Robustness of Deep Learning Based Face Recognition Under Morphing Attacks

Iurii Medvedev
https://visteam.isr.uc.pt/team/iurii-medvedev/

Institute of Systems and Robotics,
University of Coimbra,
Portugal

Nuno Gonçalves
https://visteam.isr.uc.pt/team/nuno-goncalves-2/

Portuguese Mint and Official Printing Office (INCM),
Lisbon,
Portugal

INTRODUCTION

Last decades with the development of deep learning techniques the evident advances have been reached in the field of face recognition. However at the same time more evolved and sophisticated techniques for performing the presentation attacks continue to appear (see Fig. 1), which require the development of new protection solutions [1].

One of such face image manipulating methods is **Face Morphing**. Image morphing techniques are used to combine information from two (or more) images into one image. Over the past decade, it has gained significant attention and has been more thoroughly investigated. As awareness of the problem has grown, numerous counterfeit documents employing face morphing techniques have been uncovered at control gates [12] (an example is presented at Fig. 2).

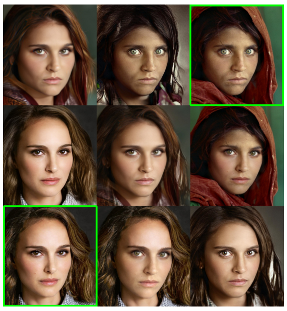

Figure 1: Example of various face morphing techniques.

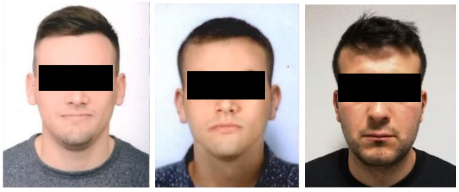

Figure 2: Real face morphing example. a) - ID image of accomplice (requester of a document); b) Face image on counterfeited ID Document; c) Face image of a person that used a counterfeited document.

Given the importance of reducing vulnerabilities in modern face recognition systems and the significant risks posed by presentation attacks, this thesis aims to contribute with the tools for combating the face morphing problem involving deep learning algorithms.

FACE MORPHING ATTACKS PROBLEM

The potential for morphing attacks to compromise identification systems was first explored in [3]. This study shows how ICAO-compliant morphed images could effectively bypass both human and automated border control checks.

The typical pipeline for creating a morphed identity document for impersonation involves several coordinated steps designed to exploit face recognition systems. First, a wanted individual collaborates with a complicit accomplice to generate a synthetic morphed face image with facial features from both parties. This morphed image is made to be realistic and compliant with official ID photo standards. Next, the accomplice uses this image to apply for and obtain a legitimate identification document, such as a passport or national ID card. Once issued, the document (though legally tied to the accomplice) can also successfully match the facial characteristics of the wanted person. As a result, the wanted individual is able to use the authentic document to travel or access secure services while evading detection.

RESEARCH OBJECTIVES AND QUESTIONS

The primary objective of this thesis is to contribute to enhancing the robustness of face recognition systems against presentation attacks, particularly those involving face morphing [10, 11]. Our research will focus on two key areas: improving the detection of morphing attacks and strengthening the robustness of deep facial feature representations against morphing.

In this context, we can outline our specific objectives as follows:

• Study data collection for face recognition, including morphed face generation.

• Improve deep learning methods for morphing detection and vulnerability analysis in ID enrollment.

• Develop morph-resistant face recognition strategies for secure document applications.

• Evaluate the effectiveness of proposed methods with custom protocols and bublic benchmarks.

MORPHING ATTACK DETECTION

Morphing Attack detection is a straightforward approach to combat their risks for facial biometric systems and it is typically categorized into two processing pipelines based on the availability of reference data: *no-reference* and *differential* approaches.

In the *no-reference* or *Single Morphing Attack Detection (SMAD)* scenario, the algorithm receives only a single face image without any corresponding trusted reference and must determine whether the image is morphed (for instance, in Enrollment pipelines).

Differential morphing attack detection (DMAD) methods rely on the availability of a trusted reference image typically captured during a live interaction with the facial biometric system allowing the comparison between the live capture and the enrolled (potentially morphed) image (for instance in Automated Border Control).

For morphing attack detection, we proposed advanced detection strategies based on multitask learning frameworks with sophisticated morph sample labeling (see Fig. 3). In our setup, images are processed by two parallel feature extractors, and their outputs are compared to evaluate whether they belong to the same identity. The underlying principle is that genuine (non-morphed) face pairs will produce highly similar features, while morphed images will result in dissimilar outputs.

On practice this implies the designing of a complex multitask learning problem and we will call such concept as *Fused Classification* further in the work.

Our methods achieved state-of-the-art performance in publicly available benchmarks for both Single-image Morphing Attack Detection (SMAD) [4] and Differential Morphing Attack Detection (DMAD) [8] scenarios.

INCREASING THE ROBUSTNESS OF FACIAL BIOMETRIC TEMPLATES TO MAD

Beyond detection, alternative strategies exist. For instance in can be approached by increasing the *Robustness* the face feature templates to morphing attacks. This approach shifts the emphasis from detecting morphs

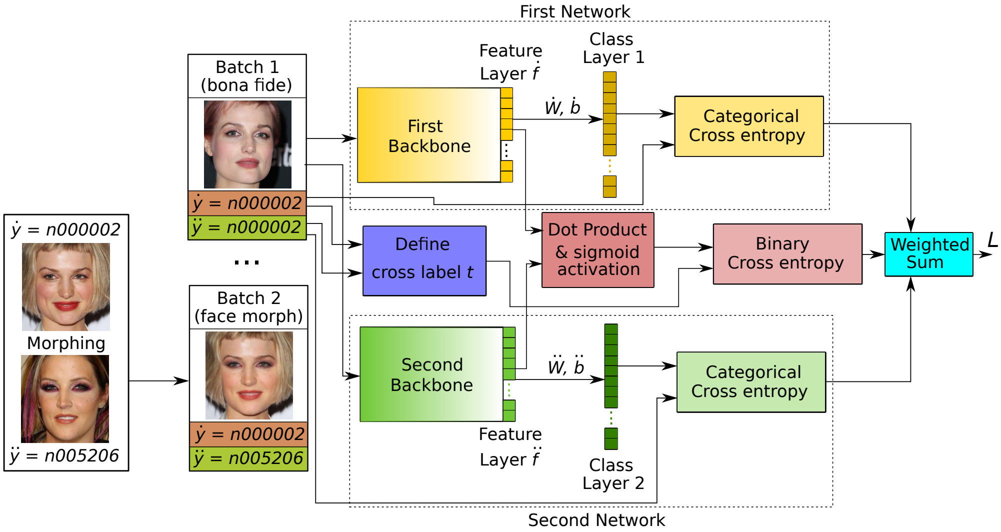

Figure 3: Schematic of the Fused Classificartion. For clarity in visualization, the presented batch contains only a single image. Labels $\dot{y}$ and $\ddot{y}$ are displayed with descriptive names for better understanding. In the actual implementation, these labels are represented by their numerical index values, which are subsequently encoded into one-hot vectors for processing.

to mitigating their impact by reducing the likelihood that morph samples can successfully match legitimate face templates.

Achieving this level of robustness requires modifications to the core of the face recognition system. Namely the target is to modify the face feature extraction mechanism, which is used for generating biometric templates. Such robustness based strategies focus on improving the discriminatory power of the templates themselves by modifying the deep face feature domain.

Our initial approach here addresses contrastive learning methods by introducing a dedicated branch for morph samples, allowing explicit control over their feature distribution [2]. Additionally, we refine traditional classification strategies through a carefully designed softmax-based margin loss, which intentionally disbalabce morph samples from bona fide ones [5].

THESIS CONTRIBUTIONS

In this thesis we approached the problem of face recognition robustness to morphing attacks from multiple angles, introducing new methods for morphing attacks detection and face image templates robustness.

In the work, many collateral contributions were presented, which appeared mainly in the process of data curation and performance assessment. This include novel datasets and benchmarks tailored to face recognition tasks [9]. These resources have been made available to the academic community and are already in use for new research projects.

This also include the developed data filtering techniques and resulting metadata for public academic face datasets [13] [6], which can facilitate more refined data preparation in future studies.

Our contributions include the creation of a dedicated benchmarks for face morphing detection [4][8] and evaluating robustness to morphing attacks [7], which is a resource that, to our knowledge, currently has no publicly available alternative and is designed to scale with future developments .

Summing up in this thesis we provided significant contributions that support the broader academic community in advancing research in face recognition and presentation attack detection.

REFERENCES

[1] Zahid Akhtar, Dipankar Dasgupta, and Bonny Banerjee. Face Authenticity: An Overview of Face Manipulation Generation, Detection and Recognition. *SSRN Electronic Journal*, 01 2019. doi: 10.2139/ssrn.3419272.

[2] W. Chen, X. Chen, J. Zhang, and K. Huang. Beyond Triplet Loss: A Deep Quadruplet Network for Person Re-identification. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1320–1329, 2017. doi: 10.1109/CVPR.2017.145.

[3] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. The magic passport. *IJCB 2014 - 2014 IEEE/IAPR International Joint Conference on Biometrics*, 12 2014. doi: 10.1109/BTAS.2014.6996240.

[4] I. Medvedev, F. Shadmand, and N. Gonçalves. Mordeephy: Face morphing detection via fused classification. In *Proceedings of ICPRAM*, pages 193–204. SciTePress, 2023. ISBN 978-989-758-626-2. doi: 10.5220/0011606100003411.

[5] Iurii Medvedev and Nuno Goncalves. Morphguard: Morph specific margin loss for enhancing robustness to face morphing attacks, 2025. URL `https://arxiv.org/abs/2505.10497`.

[6] Iurii Medvedev and Nuno Gonçalves. Improving performance of facial biometrics with quality-driven dataset filtering. In *2023 IEEE 17th International Conference on Automatic Face and Gesture Recognition (FG)*, pages 1–8, 2023. doi: 10.1109/FG57933.2023.10042579.

[7] Iurii Medvedev and Nuno Gonçalves. Morfacing: A benchmark for estimation face recognition robustness to face morphing attacks. In *2024 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–10, 2024. doi: 10.1109/IJCB62174.2024.10744449.

[8] Iurii Medvedev, Joana Alves Pimenta, and Nuno Gonçalves. Fused classification for differential face morphing detection. In *2024 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*, pages 1043–1050, 2024. doi: 10.1109/WACVW60836.2024.00114.

[9] Iurii Medvedev, Farhad Shadmand, and Nuno Gonçalves. Young labeled faces in the wild (ylfw): A dataset for children faces recognition. In *2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition (FG)*, pages 1–10, 2024. doi: 10.1109/FG59268.2024.10582021.

[10] Raghavendra Ramachandra and Christoph Busch. Presentation Attack Detection Methods for Face Recognition Systems: A Comprehensive Survey. *ACM Comput. Surv.*, 50(1), March 2017. ISSN 0360-0300. doi: 10.1145/3038924. URL `https://doi.org/10.1145/3038924`.

[11] U. Scherhag, R. Raghavendra, K. B. Raja, M. Gomez-Barrero, C. Rathgeb, and C. Busch. On the vulnerability of face recognition systems towards morphed face attacks. In *2017 5th International Workshop on Biometrics and Forensics (IWBF)*, pages 1–6, 2017. doi: 10.1109/IWBF.2017.7935088.

[12] Matjaž Torkar. Morphing cases in slovenia. NIST IFPS, 2022. Ministry of the Interior Police, Slovenia.

[13] João Tremoço, Iurii Medvedev, and Nuno Gonçalves. QualFace: Adapting Deep Learning Face Recognition for ID and Travel Documents with Quality Assessment. In *2021 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–6, 2021.