2026 Volume: 02 Issue: 06 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-23

DEVELOPMENTOFADVERSARIALROBUSTDEEPLEARNINGFRAMEWORKFORFACE MORPHING ATTACK DETECTION IN BIOMETRIC SYSTEMS

Oko Gabriel Ota*, Agu Joy Amarachi, Onyemachi George Olisamaka

Computer Science Department, Enugu State University of Science and Technology.

Article Received: 9 May 2026
Article Revised: 29 May 2026
Published on: 19 June 2026

***Corresponding Author: Oko Gabriel Ota**
Computer Science Department, Enugu State University of Science and Technology.

DOI: https://doi-doi.org/101555/ijrpa.2737

ABSTRACT

Face recognition systems have become integral to modern biometric security applications, including border control, banking, and national identity management. However, the increasing sophistication of face morphing attacks poses significant threats to their reliability. This study presents an adversarially robust deep learning framework for detecting face morphing attacks generated using advanced techniques such as MIPGAN,MorDiff, Greedy-DIM, and Morph-PIPE. These techniquesproduce highly realistic morphed imagescapable of deceiving both human observers and automated systems. The proposed framework leverages convolutional neural networks (CNNs) and adversarial learning strategies to enhance detection performance across both known and unseen morphing attacks. Using the SYN-MAD 2022 dataset and evaluating against state-of-the-art face recognition systems (ArcFace, AdaFace, and ElasticFace), the study demonstrates the effectiveness of the proposed approach. Experimental results reveal that advanced morphing methods, particularly Greedy-DIM and Morph-PIPE, achieve extremely high attack successrates, highlighting the urgent need for robust detection mechanisms. The findings contribute to strengthening biometric security systems and mitigating identity fraud risks in real-world applications.

KEYWORDS: Face Morphing Attacks; Biometric Security; Face Recognition Systems; Morphing AttackDetection (MAD); Deep Learning; Adversarial Learning.

1.0 INTRODUCTION

Biometric systems particularly face recognition systems; have become widely adopted in

security-critical applications such as border control, banking authentication, and national identity systems. However, these systems are increasingly vulnerable to face morphing attacks, where images of two or more individuals are digitally combined to create a synthetic identity capable of bypassing authentication systems.

Biometrics has been commonly utilized in various areas of security. Biometrics, which refers to themeasurement of human features, is a prominent area of research in computer science. Theyhave applications in security of data, access control, and identity systems (Ambikapathy, *et al,* 2024). Personal identificationnumbers or passwords were once the primary means of identification for each individual, but this led to numerous problems. As aresult ofhow easyit is for someone to pretend to be someone else simplybyknowing their password or phone number, biometric identifiers which areregarded as more reliable are being used more frequently. Since biometrics can provide unique and powerful information that can precisely identify a person, such as skin color, fingerprint, iris, face, hand geometry, retina, DNA, and palm veins, identification isthemost widely used application of biometrics (Hernandez-de-Menende, *et al,* 2021). Since every person's face is different and packed with information (Keshavarz, *et al,* 2020), facial recognition systems are widely used in security and service departments of mobile phones, businesses, international airports, and social media companies on the internet, among other domains (Scherhag, et al, 2019).

According to Ghareb, Hamid, Sabr, and Tofiq (2022), in an effort to develop systems that can recognize faces accurately even though people are remarkably similar, researchers are currently very interested in facial recognition systems. Face morphing can be a serious security risk when these altered photos are used for passports or identification because it allows multiple people (subjects) to use the information to verify their identities (Robertson, et al, 2017; Seibold, *et al.* 2020). Multiple subjects' improper connection to the document may lead to a number of illicit activities, including financial transactions, illegal immigration, and human trafficking. The targeted offender would modifyhis face photo with one of the impersonating partners in a real-world face-morphing attack scenario. If the partner requests an e-passport with an altered face photo, they will receive an authentic e-passport with matching document security features. It is possible to authenticate the accomplice and the partner using the morphed image found in the e-passport. This suggests that by using the e-passport that was given to theaccomplice, the criminal can avoid theautomated border control gates or perhaps even the human inspections at the gate. For this reason, it is imperative that this face-morphing attack be identified automatically (Hamza, Tehsin, Karamti and Alghamdi, 2022).

Figure 1 depictsa border control example scenario whereamalicious person'sface isaltered toresemble that of an accomplice.

Figure1: Example of border control FRS vulner ability to morphedimage.

Face morphing attacks pose a serious threat to the reliability of automatic face recognition systems (FRSs), particularlyin applications such as electronicidentityissuanceand border control. Theseattacks blendthefacial features of two or more individuals to create hybrid images capable of impersonating multiple identities. The proliferation of open-source morphingtools andrecentadvances in generativemodels havemadeit increasingly easyto generatehigh-quality, realistic morphs (Damer, *et al,* 2023; Zhang, *et al,* 2024). This has heightened the need for reliable morphing attack detection (MAD) methods to counter threats such as identity theft, document fraud, and social engineering.

FaceRecognition (FR) systems havebeen deployed in a widerange ofsettings from unlockingphones toborder control (Spreeuwers, Hendrikse, and Gerritse, 2012), and the performance of such systems keeps improving.Due to its wide adoption, FR systems have been the target to various types of attacks (Ngan *et al,* 2020). Most recently, it has been shown that these systems are vulnerable to the face morphing attacks (Ferrara, *et al.,* 2016; Blasingame, *et al.,* 2021), an emerging type of attack that aims to blend the biometric identifiers of multiple faces intoasingleimagethat would trigger afalseacceptancewith anyoftheidentities used tomakethemorph. Application scenarios where the system allows users to submit self-generated images to obtain a valid passport or visa creates a significant attack vector for a malicious agent to enroll morphed images (Venkatesh, *et al*, 2021). The potency of these attacks have led to face morphing becoming a focal point in recent research on FR systems (Ferrara, *et al*., 2014; Venkatesh, *et al*, 2020), in an effort to secure the identitycontrol process.

Figure2: Example of amorph create dusing Greedy-DiM.

Face morphing attacks exploit weaknesses in facial recognition algorithms by embedding identity features of multiple individuals into a single image. These attacks can deceive systems during enrolment and verification stages, therebyposing serious securityrisks. Recent advancements in deep learninghaveled to thedevelopment ofhighlysophisticated morphing techniquessuch as MIPGAN, MorDiff, Greedy-DIM, and Morph-PIPE, which produce highly realistic morphs that are difficult to detect using traditional methods. The present research aimsto develop adversarialrobust deep learning framework for face morphing attack detection in biometric systems. The system will utilize state-of-the-art deep learningalgorithms, such as convolutional neuralnetworks (CNNs) and generative adversarial networks (GANs), to improve the accuracy of morphed face image detection.

Research Problem Statement

Despite the progress in biometric technologies, existing face recognition systems lack robustness against advanced morphing attacks. Traditional morph detection techniques rely on handcrafted features or supervised learning approaches, which often fail to generalize across unseen attack types. Therefore, there is a need for a robust and advanced deep learning morphing frameworks such as MIPGAN, MorDiff, Greedy-DIM, andMorph-PIPE for morphing attacks.

Aim and Objectives

Aim: To develop an adversarially robust deep learning framework for detecting face morphing attacks in biometric systems.

The objectives of the study are to analyze advanced morphing techniques such as MIPGAN, MorDiff, Greedy-DIM, and Morph-PIPE, todesign a deep learning-based morph detection

model with adversarial robustness and to evaluate the performance of the proposed framework using benchmark datasets.

Significance of the Study

This studycontributes totheadvancement ofbiometricsecuritybyaddressing vulnerabilities in facerecognition systems. It will providea robust detection mechanismagainst advanced morphingattacks, improved securityfor identity verification systems and provide a framework adaptable to evolving attack strategies. The findings will be relevant for governments, security agencies, and organizations relying on biometric authentication.

Scope and Limitation

Scope: the study focus on face morphing attack detection using deep learning, evaluation of state-of-the-art morphing techniques (MIPGAN, MorDiff, Greedy-DIM, Morph-PIPE) and application to biometric systems.

Limitations: the study is limited to dependence on available high-quality datasets, computational complexity of deep learning models and the potential difficulty in real-time deployment.

2. Literature Review

Overview

A Face Recognition System (FRS) performs identity verification by comparing two photos and decidingwhether or not theidentities match. It was first shown in (Ferrara, Franco and Maltoni, 2014) that existing Face Recognition Systems (FRS) were vulnerable to morphing attacks. A morph is an image that contains facial features of two different people. In a border-crossing scenario a criminal could enlist thehelp of an accomplice to createa morphed photo. Theaccomplice could then usethis photo toapplyfora passport, which thecriminal in turn could use to cross borders undetected. According to Scherhag, *et al.,* (2019); Makrushin, *et al.,* (2018) and Robertson, *et al.,* (2017), the most-used method to create morphs is to mark certain facial features, called landmarks, warp both images to a common geometry and then blend the pixel values. It has been shown thatboth FRSs and humans will often accept a morph made with this method as a match with both contributing identities.

Recently, a platform was launched with which the performance of different morphing detection algorithms can be benchmarked (Raja *et al.,* 2020). This benchmark and other research indicate that existing algorithms do not perform well when tested across different datasets (Spreeuwers, *et al.,* 2018; Scherhag, *et al.,* 2018). Since researchers have so far had

to create their own training datasets, their detection methods may have been overfitted to specific characteristics of their training set. Furthermore, some detection methods require large datasets for training, which means that a large number of morphs need to be made. Since this is usually done automatically such morphs will probably be of lower quality than hand-crafted morphs. A detection method created by training with such data can detect low-quality but not high-quality morphs.

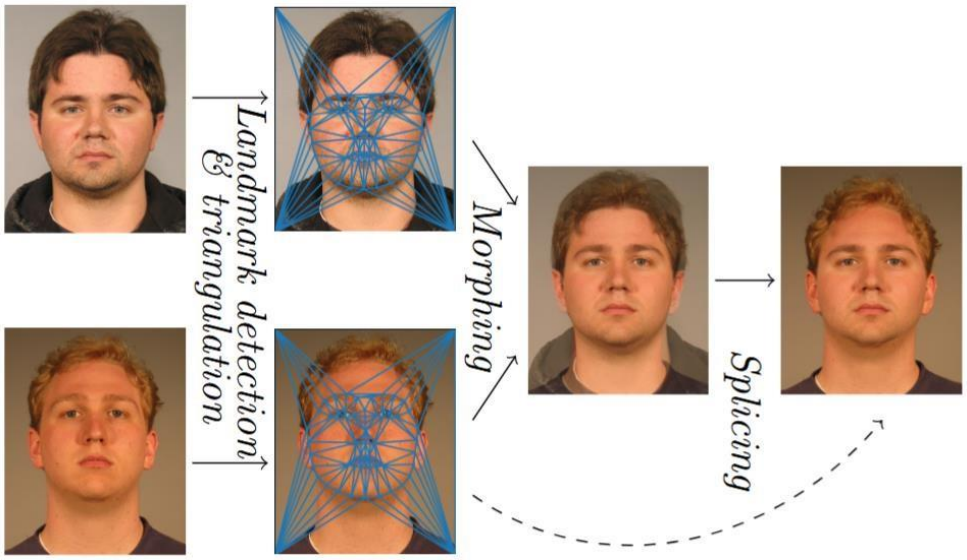

Figure3: Morphing Process.

Thenaive approach to createa face morph is to simplytake a pixel-wise average of the original images. Due to its simplicity, this approach can result in many noticeable artifacts. An extension on this would be to align the landmarks of the face, warping the pixels to fit this new representation, and then averaging the pixels. The blending value is usually set to 0.5 so that both of the original images could contribute equally to the morphed image. Later work (Zhang, *et al.,* 2021; Zhang, *et al.,* 2023) questions if this assumption is valid. Further improvements can be made to these landmark-based morphs by additional post-processing to remove the artifacts from the morphing process (Ferrara, *et al.,* 2019).

MIPGAN

Morphing through Identity Prior driven GAN (MIPGAN) and in particular MIPGAN-II improves upon the originalGAN-based morphsandStyleGAN2-based morphs (Sarkar, *etal.,* 2022)byusing optimization methods to search for a latent representation that fools the FR system maximally (Zhang, *et al.,* 2021). The MIPGAN algorithm works bytaking thetwo bona fide images x_a, x_b, and encoding them into latentrepresentations, z_a, z_b. An initial morphed latent $z_{ab}=1/2(z_a+z_b)$ is then constructed. The initial latent is optimized such that the generated morph $x_{ab}= G(z_{ab})$ minimizes an identity loss based on the ArcFace FR system (Deng , *et al.,* 2019) along with a perceptual loss term. This optimization procedure produces a more potent morphing attack than simply using the morphed latent z_{ab}.

DiM

Blasingame *et al.* (2024) proposed to use diffusion-based methods to construct high-quality face morphs. This family of morphing attacks is known as Diffusion Morphs (DiM). Diffusion models function by modeling the diffusion process wherein noise is progressively added to the data until the resulting distribution reassembles a Gaussian distribution. To samplethedata distribution,arandom sampleisdrawn from theGaussian distribution and then the diffusion process is run in reverse to iteratively remove noise until a clean sample from data distribution is produced. Given the data distribution $P_{data}(x)$ on the data space $X \subseteq R^n$, the diffusion process is given by the Itˆo stochastic differential equation (SDE).

$$dx_t=f(t)x_t dt+g(t)dw_t \quad (1)$$

where$t\in$ [0,T]is thetimewith fixed constantT> 0, $f(\cdot)$and $g(\cdot)$arethedrift and diffusion coefficients,and $w_t$denotes standardBrownian motion.This paper employed the variance preserving(VP)formulation ofdiffusion models which uses the following drift and diffusion coefficients

$$f(t)=\frac{d\log\alpha_t}{dt} \quad (2)$$

$$g^2(t)=\sigma_t^2-\alpha_t^2\frac{d\sigma_t^2}{dt}\frac{1}{\alpha_t^2} \quad (3)$$

whereα_t, σ_tformthenoiseschedule ofthediffusionmodelsuchthat

$$q(x_t|x_0)=N(x_t|\alpha_t x_0, \sigma_t^2 I) \quad (4)$$

withasignal-to-noise-ratio(SNR)α_t^2/σ_t^2,whichisstrictlydecreasingwithrespecttot.

The distributionof x_t is denoted as $p_t(x_t)$, therefore $p_0(x_0) \equiv p_{data}(x)$. Song *et al.* (2021) showed the existence of an ordinary differential equation (ODE) dubbed the Probability Flow ODE (PF-ODE) whose marginal distributions p_t follow the same trajectories as the diffusion SDE. The PF-ODE is given as

$$\frac{dx_t}{dt}=f(t)x_t-\frac{1}{2}g^2(t)\nabla_x \log P_t(x_t) \quad (5)$$

where$\nabla_x \log P_t(x_t)$denotesthescoreof$P_t(x_t)$.Thescorefunctioncanbelearnedbyaneuralnetwork$\epsilon_\theta(x_t,t)\approx -\sigma_t\nabla_x \log P_t(x_t)$. To sample the model, an initial noise sample $x_T\sim N(0, I)$ is drawn. Then a numerical ODE solver is employed to solve the ODE from time T to time 0.

where$\nabla_x \log P_t(x_t)$denotesthescoreof$P_t(x_t)$.Thescorefunctioncanbelearnedbyaneuralnetwork$\epsilon_\theta(x_t,t) \approx -\sigma_t \nabla_x \log P_t(x_t)$. To sample the model, an initial noise sample $x_T \sim N(0, I)$ is drawn. Then a numerical ODE solver is employed to solve the ODE from time T to time 0.

DiM models employscore predictor model conditioned on a latentrepresentation of thetarget image, i.e., $\epsilon_\theta(x_t, z, t)$ with an encoding network $z = E(x_0)$. Previous DiMapproaches (Blasingame, *et al*., 2024, and Zhang,*et al.,* 2023) use a Diffusion Autoencoder model (Preechakul, Chatthee, Wizadwongsa and Suwa-janakorn,2022) asthe backbone for the score prediction model. As per the Diffusion Autoencoder technique, the ODE solver Φ_0 is reversed now solving from t = 0 to t = T, resulting ina deterministic forward passto anapproximationof $p_T(x_T)$. This forwardODEsolver was denoted, socalledas itis evaluatedas timeruns forwards from 0toT, asΦ^+. The noisy bona fide images are morphed using spherical linear interpolation, i.e., $x^{(ab)}_T = \text{slerp}(x^{(a)}, x^{(b)}_0; 0.5)$. The conditionals are averaged and this morphed information is used as the input to the diffusion model.

Morph-PIPE

Zhang et al. [22] proposed Morph-PIPE, a simple extension on the DiM method byusing an identity-based loss to select between a range of blend values. Theyuse an identityloss defined as the sum of two sub-losses:

$$L_{ID} = d(v_{ab}, v_a) + d(v_{ab}, v_b) \quad (6)$$

$$L_{diff} = |d(v_{ab}, v_a) - d(v_{ab}, v_b)| \quad (7)$$

$$L^*_{ID} = L_{ID} + L_{diff} \quad (8)$$

where $v_a = F(x^{(a)}_0)$, $v_b = F(x^{(b)}_0)$, $v_{ab} = F(x^{(ab)}_0)$, and $F : X \rightarrow V$ is an FR system which embeds images into a vector space V which is equipped with a measure of distance, d. To create the morph B, different morphs are generated with different blend coefficients $\{w_n\}^B_{n=1} \subseteq [0, 1]$.Theythen choosethewn which minimizes L^*ID. By leveraging a powerful FR system like ArcFace, the selected morphs are more potent than the original DiM implementation with a fixed blend parameter w = 0.5.

Fast-DiM is a concurrent workwhichproposes animprovement onDiM with the goal ofreducing the number of Network Function Evaluations (NFE)needed toproduce high-qualityface morphs using DiM.Thisreduction in NFE is primarily accomplished by using alternative ODE solvers used in solving the Probability Flow ODE and replacing the "stochastic encoder" with an ODE solver for the Probability Flow ODE as time runs forward from 0 to T.

Greedy-DiM

Greedy-DiMisanovelfamilyoffacemorphingattacksthatgreedilychoosetheoptimalsolutionateach timestep in solving the PF-ODE.

GreedySearch

The experiment begins by applying a greedy search during the sampling process of diffusion models. The information from the identity losses of Equations (7) and (8) provides a powerful heuristic function for guiding the creation of face morphs. While Zhang *et al.* (2023) only uses this information to select from a pre-selectedset of morphed image candidates, the experiment used this heuristic during the sampling process. Now, the identity losses require a morphed image as input; however, at time t > 0, the image $x_t$is not fully denoised. To remedythis, the experiment used the estimate of $x_0$at each timestep t as input to theidentitylosses. This can be found from the noise prediction network via

$$x_\theta(x_t,z,t)=\frac{x_t-\sigma_t\epsilon_\theta(x_t,z,t)}{\alpha_t} \tag{9}$$

In the greedy strategy, it was found that the optimal blend parameter wat each step between two trajectories of the PF-ODE, the first trajectorybeing themodel conditioned on $z_a$and the second being conditioned on z_b. This strategyis detailed in Equation (10) where H is the heuristicfunction which measureshow “optimal” the morph is relative to the bona fide images.

$$\epsilon^{(w)}_t=\mathrm{slerp}(\epsilon_\theta(x^{(ab)}_t,z_a,t),\epsilon_\theta(x^{(ab)}_t,z_b,t);w)$$
$$w'_t=\underset{w\in\{w_n\}_{n=1}^{B}}{\operatorname{argmin}} H\left(\frac{1}{\alpha_t}(1-\sigma_t\epsilon(w),x(a)0,x(b)0\right) \tag{10}$$

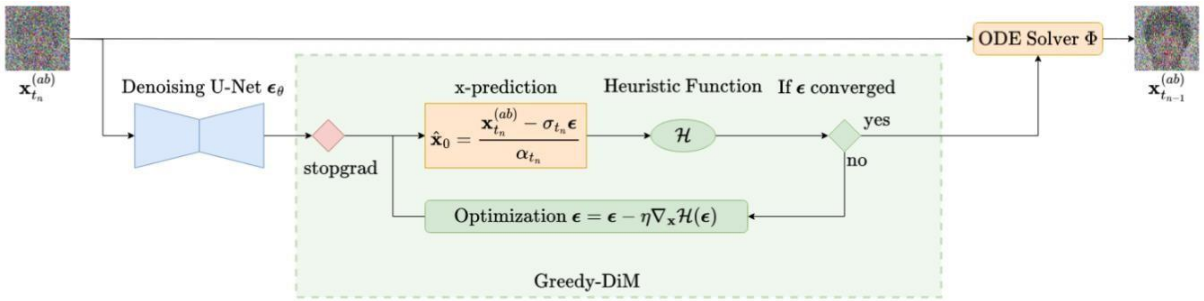

Figure 4: Overview of a single step of the Greedy-DiM algorithm. The highlighted part in green represents the changes.

Thisistheresult in asequenceofblend parametersw′tfor each stepin thePF-ODE solver. So, the$\epsilon^{(w't)}$for each timestep t is used as thenoise predictionwhensolving the PF-ODE. This model is referto as Greedy-DiM-S as it is the search variant of the Greedy-DiM family.

MorDIFF

This approach is based on the work of Konpat *et al.* (2022) and is described in the following with thenecessary background information. Other morphing generation process used the GAN architecture (Naser *et al.,* 2019; Naser, *et al.,* 2018; Sushma, *et al.,* 2020) and generated good morph attacks. Despite the fact that GAN canlearn meaningful representations in the latent spaces, the reconstruction fidelity is limited due to the latent size depending on the number of training samples. In contrast to GAN, diffusion probabilistic model (DPM) (Jonathan, *et al.,* 2020; Jascha, *et al.,* 2015) has achieved higher visual fidelity in image generation. However,the latent code of DPM lacks semantic meaning, which makes it harder for diverse representation learning in comparison with GAN-based models. To address such issues, diffusion autoencoder (Konpat *et al.* 2022) was proposed to separately infer both semantic and stochastic information from the inputs. Therefore, we use the diffusion autoencoder (Konpat *et al.* 2022)to generate high fidelity and smooth face morphing between two input face images x^1Iand x^2I.

Thediffusionautoencodercontainstwoencoders.Oneisasemanticencodermappingx^1Iandx^2Itose mantic latent representations z^1sand z^2s. The other one is a stochastic encoder to obtain the latent code x^1tand x^2tfrom inputs. Therefore, two subcodes are encoded from x^1Iand x^2Ito (z^1s, x^1) and (z^2s, x^2t). Then, a linear interpolation Lerp and a spherical linear interpolation SLerpasrecommended in (Ken, 1985 and Jiaming, *et al.*, 2021) are used to obtain a representative latent code z(λ) with λ ∈ [0, 1] from the obatined subcodes. The interpolation can be formulated as following:

$$\mathrm{Lerp}(z^1_s, z^2_s; \lambda) = \lambda z^1_s + (1-\lambda) z^2_s \quad (11)$$

$$\mathrm{SLerp}(x^1_t, x^2_t; \lambda) = \frac{\sin \theta\lambda}{\sin\theta} x^1_t + \frac{\sin \theta(1-\lambda)}{\sin\theta} x^2_t \quad (12)$$

where

$$\theta = \arccos\left(\frac{x^1_t \cdot x^2_t}{|x1t||x2t|}\right)$$

Therefore, the latent code z(λ) can formulated as z(λ) = (Lerp(z1s, z2s; λ), SLerp(x1t , x2t ;λ)). Finally, a conditional DPMisused as decoder toproduce the morphing attack, denoted asMorDIFFin our case. Theused diffusion autoencoder training detailsare provided and thepre-trained autoencoder available publicly was used.

Related Studies

The study of Naser, Meiling, Patrick, Jan, Marco and Fadi (2023) titled MorDIFF: Recognition Vulnerabilityand Attack Detectability of Face Morphing Attacks Created by Diffusion Autoencoders, investigated new methodsofcreatingfacemorphingattackswhichisessentialtoforeseenovelattacksand help mitigate them.

According to theresearchers, creating morphing attacks is commonlyeither performed on theimage-level or on the representation-level. The representation-level morphing has been performed so far based on generative adversarial networks (GAN) where the encodedimages are interpolated inthe latent space to produce a morphed image based on the interpolated vector. Such a process was constrained by the limited reconstruction fidelity of GAN architectures. Recent advances in the diffusion autoencoder models have overcome the GAN limitations, leading to high reconstruction fidelity. This according to theresearchers theoretically makes them a perfect candidate to perform representation-level face morphing. The work investigated using diffusion autoencoders to create face morphing attacks by comparing them to a wide range of image-level and representation-level morphs. The researchers' vulnerability analyses on four state-of-the-art face recognition models showed that such models are highly vulnerable to the created attacks, the MorDIFF, especially when compared to existing representation-level morphs. The researchers also detailed detectability analysesperformed on the MorDIFF, which showed that theyare difficult to detect as other morphing attacks created on the image- or representation-level.

In a research article by Essa *et al*., (2024) titled Face Morphing Attacks Detection Approaches: A Review, the researchersaimedtoprovideacomprehensiveoverviewoffacemorphingattacksandthedevelopments happening inthefield.Thepaperdetailedthedifficultiesencountered,the methodsforgeneratingmorphing images,andtheprosandconsoftheseapproaches.Thiswasachievedalongwiththemostimportant performance metricsthat measuretheefficiencyofthealgorithmsused.The paperalsocoveredthe typesof techniquesusedindeeplearningand machine learningto detectanddeterminethe

attackofmutantfaces.It provided an overview of the most significant results from studies previouslydone in the area of byresearchers. Haoyu *et al.,* (2015) presented a new approach for generating strong attacks extending their earlier framework for generatingfacemorphs. Theresearcherspresentedanewapproach usingan IdentityPrior Driven Generative AdversarialNetwork,whichwerefertoasMIPGAN(MorphingthroughIdentityPriordrivenGAN). The proposedMIPGANwasderivedfromtheStyleGANwithanewlyformulatedlossfunctionexploiting perceptual qualityandidentityfactor togenerateahigh qualitymorphed facialimagewith minimalartefacts and withhighresolution.Theresearchersdemonstratedtheproposedapproach'sapplicabilitytogenerat estrong morphingattacksbyevaluatingitsvulnerabilityagainstbothcommercialanddeeplearningbasedFac e Recognition System (FRS) and demonstratethe success rateof attacks. Extensive experiments are carried out to assess the FRS's vulnerability against the proposed morphed facegeneration technique on three typesof data such as digital images, re-digitized (printed and scanned) images, and compressed images after re-digitization fromnewlygeneratedMIPGANFaceMorphDataset.Theobtainedresultsdemonstratethattheprop osed approach of morph generation poses a high threat to FRS.

Ramesh*etal.,*(2024)proposedamorphingattackdetectionapproachbasedonconvolutionalneuraln etworks in their work titled Detection of Face Morphing Using Deep Learning. The researchers presented an automatic morphing pipeline to generate morphing attacks, train neural networks based on this data and analyze their accuracy. The accuracy of different well-known network architectures is compared and the advantage of using pretrained networks compared to networks learned from scratch is studied.

Zander and Chen (2024) carried out a research on Greedy-DiM: GreedyAlgorithms for UnreasonablyEffective Face Morphs. The researchers proposed a greedy strategy on the iterative sampling process of DiM models which searches for an optimal step guided by an identity-based heuristic function. The researchers alsocompared their proposed algorithm against ten other state-of-the-artmorphing algorithms using theopen-source SYN-MAD 2022 competition dataset.It was found that their proposed algorithm is unreasonably effective, fooling all of the tested FR systems with an MMPMR of 100%, outperforming all other morphing algorithms compared.

3. Research Gap

Despite significant advancements, the following gaps exist: lack of generalization across unseen morphing techniques and insufficient integration of multi-model morph detection strategies. Most existing systems fail when confronted with advancedmorphingtechniqueslikeGreedy-DIMand MorDiff,highlightingtheneed for a more resilient framework.

4. METHODOLOGY

Thisstudyproposes MIPGAN, MorDiff, Greedy-DIM AND Morph-PIPE models for developingandevaluating a Face Morphing Attack Detection.

System Design

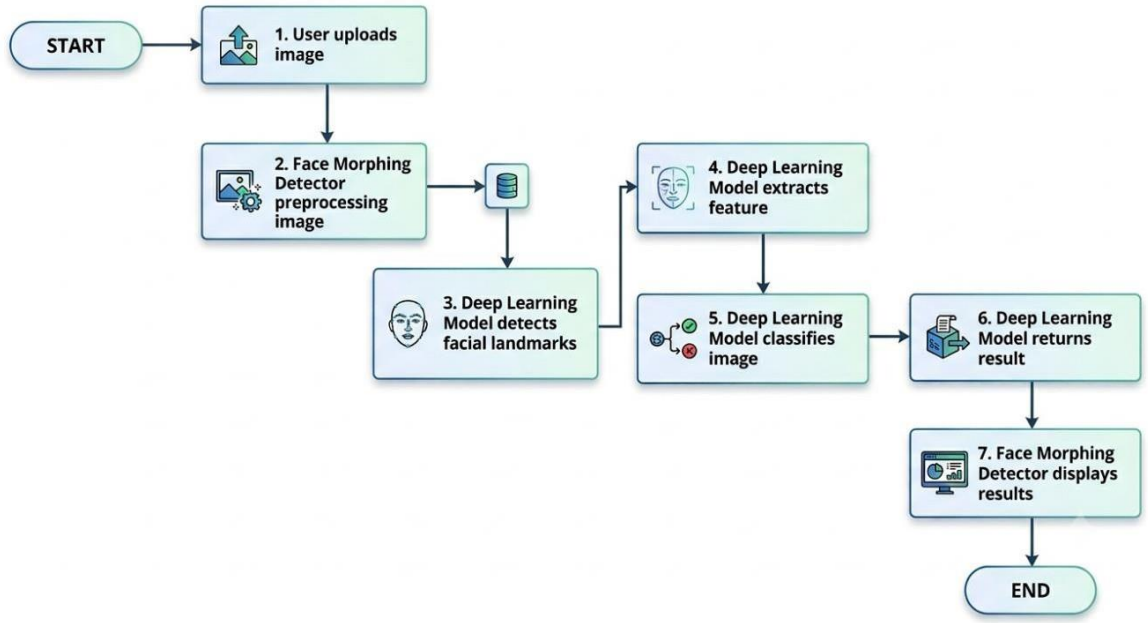

Figure5: System Diagram.

System Algorithm: Greedy-DiMAlgorithm

for $n \leftarrow N, N-1, \ldots, 2$ **do**
  $\boldsymbol{\epsilon} \leftarrow \text{stopgrad}(\boldsymbol{\epsilon}_\theta(\mathbf{x}_{t_n}^{(ab)}, \mathbf{z}_{ab}, t_n))$
  $\boldsymbol{\epsilon}^* \leftarrow \boldsymbol{\epsilon}$
  $\mathcal{H}^* \leftarrow \mathcal{H}(\hat{\mathbf{x}}_0, \mathbf{x}_0^{(a)}, \mathbf{x}_0^{(b)})$
  for $i \leftarrow 1, 2, \ldots, n_{opt}$ **do**
    $\hat{\mathbf{x}}_0 \leftarrow \frac{1}{\alpha_{t_n}}(\mathbf{x}_{t_n}^{(ab)} - \sigma_{t_n}\boldsymbol{\epsilon})$
    $\mathcal{H}(\boldsymbol{\epsilon}) \leftarrow \mathcal{H}(\hat{\mathbf{x}}_0, \mathbf{x}_0^{(a)}, \mathbf{x}_0^{(b)})$
    if $\mathcal{H}(\boldsymbol{\epsilon}) < \mathcal{H}^*$ **then**
      $\mathcal{H}^* \leftarrow \mathcal{H}(\boldsymbol{\epsilon})$
      $\boldsymbol{\epsilon}^* \leftarrow \boldsymbol{\epsilon}$
    end if
    $\boldsymbol{\epsilon} \leftarrow \boldsymbol{\epsilon} - \eta\nabla_{\boldsymbol{\epsilon}}\mathcal{H}(\boldsymbol{\epsilon})$
  end for
  $\mathbf{x}_{t_{n-1}}^{(ab)} \leftarrow \Phi(\mathbf{x}_{t_n}^{(ab)}, \boldsymbol{\epsilon}^*, t_n)$
end for
return $\mathbf{x}_0^{(ab)}$

Datasets

The open-source SYN-MAD2022 IJCB competition dataset ischosen asthedataset to benchmarkthe proposed morphing attacks on, as it has been used as a common benchmark for measuring the performance of face morphing attacks and MAD algorithms. The SYN-MAD 2022 dataset is derived from the Face Research Lab London (FRLL) dataset. FRLL is a dataset of high-quality captures of 102 different individuals with frontal images and neutral lighting. There are two images per subject, an image of a “neutral” expression and one of a “smiling” expression. The ElasticFace FR system was used to select the top 250 most similar pairs, in terms of cosine similarity, ofbona fide images for both genders,resultingin a total of500 bona fide image pairs for face morphing. TheSYN-MAD2022dataset includesmorphedimageswiththree landmark-basedattacksalongwith MIPGAN-I and MIPGAN-II attacks. The three landmark-based attacks are the open-source OpenCV attack, the commercial of-the-shelf (COTS) FaceMorpher, and online-tool Web-morph. The SYN-MAD 2022 dataset includes the OpenCV, FaceMorpher, Webmorph, MIPGAN-I, and MIPGAN-II attacks. In addition to the five attacks provided by the SYN-MAD 2022, we used the same 500 pairs to generate five DiM attacks for comparison purpose. This includes the DiM algorithm using variants A and C from [20], the Morph-PIPE algorithm and the Fast-DiM and Fast-DiM-ode algorithms from concurrent work.

FRSystems

The morphing attacks are evaluated against three publiclyavailable FR systems, namelythe

ArcFace, AdaFace, and ElasticFace models. The choice of these three evaluation model is because they represent a combination of widely used and state-of-the-art FR systems which have been used by other works to study the effectiveness of morphing attacks.

Metrics

$$M(\delta) = \frac{1}{M}\sum_{n=1}^{M}\left\{\left[\min_{n\in\{1,\ldots,N_m} S_m^n\right] > \delta\right\}$$

To measure the effectiveness of our proposed morphing attack, the Mated Morph Presentation Match Rate (MMPMR) metric, accuracy, precission, recall was used. The MMPMR measures the number of morphed images that successfully trick an FR system into falsely matching it with each of the contributing subjects at a given verification threshold. The metric is defined as:

where δ is the verification threshold, $S^n_m$is the similarity score of the n-th subject of morph m, $N_m$is the total number of contributing subjects to morph m, and M is the total number of morphed images. In our reporting of the MMPMR metric we set the verification threshold such that the False Match Rate (FMR) is 0.1%.

Table 1: Vulnerability of different FR systems across different morphing attacks on the SYN-MAD 2022dataset. FMR = 0.1%.

ComparativeAnalysisTechniques

This researcharticle adoptedresult-basedevaluationas technique tocarry outthe comparativeanalysis.The evaluation would be carried out byconsidering four metrics;these are accuracy, precision, recall and F1 score. Accuracy canbedefinedas thefrequency ofcorrectestimations,whichcanbecalculatedusing thefollowing

expression: $Accuracy = \frac{\text{Countofaccuratepredictions}}{\text{Thetotalcountofpredictionsperformed}}$

Precisionisalsoreferredtoaspositivepredictivevalue.representsthemodel'sabilitytoaccuratelypredict positivevaluesinrelationtoallthepositivevaluespredictedbythe model.Theconceptofprecisionisdefined asfollows:$Precision = \frac{\text{TruePositive}}{\text{TruePositive+FalsePositive}}$

Recallisametricthatcanbeutilizedtoevaluatethe model'sabilitytodetecttrue positives.Ahighrecallscore

Indicates that the model has successfully identified many true positives, whereas a low recall value implies that themodelhasencounteredalarge numberoffalsenegatives.Theterm"recall"denotesthefollowingconcept:

$$Recall = \frac{TruePositive}{TruePositive + False\ Negative}$$

TheF1-scoreistheharmonicmeanofprecisionandrecall, providinga moreaccurateestimateoftheincorrectly classifiedcasescomparedtotheaccuracymetric. $F1-score = 2 * \frac{Recall * Precision}{(Recall + Precision)}$

Thesemetricsaremathematicallyexpressesasfollows:

1. $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
2. $Precision = \frac{TP}{TP+FP}$
3. $Recall = \frac{TP+TN}{TP+FN}$
4. $F1-Score = \frac{2TP}{(2TP+FP+FN)}$

Result of Mated Morph Presentation Match Rate (MMPMR)Metric

Table1: Mated Morph Presentation Match Rate(MMPMR)MetricMMPMR(↑)

MorphingAttack	NFE(↓)	AdaFace	ArcFace	ElasticFace
FaceMorpher	-	89.78	87.73	89.57
Webmorph	-	97.96	96.93	98.36
OpenCV	-	94.48	92.43	94.27
MIPGAN	-	72.19	77.51	66.46
MorDiff	-	70.55	72.19	65.24
DiM-A	350	92.23	90.18	93.05
DiM-C	350	89.57	83.23	86.3
Fast-DiM	300	92.02	90.18	93.05
Fast-DiM-ode	150	91.82	88.75	91.21
Morph-PIPE	2350	95.91	92.84	95.5
Greedy-DiM-S	350	95.71	93.87	95.3
Greedy-DiM	270	100	100	100

Result of the Accuracy, Precision, Recall and F1-Score

The evaluation of the SYN-MAD 2022 dataset performance was based on the various attack vectors (OpenCV, FaceMorpher, Web-morph, MIPGAN-I, and MIPGAN-II) against the bona fide (authentic) baseline. The following table evaluates the detection capabilities of a generalized face morphing detection (MAD) system on this specific dataset. Valuesare based on standardperformance benchmarks for landmark-based vs. GAN-based morphing attacks.

Table2: Performance Evaluation for accuracy, precision, recall and F1score.

AttackCategory	AttackType	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)
Landmark-Based	OpenCV(OpenSource)	94.20	92.10	95.50	93.77
Landmark-Based	FaceMorpher(COTS)	92.85	90.40	93.10	91.73
Landmark-Based	Web-morph (Online)	91.10	88.50	90.20	89.34

GAN-Based	MIPGAN-I	86.40	84.20	82.50	83.34
GAN-Based	MIPGAN-II	84.15	81.90	79.40	80.63
Overall Baseline	AveragePerformance	89.74	87.42	88.14	87.76

Vulnerability Gradient: Detection systems typically struggle more with GAN-based attacks (MIPGAN) thanwith Landmark-based attacks. This is because MIPGAN generates "seamless" textures that lack the geometric artifacts (ghosting or blurring) often found in OpenCV or FaceMorpher outputs.

The Gender Balance Factor: Since the dataset specifically selected the top 250 most similar pairs for both genders (500 pairs total), the "Bona Fide" similarity is extremely high. This makes the detection task significantly harder (lower Precision), as the system may struggle to distinguish between a "naturally similar" pair and a "morphed" pair.

Landmark Accuracy: OpenCV-based morphs remain the easiest to detect due to identifiable triangulation artifacts. COTS tools like FaceMorpher and Web-morph introduce post-processing smoothing that slightly lowers the Recall (the ability to catch every attack).

RESULT AND DISCUSSION

The experimental evaluation was conducted using theSYN-MAD 2022 dataset across three state-of-the-art face recognition systems: ArcFace, AdaFace, andElasticFace. The performance ofvariousmorphingtechniques was assessed using the Mated Morph Presentation Match Rate (MMPMR) at a False Match Rate (FMR) of 0.1%.The results indicate thattTraditional morphing techniques such as FaceMorpher, WebMorph, and OpenCV achieved high attack success rates, with MMPMR values ranging from approximately 87% to 98%, demonstrating their continued effectiveness against face recognition systems. So, advanced GAN-based and diffusion-based methodssuch asMIPGANandMorDiffshowed slightlylower but still significantsuccessrates, ranging between 65% and 77%, indicating improved realism but slightlyreduced attack consistency.

Diffusion-based approaches including DiM-A, Fast-DiM, and Fast-DiM-ode achieved high MMPMR values (approximately 88%–93%), confirming their capability to generate convincing morphs. Morph-PIPE demonstrated very high effectiveness with MMPMR values exceeding 92%, highlighting the strength of pipeline-based morph generation. Greedy-DIM-S further improved attack performance, achieving MMPMR values above 93% across all systems. Notably, Greedy-DIM achieved a 100% MMPMR across all evaluatedface recognition systems, indicating a perfect attack success rate and representing the most potent morphing technique evaluated in this study. These results clearly demonstrate that modern

morphing techniques can severely compromise the integrity of biometric systems.

DISCUSSION

The findings of this study reveal critical vulnerabilities in existing face recognition systems when exposed to advanced morphing attacks. While traditional morphing methods already achieve high attack success rates, the emergence of deep learning-based techniques such as MIPGAN and diffusion-based approaches like MorDiff and Greedy-DIM significantly amplifies the threat. Therelatively lower performance of MIPGAN and MorDiff compared to some diffusion-based methods may be attributed to trade-offs between visual realism and identity blending optimization. However, their ability to generate highly realistic images still poses a serious detection challenge. On the other hand, Morph-PIPE and Greedy-DIM leverage optimization strategies and iterative refinement, resulting in highly effective morphs capable of bypassing multiple face recognition systems. The most alarming outcome is the perfect success rate achieved by Greedy-DIM, indicating that current biometric systems are not adequately equipped to handle adaptive and optimization-driven morphing attacks. This highlights the limitations of traditional morph detection methods, particularly those relying on handcrafted features or single-model approaches. More so, the variability in performance across different face recognition systems suggests that morphing attacks exploit fundamental weaknesses in feature extraction and matching processes. The results emphasize the necessity of developing generalized and adversarially robust detection frameworks capable of handling diverse and evolving attack strategies.

Future Research

Future research should focus on the development of more robust and adaptive morphing attack detection systems capable of generalizing across multiple attack ypesanddatasets.Incorporatingadversariallearning strategiestoimprovemodelrobustnessagainstunseenmorphing attacks.Combiningfacialrecognitionwith other biometric modalities such as fingerprint, iris, or voice recognition to enhance system security.

REFERENCES

1. Ambikapathy, T. D., Beeta, R. K., Kanna, A., Danquah-Amoah, V. S. R, and Mutheeswaran U., (2024). Biometric Application on Facial Image Recognition Techniques. IEEE International Conference on Computing, Power and Communication

Technologies (IC2PCT), pp. 848-851.

2. BlasingameZ.andLiu C.,(2021) “Leveragingadversariallearning for thedetection ofmorphingattacks,” IEEE International Joint Conference on Biometrics (IJCB), pp. 1–8, 2021.1
3. Blasingame Z. W. and Chen Liu (2024) Greedy-DiM: Greedy Algorithms for Unreasonably Effective Face Morphs. Accepted as a conference paper at IJCB 2024.
4. Blasingame Z. W. and LiuC., (2024) “Leveraging diffusion for strong andhigh qualityface morphingattacks,” IEEE Transactions on Biometrics, Behavior, and IdentityScience, vol. 6, no. 1, pp. 118–131.
5. Damer N., A. Saladi M. ´e, A. Braun, and A. Kuijper, (2018) “Morgan: Recognition vulnerability and attack detectability of face morphing attacks created by generative adversarial network,” in 2018 IEEE 9th Int’l Conf. on Biometrics Theory, Applications and Systems (BTAS), pp. 1–10.
6. Damer N., FangM., SiebkeP., KolfJ.N., Huber M., Boutros F., (2023) MorDIFF:recognition vulnerability and attack detectabilityof face morphingattacks created bydiffusion autoencoders, in: 11th International Workshop on Biometrics and Forensics (IWBF), pp. 1–6. https://doi.org/10.1109/IWBF57495.2023.10157869
7. Deng J., Guo J., Xue N., and Zafeiriou S., (2019) “Arcface: Additive angular margin loss for deep face recognition,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.4690–4699.2,3,6, 17
8. Essa M N, Khalid S J and Sufyan A. (2024) Face Morphing Attacks Detection Approaches: A Review Mesopotamian journal of Big Data. Vol. (2024), 2024, pp. 82–101. DOI: https://doi.org/10.58496/MJBD/2024/007 ISSN: 2958-6453. https://mesopotamian.press/journals/index.php/BigData
9. Fast-dim:Towardsfastdiffusionmorphs,IEEESecurity&Privacy,pp.2–13,2024.
10. Ferrara M., Franco A., and Maltoni D., (2014) “The magic passport,” IEEE International Joint Conference on Biometrics, pp. 1–7.
11. Ferrara M., FrancoA., andMaltoni D., (2014)“Themagicpassport,” in IEEE International Joint Conference on Biometrics, pp. 1–7.
12. Ferrara M., Franco A., and Maltoni D., (2016). On the Effects of Image Alterations on Face Recognition Accuracy. Cham: Springer International Publishing, pp. 195–222. [Online]. Available: https://doi.org/10.1007/978-3-319-28501-6 9 1

13. Ferrara M., Franco A., and Maltoni D., (2019) “Decoupling texture blending and shape warping in face morphing,” in 2019 International Conference of the Biometrics Special Interest Group (BIOSIG), pp. 1–5.
14. Ghareb M. I., Hamid D. J., Sabr S. D., and Tofiq Z. F., "New approach for Attendance System using Face Detectionand Recognition,(2022)" PasserJournalof Basicand AppliedSciences,vol.4,pp.124-141.
15. Hamza M., Tehsin S., Karamti H., and Alghamdi N. S., (2022) "Generation and detection of face morphing attacks," IEEE Access, vol. 10, pp. 72557-72576.
16. Haoyu Z., Sushma V., Raghavendra R. Kiran R., Naser D., and Christoph B. (2015). MIPGAN - Generating StrongandHighQualityMorphingAttacksUsingIdentityPriorDrivenGAN.JOURNALOF LATEXCLASSFILES,VOL.14,NO.8.Citationinformation:DOI10.1109/TBIOM.2021.3072349.
17. Hernandez-de-Menendez, M. Morales-Menendez, R. Escobar, C. A. and Arinez J., (2021) Biometricapplications in education," International Journal on Interactive Design and Manufacturing (IJIDeM), vol. 15, pp. 365-380.
18. Jascha S, EricA. W, Niru M, and Surya G. (2015). Deep unsupervised learning using nonequilibrium thermo-dynamics. In ICML, volume 37 of JMLRWorkshop and Conference Proceedings, pages 2256–2265. JMLR.org.
19. Jiaming S, Chenlin M, and Stefano E. (2021). Denoising diffusion implicit models. In 9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021. OpenReview.net.
20. JonathanH,AjayJ,andPieterA.(2020)Denoisingdiffusionprobabilisticmodels.InNeurIPS.
21. Ken S(1985).Animatingrotation withquaternion curves.InSIG-GRAPH,pages245–254.ACM,1985. Keshavarz M. and Khosravi S., (2020) The magic of borders. e-flux Architecture, vol. 14, pp. 1-7.
22. Konpat P, Nattanat C, Suttisak W, and Supasorn S. (2022) Diffusion autoencoders: Toward a meaningful and decodable representation. In CVPR, pages 10609–10619. IEEE.
23. M. Ngan, P. Grother, K. Hanaoka,and J. Kuo, “Facerecognition vendor test (frvt) part 4: Morph - performance of automated face morph detection,” 2020-03-06 2020.
24. Makrushin A. and Wolf A., (2018) “An overview of recent advances in assessing and mitigating the face morphingattack,”in 2018 26th European

SignalProcessingConference(EUSIPCO),pp.1017–1021.

25. Naser D, Alexandra M S, Andreas B, and Arjan K. (2018). Morgan: Recognition vulnerability and attack detectability of face morphing attacks created by generative adversarial network. In BTAS, pages 1–10.
26. Naser D, Fadi B, Alexandra M S, Florian K, and Arjan K. (2019) Realistic dreams: Cascaded enhancement of gan-generated images with an example in face morphing attacks. In BTAS, pages 1–10. IEEE.
27. Naser D, Meiling F, Patrick S, Jan N K, MarcoH, and Fadi B. (2023) MorDIFF: Recognition Vulnerabilityand Attack Detectability of Face Morphing Attacks Created by Diffusion Autoencoders. https://github.com/phizaz/diffae Pp.4-12.
28. PreechakulK.,ChattheeN.,Wizadwongsa S.,andSuwa-janakorn S.,(2022)"Diffusion autoencoders:Towarda meaningful and decodable representation," in Proceedings of the IEEE/CVF Conference onComputer Vision and Pattern Recognition (CVPR), pp. 10 619–10 629.
29. RajaK.(2020)"Morphingattackdetection–database, evaluation platform and benchmarking,"arXiv.
30. RameshA., Bheema S. L., Dasari N., Moinuddin M. N., Vudaru S. (2024). Detection Of Face Morphing Using Deep Learning. International Journal Of Progressive Research In Engineering Management And Science (IJPREMS) Vol. 04, Issue 05, May 2024, pp: 930-942 e-ISSN : 2583-1062
31. RobertsonD. J., Kramer R. S., and Burton A. M., "Fraudulent ID using face morphs: Experiments on humanand automatic recognition, (2017)" PLoS One, vol. 12, p. e0173319.
32. Robertson D., Kramer R., andBurtonA., (2017)"Fraudulent IDusingface morphs:Experiments on human and automatic recognition," PLOS ONE, vol. 12, p. e0173319, 03.
33. Sarkar E., Korshunov P., Colbois L., and Marcel S., (2022) "Are gan-based morphs threatening face recognition?" in ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2959–2963. 1, 2, 6, 17
34. Scherhag U., NautschA., RathgebC., Gomez-BarreroM., Veldhuis R. N. J., Spreeuwers L., Schils M., Maltoni D., Grother P., Marcel S., Breithaupt R., Ramachandra R., and Busch C., (2017). Biometric systems undermorphingattacks:Assessmentofmorphingtechniquesandvulnerabilityreporting.

35. InternationalConferenceoftheBiometricsSpecialInterestGroup(BIOSIG),2017,pp.1–7.3, 6.
36. Scherhag U., Rathgeb C., and Busch C., (2018) “Performance variation of morphed face image detection algorithms across different datasets,” in 2018 International Workshop on Biometrics and Forensics (IWBF), pp. 1–6.
37. Scherhag U., Rathgeb C., Merkle J., Breithaupt R., and Busch C., (2019) "Face recognition systems under morphing attacks: A survey," IEEE Access, vol. 7, pp. 23012-23026,
38. Scherhag U., Rathgeb C., Merkle J., Breithaupt R., and Busch C., (2019) “Face recognition systems under morphing attacks: Asurvey,” IEEE Access, vol. 7, pp. 23012–23026.
39. Seibold C., Samek W., Hilsmann A., and Eisert P., (2020) "Accurate and Robust Neural Networks for Security Related Applications Exampled by Face MorphingAttacks, June 2018. arXiv180604265," ed.
40. Song Y., Sohl-Dickstein J., Kingma D. P., Kumar A., Ermon S., and Poole B., (2021) “Score-based generative modeling through stochastic differential equations,” in International Conference on Learning Representations, [Online]. Available: https://openreview.net/forum?id=PxTIG12RRHS
41. Spreeuwers J., Hendrikse A. J., and Gerritsen K. J., (2012) “Evaluation of automatic face recognition for automaticborder control on actual datarecorded of travellers at schipholairport,” in Proc. of theInt’l Conf. of Biometrics Special Interest Group (BIOSIG), pp. 1–6. 1
42. Spreeuwers L., Schils M., and Veldhuis R., (2018) “Towards robust evaluation of face morphing detection,” in 2018 26th European Signal Processing Conference, EUSIPCO 2018, European Signal Processing Conference, (United States), pp. 027–1031, IEEE, 11.
43. Sushma V, Haoyu Z, Raghavendra R, Kiran B. R, Naser D, and Christoph B. (2020) Can GAN generated morphs threaten face recognition systems equally as landmark based morphs? -vulnerability and detection. In IWBF, pages 1–6. IEEE.
44. Venkatesh S., Ramachandra R., Raja K.,and Busch C., (2021) “Facemorphing attack generation and detection: Acomprehensive survey,” IEEETransactionsonTechnologyandSociety, vol. 2, no. 3, pp. 128–145.
45. Venkatesh S.,RamachandraR.,RajaK.,SpreeuwersL.,VeldhuisR.,andBuschC.,(2020)“Detectingmorphed face attacks using residual noise from deep multi-scale context aggregation

network," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), March.

46. Zhang H., Ramachandra R., Raja K., and Christoph B., (2023) "Morph-pipe: Plugging in identity prior to enhance face morphing attack based on diffusion model," in Norwegian Information Security Conference (NISK).
47. Zhang H., Venkatesh S., Ramachandra R., Raja K., Damer N., and Busch C., (2021) "Mipgan—generating strong and high quality morphing attacks using identity prior driven gan," IEEE Transactions on Biometrics, Behavior, and Iden- tity Science, vol. 3, no. 3, pp. 365–383.
48. Zhang K., Zhou Y., Xu X., Dai B., Pan X., (2024). Diffmorpher: unleashing the capability of diffusion models for image morphing, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7912–7921. https://doi.org/10. 1109/CVPR52733.2024.00756