Doctoral theses at NTNU, 2022:299

Sushma Krupa Venkatesh

Robust Algorithms For Face Morphing Attack Detection

Database, Vulnerability and Detection

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of Philosophiae Docto
Faculty of Information Technology and Electrical Engineering
Department of Information Security and Communication Technology

NTNU
Norwegian University of Science and Technology

Sushma Krupa Venkatesh

Robust Algorithms For Face Morphing Attack Detection

Database, Vulnerability and Detection

Thesis for the Degree of Philosophiae Doctor

Gjøvik, October 2022

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Information Security and Communication Technology

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Information Technology and Electrical Engineering

Department of Information Security and Communication Technology

ISBN 978-82-326-6115-2 (printed ver.)
ISBN 978-82-326-5932-6 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2022:299

Printed by NTNU Grafisk senter

Declaration of Authorship

signature:

.....

(Sushma Venkatesh)

Gjøvik, Date: 6th September 2022

Abstract

Biometrics has emerged as a promising technology for automated recognition of individuals. The stored biometric characteristics are used to recognise an individual and hence biometrics technology plays a major role in security-related applications and stands in the front-line for authentication of data subjects. Biometrics has a wide range of applications in law enforcement, surveillance, banking, border control, medical records, time and attendance tracking etc. As the inherent biometrics characteristics don't undergo any changes, biometric technology has shown the best performance for authentication of data subjects.

Though biometrics technology is promising for person authentication, attackers may employ various techniques like presentation attacks and adversarial attacks to impersonate an enrolled individual with interest to obtain unauthorised access to the system. In addition to these attacks, in relatively recent times, biometric are attacked in the facial image enrolment stage, especially in the ID related applications, by performing face morphing. Facial morphing is initially performed for entertainment purposes, but gradually it has been used to attack face recognition systems. The face morphing process combines two different facial identities to generate a single facial image with the facial representations of both identities. An attacker may use the morphed facial image to enrol it in the ID documents (like driving license, passport). Since the morphed facial image shows a high resemblance to both facial identities, the ID document can be claimed by both identities. This indicates the severity of facial morphing and the necessity of morphing attack detection mechanisms to avoid the security lapse.

Hence the primary objective of this thesis is the development of face morphing attack detection techniques using hand-crafted and deep learning approaches. During this doctoral work, morphing attack detection approaches are developed for both digital and print-scan datasets. To empirically evaluate the performance of

the newly generated morphing attack detection approaches, various face morphing databases are generated using landmark and deep learning-based GAN techniques. Furthermore, the vulnerability of face recognition systems to face morphing attacks with ageing co-variate is evaluated. To this extent, this doctoral thesis contributes with the novel morphing attack detection approaches.

Acknowledgement

I would like to express my sincere gratitude to NTNU and Department of Information Security and Communication Technology for their funding support to undertake this research program. Further I would like to extend my gratitude to the administration team for their support. I appreciate the timely support extended by the IT team and a big thanks to the whole team. I would like to thank my colleagues for their great support, timely feedback and helping in achieving a positive work environment. Finally I would like to thank the supervisors Prof. Christoph Busch and Prof. Kiran Raja for their support.

Contents

List of Tables

List of Figures

List of Abbreviations

ABC Automatic Border Control Gate

APCER Attack Presentation Classification Error Rate

BPCER Bona fide Presentation Classification Error Rate

BSIF Binarized Statistical Image Features

CNN Convolutional Neural Network

D EER Detection Equal Error Rate

D-MAD Differential Morphing Attack Detection

DCNN Deep Convolutional Neural Network

DET Detection Error Trade-off

FAR False Accept Rate

FMMPMR Fully Mated Morph Presentation Match Rate

FMR False Match Rate

FRS Face Recognition System

GAN Generative Adversarial Network

HOG Histogram of Gradients

ICAO International Civil Aviation Organisation

ISO International Organisation for Standardisation

MAD Morphing Attack Detection

MAP Morphing Attack Potential

MMPMR Mated Morph Presentation Match Rate

NIST National Institute of Standards and Technology

P-CRC Probabilistic Collaborative Representation Classifier

S-MAD Single Image Morphing Attack Detection

SOTA State Of The Art

SRKDA Spectral Regression Kernel Discriminant Analysis

SVM Support Vector Machine

WD Wavelet Denoising

Part I

Overview

Chapter 1

Introduction

Biometrics is used to automatically verify/identify an individual based on the physiological and behavioural characteristics [16]. Among various biometric characteristics, face biometric systems are widely deployed in various security related applications by considering the usability (non-intrusive capture) and the reliable verification performance [17, 18]. The rapid development in the field of deep learning has led to cutting edge Face Recognition Systems (FRS) that result in accurate recognition systems [19, 20]. The FRS are extensively integrated with the border control applications to enable passport control. Because face biometrics are reliable for both border control officers (human observers) and Automatic Border Control (ABC) gates to verify the identity.

Though FRS are reliable in achieving high recognition performance, they are vulnerable to direct (e.g. presentation attack/spoofing attack) and indirect attacks (e.g. template attacks, database injection attacks). The direct attacks are performed at the sensor level [21] in which the attacker can use various artefacts (e.g. facial photo, silicon face mask, printed iris, synthetic fingerprint) to impersonate a target. The indirect attacks are performed at various functional blocks of FRS to obtain non-legitimate access to the internal system operation [21]. In addition to these attacks on the FRS, one such attack that can be carried out during the face image enrolment, especially in security applications such as border control, is the face morphing attack [22].

The face morphing attack has gained significant interest in the biometric research community due to its relevance for high-security applications (e.g. passport control). The face morphing process involves manipulating the face image of two (or more) different data subjects by blending to generate a single morphed face image. A person with malicious intent may enrol the morphed facial image in the

identity documents like e-passport/eMRTD (electronic Machine Readable Travel Document). Eventually, it may lead to a false identity claim of the same eMRTD by multiple contributing data subjects[1]. Such an attack leads to a violation of sole proprietorship of an identification document that eventually leads to inadequacy in security.

1.1 Motivation and Problem Statement

Face biometrics is an integral part of the identification documents including passport/eMRTD. The passport application protocol varies across countries, especially the way in which facial images are acquired during the application process. The majority of the countries accept the printed face image from the applicant for the eMRTD enrolment. Further, the printed face image will be re-digitized/scanned to the eMRTD. However, countries like New Zealand, the UK, Estonia and Ireland accept a digital facial image for passport renewal that can be uploaded directly in the web portal [13, 23, 24, 22]. However, if a person has a criminal record, obtaining a genuine eMRTD is challenging. Therefore the applicant may use a malicious approach by using the morphing process to generate a morphed face image by blending his/her face image with a look-alike accomplice. Since the morphed face image shows higher similarity to both data subjects (malicious person and accomplice), the accomplice can submit it to the passport office with a malicious intention to enrol it in the eMRTD. Even though the border control officials thoroughly investigate any possible image manipulation, one may fail to detect morphed facial images because of (1) the absence of visible artefacts (2) high resemblance to the applicant, especially after careful post-processing. Additionally, several studies suggest that identifying an unfamiliar face is a challenging task that may deceive the border control officers to detect the morphed face images [25, 26, 27, 28, 29]. Once the morphed face image is enrolled in the eMRTD, both data subjects (malicious person and accomplice) can claim the same eMRTD to cross the border control.

A real-life risk of facial morphing attack was demonstrated in a case study reported in [30, 31]. An activist's facial image was morphed with the EU representative 'Federica Mogherini's' facial image to generate a morphed facial image of a non-existing person. The morphed face image is used by an activist to successfully obtain the eMRTD by deceiving the passport application protocol. This case clearly indicate the risk of morphing attacks that can risk the national security. Another similar case is also reported in [32], in which a person applying for the Dutch

[1] Although multiple identities can be combined to create a morphed image, it is common to combine two identities in practice due to practical applicability. This thesis restricts the scope of studying morphing attacks when two identities are used to create a morphed image

passport was identified to have submitted a morphed face image. The submitted morphed face images were generated with another person, an asylum seeker in another country.

Considering the threat posed by morphing attacks, developing reliable face Morphing Attack Detection (MAD) techniques are essential. Therefore, this doctoral work is dedicated to developing reliable and automated face morphing attack detection techniques. In general, morphing attack detection techniques are of two types (i) Single image-based morph attack detection (S-MAD)/ No-reference based morph attack detection[2] (ii) Differential image-based morph attack detection (D-MAD)/Reference-based morph attack detection.

1.2 Research Objectives

Based on the motivation discussed above, the following research objectives are formulated in this doctoral thesis.

- Generation of new databases (digital/print-scan) to facilitate the empirical evaluation of new Morphing Attack Detection (MAD) algorithms development in this doctoral work.
- To benchmark the vulnerability of FRS for different types of facial morph generation techniques and to analyze the factors (for example, morphing factor, ageing) that can contribute to circumvent the FRS to the highest degree.
- To develop novel algorithms to detect face morphing attacks reliably, especially in the Single Image Morphing Attack Detection (S-MAD) or no-reference based scenario.

1.3 Research Questions

The following research questions are formulated to investigate in this doctoral study.

1.3.1 Research Question 1 (RQ1): MAD using residual noise

During the process of morphing, two different facial images are combined together by blending the corresponding pixels of the two facial images. This process may generate additional noise due to geometric distortion. With the intuition of having additional noise stemming from the morphing process, the following research questions are formulated.

[2] S-MAD and No-reference based MAD is used interchangeably throughout the thesis

1. What is the best-suited approach to effectively detect face morphing attacks by quantifying residual noise in the digital images? Does quantifying the residual noise resulting from the morphing process help in detecting face morphing attacks?
    (a) Which deep learning architectures can be designed to quantify the residual noise resulting from morphing?
    (b) What is the performance gain achieved using residual noise-based attack detection compared to SOTA morph attack detection schemes?

1.3.2 Research Question 2 (RQ2): Features and robustness for MAD

In some countries like the United Kingdom (UK), printed facial images are submitted to the passport office by an applicant are re-digitized using a scanner to enrol in the eMRTD. Along the process of re-digitization, additional noise may be introduced due to the print-scan process, making the morphing attack detection challenging. The additional noise introduces another challenge in detecting morphing attacks efficiently. The following research question is formulated to reliably detect morphing attacks in print-scan images with this motivation.

1. What kind of novel features (texture-based/time frequency-based/deep features/ensemble) can be devised to reliably identify morphing attack when no reference image is available (i.e., S-MAD) in a print-scan scenario?
    (a) What is the best performing SOTA method to reliably detect no-reference morphing attacks in a print-scan scenario?
    (b) What image features (texture-based/time frequency-based/deep features/ensemble) can provide reliable morphing attack detection, especially in a no-reference scenario?
    (c) Does the hand-crafted feature analysis approach generalize across cross datasets when compared to deep learning features?
    (d) Does the morphing factor employed to generate a morphing image influence the performance of morph attack detection?

1.3.3 Research Question 3 (RQ3): Influence of face age progression on vulnerability and morphing attack detection

Generally, the eMRTD issued to an applicant has a life span of 10 years. Furthermore, the facial image enrolled in the eMRTD is also valid for 10 years. However, the eMRTD holder undergoes ageing with time. Hence, the facial biometric characteristics change with the appearance of wrinkles, saggy skin, and addition/reduction of fat on facial muscles due to weight gain/loss. Therefore, it is

interesting to empirically study the influence of ageing through the following research question.

1. What is the impact of ageing on morphing attack potential with respect to FRS?

   (a) Does the blending/morphing factor show any diverse effect on the FRS vulnerability and no-reference MAD performance?

   (b) Do existing MAD techniques in the literature scale up to detecting such morphing attacks with ageing co-variate?

1.3.4 Research Question 4 (RQ4): Deep learning based facial morph generation

The morphed facial image may get verified with the contributing subjects involved in the morphing procedure only when the morphed image has sufficient high quality. However, the success rate of attacks using a morphed facial image on FRS depends on the quality of the morphs generated. As traditionally employed, landmark-based morphs generate several ghosting artefacts during the morphing process and require manual intervention to assure high quality. Therefore, it is essential to generate fully automated morphs and investigate the attack potential of such morphing attacks on FRS.

1. Can deep learning-based image synthesis using Generative Adversarial Networks (GAN) be used to generate high-quality face morphs?

   (a) Does modifying the information at latent space of StyleGAN lead to the generation of high quality morphed image?

   (b) Does the StyleGAN based morph generation circumvent the FRS to a higher degree when compared to previous GAN based morph generation (MorGAN)?

   (c) Does the StyleGAN based morph generated image can be successfully detected using SOTA MAD algorithms?

1.4 Research Methodology

The following research methodology is followed to achieve the research objectives based on the aforementioned research questions.

- **Data Collection: Face morphing datasets**

  The field of face morphing lacks a diverse and large scale dataset that can facilitate the development of reliable MAD approaches. With an objective to generate reliable MAD techniques and analysis of the factors influencing the vulnerability of FRS, this doctoral work contributed to the generation of face morph databases. Morph databases were generated in digital and print-scan scenarios based on the experimental requirement to complement the real-life scenario. Chapters 5 and 6 detail the morph databases generated using deep learning based approaches such as StyleGAN and MIPGAN techniques. Chapter 7 presents the first morphing database generated with ageing co-variate. The ageing database is generated to analyze the influence of ageing factor on FRS with morphing attacks. Chapter 10 details the face morphing data generation using landmark-based approaches that also includes re-digitized data generated to analyze the impact of different printers for morphing attack detection.

  This research work is undertaken using the public databases (FRGC [33], PUT face database [34] and MORPH II database [35]). These databases are employed to generate different morphing types (Landmark, GAN) in various mediums (digital, print-scan, print-scan compression). Facial images in Figures 2.1, 2.2, 2.6, 2.7 were captured for illustration purpose and the data subjects have consented for using it in the publication medium.

- **Vulnerability analysis**

  Understanding the factors that make the FRS vulnerable is essential, to get a better perspective for generating reliable MAD techniques. Hence, this thesis focus on the factors (ex: morphing factor, ageing factor and morph generation type) that affect the vulnerability of FRS. Chapter 7 evaluates the influence of ageing for morphing attacks, and it extensively evaluates the impact of different morphing factors on such attacks. Chapter 5 and 6 performs an evaluation of the vulnerability generated by different morphing types.

  A new vulnerability metric is proposed to compute the proportion of morphed facial images verified with its contributing subjects to perform the vulnerability analysis effectively. Chapter 7 details the new vulnerability metric and its mode of operation for vulnerability assessment.

- **Morphing Attack Detection approach** The final step is to detect morphed facial images by extracting the features present in the facial image. Based on the existing literature, morphing attacks can be detected by employing a handcrafted/machine learning or deep learning approach. Motivated by

the existing MAD approaches and its detection performance in the literature, this thesis proposes both handcrafted and deep learning approaches for MAD. Chapter 10 presents the machine learning-based morph detection approach especially for the print-scan datasets. Further, Chapters 8 and 9 present the deep learning-based morphing detection approach that has shown best performance on several digital datasets.

1.5 List of Research Publications

This section provides a list of research publications earned during this doctoral study that contributes to this thesis.

- S. Venkatesh, R. Raghavendra, K. Raja, L. Spreeuwers, R. Veldhuis, and C. Busch. ¨Morphed face detection based on deep color residual noise¨. In 9th Intl. Conf. on Image Processing Theory, Tools and Applications (IPTA), IEEE, November 2019 [4].
- S. Venkatesh, R. Raghavendra, K. Raja, L. Spreeuwers, R. Veldhuis, and C. Busch. ¨Detecting morphed face attacks using residual noise from deep multi-scale context aggregation network¨. In 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), pages 269–278, IEEE, March 2020 [5].
- S. Venkatesh, K. Raja, R. Raghavendra , and C. Busch. ¨On the influence of ageing on face morph attacks: Vulnerability and Detection¨. In International Joint Conference on Biometrics (IJCB), pages 1–8, IEEE, September 2020 [3].
- S. Venkatesh, H. Zhang, R. Raghavendra, K. Raja, N. Damer, and C. Busch. ¨Can GAN generated morphs threaten face recognition systems equally as landmark based morphs? - Vulnerability and Detection¨. In the 2020 8th International Workshop on Biometrics and Forensics (IWBF), pages 1–6, 2020 [1].
- S. Venkatesh, R. Raghavendra, K. Raja, and C. Busch. ¨Single image face morphing attack detection using ensemble of features¨. In 2020 IEEE 23rd International Conference on Information Fusion (FUSION), pages 1–6, IEEE, 2020 [6].
- H. Zhang, S. Venkatesh, R. Raghavendra, K. Raja, N. Damer, and C. Busch. ¨MIPGAN — Generating strong and high quality morphing attacks using identity prior driven GAN¨. IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(3):365–383, 2021 [2].

- S. Venkatesh, R. Raghavendra, K. Raja, and C. Busch. ¨Face morphing attack generation and detection: A comprehensive survey¨. IEEE Transactions on Technology and Society, 2(3):128–145, March 2021 [7] .

1.5.1 List of Additional Publications

Additional publications not part of this doctoral thesis but are published during the doctoral study are listed below.

- N. Damer, K. Raja, M. Sussmilch, S. K. Venkatesh, F. Boutros, M. Fang, F. Kirchbuchner, R. Raghavendra, and A. Kuijper. ¨ReGenMorph: Visibly realistic GAN generated face morphing attacks by attack re-generation¨. International Symposium on Visual Computing ISVC, 2021 [36].

- S. Venkatesh, R. Raghavendra, and K. Raja. ¨Face morphing of newborns can be threatening too: Preliminary study on vulnerability and detection¨. In 2021 IEEE International Joint Conference on Biometrics (IJCB), pages1–8, 2021 [37].

- K. Raja, M. Ferrara, A. Franco, L. Spreeuwers, I. Batskos, F. de Wit, M. Gomez-Barrero, U. Scherhag, D. Fischer, S. K. Venkatesh, J. M. Singh,G. Li, L. Bergeron, S. Isadskiy, R. Raghavendra, C. Rathgeb, D. Frings,U. Seidel, F. Knopjes, R. Veldhuis, D. Maltoni, and C. Busch. ¨Morphing attack detection - database, evaluation platform, and benchmarking¨. IEEE Transactions on Information Forensics and Security, 16:4336–4351, 2021 [38].

- L. Qin, F. Peng, S. Venkatesh, R. Raghavendra, M. Long, and C. Busch. ¨Low visual distortion and robust morphing attacks based on partial face image manipulation¨. IEEE Transactions on Biometrics, Behavior and Identity Science, 3(1):72–88, 2020 [39].

1.5.2 Additional publications in various topics

- S. Venkatesh, 'Multi-spectral Finger based User Verification using Off-the-Shelf Deep Features'. IEEE International Conference on Imaging systems and techniques (IST 2022) [40].

- R. Raghavendra, S. Venkatesh, K. Raja, and C. Busch. ¨Handwritten signature and text based user verification using smartwatch¨. 2020 25th International Conference on Pattern Recognition (ICPR), 2021, pages 5099–5106, doi: 10.1109/ICPR48806.2021.9412048. [41].

- S. Venkatesh, R. Raghavendra, and P. Bours. ¨Video based deception detection using deep recurrent convolutional neural network¨. In N. Nain, S. K.Vipparthi, and B. Raman, editors, Computer Vision and Image Processing (CVIP), pages 163–169, Singapore, 2020. Springer Singapore. [42].

- R. Raghavendra, J. Singh, S. Venkatesh, K. Raja, and C. Busch. ¨ Face presentation attack detection using multi-classifier fusion of off-the-shelf deep features¨. In Proc. of the 4th Intl. Conf. on Computer Vision and Image Processing (CVIP), 2019 [43].

- J.M. Singh, S. Venkatesh, K. Raja, R. Raghavendra, and C. Busch. ¨Detecting finger-vein presentation attacks using 3D shape diffuse reflectance decomposition¨. In 2019 15th International conference on Signal-Image Technology Internet Based Systems (SITIS), pages 8-14, 2019 [44].

- N. Vetrekar, R. Raghavendra, K. Raja, S. Venkatesh, R. Gad, and C. Busch. ¨Visible to band gender classification: An extensive experimental evaluation based on multi-spectral imaging¨. In 2019 15th International Conference on Signal-Image Technology Internet-Based Systems (SITIS),pages 120–127, 2019 [45].

- K. Raja, R. Raghavendra, S. Venkatesh, M. Gomez-Barrero, C. Rathgeb,and C. Busch. ¨A study of hand crafted and naturally learned features for fingerprint presentation attack detection. In Handbook of Biometric Anti-Spoofing¨. Springer Intl. Publishing, January 2019 [46]. .

- R. Raghavendra, K. Raja, S. Venkatesh, and C. Busch. ¨Design and development of low-cost sensor to capture ventral and dorsal finger vein for biometric authentication¨. IEEE Sensors Journal, 19(15):6102–6111, August 2019 [47].

1.6 Scope of thesis

The scope of the thesis is listed below and also illustrated in the Figure 1.1

- The development of MAD algorithms using both hand-crafted and deep learning features.

- The performance of the developed MAD algorithms are benchmarked with the existing MAD techniques in the literature.

- Development of novel morph generation techniques to create high quality face morphing attacks.

- Generation of new databases to benchmark the novel MAD techniques developed in this thesis.
- Benchmarking the vulnerability of both Commercial Off-The Shelf (COTS) and deep learning-based FRS for the newly generated face morphing database.

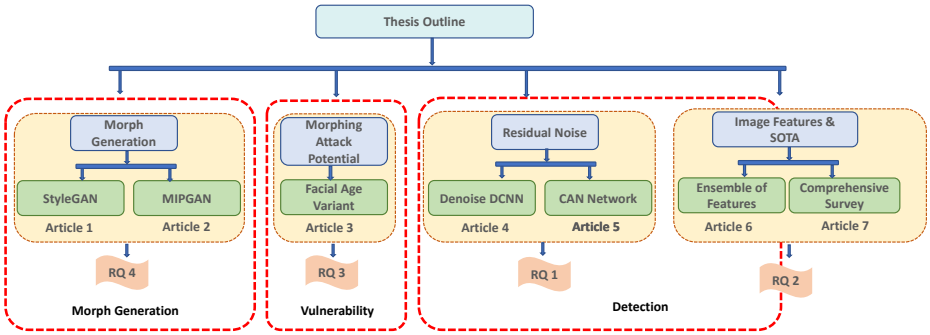


Figure 1.1: Thesis outline addressing the research questions within the scope of this dissertation

The overview of the tasks carried out during this doctoral work is presented in Figure 1.1. This thesis is divided into four parts: Part I presents the overview of the thesis. Parts II (morph generation), III (vulnerability analysis) and IV (morphing attack detection) present the research articles published during the course of this doctoral work. In Part I, Section 1 presents the introduction and the various real-life instance that serve as a motivation to define the problem statement and the research objective. Further, the research questions are formulated in Section 1.3 to address in this thesis followed by research methodology in Section 1.4. In the course of this doctoral work, the research publications achieved by addressing the research questions are listed in Section 1.5. Additionally, the research publications earned in collaborative research in morphing and other topics are listed in Section 1.5.1.

Chapter 2 presents the background of face morphing and the existing literature. Further, the evaluation metrics employed for the morphing attack detection and vulnerability analysis are presented in Section 2.5. Finally, the availability of public evaluation and benchmarking platforms are presented in Section 2.6. Furthermore, a summary of the research articles published in this doctoral work is presented in Chapter 3.

Finally research articles published in this doctoral work are presented in the Part II (morph generation), Part III (vulnerability analysis) and Part IV (face morphing attack detection). Articles 5 and 6 present the high quality morph image synthesis using deep learning based GAN approach that address the research question 1.3.4. Article 7 details about the influence of ageing on morph attacks on FRS that address the research question 1.3.3. Articles 8 and 9 present the deep learning based morph attack detection approach on digital datasets that addresses the research question 1.3.1. Article 10 presents the morphing attack detection approach on re-digitised dataset by employing hand crafted approach that addresses the research question 1.3.2. Finally the article 11 summarises the comprehensive development in the field of face morphing generation and detection.

Chapter 2

Background and Related Work

This chapter is an updated version of our article [7] that details the existing research work in this area. This Chapter gives an overview of the different morph generation tools, State-Of-The-Art(SOTA) approaches employed for morphing attack detection, evaluation metrics for morph detection and vulnerability analysis.

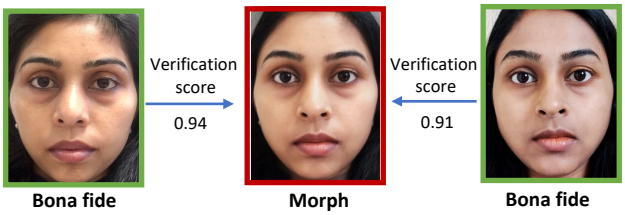
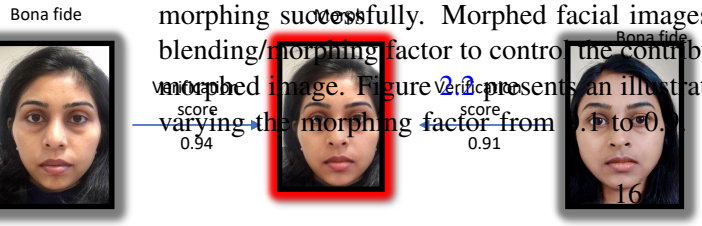


Figure 2.1: Illustration of facial morphing (Figure is taken and adopted from S.Venkatesh et al., [7])

Face morphing is performed with an intention to deceive the human observers or automatic FRS. Figure 2.1 shows the two facial images blended using the morphing process to generate a single morphed facial image. The morphed facial image possesses high similarity to the contributing facial images used for morphing. Hence the human observer evaluation and automatic FRS may fail to detect morphing successfully. Morphed facial images can be generated by varying the blending/morphing factor to control the contribution of a subject towards the final morphed image. Figure 2.2 presents an illustration of facial morphs generated by varying the morphing factor from 0.1 to 0.9. Variation in the morphing factors



indicates the contribution of weights of the corresponding facial image to generate

Figure 2.2: Facial morphs generated by varying morphing factors

2.2 Face Morphing Generation Tools

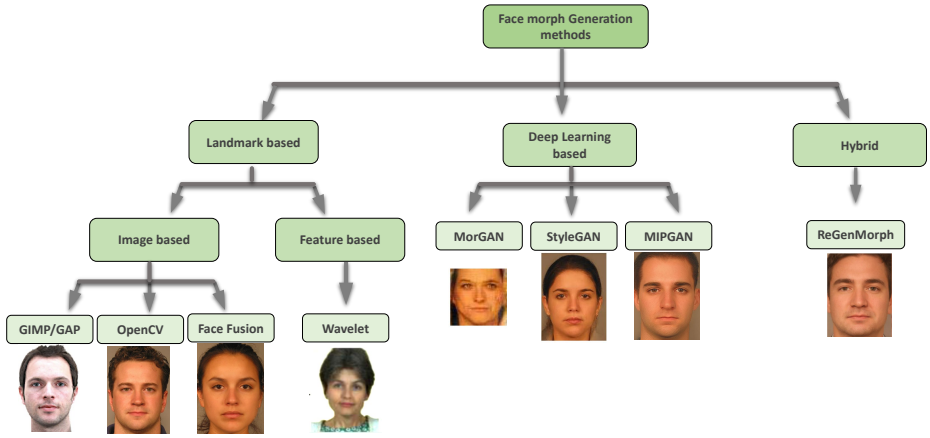


Figure 2.3: Facial morphs generated using different morph generation methods (Figure is taken and adopted from S.Venkatesh et al., [7])

9

Facial morphs are generated using various techniques. One widely used morph generation technique is the Landmark based technique, where the point correspondence between the facial images is obtained to perform morphing. Conventionally landmark based approach is extensively employed. However with the progress in the deep learning, morphed image synthesis based on deep learning based GAN has evolved.

2.2.1 Landmark based morph generation

Landmark based approach is conventionally employed for morph generation. There exists different methods for face morphing using landmark technique [48, 49, 50, 51, 52, 53]. Generally landmark based facial morphs follows four steps.

- **Correspondence:** First step is to find the corresponding points or the key points between the two facial images, also known as landmark points. The landmark points are the coordinates of the facial salient points specifically from the eyes and nose region. The manual landmark detection is accurate but it is time consuming and tedious [22] [54]. However the automatic landmark detection is less time consuming and convenient than manual processing. Various methods for automatic landmark detection includes D-lib landmark detector [55, 56], elastic bunch graph models [57], active shape models [58].

- **Warping:** In the next step the detected landmark points from the two facial images are positioned by moving the pixels based on the nearest landmark point [59]. Several approaches for warping are proposed that include Free Form Deformation which is a grid based or mesh based approach [60], image metamorphosis by field morphing approach [61], image deformation approach based on moving least squares [62], mass spring based on image deformation model [63]

- **Blending:** Once the corresponding pixels of both the facial images are positioned, they are combined by blending process. Linear blending is the most prominently used blending process, where the color values of both the facial images are combined at each pixel point. Additionally the weighted linear function makes the flexibility of changing the weights of the corresponding facial images based on the requirement. $I_M = (1-\alpha) \times I_1 + \alpha \times I_2$, where I_M is the morphed image, I_1 and I_2 are the facial images of the two different identities that are intended to be morphed and α is the morphing factor.

- **Post-processing** During the process of morphing, especially during warping and blending the pixels are re-positioned or some pixels might be lost during the morphing process and hence leading to ghosting artefacts on the morphed image as shown in the Figure 2.4. To achieve a high quality morphed facial image, it is essential to minimise the artefacts by performing certain post-processing operations. Various post-processing operations are outlined in [64, 65, 66] that includes image enhancement operations by varying the brightness, contrast and sharpness. Further image smoothing is

performed by Gaussian filtering. The ghosting artefacts can also be minimised by manual post-processing that includes image retouch, edge correction, ir
based
morph
tools tl

Landmark issues

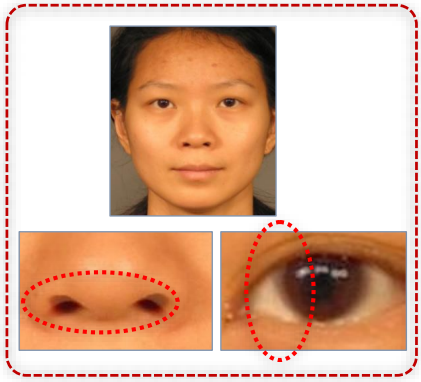


Figure 2.4: Ghosting artefacts generated in landmark based morph generation

2.2.2 Deep learning based morph generation

Advancement in the deep learning techniques has facilitated the morphed facial image synthesis using Generative Adversarial Network (GAN). The GAN architecture synthesizes a facial image similar to real person in the image by employing generator and discriminator network. The first GAN architecture that has synthesized the morphed facial image is the MorGAN architecture. It generates the morphed facial image with dimension 64 $\times$ 64 [15]. The MorGAN morph generation is further improved that has resulted in facial morphs with dimension 120 $\times$ 120 [72]. Another GAN architecture based on StyleGAN synthesizes the facial image in the latent space that has resulted in morphed facial image with a dimension 1024 $\times$ 1024. Further the two versions of StyleGAN model (StyleGAN 1 and StyleGAN 2) is employed with identity loss functions that has developed morph images generated using identity prior network MIPGAN 1 and MIPGAN 2. Another hybrid morph generation technique that avoids the introduction of GAN artefacts during the image synthesis in the latent space is the ReGenMorph [36] that is a combination of landmark and GAN based morph generation. Figure 2.3 shows an illustration of the morphed facial images generated using different GAN approaches.

2.3 Digital and Print-Scan Morph Generation

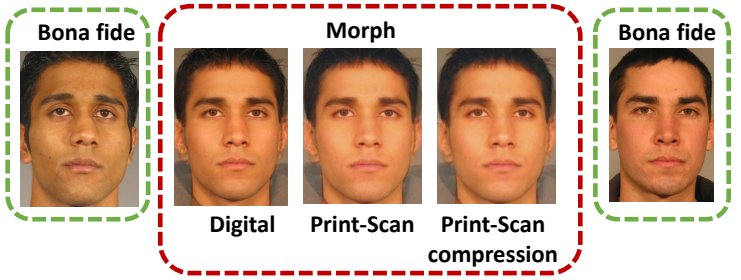


Figure 2.5: Illustration of facial morph generated in different scenarios (i) Digital (ii) Print-scan (iii) Print-Scan Compression (Figure is taken and adopted from Zhang et al., [2])

2.4 Related Work

This section discusses the existing literature for the approaches employed for detecting morphed facial images. Based on the SOTA, morphing attack detection techniques can be generalised into Single image based MAD and differential image based MAD. This section is derived from our article [7]. Reader may come across redundant information.

2.4.1 Single Image Morphing Attack Detection (S-MAD)

In this scenario, morphing attack detection has to be performed on a given single image that lacks any additional information. Hence, single image MAD is highly challenging. Figure 2.6 illustrates the real life passport application scenario where a facial image is submitted by the applicant for enrolment in the passport. This facial image intended for enrolment has to be thoroughly investigated to check the existence of morphing. As discussed in the Section 2.2, morphed facial images

can be generated in digital or print-scanned format depending on the requirement of the country's passport application procedure. Hence, it is essential to have robust techniques for facial morphing attack detection in both digital and print-scan scenario. Several morphing attack detection techniques are existing in the literature that can reliably detect morphing and are listed in the Table 2.1 below. Based on the existing literature the S-MAD techniques can be generalised into texture based, quality based, residual noise based, deep learning based, wavelet based and hybrid S-MAD approaches.

Texture based S-MAD: employs the widely used texture features like Local Binary Pattern (LBP) [74], Local Phase Quantization (LPQ) [75], Binary Statistical Image Features (BSIF) [76], Histogram Of Gradients (HoG). Further the SIFT and SURF texture features were also employed in several works [77, 78]. Considering the efficiency of texture feature for extracting the salient features, it is widely employed for morphing detection. However, the generalizability of texture based approaches are still challenging.

Quality based S-MAD: In general, the morphing process deteriorates the quality of the morphed facial image. The variation of facial images before and after morphing is analysed with respect to reflection, compression artefacts, edge distortion, Photo Response Non-Uniformity (PRNU) [79, 80, 81, 82]. Hence several works report the image degradation approach to detect face morphing.

Residual noise based S-MAD: During morphing process, additional artefacts are introduced due to the pixel discontinuity in the blending and warping process. Hence the morphed facial image posses morphing artefacts/morphing noise that are non-exiting in the bona fide facial image. Several works report the idea of extracting the residual noise from the given facial image [4] [5]. Residual noise based S-MAD shows good performance on the digital datasets with improved generalizability. But the impact of residual noise extraction for print-scan dataset is not studied.

Deep learning based S-MAD: Rapid evolution of deep learning based approaches and it's efficiency for image classification tasks has motivated to employ the deep CNN based techniques for MAD. Hence researchers have employed various pretrained deep CNN networks for MAD that includes AlexNet, VGG18, VGG19, GoogleNet, ResNet18, ResNet50, ResNet150 [83, 84, 85, 86, 87, 88, 89, 64, 5]. The deep learning approaches are explored on both digital and print-scan data and it shows better performance than hand crafted approaches.

Hybrid S-MAD: In general a combination of different approaches shows best result compared to individual approaches. As different approaches employs diverse techniques to extract features and to perform classification, fusion of different ap-

proaches in various levels (score, feature) shows reliable performance. Several works report the hybrid approach employed for morph detection [6] [84] [86] [87] [90] [91] [92].

Wavelet based S-MAD: Wavelet based sub-band selection approach is performed to analyse the texture feature variation in the bona fide and morphed facial images. Several works report the wavelet based morph detection on digital images [93, 94]. The efficiency of wavelet based MAD for print-scan data needs to be investigated.

Further the different approaches employed for S-MAD along with its advantages and disadvantages are provided in our survey article [7]. Reader is referred to

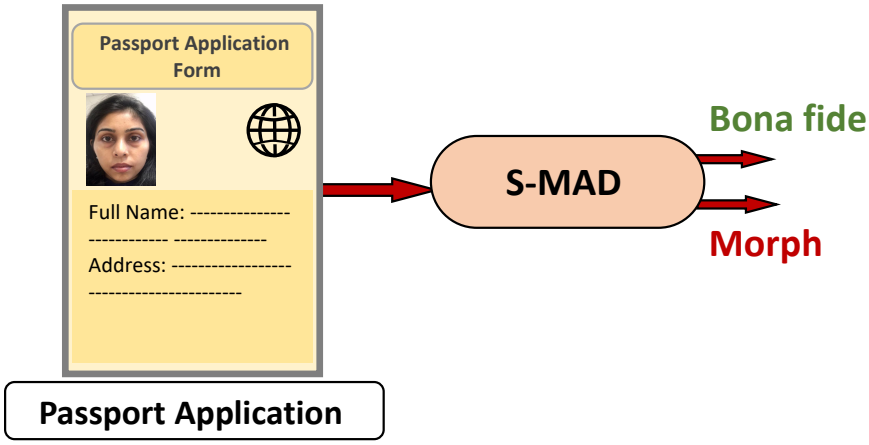


Figure 2.6: Illustration of passport application scenario for S-MAD (Figure adopted from our article [7]

2.4.2 Differential Morphing Attack Detection (D-MAD)

In the D-MAD scenario, as the name suggests there exists an additional live captured image to detect morphing in the given facial image. As the given facial image can be compared with the trusted live captured image, the challenge posed in the S-MAD is reduced in case of D-MAD. Figure 2.7 illustrates the real life border control scenario for D-MAD. In this scenario, to detect morphing in the given image (facial image for enrolment in passport), an additional reference image from the trusted live source eg: Automatic Border Control Gate (ABC) gate is available [112]. In the literature, there exists several approaches to detect morphing in differential scenario. Based on the morph generation type, it is essential to have D-MAD approaches for both digital and print-scan scenario. Table 2.2 presents the

Table 2.1: State-of-the-art S-MAD

Reference	Approach	Algorithm	Database
Raghavendra et al. [54]	Quantised DCT co-efficients	Local Binary Pattern (LBP)-SVM, Binary Statistical Image Features (BSIF)-SVM, Image Gradient (IG)-SVM	Digital
Makrushin et al. [95]	Quantised DCT co-efficients	Benford features	Digital
Neubert et al. [96]	Image degradation approach	Corner feature detector	Digital
Seibold et al. [64]	Deep learning-based approach	VGG19, Google Net, Alex Net	Digital
Raghavendra et al. [13]	Texture-based approach	LBP, LPQ, BSIF, colour textures	Print-scan
Asaad et al. [97]	Texture-based approach	Topological data analysis approach	Digital
Scherhag et al. [98]	Texture- and frequency-based approach	LBP, LPQ, BSIF, 2DFFT with SVM classifier	Digital Print-scan
Raghavendra et al. [83]	Deep CNN-based approach	Feature fusion of fully connected layers of VGG19 and Alex Net	Digital Print-scan
Kraetzer et al. [99]	Image life cycle model	Keypoints (SIFT, SURF, ORB, FAST, AGAST) and loss of edge operators (Canny and Sobel)	Digital
Hildebrandt et al. [81] [100]	StirTrace-based approach	Multi-compression anomaly detection	Digital
Debiasi et al. [82]	Image degradation	Photo Response Non-uniformity (PRNU)	Digital
Raghavendra et al. [90]	Steerable features	Luminance component extraction	Print-scan
Hildebrandt et al. [81]	StirTrace	StirTrace face morph forgery detection	Print-scan
Seibold et al. [79]	Image degradation	Specular reflection	Digital
Makrushin et al. [101]	Quantised DCT co-efficients	Benford features extracted from quantised DCT co-efficients	Digital
Neubert et al. [102]	Morph pipeline footprint detector	Benford features extracted from quantised DCT co-efficients	Digital
Spreeuwers et al. [103]	Texture-based approach	LBP-SVM, Down-up sampling	Digital
Scherhag et al. [88]	Feature difference-based approach	Pre-processing and feature extraction using texture descriptors , keypoint extractors, gradient estimators and deep learning-based method	Digital
Damer et al. [84]	Multi-detector fusion	LBPH, Transferable deep-CNN	Digital
Ferrara et al. [85]	Deep learning	AlexNet, VGG19, VGG-Face16, VGG-Face2	Print-scan
Scherhag et al. [87]	Multi-algorithm fusion	Texture descriptors (LBP, BSIF), Keypoint extractors (SIFT, SURF), gradient estimators (HoG), Deep neural network	Digital
Debiasi et al [104]	PRNU	PRNU DFT magnitude histogram and PRNU DFT energy	Digital
Seibold et al. [105]	Complex multi-class pre-training	VGG-19 network	Digital
Damer et al. [106]	Texture and deep learning based	Anomaly detection using LPQ and VGG features	Digital
Venkatesh et al. [4]	Colour denoising-based approach	Denoising Deep Convolutional Neural Network	Digital
Scherhag et al. [80]	PRNU	Spectral features and spatial features	Print-scan
Makrushin et al. [86]	Dempster-Shafer Theory	KeyPoints (SIFT, SUFT, FAST, ORB, AGAST, High Dim LBP, GoogleNet, VGG19	Digital
Raghavendra et al. [91]	Scale space approach	Colour scale space features	Print-scan
Neubert et al. [107]	Frequency and spatial domain feature space approach	Discrete Feature Transformation (DFT) , SURF, SIFT, ORB, FAST, AGAST, Canny edge, SobelX, SobelY)	Digital
Seibold et al. [89]	Style Transfer-based approach	LBP, BSIF, Image degradation, Deep neural network (VGG19)	Digital
Venkatesh et al. [5]	Colour denoising-based approach	Context Aggregation Network	Digital
Venkatesh et al. [6]	Ensemble-of-features-based approach	LBP, HoG, BSIF	Print-scan
Seibold et al. [108]	Interpretability based on DNN	Focused Layer-wise Relevance Propagation (FLRP)	Digital
Aghdaie et al. [94]	Wavelet based approach	Attention based DNN	Digital
Aghdaie et al et al. [93]	Wavelet based approach	Discriminative 2D Discrete Wavelet Transform (2D-DWT)	Digital
Abisoye et al. [109]	Texture based approach	LBP, Neighborhood Component Analysis (NCA) for feature extraction. Classification using K-Nearest Neighbor (KNN), Decision Tree Classifier (DTC), Naive Bayes (NB)	Post-processed digital
Damer et al. [110]	Pixel based approach	Pixel Wise MAD (PW-MAD) using DenseNet-121 architecture	Digital Print-Scan
Tapia et al. [111]	Combination of intensity, texture and shape based approach	texture features using LBP, BSIF, shape feature using inverse HoG	Digital
Lorenz et al. [92]	Fusion approach	Deep features (ArcFace, FaceNet) and texture features (BSIF, LBP)	Digital

existing D-MAD techniques for both digital and print-scan data. Based on the existing literature, the D-MAD techniques can be generalised into feature difference based D-MAD and Demorphing.

Feature Difference: In a morphed facial image due to distortions during morphing process, the image feature undergoes variation. Analysing the variation in feature vectors with respect to texture, gradient, landmark, histogram and deep features of the given image to be probed and the reference image indicates the existence of morphing [113, 84, 114, 115, 116]. As the differential MAD has two facial images i.e, a suspected morphed facial image and a reference image live captured from a trusted source, subtracting the features.

Demorphing: Face morphing is achieved by following various procedures (e.g. warping, blending, post-processing). Investing this idea of face morphing on a given facial image in reverse order may reveal various components used to perform morphing. Demorphing procedure is employed by several works reported so far for D-MAD [112, 117].

Wavelet decomposition: Difference features in the suspected morphed facial image and the reference facial image are investigated by wavelet decomposition in sub-band level. Differential face morph detection by employing the wavelet based

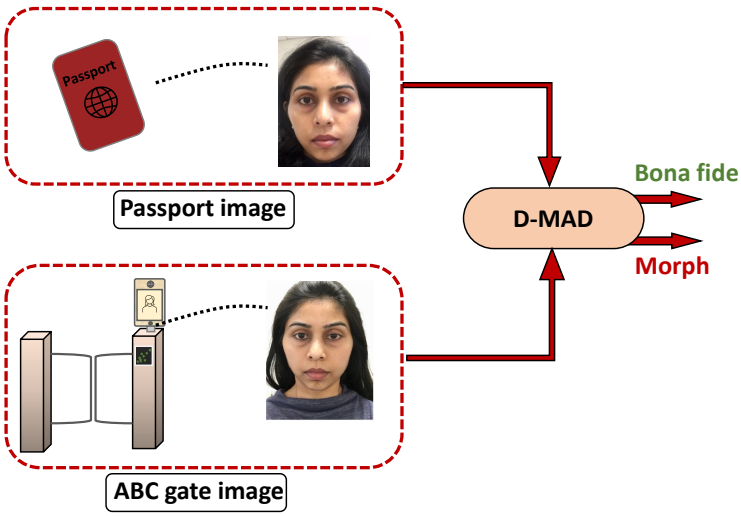


Figure 2.7: Illustration of border crossing scenario for D-MAD (Figure adopted from our article [7]

Table 2.2: State-of-the-art D-MAD

Reference	Approach	Algorithm	Database
Ferrara et al. [119]	Demorphing	Demorphing by image subtraction	Print-scan
Ferrara et al. [119]	Demorphing	Face verification	Digital
Scherhag et al. [114]	Landmark based approach	Distance-based and angle-based feature extraction with Random Forest, SVM without kernel and SVM with radial basis function classifier	Digital
Scherhag et al. [88]	Feature difference-based approach	Pre-processing and feature extraction using texture descriptors, keypoint extractors, gradient estimators and deep learning-based method	Digital
Damer et al. [84]	Multi-detector fusion	LBPH, Transferable deep-CNN	Digital
Singh et al. [116]	Deep learning	SfS Net, AlexNet	Digital Print-scan
Damer et al. [113]	Landmark shift	Landmark detection, shift representation	Digital
Peng et al. [117]	Face restoration by demorphing GAN	Symmetric dual-network architecture	Digital
Scherhag et al. [115]	Deep Face Representation	ArcFace Network, FaceNet algorithm	Digital Print-scan
Seibold et al. [105]	Deep Learning	Layer-wise Relevance Propagation (LRP)	Digital Print-scan
Ortego et al. [112]	Demorphing, Deep CNN-based	Auto-encoders	Digital Print-scan
Soleymani et al. [120]	Deep learning	Siamese network	Digital
Soleymani et al. [121]	Deep learning	Appearance and landmark disentanglement	Digital
Autherith et al. [122]	Analysis of geometric facial features	Facial anthropometry-based facial feature comparison	Digital
Chaudhary et al. [118]	Wavelet decomposition	Wavelet sub-band selection, Kullback Liebler Divergence (KLD)	Digital
Sudipta et al. [123]	Implicitly disentangle identities	Information theoretic framework using conditional GAN	Digital
Borghi et al. [124]	Deep learning based	Double Siamese network	Digital

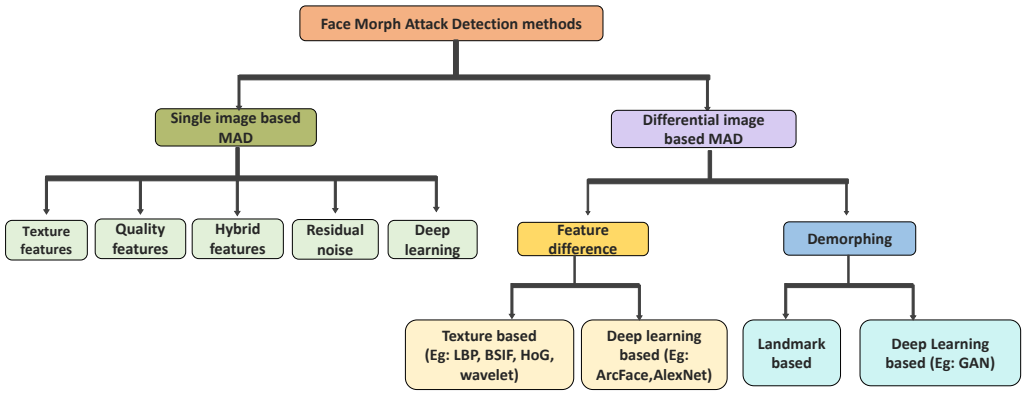


Figure 2.8: Taxonomy of Morphing attack detection approaches for S-MAD and D-MAD (Figure adopted from our article [7]

2.5 MAD Evaluation Metrics

The metrics employed for vulnerability analysis and morphing attack detection performance are presented in the following section.

2.5.1 Detection metrics

Based on the ISO/IEC 30107-3, the standardised metrics for evaluation of MAD approaches are given in [21] namely Attack Presentation Classification Error Rate and Bona fide Presentation Classification Error Rate. In the following section we discuss this metrics that are taken from ISO/IEC 30107-3 [21].

Attack Presentation Classification Error Rate (APCER) Attack presentation classification error rate defines the number of attack images (in this case morph images) misclassified as bona fide images. APCER is presented in the equation 2.1, where N_{PAIS} is the number of attack presentations (morph presentation) and Res_i will be 1, if i^{th} presentation is classified as attack presentation and 0 if it is classified as bona fide presentation [21].

$$APCER = 1 - (\frac{1}{N_{PAIS}}) \sum_{i=1}^{N_{PAIS}} Res_i, \tag{2.1}$$

Bona fide Presentation Classification Error Rate (BPCER) Bona fide presentation classification error rate is the number of bona fide presentations misclassified as attack presentations (in this case morph presentations). BPCER can be represented in equation 2.2. Where N_{BF} is the number of bona fide presentations. Res_i will be 1 if the i^{th} presentation is classified as attack presentation and 0 if it is classified as bona fide presentation [21].

$$BPCER = (\frac{\sum_{i=1}^{N_{PAIS}} Res_i}{N_{BF}}), \tag{2.2}$$

2.5.2 Vulnerability metrics

Vulnerability of the FRS are analysed using the following vulnerability metrics (i) Mated Morph Presentation Match Rate (MMPMR)(ii) Fully Mated Morph Presentation Match Rate (FMMPMR). Reader can refer to the article [7] for more information about the vulnerability assessment in Chapter 11.

Mated Morph Presentation Match Rate (MMPMR) MMPMR vulnerability metrics defines the proportion of morphed images getting verified with contributing subjects [125]. MMPMR is represented in the equation 2.3.

$$MMPMR(\tau) = \frac{1}{M} \cdot \sum_{m=1}^{M} \left\{ \left[\min_{n=1,\ldots,N_m} S_m^n \right] > \tau \right\}, \tag{2.3}$$

Where M is the number of morphed images and N_m is the number of contributing subjects to morph m. τ represents the threshold of the FRS at the designated False Match Rate (FMR). Further the comparison score for the morph m of the subject n^{th} is represented as S_m^n.

Fully Mated Morph Presentation Match Rate (FMMPMR) FMMPMR vulnerability metric is proposed by [3], that defines the proportion of morphed images that get verified with the contributing subjects. Additionally it takes into account all the attempts that the morphed image gets verified with a pairwise comparison to the contributing subjects. Hence FMMPMR follows strict protocol to evaluate the attack strength of the morph image. FMMPMR is represented in the equation 2.4.

$$FMMPMR = \frac{1}{P} \sum_{M,P} (S1_M^P > \tau) AND (S2_M^P > \tau) \ldots AND (Sk_M^P > \tau) \tag{2.4}$$

Where $P = 1, 2, \ldots, p$ represents the number of attempts the contributing subjects are compared against the M^{th} morphed image. $K = 1, 2, \ldots, k$ represents the total number of subjects contributing to generate morph. τ represents the threshold which is set according to the FRONTEX requirement [126] to FMR = 0.1% and Sk_M^P represents the comparison score for the contributing subject K for the P^{th} attempt.

Morphing Attack Potential (MAP) MAP [127] metric aims at assessing the attack potential of the dataset M that consists of the morphed images to analyze the impact of two different factors (i) number of attempts the morphed image gets verified with the (variable number of) probe images captured at the ABC gate (ii) and generalising over a (variable number of) different FRS. MAP is represented in the equation below.

$$V = \frac{|M \in \mathbb{M} : C_V(M) = true|}{|\mathbb{M}|} \tag{2.5}$$

Where, V represents the vulnerability and M represents the morphed images. The vulnerability represented as V reports the proportion of morphed images $\mathbb{M}$, where the condition $C_V(M)$ is met.

Relative Morph Match Rate (RMMR) RMMR metric provides the relative measure by combining the recognition accuracy with the vulnerability measure [125]. Following the vulnerability metrics MMPMR and FMMPMR discussed earlier, RMMR can be defined as follows.

$$RMMR(\tau)_{MMPMR} = 1 + (MMPMR(\tau)) \quad [1 \quad FNMR(\tau)] \tag{2.6}$$

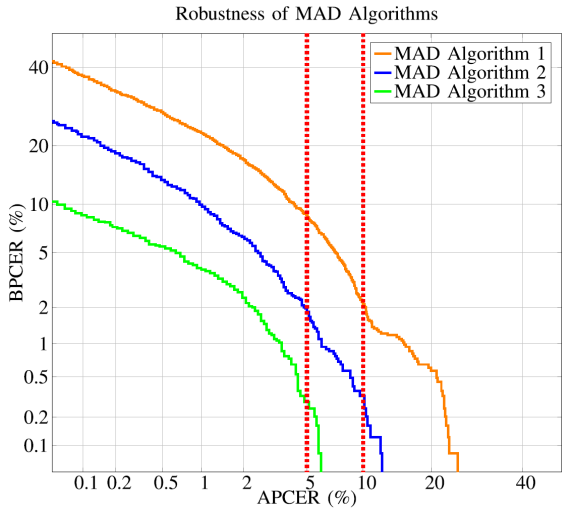


Figure 2.9: Sample illustration of the detection accuracy of MAD algorithms at different operating points with a detection error tradeoff curve (DET). As noted from the figure, MAD Algorithm 3 performs best at a chosen APCER of 5% or 10%.(Figure adopted from our article [7])

Detection Equal Error Rate (D-EER) In general, D-EER denotes a common point where both the errors APCER and BPCER are equal.

Detection Error Trade-off curve (DET) The error curve plotted corresponding to APCER and BPCER at different operational points are shown in the Figure 2.9. If one error is minimised for instance APCER is reduced, then the BPCER increases and vice versa. The error rates can be obtained by fixing the values for instance APCER = 5 $\%$.

2.6 Public benchmarking platforms

The public benchmarking platform provides a medium where one can evaluate the algorithm and assess the status. One can utilize the reliable and trustworthy infrastructure for assessment that includes sequestered datasets, protocols for evaluation and computational environment. There are two benchmarking platforms available for both S-MAD and D-MAD evaluation (i) NIST FRVT Part 4:Performance of Automated Face Morph Detection [128] and (ii) Bologna-SOTAMD [38].

NIST FRVT Part 4: Performance of Automated Face Morph Detection NIST FRVT benchmarking platform is put forward in the year 2018 as a common platform for assessment of algorithms developed for both S-MAD and D-MAD. The evaluation platform uses the datasets generated using different morphing approaches with the objective of identifying the low quality morphing that are generated using freely available open source tools, high quality morphing that are generated using the automatic morphing tools by incorporating additional post-processing steps to mask the artefacts and automated morphing that are generated using automatic morphing tools alone without manual intrusion. The NIST evaluation report [128] indicates the various algorithms presented by different institutes. Based on the evaluation report, morph detection is still a challenging task as there is no algorithm that shows best performance to detect morphing following the operational requirement of FRONTEX [126].

Bologna SOTAMD: D-MAD The Bologna public benchmarking was launched in the year 2019 and provides a common assessment platform for D-MAD techniques. The datasets employed for Bologna D-MAD benchmarking are collected using real ABC gates as a part of the European project SOTAMD [129]. Face morph generation is performed using both commercial and open-source morphing softwares. The evaluation report indicates that the existing D-MAD approaches are not reliable enough to detect morphing based on the FRONTEX operation requirement. More information about the D-MAD techniques that are evaluated can be obtained in [38].

Bologna SOTAMD: S-MAD The Bologna public benchmarking platform started in the year 2020 with the objective of providing a common assessment for both S-MAD and D-MAD algorithms. The datasets were constructed using high quality facial images similar to the real passports and the morphed images are constructed using the commercial and open-source morphing softwares. Further the morphed facial images are post-processed to mask the artefacts generated during the morphing process using automatic software and manual post-processing. More information on the Bologna benchmarking platform is provided in [38].

Summary of public benchmarking platforms Evaluation reports of the public benchmarking platforms suggest the requirement of reliable techniques for both S-MAD and D-MAD thus indicating the challenging task of detecting morphed facial images.

Chapter 3

Summary of Published Articles

This chapter summarises the research articles published during this doctoral study. The Articles 3.1 and 3.2 focus on the morph generation and address the research question 1.3.4. Articles 3.3 works on the vulnerability assessment and address the research question 1.3.3. Furthermore the Articles 3.4 , 3.5 and 3.6 focus on the morphing attack detection on digital and print-scan scenario and address the research questions 1.3.1 and 1.3.2 respectively. Finally the Article 3.7 presents a comprehensive survey of face morphing attack generation and detection, that partly address the research question 1.3.2.

3.1 Article 1: Can GAN Generated Morphs Threaten Face Recognition Systems Equally as Landmark Based Morphs?- Vulnerability and Detection [1]

This article is published at International Workshop on Biometrics and Forensics (IWBF) 2020

Most of the research works in the literature have developed MAD techniques for landmark- based morphing approaches. Morphing process induces noise, and hence the final morphed facial image exhibits ghosting artefacts as shown in Figure 3.1, that needs to be post-processed to improve the quality. Hence this research work aims to generate a high quality morphed facial image using deep learning techniques. With the advancement in the field of deep learning, morphed facial images were synthesized by employing Generative Adversarial Networks (GAN) technique [15] that does not require additional post-processing steps. Hence, it is essential to investigate to what extent can the GAN based morphed facial image challenge the FRS by generating vulnerability.

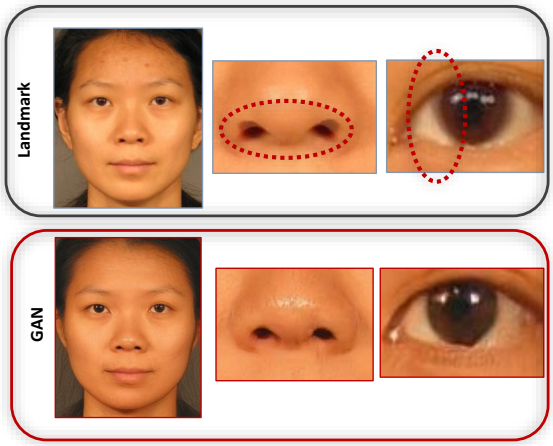


Figure 3.1: Illustration of facial morphs generated using (a) Landmark and (b) StyleGAN based morphing approaches

The existing work in the literature employs MorGAN to generate morphed facial images with resolution (64 $\times$ 64 pixels), however that does not comply with the ICAO standards [130]. With the motivation to improve the quality of the morphed facial image by employing GAN technique, this work generates a high quality morph (1024 $\times$ 1024 pixels) by employing StyleGAN complying to ICAO standards with less visual artefacts compared to landmark based morphs. We have proposed a method to perform morphing in the latent space of StyleGAN-1 architecture by performing weighted SUM fusion. Hence we employ the weight 0.5 for both the facial identities in the latent space which is then synthesized using synthesis network from StyleGAN. Although the morphs generated are of high quality, this work investigates if the StyleGAN based morphs can scale up to threaten the FRS similar to the landmark based morphs. Hence a new StyleGAN based morph dataset is developed that is derived from FRGC-V2 dataset.

To effectively analyze the vulnerability of FRS for the proposed StyleGAN based morphs, two different FRS are employed and the vulnerability is compared with the landmark based morph and MorGAN based morph. Further the performance is evaluated by employing four different MAD techniques existing in the literature. The experimental results indicates that, though the StyleGAN morphs possess visibly less artefacts, it has reduced capacity to make the FRS vulnerable when compared to landmark based morphs that may be due to the lower similarity with the geometrical features. Further the experimental results indicate that a morphing attack is less challenging to detect in case of StyleGAN and MorGAN morphs when compared with the landmark based morphs. But the StyleGAN based morphs

shows reasonable success rate in making the FRS vulnerable when compared with the MorGAN approach that may be attributed to the improved spatial resolution of StyleGAN with 1024×1024 pixels. Overall the reduced challenge to detect GAN based morphs may be attributed to the inherent noise generated during the GAN based morphing process.

3.2 Article 2: MIPGAN- Generating High Quality Morphing Attacks Using Identity Prior Driven GAN [2]

This article is published at IEEE Transactions on Biometrics, Behavior and Identity Science 2021

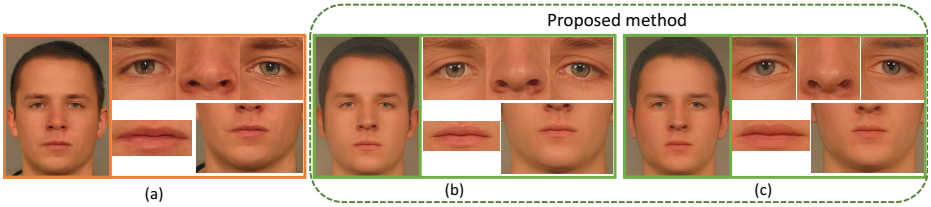


Figure 3.2: Illustration of facial morphs generated using (a) StyleGAN (b) MIPGAN-I (c) (b) MIPGAN-II approaches (Figure adopted from our article [2])

Motivated by our previous work on StyleGAN based morph generation, we extend this work to generate realistic high quality morphs by employing a new loss function to preserve the identity. Though the StyleGAN based morphs are of high quality with good perceptional resemblance with the original identity, it is challenging to generate vulnerability to circumvent the FRS to higher degree that may be a result of loss in the identity information in the synthesized images. Hence, this research work is undertaken to generate realistic facial image as shown in the Figure 3.2 by employing a novel loss function that includes identity priors. We refer this novel approach as MIPGAN (Morphing through Identity Prior GAN). We explore two versions of StyleGAN (I and II) and we term them as MIPGAN-I and MIPGAN-II. The proposed loss function includes four components (i) Identity prior computed using ArcFace FRS (ii) Perceptual loss (iii) ID- Diff (iv) MS-SSIM.

To analyze the attack potential of MIPGAN generated morphs in comparison with landmark and StyleGAN based morph generation, new morphing dataset is generated based on MIPGAN-I/II. To effectively evaluate the attack potential of MIPGAN based morph, datasets are generated in three different scenarios that includes (i) Digital (ii) Print-scan and (iii) Print-scan compression. Further the vulnerability of FRS are empirically investigated by employing five different FRS that includes both COTS and deep learning based FRS on five different datasets employing different morph generation techniques (Landmark-I, Landmark-II, StyleGAN, MIPGAN-I, MIPGAN-II).

The experimental results indicates that the MIPGAN based morphs shows vulnerability to all five FRS to a highest degree compared with StyleGAN based morphs.

Based on the experimental results deep learning based FRS shows highest vulnerability (MMPMR and FMMPMR) for the MIPGAN based morphs. Among the MIPGAN morphs generated, MIPGAN-I shows marginally better performance than MIPGAN-II. Further the experiments are performed to analyze the perceptual image quality using both human observer study and quality metrics such as SSIM and PSNR on the MIPGAN generated morphing images. The experimental results indicate the slightly better perceptual image quality of the morphs generated using MIPGAN compared to StyleGAN. Lastly we benchmark the performance of existing MAD on MIPGAN I/II.

3.3 Article 3: On the Influence of Ageing on Face Morph Attacks: Vulnerability and Detection [3]

This article is published at IEEE International Joint Conference on Biometrics (IJCB) 2020

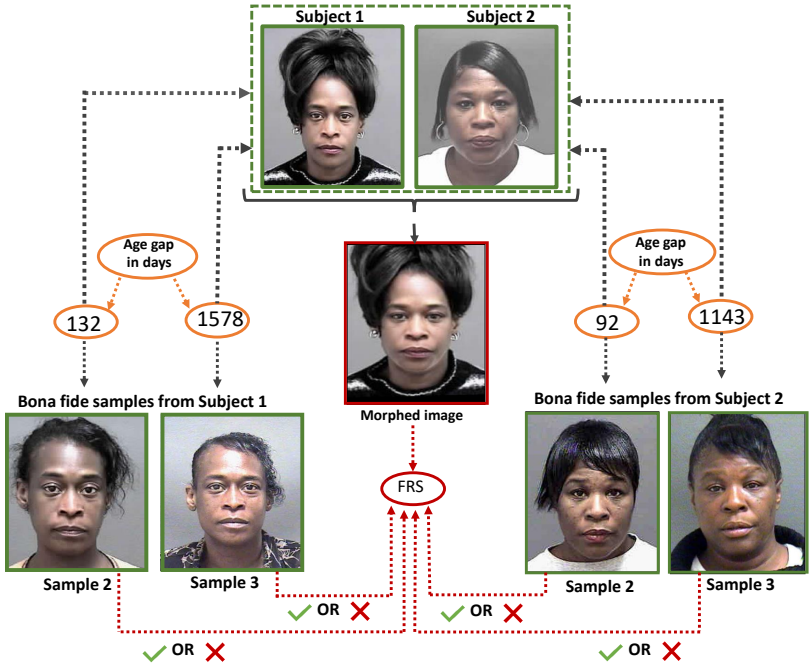


Figure 1: Intro Figure

Figure 3.3: Illustration of morph image quality with ageing co-variate (Figure adopted from our article [3])

The facial image enrolled in the eMRTD is retained for 10 years aligning with the validity of the eMRTD. However the facial biometric characteristics progresses with age resulting in addition of facial wrinkles, saggy skin, fat loss or fat deposition. Hence this work performs an analysis on the vulnerability of Commercial Off The Shelf (COTS) FRS when a morphed facial image is enrolled in the eMRTD and one of the contributing identity is probed after a period of time as the facial features have undergone age progression as illustarted in Figure 3.3.

To effectively analyze the influence of ageing, new datasets are constructed with ageing co-variate and is derived from the publicly available dataset. The datasets are developed with two different bins consisting of facial images with different age groups. The first dataset bin MorphAge-I consists of facial images with the

age group between 1 to 2 years and the second dataset bin MorphAge-II consists of facial images between the age group 2 to 5 years. To investigate the impact of different morphing factors, the facial images are morphed using three different morphing factors.

Further the vulnerability of FRS are investigated for the two different ageing dataset groups by employing two different COTS FRS. Additionally the impact of three different morphing factors for the vulnerability of the FRS is investigated on both age groups. A new vulnerability metric FMMPMR is also introduced that efficiently measures the vulnerability of the FRS when a morphed facial image is enrolled and corresponding facial images are probed by considering the number of attempts. This research work investigates if the existing MAD algorithms can scale up to detect such morphing attacks with ageing co-variate by employing 5 different MAD techniques existing in the literature.

The experimental results obtained from this research suggests that one of the COTS FRS is highly vulnerable for both age groups but the second COTS FRS indicates reasonable vulnerability. The vulnerability of FRS in MorphAge-I dataset is reduced to some extent in case of MorphAge-II that indicates, it is less likely that the morphed image gets verified against the probe images after certain ageing. Among the three different morphing factors 0.3, 0.5 and 0.7 employed in this research work, the experimental results indicates the higher vulnerability of FRS for morphing factor 0.5 when compared with the morphing factor 0.3 and 0.7. However, the MAD experiments on the two datasets indicates minor/negligible variation in the performance after ageing. Hence, the existing MAD algorithms can efficiently detect morphing even after ageing.

3.4 Article 4: Morphed Face Detection Based on Deep Color Residual Noise [4]

This article is published at 9th International Conference on Image Processing Theory, Tools and Applications (IPTA) 2019

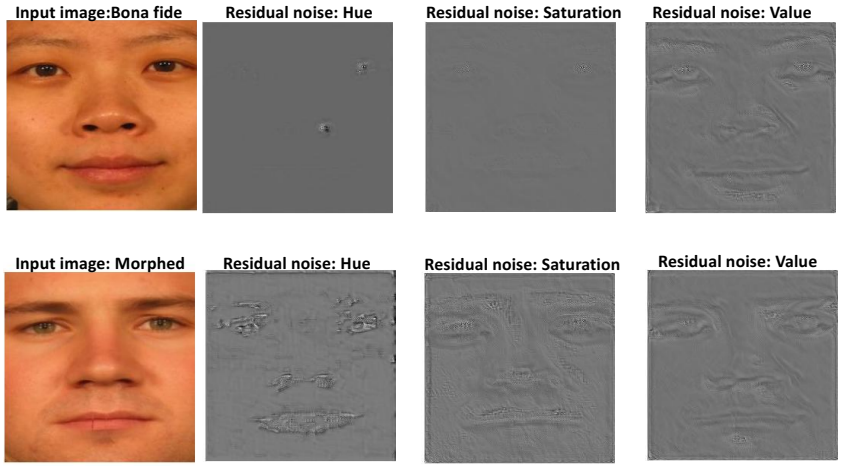


Figure 3.4: Illustration of residual noise from (a) Bona fide image (b) Morphed face image (figure taken from our article [4])

The morphed facial image is obtained by blending two different facial identities that imparts pixel distortion or missing pixels or noise due to double compression that can be termed as morphing noise. With the aim of detecting the existence of facial image morphing, this work investigates a novel technique to quantify the morphing noise from the given facial image. To this extent the proposed approach will estimate the residual noise by performing the differential operation between the given facial image and the denoised version of the the given facial image.

The proposed technique is designed to quantify the residual noise especially in the color space to achieve the discriminating features towards reliable face MAD. The denoising operation is performed using Denoise Deep Convolutional Neural Network (De-DCNN) [131] in HSV (Hue, Saturation and Value) color space. In the next step, the residual noise image is computed by subtracting the given image with it's denoised counterpart in HSV color space independently (see Figure 3.4). Then the computed residual noise are further processed to extract the features using Pyramid LBP. In this work, we have used Pyramid LBP with 3 level decomposition by considering computation versus detection accuracy. Finally Spectral Regression Kernel Discriminant Analysis (SRKDA) performs the classification to

decide if the given image is bona fide or morph.

To effectively evaluate the performance of the proposed method, extensive experiments are performed to benchmark the denoising MAD approach with 13 different SOTA MAD approaches existing in the literature. Further the performance of the proposed MAD approach is evaluated on three different databases by following different experimental protocols. (i) Firstly the performance is evaluated on three different datasets individually, (ii) Secondly, the performance is evaluated on large scale dataset by combining all three datasets together (iii) Finally, the generalizability of the proposed approach is evaluated by cross dataset evaluation. The experimental results obtained from intra-dataset and inter-dataset evaluation indicates the best performance of the proposed method in all three datasets when benchmarked with the 13 different SOTA MAD approaches.

3.5 Article 5: Detecting Morphed Face Attacks Using Residual Noise from Deep Multi-Scale Context Aggregation Network [5]

This article was published at IEEE Winter Conference on Applications of Computer Vision (WACV) 2020

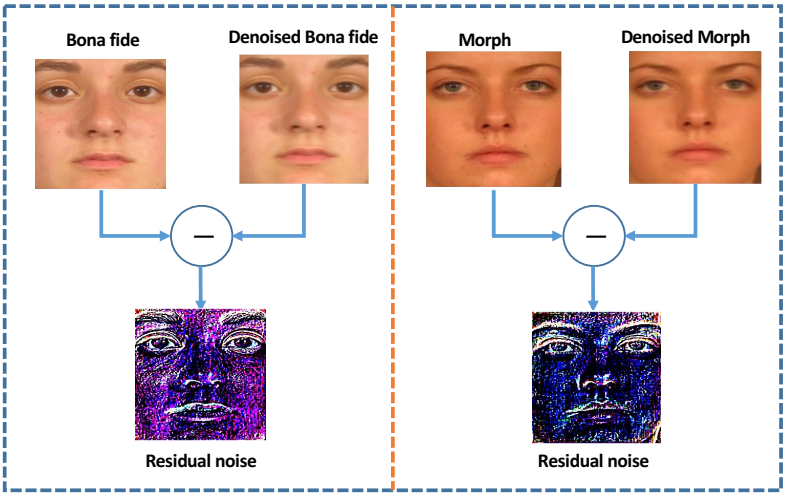


Figure 3.5: Illustration of residual noise from (a) Bona fide image (b) Morphed face image (figure taken from our article [5])

The promising result obtained using residual noise has led to the robust MAD. In this work we extend the previous work in two directions. (i) Improving the denoising operation by combining the best denoising approaches. Further this combination of denoising approaches are realised using, novel Context Aggregation Network (CAN). (ii) To overcome the need of color spaces to reduce the complexity.

In this work, we have employed the combination of four different denoising approaches to extract the morphing noise. The four different denoising approaches include Wavelet Denoising (WD), Block Matching and 3D filtering (BM3D), Multi-resolution Bilateral Filtering (MBF) and Denoising Convolutional Neural Networks (DnCNN) to extract morphing noise. The final denoised image is obtained by performing the best sub-band selection based on energy. The whole operation of denoising is realised using deep CNN architecture called deep Multi-Scale Context Aggregation Network (MS-CAN). In order to improve the generalizability of MS-CAN approach, we train our proposed MS-CAN for the general denoising

operation using IAPR-TC 12 dataset that includes natural images (consists of images of people, buildings and natural scenes) that are contaminated with Gaussian noise and salt and pepper noise. Given the face image, the proposed MS-CAN is used to obtain the residual noise. Further the deep features are extracted using pre-trained AlexNet and finally Probabilistic- Collaborative Representation Classifier (P-CRC) is used to perform classification.

The performance of the proposed method is benchmarked with 14 different deep learning and non-deep learning SOTA MAD techniques. To effectively validate the performance of the proposed method empirically, three different morphed face datasets are employed. To this extent, experiments are conducted on (i) the three datasets individually, (ii) merged dataset by combining the three datasets together and (iii) cross dataset evaluation to investigate the generalizability of the proposed method. The experimental results indicate that the proposed method is outperforming in all three datasets independently when compared with the existing deep learning and hand crafted SOTA MAD approaches.

3.6 Article 6: Single Image Face Morphing Attack Detection Using Ensemble of Features [6]

This article is published in IEEE 23rd International Conference on Information Fusion (FUSION) 2020

Conforming to the real life scenario where the applicant submits the printed facial image that will be scanned (re-digitized) to enrol in the eMRTD, the morphed facial image has to be re-digitised that may generate additional print-scan noise thereby masking the morphing noise. This leads to the challenging task of MAD on print-scan data in single image based scenario. Hence this research work analyzes the performance of proposed method with respect to re-digitised images in single image based scenario.

This research work undertakes the challenging task to detect morphing in the print-scanned facial images by employing ensemble of features. The given facial image is decomposed into two different color spaces (YCbCr and HSV) to obtain large scale complementary features that effectively serve as a cue to detect morphing. Further the scale space features are extracted individually by employing 3 level decomposition. In the next step, three different feature extraction techniques that includes LBP (Local Binary Pattern), BSIF (Binarized Statistical Image Filtering) and HoG (Histogram of Gradients) are employed to extract the texture features. Further the features are classified independently using Probabilistic Collaborative Representation Classifier (P-CRC) to obtain the morphing scores. Finally the independent morphing scores from the three feature extraction techniques are fused using SUM rule to decide if the given image is bona fide or morph.

To effectively analyze the performance of the proposed method, a new print-scan dataset is generated in addition to an existing print-scan dataset. The first dataset is the existing dataset that is re-digitised using RICOH office printer. Additionally a new re-digitised dataset is developed by using high quality photo printer (Epson expression photo XP860). Both the datasets have the same images to have a fair and balanced comparison of the proposed method. The experimental results indicates the best performance of the proposed method on both print-scan datasets when a comparative analysis is performed on the 16 different MAD algorithms existing in the literature. In general the performance of existing MAD algorithms is degraded on the newly generated dataset and indicates the challenging MAD due to the high quality print-scan process.

3.7 Article 7: Face Morphing Attack Generation and Detection: A Comprehensive survey [7]

This article is published in the journal IEEE Transactions on Technology and Society, 2021

This article presents an overview of the developments in the field of facial image morphing and morphing attack detection techniques [7]. Facial image morphing has captured the attention of the research community considering the threat caused due to the enrolment of morphed facial images in the eMRTD [22]. This article can serve as a starting point for the beginner to understand the development in the field of facial image morphing.

This article presents an overview of the existing morph generation types, available databases (public, private and sequestered), SOTA MAD techniques. It provides a technical insight in the reference-based and no-reference based morphing attacks conforming to the border control scenario and reports the existing MAD techniques (both reference-based and no-reference based techniques). Further exclusive discussion on different performance metrics for both benchmarking the vulnerability of FRS and detection of morphing attacks are discussed. The public platforms available for evaluation and benchmarking the MAD approaches are summarised in this article. During the course of work on facial image morphing, several technical challenges are observed that are reported as open challenges and risks that are faced in the morphing attack detection.

Chapter 4

Conclusions and Future Work

This thesis is aimed to develop reliable single image-based morphing attack detection algorithms. To this extent, extensive work is undertaken to address the research questions formulated in section 1.3. The research questions are answered by generating relevant publications included in Parts II, III and IV. This thesis mainly comprises the contributions from the seven publications achieved during this doctoral study. The overall thesis contribution includes the generation of face morphing databases, vulnerability assessment and development of MAD approaches. While addressing the research questions, several conclusions were drawn within the scope of this thesis that are discussed below.

4.1 Conclusions on each research question

Research Question 1

1. What is the best-suited approach to effectively detect face morphing attacks by quantifying residual noise in the digital images? Does quantifying the residual noise resulting from the morphing process help in detecting face morphing attacks?
    (a) What deep learning architectures can be designed to quantify the residual noise resulting from morphing?
    (b) What is the performance gain achieved using residual noise-based attack detection compared to SOTA morphing attack detection schemes?

Conclusion

- With the novel idea of quantifying the residual noise for S-MAD, in this thesis, two novel approaches based on deep learning are proposed. The first

work Denoise DCNN is developed over a deep CNN based denoising network to extract the residual noise in the color channels. The second approach is based on the novel Multi-Scale Context Aggregation Networks (MS-CAN) to compute the residual noise. Both of the proposed methods are extensively evaluated using three different datasets and compared with several state-of-the-art S-MAD techniques. Obtained results demonstrate the consistent and reliable performance of the proposed models. Among the two deep learning models, the MS-CAN network shows the best MAD performance. The proposed MS-CAN approach shows the reliable detection performance on all three databases (database-1 D-EER = 3.83%, database-2 D-EER = 4.85% and database-3 D-EER = 9.71%) when compared with 14 different state-of-the-art techniques.

Thus, based on the obtained results, it can be concluded that quantifying a residual noise is a promising approach to detecting morphing attacks reliably. Further, the use of residual noise also indicated reliable performance when tested for generalisation, especially on the digital medium.

Research Question 2

1. What kind of novel features (texture-based/time frequency-based/deep features/ensemble) can be devised to reliably identify morphing attacks when no reference image is available (i.e., S-MAD) in a print-scan scenario?
   (a) What is the best performing SOTA method to reliably detect no-reference morphing attacks in a print-scan scenario?
   (b) What kind of image features (texture-based/time frequency-based/deep features/ensemble) can provide reliable morphing attack detection, especially in a no-reference scenario?
   (c) Does the hand-crafted feature analysis approach generalize across datasets compared to deep learning features?
   (d) Does the morphing factor employed to generate a morphing image influence the performance of morphing attack detection?

Conclusion

- Detection of re-digitized morphing images is very challenging due to additional noise introduced during re-digitization. In this thesis, a novel approach based on the ensemble of features is proposed to address the challenges with re-digitized morphing images. The proposed method is compared with 16 different state-of-the-art methods on two different datasets.

The obtained results indicate the best performance of the proposed method (database-1 D-EER = 5.99% and database-2 D-EER = 5.64%). Thus, the use of ensemble of hand-crafted features has demonstrated the robustness toward noise resulting from the re-digitizing process. Therefore, our comprehensive analysis indicates that it is essential to employ more than a single set of hand-crafted feature to achieve reliable MAD, especially with print-scan datasets.

Research Question 3

1. What is the impact of ageing on morphing attack potential with respect to FRS?

   (a) Does the blending/morphing factor shows any diverse effect on the FRS vulnerability and no-reference MAD performance?

   (b) Do existing MAD techniques in the literature scale up to detecting such morphing attacks with ageing co-variate?

Conclusion

- Face ageing is an important influencing factor that needs to be investigated to shed light on the attack potential of morphed images with the lifetime of ID documents. To this extent, we have devised extensive experiments to benchmark both vulnerability and detection with two different age groups. The quantitative results obtained on two different age group database bins (MorphAge-1 with less ageing up to 1-2 years and MorphAge-2 with more ageing up to 2-5 years) indicates the reduced vulnerability of the COTS FRS. The empirical evaluation performed on three different morphing factors (0.3, 0.5, 0.7) indicates that morphed face images with a morphing factor 0.5 indicate the highest vulnerability. Further, benchmarking the morphing attack detection performance using the SOTA MAD approaches on the two database bins indicates a reduced effect of facial ageing. It is interesting to note that the vulnerability is reduced with age, and both morphing factors and ageing do not influence detection performance. One of the reasons for consistent detection performance can be attributed to the single image MAD. However, the detection performance may vary in the D-MAD scenario.

Research Question 4

1. Can deep learning-based image synthesis using Generative Adversarial Networks (GAN) be used to generate high-quality face morphs?

(a) Does modifying the information at latent space of StyleGAN lead to the generation of high quality morphed image?

(b) Does the StyleGAN based morph generation circumvent the FRS to higher degree when compared to previous GAN based morph generation (MorGAN)?

(c) Can the StyleGAN based morph generated image be successfully detected using SOTA MAD algorithms?

Conclusion

- Conventional morphing generation techniques based on landmarks often results in ghosting effects that demand post-processing. Therefore, the automatic generation of perceptually high-quality morph images is essential. To this extent, in this thesis, we proposed the technique using StyleGAN architecture to automatically generate high quality morphed face images. Even though the use of StyleGAN can generate perceptually high quality morphed images, the attack potential is degraded due to the loss of identity information. To overcome this drawback, we proposed a novel morph generation technique using an identity prior driven network (MIPGAN) that can preserve the identity information while generating a perceptually high quality morphed image. The identity while synthesising the morphed face images are preserved using a novel loss function that controls both the identity factor and perceptual quality of the morphed image. Extensive experiments are performed that have indicated the attack potential of the MIPGAN based morphed images to both human observers and COTS FRS.

4.2 Future works

While the findings from this thesis answered the initially formulated research objectives, the morphing attacks have evolved due to the advancement in morph generation methods (deep learning). The evolution of attacks and the findings from initial works set the foundations for future research directions. In this section, a number of interesting directions are listed.

- **Unavailability of large scale database with variation:**
  - Generally, attack detection on biometric-based applications are data-driven. Hence, large scale data is essential to train and test the model to achieve reliable performance. Several studies in the existing literature on morphing indicates the availability of various morphing data-

base, but they are not publicly available due to privacy and GDPR concerns. Hence, a systematic evaluation of morphing attack detection approaches on a large scale database is still underway.

- *Unavailability of morphing databases with diverse morph generation types:* MAD approaches must be able to achieve reliable detection performance irrespective of the morph generation type employed. Hence it is essential to generate a morph database by employing different morph generation types, including landmark and deep learning-based morph generation.
- *Unavailability of morphing database with different scenarios (digital, print-scan and print-scan compression):* Based on the existing studies, it is required to generate face morphing dataset in digital, print-scan and print-scan compression scenarios conforming to the passport application protocol in different countries.

- **Synthetic data generation**: With the recent research interest moving towards the generation of synthetic face database, one can consider generating a large-scale morphing database that overcomes the current challenge of lack of face morphing database.

- **Careful selection of data subject for morphing:** While generating a facial morph database, it is essential to take additional care to select the look-alike facial identities to generate a high quality morphed facial image. This will eventually challenge both human observers and automatic morphing attack detection and pave the way for the generation of reliable MAD approaches.

- **Algorithmic bias in MAD:** A reliable MAD approach has to detect morphing irrespective of gender, age, ethnicity and other demographic factors. While not many works are reported in this direction, a relatively recent article address the variation in detection performance for different ethnic backgrounds [132]. The influence of demographic factors should be carefully studied in future works.

- **Generalizability:** The risk of facial morphing is high in border control areas, and hence it is crucial to have MAD techniques that can achieve reliable performance. Earlier studies indicate that the existing MAD approaches achieve the best performance on a known set of data and thereby indicating the existing challenge of generalizability on unknown datasets [4, 5].

- **Variation with face co-variate:** The role of face co-variate (for ex: face ageing, ethnicity, gender, quality of the image captured and beautification of facial images) needs a systematic investigation. Work on facial ageing

and its influence on morphing attacks was reported in [3]. Further, the role of beautification needs a systematic investigation as beautification software are commonly employed to enhance or sharpen the captured facial image. One of the face co-variates that may have a larger influence on the D-MAD scenario is the image quality. As the facial images are captured in the ABC scenario, one may expect a variation in the quality of the image captured. Along the lines, the influence of occlusion due to hair and glasses needs a systematic study.

- **Performance metrics:** There is a need for standardized metrics to report the vulnerability of FRS to morphing attacks. Currently, various metrics are employed to report the vulnerability of FRS, including FMMPMR, MMPMR and MAP. Considering the real life scenario in border control, improved vulnerability metrics especially to account Failure To Accept Rate (FTAR) is required. And a generalised metric that consider different morphing generation types can also be anticipated.

Part II

Published Articles: Morph Generation

Chapter 5

Article 1: Can GAN Generated Morphs Threaten Face Recognition Systems Equally As Landmark Based Morphs? - Vulnerability and Detection

Sushma Venkatesh, Haoyu Zhang, Raghavendra Ramachandra, Kiran Raja, Naser Damer, and Christoph Busch. *Can GAN generated morphs threaten face recognition systems equally as landmark based morphs? - Vulnerability and Detection*. In 2020 8th International Workshop on Biometrics and Forensics (IWBF), pages 1–6, April 2020

5.1 Abstract

The primary objective of face morphing is to combine face images of different data subjects (e.g. an malicious actor and an accomplice) to generate a face image that can be equally verified for both contributing data subjects. In this paper, we propose a new framework for generating face morphs using a newer Generative Adversarial Network (GAN) - StyleGAN. In contrast to earlier works, we generate realistic morphs of both high-quality and high resolution of 1024×1024 pixels. With the newly created morphing dataset of 2500 morphed face images, we pose a critical question in this work. *(i) Can GAN generated morphs threaten Face Recognition Systems (FRS) equally as Landmark based morphs?* Seeking an answer, we benchmark the vulnerability of a Commercial-Off-The-Shelf FRS (COTS) and

a deep learning-based FRS (ArcFace). This work also benchmarks the detection approaches for both GAN generated morphs against the landmark based morphs using established Morphing Attack Detection (MAD) schemes.

5.2 Introduction

Due to the widespread deployment of biometric-based identification and verification of individuals, it is essential to observe a biometric characteristic that is reliable, user-friendly and easy to capture. Face biometrics is well suited for this purpose due to its popularity and widespread use for biometric authentication. Moreover we consider the ease of capturing from a distance in a non-intrusive manner and also the recently achieved high recognition accuracy. These properties further enable that face recognition is to be used in various applications that are attributed with high security requirements like border control. However, biometric face recognition systems (FRS) are known to be highly vulnerable to presentation attacks (aka., spoofing attacks) against the capture device [133]. In addition face recognition system can be deceived during the enrolment process by providing manipulated images [22].

Among the different types of attacks against FRS, face morphing has gained momentum because of the high impact it poses on border control security. The morphing process enables a malicious actor to generate a morphed image by using an accomplice's face image in a seamless manner [22]. The process introduces a significant threat to the border control scenario as it is easy to obtain a passport document with a morphed image. This fact is also due to the limitations of current passport issuance protocols in which digital images are submitted in a self-supervised manner by an applicant for passport renewal through web services in countries like New Zealand, Estonia and Ireland. In other countries there exists no live enrolment in the passport renewal process, on the contrary the facial image is provided by an applicant in printed form and is subsequently scanned and re-digitized. This leaves an opportunity for the applicant to morph the face image prior to submitting it in the passport application.

5.2.1 Morph generation process and limitations

Early work on face morphing attacks [22] demonstrated the vulnerability of FRS with respect to morphed facial images, while to the same extend human experts could be fooled [134] [135]. Following the recent works towards detecting morph attacks on both digital image and re-digitized (print-scanned) image [54, 83, 13] we must state that this area of research is still in a premature state. The crucial part of a morphing attack is the generation of high quality morphed facial image, which is ICAO compliant and can attack a deployed FRS with high probability.

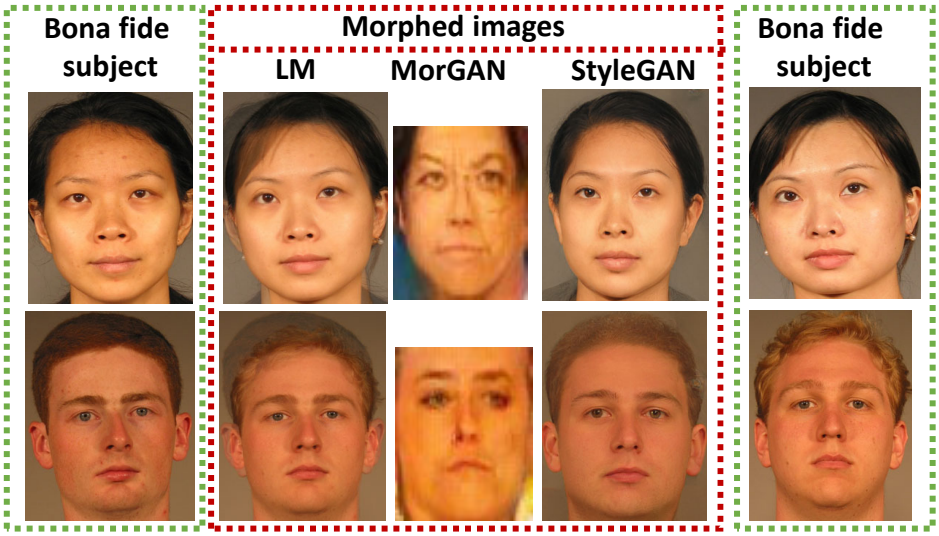


Figure 5.1: Comparison of morphed images generated using LandMark (LM) , MorGAN and StyleGAN

In the literature, there exist two different ways of generating morphed face images namely (a) Landmark based morphed face generation and (b) Deep learning-based morphed face generation. In landmark-based morph generation, given two images, the landmarks of both facial images are obtained and the Delaunay triangulation is generated for both images. Subsequently alpha blending is performed to obtain a single morphed image based on averaged Delaunay triangles. The majority of the recently published literature is based on open-source morphing tools [13] which are based on landmark constrained Delaunay triangles.

A deep learning-based approach in contrast involves synthesizing a morphed face image by using a Generative Adversarial Network (GAN). Limited works are reported in the literature on using GAN for morph generation[15]. The first reported work in this direction is based on the MorGAN [15] in which morphed images are generated corresponding to a image resolution of 64×64 pixels. Recently, the morphed images generated using MorGAN were super-resolved to have a incrementally larger dimension of 120×120 pixels [72]. It is important to note that the images generated using both approaches incorporating GAN [15] [72] are not ICAO compliant and hence have very limited use in real-life attacks. Irrespective of the morph image generation approach, it is essential that one needs to generate a high-quality image that can pose a high threat potential, when presented to a human expert in the control procedure while the passport issuance is carried out or to

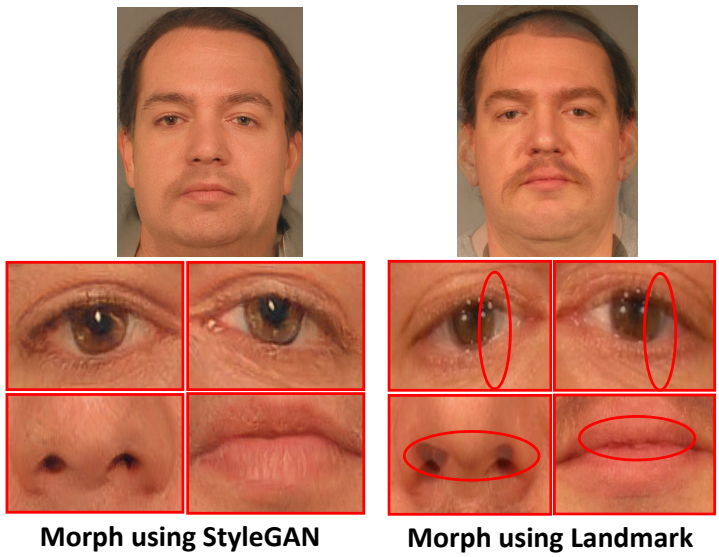


Figure 5.2: Illustration of minimal artifacts in morphed images generated using StyleGAN versus landmark based face morphing.

Motivated to address the limitation of low quality images generated by the previous GAN architectures, in this work, we present a new approach to generate high quality morph images. The recent improvement made in GAN architectures has enabled us to generate a high quality facial images with a resolution of 1024×1024 pixels using StyleGAN [136]. This is achieved by embedding the images into latent space which is further optimized to synthesize the high quality and high resolution image [137]. As illustrated in Figure 5.1 the morphed images generated using StyleGAN can be observed to be superior in terms of quality, resolution and visual depiction.

Further, as noted from Figure 5.2, a number of artifacts can be easily handled in an automatic manner with the newly proposed approach, which is capable of suppressing visual artifacts. The clear superiority of the newly proposed approach can be noted around the iris regions, where double edges are inherently dealt with. While it well known fact that landmark based morphs threaten FRS to a high degree [13, 22], one can easily conclude the amount of extra time and resources that is anticipated to make the morphs visually appealing by removing the artifacts.

While the superior quality of face images can be achieved through the newly proposed approach and eventually reaching compliance to ICAO standards, we raise some fundamental questions.

- Despite the high quality of morphed images do they scale up to threaten a FRS in the same manner as the landmark based morphs, which typically exhibit large artifacts?
- To what degree can current MAD mechanisms detect such GAN based attacks on FRS, when the processing is limited to the digital domain?

In the course of answering the above questions, we can summarize the contributions of this work as follows:

- A new approach to generate morphed face images using the StyleGAN is presented.
- A new face morphing dataset comprising of $2500 \times 3 = 7500$ morphed images is generated using the StyleGAN and MorGAN approach. In order to compare the new approach using the GAN methodology, this work also constructs a corresponding landmark based morph dataset.
- To quantify the threats from GAN based morphed face images, a comprehensive vulnerability analysis is conducted using both, a commercial FRS (COTS) and an open-source FRS (ArcFace).
- In order to give an insight into the detection challenges of such attacks, this work also reports a detailed evaluation of MAD mechanisms on both GAN based and landmark based morphed face images.

In the rest of the paper, Section 5.3 describes the morph generation process proposed in this work using StyleGAN. Section 5.4 provides the details regarding the quantitative experiments indicating the vulnerability of FRS and the detection challenge. With remarks on future works in this direction, we draw the conclusion in Section 5.5.

5.3 Morphed Face generation using StyleGAN

In this section, we present the StyleGAN based face morph generation to achieve high quality face morphs. Figure 5.3 depicts the block diagram of the proposed framework for the morphed face generation using a StyleGAN architecture[136]. Given the latent code L_1 of the faces, the StyleGAN [137] maps the inputs to an intermediate latent space (W) through the mapping network. The mapping layer consists of 8 fully connected layers that are serially connected. In this work, we force a strategy to embed the face image into the latent space (W_f), which is inspired by earlier work [137]. This process enables us to synthesize the data-subject-specific morphed face. The embedded latent space for a particular face is

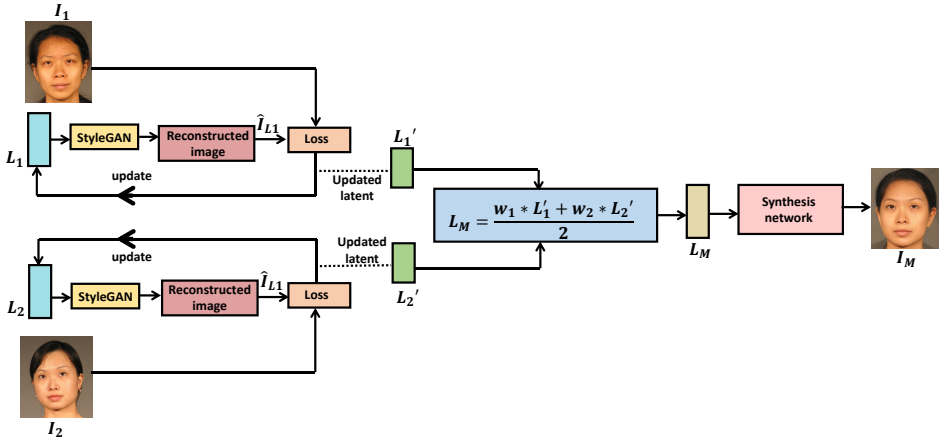


Figure 5.3: Block diagram of the morphed face image using StyleGAN

then passed through the synthesis network consisting of 18 layers, in order to control the adaptive instance normalization (AdaIN).As a direct result, we obtain the representation in 18 multiple latent spaces, each with a dimension of 512, which is further concatenated. For a given face image I_1, the embedding is carried out by optimizing a loss function that measures the similarity between I_1 and the reconstructed image $\hat{I}_{L_1}$ using the corresponding latent code L_1. To maintain the perceptual fidelity a loss is computed as the weighted combination of VGG-16 perceptual loss [137] as given below:

$$PL = min \sum_{i=1}^{4} \frac{\lambda_i}{N_i} ||F_j(\hat{I}_{L_1})^{\smile} F_j(I1)||_2^2 \tag{5.1}$$

Where, F_j is the feature output of VGG-16 layer $conv1_1$, $conv1_2$, $conv3_2$ and $conv4_2$ respectively, $\lambda_i = 1$ and N_j is the number of scalars in the j^{th} layer. The optimization is carried out using Adam optimizer with a $\beta_1 = 0.5$.

We have selected the perceptual loss based on the visual quality of the morphed image that can reflect the suitability for border control applications. Let the final reconstructed image correspond to I_1 and L_1 be $\hat{I}_1$ and the corresponding updated latent code be $L_1^{'}$. We follow the same procedure mentioned above for the second image I_2 to get again a reconstructed image $\hat{I}_2$ and the corresponding updated latent code denoted $L_2^{'}$. The morphing operation is carried out by averaging the latent code as follows:

$$L_M = \frac{w_1 * L_1^{'} + w_2 * L_2^{'}}{2} \tag{5.2}$$

Finally, L_M is passed through the synthesis network to generate the morphed image that has a resolution of 1024×1024 pixels, where w_1 and w_2 indicate the weights, which we have chosen to be $w_1 = w_2 = 0.5$.

5.3.1 Differences of proposed approach with earlier works

In contrast to earlier works [15], to avoid the bias of morph generation with known set (closed-set), the StyleGAN is trained using the disjoint face dataset from FFHQ dataset[136] consisting of high quality face images. As it can be observed from Figure 5.1, the morphed face images generated using StyleGAN have higher perceptual fidelity as compared to MorGAN based morphed images and are equally comparable to landmark based morphed generation. It can be noticed that, the MorGAN [15] based morph generation indicates low-quality images that are not ICAO complaint rendering them not suitable for passport applications. As a secondary note, the MorGAN based images also indicate a poor visual similarity to the contributing subjects, while landmark based morphs exhibit stronger artifacts that are clearly visible in Figure 5.1.

Intrigued by the high fidelity of morphed face images, we take a detailed analysis guided by a sample image to compare it against the landmark based morph generation. As observed in Figure 5.2, the ghosting artifacts in landmark based morphing can be prominently seen due to the misalignment of landmarks leading to several artifacts, especially in the ocular, mouth and nose region. It is interesting to observe that the proposed StyleGAN based morph generation did not create any perceptual noise.The example demonstrates the high quality of the generated image, when compared to the MorGAN based approach.

While contrary to a landmark based morphed faces, the proposed StyleGAN based morph generation does not indicate a strong geometrical resemblance as it is the case for a landmark based approach. Motivated by such visual observations and superior quality of morph images achieved with the proposed approach and accounting for the lower geometrical resemblance of contributing subjects, we conduct a detailed analysis of threats to FRS as detailed in the next section.

5.4 Experiments and Results

In order to measure the impact of the proposed approach of morph generation, we first create a new dataset of morphed images created from 140 unique data subjects. With the newly generated morph dataset, we first investigate and report the vulnerability of FRS and compare it with the vulnerability reported in similar earlier work using MorGAN [15] and traditional landmark based morphing. Further, we also analyze the detection potential of morphed faces generated using the proposed framework with StyleGAN.

5.4.1 Database Generation

We introduce a new morphed face database created from 140 individuals that include 47 female and 93 male data subjects. The facial images are derived from the FRGC-V2 face database [33]. The newly generated database is sub-divided into two sets for training and testing that consists of independent data subjects with no overlap between the splits. The training set consists of 690 bona fide images and 1190 morphed images. The testing set consists of 580 bona fide and 1310 morphed images. To effectively analyze the vulnerability and provide a comparison to earlier works, we have generated morph images using three different techniques, which include (i) Landmark-Based (ii) MorGAN and (iii) proposed StyleGAN approach. Care is exercised to generate morphed images with similar facial appearance within same gender category. Additionally, to guarantee high quality of the newly generated dataset constraints of high quality illumination and no pose is imposed before creating the morphs. The guidelines laid out in earlier works [13] [125] are followed to obtain a database of high relevance for morphing attack detection.

5.4.2 Evaluation Metrics for Vulnerability Analysis

We measure the vulnerability of FRS following the guidelines of Frontex and setting the operating threshold to FAR = 0.1 $\%$ (for both FRS). We further follow the realistic evaluation protocol where the morph image is created by using two face images corresponding to a malicious actor and and an accomplice. We compute the vulnerability by enrolling a given morphed face image $M_{I_{1,2}}$ and probing the corresponding contributing subjects I_1 and I_2 with an image from a different FRGC-session. We further obtain the comparison scores S_1 and S_2 for both images I_1 and I_2 against the morphed image. The morphed image $M_{I_{1,2}}$ is only considered a threat if and only if the comparison scores S_1 and S_2 succeed to cross the preset threshold at FAR = 0.1$\%$. If the condition is not met, we simply consider that the morphed image is not a real threat as the comparison scores are not able to successfully verify the morphed image against both contributing subjects making the morphing attack not realistic. We term this new metric as *Fully Mated Morphed Presentation Match Rate (FMMPMR)* and compute it in general form as:

$$FMMPMR = \frac{1}{P}\sum_{M,P}(S1_M^P > \tau)AND(S2_M^P > \tau)\ldots AND(Sk_M^P > \tau) \tag{5.3}$$

Where $P = 1, 2, \ldots, p$ represent the number of attempts made by presenting

all the probe images from the contributing subject against M^{th} morphed image, $K = 1, 2, \ldots, k$ represents the number of contributing data subjects to the constitution of the generated morphed image (in our case $K = 2$), Sk_M^P represents the comparison score of the K^{th} contributing subject obtained with P^{th} attempt (in our case the P^{th} probe image from the dataset) corresponding to M^{th} morph image and τ represents the threshold value corresponding to FAR = 0.1%.

When compared to the existing metric MMPMR [125], the FMMPMR considers the number of attempts (that are assessed jointly with contributing subjects) with regards to each face morphed images and thus reflect the realistic vulnerability of a FRS. The MMPMR [125] is designed to measure vulnerability only on the morphed image in a joint set rather than a number of attempts on each morphed image. Hence, the MMPMR fails to reflect the number of attempts (by contributing subjects) made against the corresponding morphing image to determine the vulnerability of FRS.

In this work, the COTS threshold is set at $\tau = 0.5$ based on the NIST FRVT test reports as recommended by the COTS provider while ArcFace FRS threshold is set at $\tau = 0.36$ base on the face recognition trials on FRGC-v2 dataset. The higher the value of the FMMPMR the higher the threat from morphed images and correspondingly a higher vulnerability of FRS towards morphed images must be stated.

5.4.3 Results from Vulnerability analysis

In this section, we present the vulnerability analysis using two different Face Recognition Systems (FRS) (i) a Commercial off the Shelf face recognition system (COTS), Cognitec FaceVACS-SDK Version 9.4.2 [1] and (ii) an Open-source deep learning based FRS (ArcFace). To effectively benchmark the results we also compare two different State-Of-The-Art (SOTA) morph generation techniques such as landmark based morph generation [14] and MorGAN based morph generation [15].

Figure 5.4 shows the scatter plot of the comparison scores obtained from two different FRS on images obtained using three different types of morphed face generation approaches. Table 7.2 indicates the quantitative values of both MMPMR and FMMPMR computed from two different FRS for all three cases of face morph generation techniques. Based on the obtained results the key observations made are listed below:

The following are the main observations from our experiments.

[1] outcome not necessarily constitutes the best the algorithm can do

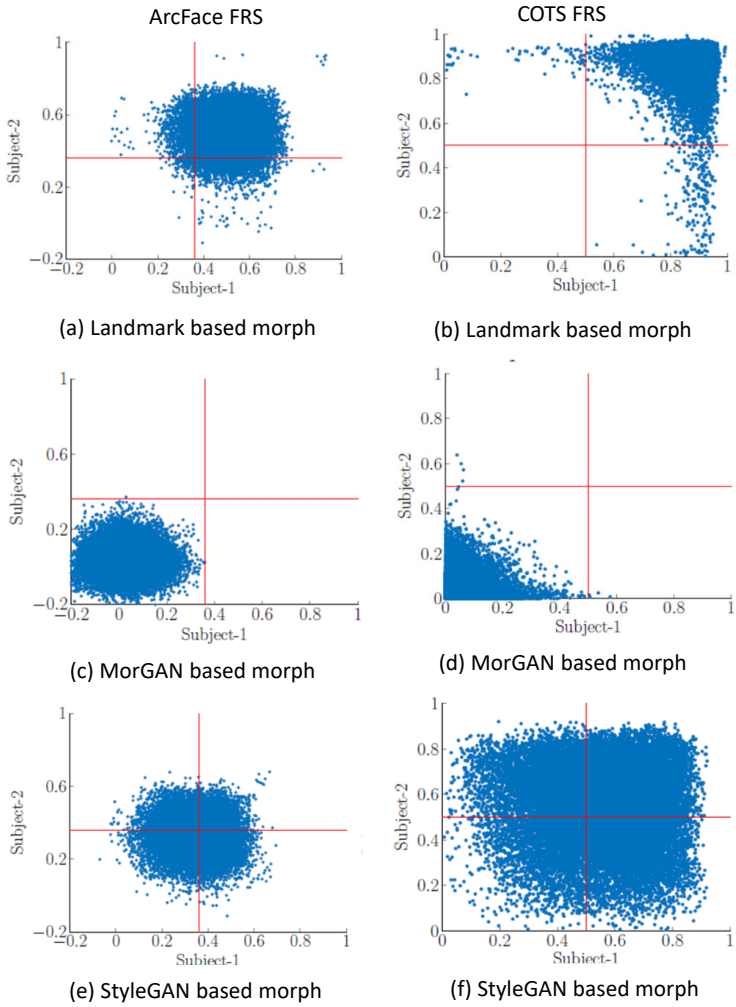


Figure 5.4: Vulnerability analysis using COTS and ArcFace. The scatter plots represents the comparison scores of morphed face image against two contributing subjects.The red lines indicate the threshold corresponding to FAR = 0.1%

Table 5.1: Vulnerability analysis FMMPMR(%) and MMPMR(%)

Morphing Type	FMMPMR (%)		MMPMR (%)	
	ArcFace	COTS	ArcFace	COTS
Landmark based Morph[14]	86.04	98.91	95.76	100
MorGAN[15]	0	0	0	0
StyleGAN	21.1	41.49	39	64.68

- Landmark based face morph generation indicates a high threat to FRS (analogously high vulnerability of FRS to such images) compared to that of two other morph generation methods. This can be attributed to the fact that the landmark based morph generation preserves both texture and geometrical structure of the morphed image corresponding to it's contributing subjects.

- The analysis of the experimental results also show that MorGAN based morph generated images do not pose a severe threat to FRS. The potential reason for this can be due to low quality generated morph image (64×64 pixels). A careful observation of the images also revealed the degradation of texture and geometry in the generated morphed images. As a caveat, we note that the MorGAN network is not re-trained (or fine-tuned) on closed dataset of contributory subjects. The conscious choice was made to investigate the generalisation of GANs for morph face image generation and study the threats.

- StyleGAN based morph generation method shows relatively higher degree of threats when compared with MorGAN. Despite higher threats, the images from proposed approach of morph generation did not compete against the landmark based methods. An introspection into this indicates the quality difference of FFHQ dataset versus the employed FRGC-V2 dataset. Specifically, the pre-trained morph generator is trained on FFHQ dataset which has very different characteristics than FRGC-V2 dataset leading the network to mimic the characteristics of the FFHQ dataset. Another aspect for the lower degree of threat is due to lack of geometric correspondence of facial structure in morphed faces when compared to that of the landmark based face morphing. The lower geometrical correspondence despite the high visual quality fails in verification stage from FRS.

- When compared to ArcFace, the COTS indicates a higher vulnerability for both landmarks and StyleGAN based morph attack detection due to the high accuracy of verifying the subjects in COTS under different data capture conditions as expected in operational scenario. Thus, COTS while making itself robust about certain degree of degraded data, also accepts the morphs to a higher degree.

- Table 7.2 also indicates the distinction between FMMPMR and MMPMR metric used to quantify the vulnerability. The MMPMR reports high values in comparison to FMMPMR as it does not account for the number of attempts per morphing image. Further, we have also measured the Relative Morph Match Rate (RMMR) [125] that can account for the True Acceptance

Rate of the FRS. Since both FRS employed in this paper have reached TAR = 100%, the RMMR is the same as the FMMPMR/MMPMR.

Table 5.2: Quantitative performance of state-of-the-art MAD techniques on StyleGAN dataset

Morphing Type	Algorithms	D-EER(%)	BPCER(%) @ APCER =5(%)	BPCER(%) @ APCER =10(%)
Landmark based Morph [14]	HoG-SVM	10.29	17.66	10
	LBP-SVM	15.42	29.15	22.98
	Color Textures	1.57	0.51	0.17
	CAN	4.8	4.63	2.4
MorGAN [15]	HoG-SVM	0	0	0
	LBP-SVM	0	0	0
	Color Textures	0	0	0
	CAN	0	0	0
StyleGAN	HoG-SVM	0.04	0	0
	LBP-SVM	0.68	0	0
	Color Textures	0	0	0
	CAN	0.36	0	0

5.4.4 Performance Metrics for MAD

The performance of Morphing Attack Detection (MAD) techniques are presented using the ISO/IEC 30107-3 metrics [21] such as Attack Presentation Classification Error Rate (APCER (%)) which defines the proportion of attack images incorrectly classified as bona fide images and Bona fide Presentation Classification Error Rate (BPCER (%)) in which bona fide images incorrectly classified as attack images [21] along with the Detection Equal Error Rate (D-EER (%)).

5.4.5 MAD Detection Performance

In this section, we report the detection performance of MAD techniques to understand the impact of different types of morphing techniques. We have therefore selected four different MAD techniques - LBP-SVM [54], HoG-SVM [5], color denoising [4], Context aggregation Network (CAN) [5] based on the recent benchmarks. Table 5.2 indicates the MAD performance on all three different morph generation techniques.

Compared to three different morph generation methods, landmark based technique indicates a relatively high challenge for the detection techniques, when compared to that of GAN based techniques. However on the same kind of morph generation approach, the recent technique based on color texture indicates the lowest error rates with D-EER(%) of 1.57(%). While it is noted that the GAN generated morphs are easier to detect, a possible reason can be attributed to the residual noise [138] that is associated with GAN in generating these morphed images. Even though

StyleGAN can generate a high quality images with a resolution of 1024×1024 pixels, the inherent noise in the generated morph images make enables to detect them. This is not the case for landmark based morph images, which do not contain such characteristic noise.

5.4.6 Limitations and Future Directions

Observing the results from the empirical evaluation of different approaches of morph generation both for threats to FRS and ability to detect the morphs, we note certain limitations in the current work as listed below.

- The GAN based morph generation does not impose the landmark correspondence leading to high quality images but not with high facial similarity in geometrical appearance to contributing subjects. This has lead to lower threat to FRS in comparison to landmark based morphs. Future works in this direction can focus on imposing such a constraint in the latent space, in order to increase the threat to FRS.

- Despite the accuracy of MAD being very high, it can be primarily attributed to digital pixel level information helping to detect the attacks. A print and scan of the the same morphed images can further reveal the real challenge in detecting the morphing attacks as the print-scan cycle looses the pixel level soft-information in the image.

The future works in this direction will lead to establishing the real threat landscape on FRS from the GAN generated morphed face images.

5.5 Conclusion

This work investigated the feasibility of generating high quality morph generation and proposed a new approach using StyleGAN. The proposed approach resulted in morphed face images with a dimension of 1024×1024 pixels and no visual artifacts. To indicate the real threat potential to FRS, the morphed face images generated from proposed StyleGAN were analyzed using a commercial FRS and an open-source FRS. Further, to provide a fair comparison to earlier works, MorGAN and Landmark based approaches were benchmarked on the same set of data by creating a new morphed face database. The set of experiments clearly indicate the that StyleGAN based morphed face images do show threats to FRS but to a much lower degree as compared to traditional landmark based morph generation techniques. While detecting the attacks stemming from GAN approaches is relatively easy in the digital domain, the real challenge of detecting them after the print-scan process is still not explored. In summary, we answer the question - *Can GAN Gen-*

erated Morphs Threaten Face Recognition Equally as Landmark Based Morphs?, our experimental results indicates with a clear no in digital domain alone.

Chapter 6

Article 2: MIPGAN—Generating Strong and High Quality Morphing Attacks Using Identity Prior Driven GAN

Haoyu Zhang, Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, Naser Damer, and Christoph Busch. *MIPGAN—generating strong and high quality morphing attacks using identity prior driven GAN*. IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(3):365–383, 2021

6.1 Abstract

Face morphing attacks target to circumvent Face Recognition Systems (FRS) by employing face images derived from multiple data subjects (e.g., accomplices and malicious actors). Morphed images can be verified against contributing data subjects with a reasonable success rate, given they have a high degree of facial resemblance. The success of morphing attacks is directly dependent on the quality of the generated morph images. We present a new approach for generating strong attacks extending our earlier framework for generating face morphs. We present a new approach using an Identity Prior Driven Generative Adversarial Network, which we refer to as MIPGAN (Morphing through Identity Prior driven GAN). The proposed MIPGAN is derived from the StyleGAN with a newly formulated loss function exploiting perceptual quality and identity factor to generate a high quality morphed facial image with minimal artefacts and with high resolution. We demonstrate the proposed approach's applicability to generate strong morphing

attacks by evaluating its vulnerability against both commercial and deep learning based Face Recognition System (FRS) and demonstrate the success rate of attacks. Extensive experiments are carried out to assess the FRS's vulnerability against the proposed morphed face generation technique on three types of data such as digital images, re-digitized (printed and scanned) images, and compressed images after re-digitization from newly generated MIPGAN Face Morph Dataset. The obtained results demonstrate that the proposed approach of morph generation poses a high threat to FRS.

6.2 Introduction

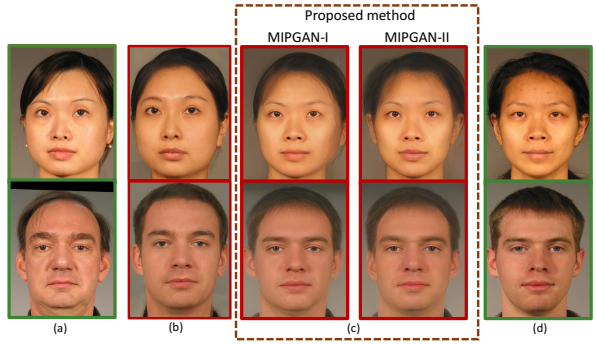


Figure 6.1: Results from StyleGAN based face morphing [1] and the proposed MIPGAN (a) Contributing subject 1 (b) StyleGAN[1] (c) Proposed method (d) Contributing subject 2

Face Recognition Systems (FRS) have provided ubiquitous ways of verifying an identity claim in many applications. FRS have been used in everyday applications from low-security applications such as smartphone unlocking to high-security applications such as identity verification in border control processes. Each of the applications mandate a chosen way of enrolment to FRS where either a supervised enrolment is carried out (for instance in on-boarding at bank premises) or unsupervised enrolment is requested (on-boarding for banking applications from home). While it provides a high degree of flexibility and convenience to users to initiate an enrolment process in an unsupervised manner, this potentially leads to a security risk: Without supervision, a data subject enrolling into the FRS can submit a face image which is manipulated, a printed face image, an image displayed from an electronic screen (e.g., iPad) or a silicone latex face mask [133]. In order to mitigate such attacks at the enrolment level, it is therefore essential to have a robust attack detection mechanism. While a number of works in recent years have

been proposed on both conducting such attacks and detecting the attacks in a robust manner for printed attacks, display attacks and mask attacks, in this work we focus on a new kind of attack referred popularly as *Morphing Attack*.

Face morphing is the process of combining two or more face images to generate a single face image that can resemble visually to all the contributing face images to a greater degree [22]. A good quality morphed face image is also effective in verifying against all contributing subjects by obtaining a comparison score that exceeds the pre-determined threshold (i.e., passes through FRS) [22, 83, 54, 38]. While morphing can be conducted using multiple face images of different subjects, the effectiveness of morphed images is reported when the face images of similar ethnicity, gender and age group are considered [13, 38, 125]. This is primarily due to the fact that a morphed image should not only defeat the FRS but should also provide high visual similarity, in order to convince a human expert in a visual comparison process.

Face morphing attacks threaten FRS due to the current practices in the ID-document application process, where the biometric enrolment is carried out in an unsupervised manner in many countries. Countries like the U.K. and New Zealand allow citizens to upload a digital face image for various applications such as passport renewal [24] and visa application [23]. The capture process for such images is unsupervised. In a similar manner, many Asian countries and European countries (e.g. in The Netherlands [139]) request the applicant to submit a scanned face image for passport/visa/identity-card applications. Given that the images are captured and submitted in an unsupervised setting, the applicant has vast opportunities to upload a morphed image with malicious intent underlining the need for robust Morphing Attack Detection (MAD) mechanisms.

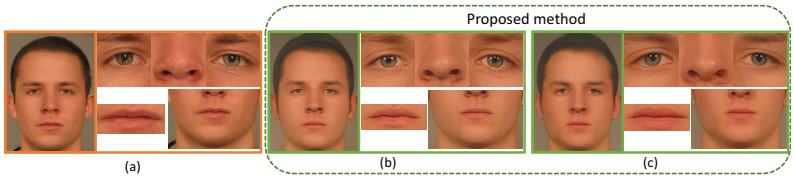


Figure 6.2: Details of segmented components in morphs generated by earlier method based on StyleGAN [1] and proposed MIPGAN (a) StyleGAN [1] (b) MIPGAN-I (c) MIPGAN-II.

6.3 Related Work

While morphing attacks have been studied in recent years, most of the attacks are conducted using the morphed images created using facial landmarks-based

approaches needing high a degree of supervision to first determine the facial landmarks, thereupon align them and then finally blend them to generate morphs. The common set of procedures for warping/blending includes Free Form Deformation (FFD) [60] [61], Deformation by moving least squares [62], deformation based on mass spring [63], Bayesian framework based morphing [140] and Delaunay triangulation based morphing [95] [141] [81] [64] [98]. Due to inadvertent artefacts caused by pixel/region-based morphing, the images need additional work in refining the signal to create highly realistic morph images. A set of post processing steps are usually included as illustrated in number of works [64] [65] [66]. Generally, some set of post processing techniques such as image smoothing, image sharpening, edge correction, histogram equalization, manual retouching, image enhancement to improve the brightness and contrast are used to eliminate the artefacts generated during the morphing process. In a parallel direction, morphed face images can also be generated using landmarks-based methods available in open-source resources like GIMP/GAP and OpenCV. Morphs generated using GIMP/GAP technique are more efficient with respect to a good quality of the resulting image (i.e., less noticeable artefacts) as pixels are aligned manually. Despite the minimal amount of effort needed for creating morphs using such approaches, a significant amount of effort needs to be dedicated to correcting artefacts. Additionally, commercial solutions like Face Fusion [142] and FantaMorph [143] can also generate good quality morphed images with limited manual intervention. Although some steps can be excluded in creating the morphs, it is very critical to meet the face image quality standards laid out by the International Civil Aviation Organization (ICAO) [144][130] for electronic Machine Readable Travel Document (eMRTD) and deployment of biometric identification applications.

6.3.1 GAN Based Face Morph Generation

In an attempt to overcome the cumbersome efforts of manually creating (semi-automated) morphed images, a fully automated approach using a Generative Adversarial Network (GAN) was proposed by Damer et al.[15]. Unlike the supervision required in the mark-up of landmarks and aligning the face images in a (partially) manual process, GAN-based techniques synthesise morphed images directly by merging two facial images in the latent space. In the work by Damer et al.[15], the proposed MorGAN architecture for morph generation basically employed a generator constituting encoders, decoders and a discriminator. The generator was trained to generate images with the dimension 64×64 pixels which is a key limiting factor of the attack, as most commercial FRS will reject images that do not meet the ICAO standard that requires a minimum Inter-Eye Distance (IED) of 90 pixels. The empirical evaluation of generated morph images using MorGAN in a vulnerability analysis against two commercial FRS indicated that those MorGAN

morphs fail to meet both quality standards and the verification threshold of the FRS [1]. Motivated to address the deficiency of the MorGAN architecture, in our recent work [1] [1] we proposed an approach based on the StyleGAN architecture [136] to increase the spatial dimension to 1024×1024 and thus to improve the face image quality. Unlike the previous approach of MorGAN [15], StyleGAN [1] achieves better spatial resolution by embedding the images in the intermediate latent space. With the increased spatial dimension of resulting morphed images from our recently proposed architecture, we not only demonstrated that the images meet quality standards but also have a reasonable success rate when attacking commercial FRS [1].

6.3.2 Limitations of GAN Based Face Morph Generation and Our Contributions

While our earlier work [1] indicated that better GAN architectures could result in superior quality morphs and could attack an FRS in general, we also acknowledge the limited threats that exist for Commercial-Off-The-Shelf (COTS) FRS, as merely a subset of morphed images was accepted. Only approximately 50% of the generated morph images were verified successfully against probe images from a contributing subject. Thus the empirical evaluation in our earlier work has shown that the attack was yet not very effective [1] for a COTS FRS[8] and an open-source FRS based on ArcFace [10]. We must state that up to now FRS are not very vulnerable to GAN-based morphing attacks unlike to landmarks-based morphing attacks. With a clear introspection into this aspect, we notice that the resulting morphed images from our earlier work [1] does not retain a high degree of facial similarity to both contributing subjects. With lower similarity to contributing subjects in terms of facial structures, the FRS do not attribute a high comparison score, as anticipated. In other words, the missing enforcement of identity information of contributing subjects will lead to a high visual quality facial image but with lower face similarity to contributing face characteristics.

In an effort to make the attacks stronger such that both subjects can be verified with a good success rate, in this work, we extend our previous architecture to generate morphs by including the identity priors before the generation of morphed faces. We now refer to this approach as *MIPGAN (Morphing through Identity Prior driven GAN)*. We propose two variants of our approach named as MIPGAN-I and MIPGAN-II based on the employed GAN being StyleGAN or StyleGAN2 respectively [136, 138]. With the inclusion of a new loss function in our proposed architecture, we increase the attack success rate against commercial-off-the-shelf (COTS) FRS and deep learning based FRS. Figure 6.1 shows the example of

[1] The preliminary work results were published at IWBF-2020 in April, 2020.

morphed face images generated using proposed MIPGAN along with outputs of both the variants. To further achieve superior quality face morphs, we also customize the newly designed loss function to account for ghosting and blurring artefacts in an end-to-end manner with no human or manual intervention eliminating the need for a high degree of interaction. As noted in Figure 6.2, the results from MIPGAN-I and MIPGAN-II is more coherent in retaining structural similarity as compared to our earlier architecture [1]. With the updated architecture to generate high-quality morphs which preserve both identity information and structural correspondence, we evaluate the applicability in creating stronger attacks by creating a large-scale dataset of morphed images by employing the face images derived from the FRGC-V2 face database [33]. The created dataset of 1270 bona fide images and 2500 morphed images is first evaluated to measure the attack success rate by verifying the morphed images against the contributing subjects using a commercial FRS from Cognitec [8]. In addition to measuring the attack success rate for digital images, we also extend our work by printing and scanning (re-digitizing) the dataset. We check the consistency of the attack success rate, unlike our earlier work which was limited to an investigation on digital images alone [1]. We also include the experiments on assessing the impact of compression (down to 15kb following ICAO guidelines) of printed and scanned face images that simulate the real-life e-passport application scenario. The key motivation to extend our work in this direction is, to mimic the passport application process that is operated in many European countries and Asian countries, which all accept printed-and-scanned facial images in the application process for an identity document (e.g. passports).

With the extensive experimental results indicating a highly satisfactory attack success rate, we also evaluate a set of MAD algorithms to benchmark the detection capabilities. To this extent, we evaluate two state-of-the-art MAD approaches on digital morphed images, re-digitized and compressed morphed images after re-digitizing. Thus, we comprehensively cover the potential morphing attacks in the digital domain and the re-digitized domain. While we note the earlier works [1] arguing that attacks in the digital domain can be detected by studying the cues such as residual noise in morphing [4], patterns of noise from morphed images, histogram features of textures or the deep features [83], we also investigate the MAD capabilities for re-digitized images which do not exhibit the similar features (residual noise) as the print-scan process eliminates the digital cues and presents another set of variations. Specifically, given the nature of the dataset in which we have only a single suspected morphed image, for which we must determine either the morph or the bona fide class, we resort to Single Image based MAD (S-MAD) approaches using two recent but robust approaches using hybrid and ensemble features [91, 4, 6, 145].

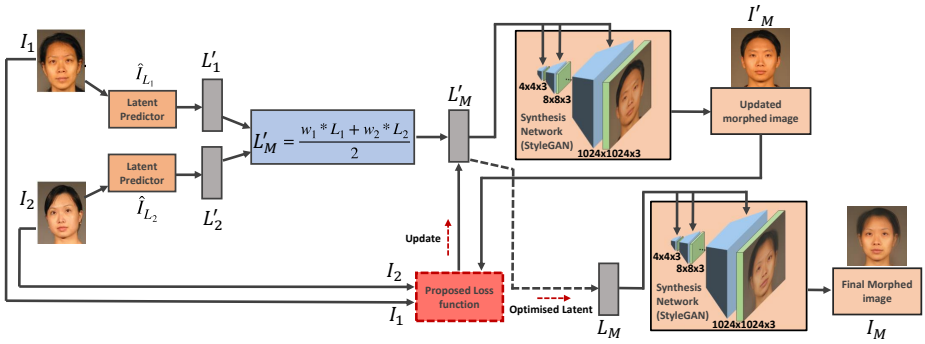


Figure 6.3: Block diagram of the proposed MIPGAN for generating high quality morphed face images

We therefore present a summary of contributions of this work as listed below:

- We present a novel approach of generating morphed face images through GAN architecture with enforced identity priors and a customized novel loss function to generate highly realistic images which we refer as *MIPGAN (Morphing through Identity Prior driven GAN)*. We present two variants of the proposed approach for generating attacks with a high success rate.
- The proposed approach (both variants) is benchmarked to measure the attack success rate by verifying COTS and deep learning based FRS through studying the vulnerability using a newly generated dataset from our proposed architecture which is referred as *MIPGAN Face Morph Dataset*.
- Human observer analysis for detecting morphs generated by the proposed and existing morphing attack methods is presented.
- Analysis of the perceptual quality metrics to illustrate the visual quality of the generated morph images is presented.
- Extensive experiments on three different data types such as (a) digital morphed images (b) print-scan morphed image (c) print-scan morphed images with compression are presented to cover the full spectrum of passport application process under morphing attacks.
- The generated images are also benchmarked against the existing MAD approaches both in digital form and the re-digitized form to provide the insights on detection challenges of SOTA approaches. We also present a generalizability study on MAD schemes by training one kind of morph genera-

tion and testing on a different kind of morph generation approach to indicate directions to future works.

In the rest of the paper, Section 6.4 describes the new architecture along with the newly designed loss function to generate high-quality morphs. Section 6.5 provides the details on the quantitative experiments indicating the vulnerability of FRS and the detection challenge. With the set of remarks and future works in this direction, we draw the conclusion in Section 6.7.

6.4 Proposed Morphed Face Generation

Figure 6.3 presents the block diagram of the proposed morphed face image generation using MIPGAN. The proposed method is based on the end-to-end optimization using a new loss function that can preserve the identity of the generated morphed face image through enforced identity priors. The proposed MIPGAN framework is designed independently on two different GAN models based on StyleGAN [136] and StyleGAN2 [138] model. We refer to the proposed scheme with StyleGAN as MIPGAN-I and with StyleGAN2 as MIPGAN-II respectively. Given the face images from the accomplice (I_1) (contributing subject 1) and the malicious (I_2) (contributing subject 2) data subjects, we predict the corresponding latent vectors $L_1^{'}$ and $L_2^{'}$ in the first step. In this work, we have employed the open-source pre-trained prediction models trained to predict the corresponding latent vector given an input image. Hence, L_1' and L_2' are predictions from the final output layer of the model, which is further reshaped. Since MIPGAN-I and MIPGAN-II are based on pre-trained StyleGAN [136] and StyleGAN2 [138] model respectively, we used two different open-source pre-trained models for prediction. Both of the prediction models employ $ResNet50$ [146] as backbone. The model for MIPGAN-I (StyleGAN) uses one convolution layer and two tree-connected layers [147] to map the output of $ResNet50$ into the final latent vector with the size of $(18, 512)$. In comparison, the model for MIPGAN-II (StyleGAN2) just uses one fully-connected layer to achieve the mapping. The predicted latent vectors thus provide the initialization for the morphed face generation that is obtained using a weighted linear average of $L_1^{'}$ and $L_2^{'}$ as follows:

$$L_M^{'} = \frac{w_1 * L_1^{'} + w_2 * L_2^{'}}{2}, \tag{6.1}$$

where w_1 and w_2 indicate the weights, which we have chosen to be $w_1 = w_2 = 1$. Equal weights are selected as shown in earlier work [3] where the morphing images generated with equal weights pose higher vulnerability to COTS FRS. Finally, $L_M^{'}$ is passed through the synthesis network (independently from StyleGAN [136] and

StyleGAN2 [138] model) to generate the corresponding morphed image $I_M^{'}$ that has a resolution of 1024×1024 pixels. The generated morphed face image $I_M^{'}$ is then optimised using the proposed loss function to generate the high quality morphed face image. In the following section, we discuss the loss function to optimism the latent vector obtained using Equation 6.1.

6.4.1 Proposed Loss Function

The proposed loss function is based on both perceptual fidelity, quality and identity factors that can facilitate high-quality face morph generation. The common issue with the GAN-based morph generation is the presence of ghost artefacts and blurring issues. We employ the perceptual loss with multiple layers to eliminate such effects as given by Eqn 6.2.

$$\begin{aligned} Loss_{Perceptual} = & \frac{1}{2}\sum_{i}\frac{1}{N_i}||F_i(I_1) - F_i(I_M^{'})||_2^2 \\ & + \frac{1}{2}\sum_{i}\frac{1}{N_i}||F_i(I_2) - F_i(I_M^{'})||_2^2, \end{aligned} \tag{6.2}$$

where N_i denotes the number of features in layer i and F_i denotes features in layer i of the perceptual network (VGG-16 in our case). For the combination of perceptual layers, we choose $conv1_1$, $conv1_2$, $conv2_2$, $conv3_3$ inspired by [148]. Compared with the original combination of layers $conv1_2$, $conv2_2$, $conv3_3$, $conv4_3$ [149], our design measures low-level features instead of high-level features like style of an image and is closer to our goal of morphing faces with high quality.

The main goal of this paper is to generate the morphed face images that can significantly attack FRS. In order to achieve this, we have introduced the identity loss function based on the feedback from FRS. We employ Arcface [10] - a deep learning based FRS because of its robust and accurate performance to obtain feedback on generated morphed face images. Specifically, we employ a pre-trained embedding extractor with $ResNet50$ as the backbone to extract the unit embedding vectors and define the identity loss by their cosine distance to improve the morph generation process as given by Eqn 6.3.

$$Loss_{Identity} = \frac{(1 - \frac{\vec{v}_1 \cdot \vec{v}_M}{\|\vec{v}_1\|\|\vec{v}_M\|}) + (1 - \frac{\vec{v}_2 \cdot \vec{v}_M}{\|\vec{v}_2\|\|\vec{v}_M\|})}{2}, \tag{6.3}$$

where $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_M$ respectively denotes the embedding vectors which are extracted from image $I_1, I_2, I_M^{'}$ respectively.

To further prove the loss function is differential for the morphed embedding vector $\vec{v}_M$, we define x_d, y_d, z_d to be the value of vector $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_M$ in dimension d

respectively and $d' \neq d$ to be other dimensions except d. The expanded identity loss function and its partial derivative are:

$$Loss_{Identity} = \frac{(1 - \frac{\sum_d x_d z_d}{\|\vec{v}_1\|\|\vec{v}_M\|}) + (1 - \frac{\sum_d y_d z_d}{\|\vec{v}_2\|\|\vec{v}_M\|})}{2}, \tag{6.4}$$

$$\begin{aligned}\frac{\partial Loss_{Identity}}{\partial z_d} = 1 &- \frac{x_d}{2\|\vec{v}_1\|}\frac{\partial}{\partial z_d}(\frac{z_d}{\sqrt{z_d^2 + \sum_{d' \neq d} z_{d'}^2}}) \\ &- \frac{y_d}{2\|\vec{v}_2\|}\frac{\partial}{\partial z_d}(\frac{z_d}{\sqrt{z_d^2 + \sum_{d' \neq d} z_{d'}^2}}),\end{aligned} \tag{6.5}$$

$$\begin{aligned}\frac{\partial}{\partial z_d}(\frac{z_d}{\sqrt{z_d^2 + \sum_{d' \neq d} z_{d'}^2}}) &= \frac{1}{\sqrt{z_d^2 + \sum_{d' \neq d} z_{d'}^2}} \\ &\quad + \frac{2z_d^2}{-2(z_d^2 + \sum_{d' \neq d} z_{d'}^2)^{\frac{3}{2}}} \\ &= \frac{\sum_{d' \neq d} z_{d'}^2}{(z_d^2 + \sum_{d' \neq d} z_{d'}^2)^{\frac{3}{2}}},\end{aligned}$$

$$\frac{\partial Loss_{Identity}}{\partial z_d} = 1 - \frac{(\frac{x_d}{2\|\vec{v}_1\|} + \frac{y_d}{2\|\vec{v}_2\|}) \sum_{d' \neq d} z_{d'}^2}{(z_d^2 + \sum_{d' \neq d} z_{d'}^2)^{\frac{3}{2}}}. \tag{6.6}$$

For any value $z_d = z'_d$, it is obvious that:

$$\begin{aligned}&\lim_{\Delta z_d \to 0} \frac{\partial Loss_{Identity}(z'_d + \Delta z_d)}{\partial z_d} \\ &= \lim_{\Delta z_d \to 0}(1 - \frac{(\frac{x_d}{2\|\vec{v}_1\|} + \frac{y_d}{2\|\vec{v}_2\|}) \sum_{d' \neq d} z_{d'}^2}{((z'_d + \Delta z_d)^2 + \sum_{d' \neq d} z_{d'}^2)^{\frac{3}{2}}}) \\ &= 1 - \frac{(\frac{x_d}{2\|\vec{v}_1\|} + \frac{y_d}{2\|\vec{v}_2\|}) \sum_{d' \neq d} z_{d'}^2}{(z_d'^2 + \sum_{d' \neq d} z_{d'}^2)^{\frac{3}{2}}} \\ &= \frac{\partial Loss_{Identity}(z'_d)}{\partial z_d}.\end{aligned}$$

Hence, for any dimension of d, the partial derivative of the identity loss function is continuous.

It is interesting to note that the identity loss based on the Arcface feature extractor model is trained to maximize the face class separability and thus is more sensitive to face attributes. Hence, only optimising the identity loss cannot achieve the same reconstruction performance as the perceptual loss but applying it on the face region can effectively control the generated attributes to be recognized as both subjects.

To solve the imbalance between different subjects, we introduce an identity difference loss as given by Eqn 6.7.

$$Loss_{ID-Diff} = |(1 - \frac{\vec{v}_1 \cdot \vec{v}_M}{\|\vec{v}_1\|\|\vec{v}_M\|}) - (1 - \frac{\vec{v}_2 \cdot \vec{v}_M}{\|\vec{v}_2\|\|\vec{v}_M\|})|. \tag{6.7}$$

With the idea of the Lagrange multiplier, it adds a constraint to the optimization process to force the cosine distance between morph embedding and each of the two reference embeddings to be the same. Since $Loss_{ID-Diff}$ is usually small with a value less than 1, we apply $L1$ loss on the difference of two cosine distance terms to avoid the vanishing gradient problem.

Finally, in order to improve the structural visibility of the generated morphed face image, we also apply the Multi-Scale Structural Similarity (MS-SSIM) loss $L_{MS-SSIM}$ to measure the similarity in structure [150]. Given two discrete non-negative signals (images in our case) x and y, luminance, contrast and structure comparison measures were given by l, c, s as computed using Eqn 6.8.

$$\begin{aligned} l(x,y) &= \frac{(2\mu_x 2\mu_y + (K_1 L)^2)}{\mu_x^2 + \mu_y^2 + (K_1 L)^2}, \\ c(x,y) &= \frac{(2\sigma_x 2\sigma_y + (K_2 L)^2)}{\sigma_x^2 + \sigma_y^2 + (K_2 L)^2}, \end{aligned} \tag{6.8}$$

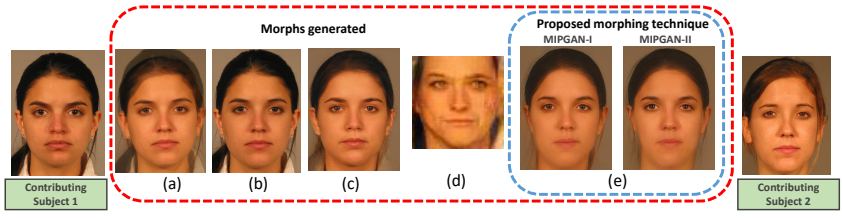


Figure 6.4: Qualitative results of proposed MIPGAN together with existing GAN based face morph generation methods (a) Landmark-I [13] (b) Landmark-II [14] (c) StyleGAN[1] (d) MorGAN [15] (e) Proposed method

where μ_x, σ_x and σ_{xy} denotes the mean of x, the variance of x and the covariance of x and y respectively. L is the dynamic range of the signal and $K_1 \ll 1, K_2 \ll 1$ are two constant scalars. The MSSSIM loss $L_{MS-SSIM}$ is further defined by Eqn 6.9.

$$\begin{aligned} MSSSIM(x,y) =& [l_J(x,y)]^{\alpha_J} \cdot \prod_{j=1}^{J} [c_j(x,y)]^{\beta_j} [s_j(x,y)]^{\gamma_j}, \\ L_{MS-SSIM} =& \frac{1}{2}(1 - MSSSIM(I_1, I'_M)) \\ &+ \frac{1}{2}(1 - MSSSIM(I_2, I'_M)), \end{aligned} \tag{6.9}$$

where $j = 1, 2, \ldots, J$ represents the j^{th} scale and α_j, β_j and γ_j are the factors of relative importance. As suggested in [150], we also set $\alpha_j = \beta_j = \gamma_j$, $\sum_{j=1}^{J} \gamma_j = 1$ and use the resulting parameters $\beta_1 = \gamma_1 = 0.0448, \beta_2 = \gamma_2 = 0.2856, \beta_3 = \gamma_3 = 0.3001, \beta_4 = \gamma_4 = 0.2363, \alpha_5 = \beta_5 = \gamma_5 = 0.1333$.

Thus, the proposed loss function can be formulated as:

$$\begin{aligned} Loss =& \lambda_1 Loss_{Perceptual} + \lambda_2 Loss_{Identity} \\ &+ \lambda_3 Loss_{MS-SSIM} + \lambda_4 Loss_{ID-Diff}, \end{aligned} \tag{6.10}$$

where λ_1, λ_2, λ_3 and λ_4 are the hyper-parameters that are set to achieve both stable and generalized convergence. In this work, we empirically set $\lambda_1 = 0.0002$, $\lambda_2 = 10$, $\lambda_3 = 1$ and $\lambda_4 = 1$.

6.4.2 Training and Optimization

The training and optimization of the proposed method are carried out on Tensorflow version 1.13 and version 1.14 for StyleGAN and StyleGAN2, respectively. The optimization is carried out using NVIDIA GTX 1070 8 GB GPU with CUDA version 10.0 and CUDNN version 7.5 and NVIDIA Tesla P100 PCIE 16 GB GPU. The Adam optimizer with hyper-parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 1 \times 10^{-8}$ as recommended in the original paper [151] is employed on this work. The list of morphing pairs is generated in advance with careful considerations to gender. During each optimization process of 150 iterations, the learning rate is initially set to $\eta = 0.03$ with an exponential decay per 6 iterations of $\eta_{new} = \eta * 0.95$.

Figure 6.4 illustrates the qualitative results of the proposed MIPGAN framework based on StyleGAN and StyleGAN2. Further, the qualitative results of the existing methods based on StyleGAN [1] and MorGAN [15] are provided alongside for the convenience of the reader in the same figure. It is interesting to note that

the proposed MIPGAN generated face morph images indicate both perceptual and geometric features correspondence to both contributing subjects (for instance, malicious actor and accomplice).

6.5 Experiments and results

This section presents and discusses the experimental protocols, datasets, and quantitative results of the proposed face morphing technique. The images generated from the proposed MIPGAN-I and MIPGAN-II architectures are compared with the state-of-the-art techniques based on both facial landmarks [13] and StyleGAN based morph generation [1]. The effectiveness of the face morphing generation is quantitatively evaluated by benchmarking the vulnerability of the COTS FRS

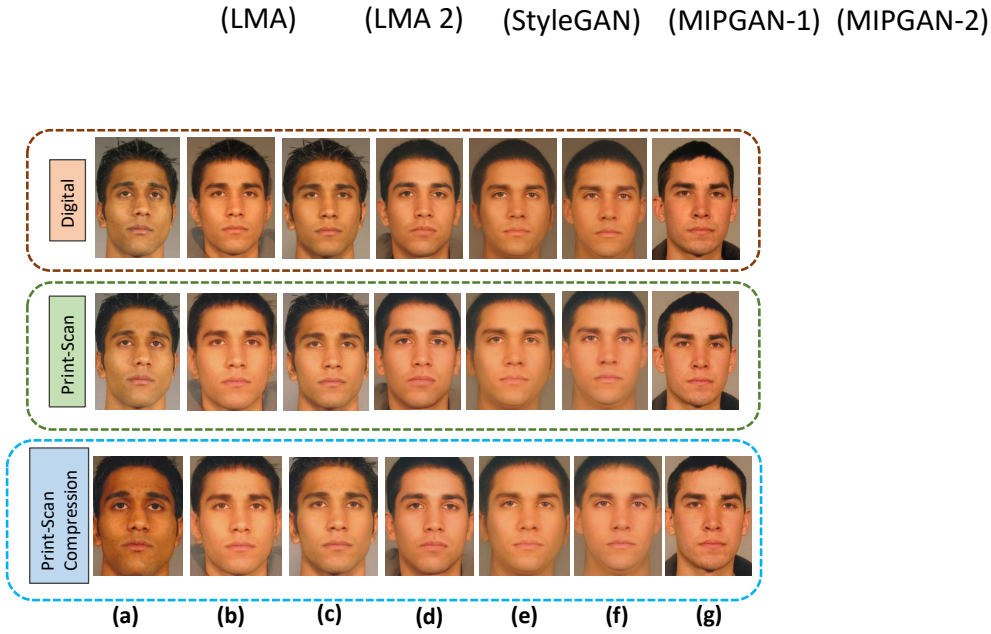


Figure 6.5: Illustration of morphing in digital, print-scan and print-scan compression data (a) Contributing subject 1 (b) Landmark-I [13] (c) Landmark-II [14] (d) StyleGAN [1] (e) MIPGAN-I (f) MIPGAN-II (g) Contributing subject 2

6.5.1 MIPGAN Face Morph Dataset

We employ the face images from FRGC-V2 face database [33] to generate the *MIPGAN Face Morph Dataset* consisting of morphed face images using both state-of-the-art and the proposed MIPGAN technique. We have selected 140 unique data subjects from the FRGC dataset by considering the high-quality face images captured in constrained conditions that resemble the passport image quality.

Among 140 data subjects, 47 data subjects are female and 93 data subjects are male. Each data subject has a variable size of 7-21 additional captured samples, resulting for the whole dataset to have 1270 samples corresponding to 140 data subjects. We employ three different face morph generation techniques based on facial landmarks constrained by Delaunay triangulation with blending [13] we term this as Landmarks-I, landmarks-based techniques with automatic post processing and color equalisation [14], we term this as Landmarks-II and StyleGAN [1]. We do not consider MorGAN [15] [72] based face morph generation as it was earlier demonstrated that MorGAN does not generate ICAO compliant images and thus makes COTS FRS not vulnerable [1]. All the samples are pre-processed to meet the ICAO standards [130] and morphing is carried out by following the guidelines outlined earlier [13] [125], i.e, careful selection of subjects based on gender and similarity score using a FRS, in order to have realistic attacks.

To effectively evaluate the proposed method's quantitative performance and the existing techniques, we create three different types of attacks from morphed images, such as **Digital morphed images**: Morphed face images that are obtained from the morph generation process in the digital domain. **Print-scanned morphed images**: The digital morphed and bona fide images are printed and then scanned (or re-digitized) to simulate the passport application process. We have employed a DNP-DS820 [152] dye-sublimation photo printer to generate the prints of the digital morphed and bona fide face images in this work. The use of a dye-sublimation photo printer guarantees high-quality photo printing (generally used for a passport application) and makes sure that printed photos are free from dotted patterns (or individual droplets of ink) that are resulting from the printing process of conventional printers. Each of these printed photos is then scanned (or re-digitized) using the Canon office scanner to have 300 dpi as suggested in ICAO standards [130]. **Print-scanned compressed morphed images**: The printed and scanned images (both morphed and bona fide) are compressed to have a size of 15kb that makes it suitable to store in the e-passport. This process reflects the real-life scenario of face image storage in passport systems. Thus, the overall dataset has 2500×3 (types of morph data) $\times$ 4 types of morph generation technique = $30,000$ morph samples and 1270×3 (types of morph data) $\times$ 4 types of morph generation technique = $15,240$ bona fide samples. Figure 6.5 illustrates the three data types of attacks that are used to evaluate the effectiveness of the proposed method and the existing methods of face morph generation. It is evident that the visual quality of the images vary largely for different attack types (for instance, the digital data attack indicates the best quality and print-scan with compression indicates the lowest quality).

Table 6.1: Quantitative evaluation of vulnerability of COTS Cognitec-FRS [8] from various morph generation approaches. Note that, since FNMR = 0 @ FMR = 0.1% for Cognitec-FRS [8] following Eq. 11.2 and 11.3, the value of RMMR is equal to MMPMR/FMMPMR. Therefore, we have not entered RMMR separately in the Table above.

Morph generation type	MMPMR/RMMR(%)	FMMPMR/RMMR(%)	MMPMR/RMMR(%)	FMMPMR/RMMR(%)	MMPMR/RMMR(%)	FMMPMR/RMMR(%)
	Digital		**Print-Scan**		**Print-Scan with compression**	
	Male		Male		Male	
Landmark-I [13]	100	98.77	97.23	97.34	97.38	96.95
Landmark-II [14]	87.29	76.86	90.32	78.23	88.78	77.14
StyleGAN [1]	63.51	41.27	60.59	39.51	57.12	35.05
MIPGAN-I	93.35	83.08	91.72	80.55	91.07	77.89
MIPGAN-II	92.22	80.45	90.74	77.67	89.16	73.47
	Female		Female		Female	
Landmark-I [13]	100	99.26	99.37	99.02	99.78	99.24
Landmark-II [14]	94.28	88.67	98.22	91.48	98.16	90.97
StyleGAN [1]	68.75	42.62	66.45	42.01	66.45	40.49
MIPGAN-I	98.57	93.11	98.16	91.22	96.12	90.52
MIPGAN-II	95.91	87.66	95.30	86.26	94.69	84.47
	Combined		Combined		Combined	
Landmark-I [13]	**100**	**98.84**	**97.64**	**97.60**	**97.84**	**97.30**
Landmark-II [14]	88.65	78.72	91.85	81.56	90.61	79.33
StyleGAN [1]	64.68	41.49	61.72	39.90	58.92	35.89
MIPGAN-I	94.36	84.65	92.97	82.23	92.29	79.88
MIPGAN-II	92.93	81.59	80.56	79.02	90.24	75.20

Table 6.2: Quantitative evaluation of vulnerability of VGGFace2 [9] FRS from various morph generation approaches. Note that, since FNMR = 0 @ FMR = 0.1% for VGGFace2 [9] following Eq. 11.2 and 11.3, the value of RMMR is equal to MMPMR/FMMPMR. Therefore, we have not entered RMMR separately in the Table above.

Morph generation type	MMPMR/RMMR(%)	FMMPMR/RMMR(%)	MMPMR/RMMR(%)	FMMPMR/RMMR(%)	MMPMR/RMMR(%)	FMMPMR/RMMR(%)
	Digital		**Print-Scan**		**Print-Scan with compression**	
	Male		Male		Male	
Landmark-I [13]	85.59	70.80	83.91	68.20	83.86	67.73
Landmark-II [14]	63.27	46.55	63.12	46.37	63.72	46.80
StyleGAN [1]	61.19	41.01	61.68	41.43	61.68	41.04
MIPGAN-I	76.96	59.24	76.96	57.16	76.07	57.31
MIPGAN-II	75.73	56.97	72.87	54.57	72.87	54.43
	Female		Female		Female	
Landmark-I [13]	96.03	83.55	93.95	82.02	93.32	81.39
Landmark-II [14]	87.76	71.85	89.39	73.82	89.80	74.27
StyleGAN [1]	80.42	59.19	79.79	59.10	78.54	58.83
MIPGAN-I	90.41	76.68	89.39	75.95	89.18	75.85
MIPGAN-II	88.98	75.42	87.96	74.54	88.37	74.90
	Combined		Combined		Combined	
Landmark-I [13]	**87.64**	**72.82**	**85.87**	**70.39**	**85.71**	**69.90**
Landmark-II [14]	68.07	50.64	68.27	50.80	68.86	51.28
StyleGAN [1]	64.92	43.91	65.20	44.25	64.96	43.88
MIPGAN-I	79.61	62.06	79.41	60.19	78.66	60.30
MIPGAN-II	78.34	59.95	75.84	57.80	75.92	57.73

Table 6.3: Quantitative evaluation of vulnerability of Arcface [10] FRS from various morph generation approaches. Note that, since FNMR = 0 @ FMR = 0.1% for Arcface [10] following Eq. 11.2 and 11.3, the value of RMMR is equal to MMPMR/FMMPMR. Therefore, we have not entered RMMR separately in the Table above.

Morph generation type	MMPMR/RMMR(%)	FMMPMR/RMMR(%)	MMPMR/RMMR(%)	FMMPMR/RMMR(%)	MMPMR/RMMR(%)	FMMPMR/RMMR(%)
	Digital		**Print-Scan**		**Print-Scan with compression**	
	Male		Male		Male	
Landmark-I [13]	99.60	98.19	97.38	96.88	97.33	96.70
Landmark-II [14]	91.09	84.62	93.45	86.42	93.60	86.02
StyleGAN[1]	70.99	55.76	73.86	58.67	73.32	58.26
MIPGAN-I	93.70	85.17	92.76	84.39	93.01	84.41
MIPGAN-II	93.65	86.45	93.55	85.30	93.25	85.06
	Female		Female		Female	
Landmark-I [13]	99.79	97.01	99.79	96.91	99.79	97.01
Landmark-II [14]	94.49	86.71	97.76	89.76	98.16	89.17
StyleGAN [1]	80.21	63.22	82.71	65.70	82.71	66.05
MIPGAN-I	97.35	89.53	97.96	91.02	97.76	91.02
MIPGAN-II	96.33	89.47	95.92	89.33	96.12	89.42
	Combined		Combined		Combined	
Landmark-I [13]	**99.68**	**98.00**	**97.88**	**96.89**	**97.84**	**96.75**
Landmark-II [14]	91.79	84.96	94.33	86.96	94.53	86.54
StyleGAN [1]	72.80	56.95	75.60	59.79	75.16	59.51
MIPGAN-I	94.45	85.94	93.81	85.46	93.97	85.48
MIPGAN-II	94.21	86.94	94.05	85.95	93.85	85.77

Table 6.4: Quantitative evaluation of vulnerability of COTS Neurotec [11] FRS from various morph generation approaches. Note that, since FNMR = 0 @ FMR = 0.1% for COTS Neurotec [11] following Eq. 11.2 and 11.3, the value of RMMR is equal to MMPMR/FMMPMR. Therefore, we have not entered RMMR separately in the Table above.

Morph generation type	MMPMR/RMMR(%)	FMMPMR/RMMR(%)	MMPMR/RMMR(%)	FMMPMR/RMMR(%)	MMPMR/RMMR(%)	FMMPMR/RMMR(%)
	Digital		**Print-Scan**		**Print-Scan with compression**	
	Male		Male		Male	
Landmark-I [13]	99.40	94.70	95.45	83.71	93.23	77.16
Landmark-II [14]	88.99	68.51	88.92	63.31	80.62	53.63
StyleGAN [1]	52.26	26.47	31.88	12.98	31.60	12.15
MIPGAN-I	58.18	32.56	32.59	25.33	57.6	53.52
MIPGAN-II	53.16	29.65	47.41	20.71	50.73	23.72
	Female		Female		Female	
Landmark-I [13]	100	99.25	100	98.11	98.74	91.18
Landmark-II [14]	94.69	85.96	97.49	84.92	95.40	78.89
StyleGAN [1]	70.60	50.13	55.20	25.72	52.39	26.19
MIPGAN-I	80.98	56.29	73.06	46.87	77.89	30.50
MIPGAN-II	74.79	49.45	69.59	42.17	70.73	46.18
	Combined		Combined		Combined	
Landmark-I [13]	**99.51**	**95.37**	**96.32**	**85.43**	**94.30**	**79.25**
Landmark-II [14]	90.16	71.17	90.59	66.67	83.50	57.38
StyleGAN [1]	55.06	29.39	36.36	14.83	35.62	14.28
MIPGAN-I	63.22	35.73	40.46	28.71	61.66	34.14
MIPGAN-II	57.47	31.45	51.72	23.54	54.94	27.46

Table 6.5: Quantitative evaluation of vulnerability of LCNN-29 [12] FRS from various morph generation approaches. Note that, since FNMR = 0 @ FMR = 0.1% for LCNN-29 [12] following Eq. 11.2 and 11.3, the value of RMMR is equal to MMPMR/FMMPMR. Therefore, we have not entered RMMR separately in the Table above.

Morph generation type	MMPMR/RMMR(%)	FMMPMR/RMMR(%)	MMPMR/RMMR(%)	FMMPMR/RMMR(%)	MMPMR/RMMR(%)	FMMPMR/RMMR(%)
	Digital		**Print-Scan**		**Print-Scan with compression**	
	Male		Male		Male	
Landmark-I [13]	96.63	89.28	95.25	89.36	94.80	88.62
Landmark-II [14]	75.09	60.72	74.64	57.81	82.43	68.32
StyleGAN [1]	83.12	66.44	85.20	69.54	84.85	68.88
MIPGAN-I	95.13	86.35	94.04	84.39	94.09	84.30
MIPGAN-II	94.93	85.14	93.94	83.14	93.75	82.63
	Female		Female		Female	
Landmark-I [13]	99.16	95.00	98.75	94.26	98.96	94.49
Landmark-II [14]	92.04	82.28	94.69	82.85	95.92	86.98
StyleGAN [1]	93.33	80.08	92.92	83.06	92.92	82.76
MIPGAN-I	97.76	92.27	96.94	91.59	96.94	91.44
MIPGAN-II	95.71	90.72	95.31	89.85	95.71	89.58
	Combined		Combined		Combined	
Landmark-I [13]	**97.16**	**90.19**	**95.96**	**90.14**	**95.64**	**89.55**
Landmark-II [14]	78.42	64.20	78.58	61.85	85.11	71.36
StyleGAN [1]	85.12	68.61	86.72	71.69	86.44	71.09
MIPGAN-I	95.68	87.30	94.64	85.55	94.68	85.45
MIPGAN-II	95.12	86.05	94.25	84.23	94.17	83.75

6.5.2 Vulnerability Analysis

This section presents the vulnerability analysis of the proposed morphed face generation techniques to quantify the impact of our efficient attacks on FRS. We quantify the attack success for five different FRS including two Commercial-off-the-Shelf (COTS) FRS and three deep-learning-based open-source FRS. The COTS FRS include the Cognitec FRS (Version 9.4.2) [8] [2] and Neurotechnology (Version 10) [11] and the set of open-source FRS includes Arcface [10], VGGFace [9] and LCNN-29 [12] . The operational threshold for all 5 FRS is set at False Match Rate (FMR) of 0.1% following the guidelines of Frontex [126].

The vulnerability is assessed using two metrics Mated Morphed Presentation Match Rate (MMPMR) [125] and Fully Mated Morphed Presentation Match Rate (FMMPMR) [1] based on the threshold provided by Cognitec FRS. For a given morph image $M_{I_{1,2}}$ obtained using two subjects, we compute the vulnerability by enrolling $M_{I_{1,2}}$ and verifying it against probe images from the corresponding contributing subjects I_1 and I_2. The obtained comparison scores S_1 and S_2 for both probe images I_1 and I_2 against the morphed image $M_{I_{1,2}}$ indicates the threat to FRS, if and only if both S_1 and S_2 cross the actual verification threshold at FMR = 0.1%. The corresponding metric *FMMPMR* [1] [3] is therefore computed as:

$$FMMPMR = \frac{1}{P}\sum_{M,P}(S1_M^P > \tau)\&\&(S2_M^P > \tau)\ldots\&\&(Sk_M^P > \tau), \quad (6.11)$$

[2] Outcome not necessarily constitutes the best the algorithm can do

where $P = 1, 2, \ldots, p$ represent the number of attempts made by presenting all probe images of the contributing subjects against the M^{th} morphed image, $K = 1, 2, \ldots, k$ represents the number of composite image constitute to generate the morphed image (in our case $K = 2$), Sk_M^P represents the comparison score of the K^{th} contributing subject obtained with P^{th} attempt corresponding to M^{th} morphed image and τ represents the threshold value corresponding to FMR = 0.1%. When compared to MMPMR, the FMMPMR will consider both pairwise comparison of contributory subjects and the number of attempts. In order to also establish the relationship with respect to earlier metrics, we also report the vulnerability using MMPMR [125].

Further, to effectively analyse the vulnerability, we also present the results using Relative Morph Match Rate (RMMR) defined as follows [125]:

$$RMMR(\tau)_{MMPMR} = 1 + (MMPMR(\tau)) - [1 - FNMR(\tau)] \tag{6.12}$$

$$RMMR(\tau)_{FMMPMR} = 1 + (FMMPMR(\tau)) - [1 - FNMR(\tau)] \tag{6.13}$$

Where, FNMR indicates the False Reject Rate (FNMR) of the FRS under consideration obtained at the threshold τ. In this work, τ represents the value corresponding to FMR = 0.1%. Since we have evaluated 5 different FRS systems, we have computed FNMR corresponding to these FRS to calculate the RMMR. Note that, in Equation 11.2 and 11.3 if FNMR = 0 then RMMR corresponds to MMPMR/FMMPMR.

The obtained success rate, or alternatively the vulnerability of FRS is provided in Table 6.1, 6.2, 6.3, 6.4 and 6.5 corresponding to to Cognitec [8], VGGFace [9], Arcface [10], Neurotechnology (Version 10) [11] and LCNN-29 [12] respectively. The vulnerability analysis is carried out on 5 different morph generation methods that include facial landmarks (Landmarks-I) with image smoothing as the post-processing operation [13], Facial landmarks (Landmarks-II) with automatic image retouching and colour equalisation [14], existing GAN based face morphing method based on StyleGAN [1] and proposed MIPGAN variants (MIPGAN-I and MIPGAN-II). Based on the obtained results, the following are the concrete observations:

- The FNMR corresponding to five different FRS is equal to 0. Therefore, the value of the RMMR is equal to MMPMR or FMMPMR. This indicates that the FRS systems are accurate on our face datasets employed in this work.

- Among the five FRS, the highest vulnerability is noted for Arcface [10], which is vulnerable to all five kinds of face morphing attack methods.

- Among COTS FRS, the Cognitec FRS indicates a higher vulnerability on all five types of face morphing attack methods compared to Neurotechnology FRS.

- Among five different morph generation methods, Landmark-I indicates the highest vulnerability on all five other FRS.

- The proposed face morphing methods MIPGAN-I and MIPGAN-II consistently indicate the highest vulnerability, when compared to the existing method based on StyleGAN [1]. This indicates the high quality of morphs generated using the proposed MIPGAN-I and MIPGAN-II methods.

- The proposed MIPGAN-I and MIPGAN-II methods also indicate a higher vulnerability than the Landmark-II technique for morph generation with four different FRS.

- Among the two different metrics (MMPMR and FMMPMR), the proposed FMMPMR indicates a lower vulnerability than MMPMR consistently as FMMPMR imposes a strict selection of attack images, unlike MMPMR.

- MIPGAN-I based morphed images show a marginally better performance in attacking FRS than images generated by MIPGAN-II.

6.5.3 Perceptual Image Quality Analysis

This section presents quantitative results of the proposed morphed image generation techniques using the perceptual image quality metrics PSNR and SSIM. Both of these metrics are computed based on the reference image. Morphed face images are generated based on parent face images from two contributory data subjects. Therefore, we used the parent face images from both contributory data subjects as the reference image against which the given morphed image is assessed and we average the obtained image quality scores for both parent images. Table 6.6 indicates the quantitative results of both PSNR and SSIM on four different types of face morph generation mechanism in the digital format. Based on the obtained results, it can be observed that:

- There is little deviation in the perceptual image quality metrics computed on all four different types of face morph generation mechanisms.

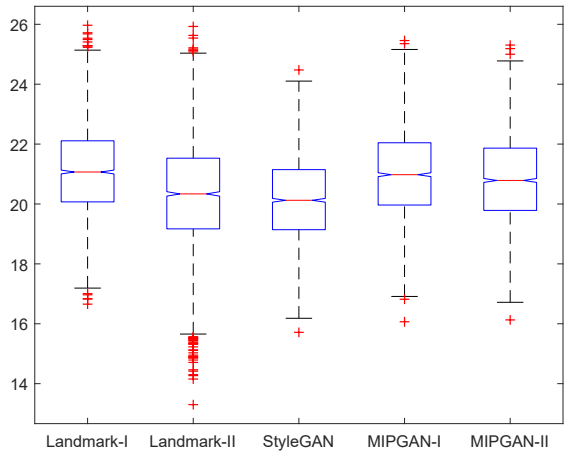


Figure 6.6: Box plots of PSNR values computed from different face morph generation methods (digital version)

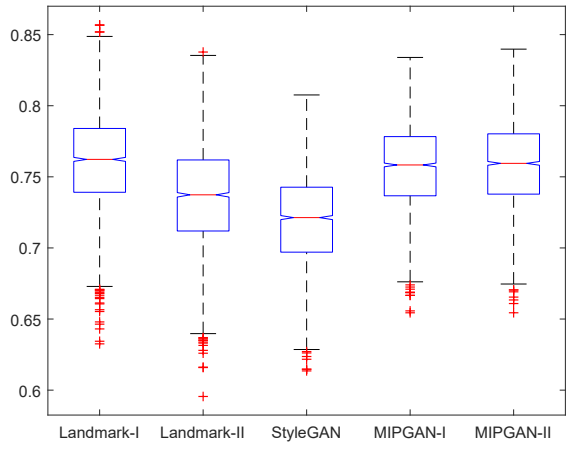


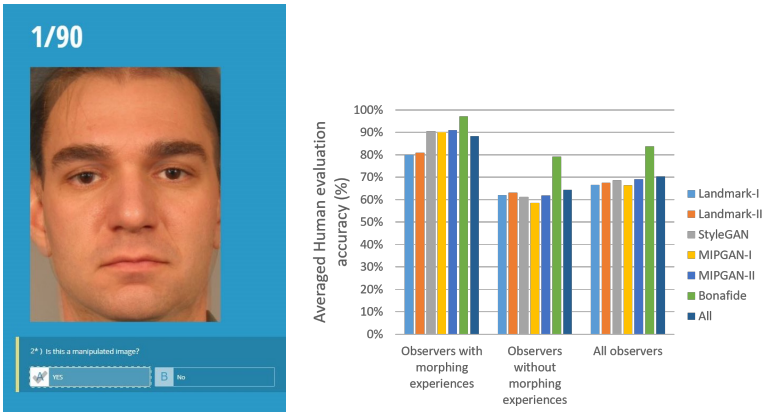
Figure 6.7: Box plots of SSIM values computed from different face morph generation methods (digital version)

- The proposed MIPGAN-I and MIPGAN-II methods indicate a slightly better image quality when compared to the StyleGAN [1] based face morphing method.
- The proposed MIPGAN-I and facial landmarks-based methods [14] indicate a similar image quality.
- Figure 6.6 and 6.7 indicate the box plots of the PSNR and SSIM quality scores. These results further indicate that the perceptual quality of the proposed MIPGAN-I and MIPGAN-II is superior to the existing state-of-the-art method based on StyleGAN [1].

Table 6.6: Morph image quality analysis using PSNR and SSIM with 95% confidence interval

Morph generation Methods	PSNR	SSIM
Landmark-I [13]	21.1111± 0.0415	0.7609±0.0009
Landmark-II [14]	20.2737±0.0523	0.7363±0.0010
StyleGAN [1]	20.1347±0.0383	0.7199±0.0008
MIPGAN-I	21.0133±0.0409	0.7573±0.0008
MIPGAN-II	20.8306±0.0409	0.7586±0.0008

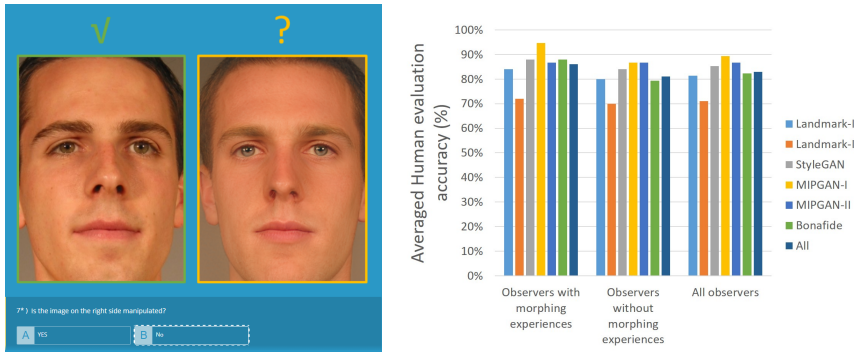
6.5.4 Human Observer Analysis



(a) (b)

Figure 6.8: (a)Example of screen shot used for human observer study (b) Quantitative results

In this section, we discuss the quantitative detection performance of human observations regarding morphed face images, which are generated using MIPGAN-I and



(a) (b)

Figure 6.9: (a)Example of screen shot used for differential human observer study (b) Quantitative results

MIPGAN-II. To this extent, we have designed and developed a Web-portal to evaluate the human morph detection performance reflecting both single image-based morphing attack detection scenario (S-MAD) and differential morphing attack detection scenario (D-MAD). We have used only digital samples of both bona fide and morphed face images as the proposed MIPGAN is used to generate the images in the digital domain. Figure 6.8 (a) shows the screenshot of the Web-portal for S-MAD in which the human observer needs to decide whether the displayed image is a morphed face image or a bona fide image by looking at one single image at a time. Correspondingly, Figure 6.9 (a) presents the screenshot for D-MAD experiment where the observer needs to detect whether the unknown image is morphed given a trusted bona fide image as a reference. We have selected a total of 90 images where 15 images are from each group corresponding to bona fide, two different types of facial landmarks based morphing such as Landmarks-I [13] and Landmarks-II [14], StyleGAN [1] based face morphing, MIPGAN-I and MIPGAN-II based face morphing. To make the testing robust, all 90 chosen images correspond to unique data subjects and there is no repetition of data subjects. To avoid gender bias by participants, we have selected a near equal distribution of male and female data subjects in each group. We have chosen 90 images considering the time constraints required to assess these images for human observers. It was important that observers do not loose focus while conducting the detection experiments.

Figure 6.8 (b) shows the quantitative results of S-MAD obtained from 56 human observers, including 14 experienced and 42 inexperienced observers. The experienced observers' group consists of researchers working in face morphing attack detection and as ID expert's in border control, while the non-experienced group

consists of students and other computer science professionals. As noticed from the Figure 6.8 (b) following are the main observations:

- Detection performance of the bona fide images indicates better detection performance by both experienced and non-experienced group when compared to the morphed face image. The experienced group indicates the detection performance with an accuracy of 97.14%, while the non-experienced group indicates the detection performance with an accuracy of 79.21%.
- Human observers with experience in face morphing demonstrate higher detection accuracy on four different face morph generation mechanisms than the inexperienced group.
- Among the four different morphing types, the experienced group indicates that the detection of the landmarks-based morphing is challenging compared to other morphing mechanisms (deep learning-based).
- Human observers with no experience in face morphing are marginally good in detecting the landmarks-based face morph images compared to other types of face morphing techniques. MIPGAN-I exhibits more challenging morph images to detect as compared to other morph generation methods.
- Based on the obtained results, it can be noted that the human observers with good experience in face morphing can detect morphed images with an accuracy of 88.25% while the human observer with no knowledge of face morphing shows the challenge to detect the morphed face images with a detection accuracy of 64.31%.
- The overall results from 56 human observers indicate that detecting morphed face images is challenging. Further, it is also interesting to note that detecting different face morphing types is also challenging.

For the quantitative results of D-MAD, 5 experienced observers and 10 inexperienced observers have participated. As shown in Figure 6.9 (b), the following observations are illustrated:

- In the scenario of D-MAD, the group with relevant experiences achieved an overall 86% accuracy, which is better than 81% for the inexperienced group. However, this difference is much less than the difference in S-MAD, which means that the reference image can help inexperienced observers to identify the morphs.

- Morphs generated by Landmark-II present a significant challenge as compared to other morph generation mechanisms in D-MAD. This may be attributed to a more natural skin texture appearance (comparing with GAN-based mechanisms) and fewer artefacts (comparing with Landmark-I) and observers focusing less on its minor artefacts in the pairwise comparison.

- It is also interesting to see that the performances of experienced observers on detecting Landmark-II (80.95% and 72.00%), StyleGAN (90.48% and 88.00%), MIPGAN-II (90.95% and 86.67%), and bona fide images (90% and 88.00%) are lower than their performance in S-MAD. We believe this is because experienced observers do not pay critical attention to tolerable difference between the trusted reference image and the unknown comparison image.

6.5.5 Ablation Study

Table 6.7: Vulnerability - Ablation study on the proposed loss function. Here, ✓ indicates the selected and ✕ indicates the not selected loss function in the ablation study

$Loss_{ID-Diff}$	$Loss_{Identity}$	$Loss_{MS-SSIM}$	$Loss_{Perceptual}$	MIPGAN-I				MIPGAN-II			
				FMMPMR		MMPMR		FMMPMR		MMPMR	
				Cognitec	ArcFace	Cognitec	ArcFace	Cognitec	ArcFace	Cognitec	ArcFace
✕	✓	✓	✓	81.82	75.87	90.69	93.47	77.83	71.98	90.1	91.18
✓	✕	✓	✓	78.07	62.15	89.17	83.77	78.39	64.51	90.04	82.54
✓	✓	✕	✓	80.82	73.33	91.81	92.66	78.73	71.79	89.58	90.55
✓	✓	✓	✕	21.37	47.85	44.18	71.95	11.92	33.12	29.47	59.56
✓	✓	✓	✓	**84.65**	**85.94**	**94.36**	**94.45**	**81.59**	**86.24**	**92.93**	**94.21**

In order to measure the impact of the loss functions in the proposed approach, we conduct an extensive ablation study. The proposed loss function combines four different entities such as: perceptual loss ($Loss_{Perceptual}$), identity loss ($Loss_{Identity}$), identity difference ($Loss_{ID-Diff}$) and Multi-Scale Structural Similarity (MS-SSIM) loss ($Loss_{MS-SSIM}$). The main contribution of our work is to use identity information, which can be considered as a specific high-level feature, to measure the loss. However, high-level features also mean that it is hard for the gradient descent algorithm to ensure a good convergence during the optimization process. Therefore, we have introduced the perceptual loss that can measure relatively low-level features in addition to MS-SSIM and identity difference loss to effectively control the optimization process to generate a high-quality morphed image. We perform the ablation study by discarding each term in the loss function iteratively. We benchmark the vulnerability using COTS FRS (Cognitec FRS (Version 9.4.2)) and the open-source ArcFace FRS, as the proposed approach is dedicated to generating high-quality morphed images.

Table 6.7 indicates the quantitative performance of the ablation study using a vulnerability analysis for both the COTS-FRS from Cognitec and for the open-source

Arcface FRS with the proposed MIPGAN-I and MIPGAN-II methods. The ablation study is carried out on the digital morphed images generated using both MIPGAN-I and MIPGAN-II Methods. Figure 6.10 and 6.11 shows the qualitative performance of the ablation study on both MIPGAN-I and MIPGAN-II, respectively. Based on the obtained results, the following are the main observations:

- Each term in our proposed loss function (see Eq. 6.10) contributes to posing a greater challenge to a FRS for both proposed MIPGAN-I and MIPGAN-II morph generation frameworks.
- Among the four other loss functions that we have used, the $Loss_{Perceptual}$ is critical in improving the proposed method's performance. Discarding the perceptual loss has resulted in a degrading performance in both qualitative (see Figure 6.10 (d) and 6.11 (d)) and quantitative results.
- The use of identity loss ($Loss_{Identity}$) also indicates the importance of im-

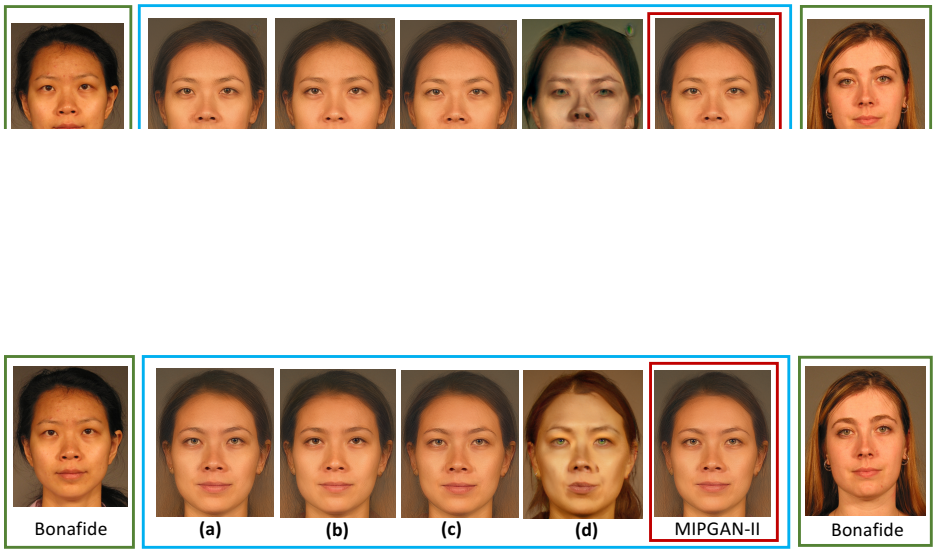


Figure 6.11: Qualitative results of ablation study using proposed MIPGAN-II (a)$Loss_{ID-Diff}$ (b) $Loss_{Identity}$ (c) $Loss_{MS-SSIM}$ (d) $Loss_{Perceptual}$

6.5.6 Hyper-parameters Study

This section presents both qualitative and quantitative results on the selection of hyper-parameters (λ_1, λ_2, λ_3, and λ_4) in the proposed loss function employed in both MIPGAN-I and MIPGAN-II. Based on the ablation study reported in Section 6.5.5, we have noticed that the perceptual loss is the vital component of our loss function (see Eq. 6.10) and the other three terms can be used as constraints during the optimization. Therefore, the first step is to study the generated morphed face images' attack strength by increasing and decreasing the value of λ_1. Among the remaining three terms, we have also noticed from the ablation study that the identity loss ($Loss_{Identity}$) is contributing more towards generating a high-quality morph compared to the other two-loss functions ($loss_{MS-SSIM}$, $Loss_{ID-Diff}$). We analyze the importance of identity loss ($Loss_{Identity}$) with respect to the other two loss functions ($Loss_{MS-SSIM}$, $Loss_{ID-Diff}$) by increasing the value of λ_3 and/or λ_3 and decreasing the value of λ_2. Further, we have also noticed from the ablation study that the loss functions $loss_{MS-SSIM}$ and $Loss_{ID-Diff}$ are less important and numerically very small. Therefore, we did not conduct studies on decreasing the values of λ_3 and λ_4. Altogether, we have tested four different cases of changing the hyper-parameter values to generate the morphed face images. These generated morphed face images are benchmarked against the proposed hyper-parameter values through the vulnerability analysis using both COTS FRS (Cognitec FRS (Version 9.4.2)) and open-source ArcFace FRS.

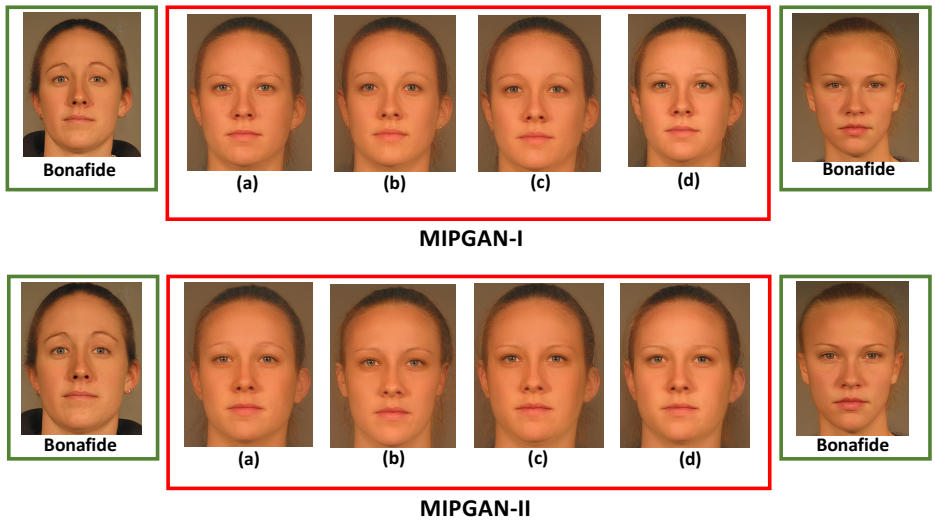


Figure 6.12: Qualitative results of Hyper-parameters study on both MIPGAN-I and MIPGAN-II (a)λ_1 (b) λ_2 (c) λ_3 (d) λ_4

Table 6.8: Quantitative results of hyper-parameters study

Proposed Morph Generators	Case-study	Hyper-parameters weights				MMPMR (%)		FMMPMR (%)	
		λ_1	λ_2	λ_3	λ_4	Cognitec	ArcFace	Cognitec	ArcFace
MIPGAN -I	1	0.0004	10	1	1	93.94	93.49	84.39	75.61
	2	0.0001	10	1	1	92.66	91.15	79.66	72.94
	3	0.0002	1	10	1	94.17	91.9	84.34	75.66
	4	0.0002	1	1	10	83.16	82.14	67.19	59.46
	Proposed weights	0.0002	10	1	1	**94.36**	**94.45**	**84.65**	**85.94**
MIPGAN -II	1	0.0004	10	1	1	91.36	91.98	81.29	76.18
	2	0.0001	10	1	1	91.69	88.29	73.91	68.16
	3	0.0002	1	10	1	90.63	90.91	80.76	75.87
	4	0.0002	1	1	10	87.22	74.33	57.43	51.91
	Proposed weights	0.0002	10	1	1	**92.93**	**94.21**	**81.59**	**86.94**

Table 6.8 shows the qualitative performance and Figure 6.12 shows the qualitative performance of the hyper-parameter study. Based on the obtained results, it can be noted that the increase in the value of λ_1 and λ_3 shows comparable results with the proposed weighting schemes. However, based on our empirical study on hyper-parameters, we noted that: if we set λ_1 and λ_2 with equal weights, then, during the optimization, the generated morph image will soon become roughly similar to both contributing subjects. This will quickly reduce identity loss ($Loss_{Identity}$) to a minimal value and loose its importance in the optimization. Hence, we set a larger factor to the identity loss compared with other loss terms measuring high-level features to ensure our most important constraint term is still effective in the later stage of optimization. Further, both λ_3 and λ_4 can make the optimization goal more comprehensive but setting a large factor will obstruct the convergence. Especially setting high values to λ_4 will end up with an image not similar to both subjects. Therefore, the selection of the proposed hyper-parameters confirms the generation of a high-quality morphed image but also aids for effective and comprehensive optimization.

6.5.7 Morphing Attack Detection Potential

Table 6.9: Quantitative performance of MAD - Training- Landmarks-I [13]

Morph Generation Type: Training	**Morph Generation Type: Testing**	**MAD Algorithms**	**Digital**			**Print-scan**			**Print-scan with compression**		
			D-EER(%)	BPCER @ APCER = 5%	BPCER @ APCER = 10%	D-EER(%)	BPCER @ APCER = 5%	BPCER @ APCER = 10%	D-EER(%)	BPCER @ APCER = 5%	BPCER @ APCER = 10%
Landmarks-I [13]	Landmarks-I [13]	Ensemble Features [6]	0	0	0	2.35	1.45	0.96	2.58	1.71	1.54
		Hybrid Features [91]	0.16	0	0	1.85	0.85	0.34	2.25	1.12	0.51
	Landmarks-II [14]	Ensemble Features [6]	49.55	92.22	88.85	41.93	81.45	76.25	42.15	83.88	77.64
		Hybrid Features [91]	49.16	99.31	97.59	44.17	86.48	80.24	46.49	88.38	81.95
	StyleGAN [1]	Ensemble Features [6]	0.22	0	0	13.36	27.44	16.46	14.77	27.27	19.38
		Hybrid Features [91]	0.16	0	0	44.96	83.7	75.47	9.44	14.57	9.14
	MIPGAN-I	Ensemble Features [6]	39.16	73.14	65.35	9.45	14.57	8.74	8.95	15.26	9.26
		Hybrid Features [91]	46.82	86.62	81.64	12.32	19.72	13.2	9.74	15.95	8.91
	MIPGAN-II	Ensemble Features [6]	34.13	70.49	61.57	5.32	6.68	2.57	6.72	8.16	4.14
		Hybrid Features [91]	44.96	83.7	75.47	5.9	8.42	3.23	5.67	6.18	2.91

Considering the success rate of the newly generated dataset, we naturally choose to evaluate the morphing attack detection performance to also validate the robustness of existing MAD mechanisms. Additionally, we investigate recent works about general face manipulation detection [153] [154] [155] and some results are

Table 6.10: Quantitative performance of MAD - Training- Landmarks-II [14]

Morph Generation Type: Training	Morph Generation Type: Testing	MAD Algorithms	Digital			Print-scan			Print-scan with compression		
				BPCER @ APCER =			BPCER @ APCER =			BPCER @ APCER =	
			D-EER(%)	5%	10%	D-EER(%)	5%	10%	D-EER(%)	5%	10%
Landmarks-II [14]	Landmarks-I [13]	Ensemble Features [6]	48.57	97.77	95.36	24.19	52.48	43.22	21.64	47.51	36.19
		Hybrid Features [91]	45.67	96.91	94.16	32.26	77.87	66.55	24.51	50.94	40.65
	Landmarks-II [14]	Ensemble Features [6]	3.62	2.22	0.68	6.32	7.97	2.42	5.57	6.41	2.42
		Hybrid Features [91]	1.53	0.17	0	5.21	5.19	3.14	5.37	5.71	3.46
	StyleGAN [1]	Ensemble Features [6]	29.67	61.92	52.48	27.18	61.57	50.6	29.18	62.14	52.48
		Hybrid Features [91]	34.76	74.44	62.95	34.8	67.23	58.14	23.17	49.22	38.25
	MIPGAN-I	Ensemble Features [6]	30.23	65.35	53.17	43.92	87.65	79.24	44.24	89.23	82.33
		Hybrid Features [91]	46.29	84.04	77.01	34.16	71.18	64.66	35.5	76.84	65.52
	MIPGAN-II	Ensemble Features [6]	27.13	58.83	45.45	33.57	77.35	65.52	40.46	84.9	75.47
		Hybrid Features [91]	46.82	83.53	75.81	35.91	77.18	65.24	36.5	79.24	68.78

Table 6.11: Quantitative performance of MAD - Training- StyleGAN [1]

Morph Generation Type: Training	Morph Generation Type: Testing	MAD Algorithms	Digital			Print-scan			Print-scan with compression		
				BPCER @ APCER =			BPCER @ APCER =			BPCER @ APCER =	
			D-EER(%)	5%	10%	D-EER(%)	5%	10%	D-EER(%)	5%	10%
StyleGAN [1]	Landmarks-I [13]	Ensemble Features [6]	0.32	0	0	16.6	28.13	19.89	13.89	22.12	17.66
		Hybrid Features [91]	0.42	0	0	15.26	26.41	17.66	14.37	22.81	16.92
	Landmarks-II [14]	Ensemble Features [6]	44.72	89.53	80.61	38.31	78.5	69.15	38.84	83.7	74.17
		Hybrid Features [91]	45.65	90.22	84.56	34.18	81.95	70.53	32.93	78.5	64.12
	StyleGAN [1]	Ensemble Features [6]	0	0	0	0	0	0	0	0	0
		Hybrid Features [91]	0	0	0	0	0	0	0	0	0
	MIPGAN-I	Ensemble Features [6]	39.97	75.98	68.78	20.21	42.14	33.44	20.73	45.28	36.53
		Hybrid Features [91]	46.45	86.79	77.87	29.34	59.19	47.51	24.87	51.62	41.18
	MIPGAN-II	Ensemble Features [6]	39.93	73.58	66.89	15.78	28.14	19.38	13.72	28.98	16.63
		Hybrid Features [91]	44.72	82.16	73.75	19.36	43.22	28.64	16.98	32.93	23.84

Table 6.12: Quantitative performance of MAD - Training- MIPGAN-I

Morph Generation Type: Training	Morph Generation Type: Testing	MAD Algorithms	Digital			Print-scan			Print-scan with compression		
				BPCER @ APCER =			BPCER @ APCER =			BPCER @ APCER =	
			D-EER(%)	5%	10%	D-EER(%)	5%	10%	D-EER(%)	5%	10%
MIPGAN-I	Landmarks-I [13]	Ensemble Features [6]	23.66	51.45	39.96	5.82	7.22	2.92	6.17	7.54	3.94
		Hybrid Features [91]	47.15	87.16	79.41	6.5	8.23	4.15	7.91	10.29	6.34
	Landmarks-II [14]	Ensemble Features [6]	35.38	82.33	68.95	41.67	95.14	83.53	43.68	96.01	85.44
		Hybrid Features [91]	28.62	75.64	61.4	44.38	95.66	85.78	38.18	90.46	78.16
	StyleGAN [1]	Ensemble Features [6]	17.72	37.22	26.58	12.19	26.24	15.26	11.82	24.69	14.23
		Hybrid Features [91]	31.16	64.32	53.85	11.99	19.2	13.72	9.93	18.15	9.94
	MIPGAN-I	Ensemble Features [6]	0	0	0	0	0	0	0	0	0
		Hybrid Features [91]	0	0	0	0	0	0	0	0	0
	MIPGAN-II	Ensemble Features [6]	2.15	0.17	0	0.68	0	0	0.64	0	0
		Hybrid Features [91]	1.36	0.34	0	0.86	0	0	0.8461	0	0

Table 6.13: Quantitative performance of MAD - Training- MIPGAN-II

Morph Generation Type: Training	Morph Generation Type: Testing	MAD Algorithms	Digital			Print-scan			Print-scan with compression		
				BPCER @ APCER =			BPCER @ APCER =			BPCER @ APCER =	
			D-EER(%)	5%	10%	D-EER(%)	5%	10%	D-EER(%)	5%	10%
MIPGAN-II	Landmarks-I [13]	Ensemble Features [6]	13.08	29.15	15.78	4.28	3.94	2.22	4.28	3.61	2.22
		Hybrid Features [91]	40.14	77.7	67.23	5.49	5.48	2.4	7.21	10.98	4.15
	Landmarks-II [14]	Ensemble Features [6]	32.37	84.9	70.32	39.2	90.12	82.32	44.17	95.49	88.73
		Hybrid Features [91]	23.88	63.8	45.62	40.22	88.9	79.2	38.96	94.28	82.14
	StyleGAN [1]	Ensemble Features [6]	12.51	22.29	15.78	13.72	29.67	18.18	14.25	31.73	20.41
		Hybrid Features [91]	24.7	49.74	41.85	12.87	26.58	14.75	11.86	26.92	15.09
	MIPGAN-I	Ensemble Features [6]	1.56	0.68	0.34	2.14	1.22	0.53	2.57	0.85	0.34
		Hybrid Features [91]	2.27	0.85	0.17	4.79	4.8	3.43	4.3	3.6	2.22
	MIPGAN-II	Ensemble Features [6]	0	0	0	0	0	0	0	0	0
		Hybrid Features [91]	0	0	0	0	0	0	0	0	0

shown in the supplementary material. In this work, we focus on single image based morphing attack detection (S-MAD) as it perfectly suits our dataset. MAD has been widely addressed in the literature by developing the techniques based on both deep learning [85], [89], [86] [112] [117] and non-deep learning [81] [79] [99] approaches. Readers can refer to [7] for an exclusive survey on face MAD. Owing to the recent works detailing the applicability of Hybrid features [91] and Ensemble features [6] in detecting morphing attacks, we choose to benchmark both Hybrid features [91] and Ensemble features [6]. While the Hybrid features [91] resort to extracting features using both scale space and color space combined with multiple classifiers, Ensemble features [6] employ a variety of textural features in conjunction with a set of classifiers. In common both approaches evaluate a wide variety of MAD mechanisms in a holistic manner supported by empirical results [91, 6]. In addition, the Hybrid features [91] mechanisms are also validated against the ongoing NIST FRVT MORPH challenge [145] with the best performance in detecting printed and scanned morph images justifying our selection of algorithm to benchmark the newly composed database.

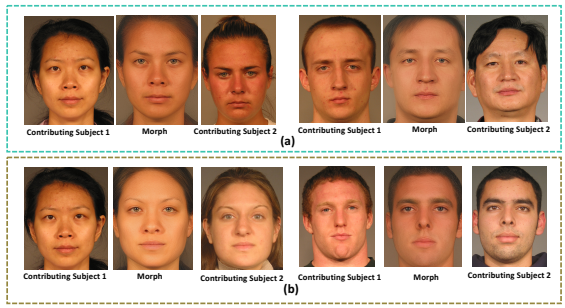


Figure 6.13: Examples of morphed images that failed to attack FRS (a) morphed face images generated using proposed MIPGAN-I (b) morphed face images generated using proposed MIPGAN-II

The reporting of MAD performance is following the ISO/IEC metrics [21] namely the Attack Presentation Classification Error Rate (APCER (%)) which defines the proportion of attack images (morph images) incorrectly classified as bona fide images and the Bona fide Presentation Classification Error Rate (BPCER (%)) in which bona fide images incorrectly classified as attack images are counted [21] along with the Detection Equal Error Rate (D-EER (%)). To evaluate the generated morphed face image's attack potential, we have sub-divided the newly generated database into two sets for training and testing that consists of independent data subjects with no overlap between the splits. The training set includes 690 bona fide images and 1190 morphed images. The testing set consists of 580 bona fide and 1310 morphed images. To effectively evaluate the performance of the MAD

reflecting a real-life scenario, we report the results on both intra (training and testing dataset from the same morph generation approach) and inter (training on one type of morphing techniques and testing on another type of morphing techniques) evaluation of MAD mechanisms. Extensive experiments are performed on digital, print-scan and print-scan with compression data types to provide an in-depth analysis of the S-MAD performance. Table 6.9, 6.10, 6.11, 6.12 and 6.13 presents the quantitative results of MAD mechanisms on morph generation methods together with the SOTA morph generation techniques. Based on the results obtained from the intra-dataset experiments, we make some concrete observations as listed below:

- The intra-dataset evaluation indicates that the morphing attacks are detected with a good success rate irrespective of the type of generation.
- In general, the attack detection success rate is high with digital data when compared to print-scan and print-scan compression.
- Among the different types of morph generation techniques, the Landmark-II based morph generation shows the highest error rates. The attack images created using StyleGAN and proposed MIPGAN can be efficiently detected using both the employed approaches with high accuracy. This can be attributed to the noises that are synthesized using GANs due to the computational modifications performed on the latent space in GAN-based morph generation methods.

In the following, we discuss the important observations based on the results obtained from inter-dataset MAD analysis:

- The performance of the MAD techniques are degraded on all five different case studies as indicated in the Table 6.9, 6.10, 6.11, 6.12 and 6.13.
- Training MAD algorithms with one type of landmarks-based method did not show the improvement in detection performance of another kind of landmarks-based morph generation method.
- When MAD mechanisms are trained using the Landmarks-I [13] method, the degraded performance is noted for all other morph generation methods except for the StyleGAN [1] based approach. This fact is also noted when we train the MAD techniques using StyleGAN [1] generated samples and test it with Landmarks-I [13] samples. Thus, the StyleGAN [1] based morph generation is easy to detect even when MAD mechanisms are not trained using the images from same morph generation scheme.

- When MAD algorithms are trained using Landmarks-II [14] samples, MAD algorithms indicate degraded performance on all other morph generation techniques.

- When MAD mechanisms are trained using the proposed MIPGAN-I generated samples. The MAD mechanisms indicate an excellent detection performance on MIPGAN-II samples. However, the detection performance of MAD methods is deceived with other morph generation techniques.

- It is interesting to note that when MAD mechanisms are trained using MIPGAN-I/MIPGAN-II, higher detection accuracy can be observed for print-scan and print-scan with compression data when compared to digital morph data. A possible reason is that the noise generated together with the morphed images using the proposed MIPGAN-I/MIPGAN-II can approximate the generated noise resulting from the print-scan and print-scan compression process.

- Based on the results of the inter-database MAD analysis, the detection of Landmarks-II [14] samples are challenging.

6.6 Limitations of Current Work and Potential Future Works

Despite this work presenting a new approach to generate strong morphing attacks, which are empirically evaluated using COTS FRS, our work has a few noted limitations. In the current scope of work, we evaluate the impact of print and scan (re-digitizing) using one printer reflecting a realistic scenario. The MAD mechanism employed in this work has not been investigated with a wide range of printers and scanners that may impact the MAD performance. While we assert that the MAD performance may not vary extremely, when tested with a wider combination of printers and scanners, that empirical evaluation is yet to be conducted in future works.

A second aspect is that the proposed approach needs pre-selection of ethnicity for generating stronger attacks. Figure 6.13 shows example morphed face images generated using the proposed method using MIPGAN-I and MIPGAN-II that fail to get verified to contributing subjects when ethnicity pre-selection is not performed [13]. We notice that the selection of contributing subjects plays an important role with the proposed method to generate stronger attacks with MIPGAN. It is our assertion that the selection of contributing subjects with similar geometric structures (particularly ethnicity and age) can improve the performance of the proposed system, but that aspect needs further investigation.

6.7 Conclusion

Addressing the limitations of generating the strong and severe morphing attacks using GAN, we have proposed a new architecture for generating face morphed images in this work. The proposed approach (MIPGAN with two variants) for devising strong morphing attacks uses identity prior driven GAN with a customized loss exploiting perceptual quality and identity factors to generate realistic images that can strongly threaten FRS. In order to validate the attack potential of the proposed morph generation method, we have created a new dataset consisting of $30,000$ morphed images and $15,240$ bona fide images. Both COTS and deep learning based FRS were evaluated empirically to measure the success rate of the new approach and vulnerability was reported indicating the applicability of the new approach and newly generated database. In a similar direction, the dataset is also validated for detection performance by studying two state-of-art MAD mechanisms. Despite the high attack detection success rate by employed MAD, we note that the morphed images generated by MIPGAN can severely threaten FRS in a present state without MAD in FRS.

Acknowledgment

This work was supported by European Union's Horizon 2020 Research and Innovation Programme under Agreement 883356.

Part III

Published Articles: Vulnerability Analysis

Chapter 7

Article 3: On the Influence of Ageing on Face Morph Attacks: Vulnerability and Detection

Sushma Venkatesh, Kiran Raja, Raghavendra Ramachandra, and Christoph Busch. *"On the influence of ageing on face morph attacks: Vulnerability and detection"*. In 2020 IEEE International Joint Conference on Biometrics (IJCB), pages 1–10. IEEE, September 2020.

7.1 Abstract

Face morphing attacks have raised critical concerns as they demonstrate a new vulnerability of Face Recognition Systems (FRS), which are widely deployed in border control applications. The face morphing process uses the images from multiple data subjects and performs an image blending operation to generate a morphed image of high quality. The generated morphed image exhibits similar visual characteristics corresponding to the biometric characteristics of the data subjects that contributed to the composite image and thus making it difficult for both humans and FRS, to detect such attacks. In this paper, we report a systematic investigation on the vulnerability of the Commercial-Off-The-Shelf (COTS) FRS when morphed images under the influence of ageing are presented. To this extent, we have introduced a new morphed face dataset with ageing derived from the publicly available MORPH II face dataset, which we refer to as MorphAge dataset. The dataset has two bins based on age intervals, the first bin - MorphAge-I dataset has 1002 unique data subjects with the age variation of 1 year to 2 years while the MorphAge-II dataset consists of 516 data subjects whose age intervals are from 2

years to 5 years. To effectively evaluate the vulnerability for morphing attacks, we also introduce a new evaluation metric, namely the Fully Mated Morphed Presentation Match Rate (FMMPMR), to quantify the vulnerability effectively in a realistic scenario. Extensive experiments are carried out using two different COTS FRS (COTS I Cognitec FaceVACS-SDK Version 9.4.2 and COTS II - Neurotechnology version 10.0) to quantify the vulnerability with ageing. Further, we also evaluate five different Morph Attack Detection (MAD) techniques to benchmark their detection performance with respect to ageing.

7.2 Introduction

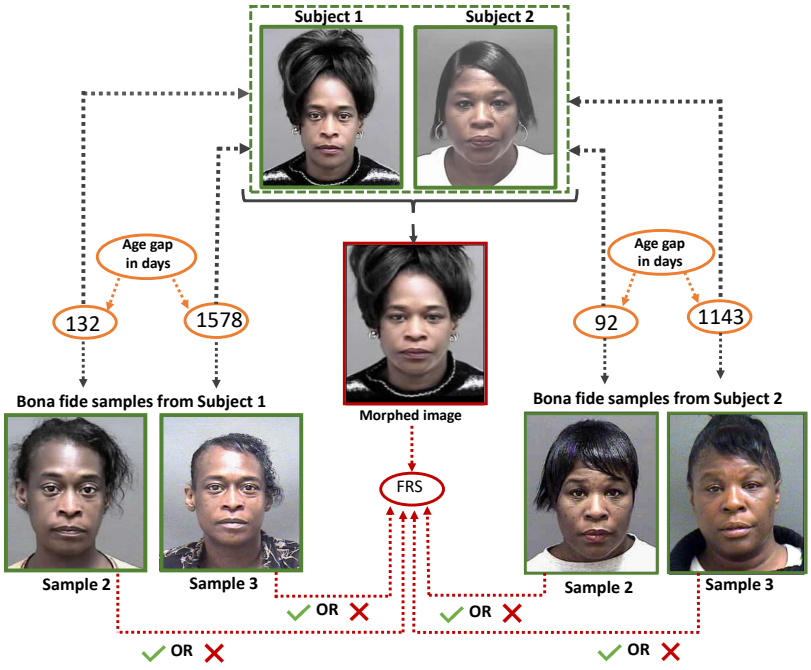


Figure 7.1: Illustration of the influence of ageing on face morphing

Facial characteristics have been well explored for identifying and verifying individuals and numerous biometric systems have been deployed in operational applications for many years [156, 17]. The preference towards face based biometric systems is founded on multiple factors such as ease of capture of facial characteristic without invasive imaging, capturing at a stand-off distance both in semi-cooperative (voluntary identification/verification) and uncooperative scenarios (surveillance) [157, 9, 158]. While many of the breakthrough articles detailing iris and vein recognition systems have shown impeccable accuracy with very low false accepts and false rejects, those systems suffer from highly constrained

image capturing processes. In order to reach the performance of such iris and vein recognition systems, face biometrics has seen benefits from recent algorithmic advancements, which was focused on features that have been engineered in a robust manner [159, 160, 161], and pre-processing that has been improvised [162] by including end-to-end learning using Deep Neural Networks (DNN) even in large scale applications [158, 163].

Such attractive and inherent advantages of the face modality have led to a wider deployment of Face Recognition Systems (FRS) in passport issuance processes, visa management, identity management and Automated Border Control (ABC). Despite the high accuracy and convenience of face biometrics, FRS systems are impeded by various factors such as ageing [164, 165], partial face availability [165] and also imperilled by various attacks that include presentation attacks (spoofing-attacks) with print, display or silicon mask attack instruments [166], make-up attacks [167, 168], coverted mask attacks [169], morphing attacks [22], database level attacks [170, 171] and comparison level attacks [172, 173]. While many of the attacks have been addressed through mitigation measures over the period of time, we focus on recently surfaced face morphing attacks [22, 54] in this work. Despite some of the recent works proposing measures to mitigate these attacks through various approaches [142, 174, 119, 54] a number of covariates are reported to impact the attack detection performance. A list of covariates impacting the performance of morphing attack detection include the techniques used to generate the morphed image [22], the configuration of the print-scan pipeline [83], factors of age and ethnicity [125] among many other unknown factors. With a clear introspection of the existing works, we observe that both the FRS vulnerability and also the Morphing Attack Detection (MAD) performance under variation of age is not studied in the context of morphing attacks, despite the fact that the issue was pointed out already in the early works in this domain [125, 119].

Starting with this observation, we focus in this work on establishing the impact of ageing on morphing attacks by carefully studying the vulnerability of FRS and MAD performance of currently reported MAD algorithms under the influence of ageing. The key motivation stems from earlier works who have disentangled the impact of ageing on face recognition systems with respect to recognition performance [164, 165, 175] and a number of works that have proposed approaches to handle the associated performance limitations [164, 176, 177, 178, 175]. We therefore provide a brief overview of impact of ageing in the subsequent section and thereafter illustrate the impact of ageing specifically for morphing. Further, we focus our work on investigating the impact using the digital images alone due to two primary factors: (i) many countries across the world allow to upload digital images via web-portal for passport renewal and visa issuance, and (ii) to align our

works with recent studies focusing on digital MAD [119].

7.2.1 Facial Ageing

Facial ageing is a commonly observed phenotype of human ageing, which is visibly seen. Despite the complexity of understanding the characteristic changes associated with the facial ageing, a number of works have reported the role of skin and soft tissues and their impact on visible changes of facial appearance [179]. Complementary works have demonstrated the role of loss of facial bone volume to contribute to facial appearance under ageing progress [180]. As it can be deduced, facial ageing being a complex process involving soft tissues and skeletal structure changes, it is influenced by many factors, such as exposure to sunlight and body weight among others. As an additional factor, large variations in facial ageing across individuals and ethnic populations can further be observed [181]. While in face recognition, the main differences in exterior facial structure making individuals distinguishable from each other allows recognition analysis to achieve high identification accuracy, a longitudinal study of the same face over a period of time has shown to challenge the accuracy [175].

7.2.2 Facial Ageing and Morphing Attacks

Under the observation of complex changes of facial appearance, which bring down the recognition accuracy of FRS unless proper measures are taken, our assertion is that the effect and impact on morphing attacks may change. For electronic Machine Readable Travel Documents (MRTDs) a typical life-cycle of 10 years is recommend [182] meaning that the drastic changes in facial appearance must be tolerated as intra-class variance during that life-cycle, while up to now the impact of morphing and its correlation with the progressing of the potentially morphed reference image in this life-cycle, has neither been considered nor investigated. Initial studies on morphing attacks have demonstrated the ability to fool a human expert (i.e. trained border guards) with morphed facial images. The changes of facial appearance, which are caused by ageing, are illustrated in Figure 9.1. Our assertion is to validate the impact of ageing and thus we formulate three specific research questions:

- How vulnerable are COTS FRS when a composite morph image is enrolled and is after a period of ageing probed against a live image from one of the contributing subjects?
- Do current Morphing Attack Detection (MAD) algorithms scale-up to detect such attacks under the influence of ageing?
- What is the impact of different alpha (or blending, morphing) factors used

to generate the morphed image under the constraint of ageing, specifically with respect to MAD?

We address each of these questions in a systematic manner through our contributions. We focus in this work to first establish the impact on FRS through an extensive empirical evaluation. While a detailed study of appearance change is more of a cognitive study, it is beyond the scope of the current work.

7.2.3 Contributions of Our Work

While the hypothesis is well justified, we also note that there exists no database with morphing and ageing according to the current literature. With such a caveat, we focus on first creating a database to facilitate and validate our assertion.

- The first key contribution is the creation of a (moderately) large-scale database of morphed faces with ageing covariate by employing the MORPH II non-commercial face dataset [35], which is hereafter referred as MorphAge Database.
- We investigate the vulnerability of FRS to such attacks by employing two widely used Commercial-Off-The-Shelf (COTS) FRS systems. This contribution not only helps in verifying our assertion but also validates the usefulness of the newly created database. Further, we also investigate the role of alpha (or blending, morphing) factor (with $\alpha = 0.3.\ 0.5$ and 0.7) while analysing the vulnerability under ageing.
- As a third contribution, we employ a set of recently reported morphing attack detection algorithms to benchmark detection performance and thereby identify the impediments if any.

In the rest of the paper, we first provide details on the newly constructed database in Section 7.3 and in Section 7.4 we investigate the vulnerability of FRS using two COTS FRS. Further, the benchmarking of morphing attack detection systems is detailed in Section 7.5 while the key observations and conclusions are reported in Section 7.7.

7.3 MorphAge Database Construction

To effectively study the influence of ageing on face morphing vulnerability and morph detection, we introduce a new dataset, which is derived from the MORPH II non-commercial dataset [35] that is publicly available. The MORPH II dataset consists of a total of 55000 unique samples captured from 13000 data subjects. The

images are captured over the time span from 2003 to 2007. The age of the subjects varies from 16 to 77 years. The dataset consists of male and female subjects with different ethnicity (African, European, Asian, Hispanic). In this work, we choose the MORPH II dataset motivated by the large number of subjects, the quality of the captured data and the variation in age for one and the same subject across different capture sessions.

The newly constructed MorphAge dataset is binned in two age groups from MORPH II dataset. The first bin - Age Group (MorphAge-I) consists of 1002 unique data subjects with a gender distribution of 143 female and 859 male subjects. For each data subject, three different samples are chosen such that the first session corresponds to the high quality data capture (younger age), second session corresponds to the aged capture of 1-8 months from first session and third session corresponds to the aged capture of same subject between 1-2 years from first session. The second bin - Age Group (MorphAge-II) is comprised of 516 unique data subjects sub-sampled from the MORPH II dataset with 62 female and 454 male data subjects. Each data subject was captured in three different sessions. The first session corresponds to the high quality data capture (younger age), the second session corresponds again to a time lapse of 1-8 months from the first session and the third session corresponds to an aged capture of 2 years up to 5 years after the first session.

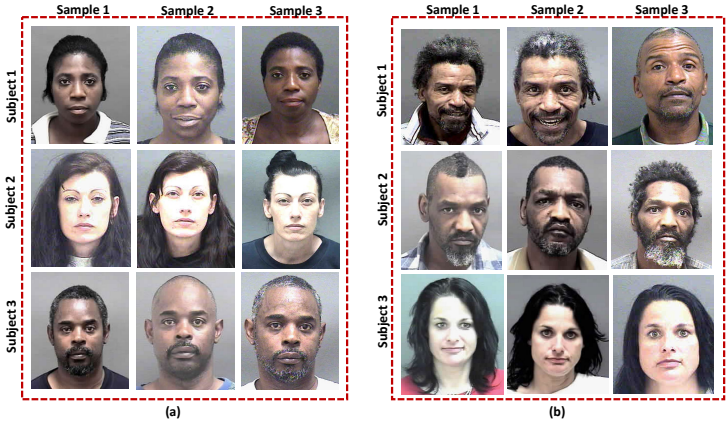


Figure 7.2: Illustration of sample images from newly constructed MorphAge dataset (a) MorphAge-I (1 year to 2 years) (b) MorphAge-II (2 years to 5 years)

Figure 2: Illustration of MorphAge I and MorphAge II

7.3.1 MorphAge-I and MorphAge-II - Bonafide Set

In both bins, i.e. MorphAge-I and MorphAge-II, we select for each data subject three samples (one of each session) such that the *first session sample is used only to generate the morphing image, the sample from the second session is used as*

Table 7.1: Statistics of bona fide and morphed images in MorphAge Database

Session	Dev	Training	Testing	Total
MorphAge-I Subset				
Session 1 (used for morphing)	251	500	251	1002
Session 2 (used for vulnerability)	251	500	251	1002
Session 3 (with age difference)	251	500	251	1002
Morphed Images	1980	6614	1944	10538
MorphAge-II Subset				
Session 1 (used for morphing)	130	257	129	516
Session 2 (used for vulnerability)	130	257	129	516
Session 3 (with age difference)	130	257	129	516
Morphed Images (with different morphing factors)	648	2310	809	3767

bona fide sample in the morph attack detection experiments and the third session to analyze the vulnerability of the commercial FRS. As seen from the Figure 7.2, the facial appearance changes significantly with the increasing age which cannot be modelled geometrically or morphologically for any particular ethnicity or age group for both the bins (MorphAge-I and MorphAge-II).

7.3.2 MorphAge-I and MorphAge-II - Morphed Image Set

To generate the morphed image datasets for the subjects represented in our newly constructed dataset, we have used the face morph generation tool from Ferrara et al. [14] [135], which is based on facial landmarks based warping and weighted linear blending to generate a high quality morphed image. We particularly, choose this technique for morphing generation over other type of generators based on GAN [15] by considering: (1) high quality of the generated morphed images, in order to establish a significant threat to the tested commercial FRS [14] (2) high quality of generated morphed image, such that the submitted images are considered compliant with the requirements in the ICAO standards and (3) feasibility to create the morphed images with various blending and warping factors.

In this work, the morphing process is carried out between only two data subjects by considering its use-case in a real-life scenario where typically one criminal morphs his/her face image with the image of an accomplice. To carefully select the pair of images for the morphing process, we use the COTS-I FRS, which is widely used in Automated Border Control installations. Through the FRS, a set of similarity scores is obtained between the probe image of a selected data subject against the

Figure 7.3: Example of generated morphed images

reference images of all data subjects. We then choose the pair of images that are successfully verified at FMR = 0.1% with high scores to retain a high degree of similarity between two constituting subjects for the morphed image. Additional care is exercised not to combine data subjects with different genders and also to separate the data subject into three independent groups such as non-overlapping training, testing and development sets [125, 142]. For a selected image pair, we generate three morphed images at three different morphing (or blending) factors $\alpha = 0.3, 0.5, 0.7$ to obtain insights with regard to the impact of ageing at different blending factors. Figure 7.3 shows the example of morphed face images with three different blending factor within our MorphAge dataset.

Table 7.1 presents the statistics of the generated dataset corresponding to the two bins - MorphAge-I and MorphAge-II. Further, in order to evaluate the MAD performance, we have divided the whole datasets into three independent and non-overlapping subsets for training, development and testing. The training subset is used purely to train the MAD techniques, the development subset is used to optimize and adjust the operating threshold for the MAD techniques and finally the testing subset is solely used to analyze the detection performance obtained at the optimal threshold.

7.4 Vulnerability Analysis

In this section, we present the vulnerability analysis of the FRS, when confronted with the morphed images under variation of age. To this extent, we employ

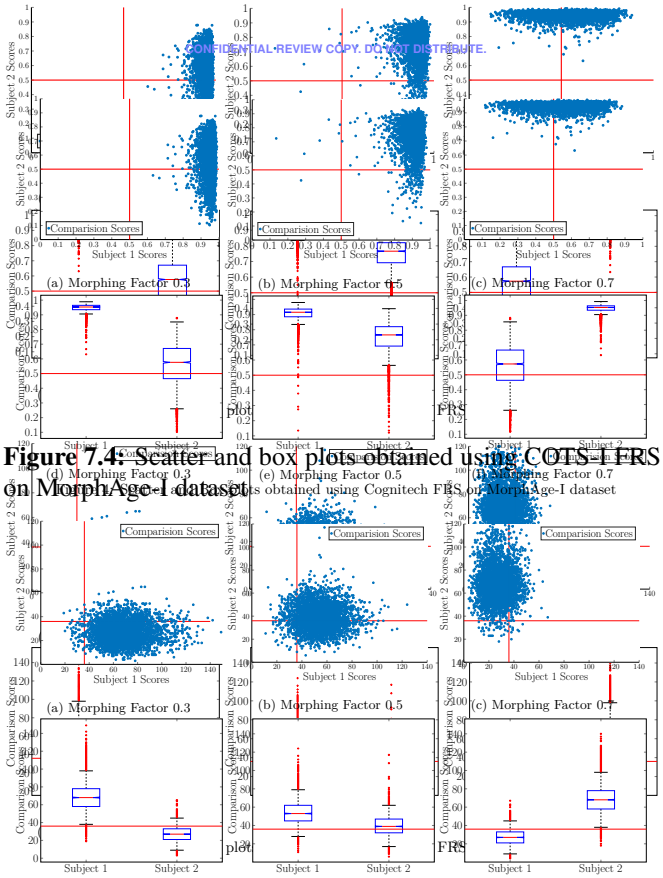


Figure 7.4: Scatter and box plots obtained using COTS-1 FRS on MorphAge-I dataset.

plots obtained using Cognitech FRS on MorphAge-I dataset

Figure 7.5: Scatter and box plots obtained using COTS-II FRS on MorphAge-I dataset.

plots obtained using Neurotech FRS on MorphAge-I dataset

I and MorphAge-II datasets. Figure 4 shows the scatter plot and box plot for MorphAge-I dataset from two COTS FRS respectively. Figure 4a,4b, 4c and Figure 5a,5b,5c provides the visualization of the compar-

(a) Morphing Factor 0.3 (b) Morphing Factor 0.5 (c) Morphing Factor 0.7

Figure 5: Scatter and Box plots obtained using Neurotech FRS on MorphAge-I dataset

I and MorphAge-II datasets. Figure 4 shows the scatter plot and box plot for MorphAge-I dataset from two COTS FRS respectively. Figure 4a,4b, 4c and Figure 5a,5b,5c provides the visualization of the compar-

both composite data subjects are probed for both FRS. In ideal conditions, the FRS is vulnerable to morphing attacks with the proviso all the obtained comparison

both composite data subjects are probed for both FRS. In ideal conditions, the FRS is vulnerable to morphing attacks with the proviso all the obtained comparison

two different COTS Face Recognition Systems (FRS) namely, COTS-I Cognitec[1] FaceVACS-SDK Version 9.4.2 and COTS-II Neurotechnology Version 10.0. To effectively measure the vulnerability of the FRS against morphed face samples, we set a realistic constraint that all contributing data subjects (in our case two) must exceed the verification threshold of the FRS. Further, in this work, we set the operating threshold of both COTS FRS to FAR = 0.1% following the guidelines of FRONTEX [126] for automated border control. Thus, we coin the new realistic constraint using a new vulnerability metric as *Fully Mated Morphed Presentation Match Rate (FMMPMR)* that can be computed as:

$$FMMPMR = \frac{1}{P}\sum_{M,P}(S1_M^P > \tau)\&\&(S2_M^P > \tau)\ldots\&\&(Sk_M^P > \tau) \quad (7.1)$$

Where $P = 1, 2, \ldots, p$ represent the number of attempts made by presenting all the probe images from the contributing subject against M^{th} morphed image, $K = 1, 2, \ldots, k$ represents the number of contributing data subjects to the constitution of the generated morphed image (in our case $K = 2$), Sk_M^P represents the comparison score of the K^{th} contributing subject obtained with P^{th} attempt (in our case the P^{th} probe image from the dataset) corresponding to M^{th} morph image and τ represents the threshold value corresponding to FAR = 0.1%.

We have employed the new metric FMMPMR considering the fact that the existing vulnerability metric MMPMR[125] accounts only for the morphed images getting verified with the contributing subjects without taking into account the number of attempts. However, the new metric FMMPMR overcomes this drawback and considers each and every attempt a morphed image gets verified with the pair of contributing subjects, i.e., reflecting the actual vulnerability of a FRS.

Table 7.2: Vulnerability analysis: FMMPMR (%)

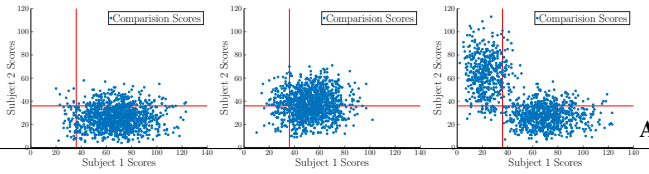
Morphing factor (α)	**FMMPMR**(%)			
	MorphAge-I		MorphAge-II	
	COTS-I	COTS-II	COTS-I	COTS-II
0.3	66.24	18.42	58.47	17.29
0.5	95.07	56.96	93.81	51.27
0.7	67.32	18.21	58.18	15.61

Table 7.2 indicates the FMMPMR (%) computed using the two COTS FRS on both bins - MorphAge-I and MorphAge-II. Figure 7.4 and Figure 7.5 shows the

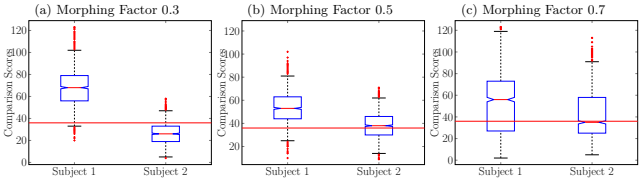
[1]Outcome not necessarily constitutes the best the algorithm can do.

Table 7.3: Experiment-I: Quantitative performance of the MAD techniques on MorphAge-I

Algorithm	**Development Set**	**Testing set**			
	EER (%)	EER (%)	BPCER (%) @ APCER (%) =		
			1	5	10
Morphing factor (α) 0.3					
LBP-SVM [54, 103, 85, 98]	28.14	35.11	84.4	68.8	56.8
BSIF-SVM [54, 98]	31.82	37.59	98.8	90	73.2
HOG-SVM [98]	32.09	33.51	84.4	63.6	53.6
AlexNet-SVM [85, 90, 64]	4.38	2	7.2	3.2	0.8
Color Denoising [4]	1.63	3.65	5.2	0.4	0.4
Morphing factor (α) 0.5					
LBP-SVM [54, 103, 85, 98]	27.82	33.76	75.2	59.2	57.2
BSIF-SVM [54, 98]	31.82	36.9	98.8	89.21	73.6
HOG-SVM [98]	30.73	34.1	81.2	63.2	56.8
AlexNet-SVM [85, 90, 64]	3.18	2.01	4.12	0	0
Color Denoising [4]	1.63	1.21	7.6	0.4	0
Morphing factor (α) 0.7					
LBP-SVM [54, 103, 85, 98]	28.86	34.92	88.4	66.8	57.2
BSIF-SVM [54, 98]	31.9	37.98	98.8	88	73.2
HOG-SVM [98]	32.98	33.38	80	62.8	57.2
AlexNet-SVM [85, 90, 64]	5.08	2.78	5.6	2	0
Color Denoising [4]	2.75	2.43	13.2	2	0.4

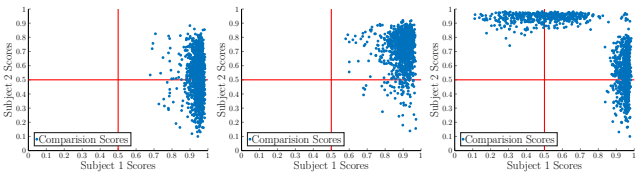


(a) Morphing Factor 0.3 (b) Morphing Factor 0.5 (c) Morphing Factor 0.7

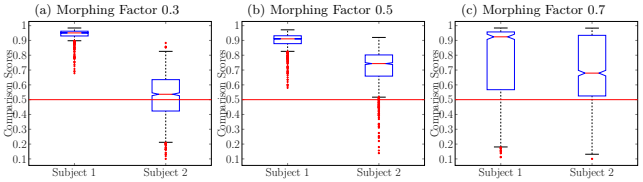


(d) Morphing Factor 0.3 (e) Morphing Factor 0.5 (f) Morphing Factor 0.7

Figure 6: Scatter and Box plots obtained using Neurotech FRS on MorphAge-II dataset



(a) Morphing Factor 0.3 (b) Morphing Factor 0.5 (c) Morphing Factor 0.7

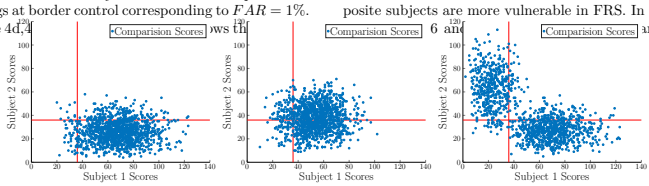


(d) Morphing Factor 0.3 (e) Morphing Factor 0.5 (f) Morphing Factor 0.7

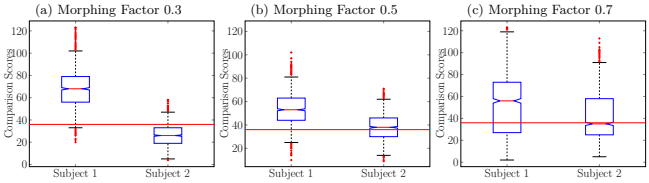
Figure 7: Scatter and Box plots obtained using Cognitech FRS on MorphAge-II dataset

Figure 7.6: Scatter and box plots obtained using COTS-I FRS on MorphAge-II dataset

scores are clustered on the top right corner of the graph. The vertical and horizontal lines indicate the threshold that is recommended by the COTS FRS in operational settings at border control corresponding to $FAR = 1\%$. Figure 4d,4 ows th 6 and and box



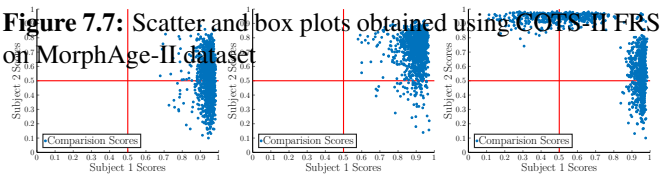
(a) Morphing Factor 0.3 (b) Morphing Factor 0.5 (c) Morphing Factor 0.7



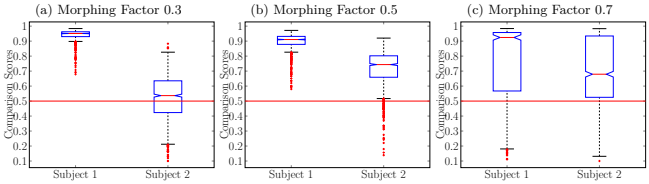
(d) Morphing Factor 0.3 (e) Morphing Factor 0.5 (f) Morphing Factor 0.7

Figure 6: Scatter and Box plots obtained using Neurotech FRS on MorphAge-II dataset

Figure 7.7: Scatter and box plots obtained using COTS-II FRS on MorphAge-II dataset



(a) Morphing Factor 0.3 (b) Morphing Factor 0.5 (c) Morphing Factor 0.7



(d) Morphing Factor 0.3 (e) Morphing Factor 0.5 (f) Morphing Factor 0.7

Figure 7: Scatter and Box plots obtained using Cognitech FRS on MorphAge-II dataset

scores are clustered on the top right corner of the graph. The vertical and horizontal lines indicate the threshold that is recommended by the COTS FRS in operational settings at border control corresponding to $FAR = 1\%$.

that can provide insights on the distributions of comparison scores corresponding to the composite images that allows us to understand which of the two composite subjects are more vulnerable in FRS. In similar

scatter plot and box plot for MorphAge-I dataset from two COTS FRS respectively. Figure 7.4(a), 7.4(b), 7.4(c) and Figure 7.5(a), 7.5(b), 7.5(c) provides the visualization of the comparison scores when the morphed image is enrolled, and both contributing data subjects are probed for both FRS. In the most serve conditions, meaning a high vulnerability of the FRS with regards to morphing attacks, we will obtain comparison scores that are clustered in the top right corner of the figure. The vertical and horizontal lines indicate the threshold that is recommended by the COTS FRS for operational settings in the border control application corresponding to $FMR = 0.1\%$. Figure 7.4(d), 7.4(e), 7.4(f) and 7.5(d), 7.5(e), 7.5(f) shows the box plot that provides insight on the distributions of comparison scores corresponding to the contributor probe images allowing us to understand which of the two probe images (of the contributing subjects) are more vulnerable for the FRS. In similar lines, Figure 7.6 and 7.7 shows the scatter plot and box plot for the MorphAge-II dataset that are computed from the two COTS FRS. Based on the obtained results the following are our main observations:

- *Intra-Age Groups:* As expected the morphed image with the morphing factor of 0.5 indicates the highest vulnerability as reflected by both COTS FRS. However, the morphing factor of 0.3 and 0.7 indicates a reduced vulnerability that can be attributed to the morphing factor weights leaning toward only one of the contributing data subjects. This fact is illustrated in Figure 7.4, 7.5, 7.6 and 7.7, where we can observe that with a morphing factor of 0.3, the subject 1 is likely to be verified. While with a morphing factor of 0.7, in most cases, subject 2 is likely to be verified rather than subject 1. While not so surprising, the morphing factor of 0.5 indicates (almost) equally both contributing subjects can be verified.

- *Inter-Age groups:* Based on the obtained results, it is also interesting to note the direct influence on the morphing factor on the vulnerability. Thus, with the morphing factor of 0.3 and 0.7, both COTS FRS shows a greater reduction value of FMMPMR on MorphAge-II dataset. This indicates morphing attacks pose lesser threats to FRS under the influence of ageing. However, with the morphed factor of 0.5, the COTS-II FRS indicates lower values of FMMPMR, while COTS-I indicates a moderate reduction in the vulnerability despite being very significant.

- Observing the box plots for the morphing factor of 0.5 from both MorphAge-I and MorphAge-II, it can be noted that, both the median and whiskers corresponding to the comparison scores from both subjects are reduced in MorphAge-II when compared to MorphAge-I. These observations, together with the quantitative value of FMMPMR, indicate the reduced threats to

morphing attacks on FRS under ageing. This fact is consistently observed for both COTS FRS and are statistically significant as observed in the box plots.

- *Role of COTS FRS:* The COTS-I FRS indicates the highest vulnerability on three morphing factors when compared to that of the COTS-II FRS. The morphing factor with 0.5 shows the highest FMMPMR with 95.07% on MorphAge-I and 93.81% on MorphAge-II with COTS-I FRS. The lowest value of FMMPMR is noted with COTS-II FRS with a morphing factor of 0.7 in the MorphAge-II dataset.

Table 7.4: Experiment-I: Quantitative performance of the MAD techniques on MorphAge-II

Algorithm	**Development Set**	**Testing set**			
	EER (%)	EER (%)	BPCER (%) @ APCER (%) =		
			1	5	10
Morphing factor (α) 0.3					
LBP-SVM [54, 103, 85, 98]	30.64	29.21	61.24	48.83	44.96
BSIF-SVM [54, 98]	33.35	39.17	58.91	51.16	48.83
HOG-SVM [98]	32.56	32.56	66.66	51.93	45.73
AlexNet-SVM [85, 90, 64]	4	5.49	7.75	4.65	4.65
Color Denoising [4]	3.15	1.7	3.1	0	0
Morphing factor (α) 0.5					
LBP-SVM [54, 103, 85, 98]	28.71	32.39	68.99	48.06	41.08
BSIF-SVM [54, 98]	32.65	39	63.56	51.16	48.83
HOG-SVM [98]	30.33	32.56	62.02	52.71	44.96
AlexNet-SVM [85, 90, 64]	2.92	3.78	6.2	3.87	3.11
Color Denoising [4]	3.77	0.75	1.55	0.77	0.77
Morphing factor (α) 0.7					
LBP-SVM [54, 103, 85, 98]	29.33	27.03	58.91	51.98	45.73
BSIF-SVM [54, 98]	34.7	33.07	58.91	51.16	48.06
HOG-SVM [98]	31.02	29.09	73.64	58.91	44.96
AlexNet-SVM [85, 90, 64]	3.15	5.5	11.62	5.42	4.65
Color Denoising [4]	3.15	0.75	3.1	0	0

7.5 Face Morph Attack Detection Performance

In this section, we benchmark the most recent digital MAD techniques on the newly created MorphAge dataset. The goal of this experiment is to understand the impact of ageing on the detection performance of the MAD techniques. To this extent, we design two different experiments to reflect the variation in the performance of the MAD techniques under the influence of ageing. **Experiment-I:**

the evaluation protocol is designed to evaluate the MAD detectors in the same age group. Thus, the MAD detectors are trained and tested with the same group data. **Experiment-II:** is designed to evaluate the performance of MAD detection with the variation in age. Thus, MAD detectors are trained with the MorphAge-I data and tested with only the MorphAge-II dataset. In both experiments, the corresponding development dataset is used to tune the parameters of the algorithm and also to compute the operating threshold at APCER = 1%, 5% and 10%. In this work, we have evaluated five different MAD schemes such as: Local Binary Pattern (LBP) LBP-SVM [54, 103, 85, 98], Binarized Statistical Image Features (BSIF) [54, 98], Histogram of Oriented Gradients (HOG) [98], AlexNet [85, 90, 64] and Color Denoising [4]. We have considered these five MAD techniques as they have indicated good performance on three different large scale digital morphing datasets [4]. The quantitative results are presented according to the ISO/IEC 30107-3 [21] metrics such as Bona fide Presentation Classification Error Rate (BPCER($\%$)) and Attack Presentation Classification Error Rate (APCER ($\%$)) along with D-EER(%).

Table 7.3 and Table 7.4 indicates the quantitative results of the MAD schemes on two different age groups MorphAge-I and MorphAge-II respectively on the Experiment-I protocol. Based on the obtained results, it can be noticed that:

- The traditional MAD methods based on LBP, BSIF, and HOG fail to indicate acceptable detection performance for both MorphAge-I and MorphAge-II dataset.
- Recently introduced MAD techniques based on AlexNet and Color denoising techniques have shown excellent performance in detecting morphing attacks.
- It is interesting to note that the MAD methods do not show any influence of the different morphing factors on the detection performance. The detection performance with different morphing factor did further not vary irrespective of the age group as well.
- Among the five benchmarked different MAD techniques, the color denoising MAD has indicated the best performance across various morphing factors (α) for both MorphAge-I and MorphAge-II.

Table 7.5 indicates the quantitative detection performance of MAD methods in Experiment-II. Based on the obtained results, it can be noted that the ageing does not influence the performance of the MAD methods. It is worth noting that, in this protocol, MAD methods are trained using only MorphAge-I dataset and are

tested on the MorphAge-II dataset with the age difference up to 5 years. Further, the data subjects in MorphAge-I and MorphAge-II do not overlap. Among the five different MAD methods, color denoising based MAD has again indicated the best performance for all three morphing factors (α). As it can be deduced, ageing does not influence the detection capabilities of MAD under the performed experimental settings.

Table 7.5: Experiment-II: Quantitative detection performance of MAD techniques on MorphAge-I v/s. MorphAge-II

Algorithm	**Development Set**	**Testing set**			
	EER (%)	EER (%)	BPCER (%) @ APCER (%) =		
			1	5	10
Morphing factor (α) 0.3					
LBP-SVM [54, 103, 85, 98]	28.14	34.19	92.24	65.89	47.28
BSIF-SVM [54, 98]	31.82	44.13	100	98.44	84.49
HOG-SVM [98]	32.09	41.86	91.47	70.54	62.01
AlexNet-SVM [85, 90, 64]	4.38	3.03	8.52	3.10	2.32
Color Denoising [4]	1.63	2.27	1.55	0.45	0
Morphing factor (α) 0.5					
LBP-SVM [54, 103, 85, 98]	27.82	33.29	86.04	66.66	48.06
BSIF-SVM [54, 98]	31.82	45.42	100	96.89	84.49
HOG-SVM [98]	30.73	37.95	85.27	67.44	57.36
AlexNet-SVM [85, 90, 64]	3.18	0.94	3.10	0.77	0.39
Color Denoising [4]	1.63	1.59	0.7	0	0
Morphing factor (α) 0.7					
LBP-SVM [54, 103, 85, 98]	28.86	32.24	93.20	65.89	49.61
BSIF-SVM [54, 98]	31.90	37.33	100	96.89	84.49
HOG-SVM [98]	32.98	33.54	88.37	68.99	58.13
AlexNet-SVM [85, 90, 64]	5.08	2.27	6.97	3.87	0.77
Color Denoising [4]	2.75	2.46	3.10	0.40	0

7.6 Discussion

Based on the observations made above from the experiments and obtained results, the research questions formulated in Section 7.2.2 are answered below.

- Q1. How vulnerable are COTS FRS when a composite morph image is enrolled and is after a period of ageing probed against a live image from one of the contributing subjects?
  - Supported by the obtained experimental results reported in Table 7.2, it is interesting to note that the value of FMMPMR is reduced to certain extent in case of MorphAge-II dataset. The morphed images are not

easily verified against the probe images after a certain degree of ageing making FRS less vulnerable.

- Q2. Do current Morphing Attack Detection (MAD) algorithms scale-up to detect such attacks under the influence of ageing?
  - Based on the experimental results reported in Table 7.4, ageing has negligible impact on the MAD and thereby the existing MAD schemes can detect the attacks even under ageing.
- Q3. What is the impact of different alpha (or blending, morphing) factors used to generate the morphed image under the constraint of ageing, specifically with respect to MAD?
  - Based on the experimental results, it is interesting to note that the morphing factors $alpha = 0.3$ and 0.7 show greater reduction in the vulnerability in both the COTS FRS with respect to ageing as reported in Table 7.2. It has to be however noted that COTS-II FRS indicates lower vulnerability when a morphing factor of 0.5 is employed.

7.7 Conclusion

We have presented an empirical study on quantifying the vulnerability of COTS FRS with regards to morphing attacks under the influence of ageing. We have introduced a new dataset with two different age groups derived from the publicly available MORPH II face dataset referred as MorphAge-I and MorphAge-II. Further, we have also introduced a new evaluation metric namely, Fully Mated Morphed Presentation Match Rate(FMMPMR) to quantify the vulnerability effectively. Extensive experiments were carried out using two different COTS FRS and three different morphing factors(with α = 0.3, 0.5 and 0.7). Based on the obtained results, it is observed that impact of ageing reduces the vulnerability from morphing attacks on COTS FRS. The reduction in the vulnerability is more prominent when the morphing factor is α = 0.3 and 0.7. However with a morphing factor of α = 0.5, the vulnerability does not change significantly with the COTS-I, while COTS-II FRS still indicates a significant reduction in the vulnerability. Extensive experiments were performed to quantify the performance variation of the MAD methods under the influence of ageing. To this extent, three different evaluation protocols are presented that show no influence of ageing on morph attack detection performance. It is also interesting to note that robust MAD methods are not sensitive to variations of the morphing factor even under the influence of ageing.

Part IV

Published Articles: Face Morphing Attack Detection

Chapter 8

Article 4: Morphed Face Detection Based on Deep Color Residual Noise

Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, Luuk Spreeuwers, Raymond Veldhuis, and Christoph Busch. *Morphed face detection based on deep color residual noise*. In 9th Intl. Conf. on Image Processing Theory, Tools and Applications (IPTA). IEEE, November 2019

8.1 Abstract

Secure access control applications like border control rely on the face based verification system by considering its reliability, usability and accuracy in-person verification. However, face recognition systems are vulnerable to morphed face attacks, in which, the morphing process combines two different facial images into a single facial image. The features extracted from the morphed face image will match to those extracted from probe images of both faces. Thus, it is essential to reliably detect the morphed face image attacks on the face recognition systems. In this work, we propose a novel approach to detect morphed face images using residual color noise. The proposed method is designed to capture the noise patterns that are a result of the morphing process. Thus, the proposed method performs first denoising using Deep Convolutional Neural Network (CNN) independently on the Hue Saturation Value (HSV) color space, and then computes the residual noise. The extracted residual noise is further processed using Pyramid Local Binary Patterns (P-LBP), which is further classified using the Spectral Regression Kernel Discriminant Analysis (SRKDA). Extensive experiments are carried out on three

different morphed face image datasets. The Morphed Attack Detection (MAD) performance of the proposed method is benchmarked with 13 different state-of-the-art techniques using the ISO IEC 30107-3 evaluation metrics. Based on the obtained quantitative results, the proposed method has indicated the best performance.

8.2 Introduction

Biometric systems employing face, fingerprint or iris recognition are widely deployed to verify the unique identity of an individual in various access control applications. Face Recognition Systems (FRS) are predominantly deployed to verify and establish an identity due to the ease of capture process in a non-invasive manner and at a distance. At the same time, face images are also used in passport based verification both for border crossing and International Civil Aviation Organization (ICAO) based identity verification amongst others. Recently well-used identity check in the airport involves an individual presenting his electronic Machine Readable Travel Document (eMRTD) to verify identity either via Automated Border Control (ABC) gates or to an immigration officer.

While the identity can be verified against a presented image on the passport, many countries issue such documents based on the printed face photo provided by the applicant. Malicious actors can therefore use such an opportunity to provide a tampered face image. One critical case of a tampered face image defeating the FRS is reported as morphed face image, which can successfully verify against multiple individuals. Morphing is an image processing technique used to combine face images of two different individuals, to obtain a single face image. Morphing poses a great threat for the identity check in passport control, as an authentic eMRTD, containing a morphed image, can be used by two different individuals. This applies to the visual inspection process by a border guard, but also to automated processing, when the verification in conducted with the commercial-off-the-shelf (COTS) FRS [119]. The challenge becomes critical, when malicious actors morph the face image against the non-blacklisted subject. This poses a potential threat to security of the border control and thereby it is essential to identify such morphed face images. Motivated by the gravity of the problem, recent research works are focused on detecting morphed images to identify a possible attack on the face recognition system. As indicated in [142], there exists two different techniques for Morphing Attack Detection (MAD): (i) No reference morphing attack detection technique (ii) Differential morphing attack detection techniques. In the former morphing detection technique, an image is analyzed individually without any reference and then classified as a bona fide image or morphed image. In the latter, an image is analyzed based on the stored reference image by using a obtained ref-

Figure 8.1: Example of the morphed face image

erence image, for instance using a live captured image (from Automatic Border Control (ABC) gate) compared against stored reference eMRTD image. In such a scenario, the eMRTD image is classified as bona fide, if the MAD-features of live captured image from the ABC-gate correspond to the extracted MAD-features of the eMRTD image. Further, the no reference morphing attack detection can be of two types depending upon the type of processing of the image data such as [54]: (a) print-scan attack detection in which the digital photo captured in the photo both (or studio) is then printed and handed over to the passport issuance center, where it is again digitised using a scanning device and subsequently stored in the eMRTD. (b) Digital attack detection, where the captured face digitally can be used directly to detect the morphing attacks. since the digital passport photos are used in many countries to renew the passport applications [54]. In this work, we focus on detecting digital face morphed attacks by considering its wide applicability in real-life applications and also it is easy to generate these attacks [54].

The digital morphed face detection is widely addressed in the literature that has resulted in several techniques that can be broadly divided into three types: (a) Texture based (b) image quality based (c) deep learning based approaches. Early works are based on hand-crafted texture-based techniques that are expected to capture the variation in micro textures during the process of morphing that facilitates morph detection. To this extent, several algorithms-based on texture features such as Binary Statistical Image Features (BSIF), Local Phase Quantization (LPQ) [54] and Local Binary Pattern (LBP) [54], [103] and it's variants are introduced to detect the digital version of morphed faces. Among these texture-based techniques, the use of LBP and BSIF has indicated consistent morphed face detection performance. The image quality-based methods are designed to quantify the variation in the compression artefacts and morphing noise introduced during the process

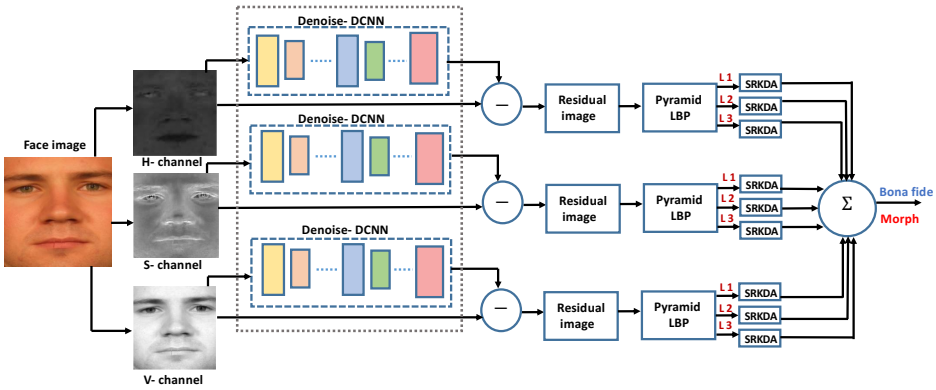


Figure 8.2: Block diagram of the proposed method

of morph generation. Several approaches are presented that includes analysis of Benford features distribution that varies after the jpeg compression [95]. spectral analysis of PRNU [82] and StirTrace [81]. Recent approaches are based on using the deep learning approaches, especially on using the pre-trained CNN architecture like AlexNet, VGG, ResNet, GoogleNet and InceptionV3 [83], [64]. Based on the several experiments reported in the literature, the deep learning-based approaches and the image degradation approaches shows the improved performance over texture-based approaches.

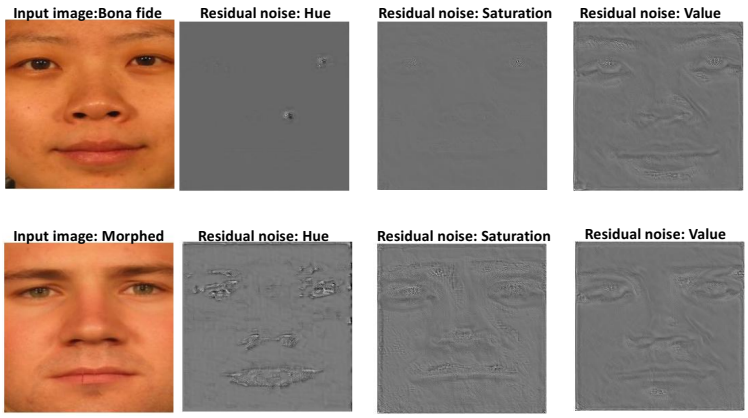


Figure 8.3: Illustration of the residual noise image computed using proposed method on (a) Bona fide image (b) Morphed image

In this work, we present a new approach for no-reference morphing face attack detection by analysing the residual noise that may be attributed due to the process of face morphing. Specifically, we compute the residual noise using a Deep-CNN

based denoising network on each of the color channels of the given face image. On the obtained residual noise of each channel, we further extract the textural features using the Pyramidal Local Binary Pattern (P-LBP) to better quantify the noise patterns. Finally, we learn a spectrally regressed kernel discriminant (SRKD) to discriminate between the bona fide and morphed image. To validate the intuition of our proposed approach, we employ three different large scale datasets compromised of bona fide and morphed images. Through empirical evaluation on these three large scale datasets, we demonstrate a superior Morphing Attack Detection (MAD) performance and further compare it against the detection performance of hand-crafted and deep-learning features. The key contributions of this work includes: (1) Novel method for morphed face detection based on the color residual noise computed based on D-CNN denoising technique. (2) Extensive experiments are carried out on three different face morphing datasets and the performance of the proposed method is benchmark with 13 different state-of-the-art techniques.

The rest of the paper is organized in the following order: Section 8.3 presents our proposed method. Section 8.4 describes the experiments and results obtained. Finally, Section 8.5 draws the conclusion.

8.3 Proposed Method

Figure 8.2 shows the functional block diagram of the proposed method for the proposed face morphing detection. The face morphing process combines two face images using mathematical operations to obtain the morphed face image. This results in the morphed image that adulterates the noisy components due to pixel discontinuities. Thus, we assert that computing these residual noises that are expected to be in high magnitude in the morphed face images when compared to that of the bona fide face images. Such irregularities can help in revealing the morphed manipulation. The proposed method is structured to compute the residual noise from the individual color spaces. Given the color (RGB) face image I_{RGB}, the first step is to decompose the image into HSV color space that can better capture the distinct characteristics of the bona fide and morphed images. For example, the morphed images may have different characteristics of edges, textures, shade and color smoothness. These characteristics can be best described by decoupling the intensity from the chroma component using HSV color space. Let the HSV color image can be represented as I_{HSV}.

In the next step, the image denoising is carried out on individual color channels. To this extent, we propose to employ, the Denoise Deep Convolutional Neural Network (De-DCNN) denoising method from [131] by considering it's denoising performance. In this work, we use the pre-trained De-DCNN that is trained using natural images with a large variety of noise [131]. We then carry out the denoising

of individual color channels to get corresponding denoised images: $I_H^{'}, I_S^{'}, I_V^{'}$. In the next step, we compute the residual noise independently on three channels as: $RN_H = I_H - I_H^{'}$, $RN_S = I_S - I_S^{'}$, $RN_V = I_V - I_V^{'}$. Figure 8.3 illustrates the residual noise computed from the HSV color space that clearly indicates the distinction between bona fide and morphed image.

In the next step, we further process the residual noise images using Pyramid-Local Binary Component (P-LBP) features to quantify the residual noise effectively. We are motivated to employ this approach by considering it's efficiency in modeling the residual noise as indicated in [183]. In this work, we use the Laplacian Pyramid with three level decomposition independently on residual noise image on which the LBP is computed. Given the residual image, the proposed method provides three sets of features computed using LBP corresponding to three level Laplacian pyramid. Thus, in total there are 9 different P-LBP features computed from 3 different residual noise images as: $RN_H^{L1}, RN_H^{L2}, RN_H^{L3}, RN_S^{L1}, RN_S^{L2}, RN_S^{L3}, RN_V * L1, RN_V * L2, RN_V * L3$. Finally, we train the morph detector based on Spectral Regression Kernel Discriminant Analysis [184] independently on nine different features using a training set. Given the test image, we compute the Morph Attack Detection (MAD) score corresponding to 9 different features as:

$MD_{f1}, MD_{f2}, MD_{f3}, MD_{f4}, MD_{f5}, MD_{f6}, MD_{f7}, MD_{f8}$, MD_{f9}. Final decision is computed by combining the MAD scores using a sum rule as: $\sum_{1=1}^{9} MD_{fi}$.

8.4 Experiments and Results

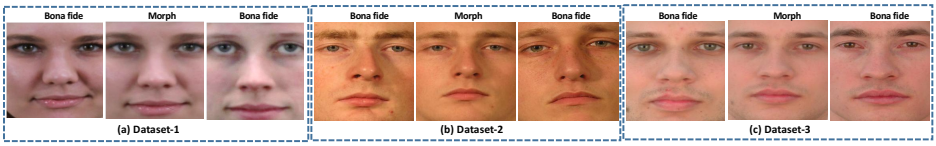


Figure 8.4: Illustration of the example images from (a) Dataset-1 (b) Dataset-2 (c) Dataset-3

In this section, we present the quantitative results of the proposed method together with 13 different State-Of-The-Art (SOTA) techniques for the morphed face image detection. Experimental results are presented using the ISO30107-3 [21] metrics such as Bona fide Presentation Classification Error Rate (BPCER(%)) and Attack Presentation Classification Error Rate (APCER (%)) along with D-EER(%). BPCER defines the proportion of bona fide presentations incorrectly classified as attack images and APCER defines attack images incorrectly classified as the bona fide images [21].

Extensive experiments are presented on three different datasets namely: **Dataset-**

Table 8.1: Quantitative performance of the MAD algorithms on Experiment-1 (individual dataset)

Algorithms	Database-1			Database-2			Database-3		
	D-EER(%)	BPCER@ APCER		D-EER(%)	BPCER@ APCER		D-EER(%)	BPCER@ APCER	
		=5%	=10%		=5%	=10%		=5%	=10%
AlexNet-SVM [90]	5.50	3.5	2.33	7.08	8.95	4.85	11	22	12
GoogleNet-SVM [90]	9.63	13.66	8.83	11.95	22.38	14.55	42.23	100	77.23
InceptionV3-SVM [90]	11.66	18.83	12.33	8.21	11.94	8.20	11.94	26	16
ResNet101-SVM [90]	5.51	6.16	4	6.48	6.10	4.74	13.76	32	22
VGG16-SVM [90] [83]	13.31	25	16.83	14.50	28.35	18.28	21.86	100	36
VGG19-SVM [90] [64]	12.49	22.66	15	12.32	22.38	14.17	24.50	52	40
BSIF-SVM [98]	26.70	53	42	12.67	25.74	14.55	20.45	44	32
Steerable pyramid - SVM [90]	26.19	65.50	50	37.97	82.08	71.64	34.00	82	70
HOG-SVM [98]	10.37	19.83	10.50	12.30	23.50	14.92	11.91	26	10
Image Gradient-SVM [54]	17.34	38	26.50	25.24	51.86	39.92	31.98	72	60
LBP-SVM [54]	18.67	39.16	28.16	9.31	14.55	8.20	22.06	62	38
PRNU [82]	26.51	57.16	44.67	39.89	78.35	70.15	35.62	74	58
LPQ-SVM [54]	17.30	43.66	28.66	13.43	26.11	16.41	20.24	56	38
Proposed Method	**3.83**	**3**	**1.5**	**4.85**	**4.85**	**3.35**	**9.71**	**14**	**8**

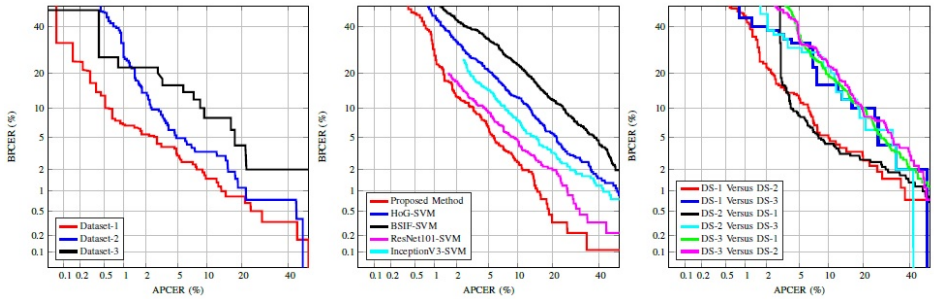


Figure 8.5: DET Curves (a) performance of the proposed method on three different datasets in Experiment-1 (b) performance of the top five MAD algorithms including the proposed method on Experiment-2 (c) performance of the proposed method in Experiment-3 (cross dataset)

1: This dataset is constructed using 179 unique subjects that are divided into two disjoint independent sets namely training (89 subjects) and testing (90 subjects). These subjects are selected using both publicly available and private face datasets. The morphing process is carried out using the open source tool mentioned in [13] that has resulted in a training set with 709 bona fide and 1255 morphed images and testing set with 918 bona fide and 1354 morphed images. Figure 8.4 (a) shows the example of the bona fide and morphed images. **Dataset-2:** This dataset is constructed using the FRGC dataset in which 568 data subjects are used to generate morphed images using an automatic method based on facial landmark and triangulation as described in [103]. This dataset has 300 bona fide and 3041 morphed images corresponding to 300 data subjects used as the training set. While the testing set is generated using 268 data subjects that correspond to 268 bona fide and 2739 morphed images. Figure 8.4 (b) illustrates the example image from this dataset. **Dataset-3:** This dataset is comprised of 100 data subjects selected from putDB dataset which is publicly available. The training dataset is comprised of 50 data subjects that are used to generate 50 bona fide and 254 morphed samples. The testing dataset is comprised of 50 data subjects that are used to generate 50 bona fide and 244 morphed images. The morphing process is based on facial landmark and triangulation as described in [103]. Figure 8.4 (c) shows the example of the bona fide and morphed images.

In this work, we have evaluated 6 deep learning based SOTA and 7 non-deep learning based techniques. In case of deep learning technique, we have used the pre-trained network and compute the corresponding features that are further classified using linear Support Vector Machines (SVM). To this extent, we have considered pre-trained CNN such as AlexNet [90], GoogleNet [90], Inception V3[90], ResNet101[90], VGG16 [90] and VGG19 [90]. In case of non-deep learning techniques, texture-based techniques such as; LBP [54], LPQ [54], BSIF [98], Steerable Pyramids [90] together with image distortion based features such as Image gradients[54] , HoG [98] and PRNU [82] are used together with linear SVM (except for PRNU) to compute the detection performance. To effectively evaluate the performance of the Morph Attack Detection (MAD) schemes, we perform three different experiments such as **Experiment-1:-** is designed to evaluate the performance of the MAD schemes when training and testing are done on the same dataset. **Experiment-2:-** is designed to evaluate the MAD schemes on the merged dataset in which all three datasets are merged into one dataset. This experiment provides an insight into the MAD performance when the dataset is increased with a number of samples. **Experiment-3:-** is designed to perform the cross-dataset comparison in which one of the datasets is used for training and another dataset is used for testing. This experiment will highlight insights on MAD techniques that tested on the unknown dataset.

Table 9.3 shows the quantitative performance of the proposed method, along with 13 different SOTA techniques on Experiment-1. It is interesting to note that (1) the MAD performance of the deep learning features shows the improved performance over non-deep learning methods on all three datasets. (2) Among three different the performance of the all MAD techniques indicate the degraded performance that can be attributed to the characteristics of the dataset. (3) The ReseNet101 and Inception V3 based features indicate better performance over other DCNN features on all three datasets. (4) Among the non-deep features, LBP and HoG schemes indicate better performance over other non-deep features on all three datasets. (5) The proposed method has indicated the best performance when compared to that of the 13 different SOTA techniques on all three different datasets. Figure 8.5 (a) shows the DET curves indicating the performance of the proposed method on all three datasets evaluated in this work.

Table 8.2: Quantitative performance of the MAD algorithms on Experiment-2 (merged dataset)

Algorithms	D-EER(%)	BPCER@ APCER	
		=5%	=10%
AlexNet-SVM [90]	9.70	17.32	9.36
GoogleNet-SVM [90]	10.87	21.35	11.98
InceptionV3-SVM [90]	8.69	14.59	7.51
ResNet101-SVM [90]	7.77	9.04	4.68
VGG16-SVM [90]	12.83	25.49	15.03
VGG19-SVM [90]	12.19	24.50	15.03
BSIF-SVM [98]	15.58	33.98	23.09
Steerable Pyramid-SVM [90]	36.78	77.88	68.08
HOG-SVM [98]	11.32	20.69	12.52
Image Gradient-SVM [54]	38.41	79.84	68.84
LBP-SVM [54]	36.58	73.42	63.98
PRNU [82]	36.88	76.84	65.35
LPQ-SVM [54]	15.03	30.28	19.82
Proposed Method	**5.34**	**6.31**	**2.50**

Table 9.4 indicates the quantitative performance of the proposed method on the Experiment-2 in which all three datasets are merged. Based on the obtained results, the deep features indicate better performance over non-deep techniques. Further, the proposed method has indicated the best performance with D-EER = 5.34% with BPCER = 6.31% @APCER = 5% and BPCER = 2.50% @APCER = 10%. These obtained results further justify the robustness of the proposed method to the increased number of samples with different image characteristics. Figure 8.5 (b) shows the DET curves indicating the performance corresponding to the top five best performing techniques including the proposed method.

Table 9.5 indicates the quantitative performance of the proposed method on the

Table 8.3: Quantitative performance of the MAD algorithms on Experiment-3 (cross Dataset)

Training Dataset	Test Dataset	Algorithms	D-EER(%)	BPCER@ APCER	
				=5%	=10%
Database-1	Database-2	Alexnet-SVM [90]	50	100	100
		Resnet101-SVM [90]	50	100	100
		HoG-SVM [98]	17.97	38.43	28.35
		Proposed method	**7.12**	**12.31**	**5.22**
Database-1	Database-3	Alexnet-SVM [90]	19.63	32	24
		Resnet101-SVM [90]	13.96	100	18
		HoG-SVM [98]	20.24	50	30
		Proposed method	**13.76**	**32**	**16**
Database-2	Database-1	Alexnet-SVM [90]	8.14	11.66	7.33
		Resnet101-SVM [90]	9.82	16.33	9.66
		HoG-SVM [98]	6.81	9	4.83
		Proposed method	**6.49**	**8.50**	**4.16**
Database-2	Database-3	Alexnet-SVM [90]	19.83	38	34
		Resnet101-SVM [90][90]	13.76	26	16
		HoG-SVM [98]	12.35	34	20
		Proposed method	**13.76**	**30**	**22**
Database-3	Database-1	Alexnet-SVM [90]	50	100	100
		Resnet101-SVM [90]	14.68	100	18.83
		HoG-SVM [98]	14.52	32	19.16
		Proposed method	**14.40**	**36.16**	**19.50**
Database-3	Database-2	Alexnet-SVM [90]	50	100	100
		Resnet101-SVM [90]	17.27	100	100
		HoG-SVM [98]	24.28	58.20	42.53
		Proposed method	**15.31**	**33.95**	**23.50**

Experiment-3 (cross-dataset evaluation). For simplicity, we have presented the results only for the top four best performing MAD techniques based on Experiment-1 and Experiment-2. Since we have three different datasets, we get six different cases in which one dataset is enrolled and the remaining two datasets are probed. Based on the obtained results, the proposed method shows improved performance when compared with the SOTA methods. However, when dataset-3 is used as the probe, the performance of the proposed method is comparable with the SOTA methods. Figure 8.5 (c) shows the DET curves of the proposed method on all six different cases of cross dataset comparison. For simplicity, we have indicated DET curves for selected techniques, however, detailed quantitative results are presented in Table 9.3, Table 9.4 & Table 9.5.

Thus, based on the extensive experiments carried out on three different datasets, the proposed method has indicated the best MAD performance when compared with 13 different SOTA techniques. Quantitative results obtained on three different experiments shows the best performance of the proposed method, which justifies the applicability of the residual noise computed based on the deep denoising technique for the robust morphed face detection.

8.5 Conclusion

In this work, we propose a novel method based on denoising to identify the presence of a morph attack. Existence of residual noise that is obtained after getting the difference of the face image with its denoised version indicates the presence of morphing. Face image in HSV color space is denoised to obtain the difference image that is obtained by subtracting the given image with its denoised version in HSV color space. Difference obtained after subtraction is the residual noise on which pyramid LBP is applied to get the spatial features with three level decomposition. Further the spatial features are classified using SRKDA classifier to reliably identify the given image as bona fide or morphed.

Extensive experiments are carried out on three different morphed face databases (digital version). We present an evaluation on 13 different algorithms based on deep learning and non-deep learning features. Among six different deep features and seven different non-deep features our proposed method based on denoising outperforms the existing techniques. Performance of the proposed method on dataset-1 gives a D-EER of 3.83% with BPCER = 1.5% at APCER = 10% and BPCER = 3% at APCER =5%. dataset-2 gives a D-EER of 4.85% with BPCER = 3.35% at APCER = 10% and BPCER = 4.85% at APCER = 5%. Finally dataset-3 presents a D-EER of 9.71% at BPCER = 8% at APCER = 10% and BPCER = 14% at APCER = 5%. Quantitative results obtained on all three datasets indicates the consistent performance that shows the robustness and reliabitlity of the proposed

method.

Chapter 9

Article 5: Detecting Morphed Face Attacks Using Residual Noise From Deep Multi-Scale Context Aggregation Network

Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, Luuk Spreeuwers, Raymond Veldhuis, and Christoph Busch. *Detecting morphed face attacks using residual noise from deep multi-scale context aggregation network*. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). IEEE, March 2020.

9.1 Abstract

Along with the deployment of the Face Recognition Systems (FRS), concerns were raised related to the vulnerability of those systems towards various attacks including morphed attacks. The morphed face attack involves two different face images in order to obtain via a morphing process a resulting attack image, which is sufficiently similar to both contributing data subjects. The obtained morphed image can successfully be verified against both subjects visually (by a human expert) and by a commercial FRS. The face morphing attack poses a severe security risk to the e-passport issuance process and to applications like border control, unless such attacks are detected and mitigated. In this work, we propose a new method to reliably detect a morphed face attack using a newly designed demising framework. To this end, we design and introduce a new deep Multi-scale Context Aggregation Network (MS-CAN) to obtain denoised images, which is subsequently used

to determine if an image is morphed or not. Extensive experiments are carried out on three different morphed face image datasets. The Morphing Attack Detection (MAD) performance of the proposed method is also benchmarked against 14 different state-of-the-art techniques using the ISO30107-3 evaluation metrics. Based on the obtained quantitative results, the proposed method has indicated the best performance on all three datasets and also on cross-dataset experiments.

9.2 Introduction

An electronic Machine Readable Travel Document (eMRTD) is a governmental document (e.g. an electronic Passport) that stores face biometric reference images corresponding to the owner of the document. When a bona fide citizen makes the application for an eMRTD in his respective country, the applicant provides a passport photo that is taken by a photographer. Depending upon the type of the application (online or in-person), the applicant submits his/her passport photo either in digital or printed form, where printed passport photos are subsequently scanned for the digitized eMRTD production process. The submitted passport photo either in *digital* or re-digitized through scanning *i.e. print-scan*) is stored in the eMRTD.

A malicious actor in such a setting can submit a morphed face image and obtain a valid eMRTD leading to exploitation of intrinsic intra-class variation tolerance of a Face Recognition Systems (FRS), which was revealed as a serious vulnerability of FRS [22]. The morphed face image generated using the face image from an attacker and a accomplice can easily be verified against both contributing subjects with existing commercial FRS. Also a human expert such as a trained border guard can be confused [185, 135, 54, 134, 29, 13, 186, 15]. This scenario becomes critical, when attackers intentionally morph their face image with a non-blacklisted subject, in order to gain access to a protected/secured area. This poses a severe threat to the security and efficacy of border control or similar applications (using eMRTD) and thereby, it is crucial to identify such morphed face images and to prevent the attacks. A sample of morphed face image and the obtained comparison scores using a commercial FRS is illustrated in Figure 9.1.

Motivated by the problem, several Morphing Attack Detection (MAD) techniques to flag digital morphed face images and print-scanned morphed face images have been proposed [135, 54, 13, 186, 15, 103, 98, 84, 104]. In this work, we focus on detecting digital morphed face images as: (i) they can be easily generated in the digital domain, (ii) digital images are used in several countries like New-Zealand, Estonia, Ireland, etc. to issue/renew the documents and (iii) the constitute a low-cost attack in digital domain. Further, it has to be noted that the digital morph image is usually uploaded to an online passport application portal by the applicant

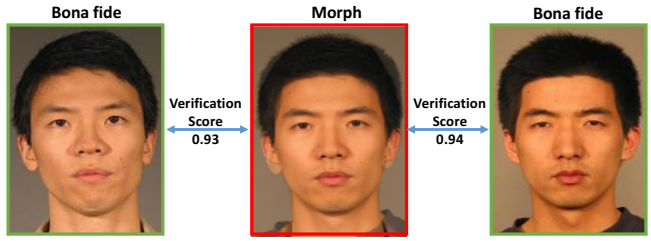


Figure 9.1: Illustration of successful verification with morphed image in a COTS Face Recognition System (FRS) operating at $FAR = 0.01\%$ (a) Subject 1 (b) Morphed face image (c) Subject 2

and there is no human control to verify the authenticity of image as in a physical passport application procedure.

Table 9.1: State-of-the-art digital MAD techniques

Reference	Algorithm Type	Algorithm
Raghavendra et al. [54]	Texture based method	Local binary pattern (SVM), Binary Statistical Image Features (BSIF), Image Gradient(IG)
Makrushin et al. [95]	Quantized DCT coefficients	Benford features
Hildebrandt et al. [81]	Stir trace based scenario	Multi-compressed Anomaly detection
Neubert [96]	Image degradation approach	Corner feature detector
Seibold et al. [64]	Deep learning based approach	VGG19, GoogleNet, AlexNet
Asaad and Sabah [97]	Texture based scenario	Topological data analysis approach
Scherhag et al. [98]	Texture and frequency based method	LBP, LPQ, BSIF, 2DFFT with SVM classifier
Debiasi et al. [104]	Image Quality	PRNU using Wavelet denoising
Raghavendra et al. [83]	Deep CNN based method	Feature fusion of fully connected layers of VGG19 and Alex Net
Damer et al. [84]	Deep and texture features	Feature fusion of LBP and Openface Net
Ferrara et al. [85]	Deep features	AlexNet, VGG19, VGG16, ResNet50
Sushma et al. [4]	Deep residual noise	Color residual noise with SRKDA

9.3 Related Works

In this section, we summarize the existing MAD techniques in Table 9.1 for a quick comprehension of the reader. As observed from Table 9.1, the most prevalent MAD techniques can be broadly divided into four algorithm types: (a) texture-based (b) image quality based (c) deep learning-based (d) hybrid features (combined/multiple features) based detection. The first work on detecting the morphed face images based on micro-textures was presented in [54]. Following this work, several other works are reported [103, 98] using the capability of micro-texture extraction techniques that can effectively capture the variations to reflect the process of morphing, which aids the morph detection task. Lately, the use of pre-trained deep CNNs with different architectures are widely studied in [64, 83, 85]. Further, the combination of deep features with handcrafted features is proposed in

[84]. Recently, the spectral analysis of Photo Response Non-Uniformity has been employed [104][80], to analyse modifications caused by the morphing procedure. For a quick overview of the existing state of the art based on morph attack detection are presented in [142]. In the recent past several approaches based on hybrid features and deep features are presented [91, 15, 117]. The combination of deep features with handcrafted features is proposed in [84]. Recently, the residual noise computed on the color channels using deep CNN based denoising is presented for the reliable face morph detection [4].

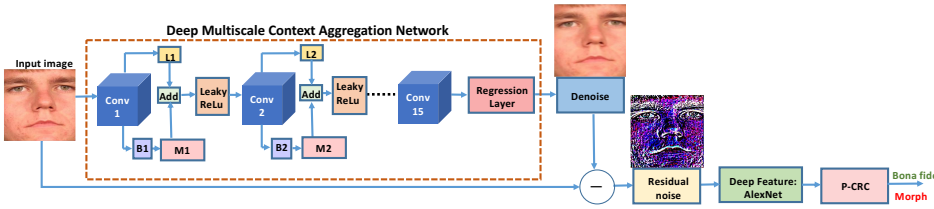


Figure 9.2: Block diagram of the proposed method. B denotes batch-normalization, M represents the scale layer that adjusts the strength of the batch-normalization, L corresponds to strength of the identity branch in batch-normalization.

9.3.1 Our Contributions

Intrigued by the effectiveness of the photo-response noise and it's success in detecting morphed attacks, we investigate to detect the noise of the morphing process using a new approach. We assert that the strategy of localizing such a noise using learning approaches lead to better detection of morphing attacks. Thus, in this work, we present a novel method for the face morphing attack detection by computing the residual noise, which can be attributed to the morphing process. The intuition behind resorting to such an approach of determining the noise using a deep learning paradigm is due to three specific reasons, where the resulting noise due to the morphing process can be: (i) random (ii) non-deterministic and abrupt (iii) sparsely distributed.

Given such properties, we first focus on commonly characterized noise in the image domain and the approaches to denoise them. The widely employed denoising approaches include Wavelet Denoising (WD) [187], Block Matching and 3D filtering (BM3D) [188], Multi-resolution Bilateral Filtering (MBF) [189] and Denoising Convolutional Neural Networks (DnCNN) [131] which can intuitively cover the possible noise in morphing process. A combination of all such denoising approaches can lead to better morphing attack detection, as asserted earlier. However, the complexity in time and parameterization of each of these approaches can lead to the cumbersome effort. In the light of the recent advancements in deep learning, we propose to aggregate the denoising approaches [187, 188, 189, 131] using

a deep Multi-scale Context Aggregation Network (MS-CAN) such that the noise in the morphed image can be easily determined, i.e., given the face image I, we obtain the denoised face image I_d using the MS-CAN. We then compute the residual noise I^r, which is employed to determine if the image I is morphed or not (bona fide). Given the residual noise image, we adapt the pre-trained off-the-shelf AlexNet to extract textural features. These features are then classified using a Collaborative Representative Classifier (CRC) to discriminate between the bona fide and morphed image.

The key contributions of this work can therefore be summarized as:

- We present a novel method for detecting morphed face images based on the deep textural features of residual noise from image.
- We introduce a deep Multi-scale Context Aggregation Network (MS-CAN) for aggregating four denoising methods to consider various kinds of noise characteristics.
- We present results and extensive experiments on three different face morphing datasets, and benchmark the results for our proposed approach with 14 different state-of-the-art techniques.

The rest of the paper is organized as follows: Section 9.4 presents the proposed method, Section 9.5 discusses the morphed face dataset used in this work, Section 9.6 discusses the quantitative performance of the state-of-the-art face Morphing Attack Detection (MAD) together with the proposed method under different evaluation protocols. Finally, Section 9.7 draws the conclusion.

9.4 Proposed Method

As noted earlier, the morphing process can involuntarily introduce noise in the resulting morphed image. The core of the proposed method is therefore to quantify the morphing noise effectively given the recent work indicating the effectiveness of noise characterization in detecting morphing attacks [104] The motivation of this work is to explore the image denoising methods to quantify the noise and thereby detect the face morphing attacks reliably. The residual noise obtained from the image can enable reliable detection of no-reference (single image) based morph images. The proposed approach for such motivation is provided in Figure 9.2, which characterizes the noise pattern. The proposed method can be visualized in two main parts: (a) aggregation of multiple denoising methods realized using MS-CAN (b) feature extraction and classification, both of which are explained in the section below.

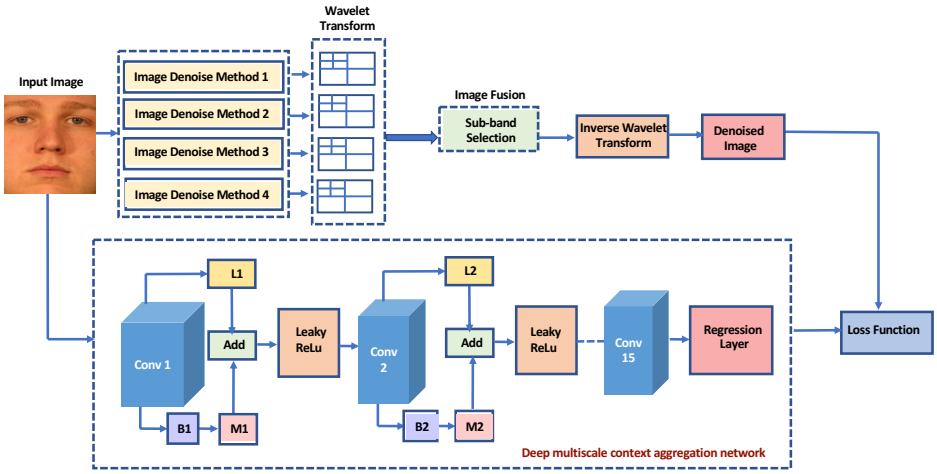


Figure 9.3: Realizing the multiple-denoising approach using a deep Multi-scale Context Aggregation Network (MS-CAN). B denotes batch-normalization, M represents the scale layer that adjusts the strength of the batch-normalization, L corresponds to strength of the identity branch in batch-normalization.

9.4.1 Aggregation of multiple denoising methods realized using MS-CAN

Figure 9.3 shows the block diagram for realizing the aggregation of multiple denoising methods through deep MS-CAN. Given the RGB color image I, the first step is to perform the denoising operation. Among several types of image denoising methods, we choose four complementary methods by considering their performance and also the mode of operation (spatial/frequency/sparse). To this extent, we have used the selected denoising methods that namely Wavelet Denoising (WD) [187], Block Matching and 3D filtering (BM3D) [188], Multiresolution Bilateral Filtering (MBF) [189] and DeNoising Convolutional Neural Networks (DnCNN) [131]. Let I_{D1}, I_{D2}, I_{D3} & I_{D4} represent the denoised images corresponding to WD, BM3D, MBF and DnCNN respectively. In the next step, we perform the aggregation to obtain a single denoised image that can represent the best of all four denoising techniques. The aggregation of best-denoised parts within the image is carried out through the wavelet-based image fusion technique, where each denoised image is decomposed into sub-bands. As the pixel values in the sub-bands from different denoising approaches are multiple. We employ the criteria for selecting the best sub-band with the highest energy values for reconstructing the final denoised image (using the inverse wavelet transform). We are motivated to use wavelet-based image fusion as it can handle multi-resolution images. Further, the image fusion strategy based on the selection of sub-bands with the highest energy allows us to retain the edge components preserved from multiple denoising

methods.

Given the denoised image I_{Di}, $\forall i = \{1, \ldots N\}$ where N represents the number of denoising methods. The corresponding wavelet decomposition (with level 2) of I_{Di} results in four different sub-band images such as approximate sub-band $\{a^N\}$, horizontal sub-band $\{h_1^N, h_2^N\}$, vertical sub-band $\{v_1^N, v_2^N\}$ and diagonal sub-band $\{d_1^N, d_2^N\}$. In the next step, we compute the energy corresponding to each sub-band that can be represented as $\{E_a^N\}$, $\{E_{h1}^N, E_{h2}^N\}$, $\{E_{v1}^N, E_{v2}^N\}$, $\{E_{d1}^N, E_{d2}^N\}$. The image fusion is performed by selecting the sub-band that corresponds to the highest energy as: $S_{h1} = h_1^{N(k)}$, where $k = \max_{i=1}^{N} \{E_{h1}^N\}$ is the index that corresponds to the highest energy. For example, if the highest energy for the horizontal sub-band h_1 is noted with the N^{th} image denoising method, then it is selected. We follow the same procedure for the remaining sub-bands to obtain $S_{h2}, S_{v1}, S_{v2}, S_{d1}, S_{d2}$ and S_a. Finally the fused denoised image I_F is obtained by taking the inverse wavelet transform.

Considering the computational effort and the parameterization of the aggregation of multiple denoising methods, we simply realize the operation of multiple denoising using a deep learning approach. It is shown in earlier works [190] [191] that, approximated image processing operations using deep MS-CAN can result in a highly accurate, robust and time-efficient technique. Inspired by such findings in [190, 191], we design our architecture in a similar fashion for our aggregated denoising approach. As indicated in Figure 9.3, the deep MS-CAN architecture consists of 15 layers of 3×3 convolution layers with exponentially increasing dilation factor. Thus, the dilation corresponding to the convolution layers are $1, 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096, 1$. Each convolution layer in the network is connected to a point-wise non-linearity using the leaky rectified linear unit (leaky-relu). Further, the adaptive normalization [191] is employed to combine both batch normalization and identity normalization. As shown in the architecture (see Figure 9.3), B_x (where $x = 1, 2, \ldots, 15$, number of layers) represents the batch-normalization, M_x represents the scale layer that adjusts the strength of the batch-normalization, L_x represents the scale layer to adjust the strength of the identity branch. We then use the additional layer to combine both M_x and L_x. The network is trained on input-output pairs that contain images from before and after the proposed denoising operation. We further employ Mean Squared Error (MSE) within regression loss function to estimate the learnability of the aggregation (approximation) operation.

$$L = \sum_{i} \frac{I_{Fi} - \hat{I_F}}{R} \tag{9.1}$$

where, R is the number of responses, I_{Fi} is the target output and $\hat{I_F}$ is the network

prediction for response i.

Training details of MS-CAN

To effectively realize the generalizability of the proposed deep MS-CAN, we train the network on natural images (including photos of people, building, natural scenes, etc.) from IAPR TC-12 [1]. We further perform the proposed multiple denoising fusion approach on this dataset to obtain the denoised image. We then train the deep MS-CAN using pairs of normal-denoised image. The Adam optimizer is used with a constant learning rate of 0.0001 and the training is carried out for 250 epochs resulting in 1.2 million iterations. We subsequently use the trained deep MS-CAN to perform the denoising operation and compute the residual noise that can be used to detect a morphing attack as shown in the Figure 9.2.

Figure 9.4 illustrates qualitative results of the residual noise computed on bona fide and morphed face images using deep MS-CAN. The variation in noise intensity between bona fide and morphed image can be observed and this asserts our intuition. These qualitative results further support our approach of detecting morphing attacks based on residual noise despite learning from general image datasets.

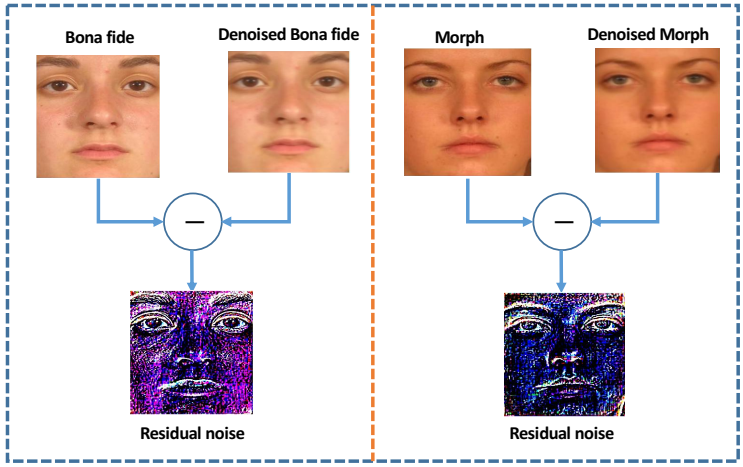


Figure 9.4: Illustration of residual noise computation using deep MS-CAN

9.4.2 Feature extraction and detection

Given the residual image, we extract the deep textural features computed using a pre-trained off-the-shelf AlexNet. We have used the features from fully connected layer $fc6$ to compute the feature from the residual noise images. These computed

[1] https://www.imageclef.org/photodata

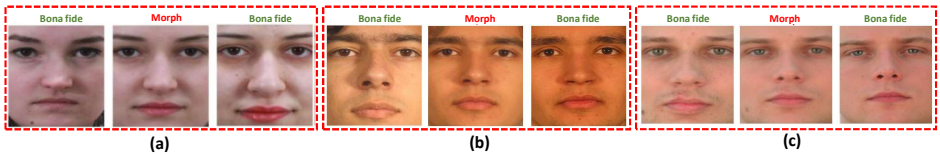


Figure 9.5: Example images from (a) Dataset-1 (b) Dataset-2 (c) (a) Dataset-3

features are then classified using a Probabilistic Collaborative Representation Classifier (P-CRC) [192]. The P-CRC used in this work utilizes the Regularized Least Square Regression (LSR) on the learned feature vectors versus the probe feature vectors [192] formulated as:

$$\widehat{F} = argmin_{\alpha} \left\| Tr_F - \mathscr{D}\alpha \right\|_2^2 + \lambda \left\| \alpha \right\|_2^2 + \frac{\psi}{K} \left\| X\alpha - X_K\alpha_K \right\|_2^2 \tag{9.2}$$

Where, Tr_F is the feature vector of the test image, $\mathscr{D}$ is the learned collaborative subspace dictionary using Tr_F, α is coefficient vector, X is the collection of the training features corresponding to K classes and λ and ψ are the regularization parameter.

Table 9.2: MAD performance on individual image denoising techniques and the proposed method

	Dataset-1			Dataset-2			Dataset-3		
		BPCER@ APCER			BPCER@ APCER			BPCER@ APCER	
Algorithms	**D-EER(%)**	**=5%**	**=10%**	**D-EER(%)**	**=5%**	**=10%**	**D-EER(%)**	**=5%**	**=10%**
BM3D [188]	15.03	40.50	22.50	25.04	55.59	42.16	14.37	32	26
WD [187]	27.96	42.83	31.50	31.35	81.34	67.53	18.01	46	34
MBF [189]	8.69	12.16	8.16	8.69	10.44	8.20	9.71	10	8
DnCNN [131]	19.82	42.83	31.50	24.96	54.85	44.77	19.83	54	38
Proposed Method	**3.24**	**3**	**1.67**	**2.63**	**1.11**	**1.11**	**7.89**	**8**	**4**

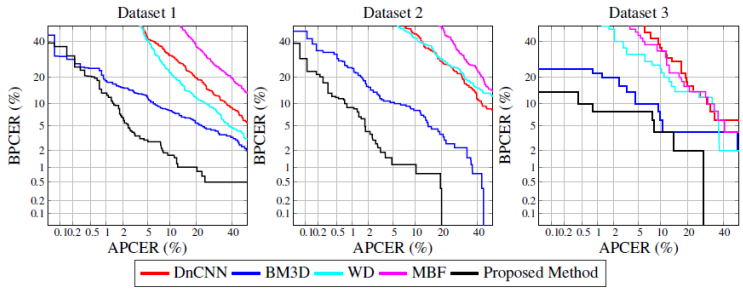


Figure 9.6: DET curves depicting MAD performance of the individual image denoising methods together with proposed method on different datasets

9.5 Face Morphing Datasets

The proposed approach is validated empirically using three different morphed face datasets employing different approaches for morphing and different composition representing the wide possible variation in morphing process as detailed below. The datasets are further used to benchmark the detection performance with State-Of-The-Art (SOTA) Morphing Attack Detection (MAD) methods.

9.5.1 Dataset-1

This database compromises 179 unique subjects that include both male and female participants from Asian and Caucasian ethnicity. This dataset is constructed using a public dataset (a subset of the FRGC face database) and a private face dataset. The whole database is divided into two partitions where the training set includes 89 disjoint and unique data subjects with multiple samples. The rest of the disjoint subjects are used in the testing set comprising 90 unique data subjects. Facial images are morphed using an open-source tool mentioned in [13]. Ultimately, the training set is composed of 709 bona fide and 1255 morphed images and the testing set is composed of 918 bona fide and 1354 morphed images. Figure 9.5 (a) shows example images from Dataset-1.

9.5.2 Dataset-2

This morphing database is a derivative of the publicly available FRGC database that compromises of 568 subjects. The entire database is divided into two partitions that include a training set of 300 unique data subjects resulting in 300 bona fide and 3041 morphed images. The testing set consists of 268 unique data subjects resulting in 268 bona fide and 2739 morphed images. Contrary to Dataset-1, the morphing process used for this dataset is based on the automatic facial landmark and triangulation, as mentioned in [103]. It has to be noted that the face morphing is performed only on the inner part of the face excluding the silhouette of the face (i.e, hair and ear region). Examples from Dataset-2 can be seen in Figure 9.5 (b).

9.5.3 Dataset-3

This database is a derivative of the publicly available PutDB database [34] that compromises 100 subjects. The entire database is divided into two different partitions consisting of 50 training and 50 testing unique data subjects. Morphing is performed based on the automatic facial landmark and triangulation as described in [103], that results in 50 bona fide and 254 morphed samples in the training set and 50 bona fide and 244 morph samples in the testing set. Similar to Dataset-2, only the inner part of the face is morphed. Figure 9.5 (c) shows example images from Dataset-3.

All three datasets are developed by following the morph data preparation steps, as discussed in [13, 125]. Since all three datasets are constructed using source face images from three different face datasets, this provides an opportunity to evaluate the generalizability of the proposed method together with the SOTA methods.

9.6 Experiments and Results

In this section, we present the quantitative results of the proposed method together with 14 different SOTA techniques for morphed face detection. Experimental results are presented using the ISO30107-3 [21] metrics such as Bona fide Presentation Classification Error Rate (BPCER(%)) and Attack Presentation Classification Error Rate (APCER (%)) along with Detection-Equal Error Rate (D-EER(%)). BPCER defines the proportion of bona fide presentations incorrectly classified as morphing attack images and APCER defines attack images incorrectly classified as bona fide images [21].

In this work, we have evaluated six deep learning-based SOTA, seven non-deep learning based techniques and one hybrid method that use both deep and handcrafted features. In case of the deep learning techniques, we have used the pre-trained network and computed the corresponding features that are further classified using a linear Support Vector Machines (SVM). To this extent, we have considered pre-trained CNN such as AlexNet [85, 90, 64], GoogleNet [90], Inception V3 [90], ResNet101 [85, 90, 64], VGG16 [85, 90, 83] and VGG19 [85, 90, 83]. The deep-learning techniques are used only as the feature extraction techniques owing to the availability of the small datasets. In case of non-deep learning techniques, texture-based techniques such as LBP [54], LPQ [54], BSIF [98], Steerable Pyramids [90] together with image distortion based features such as Image gradients [54], hybrid method [84], HoG [98] and PRNU [104] with linear SVM (except for PRNU) to compute the detection performance. To effectively evaluate the performance of the Morphing Attack Detection (MAD) schemes, we perform three different experiments such as **Experiment-1:-** designed to evaluate the performance of the MAD schemes when training and testing is carried out on the same dataset. **Experiment-2:-** designed to evaluate the MAD schemes on the merged dataset in which all three datasets are merged to one single dataset. This experiment provides an insight into the MAD performance when the dataset is increased with respect to the number of bona fide and morphed samples. **Experiment-3:-** designed to perform the cross-dataset comparison in which one of the datasets is used for training and another dataset is used for testing. This experiment will provide insights on MAD techniques that are capable to operate on unknown data.

Table 9.2 indicates the performance of the proposed method and individual image denoising methods used to build the proposed method. Figure 9.6 shows the DET

curves for all three different morphed face datasets. To have a fair comparison, we have used the same feature extraction and comparison schemes on individual denoising schemes. Based on the results, it can be noted that the proposed method has indicated the best detection performance on all three datasets demonstrating a good robustness. The superior performance of the proposed method can be attributed to (a) the aggregated denoising and fusion scheme based on the best sub-band selection (b) the robustness of MS-CAN in obtaining the noise of images trained using natural images.

Table 9.3 indicates the performance of the proposed method together with 14 different state-of-the-art methods for Experiment-1 on all three datasets. Based on the obtained results, it can be noted that (a) the use of deep-features indicate a better performance on all three datasets when compared to non-deep feature based techniques. (2) Among the deep features, the AlexNet and ResNet101 have indicated an improved performance over other deep features. (3) Among the non-deep features, HoG-SVM has indicated the best performance. (4) The proposed method shows overall the best performance when compared to 14 different SOTA techniques on all three different datasets.

Table 9.4 presents the quantitative results of the proposed and existing methods for Experiment-2. Based on the obtained results, the deep features indicate better performance over non-deep techniques. Further, the proposed method has indicated the best performance with D-EER = 4.96% with BPCER = 5.01% @APCER = 5% and BPCER = 3.05% @APCER = 10%. These obtained results further justify the robustness of the proposed method to the increased number of samples with different image characteristics.

Table 9.5 indicates the quantitative performance of the proposed method for Experiment - 3 (cross-dataset evaluation). For simplicity, we have presented the results only for the top four best performing MAD techniques based on Experiment-1 and Experiment-2. Since we have three different datasets, we get six different cases in which one dataset is enrolled and the remaining two datasets are probed. Based on the obtained results, the proposed method shows superior performance when compared with the SOTA methods in all six cases.

Thus, based on the extensive experiments carried out on three different datasets with three different performance evaluation experiments, we can conclude a superior performance over 14 different state-of-the-art methods. The evaluation results demonstrate the robustness of the proposed method, which is attributed to the proposed deep MS-CAN architecture. Further, realizing the proposed method using MS-CAN not only improved the robustness but also significantly improved computational cost by a factor of 4, as four denoising operations learnt as a single

coherent operation.

9.7 Conclusion

We have presented a novel method to detect face morphing attacks in a reliable manner. The proposed method is based on quantifying the residual noise resulting from the effect of the morphing process. The morphing noise is quantified using a aggregation of multiple denoising methods approximated using a deep Multi-Scale Context Aggregation Network (MS-CAN). We then process the residual noise from deep MS-CAN to extract deep features computed using a pretrained AlexNet. The final decision is computed using the Probabilistic Collaborative Representation Classifier (P-CRC) learnt using the extracted features. Extensive experiments are carried out using three different morphed face datasets with three different performance evaluation protocols. The performance of the proposed method is benchmarkd with the 14 different existing methods. The results have shown that the proposed method significantly outperforms existing methods on all three datasets for three different performance evaluation protocols.

Table 9.3: Quantitative performance of the MAD algorithms on Experiment-1 (individual dataset)

Algorithms	Dataset-1			Dataset-2			Dataset-3		
	D-EER(%)	BPCER@ APCER		D-EER(%)	BPCER@ APCER		D-EER(%)	BPCER@ APCER	
		=5%	=10%		=5%	=10%		=5%	=10%
AlexNet-SVM [85, 90, 64]	5.50	3.5	2.33	7.08	8.95	4.85	11	22	12
GoogleNet-SVM [90]	9.63	13.66	8.83	11.95	22.38	14.55	42.23	100	77.23
InceptionV3-SVM [90]	11.66	18.83	12.33	8.21	11.94	8.20	11.94	26	16
ResNet-SVM [85, 90, 64]	5.51	6.16	4	6.48	6.10	4.74	13.76	32	22
VGG16-SVM [85, 90, 83]	13.31	25	16.83	14.50	28.35	18.28	21.86	100	36
VGG19-SVM [85, 90, 64]	12.49	22.66	15	12.32	22.38	14.17	24.50	52	40
BSIF-SVM [54] [98]	26.70	53	42	12.67	25.74	14.55	20.45	44	32
Steerable pyramid-SVM [90]	26.19	65.50	50	37.97	82.08	71.64	34.00	82	70
HOG-SVM [98]	10.37	19.83	10.50	12.30	23.50	14.92	11.91	26	10
Image Gradient-SVM [54]	17.34	38	26.50	25.24	51.86	39.92	31.98	72	60
LBP-SVM [54, 103, 85, 98]	18.67	39.16	28.16	9.31	14.55	8.20	22.06	62	38
PRNU [104]	26.51	43	55.66	39.89	96.26	92.91	35.62	94	94
LPQ-SVM [54]	17.30	43.66	28.66	13.43	26.11	16.41	20.24	56	38
Deep Residual Noise [4]	3.83	3	1.5	4.85	4.85	3.35	9.71	14	8
Proposed Method	**3.24**	**3**	**1.67**	**2.63**	**1.11**	**1.11**	**7.89**	**8**	**4**

Table 9.4: Quantitative performance of the MAD algorithms on Experiment-2 (merged dataset)

Algorithms	D-EER(%)	BPCER@ APCER	
		=5%	=10%
AlexNet-SVM [85, 90, 64]	9.70	17.32	9.36
GoogleNet-SVM [90]	10.87	21.35	11.98
InceptionV3-SVM [90]	8.69	14.59	7.51
ResNet-SVM [85, 90, 64]	7.77	9.04	4.68
VGG16-SVM [85, 90, 83]	12.83	25.49	15.03
VGG19-SVM [85, 90, 64]	12.19	24.50	15.03
BSIF-SVM [54, 98]	15.58	33.98	23.09
Steerable Pyramid-SVM [90]	36.78	77.88	68.08
HOG-SVM [98]	11.32	20.69	12.52
Image Gradient-SVM [54]	38.41	79.84	68.84
LBP-SVM [54, 103, 85, 98]	36.58	73.42	63.98
PRNU [104]	36.88	96.84	94.11
LPQ-SVM [54]	15.03	30.28	19.82
Deep Residual Noise [4]	5.35	6.31	2.50
Proposed Method	**4.96**	**5.01**	**3.05**

Table 9.5: Quantitative performance of the MAD algorithms on Experiment-3 (cross Dataset) - D1-Dataset 1, D2-Dataset 2, D3- Dataset 3

Train	Test	Algorithms	D-EER(%)	BPCER @ APCER	
				=5%	=10%
D1	D2	AlexNet-SVM [85, 90, 64]	50	100	100
		Deep Residual Noise [4]	**7.12**	**12.31**	**5.22**
		HoG-SVM [98]	17.97	38.43	28.35
		Proposed method	10.44	16.04	10.44
D1	D3	AlexNet-SVM [85, 90, 64]	19.63	32	24
		Deep Residual Noise [4]	13.76	32	16
		HoG-SVM [98]	20.24	50	30
		Proposed method	**11.94**	**28**	**14**
D2	D1	AlexNet-SVM [85, 90, 64]	8.14	11.66	7.33
		Deep Residual Noise [4]	6.49	8.50	4.16
		HoG-SVM [98]	6.81	9	4.83
		Proposed method	**4.66**	**4.66**	**2.88**
D2	D3	AlexNet-SVM [85, 90, 64]	19.83	38	34
		Deep Residual Noise [4]	13.76	30	22
		HoG-SVM [98]	12.35	34	20
		Proposed method	**11.94**	**18**	**14**
D3	D1	AlexNet-SVM [85, 90, 64]	50	100	100
		Deep Residual Noise [4]	14.40	36.16	19.50
		HoG-SVM [98]	14.52	32	19.16
		Proposed method	**8.62**	**10.83**	**7.67**
D3	D2	AlexNet-SVM [85, 90, 64]	50	100	100
		Deep Residual Noise [4]	15.31	33.95	23.50
		HoG-SVM [98]	24.28	58.20	42.53
		Proposed method	**10.03**	**17.16**	**10.07**

Chapter 10

Article 6: Single Image Face Morphing Attack Detection Using Ensemble of Features

Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, Christoph Busch. *Single Image Face Morphing Attack Detection Using Ensemble of Features* In 23rd International Conference on Information Fusion (FUSION), pages 1–6, IEEE, 2020.

10.1 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.

10.2 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 10.1) can be used to verify two or more identities with one single morphed reference image as demonstrated in earlier works [13, 22, 135]. Not only do such morphed images bypass the FRS, but also fool the human observers including trained border guards [135] posing a severe threat to border control processes. This situation makes morphing attacks a relevant risk that constitutes a significant challenge [22].

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 [91]: (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.

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 [82]. The majority of MAD algorithms are also limited to operate on the digital data format of morphing images[54, 101, 81, 95, 79]. The popular methods in this direction includes texture based schemes like: Local Binary Patterns (LBP) [54], Binarized Statistical image features (BSIF) [54], Frequency features [99], Image degradation measure features using StirTrace algorithm [81], JPEG compression artefacts [95], PRNU [82], Benford's features [101], Specular reflection [79].

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

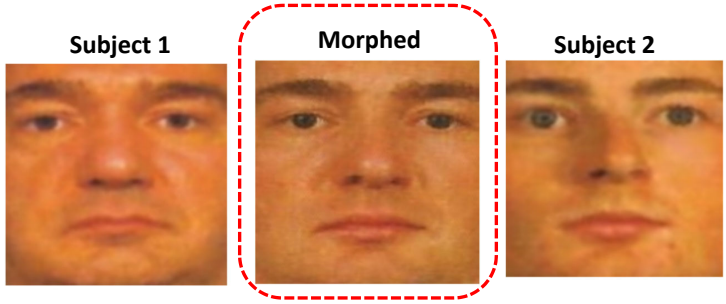


Figure 10.1: Example of the face morphing

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) [98], Binarized Statistical Image Features (BSIF) [54], color textures, deep learning approaches [64, 83], high frequency texture features [90, 91].

A set of other works in this direction have also explored multiple feature extraction techniques to detect morphing attacks. In [83], 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 [13] a hybrid feature fusion was presented for reliable face morphing attack detection. A framework proposed in [100], 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 [193], 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 [91] , 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 [86] where individual morph attack detectors were combined using Dempster-Shafer theory to improve the reliability of face morphing attacks

[86]. 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[84]. 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 [21] 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 10.3, discussion on experimental results are presented in Section 10.4 and Section 10.5 draws the conclusion.

10.3 Proposed Face Morphing Attack Detection Technique

Figure 10.2 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 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

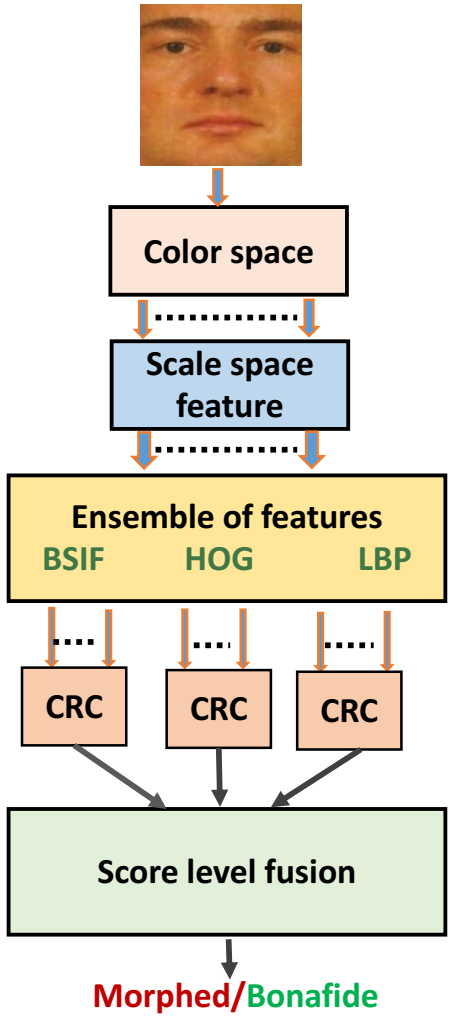


Figure 10.2: Block diagram of the proposed method

Figure 10.3: 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.

these two color spaces as earlier work[90] 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 [91]. 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 [192]. The P-CRC used in this work utilizes the Regularized Least Square Regression (LSR) on the learned feature vectors versus the probe feature vectors [192] formulated as:

$$\widehat{F} = argmin_{\alpha} \left\| Tr_F - \mathscr{D}\alpha \right\|_2^2 + \lambda \left\| \alpha \right\|_2^2 \tag{10.1}$$

where the Tr_F is the feature vector of the probe image, $\mathscr{D}$ 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.

Table 10.1: 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**

10.4 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 [21] 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 pro-

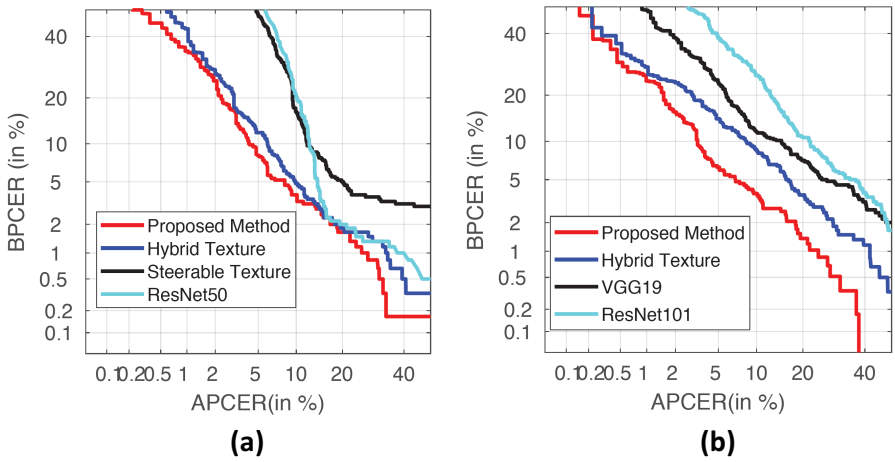


Figure 10.4: DET curves on (a) Dataset-1 (b) Dataset-2

experiments are carried out on the two different datasets such as *Dataset-1:* This is the private database [90] with 1309 bona fide images and 2608 morphed face images. This dataset is collected using RICOH office printer following the procedure mentioned in [90]. *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 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 10.3 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 [54] 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 10.1 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 [90] [64] [83] 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 [90] and Hybrid textures [91] 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 10.4 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.

10.5 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.

Chapter 11

Article 7: Face Morphing Attack Generation and Detection: A Comprehensive Survey

Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, and Christoph Busch. *Face morphing attack generation and detection: A comprehensive survey*. IEEE Transactions on Technology and Society, 2(3):128–145, March 2021

11.1 Abstract

Face recognition has been successfully deployed in real-time applications, including secure applications such as border control. The vulnerability of face recognition systems (FRSs) to various kinds of attacks (both direct and indirect attacks) and face morphing attacks has received great interest from the biometric community. The goal of a morphing attack is to subvert an FRS at an automatic border control (ABC) gate by presenting an electronic machine-readable travel document (eMRTD) or e-passport that is obtained based on a morphed face image. Since the application process for an e-passport in the majority of countries requires a passport photograph to be presented by the applicant, a malicious actor and an accomplice can generate a morphed face image to obtain the e-passport. An e-passport with a morphed face image can be used by both the malicious actor and the accomplice to cross a border, as the morphed face image can be verified against both of them. This can result in a significant threat, as a malicious actor can cross the border without revealing the trace of his/her criminal background, while the details of the accomplice are recorded in the log of the access control system. This survey aims to present a systematic overview of the progress made in the area of

face morphing in terms of both morph generation and morph detection. In this article, we describe and illustrate various aspects of face morphing attacks, including different techniques for generating morphed face images and state-of-the-art morph attack detection (MAD) algorithms based on a stringent taxonomy as well as the availability of public databases, which allow us to benchmark new MAD algorithms in a reproducible manner. The outcomes of competitions and benchmarking, vulnerability assessments, and performance evaluation metrics are also provided in a comprehensive manner. Furthermore, we discuss the open challenges and potential future areas that need to be addressed in the evolving field of biometrics.

11.2 Introduction

Biometrics is a technique for recognizing an individual based on unique biological (e.g., face, fingerprint, iris) or behavioural (e.g., gait, keystroke style) characteristics [156] [194]. With the drastic improvement in deep learning techniques, biometric-based person identification and verification has emerged as a popular technique that can be widely used for many secure access control applications. The ease of capture and the suitability of face biometric characteristics have further driven face recognition as a popular biometric modality in such applications. Face recognition systems (FRSs) are widely deployed for various applications, especially in secure access control for person identification and verification purposes. Among several other applications, such as healthcare, law enforcement, and e-commerce (banking), one of the most relevant applications is the border control process, where the facial characteristics of a traveler are compared with a reference in a passport or visa database to verify the claimed identity.

Although an FRS effectively distinguishes an individual from other subjects, the FRS's risk of being attacked to mislead or conceal an actual identity is a major concern. As with all applications, the FRS is prone to various attacks, such as presentation attacks, which have the goal of subverting the FRS by presenting an artifact [133], where various types of attacks, such as electronic display attacks, print attacks, replay attacks and 3-D face mask attacks, can be used [133] [195] [196] [197] [198] [199] [200] [201] [202] [203]. In addition to these attacks, the morphing attack has emerged in the recent past as a severe threat to the enrolment process that successfully undermines FRS capabilities [22]. Face morphing is defined as "a seamless transition of a facial image transforming a facial image into another" [204] in the context of biometrics; two or more facial images can be combined to resemble the contributing subjects. Morphing attacks raise a major concern, as the morphed image represents the facial characteristics of both individuals contributing to the morphing process (for instance, an accomplice and a

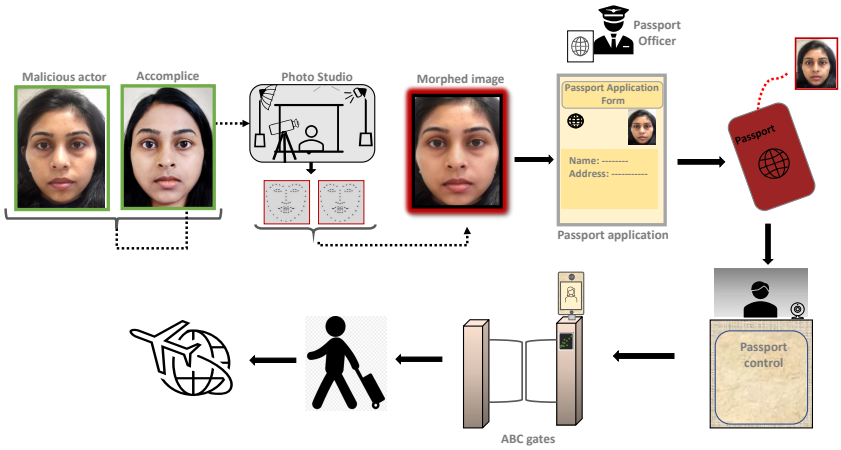


Figure 11.1: Example scenario illustrating the vulnerability of FRSs to morphed images in border control.

malicious actor). Ultimately, the resulting morphed facial image can successfully be verified with probe images from both contributing subjects, making it practically usable for various malicious actions. Therefore, this attack breaks the rule of single ownership; for instance, an identification document such as a passport or electronic machine-readable travel document (eMRTD) [205] has a unique link to the data subject for whom the document was issued. The facial image stored in the eMRTD or passport is compared with the person claiming identity document ownership while crossing the border. If the enrolled facial image is determined as a match with the live image, the data subject can cross the border. Thus, an individual with malicious intent can exploit the face morphing attack and obtain illegal access. Hence, a malicious person can easily cross a border using an eMRTD or passport if he/she has contributed to the morphed image that was used in the passport application process.

Fig 11.1 illustrates an example scenario in border control where the facial image of a malicious person is morphed with that of a look-like accomplice. As several morphing software programs are freely available, even a nontechnical person can perform morphing with ease. The accomplice can submit the generated morphed image for passport enrolment at the passport issuance office. As the morphed image's facial features resemble those of the applicant's face, the passport officer approves the application. Ultimately, a malicious person can successfully use the genuine passport, allowing him/her to achieve all foreseeable purposes (e.g., crossing a border).

In most countries, the applicant submits a printed facial image to the passport

office, allowing the possibility of providing a morphed image after printing and scanning. However, some countries, such as New Zealand, Estonia and Ireland, also accept a digital facial image for passport renewal [23]. Hence, an applicant can submit a digital facial image to the Web portal. This practice raises a further severe concern, as there is no trusted supervision while uploading the digital facial image, and this opens the possibility of uploading a morphed image. The B1/B2 visa application for the United States also allows the applicant to upload a digital facial image through the Web portal [206]. An applicant can use this opportunity to upload a morphed image with the intent to perform illegal activity.

All such vulnerabilities of FRSs have made morphing research crucial in recent years to avoid probable security lapses. Thus, several research projects have been funded by the European Union and national research councils (e.g., SWAN [207], ANANAS [208], SOTAMD [129] and iMARS [209]) to focus extensively on developing morph attack detection (MAD) algorithms. Motivated by the momentum of the problem of morphing and its criticality, a dedicated conference has been initiated by Frontex, the European Border, and Coast Guard Agency [210], where a MAD interest group gathered to discuss the challenges and advancements of MAD techniques [211]. Furthermore, the U.S. National Institute of Standards and Technology (NIST) is, in parallel, conducting testing of MAD technology within the framework of the Face Recognition Vendor Test (FRVT) under Part 4: MORPH - Performance of Automated Face Morph Detection [212]. Both industrial and academic institutions are invited to submit their MAD algorithms to benchmark the accuracy [212]. Similarly, the University of Bologna, as part of the SOTAMD project [213], introduced a parallel face morphing evaluation platform to benchmark the performance of the MAD techniques on a sequestered dataset.

The rest of this article is organised as follows: Section 11.3 presents a brief introduction to face morphing attacks, and Section 11.4 discusses the face morph generation techniques. Section 11.5 describes face morphing datasets, including private and public datasets, Section 11.6 discusses human perception capabilities in detecting morphed face images, Section 11.7 presents various automatic morphing attack detection techniques, Section 11.8 presents the performance metrics that are widely used to benchmark the performance of MAD methods as well as the vulnerability of generated morphed images, Section 11.9 discusses the public evaluation and benchmarking of MAD, Section 11.10 discusses open challenges and potential future work and Section 11.11 gives the conclusion.

11.3 Face Morphing Attack

The morphing process can be defined as a special effect that transforms one image into another image. Fig 11.2 illustrates the facial morphing process, where two

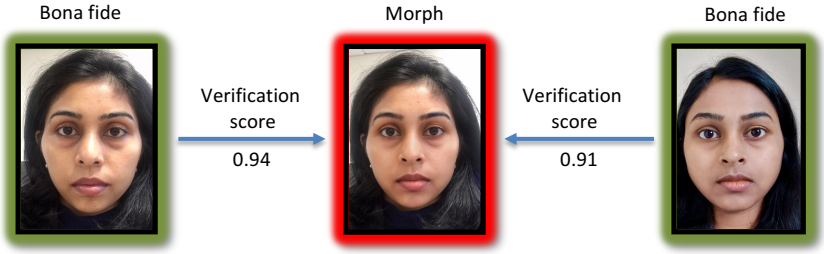


Figure 11.2: Impact of face morphing on an FRS. As noted in the figure, the morphed image can be verified equally against both contributing subjects with a high similarity score from the FRS (1 being high similarity).

facial images are combined to generate a single morphed image. Morphing can be achieved easily by using one of the numerous and freely available tools, such as MorphThing [214], 3Dthis Face Morph [67], Face Swap Online [215], Abrosoft FantaMorph [143], FaceMorpher [216], and MagicMorph [217]. The morphed image possesses near-identical features to those of both subjects contributing to generating the morph when subject preselection is applied (e.g., look-alike mode) [13].

Furthermore, when processed with care, the morphed image does not possess many visible artifacts, and thus, a human observer may fail to detect image manipulation based on morphing. In practice, this leads to a situation where a passport officer may not be able to detect the morphing attack despite being an expert in facial comparison [29, 134]. This makes it reasonable for a criminal with malicious intent to be able to use a passport enrolled with a morphed image and cross a border without challenge. Figure 11.1 illustrates the vulnerability of FRSs when attacked with morphed images in a border control scenario.

11.4 Face Morph Attack Generation

Face morphing has been widely used for more than a decade, especially in the video animation industry [218], but the attack potential against FRSs has been noted recently [22]. Morphs can be generated using various techniques, from simple image warping to recent generative adversarial networks (GANs) [48, 51, 52, 219, 220, 56, 53, 49, 2]. The most widely used morph generation methods are based on the landmark-based technique [221, 54, 14, 80], where morphing is carried out by combining the images with respect to corresponding landmarks. Recent works eliminate the constraints of landmarks by simply relying on deep network architectures [15, 2]. Figure 11.3 shows a taxonomy of face morphing generation methods that indicates the broad classification of the available techniques as (a)

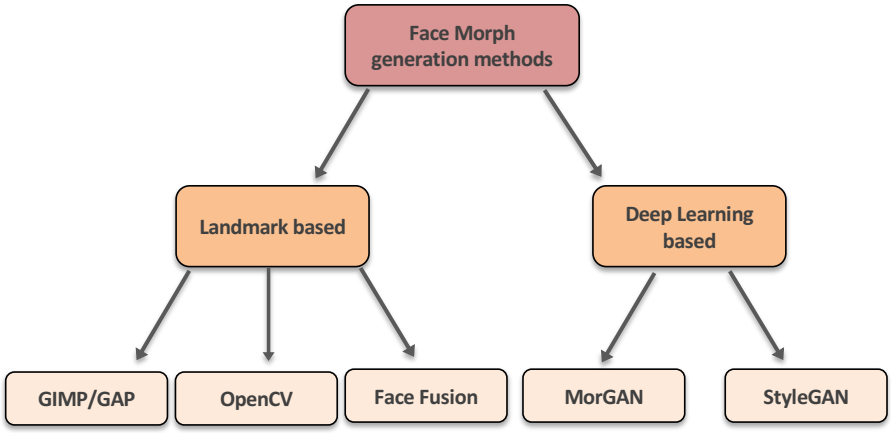


Figure 11.3: Taxonomy of face morph generation techniques

Table 11.1: Face Morphing Generation Methods: Advantages and Limitations

Face Morph Generation Method	Advantages	Limitations
Facial Landmark-based	- Availability of open-source tools. - Generates high quality morphing images. - Successfully deceives the COTS FRS. - Easy and seamless generation of morphed images by an automatic process.	- Requires manual intervention to ensure high-quality face morphing generation. - Needs post-processing to reduce ghosting effects and double edges. - Data subject selection is crucial to deceive the COTS FRS.
Deep Learning-based	- No need for manual intervention. - Seamless generation with acceptable image quality. - Does not show double edges in the generated images. - Reasonably successful in deceiving the COTS FRS. - Several open-source tools.	- Requires a complex learning procedure. - Does not always generate high-quality morphed images. - Highly prone to geometric distortions. - Requires careful pre-selection of data subjects based on age, gender and ethnicity.

landmark-based and (b) deep learning-based approaches.

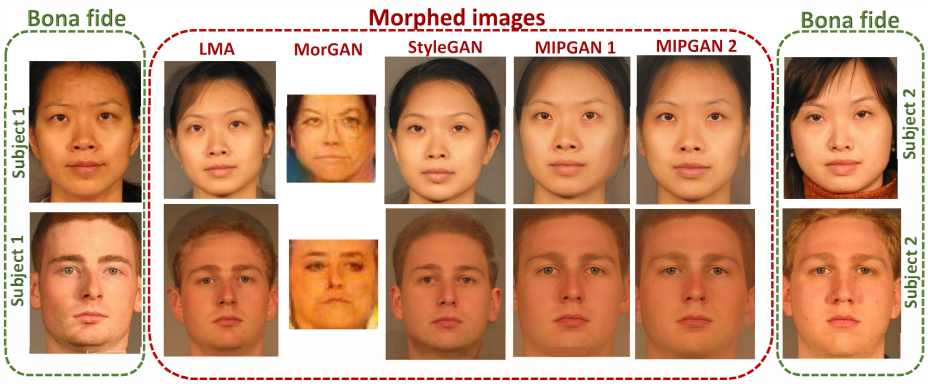


Figure 11.4: Illustration of face morph images generated using different methods

11.4.1 Landmark-based Morph Generation

Landmark-based morph generation works by obtaining landmark points on facial regions, e.g., the nose, eye, and mouth. These landmark points obtained from both faces are warped by moving the pixels to different, more averaged positions. Different procedures for warping exist, including free-form deformation (FFD) [222, 223], deformation by moving least squares [62], deformation based on mass-spring models [63], and Bayesian framework-based morphing [140]. Ruprecht et al. [59] proposed performing warping by moving the pixel points of both contributing subjects to the nearest landmark point. Delaunay triangulation was later proposed, where the pixels of both contributing facial images are distorted and moved to different directions to generate triangles [95, 141, 81, 64, 98, 88]. Images that are to be morphed are blended by considering the blending factors or the morphing factor. Face morphing applications employ a morphing factor of 0.5 to generate high-quality and useful morphs that can resemble both contributing subjects equally, to which the COTS FRS is vulnerable [54, 13, 83]. As the morphing process translates landmarks and the associated texture, there may be some misaligned pixels that contribute to noise generating artifacts and ghost-like images and making the images unrealistic in appearance (i.e., easy for a human observer to detect). Hence, certain post-processing steps, such as image smoothing, image sharpening, edge correction, histogram equalization, manual retouching, and image enhancement improve the brightness and contrast and can reduce or minimise the artifacts generated during the morphing process [64, 65, 66].

Face morph generation using open-source resources such as GIMP/GAP and OpenCV also relies upon landmarks. While open-source software based on GIMP/GAP and

OpenCV can generate morphs, significant effort must be made to post-process the generated images to eliminate artifacts. Several commercial solutions, such as Face Fusion [142] and FantaMorph [143], can also be used to generate large-scale morphed images with reasonable post-processing effort. The reader is further referred to Scherhag et al. [142], where all the publicly available morphing tools (both open-source and commercial) are listed.

11.4.2 Deep Learning-based Morph Generation

Recent improvements in deep learning-based techniques have given rise to morph generation approaches based on generative adversarial networks (GANs) [15] [1]. In general, GAN-based methods synthesize morphed images that are generated by sampling two facial images in the latent space of the deep learning network. The MorGAN architecture for morph generation basically employs a generator that consists of encoders, decoders and a discriminator. The generator is trained to generate images with dimensions of 64×64 pixels. Another recent approach based on StyleGAN architecture [137, 1] has improved the morph generation process both by increasing the spatial size to 1024×1024 and by increasing face quality. The pretrained StyleGAN achieves this by embedding the images in the intermediate latent space. The use of identity priors to enable high-quality morphed face generation was also proposed in [2] and illustrates the increased threat to FRSs by GAN-based morphs. Fig 11.4 provides sample facial morphs generated using the landmark-based technique and MorGAN- and StyleGAN-based methods. It can be noted from Fig 11.4 that deep learning-based approaches, especially with MIPGAN-I and MIPGAN-II, indicate a superior quality of the morphed face image compared to that of landmark-based morphed face generation.

11.5 Databases for Morphing Attack Detection

Given various kinds of attack generation mechanisms and the relevant attack potential determination metrics, many datasets have been generated, ranging from public to sequestered datasets with various attack strengths. This section summarizes the different face morph databases that are used in the existing works. A summary of the different datasets is provided in Table 11.2 from existing works that are typically used to benchmark both the vulnerability of FRSs and the performance of MAD techniques.

The first face morph database was introduced by Ferrara et al. [22], in which the authors employed landmark-based face morph generation using GIMP/GAP tools. This dataset has a small set of digital images consisting of only 14 morphed images generated from 8 bona fide subjects, including both male and female participants. The morphed images in this database are only in digital format and the database is

not available publicly. This dataset was extended by Ferrara et al. [135] using the landmarks and GIMP/GAP tools. The extended dataset consists of approximately 80 morphed face images, with 10 male and 9 female participants. The database is in digital form and is not publicly available.

The first large database with different ethnicities (Caucasian, Asian, European, American, Latin American, and Middle Eastern) was introduced by Raghavendra et al. [54] and employs facial landmarks and the GIMP/GAP morph generation technique using the GNU image manipulation tool. This database consists of 450 morphed face images generated using 110 subjects of different ethnic backgrounds. This database contains only digital images and has not been made public.

Table 11.2: Public and Private Face Morph Image Databases

Reference	Morph Generation Type	Morph Generation Method	Digital/Print-scan	Bona fide & Morph	Public/Private
Ferrara et al [22]	Landmark-based	GIMP/GAP	Digital	Morph: 14	Private
Ferrara et al [135]	Landmark-based	GIMP/GAP	Digital	Morph: 80	Private
Raghavendra et al [54]	Landmark-based	GIMP/GAP	Digital	Morph: 450	Private
Makrushin et al [95]	Landmark-based	Automatic generation (dlib landmark)	Digital	Complete morph: 1326, Splicing morph: 2614	Private
Scherhag et al [98]	Landmark-based	GIMP/GAP	Digital and Print-scan	Bona fide: 462 Morph: 231	Private
Raghavendra et al [13]	Landmark-based	GIMP/GAP	Digital and Print-scan	Bona fide: 1000 Morph: 1423+1423	Private
Raghavendra et al [83]	Landmark-based	GIMP/GAP	Digital and Print-scan	Morph: 362	Private
Gomez-Barrero et al [224]	-	-	Digital	Morph: 840	Private
Dunstone [225]	-	-	Digital	Morph: 1082	Public
Ferrara et al [119]	Landmark-based	Sqirlz morph	Digital and Print-scan	Morph: 100	Private
Damer et al [15]	GAN-based	GAN	Digital	Morph: 1000	Private
Raghavendra et al [90]	Landmark-based	GIMP/GAP	Digital and Print-scan	Bona fide: 1272 Morph: 2518	Private
Scherhag et al [80]	Landmark-based	OpenCV, FaceFusion, Face Morpher	Digital and Print-scan	Bona fide: 984+984+529 Morph: 964+964+529	Private
Ferrara et al [14]	Landmark-based	Triangulation	Digital	Morph: 560	Private
Scherhag et al [115]	Landmark-based	OpenCV, FaceFusion, Face Morpher, UBO morpher	Digital and Print-scan	Bona fide: 791+3298 Morph: 791+3246	Private
Singh et al [116]	Landmark-based	OpenCV	Digital and Print-scan	Morph: 588	Private
Venkatesh et al [3]	Landmark-based	UBO morpher	Digital	Morph: 10538+3767	Private
Venkatesh et al [1]	GAN-based	StyleGAN	Digital	Bona fide: 1270 Morph: 2500	Private
Raja et al [38]	Landmark-based	UBO morpher	Digital and Print-scan	Bona fide: 300+1096 Morph: 2045+3073	Sequestered
NIST-FRVT-MORPH et al [128]	Landmark-based	Automatic generation	Digital and Print-scan	Low-quality morph: 1183 Automated morph: 39113 High-quality morph: 492	Sequestered

Makrushin et al. [95] employed automatic morph generation tools to generate high-quality morph images. They employed a triangulation method based on 68 facial landmarks extracted using the dlib library [226]. Two different morph generation techniques, namely, complete morph (consisting of the facial geometry of both facial images) and splicing morph (the pixels representing the face are clipped out from the input faces), were used. A splicing morph is generated to address the

pixel discontinuity caused by warping two images in complete morphs. This database consists of approximately 1326 complete morphs and 2614 splicing morphs generated from 52 data subjects consisting of 17 females and 35 males. This database consists of face morph images in digital format only and has not been made public.

The first print-scan face morph database was presented by Scherhag et al. [98]. The authors employed the landmark-based GIMP/GAP technique for morph generation. This database consists of 231 morphed images generated from 462 bona fide images. This database is private and contains digital and print-scan (or re-digitized) images, for which HP Photosmart 5520 and Ricoh MPC 6003 SP printers were employed.

Raghavendra et al. [13] later introduced a new face morph dataset consisting of both digital and print-scan images. The face morphs were generated using an automatic tool, OpenCV, that is publicly available. This database generates morphed face images along with averaged face images and hence has a set of $1423 + 1423$ morphed face images. Along with the database, Raghavendra et al. [13] provided an evaluation protocol by defining independent sets for development, training and testing partitioning. The print-scan morphed face images were obtained by employing a Ricoh MPC 6003 SP printer. This database is private. This dataset was extended to 2518 morphed face images and 1273 bona fide images [90].

Gomez et al. [224] introduced a new face morph dataset that consists of 840 morphed face images generated from 210 subjects. This database is private and has only digital morphed face images.

Ferrara et al. [119], [85] introduced a face morph database based on the Sqirlz morphing technique. This dataset has 100 morphed images in both digital and print-scan forms. This database has not been made public for research purposes. Scherhag et al. [80] introduced a face morphing dataset that was generated using different morphing tools, such as OpenCV, FaceFusion and FaceMorpher. This is a private database that consists of both digital and print-scan samples of morphed images and is composed of $964 + 964 + 529$ morphed face images generated from subjects contained in the FRGCv2 and FERET databases. Another database by Scherhag et al. [115] employs landmark-based morph generation techniques that include OpenCV, FaceMorphed, FaceFusion and the UBO morphing method. This database consists of approximately 791+3246 morphed face images from the FERET and FRGCv2 databases. This private database consists of morphed face images in both digital and print-scan formats. Another database by Ferrara et al. [14] employs triangulation with the dlib landmark method of morph generation. This is a private database that consists of 560 digital morphed face images. The

only publicly available morphed face dataset was introduced by Biometix [225], which consists of 1082 morphed face images in digital form. However, information on the morphed image generation method involved is not available.

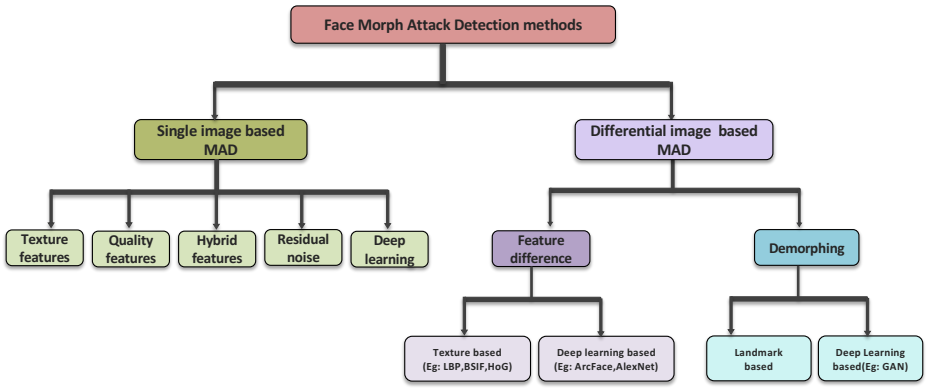


Figure 11.5: Taxonomy of MAD techniques

Singh et al [116] provided another database that employs the OpenCV-based morph generation technique to generate facial morphs. This was the first dataset introduced for probe images captured live from ABC gates with different lighting conditions, which is relevant for differential morphing attack detection. This database consists of both digital and print-scan enrolment images generated using an EPSON XP-860 printer and scanner. This dataset consists of 90 morphed face images and is not available to the public.

Damer et al. [15] introduced the first face morphing database consisting of deep learning-based morph images. The generated deep learning-based database is compared with landmark-based morphs. The authors employed 68 landmark points extracted from dlib for landmark-based morph generation and GAN architecture for deep learning-based morph generation. This database consists of 1000 morphed face images; however, the GAN-based morphs are of size 64×64, which does not meet ICAO standards. This database is private and has only digital morphed face data. Another database by Venkatesh et al. [1] employs deep learning-based morph generation. The authors employed the StyleGAN network to generate synthetic morph images by mapping the input images into the latent space. This database consists of 2500 morphed images generated using 1270 bona fide images. It has only digital morphed face images and is not publicly available.

Venkatesh et al. [3] introduced another database that consists of morphed face images under ageing as the first database of its kind. The authors employed the UBO morphing method from the University of Bologna that employs dlib and

68 landmark points for morph generation [14]. This database consists of 14305 (10538+3767) morphed face images aged by 2 to 5 years. This database has morphed face images in digital form and is not open to the public.

Raja et al. [38] presented the sequestered Bologna-SOTAMD face morphing dataset used in a recent public competition and benchmarking on the Bologna Online Evaluation Platform (BOEP), following the FVC-onGoing series [227]. The dataset comprises images from 150 data subjects collected in three different geographic locations with varying ethnicity, gender and age. Face morphing is carried out using six different techniques followed by automatic and manual post-processing to override the artifact results from face morphing. The dataset also includes printed and scanned versions with different printers, and the enrolment images follow the ICAO standards for passport images. The probe images are taken from various ABC gates and gate emulations. The database consists of 5748 morphed face images and 1396 bona fide face images.

11.5.1 Discussion

Although there exist several morphing datasets, the majority of them are private due to data protection regulations and licensing conditions. Even for publicly available face databases that are used to create face morphing datasets, the licensing conditions limit the redistribution of the generated morphed face datasets; therefore, most of the above datasets are not openly available. For the time being, the best way to compare new morphing detection methods with already published approaches is to submit the methods to the two ongoing benchmarks, either the SOTAMD benchmark at the university of Bologna, which was reported by Raja et al. [38], or the U.S. NIST-FRVT-MORPH benchmark, which was reported by Mei et al. [128]. Note that in both cases, a sequestered dataset is used.

11.6 Human Perception and Morphed Face Detection

The threat of morphing attacks is known for border crossing and ID management scenarios. Therefore, the success of a morphing attack depends on deceiving human observers, particularly ID experts and border guards. A practical scenario for a border crossing includes border guards, who compare the passport of a traveler containing a photograph (printed from a data page or digitally extracted from a chip) with the physical appearance of the traveler. Thus, the border guard considers the facial similarity of the traveler to the reference data in the passport to make his/her final decision. Several studies in the literature have indicated the effectiveness of morphed images in deceiving expert human observers [135, 134, 228, 28, 29, 95, 229, 230, 231]. Early investigations on human perception analysis of morphed images were reported by Jäger et al. [231], where

different experiments were performed to benchmark the ability of human observers to detect face morphing and its dependency on various parameters (i.e., different alpha/morphing factors). While this was an interesting study, the human observers in the experiments were students who were not trained to compare human faces. Furthermore, this work was based on only a single image and did not provide any reference images for the human observers. A similar analysis was provided by Kramer et al. [29], where a single image was provided before requesting a decision on morphing. Despite being different in terms of the underlying benchmarking mechanism, both works reported difficulty in detecting morphed face images for human observers.

Investigating the impact of morphing on FRSs and human observers simultaneously, Ferrara et al. [22] studied the detection ability of human observers and correlated it with automatic FRSs. Unlike the previous work, the human observers in the work of Ferrara et al. [22] included both trained border guards and nonspecialists who were asked to compare a morphed face image with a bona fide face image to make the decision. The analysis reported a challenge in detecting morphs even when the examiner, for instance, a border guard, was trained. Robertson et al. [134] further studied the morph detection ability of humans by comparing live face images to morphed face images with and without rudimentary training. The study reported improved performance in morph detection by human observers when provided with rudimentary training in detecting artifacts [28]. Similarly, Kramer et al. [29] investigated the role of face image quality (of the morphed image produced) on human perception and concluded that high-quality morphed images are more difficult for humans to detect.

A similar Web-based experiment simulating border control was presented by Makrushin et al. [229], who studied human perception analysis by both skilled and unskilled humans and further extended [230] to obtain more unbiased and realistic images. In both cases, skilled humans (who have knowledge of morphed face images) show the best performance in detecting a morphed face image. Summarizing the works on human perception analysis, it is noted that both skilled and unskilled human observers often fail to detect morphed face images. However, it is also noted that considerable training of human observers can improve morphed face detection [230, 28].

11.7 Face Morph Attack Detection Techniques

Noting the limitations of human observers, a number of automatic MAD approaches have been proposed in the recent past. In this section, we summarize the MAD techniques since the introduction of face morphing attacks on FRSs [22]. The available MAD techniques can be classified into two major types: (1) single image-

based MAD (S-MAD) and (2) differential image-based MAD (D-MAD). Figure 11.5 shows the taxonomy of approaches in both MAD categories reported to date.

11.7.1 Single Image-based MAD (S-MAD)

The goal of S-MAD is to effectively detect a face morphing attack based on a single image presented to the algorithm. Fig 11.6 illustrates a real-life example for S-MAD in a passport application scenario, where a facial image is submitted by the applicant for biometric enrolment in the passport application process. This submitted image is checked to potentially detect a morph of a suspect image. The passport application can be initiated by the applicant either physically or when submitting his/her facial image through a Web service [23, 24, 232, 139]. Thus, depending on the use case, the morphed image can be one of two types: (a) digital or (b) re-digitized (also commonly referred to as print-scan). S-MAD is challenging, as it is expected to be robust to image quality variations, different types of sensors (cameras), different types of morph generation tools and different types of print-scan processes (e.g., the equipment and parameter set chosen for the printing

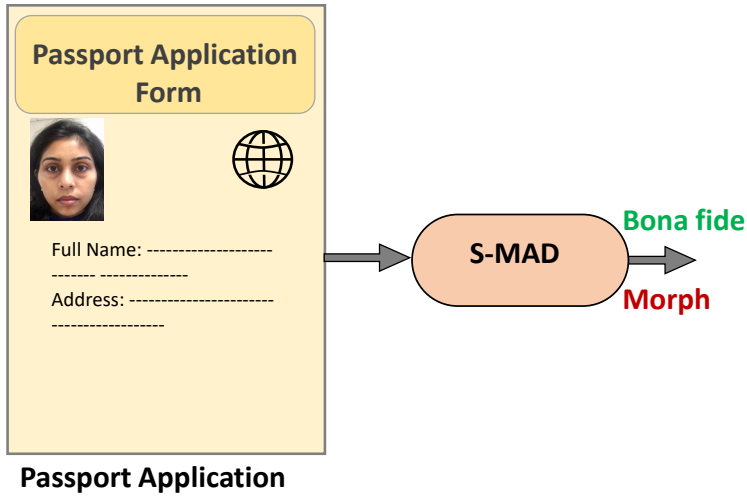


Figure 11.6: Example illustrating single image-based morph attack detection in a passport application scenario.

As shown in Figure 11.5, the existing S-MAD techniques can be further classified into five subtypes based on the features employed: (1) texture feature-based S-MAD, (2) quality-based S-MAD, (3) residual noise-based S-MAD, (4) deep learning-based S-MAD, and (5) hybrid approaches for S-MAD. Table 11.3 summarizes the existing S-MAD techniques. In the next section, we briefly discuss the existing S-MAD techniques for the convenience of the reader.

Table 11.3: State-of-the-art S-MAD

Reference	Detection Type	Approach	Algorithm	Database
Raghavendra et al. [54]	S-MAD	Texture-based approach	Local Binary Pattern (LBP)-SVM, Binary Statistical Image Features (BSIF)-SVM, Image Gradient (IG)-SVM	Digital
Makrushin et al. [95]	S-MAD	Quantised DCT co-efficients	Benford features	Digital
Neubert et al. [96]	S-MAD	Image degradation approach	Corner feature detector	Digital
Seibold et al. [64]	S-MAD	Deep learning-based approach	VGG19, Google Net, Alex Net	Digital
Raghavendra et al. [13]	S-MAD	Texture-based approach	LBP, LPQ, BSIF, colour textures	Print-scan
Asaad et al. [97]	S-MAD	Texture-based approach	Topological data analysis approach	Digital
Scherhag et al. [98]	S-MAD	Texture- and frequency-based approach	LBP, LPQ, BSIF, 2DFFT with SVM classifier	Digital Print-scan
Raghavendra et al. [83]	S-MAD	Deep CNN-based approach	Feature fusion of fully connected layers of VGG19 and Alex Net	Digital Print-scan
Kraetzer et al. [99]	S-MAD	Image life cycle model	Keypoints (SIFT, SURF, ORB, FAST, AGAST) and loss of edge operators (Canny and Sobel)	Digital
Hildebrandt et al. [81] [233]	S-MAD	StirTrace-based approach	Multi-compression anomaly detection	Digital
Debiasi et al. [104]	S-MAD	Image degradation	Photo Response Non-uniformity (PRNU)	Digital
Raghavendra et al. [90]	S-MAD	Steerable features	Luminance component extraction	Print-scan
Hildebrandt et al. [81]	S-MAD	StirTrace	StirTrace face morph forgery detection	Print-scan
Seibold et al. [79]	S-MAD	Image degradation	Specular reflection	Digital
Makrushin et al. [101]	S-MAD	Quantised DCT co-efficients	Benford features extracted from quantised DCT co-efficients	Digital
Neubert et al. [102]	S-MAD	Morph pipeline footprint detector	Benford features extracted from quantised DCT co-efficients	Digital
Spreeuwers et al. [103]	S-MAD	Texture-based approach	LBP-SVM, Down-up sampling	Digital
Scherhag et al. [88]	S-MAD D-MAD	Feature difference-based approach	Pre-processing and feature extraction using texture descriptors , keypoint extractors, gradient estimators and deep learning-based method	Digital
N Damer et al. [84]	S-MAD	Multi-detector fusion	LBPH, Transferable deep-CNN	Digital
Ferrara et al. [85]	S-MAD	Deep learning	AlexNet, VGG19, VGG-Face16, VGG-Face2	Print-scan
Scherhag et al. [87]	S-MAD	Multi-algorithm fusion	Texture descriptors (LBP, BSIF), Keypoint extractors (SIFT, SURF), gradient estimators (HoG), Deep neural network	Digital
Debiasi et al [104]	S-MAD	PRNU	PRNU DFT magnitude histogram and PRNU DFT energy	Digital
Seibold et al. [105]	S-MAD	Complex multi-class pre-training	VGG-19 network	Digital
Damer et al. [106]	S-MAD	Texture and deep learning based	Anomaly detection using LPQ and VGG features	Digital
Venkatesh et al. [4]	S-MAD	Colour denoising-based approach	Denoising Deep Convolutional Neural Network	Digital
Scherhag et al. [80]	S-MAD	PRNU	Spectral features and spatial features	Print-scan
Makrushin et al. [86]	S-MAD	Dempster-Shafer Theory	KeyPoints (SIFT, SUFT, FAST, ORB, AGAST, High Dim LBP, GoogleNet, VGG19	Digital
Raghavendra et al. [91]	S-MAD	Scale space approach	Colour scale space features	Print-scan
Neubert et al. [107]	S-MAD	Frequency and spatial domain feature space approach	Discrete Feature Transformation (DFT) , SURF, SIFT, ORB, FAST, AGAST, Canny edge, SobelX, SobelY)	Digital
Seibold et al. [89]	S-MAD	Style Transfer-based approach	LBP, BSIF, Image degradation, Deep neural network (VGG19)	Digital
Venkatesh et al. [5]	S-MAD	Colour denoising-based approach	Context Aggregation Network	Digital
Venkatesh et al. [6]	S-MAD	Ensemble-of-features-based approach	LBP, HoG, BSIF	Print-scan

Texture Feature-based S-MAD The first work on using texture features was presented by Raghavendra et al. [54]. Following the initial work, several approaches were proposed, as indicated in Table 11.3. The popular texture-based methods include local binary patterns (LBPs) [74], local phase quantization (LPQ) features [75] and binarized statistical image features (BSIFs) [76]. Furthermore, these texture features were extracted for different color channels [13] to obtain a robust detection performance. Variants of LBPs and BSIFs as well as histogram of oriented gradients (HOG) features, scale-invariant features (SIFT) [77] and speed-up robust features (SURF) [78, 98, 99, 86, 87] have also been widely explored in the reported works. The use of micro-texture-based methods has shown reasonable performance on both digital and print-scan types of S-MAD. While superior accuracy has been reported for digital S-MAD with texture-based features, the main limitation of these techniques is in their generalizability across different image qualities, imaging sensors and print-scan processes [38].

Quality-based S-MAD The quality-based techniques largely analyse image quality features by quantifying the image degradation to identify a given image as morphed or bona fide [81, 79, 80, 82]. Several features, such as double-compression artifacts, photograph response non-uniformity (PRNU), corner and edge distortions, reflection analysis and meta information in the images, are commonly used to detect distortion in a morphed image. Although these techniques have shown good performance on digital data, they have limited performance on print-scan data. However, the generalization ability of these techniques has yet to be studied for different print and scan versions in the current literature [80, 38].

Residual Noise-based S-MAD Residual noise-based methods are designed to analyze pixel discontinuities that may be greatly impacted by the morphing process. The basic idea of this approach is to extract noise patterns by subtracting the given image from the denoised version of the same image. The noise patterns obtained are further analyzed to detect morphing. The first work in this area was introduced in [4] based on CNN-based denoising on color channels. Furthermore, the residual noise is effectively captured using the deep CNN approach [5]. The use of residual noise has shown considerably good performance in terms of generalization capabilities across different digital datasets. However, these techniques have not been evaluated on print-scan face morphed datasets.

Deep Learning-based S-MAD The success of deep learning approaches for image classification tasks has motivated researchers to embrace deep convolutional neural networks (CNNs) for face MAD. All existing works are based on pre-trained networks and transfer learning. The first work in this direction was based on using pretrained networks such as AlexNet and VGG18, in which the features are fused

and classified to detect a morphing attack [83]. Following this, several deep CNN pre-trained networks, such as AlexNet, VGG19, VGG-Face16, GoogleNet, ResNet18, ResNet150, ResNet50, VGG-Face2 and OpenFace [85], [84], [5], [89], [87], [88], [64], [86], have been explored. Although deep CNNs have shown better performance than hand-crafted texture descriptor-based MAD methods on both digital and print-scan data, the generalization capability of these approaches is limited across different print and scan datasets [115].

Hybrid S-MAD Hybrid approaches are based on multiple feature extractors or classifiers that are combined to detect face morphing attacks. Several approaches have been proposed that combine features, morphing detection scores or decision scores [6], [84], [86], [87], [90], [91]. As these approaches combine more than one feature extractor and classifier, the MAD performance is generally superior to

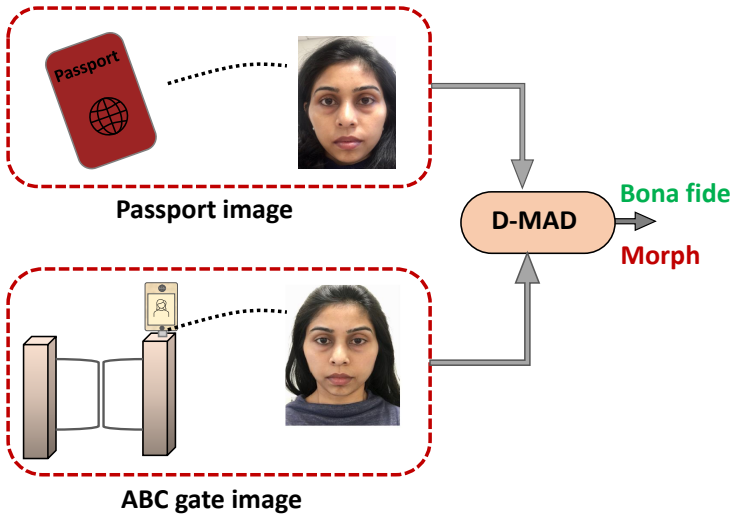


Figure 11.7: Example illustrating differential image-based morphing attack detection (D-MAD) in a passport control scenario

11.7.2 Differential Image-based MAD (D-MAD)

The objective of D-MAD approaches is to make a decision regarding whether a suspect image is morphed or bona fide when a corresponding image captured in a

Table 11.4: S-MAD Techniques: Advantages and Limitations

Feature Type	Advantages	Limitations
Texture Features	- Easy to implement. - Low computational cost. - Good performance when trained and tested with the same morph data types (digital/print-scan). - Effective on digital morph face data	- Lacks generalisation capabilities across both image resolution and morph data type (digital/print-scan). - Sensitive to image resolution. - Degraded performance with print-scan data.
Image Quality Features	- Easy to implement. - Low computational cost. - Less sensitive to accurate segmentation of the face region. - Can be used with different morph data types (digital/print-scan).	- Lacks generalisation across both image resolution and morph data types (digital/print-scan). - Sensitive to compressed data. - Not a reasonable performance across different face morph data types (digital/print-scan).
Hybrid Features	- Good detection performance across different morph data types (digital/print-scan). - Good detection performance when trained and tested with the same morph data type (digital/print-scan). - Reasonable generalisability performance for different morph data types (digital/print-scan).	- Difficult to implement, as it requires hyper-parameter tuning. - High computational cost. - Requires optimisation of several hyper-parameters.
Residual Noise Features	- Easy to implement. - Low computational cost. - Highly accurate detection performance on digital morph data type. - Less sensitive to face region. - Generalisation ability across different image resolutions.	- Applicable only to the digital morph data type. - Promising results for high-resolution images. - Sensitive to image compression.
Deep CNN features	- Good performance when trained and tested with the same morph data type (digital/print-scan). - No need to train CNN from scratch, as deep CNN shows good detection performance.	- High computational cost. - Lacks generalisation across different face morph data types (digital/print-scan). - Training CNN from scratch requires large database.

Table 11.5: State-of-the-art D-MAD

Reference	Detection Type	Approach	Algorithm	Database
M Ferrara et al. [119]	D-MAD	Demorphing	Demorphing by image subtraction	Print-scan
M Ferrara et al. [119]	D-MAD	Demorphing approach	Face verification	Digital
U Scherhag et al. [114]	D-MAD	Landmark-based approach	Distance-based and angle-based feature extraction with Random Forest, SVM without kernel and SVM with radial basis function classifier	Digital
U Scherhag et al. [88]	S-MAD + D-MAD	Feature difference-based approach	Pre-processing and feature extraction using texture descriptors, keypoint extractors, gradient estimators and deep learning-based method	Digital
N Damer et al. [84]	D-MAD	Multi-detector fusion	LBPH, Transferable deep-CNN	Digital
J M Singh [116]	D-MAD	Deep learning	SfS Net, AlexNet	Digital + Print-scan
N Damer et al. [113]	D-MAD	Landmark shift	Landmark detection, shift representation	Digital
F Peng et al. [117]	D-MAD	Face restoration by demorphing GAN	Symmetric dual-network architecture	Digital
U Scherhag et al. [115]	D-MAD	Deep Face Representation	ArcFace Network, FaceNet algorithm	Digital + Print-scan
C Seibold et al. [234]	D-MAD	Deep Learning	Layer-wise Relevance Propagation (LRP)	Digital
D Ortego et al. [112]	D-MAD	Demorphing, Deep CNN-based	Auto-encoders	Digital + Print-scan
S Soleymani et al. [120]	D-MAD	Deep learning	Siamese network	Digital
S Soleymani et al. [121]	D-MAD	Deep learning	Appearance and landmark disentanglement	Digital
S Autherith et al. [122]	D-MAD	Analysis of geometric facial features	Facial anthropometry-based facial feature comparison	Digital

Table 11.6: D-MAD Techniques: Advantages and Limitations

Algorithm Type	Advantages	Limitations
Feature difference	- Easy to implement. - Reasonable detection performance across varying image quality and resolution.	- High computational cost. - Detection performance is sensitive to the type of image data and features. - Detection performance is sensitive to the segmentation of the face region.
Demorphing	- Easy to implement. - Moderate computational time. - High detection accuracy with constrained conditions. - Can visualise the demorphed face if the suspect image is morphed	- Performance is sensitive to the facial pose and imaging conditions. - Requires constrained image data. - Fails with facial pose and lighting variations. - Prior knowledge of the blending factor (or alpha factor) is required.

trusted environment is available. The D-MAD technique is well suited to the border crossing scenario, where the suspected morph image can be obtained from the passport and can then be compared against the live captured face image (or trusted image) from the ABC gates [112]. Fig 11.7 illustrates the application of D-MAD, specifically in a border control scenario. A taxonomy of D-MAD techniques is presented in Figure 11.5, and they can be divided into two broad types: (1) feature difference-based D-MAD and (2) demorphing. Table 11.5 summarises the existing D-MAD techniques, which are briefly discussed as follows

Feature Difference-based D-MAD The basic idea of this approach is to subtract the features computed on both the suspected morph image and a live image captured in a trusted environment. The features are further classified by computing the difference in the feature vectors to detect a morphing attack. To this end, several feature extraction techniques are studied, which involve texture information, 3D information, gradient information, landmark points and deep feature information [115], [116], [84], [114], [113]. Based on the reported results, the deep CNN features have shown the best performance [115]. The majority of the existing works are reported for use cases with digital images, except for a recent work in which a print and scan dataset was explored with improved results [115, 212].

Demorphing Face demorphing techniques invert the morphing procedure and reveal the component images that are used to generate the morphed image. The first proposal in this area was that of Ferrara et al. [119], which was designed to work with landmark-based morph generation. Recent work along these lines is based on using deep CNNs [112] [117]. These techniques are robust when the image quality is good; however, the detection performance degrades when a face image is captured in real-life conditions with pose and lighting variations that are commonly encountered in ABC gates. Table 11.6 presents the advantages and limitations of existing D-MAD techniques.

11.8 Performance Metrics

In this section, we discuss the performance evaluation metrics that are widely used in the literature and publicly available competitions to benchmark the performance of MAD techniques.

11.8.1 Vulnerability Assessment of FRSs

For a morphed image to be deemed a significant threat to an FRS, it is necessary to establish the threat potential. Most works determine the threat potential by measuring the vulnerability of FRSs. We therefore provide a brief overview of suitable metrics for establishing the relevance of morph attacks through vulner-

ability metrics. The goal of face vulnerability analysis is to measure whether the generated morphed face image can be verified against all contributory data subjects. Thus, when a morphed face image is enrolled into an FRS and probed with

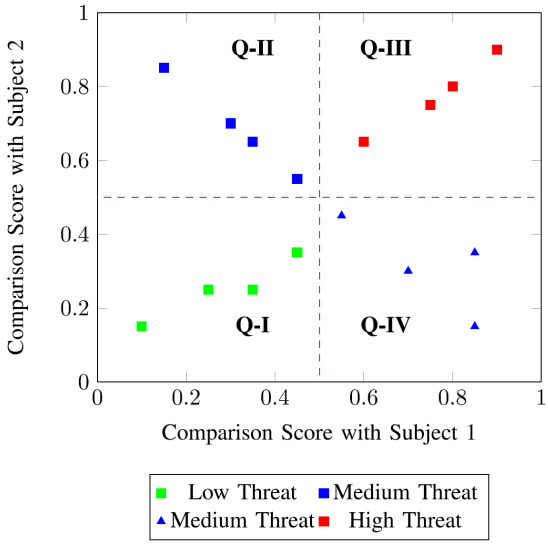


Figure 11.8: Threats of morphed images with respect to comparison scores against both contributing subjects. The figure illustrates that morphed images crossing the threshold of 0.5 (i.e., those lying in quadrant Q-III) are effective attacks with a more severe threat to the FRS than those in Q-II and Q-IV.

Fig 11.8 illustrates an example of the vulnerability plots that represent the scattered data of comparison scores from FRSs. The sample vulnerability plot is simplified for visualisation purposes to provide an illustration of the vulnerability analysis. Fig 11.8 can be interpreted using four different quadrants. The first quadrant (bottom left quadrant) $Q-I$ indicates that the morphed image is not verified as belonging to either of the two contributing data subjects. Thus, a large number of comparison scores in the first quadrant indicates that the morph generation method is not strong enough to deceive the COTS FRS (in other words, the morphed image is not a severe threat). The second quadrant (top left quadrant) $Q-II$ indicates that the morphed image can be verified as data subject-2 (one of the contributing subjects) only. Therefore the morphed images pose an intermediate-strength threat. The third quadrant (top right quadrant) $Q-III$ indicates that the morphed image

is verified as both contributing data subjects (subject-1 and subject-2). Thus, the larger the number of comparison scores in this quadrant, the greater the threat and vulnerability of the analysed FRS with respect to morphed images. The fourth quadrant (bottom right quadrant) $Q-IV$ indicates that the morphed image can be verified as data subject-1 only. Therefore, the morphed images again pose an intermediate-strength threat to the FRS.

To mathematically quantify the vulnerability of an FRS to morphed face images, the metrics below have been developed and adapted in the literature.

Mated Morph Presentation Match Rate (MMPMR) This metric was initially proposed by Scherhag et al. in [125]. It defines the proportion of morphed images verified with its contributing images.

$$MMPMR = \frac{1}{M} \cdot \sum_{m=1}^{M} \left\{ \left[\min_{n=1,\ldots,N_m} S_m^n \right] > \tau \right\}, \tag{1}$$

where M is the number of morphed images and N_m is the total number of subjects contributing to morph m. S_m^n is the comparison score for mated morph for morph m of the n^{th} subject, and τ is the threshold of the FRS at a chosen False Match Rate (FMR).

The rationale of MMPMR is that a morphing attack succeeds if all contributing subjects are verified successfully against the morphed image. MMPMR considers multiple comparisons, which are related to multiple authentication attempts. This may not always be the case. A successor of the MMPMR metric named the fully matched morph presentation match rate (FMMPMR) was introduced by Venkatesh et al. [3] to address the quadrants employed for vulnerability assessment, as shown in Figure 11.8. The details of the FMMPMR are provided below.

Fully Mated Morph Presentation Match Rate (FMMPMR) This metric defines the proportion of morphed images verified with their contributing subjects again under the condition that the morphed image is verified successfully against both contributing subjects [3]. This metric further takes into account both pairwise comparisons of contributing subjects and the number of attempts compared to MMPMR and is described as follows:

$$FMMPMR = \frac{1}{P} \sum_{M,P} (S1_M^P > \tau) AND (S2_M^P > \tau) \ldots AND (Sk_M^P > \tau) \tag{11.1}$$

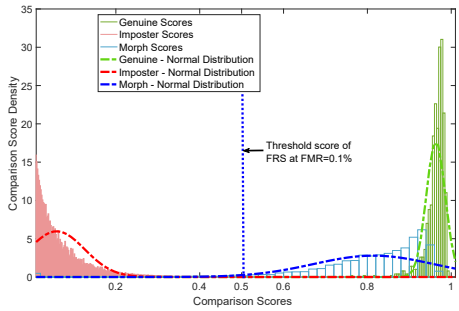


Figure 11.9: Illustration of morph attacks in conjunction with the strength of the FRS. As noted from the figure, the genuine and impostor distributions of the comparison scores are clearly separated, indicating the strength of the FRS while indicating the vulnerability to morph attacks, as most of them cross the pre-defined threshold of 0.5 at a chosen $FMR = 0.1\%$.

where $P = 1, 2, \ldots, p$ represents the number of attempts made by comparing all probe images from the contributing subject against the M^{th} morphed image, $K = 1, 2, \ldots, k$ represents the number of data subjects contributing to the constitution of the generated morphed image (in our case, $K = 2$), Sk_M^P represents the comparison score of the K^{th} contributing subject obtained in the P^{th} attempt (in this case, the P^{th} probe image from the dataset) corresponding to the M^{th} morph image and τ represents the threshold value corresponding to FMR = 0.1%. The FMMPMR metric verifies the morphed image with its contributing subjects and takes into account the number of attempts. It is therefore a relevant and realistic metric to quantify the vulnerability and establish the true attack strength of a morph generation method.

Joint Evaluation of an FRS and Vulnerability to Morph Attacks In addition to Fig 11.8, we note that the FRS can have a high recognition accuracy (i.e., biometric performance) but can also have a high vulnerability to morphing attacks. It is therefore essential to first evaluate the biometric performance of the FRS according to international standard ISO/IEC 19795-1 [235] and subsequently evaluate its vulnerability by using the preset threshold (e.g., FMR = 0.1%). We illustrate a chosen COTS FRS in Fig 11.9, where one can see the success of the morphing attack for a selected threshold (τ) of 0.5, corresponding to $FMR = 0.1\%$. We conclude that the real strength of an FRS cannot be established unless good recognition performance and robustness with respect to morphing attacks are analysed and reported. For this reason, Scherhag et al. [125] established a relative measure that combines the recognition accuracy with vulnerability measures, and this metric is referred to as the relative morph match rate ($RMMR(\%)$). Specifically, when τ is employed

to obtain either the MMPMR or FMMPMR, as discussed earlier, the RMMR can be defined as follows [125]:

$$RMMR(\tau)_{MMPMR} = 1 + (MMPMR(\tau)) - [1 - FNMR(\tau)] \tag{11.2}$$

$$RMMR(\tau)_{FMMPMR} = 1 + (FMMPMR(\tau)) - [1 - FNMR(\tau)] \tag{11.3}$$

where $FNMR$ indicates the false rejection rate ($FNMR$) of the FRS under consideration obtained at the threshold τ.

11.8.2 MAD Performance Metrics

The robustness of MAD algorithms is measured using the performance metrics defined in the International Standard ISO/IEC 30107-3 [21] and is applicable to reporting the morphing attack detection performance. Since the MAD performance can be visualised as a binary classification problem, the following metrics are widely used to benchmark MAD algorithms:

- **Attack presentation classification error rate (APCER):** Defines the proportion of attack samples incorrectly classified as bona fide face images.
- **Bona fide presentation classification error rate (BPCER):** Defines the proportion of bona fide images incorrectly classified as attack samples.

However, it is not possible to optimise both the APCER and BPCER jointly; it is thus natural to set (or fix) either the BPCER or APCER and report the result with a dependency of the other metric (either the APCER or BPCER). Most works have reported results by setting a pre-defined security level (e.g., indicating the maximum proportion of morph accepts they can tolerate) and then fixing the APCER accordingly at values of $@1\%$, 5% or 10% [83, 212, 38]. As shown in Fig 11.10, MAD Algorithm 3 would be preferred at a given APCER of 5% or 10% in the benchmark compared with the other two algorithms.

11.8.3 Joint Evaluation of MAD Algorithms and Vulnerability

In a real-life scenario, an FRS may operate with a MAD subsystem in integrated processing. For a successful attack, it is therefore important that the morphed face image can invade the enrolment process and can match to probe images from the contributing subjects. To quantify the vulnerability in the presence of a MAD, a metric called the Actual mated Morph Presentation Match Rate (AMPMR) was recently proposed in [39] and can be written as follows:

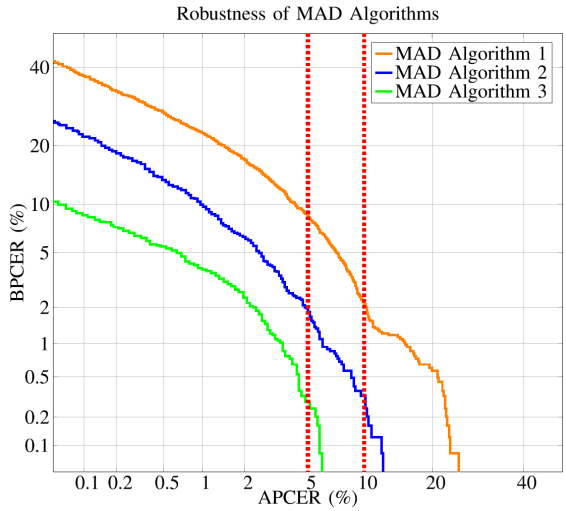


Figure 11.10: Sample illustration of the detection accuracy of MAD algorithms at different operating points with a detection error tradeoff curve (DET). As noted from the figure, MAD Algorithm 3 performs best at a chosen APCER of 5% or 10%.

$$AMPMR(th_{fa}, th_{mad}) = \frac{1}{N}\sum_{i=1}^{N}(((\min_{j=1,..Mi} SC_{ij}) > th_{fa})AND(SC_{mad-i} > th_{mad})) \tag{11.4}$$

where the total number of morphed images is denoted by N. SC_{ij} is the face recognition score of the i^{th} morphed image when compared to a probe sample of the j^{th} contributor. Mi is the number of contributors to the morphed image. SC_{mad-i} is the MAD score of the i^{th} sample. Based on these metrics, higher values of the AMPMR indicate more severe vulnerability.

11.9 Public Evaluation and Benchmarking

In this section, we summarize evaluations that publicly benchmark morphing attack detection performance. At the time of this writing, there are two such benchmarks: 1) The Face Recognition Vendor Test (FRVT) Part 4: MORPH - Performance of Automated Face Morph Detection [128] and 2) Bologna-SOTAMD: Evaluation of Differential MAD and Single Image MAD [38]. These benchmarks have provided a common platform that includes datasets, evaluation protocols and the computational environment. The platforms provide a trustworthy assessment of submitted algorithms. Below, we briefly describe the databases used in each platform and the performance achieved by various algorithms that are presented.

11.9.1 NIST-FRVT Part 4: MORPH - Performance of Automated Face Morph Detection

The FRVT MORPH test was introduced in June 2018 to provide a common platform for independent testing of MAD face technologies and to ensure a common assessment methodology. The dataset used in the evaluation was created using different morphing methods with the objective of identifying low-quality morphing (generated using freely available tools), automated morph generation (generated using an automatic tool without human intervention) and high-quality morph generation (generated with commercial morphing software and additional post-processing that is carried out to mask potential artifacts). The evaluation is carried out for both S-MAD and D-MAD techniques. However, the probe data used in the D-MAD evaluation are not effectively obtained from ABC gates. Several algorithms are evaluated, and the majority of the participants in the competition to date are from academic institutions. Most of the submissions for S-MAD are based on texture features, while for D-MAD, both face demorphing and differential feature-based techniques are evaluated. Based on the recent evaluation report, it can be noted that

1. None of the algorithms has indicated a reliable detection performance meeting the FRONTEX operational requirement [126], and thus, face morphing attack detection remains a challenging task.

2. The quality of morph generation has a direct impact on the performance of both S-MAD and D-MAD techniques.

In the S-MAD category, the use of hybrid features [91] has shown better performance than other MAD methods, while among the methods in the D-MAD category, the approach of latent feature differences based on ArcFace features [115] has attained the best detection performance.

11.9.2 Bologna-SOTAMD: Differential Morph Attack Detection

The Bologna-SOTAMD benchmark was opened for evaluation in 2019 and provided a common evaluation platform to benchmark D-MAD techniques. The Bologna-SOTAMD D-MAD benchmark consists of a database collected in the European SOTAMD project [129] using real ABC gates. The morphing was carried out using automated approaches with both open-source and commercial software. Several MAD techniques have been benchmarked, which include both face demorphing and feature difference methods, and the details of the evaluation protocol and the performance of various submitted algorithms can be found in [38]. Among the multiple algorithms evaluated, it can be noted that the existing D-MAD tech-

niques are not robust enough to detect face morphing attacks in accordance with the FRONTEX operational requirement [126], highlighting the challenge of MAD again. The use of the feature difference-based D-MAD technique shows better performance than face demorphing techniques. The best result, a detection equal error rate (D-EER) of 3.36 %, has been reported on digital data, and D-EER = 3.36 % has been reported on print-scan data.

11.9.3 Bologna-SOTAMD: Single Morph Attack Detection

The Bologna server has also hosted a public benchmark for S-MAD since 2020. The S-MAD dataset was constructed using high-quality passport images similar to those used in real passports. The morphed images were generated using both commercial (FantaMorph, FaceFusion) and open-source (triangulation with facial landmarks) face morphing software. Post-processing was carried out using automatic and manual processes to reduce the artifacts generated using the face morphing software. For more information on the database and evaluation protocol, see [38]. As evaluation started only recently, few algorithms have been benchmarked on the Bologna S-MAD platform. The baseline performance reported a D-EER of 37.10% and 38.99% on print-scan and digital morphed images. These preference measures indicate the challenges in detecting face morphing images using S-MAD techniques.

11.9.4 Discussion of Public Evaluation

Based on the above discussion of publicly available benchmarks and competitions, it can be noted that the reliable detection of face morphing attacks remains challenging. The performance of S-MAD is severely degraded compared to that of D-MAD. This can be attributed to the availability of additional information (another image) that can be used to make the final decision. The interesting outcomes of these competitions indicate that the use of the hybrid feature-based S-MAD technique has shown improved generalizability across various morph generation methods. At the same time, the feature difference method used in the context of D-MAD has shown a more robust performance on both benchmarks.

11.10 Open Challenges

The research topic of face morphing and detection has received great interest from both research and governmental stakeholders. This has resulted in intensive research activities around studying the vulnerability of COTS FRSs and the development of several MAD techniques to reliably detect such attacks. However, there are still several challenges and open issues that need to be addressed. In the next section, we present these challenges and open issues in the field of face morphing attack detection.

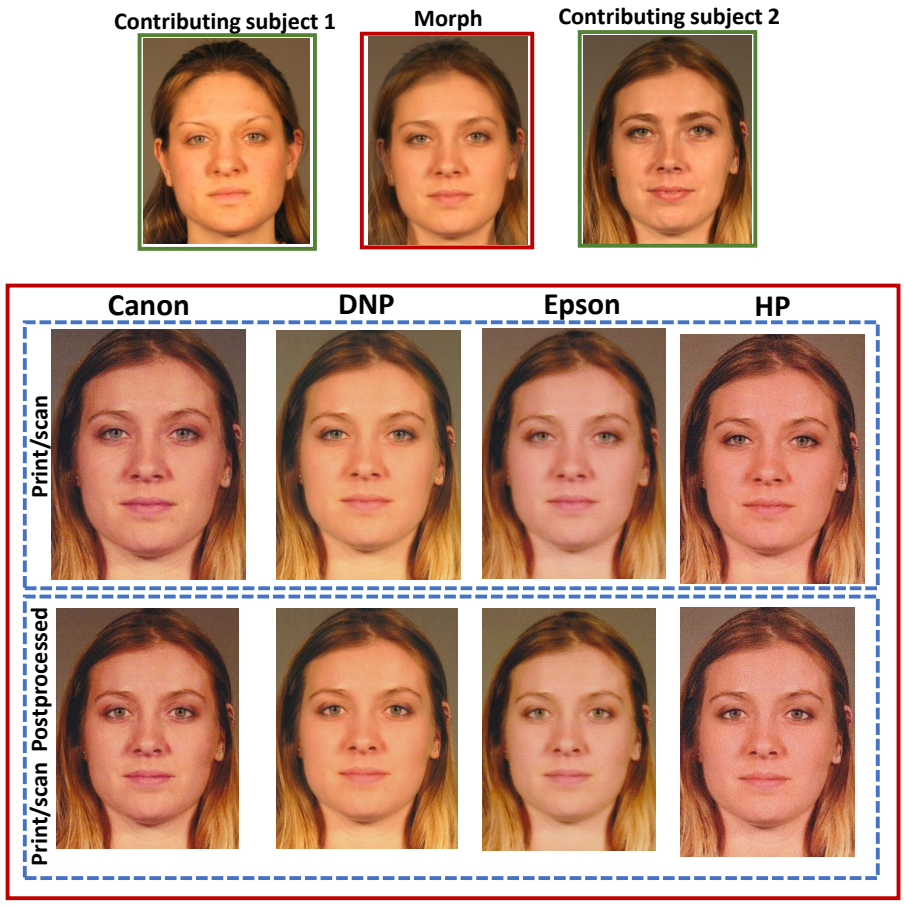


Figure 11.11: Example of print-scan images and post-processed images. The variation in the data quality across different printers and scanners is notable, which challenges the MAD algorithms.

Unavailability of Large-Scale public datasets with variation The unavailability of large-scale face morphing datasets reflecting real-life scenarios hinders the development of robust MAD. Furthermore, considering the different modes of morph attacks (digital and print-scan), it is necessary to generate and evaluate MAD algorithms on digital and print-scan datasets. However, the generation of large print-scan databases is quite expensive and tedious. Additionally, these databases cannot be shared publicly due to licensing restrictions or to privacy and General Data Protection Regulation (GDPR) [236] concerns. Hence, there is a limitation in accessing the existing morphing databases. Although the publicly available benchmarks now host large-scale databases, those datasets can only test submitted MAD algorithms. They cannot aid the further development of MAD algorithms. However, the systematic generation of morphed face images with various types of morphing software combined with different types of print-scan processes must result in large-scale databases that will become available for researchers in order to achieve significant progress in MAD.

Generalizability of MAD techniques The generalization of MAD techniques is crucial in achieving reliable performance in real-life border control scenarios. However, the existing MAD techniques are evaluated only on known types of face morph generation techniques and known sources of re-digitisation (printer and scanner types), except in NIST-FRVT Part 4: MORPH - Performance of Automated Face Morph Detection and Bologna-SOTAMD. Fig 11.11 illustrates the variation in morphed image quality due to different types of printers and scanners. The performance reported in the benchmarking study [128], [227], [38] also indicated the degraded performance of MAD techniques on both D-MAD and S-MAD when tested on unknown sources of generation. More significant degradation is noted with S-MAD methods, which is attributed to learning-based systems that can learn a decision policy based on known data. These factors limit the applicability of learning-based MAD techniques if they are not trained on a large-scale dataset with all real-life variants. Thus, it is essential to devise a MAD approach that is robust in detecting face morphing attacks.

Selection of Data Subjects for Morphing In earlier studies, morphed images were generated by randomly selecting the contributing data subjects. It is a well-established assumption that a morphing attack will be more successful with both human observers and machines (FRSs) if the candidate data subjects are selected based on a look-like measure. Some recent works [237], [13], [238], [3] describe the selection of data subjects in the morphing process. However, in the systematic study of these existing methods regarding the impact on FRS vulnerability, the detection performance of both human observers and automatic MAD detection methods still needs to be investigated.

Variation with Face Co-Variates The critical aspect that is not systematically studied with MAD is the role of face co-variates, which include age, gender, ethnicity, identification factors, image post-processing and image quality. A preliminary study on the effect of ageing on morphing vulnerability and detection was presented in [3] and has revealed the influence of ageing on face morphing vulnerability. The variation of face co-variates has a greater influence on S-MAD techniques, while for D-MAD techniques, the imaging quality plays a vital role. As the images are captured using an ABC gate in D-MAD, the influence of varying illumination due to day and night light settings needs to be investigated. Additionally, images captured live at an ABC gate may be acquired with eyeglasses or hair occlusions, and this has not yet been investigated. Thus, it is essential to benchmark both D-MAD and S-MAD techniques in a real-life scenario with all those co-variates. Another aspect that has not yet been investigated with regard to its impact on MAD is potential face beautification. It is expected that face images are beautified prior to applying for a passport in many countries [239]. As the beautification process changes the image properties, it is essential to understand both vulnerability and MAD for this particular problem.

Performance Metrics Considering that face morphing attack detection is emerging as a new operational problem, there has been only a slow convergence towards harmonized testing and reporting. Publicly available benchmarking and competitions have employed ISO/IEC metrics [21] to benchmark the detection performance of MAD techniques. However, there are no standardized metrics yet to evaluate the vulnerability of FRSs with respect to morphing attacks. Furthermore, the available vulnerability metrics, such as FMMPMR and MMPMR, are not feasible for use in operational scenarios, including ABC gates and passport application scenarios. Therefore, there is a strong need for a standardized vulnerability evaluation metric incorporating experience from both practitioners and researchers working on face MAD. The availability of an international standard using ISO/IEC, together with commonly used vulnerability metrics, is discussed in Section 11.8, and this needs further effort.

Component-Based Morphing Almost all literature has studied face morphing as a holistic problem with full-face image morphing. Qin et al. [39] introduced partial face morphing, including a preliminary study on morphing only specific regions of the face. Extensive experiments indicate that partial morphing of the eye and nose poses a severe threat to commercial face recognition systems [39]. However, the systematic evaluation of high-quality face images has yet to be studied together with the impact on human expert observers (for example, border guards and super-recognizers).

Identical Twins and Look-Alikes The influence of morphing on identical twins and look-alikes is an interesting problem that needs systematic study within the scope of morphing. The vulnerability of FRSs to face morphing images generated from identical twins and similar subjects needs to be studied on large-scale databases.

User convenience The design of user-convenient (or user-friendly) MAD systems plays a crucial role in making detection subsystems deployable in real-time applications. Thus, there is a need to design face MAD systems that allow minimal user intervention (from both operators and applicants). This fact needs to be considered when designing D-MAD techniques that are tailored for ABC systems.

11.11 Conclusion

FRSs have gained a large amount of trust for security-related applications. However, morphing attacks on FRSs can be a hindrance to establishing a secure society. Furthermore, various morphing attack detection techniques have been proposed by several researchers to effectively detect morphed images. However, improvements in deep learning and machine learning techniques have resulted in the generation of high-quality morphs using various new techniques. Hence, generalizing MAD methods is still predicted to be far in the future considering the basic challenge of obtaining large public databases with variations and different morph generation techniques. In this article, we detail the advancement of different types of morph generation techniques. Along with a brief overview of the different types of morphing attack detection techniques, the corresponding performance metrics are reported. We also provide a brief discussion of the challenges faced in this field in developing a robust technique to detect morphs, which serves as a reference for future work.

Acknowledgment

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under Grant 883356. This text reflects only the author's views and the Commission is not liable for any use that may be made of the information contained therein.

Bibliography

[1] S. Venkatesh, H. Zhang, R. Raghavendra, K. Raja, N. Damer, and C. Busch. Can GAN generated morphs threaten face recognition systems equally as landmark based morphs? - Vulnerability and Detection. In *2020 8th International Workshop on Biometrics and Forensics (IWBF)*, pages 1–6, 2020.

[2] H. Zhang, S. Venkatesh, R. Raghavendra, K. Raja, N. Damer, and C. Busch. MIPGAN—Generating strong and high quality morphing attacks using identity prior driven GAN. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 3(3):365–383, 2021.

[3] S.Venkatesh, K. Raja, R. Raghavendra, and C. Busch. On the influence of ageing on face morph attacks: Vulnerability and Detection. In *International Joint Conference on Biometrics (IJCB)*, pages 1–8. IEEE, September 2020.

[4] S. Venkatesh, R. Raghavendra, K. Raja, L. Spreeuwers, R. Veldhuis, and C. Busch. Morphed face detection based on deep color residual noise. In *9th Intl. Conf. on Image Processing Theory, Tools and Applications (IPTA)*. IEEE, November 2019.

[5] S. Venkatesh, R. Raghavendra, K. Raja, L. Spreeuwers, R. Veldhuis, and C. Busch. Detecting morphed face attacks using residual noise from deep multi-scale context aggregation network. In *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 269–278. IEEE, March 2020.

[6] S. Venkatesh, R. Raghavendra, K. Raja, and C. Busch. Single image face morphing attack detection using ensemble of features. In *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, pages 1–6. IEEE, 2020.

[7] S. Venkatesh, R. Raghavendra, K. Raja, and C. Busch. Face morphing attack generation and detection: A comprehensive survey. *IEEE Transactions on Technology and Society*, 2(3):128–145, March 2021.

[8] Cognitec Systems GmbH. Facevacs technology - version 9.4.2. https://www.cognitec.com/facevacs-technology.html, 2020.

[9] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman. Vggface2: A dataset for recognising faces across pose and age. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pages 67–74. IEEE, 2018.

[10] J. Deng, J. Guo, and S. Zafeiriou. ArcFace: Additive angular margin loss for deep face recognition. In *Conf. on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[11] Neurotechnology. VeriLook SDK. http://www.neurotechnology.com/verilook.html.

[12] X. Wu, R. He, Z. Sun, and T. Tan. A light CNN for deep face representation with noisy labels. *IEEE Transactions on Information Forensics and Security*, 13(11):2884–2896, 2018.

[13] R. Raghavendra, K. Raja, S. Venkatesh, and C. Busch. Face morphing versus face averaging: Vulnerability and detection. In *Proc. Intl. Joint Conf. on Biometrics (IJCB)*, 2017.

[14] M. Ferrara, A. Franco, and D. Maltoni. Decoupling texture blending and shape warping in face morphing. In *Intl. Conf. of the Biometrics Special Interest Group (BIOSIG)*. IEEE, September 2019.

[15] N. Damer, Y. Wainakh, V. Boller, S. von den Berken, P. Terhörst, A. Braun, and A. Kuijper. MorGAN: Recognition vulnerability and attack detectability of face morphing attacks created by generative adversarial network. In *Proc. of the 9th IEEE Intl. Conf. on Biometrics: Theory, Applications, and Systems (BTAS)*. IEEE, 2018.

[16] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 2382-37:2012 Information Technology - Vocabulary - Part 37: Biometrics*. International Organization for Standardization, 2012.

[17] S. Li and A. Jain. *Handbook of Face Recognition*. Springer-Verlag, 2 edition, 2011.

[18] W.-Y. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys (CSUR) archive*, 35(4):399–458, December 2003.

[19] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition. *British Machine Vision Association*, pages 1–12, 2015.

[20] M. Wang and W. Deng. Deep face recognition: A survey. *ArXiv*, September 2018.

[21] 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.

[22] M. Ferrara, A. Franco, and D. Maltoni. The magic passport. In *2014 IEEE Intl. Joint Conf. on Biometrics (IJCB)*, pages 1–7, September 2014.

[23] Department of Internal Affairs (DIA), NZ. https://www.passports.govt.nz/passport-photos/passport-photo-requirements/.

[24] GOV.UK. https://www.gov.uk/photos-for-passports/photo-requirements, 2020.

[25] V. Bruce, Z. Henderson, K. Greenwood, P. JB. Hancock, A. M. Burton, and P. Miller. Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, 5(4):339, 1999.

[26] A. Towler, R.I. Kemp, and D. White. Unfamiliar face matching systems in applied settings. *Face processing: systems, disorders and cultural differences. New York: Nova Science Publishing, Inc*, 2017.

[27] P. D. Josh and V. Tim. CCTV on trial: Matching video images with the defendant in the dock. *Applied Cognitive Psychology*, 23:482–505, 2009.

[28] D. Robertson, A. Mungall, D. Watson, K. Wade, S. Nightingale, and S. Butler. Detecting morphed passport photos: a training and individual differences approach. *Cognitive Research: Principles and Implications*, 3(1), June 2018.

[29] R. S. Kramer, M. O. Mireku, R.T. Flack, and K. L. Ritchie. Face morphing attacks: Investigating detection with humans and computers. *Cognitive research: principles and implications*, 4(1):1–15, 2019.

[30] S. Spelsberg. Two faces, one document. https://taz.de/Peng-Kollektiv-faelscht-Passbilder/!5534868/, 2018.

[31] T. Raphael and H. Judith. Activists smuggle photomontage into passport. https://www.spiegel.de/netzwelt/netzpolitik/biometrie-im-reisepass-peng-kollektiv-schmuggelt-fotomontage-in-auswe html, 2018.

[32] Fons Knopjes. State of the art of Morphing Detection. https://www.icao.int/Meetings/TRIP-Symposium-2019/PublishingImages/Pages/Presentations/State%20of%20the%20art%20of%20Morphing%20Detection.pdf/, Accessed: 2022.

[33] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, Jin Chang, K. Hoffman, J. Marques, Jaesik Min, and W. Worek. Overview of the face recognition grand challenge. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pages 947–954 vol. 1, June 2005.

[34] A. Kasiński, A. Florek, and A. Schmidt. The PUT face database. *Image Processing & Communications*, 13(3-4):59–64, 2008.

[35] G. Bingham, K. Kempfert, B. Yip, J. Fabish, M. Ferguson, C. Nansalo, K. Park, R. Towner, T. Kling, Y. Wang, et al. Preliminary studies on a large face database morph-ii.

[36] N. Damer, K. Raja, M. Sussmilch, S. Venkatesh, F. Boutros, M. Fang, F. Kirchbuchner, R. Raghavendra, and A. Kuijper. ReGenMorph: Visibly realistic GAN generated face morphing attacks by attack re-generation. *Inernational Symposium on Visual Computing ISVC*, 2021.

[37] S. Venkatesh, R. Raghavendra, and K. Raja. Face morphing of newborns can be threatening too : Preliminary study on vulnerability and detection. In *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–8, 2021.

[38] K. Raja, M. Ferrara, A. Franco, L. Spreeuwers, I. Batskos, et al. Morphing attack detection - database, evaluation platform and benchmarking. *IEEE Trans. on Information Forensics and Security*, November 2020.

[39] L. Qin, F. Peng, S. Venkatesh, R. Raghavendra, M. Long, and C. Busch. Low visual distortion and robust morphing attacks based on partial face image manipulation. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 3(1):72–88, 2020.

[40] S. Venkatesh. Multi-spectral finger based user verification using off-the-shelf deep features. In *IEEE International Conference on Imaging systems and techniques (IST 2022)*, 2022.

[41] R. Raghavendra, S. Venkatesh, K. Raja, and C. Busch. Handwritten signature and text based user verification using smartwatch. In *25th International Conference on Pattern Recognition (ICPR)*, pages 5099–5106, 2021.

[42] S. Venkatesh, R. Raghavendra, and P. Bours. Video based deception detection using deep recurrent convolutional neural network. In Neeta Nain, Santosh Kumar Vipparthi, and Balasubramanian Raman, editors, *Computer Vision and Image Processing*, pages 163–169, Singapore, 2020. Springer Singapore.

[43] R. Raghavendra, J. Singh, S. Venkatesh, K. Raja, and C. Busch. Face presentation attack detection using multi-classifier fusion of off-the-shelf deep features. In *Proc. of the 4th Intl. Conf. on Computer Vision and Image Processing (CVIP)*, 2019.

[44] J. M. Singh, S. Venkatesh, K. B. Raja, R. Raghavendra, and C. Busch. Detecting finger-vein presentation attacks using 3D shape amp; diffuse reflectance decomposition. In *2019 15th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)*, pages 8–14, 2019.

[45] N. Vetrekar, R. Raghavendra, K. Raja, S. Venkatesh, R. Gad, and C. Busch. Visible to band gender classification: An extensive experimental evaluation based on multi-spectral imaging. In *2019 15th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)*, pages 120–127, 2019.

[46] K. Raja, R. Raghavendra, S. Venkatesh, M. Gomez-Barrero, C. Rathgeb, and C. Busch. A study of handcrafted and naturally learned features for fingerprint presentation attack detection. In *Handbook of Biometric Anti-Spoofing*. Springer Intl. Publishing, January 2019.

[47] R. Raghavendra, K. Raja, S. Venkatesh, and C. Busch. Design and development of low-cost sensor to capture ventral and dorsal finger vein for biometric authentication. *IEEE Sensors Journal*, 19(15):6102–6111, 2019.

[48] G. Wolberg. Image morphing: a survey. *The visual computer*, pages 360–372, 1998.

[49] A. Patel. Image morphing algorithm: A survey. 2015.

[50] G. Wolberg. *Digital image warping*, volume 10662. IEEE computer society press Los Alamitos, CA, 1990.

[51] N. Arad, N. Dyn, D. Reisfeld, and Y. Yeshurun. Image warping by radial basis functions: Application to facial expressions. *CVGIP: Graphical models and image processing*, 56(2):161–172, 1994.

[52] S. Lee, G. Woberg, K-Y. Chwa, and S.Y. Shin. Image metamorphosis with scattered feature constraints. *IEEE Transactions on Visualization and Computer Graphics*, 2(4):337–354, December 1996.

[53] J. Liao, R. S. Lima, D. Nehab, H. Hoppe, P.V. Sander, and J. Yu. Automating image morphing using structural similarity on a halfway domain. *ACM Trans. Graph.*, 33(5):168:1–168:12, September 2014.

[54] R. Raghavendra, K. Raja, and C. Busch. Detecting morphed face images. In *2016 IEEE 8th Intl. Conf. on Biometrics: Theory, Applications and Systems (BTAS)*. 8th IEEE Intl. Conf. on Biometrics: Theory, Applications and Systems (BTAS-2016), IEEE, September 2016.

[55] E. D. King. Dlib-ml: A machine learning toolkit. *J. Mach. Learn. Res.*, pages 1755–1758, 2009.

[56] O. Çeliktutan, S. Ulukaya, and B. Sankur. A comparative study of face landmarking techniques. *EURASIP Journal on Image and Video Processing*, 2013(1):1–27, 2013.

[57] L. Wiskott, J-M. Fellous, N. Krüger, and C. Von Der Malsburg. Face recognition by elastic bunch graph matching. In *International Conference on Computer Analysis of Images and Patterns*, pages 456–463. Springer, 1997.

[58] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models-their training and application. *Computer vision and image understanding*, pages 38–59, 1995.

[59] D. Ruprecht and H. Muller. Image warping with scattered data interpolation. *IEEE Computer Graphics and Applications*, 15(2):37–43, 1995.

[60] S-Y Lee, K-Y Chwa, S Y Shin, and G Wolberg. Image metamorphosis using snakes and free-form deformations. In *SIGGRAPH*, volume 95. Citeseer, 1995.

[61] T. Ucicr. Feature-based image metamorphosis. *Computer graphics*, 26:2, 1992.

[62] S. Schaefer, T. McPhail, and J. Warren. Image deformation using moving least squares. In *ACM transactions on graphics (TOG)*, pages 533–540. ACM, 2006.

[63] D. W. Choi and C. J. Hwang. Image morphing using mass-spring system. 2011.

[64] C. Seibold, W. Samek, A. Hilsmann, and P. Eisert. Detection of face morphing attacks by deep learning. In *Digital Forensics and Watermarking*, pages 107–120. Springer Intl. Publishing, 2017.

[65] D. Bitouk, N. Kumar, S. Dhillon, P. Belhumeur, and K. S. Nayar. Face swapping: Automatically replacing faces in photographs. *ACM Trans. Graph.*, 27(3):39:1–39:8, 2008.

[66] Y. Weng, L. Wang, X. Li, M. Chai, and K. Zhou. Hair interpolation for portrait morphing. *Computer Graphics Forum*, 32(7):79–84, October 2013.

[67] 3D this face morph. https://3dthis.com/morph.htm, 2020. Accessed: October 2020.

[68] Face swap online. https://faceswaponline.com/, 2020. Accessed: October 2020.

[69] Magic morph 1.95. https://www.malavida.com/en/soft/magic-morph/, 2020. Accessed: October 2020.

[70] Morph thing. https://www.morphthing.com/, 2020. Accessed: October 2020.

[71] Face morpher. http://www.facemorpher.com/, 2020. Accessed: October 2020.

[72] N. Damer, F. Boutros, A.M Saladié, F. Kirchbuchner, and A. Kuijper. Realistic dreams: Cascaded enhancement of GAN-generated images with an example in face morphing attacks. In *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–10, 2019.

[73] Official Journal of the European Union. Laying down the specifications for the quality, resolution and use of fingerprints and facial image for biometric verification and identification in the entry/exit system (ees). Technical report, February 2019.

[74] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51–59, 1996.

[75] V. Ojansivu and J. Heikkilä. Blur insensitive texture classification using local phase quantization. In *2008 International Conference on Image and Signal Processing (ICISP)*, pages 236–243. Springer Berlin Heidelberg, 2008.

[76] J. Kannala and E. Rahtu. BSIF: Binarized statistical image features. In *2012 21st Intl. Conf. on Pattern Recognition (ICPR)*, pages 1363–1366, 2012.

[77] G. D. Lowe. Object recognition from local scale-invariant features. In *IEEE Intl. Conf. on Computer Vision (ICCV 1999)*, volume 2, pages 1150–1157. IEEE Computer Society, 1999.

[78] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 10(3):346–359, 2008.

[79] C. Seibold, A. Hilsmann, and P. Eisert. Reflection analysis for face morphing attack detection. In *Proc. of the 26th European Signal Processing Conf. (EUSIPCO)*, 2018.

[80] U. Scherhag, L. Debiasi, C. Rathgeb, C. Busch, and A. Uhl. Detection of face morphing attacks based on PRNU analysis. *Trans. on Biometrics, Behavior, and Identity Science (TBIOM)*, 2019.

[81] 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 *2017 5th Intl. Workshop on Biometrics and Forensics (IWBF)*, pages 1–6. IEEE, April 2017.

[82] L. Debiasi, U. Scherhag, C. Rathgeb, A. Uhl, and C. Busch. PRNU-based detection of morphed face images. In *6th Intl. Workshop on Biometrics and Forensics*, pages 1–6, 2018.

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

[84] N. Damer, S. Zienert, Y.Wainakh, A. Saladie, F. Kirchbuchner, and A. Kuijper. A multi-detector solution towards an accurate and generalized detection of face morphing attacks. In *2019 22th Intl. Conf. on Information Fusion (FUSION)*. IEEE, July 2019.

[85] M. Ferrara, A. Franco, and D. Maltoni. Face morphing detection in the presence of printing/scanning and heterogeneous image sources. *CoRR*, abs/1901.08811, 2019.

[86] 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 Conf. (EUSIPCO)*. IEEE, September 2019.

[87] U. Scherhag, C. Rathgeb, and C. Busch. Morph detection from single face images: a multi-algorithm fusion approach. In *Intl. Conf. on Biometric Engineering and Applications 2018 (ICBEA)*, pages 1–7, 2018.

[88] U. Scherhag, C. Rathgeb, and C. Busch. Towards detection of morphed face images in electronic travel documents. In *13th IAPR Workshop on Document Analysis Systems (DAS)*, pages 187–192, 2018.

[89] C. Seibold, A. Hilsmann, and P. Eisert. Style your face morph and improve your face morphing attack detector. In *2019 Intl. Conf. of the Biometrics Special Interest Group (BIOSIG)*. IEEE, September 2019.

[90] R. Raghavendra, S. Venkatesh, K. Raja, and C. Busch. Detecting face morphing attacks with collaborative representation of steerable features. In *Intl. Conf. on Computer Vision and Image Processing (CVIP)*, September 2018.

[91] R. Raghavendra, S. Venkatesh, K. Raja, and C. Busch. Towards making morphing attack detection robust using hybrid scale-space colour texture features. In *IEEE 5th Intl. Conf. on Identity, Security, and Behavior Analysis (ISBA)*. IEEE, January 2019.

[92] S. Lorenz, U. Scherhag, C. Rathgeb, and C. Busch. Morphing attack detection: A fusion approach. In *2021 IEEE 24th International Conference on Information Fusion (FUSION)*, pages 1–7, 2021.

[93] P. Aghdaie, B. Chaudhary, S. Soleymani, J. Dawson, and N.M. Nasrabadi. Detection of morphed face images using discriminative wavelet sub-bands. *arXiv preprint arXiv:2106.08565*, 2021.

[94] P. Aghdaie, B. Chaudhary, S. Soleymani, J. Dawson, and N.M. Nasrabadi. Attention aware wavelet-based detection of morphed face images. In *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–8, 2021.

[95] A. Makrushin, T. Neubert, and J. Dittmann. Automatic generation and detection of visually faultless facial morphs. In *Proc. of the 12th Intl. Joint Conf. on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 6: VISAPP, (VISIGRAPP 2017)*, pages 39–50, 2017.

[96] T. Neubert. Face morphing detection: An approach based on image degradation analysis. In *Digital Forensics and Watermarking*, pages 93–106. Springer International Publishing, 2017.

[97] A. Asaad and S. Jassim. Topological data analysis for image tampering detection. In Christian Kraetzer, Yun-Qing Shi, Jana Dittmann, and Hyoung Joong Kim, editors, *Digital Forensics and Watermarking*, pages 136–146, Cham, 2017. Springer International Publishing.

[98] U. Scherhag, R. Raghavendra, K. B. Raja, M. Gomez-Barrero, C. Rathgeb, and C. Busch. On the vulnerability of face recognition systems towards morphed face attacks. In *5th Intl. Workshop on Biometrics and Forensics (IWBF)*, pages 1–6, 2017.

[99] C. Kraetzer, A. Makrushina, T. Neubert, M. Hildebrandt, and J. Dittmann. 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)*, pages 21–32, 2017.

[100] 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*, June 2018.

[101] A. Makrushin, C. Kraetzer, T. Neubert, and J. Dittmann. Generalized benfords law for blind detection of morphed face images. In *Proc. of the 6th ACM Workshop on Information Hiding and Multimedia Security - IH&MMSec '18*. ACM Press, 2018.

[102] T. Neubert, C. Kraetzer, and J. Dittmann. Reducing the false alarm rate for face morph detection by a morph pipeline footprint detector. In *2018 26th European Signal Processing Conf. (EUSIPCO)*. IEEE, September 2018.

[103] L. Spreeuwers, M. Schils, and R. Veldhuis. Towards robust evaluation of face morphing detection. In *Proc. of the 26th European Signal Processing Conf. (EUSIPCO)*, 2018.

[104] L. Debiasi, U. Scherhag, C. Rathgeb, A. Uhl, and C. Busch. PRNU variance analysis for morphed face image detection. In *Proc. of 9th Intl. Conf. on Biometrics: Theory, Applications and Systems (BTAS 2018)*, 2018.

[105] C. Seibold, W. Samek, A. Hilsmann, and P. Eisert. Accurate and robust neural networks for security related applications exampled by face morphing attacks. *Computer Vision and Pattern Recognition*, pages 1–16, 2018.

[106] N. Damer, J. H. Grebe, S. Zienert, F. Kirchbuchner, and A. Kuijper. On the generalization of detecting face morphing attacks as anomalies: Novelty vs. outlier detection. In *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–5, 2019.

[107] T. Neubert, C. Kraetzer, and J. Dittmann. A face morphing detection concept with a frequency and a spatial domain feature space for images on emrtd. In *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*, IH&MMSec'19, page 95–100, New York, NY, USA, 2019. Association for Computing Machinery.

[108] C. Seibold, A. Hilsmann, and P. Eisert. Focused LRP: Explainable AI for face morphing attack detection. In *WACV (Workshops)*, pages 88–96, 2021.

[109] O. A.Abisoye and A. D. Mohammed. Face morphing attack detection in the presence of post-processed image sources using neighborhood component analysis and decision tree classifier. In *Information and Communication Technology and Applications: Third International Conference, ICTA 2020, Minna, Nigeria, November 24-27, 2020, Revised Selected Papers*, volume 1350, page 340. Springer Nature, 2021.

[110] N. Damer, N. Spiller, M. Fang, F. Boutros, F. Kirchbuchner, and A. Kuijper. PW-MAD: Pixel-wise supervision for generalized face morphing attack detection. In George Bebis, Vassilis Athitsos, Tong Yan, Manfred Lau, Frederick Li, Conglei Shi, Xiaoru Yuan, Christos Mousas, and Gerd Bruder, editors, *Advances in Visual Computing*, pages 291–304, Cham, 2021. Springer International Publishing.

[111] J. Tapia and C. Busch. Single morphing attack detection using feature selection and visualization based on mutual information. *IEEE Access*, 9:167628–167641, 2021.

[112] D. Ortega-Delcampo, C. Conde, D. Palacios-Alonso, and E. Cabello. Border control morphing attack detection with a convolutional neural network de-morphing approach. *IEEE Access*, 8:92301–92313, 2020.

[113] N. Damer, V. Boller, Y. Wainakh, F. Boutros, P. Terhörst, A. Braun, and A. Kuijper. Detecting face morphing attacks by analyzing the directed distances of facial landmarks shifts. In Thomas Brox, Andrés Bruhn, and Mario Fritz, editors, *Pattern Recognition*, pages 518–534, Cham, 2019. Springer International Publishing.

[114] U. Scherhag, D. Budhrani, M. Gomez-Barrero, and C. Busch. Detecting morphed face images using facial landmarks. In *Intl. Conf. on Image and Signal Processing (ICISP)*, 2018.

[115] U. Scherhag, C. Rathgeb, J. Merkle, and Christoph Busch. Deep face representations for differential morphing attack detection. *IEEE Trans. on Information Forensics and Security*, 2020.

[116] J. Singh, K. Raja, R. Raghavendra, and C. Busch. Robust morph-detection at automated border control gate using deep decomposed 3D shape & diffuse reflectance. In *Proc. of the 15th Intl. Conf. on Signal Image Technology & Internet Based Systems (SITIS)*, November 2019.

[117] F. Peng, L. Zhang, and M. Long. FD-GAN: Face de-morphing generative adversarial network for restoring accomplice's facial image. *IEEE Access*, June 2019.

[118] B. Chaudhary, P. Aghdaie, S. Soleymani, J. Dawson, and N.M. Nasrabadi. Differential morph face detection using discriminative wavelet sub-bands. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 1425–1434, June 2021.

[119] M. Ferrara, A. Franco, and D. Maltoni. Face demorphing. *IEEE Trans. on Information Forensics and Security*, 13(4):1008–1017, April 2018.

[120] S. Soleymani, B. Chaudhary, A. Dabouei, J. Dawson, and N. Nasrabadi. Differential morphed face detection using deep siamese networks. In *International Conference on Pattern Recognition*, pages 560–572. Springer, 2021.

[121] S. Soleymani, A. Dabouei, F. Taherkhani, J. Dawson, and N. Nasrabadi. Mutual information maximization on disentangled representations for differential morph detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 1731–1741, January 2021.

[122] S. Autherith and C. Pasquini. Detecting morphing attacks through face geometry features. *Journal of Imaging*, 6(11), 2020.

[123] S. Banerjee and A. Ross. Conditional identity disentanglement for differential face morph detection. In *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–8, 2021.

[124] G. Borghi, E. Pancisi, M. Ferrara, and D. Maltoni. A double siamese framework for differential morphing attack detection. *Sensors*, 21(10), 2021.

[125] U. Scherhag, A. Nautsch, C. Rathgeb, M. Gomez-Barrero, R. N. J. Veldhuis, L. Spreeuwers, M. Schils, D. Maltoni, P. Grother, S. Marcel, R. Breithaupt, R. Raghavendra, and C. Busch. Biometric systems under morphing attacks: Assessment of morphing techniques and vulnerability reporting. In *Intl. Conf. of the Biometrics Special Interest Group BIOSIG 2017*, pages 1–7, 2017.

[126] FRONTEX. Best practice technical guidelines for automated border control ABC systems, 2015.

[127] M. Ferrara, A. Franco, D. Maltoni, and C. Busch. Morphing attack potential. In *Proceedings of 10th International Workshop on Biometrics and Forensics (IWBF-2022)*, 2022.

[128] N. Mei, P. Grother, K. Hanaoka, and J. Kuo. Face Recognition Vendor Test (FRVT) Part 4: Performance of Automated Face Morph Detection. Technical report, National Institute of Standards and Technology, July 2021.

[129] State of the art morphing detection SOTAMD. https://www.ntnu.edu/iik/sotamd. Accessed: May 2020.

[130] International Civil Aviation Organization. Machine readable passports – part 9 – deployment of biometric identification and electronic storage of data in eMRTDs. http://www.icao.int/publications/Documents/9303_p9_cons_en.pdf, 2015.

[131] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017.

[132] R. Raghavendra, K. Raja, and C. Busch. Algorithmic fairness in face morphing attack detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 410–418, 2022.

[133] R. Raghavendra and C. Busch. Presentation attack detection methods for face recognition systems: A comprehensive survey. *ACM Comput. Surv.*, 50(1):1–37, 2017.

[134] D. Robertson, R. Kramer, and A. Burton. Fraudulent ID using face morphs: Experiments on human and automatic recognition. *Plos One*, March 2017.

[135] M. Ferrara, A. Franco, and D. Maltoni. On the effects of image alterations on face recognition accuracy. In *Face Recognition Across the Imaging Spectrum*, February 2016.

[136] T. Karras, S. Laine, and T. Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4401–4410, 2019.

[137] R. Abdal, Y. Qin, and P. Wonka. Image2StyleGAN: How to embed images into the StyleGAN latent space? In *Intl. Conf. on Computer Vision (ICCV)*. IEEE, October 2019.

[138] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila. Analyzing and improving the image quality of stylegan. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8107–8116, 2020.

[139] Photo for a Passport or Identity-card, Netherlands. https://www.netherlandsworldwide.nl/countries/iran/living-and-working/photo-for-a-passport-or-identity-card, 2020. Accessed: October 2020.

[140] M. Bichsel. Automatic interpolation and recognition of face images by morphing. In *Proc. of the Second Intl. Conf. on Automatic Face and Gesture Recognition*, pages 128–135. IEEE Comput. Soc. Press, 1996.

[141] J. Wu. Face recognition jammer using image morphing. *Dept. Elect. Comput. Eng., Boston Univ., Boston, MA, USA, Tech. Rep. ECE-2011*, 2011.

[142] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch. Face recognition systems under morphing attacks: A survey. *IEEEAccess*, 2019.

[143] Abrosoft fantamorph. FantaMorph,Abrasoft:http://www.fantamorph.com/, 2020. Accessed: May 2020.

[144] International Civil Aviation Organization. Machine readable passports – part 1 – introduction. http://www.icao.int/publications/Documents/9303_p1_cons_en.pdf, 2015.

[145] NIST. FRVT morph web site. https://pages.nist.gov/frvt/html/frvt_morph.html.

[146] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[147] O. Richter and R. Wattenhofer. Treeconnect: A sparse alternative to fully connected layers. In *2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 924–931. IEEE, 2018.

[148] R. Abdal, Y. Qin, and P. Wonka. Image2stylegan++: How to edit the embedded images? In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8293–8302, 2020.

[149] J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision*, pages 694–711. Springer, 2016.

[150] Z. Wang, E. P. Simoncelli, and A.C. Bovik. Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, volume 2, pages 1398–1402. IEEE, 2003.

[151] D P. Kingma and J. Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2015.

[152] DNP printer. http://dnpphoto.com/en-us/Products/Printers/DS820A, 2020. Accessed: October 2020.

[153] A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner. Faceforensics++: Learning to detect manipulated facial images. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1–11, 2019.

[154] A. Jain, P. Majumdar, R. Singh, and M. Vatsa. Detecting GANs and retouching based digital alterations via DAD - HCNN. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 672–673, 2020.

[155] P. Majumdar, A. Agarwal, R. Singh, and M. Vatsa. Evading face recognition via partial tampering of faces. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 11–20, 2019.

[156] A. Jain, P. Flynn, and A. Ross. *Handbook of Biometrics*. Springer, July 2007.

[157] B. Klare, B. Klein, E. Taborsky, A. Blanton, J. Cheney, et al. Pushing the frontiers of unconstrained face detection and recognition: IARPA Janus Benchmark A. In *Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 1931–1939. IEEE, June 2015.

[158] I. Kemelmacher-Shlizerman, S. M. Seitz, D. Miller, and E Brossard. The megaface benchmark: 1 million faces for recognition at scale. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4873–4882, 2016.

[159] J. Lu, V. E.Liong, X. Zhou, and J. Zhou. Learning compact binary face descriptor for face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(10):2041–2056, 2015.

[160] J. Lu, V. E. Liong, and J. Zhou. Simultaneous local binary feature learning and encoding for face recognition. In *Proceedings of the IEEE international conference on computer vision*, pages 3721–3729, 2015.

[161] J. Chen, V. M. Patel, L. Liu, V. Kellokumpu, G. Zhao, M. Pietikäinen, and R. Chellappa. Robust local features for remote face recognition. *Image and Vision Computing*, 64:34–46, 2017.

[162] A. Punnappurath, A. N. Rajagopalan, S. Taheri, R. Chellappa, and G. Seetharaman. Face recognition across non-uniform motion blur, illumination, and pose. *IEEE Transactions on image processing*, 24(7):2067–2082, 2015.

[163] R. Ranjan, V. M. Patel, and R. Chellappa. Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(1):121–135, 2017.

[164] U. Park, Y. Tong, and A.K Jain. Age-invariant face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 32(5):947–954, 2010.

[165] A. K. Jain, B. Klare, and U. Park. Face recognition: Some challenges in forensics. In *Face and Gesture 2011*, pages 726–733. IEEE, 2011.

[166] S. Marcel, M.S. Nixon, and S.Z. Li. *Handbook of biometric anti-spoofing*, volume 1. Springer, 2014.

[167] T. Y. Wang and A.Kumar. Recognizing human faces under disguise and makeup. In *2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*, pages 1–7. IEEE, 2016.

[168] A. Bharati, R. Singh, M. Vatsa, and K. W. Bowyer. Detecting facial retouching using supervised deep learning. *IEEE Transactions on Information Forensics and Security*, 11(9):1903–1913, 2016.

[169] N. Erdogmus and S. Marcel. Spoofing 2D face recognition systems with 3d masks. In *2013 International Conference of the BIOSIG Special Interest Group (BIOSIG)*, pages 1–8. IEEE, 2013.

[170] A. K. Jain, A. Ross, and U. Uludag. Biometric template security: Challenges and solutions. In *2005 13th European signal processing conference*, pages 1–4. IEEE, 2005.

[171] B. Biggio, L. Didaci, G. Fumera, and F. Roli. Poisoning attacks to compromise face templates. In *2013 International Conference on Biometrics (ICB)*, pages 1–7. IEEE, 2013.

[172] J. Galbally, C. McCool, J. Fierrez, S. Marcel, and J. Ortega-Garcia. On the vulnerability of face verification systems to hill-climbing attacks. *Pattern Recognition*, 43(3):1027–1038, March 2010.

[173] A. Jain, K. Nandakumar, and A. Ross. 50 years of biometric research: Accomplishments, challenges, and opportunities. *Pattern Recognition Letters*, 79:80–105, 2016.

[174] M. Ferrara, A. Franco, and D. Maltoni. Face demorphing in the presence of facial appearance variations. In *Proc. of the 26th European Signal Processing Conf. (EUSIPCO)*. IEEE, September 2018.

[175] L. Best-Rowden and A. K. Jain. Longitudinal study of automatic face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 40(1):148–162, 2017.

[176] Z. Li, U. Park, and A.K. Jain. A discriminative model for age invariant face recognition. *IEEE transactions on information forensics and security*, 6(3):1028–1037, 2011.

[177] D. Gong, Z. Li, D. Lin, J. Liu, and X. Tang. Hidden factor analysis for age invariant face recognition. In *Proceedings of the ieee international conference on computer vision*, pages 2872–2879, 2013.

[178] Y. Li, G. Wang, L. Nie, Q. Wang, and W. Tan. Distance metric optimization driven convolutional neural network for age invariant face recognition. *Pattern Recognition*, 75:51–62, 2018.

[179] J.P. Farkas, J. E. Pessa, B. Hubbard, and R.J. Rohrich. The science and theory behind facial aging. *Plastic and Reconstructive Surgery Global Open*, 1(1), 2013.

[180] R. B Shaw Jr, E.B. Katzel, P.F. Koltz, M. J. Yaremchuk, J.A. Girotto, D. M. Kahn, and H.N. Langstein. Aging of the facial skeleton: aesthetic implications and rejuvenation strategies. *Plastic and reconstructive surgery*, 127(1):374–383, 2011.

[181] G. Panis, A. Lanitis, N. Tsapatsoulis, and T.F. Cootes. Overview of research on facial ageing using the FG-NET ageing database. *Iet Biometrics*, 5(2):37–46, 2016.

[182] M. Erickson. Passport canada's citizen engagement process and the new 10-year epassport. 2013.

[183] Y. Zhang, S. Fang, Y. Xie, and T. Xu. Fake fingerprint detection based on wavelet analysis and local binary pattern. In *Biometric Recognition*, pages 191–198, 2014.

[184] D. Cai, X. He, and J. Han. Speed up kernel discriminant analysis. *The VLDB Journal—The International Journal on Very Large Data Bases*, 20(1):21–33, 2011.

[185] J. TA. Andrews, T. Tanay, and L.D. Griffin. Multiple-identity image attacks against face-based identity verification. *arXiv preprint arXiv:1906.08507*, 2019.

[186] Facial landmark based face morphing. Open CV. https://www.learnopencv.com/face-morph-using-opencv-cpp-python/.

[187] D. L. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, May 1995.

[188] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, Aug 2007.

[189] M. Zhang and B. K. Gunturk. Multiresolution bilateral filtering for image denoising. *IEEE Transactions on Image Processing*, 17(12):2324–2333, Dec 2008.

[190] F. Yu and V. Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015.

[191] Q. Chen, J. Xu, and V. Koltun. Fast image processing with fully-convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2497–2506, 2017.

[192] 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.

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

[194] R. Choras. Multimodal biometrics for person authentication. In *Digital Identity*. IntechOpen, 2019.

[195] J. Galbally, S. Marcel, and J. Fierrez. Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition. *IEEE Trans. on Image Processing*, 23(2):710–724, 2014.

[196] Shan Jia, Guodong Guo, and Zhengquan Xu. A survey on 3d mask presentation attack detection and countermeasures. *Pattern Recognition*, 98:107032, 2020.

[197] A. Anjos and S. Marcel. Counter-measures to photo attacks in face recognition: A public database and a baseline. In *2011 International Joint Conference on Biometrics (IJCB)*, pages 1–7, 2011.

[198] N. Erdogmus and S. Marcel. Spoofing face recognition with 3d masks. *IEEE Transactions on Information Forensics and Security*, 9(7):1084–1097, 2014.

[199] A. Costa-Pazo, S. Bhattacharjee, E. Vazquez-Fernandez, and S. Marcel. The replay-mobile face presentation-attack database. In *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–7, 2016.

[200] A. Agarwal, D. Yadav, N. Kohli, R. Singh, M. Vatsa, and A. Noore. Face presentation attack with latex masks in multispectral videos. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 275–283, 2017.

[201] K. Patel, H. Han, A. K. Jain, and G. Ott. Live face video vs. spoof face video: Use of moiré patterns to detect replay video attacks. In *2015 International Conference on Biometrics (ICB)*, pages 98–105, 2015.

[202] N. Evans, S. Z. Li, S. Marcel, and A. Ross. Guest editorial: Special issue on biometric spoofing and countermeasures. *IEEE Transactions on Information Forensics and Security*, 10(4):699–702, 2015.

[203] I. Chingovska, A.R. Dos Anjos, and S. Marcel. Biometrics evaluation under spoofing attacks. *IEEE transactions on Information Forensics and Security*, 9(12):2264–2276, 2014.

[204] Science Daily, Morphing. https://www.sciencedaily.com/terms/morphing.htm. Accessed: May 2020.

[205] Passport,Wikipedia. https://en.wikipedia.org/wiki/Passport#National_conditions. Accessed: May 2020.

[206] United States Visa. https://www.ustraveldocs.com/no/no-niv-photoinfo.asp, 2020. Accessed: October 2020.

[207] Secure access control over wide area network. https://www.ntnu.edu/iik/swan, 2020. Accessed: May 2020.

[208] Federal Ministry of Education and Research . https://www.bmbf.de/en/index.html. Accessed: May 2020.

[209] iMARS. https://cordis.europa.eu/project/id/883356, 2020. Accessed: October 2020.

[210] International conference on biometrics for borders: Morphing and morphing attack detection methods. https://frontex.europa.eu/research/invitations/international-conference-on-biometrics-for-borders-morphing-and-morph Accessed: May 2020.

[211] FRONTEX. Best practice technical guidelines for automated border control ABC systems, 2015.

[212] M. Ngan, P. Grother, K. Hanaoka, and J. Kuo. Face recognition vendor test (FRVT) part 4: Morph-performance of automated face morph detection. *National Institute of Technology (NIST), Tech. Rep. NISTIR*, 8292, 2020.

[213] D. Maio, D. Maltoni, R. Cappelli, A. Franco, and M. Ferrara. FVC-Ongoing–benchmark area: face morphing challenge. 2018.

[214] Morph thing. https://www.morphthing.com/, 2020. Accessed: October 2020.

[215] Face swap online. https://faceswaponline.com/, 2020. Accessed: October 2020.

[216] Face morpher. http://www.facemorpher.com/, 2020. Accessed: October 2020.

[217] Magic morph 1.95. https://downloads.tomsguide.com/magic-morph,0301-6817.html, 2020. Accessed: October 2020.

[218] Cartoon brew. https://www.cartoonbrew.com/vfx/10-unforgettable-morphs-film-tv-music-videos-144036.html. Accessed: May 2020.

[219] V. Blanz and T. Vetter. A morphable model for the synthesis of 3D faces. *Proc. of the SIGGRAPH99*, pages 187–194, 1999.

[220] V. Zanella, G. Ramirez, H. Vargas, and L.V. Rosas. Automatic morphing of face images. In Mikko Kolehmainen, Pekka Toivanen, and Bartlomiej Beliczynski, editors, *Adaptive and Natural Computing Algorithms*, pages 600–608, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[221] GNU image manipulation program (GIMP). https://www.gimp.org, 2016. Accessed: 2014-08-19.

[222] S-Y. Lee, K-Y. Chwa, S.Y. Shin, and G. Wolberg. Image metamorphosis using snakes and free-form deformations. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH 95, page 439–448, 1995.

[223] T. Ucicr. Feature-based image metamorphosis. *Computer graphics*, 26:35–42, 1992.

[224] M. Gomez-Barrero, C. Rathgeb, U. Scherhag, and C. Busch. Is your biometric system robust to morphing attacks? In *Proc. 5th Intl. Workshop on Biometrics and Forensics (IWBF)*, pages 1–6. IEEE, 2017.

[225] New face morphing database for vulnerability research. https://www.linkedin.com/pulse/new-face-morphing-dataset-vulnerability-research-ted-dunstone. Accessed: May 2020.

[226] Dlib programming library. http://dlib.net/, 2020. Accessed: October 2020.

[227] Bologna online evaluation platform (boep): Differential morph attack detection. https://biolab.csr.unibo.it/fvcongoing/UI/Form/BenchmarkAreas/BenchmarkAreaDMAD.aspx. Accessed: May 2020.

[228] D. White, R.I. Kemp, R. Jenkins, M. Matheson, and A.M. Burton. Passport officers' errors in face matching. *PloS one*, 9(8):e103510, 2014.

[229] A. Makrushin., T. Neubert., and J Dittmann. Humans vs. algorithms: Assessment of security risks posed by facial morphing to identity verification at border control. In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 4: VISAPP,*, pages 513–520. INSTICC, SciTePress, 2019.

[230] A. Makrushin, D. Siegel, and J. Dittmann. Simulation of border control in an ongoing web-based experiment for estimating morphing detection performance of humans. In *Proceedings of the 2020 ACM Workshop on Information Hiding and Multimedia Security*, IHMMSec '20, page 91–96, New York, NY, USA, 2020. Association for Computing Machinery.

[231] T. Jäger, Kerstin H Seiler, and A. Mecklinger. Picture database of morphed faces (mofa) : technical report. 2005.

[232] OCI services, india. https://ociservices.gov.in/Photo-Spec-FINAL.pdf. Accessed: May 2020.

[233] 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.

[234] S. Clemens, S.Wojciech, A. Hilsmann, and P. Eisert. Accurate and robust neural networks for face morphing attack detection. *Journal of Information Security and Applications*, 53:102526, 2020.

[235] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 19795-1:2006. Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*. International Organization for Standardization, March 2006.

[236] Official Journal of the European Union. REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN#d1e2513-1-1, 2016.

[237] A. Röttcher, U. Scherhag, and C. Busch. Finding the suitable doppelgänger for a face morphing attack. In *International Joint Conference on Biometrics (IJCB)*, pages 1–8, September 2020.

[238] N. Damer, A. Saladie, S. Zienert, Y. Wainakh, P. Terhoerst, et al. To detect or not to detect: The right faces to morph. In *2019 Intl. Conf. on Biometrics (ICB)*. IEEE, June 2019.

[239] C. Rathgeb, C. Satnoianu, N. Haryanto, K. Bernardo, and C. Busch. Differential detection of facial retouching: A multi-biometric approach. *IEEE Access*, 8:106373–106385, June 2020.

ISBN 978-82-326-6115-2 (printed ver.)
ISBN 978-82-326-5932-6 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU
Norwegian University of
Science and Technology