Detection of morphed faces from single images: a multi-algorithm fusion approach

Ulrich Scherhag, Christian Rathgeb, Christoph Busch
da/sec - Biometrics and Internet Security Research Group
Hochschule Darmstadt, Germany
{ulrich.scherhag,christian.rathgeb,christoph.busch}@h-da.de

ABSTRACT

The vulnerability of face, fingerprint and iris recognition systems to attacks based on morphed biometric samples has been established in the recent past. However, so far a reliable detection of morphed biometric samples has remained an unsolved research challenge. In this work, we propose the first multi-algorithm fusion approach to detect morphed facial images. The FRGCv2 face database is used to create a set of 4,808 morphed and 2,210 bona fide face images which are divided into a training and test set. From a single cropped facial image features are extracted using four types of complementary feature extraction algorithms, including texture descriptors, keypoint extractors, gradient estimators and a deep learning-based method. By performing a score-level fusion of comparison scores obtained by four different types of feature extractors, a detection equal error rate (D-EER) of 2.8% is achieved. Compared to the best single algorithm approach achieving a D-EER of 5.5%, the D-EER of the proposed multi-algorithm fusion system is almost twice as low, confirming the soundness of the presented approach.

CCS CONCEPTS

• **Security and privacy → Biometrics**;

KEYWORDS

Biometrics, Morphing Detection, Algorithm Fusion

ACM Reference Format:
Ulrich Scherhag, Christian Rathgeb, Christoph Busch. 2018. Detection of morphed faces from single images: a multi-algorithm fusion approach. In *Proceedings of International Conference on Biometric Engineering and Applications (ICBEA'18).* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In past years, researchers have pointed out diverse potential vulnerabilities of biometric recognition systems. Proposed attacks, which aim at gaining unauthorized access to the system, can be coarsely categorized into presentation attacks and software-based attacks [1]. Presentation attacks refer to a presentation of an attack instrument (e.g. print outs or electronic displays [2]) to the biometric capture device with the goal of interfering with the operation of the biometric recognition system [3]. To launch software attacks, e.g. substitution attacks or overriding one of the inner modules of the system, an attacker requires knowledge about the inner modules of the biometric system together with access to some of the system components.

(a) Subject 1 **(b) Morph** **(c) Subject 2**

Figure 1: Examples for bona fide and morphed cropped face images

Recently, attacks on face and fingerprint recognition systems based on morphed biometric images and templates have been presented [4–7]. Morphing techniques can be used to create artificial biometric samples, which resemble the biometric information of two (or more) individuals in image and feature domain. If morphed biometric images or templates are infiltrated to a biometric recognition system the subjects contributing to the morphed image will both (or all) be successfully verified against that single enrolled template. Hence, the unique link between individuals and their biometric reference data is not warranted. Fig. 1 shows an example of morphing two faces in the image domain.

Such attacks pose severe security threats to biometric systems, in particular to the issuance and verification process of electronic travel documents [4]: black-listed criminal offenders can use an authentic passport complying with all physical safety features to enter a country with the identity of an accomplice when performing three basic steps: (1) find a rather lookalike accomplice, (2) morph passport face photos of both, possibly utilizing free software available on the internet, and (3) the accomplice applies for a passport; the passport manufacturer will then issue an authentic passport equipped with the morphed biometric reference image and other identity attributes of the accomplice, which can be used to enter a country by both subjects. Different commercial face recognition systems have been found to be highly vulnerable to this type of attack [4]. Due to a high intra-class variability in human faces, face recognition systems are operated at false match rates (FMRs) as high as 0.1% to achieve acceptable false non-match rates (FNMRs). That is, an automated detection of morphed face images is vital to retain the security of operational face recognition systems.

ICBEA'18, May 2018, Amsterdam, The Netherlands

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

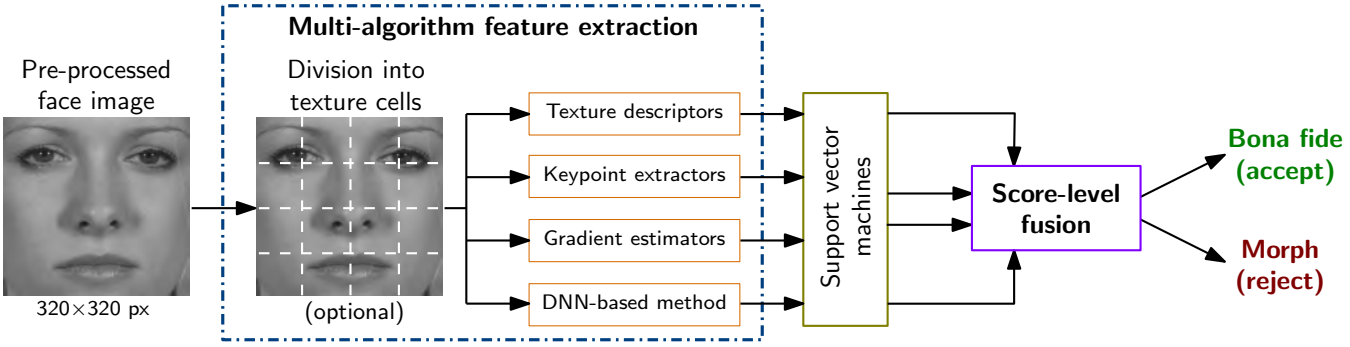


Figure 2: Overview of the proposed multi-algorithm fusion approach to detect morphed facial images

In this work, we conduct comprehensive evaluations on a dataset of 2,210 ICAO compliant face images of the FRGCv2 face database [8] from which we automatically generate a total of 4,808 morphed face images of high quality. It is demonstrated that a commercial of-the-shelf (COTS) face recognition system [9] is highly vulnerable to the above mentioned attack using the generated morphed face images. In order to prevent from such attacks we propose a multi-algorithm fusion approach to detect morphed face images. Four different types of feature extraction algorithms are employed: texture descriptors, keypoint extractors, gradient estimators and a deep learning-based method; it is shown that the detection performance can be substantially improved in a multi-algorithm score-level fusion, which maximizes the discriminativity of processed information. Compared to the best single algorithm approach achieving a D-EER of 5.5%, the proposed score-level fusion of comparison scores obtained by four different feature extractors yields a D-EER of 2.8%.

The remainder of this paper is organized as follows: related works are briefly summarized in Sect. 2. The proposed system is described in detail in Sect. 3. Experimental results are reported and discussed in Sect. 4. Finally, conclusions are drawn in Sect. 5.

2 RELATED WORK

Attacks based on morphed biometric samples were first introduced by Ferrara et al. [4]. Motivated by security gaps in the issuance process of electronic travel documents, the authors showed that commercial face recognition software tools are highly vulnerable to such attacks, i.e. different instances of images of either subject are successfully matched against the morphed image. In their experiments, decision thresholds yielding a FMR of 0.1% have been used, according to the guidelines provided by the European Agency for the Management of Operational Cooperation at the External Borders (FRONTEX) [10]. In a further study, the authors show that morphed face images are realistic enough to fool human examiners [11]. Scherhag et al.[5] reported moderate detection performance for benchmarking several general purpose texture descriptors used in conjunction with machine learning techniques to detect morphed face images. With respect to the above attack scenario, it is stressed that a detection of morphed face images becomes even more challenging if images are printed and scanned. Hildebrandt et al.[7] suggest to employ generic image forgery detection techniques, in particular multi-compression anomaly detection, to reliably detect morphed facial images. Kraetzer et al.[12] evaluate the feasibility of detecting facial morphs with keypoint descriptors and edge operators. The benefits of deep neural networks for detecting morphed images has been recently investigated by Ramachandra et al.[13]. Ferrara et al.[6] also presented two different methods to morph fingerprints in image and feature domain. For a decision threshold corresponding to a FMR of 0.1%, it is shown that commercial fingerprint recognition systems are also highly vulnerable to such attacks. Since fingerprint enrolment is usually done live in the issuance process of electronic travel documents, the authors argue that manufactured fake fingertips may be presented.

Gomez-Barrero et al.[14] proposed the first theoretical framework for measuring the vulnerability of biometric systems to attacks. Evaluations are conducted for diverse biometric systems where expected comparison scores of attacks based on morphed images or templates are directly derived from the mated and non-mated distributions of a face, fingerprint and iris recognition system. The authors identified key factors which take a major influence on a system's vulnerability to such attacks, e.g. the shape of genuine and impostor score distributions or the FMR the system is operated at. Since there is no standardised manner to evaluate the vulnerability of biometric systems to attacks based on morphed images or templates, Scherhag et al.[15] introduced new metrics for vulnerability reporting (see Sect. 4), which strongly relate to the metrics defined in [16]. In addition, the authors provide recommendations on the assessment of morphing techniques. It is emphasized that unrealistic assumptions with respect to the quality of morphed biometric samples might cloud the picture regarding the performance of detection algorithms. In summary, it becomes clear that research on attacks based on morphed biometric samples is still in statu nascendi. Nonetheless, at the time of this writing we see an increasing interest in this topic and the results of ongoing activities of different research labs are expected to be presented across diverse platforms in the near future. Eventually, it is important to note that so far there is no publicly available database of morphed face images and no publicly available morph detection algorithms.

(a) Subject 1 (b) Morph (c) Subject 2

Figure 3: Example for BSIF responses

(a) Subject 1 (b) Morph (c) Subject 2

Figure 4: Example for SURF keypoint detection

(a) Subject 1 (b) Morph (c) Subject 2

Figure 5: Example for sharpness features (two dimensions)

3 PROPOSED SYSTEM

The proposed system, which is depicted in Fig. 2, comprises three key modules, (1) face pre-processing, (2) multi-algorithm feature extraction and (3) score-level fusion and decision; in the following subsections, all modules are described in detail.

3.1 Face pre-processing

In the pre-processing stage the face of a subject is detected and normalized according to eye coordinates detected by the *dlib* landmark detector [17]. Subsequently, the normalized region is cropped to 320×320 pixels to ensure that the detection algorithm is only applied to the facial region. Finally, the cropped face part is converted to a grayscale image.

3.2 Multi-algorithm feature extraction

At feature extraction the pre-processed face image is optionally divided into multiple cells to retain local information. That is, feature extractors are applied separately on texture cells and the final feature vector is formed as a concatenation of feature vectors extracted from each cell. We employ the following four types of feature extraction methods, where up to two algorithms are considered per type:

(1) *Texture descriptors*: Local Binary Patterns (LBP) [18] and Binarized Statistical Image Features (BSIF) [19] are extracted from cropped face images. For details on these texture descriptors the reader is referred to [18, 19]. While LBP simply processes neighbouring pixel values of each pixel, BSIF utilizes specific filters learned from a set of images. Obtained feature values are stored in a corresponding histograms. The use of generic texture descriptors has shown to be successful in diverse texture classification problems. An example of BSIF applied to the images of Fig. 1 is depicted in Fig. 3. By testing different spatial sampling rates the best configuration for LBP and BSIF was determined, thus, a 3×3 LBP-patch and 5×5 BSIF filter set extracting 12 bit per pixel are employed.

(2) *Keypoint extractors*: Scale Invariant Feature Transform (SIFT) [20] and Speeded Up Robust Features (SURF) [21] extract sets of local keypoints. For details on keypoint detection, the extraction of keypoint descriptors and keypoint matching the reader is referred to [20, 21]. Keypoint extractors are employed, since morphed (averaged) images are expected to contain fewer key locations, which are defined as maxima and minima of the result of difference of Gaussians function. That is, the amount of detected keypoints is used as descriptive feature. Fig. 4 shows an example of SURF keypoints detected in the images of Fig. 1.

(3) *Gradient estimators*: Histogram of Gradients (HOG) and sharpness features are extracted from the normalized grayscale images. For further details to HOG the reader is referred to [22]. As a sharpness feature the mean of the gradient in two dimensions are calculated. The use of gradient-based methods is motivated by fact that due to the morphing process high frequency changes are reduced and, hence, the steepness of gradients is decreased. An example of sharpness features extracted from the images of Fig. 1 is depicted in Fig. 5.

(4) *Deep learning-based method*: we employ the *OpenFace* [23] algorithm in which rescaled images of 96×96 pixels are fed to the default pre-trained Deep Neural Network (DNN) to obtain a 128 dimensional face representation. This algorithm is applied to the pre-processed face image (no division into texture cells is applied). The use of Deep Facial Features (DFF) is motivated by recent advances in face recognition.

Since the above listed types of feature extraction techniques process images entirely different it is expected that they complement each other. Hence, it is expected that combinations of different types of feature extractor improve the performance of a detection subsystem in a multi-algorithm score-level fusion scenario.

Table 1: Training and test set used for experimental evaluations (a detailed list of used images will be published on the following website: currently-blinded-for-review)

	Training set				Test set			
Gender	No. of subjects	No. of images	Bona fide images	Morphed images	No. of subjects	No. of images	Bona fide images	Morphed images
Male	59	2,819	1,166*	1,653	58	2,210	499	1,711
Female	40	2,073	1,332*	741	39	1,165	462	703
All	99	4,892	2,498*	2,394	97	3,375	961	2,414

* in the training set bona fide images are horizontally mirrored

3.3 Score-level fusion and decision

In the training stage feature vectors are extracted for each algorithm and support vector machines (SVMs) are trained to distinguish between bona fide and morphed face images using a disjoint training set. For a given face image the SVMs of each single algorithm generate a normalized attack detection score in the range $[0, 1]$. In the fusion stage the sum-rule fusion of normalized scores is applied. In the context of biometric fusion, score-level fusion using the sum-rule with proper normalization has been observed to result in competitive biometric performance [24].

4 EXPERIMENTS

In the following subsections, we describe the experimental setup, conduct a vulnerability assessment of a COTS face recognition system to attacks based on the generated morphed face images and report and discuss the detection performance of proposed system.

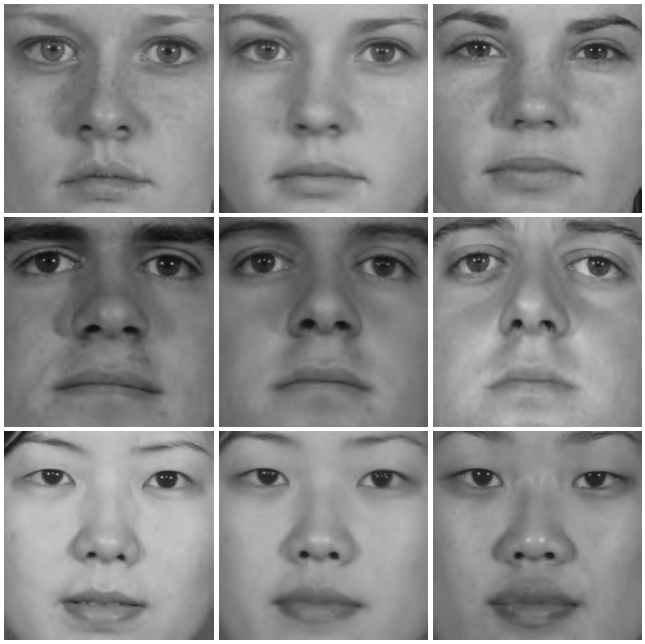
4.1 Experimental setup

Experiments are performed on a subset of the FRGCv2 face database. A total number of 2,210 frontal faces with neutral expression have been manually chosen and ICAO compliance has been verified, i.e. the distance between the eyes of a face has to be at least 90 pixels [25]. Based on this subset 4,808 morphed faces have been automatically generated for pairs of subjects of same gender using the *OpenCV* library. Further example images of bona fide and morphed face images are shown in Fig. 6. The division of images into training and test sets which has been chosen to obtain a balance between bona fide and morphed images during training is listed in Table 1.

4.2 Vulnerability assessment

The vulnerability of a COTS face recognition system to attacks based on the generated morphed face images is assessed according to the metrics specified in [15], in particular, in terms of Mated Morph Presentation Match Rate (MMPMR). This metric is an adaptation of the general Impostor Attack Presentation Match Rate (IAPMR) introduced in ISO/IEC 30107-3 [16] which is defined as the proportion of attack presentations using the same presentation attack instrument species in which the target reference is matched. However, in the adaptation the MMPMR covers the fact that not one target subject (contained in the morphed reference) is matched - but both subjects who earlier contributed to the morphed image are expected to be matched if the morphing attack is considered to be successful.

When employing the default decision threshold of the COTS face recognition system a MMPMR of 1 is obtained. This means all face images of subjects contributing to a morphed face image are successfully matched against it, hence, the attacks reveal a success chance of 100%.



(a) Subject 1 (b) Morph (c) Subject 2

Figure 6: Examples of bona fide and morphed face images of subjects of same gender, ethnicity and age group

4.3 Performance evaluation

The performance of the detection algorithms used in this work is reported according to metrics defined in ISO/IEC 30107-3 [16]. The bona fide Presentation Classification Error Rate (BPCER) is defined as the proportion of bona fide presentations incorrectly classified as attack presentations in a specific scenario. The Attack Presentation Classification Error Rate (APCER) is defined as the proportion of attack presentations using the same presentation attack instrument species incorrectly classified as bona fide presentations in a specific scenario. Further, the BPCER-10 and BPCER-20 represent the operation points yielding an APCER of 10% and 5%, respectively. Additionally, to be comparable to published works, the Detection Equal Error Rate (D-EER) will be reported.

Performance rates of the best two configurations per algorithm are listed in Table 2. The corresponding detection error trade-off

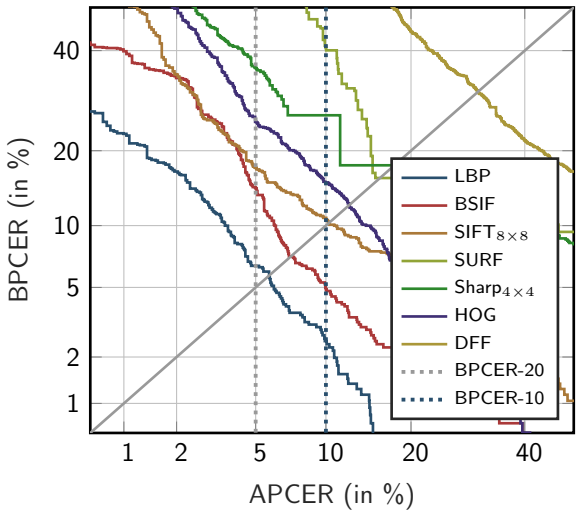


Figure 7: DET-plot of single algorithms

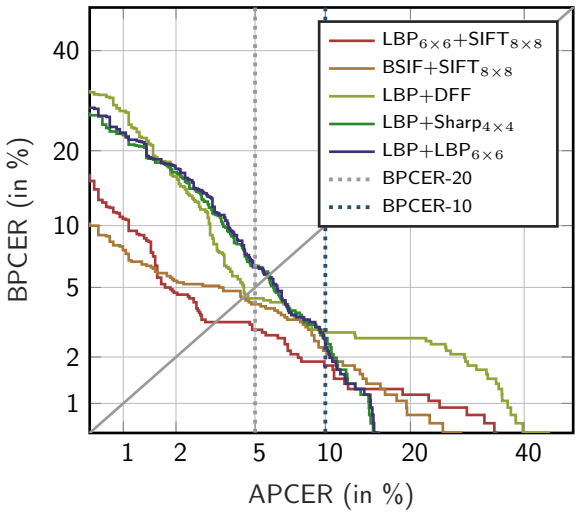


Figure 8: DET-plot of fusions of two algorithms

Table 2: Single algorithm performance

Algorithm	D-EER	BPCER-10	BPCER-20
LBP	5.5 %	2.2 %	6.2 %
$LBP_{6\times6}$	5.6 %	2.1 %	6.4 %
BSIF	7.1 %	4.8 %	13.8 %
$BSIF_{2\times2}$	7.9 %	6.0 %	14.8 %
$SIFT_{8\times8}$	10.2 %	10.4 %	17.1 %
$SIFT_{10\times10}$	10.9 %	11.8 %	18.7 %
SURF	15.7 %	40.1 %	68.6 %
$SURF_{2\times2}$	18.9 %	29.4 %	54.2 %
$Sharp_{4\times4}$	16.4 %	20.4 %	35.4 %
$Sharp_{3\times3}$	19.4 %	44.1 %	70.3 %
HOG	12.1 %	15.0 %	24.9 %
$HOG_{3\times3}$	26.3 %	61.6 %	78.6 %
DFF	30.6 %	64.4 %	78.4 %

Table 3: Performance of fusions of two algorithms

Rank	Algorithm	D-EER	BPCER-10	BPCER-20
1	$LBP_{6\times6}$ + $SIFT_{8\times8}$	3.1 %	1.6 %	2.7 %
2	BSIF + $SIFT_{8\times8}$	4.3 %	2.1 %	4.0 %
3	LBP + DFF	4.5 %	2.8 %	4.2 %
4	LBP + $Sharp_{4\times4}$	5.5 %	2.3 %	6.2 %
5	LBP + $LBP_{6\times6}$	5.5 %	2.2 %	6.2 %

(DET) curves are depicted in Fig. 7. Optional divisions of the pre-processed face image into texture cells are indicated accordingly. Competitive detection rates are achieved for texture descriptors where LBP achieves the best performance of D-EER=5.5%. Moderate detection accuracy is achieved for keypoint extractors and gradient estimators. Applying the default net which is designed for recognition purposes DFF reveals the highest D-EER. However, it is expected that application-specific training will significantly improve deep learning-based approaches with the potential drawback of data-overfitting. The above listed configurations of algorithms will be used in all fusion experiments and the five best performing combinations of algorithms will be reported.

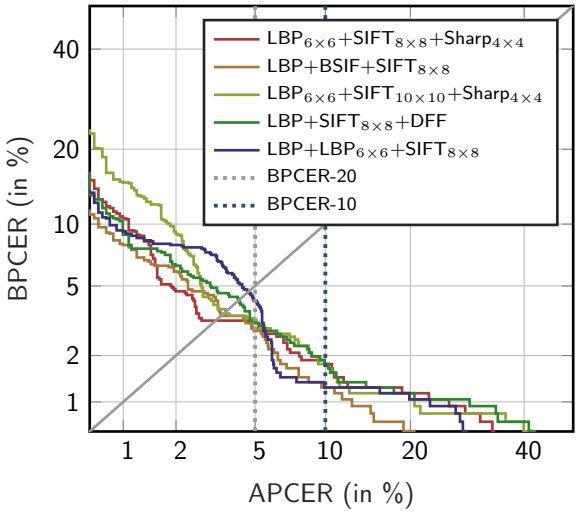


Figure 9: DET-plot of fusions of three algorithms

Detection performance rates of combinations of two algorithms are shown in Table 3 and according DET curves are plotted in Fig. 8. We observe that a combination of different types of feature extractors significantly improves D-EERs. This is emphasized by the observation that a combination of the worst performing DFF and LBP is ranked third of all combinations of pairs of feature extractors. Note that, a combination of LBP and $LBP_{6\times6}$ leads to the same detection performance as the single use of LBP. Table 4 summarized the five best performing combinations of three algorithms. The corresponding DET plot is shown in Fig. 9. Slight performance gains in terms of D-EER are observed compared to combinations of two algorithms. Again, a combination of three different types of algorithms obtains the best detection performance. Interestingly, the fusion ranked second combines both texture descriptors which suggests that the learned filters of BSIF complement LBP. It is important to note that improvements in terms of BPCER-10 and BPCER-20 are more pronounced compared to combinations of two algorithms.

Table 4: Performance of fusions of three algorithms

Rank	Algorithm	D-EER	BPCER-10	BPCER-20
1	$LBP_{6\times6}$ + $SIFT_{8\times8}$ + $Sharp_{4\times4}$	3.1 %	1.6 %	2.6 %
2	LBP + BSIF + $SIFT_{8\times8}$	3.4 %	1.2 %	2.6 %
3	$LBP_{6\times6}$ + $SIFT_{10\times10}$ + $Sharp_{4\times4}$	3.6 %	1.6 %	3.0 %
4	LBP + $SIFT_{8\times8}$ + DFF	4.0 %	1.7 %	3.1 %
5	LBP + $LBP_{6\times6}$ + $SIFT_{8\times8}$	4.2 %	1.2 %	3.2 %

Table 5: Performance of fusions of four algorithms

Rank	Algorithm	D-EER	BPCER-10	BPCER-20
1	LBP + $BSIF_{2\times2}$ + $SIFT_{8\times8}$ + DFF	2.8 %	0.7 %	1.8 %
2	LBP + $LBP_{6\times6}$ + $SIFT_{8\times8}$ + $SIFT_{10\times10}$	3.0 %	1.3 %	2.2 %
3	LBP + BSIF + $SIFT_{8\times8}$ + $Sharp_{4\times4}$	3.5 %	1.2 %	2.6 %
4	LBP + $SIFT_{8\times8}$ + $SIFT_{10\times10}$ + HOG	3.9 %	2.1 %	3.3 %
5	LBP + $SIFT_{8\times8}$ + $Sharp_{4\times4}$ + DFF	3.9 %	1.6 %	3.1 %

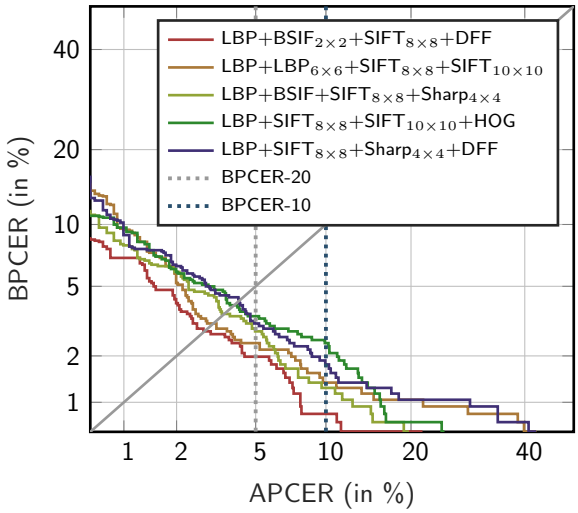


Figure 10: DET-plot of fusions of four algorithms

Also, characteristics of DET curves suggest more robustness (less BPCER oscillation across operation points) for combinations of three algorithms. Finally, performance rates of combinations of four algorithms are listed in Table 5 and resulting DET plots are shown in Fig. 10. Additional reductions of error rates are obtained in particular, in terms of BPCER-10 and BPCER-20. Compared to the best single algorithm approach the best combination of four algorithms, which again combines both employed texture descriptors, almost halves the D-EER to 2.8% and reduces the BPCER-10 and BPCER-20 to 0.7% and 1.8%, respectively.

5 CONCLUSION AND FUTURE WORK

To the authors' knowledge this work presents the first multi-algorithm fusion approach to detect morphed face images. It is shown that substantial improvements in detection performance can be achieved when different types of feature extractors are combined employing a normalized score-level fusion based on the sum-rule. Based on a subset of the publicly available FRGCv2 face database morphed face images are generated and it is shown that a COTS face recognition system is highly vulnerable to attacks based on those images. With the presented multi-algorithm fusion approach a practical D-EER of 2.8% is achieved. In the proposed system the combination of multiple algorithms can be parallelized easily in order to maintain computational workload. In general a detection mechanism relying on numerous algorithms can be expected to be exhibit high robustness. Nonetheless, transferability of the presented approach with respect to different datasets and morphing techniques yet needs to be analysed in more extensive future studies. Moreover, the potential of a weighted fusion of obtained detection scores as well as an analysis of further classifiers, e.g. random forest, could be subject to future work.

ACKNOWLEDGEMENT

This work was supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within CRISP (www.crisp-da.de).

REFERENCES

[1] N. K. Ratha, J. H. Connell, and R. M. Bolle. 2001. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal* 40, 3 (2001), 614–634.
[2] S. Marcel, M. Nixon, and S. Z. Li. 2014. *Handbook of Biometric Anti-Spoofing*. Springer-Verlag New York, Inc.
[3] ISO/IEC TC JTC1 SC37 Biometrics. 2016. *ISO/IEC IS 30107-1. Information Technology – Biometrics presentation attack detection – Part 1: Framework*. International Organization for Standardization.
[4] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. 2014. The Magic Passport. In *Proc. Int. Joint Conf. on Biometrics (IJCB)*. 1–7.
[5] U. Scherhag, R. Raghavendra, K. B. Raja, M. Gomez-Barrero, C. Rathgeb, and C. Busch. 2017. On the Vulnerability of Face Recognition Systems Towards Morphed Face Attacks. In *Proc. Int. Workshop on Biometrics and Forensics (IWBF)*. 1–6.
[6] M. Ferrara, R. Cappelli, and D. Maltoni. 2017. On the Feasibility of Creating Double-Identity Fingerprints. *IEEE Trans. on Information Forensics and Security* 12, 4 (2017), 892–900.
[7] M. Hildebrandt, T. Neubert, A. Makrushin, and J. Dittmann. 2017. Benchmarking face morphing forgery detection: Application of stirtrace for impact simulation of different processing steps. In *Proc. Int. Workshop on Biometrics and Forensics (IWBF)*. 1–6.
[8] P. Jonathon Phillips, Patrick J. Flynn, Todd Scruggs, Kevin W. Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, and William Worek. 2005. Overview of the Face Recognition Grand Challenge. In *Proc. of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (CVPR '05)*. 947–954.
[9] Cognitec. [n. d.]. FaceVACS-SDK. http://www.cognitec.com. ([n. d.]). http://www.cognitec.com Accessed: 2017-04-28.

[10] 2012. FRONTEX – Research and Development Unit: Best Practice Technical Guidelines for Automated Border Control (ABC) Systems. (2012). Version 2.0.
[11] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. 2016. On the Effects of Image Alterations on Face Recognition Accuracy. In *Face Recognition Across the Imaging Spectrum*, Thirimachos Bourlai (Ed.). Springer International Publishing, 195–222.
[12] Christian Kraetzer, Andrey Makrushin, Tom Neubert, Mario Hildebrandt, and Jana Dittmann. 2017. Modeling Attacks on Photo-ID Documents and Applying Media Forensics for the Detection of Facial Morphing. In *Proc. Workshop on Information Hiding and Multimedia Security (IH& MMSec)*. 21–32.
[13] R. Ramachandra, K. Raja, S. Venkatesh, and C. Busch. 2017. Transferable Deep-CNN features for detecting digital and print-scanned morphed face images. In *2017 IEEE Conf. on Computer Vision and Pattern Recognition Workshop (CVPRW)*.
[14] M. Gomez-Barrero, C. Rathgeb, U. Scherhag, and C. Busch. 2017. Is Your Biometric System Robust to Morphing Attacks?. In *Proc. Int. Workshop on Biometrics and Forensics (IWBF)*. 1–6.
[15] Ulrich Scherhag, Andreas Nautsch, Christian Rathgeb, Marta Gomez-Barrero, Raymond Veldhuis, Luuk Spreeuwers, Maikel Schils, Davide Maltoni, Patrick Grother, Sebastien Marcel, Ralph Breithaupt, R. Raghavendra, and Christoph Busch. 2017. Biometric Systems under Morphing Attacks: Assessment of Morphing Techniques and Vulnerability Reporting. In *Int. Conf. of the Biometrics Special Interest Group (BIOSIG)*. 1–12.
[16] ISO/IEC JTC1 SC37 Biometrics. [n. d.]. *ISO/IEC IS 30107-3:2017, IT – Biometric presentation attack detection – Part 3: Testing and Reporting*.
[17] Davis E. King. 2009. Dlib-ml: A Machine Learning Toolkit. *Journal of Machine Learning Research* 10 (2009).
[18] Shengcai Liao, Xiangxin Zhu, Zhen Lei, Lun Zhang, and Stan Z. Li. 2007. Learning Multi-scale Block Local Binary Patterns for Face Recognition. In *Proc. Int'l Conference on Biometrics (ICB'07)*. 828–837.
[19] J. Kannala and E. Rahtu. 2012. BSIF: Binarized statistical image features. In *Proc. Int'l Conf. on Pattern Recognition (ICPR'12)*. 1363–1366.
[20] David G. Lowe. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vision* 60, 2 (2004), 91–110.
[21] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. 2008. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding* 110, 3 (2008), 346 – 359.
[22] Chang Shu, Xiaoqing Ding, and Chi Fang. 2011. Histogram of the oriented gradient for face recognition. *Tsinghua Science & Technology* 16, 2 (2011), 216–224.
[23] Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. 2016. *OpenFace: A general-purpose face recognition library with mobile applications*. Technical Report. CMU School of Computer Science.
[24] Anil K. Jain, B. Klare, and Arun A. Ross. 2015. Guidelines for Best Practices in Biometrics Research. In *In Proc. Int'l Conf. on Biometrics (ICB'15)*. 1–5.
[25] International Organization for Standardization. 2005. *Information technology – Biometric data interchange formats – Part 5: Face image data*. ISO/IEC 19794-5:2005 consolidated. JTC 1/SC 37.