Machine Intelligence Research
www.mi-research.net
DOI: 10.1007/s11633-024-1533-1

Improving Accomplice Detection in the Morphing Attack

Nicoló Di Domenico[1] Guido Borghi[2] Annalisa Franco[1] Davide Maltoni[1]

[1] Department of Computer Science and Engineering, University of Bologna, Cesena 47521, Italy

[2] Department of Education and Humanities, University of Modena and Reggio Emilia, Reggio Emilia 42121, Italy

Abstract: This paper addresses a critical security challenge in the field of automated face recognition, i.e., morphing attack. The paper introduces a novel differential morphing attack detection (D-MAD) system called ACIdA, which is specifically designed to overcome the limitations of existing D-MAD approaches. Traditional methods are effective when the morphed image and live capture are distinct, but they falter when the morphed image closely resembles the accomplice. This is a critical gap because detecting accomplice involvement in addition to the criminal one is essential for robust security. ACIdA's impact is underscored by its innovative approach, which consists of three modules: One for classifying the type of attempt (bona fide, criminal, or accomplice verification attempt), and two others dedicated to analyzing identity and artifacts. This multi-faceted approach enables ACIdA to excel in scenarios where the morphed image does not equally represent both contributing subjects — a common and challenging situation in real-world applications. The paper's extensive cross-dataset experimental evaluation demonstrates that ACIdA achieves state-of-the-art results in detecting accomplices, a crucial advancement for enhancing the security of face recognition systems. Furthermore, it maintains strong performance in identifying criminals, thereby addressing a significant vulnerability in current D-MAD methods and marking a substantial contribution to the field of facial recognition security.

Keywords: Biometrics, face recognition, morphing attack, morphing attack detection (MAD), differential MAD (D-MAD).

Citation: N. Domenico, G. Borghi, A. Franco, D. Maltoni. Improving accomplice detection in the morphing attack. *Machine Intelligence Research*. http://doi.org/10.1007/s11633-024-1533-1

1 Introduction

The recently introduced morphing attack[1] poses a significant security threat in face verification-based applications, systems are usually exploited for instance in the automatic border control (ABC) gates in international airports[2]. Indeed, through this attack, it is possible to obtain a regular and legal document that presents a morphed photo, i.e., a hybrid face image that hosts two different identities, and that can be therefore shared between two subjects. In this way, a criminal can bypass official controls using the identity of an accomplice without any criminal record[3].

In this context, the development of effective morphing attack detection (MAD) systems[4], i.e., methods able to automatically detect the presence of morphing in input images, are strongly demanded by private and public institutions. Generally, two families of MAD approaches are investigated in the literature[5]: Single-image MAD (S-MAD) methods, which usually rely on the detection of the presence of visible or invisible morphing-related artifacts in the single input image, and differential MAD methods (D-MAD), which usually compare the identity of the two facial images – the document and the trusted live acquisition image – received as input[6-8].

Generally, the evaluations reported in many works of D-MAD literature consider morphed images created with an equal presence of the two contributing subjects[9-13]. This choice maximizes the probability of fooling automated verification systems with both contributing subjects. However, a successful morphing attack also requires fooling the human examiner at the document enrollment stage, typically presenting a morphed photo very similar to the document applicant[14, 15]. Therefore, the morphed image should be created with a stronger presence of the accomplice, so that the human examiner is less inclined to notice significant differences.

In this operational scenario, D-MAD methods based on the comparison of identity features[6-8, 14] lose their effectiveness. Specifically, as reported in Fig. 1, the error rate is limited when the input images are sufficiently diverse (e.g., the morphed image is compared with the criminal), while the error tends to grow with the increasing of the identity similarity (e.g., the morphed image is compared with the accomplice).

Therefore, in this work, we introduce ACIdA, a novel modular D-MAD approach designed to effectively address accomplice verification attempts, delivering robust MAD performance with both contributing starting sub-

Research Article
Manuscript received on June 7, 2024; accepted on November 14, 2024
Recommended by Associate Editor Maoguo Gong
Colored figures are available in the online version at https://link.springer.com/journal/11633

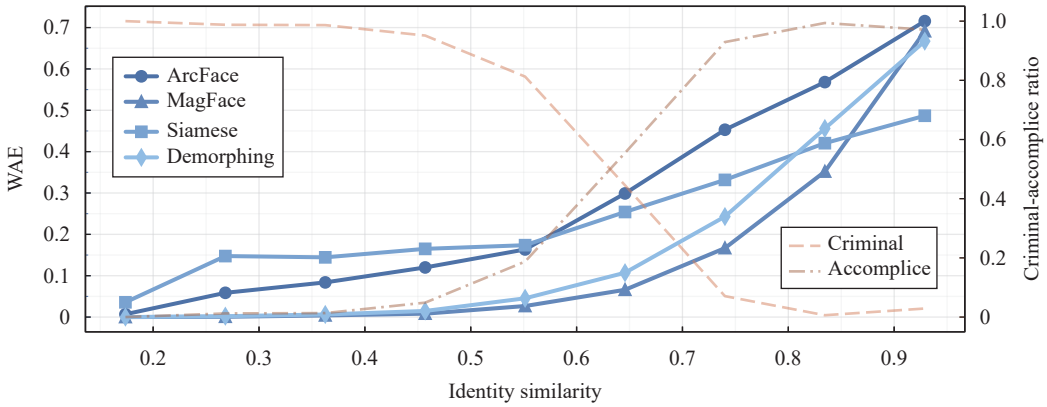


Fig. 1 Performance of literature D-MAD methods across different identity similarity scores between the document and live images. On x-axis, values close to 1 indicate a high similarity. The analysis shows that all methods are negatively influenced by increasing identity similarity, which corresponds to a higher presence of the accomplice in the input images instead of the criminal (both percentages are represented with dotted lines). The criminal-accomplice ratio is computed and represented with dotted lines. The analyzed D-MAD methods include ArcFace[6], MagFace[7], Demorphing[14], and Siamese[8], with further details provided in the referenced section. (Colored figures are available in the online version at https://link.springer.com/journal/11633)

jects. To evaluate the proposed system, we introduce and explore a novel scenario where D-MAD is employed to detect both criminal and accomplice verification attempts, as visually summarized in Fig. 2.

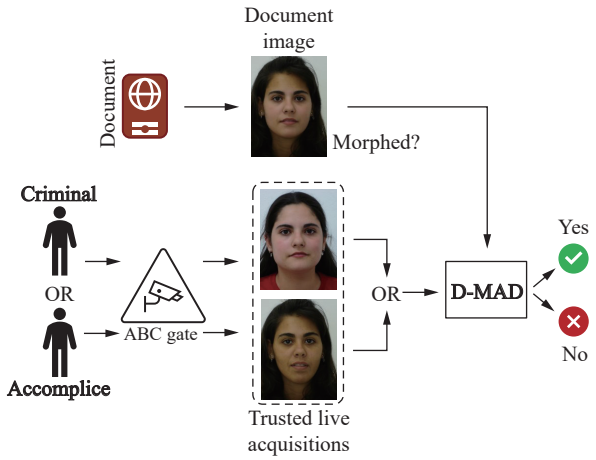


Fig. 2 Visualization of the introduced MAD scenario: The document image, eventually morphed, is compared with a trusted live acquisition obtained at the gate in which the criminal or the accomplice is passing. The proposed system, ACIdA, is focused on document images more similar to the accomplice. (Colored figures are available in the online version at https://link.springer.com/journal/11633)

Successfully tackling this task would impact the MAD task and broaden the operational context of D-MAD methods. Firstly, D-MAD methods could be exploited at the enrollment stage to avoid a morphed image being included in the official document requested by the accomplice. Secondly, the tested scenario would allow the use of MAD methods in ABC gate to identify not only the criminal but also the accomplice, who has stained itself with a criminal and punishable action during the document-issuing procedure. It is also worth considering that since in the morphing attack the accomplice applies for a valid passport, he will have to use that passport for several years if he needs to travel, thus making of interest the detection of the accomplice's attempt. We observe this scenario represents an interesting and challenging research field, that needs to be addressed in future MAD-related work.

From an implementation point of view, ACIdA is based on three modules that focus on different aspects of the introduced scenario: the attempt classification (AC) module, responsible for the classification of the identity verification attempt in three different classes (i.e., criminal, accomplice or bona fide attempt); the identity-artifact module (IdA), based on a combination of artifacts detection and identity analysis; the identity module (Id), that is based only on identity features.

The rationale behind this proposal is that when the identity information is very certain (as in the case of bona fide or criminal attempts) the D-MAD score should mainly rely on this aspect, while when the identity information is uncertain (as for the accomplice attempts), the D-MAD should also consider and evaluate the possible presence of artifacts to better detect morphing.

The underlying idea considered in this paper has been first introduced in [16] where we showed that the combinations of identity features and artifact analysis improve the accomplice detection. Then, this paper further extends that initial work, including the following original contributions:

1) We propose ACIdA, a deep learning-based D-MAD method expressively conceived to detect both criminal and accomplice attempts in the face morphing attack. Specifically, with respect to [16], we introduce a new module to provide the attempt classification, and other two modules focus on identity and artifact features. All these modules cooperate to output the final score through a weighted prediction.

2) We test ACIdA on a D-MAD scenario in which

document images are compared with both accomplice and criminal contributing subjects and the morphing process is not necessarily symmetric (i.e., it does not include an equal contribution of the two subjects). From a practical point of view, this approach broadens the scope of application of current D-MAD methods to scenarios, such as the enrollment one, that have been relatively unexplored until now. We also highlight, from a quantitative point of view, the lack in terms of accuracy of identity-based D-MAD methods currently available in the literature.

3) We perform an extensive and cross-dataset evaluation, highlighting several experimental aspects of the proposed method. Moreover, we show the generalization capability of the proposed method in a different D-MAD scenario through the FVC-onGoing[1] platform.

For the sake of reproducibility and the comparison with future works, we publicly release the code, the list of subject pairs used for morphing generation for training and testing sets, and the implementation details of the proposed method[2].

1.1 Issues with existing D-MAD methods

Differential morphing attack detection (D-MAD) methods, also referred to as two-image or pair-based approaches, employ a pair of facial images as input. The objective of these methods is to determine whether the document image has undergone morphing. The second image, which is a verified live capture, serves as a crucial reference for the task, facilitating the analysis of inconsistencies in identity, texture, or color when compared to the first image, as analyzed in Section 2.

In practice, these two images can be obtained during passport issuance, where the first image is the provided photo and the second image is captured in real-time. Similarly, during controls at ABC gates, the live image is acquired through automated face verification procedures, while the probe image is retrieved from the electronic machine readable travel document (eMRTD).

Analyzing the results obtained on two important sequestered datasets, i.e., NIST FATE MORPH[3] and FVC-onGoing[17], it can be noted that the most promising D-MAD algorithms are based on the comparison of the identity features obtained through pre-trained deep learning architectures. We observe that these approaches can be limited in accuracy, in particular, when morphed images are compared with the accomplice and, in general, with look-alike subjects.

These hypotheses are validated through our experiments, the results of which are presented in Fig. 1. Specifically, we evaluate the performance of various D-MAD systems from the literature across multiple image pairs with increasing similarity between the subjects, i.e., the document image and the live acquisition. Identity similarity is quantified by computing the cosine similarity of identity features extracted using the method described in [18]. For a thorough analysis, we also calculate the percentage of criminal and accomplice pairs relative to the total number of attempts as a function of similarity.

In order to summarize the MAD performance, results are reported through the weighted average error (WAE) metric (detailed in Section 4.4), i.e., the average of error-based metrics commonly used in the MAD task. As shown, all reported methods suffer the increasing similarity (values close to 1 on x-axis), which corresponds to an increasing number of verification attempts involving the accomplice subject instead of the criminal one (represented with dotted lines). The trend of the method[8] differs from other algorithms because, despite slightly lower performance for low similarity values, it exhibits greater robustness as the similarity value increases. One possible explanation derives from the fact that this method utilizes both identity-related features and features associated with the presence of artifacts to compute the final morphing score (see Section 2 for more details).

Furthermore, an additional limitation of current D-MAD methods lies in the robustness of pre-trained networks on large and diverse datasets for the face recognition task, which makes these methods potentially insensitive to clearly visible artifacts in morphed images. An example of this is depicted in Fig. 3, in which the SoA D-MAD method described in [6] wrongly predicts these images as bona fide, even though visible artifacts are present. From here, also taking into account the previous considerations, the intuition was born to include artifact detection in the proposed ACIdA method, as detailed in Section 3.2.

Fig. 3 Morphed images that successfully fool an identity-based D-MAD system, but exhibit visible artifacts related to the morphing procedure. This observation leads to the intuition to exploit artifact detection techniques in the proposed ACIdA system, as discussed in Section 1.1. (Colored figures are available in the online version at https://link.springer.com/journal/11633)

2 Related work

Given the importance of countering the face morphing attack, a variety of D-MAD methods have been introduced in the literature[19, 20]. From a general point of

[1] https://biolab.csr.unibo.it/fvcongoing/

[2] https://github.com/ndido98/acida

[3] https://pages.nist.gov/frvt/html/frvt_morph.html

view, these methods can be classified depending on the type of features they compute starting from the two input images: artifact-based, identity-based, demorphing-based and, finally, hybrid approached referred to as mixed feature-based.

2.1 Artifact-based D-MAD

In this category, D-MAD methods extract from the input images general-purpose features, both through classic computer vision techniques and deep learning-based architectures. The assumption is that the comparison of these features can highlight the anomalous traces – i.e., the artifacts – left by the morphing procedure.

As hand-crafted features, the most used are binarized statistical image features (BSIF)[21], local binary patterns (LBP)[22], and histogram of oriented gradients (HOG)[23], in combination with machine learning classifiers. These methods have achieved partially satisfactory results. The approach described in [24], involves computing a histogram of LBP for both images and subtracting them. The resulting 256-dimensional feature vector is then used to train a support vector machine (SVM) that generates the morphing score.

The article in [25] suggests using undecimated 2D discrete wavelet transform data to feed a Siamese neural network, which highlights the differences between genuine and altered images. While the findings presented are compelling, the absence of implementation details[4] limits the ability to reproduce the results.

Recently, various works have tackled the D-MAD task, using a variety of different deep learning architectures, such as the fusion of the output of different backbones[26, 27] or the feature-wise supervision on fine-grained classification[28]. The work of Singh and Ramachandra[29] describes a method based on the fusion of several deep features computed from six different convolutional neural networks (CNNs), trained on the ImageNet dataset, and merged through a spherical interpolation, referred as spherical linear interpolation (SLERP). Differently from ours, this method is not based on identity features, and it is specifically conceived for the on-the-fly D-MAD task with different camera resolutions and acquisition distances. In [12] Soleymani et al.[12] proposed a Siamese network based on the Inception ResNet architecture. After the first alignment stage, the embedding of the two input images is extracted; a contrastive loss is used during the training phase.

Finally, other methods compare specific facial elements, on which the Euclidean and angle distances are computed, as discussed in [30]. However, the performance of these techniques is heavily dependent on the accuracy of the facial landmark predictor, so they are not included in the current analysis. The method described in [31] proposes to analyze the locations of fundamental facial landmarks, in order to capture inconsistencies in the facial geometry introduced by the morphing process.

2.2 Identity-based D-MAD

Identity-based D-MAD methods, as the name suggests, are based on the extraction of features related to the identity of the depicted subjects in the input images. The assumption is that it is possible to detect the morphing attack through a sort of face verification procedure.

One of the most effective and accurate identity-based D-MAD approaches is presented in the recent work of Scherhag et al.[6], in which authors propose a convolutional neural network (CNN) architecture, specifically a ResNet50[32], trained with an angular margin loss referred to as ArcFace[33], originally designed for the face recognition task. This architecture is utilized to extract feature embeddings from the input images. According to Scherhag et al.[6], the network is pretrained and no additional training procedures are performed specifically for the morphing detection task: This ensures that the deep learning model does not suffer from overfitting with training datasets limited in size and variety. Finally, the extracted features are then subtracted and fed into an SVM for the final classification process. A recent evolution of this system has been proposed in [7], in which the backbone is a ResNet trained through the MagFace[18] loss function, an adaptive mechanism to learn a well-structured within-class feature distribution relying on the magnitude of vectors that have achieved SoA performance on the face recognition task. Another method[34] extracts identity features and learn to implicitly disentangle identities from the morphed image conditioned on the live trusted acquisition using a conditional generative adversarial network (GAN).

2.3 Demorphing-based D-MAD

In this category, the main idea is to reverse the morphing procedure to restore the identity of the legitimate document owner. In this manner, the comparison between the morphed image and the restored one can reveal the presence of the morphing attack.

A seminal work is described in [14], where the reverse morphing procedure (referred to as "demorphing") is applied to the input images in order to reveal the real identity hidden in the morphed image. This method works in combination with commercial-on-the-shelf (COTS) face verification systems and is interesting since it does not require any training procedure. In this case, the main issue is represented by the fact that the morphing process is rarely a simple linear combination as assumed by Ferrara et al.[14]. Moreover, the entire process is based on the accuracy of the estimation of the position of facial landmarks and even small localization errors could negatively

[4] At the time of writing, the method is being patented

influence the effectiveness of the whole pipeline. Similar approaches based on the generative adversarial network (GAN) paradigms have been proposed[35-38] in order to restore the accomplice's facial image hidden in the morphed one, with partial generalization capabilities. More recently, diffusion autoencoders have been applied to the demorphing task[39] to improve the resulting image quality and restoration accuracy. Finally, Long et al.[40] introduce a novel network based on transformer feature interaction to restore the accomplice's face.

2.4 Mixed feature-based D-MAD

D-MAD methods of this category are hybrid, since they are based on a combination of the identity and artifact-related features.

The work described in [8] proposes a deep learning-based Siamese network. This network not only focuses on features related to the identity of the input subjects but also on the presence of artifacts caused by the morphing process. The network has two separate branches that analyze the same input image, and their outputs are then combined using a fully connected layer to generate the final morphing score. In [16] the authors investigate the use of both D-MAD and S-MAD features to improve detection performance. Different ways to combine these features are analyzed. The analysis of the mixed feature-based D-MAD highlights the importance of combining features related to artifacts, typical of S-MAD methods, with features related to the subject identity, more typical of the D-MAD field: In these terms, only a few works explore this topic. In the proposed method, the combination of these two kinds of features is not fixed at the time of inference but is dynamically established based on the diversity of the input subjects, i.e., the subject depicted in the document image and the one captured through the live acquisition, in order to improve the accuracy.

3 Proposed method

The underlying idea of the development of our method is to process every possible attempt – criminal, accomplice, and bona fide – with a specific module based on different features.

Our insight is that for criminal and bona fide verification attempts, identity-related features are very effective, as confirmed by the good results achieved by [6, 18]. Differently, for accomplice verification attempts, due to the greater similarity between subjects, the discriminative power of identity features is reduced and we believe the combination with artifact analysis can improve the MAD performance. To accomplish this paradigm, an initial selector of the attempt type is needed. Then, we build a classifier that outputs the probability of the attempt type used in the final MAD score computation. The effectiveness of the use of different modules is experimentally confirmed by the ablation study reported in Section 4.6.

A general overview of the proposed method is depicted in Fig. 4. As shown, the framework is divided into three main modules: the attempt classification (AC) module, assigned to the classification of the pair attempt provided in input, the identity-artifact (IdA) block, specialized in the detection of morphed images combining identity and artifact information, and the identity block (Id), that relies only on identity analysis to detect morphing attacks.

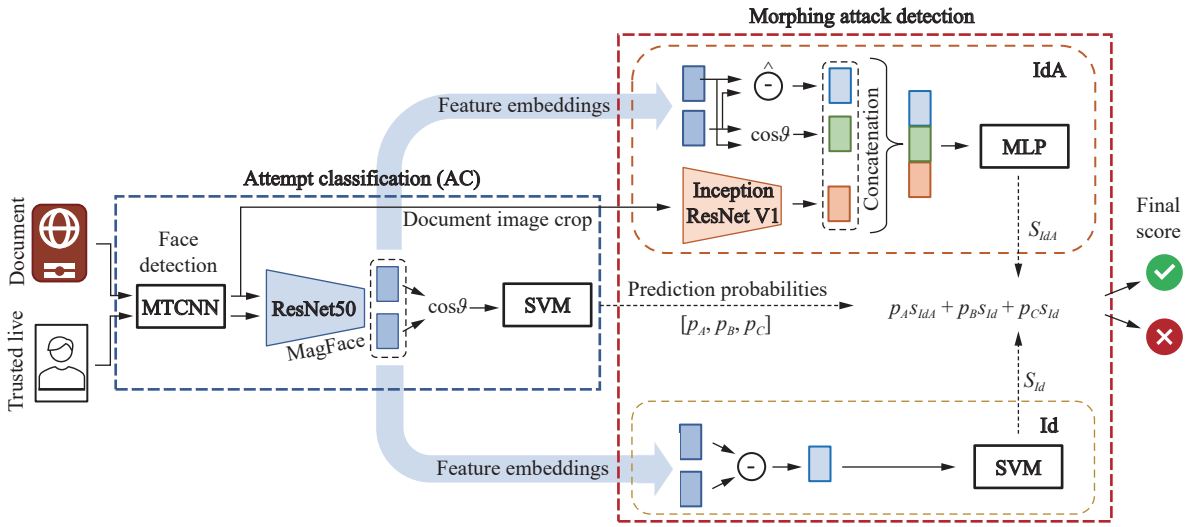


Fig. 4 General overview of ACIdA. The method is composed of three different modules: the attempt classification (AC) module, that determines if the document image is compared with the criminal or the accomplice trust live image (see Section 3.1); the identity (Id) module, an identity comparison-based MAD system (see Section 3.3) and the identity-artifact (IdA) module, that integrates both information about identity and artifact detection (see Section 3.2). The score of these two MAD modules is combined through a weighted sum to produce the final output of the system. (Colored figures are available in the online version at https://link.springer.com/journal/11633)

The final output score is obtained through a weighted sum, formally defined as:

$$S = p_A \times S_{IdA} + p_B \times S_{Id} + p_C \times S_{Id}. \quad (1)$$

Let p_A, p_B, p_C denote the probabilities that the document image is compared with the accomplice, bona fide, and criminal subjects, respectively; these probabilities are generated by the SVM classifiers within the AC module. Additionally, S_{IdA} and S_{Id} represent the outputs produced by the classifiers of the IdA and Id modules, respectively. Through this procedure, the final output is obtained as a weighted sum of all contributions generated by the different modules, as detailed below.

3.1 Attempt classification module

The attempt classification module is responsible for preparing the input for all the other modules and for the pair typology classification itself, i.e., the prediction if the subject depicted in the live acquisition (attempt) is the criminal or the accomplice (in case of morphed document image), or the same subject (bona fide). The input is represented by the probe image, i.e., the one contained in the document, and the trusted live acquisition. Firstly, these images are fed into the MTCNN face detector[41] that, as the name suggests, crops the face excluding the large part of the background, preparing the input images for the feature extraction procedure. The resulting crops are then used as input to extract features through a backbone that outputs two feature embeddings of size 512. We employ a frozen iResNet[42] architecture as backbone, trained using the magnitude and angular MagFace loss[18], originally conceived for the face recognition task. This training process has been conducted on vast-scale datasets, namely MS-Celeb-1M[43] and VGG-Face2[44], which consist of trillions of face pairs. The two produced embeddings are combined through the cosine similarity and are finally used as input for an SVM classifier that outputs 3 possible different classes: "bona fide", "accomplice" and "criminal" along with their probabilities used for the final classification, as previously detailed.

3.2 Identity-artifact (IdA) module

As the name suggests, the idea behind this module is to combine information belonging to two different tasks, i.e., face verification and artifact detection, following the considerations reported in Section 1.1.

This module receives as input the extracted feature embeddings and the document image crops. As mentioned, these embeddings are extracted using a backbone architecture trained for the face recognition task, and thus we assume these vectors represent the identity of the faces available in the input images. These identity feature embeddings are then combined in two different ways: In the first, a subtraction followed by a min-max normalization (that rescales each component in the range $[0, 1]$) is exploited. In the second, the cosine similarity, a metric widely used in the check of the identity distance in the face verification task, is computed. In this manner, both the two embeddings produced contain information regarding the difference of the identities provided in input, an essential information for the D-MAD task[6, 7].

A third feature embedding containing information about the presence of artifact in the document image is computed as follows. The document image crops are fed into an Inception-ResNet V1[45] architecture, starting from the weights obtained with the VGG-Face2[44] datasets. A fine-tuning procedure is conducted on ICAO-compliant and JPEG-compressed images for the artifact detection task, following the findings reported in [46]. Finally, a 512-dimensional feature embedding is obtained removing the final last fully connected layer of the adopted architecture.

These three outputs are finally combined through a concatenation, that has revealed good performance in the preliminary work[16], obtaining a 1 025-dimensional feature vector, used as input to a multi-layer perceptron (MLP) architecture exploited as a classifier that produces in output a score in the range $[0, 1]$.

3.3 Identity (Id) module

For the identity module, we draw inspiration from the solution presented in [6], which can be considered as the current state-of-the-art D-MAD method, as evidenced by the results published on the FVC-onGoing platform[17].

We employ an iResNet[42] network that has been trained for the purpose of face recognition using the MagFace loss[18]. As the input consists of two images, this module generates two distinct feature embeddings of size 512. These embeddings are then combined through subtraction, resulting in a single final feature vector of the same size that is fed into an SVM classifier, trained to output a probability in the $[0, 1]$ range that represents whether the probe image is the result of a morphing process. Meng et al.[18] demonstrate that the MagFace loss function yields robust embeddings by maximizing the geodesic distance between different identities. Consequently, the produced embeddings exclusively contain information pertaining to the input face identity.

4 Experimental results

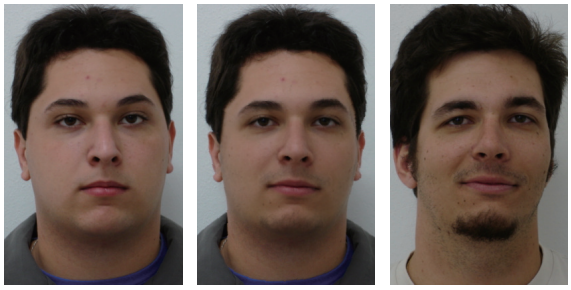
4.1 Investigated scenario

In the investigated scenario, we consider different types of verification attempts: 1) bona fide, in which an unmanipulated document image is compared to a live capture of the same subject; 2) the attempts where a

morphed document image is compared to a live capture of the criminal subject (morphed-criminal); 3) the attempts in which a morphed document image is compared to a live capture of the accomplice subject (morphed-accomplice).

The experimental results are therefore categorized into three distinct sub-benchmarks, based on the types of attempts considered:

1) Accomplice: This benchmark includes bona fide and morphed-accomplice verification attempts where the live image is from the accomplice. An example of the input pair of this benchmark consists in Figs. 5(a)–5(b).



(a) Subject 1 (b) Morphed (c) Subject 2

Fig. 5 In a D-MAD approach, the morphed image (Fig. 5(b)) can be compared with the criminal (Fig. 5(c)) or with the accomplice (Fig. 5(a)).

2) Criminal: Conversely, this benchmark encompasses bona fide and morphed-criminal attempts where the live image belongs to the criminal subject (e.g., Figs. 5(c)–5(b)).

3) Both: In this case, this benchmark is obtained through a union of the previous ones and it is useful to understand the generalization capabilities of the MAD method. It is worth noting the combination of the two sets of previous attempts allows for evaluating the global performance of the system since in real-world scenarios both types of attempts can occur and the system, of course, is not aware of the image pair type received in input.

In the subsequent analysis, particular emphasis is placed on the performance achieved in the "accomplice" scenario. As previously mentioned, due to the greater resemblance between the subjects depicted in both images, this scenario is generally regarded as more challenging for D-MAD methods based on identity analysis.

Finally, it is important to note that the "criminal" case represents one of the common benchmarks for new D-MAD methods, so these results can be used as a reference for previous work.

4.2 Datasets

Progressive morphing database (PMDB)[14]. A total of 1 108 morphed images are obtained by utilizing three commonly used datasets in the MAD field, namely AR[47], face recognition grand challenge (FRGC)[48], and Color FERET[49]. These images have been generated using a publicly available morphing algorithm outlined in a previous study[14]. The morphing process involved a cohort of 280 individuals, consisting of 134 males and 146 females. It is noteworthy that these morphed images have not been subjected to manual retouching procedures to enhance their visual quality, thus potentially exhibiting artifacts such as blurred areas or ghost effects. However, it is important to mention that the background replacement performed by the morphing is artifact-free.

Idiap morph[50] is a collection of multiple datasets publicly accessible, comprising five subsets generated using different morphing algorithms. In our analysis, we focus on OpenCV, FaceMorpher, and StyleGAN[51]. These algorithms utilize face images from the FERET, FRGC, and face research lab London set[52] datasets as input data. The overall visual quality of the morphed images created with OpenCV and FaceMorpher is negatively affected by the presence of various artifacts present in both the background and the foreground of the images. Differently, morphed faces generated through StyleGAN exhibit fewer visible visual artifacts, but common textures associated with GANs are still discernible.

MorphDB[14]. This dataset is constructed using images sourced from the Color FERET[49] and FRGC[48] datasets. It comprises 100 morphed images generated through the sqirlz morph 2.1 algorithm. This dataset offers valuable material as all morphed images have undergone manual retouching, resulting in good final visual quality.

FEI Morph. This dataset has been generated using images from the FEI face database[53], which consists of 200 subjects evenly divided between males and females. The faces within the database predominantly represent individuals aged between 19 and 40 years old, showcasing distinct appearances, hairstyles, and accessories. This dataset comprises a total of 6 000 morphed images, generated through the utilization of three different morphing algorithms: FaceFusion[5], University of Twente (UTW)[4], and Norwegian University of Science and Technology (NTNU)[4]. These algorithms have been used with two different morphing factors, specifically 0.3 and 0.5. The need to introduce this new dataset arises from the necessity to faithfully replicate the new scenario introduced, in which, specifically, the morphed image appears particularly similar to the accomplice and the goal is to detect morphing even when the live acquisition comes from the most similar subject. The dataset is publicly released[6].

4.3 Experimental protocol

In all our experiments, we conduct a cross-dataset evaluation to assess the effectiveness of our method even in the presence of possible new morphing algorithms. In-

[5] http://www.wearemoment.com/FaceFusion/

[6] https://miatbiolab.csr.unibo.it/fei-morph-dataset

deed, the training and validation of the proposed approach are performed on progressive morphing database (PMDB), MorphDB, and Idiap morph datasets, while the proposed MAD method is tested on the FEI morph dataset, in a cross-database and cross-algorithm evaluation: Indeed, the FEI morph dataset is completely disjoint from the training ones. In other words, the testing set has no common elements with the training datasets, neither in terms of contributing subjects (for the morphing generation) nor in terms of the morphing algorithms. It is important to note that the test on the FEI morph dataset is fully reproducible for future comparisons since the FEI face database[53] is publicly available as well as the morphed images used in our experiments. It is important to note that a cross-dataset evaluation procedure is relevant due to the scarcity of publicly available datasets with diverse samples for each subject and morphing algorithms. Besides, privacy issues play a crucial role in hindering the public release of such datasets.

As competitors, we select the state-of-the-art approaches following the ranking certified by the independent platforms NIST FATE-MORPH[7] and FVC-onGoing[8] for the D-MAD task. Thus, despite being limited in number and a few years old, they are highly representative as competitors in the literature. Finally, we also include competitors not based on neural networks, in order to variegate the final analysis.

4.4 Metrics

In the evaluation and comparison of MAD systems, various metrics are typically employed to assess their performance[4]. Two commonly used metrics are referred to as bona fide presentation classification error rate (BPCER) and the attack presentation classification error rate (APCER). The BPCER measures the proportion of bona fide images that are incorrectly classified:

$$\mathrm{BPCER}(\tau) = \frac{1}{N}\sum_{i=1}^{N} H(b_i - \tau). \tag{2}$$

Conversely, the APCER represents the proportion of morphed images that are erroneously labeled as bona fide:

$$\mathrm{APCER}(\tau) = 1 - \left[\frac{1}{M}\sum_{i=1}^{M} H(m_i - \tau)\right]. \tag{3}$$

In both definitions, τ represents the score threshold at which the detection scores for bona fide and morphed images (b_i, m_i) are compared. The function $H(x)$ is defined as a step function, which returns 1 if x is greater than 0 and 0 otherwise. The BPCER is typically evaluated with respect to a specified APCER value, here referred to as $\mathbf{B}_{0.05}$ and $\mathbf{B}_{0.01}$. These values correspond to the lowest BPCER achievable while maintaining an APCER of 10% and 5%, respectively. In an ideal scenario, an MAD algorithm deployed in a real-world setting would aim to achieve a low APCER (allowing only a minimal number of criminals to bypass detection) of around 0.1%, while simultaneously maintaining an acceptable BPCER (generating few false positives) of approximately 5%[54].

The equal error rate (EER) is a commonly reported metric, representing the error rate at which the BPCER and APCER are equal. It is typically presented as a single value, providing a summary measure of the system's performance. Besides, APCER and BPCER metrics can be condensed in the detection error trade-off (DET) curve, reported as well to improve the understanding and the comparison of the experimental analysis.

In this paper, we introduce and exploit the metric weighted average error metric (WAE) in order to summarize all the performance indicators in a single value (as done in Fig. 2). This metric is formally defined as:

$$\mathrm{WAE} = w_E E^{\mathrm{T}} \tag{4}$$

where E is the set of error-based metric values $E = [\mathrm{EER}, \mathbf{B}_{0.1}, \mathbf{B}_{0.05}, \mathbf{B}_{0.01}]$ and $w_E = [0.3, 0.1, 0.2, 0.4]$. These weights are chosen by assigning the majority of the weight to the most common real-world operating point (i.e., $\mathbf{B}_{0.01}$), followed by the EER, as it is useful for evaluating the performance of the system at a glance, and finally the other two chosen operating points (i.e., $\mathbf{B}_{0.05}$ and $\mathbf{B}_{0.1}$).

4.5 Training procedure

In the IdA module, the MLP has an architecture composed of 3 hidden layers of size 250, 125 and 64, with ReLU activation and a single output neuron with a sigmoid function. For the training, we adopt the Adam[55] optimizer with an initial learning rate of 10^{-5}. For the training of the Inception-ResNet, we adopt the stochastic gradient descent (SGD) optimizer with a learning rate of 10^{-3} and an early-stopping procedure (patience of 5 epochs with a minimum improvement of 10^{-4}). No momentum decay is exploited.

In the other modules, SVM classifiers share the same details: They implement the radial basis function (RBF) kernel, and are trained with the regularization parameter $C = 1.0$, and the kernel coefficient $\gamma = 10^{-3}$.

4.6 Results

The results of the proposed MAD method compared with the literature are reported in Table 1, while the related DET curve is reported in Fig. 6.

As shown, our method is able to overcome the com-

[7] https://pages.nist.gov/frvt/html/frvt_morph.html

[8] https://biolab.csr.unibo.it/fvcongoing/UI/Form/Home.aspx

Table 1 Morphing detection scores obtained on the FEI morph test dataset across different differential MAD (D-MAD) competitors presented in the literature. Results are reported in terms of equal error rate (EER), the lowest BPCER related to APCER $\leq 1\%$ and $\leq 5\%$, respectively (see Section 4.4). As reported, the proposed method achieves the best accuracy across all the competitors both in the "accomplice" and in the "criminal" benchmarks. The generalization capabilities are confirmed in the "both" case, in which our model effectively handles different attempt typologies.

Methods	Year	Accomplice			Criminal			Both		
		EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$	EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$	EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$
Demorphing[14]	2017	0.160	0.377	0.618	0.027	0.015	0.050	0.111	0.235	0.522
SVM+LBP[24]	2018	0.200	0.327	0.623	0.185	0.335	0.697	0.192	0.330	0.628
DFR[6]	2020	0.178	0.482	0.770	0.068	0.090	0.295	0.128	0.347	0.690
Siamese[8]	2021	0.153	0.347	0.563	0.061	0.095	0.370	0.115	0.257	0.515
MagFace[7]	2023	0.112	0.230	0.465	0.028	0.013	0.052	0.083	0.135	0.393
ACIdA (ours)	2023	**0.102**	**0.192**	**0.333**	**0.023**	**0.010**	**0.027**	**0.070**	**0.105**	**0.280**

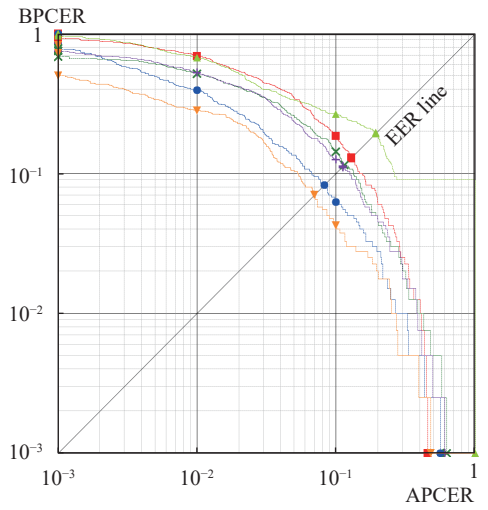


Fig. 6 Detection error trade-off (DET) curves computed on the FEI morph dataset considering several literature competitors. Competitor reported: ACIdA in orange, Meng et al.[18] in blue, Deng et al.[33] in red, Borghi et al.[8] in dark green, Scherhag et al.[24] in light green and Ferrara et al.[14] in purple. (Colored figures are available in the online version at https://link.springer.com/journal/11633)

petitors in all the investigated benchmarks. In particular, our method not only achieves better performance in comparison with D-MAD based only on identity comparison[6, 7, 14] but also with methods that include the analysis of artifacts introduced by the morphing procedure on input images[8, 24]. The results show that the error rates were higher in the accomplice scenario, indicating that this scenario presents unique challenges that require further investigation in future D-MAD studies. The proposed method demonstrated excellent accuracy in the "criminal" benchmarks, confirming its effectiveness in a more conventional D-MAD setting. Furthermore, the "both" case results suggest that incorporating an attempt classification module, which can apply different solutions for criminals and accomplices, is a crucial factor in improving overall performance, as explored in the subsequent analysis.

In the second part of our analysis, we test how the performance of the proposed system is influenced when two different score fusion techniques are employed to create the final MAD score, namely "weighted" and "selection". The first strategy produces the final morphing score by summing the ones produced by the different modules, each weighted by the probabilities returned by the attempt classification module, as described in (1) and depicted in Fig. 4; the "selection" approach directly returns the score produced by the module whose associated probability computed by the attempt classification module is the highest, as shown in (5), where $p_{\max} = \max(p_A, p_B, p_C)$.

$$S = \begin{cases} S_{Id}, & \text{if } p_B = p_{\max} \vee p_C = p_{\max} \\ S_{IdA}, & \text{if } p_A = p_{\max}. \end{cases} \tag{5}$$

Experimental results show that the weighted strategy is overall more effective than employing the selection fusion technique, respectively totaling on the global test set EER $= 0.070$ versus 0.076, $\mathbf{B}_{0.05} = 0.105$ versus 0.125, and $\mathbf{B}_{0.01} = 0.280$ versus 0.385. Therefore, the weighted sum fusion strategy is adopted in our framework.

In Table 2, we include an analysis of the feature embedding source: Indeed, several works on the face recognition task have been recently introduced in the literature[18, 33, 57, 58], constantly improving the accuracy, and can be exploited in our proposed MAD pipeline. In particular, we focus our analysis on recent and state-of-the-art methods. According to the analysis provided by the Deepface framework[9], we select best-performing algorithms on the labeled faces in the wild[59] dataset. In addition, we include in our analysis the recently introduced and promising MagFace[18] approach. It is important to note that one of the top-performing methods, ArcFace[33], has already been effectively utilized in the field of D-MAD[6]. As reported, MagFace achieves the best performance, confirming that superior accuracy in the face recognition task leads to better performance also in the MAD task, especially if based on the identity compar-

[9] https://github.com/serengil/deepface

Table 2 Morphing detection scores obtained by the proposed system on the FEI morph dataset using different feature embeddings originally developed for the face recognition task. In bold the best results, underlined the second ones.

Feat. Embed.	Accomplice			Criminal			Both		
	EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$	EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$	EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$
ArcFace[33]	<u>0.115</u>	<u>0.255</u>	<u>0.507</u>	<u>0.056</u>	<u>0.063</u>	<u>0.115</u>	<u>0.093</u>	<u>0.142</u>	<u>0.385</u>
DLib[56]	0.176	0.375	0.587	0.103	0.155	0.240	0.148	0.255	0.517
SFace[57]	0.213	0.480	0.640	0.102	0.167	0.327	0.166	0.363	0.585
Facenet[58]	0.129	0.308	0.510	0.061	0.075	0.142	0.097	0.175	0.435
MagFace[18]	**0.102**	**0.192**	**0.333**	**0.023**	**0.010**	**0.027**	**0.070**	**0.105**	**0.280**

Table 3 Ablation analysis results obtained on the FEI morph dataset. As reported, the single modules of the proposed system – IdA (Section 3.2) and Id (Section 3.3) – are separately tested. In addition, we compute the results obtained using only the Inception-ResNet architecture of the IdA module.

Module		Accomplice				Criminal				Both			
IdA	Id	EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$	$\bar{\Delta}_E$	EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$	$\bar{\Delta}_E$	EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$	$\bar{\Delta}_E$
√	√	0.102	0.192	0.333		0.023	0.010	0.027		0.070	0.105	0.280	
√		0.120	0.207	0.455	–16%	0.117	0.195	0.415	–90%	0.120	0.203	0.440	–42%
√*		0.125	0.230	0.480	–22%	0.125	0.230	0.480	–91%	0.125	0.230	0.480	–47%
	√	0.112	0.230	0.465	–18%	0.028	0.013	0.052	–30%	0.083	0.135	0.393	–22%

* Only the Inception-ResNet architecture of IdA module is tested (see Fig. 4).

ison.

In Table 3, the ablation study of the proposed method is reported. In particular, we test the performance of the system using only the IdA or the Id module and exploiting the previously determined best feature (i.e., MagFace[18]). We also report the results exploiting only the Inception-ResNet architecture, which computes the morphing score relying only on the document image, in an S-MAD fashion. In addition to the standard error metrics, we summarize the relative error rate reduction achieved by the proposed system with respect to each single module. This metric is reported with the symbol $\bar{\Delta}_E$, and is formally computed as:

$$\Delta_E = \frac{E_{\text{ACIdA}} - E_M}{E_M} \tag{6}$$

where M is the module Id, or IdA or the Inception-ResNet architecture of IdA. The Δ_E value is computed for the three error indicators (EER, $\mathbf{B}_{0.05}$, $\mathbf{B}_{0.01}$) and then averaged to obtain $\bar{\Delta}_E$. Experimental results validate that the implementation of both modules is crucial for achieving optimal performance. As anticipated, the second module, i.e., Id, demonstrates a notable capability in detecting morphed images within the "criminal" benchmark. Meanwhile, the IdA module yields noteworthy results, particularly in terms of BPCERs ($\mathbf{B}_{0.05}$, $\mathbf{B}_{0.01}$), in the "accomplice" benchmark. From a general point of view, the whole ACIdA system achieves a noticeable reduction of the error, as expressed by the metric $\bar{\Delta}_E$. The error reduction is noticeable in the "criminal" benchmark (up to 90%) where very low error rates are measured for ACIdA (EER = 2.3%). In the "accomplice" benchmark, we note a consistent error reduction of about 20%, leading to an overall performance of about –37% in the "both" benchmark.

In conclusion of our experimental analysis, we focus our investigation on the attempt classification (AC) module.

In the first experiment, reported in Table 4, we test a variety of different classifiers commonly used in biometrics. Interestingly, experimental results reveal that the SVM is the best choice, also with respect to deep learning-based solutions (i.e., MLP).

Table 4 Comparison of several classifiers of the attempt classification (AC) module, using as input feature embeddings extracted using the method described in [18].

Classifier	Accuracy	F1-score
SVM	0.650	0.634
Random forest	0.575	0.554
AdaBoost	0.603	0.580
KNN (K=5)	0.575	0.545
Decision tree	0.572	0.554
MLP	0.639	0.634

In addition, we analyze the impact of the classification accuracy of the adopted SVM classifier on the whole MAD pipeline. Indeed, in order to have an upper bound result, we replace in the AC module, the SVM classifier with an oracle, i.e., a classifier is able to perfectly predict the classes – criminal accomplice and bona fide – of the

Table 5 Classification performance and morphing detection scores obtained on the FEI morph dataset. In particular, we highlight the impact of the attempt classification accuracy on the whole proposed system, with respect to the performance of an "oracle" classifier reported on the top lines.

Classification attempt (CA)			Morphing attack detection (IdA - Id)								
Classifier	Accuracy	BF to	Accomplice			Criminal			Both		
			EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$	EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$	EER	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$
Oracle	100.0%	IdA	0.120	0.207	0.455	0.208	0.522	0.625	0.146	0.480	0.598
		Id	0.006	0.002	0.005	0.028	0.013	0.052	0.018	0.005	0.040
SVM	66.9%	IdA	0.171	0.495	0.560	0.046	0.027	0.555	0.118	0.445	0.555
		Id	**0.102**	**0.192**	**0.333**	**0.023**	**0.010**	**0.027**	**0.070**	**0.105**	**0.280**

input attempts. In addition, we analyze how the overall performance of the system is affected when the morphing score of bona fide pairs is computed using S_{IdA} instead of S_{Id} (see (1)).

Results are reported in Table 5.

Firstly, the accuracy of the SVM classifier with respect to the oracle reveals that there is a margin for improvement in the attempt classification task. Closing this gap can significantly enhance the overall performance, as demonstrated by the impressive MAD results reported in the oracle case, with an $\mathrm{EER} = 0.006$.

Secondly, it is possible to note that a substantial improvement in the overall MAD performance is achieved by properly selecting the module used in the weighted average score computation for bona fide images. Specifically, the use of the Id module with bona fide images represents the best solution, since analyzing artifacts produced by the morphing procedure is not useful – bona fide images are free of visible or not visible artifacts – and it could even be counterproductive. Indeed, in the MAD literature, the S-MAD task, and specifically the artifact detection task, is more challenging with respect to the D-MAD one[4].

Finally, the proposed system is evaluated on the FVC-onGoing platform. However, publicly available platforms like NIST FATE MORPH and FVC-onGoing do not provide a suitable testing environment for the scenario described in this paper. This is because they typically do not consider the accomplice′s attempts when the morphed image is heavily biased towards the accomplice, and they do not present the results for the criminal and accomplice attempts separately. Results are reported in Table 6: We observe that the proposed method is able to achieve good accuracy in any case, in particular with respect to the previous work[16]. In this criminal-based scenario, DFR[6] remains the preferable option to use to achieve good performance, since it is based on highly effective identity features when the comparison involves only the criminal or the morphed images equally represents the two contributing subjects (i.e., the morphing factor is 0.5).

Finally, we draw some analysis of the computing performance of the proposed system. ACIdA is able to process images in about 1.8 seconds using about 3.8 GB of RAM. This inference time is obtained in a computer equipped with an AMD EPYC 7 282 processor and using a single Quadro RTX 5 000 GPU; the operating system is virtualized on a virtual machine using only 8 cores of the available 16. We observe the processing time and resource requirements are in line with real-world scenarios, such as automated board control gates in international airports.

Table 6 Comparison of the results on the sequestered DMAD-SOTAMD_D-1.0 benchmark through the FVC-onGoing platform. As shown, the proposed method is able to generalize on a different scenario, obtaining a good accuracy w.r.t. the other D-MAD methods available in the literature.

Algorithms	EER	$\mathbf{B}_{0.1}$	$\mathbf{B}_{0.05}$	$\mathbf{B}_{0.01}$
DFR[6]	4.54	2.00	3.93	18.87
Demorphing[14]	14.17	17.20	22.77	65.57
Siamese[8]	23.37	35.03	48.97	93.60
MBLBP[60]	33.47	52.80	59.93	74.80
WL[61]	37.13	71.67	83.27	95.67
BSIF[21]	45.93	78.30	84.13	93.83
DN[62]	52.03	89.70	94.70	98.57
Laplace[63]	55.13	96.70	98.67	99.87
R-DMAD[16]	10.23	10.33	19.67	47.47
ACIdA	7.84	7.57	12.60	26.23

5 Conclusions

ACIdA, a novel modular D-MAD method was proposed, and extensive experimental validation demonstrated its ability to achieve state-of-the-art results, surpassing competing methods in the investigated scenario. These results suggest that an analysis of possible morphing artifacts in combination with the use of identity features can increase the robustness of D-MAD approaches when dealing with a high subject similarity. ACIdA has been tested in a new scenario, that aims to expand the scope of application of current D-MAD systems, by dealing with both comparisons of the trusted live images with the criminal and the accomplice. Our preliminary analys-

is indicates that this scenario is challenging and raises new interesting research aspects in the D-MAD field, especially for identity-based systems. The findings of this study contribute to the advancement of MAD systems, enhancing the security and reliability of FRS in the face of morphing attacks. Further research can build upon these results to develop even more robust and efficient MAD systems in the future. In particular, new experimentation and research are needed to improve the classification accuracy of the attempt that, as shown, is a key element for the final performance of the proposed system.

It is worth noting that unfortunately, the publicly available datasets for the MAD task are limited in number in the literature, even due to privacy concerns[64], and this entails significant difficulties for the aforementioned analyses. The existing datasets, indeed, are generally characterized by limited variations in terms of environmental factors (i.e., lighting conditions and background scenarios), and some demographic and ethnic groups are not sufficiently represented in the data. This lack of diversity and representativeness in the available datasets poses challenges for the development and evaluation of robust and unbiased MAD models, as they may not capture the full range of real-world scenarios and individual variations.

Therefore, more extensive and diverse data collection efforts will need to be undertaken in this direction to address such critical shortcomings. Researchers may need to explore innovative approaches to data acquisition, such as collaborating with various stakeholders, leveraging crowdsourcing platforms, or using synthetic data[65, 66] generation techniques, while ensuring the ethical and privacy-preserving handling of sensitive information. The development of benchmark datasets that better reflect the diversity of human populations and environmental conditions would greatly benefit the advancement of the MAD field, enabling more comprehensive testing and validation of the developed algorithms.

Finally, we propose incorporating the examined scenario into web platforms that offer D-MAD tests. Currently, these platforms often provide tests with only the criminal comparison or fail to directly display the performance of the accomplice, criminal, and combined benchmarks. Expanding these platforms to include a broader range of benchmarks and scenarios, while also providing more transparency and detailed feedback on the model′s performance, would greatly contribute to the understanding and advancement of the MAD task. This could help researchers and developers better evaluate the strengths, limitations, and biases of their approaches, ultimately leading to more robust and reliable malicious activity detection systems.

Acknowledgement

This project was supported by the European Union′s Horizon 2020 research and innovation program, Europe (No. 883356). Disclaimer: this text reflects only the author′s views, and the Commission is not liable for any use that may be made of the information contained therein.

Declarations of conflict of interest

The authors declared that they have no conflicts of interest to this work.

References

[1] M. Ferrara, A. Franco, D. Maltoni. The magic passport. In *Proceedings of IEEE International Joint Conference on Biometrics*, Clearwater, USA, 2014. DOI: 10.1109/BTAS.2014.6996240.

[2] U. Scherhag, A. Nautsch, C. Rathgeb, M. Gomez-Barrero, R. N. J. Veldhuis, L. Spreeuwers, M. Schils, D. Maltoni, P. Grother, S. Marcel, R. Breithaupt, R. Ramachandra, C. Busch. Biometric systems under morphing attacks: Assessment of morphing techniques and vulnerability reporting. In *Proceedings of International Conference of the Biometrics Special Interest Group*, Darmstadt, Germany, 2017. DOI: 10.23919/BIOSIG.2017.8053499.

[3] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, C. Busch. Face recognition systems under morphing attacks: A survey. *IEEE Access*, vol. 7, pp. 23012–23026, 2019. DOI: 10.1109/ACCESS.2019.2899367.

[4] K. Raja, M. Ferrara, A. Franco, L. Spreeuwers, I. Batskos, F. de Wit, M. Gomez-Barrero, U. Scherhag, D. Fischer, S. K. Venkatesh, J. M. Singh, G. Q. Li, L. Bergeron, S. Isadskiy, R. Ramachandra, C. Rathgeb, D. Frings, U. Seidel, F. Knopjes, R. Veldhuis, D. Maltoni, C. Busch. Morphing attack detection-database, evaluation platform, and benchmarking. *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 4336–4351, 2021. DOI: 10.1109/TIFS.2020.3035252.

[5] G. Borghi, G. Graffieti, A. Franco, D. Maltoni. Incremental training of face morphing detectors. In *Proceedings of the 26th International Conference on Pattern Recognition*, Montreal, Canada, pp. 914–921, 2022. DOI: 10.1109/ICPR56361.2022.9956395.

[6] U. Scherhag, C. Rathgeb, J. Merkle, C. Busch. Deep face representations for differential morphing attack detection. *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3625–3639, 2020. DOI: 10.1109/TIFS.2020.2994750.

[7] R. Kessler, K. Raja, J. Tapia, C. Busch. Towards minimizing efforts for morphing attacks – deep embeddings for morphing pair selection and improved morphing attack detection. *PLoS One*, vol. 19, no. 5, Article number e0304610, 2024. DOI: 10.1371/journal.pone.0304610.

[8] G. Borghi, E. Pancisi, M. Ferrara, D. Maltoni. A double Siamese framework for differential morphing attack detection. *Sensors*, vol. 21, no. 10, Article number 3466, 2021. DOI: 10.3390/s21103466.

[9] S. Soleymani, A. Dabouei, F. Taherkhani, J. Dawson, N. M. Nasrabadi. Mutual information maximization on disentangled representations for differential morph detection. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, Waikoloa, USA, pp. 1730–1740, 2021. DOI: 10.1109/WACV48630.2021.00177.

[10] L. Pellegrini, G. Borghi, A. Franco, D. Maltoni. Detecting

morphing attacks via continual incremental training. In *Proceedings of IEEE International Joint Conference on Biometrics*, Ljubljana, Slovenia, 2023. DOI: 10.1109/IJCB57857.2023.10449306.

[11] R. Ramachandra, G. Q. Li. Residual colour scale-space gradients for reference-based face morphing attack detection. In *Proceedings of the 25th International Conference on Information Fusion*, Linköping, Sweden, 2022. DOI: 10.23919/FUSION49751.2022.9841318.

[12] S. Soleymani, B. Chaudhary, A. Dabouei, J. Dawson, N. M. Nasrabadi. Differential morphed face detection using deep Siamese networks. In *Proceedings of the 25th International Conference on Pattern Recognition*, Milano, Italy, pp. 560–572, 2021. DOI: 10.1007/978-3-030-68780-9_44.

[13] T. Neubert, A. Makrushin, M. Hildebrandt, C. Kraetzer, J. Dittmann. Extended *StirTrace* benchmarking of biometric and forensic qualities of morphed face images. *IET Biometrics*, vol. 7, no. 4, pp. 325–332, 2018. DOI: 10.1049/iet-bmt.2017.0147.

[14] M. Ferrara, A. Franco, D. Maltoni. Face demorphing. *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 4, pp. 1008–1017, 2018. DOI: 10.1109/TIFS.2017.2777340.

[15] C. Rathgeb, R. Tolosana, R. Vera-Rodriguez, C. Busch. *Handbook of Digital Face Manipulation and Detection: From DeepFakes to Morphing Attacks*, Cham, Switzerland: Springer, 2022. DOI: 10.1007/978-3-030-87664-7.

[16] N. Di Domenico, G. Borghi, A. Franco, D. Maltoni. Combining identity features and artifact analysis for differential morphing attack detection. In *Proceedings of the 22nd International Conference on Image Analysis and Processing*, Udine, Italy, pp. 100–111, 2023. DOI: 10.1007/978-3-031-43148-7_9.

[17] B. Dorizzi, R. Cappelli, M. Ferrara, D. Maio, D. Maltoni, N. Houmani, S. Garcia-Salicetti, A. Mayoue. Fingerprint and on-line signature verification competitions at ICB 2009. In *Proceedings of the 3rd International Conference on Advances in Biometrics*, Alghero, Italy, pp. 725–732, 2009. DOI: 10.1007/978-3-642-01793-3_74.

[18] Q. Meng, S. C. Zhao, Z. D. Huang, F. Zhou. MagFace: A universal representation for face recognition and quality assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, USA, pp. 14220–14229, 2021. DOI: 10.1109/CVPR46437.2021.01400.

[19] M. O. Kenneth, B. A. Sulaimon, S. M. Abdulhamid, L. C. Ochei. A systematic literature review on face morphing attack detection (MAD). *Illumination of Artificial Intelligence in Cybersecurity and Forensics*, S. Misra, C. Arumugam, Eds., Cham, Switzerland: Springer, pp. 139–172, 2022. DOI: 10.1007/978-3-030-93453-8_7.

[20] M. Hamza, S. Tehsin, M. Humayun, M. F. Almufareh, M. Alfayad. A comprehensive review of face morph generation and detection of fraudulent identities. *Applied Sciences*, vol. 12, no. 24, Article number 12545, 2022. DOI: 10.3390/app122412545.

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

[22] T. Ojala, M. Pietikainen, D. Harwood. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In *Proceedings of the 12th International Conference on Pattern Recognition*, Jerusalem, Israel, pp. 582–585, 1994. DOI: 10.1109/ICPR.1994.576366.

[23] N. Dalal, B. Triggs. Histograms of oriented gradients for human detection. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, USA, pp. 886–893, 2005. DOI: 10.1109/CVPR.2005.177.

[24] U. Scherhag, C. Rathgeb, C. Busch. Towards detection of morphed face images in electronic travel documents. In *Proceedings of the 13th IAPR International Workshop on Document Analysis Systems*, Vienna, Austria, pp. 187–192, 2018. DOI: 10.1109/DAS.2018.11.

[25] B. Chaudhary, P. Aghdaie, S. Soleymani, J. Dawson, N. M. Nasrabadi. Differential morph face detection using discriminative wavelet sub-bands. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Nashville, USA, pp. 1425–1434, 2021. DOI: 10.1109/CVPRW53098.2021.00158.

[26] I. Medvedev, J. A. Pimenta, N. Gonçalves. Fused classification for differential face morphing detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision Workshops*, Waikoloa, USA, pp. 1043–1050, 2024. DOI: 10.1109/WACVW60836.2024.00114.

[27] E. Shiqerukaj, C. Rathgeb, J. Merkle, P. Drozdowski, B. Tams. Fusion of face demorphing and deep face representations for differential morphing attack detection. In *Proceedings of International Conference of the Biometrics Special Interest Group*, Darmstadt, Germany, 2022. DOI: 10.1109/BIOSIG55365.2022.9897023.

[28] L. Qin, F. Peng, M. Long. Face morphing attack detection and localization based on feature-wise supervision. *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 3649–3662, 2022. DOI: 10.1109/TIFS.2022.3212276.

[29] J. M. Singh, R. Ramachandra. Reliable face morphing attack detection in on-the-fly border control scenario with variation in image resolution and capture distance. In *Proceedings of IEEE International Joint Conference on Biometrics*, Abu Dhabi, UAE, 2022. DOI: 10.1109/IJCB54206.2022.10007987.

[30] U. Scherhag, D. Budhrani, M. Gomez-Barrero, C. Busch. Detecting morphed face images using facial landmarks. In *Proceedings of the 8th International Conference on Image and Signal Processing*, Cherbourg, France, pp. 444–452, 2018. DOI: 10.1007/978-3-319-94211-7_48.

[31] S. Autherith, C. Pasquini. Detecting morphing attacks through face geometry features. *Journal of Imaging*, vol. 6, no. 11, Article number 115, 2020. DOI: 10.3390/jimaging6110115.

[32] K. M. He, X. Y. Zhang, S. Q. Ren, J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA, pp. 770–778, 2016. DOI: 10.1109/CVPR.2016.90.

[33] J. K. Deng, J. Guo, N. N. Xue, S. Zafeiriou. ArcFace: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, USA, pp. 4685–4694, 2019. DOI: 10.1109/CVPR.2019.00482.

[34] S. Banerjee, A. Ross. Conditional identity disentanglement for differential face morph detection. In *Proceedings of IEEE International Joint Conference on Biometrics*,

Shenzhen, China, 2021. DOI: 10.1109/IJCB52358.2021.9484355.

[35] F. Peng, L. B. Zhang, M. Long. FD-GAN: Face de-morphing generative adversarial network for restoring accomplice's facial image. *IEEE Access*, vol. 7, pp. 75122–75131, 2019. DOI: 10.1109/ACCESS.2019.2920713.

[36] S. Banerjee, P. Jaiswal, A. Ross. Facial de-morphing: Extracting component faces from a single morph. In *Proceedings of IEEE International Joint Conference on Biometrics*, Abu Dhabi, UAE, 2022. DOI: 10.1109/IJCB54206.2022.10007977.

[37] D. Ortega-Delcampo, C. Conde, D. Palacios-Alonso, E. Cabello. Border control morphing attack detection with a convolutional neural network demorphing approach. *IEEE Access*, vol. 8, pp. 92301–92313, 2020. DOI: 10.1109/access.2020.2994112.

[38] J. Cai, Q. Q. Duan, M. Long, L. B. Zhang, X. L. Ding. Feature interaction-based face de-morphing factor prediction for restoring accomplice's facial image. *Sensors*, vol. 24, no. 17, Article number 5504, 2024. DOI: 10.3390/S24175504.

[39] M. Long, Q. T. Yao, L. B. Zhang, F. Peng. Face demorphing based on diffusion autoencoders. *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 3051–3063, 2024. DOI: 10.1109/TIFS.2024.3359029.

[40] M. Long, Q. Q. Duan, L. B. Zhang, F. Peng, D. Y. Zhang. Trans-FD: Transformer-based representation interaction for face de-morphing. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 6, no. 3, pp. 385–397, 2024. DOI: 10.1109/TBIOM.2024.3390056.

[41] K. P. Zhang, Z. P. Zhang, Z. F. Li, Y. Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016. DOI: 10.1109/LSP.2016.2603342.

[42] I. C. Duta, L. Liu, F. Zhu, L. Shao. Improved residual networks for image and video recognition. In *Proceedings of the 25th International Conference on Pattern Recognition*, Milan, Italy, pp. 9415–9422, 2021. DOI: 10.1109/ICPR48806.2021.9412193.

[43] Y. D. Guo, L. Zhang, Y. X. Hu, X. D. He, J. F. Gao. MS-celeb-1M: A dataset and benchmark for large-scale face recognition. In *Proceedings of the 14th European Conference on Computer Vision*, Amsterdam, The Netherlands, pp. 87–102, 2016. DOI: 10.1007/978-3-319-46487-9_6.

[44] Q. Cao, L. Shen, W. D. Xie, O. M. Parkhi, A. Zisserman. VGGFace2: A dataset for recognising faces across pose and age. In *Proceedings of the 13th IEEE International Conference on Automatic Face & Gesture Recognition*, Xi'an, China, pp. 67–74, 2018. DOI: 10.1109/FG.2018.00020.

[45] C. Szegedy, W. Liu, Y. Q. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich. Going deeper with convolutions. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Boston, USA, 2015. DOI: 10.1109/CVPR.2015.7298594.

[46] G. Borghi, N. Di Domenico, A. Franco, M. Ferrara, D. Maltoni. Revelio: A modular and effective framework for reproducible training and evaluation of morphing attack detectors. *IEEE Access*, vol. 11, pp. 120419–120437, 2023. DOI: 10.1109/ACCESS.2023.3328227.

[47] A. Martinez, R. Benavente. The AR face database: CVC technical report, 24, 1998. [Online], Available: https://portalrecerca.uab.cat/en/publications/the-ar-face-database-cvc-technical-report-24

[48] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek. Overview of the face recognition grand challenge. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, USA, pp. 947–954, 2005. DOI: 10.1109/CVPR.2005.268.

[49] P. J. Phillips, H. Wechsler, J. Huang, P. J. Rauss. The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, vol. 16, no. 5, pp. 295–306, 1998. DOI: 10.1016/S0262-8856(97)00070-X.

[50] E. Sarkar, P. Korshunov, L. Colbois, S. Marcel. Are GAN-based morphs threatening face recognition? In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, Singapore, pp. 2959–2963, 2022. DOI: 10.1109/ICASSP43922.2022.9746477.

[51] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, T. Aila. Analyzing and improving the image quality of StyleGAN. In *Proceedings of IEEE/CVF conference on computer vision and pattern recognition*, Seattle, USA, pp. 8107–8116, 2020. DOI: 10.1109/CVPR42600.2020.00813.

[52] L. DeBruine, B. Jones. Face research lab London set, 2017. https://doi.org/10.6084/m9.figshare.5047666.v5

[53] C. E. Thomaz, G. A. Giraldi. A new ranking method for principal components analysis and its application to face image analysis. *Image and Vision Computing*, vol. 28, no. 6, pp. 902–913, 2010. DOI: 10.1016/j.imavis.2009.11.005.

[54] M. Ferrara, A. Franco. Morph creation and vulnerability of face recognition systems to morphing. *Handbook of Digital Face Manipulation and Detection: From DeepFakes to Morphing Attacks*, C. Rathgeb, R. Tolosana, R. Vera-Rodriguez, C. Busch, Eds., Cham, Switzerland: Springer, pp. 117–137, 2022. DOI: 10.1007/978-3-030-87664-7_6.

[55] D. P. Kingma, J. Ba. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations*, San Diego, USA, 2015. DOI: 10.48550/arXiv.1412.6980.

[56] D. E. King. Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.

[57] Y. Y. Zhong, W. H. Deng, J. N. Hu, D. Y. Zhao, X. Li, D. C. Wen. SFace: Sigmoid-constrained hypersphere loss for robust face recognition. *IEEE Transactions on Image Processing*, vol. 30, pp. 2587–2598, 2021. DOI: 10.1109/TIP.2020.3048632.

[58] F. Schroff, D. Kalenichenko, J. Philbin. FaceNet: A unified embedding for face recognition and clustering. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Boston, USA, pp. 815–823, 2015. DOI: 10.1109/CVPR.2015.7298682.

[59] G. B. Huang, M. Mattar, T. Berg, E. Learned-Miller. Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. In *Proceedings of Workshop on faces in 'Real-Life' Images: Detection, Alignment, and Recognition*, Marseille, France, 2008.

[60] U. Scherhag, R. Ramachandra, K. B. Raja, M. Gomez-Barrero, C. Rathgeb, C. Busch. On the vulnerability of face recognition systems towards morphed face attacks. In *Proceedings of the 5th International Workshop on Biometrics and Forensics*, Coventry, UK, 2017. DOI: 10.1109/IWBF.2017.7935088.

[61] N. Damer, V. Boller, Y. Wainakh, F. Boutros, P. Terhörst,

A. Braun, A. Kuijper. Detecting face morphing attacks by analyzing the directed distances of facial landmarks shifts. In *Proceedings of the 40th German Conference on Pattern Recognition*, Stuttgart, Germany, pp. 518–534, 2019. DOI: 10.1007/978-3-030-12939-2_36.

[62] J. M. Singh, R. Ramachandra, K. B. Raja, C. Busch. Robust morph-detection at automated border control gate using deep decomposed 3D shape & diffuse reflectance. In *Proceedings of the 15th International Conference on Signal-Image Technology & Internet-Based Systems*, Sorrento, Italy, pp. 106–112, 2019. DOI: 10.1109/SITIS.2019.00028.

[63] C. Kraetzer, A. Makrushin, T. Neubert, M. Hildebrandt, J. Dittmann. Modeling attacks on photo-ID documents and applying media forensics for the detection of facial morphing. In *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security*, Philadelphia, USA, pp. 21–32, 2017. DOI: 10.1145/3082031.3083244.

[64] M. Robledo-Moreno, G. Borghi, N. Di Domenico, A. Franco, K. Raja, D. Maltoni. Towards federated learning for morphing attack detection. In *Proceedings of IEEE International Joint Conference on Biometrics*, Buffalo, USA, 2024. DOI: 10.1109/IJCB62174.2024.10744518.

[65] P. Melzi, R. Tolosana, R. Vera-Rodriguez, M. Kim, C. Rathgeb, X. M. Liu, I. DeAndres-Tame, A. Morales, J. Fierrez, J. Ortega-Garcia, W. S. Zhao, X. Y. Zhu, Z. Y. Yan, X. Y. Zhang, J. L. Wu, Z. Lei, S. Tripathi, M. Kothari, M. H. Zama, D. Deb, B. Biesseck, P. Vidal, R. Granada, G. Fickel, G. Führ, D. Menotti, A. Unnervik, A. George, C. Ecabert, H. O. Shahreza, P. Rahimi, S. Marcel, I. Sarridis, C. Koutlis, G. Baltsou, S. Papadopoulos, C. Diou, N. Di Domenico, G. Borghi, L. Pellegrini, E. Mas-Candela, Á. Sánchez-Pérez, A. Atzori, G. Fenu, F. Boutros, M. Marras, N. Damer. FRCSyn challenge at WACV 2024: Face recognition challenge in the era of synthetic data. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, Waikoloa, USA, pp. 892–901, 2024. DOI: 10.1109/WACVW60836.2024.00100.

[66] G. Tarollo, T. Fontanini, C. Ferrari, G. Borghi, A. Prati. Adversarial identity injection for semantic face image synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, USA, pp. 1471–1480, 2024. DOI: 10.1109/CVPRW63382.2024.00154.

Nicolò Di Domenico received the B. Sc. and the M. Sc. degrees in computer science from University of Bologna, Italy in 2020 and 2023, respectively. He is a Ph.D. degree candidate from the University of Bologna, Italy. After his graduation, he is currently part of the Biometric Systems Laboratory at the University of Bologna, Italy, under the supervision of Prof. Davide Maltoni.

His research interests include biometrics, face analysis, and morphing attack detection.

E-mail: nicolo.didomenico@unibo.it
ORCID iD: 0009-0006-0223-1680

Guido Borghi received the M. Sc. degree in computer engineering and the Ph. D. degree in information and communication technologies from the University of Modena and Reggio Emilia, Italy in 2015 and 2019, respectively. He is an associate professor within the University of Modena and Reggio Emilia, Italy.

His research interests include computer vision and deep learning techniques applied to intensity and depth images for face analysis, biometrics, driver monitoring and human computer interaction.

E-mail: guido.borghi@unimore.it
ORCID iD: 0000-0003-2441-7524

Annalisa Franco received the Ph. D. degree in electronics, computer science and telecommunications engineering, University of Bologna, Italy in 2004, for her work on multidimensional indexing structures and their application in pattern recognition. She is an associate professor at the Department of Computer Science and Engineering, University of Bologna, Italy. She is a member of the Biometric System Laboratory at Computer Science - Cesena. She authored several scientific papers and served as referee for a number of international journals and conferences.

Her research interests include pattern recognition, biometric systems, image databases and multidimensional data structures. Recent research activity is mainly focused on face recognition in the context of electronic identity documents.

E-mail: annalisa.franco@unibo.it (Corresponding author)
ORCID iD: 0000-0002-6625-6442

Davide Maltoni is a full professor at University of Bologna (Dept. of Computer Science and Engineering - DISI). He is co-director of the Biometric Systems Laboratory (BioLab), which is internationally known for its research and publications in the field. Several original techniques have been proposed by BioLab team for fingerprint feature extraction, matching and classification, for hand shape verification, for face location and for performance evaluation of biometric systems. Davide Maltoni is co-author of the Handbook of Fingerprint Recognition published by Springer, 2009 and holds three patents on Fingerprint Recognition. He has been elected IAPR (International Association for Pattern Recognition) Fellow 2010.

His research interests include pattern recognition, computer vision, machine learning and computational neuroscience.

E-mail: davide.maltoni@unibo.it
ORCID iD: 0000-0002-6329-6756