

Face Re-morphing: Differential Morphing Attack Detection via Feature-Space Similarity Changes

Jie Jin¹ Masakatsu Nishigaki¹ Tetsushi Ohki^{1,2}

¹Graduate School of Science and Technology, Shizuoka University ²RIKEN AIP

3-5-1 Johoku, Chuo-ku, Hamamatsu, Shizuoka 432-8011

1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

jie.jin@sec.inf.shizuoka.ac.jp, {nisigaki, ohki}@inf.shizuoka.ac.jp

Abstract

Face morphing attacks pose a serious threat to face recognition systems because a single morphed document image can be matched to multiple contributors. Differential morphing attack detection (D-MAD) addresses this threat by comparing a document image with a trusted live image, but existing methods often rely on static feature differences, constituent-face reconstruction, or multi-cue fusion.

This paper proposes Face Re-morphing, a D-MAD method that uses the feature-space response to an additional morphing operation as a detection cue. Given a document image and a trusted live image, the proposed method generates a re-morphed image and uses the change between the document–live and live–re-morphed cosine similarities as the detection score.

Experiments on FRLL-Morphs and FEI Morph show that the proposed cue is effective across different morphing conditions, re-morphing methods, and face recognition models. Comparisons with existing methods show favorable results on AMSL and indicate that the proposed method performs well under the Criminal condition on FEI Morph Version 1, particularly when using MorDIFF. These results indicate that re-morphing-induced similarity change provides a complementary cue for D-MAD.

1. Introduction

Face recognition systems are widely used for person identification and identity verification, and have been deployed in many security systems and applications. In particular, border control systems require high robustness and reliability, since authentication is performed by referring to facial images stored in passports or visas. However, previous studies have shown that face recognition systems can be bypassed by using morphed face images created from two different individuals, making morphing attacks a seri-

ous threat to face recognition systems [10, 5].

In a morphing attack, an attacker presents a morphed image at enrollment. This image is created by combining the facial images of the attacker and an accomplice. If the morphed image is accepted as a legitimate document image, multiple individuals who contributed to the morph may be successfully matched against the same document image. This means that a passport or identity document, which should correspond to only one person, can effectively be shared by multiple individuals, thereby undermining the one-to-one correspondence between a registered identity and a single subject that underlies border control and identity verification. In particular, an accomplice with no criminal record may obtain a legitimate document using a morphed image, and a different person may later use that document. This poses a serious risk to face-recognition-based operations, including automated border control.

Morphing Attack Detection (MAD) has been studied as a countermeasure against such attacks. MAD can be broadly categorized into Single-image MAD (S-MAD), which determines whether a single document image is morphed, and Differential MAD (D-MAD), which utilizes both the document image and a live image captured in a trusted acquisition environment [20]. Although S-MAD can be applied when only a single image is available, detection may become difficult for high-quality morphed images with few visible artifacts. In contrast, D-MAD can utilize both the document image and the live image, making it suitable for scenarios in which a trusted live image is available.

Existing D-MAD methods have evolved into approaches that directly exploit the difference between the document image and the live image, approaches that make decisions through the reconstruction of constituent faces contained in a morph image, and approaches that integrate multiple cues [27, 11, 3]. However, these methods often depend on how differences are modeled, on reconstruction quality, or on how multiple cues are integrated, which leaves challenges

in achieving stable detection across different morphing conditions.

In this study, we propose Face Re-morphing, a D-MAD method that exploits changes in cosine similarity caused by applying an additional morphing operation between the document image and the live image. Such changes can serve as a discriminative cue because their behavior is expected to differ depending on whether the document image is a bona-fide image or a morphed image. For a bona-fide image, the document image is composed of the facial characteristics of a single subject, and thus the cosine similarity change induced by Re-morphing with the live image is expected to remain within a relatively natural range. In contrast, when the document image is a morph, it already contains facial characteristics from multiple subjects, so the cosine similarity change produced by Re-morphing with the live image is expected to show a different tendency from that observed for bona-fide images. The proposed method utilizes this difference for morphing attack detection.

The main contributions of this paper are twofold. First, we introduce Face Re-morphing as a new detection concept for D-MAD, in which existing face morphing techniques are utilized as part of the detection process. Second, using the cosine similarity change before and after re-morphing as the detection cue, we empirically demonstrate that the similarity change exhibits a consistent discriminative tendency across different morphing methods and feature extractors.

2. Related Work

2.1. Feature Difference-Based D-MAD

Feature difference-based D-MAD approaches identify morphing attacks by measuring differences between the document image and the live image. These methods typically leverage cues including facial landmark misalignments, shape variations, deep embeddings, and identity matching scores.

As a representative example, Scherhag et al. [27] proposed a D-MAD method based on deep face representations. In their method, the difference between deep facial features extracted from the document image and the live image is fed into a classifier to distinguish bona-fide images from morphed ones. In addition, this framework has been extended to address limitations in training data. For example, IDSwapMAD proposed by Banas et al. [1] preserves the differential comparison scheme based on AdaFace features and an SVM. At the same time, it adapts the framework to more privacy-friendly training conditions.

While these methods bypass explicit face reconstruction by directly exploiting recognition feature spaces, their static comparisons remain sensitive to specific models and environmental conditions, often failing to capture morph-specific behaviors.

2.2. Face De-Morphing-Based D-MAD

Face de-morphing-based D-MAD detects morphing attacks by applying reverse morphing or separation processes to the document image and the live image. The decision is then made based on the reconstructed constituent face or the resulting similarity patterns.

A starting point of this line of research was the face de-morphing method proposed by Ferrara et al. [11], which attempted to reconstruct the contributing face by removing the bona-fide component. Following early reconstruction efforts, recent advancements have leveraged deep generative models; for instance, Long et al. [19] improved reconstruction accuracy and image quality using a Diffusion Autoencoder. Furthermore, Hoang et al. [12] addressed unknown mixing conditions by utilizing a sequence of similarity scores obtained with multiple de-morphing factors.

Overall, face de-morphing-based D-MAD represents a major direction in D-MAD research. At the same time, its detection performance tends to depend heavily on reconstruction quality and the choice of de-morphing conditions.

2.3. Hybrid-Based D-MAD

Hybrid-based D-MAD aims at more robust detection by integrating two types of information. One is identity discrepancy obtained from the comparison between the document image and the live image. The other is morphing-related artifacts contained in the document image itself.

Borghi et al. [3] proposed a double Siamese framework. It consists of one branch for extracting identity-related features and another branch for detecting morphing artifacts. In this way, the method jointly exploits differential cues and single-image artifacts. Similarly, Liu et al. [18] combined triplet-based metric learning with an artifact extractor. The former emphasizes discrepancies in the identity feature space. The latter analyzes morphing-related traces. Both types of information are then fused for the final decision.

These methods are effective because they combine feature difference-based D-MAD with artifact analysis. However, their model design tends to be more complex. Their final performance is also strongly influenced by how and when the different sources of information are integrated.

As outlined above, existing D-MAD research has evolved primarily along three directions: difference-based comparison, constituent-face reconstruction, and the integration of multiple cues. In contrast, the method investigated in this study is distinguished by incorporating existing morphing techniques into the detection process itself and by using the change in comparison scores before and after re-morphing between the document image and the live image as a discriminative cue. Accordingly, the proposed method can be positioned as a D-MAD approach that departs from the conventional paradigms of static difference comparison and constituent-face reconstruction.

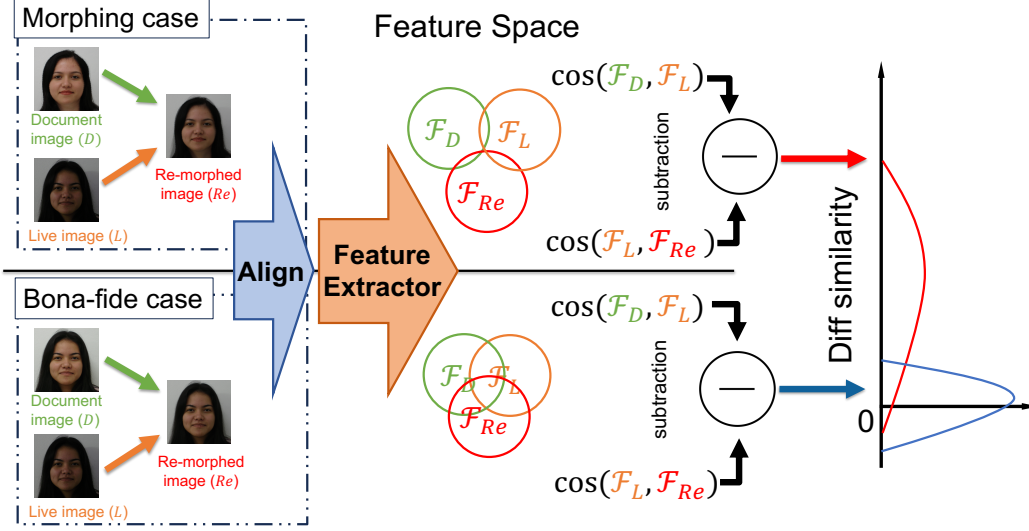


Figure 1. Overview of the proposed method. The proposed framework compares the cosine similarity between the document image and the live image with that between the live image and the re-morphed image, and uses their difference as the detection score.

3. Proposed Method

3.1. Overview

This study is based on the hypothesis that bona-fide and morphed cases exhibit different patterns of facial feature similarity change in the feature space when a re-morphed image is generated by applying an additional morphing process to a document image and a live image. The proposed method determines whether a morphing attack is present by using, as a discriminative cue, the difference between the cosine similarity of the document image and the live image and that of the re-morphed image and the live image. Figure 1 illustrates the overall framework of the proposed method.

3.2. Re-morphed Image Generation

In this study, the re-morphing process refers to applying a morphing operation, within the Face Re-morphing framework, to the document image D to be verified and the live image L acquired in a trusted environment. First, face alignment is applied to D and L , and a re-morphed image Re is generated by applying a morphing technique to the aligned images.

This study adopts three generation methods with different underlying principles for the re-morphing process. MorDIFF is a representation-space morphing technique based on a Diffusion Autoencoder, which generates morphed images by interpolating the latent representations of the input images. OpenCV is a classical image-level morphing method based on corresponding facial landmarks, which synthesizes images through Delaunay-triangulation-based warping and alpha blending. StyleGAN2-ADA is a

GAN-based method that generates morphed images by manipulating latent representations in the GAN latent space. By employing these morphing techniques, we examine how the choice of the re-morphing method affects the proposed detection framework.

3.3. Feature Extraction and Score Computation

The proposed method feeds each preprocessed image into an identical facial feature extractor $\mathcal{F}$ to obtain the corresponding feature vectors $\mathcal{F}_D$, $\mathcal{F}_L$, and $\mathcal{F}_{Re}$. Using these feature vectors, we calculate the baseline cosine similarity s_{base} between D and L , as well as the cosine similarity s_{re} between L and Re , as follows:

$$s_{base} = \cos(\mathcal{F}_D, \mathcal{F}_L), \quad s_{re} = \cos(\mathcal{F}_L, \mathcal{F}_{Re}) \quad (1)$$

The proposed method uses the difference between these two similarities as the final detection score:

$$\Delta s = s_{re} - s_{base} \quad (2)$$

This score Δs represents the amount of variation introduced by the re-morphing process relative to the original similarity between the document image D and the live image L . Based on this difference, the proposed method performs morphing attack detection by using the cosine similarity change induced by the dynamic operation of re-morphing.

3.4. Theoretical Justification

Let $\mathcal{F}_A$ and $\mathcal{F}_C$ denote the facial embeddings of the accomplice and criminal identities, respectively. In the following, we analyze the Criminal scenario where $L = C$

without loss of generality. The same argument applies symmetrically to the Accomplice scenario.

Bona-fide case (D and L belong to the same identity I). The embeddings $\mathcal{F}_D$ and $\mathcal{F}_L$ correspond to two samples of the same identity and are therefore expected to be close in the embedding space, i.e.,

$$\mathcal{F}_D \approx \mathcal{F}_L \approx \mathcal{F}_I.$$

Since modern face recognition models exhibit strong intra-class invariance, the image generated by re-morphing two samples of the same identity is also likely to be recognized as identity I , and thus $\mathcal{F}_{Re}$ is expected to remain near the same identity cluster. Therefore, both $\cos(\mathcal{F}_D, \mathcal{F}_L)$ and $\cos(\mathcal{F}_L, \mathcal{F}_{Re})$ reflect intra-class similarity, yielding

$$\Delta s = \cos(\mathcal{F}_L, \mathcal{F}_{Re}) - \cos(\mathcal{F}_D, \mathcal{F}_L) \approx 0$$

regardless of moderate intra-class variation.

Morphing case (D is a morph of A and C , and $L = C$). The embedding $\mathcal{F}_D$ can be approximated as lying in the region between $\mathcal{F}_A$ and $\mathcal{F}_C$. This property has been empirically observed in prior morphing attack studies, where morph images are matched against both contributing identities at high similarity scores [10, 27].

Consequently, $\cos(\mathcal{F}_D, \mathcal{F}_L)$ reflects a similarity lower than typical intra-class similarity. When re-morphing is applied between D and L , the resulting embedding $\mathcal{F}_{Re}$ tends to move closer to $\mathcal{F}_C$, because L contributes a full identity component while D contains only a partial contribution from C . Therefore,

$$\cos(\mathcal{F}_L, \mathcal{F}_{Re}) > \cos(\mathcal{F}_D, \mathcal{F}_L)$$

typically holds, yielding a positive Δs in many cases, with the magnitude expected to increase as the embedding distance between $\mathcal{F}_A$ and $\mathcal{F}_C$ becomes larger.

The same argument applies symmetrically to the Accomplice scenario where $L = A$. In this case, re-morphing pulls $\mathcal{F}_{Re}$ toward $\mathcal{F}_A$, again yielding a positive Δs .

4. Experiments

4.1. D-MAD Scenarios

In the D-MAD setting, a document image is compared with a trusted live image to distinguish bona-fide pairs from morphing attack pairs. A bona-fide pair consists of a bona-fide document image and a live image of the same identity, whereas an attack pair consists of a morph document image and a live image of one of the two individuals used to generate the morph. The three conditions differ according to which individual provides the live image in the attack pair.

Accomplice. The morph document image is compared with a live image of the accomplice, who submits the

morph image when applying for the document. This condition mainly represents the enrollment or document-issuance stage, where the morph image must be sufficiently similar to the accomplice to be accepted. Detection performance is evaluated using the bona-fide pairs and the accomplice-side attack pairs.

Criminal. The morph document image is compared with a live image of the criminal, who subsequently attempts to use the issued document. This condition mainly represents the operational identity-verification stage, where the criminal attempts to be verified against the morph document image. Detection performance is evaluated using the bona-fide pairs and the criminal-side attack pairs.

Both. The Accomplice and Criminal cases are evaluated together against the bona-fide pairs. This condition therefore provides an overall evaluation that considers both possible users of the morph document, rather than averaging the results obtained separately for the two conditions.

4.2. Datasets

In this study, the FRLL-Morphs dataset [24, 23] and the FEI Morph dataset [7, 29] were used for evaluation.

FRLL-Morphs is a family of morphing face datasets constructed from the Face Research Lab London Set [6]. We evaluated five morphing conditions: AMSL [21], FaceMorpher¹, OpenCV², StyleGAN [15], and WebMorph. All conditions share the same 102 bona-fide comparisons. The numbers of Accomplice and Criminal comparisons are 2,175 each for AMSL, 1,222 each for FaceMorpher and StyleGAN, and 1,221 each for OpenCV and WebMorph.

The FEI Morph dataset is constructed from the FEI Face Database and consists of Version 1 and Version 2. Version 1 [22] contains morphs generated by FaceFusion, FaceMorph, and a triangulation variant based on STASM landmarks, while Version 2 [2, 22] additionally includes FaceMorpher, UBOMorph, a SURYS method, and a method proposed by the University of Twente. Both versions share 400 bona-fide comparisons. The numbers of Accomplice and Criminal comparisons are 12,000 each for Version 1 and 28,000 each for Version 2. Therefore, Version 2 is used for the main analysis, as it subsumes the morphing methods in Version 1 and provides a more comprehensive evaluation set.

The pair information used in our experiments follows the public ACIdA repository [9]. Since the proposed method is training-free, all available pairs were used for evaluation. This public pair definition avoids manual selection of advantageous test pairs and provides a consistent basis for comparison with reported benchmark results.

¹Yao Pang, "FaceMorpher," GitHub repository. Available: <https://github.com/yaopang/FaceMorpher>.

²Satya Mallick, "Face Morph Using OpenCV - C++ / Python," LearnOpenCV, Mar. 11, 2016. Available: <https://learnopencv.com/face-morph-using-opencv-cpp-python/>.

4.3. Re-morphing Parameters

For each morphing method used for re-morphing, the experimental settings followed the corresponding official implementation. For a fair comparison, the two input face images were equally weighted in all methods by setting the interpolation coefficient to $\alpha = 0.5$. The sensitivity to α is analyzed in the supplementary material.

For MorDIFF [4, 13], the official implementation based on a Diffusion Autoencoder was used. The generation model employed an FFHQ-pretrained configuration at 256 resolution. As input preprocessing, FFHQ-style face alignment based on 68-point dlib landmarks was applied, and the final aligned image size was set to 256. Padding was enabled during alignment. The number of steps for stochastic encoding was set to $T = 250$, and the number of steps for final image reconstruction was set to $T = 20$. Under these settings, the re-morphed image was generated from the intermediate representation of the two input face images.

For OpenCV, the same face alignment strategy as in MorDIFF was applied, and the final aligned image size was also set to 256. Padding was enabled during alignment. After alignment, auxiliary points along the image boundary were added to the landmarks, Delaunay triangulation was performed, and each triangular region was synthesized using affine transformation followed by linear interpolation between the two images. Facial landmarks were detected using the dlib 68-point shape predictor, and the number of upsampling steps during face detection was set to 1.

For StyleGAN2-ADA [14], faces were first detected using 68-point dlib landmarks, and FFHQ-style face alignment was applied to obtain aligned input images at 1024 resolution. Each input image was then projected into the W space using the StyleGAN2-ADA projector, and the resulting latent representations were denoted by w_a and w_b , respectively. These latent representations were linearly interpolated as

$$w_{\text{mix}} = (1 - \alpha)w_a + \alpha w_b \quad (3)$$

and the interpolated latent representation was fed into the generator to synthesize the re-morphed image. The number of optimization steps for inversion into the W space was set to 1000 for each image. The rationale for selecting $\alpha = 0.5$ is discussed in the supplementary material.

4.4. Face Feature Extractor

CVLFace [16, 17], an existing face recognition toolkit, was used as the face feature extractor. CVLFace provides multiple face recognition backbones and training settings through a unified interface, and in this study, four models trained on WebFace4M, namely IR101, IR50, ViT KPRPE (ViT-k), and ViT, were used. For input preprocessing, the Face Alignment App included in CVLFace was used together with its corresponding face alignment model.

4.5. Evaluation Metrics

The performance of the proposed method was evaluated using Bona-fide Sample Classification Error Rate (BSCER) and Morph Attack Classification Error Rate (MACER). BSCER represents the proportion of bona-fide samples that are incorrectly classified as morphs, whereas MACER represents the proportion of morph samples that are incorrectly classified as bona-fide. These metrics were calculated following the evaluation framework of ISO/IEC 20059 for biometric systems against morphing attacks.

Let the number of bona-fide samples be denoted by N_{BF} , the number of morph samples by N_{M} , and the classification result for each sample by Res_i . Here, Res_i takes the value 1 if the i -th sample is classified as a morph, and 0 if it is classified as bona-fide. Under this definition, BSCER and MACER are given by

$$\text{BSCER} = \frac{\sum_{i=1}^{N_{\text{BF}}} \text{Res}_i}{N_{\text{BF}}} \quad (4)$$

$$\text{MACER} = \frac{\sum_{i=1}^{N_{\text{M}}} (1 - \text{Res}_i)}{N_{\text{M}}} \quad (5)$$

In addition, the Equal Error Rate (EER) was calculated as the error rate at the operating point where BSCER and MACER are equal. Furthermore, the BSCER value obtained when fixing MACER at 1% was reported as $B_{0.01}$. This enables evaluation of the misclassification rate for bona-fide samples under an operating condition where the false acceptance rate for morph samples is constrained to 1%.

5. Experimental Results

In this section, the detection performance of the proposed method on FRLL-Morphs and FEI Morph is evaluated from four perspectives. (1) Comparison across feature extractors: the effect of the feature extractor on detection performance is analyzed. Since the proposed method identifies morphing attacks based on the relationships among features extracted from the document image, the live image, and the re-morphed image generated from them, it is important to examine whether consistent tendencies can be observed across different feature extractors. (2) Comparison across re-morphing methods: the effect of the re-morphing method on detection performance is compared. In this comparison, the average performance over different feature extractors is used to analyze how the re-morphing methods, namely OpenCV (OCV), MorDIFF (MD), and StyleGAN2-ADA (SG2-A), affect D-MAD performance. (3) Comparison across detection scores: the proposed score Δs is compared with s_{re} , which uses only the similarity between the re-morphed image and the live image. This comparison

verifies the effectiveness of the difference-based score using the similarity change before and after re-morphing. (4) Comparison with existing methods: the proposed method is compared with existing methods. This comparison is conducted only under conditions where direct comparison with previous studies is possible, and AMSL in FRLL-Morphs and FEI Morph, for which published benchmark results are available, are used to clarify the relative position of the proposed method.

5.1. Comparison of Feature Extractors

This subsection compares the performance of different feature extractors on FRLL-Morphs (FRLL-M) and FEI Morph Version 2 (FEI-M2). The results are shown in Table 1.

On FRLL-M, high detection performance was obtained regardless of the feature extractor. In the OCV setting, performance was largely consistent across feature extractors, with EERs ranging from 0.00099 to 0.0014 across the three conditions, while $B_{0.01}$ was 0 for all extractors and conditions, indicating only minor variation across feature extractors. MD also showed stable performance, with the mean EER and mean $B_{0.01}$ remaining at most 0.0098. These results indicate that, on FRLL-M, the proposed method does not strongly depend on a specific feature extractor.

On FEI-M2, the overall error rates were higher than those on FRLL-M, indicating a more challenging condition. Nevertheless, the overall tendencies were consistent across feature extractors. The Criminal condition yielded low error rates, whereas the Accomplice condition showed degraded performance. The standard deviation of EER was relatively small in many conditions, indicating that the basic discriminative tendency of the proposed method was preserved across feature extractors. On the other hand, some variations were observed in $B_{0.01}$, suggesting that the feature extractor may affect strict operating points such as the low-MACER region.

Overall, the proposed method maintained consistent performance tendencies across different feature extractors, demonstrating that the similarity change caused by re-morphing can be captured in multiple face recognition feature spaces.

5.2. Comparison of Re-morphing Methods

This subsection compares the effect of the morphing method used for re-morphing. The average values over different feature extractors are reported in Table 2. To further analyze the score distributions, Figure 2 shows the Δs distributions obtained using IR101 as a representative example. The results for the other feature extractors are provided in the supplementary material.

On FRLL-M, clear performance differences were observed among the re-morphing methods. OCV achieved

Table 1. Detection performance using different feature extractors on FRLL-Morphs and FEI Morph Version 2.

Setting	Accomplice		Criminal		Both	
	EER	$B_{0.01}$	EER	$B_{0.01}$	EER	$B_{0.01}$
FRLL-M / MD / IR101	1.1e-2	2.0e-2	3.0e-3	0	7.6e-3	0
FRLL-M / MD / IR50	8.4e-3	0	2.0e-3	0	5.2e-3	0
FRLL-M / MD / ViT-k	9.8e-3	9.8e-3	2.3e-3	0	7.9e-3	0
FRLL-M / MD / ViT	9.8e-3	9.8e-3	2.1e-3	0	7.2e-3	0
FRLL-M / MD / mean	9.7e-3	9.8e-3	2.3e-3	0	7.0e-3	0
FRLL-M / MD / std	1.0e-3	8.0e-3	4.4e-4	0	1.2e-3	0
FRLL-M / OCV / IR101	9.9e-4	0	9.9e-4	0	9.9e-4	0
FRLL-M / OCV / IR50	9.9e-4	0	9.9e-4	0	9.9e-4	0
FRLL-M / OCV / ViT-k	9.9e-4	0	9.9e-4	0	9.9e-4	0
FRLL-M / OCV / ViT	1.4e-3	0	9.9e-4	0	1.2e-3	0
FRLL-M / OCV / mean	1.1e-3	0	9.9e-4	0	1.0e-3	0
FRLL-M / OCV / std	2.1e-4	0	0	0	1.1e-4	0
FRLL-M / SG2-A / IR101	6.9e-2	2.8e-1	5.5e-2	2.4e-1	5.9e-2	2.5e-1
FRLL-M / SG2-A / IR50	7.9e-2	2.9e-1	5.2e-2	2.6e-1	6.9e-2	2.9e-1
FRLL-M / SG2-A / ViT-k	8.0e-2	3.7e-1	5.7e-2	2.7e-1	6.7e-2	3.4e-1
FRLL-M / SG2-A / ViT	8.8e-2	3.6e-1	6.9e-2	3.1e-1	7.8e-2	3.5e-1
FRLL-M / SG2-A / mean	7.9e-2	3.3e-1	5.8e-2	2.7e-1	6.8e-2	3.1e-1
FRLL-M / SG2-A / std	8.0e-3	4.6e-2	7.3e-3	3.2e-2	8.1e-3	4.6e-2
FEI-M2 / MD / IR101	9.4e-2	8.1e-1	1.3e-2	1.8e-2	6.6e-2	6.0e-1
FEI-M2 / MD / IR50	1.2e-1	8.6e-1	1.8e-2	5.0e-2	8.5e-2	6.9e-1
FEI-M2 / MD / ViT-k	1.1e-1	7.6e-1	6.3e-3	2.5e-3	7.0e-2	5.9e-1
FEI-M2 / MD / ViT	1.1e-1	7.6e-1	7.2e-3	5.0e-3	6.6e-2	6.0e-1
FEI-M2 / MD / mean	1.1e-1	8.0e-1	1.1e-2	1.9e-2	7.2e-2	6.2e-1
FEI-M2 / MD / std	1.3e-2	4.8e-2	5.5e-3	2.2e-2	9.2e-3	5.0e-2
FEI-M2 / OCV / IR101	1.1e-1	5.7e-1	1.1e-2	1.3e-2	8.1e-2	3.5e-1
FEI-M2 / OCV / IR50	1.4e-1	8.7e-1	2.2e-2	1.3e-1	1.1e-1	7.1e-1
FEI-M2 / OCV / ViT-k	1.1e-1	3.8e-1	6.0e-3	5.0e-3	7.5e-2	3.1e-1
FEI-M2 / OCV / ViT	1.1e-1	4.1e-1	8.1e-3	2.5e-3	7.8e-2	3.4e-1
FEI-M2 / OCV / mean	1.2e-1	5.6e-1	1.2e-2	3.7e-2	8.6e-2	4.3e-1
FEI-M2 / OCV / std	1.6e-2	2.3e-1	7.0e-3	6.1e-2	1.6e-2	1.9e-1
FEI-M2 / SG2-A / IR101	2.3e-1	8.6e-1	4.8e-2	1.6e-1	1.7e-1	7.7e-1
FEI-M2 / SG2-A / IR50	2.5e-1	8.7e-1	5.5e-2	1.9e-1	1.7e-1	8.0e-1
FEI-M2 / SG2-A / ViT-k	2.5e-1	9.1e-1	4.3e-2	1.6e-1	1.6e-1	8.4e-1
FEI-M2 / SG2-A / ViT	2.5e-1	8.9e-1	3.5e-2	1.9e-1	1.7e-1	8.3e-1
FEI-M2 / SG2-A / mean	2.4e-1	8.8e-1	4.5e-2	1.7e-1	1.7e-1	8.1e-1
FEI-M2 / SG2-A / std	8.4e-3	2.1e-2	8.2e-3	1.9e-2	4.6e-3	3.4e-2

Table 2. Average detection performance of different re-morphing methods on FRLL-Morphs and FEI Morph Version 2.

Setting	Accomplice		Criminal		Both	
	EER	$B_{0.01}$	EER	$B_{0.01}$	EER	$B_{0.01}$
FRLL-M / mean / MD	9.7e-3	9.8e-3	2.3e-3	0	7.0e-3	0
FRLL-M / mean / OCV	1.1e-3	0	9.9e-4	0	1.0e-3	0
FRLL-M / mean / SG2-A	7.9e-2	3.3e-1	5.8e-2	2.7e-1	6.8e-2	3.1e-1
FEI-M2 / mean / MD	1.1e-1	8.0e-1	1.1e-2	1.9e-2	7.2e-2	6.2e-1
FEI-M2 / mean / OCV	1.2e-1	5.6e-1	1.2e-2	3.7e-2	8.6e-2	4.3e-1
FEI-M2 / mean / SG2-A	2.4e-1	8.8e-1	4.5e-2	1.7e-1	1.7e-1	8.1e-1

the best performance, with mean EERs of approximately 0.001 and $B_{0.01}$ values of 0 across all conditions, while MD also demonstrated strong detection performance. In contrast, SG2-A produced substantially higher error rates. A similar tendency was observed on FEI-M2: MD and OCV performed favorably, whereas SG2-A yielded the highest error rates. MD achieved the best EER, while OCV out-

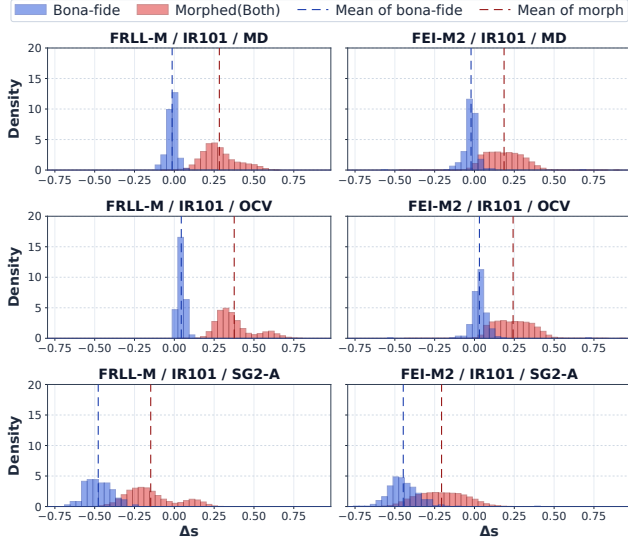


Figure 2. Distributions of Δs for bona-fide and morphed samples using IR101 on FRLL-Morphs and FEI Morph Version 2.

performed MD in $B_{0.01}$ under the Accomplice and Both conditions. Across both datasets, the Criminal condition consistently yielded lower error rates than the Accomplice condition, especially for SG2-A.

As shown in Figure 2, for MD and OCV, the Δs distribution of bona-fide samples was concentrated around zero, whereas that of morphed samples shifted toward the positive direction. This tendency is consistent with the assumption of the proposed method. In particular, the high performance of OCV may be attributed to the stability of landmark-based geometric interpolation, which can reflect the relationship between the input images in a stable manner.

In contrast, SG2-A showed a different distributional tendency. Since SG2-A requires inversion of the input images into the GAN latent space and subsequent image synthesis from the interpolated latent representation, identity drift may occur due to errors in the inversion and synthesis processes. As a result, $\mathcal{F}_{Re}$ does not necessarily move closer to $\mathcal{F}_L$, and the expected direction of similarity change may be weakened. This may contribute to the distributional overlap and the degradation in detection performance.

These results suggest that the proposed method is sensitive not only to artifacts caused by morphing, but also to how consistently the identity relationship is preserved in the feature space through re-morphing. Therefore, in Face Re-morphing, the choice of the morphing method is important not only for visual quality but also for stable identity changes in the feature space. A further discussion of the relationship between identity preservation and performance across re-morphing methods is provided in the supplementary material.

Table 3. Comparison of EERs for s_{base} , Δs , and s_{re} across datasets and re-morphing methods. Since s_{base} is independent of the re-morphing method, it is reported once per dataset.

Setting	Score	Accomplice	Criminal	Both
FRLL-M / mean / –	s_{base}	5.3e-3	1.2e-3	4.3e-3
FRLL-M / mean / MD	Δs	9.7e-3	2.3e-3	7.0e-3
	s_{re}	1.5e-1	1.1e-1	1.2e-1
FRLL-M / mean / OCV	Δs	1.1e-3	9.9e-4	1.0e-3
	s_{re}	7.9e-2	4.4e-2	6.2e-2
FRLL-M / mean / SG2-A	Δs	7.9e-2	5.8e-2	6.8e-2
	s_{re}	2.4e-1	2.0e-1	2.2e-1
FEI-M2 / mean / –	s_{base}	1.5e-1	1.7e-2	1.1e-1
FEI-M2 / mean / MD	Δs	1.1e-1	1.1e-2	7.2e-2
	s_{re}	3.4e-1	1.3e-1	2.5e-1
FEI-M2 / mean / OCV	Δs	1.2e-1	1.2e-2	8.6e-2
	s_{re}	3.1e-1	9.9e-2	2.3e-1
FEI-M2 / mean / SG2-A	Δs	2.4e-1	4.5e-2	1.7e-1
	s_{re}	4.2e-1	2.4e-1	3.4e-1

5.3. Comparison of Detection Scores

This subsection examines the effectiveness of the proposed score Δs , defined in Section 3.3, by comparing it with s_{base} and s_{re} . The results are shown in Table 3.

Overall, Δs achieved substantially lower EERs than using s_{re} alone. On FRLL-M, Δs achieved EERs in the range of 0.00099–0.079, whereas those of s_{re} ranged from 0.044–0.24. On FEI-M2, the corresponding ranges were 0.011–0.24 and 0.099–0.42, respectively. These results indicate that the similarity between the re-morphed image and the live image alone is insufficient to separate bona-fide and morphing cases.

Compared with s_{base} , however, the effectiveness of Δs depended on the re-morphing method. For OCV, Δs outperformed s_{base} under all three conditions on both datasets. For MD, Δs improved the EER under all conditions on FEI-M2, reducing it from 0.15, 0.017, and 0.11 to 0.11, 0.011, and 0.072, respectively, but did not outperform s_{base} on FRLL-M. In contrast, SG2-A produced higher EERs than s_{base} under all conditions on both datasets.

These results show that subtracting s_{base} can reveal similarity changes that are not captured by s_{re} alone, particularly when MD or OCV is used for re-morphing. However, the degradation observed with SG2-A indicates that the effectiveness of Δs depends on the re-morphing method.

5.4. Comparison with Existing Methods

This subsection compares the proposed method with existing methods on AMSL and FEI Morph Version 1 (FEI-M1), where direct comparison with previous studies is possible. The existing results on AMSL and FEI-M1 were taken from Borghi et al. [3] and Liu et al. [18], respectively. As described in Section 4.1, our evaluation uses publicly

Table 4. Comparison between the proposed method and existing methods on the AMSL dataset. Bold and underlined values indicate the best and second-best results, respectively.

Method	Accomplice		Criminal		Both	
	EER	B _{0.01}	EER	B _{0.01}	EER	B _{0.01}
[26]	.059	.833	.089	.980	.075	.966
[25]	.069	.598	.059	.667	.069	.662
[28]	-	-	-	.031	-	-
[27]	.003	<u>.020</u>	.000	.000	<u>.001</u>	.020
[3]	<u>.001</u>	<u>.020</u>	<u>.001</u>	.059	<u>.001</u>	.059
Ours						
AMSL / mean / MD	.014	.032	<u>.001</u>	.000	.009	<u>.005</u>
AMSL / mean / OCV	.000	.000	.000	.000	.000	.000
AMSL / mean / SG2-A	.104	.365	.023	<u>.056</u>	.070	.292

Table 5. Comparison between the proposed method and existing methods on FEI Morph Version 1. Bold and underlined values indicate the best and second-best results, respectively.

Method	Accomplice		Criminal		Both	
	EER	B _{0.01}	EER	B _{0.01}	EER	B _{0.01}
[11]	.160	.618	.027	.050	.111	.522
[26]	.200	.623	.185	.697	.192	.628
[27]	.178	.770	.068	.295	.128	.690
[3]	.153	.563	.061	.370	.115	.515
[7]	.125	.468	.125	.440	.125	.445
[8]	.102	<u>.333</u>	<u>.023</u>	<u>.027</u>	.070	<u>.280</u>
[18]	.078	.283	.037	.115	.059	.230
Ours						
FEI-M1 / mean / MD	<u>.099</u>	.724	.010	.015	<u>.066</u>	.538
FEI-M1 / mean / OCV	.105	.465	.010	.034	.079	.371
FEI-M1 / mean / SG2-A	.242	.882	.039	.128	.167	.811

defined pair information and does not involve manual selection of test pairs that could favor the proposed method. However, since the reported results of existing methods may have been obtained under pair settings and experimental protocols that are not fully identical to ours, the comparison should be interpreted as a reference comparison based on published benchmark results. For consistency with prior studies, the proposed method is reported using the same decimal precision. The results are shown in Tables 4 and 5.

On AMSL, the OCV setting achieved the best performance under the Accomplice, Criminal, and Both conditions, with EER and B_{0.01} values of 0.000. This indicates that no errors were observed under this evaluation protocol. We interpret this result as near-perfect performance on the AMSL benchmark rather than evidence of universal separability. This near-perfect result is consistent with the general distributional tendency observed in Figure 2 for OCV-based re-morphing on FRLL-Morphs, where bona-fide and morphed samples show clear separation across morphing conditions. Note that Figure 2 aggregates all FRLL-Morphs morphing conditions and is therefore not specific to AMSL. Since all pairs follow the public ACIdA protocol and no manual pair selection was performed, this result reflects the behavior of the proposed score under this benchmark set-

ting. MD also showed competitive performance, whereas SG2-A yielded higher error rates.

On FEI-M1, the proposed method was particularly effective under the Criminal condition. Both MD and OCV achieved an EER of 0.010, outperforming the existing methods, and MD achieved the best B_{0.01} value of 0.015. However, under the Accomplice and Both conditions, the best existing method [18] outperformed the proposed variants. Thus, the overall performance on FEI-M1 is mainly limited by the Accomplice condition.

This difference may be partly related to the morphing factor. AMSL uses a fixed morphing factor of 0.5, where the identity components of the accomplice and criminal are relatively balanced. In contrast, FEI Morph uses morphing factors of 0.3 and 0.5, where the factor denotes the criminal’s contribution. At a factor of 0.3, the accomplice contribution becomes dominant. Thus, under the Accomplice condition, the morph image can be closer to the live accomplice image, resulting in a higher s_{base} and a smaller change after re-morphing. This may increase the overlap between the Δs distributions of bona-fide and morphing cases.

Overall, the proposed method outperformed existing methods on AMSL and showed favorable results under the Criminal condition on FEI-M1. Improving discrimination under the Accomplice condition remains future work.

6. Conclusion

This paper proposed Face Re-morphing, a D-MAD method that uses the cosine similarity change caused by an additional morphing operation between the document image and the live image. The proposed method generates a re-morphed image and uses the difference between the document–live and live–re-morphed similarities as the detection score, thereby incorporating the morphing operation itself into the detection process.

Experiments on FRLL-Morphs and FEI Morph showed that the proposed cue is effective across different morphing conditions, re-morphing methods, and face feature extractors. The proposed method achieved particularly strong performance on AMSL and low error rates under the Criminal condition on FEI Morph, demonstrating that re-morphing-induced similarity change is a useful cue for D-MAD.

The results also indicate that the method depends on the re-morphing strategy and its interaction with the face feature space. This study focused on digital-image evaluation, and future work will examine robustness under print–scan and other operational acquisition processes.

Acknowledgment

This work was supported in part by JSPS KAKENHI JP23H00463, JP23K28085, and JST Moonshot R&D Grant Number JPMJMS2215.

References

- [1] A. Banas, C. Rathgeb, J. Merkle, and M. Schaubert. Idswap-mad: towards privacy-friendly training of differential face morphing attack detection. *Journal on Information Security*, 2026.
- [2] I. Batskos, L. Spreeuwers, and R. Veldhuis. Visualizing landmark-based face morphing traces on digital images. *Frontiers in Computer Science*, 5:981933, 2023.
- [3] G. Borghi, E. Pancisi, M. Ferrara, and D. Maltoni. A double siamese framework for differential morphing attack detection. *Sensors*, 21(10):3466, 2021.
- [4] N. Damer, M. Fang, P. Siebke, J. N. Kolf, M. Huber, and F. Boutros. Mordiff: Recognition vulnerability and attack detectability of face morphing attacks created by diffusion autoencoders, 2023.
- [5] N. Damer, A. M. Saladie, A. Braun, and A. Kuijper. Morgan: Recognition vulnerability and attack detectability of face morphing attacks created by generative adversarial network. In *2018 IEEE 9th international conference on biometrics theory, applications and systems (BTAS)*, pages 1–10. IEEE, 2018.
- [6] L. DeBruine and B. Jones. Face research lab london set, May 2017.
- [7] N. Di Domenico, G. Borghi, A. Franco, and D. Maltoni. Combining identity features and artifact analysis for differential morphing attack detection. In *International Conference on Image Analysis and Processing*, pages 100–111. Springer, 2023.
- [8] N. Di Domenico, G. Borghi, A. Franco, and D. Maltoni. Dealing with subject similarity in differential morphing attack detection. *arXiv preprint arXiv:2404.07667*, 2024.
- [9] N. Di Domenico, G. Borghi, A. Franco, and D. Maltoni. Improving accomplice detection in the morphing attack. *Machine Intelligence Research*, 22(3):484–498, 2025.
- [10] M. Ferrara, A. Franco, and D. Maltoni. The magic passport. In *IEEE international joint conference on biometrics*, pages 1–7. IEEE, 2014.
- [11] M. Ferrara, A. Franco, and D. Maltoni. Face demorphing. *IEEE Transactions on Information Forensics and Security*, 13(4):1008–1017, 2017.
- [12] T. T. Hoang, B. M. S. Islam, and H. Ahn. Face morphing attack detection using similarity score patterns between demorphed and live images. *Electronics*, 14(19):3851, 2025.
- [13] M. Huber, F. Boutros, A. T. Luu, K. B. Raja, R. Ramachandra, N. Damer, P. C. Neto, T. Gonçalves, A. F. Sequeira, J. S. Cardoso, J. Tremoço, M. Lourenço, S. Serra, E. Cermeño, M. Ivanovska, B. Batagelj, A. Kronovsek, P. Peer, and V. Struc. SYN-MAD 2022: Competition on face morphing attack detection based on privacy-aware synthetic training data. In *IEEE International Joint Conference on Biometrics, IJCB 2022, Abu Dhabi, United Arab Emirates, October 10-13, 2022*, pages 1–10. IEEE, 2022.
- [14] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila. Training generative adversarial networks with limited data. In *Proc. NeurIPS*, 2020.
- [15] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila. Analyzing and improving the image quality of StyleGAN. In *Proc. CVPR*, 2020.
- [16] M. Kim, A. K. Jain, and X. Liu. Adaface: Quality adaptive margin for face recognition. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 18729–18738, 2022.
- [17] M. Kim, Y. Su, F. Liu, A. Jain, and X. Liu. Keypoint relative position encoding for face recognition. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 244–255, 2024.
- [18] C. Liu, M. Ferrara, A. Franco, G. Borghi, and D. Zhong. Differential morphing attack detection via triplet-based metric learning and artifact extraction. In *2024 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–7. IEEE, 2024.
- [19] M. Long, Q. Yao, L.-B. Zhang, and F. Peng. Face demorphing based on diffusion autoencoders. *IEEE Transactions on Information Forensics and Security*, 19:3051–3063, 2024.
- [20] E. M. Namis, K. S. Jasim, and S. Al-Janabi. Face morphing attacks detection approaches: A review. *Mesopotamian Journal of Big Data*, 2024:82–101, 2024.
- [21] T. Neubert, A. Makrushin, M. Hildebrandt, C. Kraetzer, and J. Dittmann. Extended stirtrace benchmarking of biometric and forensic qualities of morphed face images. *IET Biometrics*, 7(4):325–332, 2018.
- [22] K. Raja, M. Ferrara, A. Franco, L. Spreeuwers, I. Batskos, F. De Wit, M. Gomez-Barrero, U. Scherhag, D. Fischer, S. K. Venkatesh, et al. Morphing attack detection-database, evaluation platform, and benchmarking. *IEEE Transactions on Information Forensics and Security*, 16:4336–4351, 2020.
- [23] E. Sarkar, P. Korshunov, L. Colbois, and S. Marcel. Vulnerability analysis of face morphing attacks from landmarks and generative adversarial networks. *arXiv preprint*, Oct. 2020.
- [24] E. Sarkar, P. Korshunov, L. Colbois, and S. Marcel. Are gan-based morphs threatening face recognition? In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2959–2963, 2022.
- [25] U. Scherhag, C. Rathgeb, and C. Busch. Morph detection from single face image: A multi-algorithm fusion approach. In *Proceedings of the 2018 2nd International Conference on Biometric Engineering and Applications*, pages 6–12, 2018.
- [26] U. Scherhag, C. Rathgeb, and C. Busch. Towards detection of morphed face images in electronic travel documents. In *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*, pages 187–192. IEEE, 2018.
- [27] U. Scherhag, C. Rathgeb, J. Merkle, and C. Busch. Deep face representations for differential morphing attack detection. *IEEE Transactions on Information Forensics and Security*, 15:3625–3639, 2020.
- [28] S. Soleymani, A. Dabouei, F. Taherkhani, J. Dawson, and N. M. Nasrabadi. Mutual information maximization on disentangled representations for differential morph detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1731–1741, 2021.

- [29] C. E. Thomaz and G. A. Giraldi. A new ranking method for principal components analysis and its application to face image analysis. *Image and vision computing*, 28(6):902–913, 2010.