Doktoravhandlinger ved NTNU, 2024:202

Jag Mohan Singh

Robust algorithms for 2D and 3D Face Morphing Attacks: Generation and Detection

Doktoravhandling

NTNU
Norges teknisk-naturvitenskapelige universitet
Avhandling for graden
philosophiae doctor
Fakultet for informasjonsteknologi og
elektroteknikk
Institutt for informasjonssikkerhet og
kommunikasjonsteknologi

NTNU
Kunnskap for en bedre verden

Jag Mohan Singh

Robust algorithms for 2D and 3D Face Morphing Attacks: Generation and Detection

Thesis for the Degree of Philosophiae Doctor

Gjøvik, May 2024

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

ISBN 978-82-326-7996-6 (printed ver.)
ISBN 978-82-326-7995-9 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2024:202

Printed by NTNU Grafisk senter

This thesis is dedicated to my daughter and my wife.

Declaration of Authorship

I, Jag Mohan Singh, hereby declare that this thesis and the work presented in it are entirely my own. Where I have consulted the work of others, this is always clearly stated. Neither this nor a similar work has been presented to an examination committee elsewhere.

signature:

....

(Jag Mohan Singh)

Gjøvik, Date: 7th May 2024

Abstract

Biometric Authentication (Biometrics) is a powerful tool that authenticates individuals using digital means, which includes biological or behavioral characteristics. Biometrics harnesses biological features such as fingerprints, face, hand geometry, speech, iris, and fingerphoto. Face and finger modalities have generated the interest of biometric researchers thanks to their ease of use and high accuracy. Face biometric modality, in particular, is easy to use as it can be acquired passively. Furthermore, Face Recognition Systems (FRS) excel in real-world environments, thanks to the advancements in deep learning. However, it's important to know that FRS are not immune to attacks. They are vulnerable to various types of attacks, including presentation attacks and morphing attacks to a large extent and deepfakes to a smaller extent.

This thesis focuses on Face Morphing Attacks (FMA), an active area of research in Biometrics. An FMA can be generated by linearly blending facial images in the color domain from two contributory data subjects. FMA has shown vulnerabilities in FRS when evaluated automatically by software or manually by human observers. Thus, FMA is a strong attack on FRS. Hence, detecting FMA is an actual problem from a security standpoint. Most FMA systems currently use full facial images from the two contributory data subjects. However, the part-based face morphing/compositing problem has received little attention, i.e., using facial parts from the two contributory data subjects to generate an FMA. Further, due to Generative Adversarial Networks (GANs), generating full photo-real synthetic faces or completing partial facial images is possible due to deep learning-based image synthesis advances. Thus, part-based facial morphing using the advances of deep learning could be a fruitful area of research.

Motivated by the challenges arising from attacks toward FRS, the thesis focus is two-fold. The first is to increase the attack strength by generating higher quality attacks and the second is to advance the mitigation measures, a.k.a countermeasures for the generated attacks. We focus on Morphing Attacks, which include generation and detection, known as Morphing Attack Detection (MAD). Further, evaluating vulnerabilities imposed by part-based facial morphing could be a novel area of research and we have performed an extensive assessment of this nascent area. Currently, the critical problem is performing robust MAD in real-world environments, which have the challenges of facial pose, expression, illumination, image quality, print-scan variations and image capture distance. This brings us to building robust classifiers for the facial morphing problem. Morphing has been evaluated on face images, i.e., 2D image data. We generalize Morphing to 3D by performing first-of-its-kind 3D Morph operations on point clouds and present the results on both generation and detection. We generate a GAN-based facial composite of face images from face images of two contributory data subjects, with an extensive evaluation of different facial regions.

Acknowledgement

I wish to express my deepest gratitude to the individuals who provided me with exceptional support and guidance during my Ph.D. study at the Norwegian University of Science and Technology (NTNU), Gjøvik. Specifically, I would like to extend my appreciation to my thesis supervisor, Prof. Raghavendra Ramachandra, for his unwavering efforts and unwavering commitment to my academic success. I am grateful for the lessons he taught me about maintaining a strong work ethics and the importance of constant effort. I would also like to thank Prof. Raghavendra Ramachandra for his continuous support and guidance throughout the submission and acceptance of my journal articles and conference papers.

Additionally, I wish to extend my sincere thanks to my colleagues at the NTNU including Pankaj Wasnik, Hareesh Mandalapu, and Haoyu Zhang, for their insightful and enriching discussions. I am grateful for the valuable insights and perspectives they shared with me during my time at the NTNU. I would like to express my appreciation to all my colleagues and faculty members, who contributed to my academic journey at NTNU. I understand that I may have inadvertently omitted someone, and I apologize for any oversight. Nonetheless, I extend my heartfelt thanks to all who helped me during my time at NTNU.

Contents

List of Tables

List of Figures

List of Abbreviations

3DMAD	3D Morphing Attack Detection
ABC	Automated Border Control
APCER	Attack Presentation Classification Error Rate
BPCER	Bona fide Presentation Classification Error Rate
BSIF	Binarised Statistical Image Features
CFIA	Composite Face Image Attacks
CNN	Convolutional Neural Network
COTS	Commercial-Off-the-Shelf
CRC	Collaborative Representation Classifier
D-EER	Detection Equal Error Rate
D-MAD	Differential Morphing Attack Detection
DET	Detection Error Tradeoff
DFR	Deep Face Representation
eMRTD	electronic Machine Readable Transport
FMA	Face Morphing Attack
FMI	Face Morphing Image
FMIA	Face Morphing Image Attack

FMMPMR	Fully Mated Morphed Presentation Match Rate
FRGC	Face Recognition Grand Challenge
FRS	Face Recognition System
FRVT	Face Recognition Vendor Test
FTAR	Failure to Acquire Rate
GAN	Generative Adversarial Network
GDPR	General Data Protection Regulation
GMAP	Generalized Morphing Attack Potential
HOG	Histogram of Oriented Gradients
ICAO	International Civil Aviation Organization
KLD	Kullback-Liebler Divergence
L-SVM	Linear Support Vector Machine
LBP	Local Binary Pattern
LFW	Labelled Faces in the Wild
LPQ	Local Phase Quantization
MAD	Morphing Attack Detection
MAP	Morphing Attack Potential
MIPGAN	Morphing through Identity Prior driven GAN
MMPMR	Mated Morph Presentation Match Rate
MSCAN	Multiscale Context Aggregation Network
NIST	National Institute of Standards Technology
OTF	On-the-fly
P-CRC	Probabilistic Collaborative Representation Classifier
PA	Presentation Attack
PAD	Presentation Attack Detection

PAI	Presentation Attack Instrument
PSNR	Peak-Signal-to Noise Ratio
RD-MAD	Robust Differential Morphing Attack Detection
S-MAD	Single Morphing Attack Detection
SLERP	Spherical Linear Operator
SRKDA	Spectral Regression Kernel Discriminant Analysis
SSIM	Structural Similarity Index Measure
SVM	Support Vector Machine
SWAN	Secure Access Control over Wide Area Network

Part I

Overview

Chapter 1

Introduction

Biometric authentication (biometrics) is the process of authenticating a person through digital means. A broadly accepted definition of biometric authentication is "the automatic recognition of individuals based on distinguishing between biological and behavioral traits. This field is a subset of the broader field of human identification" [25]. Examples of biometric technologies include fingerprint recognition, facial recognition, hand geometry, speaker recognition, and iris recognition. The applications of these techniques include driver license authentication, searching for known card cheats in casinos, home incarceration programs and confidentiality of healthcare data [26].

The biometric system should avoid unauthorized access to the device and only allow a genuine user (bona fide), as it contains user-specific information. Recently, deep-learning techniques such as Facenet by Schroff et al. [27] have been used for high-quality face recognition to convert a face image into a 512-dimension feature vector using a deep convolutional network and face similarity is then computed using Euclidean distance. The Facenet achieved an accuracy of 99.63% on a widely used public dataset called Labeled Faces in the Wild (LFW). The Facenet was trained using 200 million images with around eight million unique identities. The large dataset limitations of Facenet were overcome in Deep Face Recognition by Parkhi et al. [28], who used a clever combination of automation and humans in the loop to create a dataset of 2.6 million images consisting of 2.62 thousand identities and achieved state-of-the-art (SOTA) results using deep learning.

However, FRS is vulnerable to face morphing attacks generated by blending face images from one or more contributory data subjects (best case is two). Thus, the face morphing image shows vulnerability towards all contributory data subjects (two is the usual case). It should be noted that face morphing images can deceive both human observers (border control guards) and software (automatic FRS).

Given the importance of face morphing, this thesis addresses its two aspects: generation and detection.
It must be mentioned that FRS is vulnerable to Presentation Attacks (PA), a.k.a spoofing attacks, which can be achieved by presenting a biometric artefact to the biometric capture device. PA can be performed by generating a Presentation Attack Instrument (PAI) that includes either a printed photo (print-photo), displaying an image (display-photo), displaying a video (replay-video), or the use of a rigid/non-rigid 3D face mask (mask-attack). Biometric researchers had thus devised Presentation Attack Detection (PAD) as a countermeasure to PA that is extensively discussed in [29], and [30]. Further, FRS is prone to deepfakes attacks, which are generated by either expression replacement or facial replacement initially performed using Generative Adversarial Networks (GANs), which have been proven a threat to FRS and thus need to be detected, as pointed out by Korshunov et al. [31].

1.1 Motivation and Problem Statement

The primary motivation of this thesis is to make Morphing Attack Detection (MAD) classifiers robust. The robustness of a MAD classifier can be defined as its ability to perform with a similar level of accuracy irrespective of changes in pose, illumination, expression, print scan artifacts, and capture distance. This is important for real-world applications of MAD. The secondary motivation was to use additional cues, such as depth, to improve the accuracy of the MAD classifier, unlike just color images in the primary motivation. In a typical scenario of automated border control (ABC), a trusted live capture image is verified against an image stored in an electronic machine-readable travel document (eMRTD or passport). The security risk becomes especially high as there is a digital upload of facial images for passports in several countries. Thus, a malicious user can upload a face-morphing image and a single eMRTD can verify as two subjects. This scenario violates the single-user, single-document rule. As a countermeasure to this security risk, the MAD classifier should be robust to pose, illumination, and expression, as a person standing in front of an ABC gate can have a face in an arbitrary pose, illumination, and expression due to real-world settings.
Another scenario that can pose a security risk is on-the-fly (OTF) capture, where surveillance cameras are mounted and a person walks in an aisle. A person can be at an arbitrary distance from the camera and has an arbitrary facial pose, expression, and illumination. Furthermore, in the OTF scenario, the trusted surveillance image must be compared with an enrollment image stored for that person. Thus, the OTF scenario is even more challenging than the ABC gate scenario. Hence, as a countermeasure to these challenges, the MAD classifier should be robust against face pose, illumination, expression, and capture distance.

Moreover, part-based (attribute-based) facial morphing images have received less attention in the literature on facial morph generation. However, due to the increasing visual fidelity of facial images generated by generative adversarial networks (GANs), facial attribute-based morphing images can be generated, showing higher vulnerability to FRS. Regarding security risk, a facial morphing image generated by a facial attribute-based morph poses a threat similar to that explained previously in both the ABC gate and the OTF scenarios encountered in real-world settings.
Finally, we consider a scenario in which face enrollment and the probe are in 3D. 3D face morph generation and detection would make the MAD classifier more robust in 3D than 2D, as a single 3D model can handle multiple facial poses, unlike 2D. It must be noted that 3D face morph generation and detection were considered for the first time in this thesis and we have done extensive state-of-the-art (SOTA) evaluation.

1.2 Research Objective

The research objective of this thesis is to evaluate and develop MAD algorithms and further evaluate the impact of 3D data on these algorithms. This is depicted as a block diagram in Figure 1.1, which shows the evaluated problems. The research objectives of this thesis are summarized as follows:

- Perform an extensive literature review of existing algorithms for MAD in general and with 3D information in particular.
- Evaluate and propose new algorithms for MAD, benchmarking these against SOTA on public datasets.
- Evaluate and propose an algorithm with high generalization capability for MAD.
- Evaluate the impact of high-quality ground truth depth data for the morphing problem.

1.3 Research Questions

The following research questions are formulated based on the study of literature, motivation, and thesis research objectives.

1.3.1 RQ1: Robustness of MAD Classifiers

RQ1 How can we improve the robustness of MAD classifiers in real-world environments that vary in the pose, expression, illumination, capture distance and image quality? (Related Chapters 5, 6, 7)

The MAD classifier's robustness is not just a theoretical concept but a practical necessity. It must demonstrate superior performance under various challenging conditions, such as pose, expression, illumination, probe capture distance and image quality. This is crucial for the MAD classifier's real-world application, particularly in the context of on-the-fly (OTF) capture scenarios. Moreover, it is essential to investigate which feature fusion or image alignment techniques could improve the classifier's performance in these challenging conditions that mimic real-world environments.

1.3.2 RQ2: Effect of postprocessing on MAD Classifier

RQ2 What is the effect of postprocessing morphing images on the performance of the MAD classifier? Furthermore, what is the impact on the generalization of the MAD classifier trained using different mediums in the presence of postprocessing morphing images? (Related Chapter 8)
FRS is susceptible to facial morphing images, and MAD methods have been employed to identify them. Despite minor artifacts, current state-of-the-art (SOTA) MAD methods rely on datasets that involve postprocessing in the nasal, oral, or ocular regions. This inspired us to investigate the impact of postprocessing on the performance of the MAD classifiers on a broader scale. Moreover, in real-world situations, MAD classifiers are expected to be applied in environments where the training and testing media comprise digital or varied-resolution printers and scanners. Therefore, we sought to examine the generalization of the MAD classifier in the context of postprocessing and diverse mediums.

1.3.3 RQ3: Generation of Facial Attribute-based Face Morphs

RQ3 How can we generate facial attribute-based face morphing that shows vulnerabilities of FRS, and are the current MAD methods suitable to detect them? (Related Chapter 9)
FRS has been demonstrated to be susceptible to facial attribute-based compositing using a non-deep learning-based method, where the entire face of one individual is combined with a single facial feature of another individual. However, the vulnerability of FRS to compositing using a few or multiple facial attributes has yet to be thoroughly examined. Thus, we aimed to create facial composites that display vulnerability in the FRS using GAN-based facial image synthesis and to rank these facial attributes according to their level of vulnerability.

1.3.4 RQ4: Generation of 3D Face Morph

RQ4 How can we generate 3D Face Morphing when ground-truth 3D data is available from the two contributory data subjects and does the generated 3D Face Morphing show vulnerabilities of FRS? (Related Chapter 10 and Chapter 11)

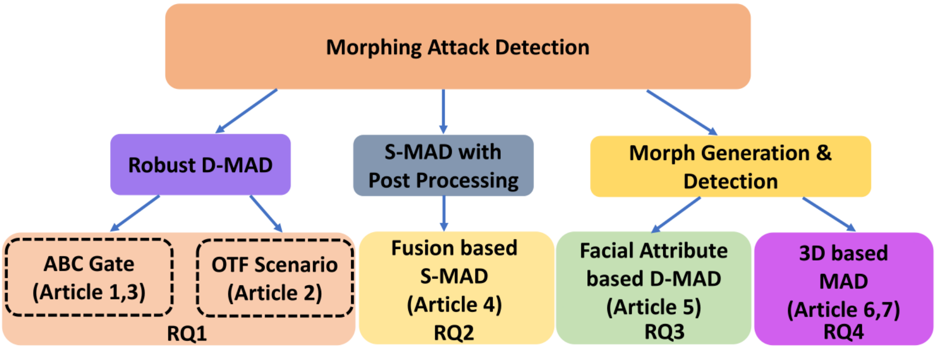

Morphing Attack Detection
Robust D-MAD
S-MAD with Post Processing
Morph Generation & Detection
ABC Gate (Article 1,3)
OTF Scenario (Article 2)
RQ1
Fusion based S-MAD (Article 4)
RQ2
Facial Attribute based D-MAD (Article 5)
RQ3
3D based MAD (Article 6,7)
RQ4

Figure 1.1: Illustration for Research Topics in Thesis

The Face Recognition System (FRS) is susceptible to the threat of 2D Face Morphing, created by combining facial images from two contributory data subjects. The widespread adoption of 3D sensors that can capture 3D face models in access control and their natural extension to 3D face recognition are the driving factors for the increased use of 3D face recognition. These factors have prompted us to investigate the problem of 3D Face Morphing, which is a novel form of attack. This research question aims to generate 3D face morphing, expose the vulnerabilities of FRS and develop methods for detecting such attacks.

1.4 Research Methodology

The thesis employs the research methodology outlined in this section to answer the research questions posed. Figure 1.1 provides a visual representation of the categorization of the research topics and associated research questions. The central theme of this thesis is the detection of morphing attacks, with the sub-topics encompassing the various types of data utilized. The systematic approach taken to tackle the research questions is summarized as follows:

- **Robust D-MAD Algorithms**
  MAD algorithms have demonstrated a high degree of accuracy when applied to individual public datasets. In particular, certain public datasets for MAD have shown near-perfect detection rates. However, it should be noted that these datasets are digital and do not contain print-scan artifacts, nor do they have feature variations in facial pose and capture distance, which are relevant factors in the context of an ABC Gate and OTF scenario. Therefore, we created datasets for these specific scenarios. A MAD algorithm is considered robust if it can perform effectively in these challenging conditions. Our proposed D-MAD algorithm was benchmarked against SOTA methods

and achieved superior performance. Further information on this section can be found in the relevant articles, which are included in Chapters 5, 6 and 7 where each chapter contains a single article.

- **S-MAD with Post Processing**
  MAD datasets have low postprocessing, but postprocessing of face-morphing images is expected in real-world environments. Thus, we generated a dataset with post-processed face-morphing images. We benchmarked SOTA on this dataset and proposed an algorithm that performed better than SOTA and is described in Chapter 8.

- **Facial Attribute-based Morphing**: Morphing is usually applied on full-face images from two contributory data subjects. The problem of facial-attribute-based morphing has received little attention. Thus, we worked on the problem of morphing using single/multiple facial attribute/s from the two contributory data subjects. This included generating a facial attribute-based dataset, evaluating its vulnerability towards FRS, and evaluating MAD algorithms on it. The details of this part are provided in Chapter 9.

- **3D Morphing**: In the existing literature, the generation of face-morphing images is typically accomplished via a linear blend of two-dimensional facial images obtained from two contributory data subjects. However, the next logical step in obtaining three-dimensional face morphing has yet to be explored. To address this gap, our work focused on the problem of generating three-dimensional face morphing from point clouds sourced from two contributory data subjects. To this end, we created a three-dimensional face dataset and proposed using the 3DMAD algorithm while assessing its vulnerability to FRS. Further details are outlined in Chapter 10 and Chapter 11.

1.5 List of Research Publications

1.5.1 List of Included Research Publications

1. Jag Mohan Singh, Raghavendra Ramachandra, Kiran B Raja and Christoph Busch. Robust morph-detection at automated border control gate using deep decomposed 3D shape & diffuse reflectance. In 2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), Sorrento (NA), Italy, pp 106-112. IEEE, doi=10.1109/SITIS.2019.00028.

2. Jag Mohan Singh, and Raghavendra Ramachandra. Reliable Face Morphing Attack Detection in On-The-Fly Border Control Scenario with Variation in

Image Resolution and Capture Distance, In IEEE International Joint Conference on Biometrics (IJCB 2022), Abu Dhabi, UAE, pp. 1-10, IEEE, doi=10.1109/IJCB54206.2022.10007987

3. Jag Mohan Singh, and Raghavendra Ramachandra. Fusion of Deep Features for Differential Face Morphing Attack Detection at Automatic Border Control Gates, In IEEE European Workshop on Visual Information Processing (EUVIP 2022), Lisbon, Portugal, pp. 1-5, IEEE, doi=10.1109/EUVIP53989.2022.9922773

4. Jag Mohan Singh, Sushma Venkatesh and Raghavendra Ramachandra.Robust Face Morphing Attack Detection Using Fusion of Multiple Features and Classification Techniques, In 26^{th} International Conference on Information Fusion 2023, South Carolina, USA, pp. 1-8, doi=10.23919/FUSION52260.2023.10224168

5. Jag Mohan Singh, and Raghavendra Ramachandra. Deep Face Attribute Composition Attacks: Generation, Vulnerability and Detection. In IEEE Access, March 2023. pp. 76468 - 76485, doi=10.1109/ACCESS.2023.3261247

6. Jag Mohan Singh, and Raghavendra Ramachandra. 3D Face Morphing Attacks: Generation, Vulnerability and Detection. In IEEE Transactions ON Biometrics, Behavior, AND Identity Science (IEEE T-BIOM) (Accepted, doi= 10.1109/TBIOM.2023.3324684).

7. Jag Mohan Singh and Raghavendra Ramachandra. 3D Face Morphing Attack Generation using Non-Rigid Registration. 18th IEEE International Conference on Automatic Face and Gesture Recognition, 2024.

1.5.2 List of Additional Research Publications

1. Raghavendra Ramachandra, Jag Mohan Singh, Sushma Krupa Venkatesh "Sound-Print: Generalised Face Presentation Attack Detection using Deep Representation of Sound Echoes" 2023 IEEE International Joint Conference on Biometrics (IJCB), Ljubljana, Slovenia pp 1-9.

2. Jag Mohan Singh, and Raghavendra Ramachandra. "DLDFD: Recurrence free 2D Convolution approach for Deep Fake Detection". 2022 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP). Online, pp. 568-574.

3. Jag Mohan Singh, Ahmad S. Madhun, Ahmed Mohammed Kedir and Raghavendra Ramachandra. "Smartphone Based Finger-Photo Verification Using

Siamese Network" 2022 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP). Online, pp. 553-559.

4. Sandip Purnapatra, Nic Smalt, Keivan Bahmani, Priyanka Das, David Yambay, Amir Mohammadi, Anjith George, Thirimachos Bourlai, Sébastien Marcel, Stephanie Schuckers, Meiling Fang, Naser Damer, Fadi Boutros, Arjan Kuijper, Alperen Kantarci, Basar Demir, Zafer Yildiz, Zabi Ghafoory, Hasan Dertli, Hazim Kemal Ekenel, Son Vu, Vassilis Christophides, Dashuang Liang, Guanghao Zhang, Zhanlong Hao, Junfu Liu, Yufeng Jin, Samo Liu, Samuel Huang, Salieri Kuei, Jag Mohan Singh and Raghavendra Ramachandra "Face Liveness Detection Competition (LivDet-Face) - 2021". 2021 IEEE International Joint Conference on Biometrics (IJCB). Shenzhen, China, pp. 1-10.

5. Jag Mohan Singh, Raghavendra Ramachandra, and Christoph Busch. "Hierarchical interpolation of imagenet features for cross-dataset presentation attack detection". 2020 3rd International Conference on Intelligent Technologies and Applications (INTAP). Gjøvik, Norway, pp. 203-214.

6. Jag Mohan Singh, Ahmed Madhun, Guoqiang Li, and Raghavendra Ramachandra. "A Survey on Unknown Presentation Attack Detection for Fingerprint". 2020 3rd International Conference on Intelligent Technologies and Applications (INTAP). Gjøvik, Norway, pp. 189-202.

7. Jag Mohan Singh, Raghavendra Ramachandra, and Patrick Bours. "Fusion of Texture and Optical Flow using Convolutional Neural Networks for Gender Classification in Videos". 2020 3rd International Conference on Intelligent Technologies and Applications (INTAP). Gjøvik, Norway, pp- 227-236.

8. Kiran Raja, Matteo Ferrara, Annalisa Franco, Luuk Spreeuwers, Illias Batskos, Florens de Wit Marta Gomez-Barrero, Ulrich Scherhag, Daniel Fischer, Sushma Venkatesh, Jag Mohan Singh, Guoqiang Li, Loïc Bergeron, Sergey Isadskiy, Raghavendra Ramachandra, Christian Rathgeb, Dinusha Frings, Uwe Seidel, Fons Knopjes, Raymond Veldhuis, Davide Maltoni, and Christoph Busch. "Morphing Attack Detection – Database, Evaluation Platform and Benchmarking". 2020 IEEE Transactions on Information Forensics and Security (TIFS). vol. 16, pp. 4336-4351.

9. Jag Mohan Singh, Sushma Venkatesh, Kiran B Raja, Raghavendra Ramachandra, and Christoph Busch. "Detecting Finger-Vein Presentation Attacks Using 3D Shape & Diffuse Reflectance Decomposition". 2019 15th

International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). Sorrento (NA), Italy, pp. 8-14.

10. Raghavendra Ramachandra, Jag Mohan Singh, Sushma Venkatesh, Kiran Raja, and Christoph Busch. "Face presentation attack detection using multi-classifier fusion of off-the-shelf deep features". International Conference on Computer Vision and Image Processing (CVIP) 2019. Jaipur, India, pp. 49-61.

11. Jag Mohan Singh, Pankaj Wasnik, and Raghavendra Ramachandra. "Hessian-based robust ray-tracing of implicit surfaces on GPU". SIGGRAPH Asia 2018 Technical Briefs. Tokyo, Japan, pp. 1-4.

1.6 Scope of Thesis

This thesis aims to develop reliable and robust MAD algorithms that can perform well in difficult situations. This thesis addresses the challenges involved in developing such algorithms based on the following constraints:

- The thesis collected datasets that addressed the scenarios of ABC Gate, OTF, facial image post-processing, faces with some or all facial attributes and faces as 3D point clouds. Since these areas need to be better handled by existing face morphing datasets, the thesis focussed on morphing attack generation in some of these dataset formats.

- The benchmarking of SOTA was done, and novel MAD algorithms were proposed, which achieved superior performance and increased robustness.

- The vulnerability of the proposed and SOTA methods towards FRS, both commercial-off-the-shelf (COTS) and deep-learning-based, was calculated. The metrics used for vulnerability calculation were from the literature, and new ones were proposed where required.

1.7 Thesis Outline

The thesis is divided into two parts, where Part I presents the introduction, research objective, research questions, research methodology and scope of thesis in Chapter 1, background and related work in Chapter 2, a summary of published articles in Chapter 3 and conclusions in Chapter 4. This is followed by Part II, which presents the research papers formatted for the thesis. Chapter 5 presents robust D-MAD where we take a trusted capture image from an Automated Border Control (ABC) Gate and compare it with an image on eMRTD. We decompose the image into a diffuse reconstructed color image and normal map for robust D-MAD. Chapter 6 presents robust D-MAD where we can capture images in an OTF scenario and perform robust D-MAD using a Hierarchical Spherical Linear Operator (SLERP). Chapter 7 presents robust D-MAD where we handle multiple lights utilizing the fusion of deep classifiers. Chapter 8 presents an analysis of D-MAD classifiers in the presence of post-processing face morphing images and offers a proposed method for D-MAD in the same scenario. Chapter 9 presents facial-attribute-based morphing generation, including an exhaustive analysis of vulnerability dependence on single/multiple facial attributes. Chapter 10 introduces a novel type of morph, 3D face Morphing generation, which uses input point clouds from two contributory data subjects, unlike 2D image data. Chapter 11 performs the 3D face morphing generation using two 3D facial point clouds directly in 3D. This is followed by Part3 III, which presents the future work in Chapter 12.

Chapter 2

Background & Related Work

FRS achieves high accuracy in real-world environments primarily owing to advancements in deep learning. Deep learning networks such as Facenet [27] and Arcface [22] have achieved high accuracy on public datasets, including labeled faces in the wild (LFW) [32] and other public datasets. However, the FRS is susceptible to direct and indirect attacks, revealing vulnerabilities. Among the attacks on FRS, morphing attacks, which involve blending two facial images from contributory data subjects, expose vulnerabilities in FRS, including automatic (based on software) and manual (based on human observers) attacks. Biometric researchers have developed countermeasures known as morphing attack detection (MAD) methods to address these vulnerabilities. These methods can be classified as single-image-based MAD (S-MAD), which requires only probe images, or differential-image-based MAD (D-MAD), which requires probe images and enrollment images. This chapter reviews the state-of-the-art works in the literature on facial morph generation and detection and discusses the different error metrics used in the literature for vulnerability analysis of morph generation methods. FRS is vulnerable to face spoofing attacks, also known as presentation attacks which can be performed by presenting a biometric artefact such as a printed photo (print-photo), displaying an image (display-photo), displaying a video (replay-video), or the use of a rigid/non-rigid 3D face mask (mask-attack) for which countermeasures have been defined by biometric researchers which are discussed briefly in this section. FRS is prone to deepfakes, which can be generated by expression swap, i.e., changing the expression of the source image by a target image or identity swap, where the genuine user's face is replaced by another person's face.

2.1 Presentation and Deepfake Attacks on FRS

2.1.1 Presentation Attacks

As mentioned earlier, presentation attacks (PA) performed by showing a biometric artefact to an automatic FRS are powerful attacks towards FRS and show its vulnerabilities. Thus, presentation attack detection (PAD) is essential and can be done by either usage of textural feature descriptors such as Local Binary Pattern (LBP) or through dynamic methods for videos based PA, which measure pulse, eye blinking, lip movement or head rotation [30]. Presentation attacks can be detected by CNNs, amongst which one is based on a deep tree network (DTN) with zero-shot face anti-spoofing method, which generalizes to 13 types of PAs [33]. Further, amongst the current challenges to PAD methods are robustness and generalization across PAD datasets with reduced bias, as pointed out in the survey by Shaheed et al. [34].

2.1.2 Deepfakes

Deepfake attacks which generated either by facial replacement which included DeepFakes [35] and FaceSwap [36] or expression replacement which included Face2Face [37] using computer graphics & visualization techniques and Neural-Textures [38] using Generative Adversarial Networks (GANs)) amongst the initial techniques. Further, with the advancement in visual fidelity of generation using GANs, newer and better architectures were used for each deepfake category. The reader is advised to look at a recent survey by Mirsky et al. [39] for newer GAN-based architectures for deepfake generation. Further, with advancements in visual fidelity due to the introduction of diffusion models and multimodal large language models (LLMs), higher quality deepfakes were being generated, as pointed out in a recent survey by Mubarak et al. [40]. Further, the generated deepfakes are strong attacks on FRS and need to be detected [31]. It must be mentioned that deepfake detection can be done using hand-crafted features based on eye-blink, disparities in lip movement, or irregularities in texture and lighting [40]. Convolutional neural networks (CNNs), either directly or in combination with long short-term memory (LSTMs), can be used for deepfake detection [40]. However, the challenges for the current deepfake detectors are performing well with blurry images, fast-moving objects, sophisticated deepfakes, or unseen data.

2.2 Facial Morph Generation

The generation of facial morphs can be accomplished through the utilization of full or partial facial images. Additionally, the methods of facial morph generation can be categorized into two types: landmark-based morph generation and deep-

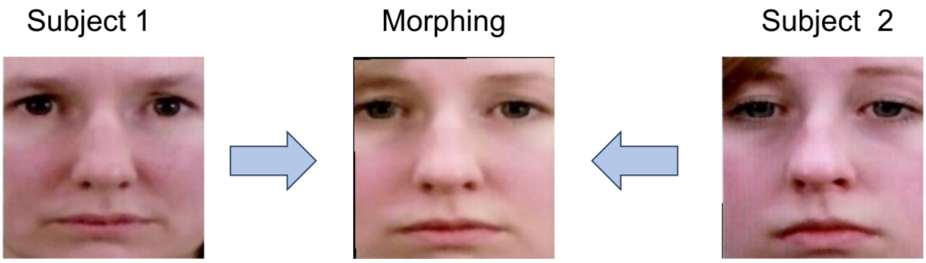


Figure 2.1: Illustration of Landmark based Morph from Morph ABC Database (Chapter 5)

learning-based morph generation. The following subsections will delve into the various facial morph-image generation techniques currently available.

2.2.1 Landmark-based facial morph generation

A landmark-based facial morph image generation consists of the following steps:

1. **Facial Alignment**: Since, the faces used are of frontal pose. The individual faces are aligned by making the eye corners horizontal.
2. **Landmark Specification and Correspondence Computation**: The different landmarks on a facial image can be specified manually or automatically (Active Shape Models [41] or more recently Dlib [42]). Once the landmarks are computed, the correspondences are established between the face images of two contributory data subjects.
3. **Delaunay Triangulation**: Once the correspondences are established, the point list of the face morphing image is generated by blending the point lists from the two contributory data subjects. Delaunay triangulation [43] of the point list of the face morphing image is performed.
4. **Warping Computation**: For each triangle of the face morphing image, an affine warping is computed between the triangle from the contributory data subjects to the triangle of the face morphing image. This results in the transformed face images from both contributory data subjects, when each triangle from both contributory data subjects is transformed.
5. **Blending Operation**: The transformed face images from two contributory data subjects are blended using the blending factor, resulting in the face morphing image which is the final result.

We now describe the different softwares used for the generation of face-morphing images in the following subsections:

Manual Landmarks: The software FotoMorph [44] and FantaMorph [45] were utilized to create face morphing images from two contributory data subjects. The landmarks were specified manually using these programs, with the number of landmarks ranging from 20 to 45 for FotoMorph and 127 for FantaMorph. Once the landmarks were specified, the software generated a morph animation, and the user was required to select the face morphing image with the minimum visible artifacts.
Automatic Landmarks: The softwares OpenCVMorph [46] and FaceMorpher [47] generate face morphing images by utilizing face images from two contributing data subjects. These images are processed to create landmarks, with 68 landmarks being automatically computed by OpenCVMorph and 77 by FaceMorpher. The process of generating the face morphing image is accomplished through the use of a script, which operates automatically.

2.2.2 Deep-Learning based facial morph image generation

In this section, we provide an overview of various deep learning-based techniques that have been employed for facial morph-image generation. Initially, GANs were utilized for this purpose, which was subsequently followed by image diffusion and transformer-based methods. In addition, a few hybrid methods have been developed that combine both GANs and Landmarks. It is important to acknowledge the developments in this field and are discussed in chronological order.

MorGAN: One of the earlier works in facial morph-image generation through deep learning is MorGAN [48]. This work employs GANs for generating facial morph-images and presents a dataset consisting of 1000 face-morphing images. Additionally, the authors conducted a vulnerability analysis of the FRS of the generated facial morphing images. In terms of detecting MorGAN attacks, it was found that LBPH outperformed CNN-based detectors. However, it is worth noting that the MorGAN generated facial morphs had a resolution of 64×64 and the image quality was slightly inferior to that of LMA morphs.
StyleGAN-based Morph: One of the early works on StyleGAN-based morph at high quality and high resolution of 1024×1024 was by Venkatesh et al. [49]. They generated a dataset of 2500 face morphing images. The proposed method first generates StyleGAN-based latents for the images from two contributory data subjects based on perceptual loss. They averaged the latents obtained from two contributory data subjects and passed them through the StyleGAN synthesis network for face morphing image generation. The authors benchmarked both GAN-generated and landmark-based morphs using established MAD methods.
Regenmorph: This method [50] works by combining LMA and GAN-based morphs. The primary motivation of Regenmorph is to increase the vulnerability towards

FRS of the generated facial morphing images. Owing to the weaknesses of MorGAN, which could not generate facial morphing images with high vulnerability towards FRS. Regenmorph generates the face image morphing by LMA and eliminates the artifacts using GANs. Thus, the method generates better quality facial morphing images when compared with MorGAN.

MIPGAN: The quality of the face morphing image was further improved in MIPGAN [51], which generated a face morphing image of 1024×1024 resolution. Further, MIPGAN improved the vulnerability towards FRS. MIPGAN generated the face morphing image by averaging StyleGAN2-based latents from both contributory data subjects, which was followed by passing it through the synthesis network of StyleGAN [12] and StyleGAN2 [52] to generate initial face morphing image. This initial face morphing image is optimized using a loss function based on identity, amongst other factors, to generate the final face morphing image.

StyleGAN2 based synthetic morphs: Sarkar et al. [53] performed an extensive evaluation of LMA and GAN-based morphs, including modified MIPGAN-II where they used pre-trained VGGFace model with Resnet50 backbone as feature extractor for identity loss instead of Resnet50 backbone and Arcface loss proposed in MIPGAN. They provided two facial morphing datasets and the first is based on landmark-based morphing, including OpenCV and FaceMorpher. The second dataset they offered includes two StyleGAN2-based methods, which consisted of synthetic morphs. The authors computed vulnerabilities of the generated facial morphs using the following FRS: Arcface [54], VGG-Face [55], Facenet [56] and ISV [57]. They concluded that LMA-based morphs show vulnerabilities towards FRS, and GAN-based morphs don't show vulnerabilities towards FRS.

Landmark Enforcement for Generative Morphing: Price et al. [58] proposed landmark enforcement, which is used to compute warped faces from both contributory data subjects to the average of individual landmarks. They further devised an approach for landmark-based loss based on StyleGAN2 for geometric constraints and explored principal component analysis (PCA) in latent space for reducing identity loss in morph generation. Further, to enhance the high-frequency component in the facial morphs they study the training of noise input for StyleGAN2.

MorDIFF: This method [59] uses diffusion autoencoders to generate the face morphing image. They performed extensive vulnerability analysis of the generated facial morphing images towards FRS. Further, MorDIFF generates more vulnerable face morphing images towards FRS than GANs. The facial morphing images generated by MorDIFF are challenging to detect, as indicated by their vulnerability analysis. The detection of MorDIFF attacks improves when training is performed on the synthetic dataset SMDD [60].

MorphGANFormer: This method [61] proposed a transformer-based alternative to generate face morphing images compared to GANs. They use four specifically designed loss functions (landmarks, biometric loss, perceptual loss and pixel-wise mean squared error) to increase the similarity between the generated facial morph images and images from both contributory data subjects. They proposed a transformer-based demorphing technique as an effective defense strategy to complement their morph generation. It needs to be pointed out that the demorphing technique can be compared to spectral unmixing of hyperspectral images, but it operates in latent-space rather than pixel-space when compared to spectral unmixing.

WALI: This method [62] generates worst-case (difficult) morphs for FRS by using concepts from Adversarially Learned Inference (ALI) and Wasserstein GANs trained with gradient penalty. It needs to be pointed out that WALI uses these methods to improve stability during training. They finetune WALI with loss functions with a specific ability to manipulate identity, resulting in more challenging morphs than landmark-based or GAN-based face morphing images. WALI has improved the quality of facial morphing images generated using MIPGAN.

Extswap: This method [63] observes that the quality of generated facial morphing images using GANs suffers due to an entangled representation. Thus, it generates SOTA face swapping by disentangling identity and attribute features in latent space, resulting in rich semantic latent space. Further, the authors performed extensive experiments that prove that their method successfully disentangles identity and attribute features. The experiments conducted by the authors during this article are of both quantitative and qualitative nature.

Approximating Optimal Generative Morphing: Colbois et al. [64] generated the optimal face morphing image by inverting the images from both contributory data subjects to generate their latent embeddings. The latent embeddings are then averaged to generate the face morphing image latent, passed through an image synthesis network to generate the face morphing image. They created the facial morphing images using several source datasets and study the effectiveness of those attacks using several FRS. Further, their method is competitive against previous deep-learning-based approaches in both black-box and white-box scenarios.

Optimal Landmark Guided Image Blending: He et al. [65] proposed to overcome the lack of identity features in GAN-based face morphing image generation by optimizing landmarks and using Graph Convolution Networks (GCNs) for combining landmark and appearance features. The authors model the landmarks as nodes in the bipartite graph, which is fully connected, and utilize GCNs to model their spatial and structural relationships. They performed extensive experiments on two public datasets showcasing the advantages of previous landmark-based and generation-based methods and generating higher-quality face morphing images, making them more vulnerable to FRS.

Summary: It needs to be pointed out that in terms of vulnerability of the generated face morphing images towards FRS, Landmark-based methods are more vulnerable than Deep-Learning-based methods [53, 48]. Landmark-based methods generated by facial morphing images have higher vulnerability towards FRS than deep-learning-based methods as they preserve the identity features more than those generated from deep-learning-based methods. Vulnerability towards FRS is affected by print-scan artifacts as they can induce quality degradation in the facial morphing images, making them less vulnerable towards FRS. Vulnerability towards FRS can be affected by the FRS being used to calculate the vulnerability as different FRS show different vulnerabilities.

2.2.3 3D Morph Generation

2D+3D Morph Generation: Liu et al. [66] formulated a 2D+3D approach for facial morph generation. The authors first detected the faces in both input facial images using Viola-Jones face detector [67], which is followed by facial interest point detection using Ramanan method [68], and the key-points are triangulated using Delaunay. Then, a forward warping is computed to the fusion (mean) vertices from both the Delaunay meshes, and the pixels inside these triangles are blended to generate the 2D blend texture. The face images are then projected to 3D using 3D Morphable Model [18] and a scaled orthographic projection (SOP(s,R,t)) is used to align the input face images and the two individual 3DMM models which results in parameters (s1,R1,t1) and (s2,R2,t2). Finally, the average 3DMM model is generated and the 2D blend texture is aligned with using mean of parameters $(\frac{s1+s2}{2}, \frac{R1+R2}{2}, \frac{t1+t2}{2})$ to generate the textured morph face. The technique would have the merit as the generated textured morphed face model would 2D identity facial features from both the contributory data subjects resulting in high vulnerability for 2D FRS. However, since the 3D face model generated using 3DMM is not physically accurate, it should show low vulnerability toward 3D FRS.

2.3 Facial Morph Detection

Morphing Attacks or facial morphing images can be used to expose vulnerabilities in FRS. Biometrics researchers have proposed morph detection as a countermeasure. Morph detection can be done in two ways: first, by using a pair of images to perform the detection, which is known as Differential Morphing Attack Detection (D-MAD), or the second way, where morph attack detection is done based on a single image where it is known as No Reference Morphing Attack Detection (NR-MAD) or Single Morphing Attack Detection (S-MAD). We want to bring the reader's attention to a survey article on face morphing [69] where authors have reviewed the current literature for S-MAD and D-MAD. According to the survey, S-MAD can be based on one of the following: 1) Texture features, 2)

Quality features, 3) Hybrid features, 4) Residual noise, or 5) Deep-learning. Further, according to the survey, D-MAD can be based on 1) Feature Difference or 2) Demorphing. The feature difference can be texture-based (LBP, BSIF, or HOG features) or deep-learning-based (Arcface or Alexnet features). Demorphing can be done using either Landmarks or Deep-Learning (GANs). For further details, the interested reader can go through the survey. We describe the literature for S-MAD and D-MAD, which are more recent than this survey, in the following subsections:

2.3.1 S-MAD

Ensemble-based Morph Detection: Kashiani et al. [70] the authors proposed the usage of ensemble-based morph detection to improve the generalization to a wide range of morphing attacks and high robustness towards adversarial attacks. The authors proposed combining convolutional neural networks (CNNs) and transformer models to take advantage of their capabilities. Further, for robustness, they employed multi-perturbation adversarial training and generated adversarial examples with high transferability for several single models. The proposed robust ensemble model performed better than SOTA, as shown in their exhaustive evaluation over several morphing attacks and face datasets.

Attention-based Morph Detection: Aghdaie et al. [71] proposed an end-to-end attention-based deep morph detector which incorporates the most-discriminative wavelet sub-bands of the input image obtained by group sparsity representation learning scheme. The most discriminative wavelet sub-bands (channels) are obtained using the attention mechanism. They employed three different attention mechanisms for this: the Convolutional Block Attention Module, the compatibility scores across spatial locations and output of their DNN highlighting the most discriminative regions, and the multi-headed self-attention augmentations. The authors evaluated their proposed approach on several morph datasets and achieved lower detection error rates than SOTA algorithms.

MAD-DDPM: This method [72] proposed a diffusion-based MAD method that only learns from characteristics of the bona fide images. They detect different forms of morphing attacks using their model as out-of-distribution samples. They performed extensive experiments over the following four datasets: 1) CASIA-WebFace, 2) FRLL-Morphs, 3) FERET-Morphs, and 4) FRGC-Morphs and compared their proposed method with discriminatively trained one-class models. The experiments done by the authors show that their proposed model performs competitively on all four datasets used in the paper.

PCA of texture patterns based Morph Detection: Dargaud et al. [73] performs S-MAD using RGB decomposition based on PCA of texture patterns. The authors mention that their method has increased explainability compared to deep-learning-based methods, as showcased by visualization of several relevant face areas used for morph detection. The authors extensively evaluated their approach in single, cross-dataset, and cross-morphed scenarios and compared it with fine-tuned MobileNetV2 architecture. The results of the evaluation show that S-MAD is challenging in cross-domain scenarios involving a wide range of morphing algorithms. The proposed method by the authors can be good or even better than the MobileNetV2 approach in the cross-domain scenarios.

MorDeephy: This method [74] performs S-MAD by learning deep facial features that carry the information about the authenticity of the features. The authors achieve this by training two backbone CNNs using bona fide and face morphing images, respectively. A single input image is passed (bona fide/face morphing) whose features are extracted from these networks and the dot product of these is used for classification. The authors also provide a public and easy-to-use face morphing detection benchmark. They achieved SOTA performance and generalized the task of face morphing detection to unseen scenarios.

Learning Residuals for Morphing Detection: Raja et al. [75] proposed an approach for S-MAD based on learning the residuals of the morphing process using an end-to-end multi-stage encoder-decoder network. The authors use cross-entropy and asymmetric losses to train their proposed network. The authors perform extensive experiments on two landmark-based and three GAN-based morphs in digital, print-scan, and print-scan compression settings. The authors achieve a near-perfect D-EER of 0% for the best case and 2.58% for the worst case in a digital domain for closed set protocol. Using three complementary Class Activation Maps (CAM) analysis methods, note the authors use CAMs for explainability.

Incremental Training for Morphing Detection: Borghi et al. [76] perform incremental training of morphing detectors motivated by that fact that MAD methods based on single images are not ready for real-world deployment due to low accuracy and generalization to datasets which are different from training set. Further, it is difficult to share datasets amongst different research groups due to privacy issues. This motivated the authors to devise an approach based on model sharing instead of data sharing which incorporates the strategies of Continual Learning. The authors proposed and released a framework based on incremental training of MADs using new data progressively which is available during different times and places. The authors use the paradigms of Learning without Forgetting (LwF) and Elastic Weight Consolidation (EWC) for Continual Learning.

Continual Learning for Morphing Detection: Pellegrini et al. [77] perform a continual learning (CL) paradigm for enabling incremental training for MAD. An underlying assumption of CL is that old data is no longer required and can be deleted. Thus, the authors explore the CL model in the scenario where the learning model is updated every time a new chunk of data is available, which can even be of variable size. The authors mention that the Learning without Forgetting (LwF) method is one of the best-performing CL methods for this scenario. Thus, the authors investigated its usage and parameterization in MAD and object classification.

IDistill: This method [78] proposed an interpretable identity distillation method that provides information on both identity separation of face morphing samples and their contribution to the final prediction. The authors learn domain information by an autoencoder and distill it into a classifier system for teaching its separate identity. The method proposed by the authors outperforms SOTA on three out of five databases used in the paper.

2.3.2 D-MAD

Feature-wise Supervision based Morphing Detection: Qin et al. [79] detects and localizes Morphing Attacks (MAs) using feature-wise supervision. The authors constructed fine-grained classification loss based on different morphing patterns and designed similarity-based and distance-based losses using the properties of D-MAD scenarios. The experimentation conducted by the authors shows that fine-grained classification loss can be used to locate the MA, whereas the D-MAD-based losses can improve the generalization capability towards unseen MAs.

Wavelet Scatter Network based Morphing Detection: Ramchandra et al. [80] performs morphing detection for face images of newborns as this task is important from the viewpoints of both security and society. The authors propose a two-layer wavelet scatter network (WSN) using 250×250 pixels with 6 rotations of wavelets per layer, resulting in 577 paths. The authors evaluate their method on 852 bona fide images and 2460 face morphing images from 42 unique newborns. The authors achieved a 10% gain in detection accuracy over existing D-MAD methods.

Fusion of Demorphing and Deep Face Representations for Morph Detection: Shiqerukaj et al. [81] have performed D-MAD using a fusion of two techniques, Demorphing and Deep Face Representations. It should be mentioned that the demorphed image is passed through FRS to obtain the score. The authors performed experiments in a cross-database scenario using high-quality facial image morphs and live bona fide captures. The authors mention that a weighted sum-based score fusion of Demorphing and Deep Face Representations results in better MAD. The authors obtained a D-EER of 4.9% compared with 5.6%, 5.8% of SOTA.

Combining Identity and Artifact Features for Morph Detection: Domenico et al. [82] mention that D-MAD approaches are based on the identity features of the face images. Thus, current D-MAD approaches could improve the images of look-alikes with similar identities. On the other hand, S-MAD approaches are based on the analysis of artifacts of the input images. This motivated the authors to fuse identity and artifact features for robust D-MAD.

2.4 Performance Metrics

We discuss the different performance metrics used for the evaluation of MAD algorithms in the biometrics literature as follows:

Attack Presentation Classification Error Rate (APCER): This metric measures the misclassification of face morphing attacks as bona fide presentations with an ideal value of 0% implying no misclassification.

Bona fide Presentation Classification Error Rate (BPCER): This metric measures the misclassification of bona fide presentations as face morphing attacks with an ideal value of 0% implying no misclassification.

Mated Morphed Presentation Match Rate (MMPMR): MMPMR is defined as the number of face morphing samples that can be verified by all the contributing data subjects by the FRS. However, MMPMR does not factor the effect of the number of attempts made into the metric computation. MMPMR is described by the following equation, which is taken from [83]

$$\text{MMPMR}(\tau) = \frac{1}{M} \sum_{m=1}^{M} \left\{ \left[\text{min}_{1 \cdots \text{N}_m} S_n^m \right] > \tau \right\} \tag{2.1}$$

where τ is the verification threshold, S_n^m is the similarity score when subject n is compared with morph m, N_m is the number of subjects used for generating morph m and M is the total number of face morphing images.

Fully Mated Morphed Presentation Match Rate (FMMPMR): FMMPMR is defined as the number of face morphing samples that the FRS can verify by all the contributing data subjects across all the attempts. Thus, it alleviates the weakness that is present in MMPMR. FMMPMR is described by the following equation, which is taken from [84]

$$FFMPR = \frac{1}{P} \sum_{M,P} (S1_M^P > \tau)(\&\&)(S2_M^P > \tau) \cdots (\&\&)(Sk_M^P > \tau) \tag{2.2}$$

where P is the number of probe images, M is the number of face morphing images, τ is the verification threshold Sk_M^P is the similarity score of k^{th} generated from comparison with M face morphing image with P^{th} attempt (or probe) image.

Morphing Attack Potential (MAP): MAP is the vulnerability metric considering multiple FRS, probe attempts and face morphing images to arrive at a matrix of vulnerabilities. MAP is described by the following equations, which are taken from [85] and part of the ISO/IEC CD 20059 [86]:

$$mc(M, \mathbb{P}, F) = |P_i \in \mathbb{P} : s_F(M, P_i) > \tau(F)| \tag{2.3}$$

where $mc(M, \mathbb{P}, F)$ returns the number of $\mathbb{P}$ probe images which are successfully verified against face morphing image (M) based on FRS (F). Further, $s_F(M, P)$

is the similarity score between face morphing image (M) and probe image (P) based on pre-fixed threshold $\tau(F)$.
Then, the authors define $fmc(M, \mathbb{P}, \mathbb{F}, r)$ as the number of FRSs in $\mathbb{F}$ for which at least r probe images in $\mathbb{P}$ are verified against the face morphing image M based on the following equation which is taken from [85]

$$fmc(M, \mathbb{P}, \mathbb{F}, r) = |F_i \in \mathbb{F} : (M, \mathbb{P}, F_i) \geq r| \tag{2.4}$$

Finally, the authors define the proposed metric MAP$[r, c]$ as the proportion of face morphing images in M for which $fmc(M, \mathbb{P}, \mathbb{F}, r) \geq c$ for both contributory data subjects. MAP$[r, c]$ is defined by the following equation taken from [85]

$$C_{MAP[r,c]}(M) = \bigwedge \begin{array}{l} fmc(M, \mathbb{P}_1, \mathbb{F}, r) \geq c \\ fmc(M, \mathbb{P}_2, \mathbb{F}, r) \geq c \end{array} \tag{2.5}$$

where $\mathbb{P}_1$ and $\mathbb{P}_2$ are the set of probe images for subjects 1 and 2, respectively.
Generalized Morphing Attack Potential (G-MAP): G-MAP is the most generic vulnerability metric, which generates a single vulnerability number considering multiple probe attempts, multiple FRSs, morph attack generation type, and face morphing images. This metric was proposed in the scope of this thesis and alleviated the weaknesses of previous metrics of MMPMR/FMMPMR/MAP by being the most generic and generating a single value (irrespective of attempts), indicating the attack potential of morphing images. It was introduced in Chapter 9 and presented in Equation 11.4.

2.5 Evaluation of MAD algorithms

Face Analysis Technology Evaluation (FATE), NIST IR 8292 NIST provides an evaluation platform for benchmarking MAD algorithms. The National Institute of Standards Technology (NIST) technical report [87] details this platform. The report's authors observed that S-MAD algorithms have reduced morph miss rates at a false detection rate of 0.01 in the submitted algorithms for both low-quality and automated datasets. However, they fail to generalize well across different unseen morphing methods. Further, the report mentions that submitted algorithms have shown promising improvements for the D-MAD algorithm and achieved a morph miss rate ranging between 9% and 36% at a false detection rate of 0.01. The authors pointed out that the better generalization of D-MAD could be attributed to using identity information of image and live probe photo, not the morphing artifacts.

Chapter 3

Summary of Published Articles

This chapter presents the summary of research articles included in this thesis. The articles included the generation of face morphing attacks and their detection using 2D and 3D data. The articles are summarised in the following sections:

3.1 Article 1: Robust Morph-Detection at Automated Border Control Gate using Deep Decomposed 3D Shape & Diffuse Reflectance (RQ1)

This article presented robust D-MAD in the ABC gate scenario where the bona fide image is taken from the trusted live capture (ABC gate) and is verified against the face image on the passport or electronic Machine Readable Travel Document (eMRTD). The proposed method decomposed both the bona fide and face images from eMRTD into diffuse reconstructed images and a normal map. The proposed method extracts Alexnet (fc7) features from the diffuse reconstructed image and quantization features (21 bits) from the normal map. The features are then passed through Linear SVM, whose scores are fused by the weighted sum rule to achieve the final score for a single camera. Further, multiple cameras are used within an ABC gate, whose individual scores are fused by the weighted sum rule for final classification. We created a morph attack database with 588 images, where bona fide images are captured in an indoor lighting environment using a Canon DSLR Camera, and the morphed and bona fide passport images are printed and scanned using an EPSON XP-860 printer and scanner. The EPSON XP-860 Printer and Scanner, which is used for scanning the attack images, is done with 300 dpi at an image resolution of 256 × 256 for detected faces. The proposed method significantly outperforms SOTA on the created dataset.

3.2 Article 2: Reliable Face Morphing Attack Detection in On-The-Fly Border Control Scenario with Variation in Image Resolution and Capture Distance (RQ1)

This article presented robust D-MAD in the OTF ABC gate scenario where the bona fide image is obtained from the ABC gate and the verification image is taken from the passport or eMRTD. The proposed method is based on the spherical linear interpolation (SLERP) and hierarchical fusion of deep features obtained from six pre-trained deep networks. The proposed method computes difference (residue) features from VGG19, Alexnet, and VGG16 in the first group and Xception, Resnet101 and Resnet50 in the second group. All the difference features are passed through individual linear SVMs to obtain six classification scores. Further, we generate two optimal pairs from each group. The residue features are SLERP interpolated, whose L1 difference is taken and passed through linear SVM to obtain two additional classification scores. Then, the eight classification scores are fused using the sum rule to get the final classification score. It must be pointed out that in the OTF scenario, both the camera resolutions and capture distances vary. We created the SCFace-Morph dataset based on selected 77 subjects from the SCFace dataset, which models the real-life scenario of ABC gates. The proposed method is extensively evaluated using three different protocols. The first protocol is designed to benchmark the performance impact of morph medium (digital or print-scan) based on camera resolution and capture distance. The second protocol benchmarks the performance impact of morph medium irrespective of camera resolution and capture distance. The third protocol benchmarks the performance is based on the camera resolution and capture distance regardless of the morph medium. We obtained superior and competitive performance across all three protocols.

3.3 Article 3: Fusion of Deep Features for Differential Face Morphing Attack Detection at Automatic Border Control Gates (RQ1)

This article presented robust D-MAD in the ABC gate scenario where trusted live capture is used as a bona fide image and eMRTD image is the probe image. The proposed method first performs pair-wise image alignment using the global affine transform computed between the two input images. This is followed by feature extraction using two deep networks, Resnet50 and Alexnet, which are passed through two linear SVMs, resulting in classification scores for each of the four cameras and three lights. We then performed a weighted fusion of scores for all the cameras from each light. The proposed performance was compared against existing SOTA with fusion and we achieved superior results where we obtained the lowest

EER=2.1% compared to the SOTA with an EER=8.6%. We generated the Morph ABC dataset based on 39 subjects with 270 face morphing images and 1549 ABC gate probe images. The first light in the dataset models a dark overcast day (180 lux), the second light models a sunrise/sunset (450 lux) and the third light models a bright day (1500 lux) to mimic real-world scenarios.

3.4 Article 4: Robust Face Morphing Attack Detection Using Fusion of Multiple Features and Classification Techniques (RQ2)

This article analyzes the effects of post-processing on face morphing images at a broader scale by introducing a new dataset of 10710 facial images before and after processing to reduce visual artefacts and generate high-quality attacks. When generated without artefacts, It must be pointed out that face morphing images can deceive both automatic FRS and human observers (border control guards). Further, the current morphing software generates ghosting artefacts, especially in the eye, nose, and mouth regions. Moreover, we proposed a novel S-MAD classifier based on an ensemble of features and classifiers. The proposed method first converts the input RGB color image into YCbCr and HSV color spaces. This is followed by decomposition into the Laplacian Pyramid. Then, in the generated scale space, multiple features, which include Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG) and Binarized Statistical Image Features (BSIF), and various classifiers, which include Support Vector Machine (SVM), Spectral Regression Kernel Discriminant Analysis (SRKDA) and Probabilistic Collaborative Representation Classifier (P-CRC) are used. Finally, two levels of hierarchical fusion are performed to make the final decision. We carried out extensive experiments on the dataset before and after post-processing. We carried out extensive experiments in two mediums: (a) Digital and (b) Print-Scan (with and without compression). Our results indicated the superior performance of the proposed S-MAD over existing SOTA present in the literature.

3.5 Article 5: Deep Face Attribute Composition Attacks: Generation, Vulnerability and Detection (RQ3)

This article proposed a novel method to generate Composite Face Image Attacks (CFIA) based on single/multiple facial attributes utilizing Generative Adversarial Networks (GANs). The proposed method first segments the two bona fide face images independently into segmented face attributes. The selected face attributes from the corresponding face images are transparently blended to generate the initial composite and face mask. This is then passed through an encoder (Resnet-34) and decoder (StyleGAN) to generate CFIA (Face Composite). We generated 526 unique CFIA samples for two contributory data subjects. We generated a dataset of 1000 individual identities, resulting in 526000 CFIA samples and 2000 unique bona fide samples. We benchmarked the attack potential of generated CFIA samples based on four automatic FRS. We introduced a new metric named Generalized Morphing Attack Potential (G-MAP) for benchmarking the generated CFIA samples. Further, we performed a human observer study and perceptual quality on a subset of the CFIA dataset. Finally, we benchmarked CFIA detection performance using three S-MAD algorithms. The dataset and source code of the proposed method were made publicly available at https://github.com/jagmohaniiit/LatentCompositionCode.

3.6 Article 6: 3D Face Morphing Attacks: Generation, Vulnerability and Detection (RQ4)

This article proposed a novel 3D face morphing method that works on point clouds in a 3D-2D-3D way, unlike previous 2D morphing approaches. The proposed method first projects the point clouds from two contributory data subjects to color images and depth maps using a single canonical view. The proposed method then computes the locally affine warping between the generated color images. It then uses the same local affine warping to transform the depth maps. The generated face morphing color image and depth map are then back-projected using the single canonical view to create the face morphing point cloud. However, the generated face morphing point shows holes from viewpoints other than the canonical viewpoint. Thus, we proposed a hole filling using multiple translated views (color images and depth maps) from the canonical view to fill the holes by using image inpainting for the color image and depth maps. We then register the generated color images with respect to the canonical view and average all the generated color images and depth maps. The averaged color image and depth map are back-projected using the canonical view to generate the final face morphing point cloud. We conducted extensive experiments on our dataset comprising 675 3D scans from 41 unique identities and 100 unique identities from the Facescape public dataset. We performed a vulnerability analysis of the proposed method using 2D and 3D FRS and conducted a human observer study. We performed a quantitative quality assessment of the generated 3D face morphing models using eight different quality metrics. Finally, we detected the generated 3D face morphing models based on pre-trained 3D deep learning models. A sample implementation of the proposed method is available at https://github.com/jagmohaniiit/3DFaceMorph.

3.7 Article 7: 3D Face Morphing Attack Generation using Non-Rigid Registration (RQ4)

This article presents a method for generating 3D face morphs from two bona fide point clouds. The proposed method in this article first selects bona fide point clouds with neutral expressions. The two input point clouds were then registered using a Bayesian Coherent Point Drift (BCPD) without optimization, and the geometry and color of the registered point clouds were averaged to generate a face-morphing point cloud. BCPD works only with geometry, so the color is added by applying the non-rigid-registration transformation to the source point cloud geometry and appending per-vertex colors would generate source aligned to target with geometry and per-vertex colors. The proposed method generates 388 face-morphing point clouds from 200 bona fide subjects. The effectiveness of the method was demonstrated through extensive vulnerability experiments, achieving a Generalized Morphing Attack Potential (G-MAP) of 97.93%, which is superior to the existing state-of-the-art (SOTA Chapter 10) with a G-MAP of 81.61%.

Chapter 4

Conclusions

This thesis's primary motivation is to make MAD classifiers robust, a crucial step towards enhancing their applicability to real-world environments. The research questions, introduced in Chapter 1 Section 1.3, form the backbone of this study. In this chapter, we delve into the conclusions drawn from these questions.

4.1 Conclusion of research questions

4.1.1 RQ1: Robustness of MAD Classifiers

RQ1 How can we improve the robustness of MAD classifiers in real-world environments that vary in the pose, expression, illumination, capture distance and image quality?

- The article in Chapter 5 presents a robust MAD classifier where the trusted live capture is captured at an ABC gate. The robustness of this MAD classifier is valid for pose, expression and illumination. The proposed method decomposes the input image into the diffuse reconstructed image and a normal map. The diffuse reconstructed image is mainly invariant to illumination. Further, the normal map is mostly invariant to pose and partially invariant to expression. Thus making the proposed MAD classifier robust for pose, expression and illumination. It needs to be pointed out that in the camera-based fusion the weights for Camera1 is 0.2, Camera2 is 0.3, Camera3 is 0.2 and Camera4 is 0.2. Thus, in our case the sum of weights is only 0.9 for camera-based fusion. However, the sum of weights for each individual camera of diffuse reconstructed image (0.7) and normal map (0.3) sum to 1. Further, we have used linear SVM as a classifier, means a constant scaling factor is applied to all scores. For e.g. lets consider weights for each camera as $w1$,$w2$,$w3$ and $w4$ and the linear SVM classifiers as $a1 \times s1 + b1$,$a2 \times s2 + b2$,$a3 \times s3 + b3$and $a4 \times s4 + b4$. Thus, the current weight-fusion score would be $w1 \times a1 \times s1 + b1 + w2 \times a2 \times s2 + b2 + w3 \times a3 \times s3 + b3 + w4 \times a4 \times s4 + b4$.This implies that a linear scaling would be applied to all the scores in our case and the final fusion accuracy should not change theoretically. The proposed method in this chapter significantly improved over SOTA [2] from an EER of 28.5 ± 0.4 to an EER of 8.6 ± 0.1.

- The article in Chapter 6 presents a comprehensive MAD classifier that captures the trusted live capture in the OTF scenario. This MAD classifier demonstrates pose, expression, illumination, capture distance, and image quality robustness. The proposed method is based on hierarchical fusion, where the first level of classifiers takes deep feature residue (D-MAD scenario) from six pre-trained deep networks as a feature. The division of six pre-trained networks into two groups and the selection of two pairs overall based on correlation and generalization ensures a comprehensive approach. Each pair is interpolated using SLERP, and the difference is used for classification. The method achieves robustness due to hierarchical fusion, optimal pair selection, and SLERP interpolation. The results of this chapter show significant improvements compared to the SOTA [8], with the EER dropping from 27.1% to 3.4% in the best case, and competitive results even when the mediums are different (EER of SOTA is 34.3% compared to proposed method EER of 26.0%).

- The article in Chapter 7 presents a superior MAD classifier that captures the trusted live capture from an ABC Gate. This MAD classifier demonstrates robustness to pose, expression, and illumination. The proposed method performs pair-wise face alignment followed by deep feature extraction and classification. It then performs weighted score fusion for all cameras for a given light. The pair-wise face alignment and weighted score fusion make the MAD classifier robust and invariant to pose and camera capture distance. The results of this chapter show that the proposed method achieves an EER of 2.1%, outperforming the SOTA [9] with an EER of 8.6% in the best case, thereby demonstrating its superior performance.

- Based on the obtained results from the proposed methods for the given research question, one could mention that it has largely been answered. However, more work could be done to improve this area.

4.1.2 RQ2: Effect of postprocessing on MAD Classifier

RQ2 What is the effect of postprocessing morphing images on the performance of the MAD classifier? Furthermore, what is the impact on the generalization of the MAD classifier trained using different mediums in the presence of postprocessing morphing images?

- The article in Chapter 8 is robust to changes in postprocessing. The article generates a postprocessed dataset for digital and print-scan mediums with and without compression. The proposed method converts the input RGB image to HSV and YCbCr color spaces. This is followed by applying the Laplacian Pyramid to extract the scale space. The extracted scale space is passed through an ensemble of features (LBP, BSIF and HOG) and classifiers (Linear SVM, SRKDA and CRC) for final classification. Using color spaces, scale space, and an ensemble of features and classifiers makes the MAD robust against postprocessing and more generalizable than SOTA. The proposed method achieved an EER=5.45% compared to SOTA [88] EER=9.82%. Based on the results obtained from the proposed method for the given research question, one could mention that we are one of the first to address this question on a large scale and more work needs to be done to improve this area.

4.1.3 RQ3: Generation of Facial Attribute-based Face Morphs

RQ3 How can we generate facial attribute-based face morphing which shows vulnerabilities of FRS and are the current MAD methods suitable to detect them?

- The article in Chapter 9) generates facial attribute-based face morphing and shows vulnerabilities towards FRS. However, current MAD methods can only partially detect them. The proposed method first generates segmented facial attributes from two contributory data subjects. This is followed by transparently blending of selected facial attributes to generate the initial composite and face mask. The CFIA (final composite face image) is generated by passing the initial composite through the encoder-decoder for face completion. The method achieves vulnerability towards FRS by using a transparent blending of facial attributes obtained from each contributory data subject. MAD methods achieve low accuracy in detecting generated CFIA as these methods work on full-face images, whereas CFIA has changed only in single/multiple facial attributes in the face image. The proposed method achieved a Generalized Morphing Attack Potential (G-MAP) of 46.9% compared to 52.4% of SOTA [3]. Based on the results, we have addressed this research question to a large extent, though more work can be done to improve further.

4.1.4 RQ4: Generation of 3D Face Morph

RQ4 How can we generate 3D Face Morphing when ground-truth 3D data is available from the two contributory data subjects and does the generated 3D Face Morphing show vulnerabilities of FRS?

- The article in Chapter 10 generates a 3D face morphing point cloud given bona fide point clouds from two contributory data subjects. The generated 3D face morphing point cloud shows vulnerability towards 2D/3D FRS. The proposed method is based on a 3D-2D-3D approach where the input bona fide point clouds are projected onto 2D color images and depth maps. The transformation is computed between the 2D color images, and the same transformation is applied to depth maps. This is followed by a blending operation to generate the 2D face morphing color image and depth map, which is then back-projected to create a 3D face morphing point cloud. The generated 3D face morphing point cloud is hole-filled to produce the final 3D face morphing point cloud. Since the blending and image registration is performed in 2D, it is much more robust due to the stability of 2D face key points. Further, corresponding facial parts are blended and 3D face morphing point clouds with identity features from both contributory data subjects are generated. Thus, the generated 3D face morphing point cloud shows vulnerability towards 2D/3D FRS used in the article. It needs to be pointed out that this article was one of the first in the biometrics domain, so SOTA was not present. However, it achieved a vulnerability of 100% compared to 3DMM [18], which generated a vulnerability of 66.67% using 3D FRS of LED3D [20].

- Further, we improved the results obtained from this proposed method with the article in Chapter 11 where we perform direct 3D morphing between two input facial point clouds based on Bayesian Coherent Point Drift (BCPD [89]) without optimization. This proposed method achieves a GMAP= 97.93% with 3D FRS compared to the previous SOTA(Chapter 10) GMAP=81.61%.

- Based on the obtained results from the proposed methods for the given research question, one could mention that it has largely been answered. However, more work could be done to improve this area.

Part II

Published Articles

Chapter 5

Article 1: Robust Morph-Detection at Automated Border Control Gate using Deep Decomposed 3D Shape & Diffuse Reflectance (RQ1)

Jag Mohan Singh, Raghavendra Ramachandra, Kiran B Raja and Christoph Busch. Robust morph-detection at automated border control gate using deep decomposed 3D shape & diffuse reflectance. In *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, Sorrento (NA), Italy, pp 106-112. IEEE

5.1 Abstract

Face recognition is widely employed in Automated Border Control (ABC) gates, which verify the face image on passport or electronic Machine Readable Travel Document (eMTRD) against the captured image to confirm the identity of the passport holder. In this paper, we present a robust morph detection algorithm that is based on differential morph detection. The proposed method decomposes the bona fide image captured from the ABC gate and the digital face image extracted from the eMRTD into the diffuse reconstructed image and a quantized normal map. The extracted features are further used to learn a linear classifier (SVM) to detect a morphing attack based on the assessment of differences between the bona fide image from the ABC gate and the digital face image extracted from the

passport. Owing to the availability of multiple cameras within an ABC gate, we extend the proposed method to fuse the classification scores to generate the final decision on morph-attack-detection. To validate our proposed algorithm, we create a morph attack database with overall 588 images, where bona fide are captured in an indoor lighting environment with a Canon DSLR Camera with one sample per subject and correspondingly images from ABC gates. We benchmark our proposed method with the existing state-of-the-art and can state that the new approach significantly outperforms previous approaches in the ABC gate scenario.

5.2 Introduction

Face recognition systems (FRS) are widely deployed at border crossings, which use Automated Border Control (ABC) gates. The deployment has ever increased since member states of the International Civil Aviation Organization (ICAO) follow ICAO's specification 9303 and store a standardized digital face image in the electronic Machine Readable Travel Document (eMRTD). However, FRS has shown to be vulnerable with respect to morphed face images - a new image as a result of a weighted linear combination of two input images, as shown in Figure 5.1. The generated morphed image challenges the FRS as it can be used to verify two unique identities (individuals), defeating the FRS's ability to verify unique subjects [90]. The challenge becomes severe as some countries issue the passport based on the digital photo uploaded by the applicant, which can provide an opportunity to upload a morphed image that can later be verified by an FRS [90, 2]. Several countermeasures have been proposed for Morphing Attack Detection (MAD). MAD can be broadly classified into No-Reference MAD (NR-MAD), which uses a single image for MAD and Differential MAD (D-MAD), which uses an image pair that includes a trusted live capture, and an image extracted from eMRTD. In addition, both MAD methods (NR-MAD and D-MAD) do or do not anticipate potential artifacts that have been introduced in the image signal with an optional print and scan process of the facial image [91]. In the rest of the paper, we present related work in Section 5.3, our proposed algorithm in Section 5.4, followed by experimental setup, and results in Section 5.5, and conclusions and future-work in Section 5.6.

5.3 Related Work

In this section, we review the related-work for D-MAD for which there are several algorithms, such as using landmark shifts proposed by Damer et. al [1], texture-descriptors based approach proposed by Scherhag et. al [2], and image subtraction based approach proposed by Ferrara et. al [7]. The authors in [1] conduct a face alignment using a common facial landmark detector [42] for each image and compute a distance-vector subsequently from landmark locations to train an SVM-RBF for morph detection. The authors in [2] also employ the face-alignment from [42],

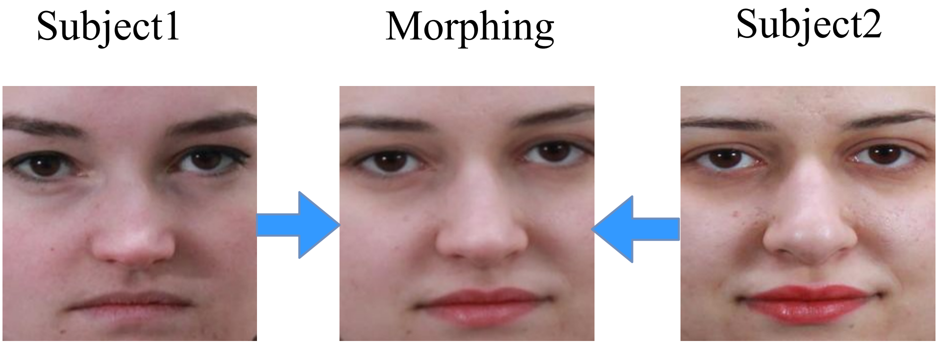


Figure 5.1: Digital morphing example from our database

followed by computing the vector differences between texture-descriptors such as LBP [92], BSIF [93], or SIFT [94]. The vector difference is then used to train an SVM-RBF for differential morph detection. One of the existing state-of-the-art (SOTA) schemes presented by authors in [7] tries to invert the morphing process using image subtraction. The authors observe that given the warping functions and alpha value, one could perfectly demorph a morphed image. However, in a practical scenario, the warping functions, and alpha value are unknown, so the authors obtain warping functions by face alignment, and prescribe $\alpha = 0.45$ for best quality demorphing. The following are the limitations of current SOTA in differential MAD, landmark shifts could occur due to pose changes, texture-descriptor features would have reduced efficacy in the presence of lighting, pose, and print-scan artifacts [90], and image subtraction methods would have reduced efficiency in the presence of lighting, pose, and print-scan artifacts as shown in Figure 5.4 some of which are also shown in [7].

In a real border control scenario, the subject is verified with the captured face image from the ABC gate, which is compared against the image stored in the eM-

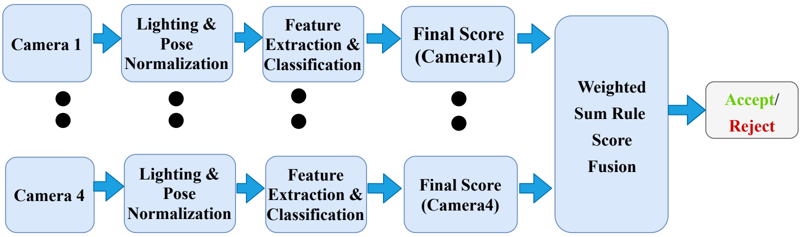

Camera 1
Lighting & Pose Normalization
Feature Extraction & Classification
Final Score (Camera1)
Camera 4
Lighting & Pose Normalization
Feature Extraction & Classification
Final Score (Camera4)
Weighted Sum Rule Score Fusion
Accept/ Reject

Figure 5.2: Pipeline of our approach showing the fusion of scores from Camera1, Camera2, Camera3 and Camera4 where each camera features are in-turn generated by fusion.

RTD. This is what we modeled in our work. We leverage this to verify if the image on eMRTD is morphed by looking at the 3D shape and reflectance for both captured images from ABC and image within the eMRTD. Specifically, we look at the normal-map and the diffuse reconstructed image, to devise a classifier that can distinguish bona fide (non-morphed) images from morphed images. We assert that the morphed image presents significantly inconsistent information within the image as compared to the non-morphed image. It has further to be noted that many ABC gates operate with multiple cameras, which enable us to reinforce the decision with fusion approaches to detect a morphing attack in a better manner, as demonstrated in our work. To the best of our knowledge, this is the first method to explore the strengths of a multi-camera capture set-up in border control operations to detect the morphing attacks. To assert our approach, we create a new database with bona fide images of 39 subjects in an ideal enrolment setting and correspondingly the probe images of the same 39 subjects, which were captured while crossing the ABC gate. The images from the 39 subjects are used to create morphed images (90).

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

- Presents a new database of morphed images and trusted live capture probe images captured in a realistic border crossing scenario with ABC gates.
- Presents a new approach employing the inherent border crossing scenario to detect the morphing attacks using a fusion of scores from a quantized normal-map approach and diffuse reconstructed image characteristics.
- Presents an extensive evaluation of state-of-art D-MAD techniques to benchmark the proposed algorithm, and demonstrate the superiority of the proposed algorithm.

5.4 Proposed Algorithm

In this section, we describe the proposed algorithm for robust morph detection at an ABC gate. In our approach, the probe face image, which is captured at the ABC gate, is compared with a face image from the eMRTD. The ABC gate face image and the digital face image from the eMRTD would likely have intensity changes due to lighting differences in the capture environments, pose changes due to the capture subject interaction, image quality differences along with the additional noises introduced in the print-scan process preceding the storing of a given digital face image in the eMRTD. Given that these changes may not optimally help in determining a morph attack, we formulate the problem of morphing attack detection first by normalizing the pose changes in the image, further to which we compute

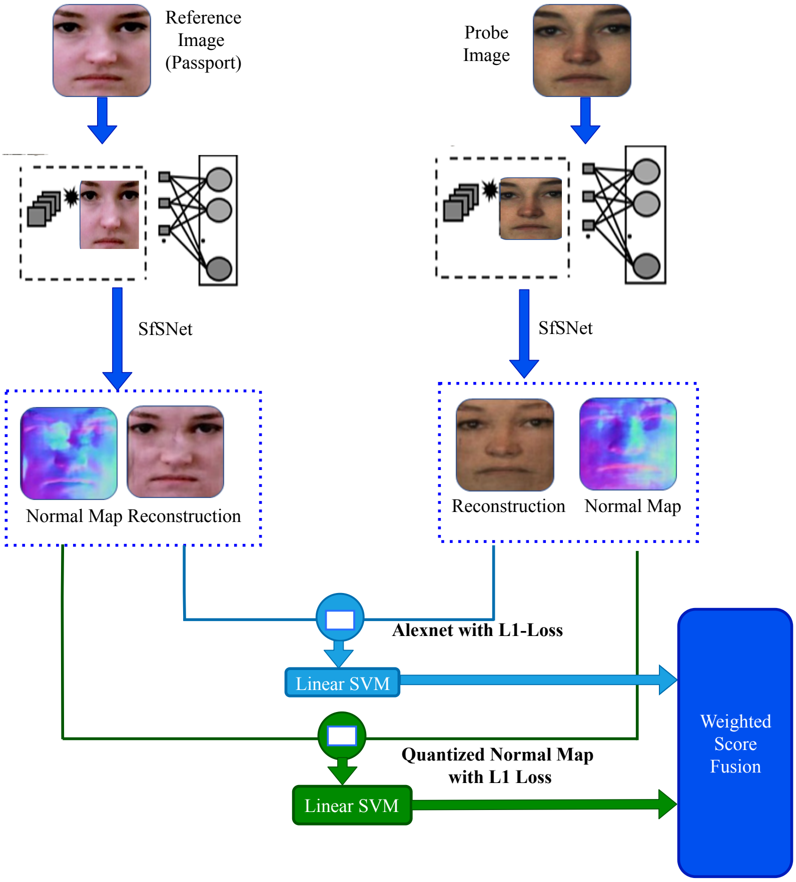


Figure 5.3: Illustration of feature extraction and classification for each camera

the features for D-MAD. The pipeline of the proposed approach is depicted in Figure 5.2, where pose normalization is carried out first. Further to this, we extract the features to learn a robust classifier, as shown in Figure 5.3 for each camera. Given the availability of multiple cameras, we further propose a weighted sum-rule score level fusion for scores from each camera. Each of the components of the proposed method is further detailed, as provided in the subsequent sections.

5.4.1 Pose Normalization

We also do pose normalization using the method from authors in [42] as the face images from ABC Gate could be in a non-frontal pose. The method we use for pose normalization is based on the key-points which are automatically detected in a face, and it makes the line joining the eye-centers horizontal.

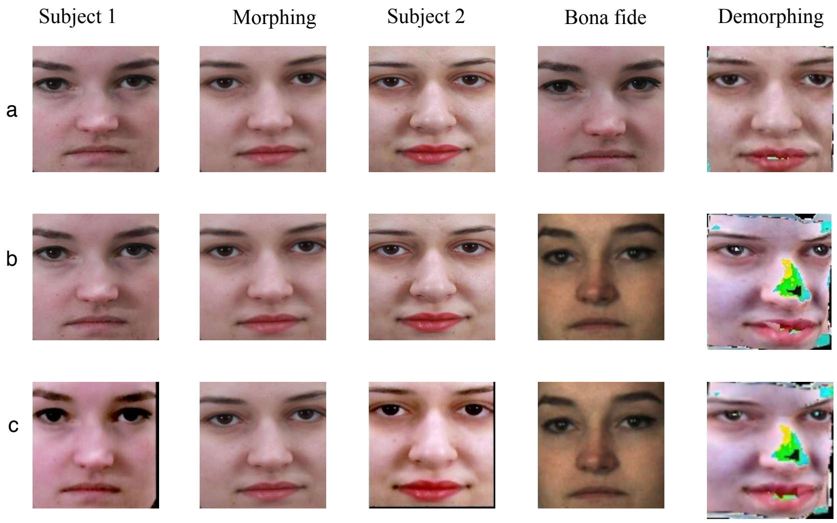


Figure 5.4: Demorphing using Image Subtraction based technqiue by Ferrara et. al [7] of subjects in different conditions fails especially in (b), and (c) where Bona fide Image is from our dataset. Rows: Digital Images in (a) Bona fide Image captured in similar lighting & (b), Print-Scan (Inkjet EPSON TM) Images (c). Results are based on our own implementation.

5.4.2 Feature Extraction and Classification

Given the images are now normalized for pose using the method described in Section 5.4.1, we proceed to extract the features. We, therefore, decompose an input image I into diffuse reconstructed image $I(p)$ and a normal map $n(p)$, which represents the shape of the face. We choose SfSNet [95], as it can decompose a single input image into the diffuse reconstructed image, normal-map, albedo-map ρ, and 2nd order spherical harmonic based lighting coefficients l_{nm}. The diffuse reconstructed image can be written as with second order spherical harmonics using [96] as follows:

$$I(p) = \rho r(n(p)) \tag{5.1}$$

where r(n(p)) which is reflectance of the material, is given by

$$r(n(p)) = \sum_{n=0}^{n=2} \sum_{m=-n}^{n} l_{nm} r_{nm}(n(p)) \tag{5.2}$$

where l_{nm} for $n = 0$ are used from the ambient coefficients identified in Section 5.4.1.

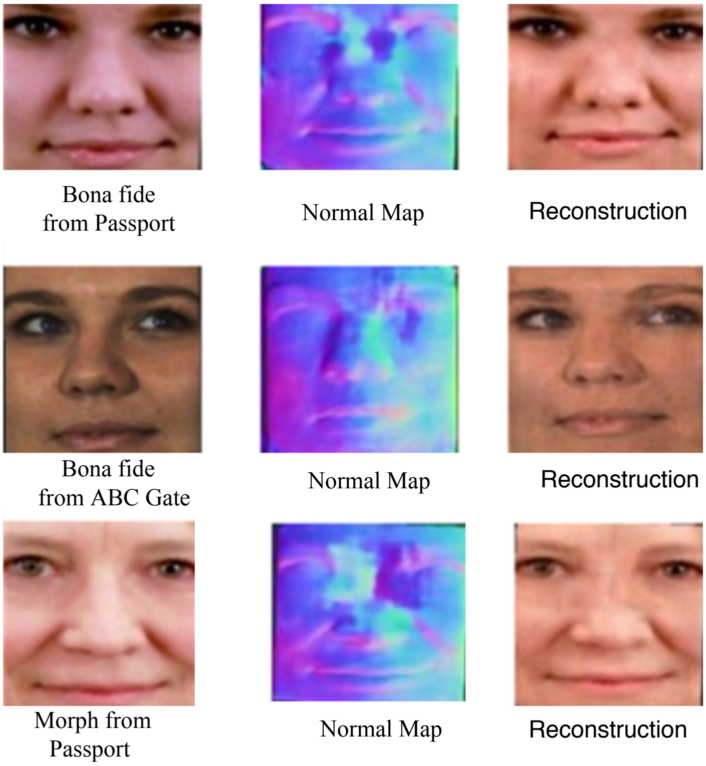


Figure 5.5: Image decomposed into Normal-Map and Reconstruction (Diffuse Reconstructed Image)

As it can be observed from Figure 5.5, the diffuse reconstructed image (pixel color differences are highlighted), and the normal-map (especially around the eye, and the nose regions) help to distinguish the bona fide and morph images, while in the non-decomposed domain they look quite similar.

Feature Extraction

We extract the features as depicted in the Figure 5.3 within the proposed algorithm shown in Figure 5.2. Owing to the robust nature of Alexnet [97] in obtaining reliable features, we employ the Alexnet to derive features from the diffusely reconstructed image. Given that the image is diffuse, we assert that it is closer in feature space than input image I. We use $fc7$ layer of Alexnet for extracting features resulting in a feature vector of 4096 elements on which we compute reconstruction-loss as L1-Loss. We compute a quantized normal map of 21-bits from the normal map, which is output by SfSNet [95] as quantization would result in the normal map being robust to small variations. This is followed by taking the simple difference as L1-Loss.

Feature Classification

Given the set of features, we train a linear SVM for diffuse reconstruction-loss, and normal-loss. The scores are fused by weighted fusion to generate the score for each camera. This is followed by a weighted sum-rule fusion of scores from each camera to achieve the final score, which can be used for the detection of morph, as shown in Figure 5.2. The weights in both fusion steps are chosen based on a greedy search optimization algorithm [98]. The weights chosen for each camera are 0.7 for the diffuse reconstructed image classifier and 0.3 for the normal-map classifier. The weights chosen for the cameras are as follows, Camera1 0.2, Camera2 0.3, Camera3 0.2, and Camera4 0.2.

5.5 Experimental Setup & Results

In this section, we provide details on our database and the corresponding experimental protocols, following the results obtained. We report the performance of the proposed D-MAD algorithm using the following metrics defined in the International Standard ISO/IEC 30107-3 [99] described as follows:

- Attack Presentation Classification Error Rate (APCER), which is the misclassification rate of morph attack presentations.
- Bona fide Presentation Classification Error Rate (BPCER), which is the misclassification of bona fide presentation as morphs.

We also report Detection Equal Error Rate (D-EER $\%$) and detection error trade-off curves, to examine the rate of change of mis-classification errors.

5.5.1 Morph ABC Database

To simulate the operational scenario with attacks in the enrolment and trusted probe images from ABC gates, we created a new database in this work. We want to point out that in a realistic operational scenario, the digital image in the eMRTD may be bona fide or morphed. First, we generate a set of enrolment images for 39 subjects captured in a realistic studio setting with multiple images using a Canon DSLR camera of 21 megapixels. Secondly, we capture the face images of the same 39 subjects in an ABC gate using a real-world equipment [100]. We employ a single image per subject from DSLR images as a bona fide passport image and treat the images, which were captured from the ABC Gate with four different cameras (one sample each) as bona fide probe images. Employing another session of DSLR images captured from the enrolment set up, we create a morphed passport image dataset using the images from 39 subjects and the approach and conditions mentioned in work by Raghavendra et al. in [90] specifically subjects not wearing glasses, and using the same gender, and ethnicity. The morphed images and bona fide images are printed and scanned using EPSON XP-860 Printer, and Scanner.

<table>
<tr><th></th><th>Bona fide Passport</th><th>Bona fide all
ABC Gate Cameras</th><th colspan="2">Morphed Passport</th></tr>
<tr><td>Train</td><td>19</td><td>237</td><td colspan="2">52</td></tr>
<tr><td>Test</td><td>20</td><td>222</td><td colspan="2">38</td></tr>
<tr><th colspan="5">Bona fide per ABC Gate Camera</th></tr>
<tr><th></th><th>Camera1</th><th>Camera2</th><th>Camera3</th><th>Camera4</th></tr>
<tr><td>Train</td><td>58</td><td>64</td><td>58</td><td>57</td></tr>
<tr><td>Test</td><td>57</td><td>63</td><td>49</td><td>53</td></tr>
</table>

Table 5.1: Dataset Details

Performance Protocol: In D-MAD, as we need two images for morph detection, we consider the bona fide passport images v/s bona fide gate images as the genuine class samples, and morph passport image v/s bona fide gate image as the attack class samples. We now go into details of the number of scores generated during training as follows: From the enrolment, we have 19 bona fide passport images, complemented with 52 morphed passport images. Further from Camera 1 in the ABC Gate we have 58 bona fide probe images, which results in $19 \times 58 = 1102$ genuine scores, and $52 \times 58 = 3016$ attack scores, Camera2 results in $19 \times 64 = 1216$ genuine scores, and $52 \times 64 = 3328$, Camera3 results in $19 \times 58 = 1102$ genuine scores, and $52 \times 58 = 3016$, and Camera4 results in $19 \times 57 = 1083$ genuine scores, and $52 \times 57 = 2964$ attack scores. The number of scores generated during testing is as follows: From the enrolment 20 bona fide passport images, complemented by 38 morphed passport images. From Camera1 in the ABC Gate we have 57 bona fide probe images, which results in $20 \times 57 = 1140$ genuine scores, and $38 \times 57 = 2166$ attack scores, Camera2 results in $20 \times 63 = 1260$ genuine scores, and $38 \times 63 = 2394$, Camera3 results in $20 \times 49 = 980$ genuine scores, and $38 \times 49 = 1862$, and Camera4 results in $20 \times 53 = 1060$ genuine scores, and $38 \times 53 = 2014$ attack scores. During fusion of scores of the four cameras, we reach 980 genuine scores, and 1862 attack scores as this are the minimum number of genuine and attack scores available in all four cameras during testing.

5.5.2 Analysis of Results

Table 5.2 presents the results of the proposed method and compares it with two state-of-the-art approaches including Landmark Shifts based Signed Distance pro-

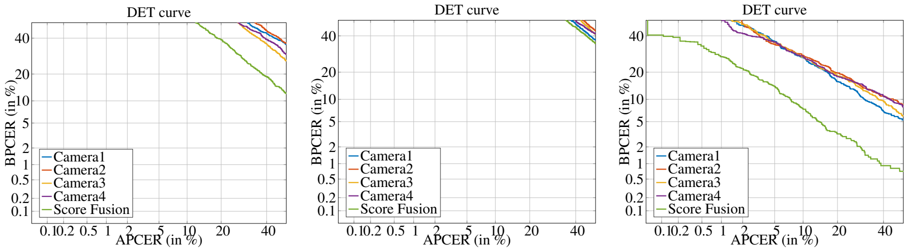
Method	**Cam**	**EER**	**BPCER20**	**BPCER10**
Signed Distance [1]	1	43.7±0.2	90.5±0.3	83.4±0.3
	2	46.7±0.3	93.5±0.4	87.8±0.2
	3	45.8±0.2	92.7±0.3	86.3±0.7
	4	45.1±0.3	91.4±0.4	82.3±0.3
	Fused	42.6±0.2	90.0±0.1	81.5±0.3
LBP & SVM [2]	1	41.7±0.4	81.1±0.6	72.4±1.0
	2	42.7±0.5	82.5±0.6	73.5±0.8
	3	38.1±0.5	83.3±0.7	71.5±0.6
	4	39.6±0.3	79.6±0.5	71.1±0.4
	Fused	28.5±0.4	67.2±0.6	54.2±0.8
Proposed Method	1	18.1±0.1	36.3±0.7	27.1±0.3
	2	19.7±0.4	34.7±0.7	28.3±0.7
	3	19.1±0.1	35.9±0.1	27.3±0.1
	4	18.8±0.1	36.1±0.1	27.5±0.3
	Fused	**8.6±0.1**	**13.9±0.4**	**7.5±0.1**

Table 5.2: Signed Distance by Damer et al. approach [1] using author's implementation, LBP and SVM by Scherhag et. al [2], and the proposed method where BPCER20 is BPCER@APCER=5%, and BPCER10 is BPCER@APCER=10%

posed by Damer et al. [1] and Texture-Descriptors based LBP-SVM by Scherag et. al [2]. As it can be noted from the Table 5.2, the proposed method outperforms existing SOTA, we achieve an EER of 8.6 ± 0.1 compared to best EER of SOTA of 28.5 ± 0.4. The results can also be seen in Figure 6.5, which presents the Detection Error Trade-off Curves, where it can be noted that fusion of scores leads to further improvement for the proposed algorithm compared to the SOTA. Despite outperforming the SOTA, we note that our proposed approach still has moderate deficiency from single cameras, as shown in Table 5.2. We make the following observations from the results:

- One can observe that in similar lighting capture environments, as shown in Figure 5.1 (row (a)), Image Subtraction based technique proposed by authors in [7] performs well, and one can generalize this argument texture descriptor based method report by authors in [2]. However, the same cannot be said for the technique proposed by authors in [1] as landmark shifts could happen due to change in pose.

- Figure 5.1 shows the degrading performance of the Image Subtraction based method proposed by authors in [7] in (rows (b), and (c)) which have lighting



(a) **(b)** **(c)**

Figure 5.6: DET Curves for (a) LBP-SVM [2], (b) Signed-Distance [1], and (c) the proposed method. DET Curves are for Scores from Camera1, Camera2, Camera3, Camera4, and Weighted Sum-Rule Fusion of scores from these individual cameras.

changes and print-scan artifacts. The advantage of using features from a diffuse reconstructed image which contains lower-order lighting terms, and normal-map are shown in Figure 5.5.

- The proposed method achieves the best D-EER compared to the existing SOTA mainly due to two factors, the use of a diffuse reconstructed image that removes the higher-order lighting components and leads to a linear light model without cast shadows as pointed out by Basri et. al [96, 101]. The second factor is the use of normal-map, which on integration gives depth-map [102], and depth-map signifies the 3D shape of the bona fide sample. The 3D shape, and consequently normal-map of the bona fide sample, should be preserved across different cameras.

5.6 Conclusion & Future Work

In this paper, we presented a novel and robust scheme to perform D-MAD in the presence of lighting, pose, and print-scan artifacts. We have constructed a new database reflecting the real-life border crossing scenario and have validated the results on our collected database. Our collected database models the real-life print-scan artifacts in the passport image and the use of camera images from the ABC gate. The proposed method outperforms the existing SOTA methods for D-MAD mainly due to the combined effect of pose normalization, use of a diffuse-reconstructed image, and normal map. In future works, the proposed algorithm shall be tested on a large scale database.

Acknowledgement

This work was carried out under the funding of the Research Council of Norway under Grant No. IKTPLUSS 248030/O70.

Chapter 6

Article 2: Reliable Face Morphing Attack Detection in On-The-Fly Border Control Scenario with Variation in Image Resolution and Capture Distance (RQ1)

Jag Mohan Singh, and Raghavendra Ramachandra. Reliable Face Morphing Attack Detection in On-The-Fly Border Control Scenario with Variation in Image Resolution and Capture Distance, In *IEEE International Joint Conference on Biometrics (IJCB)) 2022)*, Abu Dhabi, UAE, pp. 1-10, IEEE

6.1 Abstract

Face Recognition Systems (FRS) are vulnerable to various attacks that are performed both directly and indirectly. Among these attacks face morphing attacks are highly potential in deceiving both automatic FRS and human observers and indicate the severe security threat especially in the border control scenario. In this work, we present a face morphing attack detection especially in the On-The-Fly (OTF) Automatic Border Control (ABC) scenario. We present a novel algorithm for Differential-MAD (D-MAD) based on the spherical interpolation and hierarchical fusion of deep features computed from six different pre-trained deep Convolutional Neural Networks (CNNs). Extensive experiments are carried out

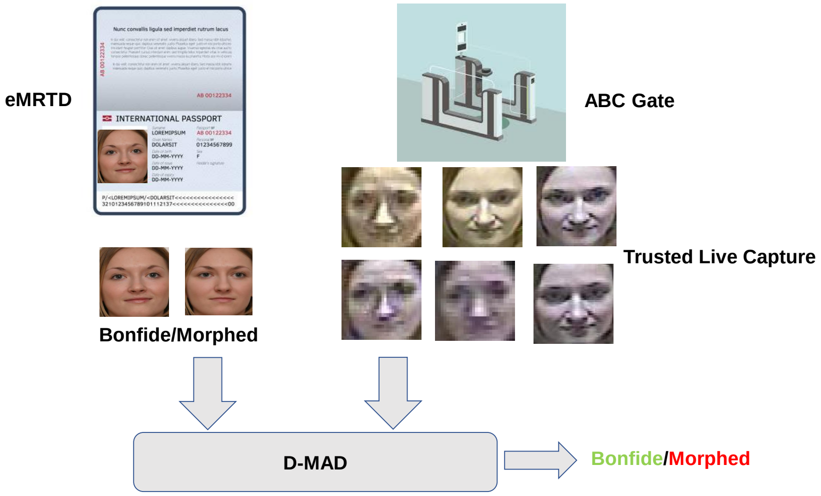


Figure 6.1: Illustration showing eMRTD presentation at an ABC Gate and D-MAD based decision.

on the newly generated face morphing dataset (SCFace-Morph) based on the publicly available SCFace dataset by considering the real-life scenario of Automatic Border Control (ABC) gates. Experimental protocols are designed to benchmark the proposed and state-of-the-art (SOTA) D-MAD techniques for different camera resolutions and the different capture distances. Obtained results have indicated the superior performance of the proposed D-MAD method when compared with the existing methods.

6.2 Introduction

Algorithm	Algorithm Classification	Brief Description
Landmark [1]	Landmark Based	Directed Landmark shifts used for classification
Feature-based [2]	Feature difference based	Feature Difference used for classification
Fusion of classifiers based [103]	Fusion of classifiers	Fusion of hand-crafted (LBPH) and CNN-based (TDCNN) features
FD-GAN [104]	FD-GAN	Image Demorphing using symmetric dual GAN
LRP [105]	LRP	Layer-Wise Relevance Propagation based on pixel-wise decision
Siamese [106]	Siamese Architecture	Siamese Architecture based on Inception ResNET v1 with weights from VGGFace2
Mutual Information Maximization [107]	Disentanglement	Disentaglement of Appearance and Landmarks based on CNN
Demorphing [108]	Image Subtraction	Inverting Morphing Equation with image-pair, and known correspondences and α
Fusion of CNN features [9]	Fusion of Classifiers	Shape (Normal-Map) and Reflectance (Diffuse Reconstruction) Decomposition: SfS-Net and Alexnet
DFR [109]	DFR	Signed Distance of Arcface and Facenet features
Demorphing [110]	Demorphing	Autoencoder-based demorphing and face simlarity analysis
Siamese [111]	Siamese	Siamese for D-MAD trained on wavelet basis chosen using Kullback-Liebler Divergence (KLD)
Double Siamese [112]	Double Siamese	Double-Siamese based D-MAD, indentity-based and artifact-based
GAN [113]	GAN	Conditional Identity Disentaglement using Conditional GAN for D-MAD
Legacy [114]	Legacy	Legacy Image and Face Verification Engine score based D-MAD

Table 6.1: State-of-the-art D-MAD techniques

Face biometrics are widely deployed in various high-security applications including border control by considering usability, high accuracy, and non-intrusive capture. The high accuracy of the face biometrics can be attributed to the advances in deep-learning-based FRS methods [27, 28, 22]. The exponential growth in face recognition applications has also increased the vulnerability to various attacks. Among different types of attacks on the Face Recognition Systems (FRS), the morphing attacks have mainly gained much interest due to their vulnerability in the border control scenario. The morphing process will perform the blending operation on the given face images (from contributory data subjects) to generate a single Face Morphing Image (FMI). Thus, the generated FMI includes the facial properties from all the contributory data subjects, thus demonstrating the vulnerability of both automatic FRS and human observers [115]. Since the morphing face images could be used to obtain the electronic Machine Readable Travel Document (eMRTD) or e-passports, the malicious person can exploit this process to cross the border through Automated Border Control (ABC) gates.

Face Morph Attack Detection (MAD) algorithms are extensively addressed in the biometric literature [116]. Available MAD algorithms can be classified into two main categories [116] (a) Single image based-MAD (S-MAD), where morph attacks are detected based on a single image (b) Differential MAD (D-MAD) algorithms, where morphing attacks are detected based on two or more images. Among these two approaches, the D-MAD-based MAD techniques have attracted biometric researchers by considering their application in the border control scenario. Figure 6.1 illustrates the D-MAD scenario in the border control application. The early work on the D-MAD approach is based on the face demorphing [108] which was followed by several existing methods that are summarised in Table 6.1. The D-MAD approaches are developed using both conventional hand-crafted features (such as LBP, HoG, LPQ) and deep features derived from pre-trained CNNs based on natural and face images [116].

The deep learning approaches based on GAN, Siamese and Double Siamese have also been proposed for D-MAD. The benchmarking of several existing D-MAD techniques is presented in [116, 117] on the data captured using ABC gates indicates the severity of the problem by showing the degraded results. The ABC gate scenarios used in [117] are based on the one-stop such that the data subject will stand still in front of the ABC gate camera. Thus, this scenario will generate constrained images less prone to the pose and environmental (external lighting) conditions. In [9], the ABC scenario based on the ’on-the-fly’ face capture, and 3D information based D-MAD method was introduced. Since ’on-the-fly’ (OTF) face capture will result in variations in face pose, expression, and lighting, the D-MAD techniques based on face demorphing and conventional features have indicated the degraded performance [9]. The experimental results with different lighting conditions indicate the further degradation of the 3D based D-MAD results. However, it is important to note that, the existing D-MAD literature did not consider the option of different cameras with varying capture resolution and capture distance impacting the detection performance. Since the D-MAD techniques are expected to work with different ABC gates with varying camera resolutions, it is necessary to devise a suitable D-MAD method for this scenario. Thus, in this work, we are motivated to consider the OTF ABC gate scenario with the various camera resolutions and different capture distances.

This work proposes a novel algorithm for a robust D-MAD especially in the OTF ABC Gate scenario with varying image resolutions and capture distances. To this extent, we introduce a novel D-MAD algorithm based on spherical interpolation and the hierarchical fusion of deep features to detect morphing attacks. The deep features are extracted using six different pre-trained deep CNN networks that are combined using a hierarchical fusion at both score level and feature level. Extensive experiments are carried out on the newly created database SCFace-Morph using the publicly available SCFace [118] database with 130 data subjects captured using both controlled and uncontrolled scenarios with different resolution cameras and different capture distances. We construct the new face morphing dataset SCFace-Morph dataset using landmark-based face morphing tools [119] and we also re-digitize (or print-scan) the face morphing images to represent the real-life scenario of the border control.

The following are the main contributions of our work:

- Proposed a novel D-MAD algorithm based on spherical interpolation and hierarchical fusion of deep features for robust face morphing detection.
- Introduced a new face morphing dataset (SCFace-Morph) constructed using the publicly available dataset (SCFace [118]) for both digital and Print-Scan (PS) morphing attacks. *To the best of our knowledge, this is the first work exploring morphing attack detection on the different camera resolutions and at various capture distances suitable for the OTF ABC scenario.* Further, the database will be made available to the semi-public together with the proposed method for the complete reproducible of the results presented in this paper.
- Extensive experiments are carried out to benchmark the performance of the proposed method with the SOTA techniques.

The rest of the paper is organized as follows: we present the proposed method in Section 6.3, experiments and results are discussed in Section 6.4 and finally conclusions and future work is discussed in Section 6.5.

Algorithm 1: Proposed Method

Input: Face Images I_1 and I_2

Output: (FS (Fused-Score))

```
Compute the features from pre-trained networks for Image ($I_1$).
for $j \leftarrow 1$ to 6 do
    $f_1^j \leftarrow$ feature from pre-trained network.
end for
Compute the features from pre-trained networks for Image ($I_2$).
for $j \leftarrow 1$ to 6 do
    $f_2^j \leftarrow$ feature from pre-trained network.
end for
Assign features to Groups as follows:
$G_1 \leftarrow \{f_i^j\}$ where $i \in \{1, 2\}$ and $j \in \{1 \dots 3\}$
$G_2 \leftarrow \{f_i^j\}$ where $i \in \{1, 2\}$ and $j \in \{4 \dots 6\}$
Compute the feature difference
for $j \leftarrow 1$ to 6 do
    $DF^j \leftarrow f_1^j - f_2^j$
end for
Train Linear-SVM using difference features and compute scores
for $j \leftarrow 1$ to 6 do
    $S_j \leftarrow$ L-SVM($DF^j$)
end for
Use the pre-computed pair of optimal features $(x1, y1)$ and $(x1, z1)$ for $G_1$
and $(x2, y2)$ and $(x2, z2)$ for $G_2$. They are computed once using
Equation 6.1.
Compute SLERP (Equation 6.3.2) based scores as follows where $i$ denotes
Group Index and $j$ denotes the pair of optimal SLERP features inside it:
for $i \leftarrow 1$ to 2 do
    for $j \leftarrow 1$ to 2 do
        $SRP_i^j \leftarrow$ SLERP($f_i^{xj}, f_i^{yj}$)
    end for
    Compute difference of SLERP features as $SRP_{D_i} \leftarrow SRP_i^1 - SRP_2^1$
    Compute score using $S_{i+7} \leftarrow$ L-SVM($SRP_{D_i}$)
end for
Generate final score by fusion using sum-rule as ($FS = \sum_{j=1}^{8} S_j$)
```

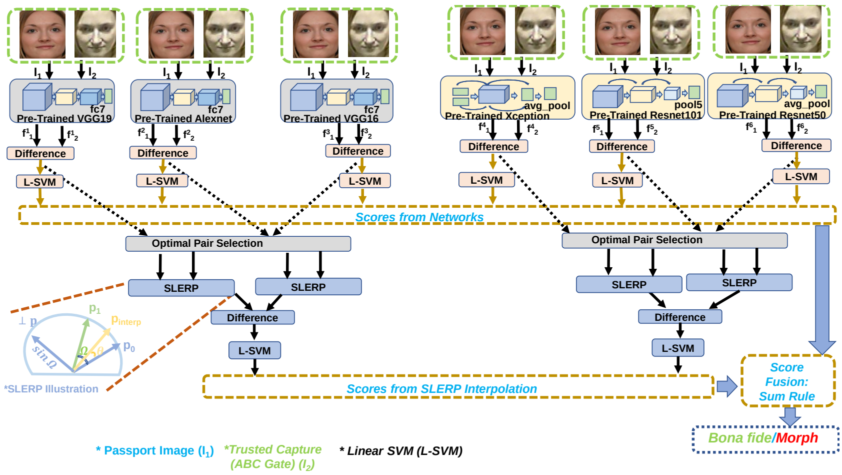


Figure 6.2: Illustration of the proposed Hybrid SLERP for Differential Morph Attack Detection (D-MAD).

6.3 Proposed Method

Figure 6.2 shows the block diagram of the proposed method for robust D-MAD, especially in the OTF border control scenario. The proposed method is designed effectively to capture the variation of the face quality in terms of environmental changes due to lighting, pose, and expression that are normally encountered with the probe image by introducing a hierarchical fusion of deep features. The novel aspect of the proposed method is the feature interpolation fusion using Spherical Linear Interpolation (SLERP) [120] tailored to D-MAD. The proposed method will take two images I_1 and I_2 corresponding to the enrolment (from e-passport) and the trusted capture (ABC Gate) face image respectively to detect the morphing attack on enrolment face image.

The primary motivation for using SLERP [120] is that it can perform the exact interpolation of quaternion vectors (representing a 3D rotation) on a 3D sphere. The deep features obtained for a face image are assumed to lie on a high-dimensional hypersphere [121]. Thus, linear interpolation of deep features would lie at a point on the line joining them, but the actual interpolation should lie on the hypersphere defining the face manifold. Thus, there would be an error due to linear interpolation of deep features when compared to spherical interpolation using SLERP, which would interpolate features exactly on the hypersphere (which is also mentioned by Buss et al. [122] for 3D Sphere). This fact is illustrated in the Figure 6.3 where the error is highlighted in red. Hence, we are motivated to use SLERP to

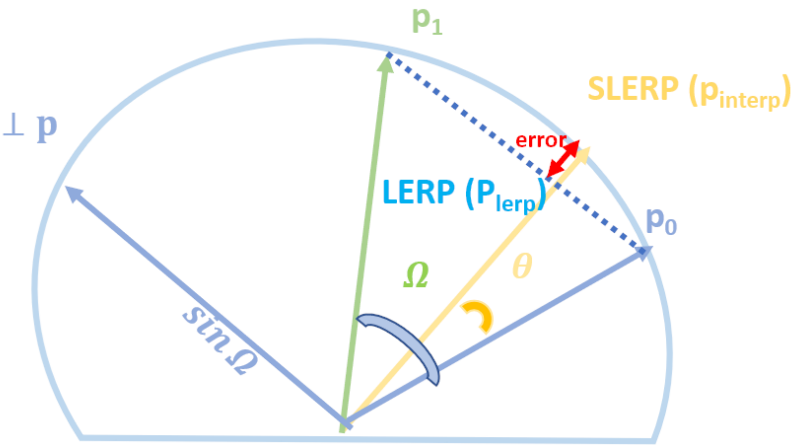


Figure 6.3: Illustration comparing SLERP and Linear Interpolation (LERP) where the error is highlighted in red.

interpolate deep features obtained corresponding to the face image.

Owing to the availability of the small size face morphing datasets, the proposed method is designed using pre-trained deep CNN networks. Hence, we propose a hierarchical fusion framework to effectively capture the complementary information from different deep features. The proposed D-MAD method can be structured in two functional blocks: (a) Deep feature extraction and (b) Hierarchical fusion.

6.3.1 Deep feature extraction

In this work, we have employed six different pre-trained deep networks that are trained on the ImageNet dataset [123]. The selected networks includes Alexnet [97], Resnet 50 [124], Resnet 101 [124], Xception [125], and VGG16 [126] and VGG19 [126]. These networks are selected based on their performance and generalization for transfer learning in various applications, including morph attack detection [116]. Further, these six network have indicated a good face morphing detection performance on various face morphing datasets [127]. Since the proposed Spherical Linear Operator for feature interpolation requires the features to have identical dimensions, we need to choose the feature extraction from the pre-trained layers having identical dimensions. Hence, we made two groups of pre-trained networks where the first group (G_1) consists of VGG-19 (fc7), VGG-16 (fc7), and Alexnet (fc7) such that each network in this group will result in a feature dimension of 4096. The second group (G_2) consists of Xception (average pool), Resnet101 (pool5), and Resnet50 (average pool) and each of these networks will result in a feature dimension of 2048. Thus, given the face image I_1 and I_2 from both the passport and the trusted environment (e.g., Automatic border Control Gates (ABC)), we compute the features from all six different CNN networks independently. Let the computed feature be: $F_i = f_1^i, f_2^i, \forall i = 1, \ldots, 6$. Where f_1^i indicates the features from the passport image (or enrolment) corresponding to i^{th} CNN network and f_2^i corresponds to the feature from trusted source corresponding to i^{th} CNN network.

6.3.2 Hierarchical fusion

In the next step, we propose the hierarchical fusion of the features extracted from six different pre-trained deep CNN networks to achieve robust face morph detection performance. The proposed fusion scheme is implemented with both score-level and feature-level fusion of the features extracted from deep CNN networks. The score fusion is designed with the conventional score level fusion in which the comparison scores obtained using Linear Support Vector Machines (L-SVM) based on the feature difference vector from six different CNN networks are combined using sum rule. Given the face images I_1 and I_2, let the computed features be f_1^i and f_2^i $\forall i = 1, \ldots, 6$. The feature difference (or residual feature computation) is performed individually for the pre-trained network $DF^i = f_1^i - f_2^i$ that is then provided to L-SVM to compute the corresponding comparison score $S^i, \forall i = 1, \ldots, 6$. The second step is designed to perform the feature level fusion using Spherical Linear Operator (SLERP) [120]. As discussed earlier in Section 6.3.1, the feature interpolation requires the identical dimension of features. Therefore, we have grouped six different networks into two main groups, and in each group, we have three different networks. In this next step for each group, we compute the optimal basis pairs for SLERP instead of using all possible random combinations. In this way, our approach can reduce the computation and combine the complementary features. Further, the feature combination is based on feature correlation computed using residual features to achieve robustness and generalization. Thus, given the face image I_1 and I_2, we compute the residual feature corresponding to G_1 which is denoted as DF^1, DF^2 and DF^3. In the next step, we compute the optimal basis pair from the triplet that can represent the complementary information to perform the residual feature combination using the SLERP method. We derive the optimal basis by computing the minimum correlation on the non-overlapping pairs generated from the given triplet, and this is indicated in

Equation6.1.

$$
\begin{aligned}
O_p = \arg\min \Big\{ & \rho(DF^1, DF^2) + \rho(DF^1, DF^3), \\
& \rho(DF^2, DF^1) + \rho(DF^2, DF^3), \\
& \rho(DF^3, DF^1) + \rho(DF^3, DF^2)) \Big\}
\end{aligned} \tag{6.1}
$$

where ρ indicates the correlation operation and O_p indicates the optimal pairs of features. Thus,

$$
\begin{aligned}
S &= \{1, 2, 3\} \\
O_p &= \{\{DF^x, DF^y\}, \{DF^x, DF^z\}\} \\
\text{s.t.} &(x \in S) \wedge (y, z \in S \setminus i) \wedge (y \neq z)
\end{aligned} \tag{6.2}
$$

where, x is the index of the optimal pairs, from the triplet $1, 2, 3$ of features and y and z are the indices of the remaining features. The first optimum pair be $O_p^1 = \{DF^{x1}, DF^{y1}\}$ and second optimum pair be: $O_p^2 = \{DF^{x1}, DF^{z1}\}$. In the next step, we compute the SLERP feature interpolation independently for each optimum pair $O_p^s, \forall s = 1, 2$ as follows:

$$
\begin{aligned}
&\text{Slerp}_1(DF^{x1}, DF^{y1}, t) = \\
&\frac{sin((1-t)\Omega)}{sin(\Omega)} \times (DF^{x1}) \\
&+ \frac{sin((t)\Omega)}{sin(\Omega)} \times (DF^{y1})
\end{aligned} \tag{6.3}
$$

where, t is the interpolation factor which is set to 0.5 as recommended in [120], and Ω is the angle between the difference features DF^{x1} and DF^{y1} and can be computed using inverse cosine on dot-product as $\Omega = \arccos DF^{x1} \cdot DF^{y1}$. We followed the similar procedure with O_p^2 to compute the Slerp$_2(DF^{x1}, DF^{z1}, t)$.

In the next step, we perform the difference between the computed SLERP features, which is then used to compute the comparison score S_7 using L-SVM for G_1. The procedures mentioned above are followed with G_2 to compute the comparison score S_8. Finally, we perform the score level fusion using the sum rule to combine the scores computed from both individual CNNs and from SLERP interpolated feature differences to obtain the final score: $FS = \sum_{i=1}^{8} S_i$ to make the final decision. The Algorithm of the proposed method is presented in 1.

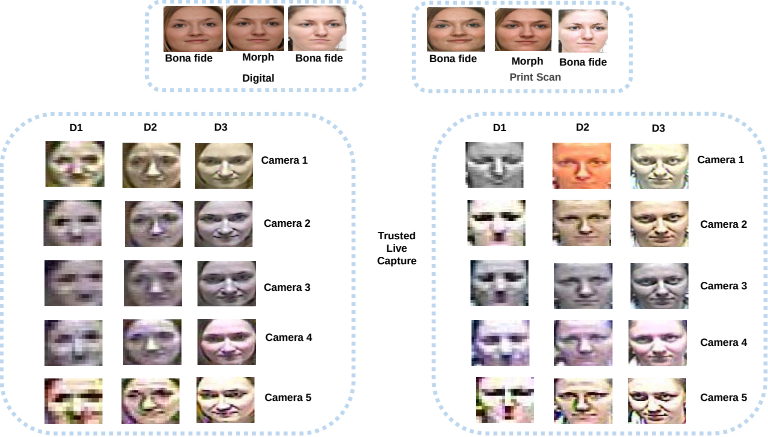


Figure 6.4: Example face images from SCFace-Morph Dataset (SCFM)

6.4 Experiments and Results

In this section, we discuss the details of the newly generated face morphing dataset based on the SCFace dataset [118], performance evaluation protocols, and quantitative performance of the proposed D-MAD method.

6.4.1 SCFace-Morph Dataset (SCFM)

This work introduces a new dataset reflecting the OTF ABC systems with different camera resolutions and capture distances. We have employed the SCFace dataset by considering its applicability to real-life OTF face recognition with varying resolutions of the cameras and capture distances. The SCFace database is comprised of 130 data subjects that are captured with eight different cameras and three different distances, which are denoted as Distance-1 (D1) is 4.2m, Distance-2 (D2) is 2.6m, and Distance-3 (D3) is 1.0m. Face biometrics are captured as the data subjects walk (without a stop) through cameras held still with the frontal face capture. Since each data subject walks through these cameras, the captured data will be of unconstrained conditions with varying face poses that can represent the real-life ABC scenario. To effectively utilize the SCFace dataset for the face morphing application, we carefully selected the 77 unique data subjects by considering the image quality following ICAO standards [128]. Further, to reflect the real-life scenario of e-passport and ABC gates, we have considered only five different cameras that can capture visible images and high-quality mug shots. The Cam1 is of resolution 540 TVL, Cam2 is of resolution 480 TVL, Cam3 is of resolution 350 TVL, Cam4 is of resolution 460 TVL, and Cam5 is of resolution 480 TVL. Thus, the mugshot images represent the face images in the e-passport, and images captured using five different cameras and three different capture distances represent the trusted capture in the D-MAD algorithm evaluation. Next, we generate the face morphing dataset, the SCFace-Morph database (SCFM), using a mugshot images from the SCFace dataset. We have selected the neutral face pose image corresponding to each data subject to perform the face morphing using the landmark-based method from [119] by considering its attack potential. Before performing the morphing, the whole dataset of 77 data subjects is divided into two independent sets with 56 subjects in the training set and 21 subjects in the testing set. The face morphing dataset is generated following the guidelines presented in [129] that resulted in 92 face morphing images in the training set and 28 face morphing images in the testing set. The total number of probe image samples corresponding to the training set is 840 and the testing set is 315. The full statistics of the SCFace-Morph dataset are summarized in Table 6.2 which specifies the number of bona fide image samples, probe image samples for each camera and capture distance and the generated face morphing images. Figure 6.4 shows the example face images from SCFace-Morph database dataset. To reflect the real-life scenario of border control, we also generate the re-digitized version of both morphing and bona fide mugshot images by performing print and scan operations. We have used Ricoh IM C6000 Color Laser multi-function printer and the scanner from same printer. Facial images are scanned to have 300 dpi to match the requirement of ICAO standards [128]. Thus, the newly generated database has both digital and print-scan (PS) medium.

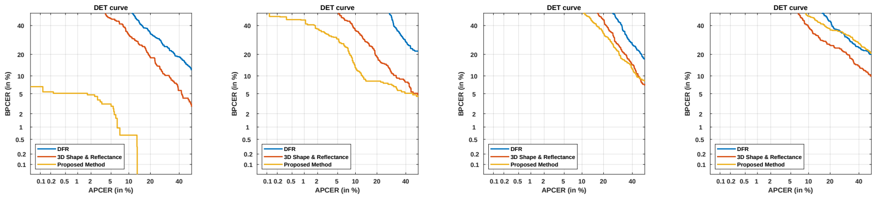
	Digital Images		
	Bona fide Passport	**Probe Images All cameras and Distances**	**Morphed Passport**
Train	56	840	92
Test	21	315	28

	Probe Images per Camera and Distance														
	Distance1					**Distance2**					**Distance3**				
	Cam1	**Cam2**	**Cam3**	**Cam4**	**Cam5**	**Cam1**	**Cam2**	**Cam3**	**Cam4**	**Cam5**	**Cam1**	**Cam2**	**Cam3**	**Cam4**	**Cam5**
Train	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56
Test	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21

Table 6.2: Statistics of SCFace-Morph Dataset (SCFM)

6.4.2 Performance evaluation Protocols

To effectively benchmark the performance of the proposed and existing D-MAD methods, we propose three different performance evaluation protocols by considering data medium (Digital/PS), camera resolution and capture distances. **Protocol 1** is designed to analyze the performance of the D-MAD techniques with both intra and inter medium experiments independently performed on camera and capture distance. Thus, Protocol 1 will benchmark the generalization of the D-MAD methods for the different morph data medium and their performance impact on



(a) (b) (c) (d)

Figure 6.5: DET Curves for Protocol 2(a) Train: Digital and Test: Digital (b) Train: Print Scan and Test: Print Scan (c) Train: Digital and Test: Print Scan (d) Train: Print Scan and Test: Digital

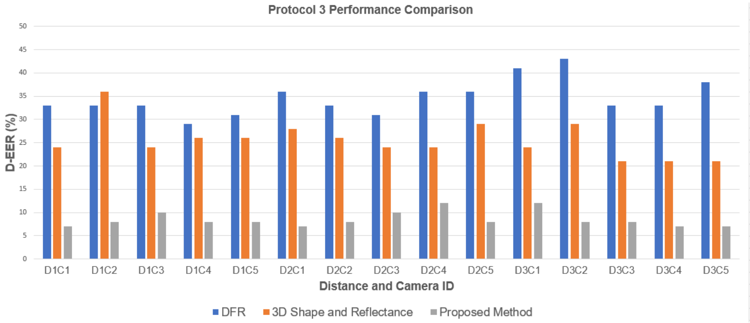


Figure 6.6: Illustration showing D-EER for Protocol 3 for DFR [8], 3D Shape and Reflectance [9], and Proposed Method

the image resolution and capture distance. **Protocol 2** is designed to benchmark the performance of the D-MAD techniques on intra and inter medium irrespective of the camera resolution and the capture distance. Thus, this protocol will use all camera and distance data to train and test the D-MAD methods. Hence, this protocol will indicate the generalization performance for a different medium. **Protocol 3** is designed to benchmark the performance of the D-MAD for individual cameras and capture distance. The D-MAD algorithms are trained and tested in this protocol by merging the Digital and PS data independently for each camera and capturing distance. This protocol will reflect the real-scenario testing as all type of data is used of training the D-MAD algorithm. This protocol will indicate the generalisation for different camera resolution and capture distance irrespective of data medium.

6.4.3 Experimental results

In this section, we present the quantitative results of the proposed method together with SOTA algorithms on D-MAD, namely Deep Feature Representation (DFR)

Algorithm:	Distance1					Distance2					Distance3				
	Cam1	Cam2	Cam3	Cam4	Cam5	Cam1	Cam2	Cam3	Cam4	Cam5	Cam1	Cam2	Cam3	Cam4	Cam5
	D-EER (%)														
	Train: Digital, Test: Digital														
DFR [8]	23.5	23.5	28.6	20.2	23.5	28.6	28.6	28.6	24.4	28.6	33.6	24.4	28.6	28.6	23.5
3D Shape and Reflectance [9]	14.3	23.5	14.3	15.2	20.2	19.3	24.4	32.7	23.5	19.3	23.5	24.4	19.3	23.5	23.5
Proposed Method	**0.9**	**5.1**	**0.0**	**8.3**	**5.1**	**5.2**	**0.9**	**5.1**	**8.3**	**9.2**	**9.2**	**4.2**	**5.1**	**5.1**	**9.2**
	Train: Print Scan, Test: Print Scan														
DFR [8]	28.6	29.5	28.6	28.6	28.6	42.9	37.8	33.6	33.6	27.7	52.1	37.8	38.7	42.0	47.9
3D Shape and Reflectance [9]	19.3	23.5	42.0	23.5	23.5	19.3	20.2	33.6	27.7	23.5	28.6	32.7	23.5	28.6	23.5
Proposed Method	**10.1**	**14.3**	**15.2**	**14.3**	**9.2**	**18.5**	**18.5**	**14.3**	**19.3**	**15.2**	**19.3**	**14.3**	**13.4**	**19.3**	**19.3**
	Train: Digital, Test: Print Scan														
DFR	28.6	28.6	28.6	32.7	29.5	33.6	34.5	33.6	32.7	28.6	42.9	32.7	28.6	37.8	33.6
3D Shape and Reflectance	23.5	37.8	32.7	29.5	**28.6**	**19.3**	38.7	**28.6**	28.6	37.8	27.7	28.6	24.4	27.7	27.7
Proposed Method	28.6	23.5	28.6	29.5	29.5	32.7	**24.4**	**28.6**	37.8	28.6	33.6	33.6	34.5	37.8	33.6
	Train: Print Scan, Test: Digital														
DFR [8]	28.6	29.5	28.6	28.6	28.6	47.0	**28.6**	**23.5**	29.5	**23.5**	37.8	**24.4**	**28.6**	32.7	47.0
3D Shape and Reflectance [9]	**24.4**	33.6	42.9	33.6	32.7	**27.7**	**28.6**	24.4	**23.5**	27.7	**32.7**	29.5	29.5	**28.6**	**24.4**
Proposed Method	32.7	**23.5**	37.8	33.6	33.6	33.6	37.8	33.6	37.8	37.8	42.9	33.6	33.6	34.5	23.5

Table 6.3: Quantitative results of proposed method and SOTA on Protocol 1

Algorithm:	Train: Digital Test: Digital			Train: Print Scan Test: Print Scan			Train: Digital Test: Print Scan			Train: Print Scan Test: Digital		
	D-EER (%)	BPCER @ APCER =		D-EER (%)	BPCER @ APCER =		D-EER (%)	BPCER @ APCER =		D-EER (%)	BPCER @ APCER =	
		5%)	10%)		5%)	10%)		5%)	10%)		5%)	10%)
DFR [8]	27.1	69.5	56.8	34.7	93.3	89.2	34.3	89.2	84.1	30.9	84.1	67.3
3D Shape and Reflectance [9]	19.4	45.4	32.7	19.4	50.2	36.2	27.6	75.2	60.6	**23.8**	57.5	38.7
Proposed Method	**3.4**	3.2	0.6	**11.5**	29.8	13.3	**26.0**	68.3	52.1	32.4	59.4	47.6

Table 6.4: Quantitative results of proposed method and SOTA on Protocol 2

Algorithm:	Distance1					Distance2					Distance3				
	Cam1	Cam2	Cam3	Cam4	Cam5	Cam1	Cam2	Cam3	Cam4	Cam5	Cam1	Cam2	Cam3	Cam4	Cam5
	D-EER (%)														
	Train: Digital and Print Scan										Test: Digital and Print Scan				
DFR [8]	32.7	32.7	32.7	28.6	30.7	35.7	32.7	30.7	35.7	35.7	40.8	42.9	32.7	32.7	37.8
3D Shape and Reflectance [9]	23.5	35.7	24.0	26.0	26.0	28.1	26.0	24.0	24.4	29.0	24.0	28.6	21.4	21.0	21.4
Proposed Method	**6.7**	**7.6**	**9.7**	**7.6**	**7.6**	**7.1**	**7.6**	**9.7**	**12.2**	**7.6**	**11.8**	**7.6**	**7.6**	**7.1**	**7.1**

Table 6.5: Quantitative results of proposed method and SOTA on Protocol 3

[109] and 3D Shape and Reflectance [9]. We choose the DFR method by considering its robust performance on the NIST FRVT benchmark [130] and 3D Shape and Reflectance is selected by considering its application in the OTF ABC based face morphing detection. The quantitative performance of the D-MAD techniques is presented using the ISO/IEC metrics [131] namely the ´´Attack Presentation Classification Error Rate (APCER (%)), which defines the proportion of attack images (face morphing 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 [131] along with the Detection Equal Error Rate (D-EER (%))´´ [132]. Table 6.3 indicates the quantitative performance of the D-MAD techniques on Protocol 1. Based on the obtained results, the proposed method has indicated improved results when the medium is preserved during training and testing (Intra evaluation). Further, the proposed method has indicated the best performance in the intra evaluation protocol irrespective of the cameras and capture distances. When the medium changes during training and testing (inter evaluation), the proposed method has indicated improved performance when training is performed on Digital and testing is performed on Print Scan.

The degraded performance of the proposed method is noted primarily in the inter evaluation when print-scan data is used for training, and digital data is used for testing. This can be attributed to the limitation of the proposed method to generalization, especially with the different image quality (because the quality of print-scan is different from that of digital) that might be due to the lack of generalized features extracted from six different pre-trained CNNs. In general, the proposed D-MAD method has better performance than the SOTA methods on Protocol 1.

Table 6.4 indicates the quantitative performance of the D-MAD techniques on Protocol 2, which is shown as DET curves in Figure 6.5. It can also be noted in this protocol that the proposed method has indicated improved performance when compared with the existing methods. The proposed method shows the best performance in the intra evaluation protocol and comparable performance with the inter evaluation protocol. The results indicate that the proposed method is robust to both camera resolutions and capture distances.

Table 6.5 shows the performance of the proposed method on Protocol 3. Based on the obtained results, it can be noted that the proposed method has indicated the best performance on both cameras and different capture distances. Further, the performance of the proposed method is not influenced by the camera type and capture distance. Figure 6.6 graphically illustrates the D-EER performance of the D-MAD techniques on Protocol 3.

Based on the series of experiments performed, it can be noted that the D-MAD algorithms are generally influenced by the camera resolution and the capture distance. Further, the data medium will strongly influence the performance of the D-MAD algorithms in the unconstrained ABC scenario.

6.5 Conclusions and Future-Work

In this paper, we have presented a novel method for robust D-MAD in the ABC gate scenario. The proposed method is developed based on the six different pre-trained deep CNN combined using hierarchical fusion. The novelty of the proposed method is in the use of spherical interpolation computed by SLERP to perform the residual feature fusion. Further, the hierarchical fusion is carried out using both score and feature level to achieve the robust D-MAD. Extensive experiments on the newly generated face morphing dataset (SCFM) based on the publicly available SCFace database. The performance of the proposed method and the existing techniques are extensively evaluated using three different protocols. The evaluation protocols are designed to benchmark the D-MAD performance on the different camera resolutions and the capture distance. The obtained results have demonstrated the improved performance of the proposed method in all three

protocols.
The future work includes improving the generalizability of the proposed method across different morphing image quality. Moreover, the proposed method will be submitted to the public benchmarks, including NIST FRVT MORPH. Further, the database will be extended to have different print-scan and morphing methods.

6.6 Supplementary Material

This supplementary material presents the additional ablation results of the proposed method. We devised two experiments such that *Experiment 1:* We report both individual and intermediate results of the proposed method. *Experiment 2:* This experiment is designed to indicate the efficacy of the proposed pair selection by performing the ablation study on the different pairs. In the following, we briefly discuss the outcome of the ablation study with both experiments.

6.7 Quantitative results of Experiment 1

Table 6.6, 6.7 and 6.8 indicates the performance of the proposed method and different components used to develop the proposed method evaluated in all three protocols respectively. It can be noted that:

- The performance of the individual network varies with the train and test data type. Typically, individual CNN networks perform better when trained and tested with the same data type.
- Fusion of individual networks indicates the improved performance over the individual CNN networks based on the proposed pair selection algorithm. This intermediate fusion result is shown as SLERP Residue 1 and SLERP Residue 2.
- The proposed method has indicated the best results compared to individual and intermediate fusion results on all three protocols.

These quantitative results indicate the improved performance of the proposed method in all three protocols.

6.8 Quantitative results of Experiment 2

The objective of this experiment is to justify the pair selection introduced as part of the proposed method. Since pair selection is made within the groups, we proposed the pair permutation as indicated in Table 6.9. Tables 6.10, 6.11 and 6.12 indicate the quantitative results of the proposed method with different pairs in all three protocols. It can be noted that:

- The proposed algorithm for the pair selection has indicated the improved performance over other possible pairs as indicated in the Table 6.10, 6.11 and 6.12.

- It can be noted that the proposed pair selection did not always show the best performance in protocol 1. However, the performance of the proposed pair is comparable. The proposed pair selection shows superior performance in average statistics, as shown in Table 6.10.

- The proposed pair selection indicates superior performance on protocols 2 and 3. Thus, these results indicated the efficacy of the proposed pair selection algorithm, which is an integral part of the proposed method to reduce the computation without compromising the detection performance.

Thus, experiments further justify the efficacy of the proposed method for reliable face morphing detection, especially in the mixed resolution and distance ABC scenario.

Algorithm:	**Distance1**					**Distance2**					**Distance3**				
	Cam1	**Cam2**	**Cam3**	**Cam4**	**Cam5**	**Cam1**	**Cam2**	**Cam3**	**Cam4**	**Cam5**	**Cam1**	**Cam2**	**Cam3**	**Cam4**	**Cam5**
	D-EER (%)														
	Train: Digital, Test: Digital														
Alexnet	14.3	14.3	14.3	18.5	19.3	18.5	14.3	14.3	14.3	15.2	19.3	19.3	15.2	14.3	10.1
VGG16	52.1	33.6	33.6	33.6	42.9	28.6	33.6	33.6	47.9	42.9	47.9	33.6	42.9	42.9	37.8
VGG19	33.6	33.6	33.6	23.5	23.5	29.5	28.6	42.9	24.4	33.6	52.1	47.9	28.6	29.5	34.5
Resnet50	14.3	18.5	15.2	14.3	14.3	20.2	19.3	23.5	23.5	14.3	9.2	14.3	9.2	19.3	23.5
Xception	9.2	9.2	9.2	13.4	9.2	9.2	9.2	8.3	9.2	13.4	15.2	9.2	10.1	9.2	9.2
Resnet101	5.1	9.2	8.3	9.2	9.2	10.1	10.1	8.3	14.3	9.2	19.3	14.3	14.3	13.4	14.3
SLERP Residue 1	19.3	18.5	15.2	23.5	19.3	27.7	19.3	19.3	19.3	19.3	27.7	23.5	33.6	24.4	28.6
SLERP Residue 2	24.4	19.3	9.2	19.3	24.4	37.8	23.5	28.6	42	28.6	23.5	28.6	24.4	28.6	23.5
Proposed Method	**0.9**	**5.1**	**0.0**	**8.3**	**5.1**	**5.2**	**0.9**	**5.1**	**8.3**	**9.2**	**9.2**	**4.2**	**5.1**	**5.1**	**9.2**
	Train: Print Scan, Test: Print Scan														
Alexnet	23.5	15.2	15.2	23.5	19.3	24.4	20.2	28.6	23.5	23.5	24.4	23.5	24.4	20.2	23.5
VGG16	38.7	28.6	34.5	33.6	33.6	32.7	37.8	37.8	33.6	33.6	47.9	37.8	38.7	37.8	27.7
VGG19	42	28.6	29.5	23.5	24.4	23.5	42.9	28.6	23.5	37.8	42.9	33.6	25.3	42.9	29.5
Resnet50	28.6	28.6	42.9	33.6	28.6	29.5	37.8	42.9	33.6	42.9	33.6	23.5	33.6	28.6	33.6
Xception	29.5	34.5	29.5	27.7	28.6	23.5	28.6	33.6	38.7	28.6	27.7	32.7	33.6	33.6	33.6
Resnet101	19.3	23.5	24.4	23.5	23.5	18.5	28.6	23.5	28.6	23.5	28.6	23.5	23.5	28.6	33.6
SLERP Residue 1	24.4	27.7	28.6	37.8	23.5	27.7	27.7	33.6	37.8	37.8	23.5	28.6	28.6	33.6	37.8
SLERP Residue 2	28.6	32.7	33.6	37.8	28.6	28.6	47.9	33.6	37.8	37.8	37.8	29.5	27.7	33.6	32.7
Proposed Method	**10.1**	**14.3**	**15.2**	**14.3**	**9.2**	**18.5**	**18.5**	**14.3**	**19.3**	**15.2**	**19.3**	**14.3**	**13.4**	**19.3**	**19.3**
	Train: Print Scan, Test: Digital														
Alexnet	28.6	27.7	23.5	28.6	23.5	23.5	23.5	32.7	33.6	32.7	23.5	33.6	33.6	37.8	33.6
VGG16	38.7	37.8	28.6	28.6	37.8	28.6	37.8	33.6	32.7	33.6	38.7	33.6	37.8	33.6	37.8
VGG19	42	38.7	38.7	43.8	37.8	32.7	37.8	42.9	33.6	33.6	47.9	37.8	38.7	33.6	37.8
Resnet50	43.8	34.5	47.9	47.9	42.9	37.8	47.9	52.1	47	37.8	37.8	37.8	48.8	42.9	37.8
Xception	53	47.9	52.1	47.9	52.1	52.1	57.1	57.1	47.9	57.1	47.9	47.9	52.1	57.1	48.8
Resnet101	42.9	52.1	47.9	47	47.9	52.1	43.8	56.3	47.9	37.8	58	52.1	47.9	58	47
SLERP Residue 1	28.6	27.7	38.7	29.5	27.7	28.6	37.8	37.8	42.9	42	28.6	42.9	38.7	34.5	37.8
SLERP Residue 2	47.9	47.9	56.3	42.9	52.1	37.8	38.7	42.9	57.1	57.1	38.7	42.9	52.1	47.9	48.8
Proposed Method	**28.6**	**23.5**	**28.6**	**29.5**	**29.5**	**32.7**	**24.4**	**28.6**	**37.8**	**28.6**	**33.6**	**33.6**	**34.5**	**37.8**	**33.6**
	Train: Print Scan, Test: Digital														
Alexnet	37.8	33.6	28.6	33.6	33.6	33.6	28.6	28.6	33.6	37.8	33.6	28.6	37.8	28.6	28.6
VGG16	52.1	42.9	43.8	38.7	47	34.5	38.7	47.9	47.9	42.9	56.3	52.1	52.1	53	38.7
VGG19	53	34.5	47	39.6	56.3	42.9	52.1	42.9	48.8	57.1	52.1	56.3	42.9	53	33.6
Resnet50	29.5	33.6	47.9	29.5	23.5	28.6	43.8	42.9	42.9	42	34.5	29.5	37.8	34.5	33.6
Xception	42	43.8	53	33.6	37.8	33.6	37.8	47.9	57.1	34.5	42.9	37.8	37.8	37.8	37.8
Resnet101	37.8	37.8	38.7	42.9	37.8	42	37.8	43.8	42.9	52.1	42.9	33.6	37.8	47.9	52.1
SLERP Residue 1	33.6	28.6	28.6	39.6	37.8	28.6	33.6	28.6	42	33.6	37.8	37.8	42.9	47.9	29.5
SLERP Residue 2	47.9	47.9	47.9	47.9	47	29.5	38.7	42.9	37.8	47	42.9	53	47	32.7	37.8
Proposed Method	**32.7**	**23.5**	**37.8**	**33.6**	**33.6**	**33.6**	**37.8**	**33.6**	**37.8**	**37.8**	**42.9**	**33.6**	**33.6**	**34.5**	**23.5**

Table 6.6: Experiment1: Ablation Study Protocol 1

Medium	Alexnet	VGG19	VGG16	Resnet50	Xception	Resnet101	Residue 1	Residue 2	**Proposed**
Train: Digital,Test: Digital	7.3	32	38.7	12	5.8	14.9	15.5	15.2	**3.4**
Train: Print Scan, Test: Print Scan	20.7	28.5	34.7	26.3	27	34.6	21.6	28.2	**11.5**
Train: Digital, Test: Print Scan	22.5	34.3	34.7	36.8	49.5	36.8	25.7	43.3	**26.0**
Train: Print Scan, Test: Digital	40.6	41	40	35.2	37.5	40.3	31.2	43.8	**32.4**

Table 6.7: Experiment1: Ablation Study Protocol 2

Algorithm:	**Distance1**					**Distance2**					**Distance3**				
	Cam1	**Cam2**	**Cam3**	**Cam4**	**Cam5**	**Cam1**	**Cam2**	**Cam3**	**Cam4**	**Cam5**	**Cam1**	**Cam2**	**Cam3**	**Cam4**	**Cam5**
	D-EER (%)														
	Train: Digital and Print Scan									**Test: Digital and Print Scan**					
Alexnet	39.3	39.4	35.9	39.6	36.5	38.8	36.3	40	40	39.9	41.1	37.4	40.2	39.7	36.9
VGG16	46.4	48.5	47.6	46.9	47.3	49.1	48.5	46.9	48.2	47.3	50.3	49	50.7	49.1	49.4
VGG19	43.6	42	41.4	40.5	37.4	37.8	40.5	38.1	40.8	41.7	46	45.5	40.3	42	37.4
Resnet50	45.5	42	46.6	45.5	38.4	40.3	43.9	47.9	46.3	46.3	40.8	38.5	43	37.6	42.9
Xception	40.2	45.1	48.7	38.7	41.7	39.7	47.3	48.7	47.3	41.8	40	42.4	43.2	46.6	42.3
Resnet101	41.2	42	41.4	45.1	40.9	42.4	40.6	46	47	47.8	44.9	43.3	43.9	47.2	48.1
SLERP Residue 1	21.4	21.4	21.4	21.4	21.4	19.3	21.4	21.9	21.9	21.4	18.9	21	21.4	21.4	21.4
SLERP Residue 2	14.3	14.3	16.8	18.9	16.8	18.9	17.3	14.7	16.8	21.9	18.9	18.9	18.9	14.3	11.8
Proposed Method	**6.7**	**7.6**	**9.7**	**7.6**	**7.6**	**7.1**	**7.6**	**9.7**	**12.2**	**7.6**	**11.8**	**7.6**	**7.6**	**7.1**	**7.1**

Table 6.8: Experiment 1: Ablation Study Protocol 3

Pair Description	SLERP Residue 1	SLERP Residue 2
Proposed Method Pair	(Alexnet,VGG16), (VGG16,VGG19)	(Resnet50,Resnet101),(Resnet101,Xception)
Pair 1	(VGG16,VGG19), (VGG19,Alexnet)	(Resnet101,Xception),(Xception,Resnet50)
Pair 2	(VGG19,Alexnet),(Alexnet,VGG16)	(Xception,Resnet50),(Resnet50,Resnet101)

Table 6.9: Description of pairs used in the proposed method and Experiment 2

Algorithm:	**Distance1**					**Distance2**					**Distance3**					**Mean D-EER%**
	Cam1	**Cam2**	**Cam3**	**Cam4**	**Cam5**	**Cam1**	**Cam2**	**Cam3**	**Cam4**	**Cam5**	**Cam1**	**Cam2**	**Cam3**	**Cam4**	**Cam5**	
	D-EER (%)															
	Train: Digital, Test: Digital															
Proposed Method	**0.9**	**5.1**	**0.0**	**8.3**	5.1	**5.2**	0.9	**5.1**	8.3	9.2	9.2	**4.2**	**5.1**	**5.1**	9.2	**5.4**
Pair 1	5.1	9.2	0.9	5.1	8.3	14.3	0	5.1	0	5.1	**8.3**	5.1	9.2	8.3	9.2	6.2
Pair 2	5.1	9.2	4.2	9.2	**4.2**	9.2	**0**	5.1	**0**	**5.1**	9.2	4.2	5.1	5.1	**8.3**	5.5
	Train: Print Scan, Test: Print Scan															
Proposed Method	**10.1**	**14.3**	**15.2**	**14.3**	**9.2**	18.5	**18.5**	**14.3**	**19.3**	15.2	19.3	14.3	**13.4**	**19.3**	19.3	**15.6**
Pair 1	15.2	14.3	23.5	19.3	9.2	14.3	20.2	19.3	27.7	15.2	20.2	**10.1**	19.3	19.3	19.3	17.8
Pair 2	13.4	15.2	23.5	19.3	13.4	**13.4**	19.3	19.3	23.5	**14.3**	**18.5**	15.2	15.2	23.5	**15.2**	17.5
	Train: Digital, Test: Print Scan															
Proposed Method	28.6	**23.5**	**28.6**	29.5	29.5	32.7	**24.4**	**28.6**	37.8	28.6	**33.6**	**33.6**	**34.5**	**37.8**	**33.6**	**31.0**
Pair 1	**23.5**	28.6	33.6	42.9	27.7	28.6	32.7	33.6	**28.6**	**24.4**	33.6	34.5	37.8	37.8	32.7	32.0
Pair 2	24.4	27.7	33.6	**28.6**	**27.7**	**24.4**	29.5	38.7	34.5	23.5	37.8	37.8	38.7	42.9	33.6	32.2
	Train: Print Scan, Test: Digital															
Proposed Method	**32.7**	**23.5**	**37.8**	**33.6**	33.6	33.6	**37.8**	**33.6**	**37.8**	**37.8**	**42.9**	**33.6**	**33.6**	**34.5**	**23.5**	**34.0**
Pair 1	37.8	33.6	42.9	28.6	29.5	33.6	37.8	33.6	42.9	34.5	42.9	33.6	38.7	42.9	33.6	36.4
Pair 2	33.6	29.5	42.9	37.8	**23.5**	**29.5**	37.8	37.8	47	37.8	47	33.6	37.8	42.9	27.7	36.4

Table 6.10: Experiment2: Protocol 1 Ablation Study with Pair 1 and Pair 2 whose description is provided in Table 6.9

D-EER (%)			
Medium	Proposed Method	Pair 1	Pair2
Train: Digital,Test: Digital	**3.4**	5	3.4
Train: Print Scan, Test: Print Scan	**11.5**	14.6	13.4
Train: Digital, Test: Print Scan	**26.0**	28.5	28.2
Train: Print Scan, Test: Digital	**32.4**	37.8	33.7

Table 6.11: Experiment2: Protocol 2 Ablation Study with Pair 1 and Pair 2 whose description is provided in Table 6.9

Algorithm:	**Distance1**					**Distance2**					**Distance3**				
	Cam1	**Cam2**	**Cam3**	**Cam4**	**Cam5**	**Cam1**	**Cam2**	**Cam3**	**Cam4**	**Cam5**	**Cam1**	**Cam2**	**Cam3**	**Cam4**	**Cam5**
	D-EER (%)														
	Train: Digital and Print Scan								**Test: Digital and Print Scan**						
Proposed Method	**6.7**	**7.6**	**9.7**	**7.6**	7.6	**7.1**	7.6	**9.7**	**12.2**	7.6	**11.8**	**7.6**	7.6	**7.1**	**7.1**
Pair 1	9.7	11.8	9.7	7.6	7.1	9.2	7.6	11.8	14.7	9.7	9.7	11.8	7.1	10.1	10.1
Pair 2	9.7	11.8	11.8	11.8	**7.1**	9.2	**7.1**	10.1	13.8	**7.1**	9.7	10.1	**7.1**	9.7	10.1

Table 6.12: Experiment2: Protocol 3 Ablation Study with Pair 1 and Pair 2 whose description is provided in Table 6.9

Chapter 7

Article 3:Fusion of Deep Features for Differential Face Morphing Attack Detection at Automatic Border Control Gates (RQ1)

Jag Mohan Singh, and Raghavendra Ramachandra. Fusion of Deep Features for Differential Face Morphing Attack Detection at Automatic Border Control Gates .In *IEEE European Workshop on Visual Information Processing (EUVIP 2022)*, Lisbon, Portugal, pp. 1-5, IEEE

7.1 Abstract

Face recognition systems (FRS) are showing increasing accuracy in an uncontrolled real world, leading to their usage in automated border control (ABC) gates. However, both automatic and manual FRS are prone to Face Morphing Attacks (FMA), which can be generated by linearly blending face images from two contributory data subjects. Differential Morphing Attack Detection (D-MAD), which compares the face image in an electronic Machine Readable Travel Document (eMRTD) with a trusted live capture from an ABC gate, is thus a significant problem to address. This paper presents Robust D-MAD (RD-MAD), which performs pair-wise pose normalization on the bona fide and probe face images based on global affine alignment. The proposed D-MAD technique is based on the comparison score level fusion of deep features extracted using off-the-shelves pre-trained deep networks such as AlexNet and ResNet. The deep features are extracted corresponding to both enroled and probe face images independently from AlexNet

and ResNet. Then the signed difference of the features is computed between enroled and probe face images independently on both AlexNet and ResNet. Two linear SVMs are trained on the signed difference features corresponding to deep networks whose comparison scores are fused using the weighted sum rule to make the final decision. Extensive experiments are performed on a challenging dataset having lighting, face pose, expression, illumination and print-scan variations. Obtained results outperform the state-of-the-art (SOTA) as we achieve an EER=2.1% compared to the SOTA with an EER=8.6%.

7.2 Introduction

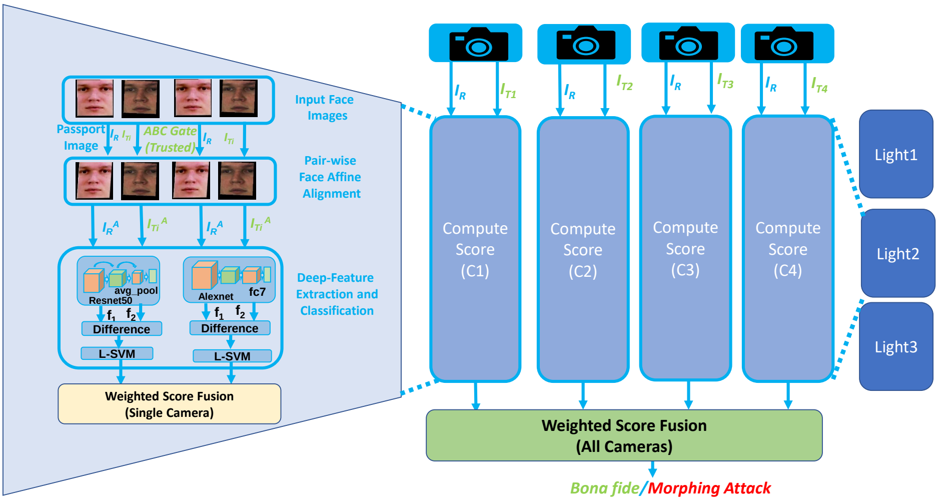


Figure 7.1: Block diagram of the proposed method

Face recognition systems (FRS) are known to achieve high accuracy in uncontrolled, real-world environments, mainly attributed to the evolution of deep-learning networks [22, 27]. The increased accuracy of FRS enabled the magnitude of face recognition applications, including border control. However, FRS is demonstrated as vulnerable to face morphing attacks generated by linear blending of face images from two contributory data subjects. To counter the face morphing attack, biometric researchers have devised Morphing Attack Detection (MAD) methods. The available MAD techniques can be divided into two types [69]: Differential Morphing Attack Detection (D-MAD) uses a reference while performing the classification and Single Morphing Attack Detection (S-MAD) is a no-reference based classification. A summary of MAD, including primary datasets, benchmarks and state-of-the-art methods, is provided in the recent survey [69]. Further, among the two morphing attack detection methods, D-MAD is more preferred by biometric researchers as it has higher real-world applicability with ABC gates. Since ABC gates provide a trusted live capture which is used a probe (reference) image and the enrollment image is from the passport, which allows the usage of reference-based D-MAD.

The process of face morphing image generation can be either based on landmarks [133] or Generative Adversarial Networks (GANs) [134]. Among these two morphing generation techniques, the landmark-based face morphing is more vulnerable and challenging to detect [134]. Therefore, in this work, we employ the landmark-based face morphing generation. The D-MAD techniques are well explored in the literature. The available D-MAD methods are broadly classified into four types: (1) landmark-based, (2) face de-morphing, (3) texture and shape based (4) deep feature-based. However, these methods are less reliable due to variation in a pose during trusted capture and also landmarks are not robust enough to capture the variation in shape due to morphing. The face de-morphing [133] technique will invert the face morphing operation such that it can retrieve the hidden face. Thus, given the morphing image and the trusted capture image, the de-morphing technique can generate the image of other data subjects by inverting the morphing process. However, de-morphing is sensitive to the pose and lighting variation and requires prior information on face morphing that is not available in real-life conditions. The third type of D-MAD technique is based on the use of handcrafted features that includes micro-texture-based features (e.g., LBP, LPQ, BSIF) [10], shape-based features (HoG) [10] and scale-space features [135]. These features have a good performance on the controlled data settings. The fourth type of DMAD is based on the deep learning techniques [136]. However, due to the limited number of datasets and samples, almost all methods use the transfer learning of the off-the-shelf deep learning compared to the other three types. The deep features have indicated better performance in detecting face morphing attacks.

In this work, we propose a novel framework for the D-MAD using the fusion of deep features extracted from the aligned facial images. Given the facial image pair corresponding to enrol and probe capture, in the first step, the images are aligned and registered using piecewise-affine registration. We then process the registered image independently to extract deep features using pre-trained CNNs such as ResNet50 and AlexNet. The signed difference is computed between the deep components extracted independently for ResNet50 and AlexNet. These features are classified using linear SVM trained separately for ResNet50 and AlexNet. Finally, the scores are fused using the weighted sum rule. We repeat this process for all four cameras independently and combine the final scores corresponding to each camera using the weighted sum rule to compute the final decision. The following are the main contributions of this work:

- We present the affine registration-based facial alignment concept to D-MAD to improve the Detection Error Rate (D-EER%).
- The proposed method is evaluated on the Morph ABC Dataset comprised of 39 data subjects that are collected in real-life scenario in on-the-fly manner with three different lighting conditions.
- We benchmark our proposed method with previous state-of-the-art (SOTA), achieving the lowest EER=2.1% compared to the SOTA with an EER=8.6% for Light1, lowest EER=4.6% compared to the SOTA with an EER=10.5% for Light2, and the lowest EER=6.8% compared to the SOTA with an EER=13.5% for Light3.

The rest of the paper, we discuss the proposed method in Section 7.3, followed by experiments and results in Section 7.4 and conclusions and future work in Section 7.5.

7.3 Proposed Method

The proposed D-MAD techniques are based on the off-the-shelf deep CNN features that can extract the robust texture-based features for face morphing attack detection. Figure 9.1 shows the block diagram of the proposed method designed to work with the automatic border control gate with four cameras. Since the data capture scenario represents the on-the-fly capture (i.e., the data subject will not stop in front of the camera but instead walk in front of the camera). The challenging issue in this scenario is the variation in the facial pose; hence it is necessary to neutralize the pose before performing the D-MAD.

Since the ABC system used in this work is based on four cameras, every time data subjects walk through the system will capture four-probe images corresponding to four different cameras. The proposed method mainly consists of the three functional blocks that are operated on each camera (1) Pair-Wise Face Alignment, (2) Deep-Feature Extraction and Classification and (3) Weighted Score Fusion. In the following, we discuss the proposed D-MAD technique.

7.3.1 Pair-wise Face Alignment

Given the reference image I_R and the trusted image captured using ABC gate $I_{T1}, I_{T2}, I_{T3}, I_{T4}$, we perform the pair-wise image alignment between the reference image and the trusted capture images independently. Since face alignment is crucial in the D-MAD, we performed the following steps to achieve near-optimal

Light 1				
Algorithm	**Cam**	**D-EER**	**BPCER20**	**BPCER10**
LBP& SVM [10]	1	41.7±0.4	81.1±0.6	72.4±1.0
	2	42.7±0.5	82.5±0.6	73.5±0.8
	3	38.1±0.5	83.3±0.7	71.5±0.6
	4	39.6±0.3	79.6±0.5	71.1±0.4
	Fused	28.5±0.4	67.2±0.6	54.2±0.8
3D Shape & Diffuse Reconstruction [9]	1	18.1±0.1	36.3±0.7	27.1±0.3
	2	19.7±0.4	34.7±0.7	28.3±0.7
	3	19.1±0.1	35.9±0.1	27.3±0.1
	4	18.8±0.1	36.1±0.1	27.5±0.3
	Fused	8.6±0.1	13.9±0.4	7.5±0.1
DFR [8]	1	30.7±8.2	81.6±0.2	63.3±0.2
	2	30.7±8.2	67.3±0.2	53.1±0.2
	3	24.9±7.7	57.1±0.1	44.9±0.1
	4	26.2±7.8	44.9±0.1	36.7±0.1
	Fused	14.0±6.2	24.5±0.1	18.4±0.1
Proposed Method	1	18.4±6.9	36.7±0.1	22.4±0.1
	2	14.3±6.2	24.5±0.2	18.4±0.2
	3	11.9±5.8	16.3±0.1	12.2±0.1
	4	3.9±3.5	2.0±0.2	2.0±0.2
	Fused	**2.1±2.5**	**2.0±0.1**	**0.0±0.0**

Table 7.1: Quantitative performance of the proposed method and SOTA on Light 1 where BPCER20 is BPCER@APCER=5%, and BPCER10 is BPCER@APCER=10%

face alignment and registration. We now describe the process of our proposed alignment method for piecewise affine alignment for a pair of images I_R, I_{Ti} as follows:

Facial Key-Point Detection

We detect the facial five key points using Dlib [42] as this would lead to the creation of large triangles during the next step of Delaunay.

Delaunay Triangulation with fewer keypoints

We compute the triangulation using the Delaunay algorithm of the facial key points detected in the previous step. It is known that Delaunay could lead to skinny triangles (slivers) according to Shewchuck et al. [137] and a few facial keypoints

Light 2				
Algorithm	**Cam**	**D-EER**	**BPCER20**	**BPCER10**
LBP & SVM [10]	1	35.9±7.4	77.6±0.1	62.7±0.1
	2	34.1±7.3	79.1±0.1	74.6±0.1
	3	40.3±7.6	88.1±0.1	64.2±0.1
	4	34.5±7.3	80.6±0.1	74.6±0.1
	Fused	34.3±7.3	62.7±0.1	53.7±0.1
3D Shape & Diffuse Reconstruction [9]	1	17.9±5.9	32.8±0.1	23.9±0.1
	2	14.9±5.5	29.9±0.1	16.4±0.1
	3	13.5±5.3	31.3±0.1	16.4±0.1
	4	19.4±6.1	34.3±0.1	26.9±0.1
	Fused	10.5±4.7	14.9±0.1	10.4±0.1
DFR [8]	1	34.5±7.3	61.2±0.1	53.7±0.1
	2	37.3±7.4	79.1±0.2	58.2±0.1
	3	37.5±7.5	79.1±0.1	68.7±0.1
	4	34.1±7.3	76.1±0.1	68.7±0.1
	Fused	28.4±6.9	52.2±0.1	44.8±0.1
Proposed Method	1	10.5±4.7	16.4±0.1	10.4±0.1
	2	13.5±5.3	26.9±0.1	13.4±0.1
	3	6.0±3.6	9.0±0.1	3.0±0.1
	4	7.5±4.1	16.4±0.1	6.0±0.1
	Fused	**4.6±3.2**	**4.5±0.1**	**3.0±0.1**

Table 7.2: Quantitative performance of the proposed method and SOTA on Light 2 where BPCER20 is BPCER@APCER=5%, and BPCER10 is BPCER@APCER=10%

lying outside the facial region are reasons for hole creation during the face image morphing process. Cheng et al. [43] have mentioned in their book that both constrained Delaunay and conformal Delaunay triangulation can be used to generate triangles with good areas. However, we have devised a simple approach of using fewer facial keypoints 5 instead of 68 to create better quality fewer triangles.

Affine Transformation

We now estimate a global affine transformation between the pair of face images using two key points from the eyes and one from the nose. The results of this alignment are shown in Figure 7.2.

Light 3				
Algorithm	**Cam**	**D-EER**	**BPCER20**	**BPCER10**
LBP& SVM [10]	1	41.2±7.2	79.2±0.1	72.2±0.1
	2	38.9±7.2	77.8±0.1	68.1±0.1
	3	34.8±7.0	81.9±0.1	75.0±0.1
	4	40.4±7.2	87.5±0.1	81.9±0.1
	Fused	33.4±6.9	79.2±0.1	68.1±0.1
3D Shape & Diffuse Reconstruction [9]	1	17.6±5.9	26.9±0.1	20.9±0.1
	2	19.4±6.1	44.8±0.1	31.3±0.1
	3	10.5±4.7	16.4±0.1	10.4±0.1
	4	17.9±5.9	32.8±0.1	22.4±0.1
	Fused	13.5±5.3	19.4±0.1	17.9±0.1
DFR [8]	1	29.2±6.7	61.1±0.1	50.0±0.1
	2	33.3±6.9	72.2±0.1	52.8±0.1
	3	33.3±6.9	73.6±0.1	65.3±0.1
	4	37.8±7.1	83.3±0.1	69.4±0.1
	Fused	27.8±6.6	55.6±0.1	38.9±0.1
Proposed Method	1	16.8±5.5	36.1±0.1	27.8±0.1
	2	16.8±5.5	33.3±0.1	23.6±0.1
	3	13.9±5.1	22.2±0.1	15.3±0.1
	4	11.1±4.6	38.9±0.1	16.7±0.1
	Fused	**6.8±3.7**	**8.3±0.1**	**4.2±0.1**

Table 7.3: Quantitative performance of the proposed method and SOTA on Light 3 where BPCER20 is BPCER@APCER=5%, and BPCER10 is BPCER@APCER=10%

7.3.2 Deep-Feature Extraction and Classification

In this step we extract features from pre-trained deep-networks namely from the ‘avg_pool’ (average pooling) layer from Resnet50 [138] and ‘fc7’ layer from Alexnet [97]. Note that the choice of networks is based on their generalization ability on Imagenet Dataset [123]. The signed difference of features is computed and passed through Linear-SVM for classification. We perform a weighted fusion of scores from Resnet50 and Alexnet with fusion weights of 0.3 and 0.7 to generate the camera-based score. Note that the weights for the fusion are chosen based on grid-search.

7.3.3 Weighted Score Fusion (All Cameras)

In this step, we perform weighted score fusion for the scores obtained from each camera to generate the final classification of Bonafide/Morphing Attack. Note the weights used for fusion of each camera-based score are $w1 = 0.1, w2 = 0.1, w3 = 0.1, w4 = 0.7$ for cameras $C1, C2, C3, C4$ for Light 1, $w1 = 0.15, w2 = 0.15, w3 = 0.35, w4 = 0.35$ for cameras $C1, C2, C3, C4$ for Light 2 and $w1 = 0.25, w2 = 0.25, w3 = 0.25, w4 = 0.25$ for cameras $C1, C2, C3, C4$, respectively. Note that the weights for the fusion are chosen based on grid-search [139].

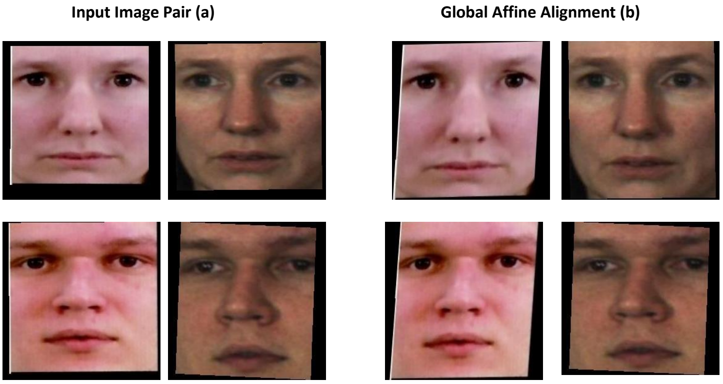


Figure 7.2: Illustration showing the global affine alignment for an input pair of images from our dataset. Notice the slant generated in the face images post alignment.

7.4 Experiments and Results

This section discusses experiments performed and results obtained using the proposed method and SOTA. The SOTA D-MAD techniques employed in this work are benchmarked on the Morph ABC Dataset [9] discussed below.

7.4.1 Morph ABC Dataset

We use the Morph ABC Dataset [9] which consists of 39 subjects, an overall 270 face morphing images and 1549 ABC gate probe images. Further, the face morphing and bona fide images were print-scanned using Epson XP-860 Printer and Scanner and more details can be found in the article by Singh et al. [9] and shown in Figure 7.3. The dataset consists of four cameras and three different lights

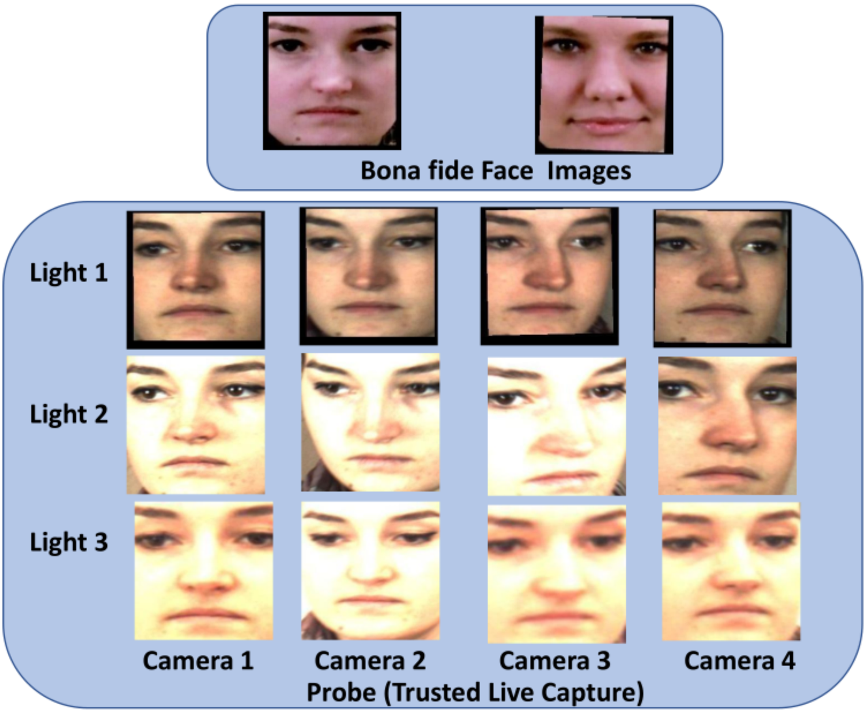
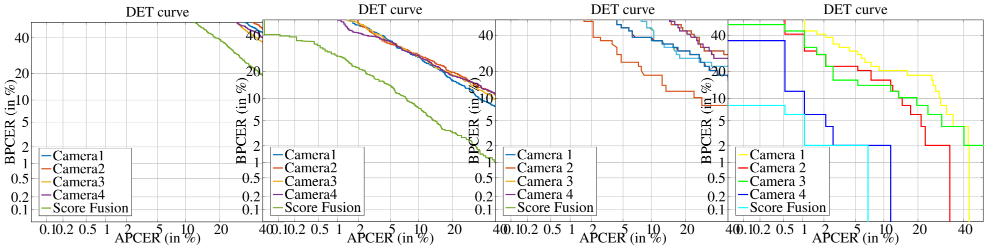
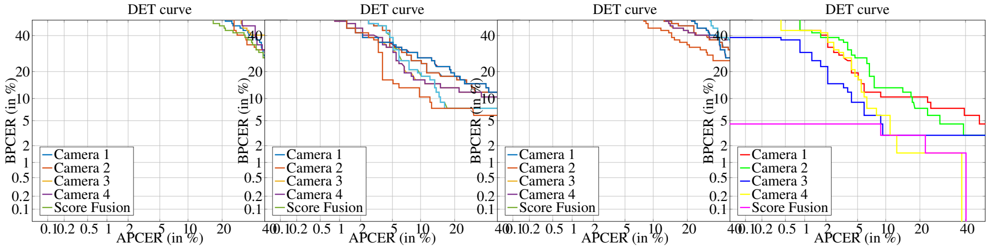


Figure 7.3: Illustration showing the Morph ABC Dataset, with real-world environment conditions of pose, expression, illumination and capture distance.



(a) **(b)** **(c)** **(d)**

Figure 7.4: DET Curves for (a) LBP-SVM [10], (b) 3D Shape & Diffuse Reconstruction [9], (c) DFR [8] (d) Proposed Method. DET Curves are for Scores from Camera1, Camera2, Camera3, Camera4, and Weighted Sum-Rule Fusion of scores from these individual cameras for Light 1.



(a) (b) (c) (d)

Figure 7.5: DET Curves for (a) LBP-SVM [10], (b) 3D Shape & Diffuse Reconstruction [9], (c) DFR [8] (d) Proposed Method. DET Curves are for Scores from Camera1, Camera2, Camera3, Camera4, and Weighted Sum-Rule Fusion of scores from these individual cameras for Light 2.

to simulate different environmental conditions, where the first light is set to an intensity of 180 lux to represent a dark overcast day. The second light is set to an intensity of 450 lux to represent the light during sunrise or sunset, which comes inside an office hallway, and finally, the third light is set to an intensity of 1500 lux to represent the light inside an office hallway during a bright day.

7.4.2 Results and Discussion

We report the metrics according to ISO/IEC 30107-3 PAD metrics [131]: Attack Potential Classification Error Rate (APCER) representing the mis-classification error rate of Attacks Samples, Bona fide Potential Classification Error Rate (BPCER) representing mis-classification error rate of Bona fide Samples. Further, we report Detection-Equal Error Rate (D-EER) and present results figuratively in the form of DET Curves. Tables 7.1, 7.2 and 7.3 shows the results of the proposed method in comparison with SOTA in tabular format and for Light1, Light2 and Light3 respectively and Figures 7.4 and Figure 7.5 shows them figuratively as DET Curves for Light1 and Light2 respectively.

We now discuss the results obtained using SOTA and the proposed method. Note that the proposed method shows much higher performance than SOTA [9], which can be attributed to two main reasons, namely facial alignment and score-fusion using the original image. SOTA method by Singh et al. [9] performs score-level fusion based on diffuse reconstructed image and normal map. However, we perform score-level fusion based on an aligned image as the generated diffuse reconstructed image and normal map are not of high quality. Further, using both Alexnet and Resnet-50 for feature extraction leads to more generalization than using Alexnet alone for feature extraction from the diffuse reconstructed image as was done in SOTA [9]. The proposed method performs well on Light 1 which represents a dark overcast day compared to Light 2, which represents sunrise/sunset or Light 3. This can be attributed to the fact that Light 1 has low lighting and thus the dynamic range of these images is low, resulting in a smoother image, as seen from Figure 7.3.

7.5 Conclusions and Future-Work

In this paper, we presented a novel method for D-MAD using a fusion of deep features which outperforms the SOTA. Further, we presented an approach for facial image alignment, which is crucial in improving the detection accuracy in D-MAD techniques. Our proposed approach for facial alignment is based on global affine alignment since the pose variation in input face images is not high. However, our alignment technique does not handle non-rigid deformations. Further, our alignment technique does not generate holes in the face morphing image. Hence, in future work, we want to work on facial alignment without generating holes but still handle non-rigid deformations.

Chapter 8

Article 4: Robust Face Morphing Attack Detection Using Fusion of Multiple Features and Classification Techniques(RQ2)

Jag Mohan Singh, Sushma Venkatesh and Raghavendra Ramachandra. Robust Face Morphing Attack Detection Using Fusion of Multiple Features and Classification Techniques. In *IEEE Fusion, 2023*, In 26^{th} International Conference on Information Fusion 2023, South Carolina, USA, pp. 1-8

8.1 Abstract

The face morphing process will combine two or more facial images to generate a single morphed facial image demonstrating Face Recognition Systems (FRS) vulnerability. The attack potential of the morphing image directly depends on the perceptual image quality, and when generated with no visible artefacts, it can deceive both human observers and automatic FRS. The current softwares for face morphing generates a morphing image with ghosting artefacts, especially in the eye region, nose and mouth area, which may serve as a potential cue to detect morphing attacks. Hence in this work, we introduce a new dataset comprising 10710 facial images before and after manual post-processing to reduce the visual artefacts and to generate high-quality attacks. Further, we propose a novel single image-based Morph Attack Detection (S-MAD) technique based on the ensemble of features and classifiers using the scale-space domain. The novel concept in the proposed method is the multi-level fusion that combines the comparison scores from dif-

ferent features and classifiers. Extensive experiments are carried out on the newly generated high-quality face images with (i) Morphs before post-processing and (ii) Morphs after post-processing. Further, the experiments are also carried out on two different mediums such as (i) Digital and (ii) Print-scan (or re-digitized) with and without compression. Extensive experimental results are performed to benchmark the detection performance with the existing S-MAD techniques. Obtained results indicate the best performance of the proposed method over existing methods.

8.2 Introduction

Biometrics has been widely studied and applied globally for person identification [140]. The trustworthiness of biometric features has gained immense popularity over multi-factor authentication. Among several other physiological modalities like a fingerprint, palmprint, finger vein and iris, face biometrics-based applications have had a wide range of applications for several decades. The face is a unique modality and humans easily identify an individual based on facial features. As identification of a person based on facial features can be achieved through the naked eye, facial biometrics has been well accepted for national ID programs and security-related applications, especially in highly secure places such as border control scenarios.

Although Face Recognition Systems(FRS) are widely installed to provide reliable person identification and recognition, it also encounters threats due to various attacks that highlight the vulnerability of FRS. Presentation attacks, adversarial attacks, and imposter attacks are some example attacks that pose a risk to the reliable performance of FRS [29, 141]. In addition to these attacks, a face-morphing attack is one such attack that can efficiently make the FRS vulnerable, especially in border control applications. Although face morphing was initially performed merely for entertainment, it has gradually transformed into a potential threat in the recent past [142]. As face morphing is achieved by blending the facial features of two or more facial identities to generate a morphing image, this will lead to the vulnerability of FRS to reliably recognize the person.

Based on the International Civil Aviation Organisation (ICAO) recommendation, the face is the prominent modality employed for person recognition and verification in the border control scenario [143, 144]. Hence all passport holders must enroll their facial image in the eMRTD to serve as an identification document for border control authority during travel. Face enrolment procedure varies with the country's passport application procedure. Scandinavian countries have installed a photo booth to perform live capture of the facial image [145]. However, most Asian countries accept printed passport-size facial images during the application process [146]. But New Zealand, Ireland and the UK have a web portal where the applicant has to upload the facial image for the passport renewal process [147, 148]. Even though the facial image undergoes manipulation and makes it easier to identify the existence of morphing, the availability of a variety of high-quality morphing software makes it challenging even for an expert human observer.

Several open-source morphing software yield superior quality morphed facial image that does not require any technical expertise [45, 149, **?**, 150, 47]. Hence a person with malicious intentions can easily generate a morphed facial image with a look-alike accomplice's facial image and successfully submit for the passport enrolment process. As it is challenging to detect unknown facial identities from the morphing image, even a trained border control official finds it difficult to detect the existence of morphing [151, 152]. Eventually, the morphing facial image will be enrolled in the eMRTD that can be claimed by both the identities involved in the morphing process. This disregards the rule of single ownership for the passport/eMRTD document and eventually creates a loophole in the security. Considering the risk of face morphing and its impact on building a secure society, extensive research has been performed to generate robust techniques for Morph Attack Detection(MAD) [153, 88, 154, 155, 156, 157, 9, 158]. Based on the MAD techniques developed by several researchers, morph attack detection techniques can be broadly classified into single image-based MAD (without reference image) and differential image-based MAD (with reference image). S-MAD techniques are applicable where single facial image-based person verification is required. In the case of the passport renewal process in Ireland, [159], since it is an online passport service, the applicant's facial image must be uploaded into the web portal. As no supervision exists while uploading the facial image into the web portal, an applicant with malicious intentions may end up uploading the morphed facial image.

Hence several researchers have investigated the problem of face morphing and developed reliable techniques. The first work on the S-MAD technique is investigated by Raghavendra et al. [160] using the texture-based approach. Since then, several S-MAD approaches have been proposed that can be broadly divided into [69] three types (1) Hand-crafted features: These techniques include the different types of features such as: texture-based [161] [162] , time-frequency based [163], color based [88], residual noise [164], image quality based [165, 166, 167] (2) Deep learning features: These includes the use of pre-trained deep CNN networks [168, 169, 170], fusion of pre-trained CNNs [171, 163], pixel based DCNN MAD [172] (3) Hybrid Features: These MAD techniques are based on using multiple features and classifiers for face morphing detection. The outcome of the multiple classifiers is combined at either feature or comparison level. Several works proposed in this category includes [173], [174], [163], [175]. Among these techniques, the hybrid approaches have indicated the best performances in detecting face-morphing attacks.

All the available State-Of-The-Art (SOTA) techniques are evaluated on the morphed datasets that are not manually and professionally post-processed. Even though the early work [176] attempts to use the manual post-processing morphs, the dataset size is tiny. In this work, we introduce a new dataset to benchmark the S-MAD techniques' performance systematically. The new dataset is constructed using different mediums such as: digital, print-scan using a DNP printer and print-scan using a Canon printer. We have used standard (Canon) and sublimation (DNP) printers to study the influence of printer noise on face morphing attack detection. The new dataset consists of a total of 10710 facial images before and after post-processing. Further, we have also proposed a new S-MAD technique based on the multi-level fusion of ensemble features and classifiers.

To efficiently evaluate the performance of the proposed MAD technique and its performance over SOTA MAD techniques, we investigate the following research questions that facilitate this study.

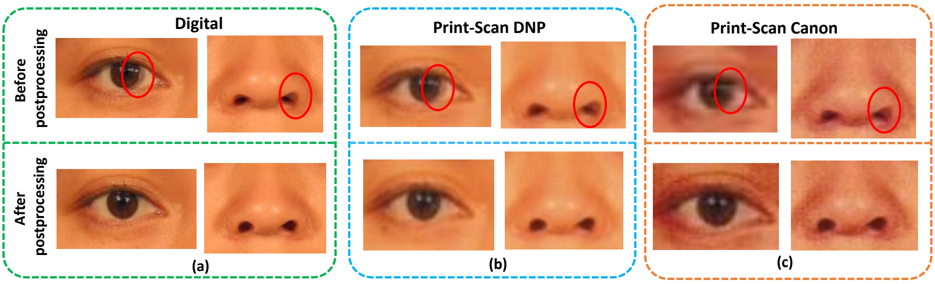


Figure 8.1: Illustration of issues of morphing before and after post-processing database from (i) Digital (ii) Print-Scan from DNP (PS-1) (iii) Print-Scan from Canon (PS-2)

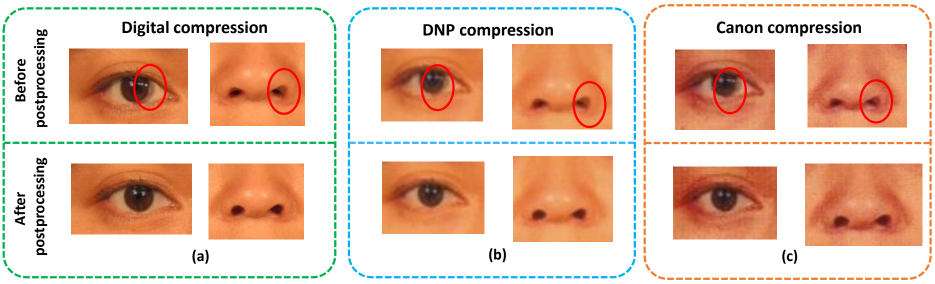


Figure 8.2: Illustration of before and after post-processing database from (i) Digital compression (ii) DNP compression (PS-1) (iii) Canon compression (PS-2)

- **Q1** Does the performance of the proposed method improves when the morph attack detection is performed on post-processed morphing images compared with the morph images before post-processing?
- **Q2** Is the proposed method generalizable for morphed facial images generated from various mediums and the morphing images before and after post-processing?

In the course of answering the research questions as mentioned above, the following are the main contributions of this work:

- We present a novel S-MAD approach based on the multi-level fusion of ensemble features and classifiers to detect face-morphing attacks reliably.
- We introduce a new dataset with manual post-processing to achieve high-quality face morphing images free from morphing noise and artefacts. The new dataset is collected using three different mediums that include both digital and two different printers.
- Extensive experiments are carried out to benchmark the detection performance of the proposed method on three different mediums with and without post-processing. Further, the influence of image compression on detection performance is also benchmarked.
- The detection performance of the proposed method is benchmarked with the existing S-MAD techniques in two different experimental protocols.

The rest of the paper is organised as follows: Section 8.3 details the newly generated dataset. Section 8.4 presents the proposed method using an ensemble of features and classifiers. Section 8.5 details the experimental protocols and corresponding results. Section 8.6 provides a discussion on the observation made from the experimental results. Finally, Section 8.6 concludes the current work.

8.3 Face Morphing Dataset

This section presents a new facial morphing dataset constructed using high-quality face images sampled from FRGC V2. The facial images are carefully selected to meet the enrolment guidelines, including zero pose, no shading on the face region, and no occlusion. The new dataset comprises 147 unique data subjects, further divided into two independent groups for training and testing. The training partition consists of 77 unique data subjects and the testing partition consists of 70 unique data subjects. In the next step, we perform the face morphing operation separately on the training and testing set. In this work, we employ the open-source face morphing tools [177, 119] based on landmarks. Further, we have used only two face images with equal weights to perform morphing based on the earlier studies [84, 176] that have indicated high vulnerability on FRS.

In general, the morphing process will result in various types of noises, especially in the eyes and nose region. These noises include double edges in the eye region and the spreading of edges in the nose region. Figure 8.1 illustrates the noises resulting from the morphing process that can be attributed to the variation in the geometry of the faces used for morphing. Even though these morphing noises are not common but exist in most cases, as shown in Figure 8.1, the morphing noises can also be predominantly observed even after the print-scan process. However, the quality of the print-scan process can also affect the visibility of edge spreading, as shown in Figure 8.1(c). Further, as noticed from Figure 8.2, even after the images are compressed to follow the guidelines of ICAO [178, 179], the morphing noises are still visible in both digital and print-scan versions. Therefore, it is essential to post-process the morphing face image to weed out these noises so that the human observer cannot identify the morphing based on these noises.

Image Type	before post-processing	after post-processing	Total
Digital images	1071	1071	2142
Print & Scan	1071X2 (printers)	1071 X 2 (printers)	4284
Print & Scan compression	1071X2 (printers)	1071X2 (printers)	4284
Total	5355	5355	10710

Table 8.1: Total number of morphing images before and after manual post-processing.

Table 8.2: Database statistics: training and testing partitions

Data Partition	**Data Type**					
	Digital		**PS-1**		**PS-2**	
	Bona fide	**Morph**	**Bona fide**	**Morph**	**Bona fide**	**Morph**
Training	689	517	689	517	689	517
Testing	583	554	583	554	583	554

Table 8.1 tabulates the statistics of the newly developed face morphing dataset with morphing samples before and after manual post-processing. The manual post-processing is carried out using Adobe Photoshop [180] to obtain professional-quality passport face images. Figure 8.1 and 8.2 illustrates the manual post-processing images in which the morphing noises are corrected to achieve the highest quality of the morphed face images. In this work, face morphing uses the alpha value (or

morphing factor) of 0.5 by considering the highest vulnerability demonstrated in several earlier works [84, 176].

We first generate the face morphing images separately for the training and testing sets. In the next step, we used two different printers, which are DNP and a Canon printer, to digitize the digital images by print & scan. The DNP printer used in this work is the dye-sublimation photo printer that can generate the highest quality passport face images and is widely deployed in photo studios. In contrast, the CANON PIXMA printer is a conventional inkjet printer used for printing passport face images. We term the data generated using the DNP printer as PS-1 and CANON PIXMA printer as PS-2, respectively. Figure 8.2 illustrates the example images from the newly developed datasets before and after manual post-processing.

8.3.1 Dataset partition: Train and Test

To effectively evaluate the Morph Attack Detection (MAD) algorithms, the whole dataset is partitioned into two independent sets: training and testing. The training set consists of 77 unique data subjects and the testing partition consists of 70 unique data subjects. The morphing images are generated by using the data subjects within each partition. Thus, the training set comprises 689 bona fide and 517 morph face images. Table 8.2 indicates the statistics of training and testing independently for morphing samples before and after manual post-processing.

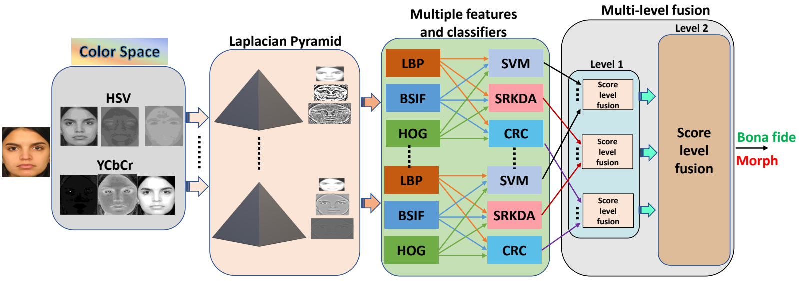


Figure 8.3: Block diagram of the proposed method

8.4 Proposed Method

Figure 8.3 shows the block diagram of the proposed method leveraged on the multi-level score level fusion of multiple features. The main objective of the proposed method is to exploit the complementary features of the different feature extractors and classifiers combined at two different levels. We assert that the use of complementary features and classification scores can capture the discriminant information useful for reliable face morph detection. The proposed method is designed using four different functional units, namely: (a) color space, (b) scale-space decomposition, (c) multiple features and classifiers (d) multi-level fusion. We discuss each of the functional units in detail in the following subsections.

8.4.1 Color space representation

Given the input image I, the first step is to extract the different color spaces using YC_bC_r and HSV. We have selected these two color spaces by considering their robustness to capture the morphing noises as is demonstrated in earlier works [88]. Thus, for the given image I, we get six different representations such as:I_{Col} = $I_H, I_S, I_V, I_Y, I_{C_b}, I_{C_r}$.

8.4.2 Scale-Space decomposition

In the next step, we extract the scale-space features on each color space image using the Laplacian pyramid [181]. The choice of Laplacian pyramid-based scale-space features extraction is made by considering the effectiveness in extracting the discriminant features compared to similar techniques such as steerable pyramids [175]. We use three-level decomposition on each color image based on their empirical evaluation. Thus, given the color image I_H, the corresponding scale-space images can be represented as I_{H1}, I_{H2}, I_{H3}. In this work, we have used six different color channels and thus, the corresponding scale-space representation will result in $6 \times 3 = 18$ sub-images that are independently processed to extract the multiple features. Let the sub-images be represented as: $SI_k = SI_1, SI_{2,} \ldots, SI_{18}, \forall k = 1, 2, \ldots, 18$.

8.4.3 Multiple features and classifiers

Multiple features and classification systems used in this work are based on three types of feature extraction and three different classifiers. Three different feature extraction techniques include Local Binary Patterns (LBP), Histogram of Gradients (HoG) and Binary Statistical Image Features (BSIF). These three features are selected by considering the complementary features that include texture features extracted using both hand-crafted and naturally learned in addition to the gradient information. These features represent the image's different characteristics, especially the pixel discontinuities, and thus can provide rich information to detect the morphing processing. Given the sub-image $SI_k, \forall k = 1, 2, \ldots, 18$, three different types of features are extracted independently.

In the next step, we employ three different types of classifiers, including linear Support Vector Machine (SVM) [182], Spectral Regression Kernel Discriminant Analysis (SRKDA) [183] and Probabilistic Collaborative Representation Classifier (P-CRC) [184]. We have considered these three classifiers by considering the high performance and robustness of various data sources[69]. Further, the non-availability of the large-scale morphing database justifies the choice of the ensemble of these three classifiers to achieve reliable morph detection. Given the features independently from the three different feature extraction techniques, we independently obtain the comparison scores from three different classification techniques.

8.4.4 Multi-level fusion

This work proposes the two-level fusion of comparison scores obtained using multiple classifiers. The first level of fusion will combine the comparison scores obtained using individual classifiers corresponding to three different feature extraction techniques. Therefore, first-level fusion has three independent fusion units corresponding to three independent classifiers. In the second level, we combine the comparison scores from the first level corresponding to individual classifiers to make the final decision. The multi-level fusion is designed based on empirical experiments that have indicated superior performance compared to serial fusion. At both levels, we have used the weighted sum rule to perform the fusion and weights are computed using the bootstrap method [185] on the development dataset and kept constant through the experiments.

8.5 Experiments and Results

In this section, we present and discuss the proposed method's quantitative results and the existing methods such as Hybrid features [88] and Ensemble features [154]. We particularly select these two existing methods as (1) these methods indicate the best performance in several reported studies [69] and one of them is benchmarked on the NIST FRVT morph [186] (2) these methods are based on the hand-crafted features thus are more appropriate to be compared with the proposed method (3) these methods are more appropriate by considering the size of the databases used in this work. The use of deep learning methods may result in overfitting due to the small datasets. The performance of the S-MAD techniques is benchmarked using ISO/IEC 30107-3 [187] metrics such as Attack Presentation Classification Error Rate (APCER (%)), Bona fide Presentation Classification Error Rate (BPCER(%)) and Detection-Equal Error Rate (D-EER(%)).

8.5.1 Experimental protocols:

To effectively evaluate the performance of the MAD algorithms using the proposed method, our experiments are categorized into three different protocols discussed as follows:

- **Experiment-1: Intra-dataset evaluation:** is performed within the same dataset type. This evaluation protocol performs training and testing on the

Table 8.3: Experiment-1: Quantitative results of MAD algorithms on different datasets

Dataset	**Post-processing**		**MAD Algorithms**	**Detection Performance**			**Detection Performance**		
				D-EER (%)	**BPCER @APCER**		**D-EER (%)**	**BPCER @APCER**	
					=5%	**=10%**		**=5%**	**=10%**
	Training	**Testing**		**without compression**			**with compression**		
Digital	Before	Before	**Proposed Method**	**0**	**0**	**0**	**0**	**0**	**0**
			Ensemble Features [154]	0.18	0	0	0.18	0	0
			Hybrid Features [88]	0	0	0	0.18	0	0
	After	After	**Proposed Method**	**0.18**	**0**	**0**	**0.36**	**0**	**0**
			Ensemble Features [154]	0.18	0	0	0.36	0	0
			Hybrid Features [88]	0.18	0	0	0.18	0	0
PS-1	Before	Before	**Proposed Method**	**0**	**0**	**0**	**3.45**	**2.47**	**1.02**
			Ensemble Features [154]	0	0	0	4.27	3.6	1.71
			Hybrid Features [88]	0	0	0	5	5.14	2.4
	After	After	**Proposed Method**	**0**	**0**	**0**	**3.09**	**2.22**	**1.02**
			Ensemble Features [154]	0	0	0	3.28	2.4	1.54
			Hybrid Features [88]	0	0	0	4.46	4.28	2.74
PS-2	Before	Before	**Proposed Method**	**10.00**	**15.01**	**10.66**	**7.72**	**11.83**	**7.2**
			Ensemble Features [154]	11.00	16.98	11.66	8.72	11.66	8.06
			Hybrid Features [88]	14.09	29.33	19.38	8.54	14.4	6.86
	After	After	**Proposed Method**	**5.74**	**6.34**	**3.75**	**5.19**	**5.14**	**2.91**
			Ensemble Features [154]	6.01	7.03	4.11	5.19	5.14	3.77
			Hybrid Features [88]	8.56	12.34	7.2	5.64	6.17	2.91

Table 8.4: Experiment-2: Quantitative performance of MAD algorithms on before post-processing data generated using different morphing types

Training data	**Testing Data**	**MAD Algorithms**	**Detection Performance**			**Detection Performance**		
			D-EER (%)	**BPCER @APCER**		**D-EER (%)**	**BPCER @APCER**	
				=5%	**=10%**		**=5%**	**=10%**
			without compression			**with compression**		
Digital	PS-1	**Proposed Method**	**31.90**	**78.38**	**67.12**	**31.64**	**77.53**	**67.58**
		Ensemble Features [154]	38.27	87.82	80.27	38.09	89.02	81.3
		Hybrid Features [88]	37.72	88.67	78.9	35.73	86.96	76.67
	PS-2	**Proposed Method**	**47.08**	**94.68**	**88.16**	**45.45**	**92.79**	**86.96**
		Ensemble Features [154]	50	97.77	93.31	50	97.42	92.28
		Hybrid Features [88]	50	96.22	93.65	50	94.85	90.73
PS-1	Digital	**Proposed Method**	**3.63**	**2.22**	**0.68**	**5.26**	**5.48**	**3.77**
		Ensemble Features [154]	8.09	14.23	6.51	8.09	15.6	6.51
		Hybrid Features [88]	8.54	13.2	7.54	20.72	43.91	33.1
	PS-2	**Proposed Method**	**20.9**	**55.74**	**41.16**	**12.18**	**23.67**	**15.95**
		Ensemble Features [154]	19.72	45.11	33.44	13.18	21.09	15.26
		Hybrid Features [88]	24.91	57.11	45.45	13.36	22.98	18.01
PS-2	Digital	**Proposed Method**	**9.45**	**16.63**	**8.74**	**9.63**	**16.98**	**9.6**
		Ensemble Features [154]	21.09	46.31	36.02	17.63	40.13	26.92
		Hybrid Features [88]	9.63	20.41	9.26	23.18	43.05	33.1
	PS-1	**Proposed Method**	**12.27**	**30.1**	**18.09**	**0.16**	**0**	**0**
		Ensemble Features [154]	14.27	28.64	19.72	0.16	0	0
		Hybrid Features [88]	19.72	42.19	31.73	0.72	0	0

Table 8.5: Experiment-2: Quantitative performance of MAD algorithms on after post-processing data generated using different morphing types

Training data	Testing Data	MAD Algorithms	Detection Performance			Detection Performance		
			D-EER (%)	BPCER @APCER		D-EER (%)	BPCER @APCER	
				=5%	=10%		=5%	=10%
			without compression			with compression		
Digital	PS-1	**Proposed Method**	**31.24**	**78.55**	**65.69**	**31.87**	**78.9**	**68.09**
		Ensemble Features [154]	34.15	83.53	76.67	37.52	87.13	77.53
		Hybrid Features [88]	38.06	89.87	80.78	37.7	88.67	81.3
	PS-2	**Proposed Method**	**35.6**	**87.3**	**71.18**	**36.15**	**87.47**	**74.09**
		Ensemble Features [154]	39.25	93.31	84.21	42.26	94.51	84.56
		Hybrid Features [88]	44.44	91.25	82.16	41.71	91.59	80.61
PS-1	Digital	**Proposed Method**	**4.09**	**3.94**	**2.91**	**6.82**	**9.43**	**5.83**
		Ensemble Features [154]	9.37	15.43	8.06	8.19	12	12.17
		Hybrid Features [88]	20.03	32.76	27.44	28.14	60.2	49.05
	PS-2	**Proposed Method**	**12.12**	**25.27**	**15.32**	**7.47**	**10.69**	**6.83**
		Ensemble Features [154]	13.02	26.75	16.63	8.19	12	6.86
		Hybrid Features [88]	24.86	50.08	42.19	10.47	16.46	10.46
PS-2	Digital	**Proposed Method**	**10.47**	**21.09**	**11.49**	**11.84**	**23.15**	**13.2**
		Ensemble Features [154]	18.48	45.11	31.9	15.93	38.59	25.38
		Hybrid Features [88]	11.1	26.92	13.89	23.13	44.94	36.87
	PS-1	**Proposed Method**	13.93	28.98	19.55	**0**	**0**	**0**
		Ensemble Features [154]	9.92	18.18	9.94	0.16	0	0
		Hybrid Features [88]	18.48	41.16	30.7	0.55	0	0

Table 8.6: Experiment-3: Quantitative performance of MAD algorithms by training after post-processing data and testing before post-processing data generated using different morphing types

			Detection Performance			**Detection Performance**		
Training data	**Testing Data**	**MAD Algorithms**	**D-EER (%)**	**BPCER @APCER**		**D-EER (%)**	**BPCER @APCER**	
				=5%	**=10%**		**=5%**	**=10%**
			without compression			**with compression**		
Digital	PS-1	**Proposed Method**	**30.72**	**80.44**	**68.95**	**31.9**	**80.44**	**67.92**
		Ensemble Features [154]	36.18	86.96	79.07	37.72	88.67	79.41
		Hybrid Features [88]	37.54	89.7	81.3	35.99	86.1	78.55
	PS-2	**Proposed Method**	**46.45**	**94.16**	**88.67**	**46.63**	**93.31**	**87.99**
		Ensemble Features [154]	50.27	97.25	93.31	51.27	96.91	92.1
		Hybrid Features [88]	50.9	94.16	90.39	52.09	95.54	91.76
PS-1	Digital	**Proposed Method**	**5.45**	**5.83**	**3.6**	**8.9**	**13.55**	**8.57**
		Ensemble Features [154]	13.1	25.55	18.52	13.18	25.72	16.12
		Hybrid Features [88]	9.82	19.72	9.6	27.09	60.89	51.11
	PS-2	**Proposed Method**	**17.9**	**37.77**	**26.7**	**11.81**	**22.29**	**12.52**
		Ensemble Features [154]	18.54	41.16	29.5	11.99	20.06	12.69
		Hybrid Features [88]	28.08	56.43	45.62	12.36	21.95	15.09
PS-2	Digital	**Proposed Method**	**11.81**	**21.44**	**14.92**	**13.72**	**31.73**	**21.09**
		Ensemble Features [154]	22.08	50.94	39.1	18.81	44.59	28.47
		Hybrid Features [88]	13.36	29.15	17.15	22.63	46.68	34.47
	PS-1	**Proposed Method**	**12.72**	**21.56**	**13.32**	**0.3**	**0**	**0**
		Ensemble Features [154]	13.18	22.81	16.46	0.16	0	0
		Hybrid Features [88]	18.99	40.13	30.7	0.72	0	0

Table 8.7: Experiment-3: Quantitative performance of MAD algorithms by training before post-processing data and testing after post-processing data generated using different morphing types

Training data	Testing Data	MAD Algorithms	Detection Performance			Detection Performance		
			D-EER (%)	BPCER @APCER		D-EER (%)	BPCER @APCER	
				=5%	=10%		=5%	=10%
			without compression			with compression		
Digital	PS-1	**Proposed Method**	**32.05**	**79.93**	**69.12**	**32.6**	**79.07**	**71.01**
		Ensemble Features [154]	36.52	87.82	80.96	38.88	89.36	80.96
		Hybrid Features [88]	39.25	89.87	83.87	37.7	88.5	80.78
	PS-2	**Proposed Method**	**37.25**	**90.05**	**75.64**	**36.15**	**88.67**	**74.95**
		Ensemble Features [154]	40.61	95.54	87.13	42.34	95.71	86.44
		Hybrid Features [88]	45.44	93.13	83.87	41.16	91.93	80.96
PS-1	Digital	**Proposed Method**	**3.46**	**2.57**	**1.02**	**5.64**	**6.68**	**3.94**
		Ensemble Features [154]	8.92	15.6	7.54	9.92	16.63	9.77
		Hybrid Features [88]	17.3	32.76	24.52	22.13	45.45	34.81
	PS-2	**Proposed Method**	**16.4**	**40.96**	**29.81**	**10.29**	**20.92**	**10.97**
		Ensemble Features [154]	17.66	42.02	31.73	10.65	18.01	11.32
		Hybrid Features [88]	24.31	50.08	41.16	11.84	19.03	13.55
PS-2	Digital	**Proposed Method**	**9.74**	**16.46**	**8.91**	**10.47**	**17.32**	**10.97**
		Ensemble Features [154]	22.41	47.51	37.9	17.66	42.53	28.47
		Hybrid Features [88]	9.74	25.9	9.6	24.04	47.68	37.56
	PS-1	**Proposed Method**	**12.48**	**23.04**	**16.26**	**0.16**	**0**	**0**
		Ensemble Features [154]	13.93	26.75	18.01	0.16	0	0
		Hybrid Features [88]	20.94	44.25	34.3	0.55	0	0

same dataset type. As shown in Table 8.3, the three dataset types (digital, PS-I and PS-II) are independently evaluated before and after post-processing. For instance, the digital dataset type before post-processing is trained and the same dataset type is tested. A similar protocol is followed for the digital dataset type after post-processing, followed by the two different print-scan dataset types PS-I (before and after post-processing) and PS-II (before and after post-processing). All experiments are carried out with and without compression.

- **Experiment-2: Inter-medium evaluation**: is performed to analyze the MAD performance of the proposed method in cross-dataset types. This protocol is designed to investigate the robustness of the proposed method when it is trained and tested on different dataset types (digital, PS-1 and PS-2) generated from different mediums (digital, print-scan with and without compression). Tables 8.4 and 8.5 indicates the two different experiments performed for cross-dataset evaluation in the inter-medium scenario. Among the three dataset types employed in this work, we train one dataset type and test it on the other two. For instance, if the digital dataset type is trained, the two different print-scan dataset types, PS-I and PS-II, are tested. The same evaluation protocol is followed for the two print-scan dataset types. To better evaluate the cross-dataset performance of the proposed method, we have performed two different experiments (i) inter-medium evaluation before post-processing and (ii) inter-medium evaluation after post-processing.

- **Experiment-3: Inter-medium varied post-processing**: is performed to evaluate the performance of MAD in cross datasets generated from various mediums (digital, print-scan with and without compression) in both before and after post-processing scenarios. Tables 8.6 and 8.7 indicates the two experiments conducted for inter-medium and varied post-processing scenario. Two different experiments were conducted to evaluate the proposed method's performance. Following the similar experimental protocol as inter-medium evaluation, the first experiment is performed by (i) training the dataset types after post-processing and testing the dataset types before post-processing. The second experiment is performed by training the dataset types before post-processing and testing after post-processing.

8.5.2 Experimental results

In this section, we present the quantitative results of the proposed method and the existing methods of the three different evaluation protocols. The quantitative results obtained from the three different protocols designed for intra-dataset evaluation, inter-medium evaluation and inter-medium with varied post-processing evaluation scenarios are tabulated in the Tables 8.3, 8.4, 8.5, 8.6, 8.7.

Results on Experiment-1: Intra-dataset evaluation

Based on the obtained results presented in Table 8.3 following are the main observations:

- The proposed method has indicated the best performance on all three data mediums before and after post-processing. Thus, the proposed method has emerged as the best-performing method before and after post-processing.

- The detection performance of the existing methods also indicates the competitive performance, especially with digital and PS-1 data mediums both before and after post-processing.

- The detection performance of the S-MAD techniques indicates the degraded performance, especially with the PS-2 data medium that can be noticed before and after post-processing data. Thus, the morph generation quality will impact the detection accuracy of both the proposed and existing S-MAD techniques.

- Performing the post-processing indicates the impact on the detection performance. In some cases, the detection performance of the proposed method and the existing methods indicates improvement. This can be attributed to the possible variations in the image quality that might have resulted from post-processing operation. However, with data compression, the performance difference is not noticeable.

- The performance of the S-MAD algorithms also varies with and without compression, irrespective of the post-processing.

Results on Experiment-2: Inter-medium evaluation

Table 8.4 and 8.5 indicates the quantitative performance of the proposed method together with existing methods in Experiment 2. Based on the obtained results following can be noted:

- The Inter-medium training and testing indicate the drastic degradation of the detection accuracy of both the proposed method and the existing methods. The degradation is noticed both before and after post-processing.

- The S-MAD algorithms degrade more when algorithms are trained with digital and tested against PS-I and PS-II. Less degradation is noted when S-MAD algorithms are trained with PS-I and tested against digital and PS-II. Similar degradation is noticed both before and after post-processing.
- The S-MAD algorithms have indicated a better detection accuracy on the print-scan compression when compared to without compression, especially on the before post-processing data. However, the S-MAD algorithms did not show much difference in the detection performance on the before post-processing data. This indicates that using the post-processing data to train and test the S-MAD algorithms might be key to achieving the generalisation in cross-medium experiments.
- Based on the experimental results in Experiment-2, the proposed method has indicated the best performance compared to existing methods on both before and after post-processing data.

Experiment-3: Inter-medium varied post-processing

In this section, we discuss the quantitative results of the proposed method and the existing S-MAD techniques, especially to study the influence of post-processing operation versus different mediums on detection accuracy. Tables 8.6 and 8.7 indicate the quantitative results of the S-MAD techniques, including the proposed method. Based on the obtained results, the following can be noted:

- The performance of the S-MAD algorithms indicates the degraded detection rate irrespective of the data post-processing type.
- In general, the performance of the S-MAD algorithms, including the proposed method, indicates the marginal improvement in the detection performance when trained using post-processed data irrespective of the data medium.
- The performance of the proposed method indicates the best performance compared with the existing methods, irrespective of the data type (before or after post-processing) used for the training. The best performance of the proposed method is when PS-1 is trained and tested on digital data before and after post-processing.

8.6 Discussion

The research questions formulated in Section 8.2 are answered below based on the extensive experiments conducted, obtained results and the observations made above.

- **Q1**. Does the performance of the proposed method improve when the morph attack detection is performed on post-processed morphed images when compared with the morph images before post-processing?
  - As noted by the obtained experimental results reported in Table 8.4 8.5, the performance of the proposed method shows a marginal improvement when used with the morph images after post-processing in Experiment-1, especially on the PS-2 data medium. However, the proposed method's performance did not significantly influence (even though the proposed method has shown little improvement in some cases) the post-processing in Experiment-2.
- **Q2**.Is the proposed method generalizable for morphed facial images generated from various mediums and also for the morphed images before and after post-processing?
  - Based on the experimental results (see Table 8.3 8.4, 8.5, 8.6, 8.7), the proposed method has indicated the best performance in two different experimental protocols.

Thus, based on the obtained results, one can attribute the improvements to using multiple features with multiple classifiers, which would increase generalization.

8.7 Conclusions and Future Work

Reliable face morphing attack detection using a single image is a challenging problem due to the variation in image quality attributed to the various source of the morph generation and digitisation processes. In this work, we proposed a new framework for S-MAD using multiple features and classifiers whose comparison scores are combined at multiple levels to detect face-morphing attacks reliably. We have also introduced a new dataset based on manual post-processing to generate high-quality face morphing images free from morphing artefacts. The dataset constructed has three different mediums: digital, Print-Scan (PS-1 re-digitised using DNP printer and PS-2 re-digitised using CANON printer) and print-scan compression. Extensive experiments are carried out using two different evaluation protocols to benchmark the performance of the proposed method together with the existing methods. The obtained results demonstrated the best performance of the proposed method in two different evaluation protocols compared with the existing methods. In future work, we could evaluate more advanced fusion techniques, benchmarking the proposed method and comparison with more SOTA approaches.

Chapter 9

Article 5: Deep Face Attribute Composition Attacks: Generation, Vulnerability and Detection (RQ3)

Jag Mohan Singh, and Raghavendra Ramachandra. Deep Face Attribute Composition Attacks: Generation, Vulnerability and Detection. In *IEEE Access, March 2023. pp. 76468 - 76485*

9.1 Abstract

Face manipulation attacks have drawn the attention of biometric researchers because of their vulnerability to Face Recognition Systems (FRS). This paper proposes a novel scheme to generate Composite Face Image Attacks (CFIA) based on facial attributes using Generative Adversarial Networks (GANs). Given the face images corresponding to two unique data subjects, the proposed CFIA method will independently generate the segmented facial attributes, then blend them using transparent masks to generate the CFIA samples. We generate 526 unique CFIA combinations of facial attributes for each pair of contributory data subjects. Extensive experiments are carried out on our newly generated CFIA dataset consisting of 1000 unique identities with 2000 bona fide samples and 526000 CFIA samples, thus resulting in an overall 528000 face image samples. We present a sequence of experiments to benchmark the attack potential of CFIA samples using four different automatic FRS. We introduced a new metric named Generalized Morphing Attack Potential (G-MAP) to benchmark the vulnerability of generated

attacks on FRS effectively. Additional experiments are performed on the representative subset of the CFIA dataset to benchmark both perceptual quality and human observer response. Finally, the CFIA detection performance is benchmarked using three different single image based face Morphing Attack Detection (MAD) algorithms. The source code of the proposed method together with CFIA dataset will be made publicly available: https://github.com/jagmohaniiit/LatentCompositionCode

9.2 Introduction

FRS demonstrates highly accurate verification rates, which has led to their widespread usage in eCommerce, online banking, surveillance and security applications. The recent advances in deep learning techniques have further increased the accuracy of the FRS [27], [22] that enabled them to be deployed in the border control applications. However, the FRS is vulnerable to various attacks, among which the face morphing attacks have gained attention due to their impact on the border control applications. Recent benchmarking results reported in NIST FRVT MOPRH [130] indicate that the higher the accuracy of the FRS, the higher the vulnerability for the morphing attacks.

One of the most widely used attacks toward FRS is the Presentation Attacks (PA), a.k.a spoofing attacks, which can be achieved by presenting a biometric artefact to the biometric capture device. PA can be performed by generating a Presentation Attack Instrument (PAI) that includes either a printed photo (print-photo), displaying an image (display-photo), displaying a video (replay-video), or the use of a rigid/non-rigid 3D face mask (mask-attack). Biometric researchers had thus devised Presentation Attack Detection (PAD) as a countermeasure to PA that is extensively discussed in [29], and [30].

The second type of widely studied attack on the FRS is the adversarial attack, which can be performed by applying a small perturbation (noise), a.k.a adversarial perturbation, to a facial image. Even though the introduced perturbation is indistinguishable to the human eye but can lead to mis-classification with high-confidence [188] and can be used to expose vulnerabilities of the FRS. Adversarial attacks have shown high vulnerability in FRS, especially on the deep learning-based FRS [189]. The white box adversarial attack requires complete knowledge of the underlying deep learning model.. Adversarial attacks could also be black-box attack performed during testing, and the attacker does not know the underlying deep-learning model. Several countermeasures to address the adversarial attacks are extensively discussed in [190], [191]. It needs to be pointed out that adversarial attacks are digital when performed on images, but they can also be performed in the physical world by using a unique eyeglass for impersonation [190].

Face morphing attacks are gaining high momentum in the biometric community. The face morphing process seamlessly combines face images from two or more subjects (also called contributory subjects) to generate a morphing image. The generated morphing image shows substantial visual similarity to both contributory subjects therefore challenging to detect by the experts (border guards and police) [69, 142, 115, 192]. Notably, the morphed images will get verified to both the contributory subjects when used with automatic FRS [69]. Therefore, the morphing attacks can be instrumented to acquire the ID documents like passports, driving licenses, bank accounts, etc. For example, a subject with criminal background can obtain a passport by collaborating with an accomplice to generate a morphing image. Then, the accomplice can apply for an ID document using the morphed image. The subject with a criminal background can use the obtained ID document to cross the border [69, 142].

Face morphing can be generated using algorithms based on facial landmarks such as Face Morpher [193] and UBO-Morph [194]. More recently, algorithms based on Generative Adversarial Networks (GANs) such as MorGAN [48], MIPGAN [51] and ReGenMorph [50] have also been used to generate face morphing images. These generated face morphing images have demonstrated the high vulnerability of FRS, especially in the passport application scenario, including automatic border control. Further, morphing attacks can deceive both human observers (border control officers) and automatic FRS in Automatic Border Control (ABC) [115, 69], [115]. Following the initial paper [142], there have been several papers on morphing detection, and the reader is advised to refer to the survey by Venkatesh et al. [69] to get a detailed overview on face morphing.

Most face morphing generation works are devised by performing the blending operation on the complete (or total) face images [69]. However, the success rate of the full-face morphing attack is high when contributory subjects are lookalikes to deceive the super-recognizer and highly trained border guards [115]. Therefore, partial face morphing was introduced in [195] where the blending operation is carried out using Poisson image editing [196]. The generated composite morphs have shown vulnerabilities of FRS based on deep-learning features such as VGG-Face [28], Arcface [22] and commercial-off-the-shelf (COTS) that includes Neurotech [197] and Cognitec [198]. Further, the human observer analysis is also discussed. However, the work presented in [195] has several limitations, including (1) it is based on landmarks, and this would lead to pixel-based artifacts due to alignment issues, and correction of these would require manual intervention [69] (2) Only a few arbitrary regions are used to generate the composite images (3) Limited only to the base regions like nose, mouth, eye and forehead. (4) Limited only to the single facial attribute composite generation (5) failure to achieve a high vulnerability of FRS.

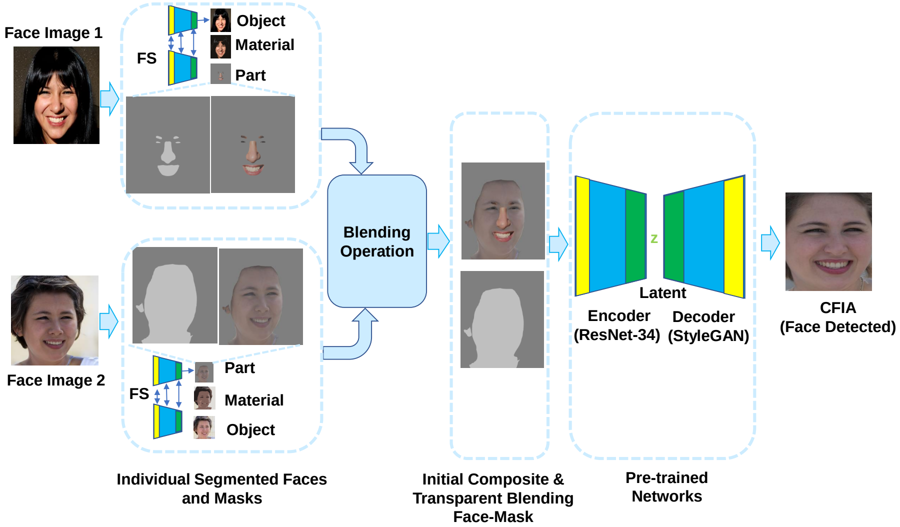


Figure 9.1: Block diagram of the proposed approach where FS is based on UPerNet Face Segmenter from Zhou et al. [11], the Encoder is based on Resnet-34 [3], and Decoder is based on StyleGAN [12] and the encoder-decoder synthesizes the final composite.

Thus, motivated by the limitations of the existing method [195], we aim to generate the composite images in a fully automatic fashion using GANs. Even though the GANs are extensively used for full face morphing attack generation [51, 49, 50], the composite (or facial attribute) based attack generation is presented for the first time in this work. The recent work by Chai et al. [3] presented a highly realistic facial image synthesis with missing regions using GAN-inversion. In this work, we modified the approach from Chai et al. [3] to generate the CFIA samples with the primary motivation to demonstrate the vulnerabilities of FRS to CFIA. Further, we exhaustively varied the regions based on facial attributes to evaluate their attack potential. The proposed method for CFIA generation is designed to consider the optimal pairing of the input images used during the compositing process to ensure high-quality CFIA generation. The CFIA samples are generated based on multiple facial attributes. Both single and multiple facial attributes are blended using the transparent (or real) value that can further improve the attack potential and challenge the detection of CFIA samples. Use of facial attributes or partial morphing will not alter the entire face and thus results in less distortion because the proposed CFIA approach will only choose the facial attributes from the contributory subjects and then synthesize the rest of the facial image using GAN. Hence, the generated CFIA images are challenging to be detected by expert border guards. The key

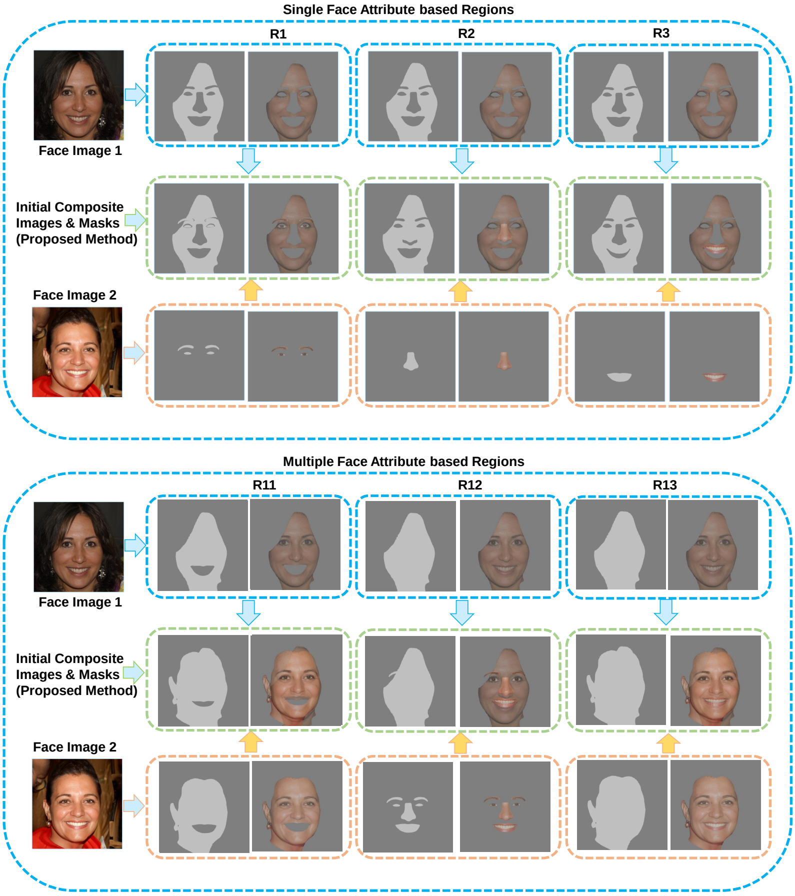

Single Face Attribute based Regions
R1
R2
R3
Face Image 1
Initial Composite Images & Masks (Proposed Method)
Face Image 2
Multiple Face Attribute based Regions
R11
R12
R13
Face Image 1
Initial Composite Images & Masks (Proposed Method)
Face Image 2

Figure 9.2: Illustration showing the comparison between the single face attribute based regions and multiple face attribute based regions for the generation of the initial composite using the proposed method.

contributions of our proposed method are as follows:

- We propose a novel framework for Composite Face Image Attack (CFIA) generation using regression and GAN-based image synthesis. The primary motivation of the proposed CFIA approach is to generate high-quality facial attack images using facial attributes with high attack potential. Further, it should be challenging to detect by both human and automatic morph detection techniques. Therefore, we generate CFIA based on single and multiple face attributes for given contributory data subjects. Further, we propose a transparent blending to improve the attack potential of the generated CFIA.

Thus, we introduce 526 different types of CFIA based on various combinations of facial attributes from contributory data subjects.

- We present a new CFIA dataset generated using 1000 unique data subjects (synthetic identities). The dataset consists of 526000 CFIA samples and 2000 bona fide samples.
- We present extensive vulnerability analysis on the newly generated CFIA dataset using deep learning-based FRS.. We also introduce an vulnerability metric called Generalized Morphing Attack Potential (G-MAP) to benchmark the attack potential effectively by considering real-life scenarios.
- We present the perceptual image quality analysis of the CFIA dataset using the Peak-Signal-to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) to benchmark the quality of the generated CFIA samples on a sub-set of the CFIA dataset with 14 unique combinations selected from the 526 combinations.
- We present the human observer study on the newly generated CFIA dataset (subset of 14 combinations) with 43 observers with and without face image manipulation detection background.
- We present extensive experiments benchmarking the performance to automatically detect the CFIA (subset of 14 combinations) using three different existing single image based face MAD techniques.
- The CFIA dataset, together with the source code of the proposed method, will be made publicly available to enable the reproducibility of the results presented in this paper `https://github.com/jagmohaniiit/LatentCompositionCode`.

In the rest of the paper we introduce the proposed method in Section 9.3, discussion on database generation methodology in presented in Section 9.4, vulnerability analysis and G-MAP is discussed in Section 9.5, Section 9.6 discuss the quantitative results of the perceptual quality evaluation, human observer study is discussed in Section 9.7, and discussion on CFIA detection (CAD) is presented in the Section 9.8. Lastly, the Section 9.9 draws the conclusions and future-work.

9.3 Proposed CFIA Generation Technique

Figure 9.1 shows the block diagram of the proposed CFIA method. The proposed CFIA method aims to automatically select single and multiple facial attribute regions from the given face images and blend them to generate a composite face image. The proposed CFIA method consists of three main functional blocks (1) generation of individual segmented faces and masks from given face images, (2) computation of the initial composite image and transparent blending face mask and (3) final CFIA generation based on pre-trained GANs.

9.3.1 Individual Segmented Faces and Masks

The proposed CFIA composite image generation is based on the different facial parts from the two contributory data subjects (e.g., skin from the first data subject and eyes from the second data subject). Therefore, we employed a high precision and accurate method to segment different facial parts. In this work, we choose the unified parsing network (UPerNet) [199] for automatic facial region segmentation, which is denoted as $\mathbb{FS}$. UPerNet [199] is based on multi-task learning and semantic segmentation to achieve high-quality results on facial segmentation and classification tasks. Thus, given the face image, UPerNet [199] provides six facial regions (or attribute) masks, including Skin (S), Eye (E), Nose (N), Mouth (M), Hair (H), and Background (B).

In this work, we have considered only two contributory face images based on real-time use-case applicability (for e.g. attacks on eMRTD or ID cards) [195, 69]. We denote the first contributory face image by F_1 and the corresponding part-based segmented masks obtained using UPerNet [199] be $SM1_i$, where $i = \{1, 2, \ldots, 6\}$ and its corresponding segmented image be $IS1_i$. Similarly, the second contributory face image be F_2 and the corresponding part-based segmented masks be $SM2_j$, where $j = \{1, 2, \ldots, 6\}$ and its corresponding segmented image be $IS2_j$. The face region segmentation process to obtain individual segments can be expressed as follows:

$$\begin{aligned} \{SM1_i, IS1_i\} &= \mathbb{FS}(F_1), \forall i = \{1, 2, \ldots, 6\} \\ \{SM2_j, IS2_j\} &= \mathbb{FS}(F_2), \forall j = \{1, 2, \ldots, 6\} \end{aligned} \tag{9.1}$$

Based on these six part-based segmentation masks (or region or facial attributes), we generate an exhaustive list of combinations from $SM1_i$ and $SM2_j$ that resulted in 526 unique CFIA samples as listed in Table 9.2. It needs to be pointed out that the selected areas are exhaustive as listed in Table 9.1. Table 9.1 mentions the CFIA Region Index, through which we give a numerical index to the output segments so that overall, it increases with the number of combinations. E.g., if we consider two combinations case, we select two regions from $SM1$ and choose a maximum of two regions out of six (in a step-wise manner) from $SM2$. Therefore, in two combinations case (see Table 9.1), we have CFIA region index 2, in which, we select 2 regions from $SM1$ and one region from $SM2$. Similarly, for CFIA region index 3 we select 2 regions from $SM1$ and 2 regions from $SM2$. We repeat

CFIA Region Index	Output Segments	Possible Pairs (Unique)
One Combinations		
1	2	$\binom{5}{1} \times \binom{5}{1} = 25(13)$
Two Combinations		
2	3	$\binom{5}{2} \times \binom{5}{1} = 50(26)$
3	4	$\binom{5}{2} \times \binom{5}{2} = 100(100)$
Three Combinations		
4	4	$\binom{5}{3} \times \binom{5}{1} = 50(50)$
5	5	$\binom{5}{3} \times \binom{5}{2} = 100(78)$
6	6	$\binom{5}{3} \times \binom{5}{3} = 100(86)$
Four Combinations		
7	5	$\binom{5}{4} \times \binom{5}{1} = 25(25)$
8	6	$\binom{5}{4} \times \binom{5}{2} = 50(50)$
9	7	$\binom{5}{4} \times \binom{5}{3} = 50(47)$
10	8	$\binom{5}{4} \times \binom{5}{4} = 25(25)$
Five Combinations		
11	6	$\binom{5}{5} \times \binom{5}{1} = 5(5)$
12	7	$\binom{5}{5} \times \binom{5}{2} = 10(10)$
13	8	$\binom{5}{5} \times \binom{5}{3} = 10(10)$
14	9	$\binom{5}{5} \times \binom{5}{4} = 5(5)$
15	10	$\binom{5}{5} \times \binom{5}{5} = 1(1)$
Six Combinations		
16	12	$\binom{6}{6} \times \binom{6}{6} = 1(1)$
Total Output Segments Possible		
607		
Total Unique Segments Possible		
526		

Table 9.1: Table showing the generation process of 526 unique CFIA combinations which are listed in detail Table 9.2.

this process for the various combinations of regions (or facial attributes), such that CFIA region index 2 results in 50 combinations corresponding to 3 output segments. Similarly, CFIA region index 3 results in 100 combinations corresponding to 4 output segments. These steps are repeated for different CFIA region indexes from 1 to 16, resulting in a total of 607 combinations. However, out of 607 combinations some of the combinations are redundant. For example, selecting face attributes from $SM1$ and $SM2$ such as SE-NM (SkinEyes-NoseMouth) can occur in two ways, firstly Skin, Eyes from $SM1$ and Nose, Mouth from $SM2$ and secondly SEN-M (SkinEyesNose-Mouth), Skin, Eyes and Nose from $SM1$ and Mouth from $SM2$ resulting in a redundant combination. Therefore, we removed all such redundant combinations and considered unique combinations. Hence, we generate 526 unique CFIA samples corresponding to two unique facial identities.

Region List						
S1-S2	**S1-S2**	**S1-S2**	**S1-S2**	**S1-S2**	**S1-S2**	**S1-S2**
E-H	H-E	H-H	H-M	H-N	H-S	M-H
N-H	S-H	S-E	S-N	S-M	S-S	EM-H
EN-H	HE-E	HE-H	HE-M	HE-N	HE-S	HM-E
HM-H	HM-M	HM-N	HM-S	HN-E	HN-H	HN-M
HN-N	HN-S	HS-E	HS-H	HS-M	HS-N	HS-S
NM-H	SE-H	SM-H	SN-H	EM-EM	EM-EN	EM-HE
EM-HM	EM-HN	EM-HS	EM-NM	EM-SE	EM-SM	EM-SN
EN-EM	EN-EN	EN-HE	EN-HM	EN-HN	EN-HS	EN-NM
EN-SE	EN-SM	EN-SN	HE-EM	HE-EN	HE-HE	HE-HM
HE-HN	HE-HS	HE-NM	HE-SE	HE-SM	HE-SN	HM-EM
HM-EN	HM-HE	HM-HM	HM-HN	HM-HS	HM-NM	HM-SE
HM-SM	HM-SN	HN-EM	HN-EN	HN-HE	HN-HM	HN-HN
HN-HS	HN-NM	HN-SE	HN-SM	HN-SN	HS-EM	HS-EN
HS-HE	HS-HM	HS-HN	HS-HS	HS-NM	HS-SE	HS-SM
HS-SN	NM-EM	NM-EN	NM-HE	NM-HM	NM-HN	NM-HS
NM-NM	NMS-E	NMS-M	NMS-N	SEE-M	SEE-N	SEH-E
SEH-M	SEH-N	SEH-S	SEN-M	SES-E	SES-M	SES-N
SME-M	SME-N	SMH-E	SMH-M	SMH-N	SMH-S	SMN-M
SMS-E	SMS-M	SMS-N	SNE-M	SNE-N	SNH-E	SNH-M
SNH-N	SNH-S	SNN-M	SNS-E	SNS-M	SNS-N	ENM-E
ENM-H	ENM-M	ENM-N	ENM-S	HEM-E	HEM-H	HEM-M
HEM-N	HEM-S	HEN-E	HEN-H	HEN-N	HEN-S	HNME
HNM-H	HNM-M	HNM-N	HNM-S	HSE-E	HSE-H	HSE-S
HSM-E	HSM-H	HSM-M	HSM-N	HSM-S	HSN-E	HSN-H
HSN-N	HSN-S	SEM-E	SEM-H	SEM-M	SEM-N	SEM-S
SEN-E	SEN-H	SEN-N	SEN-S	SNM-E	SNM-H	SNM-M
SNM-N	SNM-S	ENM-HE	ENM-HM	SEN-EM	SEN-EN	ENM-HN
ENM-HS	HEM-EM	HEM-EN	HEM-HE	HEM-HM	HEM-HN	HEM-HS
HEM-NM	HEM-SE	HEM-SM	HEM-SN	HEN-EM	HEN-EN	HEN-HE
HEN-HM	HEN-HN	HEN-HS	HEN-NM	HEN-SE	HEN-SM	HEN-SN
HNM-EM	HNM-EN	HNM-HE	HNM-HM	HNM-HN	HNM-HS	HNM-NM
HNM-SE	HNM-SM	HNM-SN	HSE-EM	HSE-EN	HSE-HE	HSE-HM
HSE-HN	HSE-HS	HSE-NM	HSE-SE	HSE-SM	HSE-SN	HSM-EM
HSM-EN	HSM-HE	HSM-HM	HSM-HN	HSM-HS	HSM-NM	HSM-SE
HSM-SM	HSM-SN	HSN-EM	HSN-EN	HSN-HE	HSN-HM	HSN-HN
HSN-HS	HSN-NM	HSN-SE	HSN-SM	HSN-SN	SEM-HE	SEM-HM
SEM-HN	SEM-HS	SEN-HE	SEN-HM	SEN-HN	SEN-HS	SNM-HE
SNM-HM	SNM-HN	SNM-HS	ENM-HEM	ENM-HEN	ENM-HNM	ENM-HSE
ENM-HSM	ENM-HSN	HEM-ENM	HEM-HEM	HEM-HEN	HEM-HNM	HEM-HSE
HEM-HSM	HEM-HSN	HEM-SEM	HEM-SEN	HEM-SNM	HEN-ENM	HEN-HEM
HEN-HEN	HEN-HNM	HEN-HSE	HEN-HSM	HEN-HSN	HEN-SEM	HEN-SEN
HEN-SNM	HNM-ENM	HNM-HEM	HNM-HEN	HNM-HNM	HNM-HSE	HNM-HSM
HNM-HSN	HNM-SEM	HNM-SEN	HNM-SNM	HSE-ENM	HSE-HEM	HSE-HEN
HSE-HNM	HSE-HSE	HSE-HSM	HSE-HSN	HSE-SEM	HSE-SEN	HSE-SNM
HSM-ENM	HSM-HEM	HSM-HEN	HSM-HNM	HSM-HSE	HSM-HSM	HSM-HSN
HSM-SEM	HSM-SEN	HSM-SNM	HSN-ENM	HSN-HEM	HSN-HEN	HSN-HNM
HSN-HSE	HSN-HSM	HSN-HSN	HSN-SEM	HSN-SEN	HSN-SNM	SEM-HEM
SEM-HEN	SEM-HNM	SEM-HSE	SEM-HSM	SEM-HSN	SEN-HEM	SEN-HEN
SEN-HNM	SEN-HSE	SEN-HSM	SEN-HSN	SNMHEM	SNM-HEN	SNM-HNM
SNM-HSE	SNM-HSM	SNM-HSN	SEN-SEM	SEN-SEN	HENM-E	HENM-H
HENM-M	HENM-N	HENM-S	HSEM-E	HSEM-H	HSEMM	HSEM-N
HSEM-S	HSEN-E	HSEN-H	HSEN-N	HSEN-S	HSNME	HSNM-H
HSNM-M	HSNM-N	HSNM-S	SENM-E	SENM-H	SENM-M	SENM-N
SENM-S	HENM-EM	HENM-EN	HENM-HE	HENM-HM	HENM-HN	HENM-HS
HENM-NM	HENM-SE	HENM-SM	HENM-SN	HSEM-EM	HSEM-EN	HSEM-HE
HSEM-HM	HSEM-HN	HSEM-HS	HSEM-NM	HSEM-SE	HSEM-SM	HSEM-SN
HSEN-EM	HSEN-EN	HSEN-HE	HSEN-HM	HSEN-HN	HSEN-HS	HSEN-NM
HSEN-SE	HSEN-SM	HSEN-SN	HSNM-EM	HSNM-EN	HSNM-HE	HSNM-HM
HSNM-HN	HSNM-HS	HSNM-NM	HSNM-SE	HSNM-SM	HSNM-SN	SENM-EM
SENM-EN	SENM-HE	SENM-HM	SENM-HN	SENM-HS	SENM-NM	SENM-SE
SENM-SM	SENM-SN	SENM-ENM	HENM-ENM	HENM-HEM	HENM-HEN	HENM-HNM
HENM-HSE	HENM-HSM	HENM-HSN	HENM-SEM	HENM-SEN	HENM-SNM	HSEM-ENM
HSEM-HEM	HSEM-HEN	HSEMH-NM	HSEMH-SE	HSEMH-SM	HSEM-HSN	HSEM-SEM
HSEM-SEN	HSEM-SNM	HSEN-ENM	HSEN-HEM	HSEN-HEN	HSEN-HNM	HSEN-HSE
HSEN-HSM	HSEN-HSN	HSEN-SEM	HSEN-SEN	HSEN-SNM	HSNM-ENM	HSNM-HEM
HSNM-HEN	HSNM-HNM	HSNM-HSE	HSNM-HSM	HSNM-HSN	HSNM-SEM	HSNM-SEN
HSNM-SNM	SENM-HEM	SENM-HEN	SENM-HNM	SENM-HSE	SENM-HSM	SENM-HSN
HENMH-ENM	HENMH-SEM	SENM-SENM	HENM-HSEN	HENM-HSNM	HENM-SENM	HSEM-HENM
HSEM-HSEM	HSEM-HSEN	HSEM-HSNM	HSEM-SENM	HSEN-HENM	HSEN-HSEM	HSENH-SEN
HSENH-SNM	HSEN-SENM	HSNM-HENM	HSNM-HSEM	HSNM-HSEN	HSNM-HSNM	HSNM-SENM
SENM-HENM	SENM-HSEM	SENM-HSEN	SENM-HSNM	HSENM-E	HSENM-H	HSENM-M
HSENM-N	HSENM-S	HSENME-M	HSENME-N	HSENMH-E	HSENMH-M	HSENMH-N
HSENMH-S	HSENMN-M	HSENMS-E	HSENMS-M	HSENMS-N	HSENMEN-M	HSENMH-EM
HSENMH-EN	HSENMH-NM	HSENMH-SE	HSENMH-SM	HSENMH-SN	HSENMS-EM	HSENMS-EN
HSENMS-NM	HSENMH-ENM	HSENMH-SEM	HSENMH-SEN	HSENMH-SNM	HSENM-SENM	HSENM-HSENM
HBSENM-HBSENM						

Table 9.2: Exhaustive List of Regions used for Composition where the compositions S1 are used for Subject 1 and S2 are used for Subject 2 where the facial attributes are B=Background, S=Skin, E=Eye, N=Nose and M=Mouth. The compositions listed in left to right order are in increasing order of Composition Region Index (for Composition Region Index, please refer Table 9.1)

.

9.3.2 Initial Composite Image and Transparent Blending Face-Mask

In the next step, we generate the initial composite image and transparent blending of face segments by applying the blending operation on the individual segmented faces ($IS1_i$ & $IS2_j$) and their corresponding masks ($SM1_i$ & $SM2_j$) from contributory data subjects (F_1 & F_2). The blending operation is carried out independently for the mask and the individual segmented faces. The blended mask m_c is generated by a simple union operation that can represent the combined facial region from $SM1_i$ and $SM2_j$ as described in Equation 9.2. The generation of the initial composite image (IC) is done in three consecutive steps shown in Equation 9.3, where first IC is initialized 0, then in the next step, IC is updated using the compositing equation with the segmented region ($IS1_i$) from the data subject F_1 as input. Finally, IC is updated using the compositing equation with the segmented region ($IS2_j$) from data subject F_2, and its segmentation masks $SM2_j$ as an input. These steps are mathematically presented in Equation 9.3.

$$m_c = SM1_i \bigcup SM2_j \tag{9.2}$$

$$\begin{aligned} IC &= 0 \\ IC &= IS1_i \\ IC &= IS2_j + (1 - SM2_j) \times IC \end{aligned} \tag{9.3}$$

Figure 9.2 shows the qualitative results of the initial composite image and the corresponding mask for both single-face attribute-based composite regions & multiple-face attribute-based composite regions.

9.3.3 Final CFIA samples Generation

Once the initial composite image and the transparent blending face mask are generated, we generate the final CFIA samples using the image inpainting based on pre-trained regressor and GAN [3]. The input composite image and its mask are passed through a pre-trained encoder ($\mathbb{E}$) and then to the decoder ($\mathbb{G}$) to generate the final composite image (FCI). The process of generating the CFIA sample is as indicated in Equation 9.4.

$$CFIA = \mathbb{D}(\mathbb{E}(IC, m_c)) \tag{9.4}$$

The encoder network ($\mathbb{E}$) selected in our work is pre-trained Resnet-34 [3], and the decoder network ($\mathbb{G}$) is a pre-trained StyleGAN-I decoder which was trained on FFHQ dataset [12]. The primary motivation for the choice of the encoder and decoder networks was that image to latent conversion is posed as a regression problem [3]. Further, it is found that Resnet-34 is suitable for regressing the latent from a face image with missing information and renders the high-quality face image. Lastly, we use the decoder ($\mathbb{D}$) based on StyleGAN-I as it provides a linear latent subspace [200]. Hence, reconstruction from the generated latent is of good perceptual quality even with missing information in the input image. Figure 9.3 shows example results corresponding to five combinations generated using the proposed method. For the simplicity, we have included the illustration for five combination and full 526 CFIA samples are included in the supplement material.

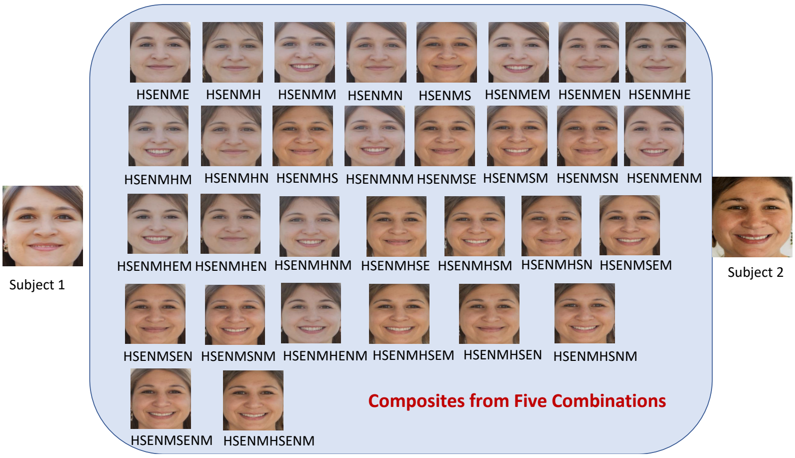


Figure 9.3: Illustration showing five combinations based composites.

9.4 CFIA Database Generation

This section presents the dataset generation process used to evaluate the proposed composite image generation. Owing to the ethical and legal challenges with face biometric datasets that will eventually limit the distribution, in this work, we generate the synthetic face images corresponding to the unique identities using StyleGAN inversion [3]. Earlier works [201, 51, 202] indicated that generating the synthetic face images have demonstrated both realness in terms of quality, uniqueness and verification accuracy. Further, synthetic face images will overcome the need for privacy and legal limitations to make the database public, which is vital for reproducible research. Figure 9.4 illustrates the CFIA dataset generation process.

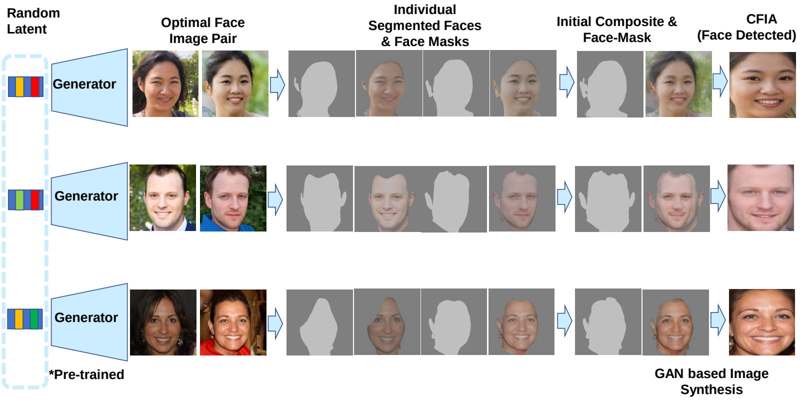


Figure 9.4: Illustration showing the CFIA dataset generation process

9.4.1 Synthetic Face Image Generation

Given a random latent vector, we use the approach from Chai et al. [3] to generate a synthetic face corresponding to unique data subjects using StyleGAN inversion. We further perturb the random latent by an ϵ amount to generate the mated face image corresponding to the given identity. The choice of ϵ is made empirically, which is small enough not to alter the identity of the generated face. However, the generation of synthetic face images with corresponding mated face images with unique identities will result in non-ICAO compliant photos with glasses, non-frontal pose, and a non-neutral face expression, as shown in Figure 9.5. Therefore, it is necessary to detect the ICAO-compliant faces for which we select faces with frontal pose automatically and remove photos with glasses and non-neutral face expressions manually.

9.4.2 Hyperparameters Selection

Hyper-parameters	**SOTA [3]**	**Proposed Method**
Frontal Pose Selection	No	**Yes**
Optimal Pairing	No	**Yes**
Epsilon(ϵ)	No	10^{-7}
Alpha(α)	1	0.5

Table 9.3: Different Hyper-parameters used for the proposed method. Note the proposed method modifies a large number of hyper-parameters compared with SOTA [3].

This section discusses the choices of the parameters associated with SOTA and

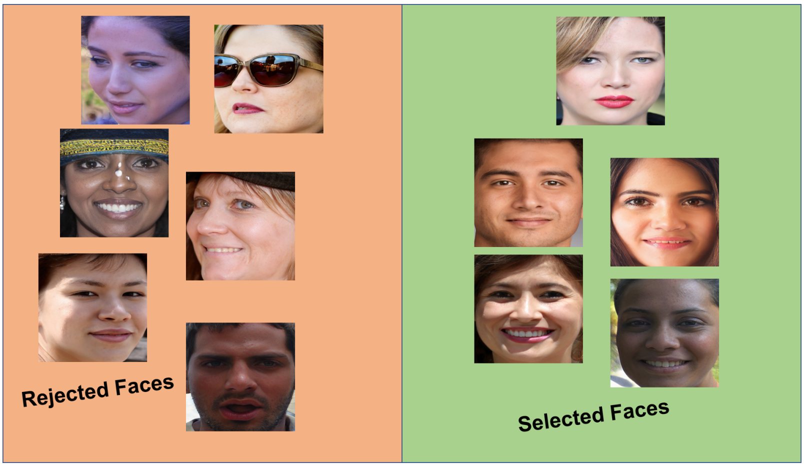


Figure 9.5: Illustration showing faces selected and rejected by our proposed frontal-pose detection Algorithm 2. Note face images with glasses and GAN-based artifacts are rejected manually.

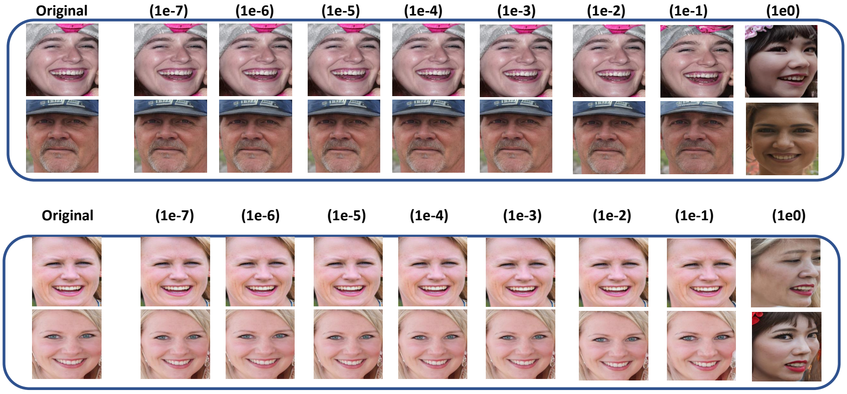


Figure 9.6: Illustration of the effect of perturbation based on epsilon (ϵ) for synthetic face generation, note artifacts start appearing when ($\epsilon = 0.1$) and results in change in identity when ($\epsilon = 1$)

the proposed method as tabulated in Table 9.3. In total, we have four different hyperparameters that are discussed as follows:

- **Epsilon(ϵ)**: The value of ϵ is empirically chosen as 10^{-7}. Since values higher than 10^{-1} lead to artifacts and a sample of different identities as shown in Figure 9.6. Thus, we choose an ϵ conservatively.
- **Alpha**: We choose α=0.5 as it is known to create the highest vulnerability towards FRS for Face Morphing Image Attack (FMIA) [69]. of segments possible is shown in Table 9.2

9.4.3 Frontal Face Pose Selection

We have developed the algorithm to automatically select the ICAO compliant face images corresponding to each unique identity as indicated in the Algorithm 2. The primary motivation behind this algorithm is that the face in a frontal pose would have similar angles between Left-Eye, Nose, and Mouth (Left Part) and Right-Eye, Nose, and Mouth (Right Part). A slight change in the face pose from a frontal face to a profile face would result in a skew, which would cause these two angles to be different. The qualitative results of the proposed frontal face selection algorithm are as shown in Figure 9.5. Since we are currently not interested in the computation of exact face pose, the heuristic works sufficiently well for our dataset, which does not consist of extreme face poses.

Algorithm 2: Non-Frontal Pose Identification

```
Input: Face Image with 5 Landmarks (Left-Eye ($LE$), Right-Eye ($RE$),
       Nose ($N$), Left-Mouth ($LM$), and Right-Mouth ($RM$)
Output: True if Face Image is Frontal
Compute the angle between the vectors of Left-Eye, Nose, and Left-Mouth,
Nose
$\theta_1 \leftarrow \arccos((\overrightarrow{LEN} \cdot \overrightarrow{LMN}))$.
Compute the angle between the vectors of Nose, Right-Eye, and Nose,
Right-Mouth
$\theta_2 \leftarrow \arccos((\overrightarrow{NRE} \cdot \overrightarrow{NRM}))$.
Compute absolute difference between the angles, as angleDiff $\leftarrow |\theta_1 - \theta_2|$
if angleDiff $\leq \tau$ then
    Face is Frontal
    return True
end if
return False
```

9.4.4 Optimal Face Pair Generation for composite image generation

Algorithm 3: Optimal Pair Finding Algorithm

```
Input: Random Image Pairs $(I_1^1, I_2^1),\cdots,(I_1^N, I_2^N)$
Output: Optimal Image Pairs $(O_1^1,O_2^1),\cdots,(O_1^N,O_2^N)$
Compute Arcface features on the input face images.
Optimal-Pair $\leftarrow$ []
for $i \leftarrow 1$ to $N$ do
    Compute Index of nearest arcface feature $j$ to $i$
    if $(j, i) \notin$ Optimal-Pair then
        Append $(i, j)$ to Optimal-Pair
    else
        Compute Index of second-nearest arcface feature $k$ to $i$
        Append $(i, k)$ to Optimal-Pair
    end if
end for
return Optimal-Pair
```

It is essential to select the look-alike data subjects to achieve the optimal attack potential with the proposed composite face image generation. We choose the optimal pairs to generate the composite face images to this extent. Given n synthetic samples, the total number of pairs possible is $((n)\times(n-1))/2$, and thus finding optimal pairs using this approach is quadratic ($O(n^2)$) as we have to compute the pair-wise distance for all pairs. The quadratic time for pair-finding is within the computing limits as our dataset now consists of 1000 unique data subjects. We have put an additional constraint in the pair-finding algorithm not to return swapped pairs, i.e., if (i,j) is the list, then (j,i) is not added to the optimal pair list. The approach for optimal pair finding is summarized in an algorithmic format in Algorithm 3 and a few optimal pairs are shown in Figure 9.4. The distance metric used in our approach is cosine-distance from Arcface [22] features.

Thus, the CFIA dataset has 1000 unique identities with 2000 bona fide samples and 526000 CFIA samples. The whole dataset will made publicly available for research purposes along with code at the following link `https://github.com/jagmohaniiit/LatentCompositionCode`.

9.5 Vulnerability Analysis

This section presents the vulnerability analysis of the proposed CFIA samples on the automatic FRS. We have benchmarked four different FRS based on deep learning. The deep learning FRS employed in this work are Arcface [54] (Model R100

Utility Features	MMPMR [83]	FMMPMR [84]	MAP [85]	G-MAP
Multiple Attempts for individual morphing Image	✓	✓	✗	✓
Pairwise comparison of probe samples	✗	✓	✓	✓
Multiple FRS	✗	✗	✓	✓
Multiple Morphing Types	✗	✗	✗	✓
Accountability for FTAR	✗	✗	✗	✓
Vulnerability as a single number	✓	✓	✗	✓

Table 9.4: Utility Features of existing and proposed vulnerability metrics

V1), VGGFace [55] (Version 2), Facenet [56] and Magface [203]. The proposed CFIA samples are generated based on the face images corresponding to two contributory subjects. Therefore, we benchmark the attack potential of CFIA by comparing the FRS scores computed from both contributory subjects against the pre-set threshold of FAR = 0.1%. The comparison scores from FRS are computed by enrolling the attack samples to FRS and then probing the face images from the contributory subjects.

In the literature, the vulnerability of FRS can be calculated using three different types of metrics namely: Mated Morphed Presentation Match Rate (MMPMR) [83], Fully Mated Morphed Presentation Match Rate (FMMPMR) [84] and Morphing Attack Potential (MAP) [85]. The MMPMR metric is based on the independent attempts, while FMMPMR employs pair-wise probe attempts of the contributory subjects. The MAP metric improves existing metrics by accommodating multiple FRS together with pair-wise probe attempts. However, the MAP metric will represent the vulnerability results in the matrix form as attempts versus multiple FRS. Hence, MAP does not quantify the vulnerability as a single number. Further, the constant number of attempts will also limit the evaluation as it enforces all enrolled attack samples to have the same number of attempts which is not true in a real-life scenario. Additionally, while computing the vulnerability, the existing metrics do not consider accommodating Failure-to-Acquire Rate (FTAR) and multiple morphing generation techniques. Even though the enroled face image (attack/CFIA/morphing or bona fide) is captured in the constrained conditions, the probe images are not essentially captured in the constrained conditions due to the nature of ID verification scenarios (for example, in border control gates, smartphone authentication, etc.). Further, the availability of different types of morphing (or attack) generation techniques (full face/partial face/facial attribute) allows an attacker to generate various attack samples. Hence, the vulnerability computation needs to accommodate different types of morphing generation. These factors motivated us to enhance the existing vulnerability metrics (MAP) to include more utility features such as (a) Dynamic attempts per morph image, (b) Accountability for FTAR, (3) Accountability for multiple morphing techniques, and (4) Single numeric value indicating the vulnerability. The enhanced vulnerability metric is termed as Generalised Morphing Attack Potential (G-MAP). Table 9.4 presents utility features of the proposed G-MAP compared to existing metrics such as MMPMR [83], FMMPMR [84] and MAP[85].

9.5.1 Mathematical Formulation of G-MAP

Let $\mathbb{P}$ denote the set of paired probe images (which can also be denoted as number of attempts), $\mathbb{F}$ denote the set of FRS, $\mathbb{D}$ denote the set of Morphing Attack Generation Type, $\mathbb{M}_d$ denote the face morphing image set corresponding to Morphing Attack Generation Type d, τ_l indicate the similarity score threshold for FRS (l), and $||$ represents the count of elements in a set during metric evaluation. The G-MAP metric is presented as below:

$$\begin{aligned} \text{G-MAP} = \frac{1}{|\mathbb{D}|}\sum_{d}^{|\mathbb{D}|}\frac{1}{|\mathbb{P}|}\frac{1}{|\mathbb{M}_d|}\min_{\mathbb{F}_l} \\ \sum_{i,j}^{|\mathbb{P}|,|\mathbb{M}_d|}\left\{\left[(S1_i^j > \tau_l) \wedge \cdots (Sk_i^j > \tau_l)\right]\right. \\ \left. \times\left[(1 - FTAR(i,l))\right]\right\} \end{aligned} \tag{9.5}$$

where, $FTAR(i,l)$ is the failure to acquire probe image in attempt i using FRS (l). The algorithm for G-MAP is presented in 4 and the code is made available in the link [204].

9.5.2 Computing G-MAP

Given the fact that G-MAP can be computed with different parameters, which include multiple probe attempts, multiple FRS and the morph attack generation types. **G-MAP with multiple probe attempts** is calculated from Equation 11.4 by setting D = 1 and F = 1 where the similarity scores $(S1_i^j)$ should be greater than threshold (τ_l) and FTAR(i,l) is calculated for each probe attempt and FRS. Thus, making **G-MAP with multiple probe attempts** identical to FMMPMR when FTAR=0. Further, **G-MAP with Multiple FRS and multiple probe attempts** is computed by taking minimum across FRS and using D=1. Finally, the full **G-MAP metric** would provide a single value indicating the vulnerability which is by taking the average as shown in Equation 11.4.

9.5.3 Quantitative evaluation of vulnerability

In this section, we present the qualitative and quantitative evaluation of the vulnerability corresponding to FRS for all 526 CFIA samples generated using a different combination of facial attributes. Since G-MAP is a function of attempts, FRS, and morphing types, this will allow one to analyse the quantitative results corresponding to (a) probe attempts independently to FRS and attack image generation type

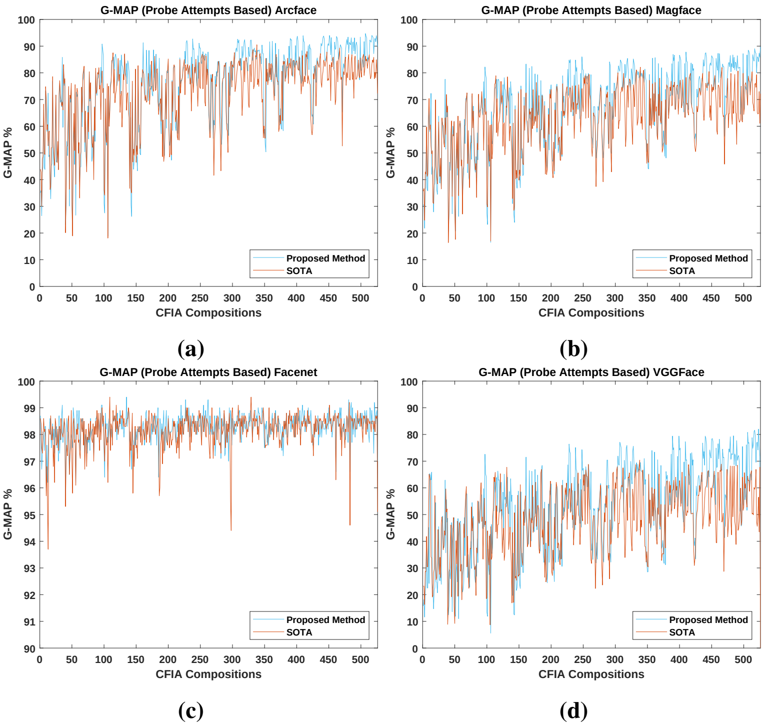


Figure 9.7: Vulnerability Plots G-MAP (Probe Attempts). X-axis indicates the number of unique CFIA generated where the index 0 corresponds to E-H, the index 1 corresponds to H-E, and the following indices in the left to right order corresponding to Table 9.1. Thus, finally, index 525 to HBSENM-HBSENM.

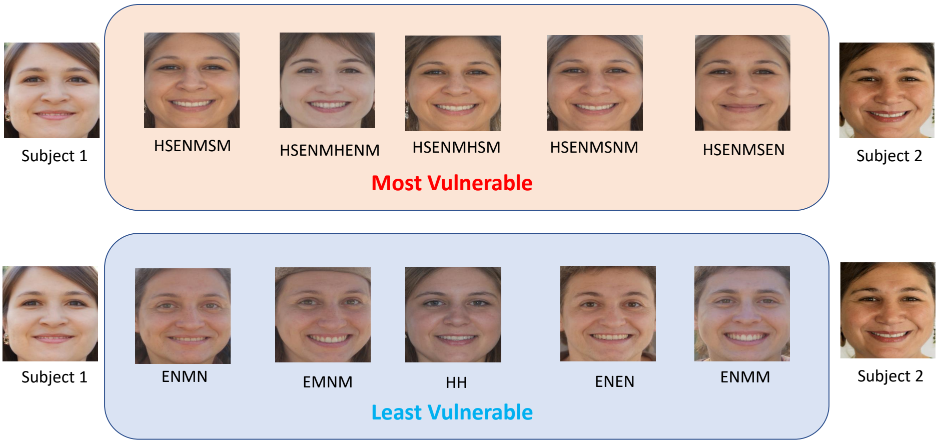


Figure 9.8: Most and Least Vulnerable CFIA Samples from the dataset.

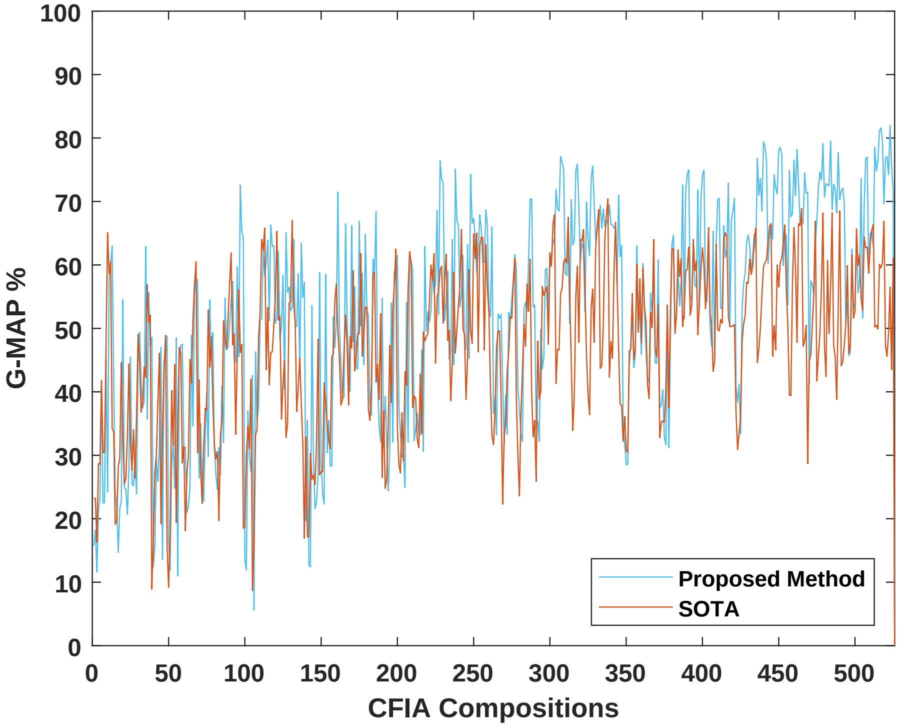


Figure 9.9: Vulnerability Plots G-MAP (Multiple FRS and multiple probe attempts-based). X-axis indicates the number of unique CFIA generated where the index 0 corresponds to E-H, the index 1 corresponds to H-E, and the following indices in the left to right order corresponding to Table 9.1. Thus, finally, index 525 to HBSENM-HBSENM.

Algorithm 4: Generalized Morph Attack Potential (G-MAP)

```
Input: Set of Probe Images $\mathbb{P}$, Set of FRS $\mathbb{F}$, Set of Morphing Attack
       Generation Type $\mathbb{D}$, Set of Morphing Attack Images in $d^{\text{th}}$ attack $\mathbb{M}_d$,
       $\tau_l$ indicate the similarity score threshold for FRS.
Output: G-MAP
Compute G-MAP Metric as follows.
for $j \leftarrow 1$ to $|\mathbb{M}_d|$ do
    for $d \leftarrow 1$ to $|\mathbb{D}|$ do
        for $l \leftarrow 1$ to $|\mathbb{F}|$ do
            for $i \leftarrow 1$ to $|\mathbb{P}|$ do
                Compute QF(i,l)=(1-FTAR(i,l))
                Compute G-MAP(d)=$\frac{1}{|\mathbb{P}|}\frac{1}{|\mathbb{M}_d|}\min_l \sum_{i,j}^{|\mathbb{P}|,|\mathbb{M}_d|}(S1_i^j > \tau_l) \wedge \cdots$
                $(Sk_i^j > \tau_l) \times QF(i,l)$
            end for
        end for
    end for
end for
Compute G-MAP= $\frac{1}{|\mathbb{D}|}G - MAP(d)$
return G-MAP
```

(b) Multiple FRS with multiple attempts independent of attack image generation type (c) Final G-MAP value as a function of attempts, multiple FRS and different types of attack image generation together with FTAR.

In this work, we first present the vulnerability of the full CFIA dataset using four different FRS such as Arcface [54] (Model R100 V1), VGGFace [55] (Version 2), Facenet [56] and Magface [203]. The vulnerability reported in this work is computed by setting the threshold of FRS at FAR = 0.1%. Figure 9.7 shows the plot of G-MAP values that are computed for multiple probe attempts independent of FRS and CFIA generation type. The composite region index started from the output segment with two regions (left extreme of x-axis in Figure 9.7) and continued till six combinations (right extreme of x-axis in Figure 9.7). Table 9.5 shows the quantitative values of G-MAP with probe attempts corresponding to four different FRS. For simplicity, we have only indicated the quantitative results to 14 combinations sampled from 526 regions. *It needs to be pointed out that these 14 regions are indicative of least, moderate and most vulnerable regions from 526 unique CFIA combinations.*

Based on the obtained results following are the main observations:

- The number of composite regions used to generate the CFIA samples plays a vital role in the vulnerability of FRS. Using a smaller number of regions (for example, 2, 3 and 4) to generate the CFIA will result in a lower vulnerability of FRS. This it can be attributed to the fact that in these regions, the blending for the generation of composite happens in a small region and the remainder of the face is generated by GAN-based image inpainting. For example, if we consider the two regions (or facial attribute) CFIA generation, then one region is taken from the contributory subject 1 and another region is taken from the contributory subject 2, from these selected regions, the whole face is generated using the GAN. This process results in the loss of identity information in the generated CFIA due to the availability of a few regions. Figure 9.8 illustrates the example of low vulnerable CFIA samples generated using two and three region combinations. The lower vulnerability is noted with both SOTA and the proposed CFIA generation.

- The CFIA samples generated using 4, 5 and 6 regions have indicated higher vulnerability of FRS. This can be attributed to the fact that the larger the number of facial attributes used from both the contributory data subjects, the higher the vulnerability of the FRS. This trend is noticed equally with both SOTA and the proposed CFIA generation. Figure 9.8 shows the CFIA samples for the top 5 highest vulnerable combinations indicating the rich identity features corresponding to both contributory subjects.

- Among the four different FRS employed in this work, the Facenet [56] indicates the higher vulnerability across different region combinations. The lowest vulnerability is noted with the VGG FRS [55].

- The proposed CFIA generation technique indicates the higher vulnerability of FRS when compared with the SOTA [3]. The higher vulnerability of FRS to the proposed technique is noted with the CFIA samples that are generated using five and six-region combinations.

- Additional experiments on Commercial-Off-The-Shelf (COTS) to indicate the importance of FTAR is included in the Appendix A.

Figure 9.9 shows the vulnerability of FRS with G-MAP computed across multiple FRS and multiple attempts for both SOTA and proposed CFIA with 526 combinations. Given CFIA sample is said to be vulnerable if the multiple probe attempts must successfully deceive the multiple FRS. Thus, the G-MAP will provide a single value indicating the vulnerability by taking the average probe attempts while accounting for FTAR. Table 9.6 indicates the G-MAP (multiple FRS and

G-MAP % (Multiple probe attempts)															
FRS	**Method**	**R1**	**R2**	**R3**	**R4**	**R5**	**R6**	**R7**	**R8**	**R9**	**R10**	**R11**	**R12**	**R13**	**R14**
Arcface (FAR=0.1%)	**SOTA [3]**	70.5	58.7	60.6	52.7	70.8	69.2	72.9	71.6	69.1	69.1	67.6	74.1	69.6	72.3
	Proposed	67.3	58.1	60.2	72.4	70.4	68.4	72.5	71.9	71.4	84.2	82.8	76.4	86.8	**89.9**
MagFace (FAR=0.1%)	**SOTA [3]**	57.6	45.0	48.7	42.7	58.0	57.3	61.0	60.7	61.1	59.0	52.6	65.1	57.7	54.0
	Proposed	67.3	58.2	60.1	72.4	70.4	68.6	72.5	72.0	71.4	84.2	82.8	76.4	86.7	**89.8**
VGGFace (FAR=0.1%)	**SOTA [3]**	65.2	64.2	62.7	63.6	64.9	63.9	66.1	65.7	67.0	67.2	65.4	65.9	68.1	67.7
	Proposed	65.4	64.4	63.0	65.9	66.4	68.1	69.1	66.0	68.5	70.5	70.5	68.6	71.0	**71.9**
Facenet (FAR=0.1%)	**SOTA [3]**	95.8	96.9	95.4	93.8	95.3	96.5	95.9	95.8	96.1	95.6	94.5	96.4	94.7	95.2
	Proposed	96.1	97.7	96.3	97.4	95.5	97.3	95.4	96.4	97.4	97.2	97.3	96.6	96.6	**97.0**

Table 9.5: Vulnerability analysis using the G-MAP metric (probe attempts-based) for the proposed method and the SOTA [3], where the description of regions is provided in Table 9.1. Where R1 is (S-E), R2 is (S-N), R3 is (S-M), R4 is (S-S), R5 is (SEN-M), R6 is (SEM-N), R7 is (SNM-E), R8 is (SEN-EM), R9 is (SEN-EN), R10 is (SEN-SEM), R11 is (SEN-SEN), R12 is (SENM-ENM), R13 is (SENM-SENM), and R14 is (HBSENM-HBSENM)).

G-MAP % (Multiple FRS and multiple probe attempts)														
Method	**R1**	**R2**	**R3**	**R4**	**R5**	**R6**	**R7**	**R8**	**R9**	**R10**	**R11**	**R12**	**R13**	**R14**
SOTA [3]	57.6	45.0	48.7	42.7	58.0	57.3	61.0	60.7	61.1	59.0	52.6	65.1	57.7	54.0
Proposed	65.4	58.1	60.1	65.9	66.4	68.1	69.1	66.0	68.5	70.5	70.5	68.6	71.0	71.9

Table 9.6: Vulnerability analysis using the G-MAP metric (Multiple FRS and multiple probe attempts-based) for the proposed method and the SOTA [3].

multiple probes) for 14 different regions (that are the same as Table 9.5) for simplicity. Based on the obtained results
following are the main observations:

- The CFIA samples generated with five and six regions combinations indicate higher vulnerability of multiple FRS. This is noted with both SOTA and the proposed CFIA technique.
- The proposed CFIA samples indicate the higher vulnerability of FRS compared to SOTA.
- Figure 9.10 shows the box plots of proposed method and SOTA computed across CFIA region index as mentioned in Table 9.1 indicates the mean and variance computed by taking the average of G-MAP values computed over all region combinations within the CFIA region index. As noticed from Figure 9.10 and Table 9.8, the combinations with less number of regions do not significantly increase the vulnerability. The combination of five regions with CFIA region index of 13, 14 and 15 indicates the higher vulnerability of FRS with the proposed CFIA technique.

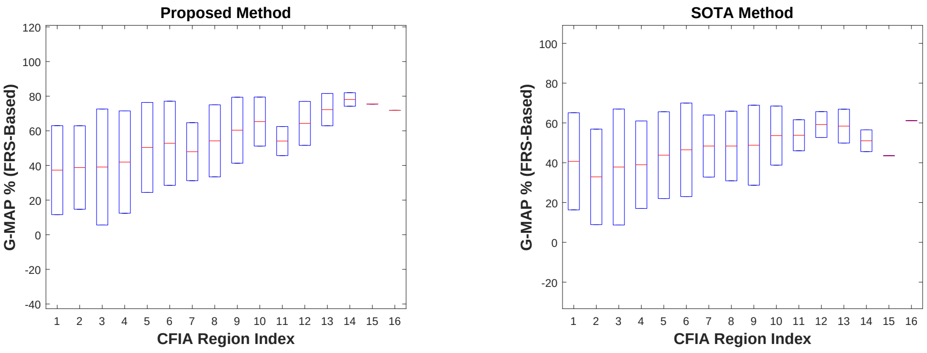
Table 9.7 indicates the vulnerability computed with full capacity of G-MAP in which multiple attempts, multiple FRS, multiple attack types and FTAR. The G-

G-MAP %	
SOTA Method [3]	Proposed Method
46.9%	52.4%

Table 9.7: G-MAP for SOTA Method and the Proposed Method computed using 526 CFIA compositions.

CFIA Region Index	Proposed Method	SOTA Method [3]
1	37.3±36.3	40.7±34.5
2	38.8±34.0	32.9±33.9
3	39.1±47.3	37.8±41.2
4	41.9±41.7	39±31.1
5	50.4±36.7	43.8±30.8
6	52.8±34.3	46.5±33.2
7	47.9±23.6	48.4±22.0
8	54.2±29.4	48.4±24.7
9	60.3±26.9	48.8±28.4
10	65.3±20.0	53.6±21.0
11	54.1±11.8	53.8±11.0
12	64.3±17.9	59.2±9.1
13	72.2±13.2	58.4±12.0
14	78.1±5.5	51.0±7.7
15	75.5±0	43.5±0
16	71.8±0	61.1±0

Table 9.8: Table showing mean and standard deviation for each CFIA region index based on SOTA [3] and the Proposed Method. (for CFIA region index please refer Table 9.1)



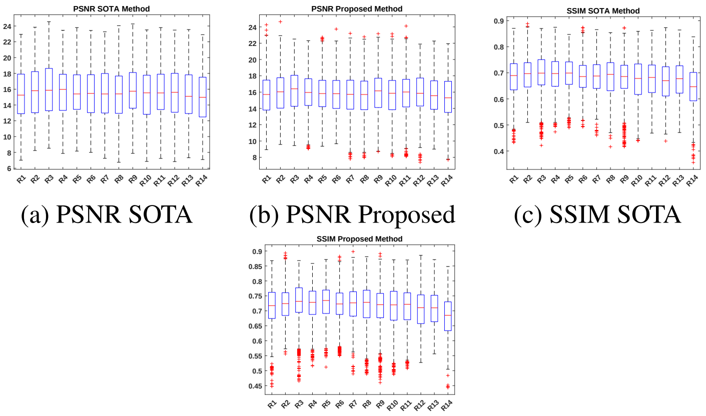
(a)

(b)

Figure 9.10: G-MAP Combinations (FRS-Based) where the number denotes the CFIA Region Index of (a) Proposed and (b) SOTA Method [3] (Table 9.1)

MAP values indicated in the 9.7 quantify the vulnerability of the proposed and SOTA for the complete CFIA dataset with 526 attack types and four different FRS. The obtained results indicate that the proposed method gives higher bounds of vulnerability for all 526 attack types.

9.6 Perceptual quality evaluation of the composite images



(a) PSNR SOTA

(b) PSNR Proposed

(c) SSIM SOTA

(d) SSIM Proposed

Figure 9.11: Box plots showing PSNR of SOTA [3] and the proposed Method for 14 regions. These 14 regions are same as indicated in Table 9.5

This section presents the quantitative analysis of the proposed CFIA samples using two perceptual image quality metrics, namely, PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index). We present the results pertaining to 14

Region	**PSNR**		**SSIM**	
R1	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
R1	15.4±10.2	15.6±7.0	0.68±0.01	0.71±0.00
R2	15.5±9.5	15.7±6.6	0.68±0.01	0.71±0.00
R3	15.5±10.2	15.6±7.0	0.68±0.01	0.71±0.00
R4	15.6±8.6	15.9±4.6	0.69±0.01	0.71±0.00
R5	15.4±10.6	15.6±7.4	0.68±0.01	0.71±0.00
R6	15.5±10.0	15.7±6.9	0.68±0.01	0.71±0.00
R7	15.5±10.6	15.7±7.4	0.68±0.01	0.71±0.00
R8	15.4±9.6	15.6±6.8	0.68±0.01	0.71±0.00
R9	15.4±8.6	15.7±6.7	0.68±0.01	0.73±0.00
R10	15.7±8.7	16.0±4.7	0.69±0.01	0.72±0.00
R11	15.6±9.9	16.0±5.0	0.69±0.01	0.72±0.00
R12	15.3±7.8	15.7±6.4	0.67±0.00	0.71±0.00
R13	15.7±10.3	16.0±5.2	0.69±0.01	0.72±0.00
R14	15.8±14.4	16.0±6.4	0.68±0.01	0.71±0.00

Table 9.9: Perceptual Image Quality Metrics PSNR and SSIM comparison for SOTA [3] and proposed Method on 14 different regions mentioned in the Table 9.5

regions out of 526 unique regions for the simplicity and these regions are same as mention in Section 9.5 and in Table 9.5. It is worth noting that, these 14 regions will represent the lower, moderate and high vulnerability of FRS. Both PSNR and SSIM are reference image-based quality metrics and thus require a pair of images for evaluation (face image from the contributory data subject and the generated face composite image). Table 9.9 indicates the quantitative analysis of the perceptual quality analysis on both SOTA [3] and the proposed CFIA method. Figure 9.11 illustrates the box plots corresponding to both SSIM and PSNR computed on all 14 regions. Following are the main observations from the obtained results:

- The PSNR metric has a higher mean-value and less variance for the proposed CFIA method compared with SOTA [3] indicating lesser noise in the face composites generated using the proposed CFIA method. This is expected as transparent blending would produce a lower contrast image, as the choice of blending-factor ($\alpha = 0.5$) would generate a pixel value lower than those from contributory data subjects as the blending equation is applied twice refer Equation 9.3. Thus, the proposed CFIA method generates a more consistent image quality irrespective of the region compared with SOTA [3].

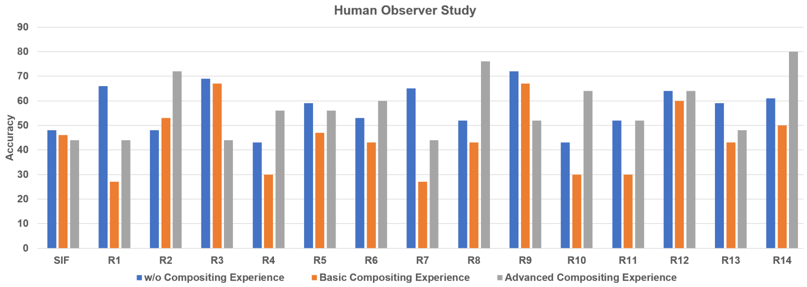
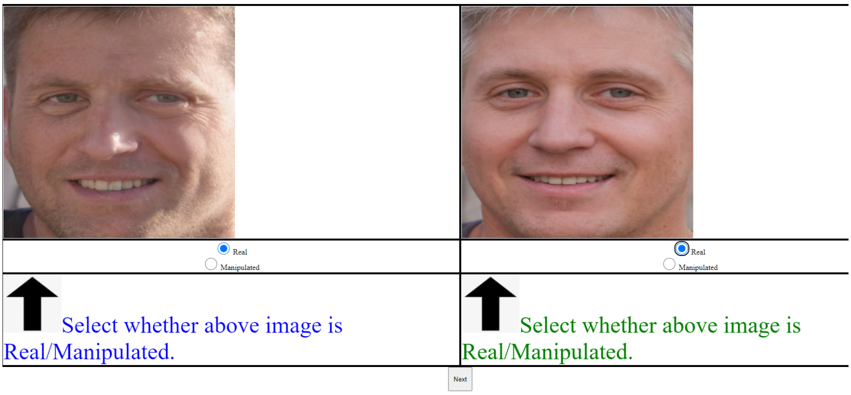


Figure 9.12: Illustration showing average accuracy quantitatively for the human observer study where bona fide or Synthetic Face Image (without any modification) is denoted as SIF.

- The SSIM metric produces a more stable value for both the proposed CFIA method and SOTA [3]. The proposed CFIA method gives a higher value for SSIM than the SOTA [3]. Since SSIM is a metric more tuned to the Human Visual System (HVS), [205] as it measures luminance distortion, contrast distortion, and loss of correlation. Thus, our proposed CFIA method generates higher-quality composites for HVS.

9.7 Human Observer Study



● Real
○ Manipulated

Select whether above image is Real/Manipulated.

● Real
○ Manipulated

Select whether above image is Real/Manipulated.

Next

Figure 9.13: Screenshot from the GUI (Full Page) of human observer web page.

We perform a Human Observer Study (HOS) of the generated composites to evaluate the detection performance by human experts. We present the results pertaining to 14 regions out of 526 unique regions for the simplicity and these regions are same as mention in Section 9.5 and in Table 9.5. It is worth noting that, these 14 regions will represent the lower, moderate and high vulnerability of FRS. The HOS is conducted using a web-based application [1] where a dedicated web page is set up with the use of PHP and HTML-CSS. In this study, GDPR norms are respected, and we only store the individual's email, gender, experience with the composite problem, and age group. We have made sure that the user remains anonymous during the study. Figure 9.13 shows the screenshot of the GUI of our website where the HOS is carried out. In this study, a human observer is shown a webpage with two images at a time where the observer has to decide independently on whether each of them is real/composite (or manipulated). The current study shows 43 image pairs, and it takes around 20 minutes to complete the study. The study includes synthetic face images and 14 different types of composites as mentioned in Table 9.5. Further, the human observer is explained in detail the stepwise instructions to perform the study. This enables people without awareness of the image manipulation problem and those with basic and advanced awareness of the composition problem to participate in the study. In the current evaluation, 51 human observers have participated and completed the study, including 40 participants without awareness, 6 with basic awareness, and 5 with an advanced awareness of the composition problem.

The quantitative results of HOS are as shown in Figure 9.12 and the following are the important observations:

- The average detection accuracy is similar for human observers without awareness of the composition problem and those with basic awareness. This can be attributed to the innate human ability to detect composites. However, the average detection accuracy for human observers with advanced awareness of the composition problem is much higher than both without awareness and basic awareness.

- The average accuracy is not very high for faces based on the composition, which utilizes a single facial attribute except for **R2** with advanced awareness. This can be attributed to the fact that a large part of the facial region needs to be inpainted in the case of single facial attribute composition.

- The average detection accuracy is high for the regions **R8**, **R10**, **R12**, and **R14**. **R8** has moderate parts of faces being used for compositing from the two contributory data subjects. The reason for high detection accuracy can be attributed to the fact that **R8** has only eyes from both the contributory data subjects but his nose and mouth from different contributory data subjects.

[1] https://folk.ntnu.no/jagms/indexCompositeUpdated.html

The same reasoning with more significant facial parts used for compositing can be extended to **R10** where the nose and mouth are from different contributory data subjects but have skin and eyes from both contributory data subjects. Now for the compositing region **R12**, the skin region is only from one contributory data subject. Thus, in all three cases, the asymmetry in the regions from the contributory data subjects aids the human observer in performing the detection at high accuracy.

- However, the average performance of the human observers for detecting a normal face image (or non-composite) is 46%. Further, it is also interesting to observe that degraded performance is noted in the advanced experience group. Thus, our analysis indicates that human observers are limited in detecting the normal face images compared to the composite face images.
- Now, for the compositing region **R14** all facial parts from the contributory data subjects are being used. Thus, the composited image can be distinguished from a synthetic face image using global image-based cues.
- In summary, we could say that either asymmetric regions or global level cues can help the human observer perform detection at high accuracy rates. However, our analysis indicates that it is very challenging for humans to detect composite attacks.

9.8 Composite Face Image Attack Detection

In this section, we benchmark CFIA detection based on a single image. Since the generation of CFIA is procedurally similar to morphing generation with transparent blending. Therefore, we have employed three different Face Morphing Attack Detection (MAD) techniques to benchmark the CFIA detection. MAD methods are selected by considering their detection performance on various morphing data sources, including NIST FRVT MORPH benchmarking. To this extent, we have chosen three different S-MAD approaches, namely: Color denoising based S-MAD (DetAlgo1) [4], Hybrid features (DetAlgo2) [5] and Residual noise-based S-MAD Network (DetAlgo3) [6]. We also report the performance of CAD algorithms on 14 different regions for the same reasons that were descried in previous section 9.5. These algorithms are briefly explained as follows:

Color denoising based S-MAD (DetAlgo1) [4]: DetAlgo1 is based on using the color information by converting the RGB image HSV color space. Then, each color channel is denoised using a Deep Convolutional Neural Network to compute the corresponding residual noise. In the next step, Pyramid LBP (P-LBP) and an SRKDA classifier for final detection.

Detection Method (Region)	D-EER (%)		BPCER @ APCER =			
			5%		10%	
R1	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	50.0	42.9	96.0	92.1	92.5	86.3
DetAlgo2 [5]	50.0	50.0	95.9	94.3	92.4	89.2
DetAlgo3 [6]	38.2	28.7	85.4	74.0	78.5	57.2
R2	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	50.0	44.6	96.0	93.0	92.3	86.0
DetAlgo2 [5]	50.0	50.0	96.2	94.4	92.3	91.2
DetAlgo3 [6]	39.5	29.3	87.1	78.5	78.9	64.8
R3	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	50.0	47.0	95.5	93.9	92.1	88.7
DetAlgo2 [5]	50.0	50.0	96.3	94.7	92.8	91.5
DetAlgo3 [6]	40.6	32.2	88.3	79.1	80.3	65.8
R4	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	49.0	39.6	94.3	90.5	89.0	81.7
DetAlgo2 [5]	50.0	50.0	96.1	92.6	92.8	88.9
DetAlgo3 [6]	42.8	32.4	89.8	77.7	82.2	64.6
R5	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	50.0	45.0	95.0	93.3	91.0	86.6
DetAlgo2 [5]	50.0	50.0	96.6	93.9	92.7	90.8
DetAlgo3 [6]	42.0	31.6	87.9	76.5	80.7	64.8
R6	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	50.0	44.1	95.5	92.5	91.8	84.6
DetAlgo2 [5]	50.0	50.0	95.3	92.7	91.0	88.9
DetAlgo3 [6]	38.0	29.3	85.3	73.2	78.5	60.2
R7	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	50.0	43.5	95.0	92.5	90.8	85.8
DetAlgo2 [5]	50.0	49.7	95.8	92.8	92.1	87.5
DetAlgo3 [6]	39.5	29.8	87.9	73.7	78.9	60.0
R8	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	50.0	44.6	96.2	94.0	92.2	85.9
DetAlgo2 [5]	50.0	49.8	96.0	92.5	92.0	87.4
DetAlgo3 [6]	41.7	30.6	87.4	75.4	81.3	61.5
R9	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	50.0	43.4	95.7	91.8	90.6	84.5
DetAlgo2 [5]	50.0	50.0	95.8	91.6	91.5	86.3
DetAlgo3 [6]	40.5	28.4	87.5	78.5	81.7	63.2
R10	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	48.2	38.5	94.2	86.6	87.9	77.9
DetAlgo2 [5]	50.0	48.0	94.7	91.8	90.7	87.4
DetAlgo3 [6]	41.6	30.2	87.0	76.2	80.9	61.9
R11	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	49.0	37.3	95.2	88.0	89.6	76.3
DetAlgo2 [5]	50.0	50.0	93.7	92.7	92.2	88.6
DetAlgo3 [6]	41.9	31.7	87.1	80.0	80.7	67.8
R12	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	50.0	43.7	95.4	90.6	91.4	83.0
DetAlgo2 [5]	50.0	48.8	94.8	91.4	91.4	86.0
DetAlgo3 [6]	41.8	30.8	87.6	76.4	80.5	64.1
R13	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	48.7	37.4	94.2	86.8	89.0	76.6
DetAlgo2 [5]	50.0	49.1	93.6	91.6	91.4	86.7
DetAlgo3 [6]	41.4	32.2	86.9	78.5	80.2	64.8
R14	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**	**SOTA [3]**	**Proposed**
DetAlgo1 [4]	50.0	42.6	97.6	90.2	94.8	83.7
DetAlgo2 [5]	50.0	49.5	97.2	95.7	93.4	88.0
DetAlgo3 [6]	46.4	34.0	92.2	83.8	83.2	72.1

Table 9.10: CFIA Attack Detection using DetAlgo1 [4], DetAlgo2 [5], and DetAlgo3 [6]

Hybrid features (DetAlgo2) [5]: DetAlgo2 is based on two different colors spaces. Given the RGB image, firstly, it is converted to HSV and YCbCr color space. In the next step, micro-texture features are computed using pyramid-LBP and passed through the SRKDA classifier. The final classification is performed using SUM rule fusion to make the final decision on detection.

Residual noise-based S-MAD Network (DetAlgo3) [6]: DetAlgo3 is based on the computing the residual noise using the Multi-Scale Context Aggregation Network (MS-CAN). The residual noise is further processed through Alexnet to obtain the classified features using the Collaborative Representative Classifier (CRC) to make the final decision to detect the attack.

To benchmark CFIA detection performance we resort to the off-the-shelf S-MAD. Three different S-MAD methods employed in this work are trained using different morph generation types (landmark-based and deep learning) and three different mediums (Digital, print-scanned, and print-scanned compression) generated using the publicly available FRGC face database. The quantitative results are presented using the ISO/IEC metrics [187] which are as follows: 1) Attack Presentation Classification Error Rate (APCER (%)) defining the percentage of attack images (morph images) incorrectly classified as bona fide images [187] , 2) Bonafide Presentation Classification Error Rate (BPCER (%)) defining the percentage of bona fide images incorrectly classified as attack images [187] and 3) Detection Equal Error Rate (D-EER (%)) [51]. The detection performance is benchmarked with both SOTA and proposed CFIA images and quantitative results are presented in Table 9.10 and bar chart with D-EER (%) on all 14 different regions. Based on the obtained results following are the main observations:

- The CFIA detection performance is degraded with all three detection algorithms.
- Among three different detection algorithms. DetAlgo3 indicates the better detection accuracy attributed to the quantification of residual noise.
- Among the 14 different regions, the degraded detection performance is noted with the R14 on all three detection algorithms.

Thus, based on the obtained results, we can conclude that the detection of CFIA attacks is very challenging and this needs more sophisticated detection algorithms to be devised for reliable detection.

9.9 Conclusions and Future-Work

In this work, we presented a new type of digital attack based on the facial attributes and we termed it as Composite Face Image Attack (CFIA). Given the facial images from the two contributory data subjects, the proposed CFIA will first segment the face images into six different attributes independently. Then, these segments are blended using a transparent mask based on both single face-attribute and multiple face attributes. These attributes are processed using the image inpainting based on pre-trained GAN to generate the final CFIA samples. In this work, given the face images from two contributory data subjects, we generate 526 different composite face images based on single and multiple face attributes. We contributed a new dataset with 1000 unique identities that will result in 526000 CFIA samples. Extensive experiments are performed to evaluate the attack potential of the newly generated CFIA using four different FRS. To effectively benchmark the vulnerability of the generated CFIA, we have introduced a generalized vulnerability metric. Further, we benchmark the detection accuracy using both human and automatic detection techniques. Our results demonstrated that the proposed CFIA could indicate the vulnerability of the FRS while it is difficult to detect using both human and automatic detection techniques. In the future work, we would like to extend the present work in several directions: 1) Generation of composites of higher quality, 2) Evaluation of the proposed method on real face images on public datasets, 3) Development of novel detection techniques.

9.10 Appendix: Role of FTAR in computing vulnerability

In this appendix, we present additional results on the vulnerability of COTS to illustrate the importance of FTAR in computing the G-MAP. The use of academic FRS does not include quality estimation to optimize the verification performance; thus, FTAR can be assumed to be zero. However, with COTS FRS (which is more practical), the captured face quality is imposed because of which the FRS seeks good-quality face images to optimize the verification performance. The requirement of good quality will result in the rejection of probe attempts deemed low-quality face capture and, thus, the failure of verification with reasonable attempts. Hence the proposed FTAR will penalise the failure to verify with a reasonable attempt.

Table 10.3 and 9.12 indicates the quantitative results of two different Commercial-Off-The-Shelf (COTS) such as Neurotechnology Version 10.0 [197] and Cognitec FaceVACS-SDK Version 9.4.2 [198] [2] in which G-MAP is computed with the multiple attempts on 14 different combinations. These 14 regions are the same as those used in the earlier sections of the papers that are representative of low, moderate and high vulnerability combinations. As noticed from the Tables 10.3 and 9.12 the G-MAP with FTAR will indicate the less vulnerability meaning that, the COTS FRS fail to perform the verification. Therefore accountability to FTAR is important to be consider for vulnerability calculation.

G-MAP % (Probe Attempts) with FTAR															
FRS	**Method**	**R1**	**R2**	**R3**	**R4**	**R5**	**R6**	**R7**	**R8**	**R9**	**R10**	**R11**	**R12**	**R13**	**R14**
Neurotech (FAR=0.1%)	**SOTA [3]**	18.2	10.4	10.9	8.1	16.2	14.6	19.0	18.8	19.3	14.2	14.0	21.6	15.0	11.3
	Proposed	13.8	10.2	9.7	17.6	13.5	14.4	16.0	17.1	19.3	22.2	22.9	21.0	23.7	**23.3**
Cognitec (FAR=0.1%)	**SOTA [3]**	31.6	22.2	23.3	19.8	27.7	28.8	33.1	34.0	37.9	30.0	24.8	41.1	25.1	21.3
	Proposed	30.5	22.9	22.3	43.9	28.1	26.9	31.9	34.7	35.3	54.6	55.1	43.0	57.6	**60.7**

Table 9.11: Vulnerability analysis using the proposed GMAP metric (probe attempts-based with FTAR) for the proposed method and the SOTA [3]

G-MAP % (Probe Attempts) without FTAR															
FRS	**Method**	**R1**	**R2**	**R3**	**R4**	**R5**	**R6**	**R7**	**R8**	**R9**	**R10**	**R11**	**R12**	**R13**	**R14**
Neurotech (FAR=0.1%)	**SOTA [3]**	54.4	33.1	35.1	32.4	50.0	44.3	59.1	58.1	59.5	51.3	50.3	65.4	55.0	45.3
	Proposed	43.1	31.6	33.1	57.8	41.8	43.5	49.7	51.9	56.9	72.2	75.2	63.9	79.5	79.2
Cognitec (FAR=0.1%)	**SOTA [3]**	31.9	22.5	23.5	20.0	28.0	29.2	33.5	34.4	38.3	30.3	25.2	41.6	25.5	21.6
	Proposed	30.8	23.2	22.6	44.4	28.4	27.2	32.2	35.1	35.7	55.2	55.8	43.4	58.3	61.4

Table 9.12: Vulnerability analysis using the G-MAP metric (Probe Attempts- without FTAR) for the proposed method and the SOTA [3]

[2] Disclaimer: These results were produced in experiments conducted by us and should; therefore, the outcome does not necessarily constitute the best the algorithm can do.

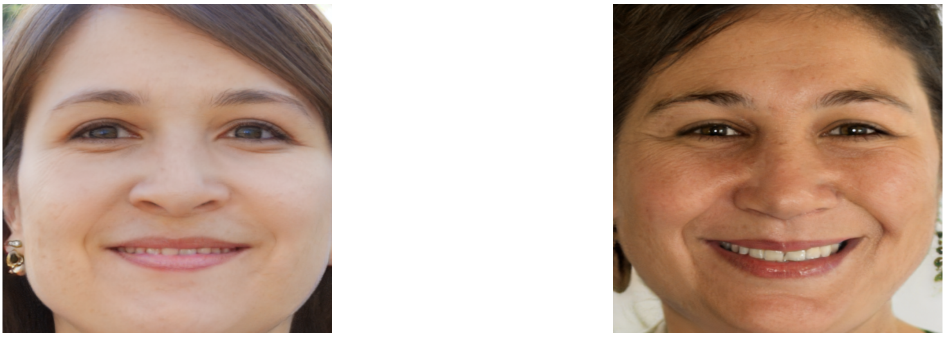


Figure 9.14: Bona fide subjects used for composition results as shown in Figures 9.15-9.16, 9.17, 9.18, 9.19, 9.20 and 9.21

9.11 Supplementary Material: Deep Composite Face Image Attacks: Generation, Vulnerability and Detection

9.11.1 Full Composition Results for two contributory data subjects.

In this section, we present the 526 composition images for bona fide images from Figure 9.14 in Figures 9.15- 9.16, 9.17, 9.18, 9.19, 9.20 and 9.21. Note the composition figures are in left to right order of the composition regions mentioned in Table 2 from the main manuscript. Further. each figure mentions the combination for the starting and ending CFIA.

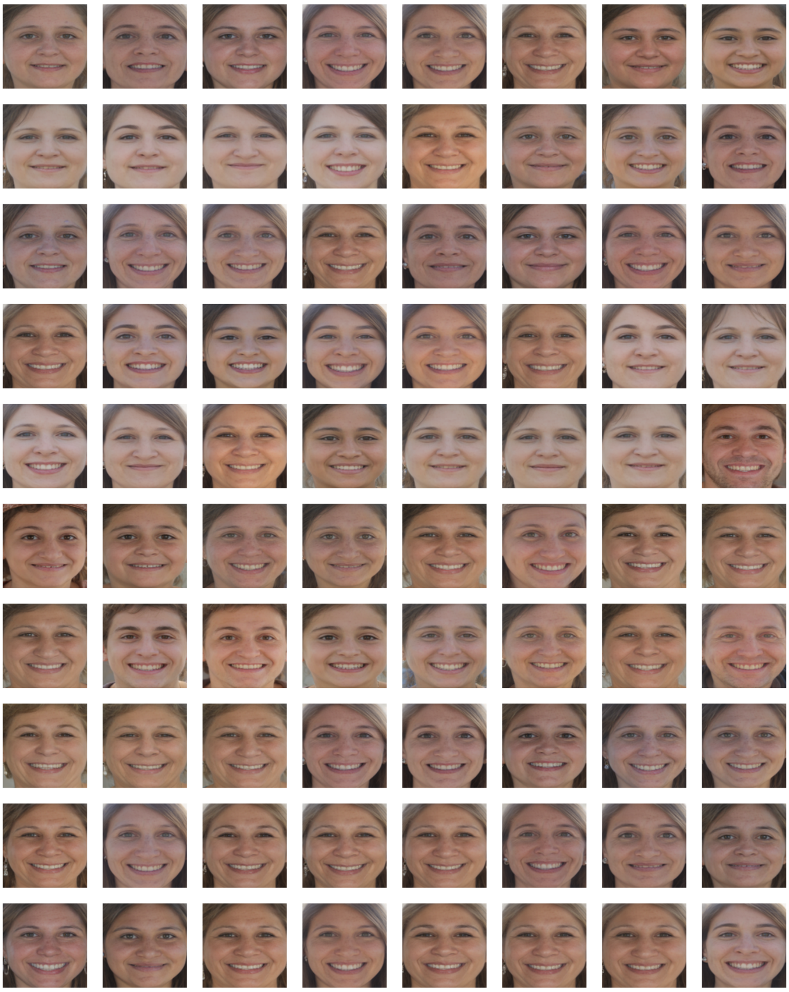


Figure 9.15: Bona fide subjects used for composition results where starting composition is E-H and ending composition is HN-EM

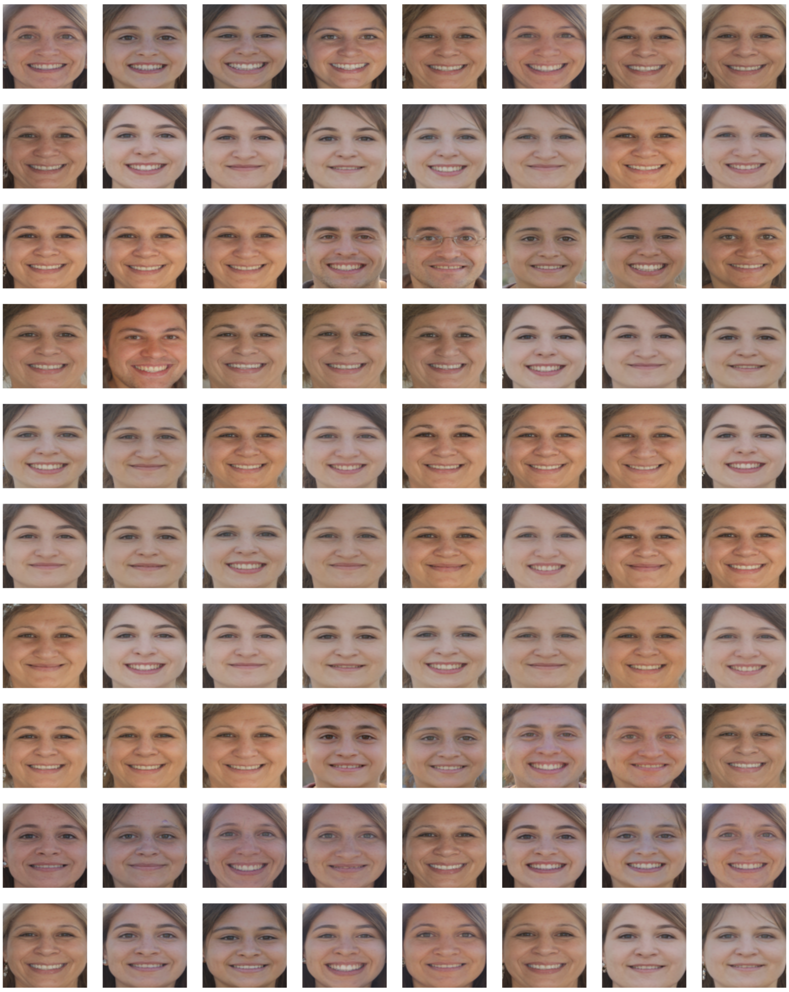


Figure 9.16: Bona fide subjects used for composition results where starting composition is HN-EN and ending composition is HSE-H

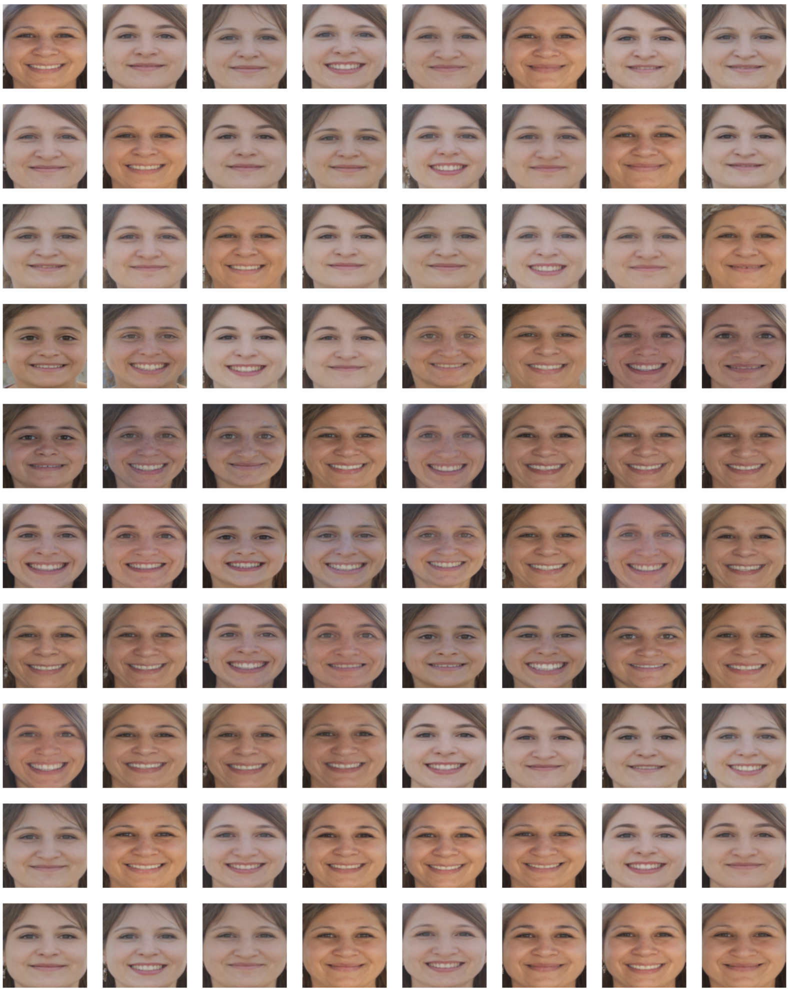


Figure 9.17: Bona fide subjects used for composition results where starting composition is HSE-S and ending composition is HSM-SN

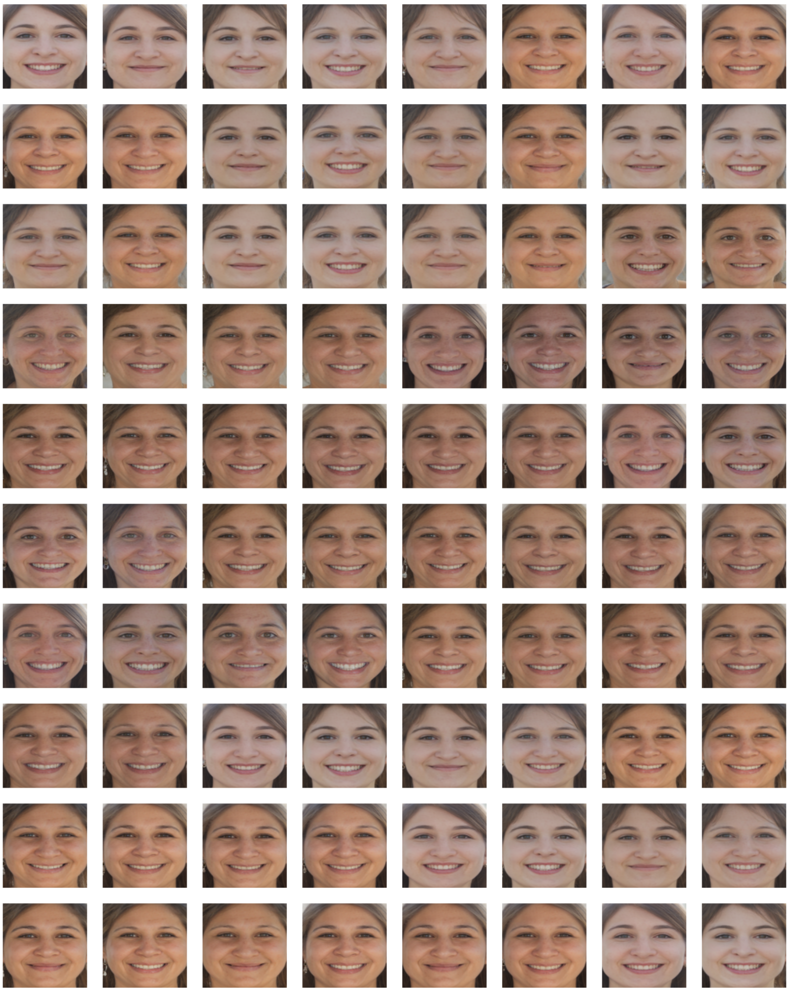


Figure 9.18: Bona fide subjects used for composition results where starting composition is HSN-EM and ending composition is HSN-HEM

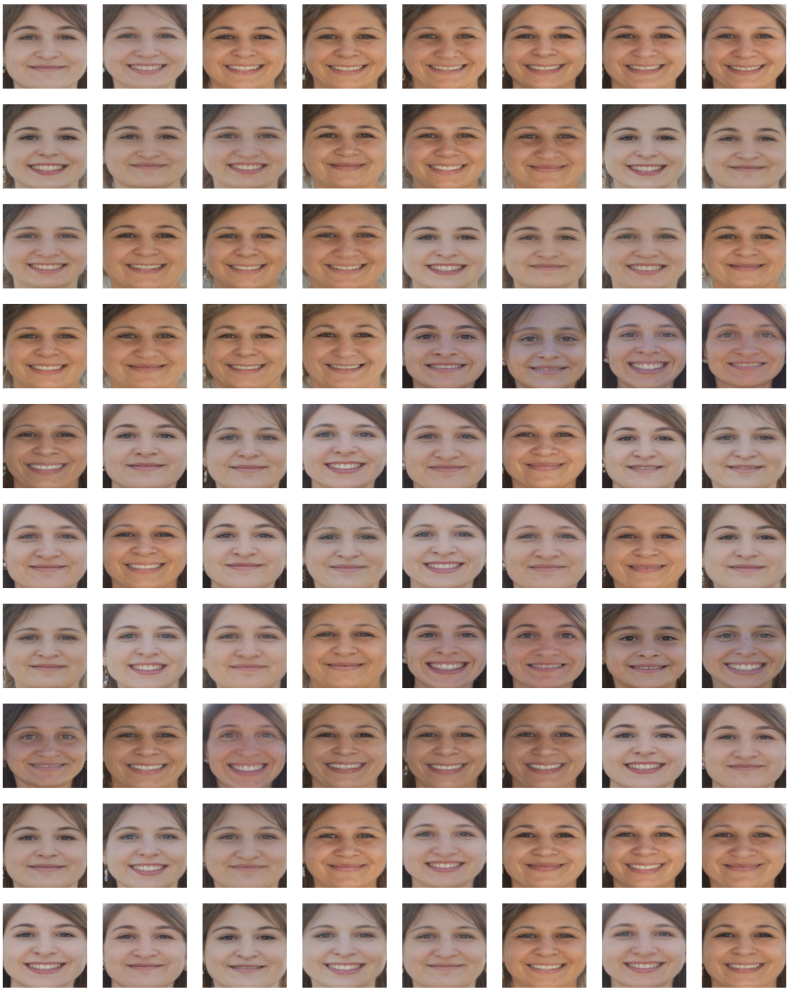


Figure 9.19: Bona fide subjects used for composition results where starting composition is HSN-HEN and ending composition is HSEN-SE

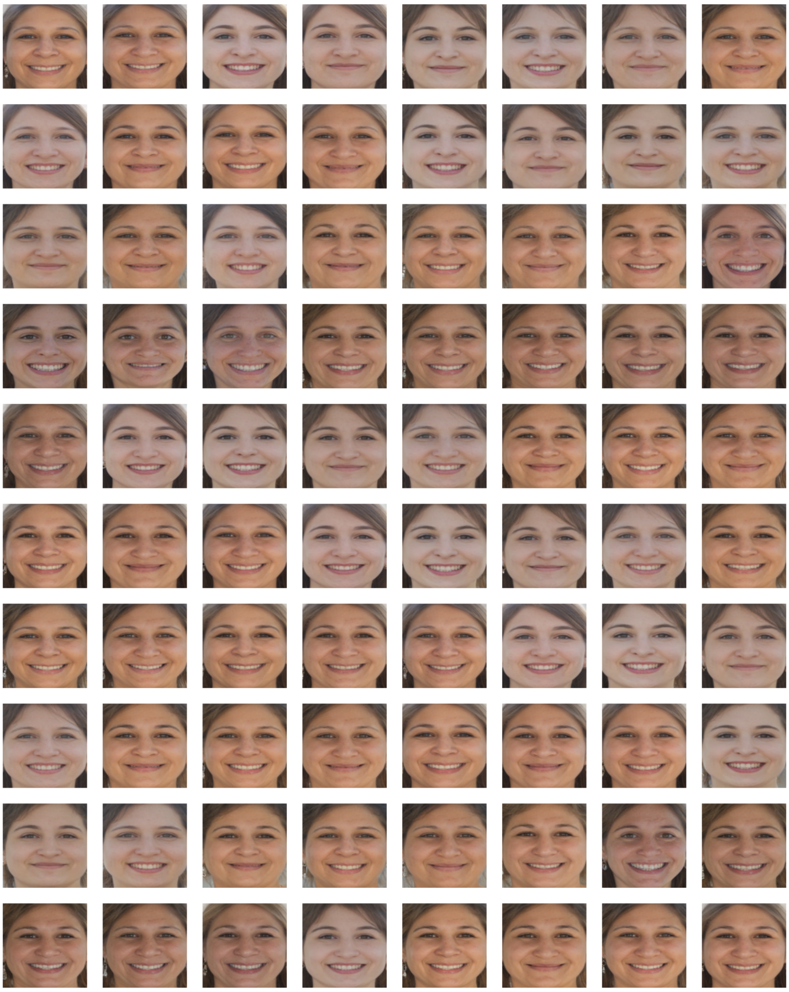


Figure 9.20: Bona fide subjects used for composition results where starting composition is HSEN-SM and ending composition is HSEM-SENM

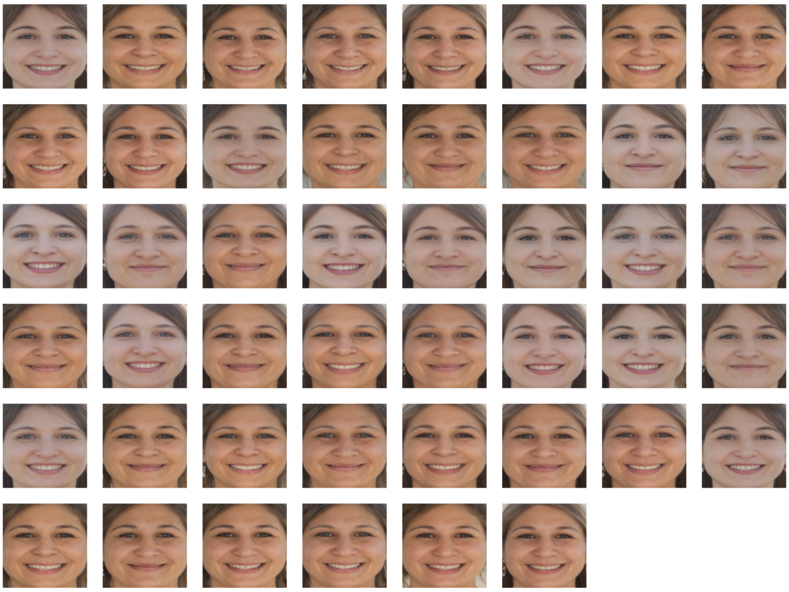


Figure 9.21: Bona fide subjects used for composition result where starting composition is HSEN-HENM and ending composition is HBSENM-HBSENM

Chapter 10

Article 6:3D Face Morphing Attacks: Generation, Vulnerability and Detection (RQ4)

Jag Mohan Singh, and Raghavendra Ramachandra. 3D Face Morphing Attacks: Generation, Vulnerability and Detection. In *IEEE TRANSACTIONS ON BIOMETRICS, BEHAVIOR, AND IDENTITY SCIENCE, DOI 10.1109/TBIOM.2023.3324684*

10.1 Abstract

Face Recognition systems (FRS) have been found vulnerable to morphing attacks, where the morphed face image is generated by blending the face images from contributory data subjects. This work presents a novel direction toward generating face-morphing attacks in 3D. To this extent, we have introduced a novel approach based on blending the 3D face point clouds corresponding to the contributory data subjects. The proposed method will generate the 3D face morphing by projecting the input 3D face point clouds to depth maps and 2D color images, followed by the image blending and wrapping operations performed independently on the color images and depth maps. We then back-project the 2D morphing color map and the depth map to the point cloud using the canonical (fixed) view. Given that the generated 3D face morphing models will result in holes due to a single canonical view, we have proposed a new algorithm for hole filling that will result in a high-quality 3D face morphing model. Extensive experiments are carried out on the newly generated 3D face dataset comprised of 675 3D scans corresponding to

41 unique data subjects and the publicly available Facescape database with 100 unique identities. Experiments are performed to benchmark the vulnerability of the proposed 3D morph generation scheme against automatic 2D, 3D FRS and human observer analysis. We also present the quantitative assessment of the quality of the generated 3D face morphing models using eight different quality metrics. Finally, we have proposed three different 3D face Morphing Attack Detection (3D-MAD) algorithms to benchmark the performance of the 3D face morphing attack detection techniques.

10.2 Introduction

Face Recognition Systems (FRS) are being widely deployed in numerous applications related to security settings such as automated border control (ABC) gates and commercial settings like eCommerce and e-banking scenarios. The rapid evolution of FRS can be attributed to the advances in deep learning FRS [27, 22], which improved accuracy in real-world and uncontrolled scenarios. These factors accelerated the use of 2D face images in electronic machine-readable documents (eMRTD), which are exclusively used to verify the owner of a passport at various ID services, including border control (both automatic and human). Because most countries still use printed passport images for the passport application process, the face morphing attack has indicated the vulnerability of both human and automatic FRS [142, 160]. Face morphing is the process of blending multiple face images based on either facial landmarks [206] or Generative Adversarial Networks [207] to generate a morphing face image. The extensive analysis reported in the literature [83, 208, 48, 49] demonstrated the vulnerability of 2D face morphing images to both deep learning and commercial off-the-shelf FRS.

There exist several techniques to detect the 2D face morphing attacks that can be classified as [69] (a) Single image-based Morph Attack Detection (S-MAD): where the face Morphing Attack Detection (MAD) techniques will use the single face image to arrive at the final decision (b) Differential Morphing Attack Detection (D-MAD): where a pair of 2D face images are used to arrive at the final decision. S-MAD and D-MAD techniques have been extensively studied in the literature, resulting in several MAD techniques. The reader is advised to refer to a recent survey by Venkatesh et al. [69] to obtain a comprehensive overview of existing 2D MAD techniques. Despite the rapid progress in 2D MAD techniques, a recent evaluation report from NIST FRVT MORPH [130] indicates the degraded detection of 2D face morphing attacks. Thus, 2D MAD attacks, especially in the S-MAD scenario, present significant challenges for reliable detection. These factors motivated us to explore 3D face morphing so that depth information may provide a reliable cue that makes morphing detection easier. 3D face recognition has been

widely studied over the past several decades, resulting in several real-life security-based applications with 3D face photo-based national ID cards [209], [210], [211], 3D face photo-based driving license cards [211] and 3D face-based automatic border control gates (ABC) [212]. The real case reported in [213] demonstrated using a 2D rendered face image from a 3D face model instead of a real 2D face photo to obtain the ID card bypassing the human observers in the ID card issuing protocol. Although most real-life 3D face applications are based on comparing 3D face models against 2D face images for verification, this is mainly because e-passports use 2D face images.
However, the use of 3D to 3D comparison will be realistic, especially in the border control scenario, as both ICAO 9303 [214] and ISO/IEC 19794-5 [215] standards are well defined to accommodate the 3D face model in the 3rd generation e-passport. The 3D face ID cards are a reality as they are being deployed in countries such as the UAE [209], which can facilitate both human observers and automatic FRS to achieve accurate, secure, and reliable ID verification. Further, the evolving technology has made it possible for 3D face imaging on handheld devices and smartphones (e.g., Apple Face ID [216] uses 3D face recognition) that can further enable remote ID verification based on 3D face verification. These factors motivated us to investigate the feasibility of generating 3D face morphing and studying their vulnerability and detection. An early attempt in [217] (master's thesis) employed the 3DMM [18] technique to generate a 3D face morphing model. However, the reported results indicate the lowest vulnerability to conventional FRS, indicating the limitation of the 3DMM..

This work presents a novel method for generating 3D face morphing using 3D point clouds. Given the 3D scans from the accomplished and malicious actors, the proposed method will project the 3D point clouds to the depth maps & the 2D color images, which are independently blended, warped, and back-projected to the 3D to obtain 3D face morphing. The motivation for projecting to the 2D for morphing is to effectively address the non-rigid registration, especially with the high volume of point clouds (85K) that needs to be registered between two unique data subjects. Further, using canonical view generation to project from 3D to 2D and back project to 3D will assure a high-quality depth even for the morphed face images, thus indicating the high vulnerability of the FRS. Therefore, this is the first framework to address the generation of 3D face morphing of two unique face 3D scans that can result in vulnerability to FRS. More particularly, we aim to answer the following research questions, which will be answered systematically in this study:

- **RQ#1:** Does the proposed 3D face morphing generation technique yield a high-quality 3D morphed model?

- **RQ#2:** Does the generated 3D face morphing model indicate the vulnerability for both automatic 3D FRS and human observers?
- **RQ#3:** Are the generated 3D face morphing models more vulnerable when compared to 2D face morphing images for both automatic 3D FRS and human observers?
- **RQ#4:** Does the 3D point cloud information be used to detect the 3D face morphing attacks reliably?

We systematically address these research questions through the following contributions:

- We present a novel 3D face morphing generation method based on the point clouds obtained by fusing depth maps and 2D color images to generate the 3D face morphing model.
- Extensive analysis of the vulnerability of the generated 3D face morphing is studied by quantifying the attack success rate to 3D FRS. The vulnerability analysis is also performed using 2D FRS (deep learning and COTS).
- Human observer analysis for detecting the 3D face morphing and 2D face morphing is presented to study the significance of depth information in detecting the morphing attack.
- The quantitative analysis of the generated 3D morphed face models is presented using eight different quality features representing color and geometry.
- We present three different 3D MAD techniques based on the deep features from point clouds to benchmark the 3D face MAD.
- A new 3D face dataset with bona fide and morphed models is developed corresponding to 41 unique data subjects resulting in 675 3D scans. We collected a new 3D face dataset as we were interested in capturing high-resolution (suitable for ID enrolment) inner face data [218] Our 3D face dataset consists of raw 3D scans (number of 3D vertices between 31289 & 201065) and processed 3D scans (number of 3D vertices between 35950 & 121088), which is much higher than existing 3D face datasets[1].
- The proposed method is benchmarked on a publicly available dataset from FaceScape and the newly constructed dataset.

[1] The reader is referred to Table 1 of 3D face datasets (inner face data only) from the survey by Egger et al. [218])

In the rest of the paper, we introduce the proposed method in Section 10.3 and experiments & results in Section 10.4. This is followed by a discussion about the different aspects of the proposed method in Section 10.5, followed by limitations & potential future-works in Section 10.6 and finally conclusions in Section 10.7.

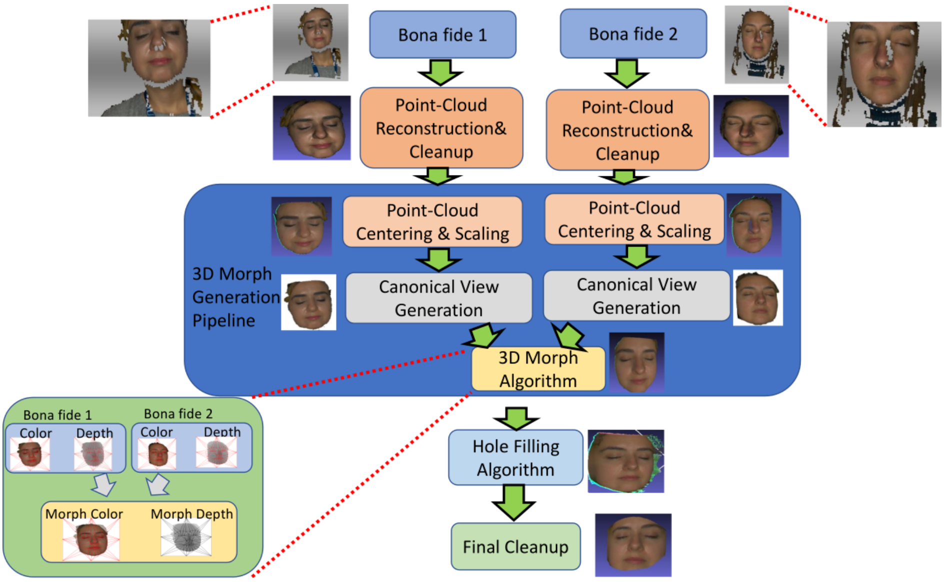

Bona fide 1
Bona fide 2
Point-Cloud Reconstruction& Cleanup
Point-Cloud Reconstruction& Cleanup
Point-Cloud Centering & Scaling
Point-Cloud Centering & Scaling
3D Morph Generation Pipeline
Canonical View Generation
Canonical View Generation
3D Morph Algorithm
Bona fide 1
Color
Depth
Bona fide 2
Color
Depth
Morph Color
Morph Depth
Hole Filling Algorithm
Final Cleanup

Figure 10.1: Block diagram of the proposed 3D face morphing generation technique

10.3 Proposed Method

Figure 11.1 shows the block diagram of the proposed 3D face morphing generation framework based on the 3D point clouds. We are motivated to employ 3D point clouds over traditional 3D triangle mesh for two main reasons. The first is that connectivity information in a 3D triangle mesh leads to **overhead storage**, processing, managing, and manipulating the triangular meshes. Thus, 3D triangle meshes will significantly increase compute and memory, making them less suitable for low-compute devices. The second reason is that the commodity scanning devices (for example, the Artec Sensor) can reproduce detailed colored point clouds that capture appearance and geometry. Thus, allowing us to generate high-quality 3D face-morphing attacks.

However, the 3D face morphing generation using point clouds introduces numerous challenges (a) Establishing a dense 3D correspondence between two different bona fide 3D point clouds that are to be morphed. Because 3D face point clouds from two different subjects are affected by various factors such as differences in input point density, reliable detection of 3D facial key points, and estimation of affine/perspective warping (b) Locally affine deformation present between two different 3D point clouds to be morphed is difficult to estimate [219, 220, 221]. (c) The misalignment of dense 3D correspondence between the two different 3D point clouds to be morphed increases with non-rigid deformation [222].

The crucial part of 3D morphing using point clouds is reliable alignment before performing the morphing operation. Given the 3D face point clouds on the source and the target face, the point cloud registration can be defined as aligning a source point cloud to a target point cloud. The point cloud registration can be grouped into three broad categories [223] namely 1) Deformation Field, 2) Extrinsic Methods and 3) Learning-based methods. Deformation Field-based techniques could be defined as the computation of deformation between the two-point clouds, which can be achieved either by assuming pointwise position [224] variables or by pointwise affine transformations [225]. Pointwise position variables methods are simplistic as they don't model deformations compared with pointwise affine transformations, which model local rotations. However, since the local transformations must be stored and computed at a per-point level, this results in high computational and memory costs. This limitation was overcome by deformation field-based methods using deformation graph embedding over the initial point set, which consists of fewer nodes than the underlying point set [226, 227]. Extrinsic methods are based on optimizing an energy function to compute the point set correspondence which usually includes an alignment term and a regularization term [226]. However, the optimization-based methods compute deterministic modeling of the transformation. Probabilistic modeling of transformation was done by Myronenko et al. [16] in their algorithm Coherent Point Drift (CPD) which assumes the source points to be centroids of equally-weighted Gaussian with isotropic covariance matrix in Gaussian Mixture Model (GMM). CPD consists of alignment and regularization terms for the transformation computation and performs non-rigid registration but has memory and compute costs. However, the main limitation of optimization-based methods is that they produce good results when the input surfaces are close. Further, they require good initialization of the correspondences and the lack of these, leads to convergence to local minima. This was overcome by learning-based data-driven methods, which are of two types (1) Supervised methods and (2) Unsupervised methods. Supervised methods require ground-truth data for training [228] but can work with varying point cloud density and underlying geometry. Unsupervised methods don't require ground-truth data and can be trained using a deformation module based on CNN, followed by an alignment module to compute the deformation [17].

However, the use of existing point cloud registration for this precise application of 3D face morphing point cloud generation will pose challenges such as: **registration using the same individual**: Point cloud registration has mainly focused on the non-rigid registration of two-point clouds from the same individual [223]. This is primarily because high-quality registration aims to produce a globally consistent 3D mesh. Thus, the registration methods have not been tested when two different point clouds are registered compared to those from the same individual. **Vertex accurate correspondence**: 3D Face Morphing requires perfect vertex correspondence between the source and target point clouds, which is challenging and has not been evaluated extensively. **Low vertex count point clouds**: Point cloud registration, especially using learning-based methods, has network architectures based on point clouds with a low number of vertices (1024). Thus, registering point clouds with many vertices (75K) has not been evaluated extensively and is therefore suitable for low-resolution face images. To effectively address these challenges, the proposed method consists of four stages, including (1) point cloud reconstruction and cleanup, (2) 3D morph generation, (3) hole-filling algorithm, and (4) final cleanup. In the following subsections, these steps are discussed in detail.

10.3.1 Point Cloud Reconstruction & Cleanup

We capture a sequence of raw 3D scans using Artec Eva sensor [229] from two data subjects to be morphed (S_1 and S_2). In this work, we consider the case of morphing two data subjects at a time because of its real-life applications, as demonstrated in several 2D face morphing works [142, 69]. We process both S_1 and S_2 by performing a series of pre-processing operations such as noise filtering, texturing, and fusion of input depth maps to generate the corresponding point clouds P_1 and P_2. These operations are carried out using Artec Eva Studio SDK filters together with the Meshlab filter [230]. The cleaned and process point clouds are qualitatively shown in Figure 11.1.

10.3.2 3D Morph Generation Pipeline

In the next step, we process the point clouds P_1 and P_2 to generate a 3D face morphing point cloud by the following series of operations which are discussed below:

Point-Cloud Centering & Scaling

We first compute the minimum enclosing spheres using the algorithm from Gärtner et al. [231] to get the two bounding spheres with centers and radii (C_1,r_1), & (C_2,r_2) corresponding to the point cloud P_1, and P_2 respectively. Note $P_1 = (v_1^1, \ldots, v_1^{\mathrm{n}1})$ where v_1^i is the i^{th} 3D vertex, and $n1$ is the number of points in the point cloud P_1, and $P_2 = (v_2^1, \ldots, v_2^{\mathrm{n}2})$ where v_2^i is the i^{th} 3D vertex, and

$n2$ is the number of points in the point cloud P_2. We then subtract the sphere center C_1 from each 3D vertex of P_1 and repeat the same operation on P_2 with C_2. Finally, the centered point clouds are scaled to the common radius, normalizing the 3D point clouds to the common scale. The resulting centered and scaled point clouds corresponding to P_1 and P_2 are denoted as PC_1 and PC_2, respectively. Figure 11.1 shows this operation's qualitative result, which shows centered and scaled 3D point clouds.

Canonical View Generation

This step performs the fine alignment by projecting the 3D face point clouds PC_1 and PC_2 to the canonical (fixed) view. This step aims to keep the view and projection matrix identical to the 3D face point clouds PC_1 and PC_2. We then project PC_1 and PC_2 to generate 2D color images and depth maps using the canonical view parameters. The generated 2D color images and depth maps are denoted as (I_1,D_1) and (I_2,D_2) that corresponds to the point clouds PC_1, and PC_2 respectively. We particularly choose the canonical view for the fine alignment because the traditional scheme of alignment, such as Iterative Closest Point (ICP) [222] doesn't provide a good alignment result when used on point clouds[220]. This can be attributed to the limitations of the ICP to function when a locally affine/non-rigid deformation exists between the point clouds[232] The qualitative results of the canonical view transformation are shown in Figure 11.1, which demonstrates the aligned 2D color images and depth maps zoomed in the inset image.

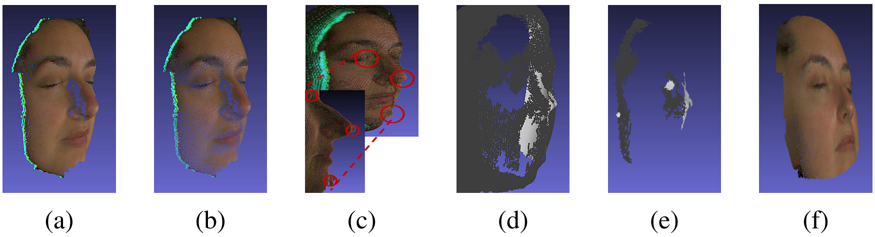


Figure 10.2: Qualitative results of the hole filling algorithms (a) Input Point Cloud with holes, (b) Point Cloud with Normals which has noise, (c) Point Cloud with Screened Poisson Reconstruction [13] where artifacts are shown in the inset, (d) Point Cloud Reconstructed with APSS [14], (e) Point Cloud Reconstructed with RIMLS [15], (f) Point Cloud Hole Filled using Proposed Method

3D Morph Generation

Given the 2D face color images (I_1,I_2) and depth-maps (D_1,D_2) corresponding to PC_1, PC_2. We perform the morphing operation as explained in the Algorithm 5. The primary idea is to perform the morphing in 2D and back-project to 3D. The primary motivation for using a 2D morph generation method is to address the

Algorithm 5: 3D Face Morphing Algorithm

```
Input ($I_1$, $I_2$, $D_1$, $D_2$, $CV$)
Output ($P_M$)
Detect Facial Keypoints on $K_1$ on $I_1$, and $K_2$ on $I_2$ using Dlib [42], and generate key-points of the morph using Equation 10.1.
Perform Delaunay Triangulation on $K_M$ which is obtained by blending $K_1$ and $K_2$ using Equation 10.1.
Estimate Affine Warping between corresponding triangles of $K_1$ & $K_M$ denoted as $w_1^M$, and for $K_2$ & $K_M$ denoted as $w_2^M$.
Apply affine warping $w_1^M$ on $I_1$ to obtain $I_{1M}$, and on $D_1$ to obtain $D_{1M}$.
Apply affine warping $w_2^M$ on $I_2$ to obtain $I_{2M}$, and on $D_2$ to obtain $D_{2M}$.
Obtain morphed color image $I_M$ using the warped keypoints from the color images $I_1$, and $I_2$ using Equation 10.1, and morphed depth map $D_M$ using Equation 10.2.
Obtain the morphed point cloud by back-projecting $I_M$, and $D_M$ to obtain the colored 3D point cloud $P_M$ with 3D coordinates $\forall i \in \{1, \cdots, n3\} (x_i, y_i, z_i) = (x_i, y_i, D_M(x_i, y_i))$ and color $\forall i \in \{1, \cdots, n3\}\text{Color}(x_i, y_i, z_i) = C_M(x_i, y_i))$ where $n3 = \min(n1, n2)$.
```

challenge of finding correspondence between PC_1 and PC_2. The underlining idea is to perform the steps of morphing (facial landmark detection, Delaunay triangulation, & warping) on 2D color images and re-use the same (facial landmark locations, triangulation, and warping) on the depth maps. In this work, we have used the blending (morphing) factor (α) as 0.5 as it is well demonstrated to be highly vulnerable in the earlier works on 2D face morphing [207]. The morphing is carried out as mentioned in the equation below:

$$\begin{aligned} I_M = \alpha \times I_1(K_1') + (1 - \alpha) \times I_2(K_2') \\ K_1' = w_1^M(K_1) \\ K_2' = w_2^M(K_2) \\ K_M = \alpha \times K_1 + (1 - \alpha) * K_2 \end{aligned} \tag{10.1}$$

where α is the blending factor, K_1 denotes 2D facial landmark locations corresponding to I_1, K_2 denotes 2D facial landmark locations corresponding to I_2, K_M

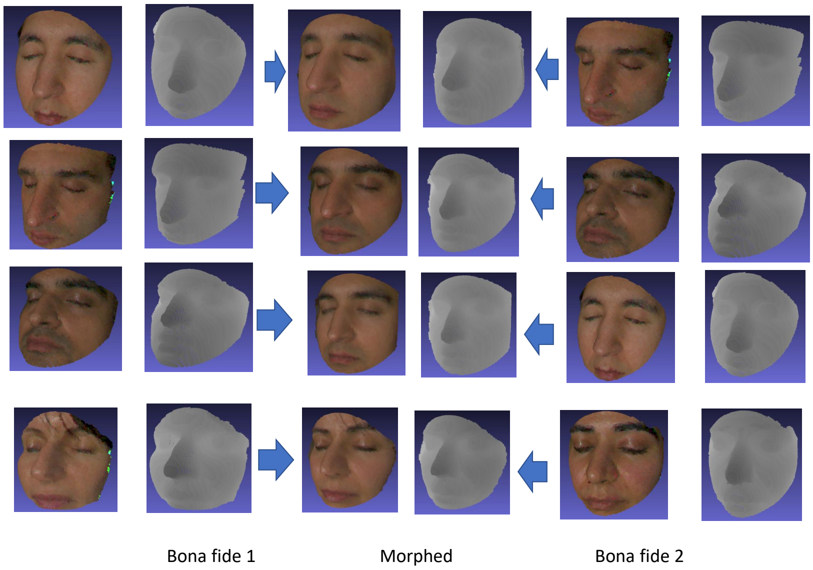


Figure 10.3: Illustration of 2D color image and depth maps for bona fide and morphs generated using the proposed method

is generated by blending K_1, & K_2, w_1^M denotes the warping function from K_1 to K_M, w_2^M denotes the warping function from K_2 to K_M, and I_M is the morphed 2D color image. Similarly, the same operations are carried out on the depth maps as shown in the equation below:

$$D_M = \alpha \times D_1(K_1') + (1-\alpha) \times D_2(K_2') \tag{10.2}$$

where D_M is the morphed depth map.

In the next step, we back-project I_M, and D_M to get the 3D face morphing point cloud $P_M = (v_M^1, \ldots, v_M^{n3})$ where $n3 = \min(n1, n2)$ is the number of vertices. Note each 3D vertex is obtained using $i = 1^{n3}(x_i, y_i, z_i) = (x, y, D_M(x, y))$ and the qualitative results is shown in Figure 11.1. However, generating the 3D face morphing will result in multiple holes due to a single canonical view. These holes are visible from other views. Therefore, we present a novel hole-filling algorithm to further improve the perceptual visual quality of the 3D face morphing.

10.3.3 Hole Filling Algorithm

In this step, we propose a new hole-filling algorithm tailored to this specific 3D face morphing generation problem. Since the holes are visible from different

Algorithm 6: Hole Filling Point Cloud

```
Input ($n4$-views)
Output ($C_{\textbf{hf}}$,$D_{\textbf{hf}}$,$P_{\textbf{hf}}$)
Generate $n$ pairs of color-maps, and depth-maps
$\{(C_1, D_1), (C_2, D_2), \ldots, (C_j, D_j), \ldots, (C_{n4}, D_{n4})\}$, translated from the
canonical view.
for $j \leftarrow 1$ to $n4$ do
    Perform Image In-painting [233] on $C_j$, and $D_j$.
    Perform Image Registration of $C_j$ with the
    canonical view-point color-map $C_{\text{CV}}$ using
    the following steps:
            Feature Computation using Oriented
    FAST and Rotated BRIEF (ORB) Descriptor [234].
            Brute-Force Matching of features using Hamming Distance.
            Homography computation using inlier
    features.
            Perspectively warp the color and depth maps using computed
    homography.
end for
Average all the registered color-maps ($C_{\text{hf}}$) and the depth-maps ($D_{\text{hf}}$).
Back-Project the averaged color-map and
depth-map from 2D to 3D to generate
hole-filled point cloud ($P_{\text{hf}}$) using the canonical view parameters.
```

views, filling the holes in these views is necessary to improve the perceptual visual quality. Note that the holes are generated when the bona fide subject is looked at from a view different from the canonical camera, especially in high curvature regions such as the nose, as such areas are not completely visible from one canonical view. Therefore, we transform the 3D face morphing point cloud P_M multiple times independently to generate P_M^j where $j = 1 \ldots n4$ and $n4$ is the number of transformations and each transformation is a 3D translation [235]. In this work, we empirically choose the number of 3D translations to 7 to balance computational cost and the visual quality achieved after the hole filling. Using more 3D translations will significantly increase the computational cost and fail to improve the visual quality. We tried the conventional approach of hole filling using 3D triangulation of 3D point cloud proposed in [13],[14],[15]. Figure 10.2 shows the qualitative results of three different SOTA triangulation algorithms that indicate non-satisfactory results. This is because 3D orientation (3D normal) estimation indicates artifacts in the 3D triangulated mesh. Therefore, filling holes directly in the

3D point cloud is challenging, as the underlying surface (manifold) is not known in advance. The errors in 3D orientation estimation make it difficult to employ the conventional 3D hole-filling approaches.

This has motivated us to devise a new approach to achieve effective hole-filling. To this extent, we project each point cloud P_M^j to the 2D face morphing color image (C_j) and its corresponding depth map (D_j). We fill the holes in C_j & D_j using steps 2 to 9 described in Algorithm 6. Finally, we obtain the hole-filled 3D face morphing point cloud (P_{hf}) as indicated in steps 10 and 11 in Algorithm 6. Figure 10.2 (e) shows the qualitative results of the proposed hole filling that indicated the superior visual quality compared to the existing methods.

10.3.4 Final Cleanup Algorithm

The final cleanup uses a clipping region outside a portion of the bounding sphere. The final result corresponding to the proposed 3D face morphing, a point cloud, is shown in Figure 10.3 for an example data subjects [2]. On the whole, the following are the main advantages of the proposed method:

- The proposed method performs the alignment based on 2D facial key points, which preserves the identity in the generated 3D face morphing attack sample.
- The proposed method results in low computation and memory compared with existing 3D-3D techniques by overcoming the 3D registration.
- The proposed method results in a high vulnerability of FRS as the identity features are preserved for contributed data subjects used to generate the morphing attack. Therefore, the proposed method can cause high-quality 3D face morphing attacks, resulting in the vulnerability of both 2D and 3D face recognition systems.
- The proposed method can handle wide variation in the 3D pose.

10.3.5 Qualitative and Quantitative Comparison of Proposed Method with SOTA

To illustrate the effectiveness of the proposed method, we selected a few SOTA methods based on non-rigid point cloud registration and methods generating a 3D face model from a 2D face image. Our current evaluation of SOTA for non-rigid point cloud registration (NRPCR) methods includes CPD by Myronenko et al. [16] and Cornet3D by Zeng et al. [17]. CPD is based on optimization and was

[2] Supporting Video is available at https://folk.ntnu.no/jagms/SupportingVideo.mp4

the SOTA method for NRPCR earlier, whereas Cornet3D is a more recent unsupervised deep learning-based method for NRPCR. Further, for evaluating methods generating a 3D face model from a 2D face image, we selected 3DMM by Blanz et al. [18] and a more recent deep-learning-based method FLAME by Li et al. [19]. 3DMM introduced the concept of the morphable model, where the parameters such as shape and texture can be controlled during 3D face synthesis. Further, 3DMM provided earlier SOTA results on 3D face generation from a 2D face image. FLAME enhanced the quality of the generated 3D face model from a 2D face image by using more controllable parameters such as pose, expression, shape and texture during the 3D face synthesis process.

Qualitative Comparison and Analysis

The results of qualitative comparison with SOTA are shown in Figure 10.4 and the quantitative vulnerability computed using MMPMR [83] and FMMPMR [49] (refer Section 10.4.3 for the definition of these metrics) is indicated in the Table 10.1. It can be noticed from Figure 10.4 that SOTA methods don't contain identity features of the 3D face morphing model to a large extent. However, CPD does contain the identity features of the 3D face morphing model but fails on the alignment of the two input point clouds, which results in double features such as eyebrows.

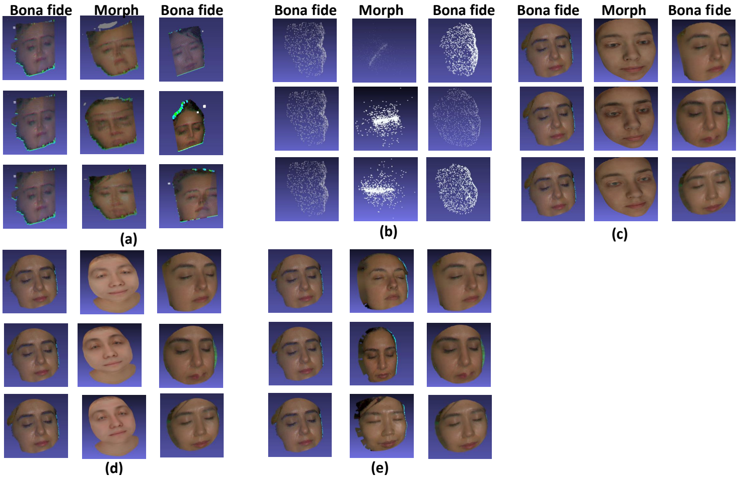
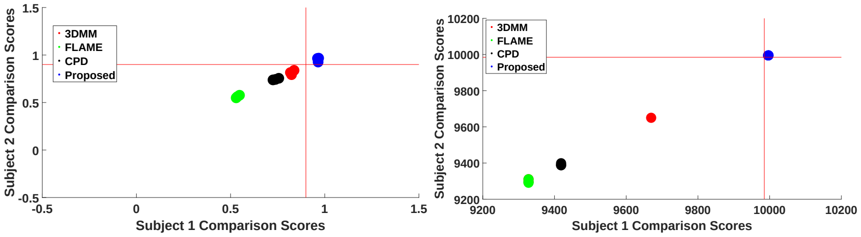


Figure 10.4: Illustration of the SOTA Comparison showing Bona fide and Morphs generated using (a) CPD [16], (b) Corrnet3D [17], (c) 3DMM [18] (d) FLAME [19], (e) Proposed Method. Note that both 3DMM and FLAME need a single image as input, and in the current evaluation, we pass a 2D rendering generated using the proposed method. Note that the proposed method shows high-quality rendering and identity features of the 2D face morphing image.

Orrnet3D produces lower-quality results, which can be attributed to the fact that the authors have yet to focus on face registration exclusively. Further, 3DMM and FLAME generate a 3D face model from a 2D face image. Thus, we passed the rendering (2D face image) of the 3D face morphing model as an input. However, these methods fail to preserve the identity features during the 3D face model generation, as seen from Figure 10.4. The generated 3D model has a low resemblance to the identity features of the face morphing image.

Table 10.1: Vulnerability of SOTA on Comparison Dataset

Feature	3DMM [18]	FLAME [19]	CPD [16]	Proposed
PointNet++ [21]	0	0	0	100%
LED3D [20]	66.67%	0	0	100%



(a) (b)

Figure 10.5: Illustration showing scatter plot of Comparison scores using Bona fide and Morphs generated using Proposed Method (a) LED3D [20] and (b) Pointnet++ [21] based where SOTA algorithms are 3DMM [18], FLAME [19], CPD [16]

Quantitative Comparison and Analysis

The results of the quantitative comparison are shown in Figure 10.5, where we have evaluated two 3D point feature extraction methods, namely LED3D [20] and Pointnet++ [21]. However, it can be seen that 3D comparison results in low values for SOTA compared to the proposed method. This can be attributed to the low-resolution of the identity-specific depth generation by the SOTA, which is also shown in Figure 10.6.

10.4 Experiments and Results

In this section, we present the discussion on extensive experiments carried out on the newly acquired 3D face dataset. We discuss the quantitative results of the various experiments, including vulnerability study on automatic FRS and human

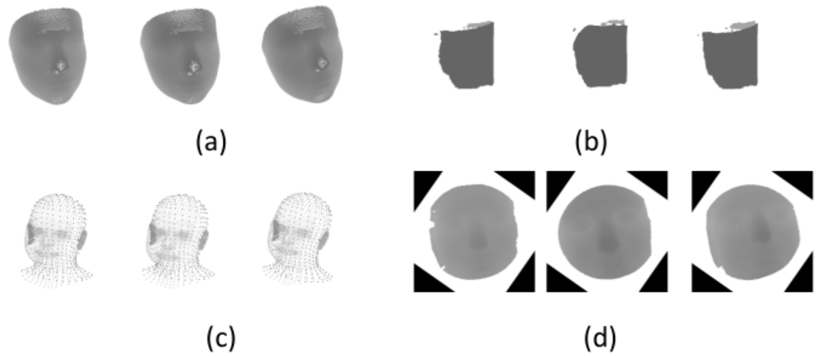
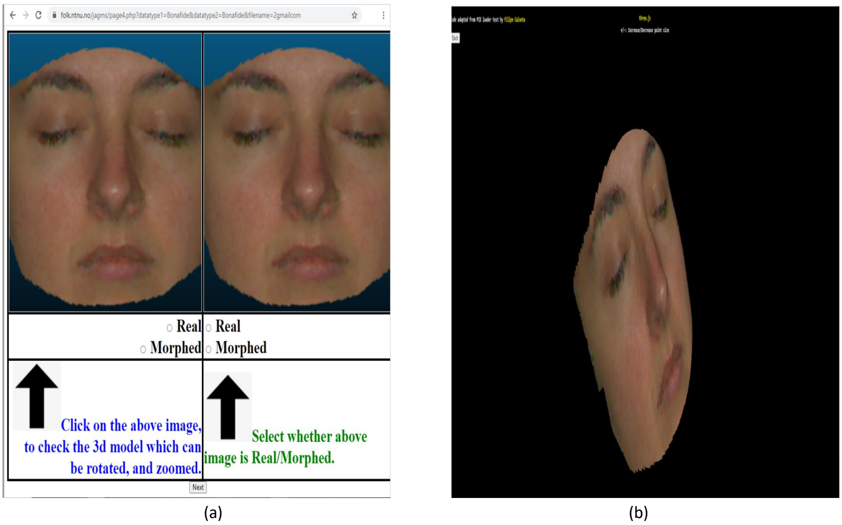


Figure 10.6: Illustration showing depth maps using SOTA and proposed method (a) 3DMM [18], (b) CPD [16], (c) FLAME [19] and (d) Proposed Method.


folk.ntnu.no/jagms/page4.php?datatype1=Bonafide&datatype2=Bonafide&filename=2gmailcom

○ Real ○ Real
○ Morphed ○ Morphed

Click on the above image, to check the 3d model which can be rotated, and zoomed.

Select whether above image is Real/Morphed.

Next

(a)

Code adapted from PCD loader test by Filipe Caixeta

three.js

+/-: Increase/Decrease point size

Back

(b)

Figure 10.7: Screenshots from the GUI of human observer web page (a) Full Page Screenshot, and (b) Screenshot of 3D model page.

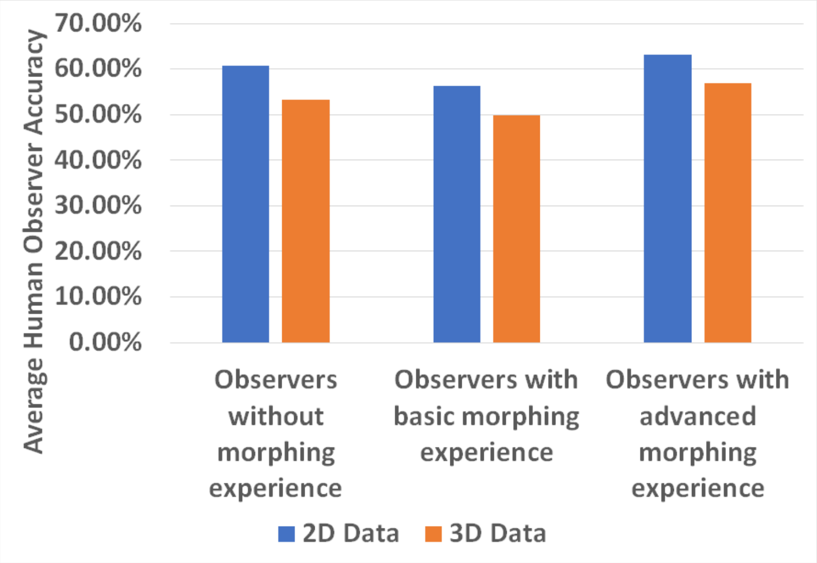


Figure 10.8: Illustration of average accuracy of human observer study, note that 2D accuracy is always higher than 3D.

observer study, quantitative quality estimation based on color and geometry of the generated 3D face morphing models and automatic detection of 3D MAD attacks.

10.4.1 3D Face Data Collection

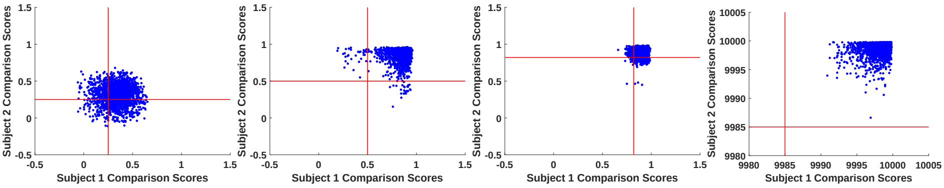
In this work, we have constructed a new 3D face dataset using the Artec Eva 3D scanner [229]. The data collection is carried out in an indoor lighting environment. The data subjects are asked to sit on the chair by closing their eyes to avoid the light's strong reflection from the 3D scanner. The 3D scanner is moved in the vertical direction to capture the 3D sequence.

We have used the Artec Studio Professional 14 for the 3D data collection and processing. We have collected the 3D face data from 41 subjects, including 28 males and 13 females. We have captured nine to ten samples for each data subject in three different sessions in three days. The statistics of the whole 3D face dataset are summarized in Table 10.2. We name our newly collected dataset as 3D Morphing Dataset (3DMD).

We may have used the existing 3D face datasets such as FRGC [236] and BU-3DFE [237]. However, the FRGC dataset provides a single depth map and a color image. Thus, a high-quality point cloud cannot be generated. Further, the dataset has a few misaligned color images and depth maps [238] that will result in a low-

Table 10.2: Statistics of newly collected 3D Morphing Dataset (3DMD)

3D face Bona fide		
Total Data Subjects	Males	Females
41	28	13
Total 3D samples	Males	Females
330	224	106
3D face Morphs		
Total 3D Morphs	Males	Females
345	278	67



(a) **(b)** **(c)** **(d)**

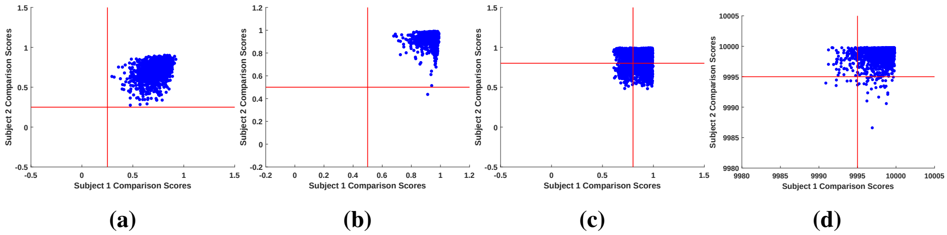
Figure 10.9: Vulnerability Plots using 2D & 3D FRS on 3D Morphing dataset (3FMD) (a) 2D face FRS using Arcface [22], (b) 2D face FRS using COTS, and (c) 3D face FRS using Led3D [20], and (d) 3D face FRS using Pointnet++ [21]

quality 3D morphing generation. The BU-3DFE [237] dataset does provide 3D models, but these are perfectly registered, and the capture conditions are identical for all the subjects. This does not model the real-world scenario of capturing 3D point clouds with changes in capture conditions that could happen during data collection. The quality of our 3D face dataset has a much higher number of 3D vertices between 35950 & 121088 for the inner face compared to previous methods [218]. These factors motivated us to generate a new 3D face dataset to enable a high-quality 3D face morphing generation suitable for the ID control scenario.

10.4.2 Human Observer Analysis

We perform the human observer analysis to evaluate the human detection performance of the generated 3D morphs. The survey is set up online[3] and is created using PHP, & HTML-CSS tools. GDPR norms are followed during the survey creation, and participants' email (used only for registration to avoid duplication), gender, &

[3] https://folk.ntnu.no/jagms



(a) (b) (c) (d)

Figure 10.10: Vulnerability Plots using 2D & 3D FRS on Facescape Dataset (a) 2D face FRS using Arcface [22], (b) 2D face FRS using COTS, and (c) 3D face FRS using Led3D [20], and (d) 3D face FRS using Pointnet++ [21]

Table 10.3: Vulnerability analysis of 2D and 3D FRS on 3D morphing dataset

Combined			**Male**		**Female**	
Algorithm	MMPMR%	FMMPR%	MMPMR%	FMMPR%	MMPMR%	FMMPR%
2D Vulnerability Analysis						
COTS	97.45%	89.78%	97.98%	90.65%	94.03%	86.36%
Arcface	63.81%	28.66%	64.92%	27.13%	59.70%	33.33%
3D Vulnerability Analysis						
LED3D [20]	81.69%	54.00%	82.67%	51.84%	77.61%	63.64%
PointNet++ [21]	95.65%	80.52%	95.32%	79.42%	95.52%	84.85%

Table 10.4: Vulnerability analysis of 2D and 3D FRS on FaceScape Dataset

Combined			**Male**		**Female**	
Algorithm	MMPMR%	FMMPR%	MMPMR%	FMMPR%	MMPMR%	FMMPR%
2D Vulnerability Analysis						
COTS	100%	99.9%	100%	99.9%	100%	100%
Arcface	100%	100%	100%	100%	100%	100%
3D Vulnerability Analysis						
LED3D [20]	88.8%	88.8%	90.5%	90.5%	84.9%	84.9%
PointNet++ [21]	95.4%	95.4%	94.1%	94.1%	97.5%	97.5%

experience with the morphing problem are only recorded. All measures are implemented with full consideration of the anonymity of participants. We have designed the GUI for the human observer study to benchmark the single image morphing detection in this work.

Figure 10.7 shows the screenshot of the web portal used for the human observer's study. The GUI is designed to display two face images simultaneously, such that one corresponds to the 2D face and another to the 3D face. Then, the human observer is prompted to independently decide these face images as either morph or bona fide. The human observers are provided with an option to rotate the 3D face in different directions to make their decision effectively. Further, the opportunities to zoom in and out of the 3D face model are also provided. We have mainly selected to present both 2D/3D face images for human evaluation simultaneously to check whether the 3D information might help detect the morphing attacks. Due to the time factor, we have used 19 bona fide and 19 morph samples independently from 2D and 3D for the human observer study. Thus, each human observer spent around 20 minutes on average to complete this study. The detailed step-wise instructions on using the web portal are available for every participant beforehand.

The human observer study uses 36 observers with and without face morphing experience. The quantitative results of the human observer study are shown in Figure 10.8. We summarize the human observer's results from the survey as follows:

- The average detection accuracy of human observers for 2D face bona fide samples is 55.83% and 42.5% in a 3D face, respectively. The average detection accuracy of human observers for morphs in 2D is 58.33% and 51.85% in a 3D face. Thus, detection accuracy is similar for bona fide and morph in 2D. However, the detection accuracy in 3D is lower for bona fide when compared with morph.
- The average detection accuracy is similar for observers without morphing experience and basic morphing experience. Human observers with advanced morphing experience have the highest average detection accuracy. The observers without morphing experience perform similarly to observers with basic morphing experience, which can be attributed to the innate human capacity to distinguish between bona fide v/s morphed.
- The survey further validates that generated 3D morphs are challenging to detect from human observations. The average detection accuracy of human observers does not exceed 63.15%, which shows that 2D and 3D morphs developed in this work are high quality and difficult to detect.

The average detection accuracy in a 2D face is higher than that in a 3D face, which can be attributed to the following reasons:

- The fact that 2D morph is more prevalent, and thus observers generally look for specific artifacts in different regions of the face, makes the task relatively easy with a 2D face.
- The aspect of what artifacts to look at in 3D is unclear to the human observers, as they are not trained for this task.
- The quality of generated 3D morphs is high, so human observers find it difficult to distinguish the 3D morphs from the 3D bona fide.

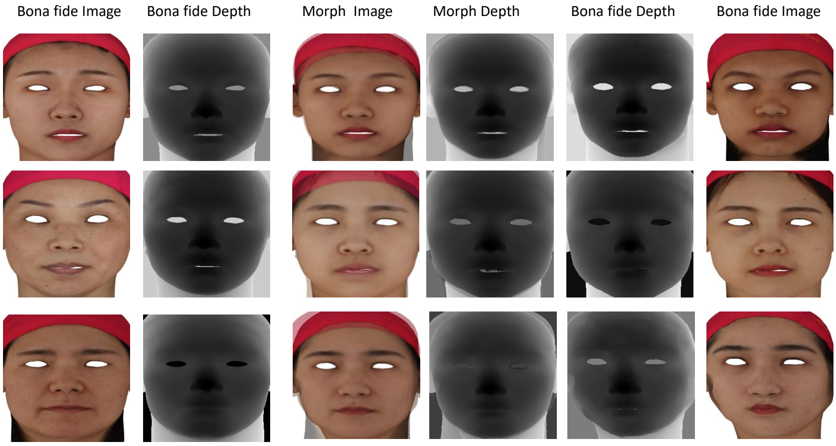


Figure 10.11: Illustration of the Color Images and Depth Maps of Bona fide Samples and Face Morphs generated using the proposed method on Facescape Dataset [23]

10.4.3 Vulnerability Study

In this work, we benchmark the performance of the automatic FRS on both 2D and 3D face models. The 2D face vulnerability is computed using the color image and the 3D face vulnerability is calculated based on depth-map/point cloud. We have used two different metrics to benchmark the vulnerability assessment that, includes Mated Morphed Presentation Match Rate (MMPMR) [83] and Fully Mated Morphed Presentation Match Rate (FMMPMR) [49]. MMPMR can be defined as the percentage of morph samples which can be verified with all the contributing data subjects [49]. However, MMPMR does not consider the number of attempts made during score computation. This is rectified in FMMPMR [49], where the morphing image sample should be verified across all the attempts. The higher value of MMPMR and FMMPMR indicates the higher vulnerability of the FRS. The vulnerability analysis is performed by enrolling the morphing image (2D/3D) and then obtaining the comparison score by probing both contributory data subjects' face images (2D/3D). To compute the vulnerability of 2D face morphing images, we have used two different FRS such as Arcface [22] and a Commercial-off-the-Shelf (COTS) FRS [4]. The 3D face vulnerability analysis uses Deep Learning-based FRS such as Led3D [20] and PointNet++ [21]. The thresholds for all FRS used in this work are set at FAR=0.1% following the guidelines of Frontex for border control [239].

[4] The name of the COTS is not indicated to respect confidentiality

Quantitative vulnerability results on 3D morphing dataset

The results are summarized in Table 10.3, and the vulnerability plots are shown in Figure 10.9. Based on the obtained results, it can be noted that (1) Both 2D and 3D FRS are vulnerable to the generated face morphing attacks (2) Among the 2D FRS, the COTS indicates the highest vulnerability compared to the Arcface FRS. (3) Among the 3D FRS the PointNet++ [21] indicates the highest vulnerability. Thus, the quantitative results of the vulnerability analysis indicate the effectiveness of the generated 3D face morphing attacks.

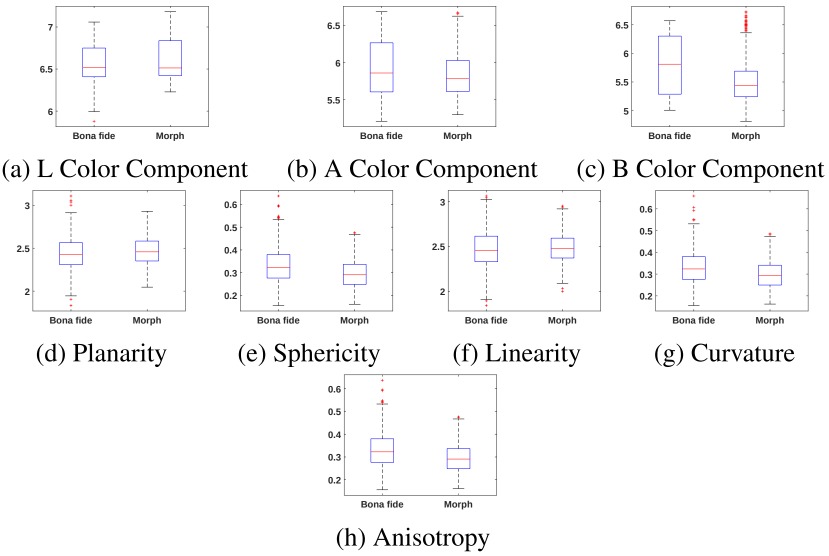
Quantitative vulnerability results on Facescape dataset

We have employed 100 unique databases with 56 male and 44 female data subjects. For each data subject, we have selected two 3D face scans. One is used to generate the 3D face morphing, and another is used as the probe image to obtain the comparison score to compute the vulnerability metrics. We then used the proposed method to get the 3D morphing models, resulting in 2486 morphing models. Figure 10.11 shows the example of the proposed 3D morphing generation samples together with the bona fide 3D scans from Facescape Dataset [23]. The quantitative vulnerability results on the Facescape dataset are indicated in Table 10.4, and the vulnerability plots are shown in Figure 10.10. Here also, it can be noticed that the proposed 3D face morphing generation samples exhibit a high vulnerability with both 2D and 3D FRS. Among 2D FRS, both COTS and Arcface indicate a similar vulnerability with MMPMR = 100%. However, among 3D FRS, PointNet++ [21] shows the highest vulnerability.

Thus, based on the vulnerability analysis reported on 3DMD and Facescape datasets with 2D and 3D FRS, the proposed 3D face morphing technique indicates a consistently high vulnerability. The vulnerability is noted high with the Facescape dataset compared to the 3D morphing dataset. The variation in the vulnerability performance across different FRS can be attributed to the type of feature extraction and classification techniques employed in individual FRS. For example, 2D face recognition systems are based on identity features, whereas 3D-based systems are based on high-resolution depth and shape.

10.4.4 Automatic 3D Face Point Cloud Quality Estimation

In this work, we estimate the visual quality based on the effectiveness of different types of features, including both color and geometry, as proposed in [240]. This study aims to quantitatively estimate the quality of the generated 3D face morphing point clouds and the bona fide 3D face point clouds to quantify the quality of the proposed morphing generation. To this extent, five different point cloud features based on geometry, namely curvature, anisotropy, linearity, planarity, sphericity, and three color information features, namely L color component, A color component, B color component, are computed to benchmark the quality based on the geometry of the generated 3D morphing models.



(a) L Color Component (b) A Color Component (c) B Color Component

(d) Planarity (e) Sphericity (f) Linearity (g) Curvature

(h) Anisotropy

Figure 10.12: Box plots showing the eight different 3D model quality estimation from 3D bona fide and 3D morph based on color and geometry

Table 10.5: Quantitative values of quality features for 3D face point clouds corresponding to 3D bona fide and morph based on color and geometry

3D Face Quality Features (mean ± std. deviation)	**Data type**	
	Bona fide	**Morphed**
L Color	6.5614±0.2191	6.6076±0.2340
A Color	5.9368±0.3547	5.8546±0.3260
B Color	5.7998±0.5074	5.5326±0.4198
Linearity	2.4708±0.2196	2.4911±0.1776
Sphericity	0.3318±0.0807	0.2936±0.0592
Anisotropy	0.3318±0.0807	0.2936±0.0592
Curvature	0.3330±0.0821	0.2965±0.0606
Planarity	2.4430±0.2176	2.4711±0.1733

Figure 10.12 shows the box plot of the eight different quality metrics for both 3D bona fide and 3D morphing point clouds. The quantitative values (mean and standard deviation) of different quality features are also shown in Table 10.5. As noted from Figure 10.9, the quality estimations, mainly based on geometry, indicate the near-complete overlapping for 3D bona fide and 3D morph. Thus, the proposed

3D face morphing generation did not degrade the depth quality. Instead, it has achieved comparable quality based on geometry from bona fide 3D models used for the morphing operation. A similar observation can also be noted with the color image quality estimation.

10.4.5 3D Face Morphing Attack Detection

In this section, we present our proposed method for a single 3D model-based MAD. Because the 3D face morphing is extensively presented in this paper for the first time, there exists no state-of-the-art to detect these attacks. Therefore, we are motivated to develop 3D MAD techniques to detect these attacks reliably. The proposed 3D MAD techniques are based on the pre-trained 3D point-based networks used to extract the features, as shown in Figure 10.13. Thus, given the 3D face point clouds, we first compute the features from the pre-trained network and in the next step, we feed the same to the linear support vector machine to make the final decision on either bona fide or morph. In this work, we have used three different pre-trained point could networks such as Pointnet [21, 241], Pointnet++ [21, 241] and SimpleView [241] independently to benchmark the 3D MAD performance. All three pre-trained CNNs are trained on ModelNet40 dataset [242].

The Pointnet [21, 241] is one of the earliest point-based classifications of deep learning networks invariant to the permutation of 3D vertices. Given the 3D face point clouds, we extract the feature from the classification task layer corresponding to the feature dimension of 4096. The Pointnet++ [21, 241] is the improved version of Pointnet [21, 241] achieved by introducing a hierarchical neural network that was applied recursively. In this work, given the 3D face point clouds, we extract the features from the classification task layer of Pointnet++ to obtain a 40-dimensional feature vector. The SimpleView [241] network is based on projecting the point clouds to multiple view depth maps. In this work, given the 3D face point clouds, we extract the features from the classification task layer of the SimpleView network to obtain a 40-dimensional feature vector.

To effectively benchmark the performance of the proposed 3D MAD, we divide the newly collected dataset into two independent sets, namely training and testing. The training set consists of 3D bona fide and morphing samples from 21 unique data subjects and the testing set consists of 3D samples from 20 unique data subjects. Thus, the training set consists of 168 bona fide and 194 morphed features and the testing set consists of 160 bona fide and 151 morphed features summarized in Table 10.6. Table 10.7 shows the quantitative performance of the proposed 3D MAD techniques. Figure 10.14 shows the performance of individual algorithms in DET. The performance is benchmarked using ISO/IEC metrics [99] defined as Attack Presentation Classification Error Rate (APCER), which is the mis-classification

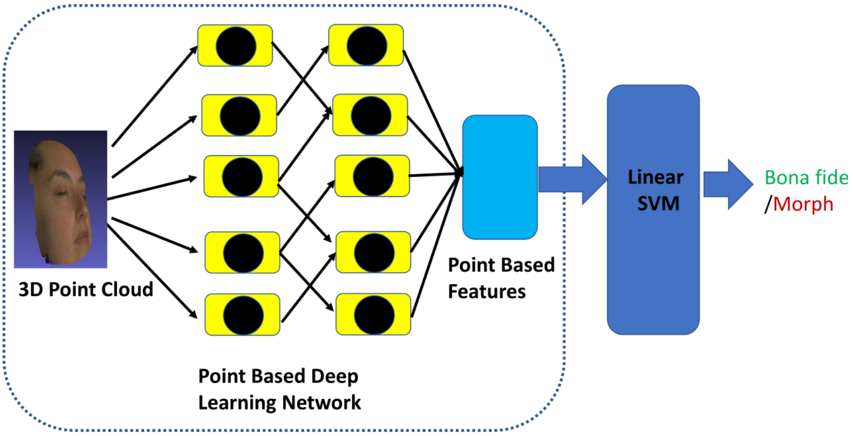


Figure 10.13: Illustration of the proposed 3D face MAD

rate of attack presentations and Bona fide Presentation Classification Error Rate (BPCER) is the mis-classification of bona fide presentation as attacks. Based on the results, the best performance is obtained with the SimpleView [241] network with a D-EER of 1.59%.

Table 10.6: Morphing Attack Detection (S-MAD) Method Protocol

Train Dataset (21 Subjects)	
Bona fide Samples	Morphing Samples
168	194
Test Dataset (20 Subjects)	
Bona fide Samples	Morphing Samples
160	151

Table 10.7: Quantitative performance of the proposed 3D MAD techniques

Algorithm Proposed Method	D-EER (%)	BPCER @ APCER =	
		5%	10%
Pointnet [21]	2.57	3.12	2.5
Pointnet++ [21]	37.33	81.87	68.12
SimpleView [241]	**1.59**	**2.5**	**0**

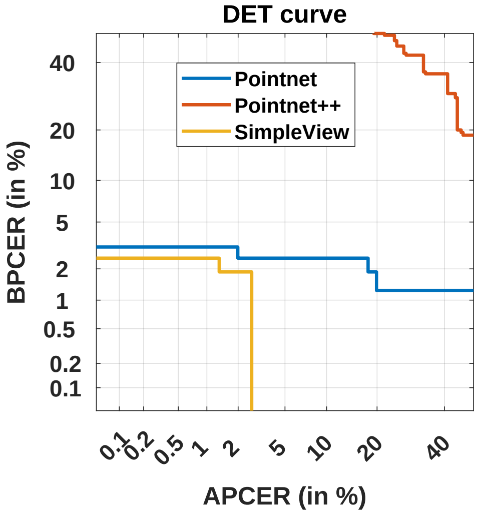


Figure 10.14: DET Curve for the Proposed 3D Morphing Detection methods.

10.5 Discussion

Based on the extensive experiments and obtained results made above, the research questions formulated in Section 10.2 are answered below.

- **RQ#1**. Does the proposed 3D face morphing generation technique yield a high-quality 3D morphed model?
  - Yes, the proposed method of generating the 3D face morphing has resulted in a high-quality morphed model almost similar to that of the original 3D bona fide. The quality analysis reported in Figure 10.12 and Table 10.5 also justifies the quality of the generated 3D morphs quantitatively as the quality values from 3D morphing show larger overlapping with the 3D bona fide. In addition, the human observer analysis reported in Section 10.4.2 also justifies the quality of the proposed 3D face morphing generation method as it is found reasonably difficult to detect based on the artefacts.
- **RQ#2**. Does the generated 3D face morphing model indicate the vulnerability for both automatic 3D FRS and human observers?
  - Yes, based on the analysis reported in Section 10.4.3, the generated 3D face morphing model indicates a high degree of vulnerability for both automatic 3D FRS and human observers.

- **RQ#3**. Are the generated 3D face morphing models more vulnerable when compared to 2D face images for both automatic 3D FRS and human observers?
    - Equally vulnerable, the 3D face morphing models are more vulnerable than their 2D counterparts, as shown in Figure 10.9 when using automatic FRS.
    - However, the vulnerability is almost comparable when evaluated by a human observer study (see Section 10.4.2), where one of the main reasons could be more prevalence of 2D morphs, which makes human observers sensitive about which artifacts to look for.
- **RQ#4**.Can the 3D point cloud information be used to detect the 3D face morphing attacks reliably?
    - Yes, on using the proposed 3D face morphing attack Detection approaches (see Section 10.4.5) the point cloud information can be used for reliable 3D morphing detection.

10.6 Limitations of Current Work and Potential Future Works

Although the work presents a new dimension for face morphing attack generation and detection, especially in 3D, this work has a few limitations. In the current scope of work, the 3D morph generation and detection are carried out on the high-quality 3D scans collected using the Artec Eva sensor. We have employed high-quality 3D face scans to achieve good enrolment quality scans that may reflect the real-life ID enrolment scenario. Thus, future works could investigate the proposed 3D morphing generation and detection techniques using low-quality (depth) 3D scans. Further, extending the study towards in-the-wild capture can also be considered in future work. As a second aspect, the analysis is carried out using 41 data subjects due to the present pandemic outbreak. However, we have also presented the results on the publicly available 3D face dataset, Facescape, with 100 unique IDs. Future work can benchmark the proposed method on large-scale datasets with different 3D resolutions. As a third aspect, cleaning noise from 3D scans is tedious and sometimes requires manual intervention. Thus, future work can develop a fully automated noise removal in 3D point clouds to easily the 3D morph generation.

10.7 Conclusion

This work presented a new dimension for face morphing attack generation and detection, especially in 3D. We have introduced a novel algorithm to generate high-quality 3D face morphing models using point clouds. To validate the attack potential of the newly generated 3D face morphing attacks, the vulnerability analysis uses 2D and 3D FRS. Further, the human observer analysis is also presented to investigate the usefulness of 3D information in morph detection. Obtained results justify the high vulnerability of the proposed 3D face morphing models. We also presented an automatic quality analysis of the generated 3D morphing models that indicate a similar quality as the bona fide 3D scans. Finally, we have proposed three different 3D MAD algorithms to detect the 3D morphing attacks using pre-trained point-based CNN models. Extensive experiments indicate the efficacy of the proposed 3D MAD algorithms in detecting 3D face morphing attacks.

Chapter 11

Article 7: 3D Face Morphing Attack Generation using Non-Rigid Registration (RQ4)

Jag Mohan Singh and Raghavendra Ramachandra. 3D Face Morphing Attack Generation using Non-Rigid Registration. *18th IEEE International Conference on Automatic Face and Gesture Recognition, 2024.*

11.1 Abstract

Face Recognition Systems (FRS) are widely used in commercial environments, such as e-commerce and e-banking, owing to their high accuracy in real-world conditions. However, these systems are vulnerable to facial morphing attacks, which are generated by blending face color images of different subjects. This paper presents a new method for generating 3D face morphs from two bona fide point clouds. The proposed method first selects bona fide point clouds with neutral expressions. The two input point clouds were then registered using a Bayesian Coherent Point Drift (BCPD) without optimization, and the geometry and color of the registered point clouds were averaged to generate a face morphing point cloud. The proposed method generates 388 face-morphing point clouds from 200 bona fide subjects. The effectiveness of the method was demonstrated through extensive vulnerability experiments, achieving a Generalized Morphing Attack Potential (G-MAP) of 97.93%, which is superior to the existing state-of-the-art (SOTA) with a G-MAP of 81.61%.

11.2 Introduction & Related Work

Face Recognition Systems (FRS) have achieved high levels of accuracy in uncontrolled, real-world environments, largely owing to advances in deep learning algorithms, as documented in the literature [27, 22]. This high level of accuracy has led to the adoption of FRS in various commercial settings including e-commerce and e-banking. In particular, facial biometrics are utilized as primary identifiers in passport scenarios to facilitate secure border control and other identification verification applications. Facial biometrics can be captured live or through a passport photo submitted during the application process for identity-document issuance protocols. These biometrics are then stored in an identity document, such as an e-passport, which can be used for verification purposes as needed. Moreover, the implementation of e-passports will facilitate effortless Automatic Border Control (ABC) by eliminating the need for manual intervention through a comparison of the live image captured at the ABC gate with that of the electronic passport.

The accelerating adoption of the FRS technology is accompanied by an increase in vulnerability to various direct and indirect attacks. Among the several types of attacks on FRS, morphing attacks have gained prominence owing to their relevance in high-security applications, such as border control. Morphing involves seamless transformation of multiple face images into a single composite face image that exhibits the geometric and textural features of the original images. Morphing facial images have the ability to deceive both human observers (including experienced border guards) [115] and automatic FRS [69] posing a threat to ID verification and ABC scenarios. In the application process for a passport, an attacker may employ a face morphing image to obtain a legitimate passport, which can subsequently be utilized to enter the country through ABC gates. These factors have motivated researchers to investigate both the generation and detection of face-morphing attacks [69].

The generation of face morphing for 2D images has been extensively explored using both handcrafted (landmarks) and deep learning techniques, such as Generative Adversarial Networks (GANs) and Diffusion Models. However, there has been less research on 3D face morphing due to the challenges of 3D facial key point registration between point clouds. Early work [24] in this area addressed the problem by converting 3D face point cloud into 2D RGB images and depth maps and using landmark-based morphing to generate morphing 2D RGB image and depth map which are back projected for generating morphing face point cloud. The process of generating a 3D face morphing point cloud using point clouds can be described as follows: Given two facial point clouds from two distinct individuals, the objective is to create a facial morphing point cloud that has an average

3D coordinate and color for the corresponding points. To the best of our knowledge, there is no existing work on directly generating 3D face morphing using 3D point clouds. Therefore, this work aims to address this gap by exploring 3D face morphing generation using 3D point clouds.

In the realm of 3D face generation, the critical component is the reliable registration of point clouds, which enables the generation of high-quality 3D morphing that can effectively deceive 3D FRS and achieve maximum attack potential. Although 3D point cloud registration has been extensively studied in the literature on object detection and classification, the challenge of registering non-rigid objects, such as human faces, remains. This is due to the lack of known 3D correspondences between the two human face point clouds and the need to estimate affine transformations for each sub-region of the face. Various point set registration techniques have been proposed, including Reducing Kernel Hilbert Space (RKHS), spline functions, Thin Plate Spline (TPS [243]), correlation-based [244], Gaussian Mixture Models (GMM [245]), Coherent Point Drift (CPD [16]), and Bayesian Coherent Point Drift (BCPD [89]). In this work, we employed BCPD because of its ability to register non-rigid objects such as human faces accurately, robustness against target rotation, and the use of non-Gaussian kernels, which results in greater efficiency than other existing methods, and the algorithm guarantees convergence. The following are the main contributions of this work:

- First work on generating 3D face morphing utilizing point clouds, leveraging the Bayesian Coherent Point Drift (BCPD) method for alignment and averaging the 3D coordinates and color from the given point clouds.
- Extensive analysis on the publicly available 3D face dataset Facescape [23] with 200 unique identities. The attack potential of the proposed 3D morphing generation is evaluated using five different 3D FRS and two different 2D FRS.
- The quantitative values of the attack potential is evaluated is Generalised Morphing Attack Potential (G-MAP) metric and compared with the existing 3D morphing generation techniques.
- The dataset and source code is available for research purpose. Link will be added in the final version.

In the rest of the paper, we present the proposed method in Section 11.3 followed by Dataset details in Section 11.4, experiments and results in Section 11.5 and Section 11.6 discuss the conclusion.

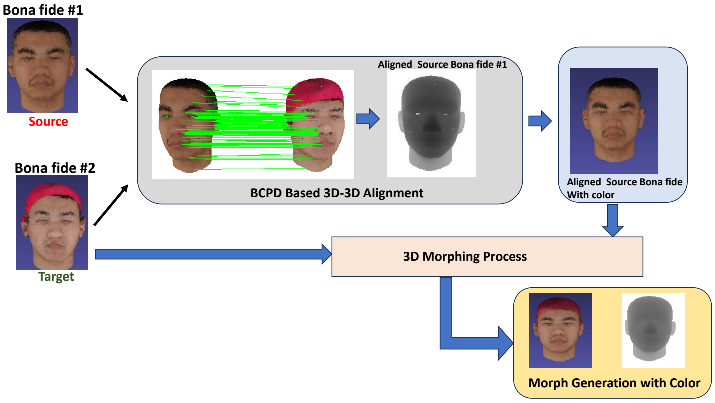

Bona fide #1
Source
Bona fide #2
Target
Aligned Source Bona fide #1
BCPD Based 3D-3D Alignment
Aligned Source Bona fide With color
3D Morphing Process
Morph Generation with Color

Figure 11.1: Illustration showing block diagram of the proposed approach

11.3 Proposed Method

Figure 11.1 shows the block diagram of the proposed method for 3D face morphing generation that can be structured in three different steps: (a) 3D point clouds alignment using Bayesian Coherent Point Drift (BCPD [89]) (b) Colorization of the aligned point clouds (c) 3D morphing point clouds generation. Given two bona fide point clouds that are to be morphed, the proposed method will generate the 3D morphing cloud points as discussed below.

11.3.1 BCPD-based 3D-3D Alignment

We adapted the BPCD algorithm [89] to perform 3D point cloud registration corresponding to two bona fide subjects. We first present the notations that are used to present the adapted BPCD algorithm, and then present the different steps of the 3D point cloud registration.

Notation

- Let Ps_1, Ps_2 denote the two input point clouds where (Ps_1) is considered as *source point cloud* and (Ps_2) is considered as the *target point cloud*.

- Let D denote the dimensionality of data, which in our case is 3 because of 3D point clouds.

- Let the source point cloud be denoted as $Ps_1 = (y_1^T, y_2^T, \cdots, y_M^T) \in \mathbb{R}^3$

- Let the target point cloud be denoted as $Ps_2 = (x_1^T, x_2^T, \cdots, x_N^T) \in \mathbb{R}^3$

- Let Cs_1 be the colors of source point cloud and Cs_2 be the colors of target

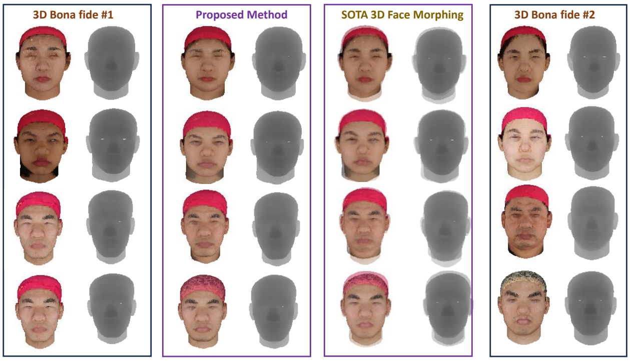


Figure 11.2: Illustration showing Bona fide Input and Morphing Face Samples generated using proposed method and SOTA [24] where SOTA shows blending artifacts in facial boundaries.

point cloud.

- Let the displacement vectors obtained during the non-rigid transformation be denoted as V=$(v_1^T, v_2^T, \cdots, v_M^T) \in \mathbb{R}^3$
- Let the similarity transform be denoted as $\rho = (s, R, T)$
- Let the multivariate normal distribution of z with mean μ and co-variance matrix S be denoted as $\phi(z; \mu, S)$
- Let the non-rigid transformation be denoted as $T(y_M) = sR(y_M + v_M) + t$
- Let $G = g_{mm'} \in \mathbb{R}^{M \times M}$ be the Gram-Matrix with $g_{mm'} = \kappa(y_M, y_M\prime)$ where $\kappa(.,.)$ is a positive-definite kernel.
- Let $P = (p_{mn}) \in [0,1]^{M \times N}$ be the probability matrix where p_{mn} represents the posterior probability that x_n corresponds to y_m.
- Let $\nu = (\nu_1, \nu_2, \cdots, \nu_m)$ denote the estimated number of target points matched with each source points i.e. $(\nu_m = \sum_{n=1}^{N} p_{mn})$
- Let Pst_1 denote the transformed source point cloud.

Initialization

The process of optimizing typically begins with the initialization of various variables. In our case, we followed the BCPD algorithm [89] to initialize the main variables, which include the Rotation Matrix (R), the Translation Vector (t), and

σ^2. The Rotation Matrix was initialized to the identity matrix, the Translation Vector was initialized to zeros, and σ^2 was initialized based on the pairwise Euclidean distance between the source and target point clouds. The specific steps involved in this initialization process are outlined below.

- $\hat{y} = y$, $\hat{v} = 0$, $\Sigma = I_M$, $s = 1$, $R = I_D$, $t = 0$, $\langle \alpha_M \rangle = \frac{1}{M}$ $\sigma^2 = \frac{\gamma}{NMD} \sum_{n=1}^{N} \sum_{m=1}^{M} ||x_n - y_m||^2$, $G = \langle g_{mm\prime} \rangle$ with $\langle g_{mm\prime} \rangle = \kappa(y_m, y_{m\prime})$

Optimization

Repeat the following steps until convergence.

- Firstly, the probability matrix P and its related variables are updated. The update is done based on existing variables. The variables updated in this step include $\langle \phi_{mn} \rangle$, p_{mn}, ν, $\nu\prime$ and $\hat{x}$.

- Next update the following terms, the displacement variable ($\hat{v}$), covariance matrix ($\sum$) and related variables ($\tilde{\nu}$, $\tilde{u}$ and $|\alpha_m|$ for all m) are updated in this step.

- Finally, update the parameters of transformation scaling (s), rotation matrix (R), translation vector (t) and the related variables (σ^2,$\hat{y}$).

Note the details about the updating of variables mentioned in previous step are available in [89] (specifically in Figure 2 of the paper). Once convergence is reached, the vertex-coordinates coordinates in the source point cloud can be transformed by the following equation:

$$Pst_1 = sR(Ps_1 + V) + T \tag{11.1}$$

11.3.2 Colorization of the aligned point clouds

The BCPD method does not transform per-vertex colors. Therefore, the transformation of Rotation, Displacement, and Translation is applied only to the point coordinates of the colored source point cloud generated by BCPD. The target point cloud is kept fixed, and the colors from the aligned point clouds are averaged to generate a face-morphing point cloud. To generate an aligned source point cloud, the coordinates of the source point cloud Ps_1 are transformed using Equation 11.1, resulting in a color-preserved aligned source point cloud Pst_1.

11.3.3 3D Morphing Process

The steps of generating the 3D face morphing point cloud (P_m) given the aligned source point cloud (Pst_1) and the target point cloud (Ps_2) is done by generating the face morphing point cloud (P_m) vertices and colors by Equation 11.2 and Equation 11.3, respectively.

$$P_m = \alpha \times Pst_1 + (1 - \alpha) \times Ps_2 \tag{11.2}$$

$$C_m = \alpha \times Cs_1 + (1 - \alpha) \times Cs_2 \tag{11.3}$$

Figure 11.2 shows the qualitative results of the 3D face morphing generation using proposed and the state-of-the-art [24]. SOTA method generates facial morphing samples in which blending is noticeable at the boundaries whereas the proposed method blending is not visible. The blending is noticeable in depth-maps generated using SOTA apart from color images. The quality of the morphs generated by both methods appears to be similar in interior regions.

11.4 Dataset Details

In this work, we employed the publicly available 3D face dataset, Facescape [23]. Facescape dataset comprised 18,760 textured 3D faces from 938 data subjects captured with 20 different expressions. In this work, we selected 200 data subjects (112 males and 88 females) with neutral and smiling expressions to perform the morphing operation using both the proposed and existing 3D face morphing generation techniques. Morphing between the data subjects was performed by following the guidelines from [239], which include gender- and ethnicity-specific subject selection for morphing. The 3D face point clouds corresponding to neutral expressions were used to generate the morphing generation, and the smiling expression was used to compute the attack potential of the proposed morphing technique. The number of face morphing samples generated were 388 in total which includes 218 male morphs and 170 female morphs.

11.5 Experiments & Results

In this section, we present a quantitative analysis of the attack potential of both the proposed and existing methods for 3D face morphing generation. This analysis was conducted using two different FRS: one based on depth maps using five pre-trained deep CNNs: Resnet-34, Inceptionv3, VGG16, Mobilenetv2 proposed in [246] and the the second FRS based on the depth maps, as proposed in [?]. Additionally, we utilized 2D FRS consisting of ArcFace and MagFace, which evaluates the attack potential of morphing attacks using color-morphed images without depth maps.

In this work, Generalized Morphing Attack Potential (G-MAP) [247] which is a quantitative measure is used to evaluate the attack potential of morphing images. G-MAP metric was designed to address the limitations of other evaluation metrics, as discussed in [247].[1] To compute the G-MAP values, the morphing sample was enrolled in the FRS, and comparison scores were computed by probing the samples of the contributing subjects. If the computed scores exceed the False Acceptance Rate (FAR) threshold, the enrolled sample is considered a successful attack. Therefore, higher values of G-MAP correspond to a higher attack potential for morphing techniques. The G-MAP metric is defined as follows [247]:

$$\begin{aligned} \text{G-MAP} = &\frac{1}{|\mathbb{D}|}\sum_{d}^{|\mathbb{D}|}\frac{1}{|\mathbb{P}|}\frac{1}{|\mathbb{M}_d|}\min_{\mathbb{F}_l} \\ &\sum_{i,j}^{|\mathbb{P}|,|\mathbb{M}_d|}\Bigg\{\left[(S1_i^j > \tau_l) \wedge \cdots (Sk_i^j > \tau_l)\right] \\ &\times\left[(1 - FTAR(i,l))\right]\Bigg\} \end{aligned} \tag{11.4}$$

where, $\mathbb{P}$ is set of probe images, $\mathbb{F}$ is the set of FRS, $\mathbb{D}$ is the set of Morphing Attack Generation Type, $\mathbb{M}_d$ is the face morphing image set for the Morphing Attack Generation Type d, τ_l indicate the similarity score threshold for FRS (l), $FTAR(i,l)$ is the failure to acquire probe image in attempt i using FRS (l), and $||$ is the number of elements in a set.

In this work, we present the quantitative results for G-MAP with multiple probe attempts (G-MAP-MA) calculated from Equation 11.4 by setting D = 1, F = 1, and FTAR = 0. We also present the G-MAP with multiple attempts and multiple FRS (G-MAP-MAMF) by taking the minimum across the FRS with D = 1 in Equation 11.4.

Table 11.1 shows the Generalized Morphing Attack Potential (G-MAP) of multiple attempts using different face recognition systems (FRS) based on the depth maps. Based on the obtained results, it can be noted that (a) the proposed 3D face morphing generation techniques indicate higher values of GMAP, and thus indicate a higher attack potential compared to the existing method [24]. Improved performance of the proposed method was observed with both 3D and 2D FRS. The improved performance can be attributed to the high-quality color and depth maps, which can result in the vulnerability of the FRS. (b) With 3D FRS, the proposed method exhibited the best performance of GMAP = 97.93% with deep CNNs and

[1] A more detailed discussion about the advantages of G-MAP compared to other metrics is provided in [247]

Table 11.1: Quantitative results of Vulnerability of FRS: GMAP-MA@FAR = 0.1%

Algorithm/Features	3D Face Morphing [24]	Proposed Method
3D FRS		
Resnet34 [246]	86.79	97.93
Inceptionv3 [246]	81.61	97.93
VGG16 [246]	96.37	97.93
Mobilenetv2 [246]	97.67	97.93
Led3D [20]	100	100
2D FRS		
Arcface [22]	95.85	100
Magface [248]	87.82	100

100% with Led3D FRS. With 2D FRS, the proposed method exhibits the best performance with GMAP = 100% on both FRS. (c) Overall, there is a slight difference between SOTA and the proposed method, where SOTA shows blending artifacts in facial boundaries compared with the proposed method, which can also be seen in Figure 11.2. This resulted in a lower G-MAP score with SOTA than with the proposed method. SOTA is based on a 3D-2D-3D approach, where blending is performed in 2D and can result in artifacts. We also compute the GMAP-MAMF, which can quantify the attack potential of the generated morphing samples across multiple attempts and the FRS. For 3D FRS, the proposed morphing generation technique indicated a GMAP of 97.93%, whereas the SOTA was 81.61%. For the 2D FRS, the proposed method indicates GMAP = 100%, whereas SOTA is 87.82%. These results justify the higher attack potential of the proposed method compared to the existing method.

11.6 Conclusions & Future-Work

In this paper, we introduced a method for directly registering 3D point clouds to generate a face-morphing point cloud based on BCPD. We evaluated the proposed 3D face morphing attack generation method on a publicly available dataset (Facescape Database) containing 200 unique data subjects. The attack potential of the proposed method was compared to that of the existing method using the G-MAP metric, and the results demonstrated the highest attack potential, as indicated by the quantitative analysis. In the present work, the poses of the subjects were predominantly near-frontal, which simplifies the registration process. Moving forward, we plan to develop a method that can handle arbitrary facial positions and lighting conditions. This approach would be more representative of real-world scenarios because it would enable data capture under a variety of lighting conditions.

Part III

Future Work

Chapter 12

Future Works

The thesis answered the formulated research questions and we concluded. Further, the thesis reached closer to its primary goal of robust MAD and, additionally, the generation of high-quality face morphing images. However, several research directions can be explored in the future, which we are listing below:

12.0.1 Attack Presentation Classification Error Rate (APCER) at low Bona fide Presentation Classification Error Rate (BPCER)

It needs to be pointed out that in the currently published articles we have calculated APCER@BPCER=0.1%. However, one could evaluate APCER@BPCER=0.01%, which minimizes false alarms with low BPCER, and calculate the attack presentations misclassification rate at this threshold which is mentioned in National Institute of Standards Technology (NIST) technical report [87].

12.0.2 GAN generated postprocessing

One future direction is to use the Spectral GAN approach by Dong et al. [249], where the authors have mitigated the spectral artifacts in GAN-generated images. The authors proposed two methods for this where the first is based on CyclicGAN with losses of L_{GAN} (forward and reverse direction), L_{cyc} (cyclic GAN loss), $L_{identity}$ (identity loss) incorporated in Cyclic GAN and the losses proposed by the authors which are L_{power} (power loss which regularizes the range of spectral power distribution) and L_{max} (max loss which ensures that maximum value of spectra remains unchanged during domain transfer). The authors subtract the mean difference between GAN-generated and real-world images in the second method based on the training dataset from the input spectrum. The authors follow both these methods by power distribution correction (PDC), which corrects the power spectrum of an input distribution based on a power distribution. Note that authors evaluate the following methods, Method1, Method1+PDC, Method2, Method2+PDC, and PDC, and used both SVM-Classifier and shallow CNN-based spectrum detector on the datasets of BigGAN, CRN, CycleGAN, IMLE, ProGAN, StarGAN, StyleGAN and StyleGAN2 where mitigation results in a reduction of accuracy % implying reduction of spectral artifacts. Thus, it can be used for automatic postprocessing face morphing images and be a potential future work.

12.0.3 Different Score Fusion Methods

Kumar et al. [250] further present linear and non-linear fusion, where non-linear fusion achieves better accuracy than the weighted sum rule for most experiments they evaluated. Kumar et al. [251] performed weighted minimum, weighted sum, and weighted product fusions where in one dataset, weighted sum achieves the best performance and in another, weighted product achieves the best performance. Kumar et al. [252] perform multibiometric fusion using linear and non-linear combinations of scores where the fusion parameters are found using Particle Swarm Optimization (PSO). The authors mention that the computational cost of PSO is manageable as the search space is low. Thus, checking different score fusion methods, significantly non-linear fusion is a potential future work.

12.0.4 High-Quality Deep-Learning-Based Morph Generation

In this thesis, we focussed on generating morphs through landmarks and CFIA using deep learning. However, increasing morph generation quality using deep-learning-based techniques is an active area of research. It should be noted that this area started with GANs and currently uses diffusion models. Increasing the morph generation quality close to or better than landmark-based morphs could be a fruitful research direction to explore further in subsequent articles.

12.0.5 Improving Generalization Accuracy of MAD

MAD methods perform well on a dataset with similar environments during training and testing. However, their accuracy drops when training and testing environments differ, e.g., training on a digital medium and testing on a print-scan medium. This thesis has worked in this direction and improved the generalization accuracy of MAD methods, but more work is required. The generalization accuracy is vital from a real-world applicability perspective when the test medium/dataset is unseen/unknown. Thus, this could be an exciting research direction.

12.0.6 Composite Attack Detection (CAD) methods

In chapter 9, we introduced CFIA. The CFIA showed vulnerability towards FRS and was challenging for human observers to detect. Further, existing MAD methods have shown low accuracy in detecting them. Thus, developing CAD methods to detect CFIA could be an exciting research direction.

12.0.7 Explainability of MAD methods

In the current literature, MAD methods perform well, given that training and testing are similar. However, little attention has been paid to explaining the decisions of MAD methods, which showcase the facial parts responsible for the decision at a basic level and provide a more detailed explanation at an advanced level. It must be pointed out that only a few works exist in this area [253]. This is important for the real-world applicability of MAD and could be an exciting research direction.

12.0.8 Arbitrary Expression 3D Face Morphing

The 3D face morphing was evaluated for neutral face expression. However, it must be assessed for arbitrary expression for real-world applicability. It must be pointed out that such a use case becomes active when a 3D camera is mounted in an ABC or OTF scenario, unlike the current 2D imaging cameras.

12.0.9 Synthetic Morph Generation

The CFIA method works on synthetic data and generates composites from them. Synthetic data has recently been gaining importance in the Face Recognition community. Synthetic data would allow for large-scale training and improve generalization. Further, high-quality synthetic depth data would be helpful in the generation of 3D face morphing. This is important as acquiring ground-truth bona fide depth maps is time-consuming, and depth sensors are required. Since deep-learning-based methods provide depth data for faces with low accuracy, there is a need to improve the depth quality of deep-learning-based methods as this would result in large-scale applicability of MAD. Further, there is a need to model the effect of print-scan on depth. The achievement of these two goals would result in a much larger-scale application of MAD methods compared to the current scale and enhance their real-world adoption. Finally, generating a synthetic face morph image and having a deep-learning-based transfer to real-world images could be an exciting new research direction.

12.0.10 Anonymity of the Morphing Dataset

The dataset collection for face image morphing is done under GDPR norms. However, the current morphing dataset can be reverse-engineered with some effort, leading to users' identities being revealed. Differential Privacy can guarantee the anonymity of datasets and thus could be helpful in this research direction.

Bibliography

[1] Naser Damer, Viola Boller, Yaza Wainakh, Fadi Boutros, Philipp Terhörst, Andreas Braun, and Arjan Kuijper. Detecting face morphing attacks by analyzing the directed distances of facial landmarks shifts. In *Proc. of the German Conference on Pattern Recognition (GCPR)*, pages 518–534, Stuttgart, Germany, October 2018.

[2] U. Scherhag, C. Rathgeb, and C. Busch. Towards detection of morphed face images in electronic travel documents. *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*, pages 187–192, April 2018.

[3] Lucy Chai, Jonas Wulff, and Phillip Isola. Using latent space regression to analyze and leverage compositionality in gans. In *International Conference on Learning Representations*, pages 1–30, Virtual, 2021.

[4] Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, Luuk Spreeuwers, Raymond Veldhuis, and Christoph Busch. Morphed face detection based on deep color residual noise. In *2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 1–6, Istanbul, Turkey, 2019.

[5] Raghavendra Ramachandra, Sushma Venkatesh, Kiran Raja, and Christoph Busch. Towards making morphing attack detection robust using hybrid scale-space colour texture features. In *2019 IEEE 5th International Conference on Identity, Security, and Behavior Analysis (ISBA)*, pages 1–8, Hyderabad, India, 2019.

[6] 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)*, pages 269–278, Snowmass Village, CO, USA, March 2020.

[7] M. Ferrara, A. Franco, and D. Maltoni. Face demorphing. *IEEE Transactions on Information Forensics and Security*, 13(4):1008–1017, April 2018.

[8] Ulrich Scherhag, Christian Rathgeb, Johannes Merkle, and Christoph Busch. Deep face representations for differential morphing attack detection. *IEEE Transactions on Information Forensics and Security*, 15:3625–3639, 2020.

[9] Jag Mohan Singh, Raghavendra Ramachandra, Kiran B Raja, and Christoph Busch. Robust morph-detection at automated border control gate using deep decomposed 3d shape & diffuse reflectance. In *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pages 106–112, Sorrento (NA), Italy, 2019. IEEE.

[10] U. Scherhag, C. Rathgeb, and C. Busch. Towards detection of morphed face images in electronic travel documents. In *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*, pages 187–192, Vienna, Austria, April 2018.

[11] Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba. Scene parsing through ade20k dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 633–641, Honolulu, Hawaii, USA, 2017.

[12] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4396–4405, Long Beach, CA, USA, June 2019.

[13] Michael Kazhdan and Hugues Hoppe. Screened poisson surface reconstruction. *ACM Trans. Graph.*, 32(3):1–13, jul 2013.

[14] Gaël Guennebaud and Markus Gross. Algebraic point set surfaces. *ACM Trans. Graph.*, 26(3):23–33, July 2007.

[15] A Cengiz Öztireli, Gael Guennebaud, and Markus Gross. Feature preserving point set surfaces based on non-linear kernel regression. In *Computer graphics forum*, volume 28, pages 493–501. Wiley Online Library, 2009.

[16] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence*, 32(12):2262–2275, 2010.

[17] Yiming Zeng, Yue Qian, Zhiyu Zhu, Junhui Hou, Hui Yuan, and Ying He. Cornet3d: unsupervised end-to-end learning of dense correspondence for 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6052–6061, Virtual, 2021.

[18] Volker Blanz and Thomas Vetter. A morphable model for the synthesis of 3d faces. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 187–194, Los Angeles, CA, USA, 1999.

[19] Tianye Li, Timo Bolkart, Michael. J. Black, Hao Li, and Javier Romero. Learning a model of facial shape and expression from 4D scans. In *ACM Transactions on Graphics, (Proc. SIGGRAPH Asia)*, volume 36, pages 194:1–194:17, Bangkok, Thailand, 2017.

[20] Guodong Mu, Di Huang, Guosheng Hu, Jia Sun, and Yunhong Wang. Led3d: A lightweight and efficient deep approach to recognizing low-quality 3d faces. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5773–5782, Long Beach, CA, USA, June 2019.

[21] Charles R. Qi, Li Yi, Hao Su, and Leonidas J. Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 5105–5114, Long Beach, California, USA, 2017. Curran Associates Inc.

[22] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4690–4699, Long Beach, CA, USA, 2019.

[23] Haotian Yang, Hao Zhu, Yanru Wang, Mingkai Huang, Qiu Shen, Ruigang Yang, and Xun Cao. Facescape: a large-scale high quality 3d face dataset and detailed riggable 3d face prediction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 601–610, Virtual, 2020.

[24] Jag Mohan Singh and Raghavendra Ramachandra. 3d face morphing attacks: Generation, vulnerability and detection. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, pages 1–1, 2023.

[25] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC TR 24741 Biometrics – Overview and application*. International Organization for Standardization, 2020.

[26] J. Wayman, A. Jain, D. Maltoni, and D. Maio. An introduction to biometric authentication systems. In J. Wayman, A. Jain, D. Maltoni, and D. Maio, editors, *Biometric Systems: Technology, Design and Performance Evaluation*, pages 1–20. Springer London, London, 2005.

[27] F. Schroff, D. Kalenichenko, and J. Philbin. Facenet: A unified embedding for face recognition and clustering. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 815–823, Boston, MA, USA, 2015.

[28] Omkar M. Parkhi, Andrea Vedaldi, and Andrew Zisserman. Deep face recognition. In Xianghua Xie, Mark W. Jones, and Gary K. L. Tam, editors, *Proceedings of the British Machine Vision Conference (BMVC)*, pages 41.1–41.12, Swansea,UK, September 2015. BMVA Press.

[29] Raghavendra Ramachandra and Christoph Busch. Presentation attack detection methods for face recognition systems: A comprehensive survey. volume 50, pages 1–37. ACM New York, NY, USA, 2017.

[30] Faseela Abdullakutty, Eyad Elyan, and Pamela Johnston. A review of state-of-the-art in face presentation attack detection: From early development to advanced deep learning and multi-modal fusion methods. *Information Fusion*, 75:55–69, 2021.

[31] Pavel Korshunov and Sébastien Marcel. Deepfakes: a new threat to face recognition? assessment and detection. *CoRR*, abs/1812.08685, 2018.

[32] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, MA, USA, October 2007.

[33] Yaojie Liu, Joel Stehouwer, Amin Jourabloo, and Xiaoming Liu. Deep tree learning for zero-shot face anti-spoofing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4680–4689, Long Beach, CA, USA, 2019.

[34] Kashif Shaheed, Piotr Szczuko, Munish Kumar, Imran Qureshi, Qaisar Abbas, and Ihsan Ullah. Deep learning techniques for biometric security: A systematic review of presentation attack detection systems. *Engineering Applications of Artificial Intelligence*, 129:107569, 2024.

[35] Deepfakes code. https://github.com/deepfakes/faceswap, 2019. (Accessed on 08/26/2021).

[36] Marek Kowalski. Faceswap code, 2016. (Accessed on 08/26/2021).

[37] Justus Thies, Michael Zollhofer, Marc Stamminger, Christian Theobalt, and Matthias Nießner. Face2face: Real-time face capture and reenactment of rgb videos. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2387–2395, Las Vegas, NV, USA, 2016.

[38] Justus Thies, Michael Zollhöfer, and Matthias Nießner. Deferred neural rendering: Image synthesis using neural textures. *Acm Transactions on Graphics (TOG)*, 38(4):1–12, 2019.

[39] Yisroel Mirsky and Wenke Lee. The creation and detection of deepfakes: A survey. *ACM Comput. Surv.*, 54(1), January 2021.

[40] Rami Mubarak, Tariq Alsboui, Omar Alshaikh, Isa Inuwa-Dutse, Saad Khan, and Simon Parkinson. A survey on the detection and impacts of deepfakes in visual, audio, and textual formats. *IEEE Access*, 11:144497–144529, 2023.

[41] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(6):681–685, 2001.

[42] Davis E. King. Dlib-ml: A machine learning toolkit. In *Journal of Machine Learning Research*, volume 10, pages 1755–1758, 2009.

[43] Siu-Wing Cheng, Tamal Krishna Dey, Jonathan Shewchuk, and Sartaj Sahni. *Delaunay mesh generation*. CRC Press Boca Raton, 2013.

[44] Image morphing software. https://fotomorph.informer.com/13.9/, 2023 (Acccessed).

[45] Image morphing software version 5. https://fantamorph.com/, 2023 (Acccessed).

[46] Image morphing software. https://learnopencv.com/face-morph-using-opencv-cpp-python/, 2023 (Acccessed).

[47] Face morpher. http://www.facemorpher.com/, 2021. Accessed: September 2021.

[48] Naser Damer, Alexandra Moseguí Saladié, Andreas Braun, and Arjan Kuijper. Morgan: Recognition vulnerability and attack detectability of face

morphing attacks created by generative adversarial network. In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–10, Los Angeles, California, USA, 2018.

[49] 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, Porto, Portugal, 2020. IEEE.

[50] Naser Damer, Kiran Raja, Marius Süßmilch, Sushma Venkatesh, Fadi Boutros, Meiling Fang, Florian Kirchbuchner, Raghavendra Ramachandra, and Arjan Kuijper. Regenmorph: visibly realistic gan generated face morphing attacks by attack re-generation. In *International Symposium on Visual Computing*, pages 251–264, Virtual, 2021. Springer.

[51] 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.

[52] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Virtual, June 2020.

[53] Eklavya Sarkar, Pavel Korshunov, Laurent Colbois, and Sébastien Marcel. Are gan-based morphs threatening face recognition? In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2959–2963, Singapore, 2022.

[54] InsightFace. Insightface: 2d and 3d face analysis project. https://github.com/deepinsight/insightface.git, 2022.

[55] RC Malli. keras-vggface. https://github.com/rcmalli/keras-vggface.git, 2022.

[56] David Sandberg. Facenet tensorflow. https://github.com/davidsandberg/facenet.git, 2022.

[57] Roy Wallace, Mitchell McLaren, Christopher McCool, and Sebastien Marcel. Inter-session variability modelling and joint factor analysis for face authentication. In *2011 International Joint Conference on Biometrics (IJCB)*, pages 1–8, Washington, DC, USA, 2011. IEEE.

[58] Samuel Price, Sobhan Soleymani, and Nasser M. Nasrabadi. Landmark enforcement and style manipulation for generative morphing. In *2022 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–10, 2022.

[59] Naser Damer, Meiling Fang, Patrick Siebke, Jan Niklas Kolf, Marco Huber, and Fadi Boutros. Mordiff: Recognition vulnerability and attack detectability of face morphing attacks created by diffusion autoencoders. In *2023 11th International Workshop on Biometrics and Forensics (IWBF)*, pages 1–6, Barcelona, Spain, 2023. IEEE.

[60] Naser Damer, César Augusto Fontanillo López, Meiling Fang, Noémie Spiller, Minh Vu Pham, and Fadi Boutros. Privacy-friendly synthetic data for the development of face morphing attack detectors. pages 1605–1616, 2022.

[61] Na Zhang, Xudong Liu, Xin Li, and Guo-Jun Qi. Morphganformer: Transformer-based face morphing and de-morphing. *arXiv preprint arXiv:2302.09404*, 2023.

[62] Una M. Kelly, Luuk Spreeuwers, and Raymond Veldhuis. Worst-case morphs: a theoretical and a practical approach. In *2022 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–5, Darmstadt, Germany, 2022.

[63] Aravinda Reddy PN, K Sreenivasa Rao, Raghavendra Ramachandra, et al. Extswap: Leveraging extended latent mapper for generating high quality face swapping. *arXiv preprint arXiv:2310.12736*, 2023.

[64] Laurent Colbois, Hatef Otroshi Shahreza, and Sébastien Marcel. Approximating optimal morphing attacks using template inversion. In *IEEE International Joint Conference on Biometric*, Ljubljana, Slovenia, September 2023.

[65] Qiaoyun He, Zongyong Deng, Zuyuan He, and Qijun Zhao. Optimal-landmark-guided image blending for face morphing attacks. In *IEEE International Joint Conference on Biometrics (IJCB 2023)*, pages 1–9, Ljubljana, Slovenia, 2023. IEEE.

[66] Yen-Cheng Liu et al. 2d + 3d face morphing. *Computer Vision, Graphics, and Image Processing*, 46(1):1–14, 2016.

[67] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–I, Kauai, HI, USA, 2001.

[68] Xiangxin Zhu and Deva Ramanan. Face detection, pose estimation, and landmark localization in the wild. In *2012 IEEE conference on computer vision and pattern recognition*, pages 2879–2886, Providence, Rhode Island, USA, 2012. IEEE.

[69] Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, and Christoph Busch. Face morphing attack generation & detection: A comprehensive survey. *IEEE Transactions on Technology and Society*, pages 128–145, 2021.

[70] Hossein Kashiani, Shoaib Meraj Sami, Sobhan Soleymani, and Nasser M. Nasrabadi. Robust ensemble morph detection with domain generalization. In *2022 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–10, Abu Dhabi, UAE, 2022.

[71] Poorya Aghdaie, Sobhan Soleymani, Nasser M Nasrabadi, and Jeremy Dawson. Attention augmented face morph detection. *IEEE Access*, 11:24281–24298, 2023.

[72] Marija Ivanovska and Vitomir Štruc. Face morphing attack detection with denoising diffusion probabilistic models. In *2023 11th International Workshop on Biometrics and Forensics (IWBF)*, pages 1–6. IEEE, 2023.

[73] Laurine Dargaud, Mathias Ibsen, Juan Tapia, and Christoph Busch. A principal component analysis-based approach for single morphing attack detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 683–692, 2023.

[74] Iurii Medvedev, Farhad Shadmand, and Nuno Gonçalves. Mordeephy: Face morphing detection via fused classification. In Maria De Marsico, Gabriella Sanniti di Baja, and Ana L. N. Fred, editors, *Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods, ICPRAM 2023, Lisbon, Portugal, February 22-24, 2023*, pages 193–204, Lisbon, Portugal, 2023. SCITEPRESS.

[75] Kiran Raja, Gourav Gupta, Sushma Venkatesh, Raghavendra Ramachandra, and Christoph Busch. Towards generalized morphing attack detection by learning residuals. *Image and Vision Computing*, 126:104535, 2022.

[76] Guido Borghi, Gabriele Graffieti, Annalisa Franco, and Davide Maltoni. Incremental training of face morphing detectors. In *2022 26th International Conference on Pattern Recognition (ICPR)*, pages 914–921. IEEE, 2022.

[77] Lorenzo Pellegrini, Guido Borghi, Annalisa Franco, and Davide Maltoni. Detecting morphing attacks via continual incremental training. pages 1–9, Ljubljana, Slovenia, 2023. IEEE.

[78] Eduarda Caldeira, Pedro C. Neto, Tiago Gonçalves, Naser Damer, Ana F. Sequeira, and Jaime S. Cardoso. Unveiling the two-faced truth: Disentangling morphed identities for face morphing detection. In *2023 31st European Signal Processing Conference (EUSIPCO)*, pages 955–959, 2023.

[79] Le Qin, Fei Peng, and Min Long. Face morphing attack detection and localization based on feature-wise supervision. *IEEE Transactions on Information Forensics and Security*, 17:3649–3662, 2022.

[80] Raghavendra Ramachandra, Sushma Venkatesh, Guoqiang Li, and Kiran Raja. Differential newborn face morphing attack detection using wavelet scatter network. pages 1–4, 2023.

[81] Elidona Shiqerukaj, Christian Rathgeb, Johannes Merkle, Pawel Drozdowski, and Benjamin Tams. Fusion of face demorphing and deep face representations for differential morphing attack detection. In *2022 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–5, Darmstadt, Germany, 2022. IEEE.

[82] Nicolò Di Domenico, Guido Borghi, Annalisa Franco, and Davide Maltoni. Combining identity features and artifact analysis for differential morphing attack detection. In *International Conference on Image Analysis and Processing*, pages 100–111. Springer, 2023.

[83] Ulrich Scherhag, Andreas Nautsch, Christian Rathgeb, Marta Gomez-Barrero, Raymond N. J. Veldhuis, Luuk Spreeuwers, Maikel Schils, Davide Maltoni, Patrick Grother, Sebastien Marcel, Ralph Breithaupt, Raghavendra Ramachandra, and Christoph Busch. Biometric systems under morphing attacks: Assessment of morphing techniques and vulnerability reporting. In *2017 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–7, Darmstadt, Germany, 2017.

[84] 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, Houston, TX, USA, 2020.

[85] Matteo Ferrara, Annalisa Franco, Davide Maltoni, and Christoph Busch. Morphing attack potential. In *2022 International Workshop on Biometrics and Forensics (IWBF)*, pages 1–6, Dubai, UAE, 2022. IEEE.

[86] ISO/IEC JTC1 SC37 Biometrics. ISO/IEC CD 20059 - methodologies to evaluate the resistance of biometric recognition systems to morphing attacks. 2023.

[87] NGan. Mei, P. Grother, K. Hanaoka, and J. Kuo. Face Analysis Technology Evaluation (FATE) Part 4: MORPH - Performance of Automated Face Morph Detection. Technical report, National Institute of Standards and Technology, April 2024.

[88] 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)*, Hyderabad, India, January 2019. IEEE.

[89] Osamu Hirose. A bayesian formulation of coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(7):2269–2286, 2021.

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

[91] R.Raghavendra, S. Venkatesh, K. Raja, and C. Busch. Towards making morphing attack detection robust using hybrid scale-space colour texture features. *Proceedings of 5th International Conference on Identity, Security and Behaviour Analysis (ISBA 2019)*, January 2019.

[92] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, July 2002.

[93] J. Kannala and E. Rahtu. BSIF: Binarized statistical image features. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, pages 1363–1366, Tsukuba, Japan, 2012.

[94] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, November 2004.

[95] Soumyadip Sengupta, Angjoo Kanazawa, Carlos D. Castillo, and David W. Jacobs. Sfsnet: Learning shape, reflectance and illuminance of faces 'in the wild'. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6296–6305, Salt Lake City, UT, USA, June 2018.

[96] R. Basri and D. W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):218–233, Feb 2003.

[97] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25, pages 1097–1105, Lake Tahoe, Nevada, USA, 2012. Curran Associates, Inc.

[98] R. Raghavendra, Kiran B. Raja, Sushma Venkatesh, and Christoph Busch. Improved ear verification after surgery - an approach based on collaborative representation of locally competitive features. *Pattern Recognition*, 83:416 – 429, 2018.

[99] ISO/IEC JTC1 SC37 Biometrics. ISO/IEC IS 30107-3. information technology - biometric presentation attack detection - part 3: Testing and reporting. 2023.

[100] Ramachandra Raghavendra and Christoph Busch. Improved face recognition by combining information from multiple cameras in automatic border control system. *2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 1–6, 2015.

[101] Ronen Basri and David Jacobs. *Illumination Modeling for Face Recognition*, pages 89–111. Springer New York, New York, NY, 2005.

[102] Yvain Quéau, Jean-Denis Durou, and Jean-François Aujol. Normal integration: A survey. *Journal of Mathematical Imaging and Vision*, 60(4):576–593, May 2018.

[103] Naser Damer, Steffen Zienert, Yaza Wainakh, Alexandra Moseguí Saladié, Florian Kirchbuchner, and Arjan Kuijper. A multi-detector solution towards an accurate and generalized detection of face morphing attacks. In *2019 22th International Conference on Information Fusion (FUSION)*, pages 1–8, Ottawa, Canada, 2019.

[104] Fei Peng, Le-Bing Zhang, and Min Long. Fd-gan: Face de-morphing generative adversarial network for restoring accomplice's facial image. *IEEE Access*, 7:75122–75131, 2019.

[105] Clemens Seibold, Wojciech Samek, Anna Hilsmann, and Peter Eisert. Accurate and robust neural networks for face morphing attack detection. *Journal of Information Security and Applications*, 53:1–14, 2020.

[106] Sobhan Soleymani, Baaria Chaudhary, Ali Dabouei, Jeremy Dawson, and Nasser M Nasrabadi. Differential morphed face detection using deep siamese networks. In *International Conference on Pattern Recognition*, pages 560–572, Milan, Italy, 2021. Springer.

[107] Sobhan Soleymani, Ali Dabouei, Fariborz Taherkhani, Jeremy Dawson, and Nasser M. 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, Waikoloa, HI, USA, January 2021.

[108] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. Face demorphing. *IEEE Transactions on Information Forensics and Security*, 13(4):1008–1017, 2017.

[109] Ulrich Scherhag, Christian Rathgeb, Johannes Merkle, and Christoph Busch. Deep face representations for differential morphing attack detection. In *IEEE Transactions on Information Forensics and Security*, volume 15, pages 3625–3639. IEEE, 2020.

[110] David Ortega-Delcampo, Cristina Conde, Daniel Palacios-Alonso, and Enrique Cabello. Border control morphing attack detection with a convolutional neural network de-morphing approach. *IEEE Access*, 8:92301–92313, 2020.

[111] Baaria Chaudhary, Poorya Aghdaie, Sobhan Soleymani, Jeremy Dawson, and Nasser 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.

[112] Guido Borghi, Emanuele Pancisi, Matteo Ferrara, and Davide Maltoni. A double siamese framework for differential morphing attack detection. *Sensors*, 21(10), 2021.

[113] Sudipta Banerjee and Arun Ross. Conditional identity disentanglement for differential face morph detection. In *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–8, Shenzhen,China, 2021.

[114] Ilias Batskos, Florens F de Wit, Luuk J Spreeuwers, and Raymond J Veldhuis. Preventing face morphing attacks by using legacy face images. *IET biometrics*, 10(4), 2021.

[115] Sankini Rancha Godage, Frøy Løvåsdal, Sushma Venkatesh, Kiran Raja, Raghavendra Ramachandra, and Christoph Busch. Analyzing human observer ability in morphing attack detection -where do we stand? *IEEE Transactions on Technology and Society*, pages 1–21, 2022.

[116] Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, and Christoph Busch. Face morphing attack generation & detection: A comprehensive survey. volume 2, pages 128–145. IEEE, 2021.

[117] Kiran Raja, Matteo Ferrara, Annalisa Franco, Luuk Spreeuwers, Ilias Batskos, Florens de Wit, Marta Gomez-Barrero, Ulrich Scherhag, Daniel Fischer, Sushma Krupa Venkatesh, et al. Morphing attack detection-database, evaluation platform, and benchmarking. *IEEE transactions on information forensics and security*, 16:4336–4351, 2020.

[118] Mislav Grgic, Kresimir Delac, and Sonja Grgic. Scface–surveillance cameras face database. *Multimedia tools and applications*, 51(3):863–879, 2011.

[119] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. Decoupling texture blending and shape warping in face morphing. In *2019 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–5, Darmstadt, Germany, 2019.

[120] Ken Shoemake. Animating rotation with quaternion curves. In *Proceedings of the 12th annual conference on Computer graphics and interactive techniques*, pages 245–254, San Francisco, CA, USA, 1985.

[121] Weiyang Liu, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. Sphereface: Deep hypersphere embedding for face recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 212–220, Honolulu, Hawaii,USA, 2017.

[122] Samuel R. Buss and Jay P. Fillmore. Spherical averages and applications to spherical splines and interpolation. *ACM Trans. Graph.*, 20(2):95–126, apr 2001.

[123] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255, Miami,Florida, USA, 2009. IEEE.

[124] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[125] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258, Honolulu, Hawaii, USA, 2017.

[126] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[127] 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 *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 269–278, Snowmass Village, CO, USA, 2020.

[128] Andreas Wolf. Portrait quality (reference facial images for mrtd). *Version: 0.06 ICAO, Published by authority of the Secretary General*, 2016.

[129] R. Raghavendra, Kiran B. Raja, Sushma Venkatesh, and Christoph Busch. Face morphing versus face averaging: Vulnerability and detection. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pages 555–563, Denver, Colorado, USA, 2017.

[130] Mei Ngan, Patrick Grother, Kayee Hanaoka, and Jason Kuo. Face recognition vendor test (frvt) part 4: Morph - performance of automated face morph detection. pages 1–45, 2020-03-06 2020.

[131] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC IS 30107-3. Information Technology - Biometric presentation attack detection - Part 3: Testing and Reporting*. International Organization for Standardization, 2023.

[132] 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.

[133] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. Face demorphing. *IEEE Transactions on Information Forensics and Security*, 13(4):1008–1017, 2017.

[134] 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.

[135] Raghavendra Ramachandra and Guoqiang Li. Residual colour scale-space gradients for reference-based face morphing attack detection. In *25th International Conference on Information Fusion*, pages 1–8, Linköping, Sweden, 2022.

[136] Clemens Seibold, Wojciech Samek, Anna Hilsmann, and Peter Eisert. Detection of face morphing attacks by deep learning. In *International Workshop on Digital Watermarking*, pages 107–120, Magdeburg,Germany, 2017. Springer.

[137] Jonathan Richard Shewchuk. Triangle: Engineering a 2d quality mesh generator and delaunay triangulator. In *Workshop on applied computational geometry*, pages 203–222, Philadelphia, PA, USA, 1996. Springer.

[138] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, Las Vegas, Nevada, USA, June 2016.

[139] R. Raghavendra and Christoph Busch. Novel image fusion scheme based on dependency measure for robust multispectral palmprint recognition. *Pattern Recognition*, 47(6):2205–2221, 2014.

[140] A. Jain, P. Flynn, and A. Ross. *Handbook of Biometrics*. Springer, July 2007.

[141] Fatemeh Vakhshiteh, Ahmad Nickabadi, and Raghavendra Ramachandra. Adversarial attacks against face recognition: A comprehensive study. *IEEE Access*, 9:92735–92756, 2021.

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

[143] International Civil Aviation Organization. Machine readable passports – part 12 – public key infrastructure for MRTDs. http://www.icao.int/publications/Documents/9303_p12_cons_en.pdf, 2015.

[144] 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, 2021. Last accessed: 2021-11-23.

[145] Tarmo Kalvet, Henrik Karlzén, Amund Hunstad, and Marek Tiits. Live enrollment for identity documents in europe: The cases of sweden, norway, kosovo, and estonia. *JeDEM - eJournal of eDemocracy and Open Government*, pages 53–73, 2018.

[146] OCI services, india. https://ociservices.gov.in/Photo-Spec-FINAL.pdf. Accessed: May 2020.

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

[148] GOV.UK. https://www.gov.uk/photos-for-passports. Accessed: May 2020.

[149] 3dthis face morph. https://3dthis.com/morph.htm, 2021. Accessed: September 2021.

[150] Morph thing. https://www.morphthing.com/, 2021. Accessed: September 2021.

[151] Robin SS Kramer, Michael O Mireku, Tessa R Flack, and Kay L Ritchie. Face morphing attacks: Investigating detection with humans and computers. *Cognitive research: principles and implications*, 4(1):1–15, 2019.

[152] D. Robertson, R. Kramer, and A. Burton. Fraudulent ID using face morphs: Experiments on human and automatic recognition. *Plos One*, pages 1–12, March 2017.

[153] Baaria Chaudhary, Poorya Aghdaie, Sobhan Soleymani, Jeremy Dawson, and Nasser 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, Virtual, June 2021.

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

[155] M. Ferrara, A. Franco, and D. Maltoni. Face demorphing. *IEEE Trans. on Information Forensics and Security*, 13(4):1008–1017, April 2018.

[156] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. Face demorphing in the presence of facial appearance variations. In *2018 26th European Signal*

Processing Conference (EUSIPCO), pages 2365–2369, Rome,Italy, 2018. IEEE.

[157] L. Spreeuwers, M. Schils, and R. Veldhuis. Towards robust evaluation of face morphing detection. In *Proc. of the 26th European Signal Processing Conf. (EUSIPCO)*, pages 1027–1031, Rome,Italy, 2018.

[158] Jag Mohan Singh and Raghavendra Ramachandra. Reliable face morphing attack detection in on-the-fly border control scenario with variation in image resolution and capture distance. In *2022 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–10, Abu Dhabi, UAE, 2022. IEEE.

[159] Passport online. https://www.dfa.ie/passportonline/, 2021. Accessed: September 2021.

[160] R. Raghavendra, Kiran B. Raja, and Christoph Busch. Detecting morphed face images. In *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–7, Niagara Falls, NY, USA, 2016.

[161] 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)*, pages 1–5, A Coruña, Spain, September 2019. IEEE.

[162] Andrey Makrushin, Christian Kraetzer, Tom Neubert, and Jana Dittmann. Generalized benford's law for blind detection of morphed face images. In *Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security*, IH&MMSec '18, page 49–54, Innsbruck, Austria, 2018. Association for Computing Machinery.

[163] Poorya Aghdaie, Baaria Chaudhary, Sobhan Soleymani, Jeremy Dawson, and Nasser M. Nasrabadi. Detection of morphed face images using discriminative wavelet sub-bands. pages 1–6, 2021.

[164] 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)*, pages 1–6, Istanbul, Turkey, November 2019. IEEE.

[165] 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, Coventry, UK, April 2017. IEEE.

[166] C. Seibold, A. Hilsmann, and P. Eisert. Reflection analysis for face morphing attack detection. In *Proc. of the 26th European Signal Processing Conf. (EUSIPCO)*, pages 1022–1026, Rome,Italy, 2018.

[167] 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)*, pages 302–317, 2019.

[168] 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)*, pages 1–11, Darmstadt, Germany, September 2019. IEEE.

[169] 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, Honolulu, Hawaii, USA, 2017.

[170] Ferrara. Matteo, Franco. Annalisa, and Maltoni. Davide. Face morphing detection in the presence of printing/scanning and heterogeneous image sources. *IET Biometrics*, 10:290–303(13), May 2021.

[171] Sushma Venkatesh, Ramachandra Raghavendra, KSBAran Raja, Luuk Spreeuwers, Raymond Veldhuis, and Christoph Busch. Detecting morphed face attacks using residual noise from deep multi-scale context aggregation network. In *The IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 269–278, Snowmass Village, CO, USA, March 2020.

[172] Naser Damer, Noémie Spiller, Meiling Fang, Fadi Boutros, Florian Kirchbuchner, and Arjan 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 - 16th International Symposium, ISVC 2021, Virtual Event, October 4-6, 2021, Proceedings, Part I*, volume 13017 of *Lecture Notes in Computer Science*, pages 291–304. Springer, 2021.

[173] Poorya Aghdaie, Baaria Chaudhary, Sobhan Soleymani, Jeremy Dawson, and Nasser M. Nasrabadi. Morph detection enhanced by structured group sparsity. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops*, pages 311–320, Waikoloa, HI, USA, January 2022.

[174] Kelsey O'Haire, Sobhan Soleymani, Baaria Chaudhary, Poorya Aghdaie, Jeremy Dawson, and Nasser M Nasrabadi. Adversarially perturbed wavelet-based morphed face generation. In *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*, pages 01–05, Jodhpur, India, 2021. IEEE.

[175] R. Raghavendra, S. Venkatesh, K. Raja, and C. Busch. Detecting face morphing attacks with collaborative representation of steerable scale-space features. In *Intl. Conf. on Computer Vision and Image Processing (CVIP)*, pages 255–265, Jabalpur, India, September 2018.

[176] Kiran Raja, Matteo Ferrara, Annalisa Franco, Luuk Spreeuwers, Ilias Batskos, Florens de Wit, Marta Gomez-Barrero, Ulrich Scherhag, Daniel Fischer, Sushma Krupa Venkatesh, et al. Morphing attack detection-database, evaluation platform, and benchmarking. *IEEE transactions on information forensics and security*, 16:4336–4351, 2020.

[177] M. Ferrara, A. Franco, and D. Maltoni. *Face Recognition Across the Imaging Spectrum*, chapter On the Effects of Image Alterations on Face Recognition Accuracy. Springer, 2016.

[178] International Civil Aviation Organization. Machine readable passports – part 1 – introduction. http://www.icao.int/publications/Documents/9303_p1_cons_en.pdf, 2015. Last accessed: 2015-11-23.

[179] Intl. 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. Last accessed: 2015-11-23.

[180] Adobe photoshop. https://www.adobe.com/no/products/photoshop.html, 2021. Accessed: December 2021.

[181] P. Burt and E. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540, 1983.

[182] V.N. Vapnik. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10(5):988–999, 1999.

[183] Deng Cai, Xiaofei He, and Jiawei Han. Speed up kernel discriminant analysis. *The VLDB Journal—The International Journal on Very Large Data Bases*, 20(1):21–33, 2011.

[184] Lei Zhang, Meng Yang, and Xiangchu Feng. Sparse representation or collaborative representation: Which helps face recognition? In *2011 International Conference on Computer Vision (ICCV)*, pages 471–478, Barcelona, Spain, 2011.

[185] R. Raghavendra and C. Busch. Novel image fusion scheme based on dependency measure for robust multispectral palmprint recognition. *Pattern Recognition*, 47(6):2205–2221, 2014.

[186] NGan. 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.

[187] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 30107-3. Information Technology - Biometric presentation attack detection - Part 3: Testing and Reporting*. International Organization for Standardization, 2023.

[188] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

[189] Ying Xu, Kiran Raja, Raghavendra Ramachandra, and Christoph Busch. Adversarial attacks on face recognition systems. In *Handbook of Digital Face Manipulation and Detection*, pages 139–161. Springer, Cham, 2022.

[190] Fatemeh Vakhshiteh, Ahmad Nickabadi, and Raghavendra Ramachandra. Adversarial attacks against face recognition: A comprehensive study. *IEEE Access*, 9:92735–92756, 2021.

[191] Naveed Akhtar and Ajmal Mian. Threat of adversarial attacks on deep learning in computer vision: A survey. *IEEE Access*, 6:14410–14430, 2018.

[192] Chris Burt. Morphing attack detection for face biometric spoofs needs more generalization, datasets. `https://bit.ly/3lsJ1K8`, 2022.

[193] Quek, Alyssa. Face Morpher. `https://github.com/alyssaq/face_morpher`, 2018. [Online; accessed 19-January-2022].

[194] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. Face demorphing. *IEEE Transactions on Information Forensics and Security*, 13(4):1008–1017, 2017.

[195] Le Qin, Fei Peng, Sushma Venkatesh, Raghavendra Ramachandra, Min Long, and Christoph 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, 2021.

[196] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. In *ACM SIGGRAPH 2003 Papers*, SIGGRAPH '03, page 313–318, San Diego, California, 2003. Association for Computing Machinery.

[197] Neurotechnology. Neurotech Verilook SDK (11.1). https://www.neurotechnology.com/verilook.html, 2019. [Online; accessed 19-January-2022].

[198] Cognitec. Cognitec Face VACS (9.6). https://www.cognitec.com/facevacs-technology.html, 2019. [Online; accessed 19-January-2022].

[199] Tete Xiao, Yingcheng Liu, Bolei Zhou, Yuning Jiang, and Jian Sun. Unified perceptual parsing for scene understanding. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 418–434, Munich, Germany, 2018.

[200] Yujun Shen, Ceyuan Yang, Xiaoou Tang, and Bolei Zhou. Interfacegan: Interpreting the disentangled face representation learned by gans. *IEEE transactions on pattern analysis and machine intelligence*, 44(4):2004–2018, 2020.

[201] Eklavya Sarkar, Pavel Korshunov, Laurent Colbois, and Sébastien Marcel. Vulnerability analysis of face morphing attacks from landmarks and generative adversarial networks. *arXiv preprint arXiv:2012.05344*, 2020.

[202] Naser Damer, César Augusto Fontanillo López, Meiling Fang, Noémie Spiller, Minh Vu Pham, and Fadi Boutros. Privacy-friendly synthetic data for the development of face morphing attack detectors. *arXiv preprint arXiv:2203.06691*, 2022.

[203] Irving Meng. Magface. https://github.com/IrvingMeng/MagFace.git, 2023.

[204] NTNU Jag Mohan Singh. Generalized morphing attack potential. https://github.com/jagmohaniiit/LatentCompositionCode, 2022.

[205] Alain Horé and Djemel Ziou. Image quality metrics: Psnr vs. ssim. In *2010 20th International Conference on Pattern Recognition*, pages 2366–2369, Istanbul, Turkey, 2010.

[206] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. Face demorphing. *IEEE Transactions on Information Forensics and Security*, 13(4):1008–1017, 2017.

[207] 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.

[208] Ramachandra Raghavendra, KiranB Raja, Sushma Venkatesh, and Christoph Busch. Face morphing versus face averaging: Vulnerability and detection. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pages 555–563, Denver, CO, USA, 2017. IEEE.

[209] Aya Al Deeb. Uae reviews features of new id card, 3d photo included. https://www.gulftoday.ae/news/2021/08/05/uae-reviews-features-of-new-id-card-3d-photo-included, 2020. [Online; accessed 16-October-2021].

[210] IDEMIA. Stereo laser image. https://www.idemia.com/wp-content/uploads/2021/02/stereo-laser-image-idemia-brochure-202007.pdf, 2020. [Online; accessed 18-October-2021].

[211] Jan Willem 'JW' ter Hennepe. 3d photo id. https://www.icao.int/Meetings/AMC/MRTD-SEMINAR-2010-AFRICA/Documentation/11_Morpho-3DPhotoID.pdf, 2010. [Online; accessed 16-October-2021].

[212] 3D Face Based ABC Systems. 3d face enrolment for id cards. http://cubox.aero/cubox/php/en_product01-2.php?product=1/, 2021. [Online; accessed 18-October-2021].

[213] S. Dent. Using a 3d render as a french id card 'photo'. https://engt.co/3EiPnQv, 2017. [Online; accessed 16-October-2021].

[214] ICAO. Machine readable travel documents. part 11: Security mechanisms for mrtds. technical report doc 9303. 2021.

[215] ISO/IEC JTC1 SC37 Biometrics. ISO/IEC 39794-5:2019 information technology — extensible biometric data interchange formats — part 5: Face image data. 2019.

[216] Wikipedia contributors. Apple Face ID. https://en.wikipedia.org/wiki/Face_ID, 2017. [Online; accessed 12-December-2021].

[217] Sanjeet Prasad Vardam. Vulnerability of 3d face recognition systems of morphing attacks, August 2021.

[218] Bernhard Egger, William AP Smith, Ayush Tewari, Stefanie Wuhrer, Michael Zollhoefer, Thabo Beeler, Florian Bernard, Timo Bolkart, Adam Kortylewski, Sami Romdhani, et al. 3d morphable face models—past, present, and future. *ACM Transactions on Graphics (TOG)*, 39(5):1–38, 2020.

[219] Yuxin Yao, Bailin Deng, Weiwei Xu, and Juyong Zhang. Quasi-newton solver for robust non-rigid registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Virtual, June 2020.

[220] Hao Li, Robert W Sumner, and Mark Pauly. Global correspondence optimization for non-rigid registration of depth scans. In *Computer graphics forum*, volume 27, pages 1421–1430. Wiley Online Library, 2008.

[221] N. Gelfand, N. J. Mitra, L. J. Guibas, and H. Pottmann. Robust global registration. In *Symposium on Geometry Processing*, pages 197–206, Vienna, Austria, 2005.

[222] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, Feb 1992.

[223] Bailin Deng, Yuxin Yao, Roberto M Dyke, and Juyong Zhang. A survey of non-rigid 3d registration. *arXiv preprint arXiv:2203.07858*, 2022.

[224] Miao Liao, Qing Zhang, Huamin Wang, Ruigang Yang, and Minglun Gong. Modeling deformable objects from a single depth camera. In *2009 IEEE 12th International Conference on Computer Vision*, pages 167–174, Kyoto, Japan, 2009.

[225] Jingyu Yang, Daoliang Guo, Kun Li, Zhenchao Wu, and Yu-Kun Lai. Global 3d non-rigid registration of deformable objects using a single rgb-d camera. *IEEE Transactions on Image Processing*, 28(10):4746–4761, 2019.

[226] Hao Li, Robert W Sumner, and Mark Pauly. Global correspondence optimization for non-rigid registration of depth scans. In *Computer graphics forum*, volume 27, pages 1421–1430. Wiley Online Library, 2008.

[227] Konstantinos Zampogiannis, Cornelia Fermüller, and Yiannis Aloimonos. Topology-aware non-rigid point cloud registration. *CoRR*, abs/1811.07014, 2018.

[228] Giovanni Trappolini, Luca Cosmo, Luca Moschella, Riccardo Marin, Simone Melzi, and Emanuele Rodolà. Shape registration in the time of transformers. *Advances in Neural Information Processing Systems*, 34:5731–5744, 2021.

[229] Artec eva sensor. https://bit.ly/3BiGnJ1, 2021. [Online; accessed 16-October-2021].

[230] Paolo Cignoni, Marco Callieri, Massimiliano Corsini, Matteo Dellepiane, Fabio Ganovelli, and Guido Ranzuglia. MeshLab: an Open-Source Mesh Processing Tool. In Vittorio Scarano, Rosario De Chiara, and Ugo Erra, editors, *Eurographics Italian Chapter Conference*, pages 129–136. The Eurographics Association, 2008.

[231] Bernd Gärtner. Fast and robust smallest enclosing balls. In *European symposium on algorithms*, pages 325–338, Prague, Czech Republic, 1999. Springer.

[232] Dirk Haehnel, Sebastian Thrun, and Wolfram Burgard. An extension of the icp algorithm for modeling nonrigid objects with mobile robots. In *IJCAI*, volume 3, pages 915–920, Acapulco, Mexico, 2003.

[233] Alexandru Telea. An image inpainting technique based on the fast marching method. *Journal of graphics tools*, 9(1):23–34, 2004.

[234] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In *2011 International Conference on Computer Vision*, pages 2564–2571, Barcelona, Spain, 2011.

[235] James D Foley, Andries Van Dam, Steven K Feiner, John F Hughes, and Richard L Phillips. *Introduction to computer graphics*, volume 55. Addison-Wesley Reading, 1994.

[236] 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)*, volume 1, pages 947–954 vol. 1, San Deigo, CA, USA, 2005.

[237] Lijun Yin, Xiaozhou Wei, Yi Sun, Jun Wang, and Matthew J Rosato. A 3d facial expression database for facial behavior research. In *7th international conference on automatic face and gesture recognition (FGR06)*, pages 211–216, Southampton, UK, 2006. IEEE.

[238] Thomas Maurer, David Guigonis, Igor Maslov, Bastien Pesenti, Alexei Tsaregorodtsev, David West, and Gerard Medioni. Performance of geometrix activeid^ TM 3d face recognition engine on the frgc data. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*, pages 154–154, San Deigo, CA, USA, 2005. IEEE.

[239] Frontex. *Best Practice Technical Guidelines for Automated Border Control (ABC) Systems*. Frontex, 2015.

[240] Zicheng Zhang. No-reference quality assessment for 3d colored point cloud and mesh models. *arXiv preprint arXiv:2107.02041*, 2021.

[241] Ankit Goyal, Hei Law, Bowei Liu, Alejandro Newell, and Jia Deng. Revisiting point cloud shape classification with a simple and effective baseline. In *International Conference on Machine Learning*, Vienna, Austria, 2021.

[242] Zhirong Wu, Shuran Song, Aditya Khosla, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets for 2.5d object recognition and next-best-view prediction. *CoRR*, abs/1406.5670, 2014.

[243] Haili Chui and Anand Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3):114–141, 2003.

[244] Yanghai Tsin and Takeo Kanade. A correlation-based approach to robust point set registration. In *Computer Vision-ECCV 2004: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part III 8*, pages 558–569. Springer, 2004.

[245] Bing Jian and Baba C Vemuri. Robust point set registration using gaussian mixture models. *IEEE transactions on pattern analysis and machine intelligence*, 33(8):1633–1645, 2010.

[246] Changyuan Jiang, Shisong Lin, Wei Chen, Feng Liu, and Linlin Shen. Pointface: Point set based feature learning for 3d face recognition. In *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–8, Shenzhen, China, 2021. IEEE.

[247] Jag Mohan Singh and Raghavendra Ramachandra. Deep composite face image attacks: Generation, vulnerability and detection. *IEEE Access*, 11:76468–76485, 2023.

[248] Qiang Meng, Shichao Zhao, Zhida Huang, and Feng Zhou. Magface: A universal representation for face recognition and quality assessment. In

Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 14225–14234, Nashville, Tennessee ,USA, 2021.

[249] Chengdong Dong, Ajay Kumar, and Eryun Liu. Think twice before detecting gan-generated fake images from their spectral domain imprints. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7855–7864, 2022.

[250] Ajay Kumar and Yingbo Zhou. Human identification using finger images. *IEEE Transactions on image processing*, 21(4):2228–2244, 2011.

[251] Ajay Kumar and Chenye Wu. Automated human identification using ear imaging. *Pattern Recognition*, 45(3):956–968, 2012.

[252] Ajay Kumar, Vivek Kanhangad, and David Zhang. A new framework for adaptive multimodal biometrics management. *IEEE transactions on Information Forensics and Security*, 5(1):92–102, 2010.

[253] Clemens Seibold, Anna Hilsmann, and Peter Eisert. Focused lrp: Explainable ai for face morphing attack detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 88–96, Waikoloa, HI, USA, 2021.

ISBN 978-82-326-7996-6 (trykt utg.)
ISBN 978-82-326-7995-9 (elektr. utg.)
ISSN 1503-8181 (trykt utg.)
ISSN 2703-8084 (online ver.)

NTNU
Kunnskap for en bedre verden