IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 1, 2020, accepted May 1, 2020, date of publication May 12, 2020, date of current version May 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2994112

Border Control Morphing Attack Detection With a Convolutional Neural Network De-Morphing Approach

DAVID ORTEGA-DELCAMPO, CRISTINA CONDE, DANIEL PALACIOS-ALONSO, AND ENRIQUE CABELLO

Escuela Técnica Superior de Ingeniería Informática, Universidad Rey Juan Carlos, Campus de Móstoles, 28933 Madrid, Spain

Corresponding author: Daniel Palacios-Alonso (daniel.palacios@urjc.es)

This work was supported in part by the Spanish Ministry of Economy Research Project Bio-Inspired Face Recognition From Multiple Viewpoints Evaluation in a Presentation Attack Detection Environment (BIOINPAD) under Grant TIN2016-80644-P, in part by the European Commission 7FP Project ABC4EU under Grant 312797, and in part by the Spanish General Directorate of Police.

ABSTRACT Currently, the use of biometric identification, automated or semiautomated, is a reality. For this reason, the number of attacks has increased in such systems. One of the most common biometric attacks is the presentation attack (PA) because it is relatively easy to perform. Automated border control (ABC) is a clear target for phishers. Concerning biometric attacks, morphing is one of the most threatening attacks because authentication systems are usually unable to correctly detect them. In this attack, a fake face is generated with the morphing and blending of two different subjects (genuine and phisher), and the image result is stored in the passport. These attacks can generate risky situations in cases of border crossings where an ABC system should perform identification tasks. This research work proposes a de-morphing architecture that is founded on a convolutional neural network (CNN) architecture. This technique is based on the use of two images: the potentially morphed image stored in the passport, and the snapshot of the person located in the ABC system. The goal of the de-morphing process is to unravel the *chip* image. If the *chip* image is a morphed one, the revealing process between the *in vivo* image and the morphed *chip* image will return a different facial identity to the person located in the ABC system, and the impostor will be uncovered *in situ*. If the *chip* image is a non-morphing image, the resulting image will be similar to a genuine passenger. Therefore, the information obtained is considered at the border crossing. The equal error rate (EER) achieved is very low compared to the literature values published to date. The accomplished outcomes endorse a robust method that provides high accuracy rates without taking into account the quality of images used. This key point is crucial to plausible deployment plans in areas such as ABC.

INDEX TERMS ABC, biometric systems, de-morphing, neural networks, MAD.

I. INTRODUCTION

The face recognition process is a well-known biometric identification challenge due to the high accuracy rates achieved and low intrusion to the subjects under identification. To approach the facial authentication process, there are several methods. Some of these approaches have high security requirements. An example is the automatic border control (ABC) system, in which the biometric trait is used to control and assure the border crossing process. Three biometric elements (iris, fingerprint and face) could be considered in ABC systems, but as a matter of fact, only the face is widely considered at airports. The authentication process of the ABC has to determine whether or not there is a coincidental match between the facial image stored in the Electronic Machine Readable Travel Document (eMRTD) and a snapshot taken *in situ* (see Fig. 1).

The ABC systems are exposed to multiple attacks or threats, for example, identity theft or fraud, which also is called spoofing. For this reason, many current research works focus their attention on anti-spoofing techniques [1]. Morphing attack is one of the most dangerous attacks because of its high difficulty to be detected. It is based on the application of morphing techniques to the facial image recorded in the

The associate editor coordinating the review of this manuscript and approving it for publication was Jinjia Zhou.

FIGURE 1. Examples of *FRAV-ABC* data set images. Passport *chip* images (top) and snap-shot images taken in the border scenario (bottom).

passport or traveler document. The morphing attack consists of manipulating and storing inside of the eMRTD a morphed image between the real owner's ID card (accomplice) and the surrogate or impostor (criminal). Then, the system should distinguish whether the traveler is who it claims to be or not. This approach relies on comparing the taken picture at the site (ABC) and the eMRTD's image that contains the potentially morphed image. If a MAD module is not present in the ABC, the usual verification response will be acceptance, considering the high similarity between the criminal face and the morphed criminal+accomplice image.

The morphing process emerged from Arts' world such as films, video clips or advertisements as an art resource to achieve awesome special effects [2], [3]. In the beginning, the process was handcrafted but this situation changed quickly due to the emergence of the first new algorithms that automated the morphing tasks [4]. It should be noted that it was an arduous task, even to experts, to distinguish two faces when they were merged [5]–[7]. Thus, the technique evolved from an art resource to a spoofing toolkit [8].

Morphing of facial images can be considered as one of the most important threats of ABC systems [9]–[11] since applications based on face recognition are likely to be deceived [12], [13]. For instance, the outcomes of the National Institute of Standards and Technology (NIST) Face Recognition Vendor Test MORPH (https://pages.nist.gov/frvt/html/frvt_morph.html) discussed that the submitted MAD algorithms lack robustness and performance when considering unseen and challenging corpora, as explained in [14]. However, other biometric features have been considered in morphing attacks such as fingerprints [15] or the iris [16]. In any case, the focus of this study is on facial morphing since this attack is the most devastating and difficult to detect in ABC systems.

In recent years, the wide use of ABC systems in airports has increased the attention and the study of the possible multiple menaces (e.g., presentation attack) as explained by the European Border and Coast Guard Agency (FRONTEX) [17], [18]. These attacks incentivize the proliferation of algorithms about presentation attack detection (PAD) [19]–[21] and especially morphing attack detection (MAD) [13], [22] because it is a difficult paradigm to be detected.

In this paper, a novel method to detect morphing attacks is explained using a reverse de-morphing approach based on convolutional neural networks. There are several differences compared with previous works [23], [24], which are explained as follows.

Ferrara *et al.*'s work consists of detecting the morphing attack, the elaboration of two corpora (PMDB and MorphDB) and the assessment of the quality of two corpora using a commercial off-the-shelf (COTS) algorithm. The key point in Ferrara's algorithm is that their algorithm depends on the prior knowledge of the generation of the morphed face, such as the morphing process and the morphing parameters. Moreover, the reconstruction faces rely on the inverse engineering process of morphing tasks using a mathematical method. Finally, this work is based on Delaunay-Voronoi triangulation but there are new approaches in which the de-morphing process is performed with neural networks. For instance, Damer *et al.* [25] and Peng *et al.* [24] propose the use of the generative adversarial network (GAN). Regarding Peng *et al.*'s work, it is based on disentangling the accomplice identity from a potentially morphed image. However, the authors developed approach is divided into two aims. The first aim consists of unraveling the criminal identity. The second aim relies on comparing the image obtained in the previous stage with the *in vivo* image obtained in the ABC gate. Therefore, the authors can conclude whether morphing attack occurred or not. Additionally, the de-morphing process of Peng *et al.* is based on a GAN, but the presented approach relies on an autoencoder architecture. Regarding MAD, another key point is that none of the approaches consider print and scan images in their studies. Finally, Peng and Ferrara's works take the pictures for their corpus in a controlled environment. Nevertheless, in this research work, 1170 *in vivo* images taken *in vivo* in the eGates or automated boarding control system are used.

The paper is organized as follows: the state of the art is presented in the following section. The dataset is then described. Since a morphing method is required, Section IV is devoted to presenting a morphing method and its adaptation to passport control in the ABC. Subsequently, (Section V), the de-morphing approach is detailed. Section VI points out the results and provides discussion. The conclusions are presented in the last section.

II. PREVIOUS WORKS

In recent years, as morphing techniques have undergone experimental investigations, an impressive improvement in several aspects such as visual quality and automation generation has been achieved. From a substantive viewpoint, morphing's corpora are designed with open source and well-known software such as the GNU Image Manipulation Program (GIMP) which has a plugin called the GIMP Animation Package (GAP) [26]. This plugin is able to merge images [10], [13], [23], [27], but most of the software uses the Delaunay-Voronoi triangulation algorithm (DVT) [28]–[33] and a swapping technique to improve the outcome achieved [34]–[39]. Moreover, some current research works use morphing pictures with generative adversarial networks (GANs) instead of using the triangulation process as mentioned previously [25].

Two MAD implementations can be found in the literature, depending on morphing attack scenarios:

a) MAD with a single image (no-reference). Only one morphed image is available.
b) MAD with two images (differential MAD). The morphed picture and another one are used. This is the typical scenario in ABC systems [10], [23], [24], [28].

The first approach, no-reference, seeks to determine the noise or the deterioration in terms of quality of the image. The picture achieved after the morphing process, however, presents low quality. For this reason, this technique is based on micro-texture analysis or spatial descriptor occurrences or spectral analysis with the Fourier transform.

On the one hand, there are research works that rely on micro-textures which use some features such as the local binary pattern (LBP - [40]) in [25], [31], [39], [41]–[43], or weighted local magnitude pattern (WLMP) which is proposed and explained in [44]. On the other hand, there are research works based on analysis of descriptors which use the scale invariant feature transform (SIFT - [45]) in [46], binarized statistical image features (BSIF - [47]) in [27], and speeded-up robust features (SURF - [48]) in [34]. Finally, others use spatial descriptors such as histogram of oriented gradients (HOG - [49]) in [30].

In addition to the structural descriptor and texture analysis, other studies assess the degradation of the image through spectral image analysis. Some researchers try to detect a possible manipulation using the last mentioned technique [35]. Others try to evaluate the noise pattern employing the photo response non-uniformity (PRNU) approach [50]) in the full image [32] or each region [33].

With the advent of deep learning in the last decades, some approaches use convolutional neural networks (CNNs) to detect the morphing process [25], [38], [39], [51]. Some of the most well-known corpora are VGG19 [52], AlexNet [53] or GAN [54], [55]. The main drawback of these kinds of corpora is the number of samples required to train models. For this reason, some research works use pretrained networks, that is, networks with precalculated weights such as FaceNet [56] or VGG-Face [57].

Differential MAD needs two images for morphing detection and often proposes solutions for similar ABC systems where two images of identities are available. For instance, Scherhag [58] seeks SIFT descriptors in the ID passport image and the *in situ* image. Once the descriptors of both images are detected, they are compared. It is important to remark that in this case, the ID passport image is not a trustworthy image but a fake one. This fake image is based on the surrogate image and the ABC person's face. The approach is similar to the previous research study in [58], but this time, the amount and position of the face landmark detected are compared [59].

However, other differential MAD approaches take greater advantage of two available images and propose that when one of the identities is removed from the morphed image, the other one remains [10], [23]. This removal process is named de-morphing.

Both differential and no-reference MAD approaches have a challenge with real-world images. The real morphed images have often been printed and scanned, and then, this final image is embedded in the passport. Given this action, MAD algorithms are no longer able to detect manipulated images. The aforementioned studies [10], [23] also analyze this problem.

III. DATASET

This research work was carried out with *FRAV-ABC* database (see Fig. 1), which was designed and developed by the research group FRAV following all conditions present in a general ABC system. Initially, a study of the facial databases available for the research community was performed, but none of them exactly fulfilled border crossing and ABC conditions. Considering the strict procedure in which a passport image is acquired and the normative restrictions (ICAO Doc 9303 [60]), a new database was acquired. To achieve it, a real airport ABC infrastructure was used that followed all of the aforementioned conditions to emulate as far as possible a real morphing attack.

The corpus was composed of 1170 individuals, 640 females and 530 males, with an age range of study participants between 18 and 74 years old. Indeed, 70% of subjects ranged between 25 and 50 years old. Each subject provided two images. The first was a *chip* image of a real passport, and the second was an *in vivo* image. *Chip* images have a resolution of 250×300, which are color images of real passports that comply with the standard regulation of the International Civil Aviation Organization (ICAO) Document 9303 for the eMRDT [61], and *in vivo* images have a resolution of 300×300 pixels, which are color images captured *in situ* at the airport by an ABC device.

The corpus was divided into two data sets: *FRAV-ABC-Train* with 1000 subjects (70%-30% as recommended in [62], [63], where 700 are used to train and 300 are for validation) and *FRAV-ABC-Test* with 170 subjects (roughly a 15% of the total). The authors designed and developed a large corpus. The way to build it is explained as follows.

On the one hand, a thousand subjects were mixed and morphed to each other, except themselves. This action provides (1000 × 1000) - 1000 combinations. Thus, the final training corpus returned 999.000 images. Note that the age, gender, and ethnicity of the subjects were not considered because the authors wanted to accomplish producing a robust data set. On the other hand, the test corpus was designed and developed in a similar way. Specifically, 170 images were combined with each other, except themselves. Therefore, the arithmetic equation returns 28730 images.

The verification process was performed with a TensorFlow implementation of the face recognizer described in [56] but reimplemented and published by [64]. Moreover, this implementation is based on ideas from [57]. This available subsystem is a facial recognition system with high accuracy rates (99.63% in LFW (Labeled Faces in the Wild [65]) and 95.12% in YTF (YouTube Faces [66])).

To follow the Spanish passport image generation procedure, 170 morphed images of the *FRAV-ABC-Test* data set were printed at 300 dots per inch (dpi) quality with a LaserJet color printer. Next, these images were scanned to build a new data set (denoted by *FRAV-ABC-Test-P&S-300*). Furthermore, a new degradation step was carried out to assess the efficiency of algorithms used in this research work. *FRAV-ABC-Test-P&S-150* was devised from the new process of printing and scanning of the same 170 images, but the images were printed at 150 dpi (see Fig. 2). In this way, a whole set of ''fake morphed'' passports with a highly realistic appearance was created. Finally, it should be highlighted the a publicly available database, *CASIAWebFace* [67], which has 500K facial images, was used. This database has been used for autoencoder face training.

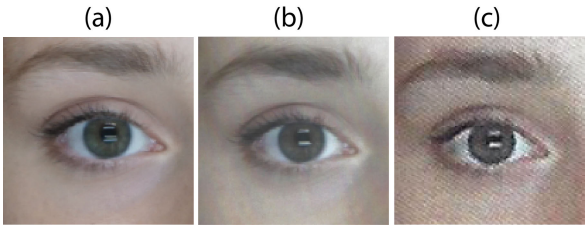

FIGURE 2. (a) *FRAV-ABC-Test* images quality. (b) *FRAV-ABC-Test-P&S-300* images quality. (c) *FRAV-ABC-Test-P&S-150* images quality.

IV. MORPHING METHODS AND SETUP

This section describes the morphing process and is split into two different parts. The first part will detail the morphing process selected and adapted to obtain a realistic image. This process of visually detection for a border guard can be an arduous task. The general facial morphing procedure has been tailored to suit the problem taken into account. Thus, the state-of-the-art algorithm and the enhancements added will be described. The second part will show the need for a morphing detection module in a face verification system.

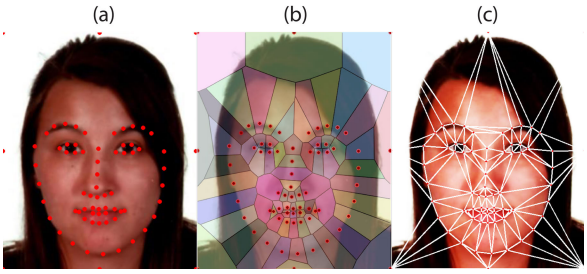

FIGURE 3. (a) Landmark face detection, (b) Delaunay computation (c) Voronoi triangulation.

A. MORPHING PROCESS

Currently, it is simple to find morphing commercial software [26], [68]–[74]. All of them provide a high-quality performance with morphed images, but it should be noticed that researchers have to perform manual manipulation and generate a large enough number of images to achieve an adequate corpus for their research works. The well-known algorithm [22], [23] used in the literature is adapted to the studied problem and explained below.

a) Given two pictures (see Figure 1(a)), 76 reference points are located in each one. There are 68 face landmarks, calculated with the Kazemi and Sullivan algorithm [59] (Dlib [75]) and eight more in the middle of image boundaries (see Figure 3 (a)).
b) The alignment process is mandatory. To carry out this task, the position and size of both images must match up at eye level.
c) Both images are triangulated by the Delaunay-Voronoi algorithm (DVT [76]) (see Figure 3 (b) and (c)) and each triangle in one image has a counterpart triangle in the other image (Figure 4).
d) Each triangle is blended in only one triangle whose vertexes are the midpoints (Figure. 4 (d)). The result of merging all triangles is the average image (Figure 5(a)), and this process is called the warping process [43], [77], [78].

The average image has ghost artifacts in peripheral regions and has low quality as an attack because it is too detectable. For that reason, it is necessary to carry out some enhancements that will be explained in the following.

Two main enhancements have been considered to obtain a realistic appearance without losing the morphing effects in the final image.

a) For cropping, in the target image, the convex hull of the face peripheral landmarks are placed in some of the source images 5 (b).
b) Using the Poisson image editing [79], the merging process is carried out. This method avoids hard seams, different capture illumination conditions or distinct skin colors (see Figure 5 (c)).

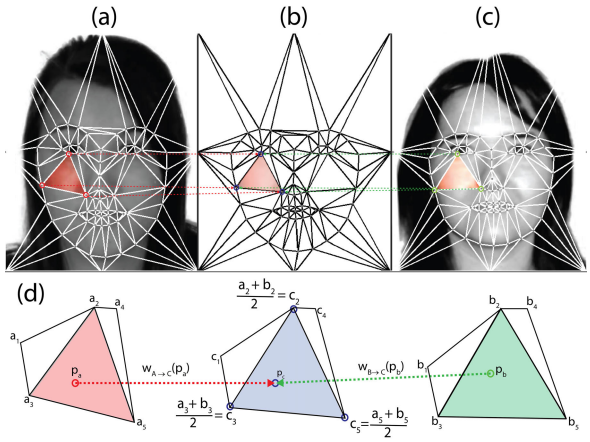

FIGURE 4. Warping the morphed target image and blending each source triangle that contained pixels.

FIGURE 5. (a)Average image, (b) Clipping and replacing the face region and (c) the fuzzy mask to enhance the result and cropping source image.

B. FACIAL VERIFICATION UNDER MORPHING ATTACKS

Facial verification systems are not prepared to deal with morphing attacks [12] and the common MADs are not effective with noisy or low-quality images [31], [41], [80]. The main problem with the verification process is the acceptance threshold because it is complicated to distinguish whether a morphing attack is produced or not due to the probability density function (PDF) that is located between the positive and negative acceptance, as shown in Fig. 6. Then, it is difficult to establish a rejection threshold for the transformation.

Fig. 6 and 7 depict the similarity scores obtained with FaceNet and one of COTS used by Ferrara et. al in [8], respectively, from different presentations. Each illustration shows three kinds of curves or areas. The first is a genuine traveler (positive presentation denoted by blue), the second is an impostor (negative presentation signified by orange), and finally, the third is a morphing attack (morphing presentation designated by a red striped line). The left plot depicts the test with *FRAV-ABC-Test*, the middle plot illustrates the test with *FRAV-ABC-Test-P&S-300* and the right plot represents the test using the *FRAV-ABC-Test-P&S-150* data set.

As observed in both figures, the scores of genuine individuals and impostors are well separated, and it is possible to define an adequate threshold to achieve high accuracy rates with FaceNet as well as with the COTS when digital images are used. However, the problem is more complex with the print and scan images (see pictures (b) and (c) in Figure 6 and Figure 7). Therefore, it is mandatory the use of MAD systems to prevent plausible attacks in both cases, open source systems (FaceNet) and COTS.

C. MAD SYSTEM UNDER PRINT AND SCAN IMAGES

The photo response non-uniformity (PRNU) system [50] is depicted in Fig. 8. The PRNU is selected as an example of existing MAD methods to be compared with the current approach considering several data sets. There are three data sets such as (a) FRAV-ABC-Test, (b) FRAV-ABC-Test-P&S-300, and (c) FRAV-ABC-Test-P&S-150). Each row is divided into two kinds of illustrations. On the left side, bona fide and morphing images are shown. On the right side, the histogram of one hundred bona fide images against one hundred morphing images are depicted. Regarding the first histogram, there exists a small difference between both pictures. However, it is difficult to unravel or distinguish the images' histograms when print and scan images are examined.

V. DE-MORPHING APPROACH CONSIDERED

With the morphing scheme described in the previous section, the de-morphing approach can be presented. The de-morphing process does not depend on the morphing scheme considered. The advantage of the previously described morphing method is that it avoids ghost effects or abrupt skin texture changes, making the de-morphing process a truly challenging situation. Morphing procedures that are not as complex can be visually detected devoting some attention to ghost artifacts in the hair or face limits and face skin color changes.

The de-morphing process is shown in Fig. 9. The input of the de-morphing process is the *in vivo* and the passport images and the output is one image. The goal of the de-morphing process is to unravel the *chip* image. If the *chip* image is a morphed image, the *in vivo* image is unlinked from the *chip*, and the output will be a new image. This new image will be quite different from the *in vivo* image since only the information from the other image used in the morphing process (impostor) will remain. If the *chip* image is a non-morphing image, the output will be similar to the *in vivo* image. Therefore, the last stage in the process will be an identity verification process between the output image and

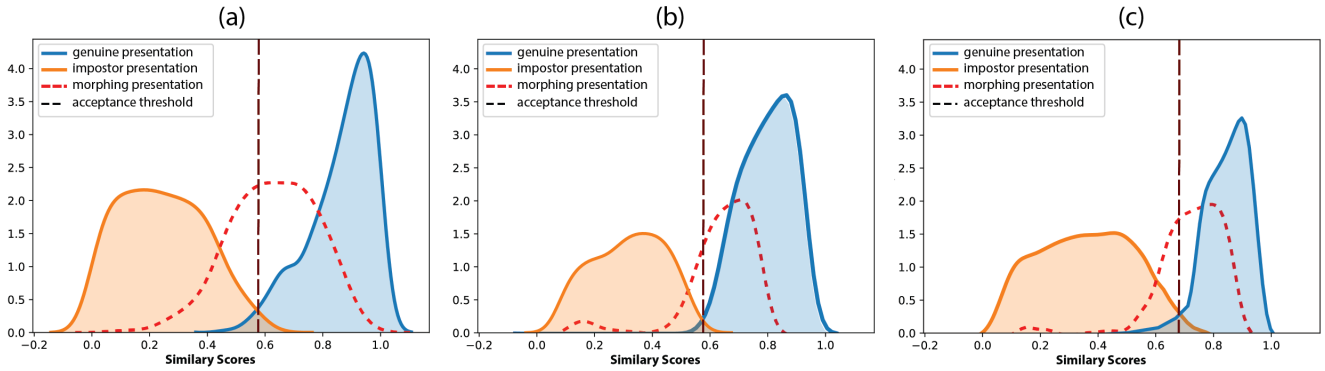

FIGURE 6. Density graph of similarity scores of FaceNet [56] calculated with *FRAV-ABC-Test* images (a), *FRAV-ABC-Test-P&S-300* (b) and *FRAV-ABC-Test-P&S-150* (c).

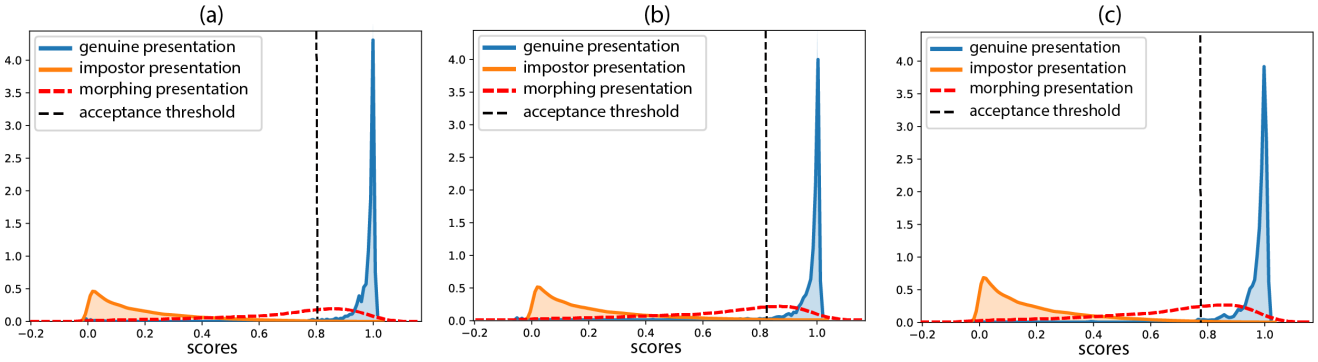

FIGURE 7. COTS Scores density calculated with *FRAV-ABC-Test* images (a), *FRAV-ABC-Test-P&S-300* (b) and *FRAV-ABC-Test-P&S-150* (c).

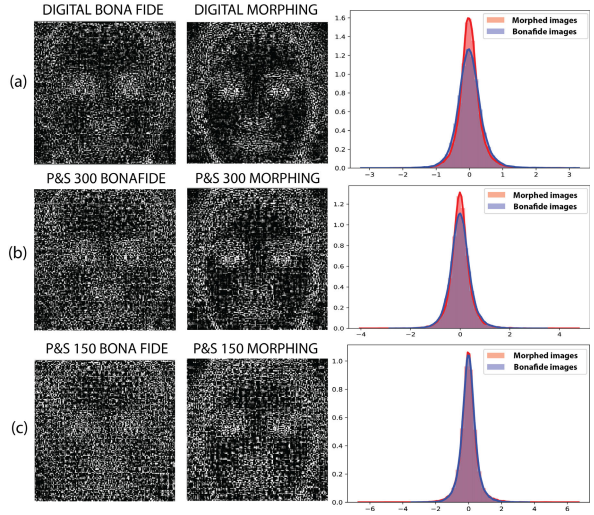

FIGURE 8. PRNU [50] map of the morphing image and bona fide image and histogram comparison of PRNU averaged values over 100 morphed images and 100 bona fide images of the data sets: (a) *FRAV-ABC-Test*, (b) *FRAV-ABC-Test-P&S-300*, (c)and *FRAV-ABC-Test-P&S-150*.

the *in vivo* image. If both images are similar and they can be assumed to have the same subject, the *chip* image is not morphed. However, if both images are not similar, that is, if researchers can assume that those images are from different subjects, it can be noticed that the original *chip* image is a morphed image. That morphed image is a mixture between the *in vivo* subject and the subject whose information has been kept in output image.

The de-morphing process has been split into two parts. The first one is a facial autoencoder for each of the input images that is followed by a decoder network. Therefore, three neural networks will be used: two encoders of the same size and architecture (one for the *in vivo* and another one for the *chip* images) and one decoder neural network.

A. DE-MORPHING PROCESS

This process tries to discover the initial pictures from two images. These images could be bona fide or genuine; in contrast, they could be fraudulent. Two pictographs are used in Figure 9 to explain the meaning of these images. The green passport means a genuine *chip* picture and the red passport means a fraudulent *chip* image in which the morphing process is carried out. Finally, the blue camera symbol is used to explain the *in vivo* picture in ABC.

In Equation 1, morphing and de-morphing processes are illustrated. Operations on sets help to understand the role that different images (*in vivo* or *chip* images) play in the process.

$$A \cap B = C;\ C - A \sim B;\ C - A \nsim A \tag{1}$$

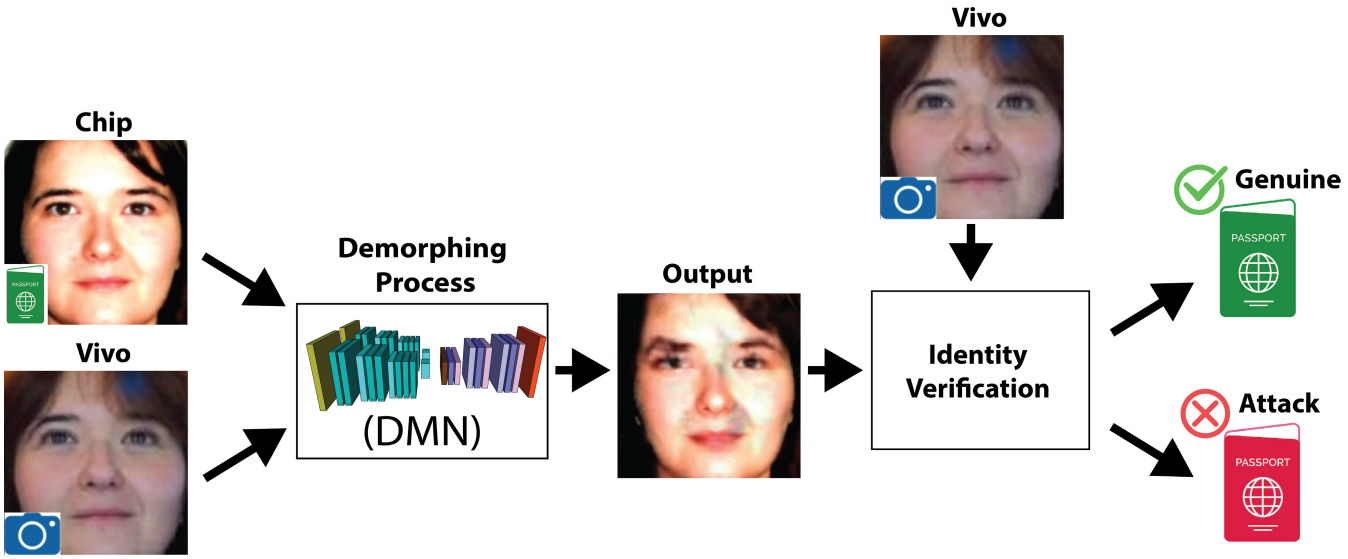

FIGURE 9. De-morphing process and identity verify process.

The intersection of two images (A and B) is the morphing process representation, where C is the morphed image. The difference between C-A represents the de-morphing process. C-A should be an image similar to B (in case C is a morphed image); additionally, C-A will not be similar to the A image. If a morphed image is presented in an ABC gate, only C and A image are obtained (C from the *chip* and A from the *in vivo* image). If C-A is not similar to A, it can be assumed that a B image has been used to compose a morphing attack.

Therefore, if the compared output snapshot and *in vivo* image are similar and the *in vivo* image is also similar to the *chip* image prior to the de-morphing process, then it might be noticed that the output is not a morphing attack, as illustrated in Fig. 10(a). However, if the output picture is not similar to the *in vivo* image (regardless of the similarity with *chip* image before the de-morphing process), then a morphing attack was performed, as depicted in Fig. 10(b).

Once the de-morphing process is performed, the similarity output is compared with the *in vivo* and *chip* images. In Fig. 11, output distributions of the classifier are shown. Two examples are depicted to illustrate the difference between a morphed and non-morphed process. When a genuine *chip* image is compared to an *in vivo* image, both plots are overlapped largely as shown in Fig. 11(a); in other words, the distance among pictures is minimum. In contrast, if the comparison is based on a morphed or manipulated *chip* image and a *in vivo* snapshot, the distance among plots is evident as shown in Fig. 11(b).

Moreover, the Figure 11 depicts the analysis of the difference between the likeness degree of the de-morphing image and a previous *chip* image with blue. The figure also shows the likeness degree of the de-morphing image and the *in vivo* image is illustrated with orange. From these two plots,

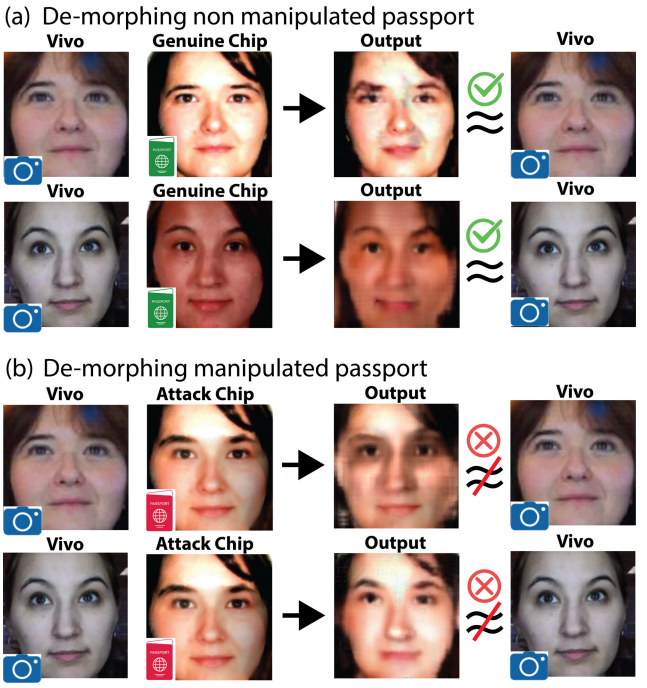

FIGURE 10. De-morphing process in passports (a) with non-morphed face image and (b) with morphed face image.

the probability density function can be calculated to detect a morphing attack, as shown in Fig 11 (c). The probability density function has been computed from the difference of similarities from de-morphing image and the passport and *in vivo* images.

The de-morphing process is performed by a convolutional network (see Fig. 12(c)), which is composed of two extraction branches of features. These features are based on an

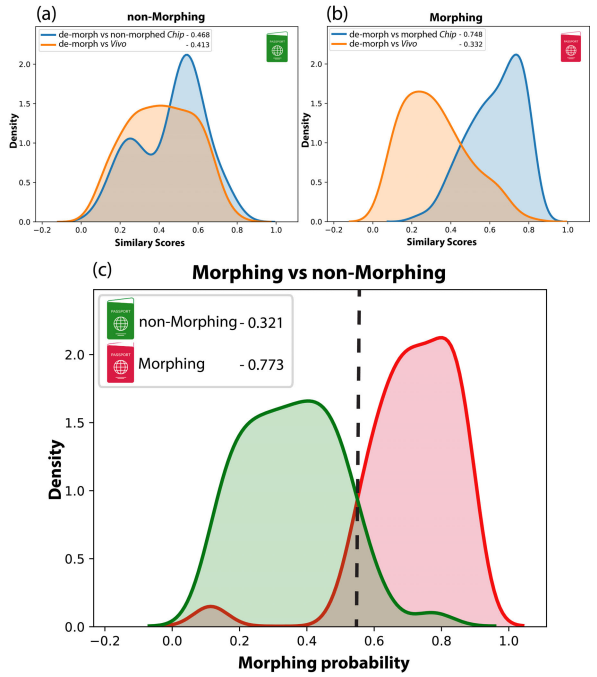

FIGURE 11. Density graph of similarity scores of FaceNet [56] calculated between de-morphing result image and *in vivo* image (a) with non-morphed face image and (b) with morphed face image. (C) Probability distribution in morphing and non-morphing passport images.

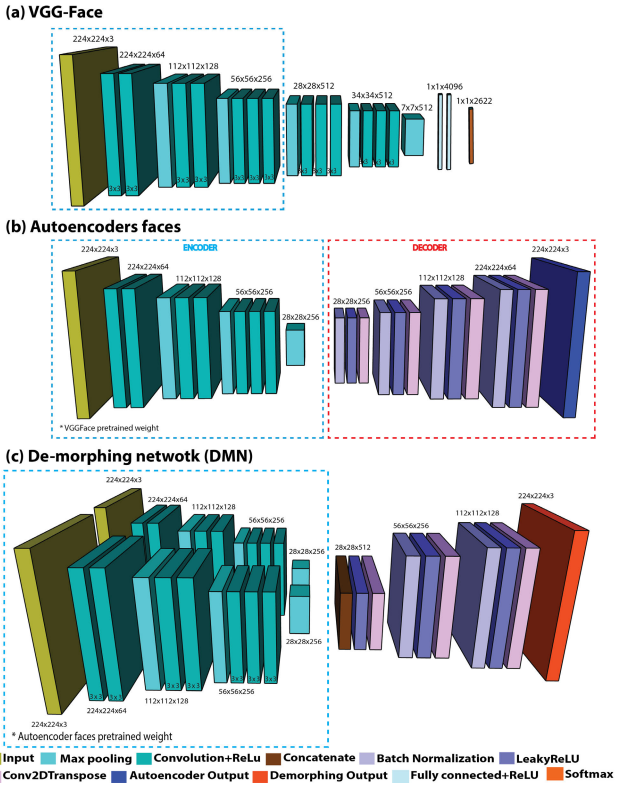

FIGURE 12. Networks architectures: (a) VGG-Face [57], (b) Autoencoder and (c) de-morphing network (DMN).

autoencoder (see Fig. 12(b). This autoencoder is divided into several layers for reconstruction of the image. As depicted in Fig. 13, the first layers of the convolutional network extract the features of input images (*chip* and *in vivo*). The transposed convolution of reconstruction layers are in charge of distinguishing whether these features are not located in both images. Indeed, if this were the case, it could be assumed that the process obtains the criminal's features.

B. AUTOENCODER

The autoencoder whose architecture is described in Fig. 12(b) is composed of two stages, encoding and decoding. Both stages belong to a specific convolutional neural network (CNN) similar to that described in [63]. In the first stage, encoding (denoted by blue striped line), reduces the initial shape image $224\times224\times3$ to $28\times28\times256$ without losing critical attribute information, as depicted in Fig. 14. This figure shows the original input images (column a) and on the right side, the output reconstructed images after the encoder process (column b). The main facial information remains during the process.

The second stage, decoding, provides the original input image ($224\times224\times3$) using transposed convolution successive layers (denoted by red striped line) [81]. The architecture of encoder layers of the autoencoder is the same as the VGG-Face first layers [57], as depicted in Fig. 12(a). VGG-Face is a CNN implementation designed to identify and verify individuals with high accuracy rates such as 98.95% when they use *Labeled Faces in the Wild* (LFW) [65] or 97.3% when they use *YouTube Faces* (YTF) [66]. It should be noted that VGG-Face provides pretrained weights with 2.6 million faces. As shown in the state-of-the-art section, these kinds of neural networks are well suited for extracting information from face images.

As explained above, VGGFaces have pretrained weights in their first layers but they were insufficient to obtain good results in the current problem. Thus, the decoder layers should be trained to achieve final high accuracy rates.

On the one hand, the autoencoder was trained with TensorFlow 1.15 library with CASIAWebFace pictures, using the same identities (faces) as inputs and outputs in every single step. The autoencoder was trained approximately with 3000 epochs or iterations, with 512 samples per batch, using mean squared error (MSE). Moreover, the option of "early stopping (patience = 500)" was used in all scenarios, that is, if the algorithm did not improve in 500 iterations, then it was stopped. The graphic card used to train the current autoencoder was a NVIDIA GeForce GTX 1050 (8 GB RAM). This study relied on the learning rate used in [82]. The learning rate for model fine-tuning starts from 0.005 and decreases to 0.001. Finally, an adaptive momentum (Adam) [83] was used as the optimization algorithm.

On the other hand, it is necessary to assess the similarity between the input and output images to test the autoencoder yield. To perform this assessment, 170 *chip* images of *FRAV-ABC-Test* corpus were processed and calculated by

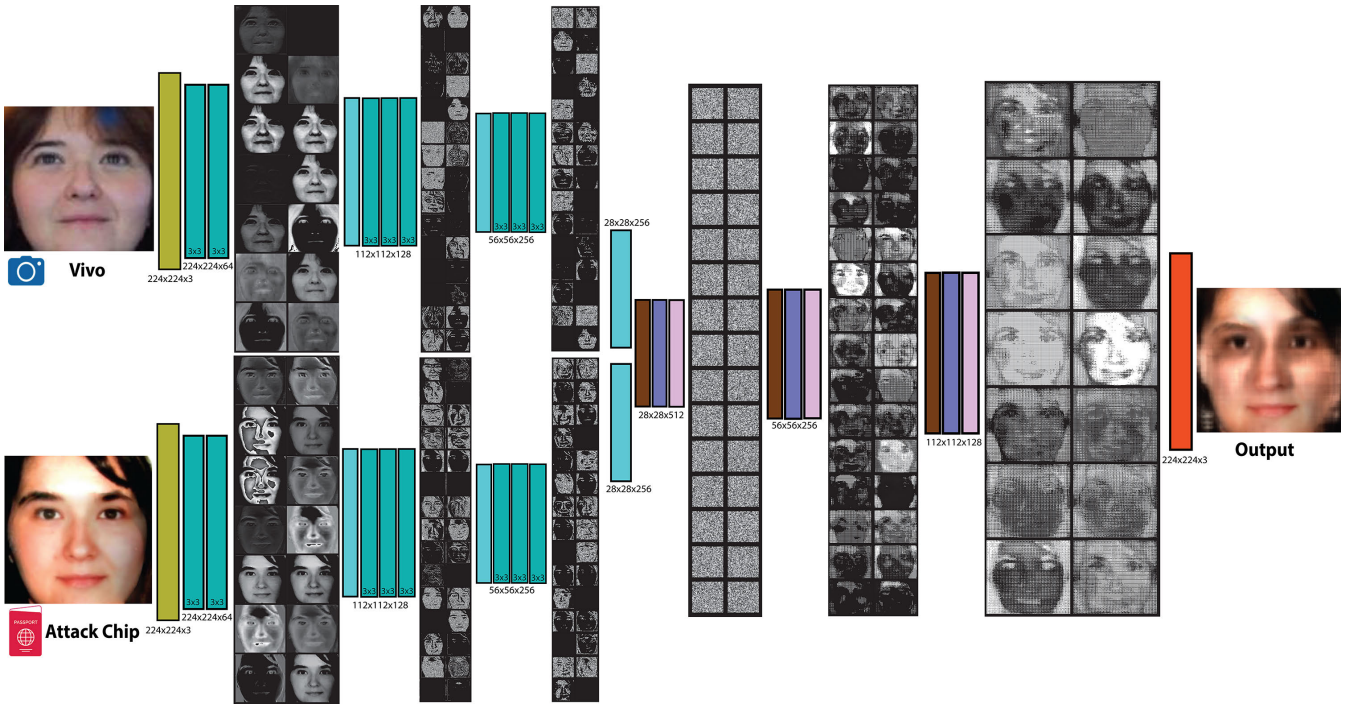

FIGURE 13. De-morphing explanation. Details of Figure 12 (c).

FIGURE 14. Autoencoder process with face image. Autoencoder process example: Original images (a) and Reconstructed images after autoencoder information extraction (b).

autoencoder and FaceNet facial verification acceptance probability. Moreover, FaceNet was used instead of the VGGFace corpus because the first one avoided noisy outcomes caused by the use of the same autoencoder's architecture.

C. DE-MORPHING FACES

Once the encoding process has ended, two images are obtained as output. Their sizes are $28\times28\times256$. After that, these images are concatenated with only one output image whose size is $28\times28\times512$. This image merges the two previous images' information. Finally, the decoder returns an output image with the original resolution ($224\times224\times3$), using transposed convolution successive layers, as depicted in Fig. 12(c).

The training process of the de-morphing neural network is based on a supervised classification algorithm like all CNNs. To obtain a robust training corpus, it is necessary to perform a large number of combinations. The training subjects were 1000. From those subjects, 700 were used as the training set and 300 as the validation set. Therefore, all combinations increase to approximately one million morphing images. The network was trained with the TensorFlow 1.15 library and was trained in 5000 epochs or iterations, with 512 samples per batch, with GeForce GTX 1050 (8 GB RAM), using mean squared error (MSE). As in [82], the learning rate for model fine-tuning starts from 0.005 and decreases to 0.001. Finally, an adaptive momentum (Adam) [83] as the optimization algorithm has been used.

VI. RESULTS AND DISCUSSION

This section presents the evaluation metrics commonly followed in the morphing attacks detection approaches and the results obtained in the presented work.

A. EVALUATION METRICS

Recently, the community has achieved a common standard ISO (IEC 30107-3:2016) [84] to evaluate PAD systems. In this standard, the capability of the attack detection is measured with the following errors: attack presentation classification error rate (APCER) and bona fide presentation classification error rate (BPCER). This measure can be defined as follows:

- **Attack presentation classification error rate (APCER)** is defined as the proportion of presentation attacks that have been classified incorrectly (as bona fide) [84] (Equation 2).
- **Bona fide presentation classification error rate (BPCER)** is defined as the proportion of bona fide presentation incorrectly classified as presentation attacks [84] (Equation 3).

$$APCER_{PAIs} = 1 - \left(\frac{1}{|PAI|}\right) \sum_{\omega=1}^{|PAI|} (RES_\omega), \tag{2}$$

where $|PAI|$ is the number of presentation attack instruments (PAI) and RES_ω takes the value 1 if the presentation ω is assessed as attack and 0 if it is evaluated as bona fide. A PAI is defined as a used object or biometric trait in a presentation attack.

$$BPCER_{PAIs} = \frac{\sum_{\omega=1}^{|BF|} RES_\omega}{|BF|}, \tag{3}$$

where $|BF|$ is the cardinality of bona fide presentations and RES_i returns the value 1 if the presentation ω is allocated as an attack and 0 if it is analyzed as bona fide.

An APCER-BPCER DET curve (detection error trade-off) and the EER (equal error rate) where both errors are identical, provides a comparison among MAD systems.

B. RESULTS

The study estimates the quality of morphing attack detection. It explores its potential application in ABC, considering the images of the *FRAV-ABC* dataset acquired in a real ABC system. This research work presents a specific set of experiments concentrated in combinations of different kinds of pictures such as *in vivo*, *chip*, and Print & Scan photos, as described in Fig. 15.

APCER and BPCER errors of three corpora (*FRAV-ABC-Test*, *FRAV-ABC-Test-P&S-300* and *FRAV-ABC-Test-P&S-150*) are shown in Fig. 15. Each curve represents results with one database. Curves closer to the origin (bottom left) present a lower EER and therefore represent better performance. The accuracy rate obtained increases to 98% in all corpora. The first corpus, *FRAV-ABC-Test*, obtained 0.78 EER and an accuracy rate of 98.7% with a similar threshold. This corpus contains the original images without any compression or downsampling. The two other corpora, *FRAV-ABC-Test-P&S-300* and *FRAV-ABC-Test-P&S-150*, obtained analogous outcomes. The second corpus, *FRAV-ABC-Test-P&S-300*, achieved an EER of 0.80 and an accuracy rate of 98.2%. The third corpus, *FRAV-ABC-Test-P&S-150*, yielded the worst result but very similar to the previous one, obtaining an EER of 1.20 and an accuracy rate of 98.1%.

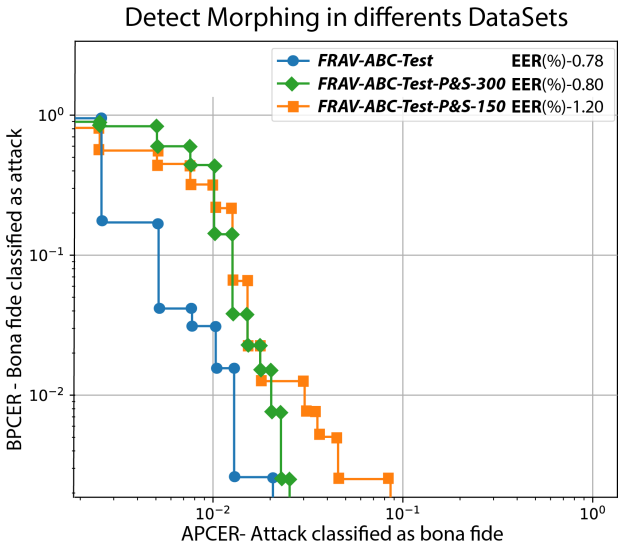

FIGURE 15. APCER and BPCER DET in *FRAV-ABC-Test*, *FRAV-ABC-Test-P&S-300* and *FRAV-ABC-Test-P&S-150*.

The best results are obtained using digital images. It can be seen that the curve obtained from digital images is closer to the origin; therefore, this curve has the lowest EER. In the case of P&S images, the performance is very similar even when using a different resolution. Since this is the procedure followed by the passport issue authorities, it should be remarked that the results are quite similar independent of the resolution considered. Increasing the resolution will not have a significant improvement in the attack detection results.

TABLE 1. Detect morphing in digital images and in printed and scanned images.

Type	Environment	Approach	D-EER (%)	ACC (%)
Digital	Lab	[13]	7.10	-
	Lab	[85]	8.20	-
	Lab	[80]	0.90	99.3
	ABC	DMN	**0.78**	**98.7**
P&S	Lab	[13]	20.70	-
	Lab	[85]	12.50	-
	Lab	[80]	6.10	93.5
	ABC (300dpi)	DMN	**0.80**	**98.2**
	ABC (150dpi)	DMN	**1.20**	**98.1**

A comparison between several MAD approaches presented in the literature and the results obtained in this study is pointed out in Table 1. Both the EER and accuracy values are depicted to properly explain the system behavior. The results are shown using both input types of images (digital or print and scan). The environment conditions in the data acquisition task have also been added to this table. The outcomes of this research work are the only results that have been acquired under real ABC conditions. The results achieved exceed the

outcomes obtained in several studies presented in the literature. On the one hand, the EER of digital images is 14% lower than the best result accomplished in the literature and one order of magnitude better than the others. The accuracy obtained is similar to the state of the art. On the other hand, the EER of print and scan images is one order of magnitude better than all others. Thus, the use of printed and scanned images with different qualities (300 or 150 dpi) is not significant. However, the accuracy is even better than the only value reported in the table.

The calculation capacity of this study is notable. On the one hand, the average time of the full de-morphing process was approximately 5 seconds. This time was calculated with test images on a personal computer. The characteristics of the test environment are as follows: Intel® Core™ i7 motherboard with 8 GB RAM. Note that Frontex recommends the time to be less than 30 seconds [18].

On the other hand, the average time of the DMN process was 3.726 seconds and 0.403 for each of the two verifications (image de-morphing vs image *in vivo* and image de-morphing vs image *chip*). Thus, the final time was 4.532 seconds ($3.726 + 3 \times 0.403$).

VII. CONCLUSION

This research work proposes a new de-morphing-based approach using a CNN to detect morphing presentation attacks in real automated border control systems. A current CNN architecture has been adapted to this specific problem. A neural network was trained with different images such as *in vivo*, *chip*, Print, Scan, and Print & Scan. A deep evaluation was carried out to check and assess the morphing attack detection capability. The assessment has been performed taking into account two images (*in vivo* and passport *chip*), which is currently the most typical situation in border control.

Regarding the experimental results, it can be concluded that the CNN paradigm is suitable for morphing attack detection, obtaining relevant outcomes. The print and scan results achieved are remarkably better than other aforementioned research works. A significant influence of the dpi scan resolution in detection attack outcomes has not been shown.

The results achieved in digital images are significantly better than ''print and scan'' samples and improve the values obtained in the literature. The comparison of outcomes was performed against three different studies, and the current research work enhanced the previous studies by one order of magnitude.

One of the most relevant aspects is the improvement of quality and visual aspects of the pictures achieved after the de-morphing process. Moreover, the de-morphing network is perfectly adapted in the ABC systems procedure. In addition to the foregoing, the hidden identity of the impostor is attained. This feature could be very useful for other future applications.

Towards this point, future work is envisioned that would increase the number of users of our own database, so that by adding samples to the database, better training performance and more reliable results from the testing procedure could be obtained. Moreover, paying more attention to feature selection for the CNN would enhance the outcomes.

REFERENCES

[1] D. del Campo, C. Conde, A. Serrano, I. Diego, and E. Cabello, ''Face recognition-based presentation attack detection in a two-step segregated automated border control E-gate-results of a pilot experience at adolfo suárez Madrid-Barajas airport,'' in *Proc. SECRYPT*, 2017, pp. 129–138.

[2] S. Lee, G. Wolberg, and S. Y. Shin, ''Polymorph: Morphing among multiple images,'' *IEEE Comput. Graph. Appl.*, vol. 18, no. 1, pp. 58–71, Jan./Feb. 1998.

[3] T. Ucicr, ''Feature-based image metamorphosis,'' *ACM SIGGRAPH Comput. Graph.*, vol. 26, no. 2, pp. 35–42, 1992.

[4] G. Wolberg, ''Image morphing: A survey,'' *Vis. Comput.*, vol. 14, nos. 8–9, pp. 360–372, Dec. 1998.

[5] J. M. Beale and F. C. Keil, ''Categorical effects in the perception of faces,'' *Cognition*, vol. 57, no. 3, pp. 217–239, Dec. 1995.

[6] D. T. Levin and J. M. Beale, ''Categorical perception occurs in newly learned faces, other-race faces, and inverted faces,'' *Perception Psychophys.*, vol. 62, no. 2, pp. 386–401, Jan. 2000.

[7] D. J. Robertson, A. Mungall, D. G. Watson, K. A. Wade, S. J. Nightingale, and S. Butler, ''Detecting morphed passport photos: A training and individual differences approach,'' *Cognit. Res., Princ. Implications*, vol. 3, no. 1, p. 27, Dec. 2018.

[8] M. Ferrara, A. Franco, and D. Maltoni, ''The magic passport,'' in *Proc. IEEE Int. Joint Conf. Biometrics*, Sep. 2014, pp. 1–7.

[9] R. Ramachandra, S. Venkatesh, K. Raja, and C. Busch, ''Towards making morphing attack detection robust using hybrid scale-space colour texture features,'' in *Proc. IEEE 5th Int. Conf. Identity, Secur., Behav. Anal. (ISBA)*, Jan. 2019, pp. 1–8.

[10] M. Ferrara, A. Franco, and D. Maltoni, ''Face demorphing in the presence of facial appearance variations,'' in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 2365–2369.

[11] A. Makrushin and A. Wolf, ''An overview of recent advances in assessing and mitigating the face morphing attack,'' in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 1017–1021.

[12] M. Gomez-Barrero, C. Rathgeb, U. Scherhag, and C. Busch, ''Is your biometric system robust to morphing attacks?'' in *Proc. 5th Int. Workshop Biometrics Forensics (IWBF)*, Apr. 2017, pp. 1–6.

[13] U. Scherhag, R. Raghavendra, K. B. Raja, M. Gomez-Barrero, C. Rathgeb, and C. Busch, ''On the vulnerability of face recognition systems towards morphed face attacks,'' in *Proc. 5th Int. Workshop Biometrics Forensics (IWBF)*, Apr. 2017, pp. 1–6.

[14] U. Scherhag, C. Rathgeb, J. Merkle, and C. Busch, ''Deep face representations for differential morphing attack detection,'' 2020, *arXiv:2001.01202*. [Online]. Available: http://arxiv.org/abs/2001.01202

[15] M. Ferrara, R. Cappelli, and D. Maltoni, ''On the feasibility of creating double-identity fingerprints,'' *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 4, pp. 892–900, Apr. 2017.

[16] C. Rathgeb and C. Busch, ''On the feasibility of creating morphed iris-codes,'' in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2017, pp. 152–157.

[17] FRONTEX. (2019). *Frontexonline*. [Online]. Available: https://frontex.europa.eu/

[18] Frontex (EU Body or Agency), ''Best practice technical guidelines for automated border control (ABC) systems,'' FRONTEX, Res. Develop. Unit, Tech. Rep., May 2016. [Online]. Available: https://op.europa.eu/en/publication-detail/-/publication/e81d082d-20a8-11e6-86d0-01aa75ed71a1 and https://op.europa.eu/s/n6qM

[19] S. Jia, G. Guo, and Z. Xu, ''A survey on 3D mask presentation attack detection and countermeasures,'' *Pattern Recognit.*, vol. 98, Feb. 2020, Art. no. 107032.

[20] R. Ramachandra and C. Busch, ''Presentation attack detection methods for face recognition systems: A comprehensive survey,'' *ACM Comput. Surv. (CSUR)*, vol. 50, no. 1, pp. 1–37, 2017.

[21] N. Damer, K. Dimitrov, R. C. Wilson, E. R. Hancock, and W. A. Smith, ''Practical view on face presentation attack detection,'' in *Proc. BMVC*, 2016, pp. 1–11.

[22] 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.

[23] M. Ferrara, A. Franco, and D. Maltoni, "Face demorphing," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 4, pp. 1008–1017, Apr. 2018.
[24] F. Peng, L.-B. Zhang, and M. Long, "FD-GAN: Face de-morphing generative adversarial network for restoring Accomplice's facial image," *IEEE Access*, vol. 7, pp. 75122–75131, 2019.
[25] N. Damer, A. M. Saladie, A. Braun, and A. Kuijper, "MorGAN: Recognition vulnerability and attack detectability of face morphing attacks created by generative adversarial network," in *Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Oct. 2018, pp. 1–10.
[26] GIMP. (1996). *Gimp (GNU Image Manipulation Program)*. [Online]. Available: https://www.gimp.org
[27] R. Raghavendra, K. B. Raja, and C. Busch, "Detecting morphed face images," in *Proc. IEEE 8th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Sep. 2016, pp. 1–7.
[28] U. Scherhag, C. Rathgeb, and C. Busch, "Towards detection of morphed face images in electronic travel documents," in *Proc. 13th IAPR Int. Workshop Document Anal. Syst. (DAS)*, Apr. 2018, pp. 187–192.
[29] U. Scherhag, C. Rathgeb, and C. Busch, "Morph deterction from single face image: A multi-algorithm fusion approach," in *Proc. 2nd Int. Conf. Biometric Eng. Appl. (ICBEA)*, 2018, pp. 6–12.
[30] U. Scherhag, C. Rathgeb, and C. Busch, "Performance variation of morphed face image detection algorithms across different datasets," in *Proc. Int. Workshop Biometrics Forensics (IWBF)*, Jun. 2018, pp. 1–6.
[31] L. Spreeuwers, M. Schils, and R. Veldhuis, "Towards robust evaluation of face morphing detection," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 1027–1031.
[32] L. Debiasi, U. Scherhag, C. Rathgeb, A. Uhl, and C. Busch, "PRNU-based detection of morphed face images," in *Proc. Int. Workshop Biometrics Forensics (IWBF)*, Jun. 2018, pp. 1–7.
[33] L. Debiasi, C. Rathgeb, U. Scherhag, A. Uhl, and C. Busch, "PRNU variance analysis for morphed face image detection," in *Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Oct. 2018, pp. 1–9.
[34] C. Kraetzer, A. Makrushin, T. Neubert, M. Hildebrandt, and J. Dittmann, "Modeling attacks on photo-ID documents and applying media forensics for the detection of facial morphing," in *Proc. 5th ACM Workshop Inf. Hiding Multimedia Secur. (IHMMSec)*, 2017, pp. 21–32.
[35] L.-B. Zhang, F. Peng, and M. Long, "Face morphing detection using Fourier spectrum of sensor pattern noise," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2018, pp. 1–6.
[36] A. Makrushin, T. Neubert, and J. Dittmann, "Automatic generation and detection of visually faultless facial morphs," in *Proc. 12th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2017, pp. 39–50.
[37] C. Seibold, A. Hilsmann, and P. Eisert, "Reflection analysis for face morphing attack detection," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 1022–1026.
[38] C. Seibold, W. Samek, A. Hilsmann, and P. Eisert, "Accurate and robust neural networks for security related applications exampled by face morphing attacks," 2018, *arXiv:1806.04265*. [Online]. Available: http://arxiv.org/abs/1806.04265
[39] L. Wandzik, G. Kaeding, and R. V. Garcia, "Morphing detection using a General- purpose face recognition system," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 1012–1016.
[40] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Gray scale and rotation invariant texture classification with local binary patterns," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2000, pp. 404–420.
[41] R. Raghavendra, K. Raja, S. Venkatesh, and C. Busch, "Face morphing versus face averaging: Vulnerability and detection," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2017, pp. 555–563.
[42] A. Asaad and S. Jassim, "Topological data analysis for image tampering detection," in *Proc. Int. Workshop Digital Watermarking*. Berlin, Germany: Springer, 2017, pp. 136–146.
[43] S. Jassim and A. Asaad, "Automatic detection of image morphing by topology-based analysis," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 1007–1011.
[44] A. Agarwal, R. Singh, M. Vatsa, and A. Noore, "SWAPPED! Digital face presentation attack detection via weighted local magnitude pattern," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2017, pp. 659–665.
[45] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[46] T. Neubert, "Face morphing detection: An approach based on image degradation analysis," in *Proc. Int. Workshop Digital Watermarking*. Berlin, Germany: Springer, 2017, pp. 93–106.
[47] J. Kannala and E. Rahtu, "Bsif: Binarized statistical image features," in *Proc. 21st Int. Conf. Pattern Recognit. (ICPR)*, Nov. 2012, pp. 1363–1366.
[48] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.
[49] C. Shu, X. Ding, and C. Fang, "Histogram of the oriented gradient for face recognition," *Tsinghua Sci. Technol.*, vol. 16, no. 2, pp. 216–224, Apr. 2011.
[50] J. Luka, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 2, pp. 205–214, Jun. 2006.
[51] C. Seibold, W. Samek, A. Hilsmann, and P. Eisert, "Detection of face morphing attacks by deep learning," in *Proc. Int. Workshop Digit. Watermarking*. Berlin, Germany: Springer, 2017, pp. 107–120.
[52] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: http://arxiv.org/abs/1409.1556
[53] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[54] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
[55] L. Debiasi, N. Damer, A. M. Saladié, C. Rathgeb, U. Scherhag, C. Busch, F. Kirchbuchner, and A. Uhl, "On the detection of GAN-based face morphs using established Morph detectors," in *Proc. Int. Conf. Image Anal. Process.* Berlin, Germany: Springer, 2019, pp. 345–356.
[56] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 815–823.
[57] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Proc. BMVC*, 2015, vol. 1, no. 3, p. 6.
[58] U. Scherhag, D. Budhrani, M. Gomez-Barrero, and C. Busch, "Detecting morphed face images using facial landmarks," in *Proc. Int. Conf. Image Signal Process.* Berlin, Germany: Springer, 2018, pp. 444–452.
[59] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1867–1874.
[60] *Machine Readable Travel Documents*, document 9303, 2006.
[61] ICAO. (2018). *ICAO-Innovation Communication Automation Digitalization Operation*. [Online]. Available: https://www.icao.int
[62] A. Ng, "Machine learning yearning: Technical strategy for AI engineers, in the era of deep learning, draft version 0.5," Harvard Bus. Publishing, Boston, MA, USA, Tech. Rep., 2016. [Online]. Available: https://www.deeplearning.ai/
[63] I. Goodfellow, Y. Casiabengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org
[64] D. Sandberg. *Public Re-Implementation of Facenet*. Accessed: Mar. 3, 2020. [Online]. Available: https://github.com/davidsandberg/facenet
[65] E. Learned-Miller, G. B. Huang, A. RoyChowdhury, H. Li, and G. Hua, "Labeled faces in the wild: A survey," in *Advances in Face Detection and Facial Image Analysis*. Berlin, Germany: Springer, 2016, pp. 189–248.
[66] L. Wolf, T. Hassner, and I. Maoz, *Face Recognition in Unconstrained Videos with Matched Background Similarity*. Piscataway, NJ, USA: IEEE Press, 2011.
[67] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch," 2014, *arXiv:1411.7923*. [Online]. Available: http://arxiv.org/abs/1411.7923
[68] AbroSoft. (2019). *FantaMorph, Version 2.0*. Accessed: Mar. 4, 2020. [Online]. Available: http://www.fantamorph.com
[69] MorpheusSoftware. (2007). *Morpheus Photo Morpher, Version v3.17*. Accessed: Mar. 4, 2020. [Online]. Available: https://www.morpheussoftware.net/
[70] MorphThing. (2019). *Morphthing*. Accessed: Mar. 4, 2020. [Online]. Available: http://www.morphthing.com/
[71] Foundry. (2016). *Nuke*. Accessed: Mar. 4, 2020. [Online]. Available: https://www.foundry.com/products/nuke
[72] SilhouetteFX. (2019). *Silhouettefx, Version v7*. Accessed: Mar. 4, 2020. [Online]. Available: https://www.foundry.com/products/nuke
[73] Luxand Technology. (2019). *FaceMorpher, Version Luxand*. Accessed: Mar. 4, 2020. [Online]. Available: http://www.facemorpher.com
[74] S. Kumar. (2019) *WinMorph, Version 3.01*. Accessed: Mar. 4, 2020. [Online]. Available: https://www.debugmode.com/winmorph/
[75] Dlibk. (2020). *Dlib C++ Toolkit*. Accessed: Mar. 4, 2020. [Online]. Available: https://dlib.net/

[76] S. V. Lokhande and S. B. Patil, "Morphing techniques for facial images-a review," *Int. J. Eng.*, vol. 2, no. 12, pp. 1106–1110, 2013.
[77] J. Wu, "Face recognition jammer using image morphing," Dept. Elect. Comput. Eng., Boston Univ., Boston, MA, USA, Tech. Rep. ECE-2011, 2011.
[78] M. Hildebrandt, T. Neubert, A. Makrushin, and J. Dittmann, "Benchmarking face morphing forgery detection: Application of stirtrace for impact simulation of different processing steps," in *Proc. 5th Int. Workshop Biometrics Forensics (IWBF)*, Apr. 2017, pp. 1–6.
[79] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," in *Proc. ACM SIGGRAPH Papers*, 2003, pp. 313–318.
[80] M. Ferrara, A. Franco, and D. Maltoni, "Face morphing detection in the presence of printing/scanning and heterogeneous image sources," 2019, *arXiv:1901.08811*. [Online]. Available: http://arxiv.org/abs/1901.08811
[81] M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus, "Deconvolutional networks," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2528–2535.
[82] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A dataset for recognising faces across pose and age," in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2018, pp. 67–74.
[83] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980
[84] *Information Technology -Biometric Presentation Attack Detection—Part 3: Testing and Reporting*, Standard 30107-3:2017, International Organization for Standardization, Geneva, Switzerland, Mar. 2017.
[85] R. Raghavendra, K. B. Raja, S. Venkatesh, and C. Busch, "Transferable deep-CNN features for detecting digital and print-scanned morphed face images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 1822–1830.

DAVID ORTEGA-DELCAMPO was born in Madrid, Spain. He received the B.S. degree in computer engineering and the M.S. degree in artificial vision from Universidad Rey Juan Carlos (URJC), in 2002 and 2015, respectively, where he is currently pursuing the Ph.D. degree in computer science. He worked as an Engineer in a private company for 14 years. His research interests include the development of several neural network approaches, morphing and de-morphing processes, pattern recognition, machine learning, image processing, biometrics, and security. Mr. Ortega-Delcampo is a member of the Face Recognition and Artificial Vision Group.

CRISTINA CONDE received the B.S. degree in physics (electronics) from the Complutense University of Madrid, in 1999, and the Ph.D. degree from the University Rey Juan Carlos, Madrid, in 2006. She was several years working in the private sector. In 2001, she joined the University Rey Juan Carlos as an Assistant Professor. For seven years, she was the Vice-Dean of Studies with the Computer Science School. Her research interests include image and video analysis, pattern recognition, and machine learning in both, classical and biologically inspires computation. She has coordinated several National and European Projects.

DANIEL PALACIOS-ALONSO was born in Madrid, Spain. He received the B.S. and M.S. degrees in computer science from the Universidad Politécnica de Madrid (UPM), in 2009, and the Ph.D. degree in advanced computation from UPM, in 2017. He worked as a Team Leader in a technological consulting firm for five years. Since 2013, he has been a member of the Neuromorphic Speech Processing Laboratory, Center for Biomedical Technology. He is currently an Assistant Professor with Universidad Rey Juan Carlos (URJC). He is also the Head of the Bioinspired Systems and Applications Group. His research interests include the stress and emotional states, neurodegenerative diseases such as Parkinson, ALS, Alzheimer's, among others, artificial vision, pattern recognition, and biomedical signal processing. He is a Reviewer of National and International Journals. Dr. Palacios-Alonso was a recipient of several Best Paper Awards (ICPRS-2016, BIOSIGNALS-2019) and the Doctoral Consortium Award of Spanish Association of Artificial Intelligence in 2013.

ENRIQUE CABELLO received the B.S. degree in physics (electronics) from the University of Salamanca, and the Ph.D. degree from the Polytechnic University of Madrid. In 1990, he joined the Computer Science Department, University of Salamanca. Since 1998, he has been with Universidad Rey Juan Carlos, where he has been the Head of the Face Recognition and Artificial Vision Group, since 2001. He is currently the Head of the Computer Science and Statistics Department. His research interests include image and video analysis, pattern recognition, and machine learning, using classic, and bio-inspired approaches.