Single Image Face Morphing Attack Detection Using Ensemble of Features

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

***Abstract*—Face morphing attacks have demonstrated a severe threat in the passport issuance protocol that weakens the border control operations. A morphed face images if used after printing and scanning (re-digitizing) to obtain a passport is very challenging to be detected as attack. In this paper, we present a novel method to detect such morphing attacks using an ensemble of features computed on the scale-space representation derived from the color space for a given image. Given the limited availability of datasets representing realistic morphing attacks, we introduce and present a new print-scan image dataset of morphed face images. Experiments are carried out on the two different datasets and compared with sixteen existing state-of-art Morphing Attack Detection (MAD) mechanism based on single image MAD (S-MAD). The proposed approach indicates a superior MAD performance on both datasets suggesting the applicability in operational scenarios.**

***Index Terms*—Biometrics, Face Recognition, Face morphing, Attacks, Vulnerability of Biometric Systems, Machine learning**

I. Introduction

Face biometrics is widely deployed in secure border control applications, where the identity of a person is verified either by an electronic passport or a national identity card. While the face picture for the passport is captured in a few countries under controlled conditions inside a trusted authoritative unit (e.g. Police Station), for the majority of countries the applicant is asked to submit a face image. The applicant can therefore provide any face picture that can resemble the applicant to higher degree through morphing techniques. Morphing techniques seamlessly transform one image contents to another image with high degree of resemblance to challenge the Face Recognition Systems (FRS). Morphed images (as shown in Figure 1) can be used to verify two or more identities with one single morphed reference image as demonstrated in earlier works [3], [4], [12]. Not only do such morphed images bypass the FRS, but also fool the human observers including trained border guards [4] posing a severe threat to border control processes. This situation makes morphing attacks a relevant risk that constitutes a significant challenge [3].

Morphing Attack Detection (MAD) has been studied for the past couple of years resulting in various algorithms that can be broadly divided in two types [16]: (1) Single Image Morphing Attack Detection (S-MAD) techniques (a.k.a as No-Reference MAD) and (2) Differential Morphing Attack Detection (D-MAD) techniques. Among these two types, the S-MAD is more challenging as the decision needs to be taken on ae single image without a trusted image available for the same subject. Added to the magnitude of challenge, reliable detection mechanisms should address not only digital data formats, but also the print-scanned (re-digitized) data formats, where the inherent/residual digital information of morphing is lost in the process of printing and scanning.

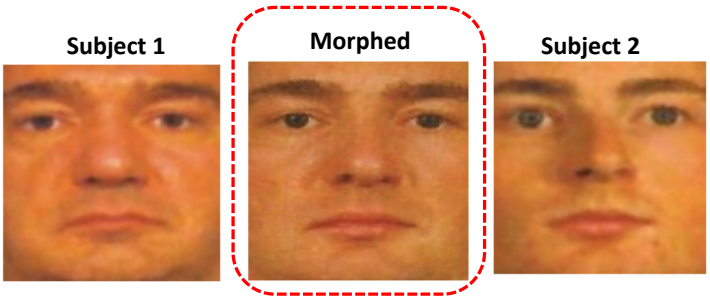

Fig. 1: Example of the face morphing

The digital format of morphing image is expected to contain residues of the morphing process and thus, most of the state-of-the-art MAD techniques are designed to capture these effects, for instance PRNU [2]. The majority of MAD algorithms are also limited to operate on the digital data format of morphing images [5], [9], [10], [13], [19]. The popular methods in this direction includes texture based schemes like: Local Binary Patterns (LBP) [13], Binarized Statistical image features (BSIF) [13], Frequency features [7], Image degradation measure features using StirTrace algorithm [5], JPEG compression artefacts [10], PRNU [2], Benford's features [9], Specular reflection [19].

Considering the fact that, most countries use print-scanned face image for issuing passport, very little focus has been given to detect such morphed images. In fact, the use of print-scanned morphed images is commonly encountered in real-life passport application as the applicant will submit the printed photo which is then digitized using a scanning process and stored in the passport. It has to be noted that detecting a printed-scanned face morph image is very challenging as the print-scan process also introduces additional noise. The commonly employed approaches to detect morphing attacks after printing and scanning include the texture-based approaches Local Binary Patterns (LBP) [17], Binarized Statistical Image Features (BSIF)

[13], color textures, deep learning approaches [14], [20], high frequency texture features [15], [16].

A set of other works in this direction have also explored multiple feature extraction techniques to detect morphing attacks. In [14], an approach based on feature fusion from pre-trained deep learning networks (AlexNet and VGG-19) is presented. Experiments presented on both digital and print-scanned dataset indicate the reliable detection of morph attacks. In [12] a hybrid feature fusion was presented for reliable face morphing attack detection. A framework proposed in [11], explores the StirTrace based approach that detects the Face Morphing Forgery (FMF) by using keypoint features, Benford features and fusion of both keypoint features and Benford features. In [18], performance variation and robustness of various morph detection algorithms on different datasets are studied. For the detection of morphs, facial images are pre-processed then features are extracted. For this the facial image is divided into 4×4 cells, then the textural features are extracted individually and further fused to obtain a final feature vector. In addition to this, keypoint extractors and gradient estimators are employed and finally compared using SVM. In [16] , the authors present a technique using color space features, where scale space texture features are extracted and classified using spectral regression classifier. The comparison score level fusion is finally carried out to detect the morphing attacks. Dempster-Shafer theory for morph attack detection was also explored for detecting morphing attacks [8] where individual morph attack detectors were combined using Dempster-Shafer theory to improve the reliability of face morphing attacks [8]. The results obtained by employing this method indicates the improved detection accuracy using multiple detectors rather than individual detectors. Another approach in this direction also employed multi-detector fusion approach [1]. It extracts both hand-crafted features using Local Binary Pattern Histogram (LBPH) and CNN based features. Further both feature types are combined using feature level fusion after z-score normalization.

The reported results indicate still high error rates, which further exemplifies the difficulty in detecting a single image morphed face image after print-scan process. Further, all the reported results are presented only on one source of the printer. Thus the generalization of the state-of-the-art techniques is not evaluated for multiple printers and scanners. In this work, we present a novel scheme based on an ensemble of features that are classified individually using Collaborative Representative Classifier (CRC) to detect reliably with an S-MAD approach a morphed attacks even after the print-scan process. We assert that the use of the multiple features can provide a complementary feature set that can be used to detect the morphed face images efficiently. Motivated by this, we explore multiple features in an ensemble classifier approach to detect morphing attacks in this work resulting in the contributions as follows: (1) Presenting a novel method based on an ensemble of features to detect based on a single image a morphing attack after print-scan process. (2) Presenting a new dataset with high-quality print-scan morphed face images to evaluate multiple state-of-art MAD mechanisms. (3) Reporting an extensive set of results on two different datasets (including the newly introduced dataset) that are generated using two different print-scan process. Each of these datasets is comprised of 1309 bona fide face images and 2608 morphed face images. (4) The performance of the proposed method is benchmarked with 16 different state-of-the-art techniques using the IS0/IEC 30107-3 [6] metrics with Bona fide Presentation Classification Error Rate (BPCER) computed at Attack Presentation Classification Error Rate (APCER) @5% and @10% together with Detection-Equal Error Rate (D-EER%).

The rest of the paper is organised as follows: the proposed method is presented in Section II, discussion on experimental results are presented in Section III and Section IV draws the conclusion.

II. Proposed Face Morphing Attack Detection Technique

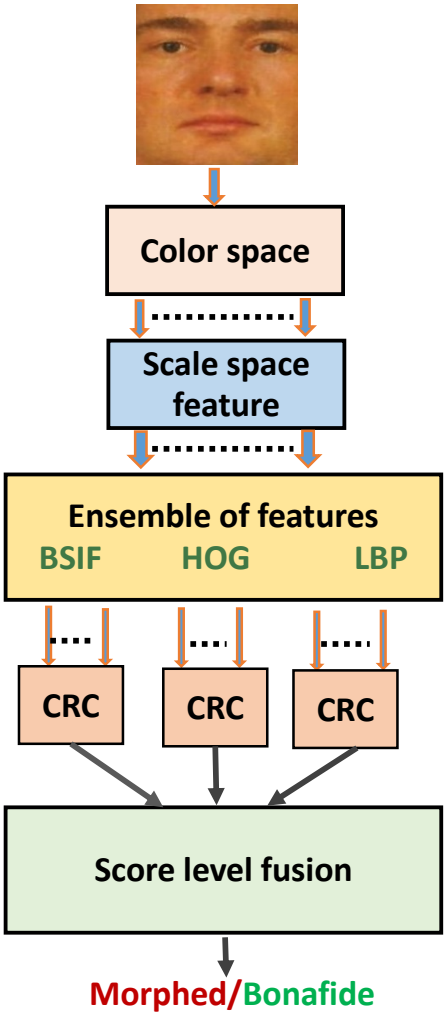

Fig. 2: Block diagram of the proposed method

Figure 2 illustrates the block diagram of the proposed approach for single image morphed face attack detection (S-MAD). The proposed method is structured using five main

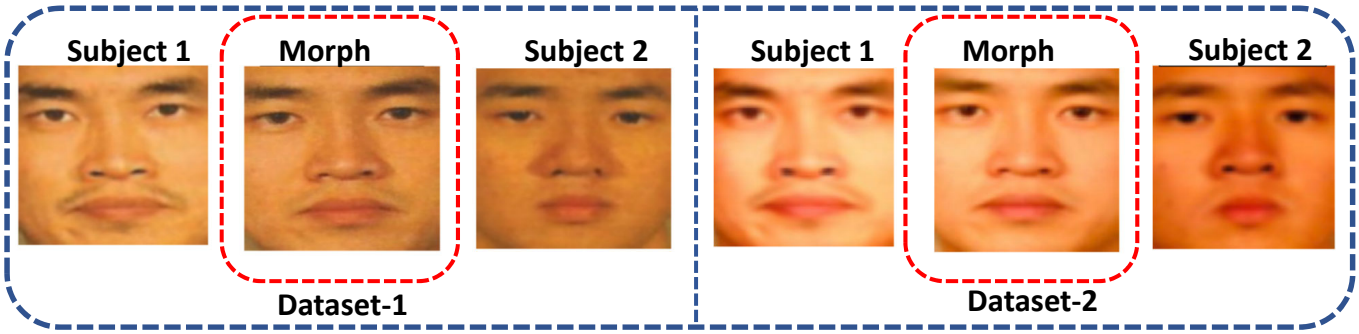

Fig. 3: Illustration of the example images (a) Dataset-1 (b) Dataset-2. The difference in quality of images across both datasets can be observed in the illustration.

functional units through which the face image is processed before taking the final decision. Given the face image I, the first step is to represent the I using two different color spaces such as YCbCr and HSV. We are motivated to use these two color spaces as earlier work [15] indicated that the use of multiple color spaces can enhance the MAD accuracy by providing complementary information. Let the extracted color space images be represented as $Ic_i, \forall i = 1, \ldots, 6$. In the second step, we extract the scale-space features that can capture the high-frequency components from each of the color space image Ic_i. In this work, the scale-space features are extracted using a Laplacian pyramid with 3 level decomposition. We are motivated to employ the Laplacian transform since they have indicated a robust (or saliency) features useful for MAD [16]. Let the scale spaces be presented as Ic_i^j $\forall j = 1, 2, 3 \& \forall i = 1, \ldots, 6$. In the next step, we perform the feature extraction on individual scale-space images Ic_i^j using multiple feature extraction techniques. As the use of more than one feature space can provide complementary features, we are motivated to use three different feature extraction techniques namely: Local Binary Patterns (LBP), Histogram of gradients (HOG) and Binarized Statistical Image Features (BSIF). These three different features are selected as morphing residues can be detected based on the local (LBP) and global (BSIF) texture features together with the pixel gradients. Thus, it is our assertion that the use of these three feature extraction techniques can capture complementary residual features that in turn can be used to detect a morphed face image. The LBP features are extracted from Ic_i^j using an image block of size 20×20 pixels with 10 pixel overlapping, the BSIF features are extracted using a filter size of 15×15 with 12 bit length which are determined empirically. Let features extracted using LBP, BSIF and HOG be denoted as: LIc_i^j, BIc_i^j and HIc_i^j respectively.

In the next step, we perform the classification of features independently to obtain the morphing scores. In this work, we employed the Probabilistic Collaborative Representation Classifier (P-CRC), which maximizes the likelihood ratio of a test sample jointly with other classes to perform the classification [21]. The P-CRC used in this work utilizes the Regularized Least Square Regression (LSR) on the learned feature vectors versus the probe feature vectors [21] formulated as:

$$\widehat{F} = argmin_{\alpha} \left\| Tr_F - D\alpha \right\|_2^2 + \lambda \left\| \alpha \right\|_2^2 \tag{1}$$

where the Tr_F is the feature vector of the probe image, D is the learned collaborative subspace dictionary using Tr_F, α is coefficient vector and λ is the regularization parameter. Let the morphing score corresponding to LIc_i^j, BIc_i^j and HIc_i^j be $SLIc_i^j$, $SBIc_i^j$ and $SHIc_i^j$ respectively.

Finally, the morphing scores obtained from P-CRC are combined using the SUM rule to compute the final score F as: $F = SLIc_i^j + SBIc_i^j + SHIc_i^j$ on which the final decision is made to accept it as bona fide or morphed image.

III. Experiments and Results

In this section, we present the experiments and results of the proposed scheme on two different datasets. We also present a comparative analysis by benchmarking 16 different state-of-the-art techniques on both datasets. Further, all results of MAD algorithms are presented following the IS0/IEC 30107-3 [6] metrics: *Bona fide Presentation Classification Error Rate (BPCER) and Attack Presentation Classification Error Rate (APCER)*. **BPCER** is defined as the proportion of bona fide presentations incorrectly classified as attacks while **APCER** is defined as the proportion of attack presentations incorrectly classified as bona fide presentations. In particular, we report the performance of the proposed method by reporting the value of BPCER while fixing the APCER to 5% and 10% according to the recommendation from IS0/IEC 30107-3 [6]. Besides, we also present the results with Detection-Equal Error Rate (D-EER%) where $BPCER = APCER$. Extensive experiments are carried out on the two different datasets such as *Dataset-1:* This is the private database [15] with 1309 bona fide images and 2608 morphed face images. This dataset is collected using RICOH office printer following the procedure mentioned in [15]. *Dataset-2:* This a new dataset collected in this work using high-quality photo printer (Epson expression photo XP860). In order to have a fair comparison, we have used the same

TABLE I: Quantitative results of the state-of-the-art and proposed method

Algorithm	Dataset-1			Dataset-2		
	D-EER(%)	BPCER @ APCER		D-EER(%)	BPCER @ APCER	
		=5%	=10%		=5%	=10%
Deep Learning Based Approaches						
AlexNet	12.64	59.66	29.50	40.02	83.50	74.83
DenseNet	12.99	70.83	35.83	15.82	33.00	22.16
GoolgleLeNet	12.99	72.50	48.50	19.34	44.00	31.16
InceptionV3	13.03	75.16	43.83	21.06	58.00	42.50
ResNet50	12.05	54.33	20.83	20.56	53.83	41.66
ResNet101	13.35	78.33	43.16	15.31	38.33	25.66
VGG16	12.49	49.50	22.16	16.33	40.83	24.83
VGG19	13.31	71.50	45.33	13.51	27.83	13.51
Non-Deep Learning Based Approaches						
BSIF-SVM	14.21	96.50	87.16	23.34	50.33	41.16
Steerable Textures	11.66	48.33	16.50	31.49	76.66	63.83
Hybrid textures	7.47	12.00	4.83	9.32	14.33	8.66
HoG-SVM	13.85	87.50	63.00	18.48	35.00	25.33
IG-SVM	29.29	67.33	59.83	34.19	78.33	64.33
Color Textures	14.01	65.50	39.66	18.71	44.83	30.50
LBP-SVM	13.47	80.66	52.16	35.01	84.50	70.00
LPQ-SVM	14.13	88.88	76.33	27.65	80.83	67.00
Proposed Method	**5.99**	**8.17**	**3.83**	**5.64**	**6.34**	**3.77**

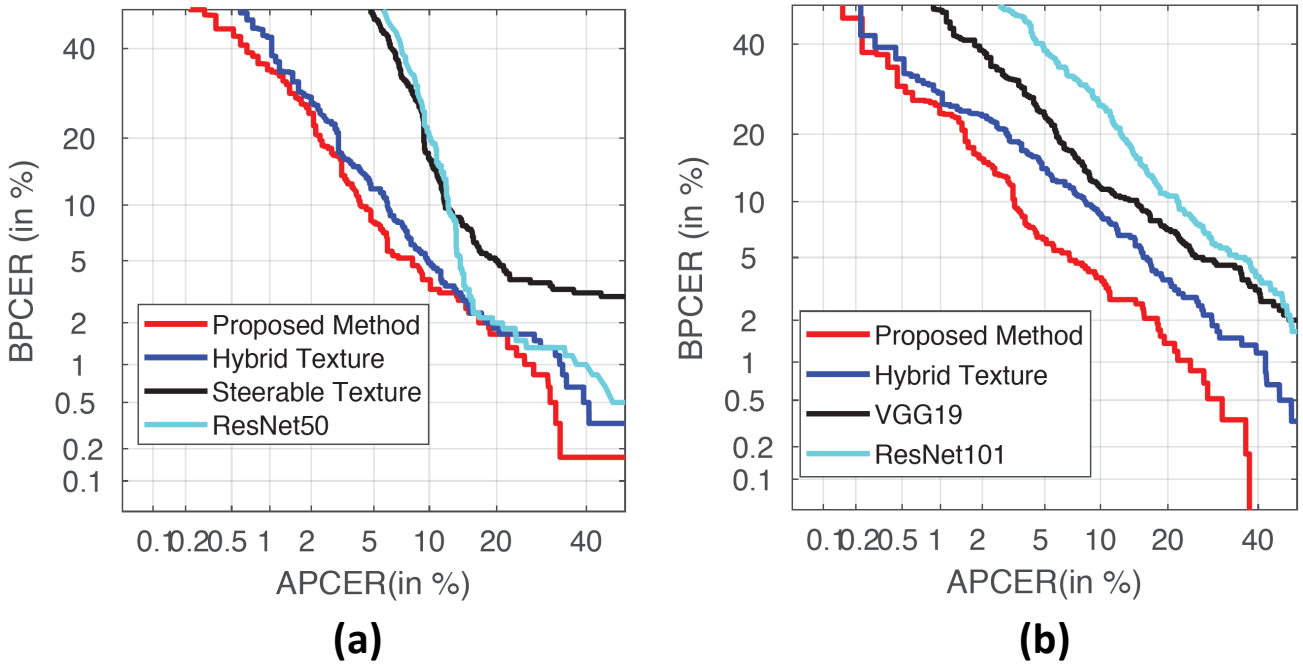

Fig. 4: DET curves on (a) Dataset-1 (b) Dataset-2

images that are used to generate Dataset-1 to generate a new dataset. Thus, this dataset also has 1309 bona fide and 2608 morphed face images. This is one of the largest semi-public database containing print and scanned morphed faces available for academic research purposes. Figure 3 shows the example of the bona fide and morphed images from both datasets.

Performance evaluation protocol: In order to effectively benchmark the performance of algorithms, we divide the whole dataset into two independent parts: Training set with 709 bona fide and 1255 morphed images. Testing set with 600 bona fide and 1353 morphed images. The disjoint partition is made based on the individual subjects as mentioned in [13] where the subjects that are contained in training set are not present in testing set. We have followed the same procedure for both datasets to make sure that the same images are in the training and testing set.

Table I indicates the quantitative performance of the proposed method together with 16 different state-of-the-art methods on both Dataset-1 and Dataset-2. The following are the important observations:

- In general, the performance of evaluated algorithms is degraded on the Dataset-2 when compared to that of Dataset-1. This certainly indicates that with newly introduced Dataset-2 it is more challenging to detect the morphed image as a result of high quality print-scan process.
- Among the available state-of-the-art pre-trained deep learning techniques [15] [20] [14] evaluated on both Dataset-1 and Dataset-2, the obtained performance is highly similar. The limited performance of the pre-trained deep learning networks can be attributed to the use of a small-scale dataset in training robust networks.
- Among the available state-of-the-art non-deep learning techniques, the recent methods based on Steerable Textures [15] and Hybrid textures [16] indicate a good performance on Dataset-1. However, the performance of these techniques degrades on the Dataset-2, indicating the poor generalization capability of previously reported approaches.
- The proposed method has indicated the best but not ideal performance on both datasets. The proposed method shows the performance of D-EER(%) = 5.99% with BPCER = 8.17% @ APCER = 5% and BPCER = 5.64% @ APCER = 10% on Dataset-1. While on Dataset-2, the proposed method has indicated a performance of D-EER(%) = 5.64% with BPCER = 6.34% @ APCER = 5% and BPCER = 3.77% @ APCER = 10%. It is interesting to note that, the detection performance of the proposed method is consistent across both datasets. This can be attributed to the complementary features extracted using the proposed approach within ensemble of features.
- Figure 4 shows the DET curves of four different methods (for the sake of simplicity) on Dataset-1 and Dataset-2. It is interesting to observe the improved performance of the proposed scheme on both datasets that can be attributed to the ensemble learning of multiple features and classifier.

IV. Conclusion

Morphed face detection in a real-life scenario with no reference and only a single morphed face image, which further has been print-scanned, remains a challenging task. In this work, we have proposed a novel scheme to reliably detect print-scanned morphed face images using an ensemble of features in a collaborative manner. Given the image, the proposed method first extracts the two different color spaces. In the next step, a scale-space representation using Laplacian transform with 2 level decomposition is performed on each of these color images to capture the high-frequency features. We then use the ensemble of features such as Local Binary Patterns (LBP), Histogram of gradients (HOG) and Binarized Statistical Image Features (BSIF). The ensemble of features is extracted independently from every high-frequency image that is in turn provided to the P-CRC classifier to obtain the individual morphing scores. Finally, the individual morphing scores are fused using a simple sum rule to make the final decision on morphing attack. Further, we have also introduced a new morphed face dataset with high-quality print-scan images that is more challenging to detect. Extensive experiments are performed on two different morphed face image dataset (including the newly introduced dataset) reflects two different print-scan process to study the scalability of previously published MAD mechanisms. The detection performance of the proposed method is benchmarked with 16 different state-of-the-art methods that include both deep learning and non-deep learning methods. The proposed method has indicated the best performance with D-EER (%) = 5.99% with BPCER = 8.17% @ APCER = 5% and BPCER = 5.64% @ APCER = 10% on Dataset-1 and D-EER(%) = 5.64% with BPCER = 6.34% @ APCER = 5% and BPCER = 3.77% @ APCER = 10% on Dataset-2. The obtained results have demonstrated the superior performance of the proposed method indicating the robustness to different type of printers and reliability of MAD. The aspects of generalizability needs further investigation with an evaluation on multiple datasets which will be carried in future works.

References

[1] N. Damer, S. Zienert, Y. Wainakh, A. M. Saladié, F. Kirchbuchner, and A. Kuijper. A multi-detector solution towards an accurate and generalized detection of face morphing attacks. In *22th International Conference on Information Fusion (FUSION)*, pages 1–8, 2019.

[2] L. Debiasi, U. Scherhag, C. Rathgeb, A. Uhl, and C. Busch. Prnu-based detection of morphed face images. In *2018 International Workshop on Biometrics and Forensics (IWBF)*, pages 1–7, 2018.

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

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

[5] M. Hildebrandt, T. Neubert, A. Makrushin, and J. Dittmann. Benchmarking face morphing forgery detection: Application of stirtrace for impact simulation of different processing steps. In *International Workshop on Biometrics and Forensics (IWBF 2017)*, pages 1–6, 2017.

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

[7] C. Kraetzer, A. Makrushin, T. Neubert, M. Hildebrandt, and J. Dittmann. Modeling attacks on photo-id documents and applying media forensics for the detection of facial morphing. In *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security*, IH&MMSec '17, pages 21–32, 2017.

[8] A. Makrushin, C. Kraetzer, J. Dittmann, C. Seibold, A. Hilsmann, and P. Eisert. Dempster-shafer theory for fusing face morphing detectors. In *2019 27th European Signal Processing Conference (EUSIPCO)*, pages 1–5, 2019.

[9] A. Makrushin, C. Kraetzer, T. Neubert, and J. Dittmann. Generalized benford's law for blind detection of morphed face images. In *Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security*, IH&MMSec '18, pages 49–54, 2018.

[10] A. Makrushin, T. Neubert, and J. Dittmann. Automatic generation and detection of visually faultless facial morphs. In *VISAPP)*, pages 39–50, 2017.

[11] T. Neubert, A. Makrushin, M. Hildebrandt, C. Kraetzer, and J. Dittmann. Extended stirtrace benchmarking of biometric and forensic qualities of morphed face images. *IET Biometrics*, 7:325–332, 2018.

[12] R. Raghavendra, K. Raja, S. Venkatesh, and C. Busch. Face morphing versus face averaging: Vulnerability and detection. In *IEEE International Joint Conference on Biometrics (IJCB)*, pages 555–563, 2017.

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

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

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

[16] R. Raghavendra, S. Venkatesh, K. Raja, and C. Busch. Towards making morphing attack detection robust using hybrid scale-space colour texture features. In *IEEE International Conference on Identity, Security and Behaviour Analysis (ISBA 2019)*, pages 1–7, 2019.

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

[18] U. Scherhag, C. Rathgeb, and C. Busch. Performance variation of morphed face image detection algorithms across different datasets. In *2018 International Workshop on Biometrics and Forensics (IWBF)*, pages 1–6, 2018.

[19] C. Seibold, A. Hilsmann, and P. Eisert. Reflection analysis for face morphing attack detection. *arXiv preprint arXiv:1807.02030*, 2018.

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

[21] L. Zhang, M. Yang, and X. Feng. Sparse representation or collaborative representation: Which helps face recognition? In *IEEE International Conference on Computer Vision (ICCV)*, pages 471–478, 2011.