the impostor is accepted and the accomplice is rejected.

Blending Factor	Acceptance / Rejection Rate [%]	Threshold
1.0	0.0 %	0.78
0.9	0.0 %	0.78
0.8	0.0 %	0.78
0.7	0.0 %	0.78
0.6	0.0 %	0.78
0.5	0.27 %	0.78
0.4	2.03 %	0.78
0.3	7.90 %	0.78
0.2	16.84 %	0.78
0.1	25.05 %	0.78
0.0	**29.86** %	**0.78**

Case 4: Complete Rejection. This case represents the complete opposite to the morphing attack, where both the accomplice and the impostor get rejected.

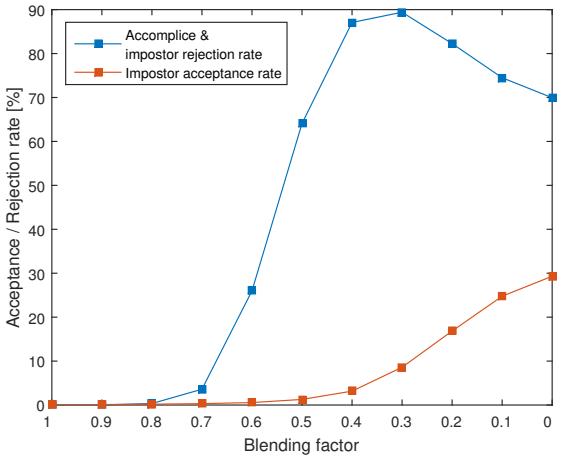


Fig. 7. Case 4: Complete rejection rates for different blending factors. A factor of 1.0 corresponds to the original image and a factor of 0.0 to a complete replacement of the original face region. Both the accomplice and the impostor are rejected in this case. The drop in the main curve (blue) is attributed to the increasing impostor acceptance rates depicted in orange.

The maximum rejection rate has its peak of about 90% at a blending factor of 0.3 (Tab. 4). After that, it drops to about 70% at 0.0. The drop is due to the increasing impostor acceptance rate, which is also shown in Figure 7.

Table 4. Case 4: The table shows rejection rates with respect to the blending factor. A factor of 1.0 corresponds to the original image and a factor of 0.0 to a complete replacement of the original face region. In this case both the impostor and the accomplice are rejected.

Blending Factor	**Rejection Rate [%]**	**Threshold**
1.0	0.0 %	0.78
0.9	0.02 %	0.78
0.8	0.34 %	0.78
0.7	3.59 %	0.78
0.6	25.99 %	0.78
0.5	64.08 %	0.78
0.4	87.08 %	0.78
0.3	**89.40** %	**0.78**
0.2	82.27 %	0.78
0.1	74.50 %	0.78
0.0	69.95 %	0.78

5.3 Examples of Successful Attacks

In Figure 8 three successful morphing attacks are shown. We only consider face morphs with a blending factor of 0.5 for this illustration and select individuals from three different ethnic groups. Our experiments revealed that about 75% of successful attacks were attributed to individuals of Asian descent. The reason for this situation may lay in our training dataset that contains mainly people of Caucasian descent. As a consequence, the model used for face recognition might be biased towards discerning Caucasian more accurately than Asians.

Fig. 8. Face images from three successful morphing attacks with a blending factor of 0.5. Accomplice image (left), morphed image (middle), impostor image (right).

5.4 Analysis of Variance

The null hypothesis is rejected if the F-value calculated from the data is greater than the critical value of the F-distribution for some desired false-rejection probability. The F-test reveals that for a blending factor of 1.0, the variance within positive and negative identity pairs is much lower than between the pairs. However, the situation changes with increasing blending factors. For a factor of 0.5, we fail to reject the null hypothesis with a significance level of 0.01 (Fig. 9). This means that starting from a blending factor of 0.5, the two groups containing positive and negative identity pairs cannot be separated anymore. A misclassification is most likely to occur at a blending factor of 0.4. The results are consistent with the evaluation presented in Section 5.1.

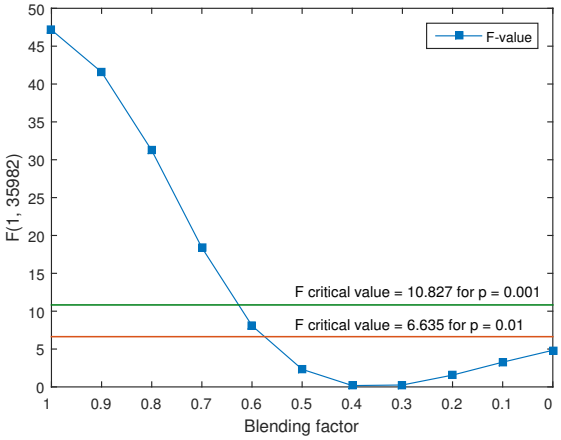


Fig. 9. F-values for increasing blending factors and critical F-values for two different significance levels. For values above 6.635, the null hypothesis that all group means are equal can be rejected with a significance level of 0.01.

Table 5. ANOVA test for an increasing blending factor. The lowest value is measured at a blending factor of 0.4. This is consistent with the result obtained for case 1 (Fig. 4).

Blending factor											
	1.0	0.9	0.8	0.7	0.6	0.5	**0.4**	0.3	0.2	0.1	0.0
$F(1, \infty)$	47.1	41.6	31.2	18.4	8.1	2.3	**0.2**	0.3	1.6	3.3	4.9

6 Conclusions

We present the results of a simulated morphing attack on a state-of-the-art face recognition system using a large dataset with varying blending factor. By defining a successful morphing attack as a situation where both accomplice and impostor get accepted, we state that the main reason for the low success rate of the attack is the rejection of the accomplice. This is a direct consequence of the compactness of identity clusters which could be attributed to good feature representation delivered by the convolutional neural network. CNNs account for less variability within the identity groups than previous methods, thus allowing for tighter thresholds and making the morphing attack less significant. However, poor image quality and low resolution can have a negative impact on cluster compactness. This may open the door for potential morphing attacks, as less compact identity groups require wider baseline thresholds.

In order to improve the robustness against morphing attacks, we plan to analyze facial features in more detail using soft biometrics and face symmetry. This includes the analysis of regions affected by the blending operation, e.g.

eyes and mouth as well as regions that were not manipulated, e.g. hair, ears and forehead. Face shape analysis is another factor that could help to reduce the number of successful attacks. The currently used registration method normalizes the original face shape, thus suppressing discriminative information that could help to repel the attack. We would like to investigate these presumptions in future work.

Acknowledgment

This work has been partially funded by the German Federal Ministry of Education and Research (BMBF) under contract number FKZ: 16KIS 0512.

References

[1] Dutta, A.: Predicting Performance of a Face Recognition System Based on Image Quality. Ph.D. thesis, University of Twente (2015), http://arxiv.org/pdf/1510.07112v1

[2] European Union, E.A.f.t.M.o.O.C.a.t.E.B.: Best Practice Technical Guidelines for Automated Border Control (ABC) Systems. FRONTEX (2015)

[3] Ferrara, M., Franco, A., Maltoni, D.: The magic passport. In: 2014 IEEE International Joint Conference on Biometrics (IJCB). pp. 1–7 (2014)

[4] Galbally, J., Marcel, S., Fierrez, J.: Biometric antispoofing methods: A survey in face recognition. IEEE Access 2, 1530–1552 (2014)

[5] Guo, Y., Zhang, L., Hu, Y., He, X., Gao, J.: Ms-celeb-1m: A dataset and benchmark for large scale face recognition. In: European Conference on Computer Vision. Springer (2016)

[6] Hadid, A., Evans, N., Marcel, S., Fierrez, J.: Biometrics systems under spoofing attack: An evaluation methodology and lessons learned. IEEE Signal Processing Magazine 32(5), 20–30 (2015)

[7] Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Tech. Rep. 07-49, University of Massachusetts, Amherst (October 2007)

[8] Jinyu Zuo, Harry Wechsler, et al: Adaptive Biometric Authentication using Nonlinear Mappings On Quality Measures and Verification Scores: 17th IEEE International Conference on Image Processing (ICIP), 2010. IEEE, Piscataway, NJ (2010), http://ieeexplore.ieee.org/servlet/opac?punumber=5641636

[9] Kazemi, V., Sullivan, J.: One millisecond face alignment with an ensemble of regression trees. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1867–1874 (2014)

[10] King, D.E.: Dlib-ml: A machine learning toolkit. Journal of Machine Learning Research 10, 1755–1758 (2009)

[11] Makrushin, A., Neubert, T., Dittmann, J.: Automatic generation and detection of visually faultless facial morphs. In: Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 6: VISAPP, (VISIGRAPP 2017). pp. 39–50. INSTICC, ScitePress (2017)

[12] Nguyen, A., Yosinski, J., Clune, J.: Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 427–436 (2015)
[13] Parkhi, O.M., Vedaldi, A., Zisserman, A.: Deep face recognition. In: British Machine Vision Conference (2015)
[14] Pérez, P., Gangnet, M., Blake, A.: Poisson image editing. In: ACM Transactions on Graphics (TOG). vol. 22, pp. 313–318. ACM (2003)
[15] R. Raghavendra , K. Raja, C. Busch: Detecting morphed facial images. Proceedings of 8th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS-2016) September 6-9, Niagra Falls, USA, (2016)
[16] Ralph Gross, Iain Matthews, Jeffrey Cohn, Takeo Kanade, Simon Baker: Multi-pie. In: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition. IEEE Computer Society (2008), https://www.microsoft.com/en-us/research/publication/multi-pie/
[17] Sandberg, D.: Tensorflow implementation of the facenet face recognizer. https://github.com/davidsandberg/facenet (2016)
[18] Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: A unified embedding for face recognition and clustering. CoRR abs/1503.03832 (2015), http://arxiv.org/abs/1503.03832
[19] Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.: Inception-v4, inception-resnet and the impact of residual connections on learning (2016), http://arxiv.org/pdf/1602.07261v2
[20] Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., Fergus, R.: Intriguing properties of neural networks (2013), http://arxiv.org/pdf/1312.6199v4
[21] Valente, J., Soatto, S.: Perspective distortion modeling, learning and compensation. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 9–16 (2015)
[22] Wen, Y., Zhang, K., Li, Z., Qiao, Y.: A discriminative feature learning approach for deep face recognition. In: Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VII. pp. 499–515 (2016), http://dx.doi.org/10.1007/978-3-319-46478-7_31
[23] Wolberg, G.: Digital image warping, vol. 10662. IEEE computer society press Los Alamitos, CA (1990)
[24] Zhang, K., Zhang, Z., Li, Z., Qiao, Y.: Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Processing Letters 23(10), 1499–1503 (Oct 2016)