Methods to Prevent Registration Using Fake Face Images

Luis Cárabe
Departamento de Ingeniería Informática
Universidad Autónoma de Madrid
Madrid, Spain
e-mail: luiscarabe@gmail.com

Eduardo Cermeño
Research Department
Vaelsys
Madrid, Spain
e-mail: eduardo.cm@vaelsys.com

***Abstract*—Face identification is increasingly being used to register and access specific applications and online services. This opens up new possibilities for malicious attacks, such as users registering multiple times with different images or impersonating other users. Morphing is often the preferred method for these attacks as it allows the physical features of a subject to be progressively modified to resemble another subject. Publications focus on impersonating this other person, usually someone who is allowed access to a restricted area or software app. However, there is no such list of authorized people in many other applications, just a blacklist of people who cannot enter, log in, or register again. In such cases, the morphing target person is not relevant as the criminal's main objective is to minimize the probability of being detected. We present a comparison of the identification rate and behavior of 5 recognizers (Eigenfaces, Fisherfaces, Local Binary Patterns Histograms, Scale-invariant Feature Transform, and FaceNet) against morphing attacks. We also show the performance that a morphing detector could achieve. We prove that the use of FaceNet along with a morphing detector is an optimal resource to maintain a high level of security, identification rate, and attack detection.**

***Index Terms*—Access control; biometrics; deep learning; FaceNet; face recognition; identification; morphing; security; spoofing attack.**

I. Introduction

Face recognition is gaining momentum. Continuous improvements in this well-known research field [1][2][4][5][12] have led to an increasing number of commercial applications. Many sectors have found this technology the perfect match for their security concerns and requirements. Face recognition is used in a wide range of processes: sign up, log in, ID verification, and more broadly in any application that needs to comply with "Know Your Customer" policies.

Like in any other biometric technology, people have tried to deceive face recognition systems [19]. We can find several approaches in the literature. For instance, using a print of a photograph of a subject might allow someone to impersonate as that subject [19][20]. A well-known technique to try to fool face recognition systems is morphing. Morphing techniques consist of generating intermediate frames between two images to achieve a smooth transition between them. If we use it on two images of different faces, we could get frames that merge features of both faces in one. Depending on the level of morphing being applied, one person will be recognized better than the other. In the context of Automated Border Control (ABC), Ferrara *et al.* [3] studied a way to take advantage of morphing to use only one photo ID to verify two different subjects successfully.

The verification process differs from the identification one because the former is a one-to-one matching with only one possible output: match or mismatch. On the contrary, the latter is a one-to-many matching where an image is presented to a face recognizer that compares it against all the stored subjects in its database and outputs the closest match or a top matches list. When morphing an original subject's image to attack a verification system, it is necessary to care about the person's identity recognized by the face recognition algorithm as it must be the target subject. Whereas in the attack to a face identification system, we only need to make sure that the original subject is not identified correctly, it is not relevant who the system thinks the image belongs to, as far as it is not the original subject. This increases the chances of a successful attack because the attacker can reduce the morphing level applied. It is not required to make it look like somebody, but change the image enough to make the face recognition system fail.

In this work, we study the behavior of different face recognition techniques with morphed images. Our aim is to find the most robust one, considering robustness as the quality of requiring a higher amount of morphing alteration to misclassify a subject. We resort to morphing detectors, algorithms designed to detect whether an image is the result of a morphing process and if they can, therefore, be used to endorse a face recognition algorithm against morphing attacks. Furthermore, we analyze the value of implementing a morphing detector along with the face identification algorithm to build a stronger solution that can be used for registration processes or similar ones.

In Section II, we present a brief review of past spoofing attacks to face recognition algorithms and spoofing detection methods. In Section III, we describe the morphing, face recognition, and morphing detection methods used in our study. In Section IV, we describe the scenario of our experiments and the implementation of the methods and database used. In sections V and VI, we present the results of the experiments and their discussion. Finally, in Section VII, we

 ISBN: 978-1-61208-862-4

make conclusions about the findings of our experiments.

II. Related work

We have divided the section into two subsections spoofing attacks and spoofing detection.

A. Spoofing attacks

Spoofing attacks can be undertaken under different approaches. Hadid *et al.* [19] and Mohammadi et al. [20] explore several databases with *presentation attacks*. These attacks consist of showing a printed image (or printed mask) to a camera with facial recognition software to fool it. Apart from this, Ferrara *et al.* [21] study the effects of geometric distortions (barrel distortion, vertical contraction, and extension) and digital beautification on face recognition accuracy. Other digital manipulation techniques can be very harmful, e.g., face synthesis, attribute manipulation, and identity or expression swap [22].

Ferrara *et al.* [3] were the first to present a successful morphing attack in a simulation of an ABC control, using two commercial face recognition software tools. They manually created morphed images to verify the two contributing subjects with the same photo. They were able to achieve that for eleven pairs of subjects in both face verification tools. Ferrara *et al.* [21] expand the experiment proving that human experts (border guard group) and non-experts, in most cases, do not detect morphed images. However, Robertson *et al.* [23] reveal that although the attack may go more unnoticed in untrained subjects, when the subjects receive morphing training, they tend to detect morphing with higher probability. Wandzik *et al.* [24] and Scherhag *et al.* [25] present more examples of verification attacks. In the first one, they carried out the experiment using FaceNet, utilizing more than 3000 pairs with 22 morphed images between each pair, working with triplets of images (impostor-accomplice-morphing). In the second one, experiments were conducted to prove face verification's vulnerability both with printed and scanned images.

B. Spoofing detection

Galbally *et al.* [33] present a survey on hardware-level and software-level methods to detect presentation attacks in images and videos. Hadiprakoso *et al.* [34] and Wu *et al.* [35] present more recent studies. In the first one, they combine a Convolutional Neural Network (CNN) analysis with face liveness detection module to be able to detect static and dynamic attacks, such as masks, photos, or video replays. In the latter, they compare the performance of some methods to detect spoofing attacks.

Focusing on morphing, the first detector was presented by Raghavendra *et al.* [36], which successfully verified all the 450 morphed face images from a database. Additional approaches can be found in [37]–[40]. In order to detect morphing successfully, the authors use different techniques, such as Fourier spectrum of sensor pattern noise, Local Binary Pattern (LBP), or a *demorphing* process. Scherhag *et al.* [16] and Raja *et al.* [41] present a review of these methods, along with others.

III. Methods

We have divided this section into three subsections: Morphing attacks, Face Recognition and Morphing Detection.

A. Morphing attacks

The first method used in our study is the morphing attack. A morphing attack is the alteration of a subject's portrait using morphing techniques leading to his misidentification.

Most of the morphing methods found in the literature [16] are based on Delaunay triangulation [18][28]. It includes three stages: feature specification, warping, and blending. In the first step, a correspondence between the two images is created by determining the face key landmarks (eyes, mouth, nose, face contour, etc) either manually or automatically (using software). Then, a Delaunay triangulation is applied using the landmarks as vertices for the non-overlapping triangles. During warping [15], the corresponding triangles of both images suffer a geometrical transformation in order to be aligned. The last step requires to merge each pixel's color value, where a linear blending is applied.

At the warping and blending steps of the process, a parameter α is taken into account. In the case of warping, it conditions how much each position of each face's landmarks contributes to the morphed image. If $\alpha = 0$, only the first image's landmarks are taken into account. If $\alpha = 1$, only the landmarks of the second image are considered. The in-between values achieve a linear combination of the positions of the landmarks of both contributing images. That is to say, if l_r represents the landmark positions of the resulting image and $l_{0,1}$ the landmark positions of the first and second images:

$$l_r = (1 - \alpha)l_0 + \alpha l_1.$$

The blending step has a similar behavior. The color of all the correlated pixels are combined using a linear transformation. $\alpha = 0$ only considers the first image and $\alpha = 1$ the second. If c_r represents the color of the pixels of the resulting image and $c_{0,1}$ the colors of the pixels of the first and second images:

$$c_r = (1 - \alpha)c_0 + \alpha c_1.$$

α is used as a quantifier of the morphing process. For example, a morphing process (amount) of 5% means that $\alpha = 0.05$. The first subject of the pair will contribute to the final image by 95% in both the landmarks' position and the pixels' value. The second subject will contribute with the remaining 5%.

B. Face recognition

A key component of any user registration system using faces is the face recognition algorithm. There are different approaches in the literature that can be classified into four families: holistic, local, hybrid and deep learning [4][5]. The local approach classifies according to specific facial features, whereas the holistic approach considers the whole face as a unit. The hybrid approach combines both techniques. Many recent advances have been made in the deep learning approach, using CNNs that offer better speed and accuracy.

 ISBN: 978-1-61208-862-4

We have selected the more promising ones with care to include at least one from each category (except hybrid, due to its high complexity [4]).

1) Holistic: In the holistic approach, we have selected Eigenfaces [6] and Fisherfaces [7]. Eigenfaces is based on the Principal Component Analysis (PCA) technique. It tries to reduce the dimensionality of the data space by projecting the face images into a subspace called feature space. It also tries to find a basis of that subspace for the dataset. This is achieved by finding the eigenvectors (referred to as eigenfaces) of the covariance matrix of the set of faces. The resulting eigenfaces form the basis of the feature space. To identify faces, the testing image is projected into this subspace using a linear combination of the eigenfaces basis.

Fisherfaces has the same objective as Eigenfaces: reduces dimensionality. Nevertheless, instead of using only an unsupervised technique (PCA), it also uses Linear Discriminative Analysis (LDA), which works with a supervised learning technique. The LDA technique attempts to model the difference between two distinct classes (individuals). That is, by using scatter matrices, it tries to find a linear combination of features that separate two or more classes. This method achieves excellent results even with severe illumination changes.

2) Local: In the local category, we have chosen Local Binary Patterns Histograms (LBPH) [30] and Scale-invariant Feature Transform (SIFT) [9]. The LBPH algorithm works by creating histograms of the binary patterns extracted by LBP [8]. Those binary patterns are obtained as follows: First, the image (in gray scale) is divided into 3x3 pixel regions. Then, for each region, the central pixel's value is taken as a reference, which will act as a threshold for the neighboring pixels. We look at the value of each pixel in the grid, if it is above the threshold (the value of the central pixel), it is assigned a 1. If it is below, a 0. Then the binary values are concatenated, and the result is assigned to the central pixel. To classify an image, it finds its closest histogram from the training database.

SIFT generates image features that are highly distinctive and invariant to certain transformations, such as translation, scaling, and rotation. To obtain those features, the algorithm first tests different image scales, looking for invariant key points. Then, among all the key points obtained, the most stable ones are selected. Meaning that those with the highest sensitivity to noise (points with low contrast) and those located on edges are discarded. Later, the algorithm assigns one or more orientations to each key point, based on the directions of the local gradient of the image, achieving rotation invariance. Finally, each key point is assigned a feature descriptor, ensuring that they are highly distinctive and invariant to lighting changes.

3) Deep learning: In the deep learning group, we have chosen FaceNet [10]. It uses convolutional layers to create a 128-dimensional embedding for every image. The FaceNet model is trained with a Triplet Loss technique. It selects combinations of three images: two images from the same subject (one image is called the anchor and the other one the positive input), and another image from a different subject (negative input). The Triplet Loss tries to minimize the anchor's embeddings distance with the positive input and maximize it with the negative input. Once the model is trained, FaceNet can compute the 128-dimensional embedding for each image in our training database. In the face identification process, FaceNet will return the subject whose embeddings are most similar to those obtained in the testing image.

As seen in [4][5][7][9][26], all of these face recognition techniques have been well studied and have good performance when using frontal views of faces.

C. Morphing detection

Apart from observing how the recognizers behave against morphing, it may be interesting to consider a morphing detector capable of classifying images as morphed or bonafide (unaltered).

We have selected a morphing detector that operates in Single Image Morphing Attack Detection scenarios (S-MAD). It refers to algorithms that only analyze one photograph to check its morphing. In contrast, Differential Morphing Attack Detection (D-MAD) groups algorithms that analyze a pair of images, one of them being a trusted unaltered photograph that the algorithm uses to verify the morphing on the other image. Our scenario falls into the first category since we only provide one image (the one that the subject uses to access) to the detector to get a morphing verification.

IV. Experiments

The experiments found in the literature do not take into consideration morphing attacks against face identification. We wish to study the approach that performs better against these attacks from two perspectives: a basic one, where we analyze the performance of the recognizers in correct subject identification, and, an advanced one where we study the ability to detect fraudulent registrations.

In the first case, from our point of view, good performance means that the algorithm can correctly identify the original subject in images that have been morphed. Since morphing is an incremental process, the most robust algorithm should be the one requiring the highest amount of morphing to force its failure. Therefore, the selection criteria should be based on the first frame where the face recognition algorithm does not recognize the original subject but another (either the target subject or any other person). The higher the alteration percentage required to avoid the correct identification by the recognizer, the more robust it has to be considered.

In our study, the original image (first contributing subject) is morphed into 100 images with $n\%$ morphing ($n \in \{1, .., 100\}$). We consider that the original image has been morphed 0%, the target image (second contributing subject) has been morphed 100%, and any other image in between has $n\%$ ($n \in \{1, .., 99\}$) as the amount of morphing.

Regarding the advanced scenario, we aim to study which recognizer is better to prevent multiple registrations of the same subject. For this purpose, we assume that a recognizer will accept a person as a new record when it has 0 identified

 ISBN: 978-1-61208-862-4

subjects with confidence above a threshold. If there are subjects identified, the person will be rejected. This gives us two rates. The False Acceptance Rate (FAR, impostors being able to register again), and the False Rejection Rate (FRR, genuine subjects not being able to be registered for the first time). As the FRR decreases and the FAR grows with the threshold, the best performing recognizer will be the one whose FRR decreases earliest and whose FAR grows latest.

Moreover, we are going to study the performance of the morphing detector applied to both perspectives.

A. Implementation

1) Morphing: For the morphing implementation, we have used the Python code presented by Patel [27], based on OpenCV functions [17][18]. To find the face landmarks, it uses Dlib's facial landmark detector [29]. Then, as we have seen, those landmarks are employed as vertices of the Delaunay triangles. Using the corresponding triangles, it performs warping and blending to obtain all the intermediate frames.

TABLE I
CLAIMED ACCURACY OF THE SELECTED RECOGNIZERS.

Category	Recognizer	Accuracy (database)
Holistic	Eigenfaces	97.5% (ORL) [26]
	Fisherfaces	92.7% (Yale) [7]
Local	LBPH	76% (FERET) [30]
	SIFT	84.03% (BANCA) [9]
Deep Learning	FaceNet	99.63% (LFW) [10]

2) Face recognizers: Table I shows the accuracy claimed for all the selected recognizers. For the first three face recognition algorithms (Eigenfaces, Fisherfaces, and LBPH), we have employed a Python implementation of Raja [31] that uses the Face library of OpenCV to cover the feature extraction and classification. Besides, a Haar cascade classifier is used for face detection. Slightly modifying the previous implementation, we have gotten a SIFT deployment, using the xfeatures2d OpenCV class to perform the SIFT feature extraction and the Scikit-learn library for classification using a Support Vector Machine (SVM). Moreover, we have used a Tensorflow implementation of FaceNet [32] written in Python. It uses a pre-trained model that employs VGGFace2 as the training dataset and the Inception-ResNet-v1 architecture, achieving an accuracy with the verification problem in the Labeled Faces in the Wild database (LFW) [11] of 99.65+-0.00252%. It also uses an SVM for classification.

The testing subjects are to be included in all the recognizers training database, what is known as closed-set identification. In order to get similar behavior in all the implementations, we introduced small changes in the code files. Every algorithm used can output the top 5 identification matches of the face presented. The parameters of the Haar cascade classifier that worked better with our database were *scaleFactor=1.001, minNeighbors=2, minSize=(90,90), outputRejectLevels=True*. Regarding the SVM used on SIFT, we have employed the settings *kernel="poly", C=10, gamma=0.0001*. We have left all the other configurations according to the original sources.

3) Morphing detector: Regarding the morphing detector, we have tried the algorithms of [37]–[39]. The one that had the best performance and integration in our scenario has been the detector presented by Raghavendra *et al.* [38], which has better results than other state-of-the-art alternatives. Although it is designed to detect morphing in printed-scanned photographs, it achieves excellent detection results in our context (see Figure 2), and therefore, it is the morphing detector used.

B. Database

1) Basic scenario: We recommend that face recognition algorithms should be trained with a database composed of N subjects, with a number of photos per subject between 5 and 20. This quantity helps to avoid imbalanced data and biased results. To test the morphing, we have chosen pairs of similar-looking subjects. This should reduce the amount of alteration required to pass from the original image (referred to as A) to the target image (referred to as B).

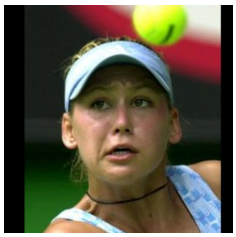
We have created a database based on LFW [11]. As seen in [5], it is a widely used unconstrained database to test state-of-the-art face recognizers. Usually, algorithms struggle with lighting, location, setting, pose, or age variations, as well as occlusions or misalignment [12]–[14]. However, over time, algorithms have improved significantly in this area.

The database has 5749 subjects, but, as stated above, we want only the ones that have between 5 and 20 images each (both numbers included). That filters the database to 366 people with a total number of 3062 images. The Haar cascade face detector does not correctly detect the subject face in 5 of the 3062 images because those images have more than one face present and the wrong face is detected. We deleted those images from the database. The deleted images are *Erika_Harold_0003*, *Hugh_Grant_0008*, *Igor_Ivanov_0014*, *Jean_Charest_0004*, and *Joe_Lieberman_0004*. That implies that Erika Harold now has four images instead of 5, considering this an exception.

TTo determine the pairs of subjects who look more alike, we have used the Similar-looking LFW database (SLLFW) [42], which offers 3000 pairs of similar-looking faces (using the images of LFW). We have picked 25 pairs of images from it, taking into account two factors: first, the individuals must be included in our 366 subjects database; second, once the similar-looking images selected are removed from the training database, the subjects need to have more than five photos to train. Figure 1 shows an example of one selected pair.

Considering all the pairs, there are 49 different images (*Renee_Zellweger_0009* appears twice). The training database of the recognizers consists of $3062 - 5 - 49 = 3008$ images of 366 subjects. In Table II, we provide all the pairs used.

2) Registration scenario: Regarding the advanced perspective, we use the same training database seen in the previous section. For the testing, we need two groups of subjects: impostors and genuine ones. Considering the first case, we have used the 49 different subjects (already registered) seen in Table II. We have randomly morphed them with people not included in the training database (LFW subjects with n images,



(a) Anna_Kournikova_0011.

(b) Jelena_Dokic_0007.

Fig. 1. Similar-looking pair.

TABLE II
SIMILAR-LOOKING PAIRS SELECTED.

No.	Original subject	Target subject
1	Amelia_Vega_0003	Norah_Jones_0015
2	Ana_Guevara_0002	Ian_Thorpe_0006
3	Andy_Roddick_0008	Richard_Virenque_0004
4	Angelina_Jolie_0002	Britney_Spears_0004
5	Anna_Kournikova_0011	Jelena_Dokic_0007
6	Ben_Affleck_0002	Ian_Thorpe_0007
7	Bill_McBride_0010	Jon_Gruden_0002
8	Bill_Simon_0011	Ron_Dittemore_0001
9	Catherine_Zeta-Jones_0001	Salma_Hayek_0001
10	Edmund_Stoiber_0004	John_Snow_0003
11	Eduardo_Duhalde_0006	George_HW_Bush_0005
12	Fidel_Castro_0018	Mohamed_ElBaradei_0003
13	Hillary_Clinton_0010	Renee_Zellweger_0009
14	Howard_Dean_0003	Kevin_Costner_0005
15	James_Blake_0006	Mark_Philippoussis_0003
16	Jason_Kidd_0003	Leonardo_DiCaprio_0003
17	Jean-Pierre_Raffarin_0001	Joschka_Fischer_0012
18	Jimmy_Carter_0006	John_Snow_0004
19	Joan_Laporta_0007	Pierce_Brosnan_0006
20	John_Kerry_0005	Robert_Redford_0002
21	Julianne_Moore_0019	Nancy_Pelosi_0002
22	Kate_Hudson_0008	Mariah_Carey_0006
23	Matthew_Perry_0007	Rubens_Barrichello_0011
24	Mike_Martz_0005	Paul_ONeill_0003
25	Renee_Zellweger_0009	Sheryl_Crow_0001

$n < 5$ *or* $n > 20$), selecting arbitrarily, for each subject, nine morphed images (between 1% and 80% of alteration) and the unaltered image. The impostors database has $49*(9+1) = 490$ images. We have selected 490 unaltered images of different subjects not included in the training database used for the genuine subjects.

3) Morphing detector: To train and test the morphing detector, we have picked the LFW subjects' images not used in the other experiments. We have split the subjects randomly into two groups, one for testing and the other one for training. Due to Matlab memory limitations (we have used Matlab Online to train the model, which provides up to 16 GB of RAM [43]), we have trained the detector using 3000 bonafide (not altered) images from the training group and 3500 morphed images. The morphed images were created randomly using pairs from the subjects included in the training group, covering all percentages between 1 and 99. Analogously, we have tested the detector using 500 bonafide images and 500 morphed images. Figure 2 represents the Receiver Operating Characteristic (ROC) curve obtained, showing the excellent performance achieved.

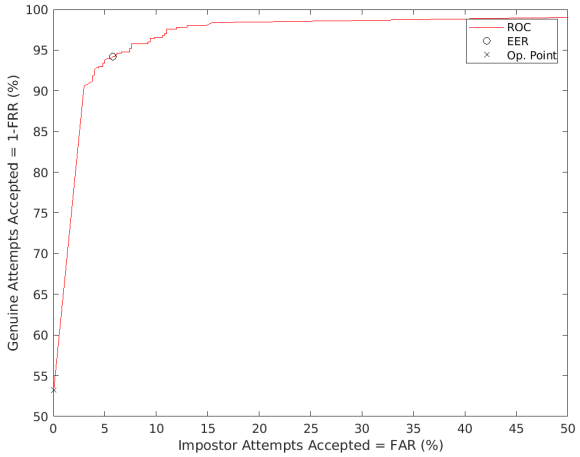


Fig. 2. ROC curve of the morphing detector.

V. RESULTS

Figure 3 shows the face identification algorithms' robustness results in the basic scenario (correct identification). It is divided into three plots. Figure 3a exhibits the face recognizers' comparison analyzing the top 1 identification matches. Figure 3b analyzing the top 3. Figure 3c the top 5. Their x-axes represent the level of morphing in the pairs. 0% morphing symbolizes the unaltered image of the first subject of the pair (original subject), 100% the second subject, and the rest of percentages the in-between morphings. Their y-axes reflect the percentage of couples who still have their original subject identified within the top analyzed for each morphing level.

It has to be observed that the identification percentages rise as we increase the top analyzed. However, the three graphs show similar robustness ranking.

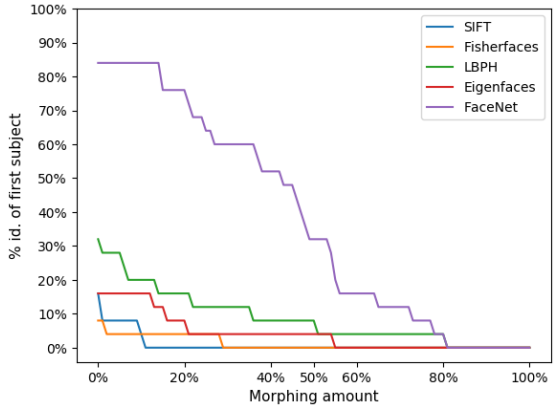
For each face recognizer, we have elaborated a table that shows the average confidence percentages outputted when the original subject is included in the top 1. The first row (*Morph*) shows the most relevant morphing percentages. Rows *1–5* show the first five identified subjects' average confidence.

For each recognizer, confidences have been normalized taking 100% as the best result obtained in our experiments (when the subject is correctly identified), and 0% as the confidence obtained in the last recognition position in images of subjects not included in the training database.

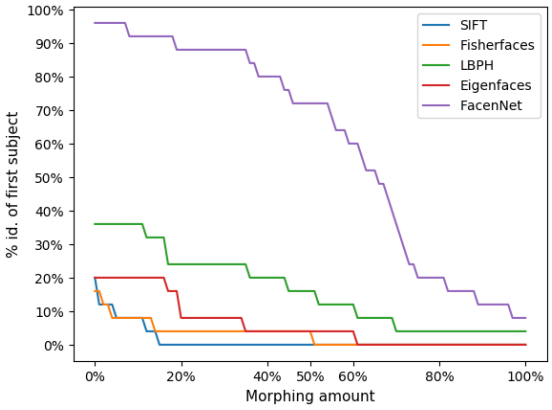
Also, we have included the FAR vs FRR plot of the advanced scenario, with and without using the morphing detector (mor. det.) to filter the accepted subjects. In the first case, we accept a subject (as a new register) when there are no identified subjects above the confidence threshold (x-axis). In case of having a morphing detector, to accept a subject, the previous condition must be met, and the morphing detector must output less than 50% of morphing confidence. Otherwise, the subject will be rejected.

In the following subsections, we describe the performance achieved by each method in both scenarios.

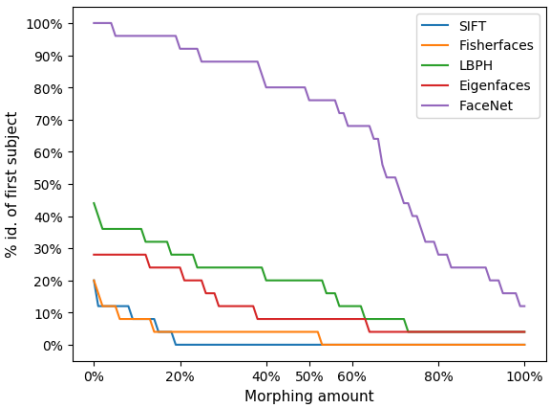
 ISBN: 978-1-61208-862-4



(a) Top 1.



(b) Top 3.



(c) Top 5.

Fig. 3. Percentage of morphed images identified as the original subject for each level of morphing.

A. *FaceNet*

It achieves the best identification scores for each top, with 0% morphing, being 84%, 96%, and 100%, respectively. FaceNet manages to maintain a high identification rate even with a considerable morphing alteration. For instance, at 50% morphing, it achieves 32%, 72% and 76% identification of the original subject for each top. Looking at the top 3 and 5, it even identifies more than 8% of the images with the original subject totally transformed (100% morphing). FaceNet takes the longest time (most significant morphing alteration) to misidentify the original subject.

TABLE III
AVERAGE CONFIDENCE PERCENTAGES OF FACENET WHEN THE ORIGINAL SUBJECT IS INCLUDED IN THE TOP 1.

Morph	**0**	**10**	**20**	**25**	**30**	**40**	**50**	**60**	**70**	**80**
1	48,5	52,6	39,3	45,1	38,8	34,7	28,9	31,1	22,5	20,8
2	19,7	19,6	20,0	18,7	19,9	22,7	21,0	19,9	20,5	19,9
3	15,2	16,0	16,0	13,6	14,7	16,3	15,0	18,8	19,0	16,6
4	13,3	12,1	13,8	12,3	12,9	13,8	13,9	15,4	18,6	16,1
5	12,2	11,4	12,9	11,6	12,2	12,5	12,7	14,5	13,0	13,7

Table III shows the average confidence of FaceNet for different morphing levels. The best result is obtained with an alteration of 10%. The average confidence is computed only with the subjects that were correctly classified in the top 1. The average confidence distance between the first and second place of the top is 16.1%.

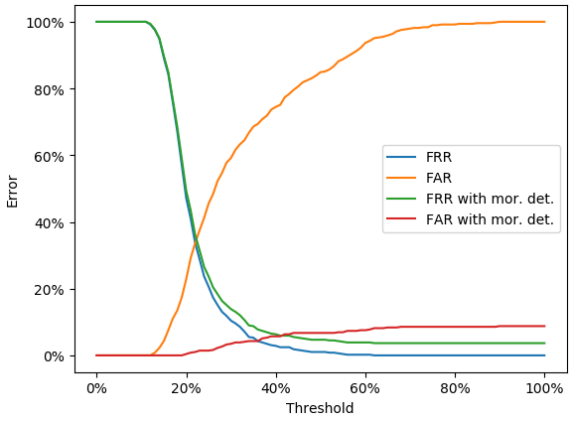


Fig. 4. FAR vs FRR of FaceNet.

Figure 4 shows how the FRR remains below 10% from the 31% threshold. Once the 55% threshold is reached, this error drops to almost 0%. Even with the morphing detector's application (which can cause extra false rejections), it performs well, adding less than 3.7% extra error. On the contrary, the

FAR rises significantly from the 15% threshold, reaching 90% error at 57%. The morphing detector strongly reduces this error, dropping it below 9% for all possible thresholds.

B. LBPH

Its best scores for each top are 32%, 36%, and 44%, respectively. At 50% morphing, it achieves 8%, 16%, and 20% identification of the original subject for each top. Looking at the top 3 and 5, it identifies 4% of the images with the original subject totally transformed (100% morphing). LBPH is the second most robust algorithm, having a distance with FaceNet of more than 50% misidentification in some cases. In general, its recognition rate decreases more slowly than FaceNet, but LBPH is always below it, getting a tie only above 78% morphing in the top 1.

TABLE IV
AVERAGE CONFIDENCE PERCENTAGES OF LBPH WHEN THE ORIGINAL SUBJECT IS INCLUDED IN THE TOP 1.

Morph	0	10	20	30	40	50	60	67	77	79
1	59,2	56,3	64,9	74,1	58,8	61,9	63,3	63,3	44,8	44,3
2	45,0	48,0	51,6	53,6	48,4	53,4	51,5	61,7	37,3	42,3
3	42,5	44,6	48,1	48,4	44,9	52,0	49,9	61,0	27,9	29,8
4	41,2	42,7	45,0	45,5	42,4	47,3	40,0	56,7	26,8	28,5
5	39,4	40,2	41,7	42,2	36,4	46,6	39,6	56,4	16,6	23,7

Table IV shows that its highest confidence peak is 74.1%, obtained with a morphing alteration of 30%. However, the number of individuals used to calculate it is lower than FaceNet since only three were correctly classified, whereas FaceNet classifies fifteen properly with the same amount of morphing. The average confidence distance between the first and second place of the top is 9.8%.

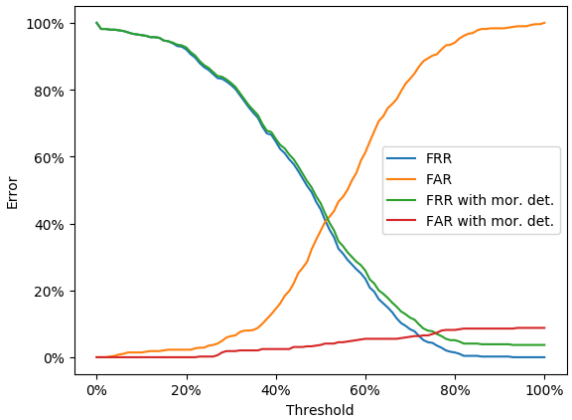


Fig. 5. FAR vs FRR of LBPH.

Figure 5 shows how the FRR remains below 10% from the 70% threshold. Once the 82% threshold is reached, this error drops to almost 0%. The application of the morphing detector adds less than 3.7% extra error. On the contrary, the FAR exceeds 10% from the 37% threshold, reaching 90% error at 75%. The morphing detector strongly reduces this error, dropping it below 9% for all possible thresholds.

C. Eigenfaces

Its best scores for each top are 16%, 20%, and 28%, respectively. At 50% morphing, it achieves 4%, 4%, and 8% identification of the original subject for each top. Looking at the top 5, it identifies 4% of the images with the original subject totally transformed (100% morphing). Eigenfaces takes the third position. In some percentages, it achieves a distance with LBPH of, at most, 16% identification. Although its performance is low, it maintains 8% and 4% identification for a long time. For example, between 38% and 100% morphing in the top 5.

TABLE V
AVERAGE CONFIDENCE PERCENTAGES OF EIGENFACES WHEN THE ORIGINAL SUBJECT IS INCLUDED IN THE TOP 1.

Morph	0	5	10	15	20	30	36	43	46	54
1	77,4	85,8	76,8	70,9	98,5	66,2	92,8	82,2	74,6	87,3
2	69,1	76,3	76,0	69,7	84,8	66,1	78,6	82,1	68,6	81,6
3	65,5	71,4	71,0	68,5	71,6	64,3	73,1	75,4	65,3	79,6
4	61,3	66,1	69,9	64,5	71,1	62,1	73,0	72,8	65,0	78,6
5	58,9	65,5	69,5	63,4	71,1	59,4	72,7	71,8	62,3	78,2

Table V shows that its highest confidence peak is 98.5%, obtained with a morphing alteration of 20%. In this case, only two subjects were correctly classified. The average confidence distance between the first and second place of the top is 6%.

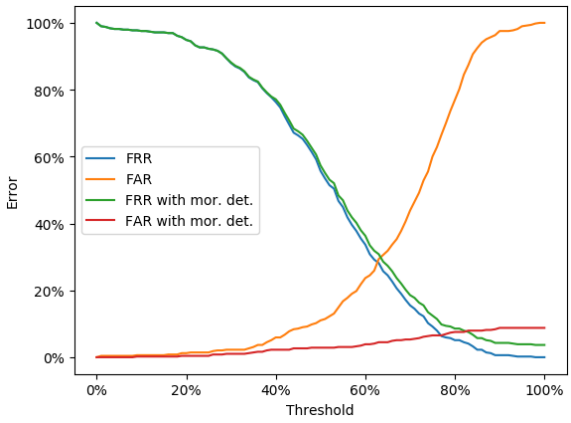


Fig. 6. FAR vs FRR of Eigenfaces.

Figure 6 shows how the FRR remains below 10% from the 75% threshold. Once the 93% threshold is reached, this error drops to almost 0%. The application of the morphing detector adds less than 3.7% extra error. On the contrary, the FAR exceeds 10% from the 49% threshold, reaching 90% error at 84%. The morphing detector strongly reduces this error, dropping it below 9% for all possible thresholds.

D. Fisherfaces

Its best scores for each top are 8%, 16%, and 20%, respectively. At 50% morphing, it achieves 0%, 4% ,and 4% identification of the original subject for each top. Fisherfaces fails to identify any original subject with 100% alteration. As with Eigenfaces, although its performance is low, in some

cases, it manages to maintain a 4% identification rate for a long range, for instance, in 14%–52% morphing in the top 5. However, at all the percentages, it has equal or lower recognition rates than Eigenfaces.

TABLE VI
AVERAGE CONFIDENCE PERCENTAGES OF FISHERFACES WHEN THE ORIGINAL SUBJECT IS INCLUDED IN THE TOP 1.

Morph	0	2	3	5	6	10	15	20	25	28
1	63,2	92,1	68,6	94,8	51,3	97,3	100	88,2	89,8	88,0
2	59,8	84,5	64,3	87,2	51,2	91,5	93,1	84,8	88,8	87,8
3	58,4	80,2	61,3	83,8	49,5	89,5	91,4	83,6	86,0	83,4
4	56,6	79,8	60,1	83,8	49,5	88,9	90,5	80,9	85,4	83,3
5	55,6	79,2	59,6	83,6	49,4	88,4	89,1	80,1	84,0	81,7

Table VI shows that its highest confidence peak is 100%, obtained with a morphing alteration of 15%, with just one person correctly classified. The average confidence distance between the first and second place of the top is 4%.

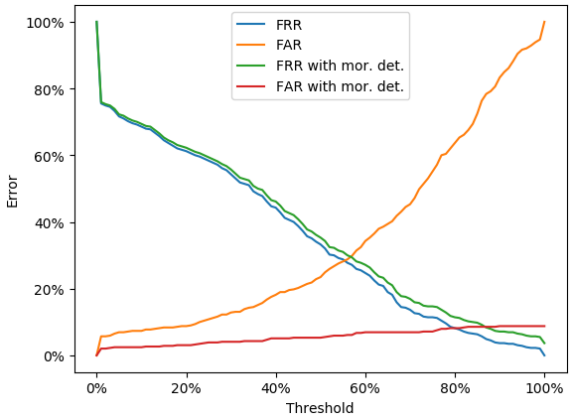


Fig. 7. FAR vs FRR of Fisherfaces.

Figure 7 shows how the FRR remains below 10% from the 78% threshold. This error drops to 0% only when the threshold is 100%. The application of the morphing detector adds less than 3.7% extra error. On the contrary, the FAR exceeds 10% from the 23% threshold, reaching 90% error at 94%. The morphing detector strongly reduces this error, dropping it below 9% for all possible thresholds.

E. SIFT

Its best scores for each top are 16%, 20%, and 20%, respectively. Once we reach 20% morphing, SIFT obtains 0% identification in all cases. Although the values achieved at 0% morphing are better than those obtained with Fisherfaces, SIFT's decrease rate is higher.

Table VII shows that its highest confidence peak is 69.4%, obtained with a morphing alteration of 8%, but only two people are correctly classified in that case. The average confidence distance between the first and second place of the top is 18.7%.

Figure 8 shows how the FRR remains below 10% from the 59% threshold. This error drops to 0% only when the threshold is 100%. The application of the morphing detector

TABLE VII
AVERAGE CONFIDENCE PERCENTAGES OF SIFT WHEN THE ORIGINAL SUBJECT IS INCLUDED IN THE TOP 1.

Morph	0	1	2	3	4	5	6	7	8	10
1	33,4	45,7	43,6	41,6	60,2	42,9	43,7	48,7	69,4	43,9
2	30,8	28,2	28,3	28,1	25,6	30,9	30,2	26,8	27,1	30,2
3	23,8	23,9	23,9	24,7	25,6	26,0	28,6	24,4	27,1	30,2
4	22,7	21,7	21,8	23,6	23,2	26,0	25,4	24,4	19,7	30,2
5	20,6	21,7	21,8	22,5	23,2	18,8	23,0	24,4	19,7	21,9

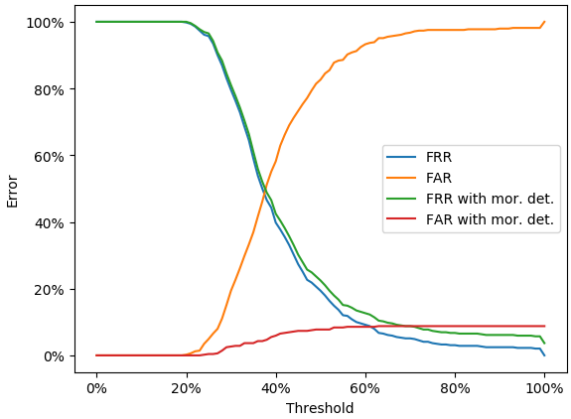


Fig. 8. FAR vs FRR of SIFT.

adds less than 3.7% extra error. On the contrary, the FAR exceeds 10% from the 28% threshold, reaching 90% error at 56%. The morphing detector strongly reduces this error, dropping it below 9% for all possible thresholds.

F. Morphing detector

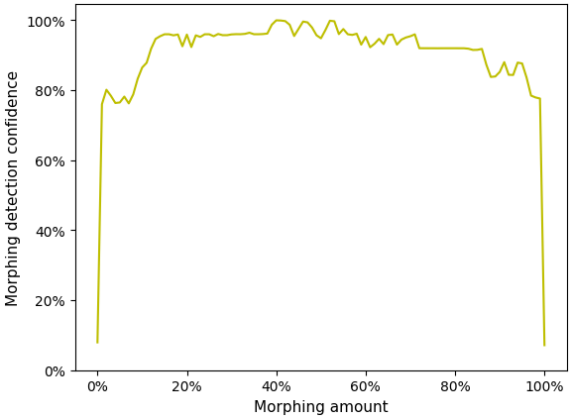


Fig. 9. Average morphing detection confidence.

Figure 9 shows the performance of the morphing detector in the morphed images used in the basic scenario. It displays the average detection rate at every quantity of morphing alteration from 0% to 100% (reflected in the x-axis), computed using all the morphed and unaltered images from the 25 similar-looking pairs.

We can observe that with the non-morphed images (0% and 100%), the detector provides less than 10% confidence. On

the contrary, between 1% and 99% of morphing alteration, it returns an average morphing confidence above 70%. Also, confidence is over 90% in 15%–85% morphing. At some morphing percentages around the maximum alteration (50%), it reaches a confidence level near 100%, proving its excellent performance.

VI. Discussion

In the basic scenario, FaceNet obtains the best performance identifying the in-between morphed images correctly. This means that in the case of a real attack on FaceNet, the attacker would need to significantly alter the image to fool the recognizer. Looking at Table VIII, we can see that analyzing the top 1, the attacker would need a 43% morphing alteration to have more than a 50% chance of the attack being successful. If we analyze the top 3, the required morphing alteration is higher than 66%. Finally, if we analyze the top 5, the alteration needed rises to 71%. FaceNet shows such good results that some attacks will fail even with the original image wholly modified (100% morphing) if we consider top 3 or top 5 lists.

TABLE VIII
Accuracy of FaceNet at different percentages of morphing. Italicized values represent the point at which the accuracy drops below 50%.

Morph	0%	43%	50%	66%	71%	100%
Top 1	84%	*48%*	32%	12%	12%	0%
Top 3	96%	80%	72%	*48%*	32%	8%
Top 5	100%	80%	76%	64%	*48%*	12%

As for the remaining recognizers, the identification rate is much worse, being extremely low in some cases. Our results for facial identification on the LFW database are notably worse than those obtained in verification. This might be expected since, for identification, we work 1 vs. N ($N = 366$ in our database), and regarding verification, we work 1 vs. 1. Thus, as mentioned in [44], the difficulty of identification is related to the number of subjects contained in the database. Some examples are Eigenfaces, in which we have obtained 16% of identification accuracy in contrast with 60.02% of verification accuracy [44], and FaceNet, with 84% and 99.6% of identification and verification accuracy, respectively [44].

The only possible alternative to FaceNet would be LBPH. When dealing with images with a considerable morphing amount (e.g., $> 75\%$), their accuracy is similar, however, LBPH offers greater distances (between the confidence of the first and second position) than with FaceNet. With both recognizers, we get the best confidence distances for 0% morphing, 28.8% for FaceNet, and 14.2% for LBPH.

We have also shown that the morphing detector has an excellent performance, outputting morphing detection confidences above 90% when the alteration is considerable (15%–85%). That would mean that most attacks that require some alteration in order to be successful would very likely be detected.

In the advanced scenario, FaceNet is the recognizer with the best FRR since it is the one that achieves an error below 10% with the lowest threshold (31%). It is followed by SIFT, which needs a threshold of 59% to achieve the same error. However, as we have seen, SIFT has low performance in the basic scenario, so it might not be recommended in a general system.

Eigenfaces is the recognizer with the best FAR since it is the one that achieves an error above 10% with the highest threshold (49%). Nevertheless, as in the case of SIFT, its performance from the basic perspective is poor, so we do not recommend its use in a general system either. Its FAR results are followed by LBPH (37% threshold), which would be a preferable option.

The inclusion of the morphing detector has a significant impact on all recognizers. It causes the FAR to always be below 9% and the FRR to grow at most 3.7%.

As the morphing detector fixes the FAR problem, FaceNet is the best algorithm in either correctly identifying subjects (basic scenario) or registering new subjects (advanced scenario). It is the best performing method in a general system.

VII. Conclusion and future work

If we want to prevent registration using fake face images, the recommended option is FaceNet or, as a second option, LBPH. Our experiments show that these techniques have significantly better results than others like Eigenfaces, Fisherfaces, or SIFT. The difference between FaceNet and any other technique is impressive. With 0% of morphing, only FaceNet presents an accuracy of over 80%. The second option, LBPH, has an accuracy below 35%, while the rest of the techniques cannot reach 20%. Even with a small amount of morphing, less than 20%, the error of more classic techniques jumps over 90%.

FaceNet is a robust technique against morphing attacks when used in combination with an S-MAD morphing detector. Both the False Rejection Rate and the False Acceptance Rate are lower than 6% when a threshold of 41% is used. This threshold can be recommended for most cases since FaceNet recognizes most attackers using images with less than 15% of morphing. Above 15%, the morphing detector can detect 95% of the potential impostors.

Therefore, we can conclude that a reasonable solution for preventing registration and login using fake face images can be built using face recognition and morphing detection state-of-the-art techniques. We have tested algorithms from different families of facial recognition techniques and found a clear difference between the one based on Deep Learning (FaceNet) and the rest. We will test newer and promising facial recognition algorithms that fall into this family of algorithms in our future work. Since the detection results are pretty robust against morphing processing, it would be interesting to challenge the solution proposed in this paper with better-designed algorithms for fooling its detection systems.

References

[1] S. F. Kak, F. M. Mustafa, and P. Valente, "A review of person recognition based on face model," *Eurasian Journal of Science & Engineeting,* vol. 4, issue 1, pp. 157–168, 2018. [Online]. Available doi: 10.23918/eajse.v4i1sip157.

 ISBN: 978-1-61208-862-4

[2] A. Shwetank, P. Neeraj, and B. Karamjit, "Future of face recognition: A review," *Second International Symposium on Computer Vision and the Internet,* vol. 58, pp. 578–585, 2015.

[3] M. Ferrara, A. Franco, and D. Maltoni "The magic passport," in *IEEE International Joint Conference on Biometrics,* Clearwater, FL, 2014, pp. 1–7. [Online]. Available doi: 10.1109/BTAS.2014.6996240.

[4] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, "Face recognition systems: A survey," *Sensors*, vol. 20, no. 2:342, 2020. [Online]. Available doi: 10.3390/s20020342.

[5] M. Wang and W. Deng, "Deep face recognition: A survey," 2018, *arXiv:1804.06655*. [Online]. Available: https://arxiv.org/abs/1804.06655.

[6] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proceedings of Computer Vision and Pattern Recognition IEEE Computer Society*, June 1991, pp. 586–591.

[7] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997. [Online]. Available doi: 10.1109/34.598228.

[8] G. Zhang, X. Huang, S. Z. Li, Y. Wang, and X. Wu, "Boosting local binary pattern (LBP)-based face recognition," in *Proc. of the 5th Chinese conference on Advances in Biometric Person Authentication,* 2004, pp. 179–186. [Online]. Available doi: 10.1007/978-3-540-30548-4_21.

[9] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli, "On the Use of SIFT Features for Face Authentication," in *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, New York, NY, USA, 2006, pp. 35–35. [Online]. Available doi: 10.1109/CVPRW.2006.149.

[10] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2015, pp. 815–823. [Online]. Available doi: 10.1109/CVPR.2015.7298682.

[11] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments," University of Massachusetts, Technical Report 07-49, Oct. 2007.

[12] A. K. Agrawal and Y. N. Singh, "Evaluation of face recognition methods in unconstrained environments," *Procedia Computer Science,* vol. 48, pp. 644–751, 2015.

[13] X. Zhang and Y. Gao "Face recognition across pose: A review," *Pattern Recognition,* vol. 42, no. 11, pp. 2876–2896, 2009. [Online]. Available doi: 10.1016/j.patcog.2009.04.017.

[14] G. H Givens *et al.* "Introduction to face recognition and evaluation of algorithm performance," *Computational Statistics and Data Analysis,* vol. 67, pp. 236–247, 2013. [Online]. Available doi: 10.1016/j.csda.2013.05.025.

[15] G. Wolberg, *Digital Image Warping,* IEEE Computer Society Press, Los Alamitos, California, 1990.

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

[17] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library.* Sebastopol, CA, USA: O'Reilly Media, 2008.

[18] S. Mallick, *Face Morph Using OpenCV - C++/Python,* Learn OpenCV, March 11, 2016. Retrieved: April, 2021. [Online]. Available: https://learnopencv.com/face-morph-using-opencv-cpp-python/.

[19] A. Hadid, N. Evans, S. Marcel, and J. Fierrez, "Biometrics systems under spoofing attack: An evaluation methodology and lessons learned," *IEEE Signal Processing Magazine,* vol. 32, no. 5, pp. 20–30, Sept. 2015. [Online]. Available doi: 10.1109/MSP.2015.2437652.

[20] A. Mohammadi, S. Bhattacharjee, and S. Marcel, "Deeply vulnerable: A study of the robustness of face recognition to presentation attacks," *Institution of Engineering and Technology Biometrics,* vol. 7, issue 1, pp. 15–26, 2018. [Online]. Available doi: 10.1049/iet-bmt.2017.0079.

[21] M. Ferrara, A. Franco, and D. Maltoni, "On the effects of image alterations on face recognition accuracy," in *Face Recognition Across the Image Spectrum.* Springer Nature, 2016, pp. 195–222.

[22] R. Tolosana, R. Vera-Rodríguez, J. Fiérrez, A. Morales, and J. Ortega-García, "DeepFakes and beyond: A survey of face manipulation and fake detection," *arXiv preprint arXiv:2001.00179*, 2020.

[23] D. J. Robertson, R. S. S. Kramer, and A. M. Burton, "Fraudulent ID using face morphs: Experiments on human and automatic recognition," *PLoS ONE,* vol. 12, no. 3:e0173319. [Online]. Available doi: doi:10.1371/journal.pone.0173319.

[24] L. Wandzik, R. V. García, G. Kaeding, and X. Chen, "CNNs under attack: On the vulnerability of deep neural networks based face recognitionto image morphing," in *Proc. 16th Int. Workshop on Digital Forensics and Watermarking,* 2017, pp. 121–135.

[25] U. Scherhag *et al.* "On the vulnerability of face recognition systems towards morphed face attacks," in *2017 5th International Workshop on Biometrics and Forensics (IWBF)*, Coventry, 2017, pp. 1–6. [Online]. Available doi: 10.1109/IWBF.2017.7935088.

[26] M. Sharif, F. Naz, M. Yasmin, M. A. Shahid, and A. Rehman, "Face recognition: A survey," *Journal of Engineering Science and Technology Review,* vol. 10, no. 2, pp. 166–177, 2017.

[27] S. Patel, *Face Morphing*, Github, 2018. Retrieved: April, 2021. [Online]. Available: https://github.com/cirbuk/face-morphing.

[28] B. Delaunay "Sur la sphère vide. A la mémoire de Georges Voronoï [On the empty sphere. In memory of Georges Voronoï]," *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et naturelles* vol. 6, pp. 793–800, 1934.

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

[30] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 28, no. 12, pp. 2037–2041, Dec. 2006. Available doi: 10.1109/TPAMI.2006.244.

[31] R. Raja, *Face Recognition with OpenCV and Python*, Github, 2017. Retrieved: April, 2021. [Online]. Available: https://github.com/informramiz/opencv-face-recognition-python.

[32] D. Sandberg, *Face Recognition using Tensorflow*, Github, 2016. Retrieved: April, 2021. [Online]. Available: https://github.com/davidsandberg/facenet.

[33] J. Galbally, S. Marcel, and J. Fiérrez, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access,* vol. 2, pp. 1530–1552, 2014. [Online]. Available doi: 10.1109/ACCESS.2014.2381273.

[34] R. B. Hadiprakoso, H. Setiawan, and Girinoto, "Face Anti-Spoofing Using CNN Classifier & Face liveness Detection," in *2020 3rd International Conference on Information and Communications Technology (ICOIACT),* Yogyakarta, Indonesia, 2020, pp. 143–147. [Online]. Available doi: 10.1109/ICOIACT50329.2020.9331977.

[35] B. Wu, M. Pan, and Y. Zhang, "A review of face anti-spoofing and its applications in China," in *International Conference on Harmony Search Algorithm,* Springer, 2019, pp. 35–43. [Online]. Available doi: 10.1007/978-3-030-31967-0_4.

[36] R. Raghavendra, K. B. Raja, and C. Busch. "Detecting Morphed Face Images," in *8th IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS),* pp. 1–8, 2016.

[37] L. Zhang, F. Peng, and M. Long, "Face morphing detection using Fourier spectrum of sensor pattern noise," in *2018 IEEE International Conference on Multimedia and Expo (ICME),* San Diego, CA, 2018, pp. 1–6. [Online]. Available doi: 10.1109/ICME.2018.8486607.

[38] R. Raghavendra, K. Raja, S. Venkatesh, and C. Busch, "Face morphing versus face averaging: Vulnerability and detection," in *2017 IEEE International Joint Conference on Biometrics (IJCB),* Denver, CO, 2017, pp. 555–563. [Online]. Available doi: 10.1109/BTAS.2017.8272742.

[39] L. Spreeuwers, M. Schils, and R. Veldhuis, "Towards robust evaluation of face morphing detection," in *2018 26th European Signal Processing Conference (EUSIPCO),* Rome, 2018, pp. 1027–1031. [Online]. Available doi: 10.23919/EUSIPCO.2018.8553018.

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

[41] K. Raja *et al.*, "Morphing attack detection - database, evaluation platform and benchmarking," *IEEE Transactions on Information Forensics and Security,* 2020. [Online] Available doi: 10.1109/TIFS.2020.3035252.

[42] W. Deng, J. Hu, N. Zhang, B. Chen, and J. Guo, "Fine-grained face verification: FGLFW database, baselines, and human-DCMN partnership," *Pattern Recognition,* vol. 66, pp.63–73, 2017.

[43] V. K. Vishnoi, *What Is The Configuration Of The System That Is Being Used In Matlab Online Web Based Version,* MATLAB Answers, January 24, 2020. Retrieved: April, 2021. [Online]. Available: https://www.mathworks.com/matlabcentral/answers/500853-what-is-the-configuration-of-the-system-that-is-being-used-in-matlab-online-web-based-version.

[44] E. Learned-Miller *et al.*, "Labeled Faces in the Wild: A Survey," in *Advances in Face Detection and Facial Image Analysis*, 2016, pp. 189–248. [Online]. Available doi: 10.1007/978-3-319-25958-1_8.