Face Morphing Attacks: What needs do be done

C. Busch, S. Caillebotte, U. Seidel, F. Knopjes, D. Maltoni, M. Ferrara, R. Veldhuis, L. Spreeuwers, K. Raja, R. Raghavendra, M. Gomez-Barrero, C. Rathgeb

Abstract: The intention of this paper is to summarize, what countermeasures are needed to mitigate the threat of electronic passports with morphed images. This paper indicates, what Frontex and the concerned national governments can and should do as short term and as long term countermeasures. The suggestions presented in this contribution are based on the discussions of the SOTAMD and iMARS project consortia which proposed a harmonized European approach to tackle the morphing attack threat with joint forces from industry, academia and governmental agencies.

Keywords: face recognition; face morphing attacks; vulnerability analysis; border control

Introduction

The problem of morphing attacks has been addressed in the biometric research community only recently, despite it was already identified back in 2004 in the presentation by Matthew Lewis and Philip Statham at the Biometrics Consortium Conference (BCC). Five years later in 2009 the morphing attack was classified as *vulnerability* of a biometric system in ISO/IEC FDIS 19792 stating: *"… a synthesised characteristic could be injected electrically during a replay attack or planted in the reference database. - feature sets comprising amalgamations of biometric features from 2 or more individuals, e.g. morphed facial images"*. However it took until 2014, before the feasibility of face morphing attacks was first demonstrated in the FIDELITY project by Matteo Ferrara, Annalisa Franco and Davide Maltoni and published in their IJCB paper "*The magic passport*" [Fer14]. Only then researchers started to investigate countermeasures for the problem [Ram16][Sch19]. In 2017 the iMARS consortium[1] was formed with the joint research resources from industry, academia and governmental agencies and is seeking now support from the EU H2020 research program. In 2019 the Dutch National Office for Identity Data and the German Bundeskriminalamt were tasked by European Commission DG Home, to investigate the State-Of-the-Art of Morphing Detection (SOTAMD)[2] by collecting an initial morphing test dataset and by evaluating currently available academic morphing attack detection solutions.

In order to maintain the control on migration of third country nationals, refugees and asylum seekers with the established procedures, Europe should immediately start an action to secure the trusted link between a MRTD and the document holder and to develop and deploy technical mechanisms that can detect a morph passport at borders. This paper describes the necessary steps that should be taken to protect European borders against the threat of morphing attacks. The following chapters describe how Frontex and government agencies of European Member States can support this process.

[1] The iMARS consortium consists of Idemia, NTNU, University Bologna, University Twente, Hochschule Darmstadt, University Leuven, Dutch National Office for Identity Data, German Bundeskriminalamt, Vision-Box, Cognitec, IBS, EAB and various end users (border control agencies)

[2] SOTAMD partners are Dutch National Office for Identity Data, German Bundeskriminalamt, University of Bologna, University of Twente, NTNU and Hochschule Darmstadt

Needs to re-establish a Trusted Link

Unfortunately, in many ICAO Members States and most European Member States, the facial image used for an electronic travel document is provided by the applicant in printed or digital form and not taken by means of 'live-enrolment' in an controlled environment such as a municipality office. Moreover, some countries even operate smartphone-based enrolments, such as the application process for the passport card in Ireland. This fundamental weakness must be stopped **immediately** and a European regulation should enforce that all Member States switch to live enrolment, as it is already operational e.g. in Norway and Sweden. Only then, with full control of the biometric capture process by a civil servant in the passport application office, trust in the link of passport holder to reference data can be assured. The iMARS consortium has proposed to define a secure ID Document application process, which is robust against enrolment fraud such that it will be made more difficult to apply for an ID document with a photograph that has been morphed or manipulated otherwise (e.g. data subjects that want to look younger or more beautiful) in order to influence the biometric recognition process, or by presenting a fraudulent document (in the case of first-time issuance or renewal). Citizens living abroad require a specific use case: the only feasible process for an EU citizen, living abroad, far away from an embassy or a consulate, where she/he can apply for an ID document, could be an application (i.e. passport renewal) from home, which would require specific precautions to prevent enrolment fraud. On the other side of the spectrum, even the seemingly secure live enrolment at a passport office requires precautions to detect a case that someone tries to enrol with a well-crafted facemask (i.e., conducting a presentation attack with a morphed face image on the mask).

The iMARS consortium proposed to define:

- Technical specifications for serving those use cases. Such specifications can be used in a new European regulation on passport application.

- The specifications could also include solutions that secure a wide range of application processes against enrolment fraud (e.g., live-enrolment with kiosk).

- Requirements for Presentation Attack Detection (PAD), to avoid for instance that a silicon mask is used against a face capture device in a live enrolment process.

Moreover, the regulation should define that for facial reference images, which are stored in the ICAO 9303 Logical Data Structure (LDS), the capture device certification scheme will be recorded in the data interchange format, as defined in the new extensible interchange format ISO/IEC 39794 [ISO39794]. This way, the future receiver of the facial reference image can have assurance that the image was captured with live enrolment and thus can be considered trustworthy.

As the passport application process in non-European states cannot be regulated, Europe should through its stakeholders initiate the discussion process, to suggest in the upcoming revision of ICAO 9303 [ICAO9303] a secondary mandatory biometric identifier (iris or fingerprint reference images). Note that ICAO 9303 already allows in Data Group 3 the storage of finger images and in Data Group 4 the storage of iris images. In fact, the USA have recently decided to include an iris image into future passports. A future European border control point could then in case of suspicion regarding a potentially morphed face image rely on fingerprint or iris recognition for traveler identification beyond doubt. For iris, this comes with the advantage that face and iris probe images could be captured with one single capture device.

Further activities of European stakeholders to initiate a new EU regulation are needed.

Need to detect automatically Morph Passports at Borders

Given the validity period of electronic passports, after the future EU-wide transition to live enrolment European border management must anticipate that European passports are presented at least for the next 10 years[3] potentially containing morphed image; as well as passports of third-country nationals beyond the reach of the EU. One of the main goals of the iMARS consortium is to propose efficient solutions for border control points to detect ID documents containing manipulated/morphed images. The Morphing Attack Detection (MAD) solution is expected to enable efficient and reliable automatic data authenticity checks and elevate the process and security of biometric technology to a level that allows operational deployment. For deployed and potentially fraudulent passports, the MAD solutions suggested by iMARS will analyse those potentially manipulated documents. iMARS will provide solutions for the border control process on the one hand based on a differential analysis, where the images stored on the ID document are compared with a trusted live image of the ID document holder, while the capture process is run under supervision.

The iMARS consortium suggests for the processes at European border the development of explicit and implicit image pair detection algorithms (differential MAD – see Figure 1): iMARS explicit image pair based detection algorithms will use image pairs (morphed and bona fide) to setup various models for differential MAD. Regarding Implicit Image Pair detection algorithms, the use of the Deep Neural Network (DNN) approach, along with other methods, will be used to reach the expected progress.

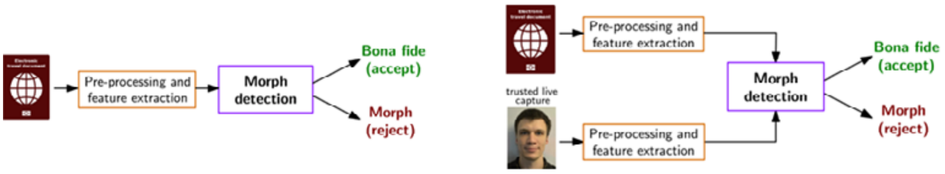

Figure 1: Morphing attack detection scenarios. Left: Single image Morphing Attack Detection, Right: Differential Morphing Attack Detection

The challenge is that MAD systems can, to date, not generalize across databases (e.g. of different sample quality) and can either reliably detect morphed images stemming from a print and scan process yet. Thus, the two pressing objectives are to:

[3] As Europe cannot impact passport application procedures in non-European countries, one should expect morphed passport to be presented at European borders way beyond that date.

- Improve the detection accuracy and its MAD's capability to generalize across databases.

- Measure detection accuracy as a function of the quality (e.g. are the 90-pixel inter-eye-distance sufficient).

The robustness of differential MAD will depend on both the different resolution of the enrolment images (i.e., in the passport), and on the quality of face image data and real-life noise (i.e., illumination, pose and shadow) that are commonly encountered in ABC systems. Specific challenges foreseen are to:

- Achieve high morphed detection rate without increasing the false rejection rate, even in the presence of image variations (i.e., pose, lighting, accessories, etc.)

- Adjust MAD for high-quality and low-quality low-resolution morphs. Get a better idea of the characteristics of these morphs in order to improve the detection performance.

A general challenge lies in the geometric transformation of the trusted live face image to the passport image geometry with a sufficient accuracy, i.e., accurate registration of the images is required. Further, the morph composition will have hyperparameters such as the percentage of the two images, contributing to the morph, in the morphing process, which needs to be estimated in the differential detection process.

Further research as suggested by the iMARS consortium is needed.

Need to detect automatically Morph Passports in Forensic Investigations

In order to support forensic investigations, the iMARS consortium suggests the development of explicit single image detection algorithms based on advanced feature extraction methods, which is especially relevant when no trusted image reference of the passport owner is available (see Figure 1). Regarding iMARS implicit single image detection, DNNs will be trained with large-scale face morph databases. As a consequence, robustness of the deep learning-based morph detector based on DNNs will be improved. Advanced machine learning techniques, e.g. transfer-learning, will be analysed, which will deal with different levels of quality, as the iMARS databases will cope with the problem of variability of face sample quality. This class of MAD solutions is also known as no-reference MAD or forensic MAD.

In a forensic investigation, the examination is based on the relatively low-resolution digital image stored in the passport, which has been processed by the authority or passport producer. While a morphed image may be visually indistinguishable to humans, the signal artefacts may be detected by MAD solutions. However, in carefully designed morphed images, the signal artefacts can be attenuated or completely suppressed. Further as the print and scan process tends to hide morphing artefacts, digital forensic tools are confronted with a challenge to detect alterations. As MAD methods mostly rely on trainable classifiers such as a Support Vector Machines (SVMs), their detection capability is linked to the quality of the training data. In addition, the artefacts have to pass the underlying feature extractor, i.e. even if there are artefacts in the image, those may not be reflected in the extracted features leading to misclassification of the MAD classifier. In explicit methods, the feature extraction is manually designed and the classifier needs to find a proper decision boundary. Thus, the number of parameters to be estimated is low compared to the number of parameters to be learned in end-to-end approaches.

Advanced feature extraction methods and image forensic techniques will be employed to improve the MAD performance. Further techniques will be applied to enhance robustness and generalizability, e.g. fusion of multiple MAD subsystems.

iMARS will also categorize and consider admissible image processing steps in contrast to typical manipulative processing steps, and will analyse and model their traces in image data. This knowledge will strengthen the best detection methods and will identify rules to facilitate the distinction between correct bona fide and manipulated images.

Consequently, for implicit morphing attack detection, the number of training data used for end-to-end learning via CNNs has to be larger than for explicit methods. In addition, to avoid learning spurious correlations, sufficient variation in the training data is mandatory. Finally, it should be noted that, given the independence in the underlying concept, the fusion with other approaches is a promising direction.

Further research as suggested by the iMARS consortium is needed.

Need to compose Test Data and establish an Online Evaluation Platform

Testing of MAD solution can't be done without appropriate data. To tackle this issue, the SOTAMD consortium has composed a database of 150 individuals. Multiple passport enrolment images have been captured over the Summer 2019 with a typical eMRTD issuance process, including print and scan from the facial images, and from at least two automated border gates facial samples have been acquired (e.g., from the German BEC testing gate in Bonn-Siegburg). This dataset can be considered as a high quality data set. The data was split into a subset that is used for the morphing process and a disjoint subset that serves as bona fide image in a differential morph detection trial. Morphed face images were generated by each academic SOTAMD partner[4] with three different selected morphing algorithms. To mimic the application process as close as possible, both bona fide and morphed images are printed using professional photo printing devices and then scanned afterwards.

The iMARS consortium suggests to augment this initial data and to contribute an additional dataset (around 10,000 digital morphed images) with multiple enrolment and border gate probe images per subject, which are captured in a variety of illumination conditions. This new data will stem from challenging operational conditions at 5 selected borders (e.g. Cyprus, France, Greece, Israel, Portugal). This data will constitute the iMARS mixed-quality dataset. Moreover, the iMARS consortium suggested to contribute high quality morphs. The main challenges obtaining good morphs of images lie in: i) the mapping of the corresponding image positions or elements, and ii) the proper fusion of the image texture information. While automated morphing strategies lead to visually appealing results in certain scenarios, the results can degrade considerably under conditions such as pose variation. Although the ISO/IEC 19794-5 standard

[4] Academic SOTAMD partners are University of Bologna, University of Twente, Norwegian University of Science and Technology and Hochschule Darmstadt [5] https://biolab.csr.unibo.it/FVCOnGoing

[ISO19794] defines face poses close to zero degrees, in practice this does not hold true in all cases. The main challenges for successful morphing strategies thus are:

- Minimizing image artefacts generated by morphing to limit as much as possible human intervention.
- Establishing the best morphing factor (also known as α factor) to maximize the probability to fool the human officer during enrolment and the automatic recognition system at the border.
- Developing automatic methods for aligning critical areas such as nostrils and irises.
- Aligning the eyes between the two images to reduce visual artefacts in the face.
- Avoiding algorithms that widely introduce visual artefacts such as shadows etc.

The University of Bologna is currently extending the existing FVConGoing platform[5] with new benchmarking services for differential morph detection and thereby developing the Bologna-Online-Evaluation-Platform (BOEP) platform. The new SOTAMD dataset will be stored in a highly protected environment, not exposed directly to the internet. It is suggested that the iMARS mixed-quality dataset is added, as soon as it becomes available. Further, it is suggested that BOEP will be extended in order to benchmark also non-reference MAD mechanisms. The datasets will be accessible through BOEP and provide open access benchmark tests. Thus, with BOEP, Frontex and the national border control agencies will be able to evaluate if the MAD State-of-the Art meets the operational requirements. The technical interfaces are by design equivalent to the benchmark portal of the NIST Face Recognition Vendor Test (FRVT) MORPH Competition [NISTFRVT]. However the functionality of BOEP will exceed FRVT-MORPH. The BOEP will provide a dedicated benchmark environment that can allow different tests on different selection pre-conditions (e.g., lookalikes, random selected pairs, or skin-color similarities, etc.). The quality and method of the morphing technique influences the ability of detection. In order to cover a broad spectrum of attacks the generated database has to cover a broad range of morphing strategies.

Hosting the data on the BOEP will enable researchers and operators to submit algorithms for online evaluation, without the need that confidential data has to travel to an evaluation lab. This will hence allow:

- Testing on lookalikes, same demographic subgroup, versus random selected pairs
- Testing with variation of morph algorithms, alpha values, and resolution.
- Testing with various facial image quality

The data that will be available on the BOEP will constitute a scenario test. In order to evaluate the impact of operational deployment, border control agencies shall be motivated by Frontex to contribute any real case data for differential MAD cases or forensic MAD cases to the University of Bologna, such that an additional benchmark with real case morph images can be offered. Specifically, for the differential MAD testing case, the collected data on the BOEP has no time gap between the passport image and the trusted image from the ABC gate. Thus, image pairs from real cases, where there might be a time gap of up to ten years, is of great interest.

Further morphing attack data as suggested by the iMARS consortium is needed.

Need to standardize Testing of MAD Solutions

When analyzing the vulnerability of face recognition systems to morphing attacks, the need to augment the metrics for evaluation of presentation attacks is obvious. The *Impostor Attack Presentation Match Rate* (IAPMR) [ISO30107] introduced in ISO/IEC 30107-3, represents a standardized metric for evaluating the impact of a presentation attack. Contrary to PAD evaluations, for a morphing attack all individuals contributing to the morph want to be successfully authenticated against the morphed facial image. This scenario cannot be evaluated using the IAMPR and thus motivated the introduction of new evaluation metrics [Sch17]. A Morphing Attack (MA) is only successful if all involved subjects have been successfully verified. Motivated by ISO/IEC 30107-3 [4], the *Mated-Morph-Presentation-Match-Rate* (MMPMR) is proposed in [Sch17] to evaluate the effect of a MAs on the overall system. This metric that has been established in the academic literature and should be further developed in an international ISO/IEC standard. The iMARS consortium suggested to anchor the MAD evaluation methodology in the by ISO/IEC 30107 multipart standard.

Standardization of measures to define the threats and the efficacy of countermeasures in a quantifiable and objective way has taken first steps. The challenge is to:

- Find consensus in the MAD research community and formulate a narrow set of relevant metrics.

- Standardise metrics to evaluate the performance of MAD methods and the vulnerability of biometric recognition systems to morphing attacks.

The iMARS consortium will initiate the ISO/IEC standardisation process of metrics to evaluate the performance of MAD methods and measures for the vulnerability of biometric recognition systems to morphing attacks. Identified methods will be validated with live images acquired on operational eGates.

An international standard for MAD testing as suggested by the iMARS consortium is needed. Border control agencies of EU Member State shall be motivated by Frontex to participate in this standardisation process.

Need to develop Face Image Quality Metrics

Assessment of face image quality is vital to capture samples that are sufficiently good in term of illumination, sharpness, or pose, such that the probe sample can verify an individual's identity accurately and reliably. The framework for biometric sample quality is well described in ISO/IEC 29794-1:2016 [ISO29794]. The essential definition is that a quality measure shall represent the quality of the source (e.g. the skin for a fingerprint recognition system) but also the fidelity of the sensor (i.e., is the image signal representing the source?). An expression of a quality score must be in the range of 0 to 100 (poor to best quality). Moreover, a quality score must be capable of predicting recognition performance (i.e. a sample with a low-quality score will likely not reach a high similarity score in a later recognition). Such correlation of quality scores and low false–non-match rate can be observed with the Error-versus-reject-curve (ERC).

The iMARS consortium suggests the development of an automatic face image quality assessment software, which can predict recognition accuracy and provide actionable feedback to the data capture subject and/or to the operational personnel. Such software will serve the needs of passport enrolment agencies but also European agencies (e.g. FRONTEX and EU-LISA) that must control quality of data in their databases. The resulting software prototype to automate image quality assessment will form the basis for a technical contribution to an international standard ISO/IEC 29794-5 (as revision of the previously existing technical report). This work will be the equivalent to the successful NFIQ2.0 metric for fingerprint images.

Once predictive face quality software is available, MAD evaluation can be adapted to the three relevant scenarios (i.e., ID Document issuance, border control, and forensic investigation) that should be used to evaluate the MAD solutions. The iMARS consortium suggested to adjust to the use cases corresponding to the different dimensions of quality: Image quality will differentiate use cases involving, on the one hand, always high quality, versus, on the other hand, a high-quality image during enrolment and a low-quality image at the border gate (due to poor illumination or pose variations caused by distracting factors at the gate). Once a predictive metric is available, the impact of face image quality on biometric recognition and MAD performance can be evaluated, e.g. the correlation of quality of acquired face images for an image pair based morphing attack detector can be measured. Moreover, the impact of face image quality on biometric recognition performance as well as morphing attack detection will be benchmarked.

Need to train operating Border Officers and Communication Personnel

The iMARS consortium is committed to deliver sustainable solutions for Border control operators. In interaction with Frontex and governmental agencies of European member states, the iMARS consortium suggests to address the usability/ergonomics requirements defined by the operators. Furthermore, the iMARS consortium suggests to develop best practices and a training curriculum for improving the officers' skills on manipulated/morphed image and document fraud detection while also respecting fundamental rights: training will lead to better mutual understanding and knowledge transfer. The iMARS consortium will design a training curriculum to reinforce end-users' skills on MAD solutions and to transmit professional expertise gained during the project. This curriculum will allow increasing their detection of manipulated images or document fraud, will enhance their skills, but also show them that the tools will not replace, but complement, their expertise. The iMARS consortium will team up with Frontex and their training procedures.

Training of operators' communication personnel is also considered to mitigate public excitement and explain attack resolving solutions against morphing attacks, once the threat is reported in the media.

Impact and Conclusion

If the needs elaborated in this contribution are implemented, then a strong impact can be expected for the security of European borders. The SOTAMD and iMARS consortium will facilitate reproducible research by employing a unified platform to allow standardised online evaluation. Application oriented specific benchmarks will be defined and developed. This will lead to a situation where error rates of differential MAD approach are reduced considerably – thus the chance that a criminal can fool an ABC system while keeping the amount of false morphing warnings will be quite limited.

The efficiency of the iMARS MAD solutions will entice practitioners to use them: the reduction of false alarm rate (Bona Fide Presentation Classification Error Rate - BPCER) of morphing attack detectors will allow the technology to be deployed at border with neglectable interference of the passenger flow and the outcomes can be successfully integrated into the existing ePassport life-cycle in a reasonable amount of time. When MAD solutions are deployed as first-line control solution, they will not be a substitute for a human expertise, which will always require that results need to be confirmed by an empowered agent in the second line in case of alarm or doubt.

The standardisation activities performed during the project will also contribute to the reproducibility of the tests performed and are of global benefit for all ICAO members. The standardisation project on face image quality ISO/IEC 29794-5 will be initiated and supported.

Within the iMARS project, new strategies to prevent cross-border crime will be proposed and implemented. Findings of the project will be consolidated in form of guidelines, which could be used by civil servants to take transparent and reliable decisions. For example, in case of doubt, biometric verification at the border shall be done with a second biometric identifier, such as fingerprint or iris reference images.

References

[Fer14] M. Ferrara, A. Franco, and D. Maltoni, The Magic Passport , in proceedings International Joint Conference on Biometrics (IJCB), Clearwater, Florida, USA, pp.1-7, October 2014..

[Sch19] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch. Face recognition systems under morphing attacks: A survey. IEEE Access, 7:23012–23026, 2019.

[ISO19794] ISO/IEC JTC1 SC37 Biometrics, ISO/IEC 19794-5:2005, Biometric data interchange format - Part 5: Face image data, 2005.

[ISO29794] ISO/IEC JTC1 SC37 Biometrics, ISO/IEC 29794-1:2016, Biometric sample quality – Part 1: Framework, 2016.

[ISO39794] ISO/IEC JTC1 SC37 Biometrics, ISO/IEC 39794-1:2019, Extensible biometric data interchange format – Part 1: Framework, 2019.

[ISO30107] ISO/IEC JTC1 SC37 Biometrics, ISO/IEC 30107-3:2017, Biometric presentation attack detection – Part 3: Testing and Reporting, 2017.

[ICAO9303] International Civil Aviation Organization, ICAO Doc 9303, Ma-chine Readable Travel Documents - Part 9: Deployment of Bio-metric Identification and Electronic Storage of Data in MRTDs (7th edition), 2015.

[NISTFRVT] U.S. NIST Face Recognition Vendor Test – Morph, https://www.nist.gov/programs-projects/frvt-morph

[Sch17] U. Scherhag, A. Nautsch, C. Rathgeb, M. Gomez-Barrero, R. N. J. Veldhuis, L. Spreeuwers, M. Schils, D. Maltoni, P. Grother, S. Marcel, R. Breithaupt und R. Ramachandra, „Biometric Systems under Morphing Attacks: Assessment of Morphing Techniques and Vulnerability Reporting," in Proceedings of the 2017 International Conference of the Biometrics Special Interest Group (BIOSIG), 2017.

[Ram16] R. Ramachandra, K. B. Raja und C. Busch, „Detecting morphed face images," in Proceedings of the 8th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2016.